Calculus 1

Thuyen Dang

November 23, 2026

Contents

Preface v

1 Functions 1
- 1.1 What is a Function? . . . 1
- 1.2 Visualizing Functions: Graphs . . . 2
- 1.3 New Functions from Old . . . 3
  - 1.3.1 Arithmetic of Functions . . . 3
  - 1.3.2 Composition of Functions . . . 4
- 1.4 Symmetries and Transformations . . . 5
  - 1.4.1 Transforming the Graph . . . 5
  - 1.4.2 Even and Odd Functions . . . 5
- 1.5 Inverse Functions . . . 6
  - 1.5.1 The Concept of the Inverse . . . 6
  - 1.5.2 Existence of Inverse and the Horizontal Line Test . . . 7
  - 1.5.3 Finding the Inverse . . . 8
  - 1.5.4 Piecewise Functions . . . 8
- 1.6 Classification of Functions . . . 9
- 1.7 Trigonometric Functions: The Geometry of the Circle . . . 10
  - 1.7.1 Radians versus Degrees . . . 10
  - 1.7.2 Definitions from the Unit Circle . . . 10
  - 1.7.3 The Complex Connection . . . 11
  - 1.7.4 Inverse Trigonometric Functions and the Art of Restriction . . . 13
- 1.8 Exponential and Logarithmic Functions . . . 16

2 Limits 19
- 2.1 The Big Picture: Why Do We Need Limits? . . . 19
- 2.2 Two Fundamental Problems Using Limits . . . 21
  - 2.2.1 The Tangent Line Problem . . . 21
  - 2.2.2 The Area Problem . . . 22
- 2.3 Finding Limits Intuitively: Graphs and Numerics . . . 23
- 2.4 Fractured Journeys: One-Sided Limits . . . 25
  - 2.4.1 Formalizing the Directional Approach . . . 26
  - 2.4.2 Divergence: To Infinity and Beyond . . . 27
  - 2.4.3 The Topologist's Sine Curve: A Frantic Oscillation . . . 28
- 2.5 Formalizing the Limit, Rigorously . . . 29
- 2.6 The Zoo of Limit Behaviors: What Could Go Wrong? . . . 34

2.7 Continuous Function . . . 35
2.7.1 Global Continuity and the Limit Cheat Codes . . . 37
2.7.2 Discontinuity: What Does It Mean to Break? . . . 39
2.7.3 Continuous Functions Commute with Limits . . . 40
2.8 The Limit Laws . . . 43
2.9 The Squeeze Theorem . . . 45
2.10 Limit at Infinity . . . 52
2.11 Asymptotes . . . 56
2.11.1 Horizontal Asymptotes: The Ends of the Universe . . . 58
2.11.2 Vertical Asymptotes: Singularities . . . 59
2.11.3 Slant (Oblique) Asymptotes: Linear Escapes . . . 60
2.12 The Intermediate Value Theorem . . . 63

3 Derivative **67**
3.1 Introduction to the Derivative . . . 67
3.1.1 From Secants to Tangents . . . 67
3.1.2 The Derivative as a Function . . . 70
3.2 The Anatomy of Non-Differentiability . . . 71
3.2.1 The Corner . . . 71
3.2.2 The Vertical Tangent . . . 72
3.2.3 The Discontinuity . . . 74
3.3 Differentiable Implies Continuous . . . 75
3.4 Differentiation Rules . . . 77
3.4.1 Linearity: The Easiest Combinations . . . 77
3.4.2 The Product Rule . . . 78
3.4.3 The Power Rule: Climbing the Ladder . . . 80
3.5 The Quotient Rule . . . 82
3.5.1 Differentiating the Reciprocal . . . 82
3.5.2 Synthesizing the Quotient Rule . . . 83
3.5.3 The Hierarchy of Rules . . . 85
3.6 Derivatives of Trigonometric Functions . . . 86
3.6.1 The Algebraic Path: Hadamard's Advice . . . 86
3.6.2 The Geometric Path: Newton's Vision . . . 88
3.7 Asymptotic Analysis: A Scenic Route to Calculus . . . 91
3.7.1 Small o and Big O . . . 91
3.7.2 Newton's Approximation in Disguise . . . 92
3.8 The Chain Rule . . . 95
3.8.1 The Chains of Leibniz . . . 97
3.8.2 The Derivative of the Exponential Function . . . 99
3.8.3 Higher Derivatives . . . 101
3.8.4 Cycles, Rotations, and Trigonometry: The Math of Oscillation . . . 101
3.9 Derivatives of Inverse Functions . . . 103
3.9.1 Derivatives of the Inverse Trigonometric Functions . . . 105
3.9.2 Derivatives of Logarithms and Radicals . . . 107
3.10 A Bag of Logarithmic Tricks . . . 108
3.11 Where do I begin? . . . 112
3.12 Implicit Differentiation . . . 114
3.12.1 Breaking Free from Explicit Formulas . . . 115

3.12.2 Self-Intersecting Curves: The Folium of Descartes . . . 116
3.12.3 Iterating the Process: Higher Derivatives . . . 118
3.13 The Physical Meaning of Differentiation: Rate of Change . . . 119
3.13.1 Velocity, Speed, and Acceleration . . . 119
3.13.2 Growth Models . . . 120
3.13.3 Related Rates . . . 121

4 Applications of the Derivative **124**
4.1 l'Hôpital's Rule: The Limits of Indeterminacy . . . 124
4.2 The Extreme Value Theorem . . . 130
4.3 The Mean Value Theorem . . . 134
4.4 First Derivative Dictates Monotonicity . . . 137
4.5 Second Derivative Dictates Concavity . . . 140
4.6 Curve Sketching . . . 143
4.7 Optimization Problems . . . 147
4.8 Differentials and Error Estimation . . . 150
4.9 The Root-Finding Problem and Newton Method . . . 152
4.9.1 Existence via Topological Theorems . . . 152
4.9.2 Discrete Dynamics and Rates of Convergence . . . 153
4.9.3 Construction of Root-Finding Algorithms . . . 154
4.9.4 Higher-Order Approximations . . . 156
4.9.5 Asymptotic Error Analysis of Newton Method . . . 158
4.9.6 Geometric Meaning: Tangent Line Projection . . . 159

5 Integration **163**
5.1 The Big Picture: Motivation of Integration . . . 163
5.2 Signed Quantities and Net Change . . . 165
5.3 Sigma Notation . . . 167
5.4 Riemann Sums . . . 171
5.5 The Definite Integral . . . 175
5.6 Properties of the Definite Integral . . . 177
5.6.1 Bounds and comparison . . . 181
5.6.2 Area as a function of its upper limit: a first glimpse . . . 183
5.7 The Fundamental Theorem of Calculus . . . 185
5.7.1 Statement . . . 186
5.7.2 The Mean Value Theorem for integrals . . . 186
5.7.3 Proof of the Fundamental Theorem . . . 187
5.7.4 Application: antiderivatives and integral evaluation . . . 190
5.7.5 Application: paying the debt of transcendental functions . . . 197
5.8 The Substitution Rule and Change of Variables . . . 203
5.8.1 Inverting the Chain Rule: Indefinite Integrals . . . 203
5.8.2 Substitution for Definite Integrals . . . 206
5.8.3 Geometric Meaning of the Change of Variables . . . 208
5.9 Bonus: Darboux Sums, Compactness, and Integrability . . . 213
5.9.1 Axiomatic Properties and Construction of Darboux Sums . . . 213
5.9.2 The Engine of Global Regularity: Heine–Borel–Cantor . . . 214
5.9.3 Convergence via Uniform Continuity . . . 216

A Midterm 1 Review **218**
A.1 Practice Problems 218
A.2 Bonus Problems 222

B Midterm 2 Review **223**
B.1 Practice Problems 223
B.2 Bonus Problems 228

C Midterm 3 Review **230**
C.1 Practice Problems 230
C.2 Bonus Problems 235

D Final Review **237**
D.1 Practice Problems 237
D.2 Bonus Problems 243

Bibliography **245**

Index **245**

Preface

> There are two ways to do great mathematics. The first way is to be smarter than everybody else. The second way is to be stupider than everybody else–but persistent.
>
> Raoul Bott

This is the first draft of the Calculus 1 notes for the Fall 2026 at Purdue University Northwest. Since this is a rough draft, it's currently on a "for your eyes and ears only" basis for MA 163.

In putting these together, I've leaned heavily on the work of [BCGS18, LT14, LT18, GM03, GM10, Pug15, Ste25, Abb15, Cli26, Tao22a, Tao22b] and the teaching materials of my colleague, Bir Kafle. If these notes make sense, it's likely thanks to them.

These notes started life as a semi-honest attempt to transcribe my own handwritten lecture scribbles. However, I quickly realized that technical details have a habit of acting like industrial-strength melatonin on students. To save us all from a collective nap, I've shifted to a more conversational style. If you're the type of person who finds comfort in the cold, austere rigors of "Baby Rudin," you're going to hate this. You'd be much better served looking elsewhere. This style is also my attempt at a tribute to the legendary teachers who actually made me love physics: Richard Feynman, David Tong, and Savdeep Sethi.

You'll also notice a lot of physics. This isn't an accident. Most of mathematics was essentially invented because we were trying to figure out how the universe works through physics.

None of the math in here is original. It's all classic material. I claim zero originality for the content, though I'll happily take full credit for every typo, blunder, and confusing explanation you find. If you spot something that looks wrong (and you will), please let me know. I'd appreciate the help.

Acknowledgements

These notes were assembled in rather indecent haste. They would almost certainly not exist—and would look considerably worse—were it not for the tireless typist Gemini Enterprise, who took on much of the heavy lifting with the LaTeX typesetting and illustration.

Chapter 1

Functions

> The shortest path between two truths in the real domain passes through the complex domain.
>
> Jacques Hadamard

1.1 What is a Function?

Before we dive into calculus, we must first master the fundamental currency of our subject: the *function*. In physics, a function might describe the trajectory of a particle through spacetime. In pure mathematics, a function acts as a deterministic bridge between sets, a mechanism for transforming one mathematical object into another.

Intuition. The Blender Analogy

Think of a function as a highly sophisticated blender. We feed inputs (ingredients) into the blender, and it processes them according to a specific internal mechanism to produce an output (a smoothie).

- Input $x = \text{Strawberry} \implies f(x) = \text{Strawberry Smoothie}$
- Input $x = \text{Avocado} \implies f(x) = \text{Avocado Smoothie}$

But what happens if we attempt to put rocks into our blender? The blender breaks! This physically motivates a crucial mathematical constraint: not everything belongs in the blender. We must define the valid inputs before we turn the machine on.

Let us elevate our intuition to a formal setting.

Definition 1.1.1 (Function)**.** A *function* f is a rule that applies to each value x in a set D and returns a *unique* value $f(x)$.

This uniqueness is paramount. A function does not roll dice; if you input the same x tomorrow, it will yield the identical $f(x)$ as it did today.

Definition 1.1.2 (Domain and Range)**.** Given a function $f\colon \mathbb{R} \to \mathbb{R}$, the *domain* of f, denoted $\mathrm{dom}(f)$, is defined as the set of all values x such that $f(x)$ is a unique and valid output:

$$\mathrm{dom}(f) \coloneqq \{x \in \mathbb{R} : f(x) \text{ is a unique real number}\}. \tag{1.1}$$

The *range* of f, denoted $\mathrm{ran}(f)$, is the set of all possible outputs produced by the function:

$$\mathrm{ran}(f) \coloneqq \{f(x) : x \in \mathrm{dom}(f)\}. \tag{1.2}$$

Note. In this course, we will only work with functions whose domains are subsets of $\mathbb{R}$, the set of all real numbers.

Example 1.1.3. Consider the squaring function $f(x) = x^2$.

- The domain is $\mathrm{dom}(f) = \mathbb{R}$, since any real number can be squared.
- The range is $\mathrm{ran}(f) = [0, \infty)$. To see this formally, let $y \in [0, \infty)$. The equation $f(x) = y \iff x^2 = y$ admits the real solutions $x = \pm\sqrt{y}$. However, if we let $y \in (-\infty, 0)$, the equation $x^2 = y$ does not have any solutions in $\mathbb{R}$. Thus, negative numbers are not in the range.

Example 1.1.4. Consider the reciprocal function $f(x) = \frac{1}{x}$. This function is undefined when the denominator is zero. Thus, the blender breaks at $x = 0$, leaving us with the domain:

$$\mathrm{dom}(f) = (-\infty, 0) \cup (0, \infty). \tag{1.3}$$

1.2 Visualizing Functions: Graphs

To truly understand a function's behavior, we must visualize it.

Definition 1.2.1 (Graph)**.** The *graph* of a function f is the set of all points in the xy-plane that satisfy the equation $y = f(x)$. Here, x represents the *independent variable* and y represents the *dependent variable*.

Example 1.2.2. For $f(x) = x^2$, the point $(1, 1)$ lies on the graph because when we substitute $x = 1$, we find $y = 1^2 = 1 \implies y = x^2$. Thus, $(1, 1) \in$ graph of $f(x) = x^2$, see Figure 1.1.

> *Moral. The Vertical Line Test*
>
> Because a function must return a *unique* output for each input, a curve represents a function if and only if it passes the Vertical Line Test (VLT): every vertical line intersects the graph at most once.

Example 1.2.3. Does the graph of the equation $x = y^2$ represent a function with variable x? No. If we pick an input $x > 0$, say $x = 4$, we find two corresponding y-values: $y = \sqrt{x}$ and $y = -\sqrt{x}$. The output is not unique. The graph is a sideways parabola that unequivocally fails the VLT, see Figure 1.2.

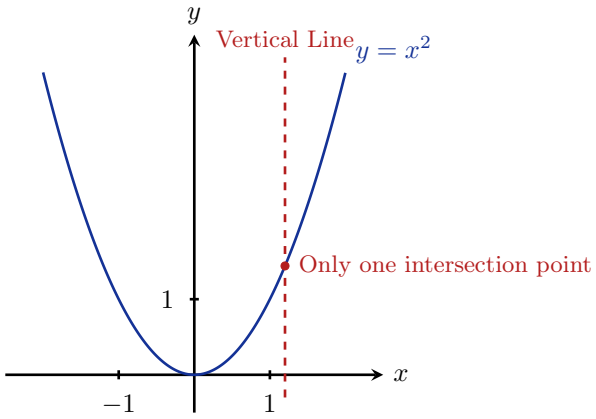


Figure 1.1 – Illustration of the vertical line test on the graph $y = x^2$

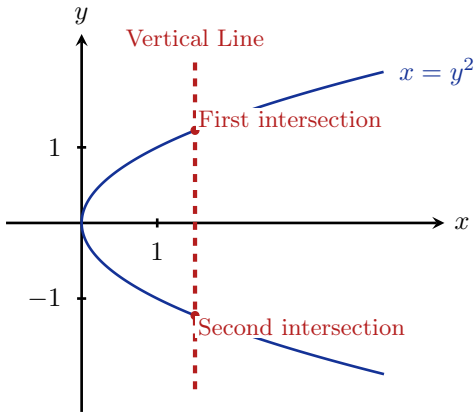


Figure 1.2 – Illustration of the vertical line test on the graph $x = y^2$

1.3 New Functions from Old

Once we have a basic vocabulary of functions, we can combine them to build increasingly complex structures.

1.3.1 Arithmetic of Functions

Given two functions f and g with ranges in $\mathbb{R}$, we define new algebraic combinations:

$$\begin{aligned} h \coloneqq f + g &\iff h(x) = f(x) + g(x) \\ h \coloneqq f - g &\iff h(x) = f(x) - g(x) \\ h \coloneqq fg &\iff h(x) = f(x)g(x) \\ h \coloneqq \frac{f}{g} &\iff h(x) = \frac{f(x)}{g(x)} \end{aligned}$$

⚠ *Caveat. The Subtlety of Domains*

One must be remarkably mindful about domains when performing algebra. Consider:

$$f(x) = \frac{1}{x-1} \quad \text{and} \quad g(x) = x - 1. \tag{1.4}$$

The domains are $\text{dom}(f) = (-\infty, 1) \cup (1, \infty)$ and $\text{dom}(g) = \mathbb{R}$. If we define $h = fg$, we compute:

$$h(x) = f(x)g(x) = \frac{1}{x-1} \cdot (x-1) = \frac{x-1}{x-1}. \tag{1.5}$$

For any $x \neq 1$, this fraction simplifies to 1. However, at $x = 1$, the original function $f(x)$ is undefined. We cannot simply brush this under the algebraic rug! We conclude that $h(x) = 1$ with restricted domain $\text{dom}(h) = (-\infty, 1) \cup (1, \infty)$.

1.3.2 Composition of Functions

Instead of adding outputs, what if we feed the output of g directly into the input of f? This mathematical assembly line is known as composition, see Figure 1.3.

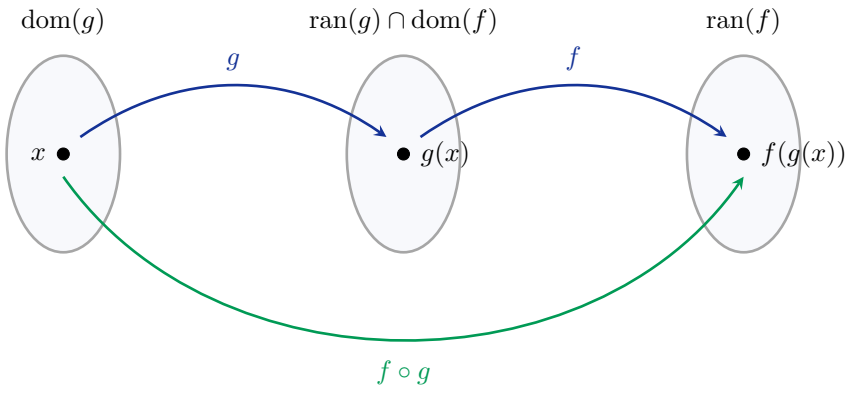


Figure 1.3 – Composition of functions $f \circ g$

Definition 1.3.1 (Composition)**.** The *composition* of f and g, denoted $f \circ g$, is defined as:

$$(f \circ g)(x) = f(g(x)). \tag{1.6}$$

Proposition 1.3.2. *The domain of $f \circ g$ is the set of all $x \in \text{dom}(g)$ such that $g(x) \in \text{dom}(f)$.*

Example 1.3.3. Let $f(x) = 3x^2 - x$ and $g(x) = \frac{1}{x}$. Then:

$$(f \circ g)(x) = f(g(x)) = f\left(\frac{1}{x}\right) = 3\left(\frac{1}{x}\right)^2 - \left(\frac{1}{x}\right) = \frac{3}{x^2} - \frac{1}{x}. \tag{1.7}$$

⚠ *Caveat. Non-Commutativity*

In general, $(f \circ g) \neq (g \circ f)$. Reversing the assembly line fundamentally changes the product. Using the previous example:

$$(g \circ f)(x) = g(f(x)) = g(3x^2 - x) = \frac{1}{3x^2 - x}. \tag{1.8}$$

1.4 Symmetries and Transformations

Mathematicians study the transformations of functions to uncover underlying structure, just as physicists study the symmetries of the universe to uncover conservation laws.

1.4.1 Transforming the Graph

Let $c > 0$ and suppose we start with an original graph $y = f(x)$.

1. *Horizontal shift c units right:* $y = f(x - c)$
2. *Horizontal shift c units left:* $y = f(x + c)$
3. *Vertical shift c units downward:* $y = f(x) - c$
4. *Vertical shift c units upward:* $y = f(x) + c$
5. *Reflection about the x-axis:* $y = -f(x)$
6. *Reflection about the y-axis:* $y = f(-x)$
7. *Reflection about the origin:* $y = -f(-x)$
8. *Scale vertically:* $y = cf(x)$. (Stretch if $c > 1$, shrink if $0 < c < 1$).
9. *Scale horizontally:* $y = f(cx)$. (Compress if $c > 1$, expand if $0 < c < 1$).

Exercise 1.4.1. Draw graphs illustrating all the 9 transformations above.

1.4.2 Even and Odd Functions

Definition 1.4.2 (Symmetry)**.** A function f is *even* if and only if for each $x \in \text{dom}(f)$, $f(x) = f(-x)$. A function f is *odd* if and only if for each $x \in \text{dom}(f)$, $f(x) = -f(-x)$.

Example 1.4.3. 1. Let $f(x) = x^4 - 2x^2 - 20$.

$$f(-x) = (-x)^4 - 2(-x)^2 - 20 = x^4 - 2x^2 - 20 = f(x). \tag{1.9}$$

Thus, f is even.

2. Let $g(x) = x^3 - 3x + 1$.

$$g(-x) = (-x)^3 - 3(-x) + 1 = -x^3 + 3x + 1 = -(x^3 - 3x - 1). \tag{1.10}$$

This is neither $g(x)$ nor $-g(x)$. Thus, g is neither even nor odd.

3. Let $h(x) = x$. We have $h(-x) = -x = -h(x)$, so h is odd.

Exercise 1.4.4. Given a function f, show that $f_e(x) \coloneqq \frac{f(x)+f(-x)}{2}$ is even, and $f_o(x) \coloneqq \frac{f(x)-f(-x)}{2}$ is odd. Moreover, verify the decomposition $f(x) = f_e(x) + f_o(x)$.

1.5 Inverse Functions

If a function f represents a process, can we reverse it? Can we undo the machine?

1.5.1 The Concept of the Inverse

Intuition. Undoing a Process

Consider a highly restricted function $f(x) = x + 3$ mapping $\text{dom}(f) = \{1, 2, 3\}$ to $\text{ran}(f) = \{4, 5, 6\}$. Using ordered pairs, we have:

$$f = \{(1, 4), (2, 5), (3, 6)\}. \tag{1.11}$$

Suppose we are given a value $y \in \text{ran}(f)$. Can we find the original $x \in \text{dom}(f)$ such that $y = f(x)$?

$$\begin{aligned} y = 4 &\implies x = 1, \\ y = 5 &\implies x = 2, \\ y = 6 &\implies x = 3. \end{aligned}$$

This backward mapping defines a new function $g : \text{ran}(f) \to \text{dom}(f)$, represented by $g = \{(4, 1), (5, 2), (6, 3)\}$. Because g "undoes" what f does, we call g the inverse of f. Notice that $x = g(y) = g(f(x))$.

Definition 1.5.1 (Inverse Function)**.** A function g is the *inverse* function of f if:

$$\begin{aligned} g(f(x)) &= x \quad \text{for each } x \in \text{dom}(f), \quad \text{and} \\ f(g(x)) &= x \quad \text{for each } x \in \text{dom}(g). \end{aligned}$$

We traditionally denote the inverse function g by the notation f^{-1} (read as "f inverse").

This definition is illustrated in Figure 1.4.

Caveat. Notation Danger

The notation $f^{-1}(x)$ means the inverse function. It unequivocally does *not* mean the reciprocal $\frac{1}{f(x)}$. This confusion has claimed the sanity of several students!

Remark 1.5.2. Several elegant facts emerge from this definition:

1. If g is the inverse of f, then f is the inverse of g.
2. The domains and ranges swap: $\text{dom}(f^{-1}) = \text{ran}(f)$ and $\text{ran}(f^{-1}) = \text{dom}(f)$.

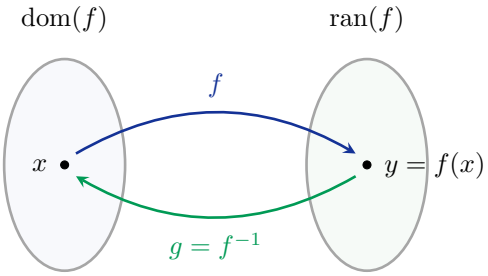


Figure 1.4 – Illustration of an inverse function $g = f^{-1}$

3. A function need not have an inverse, but if it does, then the inverse is unique.
4. If the point (a, b) is in the graph of $y = f(x)$, then (b, a) is in the graph of $y = f^{-1}(x)$, see Figure 1.5.

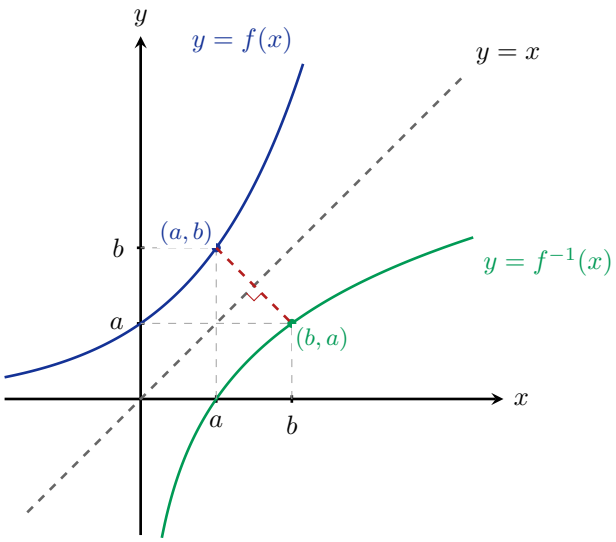


Figure 1.5 – Symmetric relationship of the graphs of $y = f(x)$ and $y = f^{-1}(x)$ reflected across $y = x$

1.5.2 Existence of Inverse and the Horizontal Line Test

When does a function actually have an inverse?

Definition 1.5.3 (1-to-1)**.** A function f is *1-to-1* if and only if $f(x_1) = f(x_2)$ implies $x_1 = x_2$ for all $x_1, x_2 \in \text{dom}(f)$.

Moral. Horizontal Line Test (HLT)

Every horizontal line intersects the graph of a 1-to-1 function at most once.

Theorem 1.5.4. *A function has an inverse if and only if it is 1-to-1.*

Example 1.5.5. Is $f(x) = x^3 - 1$ a 1-to-1 function? Suppose $f(x_1) = f(x_2)$. Then:

$$x_1^3 - 1 = x_2^3 - 1 \implies x_1^3 = x_2^3 \implies \sqrt[3]{x_1^3} = \sqrt[3]{x_2^3} \implies x_1 = x_2.$$

Therefore, f is 1-to-1, and by our theorem, f^{-1} exists.

1.5.3 Finding the Inverse

To find the inverse of $y = f(x)$:

1. Check if the inverse exists (is it 1-to-1 via the HLT or definition?).
2. Solve for x as a function of y: $x = g(y) = f^{-1}(y)$.
3. Interchange x and y to obtain standard formatting: $y = f^{-1}(x)$.
4. Define $\text{dom}(f^{-1}) = \text{ran}(f)$.
5. Verify that $f(f^{-1}(x)) = x$ and $f^{-1}(f(x)) = x$.

Example 1.5.6. For $f(x) = x^3 - 1$, we know it is 1-to-1. We solve:

$$y = x^3 - 1 \iff y + 1 = x^3 \iff \sqrt[3]{y+1} = x. \tag{1.12}$$

Interchanging variables yields $y = \sqrt[3]{x+1}$, so $f^{-1}(x) = \sqrt[3]{x+1}$.

Story. The Limits of Algebraic Formulas

Consider $f(x) = x^5 + x - 1$. Its graph is strictly increasing, so it passes the HLT and is entirely 1-to-1. Can we find the inverse? We attempt to solve:

$$y = x^5 + x - 1 \iff x^5 + x - 1 - y = 0 \tag{1.13}$$

Can we isolate x using roots, as we did for $x^3 - 1$? Galois theory (in abstract algebra) dictates a profound truth: the equation (1.13) *cannot* be solved by radicals! The inverse function absolutely exists, but we are fundamentally barred from writing it down using elementary algebraic formulas. The good new is that we can still visualize it: the graph of the inverse is just the reflection of the original function about the line $y = x$.

1.5.4 Piecewise Functions

Sometimes, a single rule is insufficient to describe a system over its entire domain. The domain is divided into smaller pieces, each governed by its own rule.

Example 1.5.7.

$$g(x) = \begin{cases} x & x \geq 0 \\ x^2 + 1 & x < 0 \end{cases} \tag{1.14}$$

Here, $\text{dom}(g) = \mathbb{R}$, but the function behaves entirely differently depending on the sign of the input, see Figure 1.6.

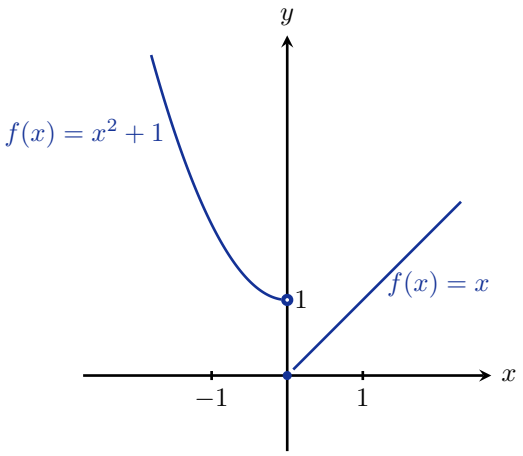


Figure 1.6 – Graph of the piecewise function $f(x)$

1.6 Classification of Functions

To navigate calculus, we classify functions into families:

- **Algebraic Functions:**
  - *Polynomials:* $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$. (e.g., $f(x) = x^2 - 1$).
  - *Radical:* Involving roots like $\sqrt[n]{x}$.
  - *Rational:* Expressions of the form $\frac{p(x)}{q(x)}$, where p and q are polynomials.
- **Transcendental Functions:** those are not algebraic, which include
  - *Trigonometric functions* $(\sin, \cos, \tan, \text{etc.})$
  - *Exponential and Logarithmic functions* $(\mathrm{e}^x, \log x, \text{etc.})$
  - *Non-elementary functions* (error function erf, Gamma function, sine integral, etc. —you will see them in Calculus II.)

Together, the Algebraic, Trigonometric, and Exponential/Logarithmic functions form what we call *Elementary functions*. Understanding this taxonomy allows us to anticipate the behavior, symmetries, and limitations of the functions we use to model the world. As we proceed deeper into mathematics, these functions will be the primary objects of calculus.

1.7 Trigonometric Functions: The Geometry of the Circle

1.7.1 Radians versus Degrees

Historically, humanity divided the circle into 360 degrees (360°). This convention is a Babylonian artifact, intimately tied to their base-60 numeral system and the approximate number of days in a year. But Nature does not care about the degrees: Nature cares about intrinsic geometry.

> *Intuition. The Radian Measure*
>
> The *radian* is the angle formed when we take a segment equal to the radius and wrap it around the circle's circumference. In a unit circle (radius 1), an angle of s radians corresponds precisely to an arc length of s. This intrinsic geometric choice is what allows the derivative (a concept you will learn soon) of $\sin x$ to be beautifully exactly $\cos x$, without any messy proportionality constants bleeding out.

The conversion rule is simple: a full circle of 360° corresponds to an arc length of 2π on the unit circle, see Table 1.1.

Degree	Radian
0°	0
30°	$\pi/6$
45°	$\pi/4$
60°	$\pi/3$
90°	$\pi/2$
120°	$2\pi/3$
360°	2π

Table 1.1 – Standard angle conversions.

To convert a degree measure into radians, we multiply by the ratio of a full circle. For instance, to compute the radian equivalent of 30°:

$$30^\circ \cdot \frac{2\pi}{360^\circ} = \frac{1}{12} \cdot 2\pi = \frac{\pi}{6}. \tag{1.15}$$

1.7.2 Definitions from the Unit Circle

Definition 1.7.1 (Trigonometric Functions)**.** Consider a point (x, y) on the unit circle $x^2 + y^2 = 1$, such that the ray extending from the origin to (x, y) subtends an angle θ with the positive x-axis, see Figure 1.7. We define the six primary trigonometric functions as follows:

$$\sin\theta = \frac{y}{1} = y, \csc\theta = \frac{1}{y},$$
$$\cos\theta = \frac{x}{1} = x, \sec\theta = \frac{1}{x},$$
$$\tan\theta = \frac{y}{x}, \cot\theta = \frac{x}{y}.$$

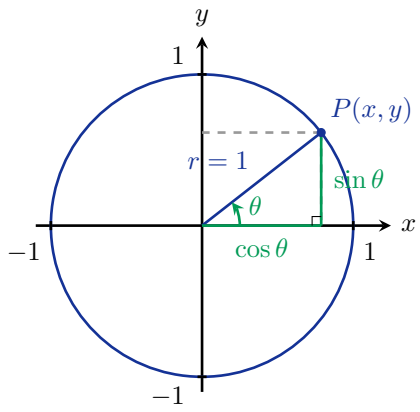


Figure 1.7 – Trigonometric functions defined on the unit circle

As an immediate corollary of these definitions, we obtain fundamental algebraic identities linking these functions.

Corollary 1.7.2 (Trigonometric Identities)**.**

$$\tan x = \frac{\sin x}{\cos x}, \quad \cot x = \frac{1}{\tan x}, \quad \csc x = \frac{1}{\sin x}, \quad \sec x = \frac{1}{\cos x}.$$

See Figure 1.8, Figure 1.9, and Figure 1.10 for the graphs of those trigonometric functions.

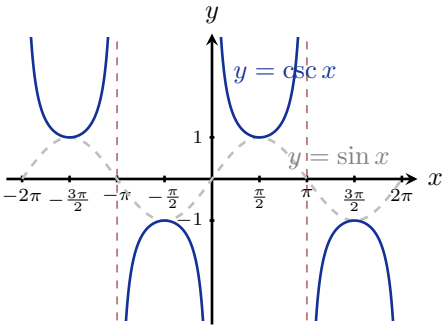


Figure 1.8 – Graphs of the Sine and Cosecant functions

1.7.3 The Complex Connection

We now introduce the imaginary unit $i = \sqrt{-1}$, which trivially satisfies $i^2 = -1$. One of the most profound realizations in modern mathematics is how the exponential function dances with the trigonometric functions when evaluated on the imaginary axis.

Story. Euler's Formula

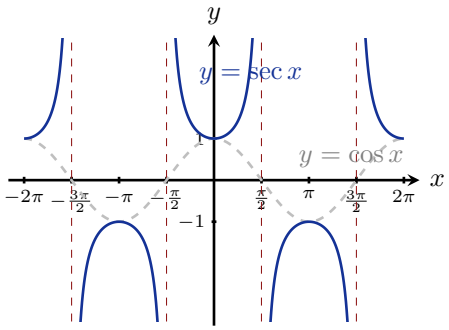


Figure 1.9 – Graphs of the Cosine and Secant functions

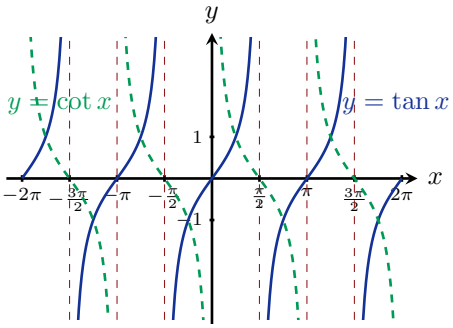


Figure 1.10 – Graphs of the Tangent and Cotangent functions

> Leonhard Euler discovered that exponential growth in the imaginary direction corresponds perfectly to rotation in the complex plane. This is crystallized in Euler's formula, which bridges the gap between the algebraic structure of exponents and the geometric structure of circles.

Let $\mathrm{e} := 2.71828182846\ldots$, which is a real irrational number usually known as the Euler's constant.[1]

Theorem 1.7.3 (Euler's Identity)**.** *For any real angle θ,*

$$\mathrm{e}^{i\theta} = \cos\theta + i\sin\theta. \tag{1.16}$$

To witness the staggering power of this identity, let us use it to effortlessly re-derive classical trigonometric identities. By passing through the complex realm, the multiplicative properties of exponents are magically transformed into the additive properties of angles. We trade geometry for algebra, and the complex plane pays the exchange rate.

Lemma 1.7.4 (The Pythagorean Identity)**.** *For any $x \in \mathbb{R}$, $\cos^2 x + \sin^2 x = 1$.*

[1] You can think of e as something that is similar to π, just not edible! The rigorous definition of e requires a little bit calculus that we will study later.

Proof. We begin with the trivial statement $1 = e^0$. We can rewrite the exponent as $ix - ix$:

$$\begin{aligned} 1 &= e^{ix-ix} = e^{ix}e^{-ix} \\ &= (\cos x + i \sin x)(\cos(-x) + i \sin(-x)). \end{aligned}$$

Recalling that cosine is an even function ($\cos(-x) = \cos x$) and sine is an odd function ($\sin(-x) = -\sin x$), we obtain:

$$\begin{aligned} 1 &= (\cos x + i \sin x)(\cos x - i \sin x) \\ &= \cos x \cos x - \cos x(i \sin x) + (i \sin x) \cos x - i^2 \sin^2 x \\ &= \cos^2 x + \sin^2 x. \end{aligned}$$

This is exactly the Pythagorean identity. □

Lemma 1.7.5 (Double Angle Formulas)**.** *For any $\theta \in \mathbb{R}$,*

$$\begin{aligned} \cos(2\theta) &= \cos^2 \theta - \sin^2 \theta, \\ \sin(2\theta) &= 2 \sin \theta \cos \theta. \end{aligned}$$

Proof. By Euler's formula, we evaluate $e^{i(2\theta)}$ directly:

$$e^{i(2\theta)} = \cos(2\theta) + i \sin(2\theta).$$

Alternatively, by the laws of exponents:

$$e^{i(2\theta)} = e^{i\theta + i\theta} = e^{i\theta} \cdot e^{i\theta}.$$

Expanding this product via Euler's formula:

$$\begin{aligned} e^{i\theta} \cdot e^{i\theta} &= (\cos \theta + i \sin \theta)(\cos \theta + i \sin \theta) \\ &= \cos^2 \theta + i \cos \theta \sin \theta + i \sin \theta \cos \theta + i^2 \sin^2 \theta \\ &= (\cos^2 \theta - \sin^2 \theta) + i(2 \sin \theta \cos \theta). \end{aligned}$$

Equating the real and imaginary parts of (1.7.3) and (1.7.3) yields the result. □

Exercise 1.7.6. Use Euler's identity to deduce the addition formulas for $\sin(\theta + 2\pi)$ and $\cos(x + y)$.

1.7.4 Inverse Trigonometric Functions and the Art of Restriction

Trigonometric functions are inherently periodic, which unfortunately implies they are wildly non-injective (many-to-one). If we wish to construct an *inverse* mapping, we encounter an immediate, glaring obstruction.

> ⚠ *Caveat. The Problem of Inverses*
>
> You cannot simply "invert" $y = \sin x$ over all of $\mathbb{R}$, because asking "what angle yields a sine of 0?" has infinitely many valid answers $(\dots, -\pi, 0, \pi, 2\pi, \dots)$. To forge a

rigorous, single-valued function, we must restrict the domain to a region where the original function is strictly one-to-one.

By convention, we lock down inverse trigonometric functions by isolating specific principal branches.

Definition 1.7.7 (Inverse Trigonometric Functions)**.**

1. **Arcsine:** We restrict $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1, 1]$. In this domain, the function is strictly one-to-one. We define the inverse sine as:
$$\text{dom}(\sin^{-1}) = [-1, 1], \quad \text{ran}(\sin^{-1}) = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]. \tag{1.17}$$
2. **Arccosine:** We restrict $\cos : [0, \pi] \to [-1, 1]$.
$$\text{dom}(\cos^{-1}) = [-1, 1], \quad \text{ran}(\cos^{-1}) = [0, \pi]. \tag{1.18}$$
3. **Arctangent:** We restrict $\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$.
$$\text{dom}(\tan^{-1}) = \mathbb{R}, \quad \text{ran}(\tan^{-1}) = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right). \tag{1.19}$$
4. **Arccotangent:** We restrict $\cot : (0, \pi) \to \mathbb{R}$.
$$\text{dom}(\cot^{-1}) = \mathbb{R}, \quad \text{ran}(\cot^{-1}) = (0, \pi). \tag{1.20}$$
5. **Arcsecant:** We restrict $\sec : \left[0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right] \to (-\infty, -1] \cup [1, \infty)$.
$$\text{dom}(\sec^{-1}) = (-\infty, -1] \cup [1, \infty), \quad \text{ran}(\sec^{-1}) = \left[0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right]. \tag{1.21}$$
6. **Arccosecant:** We restrict $\csc : \left[-\frac{\pi}{2}, 0\right) \cup \left(0, \frac{\pi}{2}\right] \to (-\infty, -1] \cup [1, \infty)$.
$$\text{dom}(\csc^{-1}) = (-\infty, -1] \cup [1, \infty), \quad \text{ran}(\csc^{-1}) = \left[-\frac{\pi}{2}, 0\right) \cup \left(0, \frac{\pi}{2}\right]. \tag{1.22}$$

Remark 1.7.8. Note that we did not provide exact algebraic formulas for these inverse functions. In the calculus, they are determined completely and uniquely by the very definition of an inverse:

$$f(g(x)) = x \quad \text{and} \quad g(f(x)) = x.$$

Moral. Beyond the Real Line

There are several equally valid ways to choose the domains for these inverses; we merely follow the most common mathematical convention. However, this seemingly arbitrary choice of domain restriction is a whisper of a much deeper, more elegant truth. In complex analysis, these alternative choices of domains correspond to different *branches* of the function. The jarring discontinuities and branch cuts we encounter are

simply the low-dimensional shadows of a beautifully smooth, multi-layered manifold—a *Riemann Surface*—projected down onto the flat plane. This topic is beyond Calculus and belongs an interesting subject named *Complex Analysis*.

Let us wrap up with some calculations. By recalling the standard unit circle, we can easily invert known values:

$$\sin^{-1}\left(\frac{\sqrt{3}}{2}\right) = \frac{\pi}{3}$$
$$\sin^{-1}\left(-\frac{\sqrt{3}}{2}\right) = -\frac{\pi}{3}$$

Intuition. Geometric Guiding

When manipulating inverse trigonometric functions, the most robust conceptual tool is not algebraic memorization, but geometric construction. An expression like $\theta = \sin^{-1}(k)$ is simply a set of instructions to draw a right-angled triangle.

Example 1.7.9. Let $\theta = \sin^{-1}\left(\frac{2}{5}\right)$. Find the exact values of $\cos\theta$ and $\tan\theta$.

Solution. By definition, the relation $\theta = \sin^{-1}(2/5)$ implies that $\sin\theta = 2/5$, with the implicit restriction that $\theta \in [-\pi/2, \pi/2]$. We can interpret θ as an angle in a right triangle where the opposite side has length 2 and the hypotenuse has length 5.

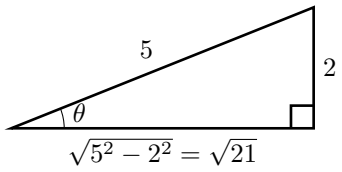


Figure 1.11 – The right triangle with $\theta = \sin^{-1}\left(\frac{2}{5}\right)$.

Using the Pythagorean theorem, the adjacent side is determined to be $\sqrt{25-4} = \sqrt{21}$. Thus, we can directly read off the remaining trigonometric ratios:

$$\cos\theta = \frac{\sqrt{21}}{5}, \qquad \tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{2/\sqrt{21}}{\sqrt{21}/5} = \frac{2}{\sqrt{21}}. \tag{1.23}$$

Why don't we take the minus sign for the square root (i.e., $\cos\theta = -\sqrt{21}/5$)?
Because $\theta = \sin^{-1}(2/5)$ is constrained to the interval $[-\pi/2, \pi/2]$ (spanning the first and fourth quadrants), a region where the cosine function is strictly non-negative! □

Example 1.7.10. Simplify the expression $\cot\left(\cos^{-1}\left(\frac{x}{4}\right)\right)$ for some $0 < x < 4$.

Solution. Let $\theta = \cos^{-1}(x/4)$, which asserts that $\cos\theta = x/4$. We construct a right triangle with an adjacent side of length x and a hypotenuse of length 4.

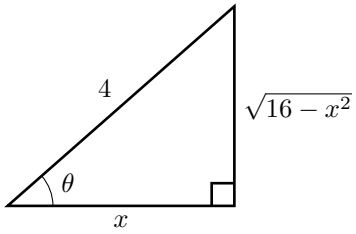


Figure 1.12 – The right triangle with $\theta = \cos^{-1}(x/4)$.

The opposite side is determined to be $\sqrt{4^2 - x^2} = \sqrt{16 - x^2}$. The cotangent is defined as the ratio of the adjacent side to the opposite side, yielding:

$$\cot\left(\cos^{-1}\left(\frac{x}{4}\right)\right) = \cot\theta = \frac{x}{\sqrt{16-x^2}}. \tag{1.24}$$

□

Exercise 1.7.11 (Homework)**.** Test your geometric intuition on the following evaluations:

(a) $\tan^{-1}\left(-\frac{1}{\sqrt{3}}\right) = ?$

(b) $\sin(\tan^{-1}(x)) = ?$

1.8 Exponential and Logarithmic Functions

Story. The Mathematics of Compounding

If trigonometric functions represent cyclic rhythms, exponential functions represent cascading growth. They are nature's favorite mechanism, arising inevitably whenever a quantity's rate of change is proportional to the quantity itself—be it populations of bacteria dividing in a petri dish, the radioactive decay of isotopes, or the compounding of wealth in finance.

Definition 1.8.1. Let $b > 0$ with $b \neq 1$. The *exponential function* with base b is defined as $f(x) \coloneqq b^x$.

Remark 1.8.2 (Properties of Exponentials). For any bases $a, b > 0$ and exponents $x, y \in \mathbb{R}$, the exponential function strictly adheres to the following operational rules:

1. $a^0 = 1$
2. $a^x a^y = a^{x+y}$
3. $(a^x)^y = a^{xy}$
4. $(ab)^x = a^x b^x$
5. $\frac{a^x}{a^y} = a^{x-y}$

6. $\left(\frac{a}{b}\right)^x = \frac{a^x}{b^x}$
7. $a^{-x} = \frac{1}{a^x}$

Example 1.8.3. Consider the dyadic exponential $f(x) = 2^x$. We can compute various discrete points to observe its rapid escalation:

- $f(0) = 2^0 = 1$
- $f(2) = 2^2 = 2 \cdot 2 = 4$
- $f\left(\frac{2}{3}\right) = 2^{2/3} = (2^2)^{1/3} = \sqrt[3]{4}$
- $f(\pi) = 2^\pi$ (an transcendental output!)

Before characterizing the global geometric behavior of b^x, let us formalize the analytic notion of a function "going up" or "going down".

Definition 1.8.4 (Increasing, Decreasing, and Monotonicity)**.** Suppose we have a function $f\colon \mathbb{R} \to \mathbb{R}$ with domain $\mathrm{dom}(f) \subset \mathbb{R}$ and range $\mathrm{ran}(f) \subset \mathbb{R}$.

(i) We say f is *increasing* if for all $x_1, x_2 \in \mathrm{dom}(f)$,

$$x_1 < x_2 \implies f(x_1) \leq f(x_2). \tag{1.25}$$

We say f is *strictly increasing* if $\leq$ is replaced by $<$.

(ii) We say f is *decreasing* if for all $x_1, x_2 \in \mathrm{dom}(f)$,

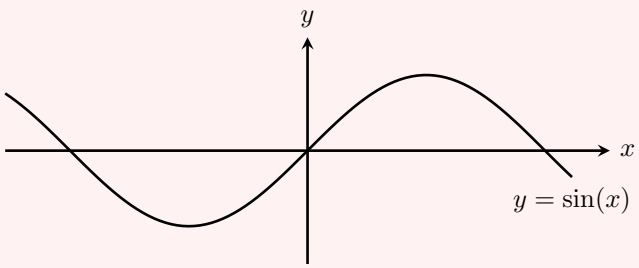
$$x_1 < x_2 \implies f(x_1) \geq f(x_2). \tag{1.26}$$

We say f is *strictly decreasing* if $\geq$ is replaced by $>$.

We say that f is *monotone* if it is either strictly increasing or strictly decreasing on its entire domain.

⚠ *Caveat. Not Everything is Monotone!*

The function $y = \sin(x)$ is emphatically *not* monotone. It constantly reverses direction, fluctuating between increasing and decreasing phases to create the waves we analyzed in the previous section.



Armed with this definition, we can classify the macro-behavior of $f(x) = b^x$:

1. $\mathrm{dom}(f) = \mathbb{R}$, and $\mathrm{ran}(f) = (0, \infty)$. (An exponential output is always strictly positive).
2. $f(0) = b^0 = 1$, which implies the y-intercept is universally at the coordinate $(0, 1)$.
3. Monotonicity is governed entirely by the base b, see Figure 1.13:
   - If $b > 1$, f is *increasing* (modeling exponential growth).
   - If $0 < b < 1$, f is *decreasing* (modeling exponential decay).

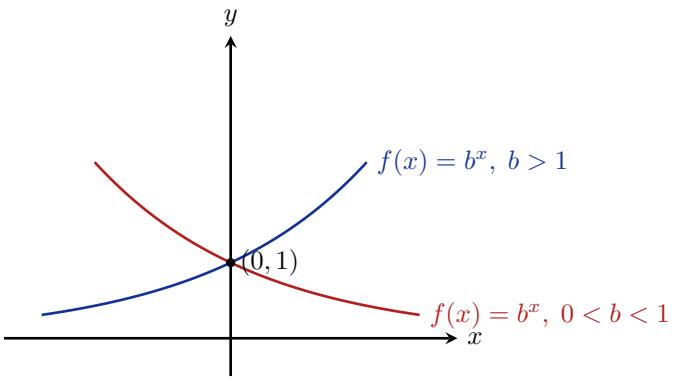


Figure 1.13 – The behaviors of exponential functions.

> *Moral. The Natural Base*
>
> Among the infinitely many bases we could choose, one reigns supreme in the calculus: Euler's number $\mathrm{e} \approx 2.71828\ldots$. The special case $b = \mathrm{e}$ gives us the *natural exponential function* $f(x) = \mathrm{e}^x$, uniquely characterized by the profound property that its rate of growth is precisely equal to its value at any given moment.

If one wishes to find the time it takes for a biological population to reach a certain size, one must invert the exponential function. This logical necessity gives birth to the logarithm.

Definition 1.8.5. For $0 < b \neq 1$, the *logarithmic function* base b, denoted by $\log_b x$, is the inverse of the exponential function b^x. It satisfies the identities:

$$\log_b(b^x) = x \quad \text{and} \quad b^{\log_b x} = x. \tag{1.27}$$

If $b = \mathrm{e}$, we typically write $f(x) = \log x$ (or sometimes $\log x$ in higher mathematics) to denote the *natural logarithmic function*.

Remark 1.8.6. Because the logarithm is the inverse of the exponential, their domains and ranges perfectly swap. For $f(x) = \log_b x$:

$$\mathrm{dom}(f) = (0, \infty), \qquad \mathrm{ran}(f) = \mathbb{R}. \tag{1.28}$$

Chapter 2

Limits

In every other Science Men prove their Conclusions by their Principles, and not their Principles by the Conclusions.

George Berkeley, *The Analyst*
(a criticism of Newton and Leibniz calculus)

Story. The Birth of the Dynamic World

Historically, calculus was born out of an urgent desire to understand two fundamentally different yet deeply intertwined phenomena: the physics of motion and the geometry of curves. Isaac Newton looked at the sky and wanted to know exactly how planets moved at any given instant. Gottfried Wilhelm Leibniz looked at geometric curves and wanted to find their exact tangents and the areas bounded beneath them. Both visionaries realized that the mathematics of their time, which dealt comfortably only with static, macroscopic quantities, was woefully inadequate. They needed a new mathematical technology to bridge the gap between the static and the dynamic. That bridge is what we now call the *limit*.

2.1 The Big Picture: Why Do We Need Limits?

Let us begin with a physical motivation. Imagine you are driving from your home to the campus. The total distance is 10 miles, and you arrive in exactly 20 minutes (which is $\frac{1}{3}$ of an hour). We can compute your average velocity over the trip:

$$v_{\text{av}} = \frac{\text{Distance}}{\text{Time}} = \frac{10}{\frac{1}{3}} = 30 \text{ miles/hr}. \tag{2.1}$$

This is a satisfying macroscopic piece of data. But suppose a police officer is sitting on the side of the highway with a radar gun. The police officer does not care about your average velocity; they want to know your *instantaneous velocity* at a single moment in time, say $t = t_p$.

If your position at time t is given by a function $s(t)$, the average velocity between time t_p and some later time t_1 is:

$$v_{\text{av}} = \frac{s(t_1) - s(t_p)}{t_1 - t_p}. \tag{2.2}$$

But if we want the velocity *exactly* at t_p, we run into a profound algebraic wall. We cannot simply plug in $t_1 = t_p$, as this yields the dreaded indeterminate form $\frac{0}{0}$.

> *Intuition. The Limit as an Approximation*
>
> The core idea of differential calculus is that the instantaneous velocity at t_p is approximately equal to the average velocity when t is *very close* to t_p. We write:
>
> $$v_{\text{int}} \approx \frac{s(t) - s(t_p)}{t - t_p}, \quad \text{where } t \text{ is close to } t_p. \tag{2.3}$$

To make this rigorous, we introduce the mathematical notation of a *limit*:

$$v_{\text{int}} = \lim_{t \to t_p} \frac{s(t) - s(t_p)}{t - t_p}. \tag{2.4}$$

Let us step away from physics and articulate what this means purely for a function.

Definition 2.1.1 (Limit of a function, intuitively)**.** Given a function $f : \mathbb{R} \to \mathbb{R}$, and $a, L \in \mathbb{R}$, we say that L is the limit of f at a, and write

$$\lim_{x \to a} f(x) = L$$

if we can make the values of $f(x)$ be arbitrarily close to L by taking x to be sufficiently close (but *not equal*) to a.

Notation 2.1.2. Instead of $\lim_{x \to a} f(x) = L$, we also frequently write $f(x) \to L$ as $x \to a$, or $f(x) \xrightarrow{x \to a} L$.

Example 2.1.3. Consider the piecewise function

$$f(x) = \begin{cases} x & \text{if } x \neq 1, \\ 2 & \text{if } x = 1. \end{cases}$$

What is $\lim_{x \to 1} f(x)$? Let us look at the graph of this function in Figure 2.1.

As x gets arbitrarily close to 1 (from either the left or the right), the value of $f(x)$ approaches 1. The fact that $f(1)$ is actually 2 is utterly irrelevant to the limit. Thus,

$$\lim_{x \to 1} f(x) = 1. \tag{2.5}$$

> ⚠ *Caveat. The Point Does Not Matter*
>
> The limit of a function as $x \to a$ depends *only* on the behavior of $f(x)$ in a punctured

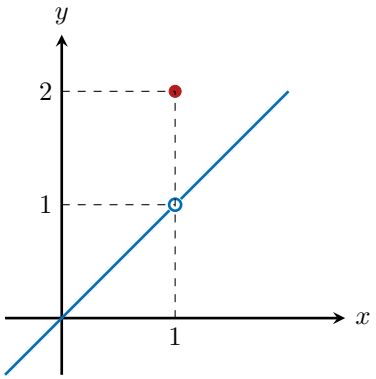


Figure 2.1 – Graph of the piecewise function $f(x)$.

neighborhood around a. The actual value $f(a)$ could be 2, 100, or completely undefined; the limit remains unchanged.

2.2 Two Fundamental Problems Using Limits

Calculus is essentially the mathematics of limits applied to two fundamental geometric problems: finding tangents (derivatives) and finding areas (integrals).

2.2.1 The Tangent Line Problem

Given a function f and a point P on its graph, find an equation of the *tangent line* to the graph at point P. Intuitively, the tangent line "touches the graph at P" without crossing it abruptly.

Key Idea: To find the equation of a line, we need a point and a slope. We already have the point P. Our task reduces to finding the slope of the tangent line, which we shall call $m_{\tan}$.

We approximate the slope of the tangent line by the slope of a *secant line*, see Figure 2.2. Choose a point $P = (c, f(c))$ and a nearby point $Q = (c + \Delta x, f(c + \Delta x))$.

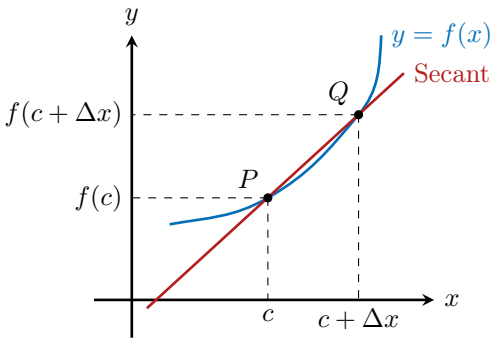


Figure 2.2 – A secant line PQ of $y = f(x)$.

The slope of the secant line passing through P and Q is standard rise over run:

$$m_{\text{sec}} = \frac{\text{rise}}{\text{run}} = \frac{f(c+\Delta x) - f(c)}{(c+\Delta x) - c} = \frac{f(c+\Delta x) - f(c)}{\Delta x}. \tag{2.6}$$

Now, invoke the magic of limits. If we let $\Delta x \to 0$, the point Q slides along the curve towards P, and the secant line gradually converges to the tangent line:

$$m_{\tan} \approx m_{\text{sec}} \quad \text{if } \Delta x \to 0. \tag{2.7}$$

Thus, we formally define:

$$m_{\tan} = \lim_{\Delta x \to 0} m_{\text{sec}} = \lim_{\Delta x \to 0} \frac{f(c+\Delta x) - f(c)}{\Delta x}. \tag{2.8}$$

This limits-based object is so foundational that we give it a special name.

Definition 2.2.1 (Difference Quotient and the Derivative)**.** Given a function f and a point a, for each $h \in \mathbb{R}$, the *difference quotient* of f at a is defined as

$$\text{DQ}(f, a)(h) \coloneqq \frac{f(a+h) - f(a)}{h}. \tag{2.9}$$

The limit of this quotient as $h \to 0$ is the *derivative* of f at a.

2.2.2 The Area Problem

Given a function $f(x) \geq 0$ for $x \in [a, b]$, find the area A bounded by the curve f, the x-axis, and the vertical lines $x = a$ and $x = b$, see Figure 2.3.

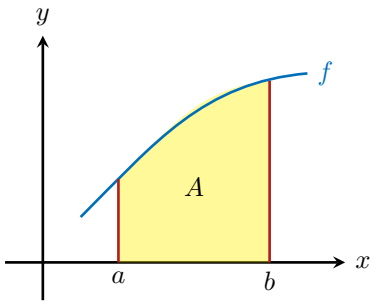


Figure 2.3 – Graph of the function $f(x)$.

We approach this systematically by generalizing from simple shapes to complex ones.

1. *Constant Function.* If $f(x) = c > 0$, the region is a simple rectangle. The area is base times height:

$$A = c(b - a). \tag{2.10}$$

2. *Piecewise Constant (Step) Function.* Suppose f is a step function, taking different constant values on different intervals:

$$f(x) = \begin{cases} c_1, & x \in [a, t) \\ c_2, & x \in [t, b] \end{cases}$$

The total area is simply the sum of the areas of the individual rectangles:

$$A = c_1(t-a) + c_2(b-t). \tag{2.11}$$

3. *General Curve Approximation.* For a general continuous curve, we approximate the area by slicing the interval $[a,b]$ into many small rectangles. Let us define intermediate points $a = t_0 < t_1 < \cdots < t_N = b$. We approximate the height of the curve over the interval $[t_i, t_{i+1}]$ by $f(t_i)$.

$$A \approx f(t_0)(t_1 - t_0) + f(t_1)(t_2 - t_1) + \cdots + f(t_{N-1})(t_N - t_{N-1}). \tag{2.12}$$

If we assume the points are *equidistant*, then the width of each sub-interval is $\Delta x = t_{i+1} - t_i$. The approximation neatly factors out:

$$A \approx \Delta x\Big(f(t_0) + f(t_1) + \cdots + f(t_{N-1})\Big). \tag{2.13}$$

> *Moral. From Discrete to Continuous*
>
> The grand synthesis occurs when we let the number of rectangles go to infinity, which forces the width $\Delta x \to 0$. We recover the exact area through a limit:
>
> $$A = \lim_{\Delta x \to 0} \Delta x\Big(f(t_0) + f(t_1) + \cdots + f(t_{N-1})\Big). \tag{2.14}$$
>
> This profound limit of a sum is what we call the *integral*.

2.3 Finding Limits Intuitively: Graphs and Numerics

To properly evaluate limits, we must be careful with domains. Before we take $x \to a$, we must ensure the function actually makes sense near a.

Example 2.3.1. Consider the function

$$f(x) = \frac{x}{\sqrt{x+1} - 1}.$$

Let us determine its domain, $\mathrm{dom}(f)$.

$$\begin{aligned} \mathrm{dom}(f) &= \{x \in \mathbb{R} : f(x) \text{ is well-defined}\} \\ &= \{x \in \mathbb{R} : f(x) \in \mathbb{R}\}. \end{aligned}$$

For $f(x)$ to be a real number, two conditions must be met:

(i) The denominator cannot be zero: $\sqrt{x+1} - 1 \neq 0$.

(ii) The expression inside the square root must be non-negative (to remain in $\mathbb{R}$): $x + 1 \geq 0 \implies x \geq -1$.

Let us analyze the first condition:

$$\begin{aligned}\sqrt{x+1}-1=0 &\iff \sqrt{x+1}=1\\ &\iff (\sqrt{x+1})^2=1^2\\ &\iff x+1=1\\ &\iff x=0.\end{aligned}$$

Thus, we require $x \neq 0$. Combining this with $x \geq -1$, we deduce the domain is:

$$\mathrm{dom}(f)=\{x : x \geq -1, x \neq 0\}=[-1,0)\cup(0,\infty).$$

Now, let us find the limit as $x \to 0$:

$$\lim_{x\to 0} f(x)=\lim_{x\to 0}\frac{x}{\sqrt{x+1}-1}.$$

There are two immediate, intuitive heuristics to approximate $\lim_{x\to 0} f(x)$: graphically and numerically.

Method 1 — The Graph. If we were to plot $f(x)$, we would trace out a gentle curve starting from $x=-1$ that passes smoothly toward the coordinate $(0,2)$, albeit with a microscopic "hole" punctured exactly at $x=0$, see Figure 2.4. We conclude $\lim_{x\to 0} f(x)=2$.

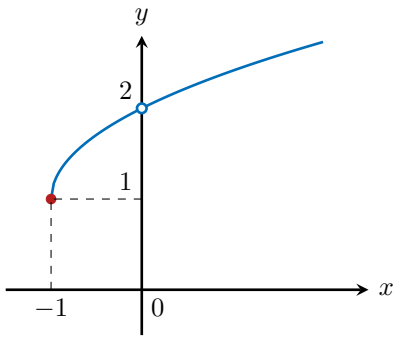


Figure 2.4 – Graph of the function $f(x)$.

Method 2 — Numerics. Alternatively, we can use a calculator to systematically probe the values of $f(x)$ as x approaches 0 from both the left (negative values) and the right (positive values), see Table 2.1.

x	-0.01	-0.0001	0	0.0001	0.01
$f(x)$	1.99499	1.99995	undefined	2.00005	2.00499

Table 2.1 – Values of $f(x)$ as x approaches 0.

The table tells a compelling story. As the x-values tighten their grip around 0, the outputs $f(x)$ march toward the value 2. We formalize this observation by writing:

$$\lim_{x\to 0} f(x)=2.$$

Moral.

The limit is oblivious to the destination itself; it only cares about the journey. The literal value of $f(a)$ has absolutely no bearing on $\lim_{x\to a} f(x)$.

2.4 Fractured Journeys: One-Sided Limits

Not all functions are so well behaved. Let us investigate a simpler, yet sharper function: the absolute value, $f(x) = |x|$. The domain of $|x|$ is the entire real line, $\mathbb{R}$. We may define it piecewise as:

$$f(x) = \begin{cases} x, & x \geq 0 \\ -x, & x < 0. \end{cases}$$

As $x \to 0$, the values of $|x|$ converge to 0 from both sides, so $\lim_{x\to 0} |x| = 0$, see Figure 2.5.

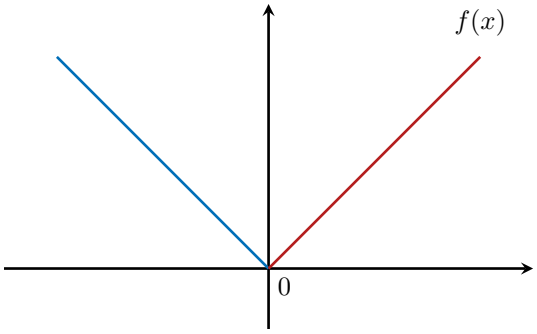


Figure 2.5 – Graph of the absolute value function $f(x) = |x|$.

But what happens if we construct a ratio that pits x against its absolute value? Consider the *signum function* (represented here purely by its ratio):

$$f(x) = \frac{|x|}{x}.$$

The domain is stripped of the origin: $\text{dom}(f) = \mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty)$. Let us rewrite $f(x)$ without the absolute value bars to expose its piecewise nature:

$$f(x) = \begin{cases} \frac{x}{x} = 1, & \text{if } x > 0 \\ \frac{-x}{x} = -1, & \text{if } x < 0. \end{cases}$$

If x is closed to 0 and positive, $f(x)$ rigidly equals 1. Conversely, if x is closed to 0 and negative, $f(x)$ rigidly equals -1, see Figure 2.6.

We are faced with a bifurcation: there are two entirely different choices for our limit L, depending solely on our direction of approach! Because the function cannot settle on a single value, we are forced to conclude that the limit *does not exist* (DNE).

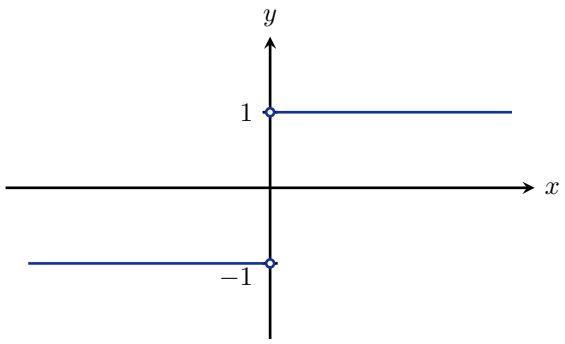


Figure 2.6 – Graph of the signum function $f(x) = \frac{|x|}{x}$

> ⚠ *Caveat. The Ambiguity of Direction*
>
> A limit only exists if the function agrees on a single destination from *all* possible directions of approach.

However, not all things are lost! We can still salvage meaningful information about the behavior of f near 0 by restricting our perspective to one direction at a time.

2.4.1 Formalizing the Directional Approach

We introduce a refinement of our limit concept to handle such fractured behavior.

Definition 2.4.1 (One-Sided Limits)**.** Given a function $f : \mathbb{R} \to \mathbb{R}$, and constants $a, L \in \mathbb{R}$:

1. *Right-sided limit:* If $f(x)$ is arbitrarily close to L for all x sufficiently close to a with $x > a$, we write:
$$\lim_{x \to a^+} f(x) = L \quad \text{or} \quad f(x) \xrightarrow{x \to a^+} L.$$
2. *Left-sided limit:* If $f(x)$ is arbitrarily close to L for all x sufficiently close to a with $x < a$, we write:
$$\lim_{x \to a^-} f(x) = L \quad \text{or} \quad f(x) \xrightarrow{x \to a^-} L.$$

Returning to our pathological example $f(x) = |x|/x$, we can now mathematically describe its behavior:

$$\lim_{x \to 0^+} \frac{|x|}{x} = 1 \quad \text{and} \quad \lim_{x \to 0^-} \frac{|x|}{x} = -1. \tag{2.15}$$

This leads us to a beautiful theorem linking local (one-sided) and global (two-sided) behavior.

Theorem 2.4.2 (Connection between one-sided and (two-sided) limits)**.** *Let f be a function defined on an open interval around a (except possibly at a). Then the limit exists and equals L:*

$$\lim_{x \to a} f(x) = L \iff \begin{cases} \lim_{x \to a^+} f(x) = L \\ \lim_{x \to a^-} f(x) = L. \end{cases} \tag{2.16}$$

Because $1 \neq -1$, the global limit $\lim_{x\to 0} \frac{|x|}{x}$ fails the test of Theorem 2.4.2.

2.4.2 Divergence: To Infinity and Beyond

Thus far, our limits have been well-behaved, finite numbers. But what if the values of the function grow without bound? Consider the inverse-square function,

$$f(x) = \frac{1}{x^2}. \tag{2.17}$$

Graphically, this function features a vertical asymptote at $x = 0$, with the curve shooting upward, see Figure 2.7.

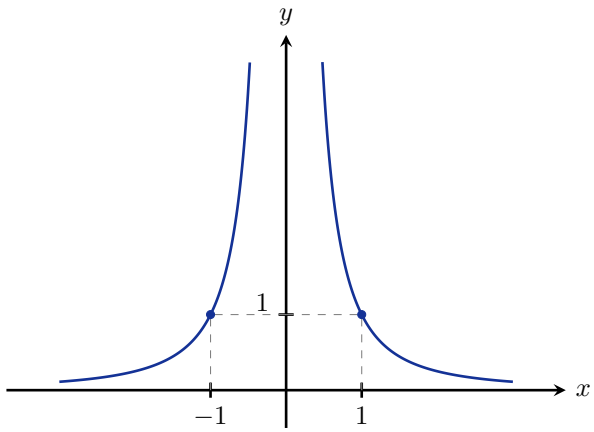


Figure 2.7 – Graph of the function $f(x) = \frac{1}{x^2}$

Let us confirm this numerically in Table 2.2:

x	-10^{-1}	-10^{-10}	0	10^{-10}	10^{-1}
$f(x)$	10^2	10^{20}	undefined	10^{20}	10^2

Table 2.2 – Values of $f(x)$ as x approaches 0.

As x approaches 0, the values of $f(x)$ explode. We encapsulate this behavior with the symbol ∞, writing:

$$\lim_{x\to 0} \frac{1}{x^2} = \infty.$$

Definition 2.4.3 (Infinite Limit)**.** Given a real-valued function f, and $a \in \mathbb{R}$, we say that ∞ is the limit of f at a, and write

$$\lim_{x\to a} f(x) = \infty \tag{2.18}$$

if we can make the values of $f(x)$ be as large as we want by taking x to be sufficiently close (but not equal) to a.

Intuition.

Saying the limit "is infinity" is a slight abuse of terminology. Infinity is not a real number; it is a direction. Writing $= \infty$ is simply mathematically precise shorthand for "diverges by growing without bound."

We must remain vigilant. Consider the simpler hyperbola $f(x) = 1/x$. Evaluating the one-sided limits by Figure 2.8 yields:

$$\lim_{x \to 0^+} \frac{1}{x} = \infty \quad \text{and} \quad \lim_{x \to 0^-} \frac{1}{x} = -\infty.$$

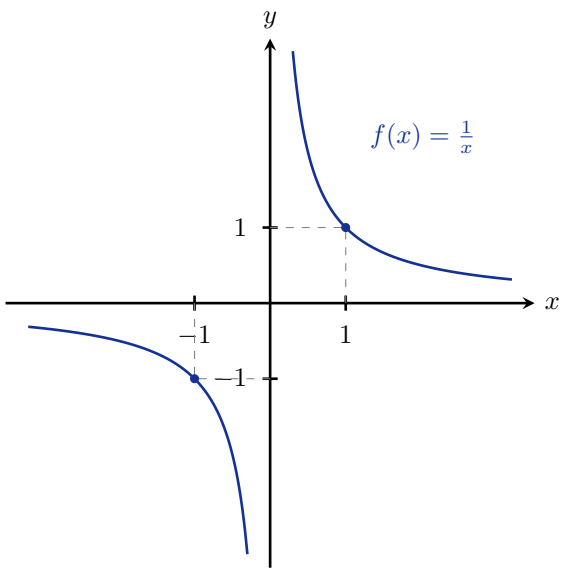


Figure 2.8 – Graph of the function $f(x) = \frac{1}{x}$

Since the two one-sided limits wildly disagree ($\infty \neq -\infty$), we conclude that the two-sided limit $\lim_{x \to 0} \frac{1}{x}$ does not exist.

2.4.3 The Topologist's Sine Curve: A Frantic Oscillation

We conclude this section with a spectacular failure of limits. Some functions do not elegantly approach a value, nor do they boldly explode to infinity. Instead, they vibrate with infinite frenzy.

Consider $f(x) = \sin\left(\frac{1}{x}\right)$, affectionately known as the *Topologist's Sine Curve*, see Figure 2.9. Its domain is $\mathrm{dom}(f) = \mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty)$.

What is $\lim_{x \to 0} f(x)$? Let us construct a sequence of x-values inching closer to zero from the right in Table 2.3.

As $x \to 0^+$, the frequency of the wave becomes infinite. No matter how closely we zoom in toward the origin, the function violently jumps between 1 and -1. It cannot settle on a single finite value, nor does it escape to infinity.

x	$\frac{2}{\pi}$	$\frac{2}{3\pi}$	$\frac{2}{5\pi}$	$\frac{2}{7\pi}$	$\cdots$
$f(x)$	1	-1	1	-1	$\cdots$

Table 2.3 – Values of $f(x)$ as x approaches 0 from the right.

Therefore, $\lim_{x\to 0^+} f(x)$ simply does not exist. Consequently, $\lim_{x\to 0} f(x)$ also perishes.

Story. The Beauty of Pathologies

In the 19th century, functions like the Topologist's Sine Curve were considered "monsters." They broke the intuitive, geometric rules that classical mathematicians implicitly trusted. But these monsters were necessary; they forced the creation of rigorous modern analysis. The failure of our geometric intuition is exactly what drives the need for precise mathematical language. We must build our castle of calculus not on the shifting sands of rough sketches, but on the bedrock of formal logic!

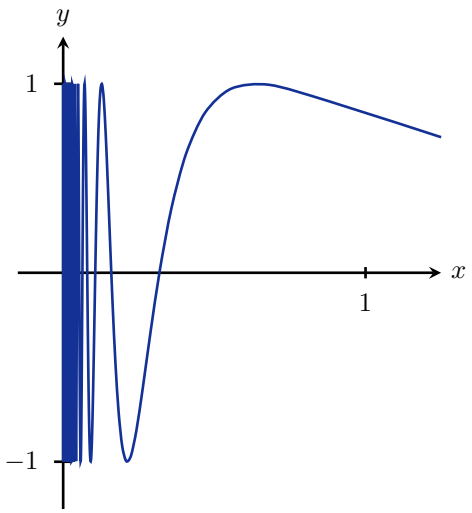


Figure 2.9 – The Topologist's Sine Curve $f(x) = \sin(1/x)$ on $(0, \infty)$

2.5 Formalizing the Limit, Rigorously

Story. The Ghosts of Departed Quantities

When Isaac Newton and Gottfried Wilhelm Leibniz first invented the calculus in the late 17th century, they relied heavily on a dangerously intuitive notion of "infinitesimals"—quantities that were magically non-zero when you needed to divide

by them, but vanished into thin air when you wanted to ignore them. The philosopher Bishop George Berkeley[1] famously skewered this logical sleight of hand in 1734, deriding these shifting variables as the "ghosts of departed quantities."
For over a century, calculus worked magnificently in physics, yet rested on shifting mathematical sands. It wasn't until the 19th century that giants like Augustin-Louis Cauchy and Karl Weierstrass stripped away the ghosts and replaced them with cold, hard inequalities. This is the story of how they did it: the δ-ε definition of a limit. The limit allows us to ask a profound, nearly philosophical question: *Where is this function going*, even if it never physically arrives?

Before we write down the formalism, let us ground ourselves in the intuition. Consider a function $f : \mathbb{R} \to \mathbb{R}$. Informally, we write

$$\lim_{x \to a} f(x) = L$$

to mean that the value of $f(x)$ becomes arbitrarily close to L as long as x is sufficiently close to a (but crucially, $x \neq a$).

Intuition. The Adversarial Game

Think of the limit as a mathematical game played between two adversaries: the *Skeptic* and the *Mathematician*.

- The Skeptic challenges the Mathematician with an error tolerance, representing a "small positive number." We traditionally call this ε (epsilon)[2]. The Skeptic says: "I demand that $f(x)$ be within a distance of ε from L. That is, $|f(x) - L| < \varepsilon$."
- The Mathematician must respond by finding a control dial, another small positive number, denoted δ (delta). The Mathematician replies: "As long as you restrict your input x to be within a distance of δ from a, your demand will be met."

Importantly, δ is allowed to depend on both a and ε, which we can write explicitly as $\delta = \delta(a, \varepsilon)$.

Let us carefully translate our English into inequalities:

1. "$f(x)$ is arbitrarily close to L within tolerance ε" translates to: $|f(x) - L| < \varepsilon$.
2. "x is sufficiently close to a (but not equal to a) within distance δ" translates to: $0 < |x - a| < \delta$.

We are now ready to formalize this concept.

Definition 2.5.1 (Rigorous Definition of a Limit)**.** We say $\lim_{x \to a} f(x) = L$ if for each $\varepsilon > 0$, there exists $\delta > 0$ depending on a and ε such that

$$|f(x) - L| < \varepsilon \quad \text{whenever} \quad 0 < |x - a| < \delta. \tag{2.19}$$

[1]The University of California, Berkeley, the city of Berkeley, California, and Berkeley College at Yale University, were all named after him.

[2]Here is a mathematical joke: A mathematician nightmare would be an ε going toward infinity when n goes toward zero.

⚠ *Caveat. The Punctured Neighborhood*

Notice the strict inequality $0 < |x - a|$. This codifies the condition that $x \neq a$. The limit of a function at a cares *only* about the behavior of the function strictly near a, completely ignoring the value of $f(a)$ itself. The function f does not even need to be defined at $x = a$.

We illustrate Definition 2.5.1 in Figure 2.10.

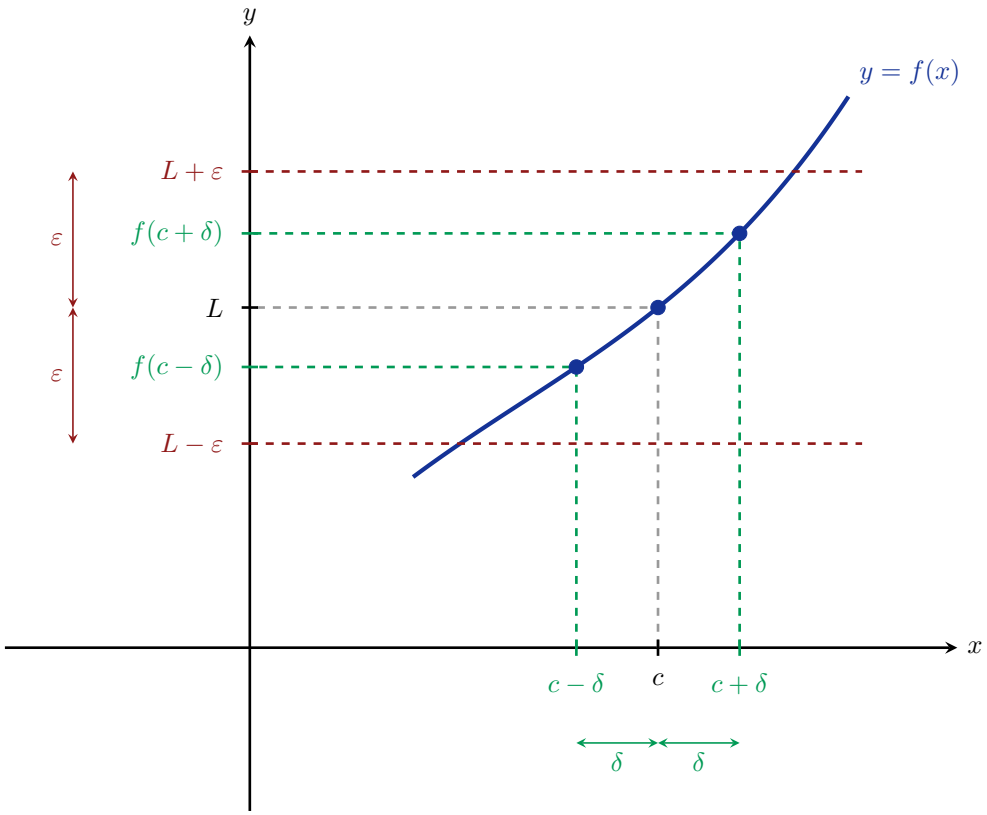


Figure 2.10 – A geometric illustration of the ε-δ limit definition

Definitions in mathematics are rarely absorbed by merely staring at them. One must get their hands dirty. Let us compute a simple limit using Definition 2.5.1.

Example 2.5.2. Use the ε-δ definition of a limit to rigorously find $\lim_{x\to 2}(3x - 2)$.

The Thinking Process Before writing a formal proof, it is highly recommended to engage in some scratch work.

1. *Educated guess:* By looking at the graph or testing numeric values close to 2, we strongly suspect that $\lim_{x\to 2}(3x - 2) = 4$. So, our target is $L = 4$ as $a = 2$.

2. *Working backwards:* Let the Skeptic choose an arbitrary $\varepsilon > 0$. We must find a $\delta > 0$ such that
$$|f(x) - L| < \varepsilon \quad \text{whenever} \quad 0 < |x - 2| < \delta.$$
Let us unpack the target inequality to forge a connection with the control inequality:
$$\begin{aligned} |(3x - 2) - 4| &< \varepsilon \\ |3x - 6| &< \varepsilon \\ 3\,|x - 2| &< \varepsilon \\ |x - 2| &< \frac{\varepsilon}{3}. \end{aligned}$$
The systematic calculation has revealed the path forward: if we simply set our dial δ to be $\varepsilon/3$, the chain of reasoning will hold in reverse. Let us crystallize this into a formal proof.

Rigorous Solution

Proof. Fix an arbitrary $\varepsilon > 0$. We choose $\delta := \frac{\varepsilon}{3}$. Because $\varepsilon > 0$, it follows that $\delta > 0$.
Suppose x is a real number satisfying $0 < |x - 2| < \delta$. Then, we have:
$$\begin{aligned} |(3x - 2) - 4| &= |3x - 6| \\ &= 3\,|x - 2| \\ &< 3\delta \\ &= 3\left(\frac{\varepsilon}{3}\right) \\ &= \varepsilon. \end{aligned}$$
Thus, by Definition 2.5.1, we conclude that $\lim_{x\to 2}(3x - 2) = 4$. □

> *Moral.*
>
> A formal limit proof is often an exercise in reverse-engineering. You start with the inequality you *want* to prove (the target inequality $|f(x) - L| < \varepsilon$) and algebraically massage it until you isolate the term you *can control* ($|x - a|$).

The linear case was rather tame because the factor relating $|f(x) - L|$ and $|x - a|$ was a constant (in our previous case, 3). In non-linear functions, this factor will depend on x, which introduces a delightful complication.

Example 2.5.3. Prove that $\lim_{x\to 1} x^2 = 1$ using the ε-δ definition.

The Thinking Process Once again, we proceed with our scratch work.

1. *Guess:* We want to show $\lim_{x\to 1} x^2 = 1$. Here $a = 1$ and $L = 1$.
2. *Working backwards:* Fix $\varepsilon > 0$. We want to find $\delta(1, \varepsilon) > 0$ such that
$$\left|x^2 - 1\right| < \varepsilon \quad \text{whenever} \quad 0 < |x - 1| < \delta.$$

Recall the algebraic identity $c^2 - d^2 = (c-d)(c+d)$ and the absolute value property $|ab| = |a|\,|b|$. We can factor the left-hand side:

$$\left|x^2 - 1\right| = |x-1|\,|x+1| < \varepsilon.$$

We have successfully isolated our controllable term, $|x-1|$. The temptation here is to simply declare $|x-1| < \frac{\varepsilon}{|x+1|}$. However, δ must be a constant that depends *only* on ε (and a), and not on the variable x. We must therefore place an upper bound on the pesky factor $|x+1|$.

3. *The Bounding Trick:* Since we are taking the limit as $x \to 1$, we only care about values of x strictly *near* 1. We are the Mathematician; we have the power to restrict δ as much as we please to make our lives easier. Let us demand that $\delta \leq 1$.

   If we enforce $|x-1| < \delta \leq 1$, then:

$$\begin{aligned} -1 < x - 1 < 1 \\ 0 < x < 2 \\ 1 < x + 1 < 3. \end{aligned}$$

   By adding 2 to the middle inequality, we see that $2 - \delta < x + 1 < 2 + \delta$. Because $\delta \leq 1$, we safely deduce that $|x+1| < 3$.

   Returning to our original bound: if $|x+1| < 3$, then $|x-1|\,|x+1| < \delta \cdot (2+\delta) < \delta \cdot 3$. To ensure this whole expression is less than ε, we simply require $\delta \leq \frac{\varepsilon}{3}$.

   To satisfy both our arbitrary bounding restriction ($\delta \leq 1$) and our error bound ($\delta \leq \frac{\varepsilon}{3}$), we take the minimum of the two.

Rigorous Solution

Proof. Fix $\varepsilon > 0$. We choose

$$\delta \coloneqq \min\left\{\frac{\varepsilon}{3}, 1\right\}.$$

(As an aside, one could conservatively choose $\delta = \min\left\{\frac{\varepsilon}{1000}, 1\right\}$, and the logic would hold perfectly well!). Since $\varepsilon > 0$, it follows that $\delta > 0$.

Suppose $0 < |x-1| < \delta$. We bound the distance between $f(x)$ and L:

$$\left|x^2 - 1\right| = |x-1|\,|x+1|\,.$$

Because $\delta \leq 1$, we have $|x-1| < 1$, which implies $-1 < x-1 < 1$, and therefore $1 < x+1 < 3$. This gives us the strict upper bound $|x+1| < 3$.

Furthermore, by our construction of δ, we know that $\delta \leq \frac{\varepsilon}{3}$. Combining these inequalities together:

$$\begin{aligned} \left|x^2 - 1\right| &= |x-1|\,|x+1| \\ &< \delta \cdot 3 \\ &\leq \left(\frac{\varepsilon}{3}\right) \cdot 3 \\ &= \varepsilon. \end{aligned}$$

Therefore, by Definition 2.5.1, $\lim_{x\to 1} x^2 = 1$. □

Mathematics is not a spectator sport. To truly internalize the concept of limits, I strongly invite you to grapple with the following exercises.

Exercise 2.5.4. Find $\lim_{x\to 4}\sqrt{x}$ by using the ε-δ definition.
Hint: Multiply by the conjugate to factor out $|x-4|$, and deploy a bounding trick similar to the one we used for the quadratic example.

Exercise 2.5.5. State the formal ε-δ definitions for 1-sided limits, namely $\lim_{x\to a^+} f(x) = L$ and $\lim_{x\to a^-} f(x) = L$.

Exercise 2.5.6. Consider the piecewise function

$$f(x) = \begin{cases} x+1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0. \end{cases}$$

Use your ε-δ definition from the previous exercise to rigorously evaluate the 1-sided limits $\lim_{x\to 0^+} f(x)$ and $\lim_{x\to 0^-} f(x)$.

Exercise 2.5.7 (The Pathological Case)**.** Use the ε-δ definition (specifically, its logical negation) to rigorously show that

$$\lim_{x\to 0} \sin\left(\frac{1}{x}\right) \quad \text{does not exist.}$$

Intuition check: What happens to the frequency of oscillation as x approaches zero? Can the Mathematician ever win the adversarial game against a sufficiently strict Skeptic?

2.6 The Zoo of Limit Behaviors: What Could Go Wrong?

Before we formalize the concept of continuity, it is highly instructive to step back and look at the big picture. Given a function $f : \mathbb{R} \to \mathbb{R}$ and a point $a \in \mathbb{R}$, we want to understand the local behavior of f near a.

The behavior of limits at a point can be broadly categorized into a hierarchy, ranging from bad to good behaviors.

1. *At least one one-sided limit does not exist.*
   Consider the wildly misbehaved function given by:

   $$f(x) = \begin{cases} \sin\left(\frac{1}{x}\right), & x > 0 \\ 0, & x \leq 0 \end{cases}$$

   As $x \to 0^+$, the argument $1/x$ rockets to infinity. This causes the sine function to oscillate furiously between -1 and 1, never settling on a single value. Consequently, $\lim_{x\to 0^+} f(x)$ simply does not exist. The function suffers a catastrophic breakdown at the origin.

2. *Both one-sided limits exist, but they are unmatched.*
   Now consider a simple step function:

   $$f(x) = \begin{cases} 1, & x > 0 \\ -1, & x < 0 \end{cases}$$

Here, the limits from both sides exist perfectly well: $\lim_{x\to 0^+} f(x) = 1$ and $\lim_{x\to 0^-} f(x) = -1$. However, they fail to agree. The graph exhibits a sudden jump, tearing the curve apart.

3. *Both one-sided limits exist and are equal.*
   Moving up the hierarchy, we find situations where the one-sided neighbors at least agree on a destination:
   $$\lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x) = L.$$
   The limit L exists, but it might not match the actual function value $f(a)$, or $f(a)$ might even be undefined.

Exercise 2.6.1. Invent 3 functions corresponding to 3 scenarios above, then draw their graphs.

Story. The Bad Popcorn

You might assume that the "bad" example in the first category represent the worst possible behavior of a function. However, there is something even worse: the Thomae's function, also known as the popcorn function

$$f(x) = \begin{cases} \frac{1}{q} & \text{if } q \text{ is the smallest natural number satisfying } x = \frac{p}{q}, \\ 0 & \text{if } x \text{ is irrational,} \end{cases}$$

see Figure 2.11. This pathological beast is continuous at *every irrational number*, yet discontinuous at *every rational number*.
The realization that such objects could exist forced mathematicians to abandon pure geometric intuition in favor of rigorous, algebraic definitions. We need a concept stricter than our "good" limit behavior to build a robust foundation for calculus. We need *continuity*.

2.7 Continuous Function

We now seek the gold standard. We want to isolate the functions whose values do not jump and do not present unexpected "holes".

Definition 2.7.1 (Continuity at a Point)**.** We say a function f is *continuous at a point* a if the following three conditions are satisfied:

1. $a \in \text{dom}(f)$ (The function must actually be defined at a).
2. $\lim_{x\to a} f(x)$ exists (The values must converge to a single destination).
3. $\lim_{x\to a} f(x) = f(a)$ (The destination must exactly match reality).

As is common in mathematics, it pays to have multiple perspectives on a single concept. Depending on the problem at hand, we can recast this definition into equivalent forms.

Proposition 2.7.2. *The following statements are equivalent to f is continuous at a:*

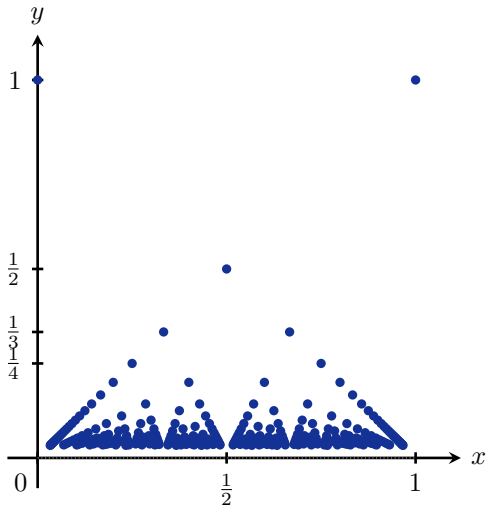


Figure 2.11 – Graph of Thomae's function (the Popcorn function) on $[0, 1]$

1. The Limit Definition: $a \in \mathrm{dom}(f)$ *and* $\lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x) = f(a)$.
2. The ε-δ Definition: $a \in \mathrm{dom}(f)$, *and for each* $\varepsilon > 0$, *there exists a* $\delta = \delta(a, \varepsilon) > 0$ *such that*
$$|f(x) - f(a)| < \varepsilon \quad \textit{whenever} \quad 0 < |x - a| < \delta.$$

⚠ *Caveat. A Subtlety in the ε-δ Formulation*

The restriction $0 < |x - a| < \delta$ is a legacy inherited from the definition of a limit (which *intentionally ignores* the behavior exactly at the point $x = a$). However, because continuity demands that the limit equals $f(a)$, the inequality $|f(x) - f(a)| < \varepsilon$ is trivially satisfied when $x = a$ (since $0 < \varepsilon$). Thus, for continuity, one often sees the simpler implication: $|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon$.

Example 2.7.3. Consider the simple parabola $f(x) = x^2$. Let us verify its continuity at $x = 1$. We showed in Example 2.5.3 that $\lim_{x\to 1} x^2 = 1$. On the other hand, direct evaluation yields $f(1) = 1^2 = 1$. Since the limit and the function value coincide, i.e., $\lim_{x\to 1} f(x) = f(1)$, we conclude that $f(x) = x^2$ is continuous at $x = 1$.

Moral. The Secret of Limits

If you know *a priori* that f is continuous at a, then the arduous task of finding the limit $\lim_{x\to a} f(x)$ is entirely equivalent to merely calculating $f(a)$. In fact, the whole business of evaluating limits in practice is largely an exercise in reducing complex expressions to known continuous functions!

For instance, consider $f(x) = \sqrt{x}$. We know it is continuous on $[0, \infty)$. Therefore, for

any $a > 0$, evaluating the limit is as effortless as substitution:

$$\lim_{x \to a} \sqrt{x} = \sqrt{a}.$$

2.7.1 Global Continuity and the Limit Cheat Codes

> The whole business of finding limits is ultimately a reduction to continuous functions.
>
> — *A guiding philosophy of Calculus*

We can extend this local, point-wise property to a global property over an entire domain.

Definition 2.7.4 (Continuity on an Interval)**.** A function f is *continuous* on an open interval (a, b) if it is continuous at every point in that interval. In particular, if $(a, b) = (-\infty, \infty)$, we say f is *continuous everywhere*.

For f to be continuous on a closed interval $[a, b]$, f must be continuous on the interior (a, b), and in addition, it must respect the boundary values via its one-sided limits:

$$\lim_{x \to a^+} f(x) = f(a) \quad \text{and} \quad \lim_{x \to b^-} f(x) = f(b).$$

> *Intuition.*
>
> Geometrically, a continuous function on an interval corresponds to a graph that you can draw as a single, unbroken curve without lifting your pen. There are no holes, jumps, or vertical asymptotes.

To practically evaluate limits using our guiding philosophy above, we need an arsenal of continuous functions.

Theorem 2.7.5. *The following families of functions are continuous everywhere on their respective* domains*:*

1. Polynomials: $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$.
2. Rational Functions: $r(x) = \frac{p(x)}{q(x)}$ *(where* $q(x) \neq 0$*).*
3. Radicals and Powers: $f(x) = \sqrt[n]{x}$*, and more generally* $f(x) = x^\alpha$ *for* $\alpha \in \mathbb{R}$.
4. Trigonometric Functions: $\sin x$, $\cos x$, $\tan x$, $\cot x$, *etc., and their inverses* $\sin^{-1} x$, $\cos^{-1} x$, *etc.*
5. Exponentials and Logarithms: e^x, b^x, $\log x$, $\log_b x$, *etc.*

An important detail is the *domain*. For example $f(x) = x^{\sqrt{2}}$ is continuous on $\mathrm{dom}(f) = [0, \infty)$, but is not defined outside this interval, let alone continuous.

Exercise 2.7.6. *For the curious minds:* Use the formal ε-δ definition to prove the continuity of these families. (Hint: A good starting point is proving that $\sin x$ and $\cos x$ are continuous everywhere via trigonometric identities).

Let us see our limit cheat code in action. What happens when our point of interest lies smoothly inside the domain?

Example 2.7.7. Evaluate $\lim_{x\to 1} \dfrac{x+1}{x^3+1}$.

Solution. Let $r(x) = \frac{x+1}{x^3+1}$. This is a rational function. We must first determine its domain. The denominator vanishes when $x^3 + 1 = 0$, which implies $x^3 = -1$, or $x = -1$. Thus, the domain is:

$$\text{dom}(r) = \mathbb{R} \setminus \{-1\} = (-\infty, -1) \cup (-1, \infty).$$

Since $x = 1 \in \text{dom}(r)$, and r is continuous on its entire domain, we can evaluate the limit purely by direct substitution:

$$\lim_{x\to 1} r(x) = r(1) = \frac{1+1}{1^3+1} = \frac{2}{2} = 1.$$

□

But what happens when the point of interest lies *outside* the domain?

Example 2.7.8. Evaluate $\lim_{x\to -1} \dfrac{x+1}{x^3+1}$.

Solution. Consider the rational function $r(x) = \frac{x+1}{x^3+1}$. As computed in Example 2.7.7, $\text{dom}(r) = (-\infty, -1) \cup (-1, \infty)$.

Crucially, notice that $x = -1 \notin \text{dom}(r)$. If we recklessly attempt to substitute $x = -1$, the term $x + 1$ generates a 0 in both the numerator and the denominator, yielding the indeterminate form $0/0$.

Our central idea is to remove the offending $x + 1$ factor from the denominator $x^3 + 1$. Recall the algebraic identity for the sum of cubes:

$$a^3 + b^3 = (a + b)(a^2 - ab + b^2).$$

Applying this, we can factor the denominator as $x^3 + 1 = (x + 1)(x^2 - x + 1)$. Therefore, we rewrite $r(x)$ as:

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

Now, we must pause. Can we simply cancel the $(x + 1)$ terms?

Yes, and the justification lies deep in the philosophy of the limit! The expression $\lim_{x\to -1}$ means that we consider values of x that are arbitrarily close to -1, *but strictly not equal to* -1. Because $x \neq -1$, it follows that $x + 1 \neq 0$. Thus, division by $(x + 1)$ is a perfectly legal operation inside the limit.

We can now cleanly reduce our function:

$$\lim_{x\to -1} r(x) = \lim_{x\to -1} \frac{1}{x^2 - x + 1}.$$

Let $g(x) = \frac{1}{x^2-x+1}$. The roots of $x^2 - x + 1 = 0$ are complex, meaning this new denominator never vanishes on the real line (equivalently, you can complete the square $x^2 - x + 1 = \left(x - \frac{1}{2}\right)^2 + \frac{3}{4} > 0$). Consequently, $\text{dom}(g) = (-\infty, \infty)$, meaning our new function g is continuous everywhere!

Having successfully reduced our problem to a continuous function evaluated comfortably inside its domain, we finish with direct substitution:

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

□

Exercise 2.7.9. Multiply the numerator and denominator by the conjugate $(\sqrt{x+1}+1)$ to eliminate the indeterminate form, and show that $\lim_{x\to 0} \frac{x}{\sqrt{x+1}-1} = 2$.

2.7.2 Discontinuity: What Does It Mean to Break?

Definition 2.7.10 (Discontinuity)**.** If a function f is not continuous at a point a, we say that f is *discontinuous* at a, and that a is a discontinuity of f.

We classify these breaks into two physical archetypes:

- *Removable Discontinuity:* A trivial flaw. If f can be made continuous by appropriately defining (or redefining) $f(a)$, the discontinuity is removable. Think of this as a missing pebble in an otherwise smoothly paved road; the overall macroscopic structure remains perfectly intact, and we can simply drop a new pebble into the hole.
- *Non-removable Discontinuity:* A fundamental structural failure. The gap cannot be bridged by redefining a single point (e.g., a jump, an oscillation, or an asymptote).[3]

> *Moral. Classification of Discontinuities*
>
> Not all discontinuities are created equal. Some are mere scratches on the surface of a function, while others are fundamental tears in its graph.
> If you can fix a discontinuous function with a single drop of ink from your pen, the discontinuity is removable. If you need to lift your pen to redraw a whole section of the curve, it is non-removable.

Example 2.7.11 (Patching a Removable Discontinuity)**.** Consider the rational function

$$r(x) = \frac{x+1}{x^3+1}. \tag{2.20}$$

The domain of r is $(-\infty,-1)\cup(-1,\infty)$. Notice that $r(-1)$ does not exist, so the function is undefined there. However, as we have seen in Example 2.7.8, the limit itself exists:

$$\lim_{x\to -1} r(x) = \frac{1}{3}. \tag{2.21}$$

If we construct a new function $f(x)$ that patches this hole,

$$f(x) \coloneqq \begin{cases} r(x), & x \neq -1 \\ \frac{1}{3}, & x = -1, \end{cases} \tag{2.22}$$

[3] In a physical fluid system, this is akin to a sudden shockwave where quantities jump instantaneously. The solution of Burger equation is a prototype of this behavior.

then for $a = -1$, we have

$$\lim_{x \to -1} f(x) = \lim_{x \to -1} r(x) = \frac{1}{3} = f(-1). \tag{2.23}$$

By definition, f is continuous at $x = -1$. Furthermore, for $a \neq -1$, $f(x) = r(x)$ is continuous on its original domain.

So f is continuous on the entirety of $\mathbb{R}$. We simply redefined $r(x) = 1/3$ at $x = -1$ to obtain a continuous function f. This perfectly illustrates that $x = -1$ is a *removable discontinuity*.

Example 2.7.12 (An Unbridgeable Gap)**.** Now, consider the hyperbola $f(x) = \frac{1}{x}$, with domain $(-\infty, 0) \cup (0, \infty)$. Let us examine the behavior near the origin:

$$\lim_{x \to 0^+} f(x) = +\infty \quad \text{and} \quad \lim_{x \to 0^-} f(x) = -\infty. \tag{2.24}$$

We clearly have $\lim_{x \to 0^+} f(x) \neq \lim_{x \to 0^-} f(x)$. The function diverges, violently tearing in opposite directions. Thus, we have a *non-removable discontinuity* at $x = 0$.

> *Story. The Ultraviolet Catastrophe*
>
> In the early 20th century, classical physics predicted that the energy radiated by a black body at short wavelengths would go to infinity–a massive, real-world divergence akin to our $1/x$ example. This *ultraviolet catastrophe* was a glaring non-removable discontinuity in the laws of physics. Max Planck resolved it by introducing quantum mechanics, fundamentally redefining the domain of our universe to fix the divergence!

2.7.3 Continuous Functions Commute with Limits

Continuity behaves beautifully under the standard operations of calculus. When we build complex functions out of simpler, continuous building blocks, the resulting structures usually inherit this smoothness.

Theorem 2.7.13 (Continuity of Composite Functions)**.** *If g is continuous at a, and f is continuous at $g(a)$, then the composition $f \circ g$ is continuous at a.*

The rigorous proof requires an $\varepsilon - \delta$ argument, which is left as an exercise for the reader. However, the heuristic of Theorem 2.7.13 is quite straightforward. Theorem 2.7.13 says that

$$\lim_{x \to a} (f \circ g)(x) = (f \circ g)(a).$$

Note that $(f \circ g)(a) = f(g(a)) = f\left(\lim_{x \to a} g(x)\right)$, we can write (2.7.3) as

$$\lim_{x \to a} f(g(x)) = f\left(\lim_{x \to a} g(x)\right). \tag{2.25}$$

> *Moral.*
>
> Continuous functions commute with limits. This is their superpower. If f is continuous, you can safely pull limits inside its argument.

Let's see this elegant machinery in action.

Example 2.7.14. Compute $\lim_{x\to 3}\sqrt{x^2-1}$.

Let $g(x)=x^2-1$ and $f(x)=\sqrt{x}$. Because the square root function is continuous on its domain, we can pull the limit right inside:

$$\lim_{x\to 3}\sqrt{x^2-1}=\sqrt{\lim_{x\to 3}(x^2-1)}=\sqrt{3^2-1}=\sqrt{8}=2\sqrt{2}. \tag{2.26}$$

Example 2.7.15. Compute $\lim_{x\to \pi/2}\mathrm{e}^{\cos(x-\pi)}$.

Let $g(x)=\cos(x-\pi)$ and $f(x)=\mathrm{e}^x$. By Theorem 2.7.13,

$$\lim_{x\to \pi/2}\mathrm{e}^{\cos(x-\pi)}=\mathrm{e}^{\lim_{x\to\pi/2}\cos(x-\pi)}=\mathrm{e}^{\cos(\pi/2-\pi)}=\mathrm{e}^{\cos(-\pi/2)}=\mathrm{e}^0=1. \tag{2.27}$$

Theorem 2.7.16 (Continuity and Arithmetic)**.** *Suppose f and g are continuous at $x=a$. Then the following functions are also continuous at a:*

1. *cf, where c is a constant.*
2. *$f\pm g$.*
3. *fg.*
4. *$\frac{1}{f}$, provided that $f(a)\neq 0$.*

Example 2.7.17. Consider the limit $\lim_{x\to 1}\frac{\log x}{x^2+1}$.

Let $g(x)=\log x$ and $f(x)=\frac{1}{x^2+1}$. Both are continuous on $(0,\infty)$. Since $1\in \operatorname{dom}\left(\frac{\log x}{x^2+1}\right)$, we simply evaluate:

$$\lim_{x\to 1}\frac{\log x}{x^2+1}=\frac{\log 1}{1^2+1}=\frac{0}{2}=0. \tag{2.28}$$

We often encounter a limit that looks completely undetermined ($0/0$ or ∞/∞). It tries to trick us into believing there is a singularity, but with a little algebraic surgery, we can reveal the smooth reality underneath.

Example 2.7.18 (Removal of a Singularity)**.** Evaluate the limit

$$\lim_{x\to 4}\frac{x^2-16}{\sqrt{x}-2}. \tag{2.29}$$

Direct substitution yields $x=4 \implies \frac{0}{0}$, which is undetermined. The core idea here is to *remove the singularity* at $x=4$ to obtain a function that is continuous at $x=4$. Notice that the problematic term is $(x-4)\to 0$. We must hunt for this hidden $(x-4)$ in both the numerator and the denominator.

There are two paths we can take to conquer this limit.

1. *Method 1: Rationalizing the Denominator.*
   For the top, we factor a difference of squares:

$$x^2-16=(x-4)(x+4).$$

For the bottom, we multiply by the conjugate, exploiting the identity $a^2 - b^2 = (a-b)(a+b)$:

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

Putting it all together, we have:

$$\begin{aligned}\lim_{x\to 4}\frac{x^2-16}{\sqrt{x}-2} &= \lim_{x\to 4}\frac{(x-4)(x+4)}{\sqrt{x}-2}\\ &= \lim_{x\to 4}\frac{(x-4)(x+4)}{\frac{x-4}{\sqrt{x}+2}}\\ &= \lim_{x\to 4}(x+4)(\sqrt{x}+2).\end{aligned}$$

2. *Method 2: Factoring the Numerator.*
Alternatively, we can be slightly more clever and recognize that $(x-4)$ itself is a difference of squares!

$$x-4 = (\sqrt{x})^2 - 2^2 = (\sqrt{x}-2)(\sqrt{x}+2). \tag{2.30}$$

Then, the limit becomes:

$$\begin{aligned}\lim_{x\to 4}\frac{x^2-16}{\sqrt{x}-2} &= \lim_{x\to 4}\frac{(x-4)(x+4)}{\sqrt{x}-2}\\ &= \lim_{x\to 4}\frac{(\sqrt{x}-2)(\sqrt{x}+2)(x+4)}{\sqrt{x}-2}\\ &= \lim_{x\to 4}(\sqrt{x}+2)(x+4).\end{aligned}$$

By algebraically cancelling the singularity, we are left with a polynomial-like expression that is fully continuous at $x=4$. We can now evaluate:

$$\lim_{x\to 4}\frac{x^2-16}{\sqrt{x}-2} = (4+4)(\sqrt{4}+2) = 8\cdot 4 = 32.$$

⚠ *Caveat.*

It is incredibly tempting to cancel out $\frac{x-4}{x-4}$ without a second thought. But remember, this division is only mathematically valid precisely because in a limit $x\to 4$, x is *never actually equal* to 4! The limit asks *what happens as we get close?*, and algebra lets us answer without breaking the math by dividing by zero.

Exercise 2.7.19 (Power Rule)**.** Show that if $\lim_{x\to a} h(x)$ exists and $\lim_{x\to a} h(x) \geq 0$ then $\lim_{x\to a}(h(x))^\beta = (\lim_{x\to a} h(x))^\beta$. *Hint: Use* (2.25)*. Why do we need* $\lim_{x\to a} h(x)$ *to be nonnegative?*

2.8 The Limit Laws

We state the Limit Laws, which are the algebraic engines powering our limit calculation.

Theorem 2.8.1 (Limit Laws)**.** *Suppose* $\lim_{x\to a} f(x)$ *and* $\lim_{x\to a} g(x)$ *exist. Let* c *and* β *be constants. Then:*

1. Sum/Difference Rule: $\lim_{x\to a}\left(f(x) \pm g(x)\right) = \lim_{x\to a} f(x) \pm \lim_{x\to a} g(x)$.
2. Product Rule: $\lim_{x\to a}\left(f(x)g(x)\right) = \left(\lim_{x\to a} f(x)\right)\left(\lim_{x\to a} g(x)\right)$.

Note that these laws also apply for 1-sided limits.

Proof. Exercise for the curious minds. Use ε-δ definition.[4] □

Example 2.8.2. Let us apply these laws to a mildly frightening expression:

$$\lim_{x\to 2} \frac{\sqrt{2x^3+9}+3x-1}{4x+1}.$$

Using the Limit Laws, since the denominator evaluated at $x=2$ gives $4(2)+1=9\neq 0$, we can simply pass the limit straight through the continuous functions:

$$\begin{aligned}\lim_{x\to 2} \frac{\sqrt{2x^3+9}+3x-1}{4x+1} &= \frac{\sqrt{2(2)^3+9}+3(2)-1}{4(2)+1}\\ &= \frac{\sqrt{16+9}+6-1}{9}\\ &= \frac{\sqrt{25}+5}{9} = \frac{10}{9}.\end{aligned}$$

Example 2.8.3 (Piecewise Discontinuities)**.** Consider a piecewise function that abruptly changes its rule:

$$f(x) \coloneqq \begin{cases} -2x+4 & \text{if } x \leq 1 \\ \sqrt{x-1} & \text{if } x > 1 \end{cases}$$

To understand its geometry at the transition point $x=1$, we evaluate the one-sided limits. Approaching from the left (where $x \leq 1$), our path follows the linear slope:

$$\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} (-2x+4) = -2(1)+4 = 2.$$

Approaching from the right (where $x > 1$), we follow the square root curve:

$$\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} \sqrt{x-1} = \sqrt{1-1} = 0.$$

Since the left limit and the right limit disagree ($2 \neq 0$), the overall two-sided limit $\lim_{x\to 1} f(x)$ does not exist. The graph experiences a severe tearing, creating a *nonremovable discontinuity* at $x=1$.

[4] There is another way to get this result, but requires a bit of multivariate calculus (Calculus III): define $h(u,v) = u+v$ and $k(x) = (f(x), g(x))$, then $f(x)+g(x) = (h\circ k)(x)$. The concept of continuity can be extended to multi-dimensional spaces, and it can be showed that h is continuous, so you can swap the limit inside $\lim_{x\to a} h(k(x)) = h(\lim_{x\to a} k(x))$, which is exactly the sum rule. The limit laws are just the limit theorem for composition of functions in disguise! (Theorem 2.7.13)

Example 2.8.4 (The 0/0 Form: An Invitation to Factor)**.** Not all singularities are so catastrophic. Consider the rational expression:

$$\lim_{x\to 2} \frac{x^2-6x+8}{x^2-4}$$

Direct substitution yields the indeterminate form $\frac{0}{0}$.

> *Moral. The Root of* $0/0$
>
> When evaluating a limit, encountering a $0/0$ form is not a sign of failure; it indicates a removable singularity. A $0/0$ form tells us that the numerator and denominator share a common root. Algebraic factorization is the key that divides out the offending zero.

We decompose the polynomials. For the numerator, we recall *Vieta's Theorem*: for a quadratic $ax^2+bx+c = a(x-x_1)(x-x_2)$, the roots x_1, x_2 satisfy $x_1+x_2 = -b/a$ and $x_1 \cdot x_2 = c/a$. We need two numbers that sum to 6 and multiply to 8. The direct substitution yields a root 2. Thus the other root, by Vieta's Theorem, must be 4. The denominator is a classic difference of squares. Thus, we rewrite the limit:

$$\lim_{x\to 2} \frac{x^2-6x+8}{x^2-4} = \lim_{x\to 2} \frac{(x-2)(x-4)}{(x-2)(x+2)}.$$

As long as $x \to 2$, we strictly have $x \neq 2$, meaning $x-2 \neq 0$. We may safely cancel the common term:

$$\lim_{x\to 2} \frac{x^2-6x+8}{x^2-4} = \lim_{x\to 2} \frac{x-4}{x+2} = \frac{2-4}{2+2} = \frac{-2}{4} = -\frac{1}{2}.$$

Example 2.8.5.

$$\lim_{x\to 1} \frac{\sqrt{x}-1}{x-1} \tag{2.31}$$

Once again, substitution yields the $0/0$ form. There are two elegant methods to dismantle this.

1. *Method 1: The Rationalization Trick.*
   We multiply the numerator and the denominator by the algebraic conjugate of the numerator.

$$\begin{aligned}\lim_{x\to 1} \frac{\sqrt{x}-1}{x-1} \cdot \frac{\sqrt{x}+1}{\sqrt{x}+1} &= \lim_{x\to 1} \frac{(\sqrt{x})^2-1^2}{(x-1)(\sqrt{x}+1)} \\ &= \lim_{x\to 1} \frac{x-1}{(x-1)(\sqrt{x}+1)} \\ &= \lim_{x\to 1} \frac{1}{\sqrt{x}+1} = \frac{1}{2}.\end{aligned}$$

2. *Method 2: The Hidden Difference of Squares.*
   Rather than building the numerator up, we break the denominator down. By recognizing

that $x - 1 = (\sqrt{x})^2 - 1^2$, we can factor it as a difference of squares:

$$\begin{aligned}\lim_{x\to 1} \frac{\sqrt{x}-1}{x-1} &= \lim_{x\to 1} \frac{\sqrt{x}-1}{(\sqrt{x}-1)(\sqrt{x}+1)} \\ &= \lim_{x\to 1} \frac{1}{\sqrt{x}+1} = \frac{1}{2}.\end{aligned}$$

Intuition. All roads lead to Rome

Whether we build the numerator up via rationalization, or break the denominator down by factoring, the underlying truth is conserved. Mathematics is deeply symmetric in this way: constructive and deconstructive approaches invariably meet at the same limit.

Exercise 2.8.6. Find $\lim_{x\to 2^+} f(x)$, $\lim_{x\to 2^-} f(x)$, and $\lim_{x\to 2} f(x)$ where

$$f(x) = \frac{x^3 - 6x^2 + 8x}{\sqrt{x-2}}.$$

Caveat. One-Sided Blindness

Before rushing into algebraic manipulation for the exercise above, remember that a function cannot have a limit where it does not exist. The real-valued term $\sqrt{x-2}$ strictly dictates that $x \geq 2$. Consequently, approaching $x = 2$ from the left ($x \to 2^-$) is analytically meaningless. The left-sided limit does not exist, which in turn means the two-sided limit cannot exist!

2.9 The Squeeze Theorem

How do we evaluate limits when algebra completely fails us? Consider the function:

$$f(x) \coloneqq \begin{cases} x \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

This function is a beautiful monstrosity. As $x \to 0$, the argument $1/x$ goes to infinity, driving the sine function through an infinite number of cycles arbitrarily close to the origin. Direct substitution gives $0 \times$ undefined, and no amount of factoring will save us.

However, we know that the sine function is bounded: $-1 \leq \sin(1/x) \leq 1$. Multiplying this entire inequality by $|x|$ yields:

$$-|x| \leq x \sin\left(\frac{1}{x}\right) \leq |x|, \tag{2.32}$$

see Figure 2.12.

To formally exploit this geometry, we need a fundamental lemma.

Lemma 2.9.1 (Limit preserves inequalities)**.** *If $h(x) \leq g(x)$ near a, then*

$$\lim_{x\to a} h(x) \leq \lim_{x\to a} g(x).$$

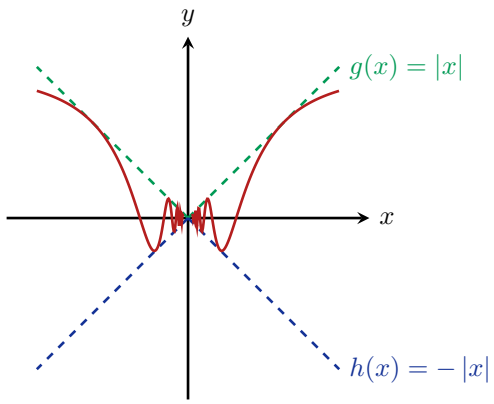


Figure 2.12 – The graph of $f(x) = x\sin(1/x)$ bounded by $g(x) = |x|$ and $h(x) = -|x|$, illustrating the Squeeze Theorem at $x = 0$.

Proof. Exercise for the reader. *Hint: use contradiction and ε-δ definition of limit.* □

Applying Lemma 2.9.1 to (2.32), we take the limit as $x \to 0$ across the entire inequality:

$$\lim_{x\to 0}(-|x|) \le \lim_{x\to 0} f(x) \le \lim_{x\to 0}|x|$$
$$0 \le \lim_{x\to 0} f(x) \le 0.$$

Trapped between two zeroes, the limit has no choice but to be 0. Furthermore, since $\lim_{x\to 0} f(x) = 0 = f(0)$, we have proven that f is continuous at the origin!

This technique is so important that it earns the title of a Theorem.

Theorem 2.9.2 (The Squeeze Theorem)**.** *If $h(x) \le f(x) \le g(x)$ for all x in an open interval containing a, except possibly at a itself, and if*

$$\lim_{x\to a} h(x) = L = \lim_{x\to a} g(x),$$

then $\lim_{x\to a} f(x)$ exists and is strictly equal to L.

Story. The Sandwich of Truth

Historically, tracking wild oscillations was a monumental task. The Squeeze Theorem (often affectionately called the *Sandwich Theorem* or the *Two Policemen Theorem*) is a masterpiece of non-constructive reasoning. Instead of wrestling directly with the chaotic behavior of $f(x)$, we delegate the heavy lifting to two simpler, bounding functions. It is a mathematical manifestation of a physical reality: if a particle is trapped in a shrinking corridor, its ultimate destination is determined entirely by the location of the walls.

We now arrive at a limit that forms the bedrock of all periodic phenomena in calculus. If you want to differentiate a sine wave—and therefore, if you want to understand light, sound, or quantum mechanics—you must pass through this limit.

Corollary 2.9.3 (Two Fundamental Limits)**.** *The following limits hold true:*

1. $\lim_{x\to 0} \dfrac{\sin x}{x} = 1$
2. $\lim_{x\to 0} (1+x)^{1/x} = \mathrm{e}$

We will provide a geometric proof of the first limit.

Proof. Let $\theta \in (0, \pi/2)$ be a small angle measured in radians. Consider the unit circle in the Cartesian plane. We construct three nested geometric regions based on the angle θ, see Figure 2.13.

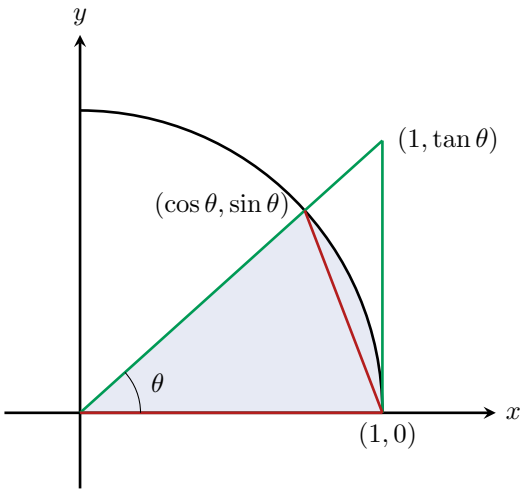


Figure 2.13 – Geometric representation of trigonometric relationships in the first quadrant of the unit circle.

1. *The Inner Triangle (Red):* Formed by the origin $(0,0)$, the point $(1,0)$, and the point $(\cos\theta, \sin\theta)$ on the unit circle. Its base is 1 and its height is $\sin\theta$.
2. *The Circular Sector (Blue Wedge):* The wedge of the unit circle swept out by the angle θ.
3. *The Outer Triangle (Green):* Formed by extending the hypotenuse until it intersects the vertical tangent at $x = 1$, meeting at $(1, \tan\theta)$. This right triangle has base 1 and height $\tan\theta$.

Visually and geometrically, each region is entirely contained within the next. Thus, their respective areas obey a strict hierarchy:

$$\text{area}(\text{inner } \triangle) \leq \text{area}(\text{sector}) \leq \text{area}(\text{outer } \triangle).$$

Substituting the standard area formulas yields:

$$\frac{1}{2}(1)(\sin\theta) \leq \frac{\theta}{2\pi}\left(\pi \cdot 1^2\right) \leq \frac{1}{2}(1)(\tan\theta).$$

Simplifying the terms, we obtain a fundamental inequality:

$$\sin\theta \leq \theta \leq \tan\theta.$$

Because $\theta > 0$, we know $\sin\theta > 0$. We may divide the entire chain by $\sin\theta$ without flipping the inequalities to obtain:

$$1 \leq \frac{\theta}{\sin\theta} \leq \frac{1}{\cos\theta}.$$

Taking the reciprocal of all terms forces the direction of the inequalities to reverse:

$$\cos\theta \leq \frac{\sin\theta}{\theta} \leq 1.$$

At this point, the algebraic and geometric path is exhausted, but the analytic path opens. We take the limit as $\theta \to 0^+$. Since the cosine function is continuous at zero, $\lim_{\theta\to 0^+} \cos\theta = 1$. The upper bound is trivially the constant 1.

By Theorem 2.9.2, our function satisfies

$$\lim_{\theta\to 0^+} \frac{\sin\theta}{\theta} = 1.$$

By noting the even symmetry of the function $\frac{\sin(-\theta)}{(-\theta)} = \frac{\sin\theta}{\theta}$, this result immediately extends to the two-sided limit, completing the proof. □

The limit

$$\lim_{x\to 0} \frac{\sin x}{x} = 1. \tag{2.33}$$

is often considered the *atomic* limit of trigonometry: it tells us that for very small angles, $\sin x$ is indistinguishable from x itself. Thus we can *heuristically* cancel them altogether as if we are factoring out the singularity in a rational function (to be *rigorous*, you need to include the $\lim_{x\to 0}$, however). We will use (2.33) to resolve more complicated limits.

Example 2.9.4. Evaluate the limit

$$\lim_{x\to 0} \frac{1-\cos x}{x}.$$

> *Intuition. Order of Vanishing*
>
> When you plug in $x = 0$, you immediately get $\frac{1-1}{0} = \frac{0}{0}$. But let's pause and consider the *rate* at which the numerator goes to zero. If you look at the graph of a cosine wave near its peak, $\cos x \approx 1 - \frac{x^2}{2}$ (you can justify this rigorously in Calculus II using the Maclaurin expansion—a special power series). Thus, the numerator $1 - \cos x$ behaves like $\frac{x^2}{2}$. It vanishes quadratically, while the denominator vanishes linearly. A quadratic zero divided by a linear zero should still be zero! Let us now make this rigorous.

To resolve this without invoking power series, we employ the half-angle trigonometric identity. Recall the double angle formula for cosine that was derived via Euler identity in

Lemma 1.7.5:

$$\begin{aligned}\cos 2t &= \cos^2 t - \sin^2 t \\ &= (1 - \sin^2 t) - \sin^2 t \\ &= 1 - 2\sin^2 t.\end{aligned}$$

Rearranging this gives us:

$$2\sin^2 t = 1 - \cos 2t.$$

Now, setting $x = 2t$ (or equivalently $t = x/2$), we obtain:

$$1 - \cos x = 2\sin^2\left(\frac{x}{2}\right). \tag{2.34}$$

We use (2.34) to rewrite our limit:

$$\lim_{x\to 0} \frac{1-\cos x}{x} = \lim_{x\to 0} \frac{2\sin^2(x/2)}{x}.$$

Let us make the substitution $u = \frac{x}{2}$. Notice that as $x \to 0$, $u \to 0$. The denominator x becomes $2u$. Thus, the limit transforms into:

$$\begin{aligned}\lim_{u\to 0} \frac{2\sin^2 u}{2u} &= \lim_{u\to 0} \frac{\sin^2 u}{u} \\ &= \lim_{u\to 0} \left((\sin u) \cdot \frac{\sin u}{u}\right).\end{aligned}$$

Because the limit of a product is the product of the limits (provided both exist), we split this apart:

$$\begin{aligned}\left(\lim_{u\to 0} \sin u\right)\left(\lim_{u\to 0} \frac{\sin u}{u}\right) &= (\sin 0)\cdot 1 \\ &= 0 \cdot 1 \\ &= 0.\end{aligned}$$

Example 2.9.5. Evaluate $\lim_{x\to 0} \dfrac{\tan x}{x}$.

Here, we decompose the tangent function into its sine and cosine constituents to leverage (2.33).

$$\begin{aligned}\lim_{x\to 0} \frac{\tan x}{x} &= \lim_{x\to 0} \left(\frac{\sin x}{x\cos x}\right) \\ &= \lim_{x\to 0} \left(\frac{\sin x}{x} \cdot \frac{1}{\cos x}\right). \\ &= \left(\lim_{x\to 0} \frac{\sin x}{x}\right)\left(\lim_{x\to 0} \frac{1}{\cos x}\right) \\ &= 1 \cdot \frac{1}{1} = 1.\end{aligned}$$

Example 2.9.6. Evaluate $\lim_{x\to 0} \dfrac{\sin 4x}{x}$.

This limit introduces a slight scaling mismatch between the argument of the sine and the denominator. The limit (2.33) demands symmetry: if we have $4x$ inside the sine, we need a $4x$ in the denominator. We enforce this by multiplying the numerator and denominator by 4:

$$\lim_{x\to 0} \frac{\sin 4x}{x} = \lim_{x\to 0} \frac{4\sin 4x}{4x}.$$

Let us define a new variable $u := 4x$. As $x \to 0$, it follows that $u \to 0$ ($\lim_{x\to 0} u = \lim_{x\to 0} 4x = 0$). The limit becomes:

$$\lim_{u\to 0} \frac{4\sin u}{u} = 4 \lim_{u\to 0} \frac{\sin u}{u} = 4 \cdot 1 = 4.$$

Exercise 2.9.7. Can you generalize this to show that $\lim_{x\to 0} \frac{\sin(ax)}{bx} = \frac{a}{b}$ for any non-zero constants a, b?

We now shift from the oscillations of trigonometry to the growth of logarithms.

Example 2.9.8. Evaluate the right-hand limit:

$$\lim_{x\to 0^+} \frac{\log(1+x)}{x}.$$

Once again, direct substitution yields $0/0$. To crack this open, we rewrite the fraction as a product, leveraging the algebraic properties of logarithms. Recall that $b\log a = \log(a^b)$. Therefore:

$$\lim_{x\to 0^+} \frac{1}{x}\log(1+x) = \lim_{x\to 0^+} \log\left[(1+x)^{\frac{1}{x}}\right].$$

Because the natural logarithm function, $\log(y)$, is perfectly continuous for $y > 0$, we are legally permitted to pull the limit *inside* the logarithm:

$$\lim_{x\to 0^+} \frac{1}{x}\log(1+x) = \log\left(\lim_{x\to 0^+} (1+x)^{\frac{1}{x}}\right) = \log(\mathrm{e}) = 1.$$

Example 2.9.9. Find the limit $\lim_{x\to 3} \dfrac{x^2-9}{|x-3|}$ (if exists).

First, we factor the numerator using the difference of squares: $a^2 - b^2 = (a-b)(a+b)$.

$$\lim_{x\to 3} \frac{(x-3)(x+3)}{|x-3|}.$$

> ⚠ *Caveat. The Absolute Value*
>
> Whenever you see an absolute value in a limit, alarm bells should ring. The function $|x-3|$ is not a single smooth entity; it is a two-faced *piecewise* function with a sharp corner at $x = 3$. You *must* split the limit into its left and right components to see the full picture.

By definition, the absolute value is resolved as:

$$|x-3| = \begin{cases} (x-3), & \text{when } x \geq 3 \\ -(x-3), & \text{when } x < 3 \end{cases}$$

Let us probe the behavior from the right ($x \to 3^+$). Here, $x > 3$, so $|x-3| = x-3$.

$$\begin{aligned} \lim_{x\to 3^+} \frac{(x-3)(x+3)}{x-3} &= \lim_{x\to 3^+} (x+3) \\ &= 6. \end{aligned}$$

Now, we approach from the left ($x \to 3^-$). Here, $x < 3$, so $|x-3| = -(x-3)$.

$$\begin{aligned} \lim_{x\to 3^-} \frac{(x-3)(x+3)}{-(x-3)} &= \lim_{x\to 3^-} -(x+3) \\ &= -6. \end{aligned}$$

Because the right-hand limit (6) does not equal the left-hand limit (-6), the two-sided limit cannot reconcile the difference. Therefore,

$$\lim_{x\to 3} \frac{x^2-9}{|x-3|} \quad \text{does not exist.}$$

Example 2.9.10. Evaluate $\lim_{x\to a} \dfrac{x^2-a^2}{\sqrt{x}-\sqrt{a}}$, where $a > 0$.

Direct substitution gives $\frac{a^2-a^2}{\sqrt{a}-\sqrt{a}} = \frac{0}{0}$. At first glance, the numerator is a polynomial while the denominator contains square roots. To bridge this gap, we must expand the numerator until we expose the term hiding in the denominator.

First, expand $x^2 - a^2$:

$$\frac{(x-a)(x+a)}{\sqrt{x}-\sqrt{a}}.$$

Now comes a clever algebraic trick. We can view $x - a$ as a difference of squares in its own right, by treating $x = (\sqrt{x})^2$ and $a = (\sqrt{a})^2$.

$$x - a = (\sqrt{x}-\sqrt{a})(\sqrt{x}+\sqrt{a}).$$

Substituting this back into our limit yields:

$$\lim_{x\to a} \frac{(\sqrt{x}-\sqrt{a})(\sqrt{x}+\sqrt{a})(x+a)}{\sqrt{x}-\sqrt{a}}.$$

The singular term $(\sqrt{x}-\sqrt{a})$ is now completely exposed in both the numerator and the denominator. We may safely cancel it (since $x \to a$ implies $x \neq a$).

$$\begin{aligned} \lim_{x\to a} \frac{x^2-a^2}{\sqrt{x}-\sqrt{a}} = \lim_{x\to a} (\sqrt{x}+\sqrt{a})(x+a) &= (\sqrt{a}+\sqrt{a})(a+a) \\ &= (2\sqrt{a})(2a) \\ &= 4a\sqrt{a}. \end{aligned}$$

Moral. Cancellation of Singularities

When a rational function evaluates to $0/0$ at $x = a$, the Factor Theorem from algebra guarantees that $(x - a)$ is a hidden factor in both the numerator and the denominator. The game is simply to pull these factors out—whether through trig identities, logarithms, or cunning factorization—and cancel them out of existence.

Exercise 2.9.11. Evaluate $\lim_{x \to a} \dfrac{x^2 - a^2}{\sqrt[3]{x} - \sqrt[3]{a}}$.

2.10 Limit at Infinity

What does it mean for a function to approach a value L as x goes to infinity?

Definition 2.10.1 (Limit at Infinity)**.** We say the limit of f at infinity is equal to L, denoted by

$$\lim_{x \to \infty} f(x) = L, \tag{2.35}$$

if for each $\varepsilon > 0$, there exists an $M = M(\varepsilon) > 0$ such that

$$|f(x) - L| < \varepsilon \quad \text{whenever} \quad x \geq M. \tag{2.36}$$

Naturally, we can reflect this concept to the other side of the real line. We define the limit at negative infinity in a similar way:

$$\lim_{x \to -\infty} f(x) = L, \tag{2.37}$$

if for each $\varepsilon > 0$, there exists an $M = M(\varepsilon) > 0$ such that $|f(x) - L| < \varepsilon$ whenever $x \leq -M$.

But what happens if the function never settles, and instead blows up completely, i.e., $L = \pm\infty$? We can no longer use the Euclidean distance $|f(x) - L|$ because infinity is not a real number. Instead, we change our winning condition accordingly: we demand that the function overcomes any finite barrier.

Definition 2.10.2 (Infinite Limit at Infinity)**.** We write $\lim_{x \to \infty} f(x) = \infty$ if for each threshold $K > 0$, there exists an $M(K) > 0$ such that

$$f(x) \geq K \quad \text{whenever} \quad x \geq M(K). \tag{2.38}$$

Exercise 2.10.3. To test your grasp of the formal machinery, write down the rigorous $\varepsilon - \delta$ (or rather, $M - K$) definitions for the following limits:

(a) $\lim_{x \to \infty} f(x) = -\infty$

(b) $\lim_{x \to -\infty} f(x) = \infty$

(c) $\lim_{x \to -\infty} f(x) = -\infty$

Treating ∞ as a number is mathematically illicit, yet informally, it is a highly effective heuristic. Before we dive into rigorous examples, let us remind ourselves of some *informal* arithmetic involving ∞.

We can split these operations into two categories: the *determined* ones (where the outcome is unconditionally clear) and the *undetermined* ones (the dangerous indeterminate forms).

The Determined Ones

1. $c \cdot \infty = \begin{cases} +\infty & \text{if } c > 0 \\ -\infty & \text{if } c < 0 \end{cases}$
2. $\infty \cdot \infty = \infty, \quad \infty \cdot (-\infty) = -\infty, \quad (-\infty) \cdot (-\infty) = \infty$
3. $\frac{1}{\pm\infty} = 0, \quad \frac{1}{0^+} = +\infty, \quad \frac{1}{0^-} = -\infty$
4. $\pm\infty + c = \pm\infty$ for any constant c.
5. $\infty + \infty = \infty, \quad -\infty - \infty = -\infty$
6. $\infty^\alpha = \begin{cases} \infty & \text{if } \alpha > 0 \\ 0 & \text{if } \alpha < 0 \end{cases}$

The Undetermined Ones

> ⚠ *Caveat.*
>
> The following expressions are mathematical battlegrounds. They cannot be evaluated by sight; they can evaluate to literally anything depending on the underlying functions!
>
> 1. $\frac{0}{0}$ and $\frac{\infty}{\infty}$ (The quotients)
> 2. $0 \cdot \infty$ (The product)
> 3. $\infty - \infty$ (The competing infinities)
> 4. $0^0, \infty^0$, and 1^∞ (The exponential paradoxes)

Let us look at a few examples, starting from the basic definition and moving toward more sophisticated techniques.

Example 2.10.4. Show that $\lim_{x\to\infty} \frac{1}{x} = 0$ using the rigorous definition.

Proof. Fix an arbitrary $\varepsilon > 0$. We need to find an $M > 0$ such that $\left|\frac{1}{x} - 0\right| < \varepsilon$ whenever $x \geq M$. We can simply choose $M = \frac{1}{\varepsilon}$. Thus, if $x \geq \frac{1}{\varepsilon}$, we immediately have $\frac{1}{x} \leq \varepsilon$, proving the limit. □

Of course, in practice, the much easier way is to lean on our informal arithmetic: $\lim_{x\to\infty} \frac{1}{x} \approx \frac{1}{\infty} = 0$.

Example 2.10.5. Evaluate the limit $\lim_{x\to\infty} x \sin\left(\frac{1}{x}\right)$.

This limit bridges the microscopic and the macroscopic. At first glance, it is a $0 \cdot \infty$ indeterminate form. We resolve this by substituting $t = 1/x$. As $x \to \infty$, the variable t approaches 0 from the right ($t \to 0^+$).

$$\lim_{x\to\infty} x \sin\left(\frac{1}{x}\right) = \lim_{t\to 0^+} \frac{\sin t}{t} = 1. \tag{2.39}$$

Isn't it surprising? The function oscillates wildly near 0, yet flattens out beautifully near ∞, see Figure 2.14 and Figure 2.15.

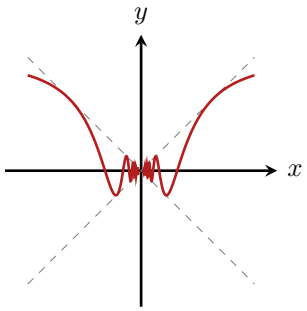


Figure 2.14 – Violent oscillations bound by $y = \pm x$ near 0.



Figure 2.15 – Flattening out asymptotically to 1 near ∞.

Remark 2.10.6. Notice that $f(x) = x\sin(1/x)$ is an even function. Therefore, its behavior as $x \to -\infty$ is identical. For large $x > 0$, the argument $1/x$ is small and positive, causing $\sin(1/x) > 0$. The function steadily approaches the asymptote $y = 1$ from below.

Proposition 2.10.7. *Let $\beta \in \mathbb{R}$. Then the behavior of power functions at infinity is*

$$\lim_{x\to\infty} x^\beta = \begin{cases} \infty & \text{if } \beta > 0 \\ 1 & \text{if } \beta = 0 \\ 0 & \text{if } \beta < 0 \end{cases} \tag{2.40}$$

Proof. Use $\varepsilon - \delta$ definition. □

Proposition 2.10.7 is the foundation to build up several interesting limits. Consider two polynomials:

$$p_n(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$$
$$q_m(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_0$$

Evaluating the limit of an isolated polynomial is trivial. We factor out the largest power of x:

$$\lim_{x\to\infty} p_n(x) = \lim_{x\to\infty} x^n \left(a_n + \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n}\right) = \infty \cdot a_n.$$

This will evaluate to ∞ if $a_n > 0$ and $-\infty$ if $a_n < 0$. But what happens when we look at the ratio of two polynomials?

$$\lim_{x\to\infty} \frac{p_n(x)}{q_m(x)} = ?$$

Intuition. Pólya's Problem Solving Strategy

This general limit is an excellent exercise to illustrate some of the techniques described by the great mathematician George Pólya in his book *How to Solve It*:

1. *Specialization:* Start with a simple case that leads to something we already know. What if the denominator is trivially $q_m(x) \equiv 1$? Then the expression reverts to a simple polynomial, and we know the *deciding factor* is exclusively the leading term, $a_n x^n$.

2. *Experimentation:* Let us factor out the dominating terms from both the numerator and the denominator, mirroring our simplified case:

$$\frac{p_n(x)}{q_m(x)} = \frac{x^n \left(a_n + \frac{a_{n-1}}{x} + \dots\right)}{x^m \left(b_m + \frac{b_{m-1}}{x} + \dots\right)} = x^{n-m} \frac{a_n + \frac{a_{n-1}}{x} + \dots}{b_m + \frac{b_{m-1}}{x} + \dots} \tag{2.41}$$

   As $x \to \infty$, all terms featuring $1/x$ vanish, leaving us with

$$\lim_{x\to\infty} \frac{p_n(x)}{q_m(x)} = \lim_{x\to\infty} x^{n-m} \cdot \left(\frac{a_n}{b_m}\right).$$

3. *Divide-and-Conquer:* We have now broken a massive, messy limit into three simple, digestible sub-problems based purely on the degrees. Applying Proposition 2.10.7,

   - $n = m$: The x terms cancel, leaving $\frac{a_n}{b_m}$.
   - $n > m$: The limit explodes to $\pm\infty$.
   - $n < m$: The denominator overpowers the numerator, dragging the limit to 0.

Moral.

It's not worth remembering concrete results or formulaic tables. Remember the *way*, not the destination! The technique of factoring out the highest power is universally applicable.

Exercise 2.10.8. Evaluate

$$\lim_{x\to\infty} \frac{2x^2 + x - 5}{3x^2 - 2x + 1}.$$

Example 2.10.9. Consider the competition between powers of x:

$$\lim_{x\to\infty} (x^2 - x) = \lim_{x\to\infty} x^2 \left(1 - \frac{1}{x}\right) = \infty \cdot 1 = \infty.$$

Conversely, the reversed battle $\lim_{x\to\infty}(x - x^2) = -\infty$. *In polynomials, the highest degree holds absolute dominion.*

Example 2.10.10. Evaluate

$$\lim_{x\to\infty}\left(\sqrt{x^2+1}-x\right).$$

This limit is of the undetermined form $\infty-\infty$. The offending terms that generate those infinity are x^2 and x. We need to cancel them before taking limit. To bring them together, the square root must be removed, thus we multiply by the algebraic conjugate

$$\begin{aligned}\lim_{x\to\infty}\left(\sqrt{x^2+1}-x\right) &= \lim_{x\to\infty}\left(\sqrt{x^2+1}-x\right)\cdot\frac{\left(\sqrt{x^2+1}+x\right)}{\left(\sqrt{x^2+1}+x\right)}\\ &= \lim_{x\to\infty}\frac{x^2+1-x^2}{\sqrt{x^2+1}+x}\\ &= \lim_{x\to\infty}\frac{1}{\sqrt{x^2+1}+x}=\frac{1}{\infty}=0.\end{aligned}$$

Exercise 2.10.11. Evaluate

$$\lim_{t\to0^+}\left(\sqrt{\frac{1}{t^2}+1}-\frac{1}{t}\right).$$

Example 2.10.12. Evaluate

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x.$$

Recall the fundamental limit: $\lim_{t\to0}(1+t)^{1/t}=\mathrm{e}$. By setting $t=1/x$, we map the infinity problem back to the origin. As $x\to\infty$, $t\to0^+$, giving us $\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x=\mathrm{e}$.

Example 2.10.13 (The Gaussian)**.** In quantum mechanics, the ground state wave function of a harmonic oscillator takes the form of a Gaussian curve, $\psi(x)\propto\mathrm{e}^{-x^2/2}$, see Figure 2.16 ($\propto$ is the notation for *propotional*: we write $f\propto g$ if there is a constant $c\neq0$ such that $f=cg$). For the particle to be localized in space, it is a physical imperative that the probability amplitude vanishes infinitely far away. Let us verify this mathematically.

$$\lim_{x\to\infty}\mathrm{e}^{-\frac{x^2}{2}}=\lim_{x\to\infty}\frac{1}{\mathrm{e}^{\frac{x^2}{2}}}.$$

Let $t=x^2/2$. As $x\to\infty$, $t\to\infty$. The exponential e^t grows without bound, meaning our function is crushed down to zero:

$$\lim_{x\to\infty}\mathrm{e}^{-\frac{x^2}{2}}=\lim_{t\to\infty}\frac{1}{\mathrm{e}^t}=0.$$

Exercise 2.10.14. Evaluate

$$\lim_{x\to\infty}\frac{x+\cos^2x}{\sqrt{x^2+1}}\qquad\text{and}\qquad\lim_{x\to-\infty}\frac{x+\cos^2x}{\sqrt{x^2+1}}.$$

2.11 Asymptotes

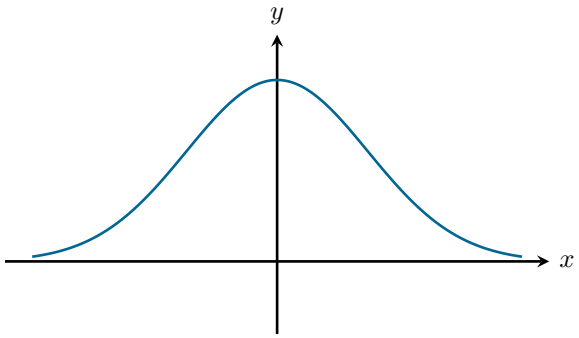


Figure 2.16 – Graph of the bell curve $e^{-x^2/2}$: the tails vanish at infinity.

Story. The Horizon of Mathematics

In physics, when you approach the event horizon of a black hole, the laws of classical mechanics begin to break down, stretching space and time into infinity. In mathematics, we have our own event horizons and infinite expanses. What happens to a function $f(x)$ when we push x to the absolute limits of its domain, or when we evaluate it at a point where its very definition fractures?
To answer this, we study *asymptotes*—the straight lines that our curved functions desperately wish to become as they march toward infinity or plunge into singularities.

In this section, we investigate *end behaviors* of a function—what it does in the asymptotic regimes. The behaviors are characterized by lines named *asymptotes*.

There are three types of asymptotes, cataloged by how limits behave at infinity or at finite points:

1. **Horizontal:** determined by *finite limits* at infinity. The function settles down to a constant value.
2. **Vertical:** determined by *infinite limits* at a finite point. The function experiences a catastrophic blow-up.
3. **Slant (or Oblique):** determined by *infinite limits* at infinity, but with a highly structured, linear growth rate.

Intuition. The Tao of Asymptotes

Before diving into the ε-δ or limit definitions, let us establish an intuition. An asymptote provides a *simple linear approximation* for a complex function in extreme regions. For $|x|$ very large, a function with a slant asymptote $y = ax + b$ is essentially saying, "I might be a complicated mess of roots, polynomials, and trigonometric functions near the origin, but out here, I am basically just a straight line."

2.11.1 Horizontal Asymptotes: The Ends of the Universe

We begin by asking what happens when we send our input variable out to the far right ($x \to \infty$) or the far left ($x \to -\infty$).

Definition 2.11.1 (Horizontal Asymptote)**.** Let f be a function defined on some interval (a, ∞). If

$$\lim_{x\to\infty} f(x) = R \tag{2.42}$$

where $R \in \mathbb{R}$, then we say the line $y = R$ is a horizontal asymptote of f. Similarly, if $\lim_{x\to-\infty} f(x) = L$ for some $L \in \mathbb{R}$, the line $y = L$ is a *horizontal asymptote* of f.

Figure 2.17 illustrates the concept of horizontal asymptote.

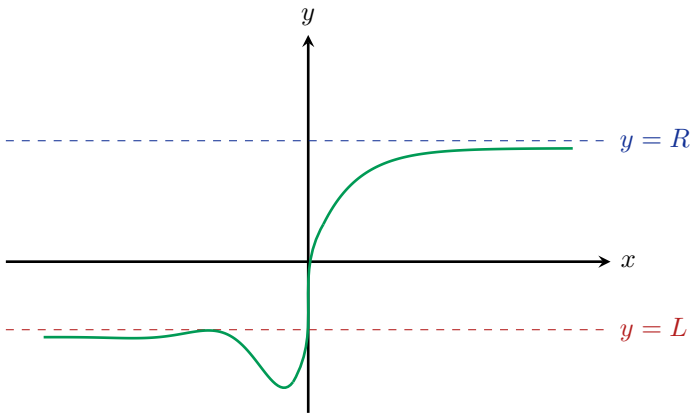


Figure 2.17 – Visualization of horizontal asymptotes, showing a function approaching $y = L$ as $x \to -\infty$ and $y = R$ as $x \to \infty$.

Example 2.11.2. Consider the function $f(x) = 2 + \frac{10}{x^2}$. Its domain is $\text{dom}(f) = (-\infty, 0) \cup (0, \infty)$. Taking the limit as $x \to \pm\infty$:

$$\lim_{x\to\infty} f(x) = 2$$
$$\lim_{x\to-\infty} f(x) = 2$$

Thus, f has exactly one horizontal asymptote: the line $y = 2$.

Example 2.11.3. Now consider something more oscillatory: $f(x) = 1 + \frac{\sin x}{\sqrt{x}}$. The domain is $\text{dom}(f) = (0, \infty)$, meaning the limit as $x \to -\infty$ does not exist. To find the limit as $x \to \infty$, we bound the oscillating numerator:

$$-\frac{1}{\sqrt{x}} \leq \frac{\sin x}{\sqrt{x}} \leq \frac{1}{\sqrt{x}}.$$

Since $\pm 1/\sqrt{x} \to 0$ as $x \to \infty$, by the Squeeze Theorem, $\lim_{x\to\infty} f(x) = 1$. Thus, $y = 1$ is the only horizontal asymptote.

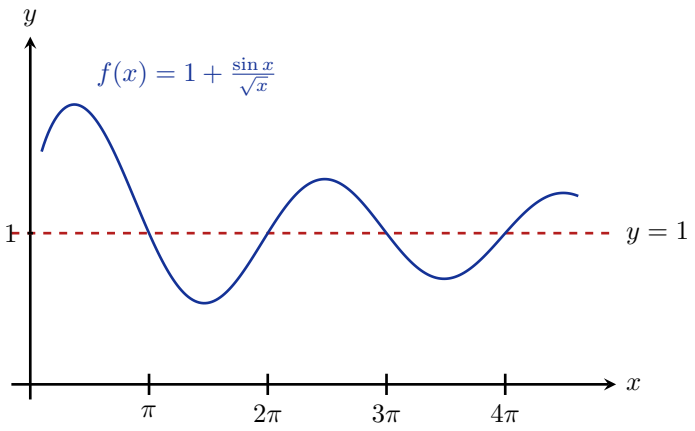


Figure 2.18 – Graph demonstrating a function repeatedly intersecting its horizontal asymptote $y = 1$.

> ⚠ *Caveat.*
>
> A common myth perpetuated in high school calculus is that "a function cannot cross its asymptote." The previous example completely shatters this notion! The function $f(x) = 1 + \frac{\sin x}{\sqrt{x}}$ crosses the line $y = 1$ infinitely many times as it decays, see Figure 2.18. An asymptote dictates *limit* behavior, not intersection properties.

Example 2.11.4. Consider $f(x) = \frac{x}{\sqrt{x^2+1}}$. Let us compute the limits at both infinities. As $x \to \infty$, the dominant term in the denominator is $\sqrt{x^2} = |x| = x$. Thus, the limit is 1. However, as $x \to -\infty$, $\sqrt{x^2} = |x| = -x$, yielding a limit of -1. Hence, there are two distinct horizontal asymptotes: $y = 1$ and $y = -1$.

Exercise 2.11.5 (The Asymptote Counting Problem)**.** We have just seen functions with 1 and 2 horizontal asymptotes. We can easily construct a function with no horizontal asymptotes (e.g., $f(x) = x$).

- Can you invent a function with 0 horizontal asymptotes, 1 horizontal asymptote, and 2 horizontal asymptotes? *Hint: use piecewise functions.*
- **Challenge:** Can a function $f : \mathbb{R} \to \mathbb{R}$ have 3 horizontal asymptotes? *Hint: Re-read the definition. How many "infinities" does the real line have?*

2.11.2 Vertical Asymptotes: Singularities

If horizontal asymptotes describe finite limits at infinity, vertical asymptotes describe *infinite* limits at a finite point. Physically, this is a singularity—a point where forces become infinite, densities diverge, and our equations demand we look closer.

Definition 2.11.6 (Vertical Asymptote)**.** If either one-sided limit diverges:

$$\lim_{x \to a^+} f(x) = \pm\infty \quad \text{or} \quad \lim_{x \to a^-} f(x) = \pm\infty, \tag{2.43}$$

then we say the vertical line $x = a$ is a *vertical asymptote* of f.

Figure 2.19 illustrates the concept of vertical asymptote.

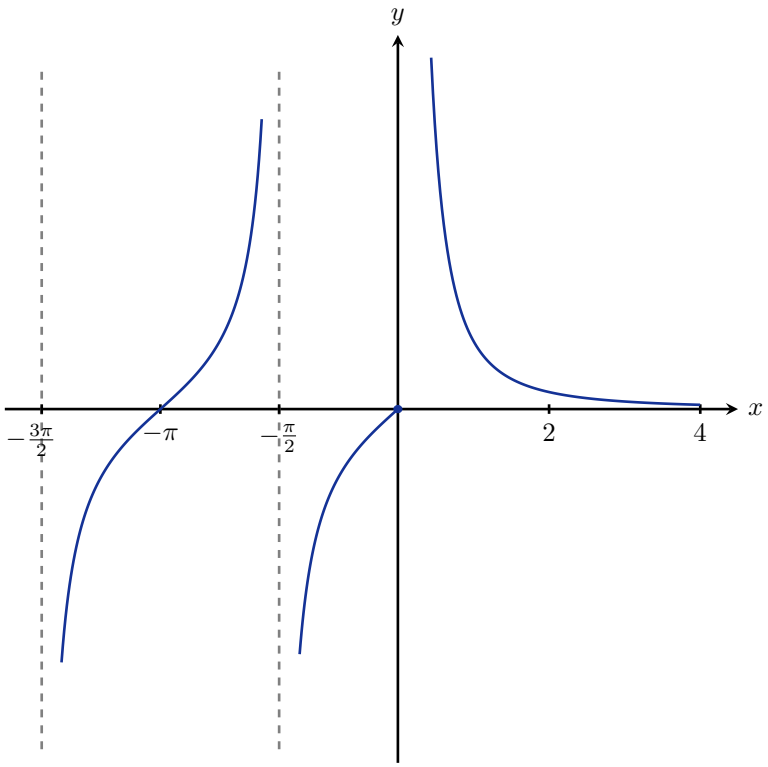


Figure 2.19 – Graph of a function with multiple vertical asymptotes.

Example 2.11.7. Let $f(x) = \frac{1}{x-1}$. It is immediately apparent that as $x \to 1^+$, $f(x) \to \infty$. Thus, $x = 1$ is a vertical asymptote.

Example 2.11.8. Consider $f(x) = \frac{\sin x}{x^2}$. As $x \to 0$, the numerator approaches 0 and the denominator approaches 0: we run into the undetermined form $0/0$. However, using $\lim_{x\to 0} \frac{\sin x}{x} = 1$, we see $\lim_{x\to 0^\pm} f(x) = \lim_{x\to 0^\pm} \frac{1}{x} = \pm\infty$. Thus, $x = 0$ is a vertical asymptote.

2.11.3 Slant (Oblique) Asymptotes: Linear Escapes

We now arrive at the most sophisticated of the three: the slant asymptote. What if a function escapes to infinity, but does so by mirroring a slanted straight line?

Definition 2.11.9 (Slant Asymptote)**.** If there exist real numbers a and b (with $a \neq 0$) such that

$$\lim_{x\to\infty} \big(f(x) - (ax+b)\big) = 0 \quad \text{or} \quad \lim_{x\to-\infty} \big(f(x) - (ax+b)\big) = 0, \tag{2.44}$$

then we say the line $y = ax + b$ is a *slant (or oblique) asymptote* of f.

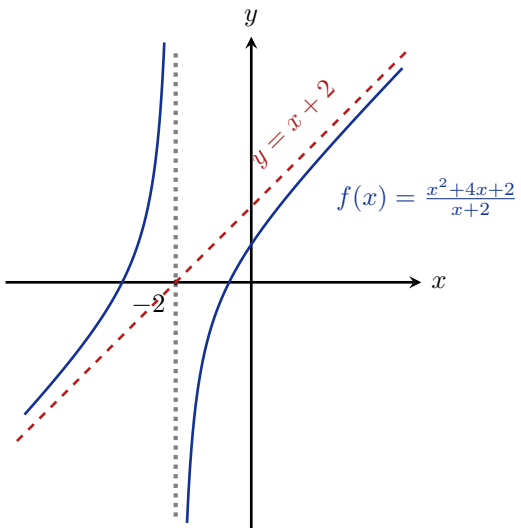


Figure 2.20 – Graph of a function and its slant asymptote and verical asymptote.

Figure 2.20 illustrates the concept of slant asymptote.

Remark 2.11.10. A horizontal asymptote is merely a special case of a slant asymptote where the slope $a = 0$. By structuring our definition this way, we unify the two concepts.

From the definition, we can extract formulas for a and b. If $f(x) - (ax + b) \to 0$, then dividing by x implies that $f(x)/x \to a$. Once a is known, we can cleanly extract b:

$$a = \lim_{x\to\infty} \frac{f(x)}{x}, \qquad b = \lim_{x\to\infty} \big(f(x) - ax\big). \tag{2.45}$$

(Similar formulas hold for $x \to -\infty$). Let us test this on a non-trivial function.

Example 2.11.11. Find the slant asymptote(s) of the function:

$$f(x) = \frac{x^2 + 2x + \sqrt{x}}{x + 1} \tag{2.46}$$

First, a quick check of the domain reveals $\mathrm{dom}(f) = [0, \infty)$. Therefore, we only need to find the limit as $x \to \infty$; the limit as $x \to -\infty$ does not exist, meaning there are no asymptotes to the left.

We will solve this using two distinct methods.

1. *Method 1: Algebraic Manipulation (Long Division)* We can rewrite the numerator to cleverly force cancellation with the denominator:

$$\begin{aligned} f(x) &= \frac{x^2 + 2x + 1 - 1 + \sqrt{x}}{x + 1} \\ &= \frac{(x + 1)^2 - 1 + \sqrt{x}}{x + 1} \\ &= x + 1 + \frac{\sqrt{x} - 1}{x + 1}. \end{aligned}$$

Now, examine the residual term as $x \to \infty$:

$$\lim_{x\to\infty} \big(f(x) - (x+1)\big) = \lim_{x\to\infty} \frac{\sqrt{x}-1}{x+1}.$$

To evaluate this rigorously, use the substitution $t = \sqrt{x}$. As $x \to \infty$, $t \to \infty$. Thus:

$$\lim_{x\to\infty} \big(f(x) - (x+1)\big) = \lim_{t\to\infty} \frac{t-1}{t^2+1} = 0.$$

Because the difference between $f(x)$ and the line $y = x+1$ vanishes at infinity, the slant asymptote is precisely $\ell(x) = x+1$.

2. *Method 2: The Analytic Limit Formulas* Using the formulas derived in (2.45), we first find the slope a:

$$a = \lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} \frac{x^2+2x+\sqrt{x}}{x(x+1)} = \lim_{x\to\infty} \frac{x^2(1+\frac{2}{x}+\frac{1}{x^{3/2}})}{x^2(1+\frac{1}{x})} = 1.$$

Next, we find the intercept b:

$$\begin{aligned} b = \lim_{x\to\infty} \big(f(x) - 1\cdot x\big) &= \lim_{x\to\infty} \left(\frac{x^2+2x+\sqrt{x}}{x+1} - x\right) \\ &= \lim_{x\to\infty} \left(\frac{x^2+2x+\sqrt{x}-x(x+1)}{x+1}\right) \\ &= \lim_{x\to\infty} \frac{x+\sqrt{x}}{x+1}. \end{aligned}$$

Applying the same substitution $t = \sqrt{x}$ (or dividing top and bottom by x) yields:

$$\lim_{t\to\infty} \frac{t^2+t}{t^2+1} = 1.$$

Thus, the slant asymptote is $\ell(x) = 1\cdot x + 1 = x+1$.

> *Moral.*
>
> Albeit it seems that Method 2 is longer and more tedious than Method 1, it is actually vastly more powerful. Method 1 (polynomial division) is merely an algebraic trick that works beautifully for rational functions. However, if transcendental functions ($e^x, \sin x$) are introduced, Method 1 completely fails, whereas the analytic limits of Method 2 hold universally.

Why do we labor over these limit calculations? Slant (and horizontal) asymptotes provide a *simple linear approximation* for a function at infinity. For very large x, computing $f(x) \approx ax + b$ is computationally cheap and analytically transparent. This is not merely an academic exercise; it is good for numerical calculation when simulating real world phenomena.

To solidify your grasp on these concepts, I leave you with a few exercises to explore the end behaviors of varied functions.

Exercise 2.11.12. Find all asymptotes (if any) of the following functions:

(a) $f(x) = x\sqrt{x^2+1}\sin\left(\frac{1}{x}\right)$ *(Hint: Recall* $\lim_{\theta\to 0}\frac{\sin\theta}{\theta} = 1$*.)*

(b) $f(x) = \frac{10x^3-3x^2+8}{\sqrt{25x^6+x^4+2}}$

(c) $f(x) = \frac{x^2+2}{\sqrt{x^2+1}}$

Exercise 2.11.13. Determine the complete end behaviors of the fundamental transcendental functions: e^x, $\log x$, and $\cos x$, strictly within their respective domains.

Exercise 2.11.14. Can you find a quadratic function $g(x) = ax^2+bx+c$ that approximates $f(x) = \frac{(x^3+3x^2+1)(1+e^{-x})}{\sqrt{x^2-1}}$ for large x? That is

$$\lim_{x\to\infty}(f(x)-g(x)) = 0.$$

2.12 The Intermediate Value Theorem

Daily life is replete with continuous processes. If you want to drive your car from a valley below sea level to a mountain peak, your altimeter cannot simply jump from a negative elevation to a positive one. At some precise moment during your journey, you must have been exactly at sea level. You cannot teleport.

In geometry, this manifests as a remarkably intuitive rule. Given two distinct points lying on different sides of a line, *any* attempt connecting the two points must cross the line, see Figure 2.21.

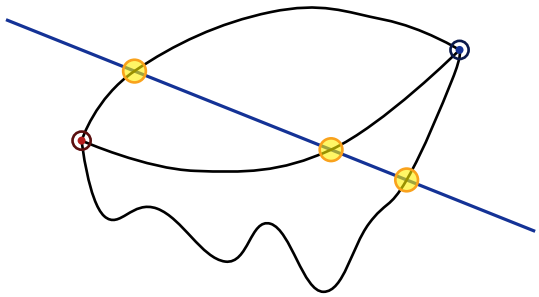


Figure 2.21 – Any continuous path connecting a point on one side of a dividing line to a point on the other side must intersect the dividing line.

> *Intuition. Translating Geometry to Calculus*
>
> We state this geometric intuition in calculus language by imagining our path as the graph of a continuous function $f(x)$. The dividing line becomes a horizontal line $y = y_0$. If our function starts below this line and ends above it, and it draws a continuous curve without picking up the pen, it is absolutely forced to intersect y_0.

To elevate this from a physical observation to a rigorous mathematical truth, we state it formally:

Theorem 2.12.1 (Intermediate Value Theorem (IVT))**.** *Let $a < b$, and let $f : [a, b] \to \mathbb{R}$ be a continuous function on the closed interval $[a, b]$. Let y_0 be a real number strictly between $f(a)$ and $f(b)$; that is, either*

$$f(a) \leq y_0 \leq f(b) \quad \text{or} \quad f(a) \geq y_0 \geq f(b).$$

Then there exists $c \in [a, b]$ such that $f(c) = y_0$.

⚠ *Caveat. An Existence Theorem*

The Intermediate Value Theorem is a classic example of an *existence theorem*. It tells you with absolute certainty that at least one number c exists, but it offers precisely zero algorithmic guidance on how to actually find it! It does not provide a formula for c.

Furthermore, we must be careful not to overstate what the theorem promises. Notice that there may be *several* values of c that satisfy $f(c) = y_0$. The theorem guarantees *at least one*, not *exactly one*, see Figure 2.22.

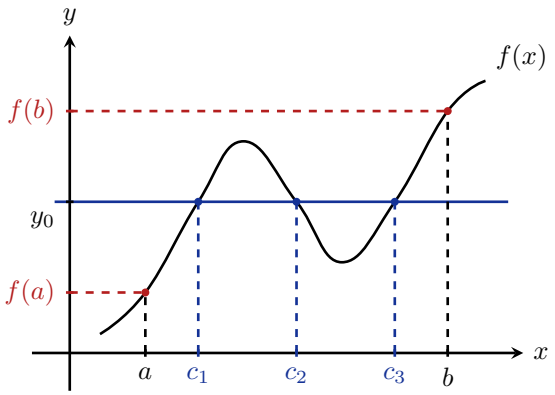


Figure 2.22 – Illustration of the Intermediate Value Theorem on $[a, b]$, showing three intermediate points c_1, c_2, c_3 where $f(c_i) = y_0$.

It is worth pausing to ask an important question: *Why continuity?*

The IVT dictates that if f takes the values $f(a)$ and $f(b)$, then it must also take all the values in between. If f is not assumed to be continuous, this logical scaffolding collapses entirely.

Consider the piecewise function $f : [-1, 1] \to \mathbb{R}$ defined by:

$$f(x) \coloneqq \begin{cases} -1 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases}$$

Here, we clearly have $f(-1) = -1$ and $f(1) = 1$. The value $y_0 = 0$ sits squarely between $f(-1)$ and $f(1)$. Yet, there is no $c \in [-1, 1]$ for which $f(c) = 0$.

Moral. No Teleportation Allowed

If a function is discontinuous, it is entirely capable of jumping past intermediate values. Continuous functions, bound by their $\varepsilon - \delta$ topological chains, cannot do so.

We now arrive at a profound meta-mathematical realization. Any student encountering this theorem for the first time might think: *Obviously a line from bottom to top hits the middle. Why do mathematicians obsess over proving this?*

In fact, any attempt to prove the IVT eventually boils down to the very definition of what the set of real numbers ($\mathbb{R}$) actually is.

Story. The Fabric of the Reals

It is somewhat funny that a result with such trivial statements seems to have a deep and difficult proof. The caveat is that there are several ways to define a real number (Dedekind cuts, equivalence classes of Cauchy sequences, etc.), but any rigorous proof of the IVT eventually converges to (yet another) theorem saying that $\mathbb{R}$ is *complete*. That is, $\mathbb{R}$ is *rich* enough to guarantee the convergence of a special kind of sequence: the Cauchy sequence. If we were doing calculus on the rational numbers ($\mathbb{Q}$), the IVT would fail spectacularly! The continuous function $f(x) = x^2 - 2$ goes from negative to positive on $[0, 2]$, but it never hits 0 in $\mathbb{Q}$ because $\sqrt{2}$ slips right through the *gaps* in the rational number line. The IVT is fundamentally a statement about the absence of microscopic holes in $\mathbb{R}$.

The true power of the IVT lies in its ability to prove the existence of solutions to horrific equations without requiring us to actually solve them.

Example 2.12.2. Show that the equation $x^5 - x + 1 = 0$ has at least 1 solution in $[-2, 0]$.

Proof. Let $f(x) = x^5 - x + 1$. Since f is a polynomial, we know from elementary analysis that f is continuous (cts) on the entirety of $\mathbb{R}$. Therefore, it is certainly continuous on the restricted interval $[-2, 0]$. We now evaluate the function at the boundary points of our interval:

$$\begin{aligned} f(0) &= 0^5 - 0 + 1 = 1 \\ f(-2) &= (-2)^5 - (-2) + 1 \\ &= -32 + 3 = -29 < 0 \end{aligned}$$

We observe that $f(-2) < 0$ and $f(0) > 0$. The value $y_0 = 0$ lies strictly between $f(-2)$ and $f(0)$. Therefore, by the Intermediate Value Theorem, there exists at least one $c \in [-2, 0]$ such that $f(c) = 0$. □

Remark 2.12.3. Notice the elegance here: we don't even need to calculate the exact values of $f(a)$ and $f(b)$ if the arithmetic is tedious. We merely need to establish their *signs*.

We leave the reader with a slightly more challenging problem that combines the IVT with the asymptotic behavior of functions.

Exercise 2.12.4 (Challenge for the curious minds)**.** Given two real numbers a and b with $b \geq 0$. Show that the equation

$$\frac{x^5 - ax + b}{x\sqrt{x^2+1}} = 2$$

has at least one real solution.

Hint: Do not attempt to solve for x algebraically! Instead, define $g(x) = \frac{x^5-ax+b}{x\sqrt{x^2+1}}$. Find the limits

$$\lim_{x\to\pm\infty} (g(x) - 2) \qquad \text{and} \qquad \lim_{x\to 0^{\pm}} (g(x) - 2)\,.$$

Then use the rigorous $\varepsilon - \delta$ definition of limits at infinity to construct a sufficiently large, finite closed interval $[-M, M]$ upon which you can apply the IVT.

Chapter 3

Derivative

Nature and Nature's laws lay hid in night:
God said, *Let Newton be!* and all was light.

Alexander Pope, *Epigram on Sir Isaac Newton*[1]

3.1 Introduction to the Derivative

3.1.1 From Secants to Tangents

We recall the Tangent Line Problem back in Section 2.2.1. Consider a sufficiently well-behaved function $f\colon \mathbb{R} \to \mathbb{R}$. We wish to determine the slope of its graph at a precise point $P = (c, f(c))$.

Because it takes two points to define a line, we artificially construct a second point nearby: $Q = (c + \Delta x, f(c + \Delta x))$. The line passing through P and Q is termed a *secant line*. Its slope is given by the standard "rise over run":

$$m_{\text{sec}} = \frac{\text{rise}}{\text{run}} = \frac{f(c + \Delta x) - f(c)}{(c + \Delta x) - c} = \frac{f(c + \Delta x) - f(c)}{\Delta x}. \tag{3.1}$$

Intuition. Zooming In

Imagine looking at a smooth curve under a microscope. As you increase the magnification around the point P, the curve appears flatter and flatter. This local linearity is the essence of differential calculus. By letting $\Delta x \to 0$, the point Q slides along the curve toward P. The secant line gradually pivots and converges to the ultimate geometric object of our desire: the *tangent line*.

[1] This statement did not stand for long. In the 20th century, J.C. Squire wrote an answer to Pope:

It did not last: the Devil howling *Ho!*
Let Einstein be! restored the status quo.

We invoke the magic of limits. As Δx vanishes, the secant slope gracefully approximates the tangent slope:

$$m_{\tan} = \lim_{\Delta x \to 0} m_{\sec} = \lim_{\Delta x \to 0} \frac{f(c+\Delta x) - f(c)}{\Delta x}. \tag{3.2}$$

One we know the tangent slope $m_{\tan}$, finding the equation of the tangent line to the graph of f at the point $(c, f(c))$ is a trivial exercise in algebra. We simply use the point-slope form:

$$y - f(c) = m_{\tan}(x - c). \tag{3.3}$$

> *Moral.*
>
> The limit of the difference quotient in (3.1) has many physical meanings and applications that go well beyond the initial geometric intuition of a tangent line. It represents the *instantaneous rate of change* of any given quantity: velocity, marginal cost, or the flow of heat, etc. This versatility is what motivated Definition 3.1.1.

Definition 3.1.1 (The Derivative at a Point)**.** Let $f\colon \mathbb{R} \to \mathbb{R}$ be a function and c be a real number. If the limit

$$\lim_{\Delta x \to 0} \frac{f(c+\Delta x) - f(c)}{\Delta x} \tag{3.4}$$

converges to some real number L, we say f is *differentiable* at c with *derivative* L. We denote this by

$$f'(c) = L.$$

If the limit does not exist—including the case where c is not even in $\mathrm{dom}(f)$—then we leave $f'(c)$ undefined and say f is *not differentiable* at c. Furthermore, if f is differentiable at any point in an open interval I, we say f is *differentiable on* I.

Remark 3.1.2 (On Notation). Calculus suffers from an abundance of historical notation. For the derivative, one might encounter:

$$f'(c), \quad \frac{\mathrm{d}f}{\mathrm{d}x}(c), \quad D_x f(c), \quad \dot{f}(c), \quad \dots$$

Each has its pros and cons depending on the tradition of the subject. The clearest notation, completely free of ambiguities, would technically require the concept of a *directional derivative*, but we shall leave that for a future chapter on multivariable mathematics. For our immediate needs, the Lagrange notation f' and the Leibniz notation $\frac{\mathrm{d}f}{\mathrm{d}x}$ will serve us well.

Remark 3.1.3. There are several equivalent ways to write the limit (3.4) in Definition 3.1.1. For example, if we use the substitution $h = \Delta x$, then we can write the derivative as:

$$f'(c) = \lim_{h \to 0} \frac{f(c+h) - f(c)}{h}, \tag{3.5}$$

provided the limit exists. Alternatively, if we let the moving point $x = c + \Delta x$, we can recast the formulation as:

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}, \tag{3.6}$$

which sometimes offers algebraic simplifications depending on the function.

Let us compute the derivatives of a few basic functions directly from the definition.

Example 3.1.4 (The Constant Function)**.** Let $f(x) = k$, where k is a constant. Let c be any real number.

$$\lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x} = \lim_{\Delta x \to 0} \frac{k - k}{\Delta x} = 0.$$

So, $f'(c) = 0$ for any real number c. Geometrically, this is obvious: a horizontal line has zero slope.

Example 3.1.5 (The Identity Function)**.** Let $f(x) = x$, and let c be a real number.

$$\lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x} = \lim_{\Delta x \to 0} \frac{(c + \Delta x) - c}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta x}{\Delta x} = 1.$$

So, $f'(c) = 1$ for any real number c. Again, perfectly aligned with our understanding that the line $y = x$ has a slope of 1.

Now let us increase the complexity slightly and observe the algebra required to resolve the indeterminate form $0/0$.

Example 3.1.6 (The Quadratic Function)**.** Let $f(x) = x^2$, and let c be a real number. We proceed via the limit definition:

$$\begin{aligned}
\lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x} &= \lim_{\Delta x \to 0} \frac{(c + \Delta x)^2 - c^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(c^2 + 2c\Delta x + (\Delta x)^2) - c^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\Delta x(2c + \Delta x)}{\Delta x}.
\end{aligned}$$

Provided $\Delta x \neq 0$ (which is strictly true under the limit operation), we can cancel the Δx terms:

$$\lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x} = \lim_{\Delta x \to 0} (2c + \Delta x) = 2c.$$

So, $f'(c) = 2c$.

Application. Let us find the tangent line to the graph of $f(x) = x^2$ at the coordinate $(2, 4)$.

$$m_{\tan} = f'(2) = 2(2) = 4.$$

Using the point-slope form (3.3):

$$\begin{aligned}
y - f(2) &= m_{\tan}(x - 2) \\
y - 2^2 &= 4(x - 2) \\
y &= 4x - 8 + 4 \\
y &= 4x - 4.
\end{aligned}$$

3.1.2 The Derivative as a Function

In the previous examples, given a real number c, if f is differentiable at c, we mapped the value $c \mapsto f'(c)$. Thus, the derivative f' can itself be viewed as a function from $\mathbb{R}$ to $\mathbb{R}$.

Since it is a function in its own right, we usually revert to the standard independent variable x. For instance, in Example 3.1.6 where $f(x) = x^2$, we write:

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

Thus, both f and f' are functions of x mapping from $\mathbb{R} \to \mathbb{R}$. This raises a subtle question:

$$\textit{Do we always have } \operatorname{dom}(f') = \operatorname{dom}(f)\textit{?}$$

> ⚠ *Caveat. The Shrinking Domain*
>
> The answer is a resounding **No**. The operation of differentiation is delicate; it requires a certain *smoothness* that the original function may lack. As a result, taking a derivative can brutally shear away points from the function's domain. In general, we only have the subset relation $\operatorname{dom}(f') \subseteq \operatorname{dom}(f)$.

Here is a counterexample. Consider the absolute value function $f(x) = |x|$. Obviously, $\operatorname{dom}(f) = \mathbb{R}$. Let us rewrite $f(x)$ as a piecewise function:

$$f(x) = \begin{cases} x & \text{if } x \in [0, \infty) \\ -x & \text{if } x \in (-\infty, 0). \end{cases}$$

The calculation in Example 3.1.5 shows that $f'(x) = 1$ on $(0, \infty)$ and $f'(x) = -1$ on $(-\infty, 0)$ (notice the open intervals here; why do we use open intervals but not closed intervals?) this tells us that at the very least, we have:

$$\operatorname{dom}(f') \supseteq (-\infty, 0) \cup (0, \infty).$$

But what happens at $x = 0$? Let us check rigorously. We must evaluate:

$$\lim_{\Delta x \to 0} \frac{f(0 + \Delta x) - f(0)}{\Delta x} = \lim_{\Delta x \to 0} \frac{|\Delta x|}{\Delta x}.$$

At this point, it is clear that we need to know the sign of Δx to evaluate the absolute value. So we must split the limit into one-sided limits:

$$\lim_{\Delta x \to 0^+} \frac{|\Delta x|}{\Delta x} = \lim_{\Delta x \to 0^+} \frac{\Delta x}{\Delta x} = 1,$$

$$\lim_{\Delta x \to 0^-} \frac{|\Delta x|}{\Delta x} = \lim_{\Delta x \to 0^-} \frac{-\Delta x}{\Delta x} = -1.$$

Because the right-hand limit (1) and the left-hand limit (-1) disagree, the two-sided limit $\lim_{\Delta x \to 0} \frac{|\Delta x|}{\Delta x}$ does not exist.

Consequently, f is not differentiable at 0. We conclude that

$$\operatorname{dom}(f') = (-\infty, 0) \cup (0, \infty), \tag{3.7}$$

which is a *strict subset* of $\operatorname{dom}(f)$.

Intuition.

The graph of $f(x) = |x|$ has a kink at $x = 0$, which prevents the function to be differentiable at 0.

Exercise 3.1.7 (On openness of the interval in the definition of derivative)**.** When we define differentiability of a function f on an interval I in Definition 3.1.1, we require I is open (i.e., I is of the form (a, b)). Why?

Exercise 3.1.8. Can we ever have $\text{dom}(f)$ be a strict subset of $\text{dom}(f')$?

Exercise for the reader: Ponder this question. By definition, can one evaluate a difference quotient at c if $f(c)$ does not even exist? Those questions about domains will come back when you study power series in Calculus II.

We conclude this section with a few exercises to build your computational muscle and geometric intuition.

Exercise 3.1.9. Find the derivative of $f(x) = \frac{2}{x}$.

Exercise 3.1.10. Find the tangent line at $x = 1$ of the function $f(x) = \sqrt{x}$.

Exercise 3.1.11. Suppose f and g are two differentiable functions on $\mathbb{R}$ and $f'(x) = g'(x)$ for all real number x. Can we conclude that $f(x) = g(x)$ for all real number x?

3.2 The Anatomy of Non-Differentiability

Let us build our intuition by examining the distinct ways where differentiability can break down.

3.2.1 The Corner

We redo the example above using a slightly different calculation. Consider the absolute value function, $f(x) = |x|$, defined for all $x \in (-\infty, \infty)$. Thus, $\text{dom}(f) = \mathbb{R}$. Let c be a real number. We calculate the derivative at c using the limit definition:

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} = \lim_{x \to c} \frac{|x| - |c|}{x - c}.$$

By definition, $|c|$ evaluates to c when $c \geq 0$, and $-c$ when $c < 0$. We split into three cases:

- *Case 1:* $c > 0$. Then for x close to c, $x > 0$, so $|x| = x$. The limit becomes

$$\lim_{x \to c} \frac{x - c}{x - c} = 1.$$

- *Case 2:* $c < 0$. Then for x close to c, $x < 0$, so $|x| = -x$. The limit becomes

$$\lim_{x \to c} \frac{-x - (-c)}{x - c} = -1.$$

- *Case 3:* $c = 0$. This is the critical point. We are evaluating:

$$\lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} \frac{|x| - |0|}{x - 0} = \lim_{x\to 0} \frac{|x|}{x}.$$

Due to the absolute value, we must split this into one-sided limits:

$$\lim_{x\to 0^+} \frac{|x|}{x} = \lim_{x\to 0^+} \frac{x}{x} = 1,$$
$$\lim_{x\to 0^-} \frac{|x|}{x} = \lim_{x\to 0^-} \frac{-x}{x} = -1.$$

Since the left and right limits disagree ($1 \neq -1$), we deduce that the limit $\lim_{x\to 0} \frac{f(x)-f(0)}{x-0}$ does not exist. Thus, f is not differentiable at 0.

> *Intuition. The Geometry of a Corner*
>
> Geometrically, $f(x) = |x|$ has a sharp *corner* at the origin, see Figure 3.1. A derivative requires the function to look locally like a single straight line when zoomed in. At $x = 0$, no matter how much you zoom in, the V-shape persists; it refuses to flatten out into a tangent line.

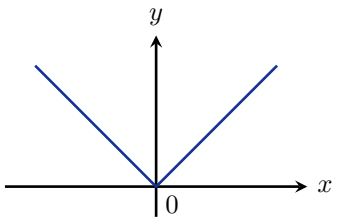


Figure 3.1 – The graph of $f(x) = |x|$ has a corner at $x = 0$.

Exercise 3.2.1. Pay attention to the divide-and-conquer argument above: why do we need 3 different cases for c? Can you use this insight to answer Exercise 3.1.7?

3.2.2 The Vertical Tangent

Consider the function $f(x) = x^{1/3}$. The domain of f is $(-\infty, \infty)$. We calculate the difference quotient:

$$\lim_{h\to 0} \frac{f(x + h) - f(x)}{h} = \lim_{h\to 0} \frac{(x + h)^{1/3} - x^{1/3}}{h}.$$

To resolve this 0/0 indeterminate form, we employ a clever algebraic trick using the difference of cubes identity:

$$(a - b)(a^2 + ab + b^2) = a^3 - b^3.$$

By multiplying by the conjugate-like factor $\left((x+h)^{2/3}+(x+h)^{1/3}x^{1/3}+x^{2/3}\right)$, we obtain:

$$\begin{aligned}
\lim_{h\to 0}\frac{(x+h)^{1/3}-x^{1/3}}{h} &= \lim_{h\to 0}\frac{\left[(x+h)^{1/3}-x^{1/3}\right]\left((x+h)^{2/3}+(x+h)^{1/3}x^{1/3}+x^{2/3}\right)}{h\left((x+h)^{2/3}+(x+h)^{1/3}x^{1/3}+x^{2/3}\right)}\\
&= \lim_{h\to 0}\frac{((x+h)^{1/3})^3-(x^{1/3})^3}{h\left((x+h)^{2/3}+(x+h)^{1/3}x^{1/3}+x^{2/3}\right)}\\
&= \lim_{h\to 0}\frac{(x+h)-x}{h\left((x+h)^{2/3}+(x+h)^{1/3}x^{1/3}+x^{2/3}\right)}\\
&= \lim_{h\to 0}\frac{1}{(x+h)^{2/3}+(x+h)^{1/3}x^{1/3}+x^{2/3}}\\
&= \frac{1}{x^{2/3}+x^{1/3}x^{1/3}+x^{2/3}}\\
&= \frac{1}{3x^{2/3}}.
\end{aligned}$$

Therefore, $f'(x)=\frac{1}{3}x^{-2/3}$. However, observe what happens at $x=0$: the denominator vanishes. Thus, $\mathrm{dom}(f')=(-\infty,0)\cup(0,\infty)$. The curve possesses a *vertical tangent line* at the origin.

> *Intuition.*
>
> Observe what happens geometrically as $x\to 0$. If you sketch the graph of $f(x)=x^{1/3}$, you will notice a curious phenomenon in Figure 3.2: As the curve passes through the origin, it briefly becomes completely vertical. The tangent line is the y-axis itself! Thus, we expect the slope of the tangent line to be infinite, and accordingly, the derivative will fail to exist at $x=0$, giving us another elegant example where $\mathrm{dom}(f')\subsetneq\mathrm{dom}(f)$ (beside the kink at 0 of $f(x)=|x|$).

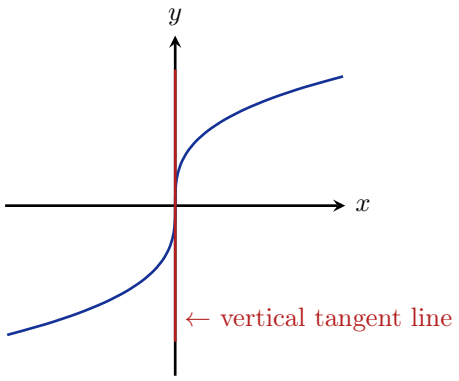


Figure 3.2 – The tangent line of the graph of $f(x)=x^{1/3}$ at $x=0$ is vertical.

3.2.3 The Discontinuity

Consider the step function:

$$f(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$$

On $(0, \infty)$, $f(x) = 1$, so $f'(x) = 0$. On $(-\infty, 0)$, $f(x) = -1$, so $f'(x) = 0$.

At $x = 0$, we calculate the limit directly:

$$\lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0} \frac{f(x) - 1}{x}.$$

Examining the one-sided limits, we have:

$$\lim_{x \to 0^+} \frac{f(x) - 1}{x} = \lim_{x \to 0^+} \frac{1 - 1}{x} = 0,$$
$$\lim_{x \to 0^-} \frac{f(x) - 1}{x} = \lim_{x \to 0^-} \frac{-1 - 1}{x} = \lim_{x \to 0^-} \frac{-2}{x} = \infty.$$

Since ∞ is not a real number, f is not differentiable at 0.

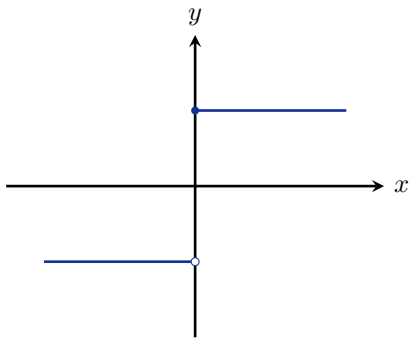


Figure 3.3 – The step function is discontinuous at $x = 0$.

From the examples above, we can isolate 3 primary scenarios where f fails to be differentiable at c:

1. f has a *corner* at c (e.g., $f(x) = |x|$).
2. f has a *vertical tangent line* at c (e.g., $f(x) = x^{1/3}$).
3. f is *not continuous* at c.

Story. The Monster in the Zoo of Functions

Looking at the graphs of our 3 examples above, they might give you a false sense of security. One might naturally form the *false belief* that a continuous function may only be non-differentiable at a few isolated points.
Lo and behold! Let me introduce the *Weierstrass monster function*. Discovered by Karl Weierstrass in the late 19th century, it is a function that is continuous *everywhere*

but differentiable *nowhere*. It is a fractal-like beast, composed of infinitely many wiggles upon wiggles, see Figure 3.4. Its mere existence shocked the mathematical establishment, proving that geometric intuition can sometimes lead us drastically astray!

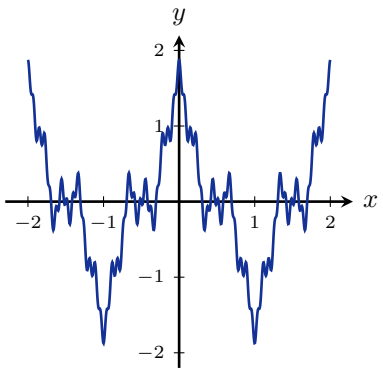


Figure 3.4 – The Weierstrass function $W(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x)$.

3.3 Differentiable Implies Continuous

Having explored how differentiability can spectacularly fail, let us now appreciate what differentiability *guarantees*.

Intuition. Newton's Approximation

Suppose f is differentiable at c. The geometric intuition is that at a tiny window about the point $(c, f(c))$, the graph of $f(x)$ behaves like its tangent line:

$$y = f(c) + m_{\tan}(x - c).$$

Equivalently,

$$f(x) \approx f(c) + f'(c)(x - c)$$

when x is close to c.

We can put this into formal $\varepsilon - \delta$ language.

Theorem 3.3.1 (Newton's Approximation)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ and $x_0 \in \mathrm{dom}(f)$. The following are equivalent:*

(i) *f is differentiable at x_0 with derivative L.*

(ii) *For each $\varepsilon > 0$, there exists $\delta = \delta(x_0, \varepsilon) > 0$ such that*

$$|f(x) - (f(x_0) + L(x - x_0))| < \varepsilon\,|x - x_0|$$

whenever $0 < |x - x_0| < \delta$.

Proof. The proof of this theorem is simply a rearrangement of the $\varepsilon - \delta$ definition of the limit:

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

□

This theorem tells us an interesting thing: if f is differentiable at x_0, then f can be well-approximated by a linear function $g(x) = f(x_0) + L(x - x_0)$ (the tangent line at x_0). But a linear function is fundamentally continuous. Do we have that f is also continuous at $x = x_0$?

The answer is yes:

Corollary 3.3.2. *Suppose f is differentiable at x_0. Then f is continuous at x_0.*

We offer two approaches to prove this fundamental fact.

Approach 1. This approach relies directly on algebraic limit laws. We compute the limit of the difference:

$$\begin{aligned}
\lim_{x \to x_0} (f(x) - f(x_0)) &= \lim_{x \to x_0} \left[\frac{f(x) - f(x_0)}{x - x_0} \cdot (x - x_0) \right] \\
&= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \cdot \lim_{x \to x_0} (x - x_0) \\
&= f'(x_0) \cdot 0 \\
&= 0.
\end{aligned}$$

This immediately yields $\lim_{x \to x_0} f(x) = f(x_0)$. Thus, f is continuous at x_0. □

Approach 2. By Newton's approximation, since the tangent line is a good approximation, the error term itself vanishes as $x \to x_0$. Thus, given $\varepsilon > 0$, whenever $|x - x_0| < \delta$ is sufficiently small, we can certainly bound the raw error by:

$$|f(x) - f(x_0) - L(x - x_0)| < \varepsilon |x - x_0|.$$

Using the triangle inequality, we can break apart the difference:

$$\begin{aligned}
|f(x) - f(x_0)| &= |(f(x) - f(x_0) - L(x - x_0)) + L(x - x_0)| \\
&\leq |f(x) - f(x_0) - L(x - x_0)| + |L| |x - x_0| \\
&< \varepsilon |x - x_0| + |L| |x - x_0|.
\end{aligned}$$

Taking the limit as $x \to x_0$,

$$\lim_{x \to x_0} |f(x) - f(x_0)| = 0.$$

Thus, f is continuous at x_0. □

> *Moral. The middle-man trick*
>
> Adding and subtracting an intermediate term is sometimes known as the *middle-man trick* and is a closed friend of any mathematician. The middle-man is very helpful when you need to build a bridge between several estimates: in the proof above, the

middle-man is the linear term $L(x - x_0)$.

So if we know a function is differentiable, it is automatically continuous. At first glance, this theorem seems a bit useless—finding $\lim_{x\to x_0} f(x)$ seems drastically easier than evaluating the thorny derivative limit $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$. Luckily, there are rules (like the Chain Rule and Product Rule) that allow us to tell immediately if a function is differentiable, bypassing the limit entirely and handing us continuity for free!

3.4 Differentiation Rules

Up until now, our toolkit for finding derivatives has been robust but primitive. We have relied entirely on the $\varepsilon - \delta$ definition. While technically sufficient, returning to the definition every time we encounter a new combination of functions is akin to building a house using only a spoon.

Recall from Chapter 1 that a complicated function is usually built from compositions, products, sums, and inverse relationships acting on old, familiar functions. In this section, we will observe how these arithmetic operations interact with differentiation. Our ultimate goal is to construct a *machinery of differentiation*—a set of rigorous, algebraic rules that can take the derivative of a horribly complicated function and systematically dismantle it into bite-sized pieces that we can easily differentiate.

3.4.1 Linearity: The Easiest Combinations

Let us start with the most straightforward arithmetic operations: addition, subtraction, and multiplication by a constant.

Theorem 3.4.1 (Derivative of sums, differences, and constant multiples)**.** *If f and g are differentiable at x, and c is any constant, then $f + g$, $f - g$, and cf are differentiable at x, and:*

$$\begin{aligned}(f+g)'(x) &= f'(x) + g'(x)\\ (f-g)'(x) &= f'(x) - g'(x)\\ (cf)'(x) &= cf'(x)\end{aligned}$$

Intuition. The Limit Under the Hood

Always remember that differentiation is, at its core, simply a limit taking place *under the hood*. Because the limit operator interacts so predictably with basic arithmetic operations, the derivative inherits these pleasant properties seamlessly.

Proof. The proofs of all three assertions are straightforward and are based on limit rules, specifically observations that relate the difference quotients of $f + g$, $f - g$, and cf to those of f and g. We also rely on the existence of the individual derivatives $f'(x)$ and $g'(x)$.

For the sum rule, we compute:

$$\begin{aligned}(f+g)'(x) &= \lim_{h\to 0} \frac{\big(f(x+h)+g(x+h)\big)-\big(f(x)+g(x)\big)}{h} \\ &= \lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h}+\frac{g(x+h)-g(x)}{h}\right) \\ &= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}+\lim_{h\to 0}\frac{g(x+h)-g(x)}{h} \\ &= f'(x)+g'(x).\end{aligned}$$

The proof that $(f-g)'(x) = f'(x) - g'(x)$ follows an identical structure. Finally, for the constant multiple:

$$\begin{aligned}(cf)'(x) &= \lim_{h\to 0}\frac{cf(x+h)-cf(x)}{h} \\ &= c\lim_{h\to 0}\frac{f(x+h)-f(x)}{h} \\ &= cf'(x).\end{aligned}$$

This completes the proof. □

Let us see this in action.

Example 3.4.2. Consider the function $h(x) = x^2+3$. We can decompose this into $f(x) = x^2$ and $g(x) = 3$. By the sum rule, Example 3.1.6, and Example 3.1.4, we find:

$$h'(x) = f'(x) + g'(x) = 2x + 0 = 2x.$$

Example 3.4.3. Similarly, if $h(x) = 2x$, we can let $f(x) = x$, meaning $h(x) = 2f(x)$. The constant multiple rule and Example 3.1.5 yield:

$$h'(x) = 2f'(x) = 2 \cdot 1 = 2.$$

Exercise 3.4.4. Do we have $(f_1 + f_2 + f_3)' = f_1' + f_2' + f_3'$?

Hint: Group $(f_1 + f_2)$ as a single function and apply the two-function sum rule.

Similarly, use this logic to convince yourself that the linearity principle holds for any finite sum of functions:

$$(f_1 + f_2 + \cdots + f_n)' = f_1' + f_2' + \cdots + f_n'.$$

3.4.2 The Product Rule

If limits play so nicely with sums, one might naively hope that the derivative of a product is simply the product of the derivatives. Let us disabuse ourselves of this notion immediately.

In our last example, suppose we redefine our constant as a function, letting $g(x) = 2$. Then our function $h(x) = 2x$ can be written as $h(x) = g(x)f(x)$. The result seems to say that:

$$h'(x) = 2f'(x) = g(x)f'(x).$$

But this clean formulation is merely an artifact of g being a constant! If g is a *bona-fide function* that changes as x varies, the fundamental symmetry of the expression $h = gf$

suggests something more profound: since g and f play the exact same role in the product, they must play the same role in the derivative. The symmetry suggests the appearance of an additional $g'(x)f(x)$ term. Indeed, we have the following.

Theorem 3.4.5 (Product Rule)**.** *If f and g are differentiable at x, then their product fg is differentiable at x, and*

$$(fg)'(x) = f(x)g'(x) + f'(x)g(x). \tag{3.8}$$

Story. Leibniz's Blunder

Even the giants of mathematics stumble. When Gottfried Wilhelm Leibniz first investigated the derivative of a product in 1672, he erroneously assumed that $(fg)'$ was simply $f'g'$. It wasn't until a few years later, after recognizing the internal inconsistencies of this assumption, that he discovered the correct formulation. Today, the product rule is affectionately known as the *Leibniz rule*.

To prove this, we must evaluate the difference quotient for the product. The algebraic hurdle is that both f and g are changing simultaneously. A classic problem-solving strategy in mathematics is to change only one variable at a time. This motivates the *middle-man trick*—adding and subtracting a mixed term—mentioned in Section 3.3.

Proof of Theorem 3.4.5. We evaluate the limit of the difference quotient:

$$(fg)'(x) = \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h}$$

We now introduce our middle-man by adding and subtracting $f(x+h)g(x)$ in the numerator:

$$\begin{aligned}
&= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x+h)g(x) + f(x+h)g(x) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} \left(f(x+h)\frac{g(x+h)-g(x)}{h} + \frac{f(x+h)-f(x)}{h} g(x) \right) \\
&= \lim_{h\to 0} f(x+h) \cdot \lim_{h\to 0} \frac{g(x+h)-g(x)}{h} + \left(\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \right) g(x) \\
&= f(x)g'(x) + f'(x)g(x).
\end{aligned}$$

□

Caveat. Differentiability Implies Continuity

In the last two steps of the proof, we distributed the limit and evaluated $\lim_{h\to 0} f(x+h) = f(x)$. This is legally permissible only because we assumed f is *differentiable* at x, which rigorously guarantees that f is also *continuous* at x.

You have seen the middle-man trick save the proof again! For those who prefer visual intuition, we have a geometric picture of them too: see Figure 3.5.

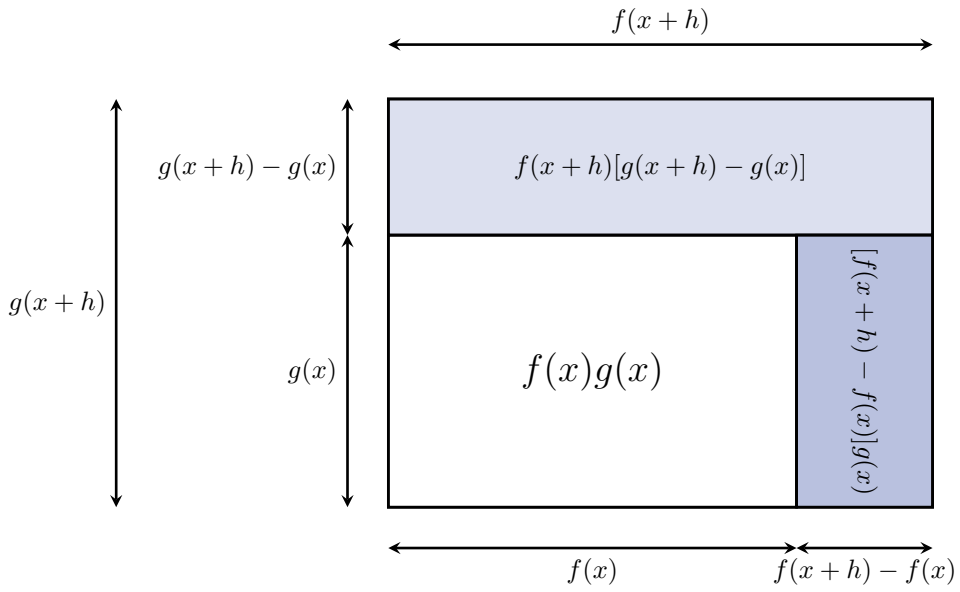


Figure 3.5 – A geometric manifestation of the middle-man trick. The total area of the large rectangle is $f(x+h)g(x+h)$, while the white area is $f(x)g(x)$. The difference, represented by the two shaded regions, perfectly maps to the algebraic decomposition in our proof.

3.4.3 The Power Rule: Climbing the Ladder

Having mastered the product rule, we are naturally led to ask: what happens when a function is multiplied by itself?

Example 3.4.6. Note that the function x^2 is a product. By the product rule:

$$(x^2)' = (x \cdot x)' = x \cdot x' + x' \cdot x = x + x = 2x.$$

Example 3.4.7. The function x^3 is also a product, so:

$$(x^3)' = (x^2 \cdot x)' = (x^2)' \cdot x + x^2 \cdot x' = (2x)x + x^2 = 2x^2 + x^2 = 3x^2.$$

A clear pattern emerges: the exponent "drops down" to become a coefficient, and the new exponent is reduced by one. This brings us to a sweeping generalization.

Proposition 3.4.8 (Power Rule)**.** *For every positive integer n,*

$$(x^n)' = nx^{n-1}. \tag{3.9}$$

> *Moral. The Domino Effect of Induction*
>
> Mathematical induction is a profoundly elegant method for proving that a statement $P(n)$ is true for all natural numbers. It is akin to climbing an infinitely tall ladder or knocking over an endless line of dominoes. To succeed, you must establish two things:
>
> 1. **The Base Case:** You can get onto the bottom rung of the ladder ($n = 1$).

2. **The Inductive Step:** From any given rung k, you are guaranteed to be able to step up to the next rung $k+1$.

Once both conditions are met, you have established the truth of the statement for all natural numbers.

Proof of Proposition 3.4.8. We use mathematical *mathematical induction*. We prove the inductive step: if the result holds for $n-1$, then it holds for n. Since the result fundamentally holds for $n = 1$, namely $(x)' = 1x^0$, its validity will then cascade for $n = 2, 3, \ldots$ This is precisely the type of step we made in Example 3.4.7, in going from knowledge of $(x^2)'$ to $(x^3)'$.

To prove the inductive step, we write $x^n = x^{n-1}x$ and apply the product rule:

$$(x^n)' = (x^{n-1}x)' = (x^{n-1})'x + x^{n-1}x'.$$

Using the assumed validity of the previous rung, $(x^{n-1})' = (n-1)x^{n-2}$, we get:

$$\begin{aligned}(x^n)' &= (x^{n-1})'x + x^{n-1}x' \\ &= (n-1)x^{n-2}x + x^{n-1} \\ &= (n-1)x^{n-1} + x^{n-1} \\ &= nx^{n-1}.\end{aligned}$$

The result holds for n, completing the proof. □

Armed with the rules of linearity and the power rule, we have effectively constructed a machine that can effortlessly devour any polynomial.

Example 3.4.9. Let p be a general polynomial function of the form:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Using the rules for differentiating a sum, a constant multiple, and the formula already verified for the derivative of x^n, we immediately obtain:

$$p'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1.$$

Example 3.4.10. If $p(x) = 2x^4 + 7x^3 + 6x^2 + 2x + 15$, then our newly minted machinery yields:

$$\begin{aligned}p'(x) &= 4(2)x^3 + 3(7)x^2 + 2(6)x + 1(2)x^0 + 0 \\ &= 8x^3 + 21x^2 + 12x + 2.\end{aligned}$$

Exercise 3.4.11. Find the derivative of $f(x) = x^2\sqrt[3]{x} - x^3 + 1$.

Exercise 3.4.12. Suppose f_1, f_2, and f_3 are differentiable on $\mathbb{R}$. Is $f_1 f_2 f_3$ differentiable? If so, what is the product rule for $(f_1 f_2 f_3)'$? Can you generalize this result to n functions $f_1, f_2, \ldots, f_n$?

3.5 The Quotient Rule

If f and g are differentiable at a point x, the familiar product rule guarantees us two beautiful facts:

1) fg is differentiable at x, and

2) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$.

The soulmate of the product is the quotient. One naturally asks: do we have $\frac{f}{g}$ is differentiable at x? Obviously, the first requirement is that we avoid division by zero: $g(x) \neq 0$.

Let's try to bring this seemingly difficult question back to familiar territory. Remember, a quotient is merely another kind of product in disguise. We can formally write

$$\frac{f}{g} = f \cdot \frac{1}{g}$$

and so, if we simply trust that the product rule will carry the day, we obtain

$$\left(\frac{f}{g}\right)' = \left(f \cdot \frac{1}{g}\right)' = f'\left(\frac{1}{g}\right) + f\left(\frac{1}{g}\right)'. \tag{3.10}$$

It becomes immediately clear that to answer the general question about $\frac{f}{g}$, one needs to start with the more specific case of $\frac{1}{g}$.

> *Intuition. Bootstrapping from the Known*
>
> Mathematics is a vast exercise in bootstrapping. When faced with a new object (the quotient), we decompose it into familiar atoms (the product) and a single new core mystery (the reciprocal).

3.5.1 Differentiating the Reciprocal

Let $h(x) = \frac{1}{g(x)}$. By the very definition of a reciprocal, this implies

$$h(x)g(x) = 1.$$

If we act boldly and differentiate both sides using product rule, we find:

$$(h(x)g(x))' = (1)' \implies h'(x)g(x) + h(x)g'(x) = 0.$$

Rearranging to isolate the unknown $h'(x)$, we arrive at:

$$\implies \left(\frac{1}{g}\right)' = h'(x) = -\frac{h(x)g'(x)}{g(x)} = -\frac{g'(x)}{g(x)^2}.$$

> ⚠ *Caveat. A Minor Mathematical Sin*

Is it over? Lo and behold, to carry out the second-to-last calculation, we implicitly assumed that $h(x) = \frac{1}{g(x)}$ is differentiable at x to begin with! We have found what the derivative *must* be, assuming it exists. To be completely rigorous, we must return to first principles.

Let us write out the rigorous limit definition for the derivative of h.

$$\begin{aligned}\lim_{t\to 0}\frac{h(x+t)-h(x)}{t} &= \lim_{t\to 0}\frac{\frac{1}{g(x+t)}-\frac{1}{g(x)}}{t}\\ &= \lim_{t\to 0}\frac{g(x)-g(x+t)}{tg(x+t)g(x)}\\ &= \left(\lim_{t\to 0}\frac{g(x)-g(x+t)}{t}\right)\left(\lim_{t\to 0}\frac{1}{g(x+t)g(x)}\right)\\ &= -g'(x)\cdot\frac{1}{g(x)^2}.\end{aligned}$$

This limit exists as long as g is differentiable at x and $g(x)\neq 0$ (which also implies g is continuous at x, hence $g(x+t)\to g(x)$ as $t\to 0$).

We have successfully proven our first major result.

Theorem 3.5.3 (Reciprocal Rule)**.** *If g is differentiable at x and $g(x)\neq 0$, then $1/g$ is also differentiable at x and*

$$\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g(x)^2}. \tag{3.11}$$

3.5.2 Synthesizing the Quotient Rule

Going back to our initial maneuvering (3.10), we can now evaluate the derivative of the general quotient:

$$\begin{aligned}\left(\frac{f}{g}\right)' &= \left(f\cdot\frac{1}{g}\right)'\\ &= f'\cdot\frac{1}{g}+f\cdot\left(\frac{1}{g}\right)'\\ &= \frac{f'}{g}+f\left(-\frac{g'}{g^2}\right)\\ &= \frac{f'g}{g^2}-\frac{fg'}{g^2}\\ &= \frac{f'g-fg'}{g^2}.\end{aligned}$$

We formalize this as our next theorem.

Theorem 3.5.4 (Quotient Rule)**.** *If f, g are differentiable at x and $g(x)\neq 0$, then f/g is also differentiable at x and*

$$\left(\frac{f}{g}\right)'(x) = \frac{g(x)f'(x)-f(x)g'(x)}{g(x)^2}. \tag{3.12}$$

Moral. The Sanity Check

Whenever you derive a grand new theorem, it is good practice to test it on a trivial case. The quotient rule must provide the exact same answer as the reciprocal rule when we set $f(x) = 1$. Let's test it: if $f(x) = 1$ for all x, then $f'(x) = 0$. By the Quotient Rule:

$$\begin{aligned}\left(\frac{1}{g}\right)' &= \frac{g(x)(1)' - 1 \cdot g'(x)}{g(x)^2} \\ &= \frac{g(x) \cdot 0 - g'(x)}{g(x)^2} = -\frac{g'(x)}{g(x)^2},\end{aligned}$$

which perfectly agrees with the Reciprocal Rule.

Now for a splendid application. We already have the power rule for positive integers: $(x^n)' = nx^{n-1}$. The reciprocal rule allows us to effortlessly extend this machinery to negative integers.

Consider x^{-n} for some positive integer n:

$$\begin{aligned}(x^{-n})' &= \left(\frac{1}{x^n}\right)' \\ &= -\frac{(x^n)'}{(x^n)^2} \\ &= -\frac{nx^{n-1}}{x^{2n}} \\ &= -nx^{(n-1)-2n} \\ &= (-n)x^{-n-1}.\end{aligned}$$

The same machinery holds! You simply bring down the power, and the new power is reduced by 1.

Story. The Horizon of Abstraction

An ambitious reader may naturally wonder: if the power rule operates so seamlessly for negative integers, does it operate similarly for rational powers? Or even irrational powers? We leave this tantalizing question for a later day. For now, let's practice our mastery with our newfound reciprocal and quotient rules.

Example 3.5.5. By the reciprocal rule,

$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}.$$

Compare this seemingly innocuous result with the general form of the reciprocal rule to appreciate its similarity.

Example 3.5.6. Applying the reciprocal rule to a quadratic denominator:

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

Example 3.5.7. Find the derivative of $h(x) = \frac{x}{x^2+1}$. We can solve this problem via two distinct paths.
Approach 1 (Product Rule + Reciprocal):

$$\begin{aligned} h'(x) &= \left(x \cdot \frac{1}{x^2+1}\right)' \\ &= (x)' \cdot \frac{1}{x^2+1} + x\left(\frac{1}{x^2+1}\right)' \\ &= \frac{1}{x^2+1} + x\left(-\frac{2x}{(x^2+1)^2}\right) \\ &= \frac{1}{x^2+1} - \frac{2x^2}{(x^2+1)^2} \\ &= \frac{(x^2+1)-2x^2}{(x^2+1)^2} = \frac{-x^2+1}{(x^2+1)^2}. \end{aligned}$$

Approach 2 (Quotient Rule directly):

$$\begin{aligned} h'(x) &= \frac{(x)'(x^2+1) - x(x^2+1)'}{(x^2+1)^2} \\ &= \frac{x^2+1-x(2x)}{(x^2+1)^2} \\ &= \frac{-x^2+1}{(x^2+1)^2}. \end{aligned}$$

Both paths lead to Rome.

3.5.3 The Hierarchy of Rules

Let us pause and reflect our intuition on all the differentiation rules established thus far. They are not merely an arbitrary list to be memorized, but a logical hierarchy born from basic principles.

1. *Sum and Scalar Multiplication:* We start with these as direct consequences of limit laws:

$$\begin{aligned} (f \pm g)' &= f' \pm g' \\ (cf)' &= cf' \quad (c \text{ is constant}). \end{aligned}$$

2. *The Product Rule:* A deeper symmetry exists here. The commutativity of the product $h = fg = gf$, combined with the scalar multiplication law, naturally implies the symmetric structure of the product rule:

$$h' = (fg)' = f'g + fg'.$$

3. *The Descendants of the Product:* As we witnessed in this chapter, the logical chain continues. The Product rule implies the Reciprocal rule, which in turn completely implies the Quotient rule:

$$\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}$$
$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}.$$

3.6 Derivatives of Trigonometric Functions

We will now take a brief hiatus from our steady march through general theorems to discuss the differentiation of concrete functions. Our machinery thus far smoothly handles the differentiation of polynomials and rational functions. However, if we survey our grand taxonomy of functions in Section 1.6, we find several wild species remaining:

- *Radicals* (e.g., roots like $\sqrt[n]{x}$)
- *Trigonometric functions* ($\sin x, \cos x, \tan x$, etc.)
- *Exponentials and Logarithms* ($\mathrm{e}^x, \log x$)
- *Non-elementary functions*

In this section, we will hunt down the derivatives of the fundamental *trigonometric functions*. We will lay down the foundation for sine and cosine, and leave the more exotic members of the taxonomy until we have derived another sophisticated tool (the Chain Rule).

3.6.1 The Algebraic Path: Hadamard's Advice

We start with the most fundamental wave of all: $f(x) = \sin x$. Let us boldly apply our first-principles definition of the derivative and observe where the machinery gets stuck:

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{\sin(x+h) - \sin x}{h}. \tag{3.13}$$

Substituting $h = 0$ inevitably brings us to the classic indeterminate form $\frac{0}{0}$. We desperately need to eliminate the h in the denominator. Remembering our fundamental limit $\lim_{h\to 0} \frac{\sin h}{h} = 1$, a strategy crystallizes: we must factor the numerator into the form $(\sin h) \times (\text{something})$.

This requires hunting for a trigonometric identity to expand $\sin(x+h)$. Which identity? Here is a powerful rule of thumb: *every trigonometric identity can be derived from Euler's identity.*

> *Story. The Complex Bridge*
>
> As the great French mathematician *Jacques Hadamard* once famously remarked:
>
> > *The shortest path between two truths in the real domain passes through the complex domain.*

We can find our real-valued addition formulas by taking a quick, elegant detour through the complex plane using Euler's formula: $e^{i\theta} = \cos\theta + i\sin\theta$.

Lemma 3.6.1 (Addition Formulas via Euler)**.** *For any $x, h \in \mathbb{R}$, we have:*

$$\cos(x+h) = \cos x \cos h - \sin x \sin h,$$
$$\sin(x+h) = \cos x \sin h + \sin x \cos h.$$

Proof. Consider the complex exponential $e^{i(x+h)}$. By Euler's formula, the left-hand side is simply:

$$e^{i(x+h)} = \cos(x+h) + i\sin(x+h). \tag{3.14}$$

On the right-hand side, we utilize the standard algebraic laws of exponents to factor the expression, applying Euler's formula to each piece:

$$\begin{aligned} e^{i(x+h)} &= e^{ix}e^{ih} \\ &= (\cos x + i\sin x)(\cos h + i\sin h) \\ &= (\cos x \cos h - \sin x \sin h) + i(\cos x \sin h + \sin x \cos h). \end{aligned} \tag{3.15}$$

Equating the real parts and the imaginary parts of (3.14) and (3.15) yields the desired identities. □

Armed with Lemma 3.6.1, we can conquer the derivative of sine.

Theorem 3.6.2 (Derivatives of Sine and Cosine)**.** *The functions $\sin x$ and $\cos x$ are differentiable for all $x \in \mathbb{R}$, and their derivatives are given by:*

$$(\sin x)' = \cos x, \qquad (\cos x)' = -\sin x.$$

Proof. Applying our addition formula to the limit definition of the derivative, we obtain:

$$\begin{aligned} \lim_{h\to 0} \frac{\sin(x+h) - \sin x}{h} &= \lim_{h\to 0} \frac{(\cos x \sin h + \sin x \cos h) - \sin x}{h} \\ &= \lim_{h\to 0} \left(\cos x \frac{\sin h}{h} + \sin x \frac{\cos h - 1}{h} \right). \end{aligned}$$

Notice that the variable governing the limit is h. Thus, we may treat $\cos x$ and $\sin x$ as constants and pull them out:

$$\lim_{h\to 0} \frac{\sin(x+h) - \sin x}{h} = \cos x \left(\lim_{h\to 0} \frac{\sin h}{h} \right) + \sin x \left(\lim_{h\to 0} \frac{\cos h - 1}{h} \right). \tag{3.16}$$

Here we use the fundamental limit $\lim_{h\to 0} \frac{\sin h}{h} = 1$. But what of the second term?

Intuition. Screaming Pythagoras

The presence of $\cos h$ and the constant 1 is screaming for the Pythagorean identity, $\cos^2 h + \sin^2 h = 1$. By multiplying the numerator and denominator by the conjugate

$(\cos h + 1)$, we can force a difference of squares.

Let us evaluate the second limit utilizing this Pythagorean trick:

$$\begin{aligned}\frac{\cos h - 1}{h} &= \frac{(\cos h - 1)(\cos h + 1)}{h(\cos h + 1)} \\ &= \frac{\cos^2 h - 1}{h(\cos h + 1)} = \frac{-\sin^2 h}{h(\cos h + 1)}.\end{aligned}$$

Taking the limit as $h \to 0$, the expression splits:

$$\lim_{h\to 0} \frac{\cos h - 1}{h} = -\left(\lim_{h\to 0} \frac{\sin h}{h}\right)\left(\lim_{h\to 0} \frac{\sin h}{\cos h + 1}\right) = -1 \cdot \frac{0}{2} = 0.$$

Substituting these evaluated limits back into (3.16), we conclude:

$$(\sin x)' = \cos x \cdot (1) + \sin x \cdot (0) = \cos x.$$

An analogous calculation confirms that $(\cos x)' = -\sin x$. □

Example 3.6.3 (Derivative of Tangent)**.** The derivatives of the remaining trigonometric functions fall out instantly from the Quotient Rule. For instance, to find the derivative of $\tan x$:

$$\begin{aligned}(\tan x)' &= \left(\frac{\sin x}{\cos x}\right)' \\ &= \frac{(\sin x)' \cos x - \sin x (\cos x)'}{\cos^2 x} \\ &= \frac{(\cos x)(\cos x) - (\sin x)(-\sin x)}{\cos^2 x} \\ &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x.\end{aligned}$$

Exercise 3.6.4. Prove the remaining differentiation rules:

$$(\cot x)' = -\csc^2 x, \qquad (\sec x)' = \sec x \tan x, \qquad (\csc x)' = -\csc x \cot x.$$

Remark 3.6.5. Since there are myriad trigonometric identities linking these functions, any derivative computation might yield multiple correct, albeit different-looking, answers. As a general rule, you should simplify as much as possible. However, do not panic if the expression you write down looks algebraically distinct from your friend's answer; both might be entirely correct! For example, in the previous computation, we could just as easily have written the final answer as $\frac{\sin^2 x + \cos^2 x}{\cos^2 x}$ without moving to $\sec^2 x$.

3.6.2 The Geometric Path: Newton's Vision

The proof of Theorem 3.6.2 gets the job done and carries a pleasant algebraic flavor—we massaged the limit via trigonometric identities into something recognizable. But it can feel a bit like symbol pushing.

Is there a geometric proof of the differentiation rules from first principles? Indeed there is. It uses nothing more than the geometric definitions of sine and cosine, and a bit of high school geometry. 2 bonus points: it explains *why* we must use radians instead of degrees, and is more faithful to the geometric vision of Isaac Newton.

> *Moral. The True Meaning of Radians*
>
> Why do calculus textbooks furiously insist on measuring angles in radians? Because on a unit circle, an angle of x radians subtends an arc of exactly length x. This direct equivalence between *angle* and *geometric length* is what allows us to relate trigonometric limits directly to the physical geometry of the circle. Degrees are an arbitrary human construct; Nature uses radians for her geometry.

Let us visualize the unit circle in Figure 3.6. Let $A_0 = (1, 0)$. An angle of x radians means the arc length from A_0 along the circle is exactly x units. Let $A_1 = (\cos x, \sin x)$; thus, $\text{area}(A_0A_1) = x$. Similarly, let $A_2 = (\cos(x+h), \sin(x+h))$, meaning the total arc length to A_2 is $x + h$. The length of the arc *between* A_1 and A_2 is therefore exactly h.

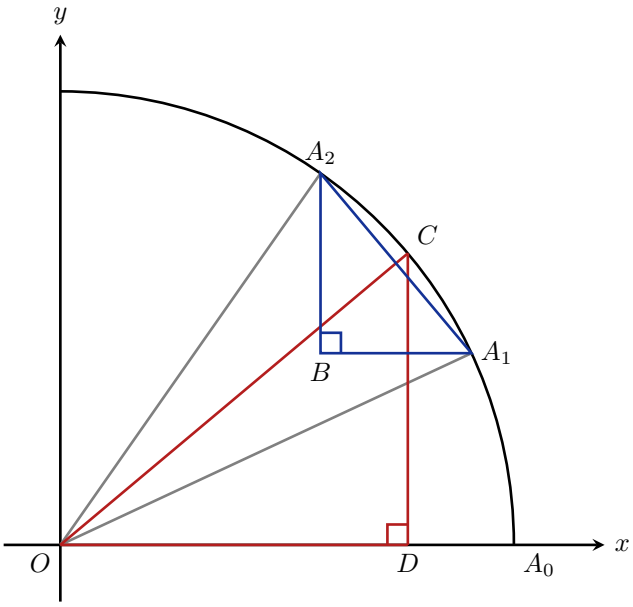


Figure 3.6 – The geometric proof of trigonometric derivatives. Notice that $\Delta BA_1A_2 \sim \Delta DCO$.

Let C be the midpoint of the arc A_1A_2. Because ΔOA_1A_2 is an isosceles triangle (bounded by radii of length 1), the ray OC acts as the perpendicular bisector of the chord A_1A_2. The coordinates of C are naturally $C = \left(\cos\left(x + \frac{h}{2}\right), \sin\left(x + \frac{h}{2}\right)\right)$.

Let D be the geometric projection of C onto the x-axis, so $D = \left(\cos\left(x + \frac{h}{2}\right), 0\right)$. Finally, let B be the intersection of the horizontal line passing through A_1 and the vertical line passing through A_2. Then $B = (\cos(x+h), \sin x)$.

We now have two right-angled triangles of interest: ΔBA_1A_2 and ΔDCO. Let us trace

the angles to see that they are, in fact, similar! In ΔDCO, the leg OD is horizontal and CD is vertical. The angle $\angle DOC$ is exactly $x + \frac{h}{2}$. In ΔBA_1A_2, the leg A_1B is horizontal and A_2B is vertical. A quick chase of complementary angles—noting that the chord A_1A_2 is perpendicular to OC—reveals that the entire triangle ΔBA_1A_2 is simply ΔDCO rotated by 90 degrees and scaled. Consequently, the angle $\angle BA_2A_1$ precisely matches $\angle DOC = x + \frac{h}{2}$.

Using the similarity $\Delta BA_1A_2 \sim \Delta DCO$, we may equate the ratios of their corresponding sides. Specifically, the ratio of the horizontal leg to the hypotenuse in both triangles must be identical:

$$\frac{BA_1}{A_1A_2} = \frac{DC}{CO}. \tag{3.17}$$

Let us carefully express these lengths in terms of our spatial coordinates:

- BA_1 is the horizontal distance between A_1 and B, which is $\cos x - \cos(x+h)$.
- DC is the vertical height of C, which is $\sin\left(x + \frac{h}{2}\right)$.
- CO is the radius of the unit circle, strictly 1.

Substituting these values into (3.17) gives:

$$\frac{\cos x - \cos(x+h)}{A_1A_2} = \sin\left(x + \frac{h}{2}\right).$$

Here comes the *coup de grâce*. We take the limit as $h \to 0$. For infinitesimally small angles, the length of the straight chord A_1A_2 is asymptotically equivalent to the length of the arc A_1A_2. Because we are working in *radians*, that arc length is identically h! Thus, as $h \to 0$, we have $A_1A_2 \approx h$.

Taking the limit yields:

$$\begin{aligned}\lim_{h\to 0} \frac{\cos x - \cos(x+h)}{h} &= \lim_{h\to 0} \sin\left(x + \frac{h}{2}\right) \\ -(\cos x)' &= \sin x \\ (\cos x)' &= -\sin x.\end{aligned}$$

This is a brilliantly geometric proof, relying on nothing more than the similarity of triangles and the true definition of radian angle measure.

Exercise 3.6.6. Find a different pair of corresponding sides in the similar triangles ΔBA_1A_2 and ΔDCO to geometrically prove that $(\sin x)' = \cos x$.

Exercise 3.6.7. Find the derivative of the following functions:

$$f(x) = \frac{x^2 \sin x}{2 + \cos x},$$
$$g(x) = \frac{2 + \tan x}{2 - \tan x}.$$

3.7 Asymptotic Analysis: A Scenic Route to Calculus

> My method to overcome a difficulty is to go round it.
>
> George Pólya

There are several roads one might take to prove theorems in calculus, such as the ever-important chain rule or derivative of exponential function. One can derive it directly by brute-forcing ε-δ arguments, or stringing together limit rules and trying hard to overcome technical hurdles with sheer determination and cleverness, see e.g., [Cli26, LT14].

In this section, we will take a detour. On our way, we will pick up a profoundly useful toolbox—one that is cherished equally by theoretical physicists and applied engineers [BO99]. Following the wisdom of Pólya, rather than smashing our heads against the wall of algebraic manipulation, we shall simply walk around it using the elegant language of *asymptotics*.

3.7.1 Small o and Big O

Before we formalize our tools, let us look at the big picture. When analyzing the behavior of functions near a point (say, x_0), we are often less interested in exact equality and far more interested in *relative sizes*. Which function dominates? Which one vanishes into obscurity?

Definition 3.7.1 (o and O Notations)**.** Let $f(x)$ and $g(x)$ be real-valued functions.

(i) We write "$f(x) = o(g(x))$ as $x \to x_0$" if

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0. \tag{3.18}$$

(ii) We write "$f(x) = O(g(x))$ as $x \to x_0$" if the ratio $\left|\frac{f(x)}{g(x)}\right|$ is bounded near x_0. That is, there exist real numbers $M > 0$ and $\delta > 0$ such that

$$|f(x)| \leq M\,|g(x)| \quad \text{for all } |x - x_0| < \delta. \tag{3.19}$$

> *Intuition.*
>
> If $f(x) = o(g(x))$, then $f(x)$ is *much smaller* than $g(x)$ near x_0. It vanishes faster than g can keep up with. If $f(x) = O(g(x))$, then $f(x)$ is *about as big* as $g(x)$ near x_0; their ratio does not blow up.

> ⚠ *Caveat. An Abuse of Notation*
>
> Note that this notation is technically an abuse of the equals sign—which may be one reason it is omitted in many introductory calculus classes! But the mathematical benefit is more than worth the notational sin.
> We do not really say that $f(x)$ is "equal" to $o(g(x))$ when we write $f(x) = o(g(x))$, since $o(g(x))$ is not a single function. Instead, $o(g(x))$ represents a *set* of functions h that satisfy $\lim_{x \to x_0} \frac{h(x)}{g(x)} = 0$. Hence, a structurally better way to write this would be

$f(x) \in o(g(x))$. Nevertheless, we will follow the tradition of physicists and engineers and use the equals sign.

To manipulate these asymptotic classes fluently, it is extremely useful to have a bridge between limits and bounds. Let us isolate a quick analytical trick.

Lemma 3.7.2. *Suppose* $\lim_{x\to x_0} h(x) = c$*, where* c *is finite. Then* $h(x)$ *is locally bounded near* x_0*. That is, there exist* $M > 0$ *and* $\delta > 0$ *such that*

$$|h(x)| < M \quad \text{for all } |x - x_0| < \delta.$$

Proof. We deploy the ε-δ definition of the limit with a concrete choice, say $\varepsilon = 1$. By definition, there exists a $\delta > 0$ such that

$$|h(x) - c| < 1 \quad \text{for all } |x - x_0| < \delta.$$

By the triangle inequality (with the middle-man c),

$$|h(x)| = |h(x) - c + c| \le |c| + |h(x) - c| < |c| + 1$$

for all $|x - x_0| < \delta$. We simply choose $M := |c| + 1$, completing the proof. □

Corollary 3.7.3. *Thus, if the limit* $\lim_{x\to x_0} \left|\frac{f(x)}{g(x)}\right|$ *exists and is finite, then*

$$f(x) = \begin{cases} O(g(x)) & \text{if the limit } > 0, \\ o(g(x)) & \text{if the limit } = 0. \end{cases}$$

Let us take a moment to build our intuition with a few examples.

Example 3.7.4.

1. $x = o(\sqrt{x})$ as $x \to 0$, because $\lim_{x\to 0} \frac{x}{\sqrt{x}} = 0$.
2. $\sqrt{x} = o(x)$ as $x \to \infty$, because $\lim_{x\to\infty} \frac{\sqrt{x}}{x} = 0$.
3. If $\lim_{h\to 0} \frac{F(h)}{h} = 0$, then $F(h) = o(h)$.
4. $\sin(2x) = O(x)$ as $x \to 0$, because $\lim_{x\to 0} \frac{\sin(2x)}{x} = 2$. (This implies the famous small-angle approximation $\sin(2x) \approx 2x$).

3.7.2 Newton's Approximation in Disguise

We are now ready to see why this notation is so powerful. We can rewrite the very foundation of differential calculus—the derivative—using the small o notation.

Proposition 3.7.5. *If* f *is differentiable at* x_0*, then*

$$f(x_0 + h) = f(x_0) + f'(x_0)h + o(h) \quad \text{as } h \to 0. \tag{3.20}$$

Proof. We begin with the standard limit definition of the derivative:

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

We can trivially rearrange this by bringing $f'(x_0)$ inside the limit (since it is a constant with respect to h):

$$\lim_{h \to 0} \left(\frac{f(x_0 + h) - f(x_0)}{h} - f'(x_0) \right) = 0$$
$$\implies \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - f'(x_0)h}{h} = 0.$$

By definition of small o (Definition 3.7.1), the numerator must be $o(h)$. Therefore:

$$f(x_0 + h) - f(x_0) - f'(x_0)h = o(h)$$
$$\implies f(x_0 + h) = f(x_0) + f'(x_0)h + o(h).$$

□

Story. The Ghosts of Departed Quantities

Interestingly, this asymptotic identity is conceptually much more aligned with the original *infinitesimal calculus* developed by Newton and Leibniz. Instead of dealing with ghostly fractions of zeroes, we view the derivative as a linear mapping. Moreover, this viewpoint is incredibly robust: higher-dimensional derivatives (the Jacobian) or even derivatives in abstract Banach spaces (the Fréchet derivative) are defined exactly this way!

Moral.

If we let $x = x_0 + h$, then Proposition 3.7.5 becomes:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0).$$

This tells us geometrically that the function $f(x)$ can be approximated by its tangent line $y = f(x_0) + f'(x_0)(x - x_0)$, up to an error that is *much smaller* than the distance $|x - x_0|$!

Let us test the raw power of this framework by deriving two classic rules of calculus. By replacing fraction manipulation with asymptotic expansions, proofs that usually require clever middle-man tricks fall out entirely naturally.

Example 3.7.6 (The Product Rule)**.** Suppose f and g are differentiable at x. Then

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

We evaluate the limit definition of the derivative for the product fg, substituting our

asymptotic expansions for $f(x+h)$ and $g(x+h)$:

$$\lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h}$$
$$= \lim_{h\to 0} \frac{\big(f(x) + f'(x)h + o(h)\big)\big(g(x) + g'(x)h + o(h)\big) - f(x)g(x)}{h}.$$

Now, we simply expand the algebraic product in the numerator. Notice how the $f(x)g(x)$ terms will instantly cancel:

$$= \lim_{h\to 0} \left[\frac{f(x)g(x) + f'(x)g(x)h + f(x)g'(x)h + \big(f'(x)g'(x)\big)h^2}{h} + \frac{o(h)\cdot o(h) - f(x)g(x)}{h} \right].$$

Dividing by h, we extract the linear terms, while bundling the remaining higher-order terms:

$$= f'(x)g(x) + f(x)g'(x) + \lim_{h\to 0} \left(\big(f'(x) + g'(x)\big)\frac{o(h)}{h} \cdot h + \frac{o(h)\cdot o(h)}{h} \right). \tag{3.21}$$

Since $f'(x)$ and $g'(x)$ are finite (as f and g are differentiable at x), and by definition $\lim_{h\to 0} \frac{o(h)}{h} = 0$, the entire last limit in (3.21) vanishes. Thus, we are left with:

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

Let us tackle division.

Example 3.7.7 (The Reciprocal Rule)**.** Suppose g is differentiable at x and $g(x) \neq 0$. Then

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

We write out the limit for the derivative of $1/g$:

$$\lim_{h\to 0} \frac{\frac{1}{g(x+h)} - \frac{1}{g(x)}}{h} = \lim_{h\to 0} \frac{g(x) - g(x+h)}{h\, g(x+h)g(x)}.$$

We now invoke our asymptotic hammer, substituting $g(x+h) = g(x) + g'(x)h + o(h)$ directly into the numerator:

$$= \lim_{h\to 0} \frac{g(x) - \big(g(x) + g'(x)h + o(h)\big)}{h\, g(x+h)g(x)}$$
$$= \lim_{h\to 0} \frac{-g'(x)h - o(h)}{h\, g(x+h)g(x)}.$$

Dividing the numerator and denominator by h, we split the limit:

$$= \lim_{h\to 0} \left(-\frac{g'(x)}{g(x+h)g(x)} - \frac{o(h)}{h} \cdot \frac{1}{g(x+h)g(x)} \right).$$

Because g is differentiable (and therefore continuous), $g(x+h) \to g(x)$ as $h \to 0$. Furthermore, $\frac{o(h)}{h} \to 0$. The second term vanishes, and the first term limits to our desired result:

$$\lim_{h\to 0} \frac{\frac{1}{g(x+h)} - \frac{1}{g(x)}}{h} = -\frac{g'(x)}{g(x)^2}.$$

Exercise 3.7.8 (The Quotient Rule)**.** Using the small o notation techniques we have just developed, derive the standard quotient rule:

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}.$$

Hint: Do not be afraid to let $o(h)$ absorb any cross-terms that divide out to zero as $h \to 0$.

3.8 The Chain Rule

Story. The Gears of Calculus

Imagine a physical machine comprised of interlocking gears. If the first gear (let's call its rotation x) drives a second gear (y) such that one turn of x produces 2 turns of y, we say the rate of change is 2. If this second gear drives a third gear (z) such that one turn of y produces 3 turns of z, what is the relationship between the first and third gears?
It is intuitively obvious: turning x once turns y twice, and each of those two turns of y turns z three times. Thus, one turn of x yields $2 \times 3 = 6$ turns of z. The rates multiply. This simple mechanical truth is the beating heart of the Chain Rule.

Before delving into infinitesimals, let us formalize this with simple linear functions. Consider a composite of two linear functions:

$$g(x) = 2x + 3,$$
$$f(y) = 3y + 1.$$

If we compose them, substituting $y = g(x)$ into f, we find:

$$f(g(x)) = 3(2x + 3) + 1 = 6x + 10.$$

Notice the slopes. The rate of change of g is $g' = 2$. The rate of change of f is $f' = 3$. The rate of change of the composite function $f \circ g$ is exactly their product:

$$(f \circ g)' = f' \cdot g' = (3)(2) = 6.$$

For continuous f and g, recall the continuity of $f(g(x))$: a small change in x causes a small change in $g(x)$, and that, in turn, causes a small change in $f(g(x))$. Now, with differentiable f and g, we are no longer just guaranteeing that the changes are small; we can precisely *quantify* these small changes.

Intuition. Linearity under the Microscope

Every differentiable function locally behaves like a linear function, $F(x+h) \approx F(x) + F'(x)h$. If we compose two differentiable functions, we are locally composing two linear functions. Since the slopes of linear functions multiply upon composition, the derivatives (which are just local slopes) must also multiply.

We now state this principle rigorously.

Theorem 3.8.1 (Chain Rule)**.** *If g is differentiable at x and f is differentiable at $g(x)$, then the composite function $f \circ g$ is differentiable at x, and its derivative is given by*

$$(f \circ g)'(x) = f'(g(x))g'(x). \tag{3.22}$$

Proof. We compute the limit of the difference quotient for $f \circ g$:

$$\lim_{h\to 0} \frac{(f\circ g)(x+h) - (f\circ g)(x)}{h} = \lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{h}.$$

To evaluate this, we use the fact that g is differentiable at x. We can express the local behavior of g using little-o notation. Namely, an increment h in the domain yields:

$$g(x+h) = g(x) + g'(x)h + o(h),$$

where $o(h)$ represents an error term that vanishes faster than h (i.e., $\lim_{h\to 0} o(h)/h = 0$).

Substituting this into the numerator of our limit, we obtain:

$$\lim_{h\to 0} \frac{f\left(g(x) + g'(x)h + o(h)\right) - f(g(x))}{h}.$$

Now, treat the quantity $k = g'(x)h + o(h)$ as a small perturbation to $g(x)$. Since f is differentiable at $g(x)$, we can similarly expand $f(g(x)+k)$ as $f(g(x)) + f'(g(x))k + o(k)$. Notice that as $h \to 0$, $k \to 0$ as well. Thus:

$$f(g(x+h)) = f(g(x)) + f'(g(x)) \underbrace{\big(g'(x)h + o(h)\big)}_{k} + o(k).$$

Because k scales linearly with h, the error term $o(k)$ is also bounded by $o(h)$. Thus, the entire expression simplifies nicely:

$$\lim_{h\to 0} \frac{f(g(x)) + f'(g(x))\big(g'(x)h + o(h)\big) - f(g(x))}{h}$$

The $f(g(x))$ terms cancel, leaving:

$$\begin{aligned} &= \lim_{h\to 0} \frac{f'(g(x))g'(x)h + f'(g(x))o(h)}{h} \\ &= \lim_{h\to 0} \left(f'(g(x))g'(x) + f'(g(x))\frac{o(h)}{h} \right). \end{aligned}$$

Since $\lim_{h\to 0} \frac{o(h)}{h} = 0$, the second term vanishes. We are left exactly with:

$$f'(g(x))g'(x).$$

This completes the proof. □

Remark 3.8.2. The use of little-o notation here is not merely pedantic; it elegantly skirts a common pothole in introductory proofs where one might naively multiply and divide by $g(x+h)-g(x)$. That naive approach fails when $g(x+h)=g(x)$ for infinitely many h near 0 (e.g., if g is constant or highly oscillatory). The linear approximation definition of the derivative handles this flawlessly.

> *Moral. Look for the Bracket!*
>
> When do we apply the chain rule? We always apply it when we can write the composition of functions with brackets:
>
> $$(f\circ g)(x)=f\big(\,g(x)\,\big).$$
>
> Look for the *bracket*! If the inside of the bracket is the *argument* of some function, it's time for chain rule. When taking the derivative, always work systematically from the outermost function to the innermost.

Example 3.8.3. Consider the function $(x^2-x+5)^4$. We can view this as $f\circ g$, where the "inner" function is $g(x)=x^2-x+5$ and the "outer" function is $f(y)=y^4$.

Using the power rule for the outer function and the polynomial rules for the inner function, the chain rule dictates:

$$\begin{aligned}\big((x^2-x+5)^4\big)' &= 4(x^2-x+5)^3\cdot(x^2-x+5)'\\ &= 4(x^2-x+5)^3(2x-1).\end{aligned}$$

3.8.1 The Chains of Leibniz

The chain rule can be used to differentiate compositions of more than two functions; indeed, this motivates its name, since we can think of this as a *chain* of functions composed with each other.

For example, given three functions $f,g,h:\mathbb{R}\to\mathbb{R}$, we can visualize the composite $F=f\circ g\circ h$ as in Figure 3.7 below.

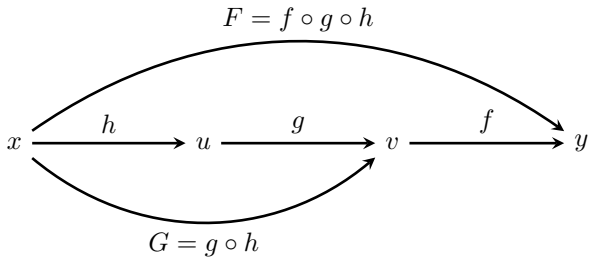


Figure 3.7 – Illustration of the chain rule of multiple functions.

Writing $G=f\circ g$, we can use the chain rule to write G' in terms of f' and g', then again

to write F' in terms of G' and h', obtaining:

$$\begin{aligned} F'(x) &= (G \circ h)'(x) \\ &= G'(h(x))h'(x) \\ &= (f \circ g)'(h(x))h'(x) \\ &= f'(g(h(x))) \cdot g'(h(x)) \cdot h'(x). \end{aligned}$$

It is perhaps easiest to view this using the alternate notation championed by Gottfried Wilhelm Leibniz:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}v}\frac{\mathrm{d}v}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}, \tag{3.23}$$

where $u = h(x)$, $v = g(u) = g(h(x))$, and $y = f(v) = f(g(h(x)))$.

> ⚠ *Caveat. The Leibniz Illusion*
>
> We reiterate that although this notation makes it very tempting to view the chain rule as just a matter of "canceling the $\mathrm{d}v$ and the $\mathrm{d}u$ from the numerator and denominator," this is *not* an actual proof of anything. The quantities $\frac{\mathrm{d}y}{\mathrm{d}v}$, $\frac{\mathrm{d}v}{\mathrm{d}u}$, and $\frac{\mathrm{d}u}{\mathrm{d}x}$ are entirely unified symbols representing limits (derivatives), not algebraic fractions.
> At this stage in your mathematical journey, the symbols $\mathrm{d}u$ and $\mathrm{d}v$ have no independent meaning. (Later, in the study of differential geometry and differential forms, we will give them strict algebraic definitions, but for now, treat the "cancellation" strictly as a mnemonic device).

Example 3.8.4. The function $F(x) = \sin(\mathrm{e}^{3x})$ can be elegantly differentiated by introducing intermediate variables: let $u = 3x$, let $v = \mathrm{e}^u = \mathrm{e}^{3x}$, and let $y = \sin(v) = \sin(\mathrm{e}^{3x})$.

Applying Leibniz's chained gears (3.23), we compute:

$$\begin{aligned} F'(x) = \frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{\mathrm{d}y}{\mathrm{d}v}\frac{\mathrm{d}v}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} \\ &= (\cos v)(\mathrm{e}^u)(3) \\ &= 3\mathrm{e}^u \cos v. \end{aligned}$$

Note that although applying (3.23) gave us an expression in terms of u, v, and x, this is not yet our final answer! It is always necessary to go one step further and substitute u and v back in terms of the original variable x. Doing so yields:

$$F'(x) = 3\mathrm{e}^{3x} \cos(\mathrm{e}^{3x}).$$

You will encounter this phenomenon—introducing variables to simplify the local geometry, then translating back to the global coordinates—in many other places as well.

Exercise 3.8.5. Find the derivative of $f(x) = \tan(\sin(x^2))$.
Hint for the reader: Do not panic at the nesting. Peel it like an onion, from outermost to innermost:

1. The derivative of $\tan(\text{blob})$ is $\sec^2(\text{blob}) \cdot (\text{blob})'$.

2. The derivative of sin(stuff) is cos(stuff) · (stuff)′.

3. The derivative of x^2 is $2x$.

Putting it all together:

$$f'(x) = \sec^2(\sin(x^2)) \cdot \cos(x^2) \cdot 2x.$$

3.8.2 The Derivative of the Exponential Function

We now embark on a heuristic argument to derive the derivative of $f(x) = \mathrm{e}^x$.

Story. The Price of Early Transcendentals

Unfortunately, we will not be able to give a fully rigorous proof at this stage. This is the price we pay for the *early transcendental* approach to calculus. Of course, with sheer determination to overcome multiple technicalities, one *can* find the derivative of e^x rigorously right now. However, the most mathematically natural way is to leave it until one learns the formal theory of sequences and series.

So far, we only have two pieces of information: e is a natural number approximating $2.718\ldots$, and it beautifully satisfies Euler's identity:

$$\mathrm{e}^{i\theta} = \cos\theta + i\sin\theta.$$

Let us calculate the derivative of this expression with respect to θ:

$$\begin{aligned}(\mathrm{e}^{i\theta})' &= (\cos\theta + i\sin\theta)' \\ &= (\cos\theta)' + i(\sin\theta)' \\ &= -\sin\theta + i\cos\theta.\end{aligned}$$

Recalling that $i^2 = -1$, we can rewrite $-\sin\theta$ as $i^2\sin\theta$:

$$\begin{aligned}(\mathrm{e}^{i\theta})' &= i^2\sin\theta + i\cos\theta \\ &= i(i\sin\theta + \cos\theta) \\ &= i\mathrm{e}^{i\theta}.\end{aligned}$$

But we ultimately want the derivative of e^x, and the above calculation only holds for a real parameter θ.

Caveat. A Leap of Faith

We are about to boldly extrapolate. We suppose that the formula we just derived, $(\mathrm{e}^{iz})' = i\mathrm{e}^{iz}$, holds for *any* complex number z. While this is a leap of faith for now, complex analysis will later vindicate our bravery.

Using this assumption, we can write $\mathrm{e}^x = \mathrm{e}^{i(-ix)}$. Applying the chain rule to this carefully constructed identity yields:

$$\begin{aligned}(\mathrm{e}^x)' &= \left(\mathrm{e}^{i(-ix)}\right)' \\ &= \left(\mathrm{e}^{iz}\right)' \cdot z' \quad \text{where } z = -ix \\ &= i\mathrm{e}^{i(-ix)} \cdot (-i) \\ &= -i^2\mathrm{e}^{-i^2x} \\ &= -(-1)\mathrm{e}^{-(-1)x} \\ &= \mathrm{e}^x.\end{aligned}$$

And there we have it: the derivative of e^x is e^x itself!

Let's us warm up with a relation between logarithm and exponential, and its application. Let a be any positive number. We want to find the derivative of a^x.

> *Intuition. The Fundamental Philosophy of Exponential*
>
> Before we formalize the derivative of a general exponential a^x, it is instructive to look at the big picture. All exponentials are, fundamentally, the natural exponential e^x wearing a disguise. By leveraging the identity $a = \mathrm{e}^{\log a}$, we can translate any problem involving arbitrary bases back into the familiar, natural language of e.

We write a as $\mathrm{e}^{\log a}$. Then $a^x = (\mathrm{e}^{\log a})^x = \mathrm{e}^{(\log a)x}$. Using the chain rule and the fact that the derivative of e^x is e^x, we get that

$$(a^x)' = (\mathrm{e}^{(\log a)x})' = (\log a)\mathrm{e}^{(\log a)x} = (\log a)a^x.$$

Notice how the change of base simply summons a scaling factor of $\log a$.

Example 3.8.6. The function e^{x^2+1} is a composition $f(g(x))$, where $g(x) = x^2 + 1$ and $f(x) = \mathrm{e}^x$. By the chain rule,

$$(\mathrm{e}^{x^2+1})' = \mathrm{e}^{x^2+1}(x^2+1)' = \mathrm{e}^{x^2+1}2x.$$

In the $\frac{\mathrm{d}}{\mathrm{d}x}$ notation, if $y = \mathrm{e}^{x^2+1}$ and $u = x^2 + 1$, then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = \mathrm{e}^u(2x) = \mathrm{e}^{x^2+1}2x.$$

More generally,

$$(\mathrm{e}^{f(x)})' = \mathrm{e}^{f(x)}f'(x)$$

for any differentiable function f. Similarly, by the chain rule,

$$(f(\mathrm{e}^x))' = f'(\mathrm{e}^x)\mathrm{e}^x$$

for any differentiable function f.

Exercise 3.8.7. Find the first derivative of the following expressions:

1. $y = 2^x$
2. $y = 10^{3x}$
3. $y = \mathrm{e}^{x^2}$.

3.8.3 Higher Derivatives

It is quite interesting that the exponential function $f(x) = \mathrm{e}^x$ has the derivative $f'(x) = \mathrm{e}^x = f(x)$. This immediately implies that we can take a derivative of the derivative of e^x, ad infinitum.

Definition 3.8.8 (Higher Derivatives)**.** When the derivative of the derivative of f exists at x, it is called the *second derivative* of f and is written in any of the following ways:

$$(f')'(x) = f''(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}f}{\mathrm{d}x}\right) = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x).$$

Similar notation is used for higher derivatives, sometimes with superscripts enclosed in parentheses to indicate the order. For instance, the n-th derivative is denoted $f^{(n)}(x)$ or $\frac{\mathrm{d}^n y}{\mathrm{d}x^n}$.

Let us look at a simple polynomial example before returning to our exponential.

Example 3.8.9. Let $f(x) = x^4$. Then:

$$f'(x) = 4x^3, \quad f''(x) = 12x^2, \quad f'''(x) = 24x, \quad f^{(4)}(x) = 24, \quad f^{(5)}(x) = 0.$$

Now, let us apply this to the exponential function:

Example 3.8.10. Let $f(x) = \mathrm{e}^x$. Then $f^{(2)}(x) = \frac{\mathrm{d}^2}{\mathrm{d}x^2}(\mathrm{e}^x) = \mathrm{e}^x$. In fact, by trivial induction:

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}(\mathrm{e}^x) = \mathrm{e}^x$$

for any positive integer n.

We say that $f(x) = \mathrm{e}^x$ is *invariant under the derivative*. It is essentially the *only* family of functions (up to scalar multiplication) with this property.

> *Moral. The Birth of a New Discipline*
>
> Because e^x is entirely characterized by its invariance under differentiation, some textbooks actually *define* e as the unique number such that $f(x) = \mathrm{e}^x$ satisfies the equation:
>
> $$f'(x) = f(x).$$
>
> This identity is an equation where the unknown is not a number, but a function, and the equation involves differentiation. Hence, mathematicians gave such equations a name: *differential equation*!

3.8.4 Cycles, Rotations, and Trigonometry: The Math of Oscillation

Higher derivatives of sine and cosine can be easily computed, and they reveal a beautiful underlying cyclic structure:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\sin x = \cos x \quad &\Longrightarrow \quad \frac{\mathrm{d}^2}{\mathrm{d}x^2}\sin x = \frac{\mathrm{d}}{\mathrm{d}x}\cos x = -\sin x \\
&\Longrightarrow \quad \frac{\mathrm{d}^3}{\mathrm{d}x^3}\sin x = \frac{\mathrm{d}}{\mathrm{d}x}(-\sin x) = -\cos x \\
&\Longrightarrow \quad \frac{\mathrm{d}^4}{\mathrm{d}x^4}\sin x = \frac{\mathrm{d}}{\mathrm{d}x}(-\cos x) = \sin x.
\end{aligned}$$

After this, the pattern repeats with every fourth derivative.

Intuition. Derivatives as Geometric Rotations

It is worth comparing this 4-cycle to the observation that:

$$\frac{\mathrm{d}^4}{\mathrm{d}x^4}\mathrm{e}^{ix} = \frac{\mathrm{d}^3}{\mathrm{d}x^3}(i\mathrm{e}^{ix}) = \frac{\mathrm{d}^2}{\mathrm{d}x^2}(-\mathrm{e}^{ix}) = \frac{\mathrm{d}}{\mathrm{d}x}(-i\mathrm{e}^{ix}) = \mathrm{e}^{ix}.$$

This admits a profound geometric interpretation. Multiplying a complex number by i corresponds to rotating it by 90 degrees ($\pi/2$ radians) counter-clockwise around the origin in the complex plane.
Doing this twice corresponds to a 180-degree rotation (multiplication by -1). Doing it four times returns us to where we started. Since differentiating the sine function has the same cyclic property, we can view the derivative operator $\frac{\mathrm{d}}{\mathrm{d}x}$ acting on trigonometric functions as a translation of the argument by $\pi/2$:

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin x = \cos x = \sin\left(x + \frac{\pi}{2}\right).$$

Because of this 4-cycle, computing arbitrarily high derivatives becomes an exercise in modular arithmetic.

Example 3.8.11. We can compute $\frac{\mathrm{d}^{53}}{\mathrm{d}x^{53}}\sin x$ by observing that $53 = 4 \cdot 13 + 1$. So the first 52 derivatives get us exactly back to $\sin x$, and we thus have:

$$\frac{\mathrm{d}^{53}}{\mathrm{d}x^{53}}\sin x = \frac{\mathrm{d}}{\mathrm{d}x}\sin x = \cos x.$$

Of particular physical importance is the fact that if $f(x) = \sin x$ or $f(x) = \cos x$ (or more generally, any linear combination $f(x) = a\sin x + b\cos x$), then:

$$f''(x) = -f(x) \quad \text{for all } x \in \mathbb{R}. \tag{3.24}$$

This is a tremendously important example of a differential equation, one which arises naturally in almost every area of physics.

Story. Newton Meets Hooke

Consider an object on a spring that, at time t, is displaced by a distance r from its equilibrium resting position. The force acting on the object is given by Hooke's Law: it has magnitude kr, where $k > 0$ is the *spring constant* measuring the stiffness of the spring. Crucially, this force is a restoring force—it is directed in the opposite direction of the displacement.

If m is the mass of the object, then Newton's second law ($F = ma$) gives us:

$$m\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = \text{force} = -kr \quad \Longrightarrow \quad \frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = -\frac{k}{m}r. \tag{3.25}$$

Let us define a natural frequency parameter $\omega \coloneqq \sqrt{k/m}$, and consider a rescaled function $f(x) = r(x/\omega)$. Using the chain rule, we can relate the derivatives of f to the derivatives of r:

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x} r\left(\frac{x}{\omega}\right) = \frac{1}{\omega} r'\left(\frac{x}{\omega}\right).$$

Differentiating a second time yields:

$$\begin{aligned} f''(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\omega} r'\left(\frac{x}{\omega}\right)\right) = \frac{1}{\omega^2} r''\left(\frac{x}{\omega}\right) \\ &= \frac{1}{\omega^2}\left(-\frac{k}{m} r\left(\frac{x}{\omega}\right)\right) = -r\left(\frac{x}{\omega}\right) = -f(x). \end{aligned}$$

This means that f must be a function that satisfies the exact differential equation we discovered in (3.24)! While a formal proof requires the theory of linear differential equations, we can confidently state that every function satisfying $f''(x) = -f(x)$ can be written in the form $f(x) = a\sin x + b\cos x$.

Returning to our physical variables, we conclude that the displacement of the mass on the spring is given by:

$$r(t) = f(\omega t) = a\sin(\omega t) + b\cos(\omega t). \tag{3.26}$$

This system is the quintessential example of a *simple harmonic oscillator*.

Exercise 3.8.12. Prove that if $r(t)$ is given by (3.26), then there exist constants $A > 0$ and $\phi \in \mathbb{R}$ (an amplitude and a phase shift) such that:

$$r(t) = A\sin(\omega t + \phi).$$

To test your intuition on higher derivatives, polynomials, and cyclic trigonometric functions, try the following

Exercise 3.8.13. Find $f^{(1026)}(x)$ given that

$$f(x) = x^{100} + \sin(2x).$$

Hint: What happens to a polynomial of degree 100 after 101 derivatives? For the sine term, don't forget the chain rule, and recall our modular arithmetic trick!

3.9 Derivatives of Inverse Functions

We are nearing the end of our quest to find a systematic way to calculate the derivatives of all basic functions. With the exponential function recently conquered, there are only three species of elementary functions left in our taxonomy: radicals, logarithms, and inverse trigonometric functions. (For non-elementary functions, one must typically resort to the rigorous definition of the derivative and rely on approximation on a case-by-case basis. But for our elementary toolkit, we demand exactness.)

In this section, we develop a unified, deeply powerful tool to resolve all three of these remaining types by exploiting the profound symmetry between a function and its inverse. Suppose we wish to find the derivative of an inverse function f^{-1}. If we assume that f^{-1} is differentiable at $b = f(a)$ and f is differentiable at a, we can write the definition of inverse:

$$y = f(f^{-1}(y))$$

then differentiate both sides with respect to y. Using the chain rule, and evaluating at $y = b$, we find:

$$1 = (f^{-1})'(b) \cdot f'(a) \implies (f^{-1})'(b) = \frac{1}{f'(a)}.$$

We can rewrite this elegantly as (note that $b = f(a)$ so $a = f^{-1}(b)$):

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))}. \tag{3.27}$$

⚠ *Caveat. The Bootstrapping Problem*

There is one logical gap in our application of implicit differentiation above: in order for this algebraic manipulation to be rigorous, we need to *know* beforehand that f^{-1} is differentiable at $b = f(a)$. One cannot simply apply the chain rule to a function whose differentiability is not yet established!

To resolve this, we present the formal theorem, which guarantees both the existence of the derivative and its explicit value.

Theorem 3.9.1 (Inverse Function Theorem)**.** *Let f be an invertible function and let $y_0 = f(x_0)$. If f is differentiable at x_0 with $f'(x_0) \neq 0$, and if f^{-1} is continuous at y_0, then f^{-1} is differentiable at y_0 and*

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Moral.

You can find the derivative of an inverse function without ever explicitly finding the formula for the inverse function itself!

Story. The Invisible Inverse

The moral above is physically and mathematically profound. Recall that finding an explicit formula for an inverse is not always possible. In several cases, we only know that the inverse exists because the original function is purely 1-to-1 and onto. For example, a monotone polynomial of degree ≥ 5 is perfectly invertible, but by Galois Theory [Ste22], there is no general algebraic formula for its roots. We are faced with a ghost function—we know it exists, but we cannot write it down. Yet, Theorem 3.9.1 allows us to find its rate of change exactly. This is the magic of calculus.

Remark 3.9.2 (On the Proof). The simplest proof of Theorem 3.9.1 operates through sequences (a staple of Calculus II), which provides a rigorous way to replace the limit of variables inside the range by something manageable, see e.g., [Tao22a]. There is also a direct ε-δ proof, which we leave as a challenging but rewarding exercise for the reader.

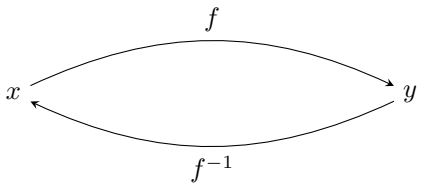


Figure 3.8 – Illustration of a function f and its inverse f^{-1}

Intuition. Leibniz's Reciprocal

To add geometric intuition, consider the mapping diagram in Figure 3.8.
The theorem essentially says $(f^{-1})'(y_0) = \frac{1}{f'(x_0)}$. This is easily remembered by utilizing the Leibniz notation for our diagram. If $y = f(x)$, then:

$$\frac{\mathrm{d}x}{\mathrm{d}y} = \frac{1}{\frac{\mathrm{d}y}{\mathrm{d}x}}.$$

To be perfectly precise with our evaluation points, this means:

$$\left.\frac{\mathrm{d}x}{\mathrm{d}y}\right|_{y=y_0} = \frac{1}{\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=x_0}}.$$

3.9.1 Derivatives of the Inverse Trigonometric Functions

Let us wield this new tool to find the derivative of the inverse sine function. Let $f(x) = \sin x$. To ensure it is invertible, we must restrict its domain to $\mathrm{dom}(f) = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Then $f'(x_0) = \cos x_0$. Thus, by the Inverse Function Theorem,

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{\cos x_0},$$

or simply:

$$(\sin^{-1})'(y_0) = \frac{1}{\cos x_0}.$$

By common practice, we usually use x for the argument of our functions, so we wish to swap y_0 for x. Therefore, we must also change x_0 accordingly. We do this before swapping for x: recall $y_0 = \sin x_0$, which means $x_0 = \sin^{-1} y_0$. Substituting this back in gives:

$$(\sin^{-1})'(y_0) = \frac{1}{\cos(\sin^{-1} y_0)}.$$

Now, exchanging variables:

$$(\sin^{-1})'(x) = \frac{1}{\cos(\sin^{-1} x)}. \tag{3.28}$$

In practice, it is usually not worth remembering the exact abstract formula for calculation (though the gist of the theorem is vital for theoretical purposes). We just start straight away

with the chain rule. Since $\sin(\sin^{-1} x) = x$, we differentiate both sides:

$$1 = \left[\sin(\sin^{-1} x)\right]' = \cos(\sin^{-1} x) \cdot (\sin^{-1})'(x).$$

Isolating the derivative immediately yields (3.28).

Remark 3.9.3 (Geometric Simplification). One can simplify (3.28) further by invoking Pythagoras. Let $\theta = \sin^{-1} x$, which means $x = \sin\theta$. We can visualize this via a right triangle in Figure 3.9.

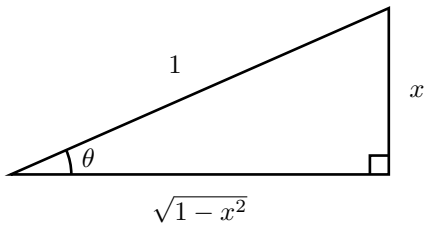


Figure 3.9 – Right triangle illustration for $\theta = \sin^{-1} x$

By the Pythagorean theorem, the adjacent side is $\sqrt{1-x^2}$. Thus,

$$\cos(\sin^{-1} x) = \cos\theta = \frac{\sqrt{1-x^2}}{1} = \sqrt{1-x^2}.$$

Substituting this back into our derivative, we obtain the beautiful algebraic expression:

$$(\sin^{-1})'(x) = \frac{1}{\sqrt{1-x^2}}.$$

Now that we have conquered the sine function, it is time for you to derive the rest of the trigonometric bestiary.

Exercise 3.9.4. Find the derivative of $\cos^{-1} x$, $\tan^{-1} x$, $\cot^{-1} x$, $\sec^{-1} x$, and $\csc^{-1} x$. For each function, explicitly state the restricted domain first.

Exercise 3.9.5 (Challenge)**.** Let $g(x) = \sin x$ with $\mathrm{dom}(g) = \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$. Is g invertible? If so, find g^{-1}.

Note that $g^{-1}(x)$ is not simply $\sin^{-1}(x)$ because of the different ranges! Thus, fixing the domain of the original function is of utmost importance.

Finally, a word of caution regarding the language we use to express these ideas.

⚠ *Caveat. The Notational Headache*

Mathematical notation is a living language riddled with historical quirks. Given a function f, the superscript -1 is a notorious trap:

- $f^{-1}(x)$ means the *inverse* function of f.
- $[f(x)]^{-1} = \frac{1}{f(x)}$ means the *reciprocal* of f. (Notice the position of the -1).

The latter is consistent with the notation $[f(x)]^n$ which means the n-th power of f.

The real bummer is $f^n(x)$. Depending on the context, it can mean the n-th power (like $\sin^2 x$, $\cos^3 x$, etc.), but it can also mean the n-fold composition:

$$f^n = \underbrace{f \circ f \circ \cdots \circ f}_{n\text{-times}}.$$

So, always look for the context!
Here is another one to watch out for: $f^{(n)}(x)$ (with parentheses around the n) means the n-th derivative of f.

Example 3.9.6. Observe the profound difference when calculating derivatives based on notational intent: Find the derivative of $f(x) = \sin^2 x$ and $g(x) = \sin(\sin x)$.

Solution. For the first function, the notation $\sin^2 x$ strictly implies $[\sin x]^2$. We apply the chain rule, treating the power function as the outer layer:

$$f'(x) = 2(\sin x) \cdot (\sin x)' = 2 \sin x \cos x = \sin(2x). \tag{3.29}$$

For the second function, $g(x) = \sin(\sin x)$, the structure is an explicit composition of functions $g(x) = (h \circ h)(x)$ where $h(x) = \sin x$. We apply the chain rule with the outer function being the sine function:

$$g'(x) = \cos(\sin x) \cdot (\sin x)' = \cos(\sin x) \cos x. \tag{3.30}$$

These derivatives are fundamentally different mathematical objects. The notation demands our vigilance. □

3.9.2 Derivatives of Logarithms and Radicals

We show how to use our knowledge of the derivative of e^x and the chain rule to compute the derivative of $\log x$. We know that the natural logarithm is the inverse function of e^x,

$$x = e^{\log x} \qquad (x > 0).$$

The expression $e^{\log x}$ is a composition of two functions, where the logarithm is applied first and then the exponential. Thus by the chain rule,

$$1 = (x)' = (e^{\log x})' = e^{\log x}(\log x)' = x(\log x)'.$$

Solving for $(\log x)'$, we get

$$(\log x)' = \frac{1}{x}, \qquad x > 0.$$

Moral.

The function $\log x$ is a rather complicated transcendental function, but its derivative is remarkably simple. The calculus of logarithms acts as a bridge, transforming the transcendental world back into the algebraic world of rational functions.

When x is negative, we have $(\log(-x))' = \frac{1}{-x}(-x)' = \frac{1}{x}$. So we see more generally that

$$(\log|x|)' = \frac{1}{x}, \qquad x \neq 0.$$

The derivative of $\log x$ can be used to compute the derivative of $\log_a x$: We have the identity $x = a^{\log_a x}$. Applying the function log to each side, we get $\log x = (\log a)(\log_a x)$. Hence

$$\log_a x = \frac{1}{\log a}\log x.$$

Differentiate to get

$$(\log_a x)' = \frac{1}{\log a}(\log x)' = \frac{1}{x\log a}.$$

We summarize in a theorem:

Theorem 3.9.7 (Derivative of Natural Exponential and Logarithm)**.**

$$\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^x) = \mathrm{e}^x, \qquad \frac{\mathrm{d}}{\mathrm{d}x}(\log|x|) = \frac{1}{x} \quad (x \neq 0).$$

Now that we know the derivatives of e^x and $\log x$, we can now establish the power rule for *any* real exponent, finally freeing ourselves from the constraints of rational numbers.

Theorem 3.9.8. *If $r \neq 0$ and $x > 0$, then x^r is differentiable and*

$$(x^r)' = rx^{r-1}.$$

Proof. For $x > 0$, $x^r = \mathrm{e}^{\log(x^r)} = \mathrm{e}^{r\log x}$, so

$$(x^r)' = (\mathrm{e}^{r\log x})' = \mathrm{e}^{r\log x}(r\log x)' = \mathrm{e}^{r\log x}\frac{r}{x} = x^r\frac{r}{x} = rx^{r-1}. \qquad \square$$

Example 3.9.9. Applying Theorem 3.9.8 to some exotic powers:

$$(x^\pi)' = \pi x^{\pi-1}, \qquad (x^\mathrm{e})' = \mathrm{e}x^{\mathrm{e}-1}.$$

Example 3.9.10. If g is a differentiable positive function, then the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}x}\log g(x) = \frac{g'(x)}{g(x)}. \tag{3.31}$$

The expression on the right-hand side of (3.31) is called the *logarithmic derivative* of g.

Example 3.9.11. If $y = \log(x^2+3x+4)$, then

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

Example 3.9.12.

$$\frac{\mathrm{d}}{\mathrm{d}x}\log(\cos x) = \frac{-\sin x}{\cos x} = -\tan x.$$

3.10 A Bag of Logarithmic Tricks

Story. The Astronomer's Lifesaver

In the early 17th century, before the dawn of mechanical calculators, astronomers and navigators spent agonizing portions of their lives multiplying gargantuan numbers to predict celestial motions. The Scottish mathematician John Napier introduced a radical parlor trick in 1614: a mathematical function that could metamorphose painful multiplication into gentle addition. This conceptual leap—the logarithm—remains one of the most powerful algebraic tools in modern calculus.

Before we engage in the algebraic gymnastics of logarithmic differentiation, we must firmly ground ourselves in the two fundamental derivatives of exponential and logarithmic functions:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^x\right) = \mathrm{e}^x, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\left(\log|x|\right) = \frac{1}{x} \quad (x \neq 0).$$

It is a central philosophy in mathematics to reduce the unfamiliar to the familiar. How do we differentiate $\log_{10} x$ or 2^x? Rather than memorizing an endless catalog of new rules, we manipulate the expressions back to our native base, e.

Example 3.10.1. Let $f(x) = 2^x$. Find $f'(x)$.

Solution. We again use the logarithm to free the variable x from the exponent:

$$\log f(x) = \log(2^x) = x \log 2.$$

Differentiating both sides implicitly with respect to x requires the chain rule on the left-hand side:

$$\begin{aligned}(\log f(x))' &= (x \log 2)' \\ \frac{1}{f(x)} f'(x) &= \log 2.\end{aligned}$$

Solving for $f'(x)$ and substituting our original function $f(x) = 2^x$ back into the expression yields:

$$f'(x) = f(x) \log 2 = 2^x \log 2.$$

□

Example 3.10.2. Let $f(x) = \log_{10} x$. Find $f'(x)$.

Solution. We begin by invoking the fundamental definition of the logarithm to unpack the function:

$$10^{f(x)} = 10^{\log_{10} x} = x.$$

By taking the natural logarithm of both sides, we project the equation into our preferred base:

$$\log\left(10^{f(x)}\right) = \log x.$$

Applying the exponent rule $\log(a^b) = b \log a$, we pull the function down from the exponent:

$$f(x) \log 10 = \log x \implies f(x) = \frac{\log x}{\log 10}.$$

Now, differentiation is trivial. Because $\frac{1}{\log 10}$ is merely a constant scalar, we apply the standard derivative for the natural logarithm:

$$f'(x) = \frac{1}{\log 10}(\log x)' = \frac{1}{\log 10} \cdot \frac{1}{x} = \frac{1}{x \log 10}.$$

□

Exercise 3.10.3. Test your intuition by generalizing the previous examples. Find the derivative of:

1. $f(x) = \log_a x$, where $a > 0$ and $x > 0$.
2. $f(x) = a^x$, where $a > 0$.

We now confront a more stubborn class of functions. What happens when the base *and* the exponent are both variables? Consider $f(x) = (x^2 + 1)^{\sin x}$. The power rule requires a constant exponent. The exponential rule requires a constant base. Here, both rules fail entirely.

> *Intuition. Log as an Algebraic Hammer*
>
> When dealing with variables in both the base and the exponent, the logarithm acts as an algebraic "lowering operator." By taking the logarithm of both sides, we convert a tricky exponentiation into a more manageable multiplication, perfectly poised for the product rule. This is the essence of logarithmic differentiation.

Example 3.10.4. Find the derivative of $f(x) = (x^2 + 1)^{\sin x}$.

Solution. Take the natural logarithm of both sides to bring the trigonometric exponent down to the baseline:

$$\log f(x) = \log \left[(x^2 + 1)^{\sin x}\right] = (\sin x) \log(x^2 + 1).$$

We now differentiate implicitly with respect to x. The left side requires the chain rule, while the right side demands the product rule.

$$\begin{aligned}
(\log f(x))' &= \left[(\sin x) \log(x^2 + 1)\right]' \\
\frac{1}{f(x)} \cdot f'(x) &= (\sin x)' \log(x^2 + 1) + \sin x \left(\log(x^2 + 1)\right)' \\
\frac{f'(x)}{f(x)} &= (\cos x) \log(x^2 + 1) + (\sin x) \cdot \frac{1}{x^2 + 1} \cdot (x^2 + 1)' \\
\frac{f'(x)}{f(x)} &= (\cos x) \log(x^2 + 1) + \frac{\sin x}{x^2 + 1}(2x).
\end{aligned}$$

Multiplying both sides by $f(x)$ yields:

$$f'(x) = f(x) \left[(\cos x) \log(x^2 + 1) + \frac{2x \sin x}{x^2 + 1}\right].$$

Finally, we substitute the original expression for $f(x)$:

$$f'(x) = (x^2 + 1)^{\sin x} \left[(\cos x) \log(x^2 + 1) + \frac{2x \sin x}{x^2 + 1}\right].$$

□

Exercise 3.10.5. Suppose $f(x)$ and $g(x)$ are differentiable functions such that $f(x) > 0$. Show your mastery of logarithmic differentiation by finding the derivative of:

$$F(x) = [f(x)]^{g(x)}.$$

⚠ *Caveat. The Tower of Powers*

Students are often tempted to differentiate $F(x) = f(x)^{g(x)}$ by treating it either as a power function (bringing down the $g(x)$) or an exponential function. It is neither! The base and the exponent are both moving targets. Logarithmic differentiation is the only robust way to tame this beast.

The classic identities of logarithms—namely $\log(ab) = \log a + \log b$ and $\log(a/b) = \log a - \log b$—are not just theoretical niceties; they are practical imperatives. Before diving into calculus, one should always ask if algebra can simplify the landscape.

Example 3.10.6. Find the derivative of $y = \frac{x^{2/3}\sqrt{x^2+1}}{(2x+1)^3}$.

Solution. We are faced with a choice of strategy.

1. *Approach 1: Quotient Rule Immediately (Good luck!).* Differentiating this heinous rational function would require nested applications of the quotient rule, product rule, and chain rule. While mathematically valid, this path is an exercise in masochism and invites countless algebra errors.

2. *Approach 2: Simplify by Log First.* We can evaluate $\frac{\mathrm{d}y}{\mathrm{d}x}$ by observing that

$$\log y = \frac{2}{3}\log x + \frac{1}{2}\log(x^2+1) - 3\log(2x+1),$$

and differentiating gives

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2}{3x} + \frac{x}{x^2+1} - \frac{6}{2x+1},$$

which we solve to get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x^{2/3}\sqrt{x^2+1}}{(2x+1)^3}\left(\frac{2}{3x} + \frac{x}{x^2+1} - \frac{6}{2x+1}\right).$$

□

☯ *Moral. Look Before You Leap*

Never underestimate the power of preprocessing your functions algebraically. The logarithm is particularly adept at shattering complex fractions into a sequence of gentle linear terms.

3.11 Where do I begin?

In this section, we attempt to address an imminent question of many calculus students:

I want to find the derivative of a given function f. Where do I start?

To answer this, we must adopt the mindset of a structural engineer. Every elementary function f, no matter how intimidating, is ultimately a combination of a small, finite set of atomic building blocks. If we understand how to differentiate the atoms, and we understand the "grammar" by which these atoms are combined, we can differentiate anything. It is a beautiful application of structural induction.

Let us look at the fundamental basis of our function spaces. Any elementary function f is a combination of what we might call basic functions. Think of these as the fundamental particles of calculus. The standard repertoire includes:

- Polynomial and rational powers: cx^n, $\sqrt[m]{x}$, ...
- Trigonometric functions: $\sin x$, $\cos x$, $\tan x$, ...
- Exponentials: e^x, a^x, ...
- Inverse trigonometric functions: $\sin^{-1} x$, $\cos^{-1} x$, ...
- Logarithms: $\log x$, $\log_{10} x$, ...

How do we build a universe of functions from just these few atoms? We combine them. If f is not a single basic function, it is built via algebraic or compositional glues.

> *Intuition. The Calculator Heuristic*
>
> Given an arbitrary function f, how do we know which rule to apply first? *Read from left to right like you are inputting the expression into a pocket calculator.* Identify the outermost operation acting on the constituent parts.

We proceed by analyzing the structure of f:

1. *Can I simplify the function?* For example, canceling common factors, log tricks, etc.
2. *Is it a basic function?*
   If yes, we apply our known catalog of fundamental derivatives.

   Example 3.11.1. Let $f(x) = 2x^3$. This is a basic power function scaled by a constant. We apply the power rule directly:

   $$f'(x) = 2 \cdot (3x^2) = 6x^2.$$

3. *Is it a combination of basic functions?*
   If the function is combination, we identify the mathematical "glue" holding it together. The operation dictates the differential rule:
   - If you see $\pm$ (Addition/Subtraction) $\longrightarrow$ **Sum Rule**

- If you see $\times$ (Multiplication) $\longrightarrow$ **Product Rule**
- If you see $\div$ (Division) $\longrightarrow$ **Quotient Rule**
- If you see $\circ$ (Nesting/Composition) $\longrightarrow$ **Chain Rule**

> *Moral.*
>
> Differentiation is a recursive algorithm. You identify the outermost operation, apply the corresponding rule, and then recursively call the derivative operator on the inner components until you hit the basic functions.

Example 3.11.2. Compute the derivative of the function:

$$f(x) = x^3 \tan(x^2 + 1).$$

Algorithmic Derivation. We read the function left-to-right, looking at the overarching structure. The primary operation bridging the two main clusters of x's is a multiplication: x^3 multiplied by $\tan(x^2 + 1)$. Therefore, our first maneuver must be the **Product Rule**.

$$\begin{aligned} f'(x) &= \left(x^3 \tan(x^2 + 1)\right)' \\ &= (x^3)' \tan(x^2 + 1) + x^3 \left(\tan(x^2 + 1)\right)'. \end{aligned} \tag{3.32}$$

Now, we have two sub-problems: $(x^3)'$ and $(\tan(x^2 + 1))'$. The first is a trivial basic function: $(x^3)' = 3x^2$.

The second term, $\left(\tan(x^2 + 1)\right)'$, is a composition. The outer function is the tangent, and the inner function is the polynomial $x^2 + 1$. Thus, we invoke the **Chain Rule**.

Substituting these back into (3.32):

$$f'(x) = 3x^2 \tan(x^2 + 1) + x^3 \Big[\sec^2(x^2 + 1)\Big] \cdot (x^2 + 1)'.$$

We are almost done. We simply need to evaluate the derivative of the innermost term, $(x^2 + 1)'$, which relies on the **Sum Rule** applied to basic functions:

$$\begin{aligned} f'(x) &= 3x^2 \tan(x^2 + 1) + \left[x^3 \sec^2(x^2 + 1)\right] \left((x^2)' + (1)'\right) \\ &= 3x^2 \tan(x^2 + 1) + \left[x^3 \sec^2(x^2 + 1)\right] (2x + 0). \end{aligned}$$

This yields our final expression for the derivative:

$$f'(x) = 3x^2 \tan(x^2 + 1) + 2x^4 \sec^2(x^2 + 1).$$

□

> ⚠ *Caveat. The Chain Rule Pitfall*
>
> Notice how in the chain rule step, the outer derivative $\sec^2(\cdot)$ evaluates its argument at the *original* inner function $(x^2 + 1)$, not the derivative of the inner function. A

common error is writing $\sec^2(2x)$. Keep the nested layers distinct in your mind!

Exercise 3.11.3 (Exercise for the reader)**.** Apply the calculator heuristic to $g(x) = \sin(e^{x\cos x})$. What is the very first rule you should apply? What is the final derivative?

⚠ *Caveat.*

At this stage of your study, it is a good practice to write down carefully every steps. Calculating in your head to save just a few lines is a recipe for mistakes and confusion, especially when you need to double-check your work again.

3.12 Implicit Differentiation

Story. The Legacy of the Inverse Function Theorem

One of the profound lessons the Inverse Function Theorem (Theorem 3.9.1) has taught us is that we can probe the local behavior of a function—such as finding its derivative—without ever knowing its explicit algebraic formula. The key to unlocking this geometric secret lies in the Chain Rule, a bridge connecting the local behavior of variables even when they are hopelessly entangled.

Let us consider a motivational example. Suppose we want to find the equation of a tangent line to the unit circle

$$x^2 + y^2 = 1$$

passing through a particular point (x_0, y_0) that lies on the circle.

Intuition. The Naive Approach

Whenever you see an equality linking variables, the immediate instinct from early algebra is to isolate y. Naturally, one may try to express y entirely in terms of x, calculate y', and then confidently apply the standard point-slope formula:

$$y - y_0 = y'(x_0)(x - x_0).$$

While computationally straightforward in theory, there are glaring issues with this brute-force approach:

1. *Branching Realities:* If one attempts to perfectly solve for y, we are immediately forced to consider two separate branches:

$$y^2 = 1 - x^2 \implies \sqrt{y^2} = \sqrt{1 - x^2} \implies |y| = \sqrt{1 - x^2}.$$

This bifurcates our beautiful, unified geometric circle into two fragmented functions:

$$y = \begin{cases} \sqrt{1 - x^2} & \text{if } y \geq 0, \\ -\sqrt{1 - x^2} & \text{if } y < 0. \end{cases}$$

Thus, formally, one needs to painstakingly manage multiple cases merely to cover a single continuous object.

2. *Artificial Singularities:* Even worse, consider the behavior at the specific point $(x_0, y_0) = (1, 0)$. The formal derivative of $y = \sqrt{1 - x^2}$ involves a division by zero and thus algebraically does not exist. Yet, geometric intuition undeniably shows that the circle possesses a tangent line at this point—it is simply a vertical line. Our algebraic formulation has artificially manufactured a singularity!

3.12.1 Breaking Free from Explicit Formulas

A far more resilient way to deal with this situation is to return to the philosophical underpinning of the Inverse Function Theorem: we forgo the explicit formula of y entirely. Instead, regarding $y = y(x)$ dynamically as a dependent variable, we treat x as our independent variable. By directly applying the derivative operator $\frac{\mathrm{d}}{\mathrm{d}x}$ to our equation, we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(x^2 + (y(x))^2\Big) = \frac{\mathrm{d}}{\mathrm{d}x}(1).$$

Applying the chain rule gives

$$2x + 2y\frac{\mathrm{d}y}{\mathrm{d}x} = 0 \implies \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{x}{y}, \quad \text{provided } y \neq 0.$$

Thus, the point-slope formula immediately grants us the tangent line:

$$y - y_0 = \left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{(x_0,y_0)} (x - x_0) \implies y - y_0 = -\frac{x_0}{y_0}(x - x_0), \quad \text{whenever } y_0 \neq 0.$$

Notice how we no longer need the isolated formulae for y, nor do we need to manually separate the space into positive and negative branches.

However, one subtle issue remains unresolved: the special points $(x_0, y_0) = (\pm 1, 0)$. What exactly did we miss here?

> ⚠ *Caveat. The Limits of Local Parameterization*
>
> By treating x as the independent variable, we implicitly assumed the curve could be parameterized locally as a function of x. At the extreme left and right edges of the circle, this assumption breaks down vertically.

Note that by the inherent *symmetry* of x and y in the equation $x^2 + y^2 = 1$, we can equivalently regard y as the independent variable and the other as the dependent variable. Thus, if we declare $x = x(y)$, we have

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

Differentiating strictly with respect to y, we get

$$\frac{\mathrm{d}}{\mathrm{d}y}\Big((x(y))^2 + y^2\Big) = \frac{\mathrm{d}}{\mathrm{d}y}(1) \implies 2x\frac{\mathrm{d}x}{\mathrm{d}y} + 2y = 0.$$

This perfectly mirrors our earlier derivation:

$$\frac{\mathrm{d}x}{\mathrm{d}y} = -\frac{y}{x}, \quad \text{provided } x \neq 0.$$

Using the point-slope formula with our axes conceptually reversed gives

$$x - x_0 = -\frac{y_0}{x_0}(y - y_0), \quad \text{whenever } x_0 \neq 0.$$

Lo and behold! If $x_0 \neq 0$ and $y_0 \neq 0$, we have suddenly derived two seemingly different formulae for the tangent line at (x_0, y_0). But since a tangent line to a circle is geometrically unique, these algebraic formulae must agree. Indeed, we observe a harmonious verification:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{x}{y} = \frac{1}{-\frac{y}{x}} = \frac{1}{\frac{\mathrm{d}x}{\mathrm{d}y}}, \quad \text{whenever } x \neq 0 \text{ and } y \neq 0.$$

> *Moral. Charts, Atlases, and the Road Ahead*
>
> The fact that we fundamentally need four local functions (the branches $y = \pm\sqrt{1-x^2}$ and $x = \pm\sqrt{1-y^2}$) to completely cover the geometry of the circle is no accident. This forms the precise starting point of defining *charts* and an *atlas* in modern Differential Geometry—a profound realization that global topological spaces cannot always be flattened perfectly onto a single coordinate grid. Later in Calculus II, we will bypass this restriction entirely using a so-called parametric form.

3.12.2 Self-Intersecting Curves: The Folium of Descartes

Implicit differentiation truly shines when dealing with intricate equations where isolating a variable is not merely inconvenient, but algebraically impenetrable. Let us explore an algebraic curve known for self-intersection.

Example 3.12.1. Find $\frac{\mathrm{d}y}{\mathrm{d}x}$ given the relation $x^3 + y^3 = xy$, see Figure 3.10.

Solution. We implicitly differentiate both sides with respect to x, keeping a careful eye on the product rule for the right-hand side:

$$3x^2 + 3y^2\frac{\mathrm{d}y}{\mathrm{d}x} = x\frac{\mathrm{d}y}{\mathrm{d}x} + y \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y - 3x^2}{3y^2 - x}. \tag{3.33}$$

Notice that in this example, we obtain an expression for $\frac{\mathrm{d}y}{\mathrm{d}x}$ that involves *both* x and y. This entanglement is entirely typical of solutions obtained via implicit differentiation. □

Example 3.12.2. Let us uncover the exact points at which the curve $x^3 + y^3 = xy$ possesses a horizontal tangent line.

Solution. From the computation in (3.33), we deduce that $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$ if and only if the numerator vanishes and the denominator does not. This means $y = 3x^2$ and $3y^2 \neq x$.

If a point (a, b) lies on the curve and satisfies $b = 3a^2$, then we substitute this geometric constraint back into the original algebraic relation $a^3 + b^3 = ab$:

$$a^3 + (3a^2)^3 - a(3a^2) = 0 \implies a^3 + 27a^6 - 3a^3 = 0 \implies 27a^6 - 2a^3 = a^3(27a^3 - 2) = 0.$$

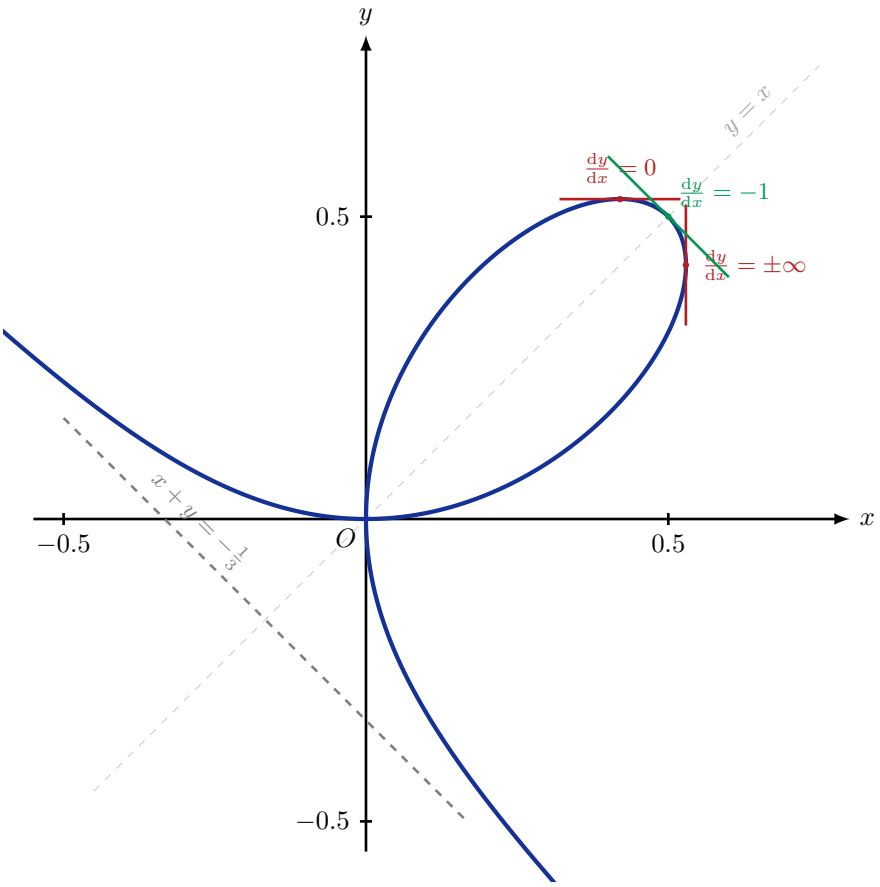


Figure 3.10 – The Folium of Descartes $x^3 + y^3 = xy$. The graph highlights points of interest for implicit differentiation, including the horizontal tangent ($\frac{\mathrm{d}y}{\mathrm{d}x} = 0$), vertical tangent ($\frac{\mathrm{d}y}{\mathrm{d}x} = \pm\infty$), and the diagonal tangent at the point of symmetry $(0.5, 0.5)$.

Solving this yields either $a = 0$ or $a = \sqrt[3]{2/27} = \frac{1}{3}\sqrt[3]{2}$.

In the first case, we have $b = 3a^2 = 0$. In the second case, we calculate $b = 3a^2 = 3 \cdot \frac{1}{9} \cdot 2^{2/3} = \frac{1}{3}2^{2/3}$.

At the origin $(a, b) = (0, 0)$, the denominator in the formula for $\frac{\mathrm{d}y}{\mathrm{d}x}$ evaluates to zero, meaning our derivative formula is undefined. However, at the point $(a, b) = \left(\frac{1}{3}2^{1/3}, \frac{1}{3}2^{2/3}\right)$, we evaluate the denominator to verify validity:

$$3b^2 - a = 3\left(\frac{1}{3}2^{2/3}\right)^2 - \frac{1}{3}2^{1/3} = 3\frac{2^{4/3}}{9} - \frac{1}{3}2^{1/3} = \frac{2}{3}2^{1/3} - \frac{1}{3}2^{1/3} = \frac{1}{3}2^{1/3} \neq 0.$$

Since the denominator cleanly resolves to a non-zero value, at this exact point we genuinely have $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$, and thus the curve manifests a true horizontal tangent line. □

Remark 3.12.3 (The Singularity at the Origin). The structural problem with the point $(0, 0)$ in the previous example is rooted in the fact that the curve crosses over itself at the origin.

Consequently, it does not have a uniquely defined tangent line there, exposing an inherent limitation in our algebraic framework when faced with geometric singularities.

Example 3.12.4. We can find the corresponding points with a vertical tangent line by exploiting the symmetry of the curve and reversing the roles of x and y.

Solution. Mimicking the logic of (3.33), we differentiate both sides of $x^3 + y^3 = xy$ with respect to y and obtain:

$$3x^2\frac{\mathrm{d}x}{\mathrm{d}y} + 3y^2 = x + y\frac{\mathrm{d}x}{\mathrm{d}y} \implies \frac{\mathrm{d}x}{\mathrm{d}y} = \frac{x - 3y^2}{3x^2 - y}.$$

A symmetric computation completely analogous to the one in Example 3.12.2 reveals that there is a vertical tangent line exactly at the mirrored point $(a, b) = \left(\frac{1}{3}2^{2/3}, \frac{1}{3}2^{1/3}\right)$. □

3.12.3 Iterating the Process: Higher Derivatives

Intuition. Bootstrapping Information

Once we have gracefully extracted the first derivative using implicit differentiation, there is nothing mathematically stopping us from treating the derivative itself as a new implicit equation and differentiating it again! This allows us to map out the concavity and deeper structural curvature without ever leaving our implicit framework.

Example 3.12.5. We can iterate this process to find higher derivatives. Let us return to the fundamental circle $x^2 + y^2 = 1$. Differentiating this gave $2x + 2yy' = 0$, which we manipulated to get $y' = -\frac{x}{y}$. We can differentiate this expression again to determine y''.

Solution. We apply the quotient rule with respect to x:

$$y'' = \frac{\mathrm{d}}{\mathrm{d}x}\left(-\frac{x}{y}\right) = -\frac{y \cdot \frac{\mathrm{d}}{\mathrm{d}x}(x) - x \cdot \frac{\mathrm{d}y}{\mathrm{d}x}}{y^2} = -\frac{y \cdot 1 - xy'}{y^2}.$$

We then substitute our known expression $y' = -\frac{x}{y}$ back into the equation to remove the y' dependence:

$$y'' = -\frac{y - x\left(-\frac{x}{y}\right)}{y^2} = -\frac{1}{y} - \frac{x^2}{y^3}.$$

□

Remark 3.12.6 (A Simplification). Notice how beautifully the innate geometry reveals itself if we take an extra second to simplify further. By finding a common denominator, we obtain

$$y'' = -\frac{y^2 + x^2}{y^3}.$$

Since our points strictly live on the unit circle, we instantly recognize the numerator as our foundational identity $x^2 + y^2 = 1$. Thus, we compactly arrive at $y'' = -\frac{1}{y^3}$. The intrinsic geometry of the space elegantly collapses the algebra!

Exercise 3.12.7 (Challenge)**.** Suppose you are given a point (x_0, y_0) that lies strictly *outside* of the unit circle $x^2 + y^2 = 1$. Using implicit differentiation to find the equations of the lines that pass through (x_0, y_0) and are tangent to the circle.

3.13 The Physical Meaning of Differentiation: Rate of Change

3.13.1 Velocity, Speed, and Acceleration

Suppose an object moves along a straight line. We can completely characterize its motion by a *position function* $s = f(t)$, which outputs the object's location at any given time t relative to a chosen reference point $s = 0$.

If we wish to understand how this object moves through space, a natural first question to ask is: *How far did it travel in a given time interval?*

The *displacement* of the object between an initial time $t = a$ and a slightly later time $t = a + \Delta t$ is precisely the difference in its positions:

$$\Delta s = f(a + \Delta t) - f(a).$$

The elapsed time for this motion is simply Δt. By taking the ratio of displacement to elapsed time, we recover an intuitive measure of how fast the object was moving on average.

Definition 3.13.1 (Average Velocity)**.** Let $s = f(t)$ be the position function of an object moving along a line. The *average velocity* of the object over the time interval $[a, a + \Delta t]$ is the displacement divided by the elapsed time:

$$v_{\text{av}} = \frac{\Delta s}{\Delta t} = \frac{f(a + \Delta t) - f(a)}{\Delta t}.$$

Geometrically, this is the slope of the secant line passing through the points $P(a, f(a))$ and $Q(a + \Delta t, f(a + \Delta t))$ on the graph of the position function.

> *Story. The Arrow of Time*
>
> Historically, the transition from *average* to *instantaneous* velocity deeply troubled classical philosophers. Zeno's arrow paradox[2] suggested that at any given indivisible "instant," an arrow in flight occupies a specific location and is therefore entirely motionless. If it is motionless at every instant, how can motion exist at all? It took the combined genius of Newton and Leibniz to realize that an "instantaneous rate" does not mean examining a single frozen snapshot; rather, it requires a mathematical limit—watching the shutter speed of our camera approach zero.

As Δt approaches 0, our measurement is calculated over smaller and smaller time intervals. The limiting value of these average velocities, when it exists, is called *instantaneous velocity*. The derivative is a mechanism we can turn twice. If velocity is the rate of change of position, what is the rate of change of velocity? We call this *acceleration*.

Definition 3.13.2 (Velocity, Speed, and Acceleration)**.** Suppose an object moves along a line with position $s = f(t)$. Then:

(i) The *velocity* at time t is $v = \frac{\mathrm{d}s}{\mathrm{d}t} = f'(t)$.

[2] See `https://en.wikipedia.org/wiki/Zeno%27s_paradoxes`. A satisfactory answer will be provided in Calculus II.

(ii) The *speed* at time t is the absolute value of velocity: $|v| = |f'(t)|$.

(iii) The *acceleration* at time t is the derivative of velocity, or the second derivative of position: $a = \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\mathrm{d}^2 s}{\mathrm{d}t^2} = f''(t)$.

> ⚠ *Caveat. Direction Matters*
>
> Velocity is a vector quantity (in one dimension, equipped with a sign). A velocity of -50 mph implies the object is moving rapidly in the negative direction. *Speed*, however, strips away this directional information. It is purely the magnitude of motion.

Example 3.13.3 (Velocity and Acceleration)**.** Suppose the position (in feet) of an object moving horizontally along a line at time t (in seconds) is modeled by the quadratic function:

$$s(t) = t^2 - 5t, \quad \text{for } 0 \leq t \leq 5.$$

We assume positive values of s correspond to positions to the right of $s = 0$. By taking derivatives, we immediately uncover the hidden dynamics of the object:

$$v(t) = s'(t) = 2t - 5$$
$$a(t) = s''(t) = 2$$

Notice that the acceleration is constant (2 ft/s^2). The object starts with a velocity of $v(0) = -5$ ft/s (moving leftward), comes to a momentary halt when $v(t) = 0 \implies t = 2.5$ s, and then begins accelerating back to the right.
Exercise for the reader: What is the total distance traveled by the object between $t = 0$ and $t = 5$? Be careful to account for the turnaround point!

3.13.2 Growth Models

Much of the change in the world around us can be classified under the broad umbrella of *growth*. Populations of species, consumer price indices, and computer networks all tend to increase in size. Mathematics is remarkably blind to the physical units we attach to our variables. The exact same formal machinery that dictates the path of a falling apple also dictates the explosive spread of a digital network.

Let $p = f(t)$ be the measure of a quantity of interest, where $t \geq 0$ represents time. The *average growth rate* of p on the interval $[a, a + \Delta t]$ is:

$$\frac{\Delta p}{\Delta t} = \frac{f(a + \Delta t) - f(a)}{\Delta t}.$$

If we let $\Delta t \to 0$, this fraction approaches the derivative $\frac{\mathrm{d}p}{\mathrm{d}t}$, which is the *instantaneous growth rate* of p with respect to time:

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \lim_{\Delta t \to 0} \frac{\Delta p}{\Delta t}.$$

Example 3.13.4 (Internet Growth)**.** The number of worldwide Internet users between 2000 and 2015 can be approximated by the mathematical model:

$$p(t) = 6t^2 + 98t + 431.2,$$

where t is measured in years after 2000, and $p(t)$ is in millions of users.

1. *Average Growth (2005 to 2010):*
   On the interval $[5, 10]$, the average rate of growth is:
   $$\frac{\Delta p}{\Delta t} = \frac{p(10) - p(5)}{10 - 5} = \frac{2011.2 - 1071.2}{5} \approx 188 \text{ million users/year.}$$

2. *Instantaneous Growth (2011):*
   The exact rate of growth at any time t is given by the derivative:
   $$p'(t) = 12t + 98.$$
   In the year 2011 ($t = 11$), the growth rate was $p'(11) = 12(11) + 98 = 230$ million users per year. Note that because $p''(t) = 12 > 0$, the growth rate itself is strictly increasing.

3. *Prediction (2020):*
   Assuming this simple polynomial model persists, the number of users in 2020 ($t = 20$) is predicted to be:
   $$p(20) \approx 6(20)^2 + 98(20) + 431.2 = 4791.2 \text{ million,}$$
   which translates to roughly 4.8 billion users.

> *Moral. A Dictionary of Derivatives*
>
> That so many words in common use denote rates of change is an eloquent testament to the importance of the derivative in how humans perceive reality. Consider this brief dictionary of familiar rates:
>
> - *Speed* $\leftrightarrow$ rate of change of distance as a function of time.
> - *Velocity* $\leftrightarrow$ rate of change of position as a function of time.
> - *Acceleration* $\leftrightarrow$ rate of change of velocity as a function of time.
> - *Angular velocity* $\leftrightarrow$ rate of change of angle as a function of time.
> - *Density* $\leftrightarrow$ rate of change of mass as a function of volume.
> - *Slope* $\leftrightarrow$ rate of change of height as a function of horizontal distance.
> - *Current* $\leftrightarrow$ rate of change of electric charge as a function of time.
> - *Marginal cost* $\leftrightarrow$ rate of change of production cost as a function of items produced.

3.13.3 Related Rates

Often, physical phenomena feature multiple variables evolving simultaneously. Because these quantities are bound together by fundamental geometric or physical laws, their rates of change cannot evolve independently. They are tethered.

To unpack these *related rates*, we rely entirely on the Chain Rule. The goal is simple: differentiate an established geometric or physical identity with respect to time t.

Example 3.13.5 (Spreading Oil Spill)**.** An oil rig springs a leak in calm seas, and the oil spreads outward in a perfectly circular patch. If the radius of the oil patch increases at a constant rate of 30 m/hr, how fast is the area of the patch increasing precisely when the patch has a radius of 100 meters?

Two variables are locked in a dance over time: the radius $r(t)$ and the area $A(t)$. The geometric law binding them is the area of a circle:

$$A(t) = \pi[r(t)]^2. \tag{3.34}$$

We are given the rate of the radius, $r'(t) = 30$, and asked to find $A'(t)$ at the instant when $r(t) = 100$. By differentiating both sides of (3.34) with respect to time t via the Chain Rule, we obtain:

$$\begin{aligned} A'(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\pi[r(t)]^2\right) \\ &= \pi\left(2r(t) \cdot r'(t)\right) \\ &= 2\pi r(t) r'(t). \end{aligned}$$

Substituting the instantaneous values into our derived equation yields:

$$A'(t) = 2\pi(100 \text{ m})\left(30\frac{\text{m}}{\text{hr}}\right) = 6000\pi\frac{\text{m}^2}{\text{hr}}.$$

The area is expanding at an astonishing rate of roughly $18,850$ m^2/hr. Notice that while the radial growth is constant, the areal growth scales linearly with the radius. The larger the spill, the faster the area grows!

> ⚠ *Caveat. Checking Units*
>
> Always include units when calculating related rates. It is an incredibly powerful, built-in error-checking mechanism. In the oil spill example, we multiplied meters (m) by meters per hour (m/hr), resulting in square meters per hour (m^2/hr). This perfectly matches our physical intuition for the rate of change of an *area*. If the units fail to match, the math is wrong.

Let us raise the stakes and consider an implicit relationship binding three separate functions of time.

Example 3.13.6 (Converging Airplanes)**.** Two small planes approach an airport. One flies due west at 120 mi/hr and the other flies due north at 150 mi/hr. Assuming they fly at the same constant elevation, how fast is the straight-line distance between the planes changing when the westbound plane is 180 miles from the airport and the northbound plane is 225 miles from the airport?

Let $x(t)$ and $y(t)$ denote the distance from the airport to the westbound and northbound planes, respectively. Because their paths are orthogonal, they form the legs of a right triangle. The distance between them, the hypotenuse, is $z(t)$. By the Pythagorean theorem:

$$[z(t)]^2 = [x(t)]^2 + [y(t)]^2. \tag{3.35}$$

Differentiating implicitly with respect to t gives:

$$2z\frac{\mathrm{d}z}{\mathrm{d}t} = 2x\frac{\mathrm{d}x}{\mathrm{d}t} + 2y\frac{\mathrm{d}y}{\mathrm{d}t}.$$

Dividing by $2z$, we isolate our target variable, the rate of change of distance between the planes:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{x\frac{\mathrm{d}x}{\mathrm{d}t} + y\frac{\mathrm{d}y}{\mathrm{d}t}}{z}. \tag{3.36}$$

Before we substitute, we must calculate the value of the hypotenuse z at the moment of interest:

$$z = \sqrt{180^2 + 225^2} = \sqrt{32400 + 50625} = \sqrt{83025} \approx 288 \text{ mi}.$$

Now, we insert the known quantities into (3.36):

$$\begin{aligned} \frac{\mathrm{d}z}{\mathrm{d}t} &= \frac{(180)(-120) + (225)(-150)}{288} \\ &\approx \frac{-21600 - 33750}{288} \approx -192 \text{ mi/hr}. \end{aligned}$$

The distance between the planes is decreasing at a rate of approximately 192 mi/hr.

> ⚠ *Caveat. The Sign of the Times*
>
> Notice the critical inclusion of the negative signs on the velocities: $\frac{\mathrm{d}x}{\mathrm{d}t} = -120$ and $\frac{\mathrm{d}y}{\mathrm{d}t} = -150$. Why? Because $x(t)$ and $y(t)$ measure the *distance to the origin.* Since the planes are approaching the airport, these distances are strictly shrinking. A common trap for the unwary student is assigning a positive rate of change to a decreasing quantity. Let the physical reality dictate the algebraic signs!

Chapter 4

Applications of the Derivative

> Mathematics is not about numbers, equations, computations, or algorithms: it is about understanding.
>
> William Paul Thurston

4.1 l'Hôpital's Rule: The Limits of Indeterminacy

Story. The Marquis and the Mercenary

What we call *l'Hôpital's Rule* was not actually discovered by the Marquis de l'Hôpital. It was discovered by the brilliant Swiss mathematician Johann Bernoulli. The Marquis, a wealthy French nobleman with a deep passion for the newly invented calculus, effectively hired Bernoulli to be his private tutor and send him mathematical discoveries. In exchange for a handsome salary, Bernoulli agreed to hand over his intellectual property rights. When l'Hôpital published the first-ever textbook on differential calculus in 1696, he included this magical rule for resolving limits. History remembered the publisher, and Bernoulli's rule has been l'Hôpital's ever since.

Calculating limits is usually a straightforward affair: we evaluate the function at the point in question. However, we often encounter situations where plugging in the value yields a result that lacks mathematical meaning. These are known as *indeterminate forms*. Let us quickly recall the primary indeterminate forms from Section 2.10:

1. $\frac{0}{0}$ and $\frac{\infty}{\infty}$ (The quotients)
2. $0 \cdot \infty$ (The product)
3. $\infty - \infty$ (The competing infinities)
4. $0^0, \infty^0$, and 1^∞ (The exponential paradoxes)

Before we deploy heavy machinery, let us remind the basic tools we used back in the day. Consider the limit:

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

Plugging in $x=2$ yields $\frac{0}{0}$. This hints us to eliminate the problematic $x-2$ factor:

$$\lim_{x\to 2} \frac{(x-2)(x+2)}{x-2} = \lim_{x\to 2}(x+2) = 2+2 = 4.$$

Similarly, consider the classical geometric limit:

$$\lim_{x\to 0} \frac{\sin x}{x} = 1.$$

The fraction evaluates to $\frac{0}{0}$. Since we cannot factor out x from $\sin x$ as in the previous example, establishing the limit originally required a highly sophisticated argument using a geometric argument combining with the Squeeze Theorem, see Corollary 2.9.3. But what if the functions are too complex for simple algebraic factorization or geometric squeezing? This is where Bernoulli's ingenious insight comes in.

Theorem 4.1.1 (L'Hôpital's Rule for $\frac{0}{0}$)**.** *Let $f(x)$ and $g(x)$ be differentiable at x_0, and suppose that:*

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0.$$

Then,

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)},$$

provided the limit on the right-hand side exists.

> *Intuition. A Tale of Linear Approximations*
>
> To truly understand *why* this works, we must look past the symbols and view the functions locally. By Newton's approximation (or the first-order Taylor expansion), any differentiable function behaves like a straight line near the point of tangency.

Proof Sketch. Let us utilize the linear approximation of f and g near x_0. Because the functions are differentiable at x_0, they are continuous, which guarantees $f(x_0)=0$ and $g(x_0)=0$.

We can expand the functions as follows, where $o(x-x_0)$ represents an error term that goes to zero faster than $(x-x_0)$ (see Section 3.7):

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + o(x-x_0)$$
$$g(x) = g(x_0) + g'(x_0)(x-x_0) + o(x-x_0)$$

Since $f(x_0)=0$ and $g(x_0)=0$, substituting these expansions into our limit yields:

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x_0)(x-x_0)+o(x-x_0)}{g'(x_0)(x-x_0)+o(x-x_0)}$$
$$= \lim_{x\to x_0} \frac{f'(x_0) + \frac{o(x-x_0)}{x-x_0}}{g'(x_0) + \frac{o(x-x_0)}{x-x_0}}.$$

As $x \to x_0$, the terms $\frac{o(x-x_0)}{x-x_0}$ vanish to 0. Thus, the expression simplifies to:

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)}.$$

□

Geometrically, if two functions cross the origin together, the ratio of their values near the origin is entirely determined by the ratio of the slopes of their tangent lines, see Figure 4.1.

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{\text{slope of } f(x) \text{ at } x_0}{\text{slope of } g(x) \text{ at } x_0}$$

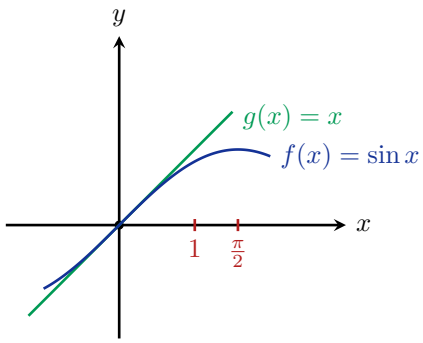


Figure 4.1 – Geometric illustration of the limit $\lim_{x\to 0} \frac{\sin x}{x} = 1$ using the tangent slopes of $f(x) = \sin x$ and $g(x) = x$ at the origin.

Armed with l'Hôpital's Rule, let us revisit our earlier examples.

Example 4.1.2. For $f(x) = x^2 - 4$ and $g(x) = x - 2$, both approach 0 as $x \to 2$.

$$\lim_{x\to 2} \frac{x^2-4}{x-2} \overset{\text{l'H}}{=} \lim_{x\to 2} \frac{(x^2-4)'}{(x-2)'} = \lim_{x\to 2} \frac{2x-0}{1-0} = \lim_{x\to 2} 2x = 4.$$

Example 4.1.3. For the trigonometric limit, $f(x) = \sin x$ and $g(x) = x$:

$$\lim_{x\to 0} \frac{\sin x}{x} \overset{\text{l'H}}{=} \lim_{x\to 0} \frac{(\sin x)'}{(x)'} = \lim_{x\to 0} \frac{\cos x}{1} = \cos 0 = 1.$$

Example 4.1.4. Let us apply the rule to a new function:

$$\lim_{x\to 1} \frac{\log x}{x^2-1} \to \frac{0}{0}.$$

Applying l'Hôpital's Rule:

$$\lim_{x\to 1} \frac{(\log x)'}{(x^2-1)'} = \lim_{x\to 1} \frac{\frac{1}{x}}{2x} = \lim_{x\to 1} \frac{1}{2x^2} = \frac{1}{2}.$$

⚠ *Caveat. L'Hôpital is NOT the Quotient Rule!*

When applying l'Hôpital, we are taking the derivative of the top and the bottom *independently*: $\lim \frac{f'(x)}{g'(x)}$. Do not confuse this with the derivative of the fraction itself:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.$$

They serve different purposes: l'Hôpital rule is for finding *limit*, quotient rule is for finding *derivative*.

A sibling of the $\frac{0}{0}$ rule is:

Theorem 4.1.5 (L'Hôpital's Rule for $\frac{\infty}{\infty}$)**.** *Let $f(x)$ and $g(x)$ be differentiable at x_0, and suppose that $\lim_{x\to x_0} f(x) = \infty$ and $\lim_{x\to x_0} g(x) = \infty$. Then*

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)},$$

provided the last limit exists.

What about the other indeterminate forms? The strategy is to mathematically bend and twist these exotic forms until they look like our comfortable $\frac{0}{0}$ or $\frac{\infty}{\infty}$ cases.

Example 4.1.6. Let us evaluate a $0 \cdot \infty$ form using log properties.

$$\lim_{x\to\infty} x \log\left(\frac{x+1}{x-1}\right) \to \infty \cdot 0.$$

We rewrite the product as a quotient to invoke l'Hôpital for $\frac{0}{0}$:

$$\lim_{x\to\infty} x \log\left(\frac{x+1}{x-1}\right) = \lim_{x\to\infty} \frac{\log\left(\frac{x+1}{x-1}\right)}{\frac{1}{x}}.$$

Now applying l'Hôpital's Rule. The derivative of the numerator requires the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}x} \log\left(\frac{x+1}{x-1}\right) = \frac{x-1}{x+1} \cdot \left(\frac{(x+1)'(x-1) - (x+1)(x-1)'}{(x-1)^2}\right) = \frac{x-1}{x+1} \cdot \frac{-2}{(x-1)^2} = \frac{-2}{x^2-1}.$$

The derivative of the denominator is $-\frac{1}{x^2}$. Substituting these back:

$$\lim_{x\to\infty} x \log\left(\frac{x+1}{x-1}\right) = \lim_{x\to\infty} \frac{\frac{-2}{x^2-1}}{-\frac{1}{x^2}} = \lim_{x\to\infty} \frac{2x^2}{x^2-1}.$$

Applying l'Hôpital once more (for the $\frac{\infty}{\infty}$ form):

$$\lim_{x\to\infty} x \log\left(\frac{x+1}{x-1}\right) \overset{\text{l'H}}{=} \lim_{x\to\infty} \frac{4x}{2x} = 2.$$

Example 4.1.7 (Factoring out the Infinity: $\infty - \infty$)**.** Suppose $\lim_{x\to x_0} f(x) = \infty$ and $\lim_{x\to x_0} g(x) = \infty$. The limit $\lim_{x\to x_0}[f(x) - g(x)]$ evaluates to $\infty - \infty$. We resolve this by forcing a quotient:

$$\lim_{x\to x_0}(f(x)-g(x)) = \lim_{x\to x_0} f(x)g(x)\left(\frac{1}{g(x)} - \frac{1}{f(x)}\right) = \lim_{x\to x_0} \frac{\left(\frac{1}{g(x)} - \frac{1}{f(x)}\right)}{\frac{1}{f(x)g(x)}} \to \frac{0}{0}.$$

To illustrate, consider the following limit that requires a little determination and multiple rounds of l'Hôpital:

$$\begin{aligned}
\lim_{x\to 0}\left(\frac{1}{x^2} - \frac{1}{\sin^2 x}\right) &= \lim_{x\to 0} \frac{\sin^2 x - x^2}{x^2 \sin^2 x} \\
&= \lim_{x\to 0} \frac{2\sin x\cos x - 2x}{2x\sin^2 x + x^2 2\sin x\cos x} \\
&= \lim_{x\to 0} \frac{\sin 2x - 2x}{2x\sin^2 x + x^2\sin 2x} \\
&= \lim_{x\to 0} \frac{2\cos 2x - 2}{2\sin^2 x + 2x2\sin x\cos x + 2x\sin 2x + 2x^2\cos 2x} \\
&= \lim_{x\to 0} \frac{\cos 2x - 1}{\sin^2 x + 2x\sin 2x + x^2\cos 2x} \\
&= \lim_{x\to 0} \frac{-2\sin 2x}{2\sin x\cos x + 2\sin 2x + 4x\cos 2x + 2x\cos 2x - 2x^2\sin 2x} \\
&= \lim_{x\to 0} \frac{-2\sin 2x}{(3-2x^2)\sin 2x + 6x\cos 2x} \\
&= \lim_{x\to 0} \frac{-4\cos 2x}{-4x\sin 2x + 2(3-2x^2)\cos 2x + 6\cos 2x + 6x\sin 2x} \\
&= \frac{-4\cdot 1}{2\cdot 3\cdot 1 + 6\cdot 1} = -\frac{4}{12} = -\frac{1}{3}.
\end{aligned}$$

Example 4.1.8 (The Log Trick: 1^∞, 0^0, ∞^0)**.** Exponential indeterminacies require us to bring the exponent down to the baseline. Suppose $\lim f(x) = 1$ and $\lim g(x) = \infty$. We are evaluating $y = \lim f(x)^{g(x)} \to 1^\infty$. Take the natural logarithm:

$$\log y = \lim_{x\to x_0} g(x)\log f(x).$$

This transforms 1^∞ into the multiplicative form $\infty\cdot 0$, which we can then write as:

$$\frac{\log f(x)}{\frac{1}{g(x)}} \to \frac{0}{0}.$$

Exercise 4.1.9. Find the following limits, paying close attention to which indeterminate form is present:

1. $\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x$
2. $\lim_{x\to 0}\frac{e^{2x}-1}{x}$

3. $\lim_{x\to-\infty} \dfrac{x^2}{\mathrm{e}^{-x}}$
4. $\lim_{x\to\infty} \mathrm{e}^{-x}\sqrt{x}$

A common pitfall for undergraduates is treating l'Hôpital's Rule as a magical hammer where every limit is a nail. Consider the following seemingly innocent limit:

$$\lim_{x\to\infty} \frac{\sqrt{x}}{\sqrt{x+1}} \to \frac{\infty}{\infty}.$$

Let us apply l'Hôpital blindly:

$$\lim_{x\to\infty} \frac{(\sqrt{x})'}{(\sqrt{x+1})'} = \lim_{x\to\infty} \frac{\frac{1}{2}x^{-1/2}}{\frac{1}{2}(x+1)^{-1/2}} = \lim_{x\to\infty} \frac{\sqrt{x+1}}{\sqrt{x}}.$$

Notice the numerator and denominator swapped. If we apply the rule again:

$$\overset{\text{l'H}}{=} \lim_{x\to\infty} \frac{\sqrt{x}}{\sqrt{x+1}}.$$

We are trapped in an infinite loop! How do we escape? By remembering our basic algebra.

> ☯ *Moral.*
>
> **Always simplify your expression first.** If we simply rewrite the original limit algebraically, the problem evaporates:
>
> $$\lim_{x\to\infty} \frac{\sqrt{x}}{\sqrt{x+1}} = \lim_{x\to\infty} \sqrt{\frac{x}{x+1}} = \lim_{x\to\infty} \sqrt{\frac{1}{1+\frac{1}{x}}} = \sqrt{\frac{1}{1+0}} = 1.$$

Note. L'Hôpital's rule also perfectly holds for one-sided limits (e.g., $x \to 0^+$). Keep this in mind as you work through these problems.

Exercise 4.1.10. Find the limits of the following indeterminate forms:

1. $\lim_{x\to1^+} \left(\dfrac{1}{\log x} - \dfrac{1}{x-1}\right)$
2. $\lim_{x\to0^+} x\log x$
3. $\lim_{x\to\frac{\pi}{2}^-} (\sec x - \tan x)$
4. $\lim_{x\to0^+} x^x$
5. $\lim_{x\to0^+} (\sin x)^x$

4.2 The Extreme Value Theorem

In the study of calculus, we often focus on the *local*—the derivative as a slope at a point, or the tangent line as a linear approximation. But in the real world—in engineering, economics, and physics—we are almost always obsessed with the *global*. Where is the greatest stress on this bridge? What is the minimum energy configuration of this molecule? What is the maximum profit?

The Extreme Value Theorem (EVT) is our bridge from the local to the global. It guarantees that if a function is *well-behaved* (continuous) on a *closed room* (a closed, bounded interval), then it *must* have reached a top and a bottom. It tells us that our search for these values is not in vain.

> *Intuition. The Mountain Climbing Metaphor*
>
> Think of the graph of f as a mountain range. If you are standing at the base a and you walk to b without teleporting (continuity), you must have reached a highest peak and a lowest valley. If the interval were open, say (a, b), you could potentially walk forever toward a boundary without ever reaching it (like $f(x) = 1/x$ on $(0, 1)$), failing to attain a maximum.

Definition 4.2.1 (Global Extrema)**.** Let $f\colon D \to \mathbb{R}$ be a real-valued function defined on a subset $D \subseteq \mathbb{R}$.

(i) We say that f attains an *absolute maximum* (or *global maximum*) at a point $c \in D$ if

$$f(c) \geq f(x) \quad \text{for all } x \in D.$$

In this case, the scalar $f(c)$ is called the *absolute maximum value* of f on D.

(ii) Similarly, f attains an *absolute minimum* (or *global minimum*) at $c \in D$ if

$$f(c) \leq f(x) \quad \text{for all } x \in D,$$

and $f(c)$ is called the *absolute minimum value* of f on D.

Collectively, the global maximum and minimum values of f are known as the *extreme values* of the function.

Example 4.2.2 (Domain Dependence of Extrema)**.** The existence of global extrema is fundamentally constrained by the topology of the underlying domain:

• Consider the identity function $f(x) = x$ restricted to the compact (i.e., closed and bounded) interval $\mathrm{dom}(f) = [-1, 1]$. Since $-1 \leq x \leq 1$ for all $x \in [-1, 1]$, f achieves:

$$\min_{x\in[-1,1]} f(x) = f(-1) = -1, \qquad \max_{x\in[-1,1]} f(x) = f(1) = 1.$$

• If we instead consider the identity function $f(x) = x$ on the unbounded domain $\mathrm{dom}(f) = (-\infty, \infty)$, the function possesses neither an absolute maximum nor an absolute minimum because $\lim_{x\to\pm\infty} f(x) = \pm\infty$.

- Consider $f(x) = x^2$ on $\mathrm{dom}(f) = (-\infty, \infty)$. Because $x^2 \geq 0 = f(0)$ for all $x \in \mathbb{R}$, we have

$$\min_{x \in (-\infty, \infty)} f(x) = f(0) = 0,$$

while no absolute maximum exists.

- On the compact domain $\mathrm{dom}(f) = [-1, 1]$, the function $f(x) = x^2$ satisfies

$$\min_{x \in [-1,1]} f(x) = f(0) = 0, \qquad \max_{x \in [-1,1]} f(x) = f(\pm 1) = 1.$$

Here, the maximum value is attained at two distinct boundary points, $c = -1$ and $c = 1$.

- For the reciprocal function $f(x) = \frac{1}{x}$ on the non-compact, punctured domain $\mathrm{dom}(f) = [-1, 0) \cup (0, 1]$, the vertical singularity yields $\lim_{x \to 0^+} f(x) = +\infty$ and $\lim_{x \to 0^-} f(x) = -\infty$. Consequently, f attains neither a global maximum nor a global minimum.

Theorem 4.2.3 (Extreme Value Theorem)**.** *Let $f \colon [a, b] \to \mathbb{R}$ be a continuous function defined on a closed and bounded interval $[a, b]$. Then f is bounded and attains both a global maximum and a global minimum on $[a, b]$; that is, there exist points $c_{\min}, c_{\max} \in [a, b]$ such that*

$$f(c_{\min}) \leq f(x) \leq f(c_{\max}) \quad \textit{for all } x \in [a, b].$$

Theorem 4.2.3 is an existence result, in the sense that it does not tell us what are the extreme values nor where to locate them on the x-axis. Nevertheless, it require a minimal assumption: continuity on a compact interval. In practice, if we know further that our function is differentiable, Theorem 4.2.5 below provides an efficient way to evaluate the extreme values. Since differentiation is inherently a local definition, thus we first need to adapt our notion of extrema in local setting.

Definition 4.2.4 (Local Extrema)**.** Let $f \colon [a, b] \to \mathbb{R}$ and let $c \in (a, b)$.

(i) The value $f(c)$ is a *local minimum* (or *minimum relative*) of f if there exists an open interval $(c_0, c_1) \subseteq [a, b]$ containing c such that

$$f(c) \leq f(x) \quad \text{for all } x \in (c_0, c_1).$$

(ii) The value $f(c)$ is a *local maximum* (or *maximum relative*) of f if there exists an open interval $(c_0, c_1) \subseteq [a, b]$ containing c such that

$$f(c) \geq f(x) \quad \text{for all } x \in (c_0, c_1).$$

> ⚠ *Caveat. Local vs. Global*
>
> A common student error is confusing *local* extrema with *global* extrema. A local maximum is a peak in a neighborhood; a global maximum is the highest point in the entire domain, see Figure 4.2.

Theorem 4.2.5 (Fermat's Theorem on Local Extrema)**.** *If f attains a local extremum (local maximum or local minimum) at an interior point $c \in (a, b)$ and the derivative $f'(c)$ exists, then*

$$f'(c) = 0.$$

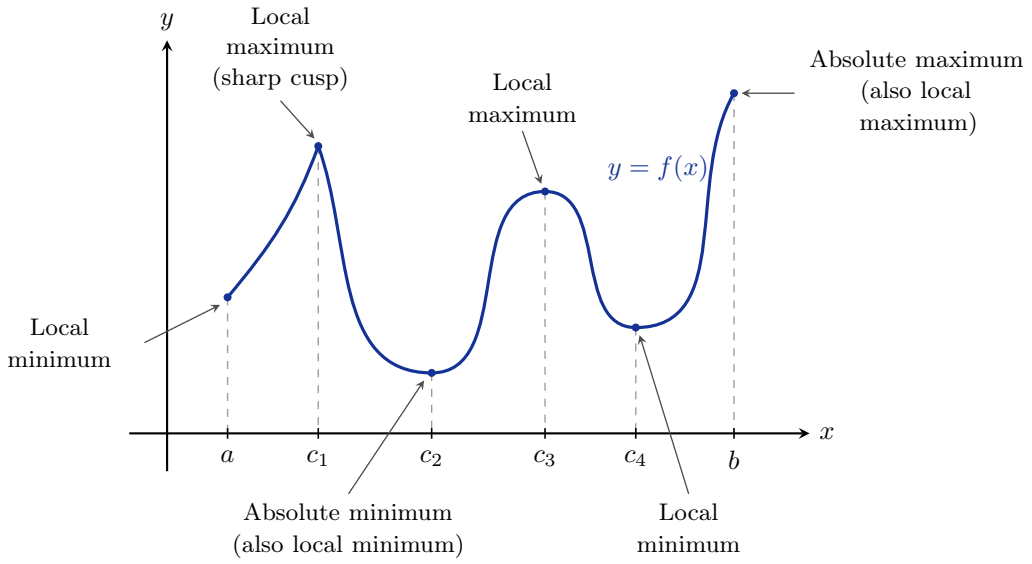


Figure 4.2 – Local and absolute extreme values of a function f on a closed interval $[a, b]$.

Definition 4.2.6 (Critical Point)**.** Let $f\colon [a,b] \to \mathbb{R}$. An interior point $c \in (a,b)$ is called a *point critical* of f if either:

$$f'(c) = 0 \quad \text{or} \quad f'(c) \text{ does not exist.}$$

⚠ *Caveat. Critical Points vs. Local Extrema*

Fermat's theorem asserts that every interior local extremum at which f is differentiable must be a critical point:

$$\text{Local extremum at } c \implies c \text{ is a critical point.}$$

The converse is strictly false:

(i) **Stationary point without an extremum:** For $f(x) = x^3$, the derivative satisfies $f'(x) = 3x^2$, so $f'(0) = 0$. However, $x = 0$ is neither a local maximum nor a local minimum, as $f(x) < 0$ for $x < 0$ and $f(x) > 0$ for $x > 0$. Instead, the origin is a horizontal inflection point.

(ii) **Non-differentiable point without an extremum:** For $f(x) = x^{1/3}$, the point $x = 0$ is a critical point because $f'(0)$ does not exist (the tangent line is vertical), yet f is strictly increasing across 0 and admits no local extremum there.

Moral. Closed Interval Method for Absolute Extrema

To locate the absolute maximum and absolute minimum of a continuous function

$f\colon [a,b] \to \mathbb{R}$:

(i) *Identify Candidates:* Find all critical points $c \in (a,b)$ of f, where $f'(c) = 0$ or $f'(c)$ fails to exist.

(ii) *Evaluate at all the Critical Points and Boundaries:* Evaluate f at all critical points in (a,b) and at the boundary endpoints $x = a$ and $x = b$.

(iii) *Compare:* Compare the computed values:

- The largest of these values is the **global maximum**.
- The smallest of these values is the **global minimum**.

Example 4.2.7 (Polynomial)**.** Find the global extrema of $f(x) = x^4 - 2x^3$ on $[-2, 2]$.

1. *Derivative:* $f'(x) = 4x^3 - 6x^2 = 2x^2(2x - 3)$.
2. *Critical Points:* $f'(x) = 0 \implies x = 0$ or $x = 3/2$. Both are in $(-2, 2)$.
3. *Evaluation:*
   - $f(-2) = (-2)^4 - 2(-2)^3 = 16 + 16 = 32$.
   - $f(0) = 0$.
   - $f(3/2) = (3/2)^4 - 2(3/2)^3 = \frac{81}{16} - \frac{54}{8} = \frac{81-108}{16} = -27/16$.
   - $f(2) = 2^4 - 2(2)^3 = 16 - 16 = 0$.

   We can organize the values into Figure 4.3, which will be extended later to include more information from the derivatives.

x	-2	0	$\frac{3}{2}$	2
f	32	0	$-\frac{27}{16}$	0

Figure 4.3 – Table of values of the function $f(x) = x^4 - 2x^3$.

Conclusion: Global max is 32 at $x = -2$; global min is $-27/16$ at $x = 3/2$.

Example 4.2.8 (Singular Derivatives)**.** Find the extreme values of $f(x) = x^{2/3}(2 - x)$ on $[-1, 2]$.

1. *Derivative:* $f'(x) = \frac{2}{3}x^{-1/3}(2 - x) + x^{2/3}(-1) = \frac{2(2-x)-3x}{3x^{1/3}} = \frac{4-5x}{3x^{1/3}}$.
2. *Critical Points:* $f'(x)$ DNE at $x = 0$. $f'(x) = 0$ at $x = 4/5$.
3. *Comparison:*
   - $f(-1) = (-1)^{2/3}(2 - (-1)) = 1 \cdot 3 = 3$.
   - $f(0) = 0$.
   - $f(4/5) = (4/5)^{2/3}(2 - 4/5) = (0.8)^{2/3}(1.2) \approx 1.03$.

x	-1	0	$\frac{4}{5}$	2
$f(x)$	1	0	≈ 1.03	0

Figure 4.4 – Table of values of the function $f(x) = x^{2/3}(2-x)$.

- $f(2) = 2^{2/3}(2-2) = 0$.

We put them in Figure 4.4:

Conclusion: Global max is 3 at $x = -1$; global min is 0 at $x = 0$ and $x = 2$.

4.3 The Mean Value Theorem

Let's motivate with an example of $f(x) = x^2 \log x$. A naive look might suggest a simple parabola-like behavior, but the logarithmic term introduces a subtle twist near the origin.

Example 4.3.1. Consider $f(x) = x^2 \log x$ on the domain $\mathrm{dom}(f) = (0, \infty)$. We compute the derivative using the product rule:

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x}(x^2)\log x + x^2 \frac{\mathrm{d}}{\mathrm{d}x}(\log x) = 2x \log x + x^2 \cdot \frac{1}{x} = 2x \log x + x = x(2\log x + 1).$$

The function $f'(x)$ is well-defined for all $x \in (0, \infty)$. Setting $f'(x) = 0$:

$$x(2\log x + 1) = 0 \implies x = 0 \text{ (excluded)} \quad \text{or} \quad 2\log x + 1 = 0.$$

Solving for x:

$$\log x = -\frac{1}{2} \implies x = \mathrm{e}^{-1/2}.$$

Evaluating the function at this point:

$$f(\mathrm{e}^{-1/2}) = (\mathrm{e}^{-1/2})^2 \log(\mathrm{e}^{-1/2}) = \mathrm{e}^{-1}\left(-\frac{1}{2}\right) = -\frac{1}{2\mathrm{e}}.$$

⚠ *Caveat.*

The calculation shows $x = \mathrm{e}^{-1/2}$ is an extreme value. However, the graph in Figure 4.5 suggests it is a global minimum. As $x \to 0^+$, $x^2 \log x \to 0$. As $x \to \infty$, $x^2 \log x \to \infty$. Since the value at the critical point is negative, it should be the global minimum. Nevertheless, to be certain that the graph does not do anything strange near infinity– which is a region we obviously can't draw, we must ask: *Does the function grow or decay at the boundaries of the domain* $(0, \infty)$*?*

To answer this question, we first need to extend our toolbox. The Mean Value Theorem (MVT) is the foundation for the next section (Section 4.4), where we will provide the rigorous justification for understanding the function's global behavior via the first derivative.

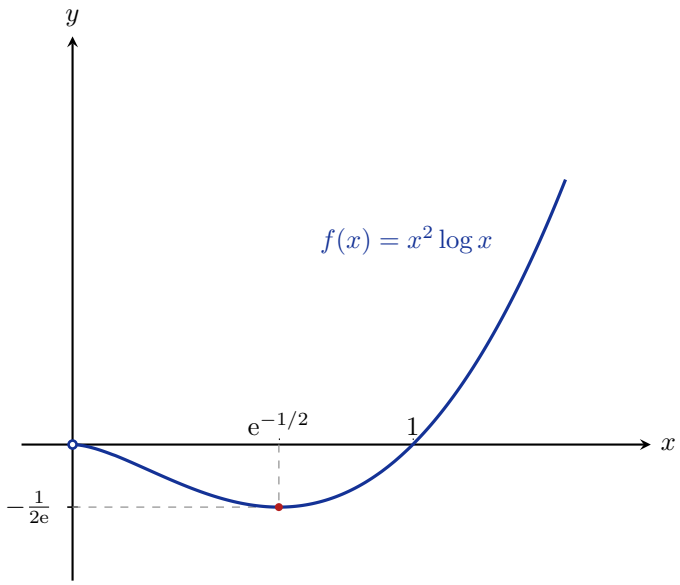


Figure 4.5 – The graph of $f(x) = x^2 \log x$ defined on $(0, \infty)$, illustrating its global minimum at $\left(e^{-1/2}, -\frac{1}{2e}\right)$ and the limit behavior at the origin.

Theorem 4.3.2 (Mean Value Theorem)**.** *Suppose f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b). Then there exists a number $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Story. A speeding problem

Think of your morning commute to campus. If you drive 18 miles in half an hour (from 7:00 AM to 8:30 AM, for example), your average velocity is 36 mph. The MVT guarantees there was at least one instant c where your speedometer read exactly 36 mph. If the speed limit is 35 mph, the math proves you were speeding!

The geometric meaning of the mean value theorem is explained in Figure 4.6. This motivate the following proof:

Proof of Theorem 4.3.2. 1. Construct the secant line $\ell(x)$ passing through $(a, f(a))$ and $(b, f(b))$.

$$\ell(x) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a).$$

2. Define the difference function $d(x) = f(x) - \ell(x)$. Note that $d(a) = 0$ and $d(b) = 0$.

3. By the Extreme Value Theorem, $d(x)$ must attain a maximum or minimum on $[a, b]$. Since $d(a) = d(b) = 0$, if d is not identically zero, the extremum must occur at some $c \in (a, b)$. At this point, $d'(c) = 0$, which implies $f'(c) = \ell'(c) = \frac{f(b)-f(a)}{b-a}$.

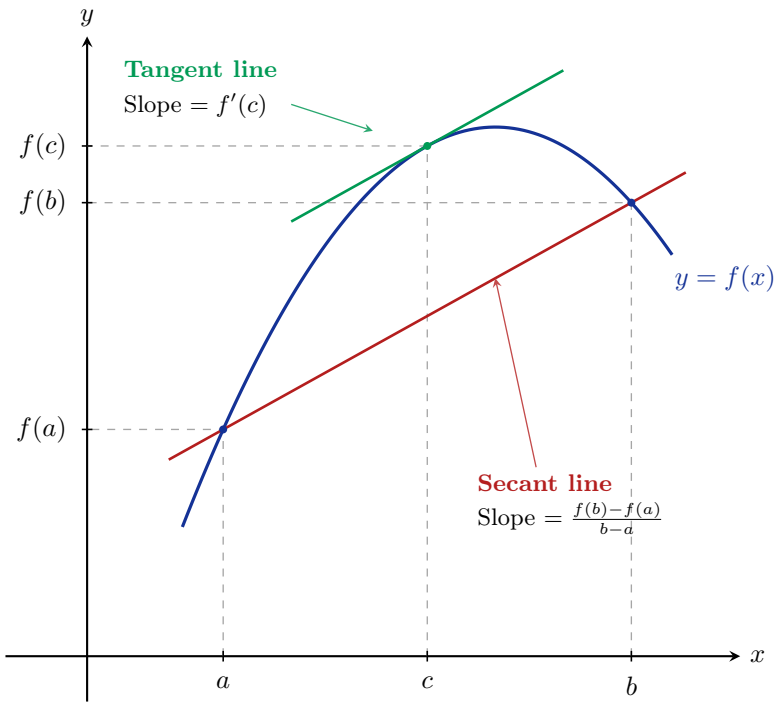


Figure 4.6 – The Mean Value Theorem states that for a continuous function f on $[a, b]$ and differentiable on (a, b), there exists a point $c \in (a, b)$ where the tangent line is parallel to the secant line.

□

> *Moral. A key of integration*
>
> If $f'(x) = 0$ for all x in an interval I, then f must be constant (prove this statement using the mean value theorem). This is the cornerstone of integration theory: the derivative identifies functions up to an additive constant C.

Here is a simple consequence of the mean value theorem:

Theorem 4.3.3 (Rolle theorem)**.** *Suppose f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b) with $f(a) = f(b)$. Then there exists a number $c \in (a, b)$ such that*

$$f'(c) = 0.$$

Proof. Exercise. □

Exercise 4.3.4. Prove that if $f'(x) = g'(x)$ for all $x \in I$, then $f(x) = g(x) + C$.

Hint: Apply the MVT to the function $h(x) = f(x) - g(x)$.

Exercise 4.3.5. What happens if $d(x)$ in the MVT proof attains its extrema at the endpoints a or b?

Hint: Consider the case where $d(x)$ is constant. What does this imply about $f(x)$?

Exercise 4.3.6. Show that $f(x) = \tan^{-1}(x)$ on $\mathbb{R}$ has no global maximum or minimum. *Hint: Consider the horizontal asymptotes at* $\pm\pi/2$.

Exercise 4.3.7. Prove that if f is continuous and $f'(x) \neq 0$ for all $x \in (a, b)$, then the global extrema must occur at the endpoints. *Hint: Use the Mean Value Theorem to show that* f *must be strictly monotonic.*

4.4 First Derivative Dictates Monotonicity

In calculus, we transition from static snapshots of functions to the dynamic understanding of how they evolve. The derivative $f'(x)$ is our primary tool for this. It answers the question: *How does the global shape of a curve emerge from the local information at each point?*

We begin by formalizing the notion of *growing*.

Definition 4.4.1 (Monotonicity)**.** Let $f : I \to \mathbb{R}$ be a function on an interval I.

1. f is *increasing* if for any $x_1, x_2 \in I$, $x_1 \leq x_2 \implies f(x_1) \leq f(x_2)$.
2. f is *strictly increasing* if $x_1 < x_2 \implies f(x_1) < f(x_2)$.
3. f is *decreasing* (resp. *strictly decreasing*) if the inequalities are reversed.

> *Intuition.*
>
> Imagine you are hiking on a mountain range. The function $f(x)$ represents your altitude at position x. The derivative $f'(x)$ is the slope of the ground beneath your feet. If $f'(x) > 0$, you are moving uphill; if $f'(x) < 0$, you are descending. The idea is that by knowing the sign of the slope everywhere, we can reconstruct the entire terrain, identifying peaks, valleys, and plateaus.

Corollary 4.4.2 (Criteria for Monotonicity)**.** *If* $f'(x) > 0$ *(reps.* $f'(x) \geq 0$*) on an interval* I*, then* f *is strictly increasing (resp. increasing) on* I*.*

If $f'(x) < 0$ *(reps.* $f'(x) \leq 0$*) on an interval* I*, then* f *is strictly decreasing (resp. decreasing) on* I*.*

Ideas of proof. Connect the local property ($f'(x)$) with the global property (the slope of an appropriate secant line) via the Mean Value Theorem. □

> *Moral. The Sign-Slope Correspondence*
>
> If the slope is positive, the function must be increasing. Why? If it weren't, there would have to be a point where it "turned back" or "dipped," which would force the derivative to be negative or zero at some point in between (by the Mean Value Theorem). Thus, the sign of the derivative is a *global* constraint imposed by *local* data.

Example 4.4.3. We back to Example 4.3.1, where we consider $f(x) = x^2 \log x$ with $\mathrm{dom}(f) = (0, \infty)$. We seek its global behavior using first derivative.

1. *Derivative calculation:* Using the product rule: $f'(x) = \frac{\mathrm{d}}{\mathrm{d}x}(x^2)\log x + x^2\frac{\mathrm{d}}{\mathrm{d}x}(\log x) = 2x\log x + x^2(\frac{1}{x}) = 2x\log x + x = x(2\log x + 1)$.

2. *Finding critical points:* f' exist everywhere on $\mathrm{dom}(f)$. We have $f'(x) = 0 \iff x(2\log x + 1) = 0$. Since $x > 0$, we solve $2\log x + 1 = 0 \implies \log x = -1/2 \implies x = \mathrm{e}^{-1/2}$.

3. *Sign analysis:* For $x < \mathrm{e}^{-1/2}$, $f'(x) < 0$ (decreasing). For $x > \mathrm{e}^{-1/2}$, $f'(x) > 0$ (increasing). We organize the calculation into Figure 4.7, with extra information from the derivative: it tells us on which interval the function decreases or increases. The double red vertical lines represent where the function does not defined (and does the derivative f').

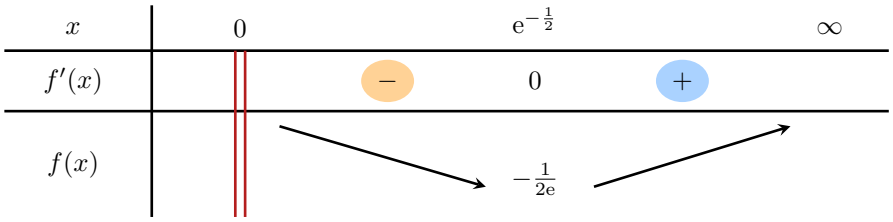


Figure 4.7 – Sign analysis table for the function $f(x) = x^2\log x$.

The function reaches a global minimum at $x = \mathrm{e}^{-1/2}$ with value $f(\mathrm{e}^{-1/2}) = (\mathrm{e}^{-1/2})^2\log(\mathrm{e}^{-1/2}) = \mathrm{e}^{-1}(-1/2) = -\frac{1}{2\mathrm{e}}$.

Example 4.4.4. Consider $f(x) = x\mathrm{e}^{-x}$ with $\mathrm{dom}(f) = (-\infty, \infty)$.

1. *Derivative:* $f'(x) = 1\cdot\mathrm{e}^{-x} + x\cdot(-\mathrm{e}^{-x}) = \mathrm{e}^{-x}(1-x)$.

2. *Critical points:* f' exist everywhere on $\mathrm{dom}(f)$. $f'(x) = 0 \implies 1 - x = 0 \implies x = 1$.

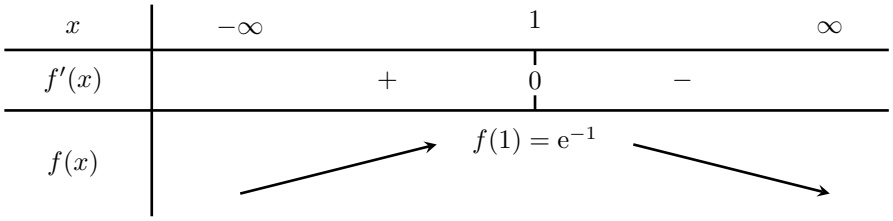


Figure 4.8 – Sign analysis table for the function $f(x)$ showing a peak at $x = 1$.

3. *Sign analysis:* At $x = 1$, we have a local maximum $f(1) = \mathrm{e}^{-1}$, see Figure 4.8.

Exercise 4.4.5. Analyze the monotonicity of $f(x) = x(x-2)^3$. *Hint: Find $f'(x)$ using the product rule. Note that $f'(x) = (x-2)^3 + 3x(x-2)^2 = (x-2)^2(x-2+3x) = (x-2)^2(4x-2)$. Locate the critical points and check the sign of $f'(x)$ in intervals $(-\infty, 1/2)$, $(1/2, 2)$, and $(2, \infty)$.*

Theorem 4.4.6 (First Derivative Test)**.** *Suppose f is continuous on an interval I, and $c \in I$ is a critical point of f.*

1. *If f' changes sign from $+$ to $-$ as x increases through c, then f has a local maximum at c.*
2. *If f' changes sign from $-$ to $+$ as x increases through c, then f has a local minimum at c.*
3. *If f' is positive on both sides near c or negative on both sides near c, then f has no local extreme value at c.*

> *Moral. Jumping over Critical Points*
>
> The test is not about the value of $f'(c)$ itself–which is zero or does not exist–but about the *change in the direction of the slope* as we cross c. It is the transition that defines the extremum.
> Thus, it is clear that figuring out the sign of f' is of utmost importance. Here is an important trick if higher derivative(s) of f exists at c. In this case, Taylor theorem (we will study in Calculus 2, which is essentially the polynomial extension of the linear approximation in Theorem 3.3.1) implies that
>
> $$f'(x) \approx a_n(x-c)^n \qquad \text{whenever } x \text{ is closed to } c,$$
>
> here $n \in \mathbb{N}$ and a_n is a constant. Thus if n is odd, then $f'(x)$ changes its sign when x jumps over c. If n is even, then $f'(x)$ keeps the same sign when x jumps over c.
> For example, in Figure 4.8, the critical point $x = 1$ is *simple*, that means $n = 1$, and f' changes sign crossing $x = 1$. ($n = 1$ is the reason we draw a single vertical line under 0 in the second line of our table.)
> Note that if $f''(c)$ does not exist, the trick does not work, and one need to pick a point inside each interval next to c to evaluate the sign of f' on such interval.

Example 4.4.7. Consider the function $f(x) = 3x^4 - 4x^3 - 6x^2 + 12x + 1$. We wish to find its intervals of increase/decrease and its local extrema.

1. *Domain:* The domain is $\text{dom}(f) = (-\infty, \infty)$.
2. *Critical Points:* The derivative is:

   $$f'(x) = 12x^3 - 12x^2 - 12x + 12$$

   Using the grouping method (or Vieta's insight):

   $$\begin{aligned} f'(x) &= 12(x^3 - x^2 - x + 1) = 12(x^2(x-1) - 1(x-1)) = 12(x-1)(x^2-1) \\ &= 12(x-1)(x-1)(x+1) = 12(x-1)^2(x+1). \end{aligned}$$

   We have f' is well-defined everywhere on $\text{dom}(f)$. $f'(x) = 0$ when $x = 1$ (double root) and $x = -1$ (single root).
3. *Sign Table:*

   From Figure 4.9, at $x = -1$, f' changes from $-$ to $+$, so we have a *local minimum* at $f(-1) = 3(-1)^4 - 4(-1)^3 - 6(-1)^2 + 12(-1) + 1 = 3 + 4 - 6 - 12 + 1 = -10$. At $x = 1$, f' does not change sign (it is positive on both sides). Thus, $x = 1$ is not an extreme (it is actually a point of *inflection*–which is studied in the next section).

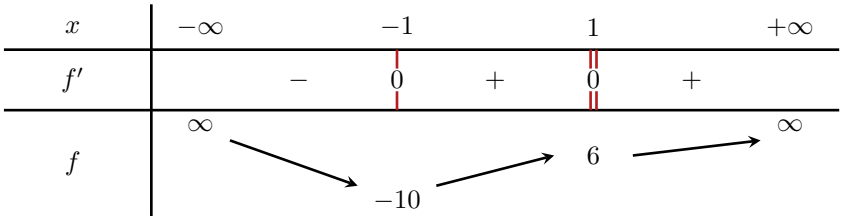


Figure 4.9 – Sign table for the quartic function f.

Exercise 4.4.8. Analyze the local extrema of $g(x) = x^{2/3}(2-x)$. Note that $g'(x)$ is undefined at $x = 0$. How does this affect your sign table?

Exercise 4.4.9 (Challenge)**.** The Fundamental Theorem of Algebra says that if f is a polynomial of degree n, then it has n complex roots. Prove that if f is a polynomial of degree n, it can have at most $n-1$ (real) critical points. Why does this limit the number of local extrema?

⚠ *Caveat.*

Always check the points where $f'(x)$ does not exist. A cusp or a vertical tangent can be a local extreme even if the derivative is not zero there!

4.5 Second Derivative Dictates Concavity

In our previous explorations of calculus, we mastered the *first derivative* f', which acts as a compass, pointing us toward the direction of *growth* or *decay*. But a compass only tells you where you are going; it does not tell you if the path is *bending* toward the heavens or *bowing* toward the earth.

Story. Equilibrium

Consider the motion of a particle. If $f(t)$ is its position, $f'(t)$ is its velocity. If the velocity is increasing, the particle is accelerating, and the trajectory bends upward. This is the essence of *concavity*. In physics, this is not merely an abstract property; it is the difference between a stable equilibrium (a potential well, where the system wants to return to the bottom) and an unstable one (a potential hill, where the system wants to roll away).

Before we touch the formal definitions, let us build a geometric intuition.

Intuition. The Tangent Line Test

Imagine placing a ruler against the graph of a function f.

- If the graph stays *above* the ruler, the function is *concave up*, see Figure 4.10.
- If the graph stays *below* the ruler, the function is *concave down*.

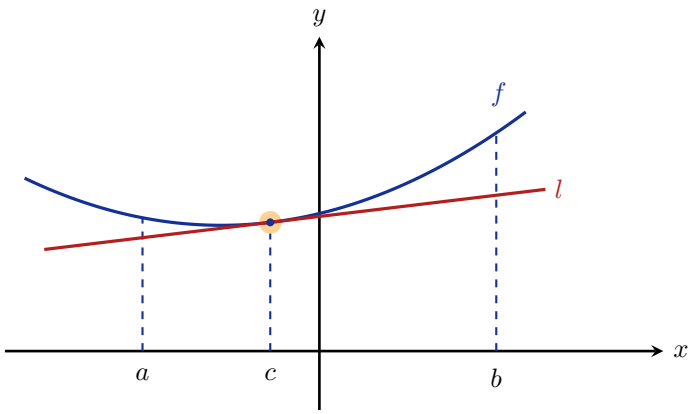


Figure 4.10 – Geometric illustration of a concave up function f using the tangent line test on an interval $[a, b]$.

Definition 4.5.1 (Concavity)**.** Let f be differentiable on an open interval I.

1. f is *concave up* (or convex) on I if f' is increasing on I.
2. f is *concave down* on I if f' is decreasing on I.

Definition 4.5.2 (Inflection Point)**.** If f is continuous at c and changes concavity at c, we say f has an *inflection point* at c.

Theorem 4.5.3 (Test for Concavity)**.** *Suppose f'' exists on an open interval I.*

- *If $f''(x) > 0$ for all $x \in I$, then f is concave up on I.*
- *If $f''(x) < 0$ for all $x \in I$, then f is concave down on I.*
- *If f'' changes sign at c, then c is an inflection point.*

⚠ *Caveat. Inflection trap*

Be careful! The condition $f''(c) = 0$ is a *necessary* but not *sufficient* condition for an inflection point. Just as a function can have a flat spot that isn't a local extremum (like $f(x) = x^3$), it can have $f''(c) = 0$ without changing concavity (like $f(x) = x^4$).

We go back to Example 4.4.7.

Example 4.5.4. Let us analyze concavity of $f(x) = 3x^4 - 4x^3 - 6x^2 + 12x + 1$.

1. *Domain:* The domain is $\text{dom}(f) = (-\infty, \infty)$.
2. *Critical Points:* The derivative is:

$$f'(x) = 12x^3 - 12x^2 - 12x + 12$$

Factorizing (with Vieta's insight):

$$f'(x) = 12(x^3 - x^2 - x + 1) = 12(x^2(x-1) - 1(x-1)) = 12(x-1)(x^2-1)$$
$$= 12(x-1)(x-1)(x+1) = 12(x-1)^2(x+1).$$

We have f' is well-defined everywhere on $\text{dom}(f)$. $f'(x) = 0$ when $x = 1$ (double root) and $x = -1$ (single root).

3. *Inflection Points:*

$$f''(x) = 12(3x^2 - 2x - 1) = 12(3x+1)(x-1)$$

Setting $f''(x) = 0$ gives $x = 1$ and $x = -1/3$.

4. *Sign Table:*

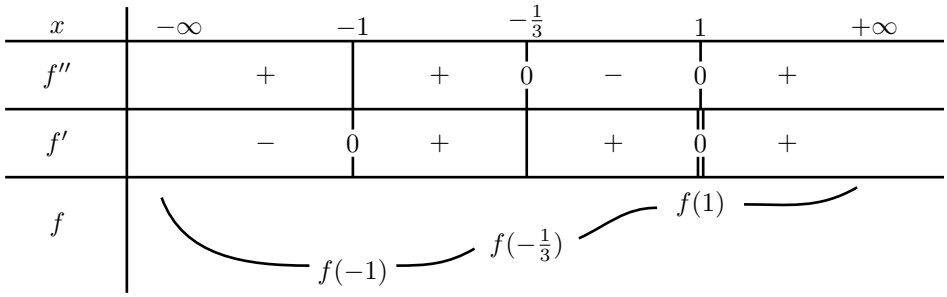


Figure 4.11 – Sign table for the quartic function f.

From Figure 4.11, at $x = -1/3$ and $x = 1$, the concavity changes, so these are *inflection points*.

Theorem 4.5.5 (Second Derivative Test)**.** *Suppose f'' is continuous on an open interval containing c, and $f'(c) = 0$.*

- *If $f''(c) > 0$, f has a local minimum at c.*
- *If $f''(c) < 0$, f has a local maximum at c.*
- *If $f''(c) = 0$, the test is inconclusive.*

Example 4.5.6. In our previous function $f(x) = 3x^4 - 4x^3 - 6x^2 + 12x + 1$:

- At $x = 1$, $f''(1) = 12(4)(0) = 0$. The test is inconclusive. (Note: Looking at the table, $x = 1$ is an inflection point, not an extremum).
- At $x = -1$, $f''(-1) = 12(-2)(-2) = 48 > 0$. Thus, $x = -1$ is a local minimum.

Exercise 4.5.7. Analyze the concavity and extrema of $f(x) = -x^3 + 9x$ on $[-4, 3]$.

Hint: Find f' and f''. Identify where $f'' = 0$ for potential inflection points and evaluate the boundary values of the interval.

Exercise 4.5.8. Consider $f(x) = x^2/(x^2 - 1)$ on $[-4, 4]$. Why does the test for concavity become difficult at $x = \pm 1$?

Hint: Examine the domain and the behavior of the limit as $x \to \pm 1$. The function is not continuous there; the second derivative test requires continuity on the interval.

4.6 Curve Sketching

Moral. Systematic Curve Sketching

Sketching the qualitative behavior of a smooth or piecewise smooth function $f\colon \mathbb{R} \to \mathbb{R}$ requires synthesizing local differential data (slopes and curvatures) with global topological and asymptotic constraints (domains, symmetries, and asymptotic lines). Rather than plotting disconnected points, we reconstruct the geometry of f systematically via:

(1) **Domain and Symmetries:** Determining $\mathrm{dom}(f)$, identifying isolated singularities, parity (even/odd), and periodicity.

(2) **Intercepts and Asymptotes:** Locating the coordinate intercepts, vertical asymptotes where $|f(x)| \to \infty$, and asymptotic lines (horizontal or oblique/slant asymptotes) governing end behaviors as $x \to \pm\infty$.

(3) **Monotonicity and Concavity:** Computing the critical points from $f'(x)$ and inflection points from $f''(x)$, subsequently compiling them into a unified sign analysis table.

Example 4.6.1. Analyze the rational function

$$f(x) \coloneqq \frac{2(x^2-9)}{x^2-4}$$

and construct its qualitative graph via a sign table.

Solution. We proceed methodically through each stage of the curve analysis.

1. *Domain and Symmetry.* The rational function f is well-defined provided its denominator does not vanish:

$$x^2 - 4 \neq 0 \iff x \neq \pm 2.$$

Consequently, the maximal domain of definition is

$$\mathrm{dom}(f) = (-\infty, -2) \cup (-2, 2) \cup (2, \infty).$$

Testing for symmetry under reflection across the origin, we observe that for every $x \in \mathrm{dom}(f)$,

$$f(-x) = \frac{2\big((-x)^2 - 9\big)}{(-x)^2 - 4} = \frac{2(x^2-9)}{x^2-4} = f(x).$$

Thus, f is an *function even*, meaning its graph is symmetric with respect to the y-axis.

2. *Intercepts and Asymptotic Behavior.*

- *x-intercepts:* Setting $f(x) = 0$ for $x \in \mathrm{dom}(f)$ yields

$$2(x^2-9) = 0 \iff x^2 - 9 = 0 \iff x = \pm 3.$$

Hence, the graph intersects the horizontal axis at $(-3, 0)$ and $(3, 0)$.

- *y-intercept:* Evaluating f at the origin gives

$$f(0) = \frac{2(0^2-9)}{0^2-4} = \frac{-18}{-4} = \frac{9}{2}.$$

Thus, the vertical intercept is located at $\big(0, \frac{9}{2}\big)$.

- *Vertical Asymptotes:* Since the denominator vanishes at $x = \pm 2$ while the numerator satisfies $2((\pm 2)^2 - 9) = -10 \neq 0$, the lines $x = -2$ and $x = 2$ may be vertical asymptotes. To verify, we examine the one-sided limits at $x = 2$:

$$\lim_{x\to 2^+} f(x) = \lim_{x\to 2^+} \frac{2(x^2-9)}{x^2-4} = \frac{2(4^+-9)}{4^+-4} = \frac{-10}{0^+} = -\infty,$$
$$\lim_{x\to 2^-} f(x) = \lim_{x\to 2^-} \frac{2(x^2-9)}{x^2-4} = \frac{2(4^--9)}{4^--4} = \frac{-10}{0^-} = +\infty.$$

By even symmetry, the behavior at $x = -2$ is reflected accordingly:

$$\lim_{x\to -2^+} f(x) = +\infty, \qquad \lim_{x\to -2^-} f(x) = -\infty.$$

- *Horizontal and Slant Asymptotes:* A general linear asymptote $y = ax + b$ satisfies

$$a := \lim_{x\to\pm\infty} \frac{f(x)}{x}, \qquad b := \lim_{x\to\pm\infty} \big(f(x) - ax\big).$$

Evaluating the slope parameter as $x \to +\infty$:

$$a = \lim_{x\to+\infty} \frac{2(x^2-9)}{x(x^2-4)} = \lim_{x\to+\infty} \frac{2x^2-18}{x^3-4x} \overset{\text{L'H}}{=} \lim_{x\to+\infty} \frac{4x}{3x^2-4} \overset{\text{L'H}}{=} \lim_{x\to+\infty} \frac{4}{6x} = 0.$$

The intercept parameter then becomes

$$b = \lim_{x\to+\infty} \big(f(x) - 0\cdot x\big) = \lim_{x\to+\infty} \frac{2x^2-18}{x^2-4} \overset{\text{L'H}}{=} \lim_{x\to+\infty} \frac{4x}{2x} = 2.$$

By symmetry, the identical computation holds as $x \to -\infty$. Therefore, $y = 2$ is a horizontal asymptote as $x \to \pm\infty$, and no non-horizontal slant asymptote exists.

3. *First Derivative and Critical Points.* Applying the quotient rule to $f(x) = \frac{2x^2-18}{x^2-4}$:

$$\begin{aligned} f'(x) &= \frac{\big(2(x^2-9)\big)'(x^2-4) - 2(x^2-9)(x^2-4)'}{(x^2-4)^2} \\ &= \frac{4x(x^2-4) - 2(x^2-9)(2x)}{(x^2-4)^2} \\ &= \frac{4x\big((x^2-4) - (x^2-9)\big)}{(x^2-4)^2} \\ &= \frac{4x(5)}{(x^2-4)^2} = \frac{20x}{(x^2-4)^2}. \end{aligned}$$

To locate the critical numbers:

(i) $f'(x) = 0 \iff 20x = 0 \iff x = 0$.

(ii) $f'(x)$ fails to exist when $(x^2-4)^2 = 0 \iff x = \pm 2$, but these points are excluded from $\mathrm{dom}(f)$.

Thus, the sole critical point in the domain is $x = 0$, corresponding to the local minimum value $f(0) = \frac{9}{2}$ on the central interval $(-2, 2)$.

4. *Second Derivative and Inflection Points.* Differentiating $f'(x) = 20x(x^2-4)^{-2}$ via the quotient rule:

$$\begin{aligned} f''(x) &= \frac{(20x)'(x^2-4)^2 - 20x\big((x^2-4)^2\big)'}{(x^2-4)^4} \\ &= \frac{20(x^2-4)^2 - 20x \cdot 2(x^2-4)(2x)}{(x^2-4)^4} \\ &= \frac{(x^2-4)\big[20(x^2-4) - 80x^2\big]}{(x^2-4)^4} \\ &= \frac{20x^2 - 80 - 80x^2}{(x^2-4)^3} \\ &= \frac{-60x^2 - 80}{(x^2-4)^3} = \frac{-20(3x^2+4)}{(x^2-4)^3}. \end{aligned}$$

Setting $f''(x) = 0$ requires $-20(3x^2+4) = 0$, which implies $3x^2+4 = 0$. Since $3x^2+4 \geq 4 > 0$ for all $x \in \mathbb{R}$, this equation possesses no real solutions (technically, one may say that it has two complex solutions $x = \pm i\frac{2}{\sqrt{3}}$, but those numbers do not belong to $\mathrm{dom}(f)$). Hence, f has no inflection points.

5. *Sign Table and Sketch.* Evaluating test points across the partitioned intervals:

- On $(-\infty, -2)$, choose $x \to -\infty$: $\lim_{x\to-\infty} f'(x) < 0$ (decreasing), $\lim_{x\to-\infty} f''(x) < 0$ (concave downward).
- On $(-2, 0)$, choose $x = -1$: $f'(-1) < 0$ (decreasing), $f''(-1) = \frac{-20(7)}{(-3)^3} > 0$ (concave upward).
- On $(0, 2)$, choose $x = 1$: $f'(1) > 0$ (increasing), $f''(1) > 0$ (concave upward).
- On $(2, \infty)$, choose $x \to \infty$: $\lim_{x\to\infty} f'(x) > 0$ (increasing), $\lim_{x\to\infty} f''(x) < 0$ (concave downward).

We summarize in Figure 4.12.

As $x \to \pm\infty$, the curve approaches the line $y = 2$ from below, crosses the x-axis at $(\pm 3, 0)$, plunges towards $-\infty$ along the outer sides of $x = \pm 2$, and forms a convex valley between $x = -2$ and $x = 2$ with a relative minimum at $\big(0, \frac{9}{2}\big)$. We sketch the function in Figure 4.13.

□

Exercise 4.6.2. Sketch the graph of each of the following functions by systematically determining their domain, intercepts, asymptotes, monotonicity, and concavity via a sign table:

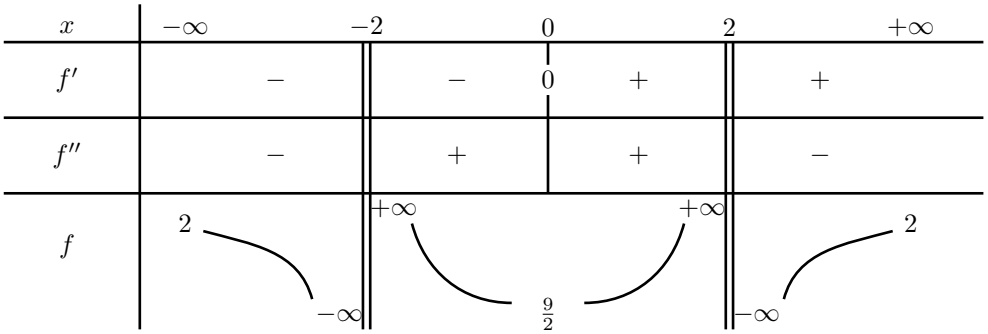


Figure 4.12 – Sign table for the function $f(x) = \frac{2(x^2-9)}{x^2-4}$.

(1) $f(x) = \mathrm{e}^{-x^2}$

(2) $f(x) = \frac{1}{8}x^{2/3}(9x^2 - 8x - 16)$

(3) $f(x) = (x-6)(x+6)^2$

(4) $f(x) = x^4 + 4x^3$

(5) $f(x) = x - 2\sqrt{x}$

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

(7) $f(x) = \frac{x^2}{x-2}$

(8) $f(x) = \frac{2x-3}{2x-8}$

(9) $f(x) = \frac{x^2}{x^2-4}$

(10) $f(x) = \arctan\left(\frac{x^2}{\sqrt{3}}\right)$

(11) $f(x) = \log(x^2+1)$

(12) $f(x) = x + 2\cos x \quad \text{on } [-2\pi, 2\pi]$

(13) $f(x) = \sin x - x \quad \text{on } [0, 2\pi]$

(14) $f(x) = 2 - 3x^{2/3}$

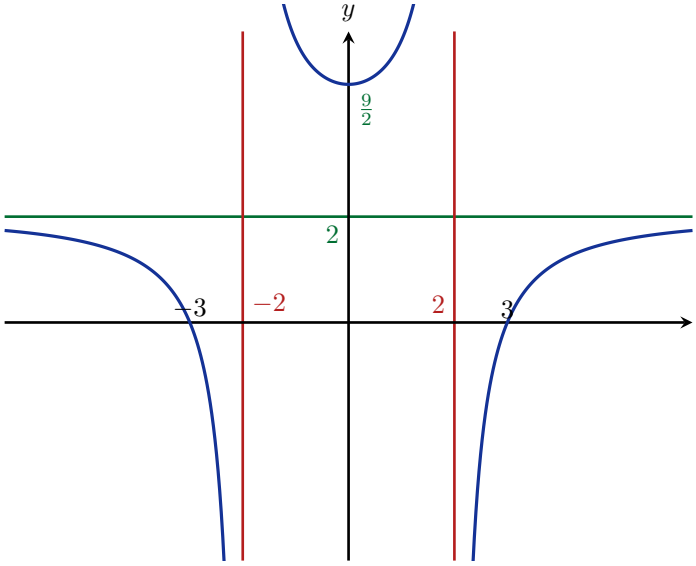


Figure 4.13 – Graph of the rational function $f(x) = \frac{2(x^2-9)}{x^2-4}$ with vertical and horizontal asymptotes.

4.7 Optimization Problems

Intuition. The Geometry of Optimization

Optimization translates geometric or physical constraints into the language of differential calculus. By expressing a quantity of interest as a single-variable function subject to bounded or physical domains, the Fermat interior extremum principle ($f'(x) = 0$) and monotonicity analysis reveal maximal configurations and fundamental analytic inequalities.

Example 4.7.1 (Isoperimetric Problem for Rectangles)**.** A rod of fixed length 20 is bent to form the perimeter of a rectangle with dimensions $x \geq 0$ and $y \geq 0$. Find the dimensions that maximize the enclosed area.

Solution. The perimeter constraint is given by

$$2(x + y) = 20 \implies x + y = 10 \implies y = 10 - x.$$

Since both dimensions must be non-negative, the permissible domain for the width x is the closed interval $[0, 10]$. The enclosed area function $A\colon [0, 10] \to \mathbb{R}$ is

$$A(x) = xy = x(10 - x) = 10x - x^2.$$

Differentiating with respect to x via the product rule:

$$A'(x) = (x)'(10 - x) + x(10 - x)' = 1(10 - x) + x(-1) = 10 - 2x.$$

Setting the derivative equal to zero to find the critical points:

$$A'(x) = 0 \iff 10 - 2x = 0 \iff x = 5.$$

We construct the sign table for $A'(x)$ on the interval $[0, 10]$ in Figure 4.14.

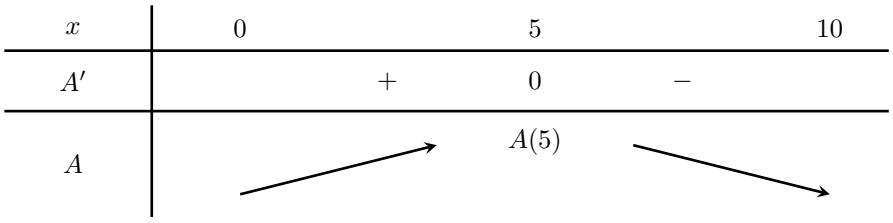


Figure 4.14 – Sign table for the area function A.

Since $A'(x) > 0$ on $(0, 5)$ and $A'(x) < 0$ on $(5, 10)$, the first derivative test guarantees that $x = 5$ achieves the absolute maximum:

$$\max_{x \in [0,10]} A(x) = A(5) = 5(10 - 5) = 25.$$

Substituting back into the constraint yields $y = 10 - 5 = 5$. Thus, the maximum area is attained when the rectangle transforms into a square with side lengths $x = y = 5$. □

Example 4.7.2 (Calculus Proof of the AM–GM Inequality for Two Variables)**.** Suppose $a, b \geq 0$. Prove the Arithmetic Mean–Geometric Mean (AM–GM) inequality:

$$\frac{a+b}{2} \geq \sqrt{ab},$$

with equality holding if and only if $a = b$.

Proof via Single-Variable Calculus. Without loss of generality, assume $0 \leq a \leq b$. If $a = 0$, the inequality simplifies to $b/2 \geq 0$, which holds trivially. Hence, we may assume $a > 0$.

To show that $\frac{a+b}{2} - \sqrt{ab} \geq 0$, consider the auxiliary function $f\colon [a, \infty) \to \mathbb{R}$ defined by

$$f(x) \coloneqq \frac{a+x}{2} - \sqrt{ax}.$$

Differentiating f with respect to x for $x \geq a$:

$$\begin{aligned} f'(x) &= \left(\frac{a+x}{2}\right)' - (\sqrt{ax})' \\ &= \frac{1}{2} - \frac{1}{2}(ax)^{-1/2} \cdot (ax)' \\ &= \frac{1}{2} - \frac{a}{2\sqrt{ax}} \\ &= \frac{1}{2}\left(1 - \frac{\sqrt{a}}{\sqrt{x}}\right) = \frac{\sqrt{x} - \sqrt{a}}{2\sqrt{x}}. \end{aligned}$$

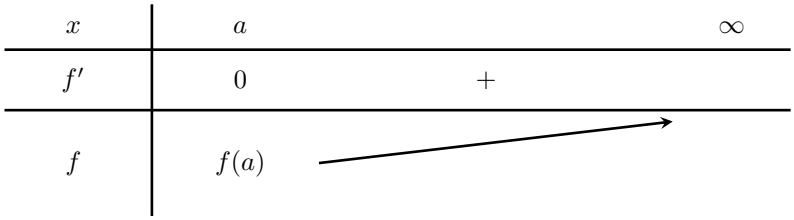


Figure 4.15 – Sign table for the function f on $[a, \infty)$.

Multiplying the numerator and denominator by the conjugate $(\sqrt{x} + \sqrt{a})$, we may also express the derivative as

$$f'(x) = \frac{x - a}{2\sqrt{x}(\sqrt{x} + \sqrt{a})}.$$

We analyze the sign of $f'(x)$ on $[a, \infty)$ in Figure 4.15.

For all $x > a$, we have $x - a > 0$, whence $f'(x) > 0$. Therefore, f is strictly increasing on $[a, \infty)$, implying that f attains its unique absolute minimum at $x = a$:

$$f(x) \geq f(a) = \frac{a + a}{2} - \sqrt{a^2} = a - a = 0 \quad \text{for all } x \geq a.$$

In particular, choosing $x = b \geq a$ yields

$$f(b) \geq 0 \iff \frac{a + b}{2} - \sqrt{ab} \geq 0 \iff \frac{a + b}{2} \geq \sqrt{ab}.$$

Equality holds if and only if $b = a$. □

Remark 4.7.3 (Algebraic Alternative). The two-variable case admits an immediate algebraic verification by expanding the non-negative square:

$$(\sqrt{a} - \sqrt{b})^2 \geq 0 \iff a - 2\sqrt{ab} + b \geq 0 \iff \frac{a + b}{2} \geq \sqrt{ab}.$$

The variational calculus perspective developed above, however, extends naturally to higher dimensions.

Exercise 4.7.4 (Three-Variable AM–GM Inequality)**.** Show that for all non-negative real numbers $a, b, c \geq 0$,

$$\frac{a + b + c}{3} \geq \sqrt[3]{abc}.$$

Hint. Fix $a, b \geq 0$ and define either the auxiliary function

$$f(x) \coloneqq \frac{a + b + x}{3} - \sqrt[3]{abx} \quad \text{for } x \geq 0,$$

or its cubic variant $g(x) \coloneqq \frac{1}{x}\left(\frac{a+b+x}{3}\right)^3 - ab$, and establish the minimum using calculus. □

Exercise 4.7.5 (General AM–GM Inequality)**.** Prove by mathematical induction, or via variational calculus, that for any finite collection of non-negative real numbers $a_1, a_2, \ldots, a_n \geq 0$,

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq (a_1 a_2 \cdots a_n)^{1/n},$$

with equality holding if and only if $a_1 = a_2 = \cdots = a_n$.

4.8 Differentials and Error Estimation

Intuition. Linearization and Differentials

Suppose a function f is differentiable at a point c. The Newton (or first-order Taylor) linear approximation asserts that for values of x close to c,

$$f(x) \approx f(c) + f'(c)(x - c).$$

Letting $\Delta x \coloneqq x - c$ denote the increment (or measurement error) in the independent variable and $\Delta y \coloneqq f(x) - f(c)$ denote the resulting change (or propagated error) in the dependent variable, we obtain the relation

$$\Delta y \approx f'(c)\,\Delta x.$$

Differentials provide an elegant operational framework for this first-order approximation, allowing us to estimate changes and propagate uncertainties using standard rules of differentiation.

Definition 4.8.1 (Differentials)**.** Let $y = f(x)$ be a differentiable function.

(i) The *differential of the independent variable* x, denoted by $\mathrm{d}x$, is an arbitrary non-zero real increment, defined to be identical to Δx:

$$\mathrm{d}x \coloneqq \Delta x.$$

(ii) The *differential of the dependent variable* y, denoted by $\mathrm{d}y$, is defined by

$$\mathrm{d}y \coloneqq f'(x)\,\mathrm{d}x.$$

Example 4.8.2 (Comparison of Δy and $\mathrm{d}y$)**.** Let $y = x^2$. Compute $\mathrm{d}y$ and the exact change Δy when $x = 1$ and $\mathrm{d}x = \Delta x = 0.01$.

Solution. Since $f(x) = x^2$, the derivative is $f'(x) = 2x$.

- *Differential estimation:* Evaluating the differential at $x = 1$ with $\mathrm{d}x = 0.01$:

$$\mathrm{d}y = f'(1)\,\mathrm{d}x = 2(1)(0.01) = 0.02.$$

- *Exact change:* On the other hand, the true change in y is

$$\begin{aligned}\Delta y &= f(x + \Delta x) - f(x)\\ &= f(1 + 0.01) - f(1)\\ &= (1.01)^2 - 1^2\\ &= 1.0201 - 1 = 0.0201.\end{aligned}$$

The approximation error $|\Delta y - \mathrm{d}y| = |0.0201 - 0.02| = 0.0001 = (\Delta x)^2$ is of second order in Δx, validating the precision of the differential approximation for small perturbations. □

Example 4.8.3 (Error Estimation in Geometric Measurements)**.** The radius of a spherical ball is measured to be $r = 0.7$ in., with an experimental measurement tolerance of ± 0.01 in.

(i) Estimate the propagated error in the calculated volume V of the ball.

(ii) Estimate the relative error and the percentage error in the volume.

Solution. Recall that the volume of a sphere of radius r is given by

$$V(r) = \frac{4}{3}\pi r^3.$$

Differentiating with respect to r yields

$$V'(r) = \frac{4}{3}\pi(3r^2) = 4\pi r^2.$$

We are given the nominal radius $r = 0.7$ in. and the measurement uncertainty $\Delta r = \mathrm{d}r = \pm 0.01$ in.

(i) *Propagated Error:*
Using the differential approximation, the propagated error ΔV in the volume satisfies

$$\begin{aligned}\Delta V \approx \mathrm{d}V &= V'(r)\,\mathrm{d}r \\ &= 4\pi r^2\,\mathrm{d}r \\ &= 4\pi(0.7)^2(\pm 0.01) \\ &= 4\pi(0.49)(\pm 0.01) \\ &= \pm 0.0196\pi \\ &\approx \pm 0.06158 \text{ in.}^3.\end{aligned}$$

(ii) *Relative and Percentage Error:*
The relative error is the ratio of the propagated error to the total volume:

$$\text{Relative error} = \frac{\Delta V}{V} \approx \frac{\mathrm{d}V}{V} = \frac{4\pi r^2\,\mathrm{d}r}{\frac{4}{3}\pi r^3} = \frac{3\,\mathrm{d}r}{r}.$$

Substituting the numerical values $r = 0.7$ and $\mathrm{d}r = \pm 0.01$:

$$\frac{\mathrm{d}V}{V} = \frac{3(\pm 0.01)}{0.7} = \pm\frac{0.03}{0.7} = \pm\frac{3}{70} \approx \pm 0.042857.$$

Consequently, the relative error is approximately ± 0.0429, corresponding to a percentage error of approximately $\pm 4.29\%$ (or roughly $\approx \pm 5\%$).

□

Example 4.8.4 (Differential Product Rule)**.** Let u and v be differentiable functions of an independent variable x. Derive the Leibniz differential identity for the product uv.

Derivation. By definition of differentials, we have

$$\mathrm{d}u = u'(x)\,\mathrm{d}x \quad \text{and} \quad \mathrm{d}v = v'(x)\,\mathrm{d}x.$$

Applying the differential operator to the product uv and employing the standard product rule for derivatives:

$$\begin{aligned} \mathrm{d}(uv) &= (uv)' \,\mathrm{d}x \\ &= \big(u\,v' + v\,u'\big)\,\mathrm{d}x \\ &= u\,v'\,\mathrm{d}x + v\,u'\,\mathrm{d}x \\ &= u\big(v'\,\mathrm{d}x\big) + v\big(u'\,\mathrm{d}x\big) \\ &= u\,\mathrm{d}v + v\,\mathrm{d}u. \end{aligned}$$

This confirms that differentials satisfy the familiar identity $\mathrm{d}(uv) = u\,\mathrm{d}v + v\,\mathrm{d}u$. □

4.9 The Root-Finding Problem and Newton Method

Intuition. From Geometric Zeros to Fixed-Point Dynamics

When attempting to sketch the graph of an ostensibly simple transcendental function such as

$$f(x) = x + 2\cos x,$$

one is immediately confronted with locating its x-intercepts:

$$x + 2\cos x = 0 \iff x = -2\cos x.$$

In general, finding the exact analytical roots or zeros of an arbitrary nonlinear equation $f(x) = 0$ is profoundly difficult. By the Abel–Ruffini theorem and Galois theory, even general polynomial equations of degree five or higher admit no closed-form solutions in radicals. Consequently, our strategy shifts from exact symbolic algebra to systematic qualitative existence criteria and high-order iterative numerical dynamics.

Problem (The Root-Finding Problem)**.** Given a continuous function $f\colon [a,b] \to \mathbb{R}$, determine a scalar $\xi \in [a,b]$ such that

$$f(\xi) = 0. \tag{4.1}$$

4.9.1 Existence via Topological Theorems

Before searching for a root, one must verify that a solution exists in the chosen interval. We recall a property of continuous functions:

Theorem 4.9.1 (Intermediate Value Theorem/Bolzano's Theorem)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function. If $f(a)$ and $f(b)$ possess opposite signs, that is,*

$$f(a)f(b) \le 0,$$

then there exists at least one root $\xi \in [a,b]$ such that $f(\xi) = 0$.

Although the Intermediate Value Theorem ensures the existence of a root, it is purely non-constructive. To develop an algorithm for locating ξ, we recast the equation $f(x) = 0$ into an equivalent *fixed-point problem*:

$$g(x) = x.$$

The simplest algebraic transformation consists of adding x to both sides, setting $g(x) \coloneqq x + f(x)$, though more sophisticated transformations can be designed as we will see below.

Theorem 4.9.2 (Brouwer's Fixed-Point Theorem in One Dimension)**.** *Let $g\colon [a,b] \to [a,b]$ be a continuous function mapping a non-empty, closed, and bounded interval into itself. Then g admits at least one fixed point $\xi \in [a,b]$; that is,*

$$g(\xi) = \xi.$$

Proof. Define the auxiliary displacement function $\varphi\colon [a,b] \to \mathbb{R}$ by

$$\varphi(x) \coloneqq x - g(x).$$

Since g is continuous and the identity map $h(x) = x$ is continuous, φ is continuous on $[a,b]$. Because the image of g lies entirely within $[a,b]$, we have $a \le g(a)$ and $g(b) \le b$. Evaluating φ at the endpoints:

$$\varphi(a) = a - g(a) \le 0 \quad \text{and} \quad \varphi(b) = b - g(b) \ge 0.$$

Therefore, $\varphi(a)\varphi(b) \le 0$. By the Intermediate Value Theorem, there exists $\xi \in [a,b]$ such that $\varphi(\xi) = 0$, which immediately implies $g(\xi) = \xi$. □

Story. Topological Consequences: Stirring Coffee and Crumpling Maps

Brouwer's Fixed-Point Theorem generalizes to any so-called *compact convex* set $K \subset \mathbb{R}^n$: every continuous map $g\colon K \to K$ has a fixed point. This yields remarkable physical interpretations:

- *Stirring a Cup of Coffee:* If you continuously stir a cup of coffee without splashing, once the liquid comes to rest, at least one particle of fluid occupies the exact three-dimensional coordinate it held before stirring began.
- *The Crumpled Map:* Take an ordinary two-dimensional map of your city lying flat on a table. If you crumple the map into a ball and drop it onto the floor within the region depicted, there is guaranteed to be at least one point on the crumpled map lying precisely directly above the actual physical location it represents.

4.9.2 Discrete Dynamics and Rates of Convergence

To numerically locate a fixed point $\xi = g(\xi)$, modern computers iterate the function g:

Definition 4.9.3 (Fixed-Point Iteration)**.** Let $g\colon [a,b] \to [a,b]$ be continuous. Given an initial guess $x_0 \in [a,b]$, the recursive sequence defined by

$$x_{k+1} = g(x_k), \qquad k = 0, 1, 2, \dots \tag{4.2}$$

is called a *iteration simple* (or *iteration fixed-point*), and the terms $\{x_k\}_{k=0}^{\infty}$ are called the *iterates*.

If the sequence $\{x_k\}$ converges to a limit ξ, the continuity of g allows us to pass to the limit on both sides:

$$\xi = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} g(x_k) = g\Big(\lim_{k\to\infty} x_k\Big) = g(\xi).$$

Thus, the limit ξ is necessarily a fixed point of g.

Definition 4.9.4 (Rate of Convergence)**.** Suppose that $\lim_{k\to\infty} x_k = \xi$. If there exist constants $\mu > 0$ and $q \geq 1$ such that

$$\lim_{k\to\infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^q} = \mu, \tag{4.3}$$

then $\{x_k\}$ is said to *with order q converge* to ξ, and q is called the *of convergence order* (or rate).

- If $q = 1$, the sequence converges *linearly linearly*, and μ must satisfy $\mu \in (0, 1)$ for convergence.
- If $q = 2$, the sequence converges *quadratically quadratically*.

If g is a contraction mapping satisfying the Lipschitz condition $|g(x) - g(y)| \leq L\,|x - y|$ with Lipschitz constant $L \in (0, 1)$, the successive error bounds satisfy:

$$|x_{k+1} - \xi| = |g(x_k) - g(\xi)| \leq L\,|x_k - \xi|\,.$$

Taking limits as $k \to \infty$ gives $0 < \mu \leq L < 1$.

Definition 4.9.5 (Asymptotic Rate of Convergence)**.** For linear convergence ($q = 1$), we define the *rate of convergence asymptotic* ρ by

$$\rho \coloneqq -\log_{10} \mu.$$

The quantity ρ quantifies the number of correct decimal digits gained per iteration.

Example 4.9.6 (Comparison of Linear and Higher-Order Convergence)**.** Let $x_k \coloneqq 1 + 10^{-k}$, converging to $\xi = 1$. Then

$$\lim_{k\to\infty} \frac{|x_{k+1} - 1|}{|x_k - 1|} = \lim_{k\to\infty} \frac{10^{-(k+1)}}{10^{-k}} = \frac{1}{10}.$$

Here $q = 1$ and $\mu = \frac{1}{10}$, so $\rho = -\log_{10}(10^{-1}) = 1$. Each successive iteration produces exactly one additional correct decimal place (e.g., $x_2 = 1.01$, $x_3 = 1.001$).

Conversely, when $q = 2$, taking logarithms base 10 of (4.3) yields

$$\log_{10} |x_{k+1} - \xi| \approx \log_{10} \mu + 2 \log_{10} |x_k - \xi| \approx 2 \log_{10} |x_k - \xi|\,.$$

Consequently, for quadratic convergence, the number of correct decimal digits approximately *doubles* with every single iteration.

4.9.3 Construction of Root-Finding Algorithms

Fixed-Point Method

Suppose g is continuously differentiable on $[a, b]$ with $\xi = g(\xi)$. By the Mean Value Theorem, there exists a point η_k between x_k and ξ such that

$$x_{k+1} - \xi = g(x_k) - g(\xi) = g'(\eta_k)(x_k - \xi).$$

Taking absolute values, dividing by $|x_k - \xi|$, and passing to the limit as $k \to \infty$ reveals

$$\lim_{k\to\infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|} = |g'(\xi)|\,. \tag{4.4}$$

Thus, $\mu = |g'(\xi)|$ and $\rho = -\log_{10} |g'(\xi)|$. The flatter the map at the fixed point (i.e., the closer $|g'(\xi)|$ is to 0), the faster the iteration converges.

Relaxation and Newton Method

A major challenge with generic fixed-point iteration is that transforming an arbitrary root-finding problem $f(x) = 0$ into a contracting map $x = g(x)$ requires careful tuning.

Definition 4.9.7 (Relaxation Method)**.** Let $f\colon [a, b] \to \mathbb{R}$ be continuous. For a fixed non-zero relaxation parameter $\lambda \neq 0$, the *method relaxation* generates iterates via

$$x_{k+1} = x_k - \lambda f(x_k), \qquad k = 0, 1, 2, \dots \tag{4.5}$$

Theorem 4.9.8. *Let $f\colon [a, b] \to \mathbb{R}$ with $f(\xi) = 0$ and $f'(\xi) \neq 0$. Suppose f is continuously differentiable in a small neighborhood $[\xi - \delta, \xi + \delta] \subset [a, b]$. Then there exist a parameter $\lambda \in \mathbb{R}$ and a neighborhood radius $h \in (0, \delta)$ such that the relaxation iterates (4.5) converge to ξ for every initial guess $x_0 \in [\xi - h, \xi + h]$.*

Proof Sketch and Derivation of Newton Method. Notice that (4.5) is a fixed-point iteration for the auxiliary map

$$g(x) \coloneqq x - \lambda f(x).$$

Differentiating g yields

$$g'(x) = 1 - \lambda f'(x).$$

To maximize the asymptotic rate of convergence, we seek to make the asymptotic ratio $\mu = |g'(\xi)|$ vanish entirely:

$$g'(\xi) = 0 \iff 1 - \lambda f'(\xi) = 0 \iff \lambda = \frac{1}{f'(\xi)}.$$

However, the true root ξ is unknown in advance. The decisive insight of Isaac Newton and Joseph Raphson was to approximate the optimal parameter dynamically at each step using the current state:

$$\lambda(x_k) \coloneqq \frac{1}{f'(x_k)}.$$

Substituting this variable relaxation parameter into the recursion yields Newton celebrated method. □

Definition 4.9.9 (Newton Method)**.** Let $f\colon [a, b] \to \mathbb{R}$ be differentiable. Given an initial guess x_0, *Newton method* (or the *Newton–Raphson method*) produces the sequence of iterates

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \qquad k = 0, 1, 2, \dots \tag{4.6}$$

under the assumption that $f'(x_k) \neq 0$ for all $k \geq 0$.

Example 4.9.10 (Computing $\sqrt{2}$)**.** Approximate the irrational number $\sqrt{2}$ using Newton method.

Solution. We translate the problem of evaluating $x = \sqrt{2}$ into the polynomial root-finding problem:

$$x^2 = 2 \iff f(x) \coloneqq x^2 - 2 = 0.$$

1. *Existence:* Evaluating f on the interval $[1, 3]$ gives

$$f(1) = 1^2 - 2 = -1, \qquad f(3) = 3^2 - 2 = 7.$$

Since $f(1)f(3) = (-1)(7) = -7 < 0$, the Intermediate Value Theorem guarantees a root $\xi \in (1, 3)$.

2. *Iteration Formula:* Differentiating yields $f'(x) = 2x$. The Newton iteration is:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} = x_k - \frac{x_k^2 - 2}{2x_k} = x_k - \frac{x_k}{2} + \frac{1}{x_k} = \frac{x_k}{2} + \frac{1}{x_k}.$$

3. *Iterative Calculations:* Starting with the initial guess $x_0 = 2 \in (1, 3)$:

$$\begin{aligned}
x_1 &= \frac{x_0}{2} + \frac{1}{x_0} = \frac{2}{2} + \frac{1}{2} = 1.5, \\
x_2 &= \frac{x_1}{2} + \frac{1}{x_1} = \frac{1.5}{2} + \frac{1}{1.5} = \frac{3}{4} + \frac{2}{3} = \frac{17}{12} \approx 1.416666666667, \\
x_3 &= \frac{x_2}{2} + \frac{1}{x_2} = \frac{17}{24} + \frac{12}{17} = \frac{577}{408} \approx 1.414215686275, \\
x_4 &= \frac{x_3}{2} + \frac{1}{x_3} \approx 1.414213562375, \\
x_5 &= \frac{x_4}{2} + \frac{1}{x_4} \approx 1.414213562373.
\end{aligned}$$

□

Notice that x_4 and x_5 agree to eleven decimal places with the true value $\sqrt{2} \approx 1.414213562373$, so although we only derived Newton method with linear convergence in mind, we actually get quadratic convergence. Is there an explanation for this important observation? We first need to step back from the numerical algorithm to expand our theoretical toolbox.

4.9.4 Higher-Order Approximations

Intuition. From Tangent Linearization to Curvature Corrections

Recall that if a function f is differentiable at a, the first-order Taylor approximation provides the tangent line linearization:

$$f(x) \approx f(a) + f'(a)(x - a).$$

While this local approximation underlies the definition of the derivative and simple relaxation schemes, it omits *curvature* (concavity) effects. To explain why Newton method converges at a quadratic rate rather than a linear one, we must sharpen the tangent approximation by incorporating a second-order correction term that accounts for $f''(x)$.

Theorem 4.9.11 (Second-Order Taylor Formula with Lagrange Remainder)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be twice continuously differentiable (that is, f'' exists and is continuous on $[a,b]$). Then there exists an intermediate point $c \in [a,b]$ such that*

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(c)}{2}(b-a)^2. \tag{4.7}$$

Proof. Since f'' is continuous on the compact interval $[a,b]$, the Extreme Value Theorem guarantees that f'' attains its absolute minimum m and absolute maximum M on $[a,b]$:

$$m \le f''(x) \le M \quad \text{for all } x \in [a,b].$$

We construct bounding functions based on these extrema:

1. *Upper Bound via Monotonicity.*
Consider the inequality $f''(x) - M \le 0$. Observe that $f''(x) - M$ is the derivative of the auxiliary function $\psi_1(x) \coloneqq f'(x) - Mx$. Because $\psi_1'(x) \le 0$, the function ψ_1 is non-increasing on $[a,b]$. Therefore, for every $x \ge a$,

$$\begin{aligned} f'(x) - Mx &\le f'(a) - Ma \\ f'(x) - f'(a) &\le M(x-a) \\ f'(x) - f'(a) - M(x-a) &\le 0. \end{aligned}$$

Next, observe that $f'(x) - f'(a) - M(x-a)$ is the derivative of

$$\phi_1(x) \coloneqq f(x) - f'(a)x - \frac{M}{2}(x-a)^2.$$

Because $\phi_1'(x) \le 0$, ϕ_1 is non-increasing on $[a,b]$. Since $x \ge a$, evaluating at x and a yields

$$\begin{aligned} f(x) - f'(a)x - \frac{M}{2}(x-a)^2 &\le f(a) - f'(a)a - \frac{M}{2}(a-a)^2 \\ f(x) - f(a) - f'(a)(x-a) - \frac{M}{2}(x-a)^2 &\le 0 \\ f(x) &\le f(a) + f'(a)(x-a) + \frac{M}{2}(x-a)^2. \end{aligned}$$

2. *Lower Bound via Monotonicity.*
Applying the identical reasoning to $m \le f''(x)$ gives the reverse inequality:

$$f(a) + f'(a)(x-a) + \frac{m}{2}(x-a)^2 \le f(x).$$

3. *Application of the Intermediate Value Theorem.*
Combining both bounds, for every $x \in [a,b]$ we have

$$f(a) + f'(a)(x-a) + \frac{m}{2}(x-a)^2 \le f(x) \le f(a) + f'(a)(x-a) + \frac{M}{2}(x-a)^2.$$

Rearranging terms, we conclude that the quantity

$$H \coloneqq \frac{2\big(f(x) - f(a) - f'(a)(x-a)\big)}{(x-a)^2}$$

lies in the range $[m, M]$. Because f'' is continuous on $[a, b]$, the Intermediate Value Theorem guarantees the existence of a point $c \in [a, b]$ such that $f''(c) = H$. Thus,

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

Choosing $x = b$, we obtain (4.7).

□

Remark 4.9.12 (Higher-Order Generalization). If f is three times continuously differentiable, an identical monotonic bounding procedure as above yields

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2 + \frac{f'''(c)}{6}(x - a)^3.$$

More generally, if f is smooth (infinite differentiable), iterating this expansion ad infinitum produces the formal *Taylor series*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x - a)^n.$$

4.9.5 Asymptotic Error Analysis of Newton Method

We now deploy Theorem 4.9.11 to establish why Newton method exhibits quadratic convergence.

Theorem 4.9.13 (Quadratic Error Equation for Newton Method)**.** *Let $f\colon [a, b] \to \mathbb{R}$ be twice continuously differentiable and let ξ be a simple root, so that $f(\xi) = 0$ with $f'(\xi) \neq 0$. If $\{x_k\}_{k=0}^{\infty}$ denotes the sequence generated by Newton method:*

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)},$$

then there exists an intermediate point c_k between x_k and ξ such that

$$x_{k+1} - \xi = \frac{1}{2}\frac{f''(c_k)}{f'(x_k)}(x_k - \xi)^2. \tag{4.8}$$

Proof. Expanding f around the iterate x_k using Theorem 4.9.11 with $a = x_k$ and $x = \xi$:

$$f(\xi) = f(x_k) + f'(x_k)(\xi - x_k) + \frac{1}{2}f''(c_k)(\xi - x_k)^2.$$

Since ξ is a root, $f(\xi) = 0$. Dividing the entire equation by $f'(x_k) \neq 0$:

$$0 = \frac{f(x_k)}{f'(x_k)} + (\xi - x_k) + \frac{1}{2}\frac{f''(c_k)}{f'(x_k)}(\xi - x_k)^2.$$

Recalling the definition of the subsequent iterate, $\frac{f(x_k)}{f'(x_k)} = x_k - x_{k+1}$. Substituting this relation into the equation gives

$$0 = (x_k - x_{k+1}) + (\xi - x_k) + \frac{1}{2}\frac{f''(c_k)}{f'(x_k)}(\xi - x_k)^2.$$

Simplifying $(x_k - x_{k+1}) + (\xi - x_k) = \xi - x_{k+1} = -(x_{k+1} - \xi)$ and noting that $(\xi - x_k)^2 = (x_k - \xi)^2$, we isolate the forward error:

$$x_{k+1} - \xi = \frac{1}{2}\frac{f''(c_k)}{f'(x_k)}(x_k - \xi)^2.$$

Taking absolute values yields

$$\frac{|x_{k+1} - \xi|}{|x_k - \xi|^2} = \frac{1}{2}\left|\frac{f''(c_k)}{f'(x_k)}\right|.$$

Assuming $f'(\xi) \neq 0$ and passing to the limit as $x_k \to \xi$ (which also forces $c_k \to \xi$), we obtain

$$\lim_{k\to\infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^2} = \frac{1}{2}\left|\frac{f''(\xi)}{f'(\xi)}\right| = \mu.$$

This confirms that the asymptotic order of convergence is $q = 2$ (quadratic convergence). □

⚠ *Caveat. Degeneracy and Pathologies of Newton Method*

The quadratic convergence identity (4.8) hinges fundamentally on 3 assumptions:

(1) *Bounded Denominator:* The derivative $|f'(x)|$ must remain bounded strictly away from zero in a neighborhood of ξ.

(2) *Simple Roots ($f'(\xi) \neq 0$):* If ξ is a multiple root ($f(\xi) = 0$ and $f'(\xi) = 0$), the quadratic acceleration deteriorates into linear convergence, see Exercise 4.9.14.

(3) *Non-Local Instabilities and Limit Cycles:* If the initial point x_0 is selected outside the basin of attraction, the tangent projections can oscillate indefinitely, alternating periodically between distinct values $\xi_1 \neq \xi_2$ rather than converging, see Example 4.9.16.

Exercise 4.9.14 (Loss of Quadratic Convergence for Multiple Roots)**.** Let $f(x) = (x-1)^2$.

(a) Show that Newton method applied to $f(x)$ generates the iterative formula

$$x_{k+1} = \frac{x_k + 1}{2}.$$

(b) Prove that the error satisfies

$$x_{k+1} - 1 = \left(\frac{1}{2}\right)^{k+1}(x_0 - 1).$$

(c) Conclude that the convergence to the multiple root $\xi = 1$ is strictly linear with asymptotic rate $\mu = \frac{1}{2}$.

4.9.6 Geometric Meaning: Tangent Line Projection

Intuition. Linear Approximation and Tangent Intercepts

Newton method replaces the potentially intractable problem of solving a non-linear equation $f(x) = 0$ with an iteratively solvable linear problem: finding the zero of the *tangent line* to the graph of f at the current state.

Proposition 4.9.15 (Derivation of Newton Iteration via Tangent Intercepts)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be differentiable, and let $x_k \in [a,b]$ be an estimate of a root ξ. Assume the derivative does not vanish at this point: $f'(x_k) \neq 0$. Then the subsequent Newton iterate*

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

is precisely the unique horizontal intercept (where $y = 0$) of the tangent line to the curve $y = f(x)$ at the point $\big(x_k, f(x_k)\big)$. See Figure 4.16.

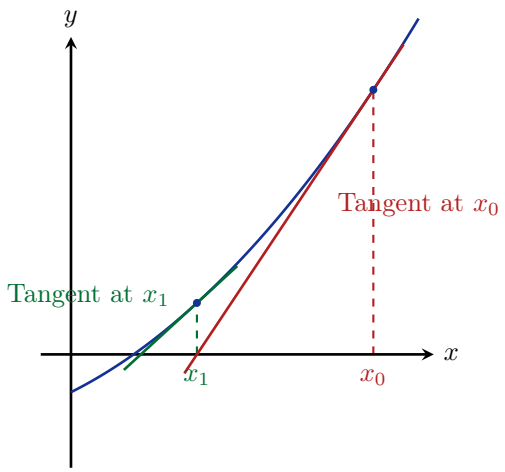


Figure 4.16 – Geometric derivation of Newton iteration.

Proof. Recall that the equation of the tangent line $L(x)$ to the curve $y = f(x)$ passing through the contact point $\big(x_k, f(x_k)\big)$ with slope $m = f'(x_k)$ is given by the point-slope formula:

$$y - f(x_k) = f'(x_k)(x - x_k). \tag{4.9}$$

Equivalently, expressing the line as a linear function of x:

$$L(x) = f(x_k) + f'(x_k)(x - x_k).$$

To approximate the root where the non-linear function vanishes ($f(x) = 0$), we instead find where its local linear approximation vanishes:

$$L(x) = 0.$$

Thus we obtain:

$$0 = f(x_k) + f'(x_k)(x - x_k).$$

Since $f'(x_k) \neq 0$, dividing both sides by the slope $f'(x_k)$ yields

$$x - x_k = -\frac{f(x_k)}{f'(x_k)},$$

so

$$x = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Defining this intercept to be the next approximation x_{k+1} reproduces the formula for Newton iteration:

$$x_{k+1} \coloneqq x_k - \frac{f(x_k)}{f'(x_k)}.$$

□

With this intuition in mind, we construct an example where Newton method fails:

Example 4.9.16 (Pathology of Newton Method: Periodic 2-Cycles)**.** Consider the cubic polynomial

$$f(x) = x^3 - 5x.$$

Determine the behavior of Newton method initiated from $x_0 = 1$.

Solution. The real roots of $f(x) = x(x^2 - 5) = 0$ are $\xi \in \{-\sqrt{5}, 0, \sqrt{5}\}$. The derivative is

$$f'(x) = 3x^2 - 5.$$

Thus, the general Newton iteration map is given by

$$x_{k+1} = g(x_k) = x_k - \frac{f(x_k)}{f'(x_k)} = x_k - \frac{x_k^3 - 5x_k}{3x_k^2 - 5} = \frac{2x_k^3}{3x_k^2 - 5}.$$

1. *First Iteration from* $x_0 = 1$. Evaluating f and f' at $x_0 = 1$:

$$f(1) = 1^3 - 5(1) = -4, \qquad f'(1) = 3(1)^2 - 5 = -2.$$

The tangent line to the curve at the point $(1, -4)$ has the equation:

$$y - (-4) = -2(x - 1) \implies y = -2x - 2.$$

The x-intercept of this tangent line yields the next iterate:

$$x_1 = 1 - \frac{-4}{-2} = 1 - 2 = -1.$$

2. *Second Iteration from* $x_1 = -1$. Evaluating f and f' at $x_1 = -1$:

$$f(-1) = (-1)^3 - 5(-1) = 4, \qquad f'(-1) = 3(-1)^2 - 5 = -2.$$

The tangent line to the curve at $(-1, 4)$ has the equation:

$$y - 4 = -2\big(x - (-1)\big) \implies y = -2x + 2.$$

The x-intercept of this tangent line yields:

$$x_2 = -1 - \frac{4}{-2} = -1 - (-2) = 1.$$

Consequently, $x_2 = x_0 = 1$. Rather than converging to any of the three roots $\{-\sqrt{5}, 0, \sqrt{5}\}$, the sequence of iterates enters an infinite periodic 2-cycle:

$$1 \longrightarrow -1 \longrightarrow 1 \longrightarrow -1 \longrightarrow \cdots$$

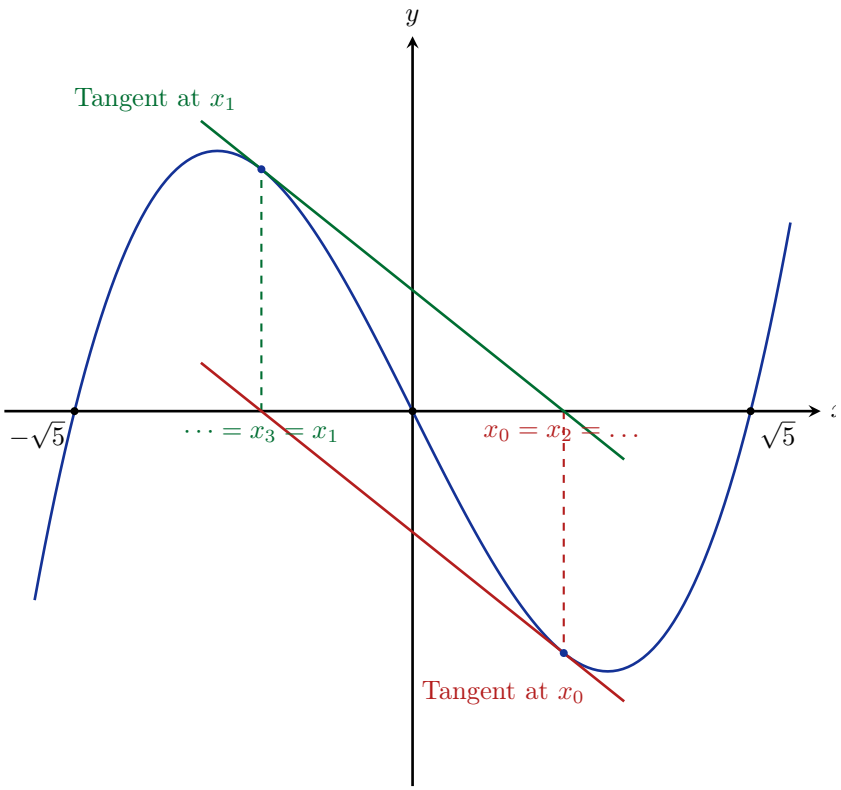


Figure 4.17 – Failure of Newton iteration via an infinite 2-cycle loop for $f(x) = x^3 - 5x$ starting at $x_0 = 1$.

□

As illustrated in Figure 4.17, the tangent line at $x_0 = 1$ intersects the x-axis at $x_1 = -1$. Because the function is odd ($f(-x) = -f(x)$) and the derivative is even ($f'(-x) = f'(x)$), the tangent slope at $x_1 = -1$ is identical to that at $x_0 = 1$ (namely $m = -2$), casting the subsequent projection straight back onto $x_2 = 1$. The algorithm is trapped in a closed loop.

Exercise 4.9.17 (Tangent 3-Cycles)**.** Construct an example of a differentiable function f and an initial guess x_0 such that the sequence of Newton iterates $\{x_k\}$ does not converge to any root, but instead enters a non-trivial 3-cycle $\{\xi_1, \xi_2, \xi_3\}$ satisfying

$$x_0 = \xi_1, \qquad x_1 = \xi_2, \qquad x_2 = \xi_3.$$

Illustrate this failure mode geometrically by sketching the intersecting tangent projections.

Chapter 5

Integration

> The whole of science is nothing more than a refinement of everyday thinking.
>
> Albert Einstein

5.1 The Big Picture: Motivation of Integration

Story. The Broken Odometer and the Inverse Problem of Motion

After three-quarters of a semester diligently carrying an eager freshman twice daily between home and campus for her calculus lectures, the odometer of her car decides to retire at the onset of a brutal Midwestern winter. Having mastered the kinematic principle that speed is the instantaneous time rate of change of distance,

$$\text{Speed} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = \frac{\mathrm{d}s}{\mathrm{d}t},$$

the student poses the inverse question: *If the odometer is broken, but the speedometer continues to report the speed $v(t) = f(t)$ at every instant of time, can one reconstruct the total distance traversed?*

When an object moves at a constant speed v over a time interval $\Delta t = t_1 - t_0$, the net distance traveled is given by elementary arithmetic:

$$\text{Distance} = \text{Speed} \times \text{Time} = v \cdot (t_1 - t_0).$$

Geometrically, this corresponds to the area of a rectangle of height v and width Δt, see Figure 5.1.

When speed varies continuously as a function of time, $v = f(t)$, we partition the total duration $[t_0, t_1]$ into n brief subintervals of duration $\Delta t = \frac{t_1 - t_0}{n}$. Across each minute duration $[t_{i-1}, t_i]$, the speed is approximately constant and equal to $f(t_i^*)$. The total distance is then

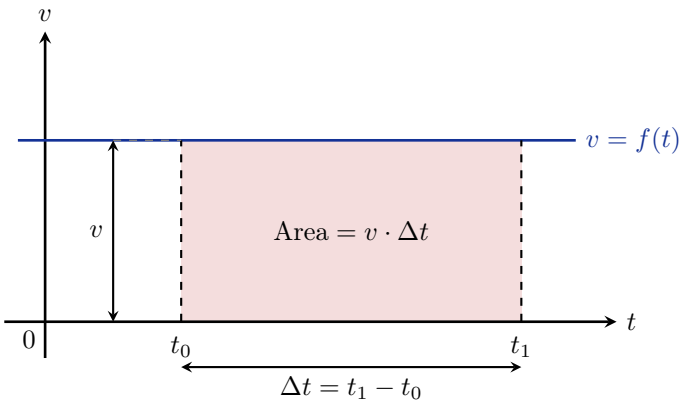


Figure 5.1 – Distance represented geometrically as the area of a single rectangle under a constant speed line.

approximated by the so-called *Riemann sum* of rectangular areas:

$$\text{Distance} \approx \sum_{i=1}^{n} f(t_i^*)\,\Delta t,$$

see Figure 5.2.

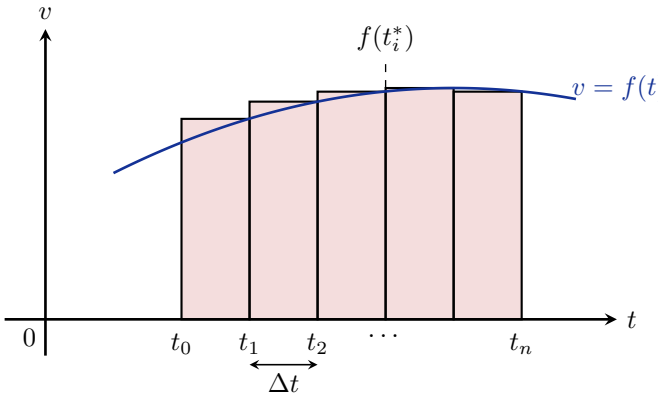


Figure 5.2 – Varying speed approximated by a Riemann sum of rectangular areas $\sum_{i=1}^{n} f(t_i^*)\,\Delta t$.

Taking the continuous limit as $\Delta t \to 0$ yields the *definite integral*:

$$\text{Distance} = \lim_{\Delta t \to 0} \sum_{i=1}^{n} f(t_i^*)\,\Delta t =: \int_{t_0}^{t_1} f(t)\,\mathrm{d}t. \tag{5.1}$$

Evaluating such limits of Riemann sums directly is typically a fool's errand. The Fundamental Theorem of Calculus (that we will study later) resolves this computational

impasse by establishing that integration is the inverse operation of differentiation:

$$\int_{t_0}^{t_1} f(t)\,\mathrm{d}t = F(t_1) - F(t_0), \tag{5.2}$$

where F is a differentiable function satisfying $F'(t) = f(t)$. So the task becomes finding such function F.[1]

5.2 Signed Quantities and Net Change

Story. Signed Speed, Direction, and the Journey to Campus

In our initial exploration of integration, we treated speed as an inherently non-negative scalar quantity. In classical mechanics, however, motion is directed: we must distinguish between advancing forward and retreating backward. This gives rise to *velocity*, which represents *signed speed*.

Let us modify the story at the beginning of the previous section to illustrate signed speed. Picture a straight road as a number line, with the positive direction pointing to the right. You leave *Home*, at position x_0 at time t_0, and drive in the positive direction past *Campus*, at position x_{Campus}, until you reach a *Friend*'s place further along at time t_{Friend}. Then you remember you have a lecture, so you turn around and drive back in the negative direction to Campus and arrive at time t_{Campus}. The questions are simple: how many miles *from home*, and how many miles *traversed*?

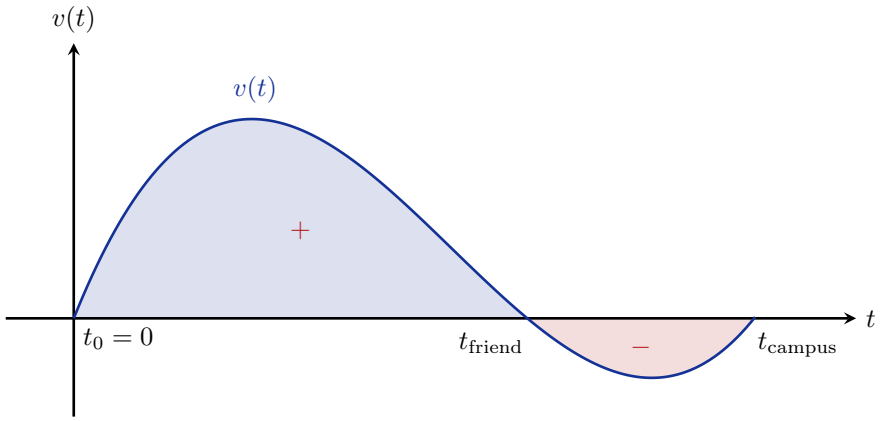


Figure 5.3 – Net change interpretation: the displacement $\int_0^{t_{\text{campus}}} v(t)\,\mathrm{d}t$ equals the area above the t-axis (positive) minus the area below it (negative).

Definition 5.2.1 (Velocity)**.** The *velocity* of an object moving along a line is its *signed speed*, which we denote by $v(t)$. The sign records the direction of motion:

- if the object moves in the positive direction at time t, then $v(t) > 0$;

[1] In fact, from now on and half of Calculus II, you will learn how to find such functions for a given f.

- if the object moves in the negative direction at time t, then $v(t) < 0$.

So speed tells you how fast you are going, and velocity also tells you *which way*. Let $s(t)$ denote the position of the object at time t.

Definition 5.2.2 (Net displacement)**.** The *net displacement* (also called the *signed distance*) travelled between times t_0 and t is

$$s(t) - s(t_0) =: \int_{t_0}^{t} v(\tau)\,\mathrm{d}\tau.$$

⚠ *Caveat. Dummy variables in integrals*

Inside the integral we have renamed the integration variable to τ. This is because the letter t is already being used for the upper limit. Writing $\int_{t_0}^{t} v(t)\,\mathrm{d}t$ is a very common mistake by inexperienced freshman: using one symbol for two different jobs. The integration variable is a *dummy variable*: it can be given any name we like. We will come back to this idea when we meet sigma notation in Section 5.3.

Now go back to our trip. Start the clock when you leave home, so $t_0 = 0$ and $x_0 = 0$. From t_0 until t_{Friend} you drive forward, so $v(t) > 0$. From t_{Friend} until t_{Campus} you drive backward, so $v(t) < 0$ (see Figure 5.3). This gives the two answers to our question.

- The *net displacement* counts motion in the negative direction *against* motion in the positive direction:

$$\begin{aligned}\text{Net Displacement} &= \int_{t_0}^{t_{\text{Campus}}} v(t)\,\mathrm{d}t \\ &= \big(\text{Area above the } t\text{-axis}\big) - \big(\text{Area below the } t\text{-axis}\big).\end{aligned}$$

- The *total distance travelled* counts every bit of motion as positive, whichever way you were going:

$$\begin{aligned}\text{Total Distance Travelled} &= \int_{t_0}^{t_{\text{Campus}}} |v(t)|\,\mathrm{d}t \\ &= \big(\text{Area above the } t\text{-axis}\big) + \big(\text{Area below the } t\text{-axis}\big).\end{aligned}$$

Intuition. Odometer versus GPS

Your car's odometer measures *total distance travelled*. It keeps going up no matter which way you drive, because it only sees the speed $|v(t)|$. A GPS showing how far you are from home measures *net displacement*. When you turn around, that number starts going down again, because driving backward undoes driving forward. At the end of the trip above, the odometer shows the full Home–Friend–Campus distance. The GPS shows only $x_{\text{Campus}} - x_0$. In summary:

(1) *Net Displacement* ($\int_{t_0}^{t_1} v(t)\,\mathrm{d}t$)*:* The algebraic sum of forward and backward motions. Periods of negative velocity cancel out portions of forward progress.

(2) *Total Distance Traveled ($\int_{t_0}^{t_1} |v(t)|\,\mathrm{d}t$):* The absolute geometric mileage recorded by a working odometer, where every direction is reckoned positively.

Moral.

An integral of a signed quantity measures *net change*: area above the axis counts as positive and area below the axis counts as negative. To measure the *total amount*, with no cancellation, integrate the absolute value.

Both quantities above come down to the same geometric problem. We need to compute the area of the region bounded by

$$\text{the graph of } f(t), \qquad t = a, \qquad t = b, \qquad \text{and the horizontal axis.}$$

For most functions we cannot find this area exactly with elementary geometry. Instead, we will approximate the region by many thin rectangles and add up their areas. To do that efficiently we need a compact way to write long sums.

5.3 Sigma Notation

Notation 5.3.1 (Sigma notation)**.** Given an ordered sequence of n real numbers $a_1, a_2, \ldots, a_n$, we write

$$\sum_{i=1}^{n} a_i \coloneqq a_1 + a_2 + \cdots + a_n.$$

Here i is the *index of summation* (also called the *dummy variable*), and a_i is the *i-th term* of the sum. The numbers 1 and n are the *lower* and *upper limits* of summation.

Story. Euler's Σ

The capital Greek letter Σ ("S", for *summa*) was introduced for sums by Leonhard Euler in his 1755 textbook *Institutiones calculi differentialis*. Euler invented or popularized much of the notation we now take for granted, including $f(x)$, e, π, and i. Notation like this lets us write down an idea in a line that would otherwise take a paragraph.

Example 5.3.2 (The name of the index does not matter)**.**

$$\sum_{k=1}^{99} k = 1 + 2 + \cdots + 99 = \sum_{i=1}^{99} i.$$

The letters i and k are dummy variables. Which letter we use does not matter, because it does not change the sum. This is the same reason why $\int_{t_0}^{t} v(\tau)\,\mathrm{d}\tau$ and $\int_{t_0}^{t} v(s)\,\mathrm{d}s$ mean the same thing.

Example 5.3.3 (Index shift)**.** The same sum can be written with different limits if we adjust the terms to match:

$$\sum_{i=1}^{6} i = 1+2+3+4+5+6 = 21,$$

$$\sum_{k=0}^{5}(k+1) = (0+1)+(1+1)+(2+1)+(3+1)+(4+1)+(5+1)$$
$$= 1+2+3+4+5+6 = 21.$$

Here are some more examples:

$$\sum_{j=3}^{7} j^2 = 3^2+4^2+5^2+6^2+7^2,$$

$$\frac{1}{\sqrt{1}}+\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}+\frac{1}{\sqrt{4}}+\frac{1}{\sqrt{5}} = \sum_{i=1}^{5}\frac{1}{\sqrt{i}} = \sum_{k=2}^{6}\frac{1}{\sqrt{k-1}},$$

$$\underbrace{1}_{a_1}+\underbrace{4}_{a_2}+\underbrace{9}_{a_3}+\underbrace{16}_{a_4}+\cdots+\underbrace{100}_{a_{10}} = \sum_{i=1}^{10} i^2,$$

$$\sum_{k=1}^{n}\frac{1}{n}\big(k^2+1\big) = \frac{1}{n}\big(1^2+1\big)+\frac{1}{n}\big(2^2+1\big)+\cdots+\frac{1}{n}\big(n^2+1\big).$$

In the last sum, n appears both as the upper limit and inside every term. That is fine: k is the index, and n is a fixed number while we add.

Proposition 5.3.4 (Linearity of finite sums)**.** *Let $\{a_i\}_{i=1}^n$ and $\{b_i\}_{i=1}^n$ be finite collections of real numbers, and let $c \in \mathbb{R}$ be a scalar. Then:*

(i) Homogeneity: $\displaystyle\sum_{i=1}^{n} c\,a_i = c\sum_{i=1}^{n} a_i$;

(ii) Additivity: $\displaystyle\sum_{i=1}^{n}(a_i \pm b_i) = \sum_{i=1}^{n} a_i \pm \sum_{i=1}^{n} b_i$.

Proof. Both statements are the familiar rules of arithmetic, applied n times. For (i), the distributive law gives

$$ca_1 + ca_2 + \cdots + ca_n = c\,(a_1 + a_2 + \cdots + a_n).$$

For (ii), since we may reorder and regroup a *finite* sum however we like, we get

$$(a_1 \pm b_1) + (a_2 \pm b_2) + \cdots + (a_n \pm b_n) = (a_1 + a_2 + \cdots + a_n) \pm (b_1 + b_2 + \cdots + b_n).$$

With $n = 2$ these identities read $ca_1 + ca_2 = c\,(a_1 + a_2)$ and $(a_1 + b_1) + (a_2 + b_2) = (a_1 + a_2) + (b_1 + b_2)$. □

⚠ *Caveat. Finite means finite*

The word *finite* in Proposition 5.3.4 matters. For infinite sums (series), we may no longer regroup and reorder terms freely. Riemann showed that the terms of some convergent series can be rearranged to add up to *any* real number you choose.

Theorem 5.3.5 (Power sum formulas for integers)**.** *For every positive integer $n \in \mathbb{N}$ and every constant $c \in \mathbb{R}$:*

(i) $\displaystyle\sum_{k=1}^{n} c = cn;$

(ii) $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2};$

(iii) $\displaystyle\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6};$

(iv) $\displaystyle\sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4} = \Big(\sum_{k=1}^{n} k\Big)^2.$

Proof. We prove (i)–(iii). Part (iv) is left as Exercise 5.3.7.

(i) The sum has n terms, each equal to c:

$$\sum_{k=1}^{n} c = \underbrace{c + c + \cdots + c}_{n \text{ terms}} = nc.$$

(ii) We want to show that $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$. Write the sum twice, once forward and once backward, and add the two lines column by column:

$$\begin{aligned}
\sum_{k=1}^{n} k &= \quad 1 \quad + \ 2 \quad + \ 3 \quad + \ \cdots \quad + \ n \\
\sum_{k=1}^{n} k &= \quad n \quad + (n-1) + (n-2) + \ \cdots \quad + \ 1 \\
\hline
2\sum_{k=1}^{n} k &= (n+1) + (n+1) + (n+1) + \ \cdots + (n+1) = n(n+1).
\end{aligned}$$

Each column adds up to $n+1$, and there are n columns. Dividing by 2 gives

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

(iii) We prove that

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \tag{5.3}$$

by induction on n.

Base step. For $n = 1$,

$$\text{LHS} = \sum_{k=1}^{1} k^2 = 1^2 = 1, \qquad \text{RHS} = \frac{1\,(1+1)\,(2\cdot 1+1)}{6} = \frac{1\cdot 2\cdot 3}{6} = 1,$$

so (5.3) holds for $n = 1$.

Induction step. Suppose (5.3) holds for some $n \geq 1$. We need to show that it also holds for $n+1$. The idea is to split off the last term so that we can use the induction hypothesis on the rest:

$$\begin{aligned}
\sum_{k=1}^{n+1} k^2 &= \bigl(1^2 + 2^2 + \cdots + n^2\bigr) + (n+1)^2 = \Bigl(\sum_{k=1}^{n} k^2\Bigr) + (n+1)^2 \\
&\overset{(5.3)}{=} \frac{n(n+1)(2n+1)}{6} + (n+1)^2 = \frac{n(n+1)(2n+1) + 6(n+1)^2}{6} \\
&= \frac{(n+1)\bigl(n(2n+1) + 6(n+1)\bigr)}{6} = \frac{(n+1)\bigl(2n^2 + 7n + 6\bigr)}{6} \\
&= \frac{(n+1)(n+2)(2n+3)}{6} = \frac{(n+1)\bigl[(n+1)+1\bigr]\bigl[2(n+1)+1\bigr]}{6}.
\end{aligned}$$

The last expression is the right-hand side of (5.3) with n replaced by $n+1$. So the identity holds for $n+1$.

By induction, (5.3) holds for every positive integer n. □

> ✠ *Story. The schoolboy and the hundred numbers*
>
> A famous story says that the young Carl Friedrich Gauss was told to add the numbers from 1 to 100 as a punishment. He gave the answer, 5050, almost at once, by pairing 1 with 100, 2 with 99, and so on. That is exactly the forward–backward argument in part (ii). The story may well be embellished, but the trick is much older. Archimedes used a formula equivalent to part (iii) in *On Spirals* (around 225 BC) to compute areas, which is the same job we are preparing these sums for. The cube identity in part (iv) is often called *Nicomachus's theorem*, after Nicomachus of Gerasa (around AD 100).

Remark 5.3.6 (Why we care). Proposition 5.3.4 and Theorem 5.3.5 together let us compute sums like the last one in Example 5.3.3 exactly:

$$\sum_{k=1}^{n} \frac{1}{n}\bigl(k^2 + 1\bigr) = \frac{1}{n}\Bigl(\sum_{k=1}^{n} k^2 + \sum_{k=1}^{n} 1\Bigr) = \frac{1}{n}\Bigl(\frac{n(n+1)(2n+1)}{6} + n\Bigr) = \frac{(n+1)(2n+1)}{6} + 1.$$

When we approximate areas under polynomial graphs by n rectangles, we will get sums of exactly this kind. Then we will let $n \to \infty$.

Exercise 5.3.7 (Sum of cubes)**.** Prove part (iv) of Theorem 5.3.5:

$$\sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4} = \Bigl(\sum_{k=1}^{n} k\Bigr)^2.$$

Bonus: Can you provide a geometric interpretation of the summation? *Hint:* Use induction, following the proof of part (iii). In the induction step, factor out $(n+1)^2$ from $\frac{n^2(n+1)^2}{4} + (n+1)^3$.

5.4 Riemann Sums

Story. From Archimedes' Parabolic Quadrature to Riemann's Discrete Slicing

Centuries before the systematic formalization of the calculus by Newton and Leibniz, Archimedes of Syracuse (c. 287–212 BCE) achieved the first rigorous integration in history: the quadrature of the parabola. Armed with only Euclidean geometry and the method of exhaustion, Archimedes inscribed infinite sequences of rectilinear geometric triangles inside curved parabolic segments, summing geometric progressions to compute curvilinear areas.

Yet for two millennia, each specific geometric shape required its own bespoke, virtuosic trick. In the mid-nineteenth century, Bernhard Riemann revolutionized this viewpoint with a radical conceptual democratization: rather than inventing tailored geometric shapes to fit the curvature of a boundary, one slices the entire underlying space into uniform vertical strips of arbitrary width Δx_k. In doing so, the intricate geometry of curved space is tamed by the elementary arithmetic of rectangles:

$$\text{Curved Area} \approx \sum_{k=1}^{n} (\text{Height}) \times (\text{Width}) = \sum_{k=1}^{n} f(x_k^*)\,\Delta x_k.$$

The modern definite integral is nothing other than the continuum limit of this discrete book-keeping.

In Section 5.2 we saw that net displacement and total distance both come down to one geometric problem: finding the area between the graph of a function, two vertical lines $t = a$ and $t = b$, and the horizontal axis. We now carry out the plan from Section 5.1. We cut the region into thin vertical strips, replace each strip by a rectangle as in Figure 5.4, and add up the areas of the rectangles using the sigma notation of Section 5.3.

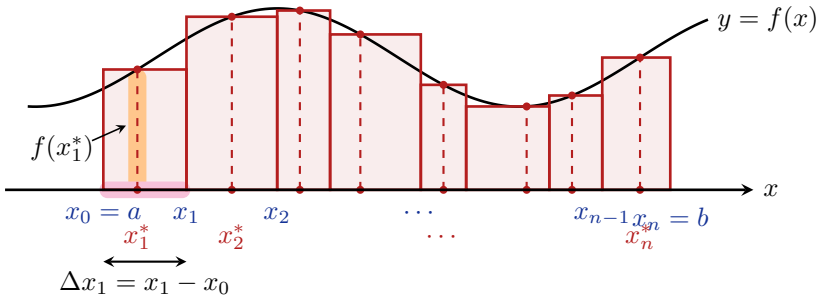


Figure 5.4 – A Riemann sum $\sum_{i=1}^{n} f(x_i^*)\,\Delta x_i$ on a non-uniform partition $a = x_0 < x_1 < \cdots < x_n = b$, with arbitrary sample points $x_i^* \in [x_{i-1}, x_i]$.

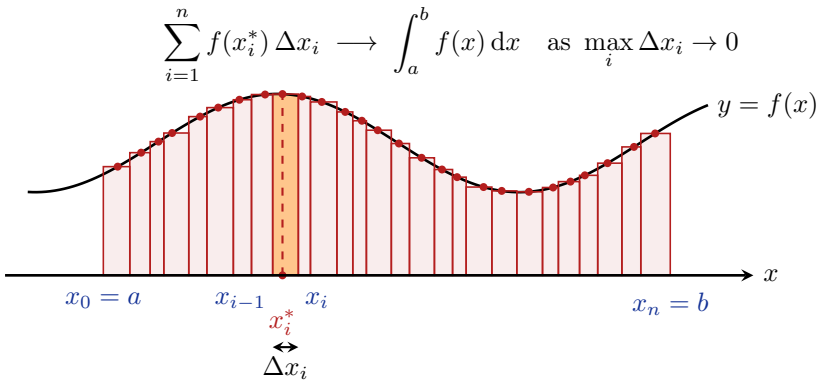


Figure 5.5 – Refining the partition: with a fine non-uniform mesh, the Riemann sum $\sum_{i=1}^{n} f(x_i^*)\,\Delta x_i$ approaches $\int_a^b f(x)\,\mathrm{d}x$ as the mesh size $\max_i \Delta x_i \to 0$.

Definition 5.4.1 (Partition)**.** Let $[a, b]$ be a closed, bounded interval. A *partition* $\mathcal{P}$ of $[a, b]$ is a finite ordered set of points

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

The partition divides $[a, b]$ into n adjacent subintervals $[x_{k-1}, x_k]$. The *of a subinterval width* of the k-th subinterval is

$$\Delta x_k \coloneqq x_k - x_{k-1}, \qquad k = 1, \ldots, n.$$

The *of a partition norm* of $\mathcal{P}$ is the length of its longest subinterval:

$$\|\mathcal{P}\| \coloneqq \max_{1 \le k \le n} \Delta x_k.$$

Definition 5.4.2 (Tagged partition)**.** A *tagged partition* is a partition $\mathcal{P}$ together with a choice of *sample point* (or *tag*) for each subinterval: an arbitrary point

$$x_k^* \in [x_{k-1}, x_k], \qquad k = 1, \ldots, n.$$

The norm $\|\mathcal{P}\|$ measures how coarse the partition is. A small norm means *every* strip is thin, and that is what makes the rectangles good approximations of the strips.

⚠ *Caveat. More points does not mean finer*

For a general partition, adding more subintervals does not by itself make the partition finer. Take the partition of $[0, 1]$ given by

$$0 < \tfrac{1}{2} < \tfrac{1}{2} + \tfrac{1}{2n} < \tfrac{1}{2} + \tfrac{2}{2n} < \cdots < 1.$$

It has $n + 1$ subintervals, but its norm is always $\frac{1}{2}$, no matter how large n is. The subinterval $[0, \frac{1}{2}]$ never shrinks. This is why the norm $\|\mathcal{P}\|$, and not the number of points n, is the right way to measure how fine a partition is.

Definition 5.4.3 (Riemann sum)**.** Let $f\colon [a,b] \to \mathbb{R}$ be a function defined on $[a,b]$. Let $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ be a tagged partition of $[a,b]$ with sample points $\{x_k^*\}_{k=1}^n$. The corresponding *Riemann sum* is

$$R\big(f, \mathcal{P}, \{x_k^*\}\big) \coloneqq \sum_{k=1}^{n} f(x_k^*)\,\Delta x_k.$$

The following special cases come up often.

- If $\Delta x_k = \dfrac{b-a}{n}$ for all $1 \le k \le n$, then $\mathcal{P}$ is called a *partition regular*. In this case the partition points are $x_k = a + k\,\dfrac{b-a}{n}$.
- If $x_k^* = x_{k-1}$ for every k, so that each tag is the left endpoint of its subinterval, then $R\big(f, \mathcal{P}, \{x_{k-1}\}\big)$ is called the *left Riemann sum*.
- If $x_k^* = x_k$ for every k, so that each tag is the right endpoint of its subinterval, then $R\big(f, \mathcal{P}, \{x_k\}\big)$ is called the *right Riemann sum*.
- If $x_k^* = \frac{1}{2}\,(x_{k-1} + x_k)$ for every k, so that each tag is the middle point of its subinterval, then $R\big(f, \mathcal{P}, \{x_k\}\big)$ is called the *middle Riemann sum*.

Each term $f(x_k^*)\,\Delta x_k$ is the area of one rectangle in Figure 5.5: its height is $f(x_k^*)$ and its width is Δx_k. The Riemann sum adds up all n of these areas.

Intuition. Riemann sums are piecewise-constant physics

Go back to the trip from Section 5.2, with $f = v$ the velocity. Over a short time interval $[t_{k-1}, t_k]$ the velocity hardly changes. So we can pretend it is constant and equal to its value $v(t_k^*)$ at a single sample time. At constant velocity, displacement is just velocity times time, so the displacement over that interval is approximately $v(t_k^*)\,\Delta t_k$. Adding these pieces gives

$$s(t_n) - s(t_0) \approx \sum_{k=1}^{n} v(t_k^*)\,\Delta t_k.$$

A Riemann sum is what you get when you replace a smoothly varying quantity by a step function and use the rule "rate $\times$ time". Which sample time you pick in each interval should matter less and less as the intervals get shorter.

Caveat. Signed rectangles

Nothing in Definition 5.4.3 says f must be positive. If $f(x_k^*) < 0$, then the term $f(x_k^*)\,\Delta x_k$ is negative, and that rectangle counts as *negative* area. So a Riemann sum approximates the *signed* area: area above the axis minus area below it. This is exactly the net-change bookkeeping from Section 5.2.

Story. Cauchy, Riemann, and the freedom to choose

In 1823 Augustin-Louis Cauchy defined the integral of a continuous function as the limit of sums in which the sample point is always the left endpoint of each subinterval. Bernhard Riemann went further in his 1854 *Habilitationsschrift* on trigonometric series, which was published only after his death, in 1868. He allowed the sample point x_k^* to be *any* point of $[x_{k-1}, x_k]$. He then asked exactly which functions give sums that settle down to a single number as $\|\mathcal{P}\| \to 0$, no matter how the tags are chosen. Giving that freedom of choice is what made the theory work for functions far more general than continuous ones.

Example 5.4.4 (Left and right sums for x^2)**.** Let $f(x) = x^2$ on $[0, 3]$. Take the regular partition $\mathcal{P}$ with $n = 3$:

$$\begin{aligned} x_0 &= 0, \\ x_1 &= 1, \qquad & \Delta x_1 &= x_1 - x_0 = 1, \\ x_2 &= 2, \qquad & \Delta x_2 &= x_2 - x_1 = 1, \\ x_3 &= 3, \qquad & \Delta x_3 &= x_3 - x_2 = 1, \end{aligned}$$

so that

$$\|\mathcal{P}\| = \max_{1 \le i \le 3} \Delta x_i = 1.$$

Left Riemann sum. The tags are the left endpoints, $x_k^* = x_{k-1}$ for $1 \le k \le 3$, that is,

$$x_1^* = x_0 = 0, \qquad x_2^* = x_1 = 1, \qquad x_3^* = x_2 = 2.$$

Hence

$$\begin{aligned} R\big(f, \mathcal{P}, \{x_{k-1}\}_{k=1}^3\big) &= \sum_{k=1}^{3} f(x_k^*)\,\Delta x_k = \sum_{k=1}^{3} f(x_{k-1})\,\Delta x_k = \sum_{k=1}^{3} x_{k-1}^2 \cdot 1 \\ &= x_0^2 + x_1^2 + x_2^2 = 0^2 + 1^2 + 2^2 = 5. \end{aligned}$$

Right Riemann sum. The tags are the right endpoints, $x_k^* = x_k$ for $1 \le k \le 3$, that is,

$$x_1^* = x_1 = 1, \qquad x_2^* = x_2 = 2, \qquad x_3^* = x_3 = 3.$$

Hence, by Theorem 5.3.5 (iii) with $n = 3$,

$$\begin{aligned} R\big(f, \mathcal{P}, \{x_k\}_{k=1}^3\big) &= \sum_{k=1}^{3} x_k^2 \cdot 1 = 1^2 + 2^2 + 3^2 = \sum_{k=1}^{3} k^2 \\ &= \frac{3\,(3+1)\,(2 \cdot 3 + 1)}{6} = 2\,(6+1) = 14. \end{aligned}$$

Remark 5.4.5 (Trapping the area). The function $f(x) = x^2$ is *increasing* on $[0, 3]$. So on each subinterval its smallest value is at the left endpoint and its largest value is at the right endpoint. The left rectangles therefore lie *under* the graph and the right rectangles lie *over* it (see Figure 5.6). The true area A is caught between the two sums:

$$5 \le A \le 14.$$

This is a very wide window. To narrow it, we refine the partition.

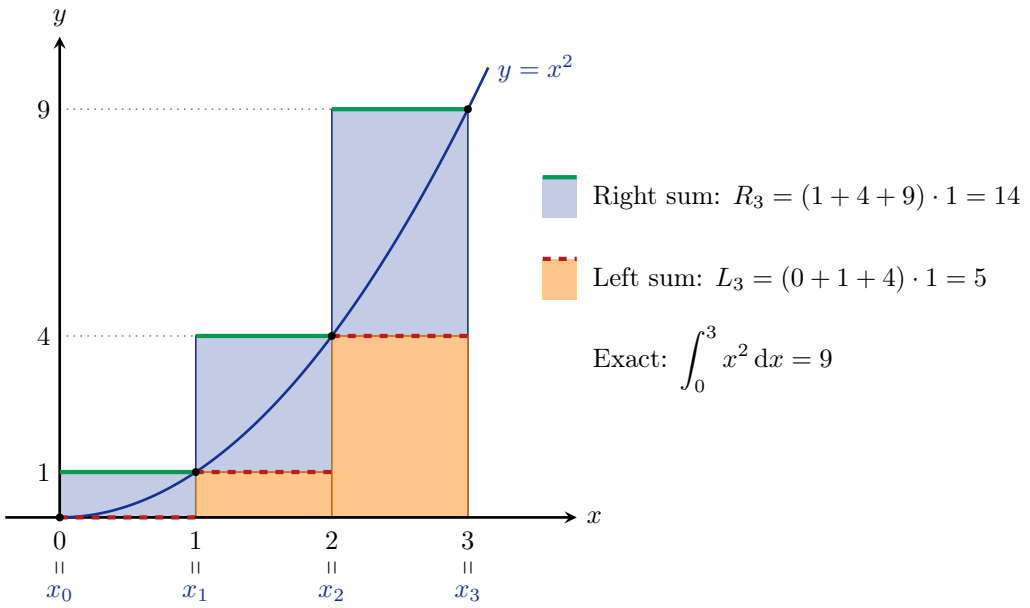


Figure 5.6 – Left and right Riemann sums for $f(x) = x^2$ on $[0, 3]$ with $n = 3$ and $\Delta x = 1$. The left sum $L_3 = 5$ (orange) underestimates and the right sum $R_3 = 14$ (whole columns, blue and orange together) overestimates the exact value $\int_0^3 x^2 \, dx = 9$, because f is increasing.

Remark 5.4.6 (Letting $n \to \infty$). Repeat Example 5.4.4 with a regular partition into n subintervals. Then $\Delta x_k = \frac{3}{n}$ and $x_k = \frac{3k}{n}$. By Proposition 5.3.4 and Theorem 5.3.5 (iii),

$$R\big(f, \mathcal{P}_n, \{x_k\}\big) = \sum_{k=1}^{n} \Big(\frac{3k}{n}\Big)^2 \frac{3}{n} = \frac{27}{n^3} \sum_{k=1}^{n} k^2 = \frac{27}{n^3} \cdot \frac{n(n+1)(2n+1)}{6} = \frac{9\,(n+1)(2n+1)}{2n^2},$$

$$R\big(f, \mathcal{P}_n, \{x_{k-1}\}\big) = \sum_{k=1}^{n} \Big(\frac{3(k-1)}{n}\Big)^2 \frac{3}{n} = \frac{27}{n^3} \sum_{j=1}^{n-1} j^2 = \frac{9\,(n-1)(2n-1)}{2n^2}.$$

For $n = 3$ these formulas give 14 and 5, which matches Example 5.4.4. As $n \to \infty$, both sums tend to $\frac{9 \cdot 2}{2} = 9$. The window closes, and the area under $y = x^2$ over $[0, 3]$ is 9. Turning this squeeze into a definition is the subject of the next section: the definite integral.

> *Moral. The mechanics of a Riemann integral*
>
> A Riemann sum turns an area problem into an addition problem. The power sum formulas of Theorem 5.3.5 turn that addition problem into a closed-form expression in n. Taking the limit $\|\mathcal{P}\| \to 0$ turns the approximation into an exact answer.

5.5 The Definite Integral

Definition 5.5.1 (Definite Integral)**.** A function $f\colon [a, b] \to \mathbb{R}$ is said to be *integrable* (in the sense of Riemann) on $[a, b]$ if there exists a number $I \in \mathbb{R}$ such that for every $\varepsilon > 0$,

there exists $\delta > 0$ with

$$\left| \sum_{k=1}^{n} f(x_k^*)\,\Delta x_k - I \right| < \varepsilon$$

for every partition $\mathcal{P}$ with mesh $\|\mathcal{P}\| < \delta$, irrespective of the choice of sample points $x_k^* \in [x_{k-1}, x_k]$. When this limit exists uniquely, we write

$$\int_a^b f(x)\,\mathrm{d}x \coloneqq \lim_{\|\mathcal{P}\|\to 0} \sum_{k=1}^{n} f(x_k^*)\,\Delta x_k, \tag{5.4}$$

and I is called the *definite integral* of f from a to b.

> *Moral. Leibniz's Notation as an Asymptotic Idealization*
>
> The symbol $\int$ introduced by Gottfried Wilhelm Leibniz is an elongated stylized letter 'S' for *summa* (Latin for "sum"). As the partition mesh $\|\mathcal{P}\| \to 0$:
>
> $$\sum \longrightarrow \int \qquad \text{and} \qquad \Delta x_k \longrightarrow \mathrm{d}x.$$
>
> The differential $\mathrm{d}x$ reminds us that we are summing infinitely many infinitesimally narrow rectangles, each of height $f(x)$ and base $\mathrm{d}x$.

Theorem 5.5.2 (Classes of Integrable Functions)**.** *Let $f\colon [a,b] \to \mathbb{R}$. If either of the following conditions is satisfied:*

(i) *f is continuous on $[a,b]$, or*

(ii) *f is bounded on $[a,b]$ and possesses at most finitely many points of discontinuity,*

then f is Riemann integrable on $[a,b]$.

The proof of this theorem is sophisticated, we present that in Section 5.9 as a bonus (non-examinable) for the curious minds.

> *Caveat. Non-Integrable Pathologies: The Dirichlet Function*
>
> Not every bounded function is Riemann integrable. Consider the classic indicator function of the rational numbers on $[0,1]$:
>
> $$f(x) \coloneqq \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$$
>
> Because both the rational numbers $\mathbb{Q}$ and irrational numbers $\mathbb{R} \setminus \mathbb{Q}$ are dense in every open interval, one can select rational tags to force $\sum f(x_k^*)\Delta x_k = 1$, or irrational tags to force $\sum f(x_k^*)\Delta x_k = 0$. The Riemann limit fails to exist, which historically compelled Henri Lebesgue to develop measure theory and Lebesgue integration.

Example 5.5.3 (Identifying the Definite Integral from a Limit of Riemann Sums)**.** Assume that

$$L \coloneqq \lim_{\|\mathcal{P}\|\to 0} \sum_{k=1}^{n} \left(3(x_k^*)^2 + 2x_k^* + 1\right) \Delta x_k$$

is the limit of a Riemann sum for a function f on the interval $[1, 3]$.

(i) Identify the underlying integrand $f(x)$.

(ii) Express the limit as a definite integral.

(iii) Provide the geometric interpretation of the integral.

Solution. Comparing the summand $f(x_k^*)\,\Delta x_k$ with the general Riemann sum structure:

$$f(x_k^*) = 3(x_k^*)^2 + 2x_k^* + 1 \implies f(x) = 3x^2 + 2x + 1.$$

Because f is a polynomial, it is continuous everywhere on $\mathbb{R}$, and hence integrable on $[1, 3]$ by Theorem 5.5.2. The limit of the Riemann sum converges to the definite integral:

$$\lim_{\|\mathcal{P}\|\to 0} \sum_{k=1}^{n} \left(3(x_k^*)^2 + 2x_k^* + 1\right) \Delta x_k = \int_1^3 \left(3x^2 + 2x + 1\right) \mathrm{d}x.$$

Because $f(x) = 3x^2 + 2x + 1 > 0$ for all $x \in [1, 3]$, this definite integral represents the exact geometric area of the planar region bounded between the parabola $y = 3x^2 + 2x + 1$, the horizontal axis $y = 0$, and the vertical lines $x = 1$ and $x = 3$.

□

5.6 Properties of the Definite Integral

Recall from the previous section what it means for a function $f\colon [a, b] \to \mathbb{R}$ to be *integrable*. The Riemann sums of Definition 5.4.3 must approach a single number as $\|\mathcal{P}\| \to 0$, whatever sample points we choose. That number is the definite integral:

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{\|\mathcal{P}\|\to 0} \sum_{k=1}^{n} f(x_k^*)\,\Delta x_k.$$

Computing this limit from scratch every time would be exhausting. In this section we collect the rules that let us work with integrals as algebraic objects. Every rule has the same proof strategy. We go back to Riemann sums, use the corresponding rule for *finite* sums (Proposition 5.3.4), and then pass to the limit. In short, the integral inherits its good behaviour from the finite sums that define it.

We will use one computation repeatedly. Let $m \in \mathbb{R}$ be a constant. Every Riemann sum of the constant function m is

$$\sum_{k=1}^{n} m\,\Delta x_k = m \sum_{k=1}^{n} (x_k - x_{k-1}) = m\,(x_n - x_0) = m\,(b - a).$$

The middle sum *telescopes*: each x_k is added once and subtracted once, except x_n and x_0. Since every Riemann sum equals $m(b - a)$, so does the limit:

$$\int_a^b m\,\mathrm{d}x = m\,(b - a). \tag{5.5}$$

This is just the area of a rectangle of height m and width $b-a$.

So far, the symbol $\int_a^b$ only makes sense when $a<b$. Two conventions extend it to every pair of limits.

Definition 5.6.1 (Oriented integrals)**.** Let $f\colon [a,b]\to\mathbb{R}$ be integrable. We define:

(i) $\displaystyle\int_c^c f(t)\,\mathrm{d}t := 0$ for every $c\in[a,b]$;

(ii) $\displaystyle\int_b^a f(t)\,\mathrm{d}t := -\int_a^b f(t)\,\mathrm{d}t.$

Intuition. Zero width and running the film backward

Convention (i) says that a strip of zero width has zero area. For convention (ii), imagine forming a Riemann sum while walking from b to a. The partition points now *decrease*, so every width $\Delta x_k = x_k - x_{k-1}$ is negative, and every term of the sum changes sign. In the language of Section 5.2, take $f=v$ to be a velocity. Swapping the limits is like playing a film of the trip backward: every forward step becomes a backward step, so the net displacement changes sign.

Proposition 5.6.2 (Properties of the definite integral)**.** *Suppose f and g are integrable functions on $[a,b]$, and let $c\in[a,b]$. Then:*

(i) Additivity in the integrand: $\displaystyle\int_a^b \big(f(x)+g(x)\big)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x$;

(ii) Homogeneity: $\displaystyle\int_a^b \lambda f(x)\,\mathrm{d}x = \lambda\int_a^b f(x)\,\mathrm{d}x$ *for every constant* $\lambda\in\mathbb{R}$;

(iii) Additivity over intervals: $\displaystyle\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x$;

(iv) Positivity: *if* $f(x)\ge 0$ *on* $[a,b]$, *then* $\displaystyle\int_a^b f(x)\,\mathrm{d}x \ge 0$.

Proof. (i) We start from the definition and use Proposition 5.3.4 (ii) on each finite sum:

$$\begin{aligned}\int_a^b \big(f(x)+g(x)\big)\,\mathrm{d}x &= \lim_{\|\mathcal{P}\|\to 0}\sum_{k=1}^{n}\big(f(x_k^*)+g(x_k^*)\big)\,\Delta x_k\\ &= \lim_{\|\mathcal{P}\|\to 0}\Big(\sum_{k=1}^{n} f(x_k^*)\,\Delta x_k + \sum_{k=1}^{n} g(x_k^*)\,\Delta x_k\Big)\\ &= \lim_{\|\mathcal{P}\|\to 0}\sum_{k=1}^{n} f(x_k^*)\,\Delta x_k + \lim_{\|\mathcal{P}\|\to 0}\sum_{k=1}^{n} g(x_k^*)\,\Delta x_k\\ &= \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x.\end{aligned}$$

The only delicate step is the third equality, where the limit of a sum becomes a sum of limits. This is allowed because *both* limits exist, and that is exactly what it means for f and g to be

integrable. We also used the same tags x_k^* for f and for g. We may do so because the limit does not depend on how the tags are chosen.

(ii) This is left as Exercise 5.6.8 (a).

(iii) Intuitively, the region over $[a,b]$ is the union of the region over $[a,c]$ and the region over $[c,b]$, so

$$\text{Area on } [a,b] = \text{Area on } [a,c] + \text{Area on } [c,b]$$

(see Figure 5.7).

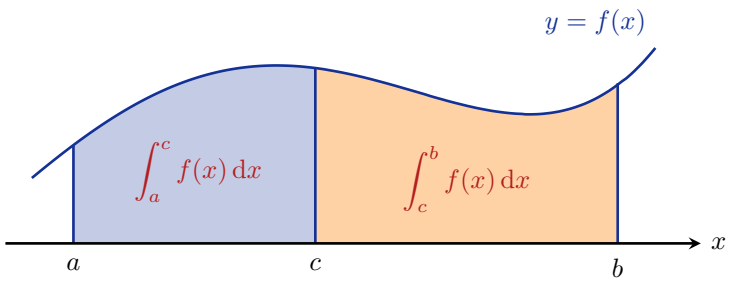


Figure 5.7 – Additivity of the Riemann integral.

To turn this picture into an argument, we cut f into two pieces. Define two auxiliary functions on $[a,b]$ (unrelated to the g in the statement):

$$g(x) := \begin{cases} f(x), & x \in [a,c], \\ 0, & x \in (c,b], \end{cases} \qquad h(x) := \begin{cases} 0, & x \in [a,c], \\ f(x), & x \in (c,b]. \end{cases}$$

The function g keeps the part of f over $[a,c]$ and switches off the rest. The function h does the opposite. Hence

$$\int_a^b g(x)\,dx = \int_a^c f(x)\,dx, \qquad \int_a^b h(x)\,dx = \int_c^b f(x)\,dx.$$

Since $f(x) = g(x) + h(x)$ for every $x \in [a,b]$, part (i) gives

$$\int_a^b f(x)\,dx = \int_a^b g(x)\,dx + \int_a^b h(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

(iv) Since $f \geq 0$ on $[a,b]$, we have $f(x_k^*) \geq 0$ for every choice of tag $x_k^* \in [a,b]$. Each width also satisfies $\Delta x_k > 0$. Therefore every Riemann sum is nonnegative:

$$\int_a^b f(x)\,dx = \lim_{\|\mathcal{P}\| \to 0} \underbrace{\sum_{k=1}^{n} \underbrace{f(x_k^*)}_{\geq 0} \underbrace{\Delta x_k}_{>0}}_{\geq 0} \geq 0.$$

The last inequality holds because a limit of nonnegative numbers is nonnegative.

□

⚠ *Caveat. What the proof of (iii) takes for granted*

The argument for (iii) uses two facts without proof. First, the cut-off functions g and h are integrable on $[a,b]$. Second, $\int_a^b g(x)\,dx = \int_a^c f(x)\,dx$. Both are believable. In a Riemann sum for g, every rectangle lying over $(c,b]$ has height zero. At most one strip straddles the point c, and its width is at most $\|\mathcal{P}\|$, so its contribution disappears in the limit. A complete proof uses the fact that integrable functions are bounded, and it belongs in a course on real analysis. Notice also why we used the half-open interval $(c,b]$. If both g and h were equal to f at $x = c$, we would get $g(c) + h(c) = 2f(c) \neq f(c)$ in general.

Moral.

The integral inherits linearity, (i)–(ii), from finite sums. It inherits additivity over intervals, (iii), from the geometry of area. It inherits positivity, (iv), from the plain fact that a sum of nonnegative numbers is nonnegative.

Proposition 5.6.2 (iii) requires the splitting point c to lie *between* a and b. Thanks to the conventions of Definition 5.6.1, this restriction can be dropped.

Corollary 5.6.3 (Additivity in any order)**.** *Let f be integrable on an interval I, and let $a, b, c \in I$ be in any order. Then*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

Equivalently, $\displaystyle\int_a^c f(x)\,dx + \int_c^b f(x)\,dx + \int_b^a f(x)\,dx = 0.$

Proof. First suppose that two of the three points coincide. If $c = a$ or $c = b$, one of the integrals on the right is 0 by Definition 5.6.1 (i), and the identity is immediate. If $a = b$, the left-hand side is 0, and the right-hand side is $\int_a^c f + \int_c^a f = 0$ by Definition 5.6.1 (ii).

Now suppose the three points are distinct. There are six possible orderings. The case $a < c < b$ is Proposition 5.6.2 (iii). Consider, for instance, the case $a < b < c$. Splitting $[a,c]$ at the point b gives

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx,$$

so

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx - \int_b^c f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

by Definition 5.6.1 (ii). The remaining four orderings are handled in the same way (Exercise 5.6.8 (c)). □

Intuition. Round trips

The equivalent form says that the integrals along the closed loop $a \to c \to b \to a$ add up to zero. With f a velocity, this is the familiar fact that a round trip has

zero net displacement, whatever route you take. This is why the conventions of Definition 5.6.1 are the *right* ones: they are exactly what makes this bookkeeping work without exceptions.

⚠ *Caveat. Not every property survives reversed limits*

Properties (i)–(iii) remain true when the limits are reversed, because each side simply changes sign. Positivity does *not*. For example,

$$\int_1^0 1\,\mathrm{d}x = -\int_0^1 1\,\mathrm{d}x = -1 < 0,$$

even though the integrand is positive. Positivity, and the comparison results below, require the lower limit to be at most the upper limit.

5.6.1 Bounds and comparison

Positivity is a small statement, but it has useful consequences.

Corollary 5.6.4 (Bounds and comparison)**.** *Let f and g be integrable on $[a,b]$, and let $m, M \in \mathbb{R}$.*

(i) *If $f(x) \geq m$ on $[a,b]$, then $\displaystyle\int_a^b f(x)\,\mathrm{d}x \geq m\,(b-a)$. Consequently, if $m \leq f(x) \leq M$ on $[a,b]$, then*

$$m\,(b-a) \leq \int_a^b f(x)\,\mathrm{d}x \leq M\,(b-a).$$

(ii) *If $f(x) \leq g(x)$ on $[a,b]$, then $\displaystyle\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x$.*

Proof. (i) Suppose $f(x) \geq m$ on $[a,b]$. Then $f(x) - m \geq 0$ on $[a,b]$, so by positivity, Proposition 5.6.2 (iv),

$$\int_a^b \big(f(x) - m\big)\,\mathrm{d}x \geq 0.$$

By Proposition 5.6.2 (i) and (5.5),

$$\int_a^b f(x)\,\mathrm{d}x + \int_a^b (-m)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + (-m)(b-a) \geq 0,$$

that is, $\int_a^b f(x)\,\mathrm{d}x \geq m\,(b-a)$. The upper bound $\int_a^b f(x)\,\mathrm{d}x \leq M\,(b-a)$ follows in the same way by applying positivity to $M - f(x) \geq 0$ (Exercise 5.6.8 (b)).

(ii) If $f(x) \leq g(x)$ on $[a,b]$, then $g(x) - f(x) \geq 0$ on $[a,b]$. By Proposition 5.6.2 (i) and (ii) (with $\lambda = -1$), followed by (iv),

$$\int_a^b g(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x = \int_a^b \big(g(x) - f(x)\big)\,\mathrm{d}x \geq 0,$$

hence $\int_a^b g(x)\,\mathrm{d}x \geq \int_a^b f(x)\,\mathrm{d}x$. □

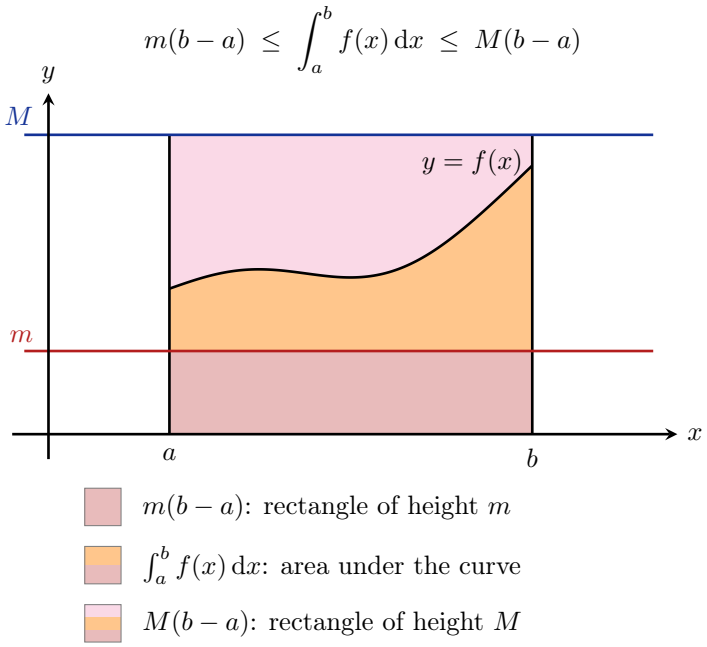


Figure 5.8 – Comparison property of the Riemann integral: if $m \le f(x) \le M$ for all $x \in [a,b]$, then $m(b-a) \le \int_a^b f(x)\,dx \le M(b-a)$.

Intuition. The speedometer estimate

Suppose that during a two-hour drive your speedometer always reads between 40 and 60 miles per hour. Without knowing anything else, you can be sure that you covered between 80 and 120 miles. This is Corollary 5.6.4 (i) with $m = 40$, $M = 60$ and $b - a = 2$. Dividing the inequality by $b - a$ gives an even more suggestive form: the *average value* $\frac{1}{b-a}\int_a^b f(x)\,dx$ of f lies between the smallest and largest values of f. We will return to this idea when we meet the Mean Value Theorem for integrals.

Example 5.6.5 (Using the rules)**.** Assume that

$$\int_0^5 f(x)\,dx = 3 \qquad \text{and} \qquad \int_0^7 f(x)\,dx = -10.$$

Evaluate the following integrals, if possible.

(i) By homogeneity,

$$\int_0^7 2f(x)\,dx = 2\int_0^7 f(x)\,dx = 2\cdot(-10) = -20.$$

(ii) Split at the point 0. It lies *outside* $[5,7]$, so we need Corollary 5.6.3 rather than

Proposition 5.6.2 (iii):

$$\int_5^7 f(x)\,\mathrm{d}x = \int_5^0 f(x)\,\mathrm{d}x + \int_0^7 f(x)\,\mathrm{d}x = -\int_0^5 f(x)\,\mathrm{d}x + \int_0^7 f(x)\,\mathrm{d}x$$
$$= -3 + (-10) = -13.$$

(iii) Combining homogeneity with Definition 5.6.1 (ii),

$$\int_7^0 6f(x)\,\mathrm{d}x = 6\int_7^0 f(x)\,\mathrm{d}x = 6\Big(-\int_0^7 f(x)\,\mathrm{d}x\Big) = 6\big(-(-10)\big) = 60.$$

(iv) The value of $\displaystyle\int_0^7 |f(x)|\,\mathrm{d}x$ *cannot* be determined from the given information.

Remark 5.6.6 (Why part (iv) cannot be computed). The data in Example 5.6.5 are *net* quantities: signed areas in which positive and negative parts cancel. The integral of $|f|$ measures the *total* area, with no cancellation. As with the GPS and the odometer in Section 5.2, the net amount does not determine the total. Here are two step functions (step functions are integrable) that both satisfy the hypotheses:

$$f_1(x) = \begin{cases} \frac{3}{5}, & 0 \le x \le 5, \\ -\frac{13}{2}, & 5 < x \le 7, \end{cases} \qquad f_2(x) = \begin{cases} \frac{53}{5}, & 0 \le x \le \frac{5}{2}, \\ -\frac{47}{5}, & \frac{5}{2} < x \le 5, \\ -\frac{13}{2}, & 5 < x \le 7. \end{cases}$$

Both have $\int_0^5 = 3$ and $\int_5^7 = -13$, hence $\int_0^7 = -10$. However,

$$\int_0^7 |f_1(x)|\,\mathrm{d}x = 3 + 13 = 16, \qquad \int_0^7 |f_2(x)|\,\mathrm{d}x = \tfrac{53}{2} + \tfrac{47}{2} + 13 = 63.$$

Comparison does give a lower bound. Apply Corollary 5.6.4 (ii) to $-|f| \le f \le |f|$, assuming $|f|$ is integrable (it always is when f is, a fact proved in real analysis). This gives $\left|\int_\alpha^\beta f\right| \le \int_\alpha^\beta |f|$ for $\alpha \le \beta$. Therefore

$$\int_0^7 |f(x)|\,\mathrm{d}x = \int_0^5 |f(x)|\,\mathrm{d}x + \int_5^7 |f(x)|\,\mathrm{d}x \ge |3| + |-13| = 16,$$

and the function f_1 shows that this bound is sharp.

5.6.2 Area as a function of its upper limit: a first glimpse

So far the limits of integration have been fixed numbers. Something remarkable happens when we let the upper limit vary.

Example 5.6.7 (The area function of $f(t) = t$)**.** Let $f(t) = t$, and for $b \ge 1$ define

$$F(b) := \int_1^b f(t)\,\mathrm{d}t = \big(\text{area under } f(t) = t \text{ on } [1, b]\big).$$

Label the points $O = (0,0)$, $A = (1,0)$, $C = (1,1)$, $B = (b,0)$ and $D = (b,b)$, as in Figure 5.9. The region under the graph is the large triangle $\triangle OBD$ with the small triangle $\triangle OAC$ removed. Hence

$$F(b) = \text{area}(\triangle OBD) - \text{area}(\triangle OAC) = \frac{1}{2} \cdot b \cdot b - \frac{1}{2} \cdot 1 \cdot 1 = \frac{1}{2}\big(b^2 - 1\big).$$

Now differentiate with respect to the upper limit:

$$\frac{\mathrm{d}}{\mathrm{d}b} F(b) = \frac{\mathrm{d}}{\mathrm{d}b}\Big(\frac{1}{2}\big(b^2 - 1\big)\Big) = \frac{1}{2} \cdot 2b - 0 = b.$$

The name of the variable does not matter, so we may call it t:

$$\frac{\mathrm{d}}{\mathrm{d}t} F(t) = t = f(t).$$

The derivative of the area function is the original function.

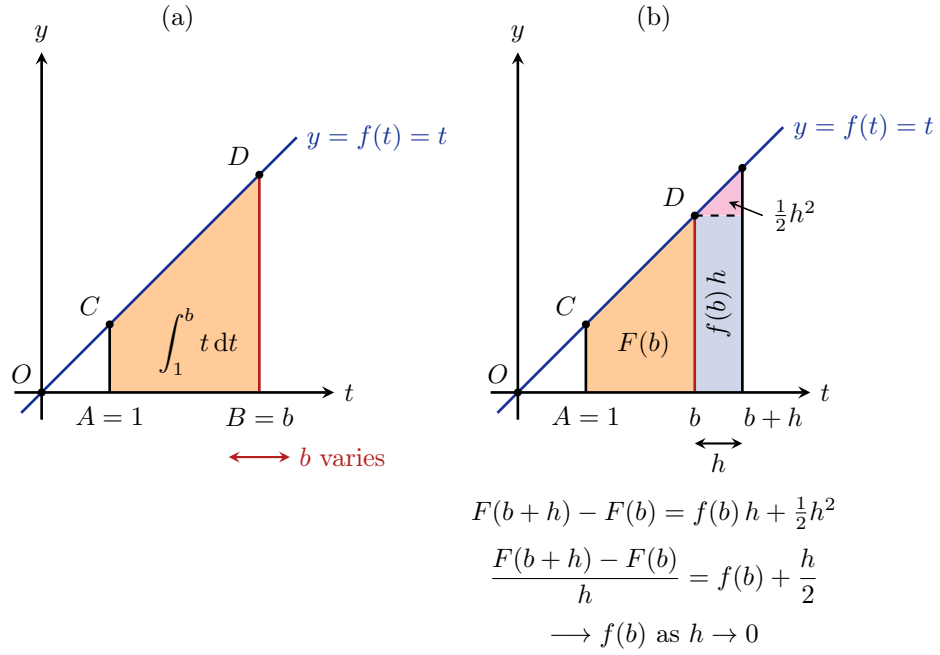


Figure 5.9 – Motivating the Fundamental Theorem of Calculus with $f(t) = t$. (a) The area function $F(b) = \int_1^b t\,\mathrm{d}t = \frac{b^2-1}{2}$ satisfies $F'(b) = b = f(b)$. (b) The extra area $F(b+h) - F(b)$ is a rectangle $f(b)\,h$ plus a triangle $\frac{1}{2}h^2$. After dividing by h, the triangle's share $\frac{h}{2}$ goes to 0.

Intuition. Why the rate of change of area is the height

Push the upper limit from b to $b+h$. The area grows by a thin strip of width h whose height is roughly $f(b)$, so $F(b+h) - F(b) \approx f(b)\,h$. Dividing by h and letting $h \to 0$

suggests that $F'(b) = f(b)$. In Example 5.6.7 we can see the error term exactly:

$$\frac{F(b+h) - F(b)}{h} = \frac{\frac{1}{2}\big((b+h)^2 - b^2\big)}{h} = b + \frac{h}{2}.$$

The extra $\frac{h}{2}$ comes from the little triangle sitting on top of the rectangular strip, and it disappears as $h \to 0$.

Moral.

Integration accumulates. Differentiation measures the rate of accumulation. Example 5.6.7 hints that the two operations undo each other. This is the content of the Fundamental Theorem of Calculus, which turns the difficult limits of Riemann sums into the much easier task of finding antiderivatives.

Story. Two problems, one theorem

For two thousand years, finding areas (quadrature) and finding tangent lines were separate problems with separate tricks. Isaac Barrow, Newton's predecessor in the Lucasian Chair at Cambridge, published a geometric form of their inverse relationship in his *Lectiones Geometricae* (1670). James Gregory had a version in 1668. It was Newton and Leibniz who recognized this link as the engine of a general method, and in doing so invented calculus.

Exercise 5.6.8. (i) Prove Proposition 5.6.2 (ii) by imitating the proof of (i). *Hint:* use Proposition 5.3.4 (i).

(ii) Complete the proof of Corollary 5.6.4 (i): show that if $f(x) \le M$ on $[a, b]$, then $\int_a^b f(x)\,\mathrm{d}x \le M\,(b-a)$.

(iii) Prove Corollary 5.6.3 for the four remaining orderings $b < a < c$, $b < c < a$, $c < a < b$, and $c < b < a$.

(iv) Check the formula $F(b) = \frac{1}{2}(b^2 - 1)$ of Example 5.6.7 directly from the definition. Use the right Riemann sum on a regular partition of $[1, b]$ into n subintervals, together with Theorem 5.3.5 (ii), and let $n \to \infty$.

5.7 The Fundamental Theorem of Calculus

In Example 5.6.7 we computed the area function

$$A(b) = \int_1^b t\,\mathrm{d}t = \frac{1}{2}\big(b^2 - 1\big)$$

by cutting one triangle out of another. Differentiating gave $A'(b) = b = f(b)$, so the derivative of the area function was the function we started with. This was not a coincidence peculiar to straight lines. It happens in general.

Observation. If $F(x)$ is defined as the integral of f from a to x,

$$F(x) \coloneqq \int_a^x f(t)\,\mathrm{d}t,$$

then $F'(x) = f(x)$.

As a slogan: *differentiation and integration are inverses of each other.* This lecture makes the slogan precise, proves it, and shows how it turns the difficult limits of Riemann sums into one-line computations.

5.7.1 Statement

Theorem 5.7.1 (The Fundamental Theorem of Calculus)**.**

(i) *Let f be a continuous function on $[a, b]$. Then f is the derivative of some differentiable function. In fact, for every $x \in [a, b]$,*

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^x f(t)\,\mathrm{d}t\right) = f(x).$$

(ii) *Let F be any function with a continuous derivative on $[a, b]$. Then*

$$F(b) - F(a) = \int_a^b F'(t)\,\mathrm{d}t.$$

Intuition. Two readings of one theorem

Part (i) is about *accumulation.* As the upper limit moves forward by a small amount h, the area grows by a thin strip of height about $f(x)$ and width h. So the rate at which area accumulates is the height of the graph. This is the heuristic from Section 5.6.2. Part (ii) is about *net change.* Take $F = s$ to be the position of a car and $F' = v$ its velocity. Then (ii) reads

$$\int_a^b v(t)\,\mathrm{d}t = s(b) - s(a).$$

In Definition 5.2.2 we wrote this relation as a *definition*, before we knew what an integral was. Now each side has its own meaning. The left side is a limit of Riemann sums, that is, of "velocity $\times$ small time" added up over the trip, as in Section 5.4. The right side is the difference between where you ended and where you started. Part (ii) says that adding up all the small displacements gives the total displacement, and now it is a theorem.

5.7.2 The Mean Value Theorem for integrals

The proof of (i) rests on a lemma that makes the bounds of Corollary 5.6.4 sharp. In Section 5.6.1 we saw that the average value of f lies between its smallest and largest values. If f is continuous, the average is actually *attained.*

Lemma 5.7.2 (The Mean Value Theorem for integrals)**.** *Let $a < b$. If f is a continuous function on $[a, b]$, then there exists $c \in [a, b]$ such that*

$$\int_a^b f(t)\,\mathrm{d}t = f(c)\,(b-a).$$

Proof. We use three facts about a continuous function on a closed, bounded interval. First, it is integrable. Second, by the Extreme Value Theorem, it attains a minimum value $m = f(x_m)$ and a maximum value $M = f(x_M)$ at some points $x_m, x_M \in [a, b]$. Third, by the Intermediate Value Theorem, it takes every value between m and M.

Since $m \le f(t) \le M$ on $[a, b]$, Corollary 5.6.4 (i) gives

$$m\,(b-a) \le \int_a^b f(t)\,\mathrm{d}t \le M\,(b-a).$$

Dividing by $b - a > 0$, we see that the *average value*

$$\mu := \frac{1}{b-a}\int_a^b f(t)\,\mathrm{d}t$$

satisfies $m \le \mu \le M$. By the Intermediate Value Theorem, applied to f on the closed interval with endpoints x_m and x_M, there is a point c between x_m and x_M, hence in $[a, b]$, with $f(c) = \mu$. Multiplying by $b - a$ gives the claim. □

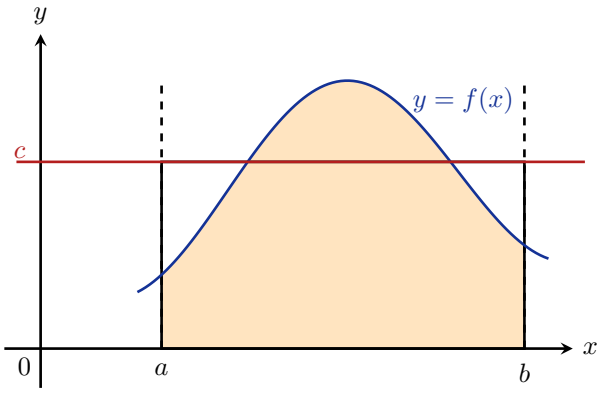


Figure 5.10 – Geometric representation of the Mean Value Theorem for Integrals, where the shaded area under the curve $f(x)$ is equal to the area of the rectangle of height c and width $b - a$.

Geometrically, Lemma 5.7.2 says that we can level off the region under the graph, filling the valleys with the hills, so that it becomes a rectangle of the same area. The height of that rectangle is a value $f(c)$ that the function actually takes (Figure 5.10).

5.7.3 Proof of the Fundamental Theorem

The ideas of the proof can be illustrated in Figure 5.11. Here is how we translate geometry to formal proof:

Proof of Theorem 5.7.1. (i) Define

$$G(x) \coloneqq \int_a^x f(t)\,\mathrm{d}t, \qquad x \in [a,b].$$

Fix $x \in [a,b]$ and let $h \neq 0$ be such that $x+h \in [a,b]$. By Corollary 5.6.3, applied to the points a, x and $x+h$, we have

$$\frac{1}{h}\big(G(x+h)-G(x)\big) = \frac{1}{h}\Big(\int_a^{x+h} f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t\Big) = \frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t.$$

We now apply Lemma 5.7.2 to the small interval between x and $x+h$.

- If $h > 0$, the lemma on $[x, x+h]$ gives a point $c_h \in [x, x+h]$ with $\int_x^{x+h} f(t)\,\mathrm{d}t = f(c_h)\big((x+h)-x\big) = h\,f(c_h)$.
- If $h < 0$, the lemma on $[x+h, x]$, together with Definition 5.6.1 (ii), gives a point $c_h \in [x+h, x]$ with $\int_x^{x+h} f(t)\,\mathrm{d}t = -\int_{x+h}^x f(t)\,\mathrm{d}t = -f(c_h)\big(x-(x+h)\big) = h\,f(c_h)$.

In both cases, c_h lies between x and $x+h$, and

$$\frac{1}{h}\big(G(x+h)-G(x)\big) = \frac{1}{h}\,h\,f(c_h) = f(c_h).$$

Since $|c_h - x| \le |h|$, the squeeze theorem gives $c_h \to x$ as $h \to 0$. Since f is continuous at x, this implies $f(c_h) \to f(x)$. Therefore

$$G'(x) = \lim_{h\to 0}\frac{1}{h}\big(G(x+h)-G(x)\big) = \lim_{h\to 0} f(c_h) = f\Big(\lim_{h\to 0} c_h\Big) = f(x).$$

At the endpoints $x = a$ and $x = b$, only $h > 0$ or only $h < 0$ is allowed, so we obtain the appropriate one-sided derivative. That is exactly what differentiability on the closed interval $[a,b]$ means. We have shown that

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(\int_a^x f(t)\,\mathrm{d}t\Big) = f(x).$$

(ii) Since F' is continuous on $[a,b]$, part (i) applies to F'. So the function

$$F_a(x) \coloneqq \int_a^x F'(t)\,\mathrm{d}t, \qquad a \le x \le b,$$

satisfies $\frac{\mathrm{d}}{\mathrm{d}x}F_a(x) = F'(x)$ on $[a,b]$. Hence the difference $\varphi \coloneqq F_a - F$ satisfies

$$\varphi'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\big(F_a(x)-F(x)\big) = 0 \qquad \text{on } [a,b].$$

A function with zero derivative on an interval is constant. Indeed, for any $x \in (a,b]$, the Mean Value Theorem of differential calculus, applied to φ on $[a,x]$, gives some $\xi \in (a,x)$ with $\varphi(x)-\varphi(a) = \varphi'(\xi)(x-a) = 0$. Thus

$$F_a(x) - F(x) = C \qquad \text{for all } x \in [a,b],$$

where C is a constant. To find C, choose $x = a$. By Definition 5.6.1 (i), $F_a(a) = \int_a^a F'(t)\,\mathrm{d}t = 0$, so

$$C = F_a(a) - F(a) = -F(a).$$

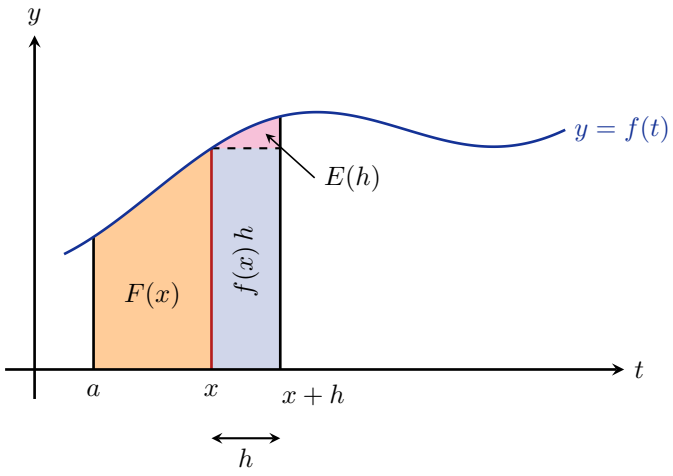
Consequently,

$$F_a(x) - F(x) = -F(a), \qquad \text{that is,} \qquad F_a(x) = F(x) - F(a) \quad \text{on } [a, b].$$

Finally, choose $x = b$. This gives $F_a(b) = F(b) - F(a)$, which says

$$\int_a^b F'(t)\,\mathrm{d}t = F(b) - F(a). \qquad \square$$

> ⚠ *Caveat. Two different Mean Value Theorems*
>
> The proof uses two theorems that are both called "the MVT". Part (i) uses the Mean Value Theorem *for integrals* (Lemma 5.7.2), which produces a point where f equals its average value. Part (ii) uses the Mean Value Theorem of *differential* calculus, which produces a point where the derivative equals the slope of a secant line. Here it is used for the fact that a zero derivative forces a function to be constant. The next remark shows that the two theorems are closely related.



$$F(x+h) - F(x) = \int_x^{x+h} f(t)\,\mathrm{d}t = f(x)\,h + E(h), \qquad E(h) = \int_x^{x+h} \big(f(t) - f(x)\big)\,\mathrm{d}t$$

$$|E(h)| \le h \max_{t\in[x,x+h]} |f(t) - f(x)| \quad \Longrightarrow \quad \frac{F(x+h) - F(x)}{h} = f(x) + \frac{E(h)}{h} \longrightarrow f(x)$$

Figure 5.11 – The key step in the Fundamental Theorem of Calculus. For continuous f and $F(x) = \int_a^x f(t)\,\mathrm{d}t$, the increment $F(x+h) - F(x)$ is a rectangle $f(x)\,h$ plus a cap $E(h)$ with $|E(h)| \le h \max_{[x,x+h]} |f(t) - f(x)|$. By continuity, $E(h)/h \to 0$, so $F'(x) = f(x)$.

⚠ *Caveat. Continuity matters*

Part (i) can fail if f is not continuous. Consider the step function

$$f(t) = \begin{cases} 0, & -1 \le t \le 0, \\ 1, & 0 < t \le 1. \end{cases}$$

It is integrable on $[-1, 1]$, and $G(x) = \int_{-1}^{x} f(t)\,\mathrm{d}t$ equals 0 for $x \le 0$ and x for $x > 0$. So $G(x) = \max\{x, 0\}$, which has a corner at $x = 0$ and is not differentiable there. The area function is still continuous, but it is only as smooth as f allows.

Remark 5.7.3 (Combining the two Mean Value Theorems). Let F have a continuous derivative on $[a, b]$. The Mean Value Theorem of differential calculus gives a point $c \in (a, b)$ with

$$F'(c) = \frac{F(b) - F(a)}{b - a} = \frac{1}{b - a} \int_a^b F'(t)\,\mathrm{d}t,$$

where the second equality is Theorem 5.7.1 (ii). Read from right to left, this is Lemma 5.7.2 for the function $f = F'$. With $F = s$ and $F' = v$, it says that at some instant your velocity equals your average velocity. If you drove 120 miles in 2 hours, then at some moment your speedometer read exactly 60. Compare this with the speedometer estimate in Section 5.6.1.

Moral.

The slogan "differentiation and integration are inverses" has two halves. Part (i): integrate, then differentiate, and you get back f exactly. Part (ii): differentiate, then integrate, and you get back F *up to a constant*, namely $\int_a^x F'(t)\,\mathrm{d}t = F(x) - F(a)$. The constant is lost because differentiation cannot see vertical shifts.

5.7.4 Application: antiderivatives and integral evaluation

Definition 5.7.4 (Antiderivative)**.** Let f be defined on an interval I. A function F is an *antiderivative* of f on I if F is differentiable on I and $F'(x) = f(x)$ for all $x \in I$.

Notation 5.7.5 (Evaluation brackets and indefinite integrals)**.**

- We write
$$F(b) - F(a) =: \big[F(x)\big]_a^b =: F(x)\Big|_a^b.$$
- We use
$$\int f(x)\,\mathrm{d}x$$
to denote the family of all antiderivatives of f, called the *indefinite integral*. On an interval, any two antiderivatives differ by a constant. So if F is one antiderivative of f, we write $\int f(x)\,\mathrm{d}x = F(x) + C$, where C is an arbitrary constant.

The claim that antiderivatives differ by a constant is proved exactly as in part (ii) of the proof above. If $F' = G' = f$ on an interval, then $(F - G)' = 0$ there, so $F - G$ is constant.

⚠ *Caveat. Intervals, and numbers versus families*

Two warnings about this notation.

- The definite integral $\int_a^b f(x)\,\mathrm{d}x$ is a *number*. The indefinite integral $\int f(x)\,\mathrm{d}x$ is a *family of functions*. Despite the shared symbol, they are different kinds of object, and Theorem 5.7.1 is precisely what connects them.
- "Differ by a constant" requires an *interval*. On $\mathbb{R}\setminus\{0\}$, both $\log|x|$ and $\log|x| + \mathbb{1}_{(0,\infty)}(x)$ are antiderivatives of $1/x$, but their difference jumps from 0 to 1 and is not constant. The zero-derivative argument breaks down because there is no interval joining -1 to 1 inside the domain.

In this language, Theorem 5.7.1 (ii) becomes the workhorse of integral calculus.

Corollary 5.7.6 (Evaluating integrals by antiderivatives)**.** *Let f be continuous on $[a,b]$, and let F be any antiderivative of f on $[a,b]$. Then*

$$\int_a^b f(x)\,\mathrm{d}x = \big[F(x)\big]_a^b = F(b) - F(a).$$

Proof. Since $F' = f$ is continuous on $[a,b]$, Theorem 5.7.1 (ii) gives $\int_a^b f(x)\,\mathrm{d}x = \int_a^b F'(x)\,\mathrm{d}x = F(b) - F(a)$. □

The choice of antiderivative does not matter. Replacing F by $F + C$ changes both $F(b)$ and $F(a)$ by C, and the difference is unchanged.

Example 5.7.7 (Triangles no longer needed)**.** Let $f(t) = t$ on $[1,b]$. Find $\displaystyle\int_1^b f(t)\,\mathrm{d}t$.

Let $F(t) = \frac{1}{2}t^2$. Then $F'(t) = \left(\frac{1}{2}t^2\right)' = t = f(t)$. By the Fundamental Theorem of Calculus,

$$\int_1^b f(t)\,\mathrm{d}t = \int_1^b F'(t)\,\mathrm{d}t \overset{\text{FTC}}{=} F(b) - F(1) = \frac{1}{2}b^2 - \frac{1}{2}\cdot 1^2 = \frac{1}{2}\big(b^2 - 1\big).$$

This agrees with Example 5.6.7, but we did not need any geometry.

Example 5.7.8 (Cancellation of signed area)**.** Let $b > 0$ and $f(t) = t$ on $[-b,b]$. With $F(t) = \frac{1}{2}t^2$ as before,

$$\int_{-b}^b f(t)\,\mathrm{d}t = \int_{-b}^b F'(t)\,\mathrm{d}t \overset{\text{FTC}}{=} F(b) - F(-b) = \frac{1}{2}b^2 - \frac{1}{2}(-b)^2 = 0.$$

Geometrically, the triangle below the axis on $[-b,0]$ cancels the congruent triangle above the axis on $[0,b]$. In the language of Section 5.2, take $v(t) = t$ as a velocity. The car moves backward while slowing down, stops at $t = 0$, and then moves forward while speeding up, retracing its path exactly. Its net displacement is zero.

What made this work is symmetry, not the particular function, see Figure 5.12. Recall that a function f on $[-b,b]$ is *function odd* if $f(-t) = -f(t)$, and *function even* if $f(-t) = f(t)$, for all $t \in [-b,b]$.

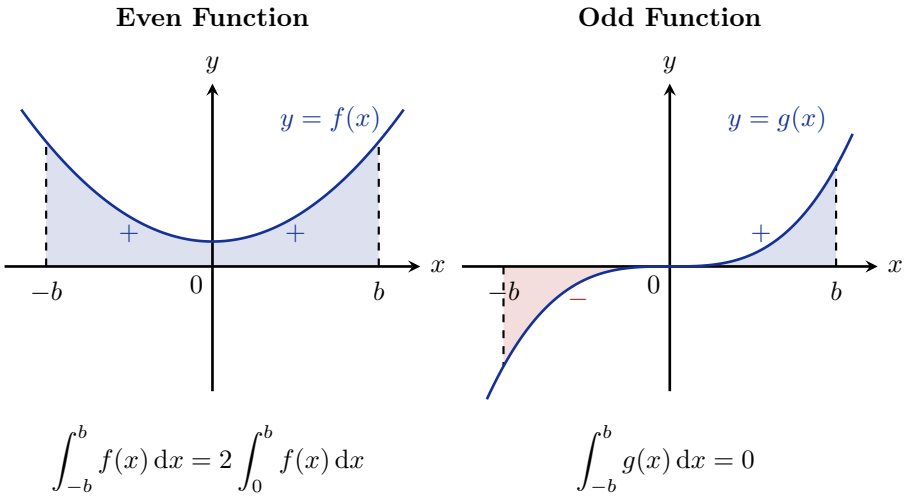


Figure 5.12 – Integrals of symmetric functions over $[-b, b]$. Left: Even function symmetry doubling the area. Right: Odd function symmetry resulting in complete cancellation.

Proposition 5.7.9 (Integrals of odd and even functions)**.** *Let $b > 0$ and let f be continuous on $[-b, b]$.*

(i) *If f is odd, then* $\displaystyle\int_{-b}^{b} f(t)\,\mathrm{d}t = 0.$

(ii) *If f is even, then* $\displaystyle\int_{-b}^{b} f(t)\,\mathrm{d}t = 2\int_{0}^{b} f(t)\,\mathrm{d}t.$

Proof. By Theorem 5.7.1 (i), the function $F(x) \coloneqq \int_{-b}^{x} f(t)\,\mathrm{d}t$ is an antiderivative of f on $[-b, b]$. Let $H(x) \coloneqq F(-x)$. By the chain rule,

$$H'(x) = -F'(-x) = -f(-x), \qquad x \in [-b, b].$$

(i) If f is odd, then $H'(x) = f(x) = F'(x)$, so $(H - F)' = 0$ on $[-b, b]$. As in the proof of Theorem 5.7.1 (ii), $H - F$ is constant. Evaluating at $x = 0$ shows that the constant is $H(0) - F(0) = 0$. Hence $F(-x) = F(x)$, and by Corollary 5.7.6,

$$\int_{-b}^{b} f(t)\,\mathrm{d}t = F(b) - F(-b) = 0.$$

(ii) If f is even, then $H'(x) = -f(x) = -F'(x)$, so $(H + F)' = 0$ on $[-b, b]$. Hence $H + F$ is constant, and evaluating at $x = 0$ shows that the constant is $2F(0)$. Hence $F(-b) = 2F(0) - F(b)$, and by Corollary 5.7.6 (used once on $[-b, b]$ and once on $[0, b]$),

$$\int_{-b}^{b} f(t)\,\mathrm{d}t = F(b) - F(-b) = 2\big(F(b) - F(0)\big) = 2\int_{0}^{b} f(t)\,\mathrm{d}t. \qquad \square$$

Intuition. Symmetry does the work

The graph of an odd function looks the same after a half-turn about the origin. Every patch of positive area on the right therefore has a matching patch of negative area on the left. The graph of an even function is a mirror image of itself across the vertical axis, so the left half contributes exactly as much as the right half. Spotting a symmetry before integrating can save a whole computation.

Caveat. Symmetric interval, integrable function

Both hypotheses in Proposition 5.7.9 are essential. The interval must be symmetric: $\int_{-1}^{2} t\,\mathrm{d}t = \frac{1}{2}(4-1) = \frac{3}{2} \neq 0$. The function must also be integrable. The function $1/t$ is odd, but "$\int_{-1}^{1} \frac{\mathrm{d}t}{t} = 0$" is meaningless, because $1/t$ is unbounded near 0 and is not Riemann integrable on $[-1, 1]$.

Example 5.7.10 (Two routes to the same answer)**.** Let $f(x) = x^2$ on $[-3, 3]$.

- *Directly.* Observe that $F(x) = \frac{1}{3}x^3$ satisfies $F'(x) = \left(\frac{1}{3}x^3\right)' = x^2 = f(x)$. By the Fundamental Theorem of Calculus,
$$\int_{-3}^{3} f(x)\,\mathrm{d}x = \int_{-3}^{3} F'(x)\,\mathrm{d}x = F(3) - F(-3) = \frac{1}{3}(3)^3 - \frac{1}{3}(-3)^3 = 9 + 9 = 18.$$
- *By symmetry.* Since f is even, Proposition 5.7.9 (ii) gives
$$\int_{-3}^{3} f(x)\,\mathrm{d}x \overset{\text{even}}{=} 2\int_{0}^{3} f(x)\,\mathrm{d}x = 2\Big(\frac{1}{3}\,3^3 - \frac{1}{3}\,0^3\Big) = 18.$$

Remark 5.7.11 (What the Fundamental Theorem saves us). The second computation contains the value $\int_0^3 x^2\,\mathrm{d}x = 9$. In Remark 5.4.6 we obtained the same number the hard way. We wrote down right and left Riemann sums on n subintervals, evaluated them with the sum-of-squares formula from Theorem 5.3.5, and let $n \to \infty$. The Fundamental Theorem replaces that whole computation with a single antiderivative, $\frac{1}{3}x^3$. The two answers must agree, and they do: both methods compute the same limit of Riemann sums.

Story. Cauchy's proof

Newton and Leibniz both knew that integration reverses differentiation, and they built calculus on that insight. For them, however, the integral was an intuitive "sum of infinitely many infinitesimal pieces", and the theorem was more a principle than a proof. In his 1823 *Résumé des leçons sur le calcul infinitésimal*, given at the École Polytechnique, Augustin-Louis Cauchy defined the integral of a continuous function as a limit of sums. He then proved part (i) using a mean value property of the integral, essentially the argument given above. With that, a principle passed down for 150 years became a theorem.

Corollary 5.7.6 reduces computing $\int_a^b f(x)\,\mathrm{d}x$ to a single question.

Problem 1 (The antiderivative problem)**.** Given a function f, can we find a function F such that $F'(t) = f(t)$?

Here are 3 partial answers:

- *In principle, yes, if f is continuous.* Theorem 5.7.1 (i) supplies one explicitly: $F(x) = \int_a^x f(t)\,\mathrm{d}t$. For computing $\int_a^b f$, however, this is circular.
- *Not always in closed form.* The function e^{-t^2}, the bell curve of probability theory, is continuous, so it has an antiderivative. Joseph Liouville's work in the 1830s showed that no antiderivative of it can be written in terms of elementary functions (polynomials, roots, exponentials, logarithms, trigonometric functions, and their compositions). This is why statisticians use tables, or software, for the normal distribution.
- *Not at all, for some discontinuous f.* Derivatives satisfy an intermediate value property (Darboux's theorem[2]), even when they are not continuous. So the step function in the caveat "Continuity matters" above has no antiderivative on $[-1, 1]$, although it is integrable.

Finding antiderivatives in closed form, when they exist, is the art of integration that occupies the coming lectures.

Exercise 5.7.12. (i) Use Corollary 5.7.6 to compute $\int_0^1 x^3\,\mathrm{d}x$. Then check your answer by computing the limit of right Riemann sums on regular partitions, using Theorem 5.3.5 (iv).

(ii) Let f be continuous on $\mathbb{R}$. Use Theorem 5.7.1 (i), Definition 5.6.1 and the chain rule to compute

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_x^b f(t)\,\mathrm{d}t \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}x}\int_a^{x^2} f(t)\,\mathrm{d}t.$$

(iii) Find all values of c guaranteed by Lemma 5.7.2 for $f(x) = x^2$ on $[0, 3]$.

(iv) Use Proposition 5.7.9 to evaluate $\int_{-2}^2 \left(x^5 - 3x^3 + x^2\right)\mathrm{d}x$ with as little computation as possible.

A Catalogue of Standard Antiderivatives

Because finding an antiderivative is differentiation performed in reverse, every differentiation rule produces a corresponding indefinite integral formula.

Proposition 5.7.13 (Standard antiderivatives)**.** *On any open interval where the respective integrands are defined, the following indefinite integral identities hold:*

(i) Power rule ($r \neq -1$): *For any real constant $r \neq -1$,*

$$\int x^r\,\mathrm{d}x = \frac{x^{r+1}}{r+1} + C.$$

[2] See `https://en.wikipedia.org/wiki/Darboux%27s_theorem_(analysis)`.

(ii) Reciprocal rule ($r = -1$): *On any interval not containing the origin,*

$$\int \frac{1}{x}\,\mathrm{d}x = \log|x| + C.$$

(iii) Natural exponential: *For any non-zero constant* $k \in \mathbb{R} \setminus \{0\}$*,*

$$\int \mathrm{e}^{kx}\,\mathrm{d}x = \frac{1}{k}\mathrm{e}^{kx} + C.$$

(iv) Trigonometric functions: *For any non-zero constant* $k \in \mathbb{R} \setminus \{0\}$*,*

$$\begin{aligned}\int \cos(kx)\,\mathrm{d}x &= \frac{1}{k}\sin(kx) + C,\\ \int \sin(kx)\,\mathrm{d}x &= -\frac{1}{k}\cos(kx) + C,\\ \int \sec^2(kx)\,\mathrm{d}x &= \frac{1}{k}\tan(kx) + C.\end{aligned}$$

Proof. By Definition 5.7.4, verifying each formula requires only differentiating the right-hand side.

(i) For $r \neq -1$, the power rule for derivatives yields

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{x^{r+1}}{r+1} + C\Big) = \frac{(r+1)x^r}{r+1} + 0 = x^r.$$

(ii) If $x > 0$, then $|x| = x$ and $\frac{\mathrm{d}}{\mathrm{d}x}(\log x + C) = \frac{1}{x}$. If $x < 0$, then $|x| = -x$, and the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}x}\big(\log(-x) + C\big) = \frac{1}{-x}\cdot(-1) = \frac{1}{x}.$$

Thus, $\frac{\mathrm{d}}{\mathrm{d}x}(\log|x| + C) = \frac{1}{x}$ on both $(-\infty, 0)$ and $(0, \infty)$.

(iii) By the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{1}{k}\mathrm{e}^{kx} + C\Big) = \frac{1}{k}\cdot k\,\mathrm{e}^{kx} = \mathrm{e}^{kx}.$$

(iv) Differentiating the trigonometric candidates:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{1}{k}\sin(kx)\Big) &= \frac{1}{k}\cdot k\cos(kx) = \cos(kx),\\ \frac{\mathrm{d}}{\mathrm{d}x}\Big(-\frac{1}{k}\cos(kx)\Big) &= -\frac{1}{k}\cdot\big(-k\sin(kx)\big) = \sin(kx),\\ \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{1}{k}\tan(kx)\Big) &= \frac{1}{k}\cdot k\sec^2(kx) = \sec^2(kx).\end{aligned}$$

□

Intuition. Reading the derivative dictionary backwards

Computing derivatives is an algorithmic procedure: given any combination of elementary functions, mechanical application of the product, quotient, and chain rules always yields a derivative.
Antidifferentiation is inverse pattern recognition. There is no automatic formula for the antiderivative of an arbitrary product or composition. One inspects the integrand to recognize which function's derivative would produce it.

Caveat. The singularity at $r = -1$

The power rule formula $\frac{x^{r+1}}{r+1}$ breaks down algebraically when $r = -1$ due to division by zero. This is not a pathology of calculus, but the exact geometric point where the algebraic powers transition into transcendental functions:

$$\lim_{r \to -1} \frac{x^{r+1} - 1}{r+1} = \log x \qquad \text{for } x > 0.$$

The natural logarithm is the unique antiderivative of x^{-1} that vanishes at $x = 1$.

Example 5.7.14 (Combining linearity with standard formulae)**.** Evaluate the definite integral

$$I = \int_1^4 \left(2\sqrt{x} - \frac{3}{x} + \mathrm{e}^{2x}\right) \mathrm{d}x.$$

Solution. We first determine an antiderivative for each term separately using Proposition 5.7.13:

$$\int 2x^{1/2} \, \mathrm{d}x = 2 \cdot \frac{x^{3/2}}{3/2} = \frac{4}{3}x^{3/2},$$
$$\int \frac{3}{x} \, \mathrm{d}x = 3 \log |x| \,,$$
$$\int \mathrm{e}^{2x} \, \mathrm{d}x = \frac{1}{2}\mathrm{e}^{2x}.$$

By linearity of the integral, an antiderivative of the full integrand on $[1, 4]$ is

$$F(x) = \frac{4}{3}x^{3/2} - 3\log x + \frac{1}{2}\mathrm{e}^{2x}.$$

Applying Corollary 5.7.6 with evaluation brackets:

$$\begin{aligned} I &= \left[\frac{4}{3}x^{3/2} - 3\log x + \frac{1}{2}\mathrm{e}^{2x}\right]_1^4 \\ &= \left(\frac{4}{3}(4)^{3/2} - 3\log 4 + \frac{1}{2}\mathrm{e}^{8}\right) - \left(\frac{4}{3}(1)^{3/2} - 3\log 1 + \frac{1}{2}\mathrm{e}^{2}\right) \\ &= \left(\frac{32}{3} - 3\log 4 + \frac{1}{2}\mathrm{e}^{8}\right) - \left(\frac{4}{3} - 0 + \frac{1}{2}\mathrm{e}^{2}\right) \\ &= \frac{28}{3} - 6\log 2 + \frac{\mathrm{e}^{8} - \mathrm{e}^{2}}{2}. \end{aligned}$$

Moral.

Do not confuse definite and indefinite integrals in multi-step problems. Find an antiderivative $F(x)$ cleanly first, check it by differentiating, and only then insert the boundary limits $[F(x)]_a^b = F(b) - F(a)$.

5.7.5 Application: paying the debt of transcendental functions

> It is the mark of an educated mind to rest unsatisfied with a definition until its logical existence is beyond dispute.
>
> Felix Klein

Intuition.

In our initial exposure to differential calculus, we defined the exponential function e^x through an uneasy compromise. We declared $\mathrm{e} \approx 2.71828\ldots$ as the mysterious base who satisfies the celebrated Euler identity $\mathrm{e}^{i\theta} = \cos\theta + i\sin\theta$, accepted the algebraic exponent laws for arbitrary irrational powers by hand-waving limits, and subsequently defined the natural logarithm $\log x$ as an afterthought by inverting e^x.
To a working mathematician or physicist, this ordering incurs serious technical debt. Where does e truly come from? Why does such a base exist, and why is a^x differentiable for irrational x?
The Fundamental Theorem of Calculus (Theorem 5.7.1) provides the machinery required to pay off this debt in full. By running the developmental sequence *in reverse*—first defining the logarithm as an accumulation of area under the hyperbola $y = 1/t$, establishing its core algebraic properties via differential identities, and then constructing the exponential function as its unique inverse—the entire theory becomes self-contained, transparent, and completely free of circular reasoning.

Differentiating Integrals: The Chain Rule Engine

Before rebuilding our library of transcendental functions, let us review the first part of the Fundamental Theorem of Calculus:

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(\int_a^x f(t)\,\mathrm{d}t\Big) = f(x).$$

In practice, the variable of differentiation rarely sits in the upper limit as a bare, solitary x. It often hides in the lower limit, or governs a non-linear upper bound. Handling these cases requires combining the FTC with oriented interval rules and the chain rule.

Example 5.7.15 (Techniques for differentiating accumulation functions)**.** Simplify each of the following expressions using the Fundamental Theorem of Calculus.

(i) $\dfrac{\mathrm{d}}{\mathrm{d}x}\displaystyle\int_1^x \sin^2 t\,\mathrm{d}t.$

Solution. The integrand $f(t) = \sin^2 t$ is continuous on $\mathbb{R}$, and the variable x appears as the isolated upper limit with a constant lower limit. Applying Part 1 of the Fundamental Theorem directly gives

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_1^x \sin^2 t\,\mathrm{d}t = \sin^2 x.$$

(ii) $\dfrac{\mathrm{d}}{\mathrm{d}x}\displaystyle\int_x^5 \sqrt{t^2+1}\,\mathrm{d}t.$

Solution. To apply the Fundamental Theorem, the variable of differentiation must sit at the upper limit. By the orientation convention (Definition 5.6.1 (ii)), $\int_x^5 f(t)\,\mathrm{d}t = -\int_5^x f(t)\,\mathrm{d}t$. Differentiating then yields

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_x^5 \sqrt{t^2+1}\,\mathrm{d}t = -\frac{\mathrm{d}}{\mathrm{d}x}\int_5^x \sqrt{t^2+1}\,\mathrm{d}t = -\sqrt{x^2+1}.$$

(iii) $\dfrac{\mathrm{d}}{\mathrm{d}x}\displaystyle\int_0^{x^2} \cos(t^2)\,\mathrm{d}t.$

Solution. Here the upper limit is not x, but a non-linear function $u(x) = x^2$. We recognize this expression as a composition of an area function with an inner polynomial map. Define

$$g(u) \coloneqq \int_0^u \cos(t^2)\,\mathrm{d}t, \qquad \text{where } u = x^2.$$

By the chain rule, $\frac{\mathrm{d}}{\mathrm{d}x}g\big(u(x)\big) = \frac{\mathrm{d}g}{\mathrm{d}u}\,\frac{\mathrm{d}u}{\mathrm{d}x}$. Part 1 of the Fundamental Theorem guarantees that

$$\frac{\mathrm{d}g}{\mathrm{d}u} = \frac{\mathrm{d}}{\mathrm{d}u}\int_0^u \cos(t^2)\,\mathrm{d}t = \cos(u^2).$$

Because $\frac{\mathrm{d}u}{\mathrm{d}x} = 2x$, substituting $u = x^2$ yields

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_0^{x^2} \cos(t^2)\,\mathrm{d}t = \big(\cos(u^2)\big)(2x) = 2x\cos(x^4).$$

Remark 5.7.16 (The Leibniz integral rule in one dimension). More generally, if $u(x)$ and $v(x)$ are differentiable functions and f is continuous, splitting the integral at an arbitrary basepoint c gives

$$\int_{v(x)}^{u(x)} f(t)\,\mathrm{d}t = \int_c^{u(x)} f(t)\,\mathrm{d}t - \int_c^{v(x)} f(t)\,\mathrm{d}t.$$

Differentiating with respect to x via the chain rule immediately produces the master formula:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{v(x)}^{u(x)} f(t)\,\mathrm{d}t = f\big(u(x)\big)\,u'(x) - f\big(v(x)\big)\,v'(x).$$

The Natural Logarithm Redefined

Recall the classical power rule for integration: for any integer $n \neq -1$,

$$\int t^n\,\mathrm{d}t = \frac{t^{n+1}}{n+1} + C.$$

The sole exponent where this algebraic formula encounters a division by zero is $n = -1$. Geometrically, however, the curve $y = 1/t$ is smooth, continuous, and strictly positive on $(0, \infty)$; the area under it across any interval $[1, x]$ is well-defined. We turn this singular exception into the definition of the logarithm.

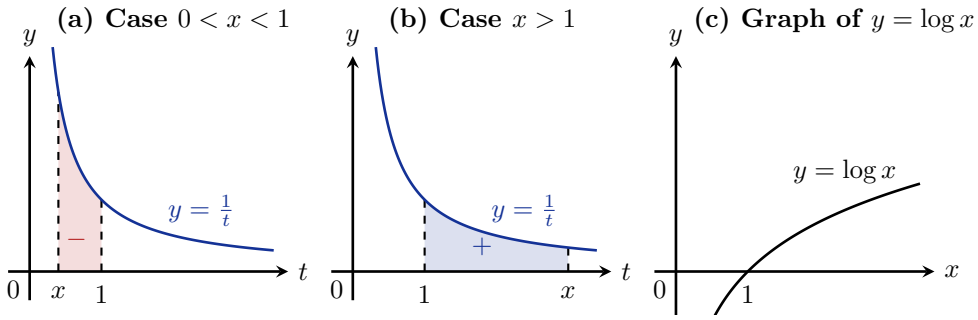


Figure 5.13 – Connecting the integral definition to the graph of $y = \log x$. Panels (a) and (b) show the signed integration area under $y = 1/t$, which directly corresponds to the negative, zero, and positive values along the logarithm curve in Panel (c).

Definition 5.7.17 (The natural logarithm)**.** For any real number $x > 0$, we define the *logarithm natural* by

$$\log x := \int_1^x \frac{1}{t}\,\mathrm{d}t. \tag{5.6}$$

By Part 1 of the Fundamental Theorem of Calculus, because $f(t) = 1/t$ is continuous on $(0, \infty)$, the function $\log x$ is differentiable on $(0, \infty)$ with

$$\frac{\mathrm{d}}{\mathrm{d}x} \log x = \frac{1}{x}. \tag{5.7}$$

Theorem 5.7.18 (Structural properties of the logarithm)**.** *The function* $\log\colon (0, \infty) \to \mathbb{R}$ *satisfies:*

(i) Normalization: $\log 1 = 0$*;*

(ii) Monotonicity: $\log x$ *is strictly increasing on* $(0, \infty)$*;*

(iii) Homomorphism property: *For all* $a, x > 0$*,*

$$\log(ax) = \log a + \log x.$$

Proof. (i) By Definition 5.6.1 (i), an integral over a degenerate interval vanishes:

$$\log 1 = \int_1^1 \frac{1}{t}\,\mathrm{d}t = 0.$$

(ii) By (5.7), $\frac{\mathrm{d}}{\mathrm{d}x} \log x = \frac{1}{x} > 0$ for all $x \in (0, \infty)$. Because its derivative is strictly positive, $\log x$ is strictly increasing.

(iii) Fix a constant $a > 0$, and consider the composite function $\varphi(x) \coloneqq \log(ax)$ defined on $(0, \infty)$. Differentiating via the chain rule gives

$$\varphi'(x) = \frac{\mathrm{d}}{\mathrm{d}x} \log(ax) = \frac{1}{ax} \cdot \frac{\mathrm{d}}{\mathrm{d}x}(ax) = \frac{1}{ax} \cdot a = \frac{1}{x}.$$

Thus, $\varphi(x)$ and $\log x$ share the exact same derivative on the connected interval $(0, \infty)$:

$$\frac{\mathrm{d}}{\mathrm{d}x}\big(\log(ax) - \log x\big) = \frac{1}{x} - \frac{1}{x} = 0.$$

As established in the proof of Theorem 5.7.1 (ii), any function whose derivative vanishes identically on an interval is constant. Therefore, there exists a constant $C \in \mathbb{R}$ such that

$$\log(ax) - \log x = C \qquad \text{for all } x > 0.$$

To determine C, evaluate this identity at $x = 1$:

$$C = \log(a \cdot 1) - \log 1 = \log a - 0 = \log a.$$

Substituting $C = \log a$ yields $\log(ax) = \log a + \log x$. □

Corollary 5.7.19 (Quotients and integer powers)**.** *For any $x, y > 0$ and any integer $n \in \mathbb{Z}$:*

(i) $\log\left(\dfrac{1}{x}\right) = -\log x$;

(ii) $\log\left(\dfrac{x}{y}\right) = \log x - \log y$;

(iii) $\log(x^n) = n \log x$.

Proof. Applying Theorem 5.7.18 (c) to $a = 1/x$ gives $\log\big(x \cdot \frac{1}{x}\big) = \log x + \log\big(\frac{1}{x}\big)$. Since $\log 1 = 0$, we find $\log\big(\frac{1}{x}\big) = -\log x$. Part (ii) follows from $\log(x/y) = \log(x \cdot y^{-1}) = \log x + \log(y^{-1}) = \log x - \log y$. Part (iii) follows by induction for $n \in \mathbb{N}$, and extends to negative integers via (i). □

The Exponential Function and the Number e

To invert $\log x$, we examine its range.

Lemma 5.7.20 (Surjectivity of the logarithm)**.** *The range of* $\log\colon (0, \infty) \to \mathbb{R}$ *is the entire real line:*

$$\lim_{x \to \infty} \log x = \infty \qquad \textit{and} \qquad \lim_{x \to 0^+} \log x = -\infty.$$

Proof. Observe that $\log 2 = \int_1^2 \frac{1}{t}\,\mathrm{d}t$. On the interval $[1, 2]$, we have $\frac{1}{t} \geq \frac{1}{2}$, so by comparison (Corollary 5.6.4 (i)),

$$\log 2 \geq \frac{1}{2}(2 - 1) = \frac{1}{2} > 0.$$

By Corollary 5.7.19 (iii), for any positive integer k, $\log(2^k) = k \log 2 \geq \frac{k}{2}$. Since log is strictly increasing, letting $k \to \infty$ forces $\lim_{x \to \infty} \log x = \infty$. Similarly, $\log(2^{-k}) = -k \log 2 \leq -\frac{k}{2} \to -\infty$ as $k \to \infty$, which implies $\lim_{x \to 0^+} \log x = -\infty$. By the Intermediate Value Theorem, log maps $(0, \infty)$ bijectively onto $(-\infty, \infty)$. □

Because $\log\colon (0,\infty) \to \mathbb{R}$ is strictly increasing and bijective, it admits a well-defined, strictly increasing inverse function.

Definition 5.7.21 (The exponential function and Euler's number)**.** We define the *exponential function*, denoted $\exp\colon \mathbb{R} \to (0,\infty)$, as the unique inverse of the natural logarithm:

$$\exp(x) = y \iff \log y = x.$$

We define the mathematical constant e as the unique real number satisfying

$$\log \mathrm{e} = 1, \qquad \text{or equivalently} \qquad \exp(1) = \mathrm{e}.$$

We write $\mathrm{e}^x \coloneqq \exp(x)$.

Theorem 5.7.22 (Structural properties of the exponential)**.** *The exponential function* $\exp(x) = \mathrm{e}^x$ *satisfies:*

(i) Normalization: $\mathrm{e}^0 = 1$*;*

(ii) Eigen-differential property: $\dfrac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^x) = \mathrm{e}^x$*;*

(iii) Additive group homomorphism: *For all* $a, x \in \mathbb{R}$*,*

$$\mathrm{e}^{a+x} = \mathrm{e}^a \mathrm{e}^x.$$

Proof. (i) Because $\log 1 = 0$ by Theorem 5.7.18 (a), inverting gives $\mathrm{e}^0 = \exp(0) = 1$.

(ii) Since $\log y$ is differentiable with non-vanishing continuous derivative $\frac{\mathrm{d}}{\mathrm{d}y} \log y = \frac{1}{y} \neq 0$ on $(0,\infty)$, the Inverse Function Theorem guarantees that its inverse e^x is differentiable on $\mathbb{R}$. Differentiating the defining identity $\log(\mathrm{e}^x) = x$ via the chain rule yields

$$1 = \frac{\mathrm{d}}{\mathrm{d}x}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\big(\log(\mathrm{e}^x)\big) = \frac{1}{\mathrm{e}^x} \cdot \frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^x).$$

Multiplying both sides by e^x confirms that

$$\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^x) = \mathrm{e}^x.$$

(iii) Fix an arbitrary constant $a \in \mathbb{R}$, and consider the quotient function

$$\psi(x) \coloneqq \frac{\mathrm{e}^{a+x}}{\mathrm{e}^x}.$$

Applying the quotient rule together with the chain rule and part (b):

$$\psi'(x) = \frac{\mathrm{e}^x \cdot \frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{a+x}) - \mathrm{e}^{a+x} \cdot \frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^x)}{(\mathrm{e}^x)^2} = \frac{\mathrm{e}^x(\mathrm{e}^{a+x} \cdot 1) - \mathrm{e}^{a+x}\mathrm{e}^x}{(\mathrm{e}^x)^2} = 0 \qquad \text{for all } x \in \mathbb{R}.$$

Because its derivative vanishes everywhere on the connected interval $\mathbb{R}$, $\psi(x)$ is constant:

$$\frac{\mathrm{e}^{a+x}}{\mathrm{e}^x} = C \qquad \text{for all } x \in \mathbb{R}.$$

To determine the constant C, evaluate at $x = 0$:

$$C = \frac{\mathrm{e}^{a+0}}{\mathrm{e}^0} = \frac{\mathrm{e}^a}{1} = \mathrm{e}^a.$$

Therefore, $\frac{\mathrm{e}^{a+x}}{\mathrm{e}^x} = \mathrm{e}^a$, which rearranges to $\mathrm{e}^{a+x} = \mathrm{e}^a \mathrm{e}^x$. □

Moral.

By defining $\log x \coloneqq \int_1^x \frac{\mathrm{d}t}{t}$, all the familiar algebraic properties of logarithms and exponentials cease to be axiomatic assertions; they become elementary consequences of the Fundamental Theorem of Calculus coupled with the fact that functions with zero derivative are constant.

Trigonometric Functions Redefined Without Triangles

The strategy deployed above is not unique to logarithms. In elementary geometry, the trigonometric functions are defined via right-angled triangles or circular arc lengths. Both geometric approaches rely on unproved spatial intuitions about angles and winding paths.

Using the Fundamental Theorem of Calculus, we can liberate trigonometry entirely from geometric diagrams by defining the inverse sine function as an integral, and subsequently defining the sine function as its inverse.

For $x \in (-1, 1)$, consider the area function

$$F(x) \coloneqq \int_0^x \frac{1}{\sqrt{1-t^2}}\,\mathrm{d}t. \tag{5.8}$$

Because the integrand $f(t) = \frac{1}{\sqrt{1-t^2}}$ is continuous on $(-1, 1)$, Part 1 of the Fundamental Theorem guarantees that F is differentiable on $(-1, 1)$, with

$$F'(x) = \frac{1}{\sqrt{1-x^2}} > 0.$$

Because $F'(x) > 0$, the function F is strictly increasing and therefore invertible.

Definition 5.7.23 (Sine without geometry)**.** Let $p \coloneqq \lim_{x\to 1^-} \int_0^x \frac{1}{\sqrt{1-t^2}}\,\mathrm{d}t$. We define the *sine function*, denoted $\sin t$, on the interval $(-p, p)$ as the unique inverse of the function $F(x)$:

$$t = \int_0^x \frac{1}{\sqrt{1-\tau^2}}\,\mathrm{d}\tau \iff x = \sin t.$$

We define the geometric number π analytically by setting $\pi \coloneqq 2p$.

Differentiating the identity $F(\sin t) = t$ using the chain rule yields

$$F'(\sin t)\cdot(\sin t)' = 1 \implies \frac{1}{\sqrt{1-\sin^2 t}}\cdot(\sin t)' = 1,$$

which proves analytically that $(\sin t)' = \sqrt{1-\sin^2 t} =: \cos t$ without drawing a single triangle!

Exercise 5.7.24. (i) Compute $\dfrac{\mathrm{d}}{\mathrm{d}x}\displaystyle\int_{\sqrt{x}}^{x^3} \mathrm{e}^{-t^2}\,\mathrm{d}t$ for $x > 0$.

(ii) Let $a > 0$ with $a \neq 1$. Define the general power function by $a^x \coloneqq \mathrm{e}^{x\log a}$. Use the chain rule to prove that $\frac{\mathrm{d}}{\mathrm{d}x}(a^x) = (\log a)\,a^x$.

(iii) Let $F(x) = \int_0^x \frac{1}{\sqrt{1-t^2}}\,\mathrm{d}t$. Prove that F is an odd function on $(-1, 1)$ by using a change of variables in the integral, and deduce that $\sin(-t) = -\sin t$.

5.8 The Substitution Rule and Change of Variables

> In mathematics, the art of proposing a question must be held of higher value than solving it; change of perspective is often the very essence of discovery.
>
> Georg Cantor

In differential calculus, the chain rule is the engine that unlocks derivatives of composite functions:

$$\frac{\mathrm{d}}{\mathrm{d}x}F\big(g(x)\big) = F'\big(g(x)\big)\,g'(x).$$

Because integration reverses differentiation, every differentiation rule must have an integral counterpart. Reversing the chain rule yields the *Substitution Rule*, also known as the method of *change of variables*.

Just as the chain rule accounts for the inner rate of change $g'(x)$, the substitution rule accounts for the local stretching or compressing of the coordinate axis when we transition from an old variable x to a new variable $u = g(x)$.

5.8.1 Inverting the Chain Rule: Indefinite Integrals

Let F be any antiderivative of a continuous function f, so that $F' = f$. If g is a differentiable function whose range is contained in the domain of f, the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}x}F\big(g(x)\big) = F'\big(g(x)\big)\,g'(x) = f\big(g(x)\big)\,g'(x).$$

Translating this identity into the language of indefinite integrals (Notation 5.7.5) yields

$$\int f\big(g(x)\big)\,g'(x)\,\mathrm{d}x = F\big(g(x)\big) + C. \tag{5.9}$$

To reveal the structure of the above equation, setting the new variable u to be the inner function,

$$u = g(x),$$

the corresponding differential represents the linear approximation of the change in u:

$$\mathrm{d}u = g'(x)\,\mathrm{d}x.$$

Substituting these formal expressions into (5.9) transforms the integral into

$$\int \underbrace{f\big(g(x)\big)}_{f(u)}\,\underbrace{g'(x)\,\mathrm{d}x}_{\mathrm{d}u} = \int f(u)\,\mathrm{d}u = F(u) + C. \tag{5.10}$$

The method replaces an intricate integral with respect to x by a simpler integral with respect to u. Once an antiderivative $F(u)$ is found, replacing u by $g(x)$ restores the final result in terms of the original variable x.

Intuition. Differentials as bookkeeping devices

Writing $\mathrm{d}u = g'(x)\,\mathrm{d}x$ is not merely formal shorthand; it accurately reflects the distortion of scale between coordinate systems. If $g'(x) = 5$, a tiny step of length $\mathrm{d}x$ on the x-axis produces a change $\mathrm{d}u$ that is five times as large on the u-axis. The factor $g'(x)$ inside the integrand compensates for this coordinate re-scaling.

Moral.

To execute a successful substitution, the integrand must contain not only a composite expression $f\big(g(x)\big)$, but also the derivative of the inner function $g'(x)$ as an algebraic factor, at least up to a non-zero multiplicative constant.

To evaluate an indefinite integral using the Substitution Rule:

1. *Identify the inner target:* Inspect the integrand for a composite structure $f\big(g(x)\big)$ where a constant multiple of $g'(x)$ appears elsewhere in the product. Set $u = g(x)$.
2. *Compute the differential:* Differentiate to find $\mathrm{d}u = g'(x)\,\mathrm{d}x$, and isolate the remaining x-terms in terms of $\mathrm{d}u$.
3. *Transform and integrate:* Re-express the entire integrand in terms of u and $\mathrm{d}u$, ensuring that no remnants of the old variable x survive. Compute the resulting integral $\int f(u)\,\mathrm{d}u$.
4. *Back-substitute:* Express the final antiderivative in terms of x by substituting $g(x)$ back in place of u.

Caveat. Incomplete substitutions

A common error is to substitute $u = g(x)$ while leaving some appearances of x inside the integrand. An intermediate expression such as $\int x^2 u^9\,\mathrm{d}u$ is completely invalid: an integral cannot be evaluated with respect to u while treating remaining functions of x as constants. Every single x and $\mathrm{d}x$ must be transformed into u and $\mathrm{d}u$ before proceeding.

Example 5.8.1 (Perfect substitutions)**.** Evaluate the following indefinite integrals and verify each result by differentiation.

(i) $\displaystyle\int 2(2x+1)^3\,\mathrm{d}x.$

Solution. The composite power $(2x+1)^3$ has inner function $u = 2x+1$. Differentiating yields $\mathrm{d}u = 2\,\mathrm{d}x$, which appears as a factor in the integrand. Substituting directly:

$$\begin{aligned}\int \underbrace{(2x+1)^3}_{u^3}\cdot\underbrace{2\,\mathrm{d}x}_{\mathrm{d}u} &= \int u^3\,\mathrm{d}u\\ &= \frac{u^4}{4}+C\\ &= \frac{(2x+1)^4}{4}+C.\end{aligned}$$

Check: $\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{4}(2x+1)^4 + C\right) = \frac{1}{4}\cdot 4(2x+1)^3 \cdot 2 = 2(2x+1)^3$.

(ii) $\displaystyle\int 10\mathrm{e}^{10x}\,\mathrm{d}x$.

Solution. The exponential e^{10x} has inner function $u = 10x$, giving $\mathrm{d}u = 10\,\mathrm{d}x$. Thus,

$$\begin{aligned}\int \underbrace{\mathrm{e}^{10x}}_{\mathrm{e}^u}\cdot\underbrace{10\,\mathrm{d}x}_{\mathrm{d}u} &= \int \mathrm{e}^u\,\mathrm{d}u\\ &= \mathrm{e}^u + C\\ &= \mathrm{e}^{10x} + C.\end{aligned}$$

Check: $\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{10x} + C) = \mathrm{e}^{10x}\cdot 10 = 10\mathrm{e}^{10x}$.

Example 5.8.2 (Introducing constant factors)**.** Find the following indefinite integrals.

(i) $\displaystyle\int x^4(x^5+6)^9\,\mathrm{d}x$.

Solution. The inner function of $(x^5+6)^9$ is $u = x^5+6$, with derivative $\frac{\mathrm{d}u}{\mathrm{d}x} = 5x^4$. The factor x^4 is present in the integrand, lacking only the scalar factor 5. We write

$$\mathrm{d}u = 5x^4\,\mathrm{d}x \implies x^4\,\mathrm{d}x = \frac{1}{5}\,\mathrm{d}u.$$

Applying the substitution rule and factoring out the constant scalar:

$$\begin{aligned}\int \underbrace{(x^5+6)^9}_{u^9}\cdot\underbrace{x^4\,\mathrm{d}x}_{\frac{1}{5}\,\mathrm{d}u} &= \int u^9\cdot\frac{1}{5}\,\mathrm{d}u\\ &= \frac{1}{5}\int u^9\,\mathrm{d}u\\ &= \frac{1}{5}\cdot\frac{u^{10}}{10} + C\\ &= \frac{1}{50}(x^5+6)^{10} + C.\end{aligned}$$

(ii) $\displaystyle\int \cos^3 x\sin x\,\mathrm{d}x$.

Solution. Rewriting the integrand as $(\cos x)^3\sin x$, we identify $u = \cos x$. Then $\mathrm{d}u = -\sin x\,\mathrm{d}x$, which implies $\sin x\,\mathrm{d}x = -\,\mathrm{d}u$. Substituting gives

$$\begin{aligned}\int \underbrace{\cos^3 x}_{u^3}\cdot\underbrace{\sin x\,\mathrm{d}x}_{-\,\mathrm{d}u} &= -\int u^3\,\mathrm{d}u\\ &= -\frac{u^4}{4} + C\\ &= -\frac{\cos^4 x}{4} + C.\end{aligned}$$

Moral. Heuristics for selecting u

Choosing the appropriate substitution often requires trial and error, guided by two primary heuristics:

- If an expression $g(x)$ appears in the integrand and a constant multiple of its derivative $g'(x)$ also appears as an algebraic factor multiplying the differential, test $u = g(x)$.
- If a complicated expression is trapped inside an algebraic root, an exponential, a trigonometric argument, or a denominator, test that inner expression as u.

5.8.2 Substitution for Definite Integrals

When evaluating a definite integral $\int_a^b f\big(g(x)\big)g'(x)\,\mathrm{d}x$, there are two distinct conceptual approaches:

(I) *Antiderivative first, limits second:* Use substitution to find the indefinite antiderivative $F\big(g(x)\big)$, and then apply the Fundamental Theorem of Calculus directly to the original limits:
$$\big[F\big(g(x)\big)\big]_a^b = F\big(g(b)\big) - F\big(g(a)\big).$$

(II) *Transforming the limits simultaneously:* Change the limits of integration from x-values to u-values at the exact moment the substitution is executed:
$$x = a \implies u = g(a), \qquad x = b \implies u = g(b).$$
Once the limits are updated, one evaluates the new u-integral directly, with no need to back-substitute into x.

The second approach is more streamlined and significantly less prone to sign errors.

Theorem 5.8.3 (Substitution Rule for Definite Integrals)**.** *Let g be a function whose derivative g' is continuous on $[a,b]$, and let f be continuous on the range of g. Then*
$$\int_a^b f\big(g(x)\big)\,g'(x)\,\mathrm{d}x = \int_{g(a)}^{g(b)} f(u)\,\mathrm{d}u. \tag{5.11}$$

Proof. Let F be an antiderivative of f on the range of g, so that $F' = f$. By the chain rule,
$$\frac{\mathrm{d}}{\mathrm{d}x}F\big(g(x)\big) = F'\big(g(x)\big)\,g'(x) = f\big(g(x)\big)\,g'(x).$$
Since f and g' are continuous, $f\big(g(x)\big)g'(x)$ is continuous on $[a,b]$. By Part 2 of the Fundamental Theorem of Calculus (Theorem 5.7.1 (ii)), the left-hand side evaluates to
$$\int_a^b f\big(g(x)\big)\,g'(x)\,\mathrm{d}x = \big[F\big(g(x)\big)\big]_a^b = F\big(g(b)\big) - F\big(g(a)\big).$$
Now consider the right-hand side of (5.11). Since $F' = f$, Part 2 of the Fundamental Theorem applies directly with respect to the variable u:
$$\int_{g(a)}^{g(b)} f(u)\,\mathrm{d}u = \big[F(u)\big]_{g(a)}^{g(b)} = F\big(g(b)\big) - F\big(g(a)\big).$$

Both sides evaluate to $F\big(g(b)\big) - F\big(g(a)\big)$, establishing the equality. □

Example 5.8.4 (Definite integral evaluations)**.** Evaluate the following definite integrals.

(i) $\displaystyle\int_0^2 \frac{\mathrm{d}x}{(x+3)^3}$.

Solution. Set $u = x + 3$, so that $\mathrm{d}u = \mathrm{d}x$. We transform the limits of integration:

$$\text{when } x = 0, \quad u = 0 + 3 = 3; \qquad \text{when } x = 2, \quad u = 2 + 3 = 5.$$

The integral becomes

$$\begin{aligned}\int_0^2 \frac{\mathrm{d}x}{(x+3)^3} &= \int_3^5 u^{-3}\,\mathrm{d}u \\ &= \Big[-\frac{1}{2}u^{-2}\Big]_3^5 \\ &= -\frac{1}{2}\Big(\frac{1}{5^2} - \frac{1}{3^2}\Big) \\ &= -\frac{1}{2}\Big(\frac{1}{25} - \frac{1}{9}\Big) \\ &= -\frac{1}{2}\Big(-\frac{16}{225}\Big) \\ &= \frac{8}{225}.\end{aligned}$$

(ii) $\displaystyle\int_2^3 \frac{x^2}{x^3 - 7}\,\mathrm{d}x$.

Solution. Set $u = x^3 - 7$, so $\mathrm{d}u = 3x^2\,\mathrm{d}x$, which gives $x^2\,\mathrm{d}x = \frac{1}{3}\,\mathrm{d}u$. Updating the limits:

$$\text{when } x = 2, \quad u = 2^3 - 7 = 1; \qquad \text{when } x = 3, \quad u = 3^3 - 7 = 20.$$

Applying the reciprocal antiderivative formula from Proposition 5.7.13 (ii):

$$\begin{aligned}\int_2^3 \frac{x^2}{x^3 - 7}\,\mathrm{d}x &= \frac{1}{3}\int_1^{20} \frac{1}{u}\,\mathrm{d}u \\ &= \frac{1}{3}\big[\log|u|\big]_1^{20} \\ &= \frac{1}{3}(\log 20 - \log 1) \\ &= \frac{1}{3}\log 20.\end{aligned}$$

(iii) $\displaystyle\int_0^{\pi/2} \sin^4 x \cos x\,\mathrm{d}x$.

Solution. Set $u = \sin x$, so $\mathrm{d}u = \cos x\,\mathrm{d}x$. Updating the limits:

$$\text{when } x = 0, \quad u = \sin 0 = 0; \qquad \text{when } x = \frac{\pi}{2}, \quad u = \sin\Big(\frac{\pi}{2}\Big) = 1.$$

Evaluating the polynomial integral in u:

$$\begin{aligned}\int_0^{\pi/2} \sin^4 x \cos x \, \mathrm{d}x &= \int_0^1 u^4 \, \mathrm{d}u \\ &= \left[\frac{u^5}{5}\right]_0^1 \\ &= \frac{1}{5} - 0 = \frac{1}{5}.\end{aligned}$$

5.8.3 Geometric Meaning of the Change of Variables

The analytical proof of Theorem 5.8.3 via Part 2 of the Fundamental Theorem is concise, but it obscures the underlying geometric mechanism. To see what change of variables actually does to area, we examine Riemann sums directly.

Suppose that the coordinate map $u = g(t)$ has a strictly positive derivative $g'(t) > 0$ on $[a, b]$, so that g is strictly increasing. It maps the interval $[a, b]$ on the t-axis bijectively onto the interval $[A, B]$ on the u-axis, where $A = g(a)$ and $B = g(b)$.

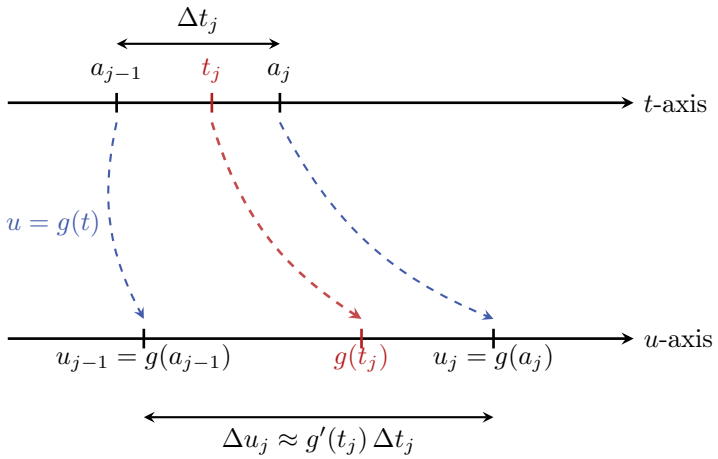


Figure 5.14 – The geometric meaning of the substitution rule: the transformation $u = g(t)$ maps the interval $[a_{j-1}, a_j]$ to $[u_{j-1}, u_j]$, scaling its width locally by the derivative factor $g'(t_j)$.

Subdivide $[a, b]$ into n subintervals via partition points

$$a = a_0 < a_1 < a_2 < \cdots < a_n = b.$$

Because g is strictly increasing, applying g to these points creates an induced partition of $[A, B]$:

$$A = u_0 < u_1 < u_2 < \cdots < u_n = B, \qquad \text{where } u_j \coloneqq g(a_j).$$

By the Mean Value Theorem of differential calculus applied to g on $[a_{j-1}, a_j]$, there exists an intermediate point $t_j \in (a_{j-1}, a_j)$ such that

$$u_j - u_{j-1} = g(a_j) - g(a_{j-1}) = g'(t_j)(a_j - a_{j-1}). \tag{5.12}$$

Equation (5.12) shows that the derivative $g'(t_j)$ acts as a local *stretching factor* between the two coordinate axes (Figure 5.14).

Now compare the approximating Riemann sums for the integrals on both sides of (5.11):

- For the integral $\int_A^B f(u)\,\mathrm{d}u$ on the u-axis, we evaluate f at the tagged sample points $g(t_j) \in [u_{j-1}, u_j]$:
$$\sum_{j=1}^{n} f\big(g(t_j)\big)\,(u_j - u_{j-1}).$$
- For the integral $\int_a^b f\big(g(t)\big)g'(t)\,\mathrm{d}t$ on the t-axis, we evaluate the composite integrand at the tags t_j:
$$\sum_{j=1}^{n} f\big(g(t_j)\big)\,g'(t_j)\,(a_j - a_{j-1}).$$

Substituting the stretching relation (5.12) reveals that these two Riemann sums are *algebraically identical term by term*:

$$\sum_{j=1}^{n} f\big(g(t_j)\big)\,(u_j - u_{j-1}) = \sum_{j=1}^{n} f\big(g(t_j)\big)\,g'(t_j)\,(a_j - a_{j-1}).$$

Because g is uniformly continuous on $[a, b]$, refining the partition on the t-axis simultaneously refines the partition on the u-axis. Passing to the limit as the mesh norm tends to zero establishes that

$$\int_A^B f(u)\,\mathrm{d}u = \int_a^b f\big(g(t)\big)g'(t)\,\mathrm{d}t.$$

> *Intuition. Stretching strips*
>
> A rectangular strip on the t-axis of height $f\big(g(t)\big)$ and width Δt corresponds to a strip on the u-axis of height $f(u)$ and width $\Delta u \approx g'(t)\,\Delta t$. The factor $g'(t)$ adjusts the base width of each approximating rectangle so that the accumulated area is preserved under the coordinate distortion.

The geometric utility of Theorem 5.8.3 becomes especially clear when g is an affine linear map:

$$g(t) = kt + c,$$

where $k \neq 0$ and c are real constants. Differentiating gives $g'(t) = k$. Applying the Substitution Rule:

$$\int_{ka+c}^{kb+c} f(x)\,\mathrm{d}x = \int_a^b f(kt + c)\,k\,\mathrm{d}t = k\int_a^b f(kt + c)\,\mathrm{d}t.$$

This single identity captures two fundamental geometric invariances of the integral:

(i) *Translation invariance ($k = 1$):* Setting $k = 1$ gives $g(t) = t + c$ and $g'(t) = 1$, which yields

$$\int_{a+c}^{b+c} f(x)\,\mathrm{d}x = \int_a^b f(t + c)\,\mathrm{d}t. \tag{5.13}$$

Shifting a function horizontally by c units and shifting the integration limits by the same amount leaves the accumulated area unchanged.

(ii) *Dilation and scaling ($c = 0$):* Setting $c = 0$ gives $g(t) = kt$ and $g'(t) = k$, which yields

$$\int_{ka}^{kb} f(x)\,\mathrm{d}x = k\int_{a}^{b} f(kt)\,\mathrm{d}t. \tag{5.14}$$

Compressing the domain horizontally by a factor of k must be balanced by multiplying the integral by k to preserve the total area.

Exercise 5.8.5. (i) Evaluate $\displaystyle\int \frac{\mathrm{e}^{\sqrt{x}}}{\sqrt{x}}\,\mathrm{d}x$.

(ii) Evaluate $\displaystyle\int_0^1 x^2\sqrt{1-x^3}\,\mathrm{d}x$ by updating the limits of integration.

(iii) Use the linear substitution $u = a + b - x$ to prove the reflection identity:

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

(iv) Apply the identity from part (c) to evaluate $\displaystyle\int_0^{\pi/2} \frac{\sin x}{\sin x + \cos x}\,\mathrm{d}x$ without finding an explicit antiderivative.

Solution to Exercise 5.8.5. (i) *Indefinite integral with an algebraic radical inside an exponential.*

We evaluate

$$\int \frac{\mathrm{e}^{\sqrt{x}}}{\sqrt{x}}\,\mathrm{d}x \qquad \text{for } x > 0.$$

The power $\sqrt{x} = x^{1/2}$ appears in the exponent. Choosing $u = \sqrt{x}$, its differential is

$$\mathrm{d}u = \frac{1}{2\sqrt{x}}\,\mathrm{d}x \implies \frac{\mathrm{d}x}{\sqrt{x}} = 2\,\mathrm{d}u.$$

Substituting into the integral:

$$\begin{aligned}\int \frac{\mathrm{e}^{\sqrt{x}}}{\sqrt{x}}\,\mathrm{d}x &= \int \underbrace{\mathrm{e}^{\sqrt{x}}}_{\mathrm{e}^u}\cdot\underbrace{\frac{\mathrm{d}x}{\sqrt{x}}}_{2\,\mathrm{d}u}\\ &= 2\int \mathrm{e}^u\,\mathrm{d}u\\ &= 2\mathrm{e}^u + C\\ &= 2\mathrm{e}^{\sqrt{x}} + C.\end{aligned}$$

Check: Differentiating the result via the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(2\mathrm{e}^{\sqrt{x}} + C\right) = 2\mathrm{e}^{\sqrt{x}}\cdot\frac{1}{2\sqrt{x}} = \frac{\mathrm{e}^{\sqrt{x}}}{\sqrt{x}},$$

confirming the computation.

(ii) *Definite integral via transformation of limits.*

We evaluate

$$\int_0^1 x^2\sqrt{1-x^3}\,\mathrm{d}x.$$

The quantity under the square root is $g(x) = 1 - x^3$. Setting $u = 1 - x^3$, the differential is

$$\mathrm{d}u = -3x^2\,\mathrm{d}x \implies x^2\,\mathrm{d}x = -\frac{1}{3}\,\mathrm{d}u.$$

We transform the integration limits directly from x to u:

$$\text{when } x = 0, \quad u = 1 - 0^3 = 1; \qquad \text{when } x = 1, \quad u = 1 - 1^3 = 0.$$

By the Substitution Rule for definite integrals (Theorem 5.8.3) and the orientation convention (Definition 5.6.1 (ii)):

$$\begin{aligned}\int_0^1 x^2\sqrt{1-x^3}\,\mathrm{d}x &= \int_1^0 \sqrt{u}\left(-\frac{1}{3}\,\mathrm{d}u\right)\\ &= \frac{1}{3}\int_0^1 u^{1/2}\,\mathrm{d}u\\ &= \frac{1}{3}\left[\frac{2}{3}u^{3/2}\right]_0^1\\ &= \frac{2}{9}(1^{3/2} - 0^{3/2})\\ &= \frac{2}{9}.\end{aligned}$$

(iii) *Proof of the reflection identity.*

We establish that for any integrable function f on $[a, b]$,

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

Consider the right-hand side. We apply the affine linear substitution

$$u = g(x) \coloneqq a + b - x.$$

The derivative is $g'(x) = -1$, yielding the differential relation

$$\mathrm{d}u = -\,\mathrm{d}x \implies \mathrm{d}x = -\,\mathrm{d}u.$$

Next, we map the boundary endpoints:

$$\text{when } x = a, \quad u = a + b - a = b; \qquad \text{when } x = b, \quad u = a + b - b = a.$$

Applying Theorem 5.8.3:

$$\begin{aligned}\int_a^b f(a+b-x)\,\mathrm{d}x &= \int_b^a f(u)\,(-\,\mathrm{d}u)\\ &= -\int_b^a f(u)\,\mathrm{d}u\\ &= \int_a^b f(u)\,\mathrm{d}u.\end{aligned}$$

Because u is a dummy variable of integration (Notation 5.3.1), renaming u back to x yields

$$\int_a^b f(a+b-x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x,$$

as claimed.

(iv) *Evaluation without an explicit antiderivative.*

We compute

$$I := \int_0^{\pi/2} \frac{\sin x}{\sin x + \cos x}\,\mathrm{d}x.$$

Here $a = 0$ and $b = \frac{\pi}{2}$, so $a + b - x = \frac{\pi}{2} - x$. By the co-function trigonometric identities:

$$\sin\Big(\frac{\pi}{2} - x\Big) = \cos x, \qquad \cos\Big(\frac{\pi}{2} - x\Big) = \sin x.$$

Applying the reflection identity proved in part (c) to $f(x) = \frac{\sin x}{\sin x + \cos x}$:

$$\begin{aligned} I &= \int_0^{\pi/2} \frac{\sin\big(\frac{\pi}{2} - x\big)}{\sin\big(\frac{\pi}{2} - x\big) + \cos\big(\frac{\pi}{2} - x\big)}\,\mathrm{d}x \\ &= \int_0^{\pi/2} \frac{\cos x}{\cos x + \sin x}\,\mathrm{d}x. \end{aligned}$$

Adding these two expressions for I together eliminates the non-trivial trigonometric denominator entirely:

$$\begin{aligned} 2I &= \int_0^{\pi/2} \frac{\sin x}{\sin x + \cos x}\,\mathrm{d}x + \int_0^{\pi/2} \frac{\cos x}{\sin x + \cos x}\,\mathrm{d}x \\ &= \int_0^{\pi/2} \frac{\sin x + \cos x}{\sin x + \cos x}\,\mathrm{d}x \\ &= \int_0^{\pi/2} 1\,\mathrm{d}x \\ &= \frac{\pi}{2} - 0 = \frac{\pi}{2}. \end{aligned}$$

Dividing by 2 produces the exact value:

$$I = \frac{\pi}{4}. \qquad \square$$

> *Intuition. The power of symmetry over primitives*
>
> Finding an antiderivative for $\frac{\sin x}{\sin x + \cos x}$ using classical methods requires either tangent half-angle substitutions or clever algebraic decomposition into symmetric and antisymmetric pieces ($\frac{1}{2}\frac{(\sin x + \cos x) + (\sin x - \cos x)}{\sin x + \cos x}$).
> The reflection substitution $u = a + b - x$ bypasses antidifferentiation entirely. By reflecting the integrand across the midpoint of the interval $x = \pi/4$, it demonstrates that the sine and cosine components contribute equally to the total area, allowing us

to evaluate the integral strictly through functional symmetry.

5.9 Bonus: Darboux Sums, Compactness, and Integrability

In Section 5.4, we defined the definite integral as a limit of general Riemann sums. While tagged Riemann sums are indispensable for physical modeling and numerical quadrature, their arbitrary choice of sample tags $x_k^* \in [x_{k-1}, x_k]$ creates substantial technical friction when establishing theorems.

To overcome this, Gaston Darboux (1875) introduced an elegant alternative: rather than allowing sample points to float unpredictably, one evaluates the extreme behavior on each subinterval. By taking the infimum and supremum of the function over each slice, we construct the sharpest possible under- and over-estimates for the area. The exact integral is trapped between these bounds. In this section, we develop this framework from first principles and deploy the fundamental property of real numbers—the *Heine–Borel theorem*—to prove that every continuous function on a closed, bounded interval is integrable.

5.9.1 Axiomatic Properties and Construction of Darboux Sums

We begin by isolating the fundamental geometric properties that any coherent theory of integration must obey.

Hypothesis 5.9.1 (Axiomatic properties of the integral)**.** For any continuous function $f\colon [a,b] \to \mathbb{R}$, the integral operator satisfies:

(i) *Interval additivity:* If $a < c < b$, then

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t. \tag{5.15}$$

(ii) *Order and bounding property:* If $m \le f(t) \le M$ for all $t \in [a,b]$, then

$$m\,(b-a) \le \int_a^b f(t)\,\mathrm{d}t \le M\,(b-a). \tag{5.16}$$

Definition 5.9.2 (Lower and upper Darboux sums)**.** Let $\mathcal{P} = \{a_0, a_1, \dots, a_n\}$ be a partition of $[a,b]$:

$$a = a_0 < a_1 < a_2 < \cdots < a_{n-1} < a_n = b.$$

Because f is continuous on each compact subinterval $[a_{i-1}, a_i]$, the Extreme Value Theorem guarantees that f attains an absolute minimum m_i and an absolute maximum M_i:

$$m_i \coloneqq \min_{t\in[a_{i-1},a_i]} f(t), \qquad M_i \coloneqq \max_{t\in[a_{i-1},a_i]} f(t).$$

We define the *lower Darboux sum* and *upper Darboux sum* of f with respect to $\mathcal{P}$ by

$$L(f,\mathcal{P}) \coloneqq \sum_{i=1}^{n} m_i(a_i - a_{i-1}), \qquad U(f,\mathcal{P}) \coloneqq \sum_{i=1}^{n} M_i(a_i - a_{i-1}).$$

Applying the interval additivity property (5.15) across all n subintervals yields

$$\int_a^b f(t)\,\mathrm{d}t = \sum_{i=1}^{n} \int_{a_{i-1}}^{a_i} f(t)\,\mathrm{d}t.$$

On each subinterval $[a_{i-1}, a_i]$, the bounding property (5.16) dictates that

$$m_i(a_i - a_{i-1}) \le \int_{a_{i-1}}^{a_i} f(t)\,\mathrm{d}t \le M_i(a_i - a_{i-1}).$$

Summing these bounds over all $i = 1, \ldots, n$ yields the squeeze inequality:

$$L(f, \mathcal{P}) \le \int_a^b f(t)\,\mathrm{d}t \le U(f, \mathcal{P}). \tag{5.17}$$

> *Intuition. Darboux sums as architectural bounding boxes*
>
> Every tagged Riemann sum $\sum_{i=1}^{n} f(t_i^*)\,\Delta a_i$ satisfies $m_i \le f(t_i^*) \le M_i$ for each strip, meaning that
>
> $$L(f, \mathcal{P}) \le R(f, \mathcal{P}, \{t_i^*\}) \le U(f, \mathcal{P}).$$
>
> The lower Darboux sum is the largest possible floor constructed from under-approximating rectangular pillars, while the upper Darboux sum is the lowest ceiling constructed from over-approximating pillars. Proving integrability reduces to showing that as we refine the partition, the ceiling and the floor converge to the exact same value.

5.9.2 The Engine of Global Regularity: Heine–Borel–Cantor

This section should be skipped during the first reading, readers only need to understand the first paragraph and the statement of Theorem 5.9.5, then move on to Section 5.9.3.

To force the gap $U(f, \mathcal{P}) - L(f, \mathcal{P})$ to zero, we need the local variations $M_i - m_i$ to be simultaneously small across *all* subintervals. The local definition of continuity states that for each point $x \in [a, b]$ and any $\varepsilon > 0$, there exists a $\delta(x) > 0$ such that

$$|t - x| < \delta(x) \implies |f(t) - f(x)| < \frac{\varepsilon}{2}.$$

However, the modulus of continuity $\delta(x)$ can shrink arbitrarily close to zero as x approaches a steep gradient, preventing us from picking a single mesh size that controls the entire interval at once.

To bridge this chasm between local continuity and global control, we invoke the topology of the real numbers.

Definition 5.9.3 (Open cover and compactness)**.** Let $K \subset \mathbb{R}$. A collection of open sets $\mathcal{U} = \{V_\alpha\}_{\alpha \in \mathcal{A}}$ is an *cover open* of K if $K \subset \bigcup_{\alpha \in \mathcal{A}} V_\alpha$. We say that K is *compact* if every open cover of K has a *finite subcover*: there exist indices $\alpha_1, \ldots, \alpha_N \in \mathcal{A}$ such that

$$K \subset V_{\alpha_1} \cup V_{\alpha_2} \cup \cdots \cup V_{\alpha_N}.$$

Theorem 5.9.4 (Heine–Borel)**.** *A subset $K \subset \mathbb{R}$ is compact if and only if it is closed and bounded. In particular, every interval of the form $[a, b]$ is compact.*

Proof. We prove that $[a,b]$ is compact via Dedekind completeness (the least upper bound property). Let $\mathcal{U} = \{V_\alpha\}_{\alpha\in\mathcal{A}}$ be an open cover of $[a,b]$. Consider the set of points up to which $[a,b]$ can be covered by finitely many members of $\mathcal{U}$:

$$S \coloneqq \big\{x \in [a,b] : [a,x] \text{ is covered by a finite subcollection of } \mathcal{U}\big\}.$$

Step 1. S is non-empty and bounded above. Clearly $a \in S$ because $a \in [a,b] \subset \bigcup_\alpha V_\alpha$, so $a \in V_{\alpha_0}$ for some α_0, covering the singleton $[a,a]$. Furthermore, $S \subset [a,b]$ is bounded above by b. By the completeness of $\mathbb{R}$, the supremum $s \coloneqq \sup S$ exists and satisfies $a \le s \le b$.

Step 2. Show that $s \in S$. Since $s \in [a,b]$, there is some open set $V_{\alpha^*} \in \mathcal{U}$ containing s. Because V_{α^*} is open, there exists $r > 0$ such that $(s-r, s+r) \subset V_{\alpha^*}$. By definition of the supremum, there exists $x_1 \in S$ with $s - r < x_1 \le s$. Because $x_1 \in S$, the interval $[a,x_1]$ is covered by finitely many sets $V_{\alpha_1}, \dots, V_{\alpha_k}$. Adjoining V_{α^*} to this collection gives a finite cover of $[a,s] \subset [a,x_1] \cup V_{\alpha^*}$. Thus $s \in S$.

Step 3. Show that $s = b$. Suppose for contradiction that $s < b$. Since $(s-r, s+r) \subset V_{\alpha^*}$, we may select a point x_2 such that $s < x_2 < \min\{b, s+r\}$. Then $[a,x_2] \subset [a,s] \cup V_{\alpha^*}$, meaning $[a,x_2]$ is also covered by our finite collection. Therefore $x_2 \in S$, directly contradicting that s is an upper bound of S. Hence $s = b$, and $[a,b]$ is covered by finitely many sets from $\mathcal{U}$. □

> ✠ *Story. The emergence of compactness*
>
> Eduard Heine introduced the property of uniform continuity in 1872, implicitly using interval-covering arguments. Émile Borel formally formulated the finite-covering theorem in his 1895 doctoral thesis for countable covers. The formulation for arbitrary open covers was established by Pierre Cousin and Henri Lebesgue. The Heine–Borel theorem transformed real analysis by demonstrating that compactness is the rigorous topological generalization of finiteness: it allows one to trade an infinite continuum of local guarantees for a single, uniform global guarantee.

Theorem 5.9.5 (Heine–Cantor)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be continuous on a compact interval $[a,b]$. Then f is* continuous uniformly*: for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$|t - s| < \delta \implies |f(t) - f(s)| < \varepsilon \qquad \text{for all } t, s \in [a,b].$$

Proof. Let $\varepsilon > 0$. By ordinary continuity, for every $x \in [a,b]$, there exists $\delta_x > 0$ such that

$$|t - x| < \delta_x \implies |f(t) - f(x)| < \frac{\varepsilon}{2}.$$

For each $x \in [a,b]$, construct the open interval $I_x \coloneqq \big(x - \frac{\delta_x}{2}, x + \frac{\delta_x}{2}\big)$. The family $\{I_x\}_{x\in[a,b]}$ forms an open cover of $[a,b]$. By Theorem 5.9.4, there exists a finite subcover:

$$[a,b] \subset I_{x_1} \cup I_{x_2} \cup \cdots \cup I_{x_N}.$$

Define the uniform modulus of continuity as the minimum of the finitely many half-radii:

$$\delta \coloneqq \min\Big\{\frac{\delta_{x_1}}{2}, \frac{\delta_{x_2}}{2}, \dots, \frac{\delta_{x_N}}{2}\Big\} > 0.$$

Now, let $t, s \in [a,b]$ with $|t-s| < \delta$. Because the sets I_{x_k} cover $[a,b]$, the point t must belong to some I_{x_j}, which means $|t-x_j| < \frac{\delta_{x_j}}{2}$. By the triangle inequality:

$$|s-x_j| \le |s-t| + |t-x_j| < \delta + \frac{\delta_{x_j}}{2} \le \frac{\delta_{x_j}}{2} + \frac{\delta_{x_j}}{2} = \delta_{x_j}.$$

Both t and s lie within δ_{x_j} of the center point x_j. Applying the triangle inequality to f:

$$|f(t)-f(s)| \le |f(t)-f(x_j)| + |f(x_j)-f(s)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \qquad \square$$

> *Moral. From local to global*
>
> It is the intuition of compactness that allows us to pass from the microscopic, local continuity of a function to its macroscopic, global control.

5.9.3 Convergence via Uniform Continuity

With Theorem 5.9.5 established, we answer the central foundational question: can the definite integral $\int_a^b f(t)\,\mathrm{d}t$ be determined to arbitrary precision solely by refining the partition?

Theorem 5.9.6 (Integrability of continuous functions)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be continuous on the compact interval $[a,b]$. Then for every precision tolerance $\varepsilon > 0$, there exists a partition $\mathcal{P}$ of $[a,b]$ such that the discrepancy between the upper and lower sums satisfies*

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon\,(b-a).$$

Consequently, the lower and upper Darboux approximations squeeze the value of $\int_a^b f(t)\,\mathrm{d}t$ to within an error tolerance strictly bounded by $\varepsilon\,(b-a)$.

Proof. Let $\varepsilon > 0$. By the Heine–Cantor theorem (Theorem 5.9.5), f is uniformly continuous on $[a,b]$. Thus, there exists $\delta > 0$ such that

$$|t-s| < \delta \implies |f(t)-f(s)| < \varepsilon \qquad \text{for all } t, s \in [a,b].$$

Choose a partition $\mathcal{P} = \{a_0, a_1, \dots, a_n\}$ of $[a,b]$ whose mesh size satisfies $\|\mathcal{P}\| = \max_{1\le i\le n}(a_i - a_{i-1}) < \delta$.

On each subinterval $[a_{i-1}, a_i]$, the Extreme Value Theorem guarantees that f attains its minimum and maximum at points $t_{\min}, t_{\max} \in [a_{i-1}, a_i]$. Since $|t_{\max} - t_{\min}| \le a_i - a_{i-1} < \delta$, uniform continuity ensures that

$$M_i - m_i = f(t_{\max}) - f(t_{\min}) < \varepsilon.$$

We now evaluate the total discrepancy between the upper and lower Darboux sums:

$$\begin{aligned} U(f,\mathcal{P}) - L(f,\mathcal{P}) &= \sum_{i=1}^{n} M_i(a_i - a_{i-1}) - \sum_{i=1}^{n} m_i(a_i - a_{i-1}) \\ &= \sum_{i=1}^{n} (M_i - m_i)(a_i - a_{i-1}) \\ &< \sum_{i=1}^{n} \varepsilon(a_i - a_{i-1}) \\ &= \varepsilon \sum_{i=1}^{n} (a_i - a_{i-1}) \\ &= \varepsilon\,(b-a). \end{aligned}$$

Because $b-a$ is a fixed positive constant, choosing $\varepsilon = \frac{\varepsilon_0}{b-a}$ makes $U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon_0$ arbitrarily small (e.g., smaller than 10^{-3} or 10^{-6}), ensuring that the upper and lower Darboux bounds converge uniquely to $\int_a^b f(t)\,\mathrm{d}t$. □

> *Moral.*
>
> Uniform continuity levels the playing field: it ensures that no subinterval fluctuates faster than any other. Because the total width of all subintervals telescopes strictly to $b-a$, bounding the local oscillation by ε automatically bounds the global error by $\varepsilon\,(b-a)$.

Exercise 5.9.7 (Classes of integrable functions)**.** Prove that if a function $f\colon [a,b] \to \mathbb{R}$ is bounded on $[a,b]$ and possesses at most finitely many points of discontinuity, then f is Riemann integrable on $[a,b]$.

Hint: Let $M = \sup_{[a,b]} |f(t)|$. If f has a single discontinuity at $c \in (a,b)$, isolate c in a tiny subinterval $[c-\eta, c+\eta]$ of total width 2η such that its contribution to $U(f,\mathcal{P}) - L(f,\mathcal{P})$ is at most $2M(2\eta) < \frac{\varepsilon}{2}$. On the remaining compact sets $[a, c-\eta]$ and $[c+\eta, b]$, the function is continuous; apply Theorem 5.9.6 to control their contributions by $\frac{\varepsilon}{2}$.

Appendix A

Midterm 1 Review

A.1 Practice Problems

Chapter 1: Functions

1. Determine whether the statement "If $f(x)$ is an even function, then $cf(ax)$ is an even function, where a and c are nonzero real numbers" is true and give an explanation or counterexample.

2. Simplify the difference quotient $\frac{f(x+h)-f(x)}{h}$ for $f(x) = \frac{x}{x+1}$.

3. Simplify the difference quotient $\frac{f(x)-f(a)}{x-a}$ for $f(x) = x^3 - 2x$.

4. Determine whether the graph of $f(x) = 3x^5 + 2x^3 - x$ is symmetric about the x-axis, the y-axis, or the origin.

5. Determine whether the graph of $f(x) = x\,|x|$ is symmetric about the x-axis, the y-axis, or the origin.

6. Find a polynomial f that satisfies $(f(x))^2 = x^4 - 12x^2 + 36$.

7. Let E be an even function and O be an odd function. Determine the symmetry, if any, of $E + O$.

8. Let E be an even function and O be an odd function. Determine the symmetry, if any, of $E \cdot O$.

9. Determine whether the statement "If f is a linear polynomial, then $f \circ f$ is a quadratic polynomial" is true and give an explanation or counterexample.

10. Find the point(s) of intersection of the parabola $y = x^2 + 2$ and the line $y = x + 4$.

11. The entire graph of f consists of the upper half of a circle of radius 2 centered at the origin and the lower half of a circle of radius 3 centered at $(5, 0)$. Find a piecewise function for f.

12. For $b > 0$ with $b \neq 1$, what are the domain and range of $f(x) = \log_b x$ and why?

13. Find the inverse $f^{-1}(x)$ of the function $f(x) = \mathrm{e}^{2x+6}$.

14. Assume that $b > 0$ and $b \neq 1$. Show that $\log_{1/b} x = -\log_b x$.

15. Draw a right triangle to simplify $\cos(2\sin^{-1} x)$ (Hint: Use $\cos 2u = \cos^2 u - \sin^2 u$).

Chapter 2: Limits and Continuity

1. Describe a process for finding the slope of the line tangent to the graph of f at $(a, f(a))$.

2. For the function $f(x) = x^3 - x$, make a table of slopes of secant lines and make a conjecture about the slope of the tangent line at $x = 1$.

3. Determine whether the statement "The value of $\lim_{x\to 3} \frac{x^2-9}{x-3}$ does not exist" is true and give an explanation or counterexample.

4. Determine whether the statement "The value of $\lim_{x\to a} f(x)$ is always found by computing $f(a)$" is true and give an explanation or counterexample.

5. Determine whether the statement "$\lim_{x\to 0} \sqrt{x} = 0$" is true and give an explanation or counterexample.

6. Determine whether the statement "$\lim_{x\to \pi/2} \cot x = 0$" is true and give an explanation or counterexample.

7. Let $f(x) = \frac{|x|}{x}$, for $x \neq 0$. Does $\lim_{x\to 0} f(x)$ exist? Explain your reasoning after first examining $\lim_{x\to 0^-} f(x)$ and $\lim_{x\to 0^+} f(x)$.

8. For a given integer a, state the values of $\lim_{x\to a^-} \lfloor x \rfloor$ and $\lim_{x\to a^+} \lfloor x \rfloor$.

9. For what values of a does $\lim_{x\to a} \lfloor x \rfloor$ exist? Explain?

10. A function f is even if $f(-x) = f(x)$. Suppose f is even, with $\lim_{x\to 2^+} f(x) = 5$ and $\lim_{x\to 2^-} f(x) = 8$. Evaluate $\lim_{x\to -2^+} f(x)$.

11. Evaluate $\lim_{x\to -2^-} f(x)$ for the function described in problem 10.

12. Assume $\lim_{x\to 1} f(x) = 8$ and $\lim_{x\to 1} g(x) = 3$. Compute $\lim_{x\to 1} \sqrt{f(x)g(x) + 3}$ and state the limit laws used to justify your computation.

13. Evaluate $\lim_{x\to -b} \frac{(x+b)^7+(x+b)^{10}}{4(x+b)}$.

14. Evaluate $\lim_{x\to -1} \frac{(2x-1)^2-9}{x+1}$.

15. Evaluate $\lim_{x\to 9} \frac{\sqrt{x}-3}{x-9}$.

16. Evaluate $\lim_{w\to 1} \left(\frac{1}{w^2-w} - \frac{1}{w-1}\right)$.

17. Evaluate $\lim_{x\to 2} \left(\frac{1}{x-2} - \frac{2}{x^2-2x}\right)$.

18. Evaluate $\lim_{x\to 0} \frac{x}{\sqrt{cx+1}-1}$, for $c \neq 0$.

19. Evaluate $\lim_{x\to 0} \frac{\sin 2x}{\sin x}$.

20. Determine whether the statement "If $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = L$, then $f(a) = g(a)$" is true and give an explanation or counterexample.

21. For $f(x) = \begin{cases} x^2+1 & \text{if } x < -1 \\ x+1 & \text{if } x \geq -1 \end{cases}$, compute $\lim_{x\to -1} f(x)$ or state that it does not exist.

22. Explain why $\lim_{x\to 2^-} \sqrt{x-2}$ does not exist.

23. Suppose a spaceship of length L_0 travels at a high speed v relative to an observer. To the observer, the ship appears to have a smaller length given by the Lorentz contraction formula $L = L_0\sqrt{1-\frac{v^2}{c^2}}$. Find $\lim_{v\to c^-} L_0\sqrt{1-\frac{v^2}{c^2}}$ and explain the significance of this limit.

24. Show that $\lim_{x\to 0} x\sin\frac{1}{x} = 0$.

25. Suppose $f(x) = \begin{cases} \frac{x^2-5x+6}{x-3} & \text{if } x \neq 3 \\ a & \text{if } x = 3 \end{cases}$. Determine a value of the constant a for which $\lim_{x\to 3} f(x) = f(3)$.

26. Suppose $g(x) = \begin{cases} x^2-5x & \text{if } x \leq -1 \\ ax^3-7 & \text{if } x > -1 \end{cases}$. Determine a value of the constant a for which $\lim_{x\to -1} g(x)$ exists and state the value of the limit, if possible.

27. Calculate $\lim_{x\to a} \frac{x^5-a^5}{x-a}$ using the factorization formula $x^n - a^n = (x-a)(x^{n-1} + ax^{n-2} + \cdots + a^{n-1})$.

28. Consider the function $F(x) = f(x)/g(x)$ with $g(a) = 0$. Does F necessarily have a vertical asymptote at $x = a$? Explain your reasoning.

29. Verify that the function $f(x) = \frac{x^2-4x+3}{x^2-3x+2}$ is undefined at $x = 1$ and at $x = 2$. Does the graph of f have vertical asymptotes at both these values of x? Explain.

30. Determine $\lim_{z\to 3} \frac{(z-1)(z-2)}{z-3}$.

31. Find all vertical asymptotes $x = a$ of the function $f(x) = \frac{x^2-9x+14}{x^2-5x+6}$. For each value of a, determine $\lim_{x\to a^+} f(x)$, $\lim_{x\to a^-} f(x)$, and $\lim_{x\to a} f(x)$.

32. Determine whether the statement "The line $x = 1$ is a vertical asymptote of the function $f(x) = \frac{x^2-7x+6}{x^2-1}$" is true and give an explanation or counterexample.

33. Determine whether the statement "The line $x = -1$ is a vertical asymptote of the function $f(x) = \frac{x^2-7x+6}{x^2-1}$" is true and give an explanation or counterexample.

34. Suppose the function g satisfies the inequality $3 - \frac{1}{x^2} \leq g(x) \leq 3 + \frac{1}{x^2}$, for all nonzero values of x. Evaluate $\lim_{x\to\infty} g(x)$ and $\lim_{x\to-\infty} g(x)$.

35. Determine $\lim_{x\to-\infty} \frac{\sqrt{40x^4+x^2}+5x^2}{\sqrt{64x^8+x^6}}$.

36. Determine $\lim_{x\to\infty} \frac{40x^5+x^2}{16x^4-2x}$ and $\lim_{x\to-\infty} \frac{40x^5+x^2}{16x^4-2x}$.

37. Determine $\lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$ for $f(x) = \frac{4x^3+1}{2x^3+2\sqrt{16x^6+1}}$.

38. Determine $\lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$ for $f(x) = \frac{\sqrt{x^6+8}}{4x^2+\sqrt{x^4+1}}$.

39. Complete the following steps for $f(x) = \frac{4x^3+4x^2+7x+4}{x^2+1}$:
(a) Find the slant asymptote of f.
(b) Find the vertical asymptotes of f (if any).
(c) Sketch a graph of the function by hand.

40. Determine whether the statement "$\lim_{x\to\infty}(x^3 - x^2) = \lim_{x\to\infty} x^3 - \lim_{x\to\infty} x^2 = \infty - \infty = 0$" is true and give an explanation or counterexample.

41. The population of a colony of squirrels is given by $p(t) = \frac{1500}{3+2e^{-0.1t}}$. Determine whether a steady state exists and give the steady-state value.

42. Suppose $f(x) = p(x)/q(x)$ is a rational function, where $p(x) = a_m x^m + \cdots + a_0$ and $q(x) = b_n x^n + \cdots + b_0$, with $a_m \neq 0$ and $b_n \neq 0$. Prove that if $m < n$, then $\lim_{x\to\{\infty\}} f(x) = 0$.

43. Use the following instructions to determine the end behavior of $f(x) = \frac{4e^x+2e^{2x}}{8e^x+e^{2x}}$.
(a) Evaluate $\lim_{x\to\infty} f(x)$ by first dividing the numerator and denominator by e^{2x}.
(b) Evaluate $\lim_{x\to-\infty} f(x)$ by first dividing the numerator and denominator by e^x.
(c) Give the horizontal asymptote(s).
(d) Graph f to confirm your work in parts (a)–(c).

44. Find the horizontal asymptotes of $f(x) = \frac{3e^{5x}+7e^{6x}}{9e^{5x}+14e^{6x}}$ using limits at infinity.

45. Complete the following steps for $f(x) = \begin{cases} 2x & \text{if } x < 1 \\ x^2 + 3x & \text{if } x \geq 1 \end{cases}$ at $a = 1$:
(a) Use the continuity checklist to show that f is not continuous at a.
(b) Determine whether f is continuous from the left or the right at a.
(c) State the interval(s) of continuity.

46. Determine the interval(s) on which $f(t) = (t^2 - 1)^{3/2}$ is continuous. At which finite endpoints of the intervals of continuity is f continuous from the left or continuous from the right?

47. Evaluate $\lim_{x\to0} \frac{e^{4x}-1}{e^x-1}$.

48. Evaluate $\lim_{x\to e^2} \frac{\log^2 x - 5\log x + 6}{\log x - 2}$.

49. Let $g(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases}$.
(a) Write a formula for $|g(x)|$.
(b) Is g continuous at $x = 0$? Explain.
(c) Is $|g|$ continuous at $x = 0$? Explain.
(d) For any function f, if $|f|$ is continuous at a, does it necessarily follow that f is continuous at a? Explain.

50. Prove that if g is continuous at a and f is continuous at $g(a)$, then the composition $f \circ g$ is continuous at a. (Hint: Write the definition of continuity for f and g separately; then combine them to form the definition of continuity for $f \circ g$.)

A.2 Bonus Problems

Chapter 1: Functions

1. Prove that if a parabola crosses the x-axis twice, the x-coordinate of the vertex of the parabola is half-way between the x-intercepts.

2. The quartic polynomial $y = f(x) = x^4 - x^2$.

(a) Graph f and find the largest intervals on which it is one-to-one.

(b) Make the substitution $u = x^2$ to solve the equation $y = f(x)$ for x in terms of y. Be sure you have included all possible solutions.

(c) Write each inverse function in the form $y = f^{-1}(x)$ for each of the intervals found in part (a).

Chapter 2: Limits and Continuity

1. Evaluate the following limit: $\lim_{x\to 0} \frac{\cos x - 1}{\cos^2 x - 1}$.

2. Evaluate the following one-sided limit: $\lim_{w\to 3^-} \frac{|w-3|}{w^2-7w+12}$.

3. Suppose $g(x) = f(1-x)$ for all x, $\lim_{x\to 1^+} f(x) = 4$, and $\lim_{x\to 1^-} f(x) = 6$. Find $\lim_{x\to 0^+} g(x)$ and $\lim_{x\to 0^-} g(x)$.

4. Find all vertical asymptotes $x = a$ of the function $f(x) = \frac{x+1}{x^3-4x^2+4x}$. For each value of a, determine $\lim_{x\to a^+} f(x)$, $\lim_{x\to a^-} f(x)$, and $\lim_{x\to a} f(x)$.

5. Determine $\lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$ for $f(x) = x - \sqrt{2x^2 - 9x}$.

6. (The Monk and the Mountain) A monk set out from a monastery in the valley at dawn. He walked all day up a winding path, stopping for lunch and taking a nap along the way. At dusk, he arrived at a temple on the mountaintop. The next day the monk made the return walk to the valley, leaving the temple at dawn, walking the same path for the entire day, and arriving at the monastery in the evening. Must there be one point along the path that the monk occupied at the same time of day on both the ascent and the descent? Explain.

Appendix B

Midterm 2 Review

B.1 Practice Problems

1. Find the line tangent to the graph of $f(x) = x^2$ at an arbitrary point a. Where does the tangent line intersect the x-axis? the y-axis?

2. The linear approximation to f at $a = 2$ is $y(x) = 5(x - 2) + 6$. Find $f(2)$ and $f'(2)$.

3. For each function f given below, find $f'(a)$ by forming the difference quotient $\frac{f(a+h)-f(a)}{h}$, and taking the limit as h tends to 0. Then find the line tangent to the graph of f at $x = a$.

(a) $f(x) = \sqrt{x}$, $a = 4$.

(b) $f(x) = mx^2 + kx$, $a = 2$.

(c) $f(x) = x^3$, $a = -1$.

4. Find a line that is tangent to the graphs of both functions $f(x) = x^2$ and $g(x) = x^2 - 2x$.

5. Find the point of intersection of the tangents to the graphs of $f_1(x) = x^2 - 2x$ and $f_2(x) = -x^2 + 1$ at the points $(2, 0)$ and $(1, 0)$, respectively.

6. Which of the functions $f(x) = 5x$, $g(x) = x^2$ would you say is more sensitive to change near $x = 3$?

7. Find all tangents to the graph of $f(x) = x^2 - x$ that go through the point $(2, 1)$. Verify that no tangent goes through the point $(2, 3)$. Can you find a geometric explanation for this?

8. Let $g(x) = x^2 \sin\left(\frac{\pi}{x}\right)$, $g(0) = 0$.

(a) Sketch the graph of g on $-1 \leq x \leq 1$.

(b) Show that g is continuous at 0.

(c) Show that g is differentiable at 0 and find $g'(0)$.

9. Show that a function that is not continuous at a cannot be differentiable at a.

10. Let $f(x) = x^{2/3}$. Does the one-sided derivative $f'_+(0)$ exist? Is f differentiable on $[0, 1]$?

11. For the function $f(x) = (x+4)^{2/3}$, find the x-values at which f is differentiable.

12. For the function $f(x) = \frac{x^2}{x^2-4}$, find the x-values at which f is differentiable.

13. For the function

$$f(x) = \begin{cases} 4 - x^2 & \text{if } x < 0 \\ \log(|x-2|+1) & \text{if } x \geq 0 \end{cases}$$

find the x-values at which f is differentiable.

14. For the function $f(x) = |x-5|$, find the x-values at which f is differentiable.

15. For the function $f(x) = |x-1|$, find the derivatives from the left and from the right at $x = 1$ (if they exist). Is the function differentiable at $x = 5$?

16. For the function

$$f(x) = \begin{cases} (x-1)^3, & x \leq 1 \\ (x-1)^2, & x > 1 \end{cases}$$

find the derivatives from the left and from the right at $x = 1$ (if they exist). Is the function differentiable at $x = 1$?

17. For the function $f(x) = (1-x)^{2/3}$, find the derivatives from the left and from the right at $x = 1$, determine whether the function is differentiable at $x = 1$.

18. Find a and b such that

$$f(x) = \begin{cases} ax^3, & x \leq 2 \\ x^2 + b, & x > 2 \end{cases}$$

is differentiable everywhere.

19. Find the derivatives of the following functions:

(a) $x^5 - 3x^4 + 0.5x^2 - 17$

(b) $\frac{x+1}{x-1}$

(c) $\frac{x^2-1}{x^2-2x+1}$

(d) $\frac{\sqrt{x}}{x+1}$

(e) $\sqrt{x^3+1}$

(f) $\left(\frac{x+1}{x}\right)^3$

(g) $\sqrt{1+\sqrt{x}}$

(h) $(\sqrt{x}+1)(\sqrt{x}-1)$

20. Calculate $\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1-x^2}{1+x^2}\right)$ by the quotient rule and $\frac{\mathrm{d}}{\mathrm{d}x}\left((1-x^2)(1+x^2)^{-1}\right)$ by the product rule and show that the results agree.

21. Find the first six derivatives $f', \ldots, f^{(6)}$ for each of these functions. In which cases will the seventh and higher derivatives be identically 0?

(a) $f(x) = x^3$

(b) $f(t) = t^3 + 5t^2$

(c) $f(r) = r^6$

(d) $f(x) = x^{-1}$

(e) $f(t) = t^{-3} + t^3$

(f) $f(r) = 6 + r + r^8$

22. Derive a formula for the derivative of a triple product by applying the product rule twice: show that if f, g, and h are differentiable at x, then fgh is too, and $(fgh)'(x) = f(x)g(x)h'(x) + f(x)g'(x)h(x) + f'(x)g(x)h(x)$.

23. Use the chain rule to find the derivatives indicated.

(a) $f(t) = 1 + t + t^2$. Find $(f(t)^2)'$ without squaring f.

(b) If $g'(3) = 0$, find $(g(t)^6)'$ when $t = 3$.

(c) If $h'(3) = 4$ and $h(3) = 5$, find $(h(t)^6)'$ when $t = 3$.

24. Find the derivatives of the following functions:

(a) $(\cos x)'$

(b) $(\cos(2x))'$

(c) $(\cos^2 x)'$

25. Find the derivatives of the following functions:

(a) $\log|\sin x|$

(b) $\mathrm{e}^{\tan^{-1} x}$

(c) $\tan^{-1}(5x^2)$

(d) $\log(\mathrm{e}^{2x} + 1)$

(e) $\mathrm{e}^{\sin x}$

(f) $\log|\tan x|$

(g) $\tan^{-1}(\mathrm{e}^{2x})$

(h) $(\mathrm{e}^{kx})^2$

(i) $1 + \sin^2 x$

(j) $1 - \cos^2 x$

26. Where are the functions $f_1(x) = |\sin x|$ and $f_2(x) = \sin|x|$ differentiable?

27. For the function $f(x) = \frac{x}{x-1}$, find $f^{(8)}(x)$.

28. Given $f^{(4)}(t) = t\cos t$, find $f^{(5)}(t)$.

29. Given $f(x) = \frac{g(x)}{h(x)}$, and $g(2) = 3$, $g'(2) = -2$, $h(2) = -1$, $h'(2) = 4$, find $f'(2)$.

30. Describe how to differentiate a piecewise function. Use your approach to find the first and second derivatives of $f(x) = x|x|$. Explain why $f''(0)$ does not exist.

31. For the function $f(x) = \frac{\cos x}{x^n}$, find the derivatives of the function f for $n = 1, 2, 3$, and 4. Use the results to write a general rule for $f'(x)$ in terms of n.

32. Let $y = 2x^3 - 6x + 10$, verify that the function satisfies the differential equation $-y''' - xy'' - 2y' = -24x^2$.

33. Given $g(5) = -1$, $g'(5) = 1$, $h(5) = 5$, $h'(5) = 2$, find $f'(5)$ for the following cases

(a) $f(x) = g(x)h(x)$.

(b) $f(x) = g(h(x))$.

(c) $f(x) = \frac{g(x)}{h(x)}$.

(d) $f(x) = h(x)\mathrm{e}^{g(x)}$.

(e) $f(x) = (g(x) + h(x))^4$.

34. Compute the first and second derivatives of:

(a) e^x

(b) e^{3x}

(c) e^{-3x}

(d) e^{-x^2}

(e) $\mathrm{e}^{-1/x}$

35. Find the linear function $\Lambda(x) = mx + b$ whose graph is the tangent line to the graph of e^{-x} at $x = 0$.

36. Compute the derivatives of the following functions:

(a) $\log|3x|$

(b) $\log(x^2)$

(c) $\log(3x) - \log x$

(d) $\mathrm{e}^{-\mathrm{e}^x}$

(e) $\log(1 + \mathrm{e}^{-x})$

37. Compute the derivatives of the following functions:

(a) $(\log x)'$

(b) $(\log(\log x))'$

(c) $(\log(\log(\log x)))'$

(d) $(\log(\log(\log(\log x))))'$

38. Find the equation of the line tangent to the graph of $\log x$ at $x = 1$.

39. Evaluate $\frac{\mathrm{d}}{\mathrm{d}x} \log\left(\frac{\sqrt{x^2+1}\sqrt{x^4-1}}{\sqrt{x^2-1}}\right)$.

40. Define two functions by $y(x) = \frac{1}{\mathrm{e}^x + \mathrm{e}^{-x}}$ and $z(x) = \mathrm{e}^x - \mathrm{e}^{-x}$. Verify the algebraic equation $z^2 = \frac{1}{y^2} - 4$ and the following differential equations:

(a) $y' = -zy^2$ and $z' = \frac{1}{y}$.

(b) Use part (a) and the chain rule to show that $y'' = y - 8y^3$.

41. Show that the normal line at any point on the circle $x^2 + y^2 = r^2$ passes through the origin.

42. For the function $y = \frac{x^2\sqrt{3x-2}}{(x+1)^2}$, $x > \frac{2}{3}$, find the derivative $\mathrm{d}y/\mathrm{d}x$.

43. For the function $y = \frac{x(x-1)^{3/2}}{\sqrt{x+1}}$, $x > 1$, find the derivative $\mathrm{d}y/\mathrm{d}x$.

44. For the equation $4x^3 + \log y^2 + 2y = 2x$, use implicit differentiation to find $\mathrm{d}y/\mathrm{d}x$.

45. For the function $y = \log\left(\frac{1+\mathrm{e}^x}{1-\mathrm{e}^x}\right)$, find the derivative $\mathrm{d}y/\mathrm{d}x$.

46. For the equation $\mathrm{e}^{xy} + x^2 - y^2 = 10$, use implicit differentiation to find $\mathrm{d}y/\mathrm{d}x$.

47. Find the point on the graph of the function $f(x) = \mathrm{e}^{2x}$ such that the tangent line to the graph at that point passes through the origin.

48. For the function $h(x) = \log_3 \frac{x\sqrt{x-1}}{2}$, find the derivative $h'(x)$.

49. For the function $y = x^{2/x}$, find the derivative $\mathrm{d}y/\mathrm{d}x$.

50. Are $f(x) = (x^x)^x$ and $g(x) = x^{(x^x)}$ the same function? Why or why not? Find $f'(x)$ and $g'(x)$.

51. For the function $y = x^{1/x}$, $x > 0$, find any asymptotes and relative extrema that may exist.

52. For the function $y = x\tan^{-1} 2x - \frac{1}{4}\log(1 + 4x^2)$, find the derivative $\mathrm{d}y/\mathrm{d}x$.

53. The line tangent to the graph of f at $(2, 7)$ has slope $\frac{1}{3}$. Find the equation of the line tangent to the graph of f^{-1} at $(7, 2)$.

54. Let $f(x) = x^3 + 2x^2 + 3x + 1$, and denote by g the inverse of f. Check that $f(1) = 7$ and calculate $g'(7)$.

55. Let $f(x) = \sqrt{x^2 - 1}$ for $x > 1$, and $g(y) = \sqrt{y^2 + 1}$ for $y > 0$.

(a) Show that f and g are inverses.

(b) Calculate f' and g'.

(c) Verify that $f'(g(y))g'(y) = 1$ and $g'(f(x))f'(x) = 1$.

56. For the function $y = 2x\sin^{-1}(2x) - \sqrt{1 + 4x^2}$, find the derivative $\mathrm{d}y/\mathrm{d}x$.

57. A particle is moving along a number line, and its position at time t is $x = f(t) = t - t^3$.

(a) Find the position of the particle at $t = 0$ and at $t = 2$.

(b) Find the velocity of the particle at $t = 0$ and at $t = 2$.

(c) In which direction is the particle moving at $t = 0$? at $t = 2$?

58. Find the acceleration $f''(t)$ of an object whose velocity at time t is given by $f'(t) = t^3 - \frac{1}{2}t^2$.

59. Suppose that at time $t = 1$, the population of a bacterial colony is 100, and that at any time, the rate at which the population grows per hour is 1.5 times the size of the population. Find the population at time $t = 3$ h.

60. Suppose that at time $t = 0$, the amount of carbon-14 in a sample of wood is 100, and that the half-life of carbon-14 is 5730 years. Find the amount of carbon-14 at time $t = 10,000$ years.

B.2 Bonus Problems

1. Find the points at which $x^2 + 4y^2 - 4x + 16y + 4 = 0$ has a vertical or horizontal tangent line.

2. Two circles of radius 4 are tangent to the graph of $y^2 = 4x$ at the point $(1, 2)$. Find equations of these two circles.

3. In this problem, all the numbers are positive. What is the relation between a and b so that the functions $f(x) = x^a$ and $g(y) = y^b$ are inverses? Show that if $\frac{1}{p} + \frac{1}{q} = 1$, then the derivatives of $\frac{x^p}{p}$ and $\frac{y^q}{q}$ are inverses.

4. A function f is called even if $f(-x) = f(x)$ for all x in the domain of f, and it is called odd if $f(-x) = -f(x)$. Let f be a differentiable function. Show that

(a) If f is even, then f' is odd.

(b) If f is odd, then f' is even.

5. Suppose f is strictly monotonic and $f'(x) \neq 0$ on an open interval. In this exercise we ask you to prove that the inverse of f is differentiable. Denote the inverse of f by g, set $y = f(x)$, and denote $f(x + h) - f(x)$ by k.

(a) Explain why $g(y + k) = x + h$.

(b) Show that if $h \neq 0$, then $k \neq 0$.

(c) Explain why $\frac{g(y+k)-g(y)}{k} = \frac{h}{k} = \frac{h}{f(x+h)-f(x)}$ tends to $\frac{1}{f'(x)}$ as k tends to 0.

6. Let f be twice-differentiable and one-to-one on an open interval I. Show that its inverse function g satisfies $g''(x) = -\frac{f''(g(x))}{[f'(g(x))]^3}$. When f is increasing and concave downward, what is the concavity of g?

7. Suppose f and g are positive functions. Obtain the product rule for $(fg)'$ by differentiating $\log(fg) = \log f + \log g$.

8. The product rule for a product of many functions can be deduced inductively from the product formula for two functions. Here is an alternative approach. Justify each step of the following argument:

(a) Suppose $y = f(x)g(x)h(x)k(x)$. Then $|y| = |f|\,|g|\,|h|\,|k|$, and when all are nonzero, $\log|y| = \log|f| + \log|g| + \log|h| + \log|k|$.

(b) $\frac{1}{y}y' = \frac{1}{f}f' + \frac{1}{g}g' + \frac{1}{h}h' + \frac{1}{k}k'$.

(c) $y' = f'ghk + fg'hk + fgh'k + fghk'$.

9. The angle between two graphs $y = f(x)$ and $y = g(x)$ is the angle between their tangents at the point of intersection. If the slopes are m_1 and m_2, then the angle of intersection θ is

$$\theta = \tan^{-1}\left(\frac{m_2 - m_1}{1 + m_1 m_2}\right).$$

(a) Derive the formula for θ.

(b) Find the angles at which the graphs of $y = x^3 + 1$ and $y = x^3 + x^2$ intersect.

10. Two graphs are orthogonal if at their point(s) of intersection, their tangent lines are perpendicular to each other. Let C and K be two real numbers. Show that the two families of curves $xy = C$ and $x^2 - y^2 = K$ are orthogonal.

Appendix C

Midterm 3 Review

C.1 Practice Problems

1. Evaluate the following limits. Use l'Hôpital's Rule when it is convenient and applicable.

(1) $\lim_{x\to 2} \dfrac{x^2-2x}{x^2-6x+8}$

(2) $\lim_{x\to 0} \dfrac{\mathrm{e}^x-1}{2x+5}$

(3) $\lim_{x\to 1} \dfrac{\log x}{4x-x^2-3}$

(4) $\lim_{x\to 0} \dfrac{\mathrm{e}^x-1}{x^2+3x}$

(5) $\lim_{x\to\infty} \dfrac{3x^4-x^2}{6x^4+12}$

(6) $\lim_{x\to\infty} \dfrac{4x^3-2x^2+6}{\pi x^3+4}$

(7) $\lim_{x\to \mathrm{e}} \dfrac{\log x-1}{x-\mathrm{e}}$

(8) $\lim_{x\to 1} \dfrac{4\tan^{-1}x-\pi}{x-1}$

(9) $\lim_{x\to 0^+} \dfrac{1-\log x}{1+\log x}$

(10) $\lim_{x\to 0^+} \dfrac{x-3\sqrt{x}}{x-\sqrt{x}}$

(11) $\lim_{x\to 0} \dfrac{3\sin 4x}{5x}$

(12) $\lim_{x\to 2\pi} \dfrac{x\sin x+x^2-4\pi^2}{x-2\pi}$

(13) $\lim_{u\to \pi/4} \dfrac{\tan u-\cot u}{u-\pi/4}$

(14) $\lim_{z\to 0} \dfrac{\tan 4z}{\tan 7z}$

(15) $\lim_{x\to 0} \dfrac{1-\cos 3x}{8x^2}$

(16) $\lim_{x\to 0} \dfrac{\sin^2 3x}{x^2}$

(17) $\lim_{x\to \pi} \dfrac{\cos x+1}{(x-\pi)^2}$

(18) $\lim_{x\to 0} \dfrac{\mathrm{e}^x-x-1}{5x^2}$

(19) $\lim_{x\to \pi/2^-} \dfrac{\tan x}{3/(2x-\pi)}$

(20) $\lim_{x\to\infty} \dfrac{\mathrm{e}^{3x}}{3\mathrm{e}^{3x}+5}$

(21) $\lim_{x\to 0} \dfrac{\mathrm{e}^x-\sin x-1}{x^4+8x^3+12x^2}$

(22) $\lim_{x\to 0} \dfrac{\sin x-x}{7x^3}$

(23) $\lim_{x\to\infty} \dfrac{\mathrm{e}^{1/x}-1}{1/x}$

(24) $\lim_{x\to\infty} \dfrac{\tan^{-1}x-\pi/2}{1/x}$

(25) $\lim_{x\to 0^+} x^{1/\log x}$

(26) $\lim_{x\to\infty} (\log_2 x-\log_3 x)$

(27) $\lim_{x\to 6} \dfrac{\sqrt[5]{5x+2}-2}{1/x-1/6}$

(28) $\lim_{x\to\pi/2}(\pi-2x)\tan x$

(29) $\lim_{x\to\infty} x^3\left(\frac{1}{x}-\sin\frac{1}{x}\right)$

(30) $\lim_{x\to\infty}(x^2\mathrm{e}^{1/x}-x^2-x)$

(31) $\lim_{x\to0^+} x^{2x}$

(32) $\lim_{x\to0}(1+4x)^{3/x}$

(33) $\lim_{\theta\to\pi/2^-}(\tan\theta)^{\cos\theta}$

(34) $\lim_{\theta\to0^+}(\sin\theta)^{\tan\theta}$

(35) $\lim_{x\to0^+}(1+x)^{\cot x}$

(36) $\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x$, for a constant a

(37) $\lim_{x\to0}(\mathrm{e}^{ax}+x)^{1/x}$, for a constant a

(38) $\lim_{z\to\infty}\left(1+\frac{10}{z^2}\right)^{z^2}$

(39) $\lim_{x\to0}(x+\cos x)^{1/x}$

(40) $\lim_{x\to0}\left(\frac{\sin x}{x}\right)^{1/x^2}$

(41) $\lim_{x\to1}\frac{x\log x-x+1}{x\log^2 x}$

(42) $\lim_{x\to1}\frac{x\log x+\log x-2x+2}{x^2\log^3 x}$

(43) $\lim_{x\to0^+} x^{1/(1+\log x)}$

(44) $\lim_{n\to\infty} n^2\log\left(n\sin\frac{1}{n}\right)$

(45) $\lim_{x\to0}\frac{a^x-b^x}{x}$, for positive constants a and b

(46) $\lim_{x\to0}(1+ax)^{b/x}$, for positive constants a and b

2. Exponential limit. Prove that $\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x=\mathrm{e}^a$, for $a\neq0$.

3. Exponentials vs. superexponentials. Show that x^x grows faster than b^x as $x\to\infty$, for $b>1$.

4. Exponential growth rates.

(a) For what values of $b>0$ does b^x grow faster than e^x as $x\to\infty$?

(b) Compare the growth rates of e^x and e^{ax} as $x\to\infty$, for $a>0$.

5. Sketch the graph of a function f that is continuous on $[a,b]$ such that f, f', and f'' have the signs indicated in the following table on $[a,b]$. There are eight different cases lettered A–H that correspond to eight different graphs.

Case	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
$\boldsymbol{f}$	+	+	+	+	−	−	−	−
$\boldsymbol{f'}$	+	+	−	−	+	+	−	−
$\boldsymbol{f''}$	+	−	+	−	+	−	+	−

Table C.1 – Sign table of f, f', and f'' across eight test cases.

6. Use the sign table to make a complete graph of f.

(1) $f(x)=x^2-6x$

(2) $f(x)=x-x^2$

(3) $f(x)=x^3-6x^2+9x$

(4) $f(x)=3x-x^3$

(5) $f(x)=x^4-6x^2$

(6) $f(x)=x^4+4x^3$

(7) $f(x) = (x-6)(x+6)^2$
(8) $f(x) = 27(x-2)^2(x+2)$
(9) $f(x) = x^3 - 6x^2 - 135x$
(10) $f(x) = x^4 + 8x^3 - 270x^2 + 1$
(11) $f(x) = x^3 - 3x^2 - 144x - 140$
(12) $f(x) = x^3 - 147x + 286$
(13) $f(x) = x - 2\sqrt{x}$
(14) $f(x) = 3\sqrt{x} - x^{3/2}$
(15) $f(x) = \dfrac{3x}{x^2-1}$
(16) $f(x) = \dfrac{2x-3}{2x-8}$
(17) $f(x) = \dfrac{x^2}{x-2}$
(18) $f(x) = \dfrac{x^2}{x^2-4}$
(19) $f(x) = \dfrac{x^2+12}{2x+1}$
(20) $f(x) = \dfrac{4x}{x^2+3}$
(21) $f(x) = \tan^{-1}\left(\dfrac{x^2}{\sqrt{3}}\right)$
(22) $f(x) = \log(x^2+1)$
(23) $f(x) = x + 2\cos x$ on $[-2\pi, 2\pi]$
(24) $f(x) = x - 3x^{2/3}$
(25) $f(x) = x - 3x^{1/3}$
(26) $f(x) = 2 - 2x^{2/3} + x^{4/3}$
(27) $f(x) = \sin x - x$ on $[0, 2\pi]$
(28) $f(x) = x\sqrt{x+3}$
(29) $f(x) = \mathrm{e}^{-x}\sin x$ on $[-\pi, \pi]$
(30) $f(x) = \dfrac{1}{\mathrm{e}^{-x}-1}$
(31) $f(x) = x + \tan x$ on $\left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right)$
(32) $f(x) = \mathrm{e}^{-x^2/2}$

7. Use analytical methods to find all local extrema of the function $f(x) = x^x$, for $x > 0$. Verify your work using a graphing utility.

8. Suppose a continuous function f is concave up on $(-\infty, 0)$ and $(0, \infty)$. Assume f has a local maximum at $x = 0$. What, if anything, do you know about $f'(0)$? Explain with an illustration.

9. Sketch the graph of a function f that is continuous on $(-\infty, \infty)$ and satisfies the following sets of conditions.

(a) $f''(x) > 0$ on $(-\infty, -2)$; $f''(-2) = 0$; $f'(-1) = f'(1) = 0$; $f''(2) = 0$; $f'(3) = 0$; $f''(x) > 0$ on $(4, \infty)$
(b) $f(-2) = f''(-1) = 0$; $f'\left(-\frac{3}{2}\right) = 0$; $f(0) = f'(0) = 0$; $f(1) = f'(1) = 0$

10. Suppose f is differentiable on $(-\infty, \infty)$ and assume it has a local extreme value at the point $x = 2$, where $f(2) = 0$. Let $g(x) = xf(x) + 1$ and let $h(x) = xf(x) + x + 1$, for all values of x.

(a) Evaluate $g(2)$, $h(2)$, $g'(2)$, and $h'(2)$.
(b) Does either g or h have a local extreme value at $x = 2$? Explain.

11. A family of double-humped functions. Consider the functions

$$f(x) = \frac{x}{(x^2+1)^n}, \quad \text{where } n \text{ is a positive integer.}$$

(a) Show that these functions are odd for all positive integers n.

(b) Show that the critical points of these functions are

$$x = \pm\frac{1}{\sqrt{2n-1}}, \quad \text{for all positive integers } n. \text{ (Start with the}$$

special cases $n = 1$ and $n = 2$.)

(c) Show that as n increases, the absolute maximum values of these functions decrease.

12. Local Extreme Value Theorem/Fermat's Theorem. If f has a local maximum value at the point c and $f'(c)$ exists, then $f'(c) = 0$.

Use the following steps to prove the theorem.

(a) Suppose f has a local maximum at c. What is the sign of $f(x) - f(c)$ if x is near c and $x > c$? What is the sign of $f(x) - f(c)$ if x is near c and $x < c$?

(b) If $f'(c)$ exists, then it is defined by $\lim_{x\to c} \dfrac{f(x) - f(c)}{x - c}$. Examine this limit as $x \to c^+$ and conclude that $f'(c) \le 0$.

(c) Examine the limit in part (b) as $x \to c^-$ and conclude that $f'(c) \ge 0$.

(d) Combine parts (b) and (c) to conclude that $f'(c) = 0$.

13. Even and odd functions.

(a) Suppose a nonconstant even function f has a local minimum at c. Does f have a local maximum or minimum at $-c$? Explain. (An even function satisfies $f(-x) = f(x)$.)

(b) Suppose a nonconstant odd function f has a local minimum at c. Does f have a local maximum or minimum at $-c$? Explain. (An odd function satisfies $f(-x) = -f(x)$.)

14. Suppose $f(2) = 6$, and $0.4 \le f'(x) \le 0.5$ for x in $[2, 2.2]$. Use the mean value theorem to estimate $f(2.1)$.

15. Suppose $g(2) = 6$ and $-0.6 \le g'(x) \le -0.5$ for x in $[1.8, 2]$. Use the mean value theorem to estimate $g(1.8)$.

16. If $h'(x) = 2\cos(3x) - 3\sin(2x)$, what could $h(x)$ be? If, in addition, $h(0) = 0$, what could $h(x)$ be?

17. If $k'(t) = 2 - 2e^{-3t}$, what could $k(t)$ be? If, in addition, $k(0) = 0$, what could $k(t)$ be?

18. Consider $f(x) = \frac{x}{x^2+1}$.

(a) Find $f'(x)$.

(b) In which interval(s) does f increase?

(c) In which interval(s) does f decrease?

(d) Find the minimum value of f in $[-10, 10]$.

(e) Find the maximum value of f in $[-10, 10]$.

19. Justify the following steps to prove the first derivative test. Suppose that $f'(x)$ is positive for $x < c$ and negative for $x > c$. We need to show that f reaches its maximum at c.

(a) Explain why f is increasing for all $x < c$, and why f is decreasing for all $x > c$.

(b) Use the continuity of f at c and the fact that f is increasing when $x < c$ to explain why $f(x)$ cannot be greater than $f(c)$ for any $x < c$. Similarly explain why $f(x)$ cannot be greater than $f(c)$ for any $x > c$.

(c) Explain why $f(c)$ is a maximum on S.

20. Suppose an experiment is carried out to determine the value of the constant m in the equation

$$y = mx$$

relating two physical quantities. Let (x_1, y_1), $(x_2, y_2), \ldots, (x_n, y_n)$ be the measured values. Find the value of m, in terms of the x_i and y_i, that minimizes E, the sum of the squares of the errors between the observed measurements and the linear function $y = mx$:

$$E = (y_1 - mx_1)^2 + (y_2 - mx_2)^2 + \cdots + (y_n - mx_n)^2.$$

21. Consider an open cardboard box whose bottom is a square of edge length x, and whose height is y. The volume V and surface area S of the box are given by

$$V = x^2 y, \qquad S = x^2 + 4xy.$$

Among all boxes with given volume, find the one with smallest surface area. Show that this box is squat, i.e., $y < x$.

22. Consider a particle of unit mass moving on a number line whose position at time t is given by $x(t) = 3t - t^2$. Find the time when the particle's position x is maximal.

23. Find the point on the graph of $y = \frac{1}{2}x^2$ closest to the point $(6, 0)$.

24. What is the largest amount by which a positive number can exceed its cube?

25. Over which intervals are the following functions concave up?

(a) $f(x) = 5x^4 - 3x^3 + x^2 - 1$

(b) $f(x) = \frac{x+1}{x-1}$

(c) $f(x) = \sqrt{x}$

(d) $f(x) = \frac{1}{\sqrt{x}}$

(e) $f(x) = \sqrt{1 - x^2}$

(f) $f(x) = \mathrm{e}^{-x^2}$

26. Are the linear approximations to $f(x) = x^2 - 3x + 5$ above or below the graph?

27. Is the secant line for $f(x) = -x^2 - 3x + 5$ on $[0, 7]$ above or below the graph?

28. Find an interval $(0, b)$ where $\mathrm{e}^{-1/x}$ is concave up. Sketch the graph on $(0, \infty)$.

29. Is e^f concave up when f is concave up?

30. Give an example of concave up functions f and g for which $f \circ g$ is not concave up.

C.2 Bonus Problems

1. The following classical curves have been studied by generations of mathematicians. Use analytical methods (including implicit differentiation) and a graphing utility to graph the curves. Include as much detail as possible.

(a) $x^{2/3} + y^{2/3} = 1$ (Astroid or hypocycloid with four cusps)

(b) $y = \dfrac{8}{x^2+4}$ (Witch of Agnesi)

(c) $x^3 + y^3 = 3xy$ (Folium of Descartes)

(d) $y^2 = x^3(1-x)$ (Pear curve)

(e) $x^4 - x^2 + y^2 = 0$ (Figure-8 curve)

2. Suppose $g'(x) \le h'(x)$ for $0 < x$ and $g(0) = h(0)$. Prove that $g(x) \le h(x)$ for $0 < x$. Then apply the result to solve the following problem:

(a) We find polynomial bounds for the cosine and sine.

i. Show that $g'(x) \le h'(x)$ for $g(x) = \sin x$ and $h(x) = x$ and deduce that

$$\sin x \le x \qquad (x > 0). \tag{C.1}$$

ii. Rewrite (C.1) as $(-\cos x)' \le \left(\frac{x^2}{2} - 1\right)'$ and deduce that $1 - \frac{x^2}{2} \le \cos x$.

iii. Continue along these lines to derive

$$1 - \frac{x^2}{2} \le \cos x \le 1 - \frac{x^2}{2} + \frac{x^4}{4!},$$

and in particular, estimate $\cos(0.2)$ with a tolerance of 0.001.

iv. Extend the previous argument to derive

$$1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} \le \cos x \le 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!}.$$

(b) Denote the exponential function e^x by $e(x)$. Use $e' = e$ to show the following.

i. For $x > 0$, $1 < e(x)$.

ii. $1 < e'(x)$ for $0 < x$, and deduce that $1 + x < e(x)$.

iii. Rewrite this as $1 + x < e'(x)$, and deduce that $1 + x + \frac{1}{2}x^2 < e(x)$ for $0 < x$.

iv. For all n and all $x > 0$, $1 + x + \frac{x^2}{2} + \cdots + \frac{x^n}{n!} < e(x)$.

3. Accept the so-called *linear approximation theorem* (see Theorem 4.9.11): for a twice continuously differentiable function f on an interval containing a and b, where $a < b$,

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(c)}{2}(b-a)^2. \tag{C.2}$$

In this problem we show (C.2) also holds for the case $a > b$. Given f'' continuous on $[a, b]$, define the function g as $g(x) = f(a + b - x)$.

(a) Show that g is defined in the interval $[a, b]$.

(b) Show that g'' is continuous in $[a, b]$.

(c) Show that

$$\begin{aligned} g(a) &= f(b), \qquad g'(a) = -f'(b), \qquad g''(a) = f''(b), \\ g(b) &= f(a), \qquad g'(b) = -f'(a), \qquad g''(b) = f''(a). \end{aligned}$$

(d) Write equation (C.2) for the function g. Then use results from part (c) to conclude that (C.2) holds for $b < a$.

4. Show, using the linear approximation theorem, that for f'' continuous on an interval containing a and b,

$$\frac{f(a) + f(b)}{2} \quad \text{differs from} \quad f\Big(\frac{a+b}{2}\Big) \quad \text{by less than} \quad \frac{M}{8}(b-a)^2,$$

where M is an upper bound for $|f''|$ on $[a, b]$.

5. Find the positive number x such that the sum of x and its reciprocal is as small as possible,

(a) Using calculus, and

(b) By the AM-GM inequality.

6. Use calculus to prove, by induction, the AM-GM inequality for n positive numbers.

7. Let $w_1, \ldots, w_n$ be positive numbers whose sum is 1, and $a_1, \ldots, a_n$ any positive numbers. Prove an extension of the AM-GM inequality:

$$a_1^{w_1} a_2^{w_2} \cdots a_n^{w_n} \le w_1 a_1 + w_2 a_2 + \cdots + w_n a_n,$$

with equality only in the case $a_1 = a_2 = \cdots = a_n$. Try an inductive proof with one of the a's as the variable.

Appendix D

Final Review

The final is comprehensive: it covers all the materials in the course. This final review contains only problems that are not included in the previous reviews.

D.1 Practice Problems

1. **Riemann sums for constant functions** Let $f(x) = c$, where $c > 0$, be a constant function on $[a, b]$. Prove that any Riemann sum for any value of n gives the exact area of the region between the graph of f and the x-axis on $[a, b]$.

2. **Riemann sums for linear functions** Assume the linear function $f(x) = mx + c$ is positive on the interval $[a, b]$. Prove that the midpoint Riemann sum with any value of n gives the exact area of the region between the graph of f and the x-axis on $[a, b]$.

3. **Shape of the graph for Riemann sums** Suppose a left Riemann sum is used to approximate the area of the region bounded by the graph of a positive function and the x-axis on the interval $[a, b]$. Describe whether the resulting approximation underestimates or overestimates the exact area in each combinations of monotonicity and concavity (there are 4 of them). Use a sketch to explain your reasoning in each case.

   Repeat the question for right Riemann sum.

4. Consider the following limits of Riemann sums for a function f on $[a, b]$. Identify f and express the limit as a definite integral.

   (a) $\lim\limits_{\Delta\to 0} \sum\limits_{k=1}^{n} (x_k^{*2} + 1)\Delta x_k$ on $[0, 2]$

   (b) $\lim\limits_{\Delta\to 0} \sum\limits_{k=1}^{n} (4 - x_k^{*2})\Delta x_k$ on $[-2, 2]$

   (c) $\lim\limits_{\Delta\to 0} \sum\limits_{k=1}^{n} x_k^* (\log x_k^*)\Delta x_k$ on $[1, 2]$

   (d) $\lim\limits_{\Delta\to 0} \sum\limits_{k=1}^{n} \left|x_k^{*2} - 1\right| \Delta x_k$ on $[-2, 2]$

5. Use geometry (not Riemann sums) to evaluate the following definite integrals. Sketch a graph of the integrand, show the region in question, and interpret your result.

(a) $\displaystyle\int_0^4 (8-2x)\,\mathrm{d}x$

(b) $\displaystyle\int_{-1}^2 (-|x|)\,\mathrm{d}x$

(c) $\displaystyle\int_0^4 \sqrt{16-x^2}\,\mathrm{d}x$

(d) $\displaystyle\int_0^4 f(x)\,\mathrm{d}x$, where

$$f(x) = \begin{cases} 5 & \text{if } x \le 2 \\ 3x-1 & \text{if } x > 2 \end{cases}$$

(e) $\displaystyle\int_{-4}^2 (2x+4)\,\mathrm{d}x$

(f) $\displaystyle\int_0^2 (1-x)\,\mathrm{d}x$

(g) $\displaystyle\int_{-1}^3 \sqrt{4-(x-1)^2}\,\mathrm{d}x$

(h) $\displaystyle\int_1^{10} g(x)\,\mathrm{d}x$, where

$$g(x) = \begin{cases} 4x & \text{if } 0 \le x \le 2 \\ -8x+16 & \text{if } 2 < x \le 3 \\ -8 & \text{if } x > 3 \end{cases}$$

6. Integrating piecewise continuous functions.

(a) Suppose f is continuous on the intervals $[a,p]$ and $(p,b]$, where $a < p < b$, with a finite jump at p. Form a uniform partition on the interval $[a,p]$ with n grid points and another uniform partition on the interval $[p,b]$ with m grid points, where p is a grid point of both partitions. Write a Riemann sum for $\int_a^b f(x)\,\mathrm{d}x$ and separate it into two pieces for $[a,p]$ and $[p,b]$. Explain why $\int_a^b f(x)\,\mathrm{d}x = \int_a^p f(x)\,\mathrm{d}x + \int_p^b f(x)\,\mathrm{d}x$.

(b) $\displaystyle\int_0^{10} f(x)\,\mathrm{d}x$, where $f(x) = \begin{cases} 2 & \text{if } 0 \le x \le 5 \\ 3 & \text{if } 5 < x \le 10 \end{cases}$

(c) $\displaystyle\int_1^6 f(x)\,\mathrm{d}x$, where $f(x) = \begin{cases} 2x & \text{if } 1 \le x < 4 \\ 10-2x & \text{if } 4 \le x \le 6 \end{cases}$

7. Powers of x by Riemann sums Consider the integral $I(p) = \int_0^1 x^p\,\mathrm{d}x$, where p is a positive integer.

(a) Write the left Riemann sum for the integral with n subintervals.

(b) It is a fact (proved by the 17th-century mathematicians Fermat and Pascal) that $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left(\frac{k}{n}\right)^p = \frac{1}{p+1}$. Use this fact to evaluate $I(p)$.

8. An exact integration formula Evaluate $\int_a^b \frac{\mathrm{d}x}{x^2}$, where $0 < a < b$, using the definition of the definite integral and the following steps.

(a) Assume $\{x_0, x_1, \ldots, x_n\}$ is a partition of $[a, b]$ with $\Delta x_k = x_k - x_{k-1}$, for $k = 1, 2, \ldots, n$. Show that $x_{k-1} \le \sqrt{x_{k-1}x_k} \le x_k$, for $k = 1, 2, \ldots, n$.

(b) Show that $\frac{1}{x_{k-1}} - \frac{1}{x_k} = \frac{\Delta x_k}{x_{k-1}x_k}$, for $k = 1, 2, \ldots, n$.

(c) Simplify the general Riemann sum for $\int_a^b \frac{\mathrm{d}x}{x^2}$ using $x_k^* = \sqrt{x_{k-1}x_k}$.

(d) Conclude that $\int_a^b \frac{\mathrm{d}x}{x^2} = \frac{1}{a} - \frac{1}{b}$.

(Source: The College Mathematics Journal, 32, 4, Sep 2001)

9. Use the fundamental theorem of calculus to calculate the derivatives.

(a) $\frac{\mathrm{d}}{\mathrm{d}x} \int_0^x t^3 \, \mathrm{d}t$

(b) $\frac{\mathrm{d}}{\mathrm{d}x} \int_0^x t^3 \mathrm{e}^{-t} \, \mathrm{d}t$

(c) $\frac{\mathrm{d}}{\mathrm{d}s} \int_{-2}^{s^2} x^3 \mathrm{e}^{-x} \, \mathrm{d}x$

(d) $h'(4)$, if $h(x) = \int_1^x \sqrt{t} \cos\left(\frac{\pi}{t}\right) \mathrm{d}t$

10. Use the fundamental theorem to calculate the integrals.

(a) $\int_1^2 t^3 \, \mathrm{d}t$

(b) $\int_0^b (x^3 + 5) \, \mathrm{d}x$

(c) $\int_0^1 \frac{1}{\sqrt{1+t}} \, \mathrm{d}t$

(d) $\int_0^7 \left(\cos t + (1+t)^{1/3}\right) \mathrm{d}t$

(e) $\int_a^b (t - \mathrm{e}^t) \, \mathrm{d}t$

11. Use the fundamental theorem to calculate the integrals.

(a) $\int_0^{\pi/4} \frac{1}{1+x^2} \, \mathrm{d}x$

(b) $\displaystyle\int_0^1 (x^2+2)^2\,\mathrm{d}x$

(c) $\displaystyle\int_1^4 \left(\frac{2}{\sqrt{x}} - \sqrt{x}\right)\mathrm{d}x$

(d) $\displaystyle\int_{-2}^{-1} (2 + 4t^{-2} - 8t^{-3})\,\mathrm{d}t$

(e) $\displaystyle\int_2^6 \left(2s + \frac{1}{s+1}\right)\mathrm{d}s$

12. Find the critical points of the function

$$f(x) = \int_{-1}^{x} t^2(t-3)(t-4)\,\mathrm{d}t,$$

and determine the intervals on which f is increasing or decreasing.

13. Determine the intervals on which the function

$$g(x) = \int_x^0 \frac{t}{t^2+1}\,\mathrm{d}t$$

is concave up or concave down.

14. Evaluate $\displaystyle\lim_{x\to 2} \frac{\int_2^x \sqrt{t^2+t+3}\,\mathrm{d}t}{x^2-4}$.

15. Maximum net area What value of $b > -1$ maximizes the integral $\int_{-1}^{b} x^2(3-x)\,\mathrm{d}x$?

16. Maximum net area Graph the function $f(x) = 8 + 2x - x^2$ and determine the values of a and b that maximize the value of the integral $\int_a^b f(x)\,\mathrm{d}x$.

17. Zero net area Consider the function $f(x) = x^2 - 4x$.

(a) Graph f on the interval $x \geq 0$.

(b) For what value of $b > 0$ is $\int_0^b f(x)\,\mathrm{d}x = 0$?

(c) In general, for the function $f(x) = x^2 - ax$, where $a > 0$, for what value of $b > 0$ (as a function of a) is $\int_0^b f(x)\,\mathrm{d}x = 0$?

18. Cubic zero net area Consider the graph of the cubic $y = x(x-a)(x-b)$, where $0 < a < b$. Verify that the graph bounds a region above the x-axis, for $0 < x < a$, and bounds a region below the x-axis, for $a < x < b$. What is the relationship between a and b if the areas of these two regions are equal?

19. An integral equation Use the Fundamental Theorem of Calculus, Part 1, to find the function f that satisfies the equation

$$\int_0^x f(t)\,\mathrm{d}t = 2\cos x + 3x - 2.$$

Verify the result by substitution into the equation.

20. Max/min of area functions Suppose f is continuous on $[0, \infty)$ and $A(x)$ is the net area of the region bounded by the graph of f and the t-axis on $[0, x]$. Show that the local maxima and minima of A occur at the zeros of f. Verify this fact with the function $f(x) = x^2 - 10x$.

21. Sine integral Show that the sine integral $S(x) = \int_0^x \frac{\sin t}{t}\,\mathrm{d}t$ satisfies the (differential) equation $xS'(x) + 2S''(x) + xS'''(x) = 0$.

22. Fresnel integral Show that the Fresnel integral $S(x) = \int_0^x \sin t^2\,\mathrm{d}t$ satisfies the (differential) equation

$$(S'(x))^2 + \left(\frac{S''(x)}{2x}\right)^2 = 1.$$

23. Symmetry of composite functions Prove that the integrand is either even or odd. Then give the value of the integral or show how it can be simplified. Assume f and g are even functions and p and q are odd functions.

(a) $\displaystyle\int_{-a}^{a} f(g(x))\,\mathrm{d}x$

(b) $\displaystyle\int_{-a}^{a} p(g(x))\,\mathrm{d}x$

(c) $\displaystyle\int_{-a}^{a} f(p(x))\,\mathrm{d}x$

(d) $\displaystyle\int_{-a}^{a} p(q(x))\,\mathrm{d}x$

24. Average value with a parameter Consider the function $f(x) = ax(1 - x)$ on the interval $[0, 1]$, where a is a positive real number.

(a) Find the average value of f as a function of a.

(b) Find the points at which the value of f equals its average value, and prove that they are independent of a.

25. Alternative definitions of means Consider the function

$$f(t) = \frac{\int_a^b x^{t+1}\,\mathrm{d}x}{\int_a^b x^t\,\mathrm{d}x}.$$

Show that the following means can be defined in terms of f.

(a) Arithmetic mean: $f(0) = \dfrac{a+b}{2}$

(b) Geometric mean: $f\left(-\dfrac{3}{2}\right) = \sqrt{ab}$

(c) Harmonic mean: $f(-3) = \dfrac{2ab}{a+b}$

(d) Logarithmic mean: $f(-1) = \dfrac{b-a}{\log b - \log a}$

(Source: Mathematics Magazine, 78, 5, Dec 2005)

26. Problems of antiquity Several calculus problems were solved by Greek mathematicians long before the discovery of calculus. The following problems were solved by Archimedes using methods that predated calculus by 2000 years.

(a) Show that the area of a segment of a parabola is 4/3 that of its inscribed triangle of greatest area. In other words, the area bounded by the parabola $y = a^2 - x^2$ and the x-axis is 4/3 the area of the triangle with vertices $(\pm a, 0)$ and $(0, a^2)$. Assume $a > 0$ is unspecified.

(b) Show that the area bounded by the parabola $y = a^2 - x^2$ and the x-axis is 2/3 the area of the rectangle with vertices $(\pm a, 0)$ and $(\pm a, a^2)$. Assume $a > 0$ is unspecified.

27. Average value of the derivative Suppose f' is a continuous function for all real numbers. Show that the average value of the derivative on an interval $[a, b]$ is $\bar{f}' = \dfrac{f(b) - f(a)}{b - a}$. Interpret this result in terms of secant lines.

28. Evaluate the following indefinite and definite integrals.

(1) $\displaystyle\int \frac{2x^2}{\sqrt{1 - 4x^3}}\,\mathrm{d}x$

(2) $\displaystyle\int (x^2 + x)^{10}(2x + 1)\,\mathrm{d}x$

(3) $\displaystyle\int x^3(x^4 + 16)^6\,\mathrm{d}x$

(4) $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{36 - 4x^2}}$

(5) $\displaystyle\int 6x^2 4^{x^3}\,\mathrm{d}x$

(6) $\displaystyle\int (x^6 - 3x^2)^4(x^5 - x)\,\mathrm{d}x$

(7) $\displaystyle\int \frac{3}{\sqrt{1 - 25x^2}}\,\mathrm{d}x$

(8) $\displaystyle\int \frac{\mathrm{e}^w}{36 + \mathrm{e}^{2w}}\,\mathrm{d}w$

(9) $\displaystyle\int x \csc x^2 \cot x^2\,\mathrm{d}x$

(10) $\displaystyle\int \sec^2(10x + 7)\,\mathrm{d}x$

(11) $\displaystyle\int 10^{4t+1}\,\mathrm{d}t$

(12) $\displaystyle\int \frac{(\sqrt{x} + 1)^4}{2\sqrt{x}}\,\mathrm{d}x$

(13) $\displaystyle\int \frac{1}{10x - 3}\,\mathrm{d}x$

(14) $\displaystyle\int \sin^{10}\theta \cos\theta\,\mathrm{d}\theta$

(15) $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{1 - 9x^2}}$

(16) $\displaystyle\int x^9 \sin x^{10}\,\mathrm{d}x$

(17) $\displaystyle\int \frac{\mathrm{d}x}{1 + 4x^2}$

(18) $\displaystyle\int \frac{2}{x\sqrt{4x^2 - 1}}\,\mathrm{d}x,\ x > \frac{1}{2}$

(19) $\displaystyle\int \frac{8x + 6}{2x^2 + 3x}\,\mathrm{d}x$

(20) $\displaystyle\int \sec 4w \tan 4w\,\mathrm{d}w$

(21) $\displaystyle\int \frac{\tan^{-1} w}{w^2 + 1}\,\mathrm{d}w$

(22) $\displaystyle\int_0^{\pi/8} \cos 2x\,\mathrm{d}x$

(23) $\displaystyle\int_0^1 2x(4 - x^2)\,\mathrm{d}x$

(24) $\displaystyle\int_1^3 \frac{2^x}{2^x + 4}\,\mathrm{d}x$

(25) $\displaystyle\int_0^{\pi/2} \sin^2\theta \cos\theta\,\mathrm{d}\theta$

(26) $\displaystyle\int_{\log\frac{\pi}{4}}^{\log\frac{\pi}{2}} \mathrm{e}^w \cos \mathrm{e}^w\,\mathrm{d}w$

(27) $\displaystyle\int_{-1}^{2} x^2 \mathrm{e}^{x^3+1}\,\mathrm{d}x$

(28) $\displaystyle\int_{\pi/4}^{\pi/2} \frac{\cos x}{\sin^2 x}\,\mathrm{d}x$

(29) $\displaystyle\int_{2/(5\sqrt{3})}^{2/5} \frac{\mathrm{d}x}{x\sqrt{25x^2-1}}$

(30) $\displaystyle\int_{0}^{4} \frac{x}{x^2+1}\,\mathrm{d}x$

(31) $\displaystyle\int_{1/3}^{1/\sqrt{3}} \frac{4}{9x^2+1}\,\mathrm{d}x$

(32) $\displaystyle\int_{0}^{1} 2\mathrm{e}^{2x}\,\mathrm{d}x$

(33) $\displaystyle\int_{0}^{2} \frac{2x}{(x^2+1)^2}\,\mathrm{d}x$

(34) $\displaystyle\int_{-2\pi}^{2\pi} \cos\frac{\theta}{8}\,\mathrm{d}\theta$

(35) $\displaystyle\int_{0}^{\pi/4} \frac{\sin x}{\cos^2 x}\,\mathrm{d}x$

(36) $\displaystyle\int_{\pi/16}^{\pi/8} 8\csc^2 4x\,\mathrm{d}x$

(37) $\displaystyle\int_{0}^{4} \frac{p}{\sqrt{9+p^2}}\,\mathrm{d}p$

(38) $\displaystyle\int_{0}^{\pi/4} \frac{\sin\theta}{\cos^3\theta}\,\mathrm{d}\theta$

(39) $\displaystyle\int_{0}^{1} \frac{v^3+1}{\sqrt{v^4+4v+4}}\,\mathrm{d}v$

(40) $\displaystyle\int_{0}^{1/8} \frac{x}{\sqrt{1-16x^2}}\,\mathrm{d}x$

(41) $\displaystyle\int_{0}^{\log 4} \frac{\mathrm{e}^x}{3+2\mathrm{e}^x}\,\mathrm{d}x$

29. Evaluate the following integrals in which the function f is unspecified. Note that $f^{(p)}$ is the pth derivative of f and f^p is the pth power of f. Assume f and its derivatives are continuous for all real numbers.

(a) $\displaystyle\int \left(5f^3(x)+7f^2(x)+f(x)\right)f'(x)\,\mathrm{d}x$

(b) $\displaystyle\int_{1}^{2} \left(5f^3(x)+7f^2(x)+f(x)\right)f'(x)\,\mathrm{d}x$, where $f(1)=4$, $f(2)=5$

(c) $\displaystyle\int \left(f^{(p)}(x)\right)^n f^{(p+1)}(x)\,\mathrm{d}x$, where p is a positive integer, $n\neq -1$

D.2 Bonus Problems

1. Express the area bounded by the ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

as a definite integral. Make a change of variables that converts this integral into one representing the area of a circle, and then evaluate it.

2. Evaluate

$$\int_{0}^{1} \sqrt{1+\sqrt{x}}\,\mathrm{d}x.$$

Hint: Let $\sqrt{x}=t$.

3. (a) Find $\dfrac{\mathrm{d}}{\mathrm{d}x}\log(\sec x+\tan x)$. For what values of x is your result valid?

(b) Explain which, if any, of the following integrals have been defined:

$$\int_1^{3/2} \sec x \, \mathrm{d}x, \qquad \int_1^{\pi/2} \sec x \, \mathrm{d}x \qquad \int_{\pi/2}^{2} \sec x \, \mathrm{d}x \qquad \int_1^2 \sec x \, \mathrm{d}x$$

4. Let g be a function whose derivative on the interval $[a, b]$ is negative. Prove that in this case, we have the following formula for changing variables in an integral:

$$\int_{g(b)}^{g(a)} f(x) \, \mathrm{d}x = \int_a^b f(g(t)) \, |g'(t)| \, \mathrm{d}t.$$

5. Use the change of variables theorem to prove:

(a) If f is any continuous function on $[a, b]$ and if $f_r(x)$ is defined on $[a + r, b + r]$ according to the rule $f_r(x) = f(x - r)$, then

$$\int_a^b f(x) \, \mathrm{d}x = \int_{a+r}^{b+r} f_r(x) \, \mathrm{d}x.$$

(b) If f is continuous on $[a, b]$ and its reflection, denoted by f_-, is defined on $[-b, -a]$ as $f_-(x) = f(-x)$, then

$$\int_a^b f(x) \, \mathrm{d}x = \int_{-b}^{-a} f_-(x) \, \mathrm{d}x.$$

6. Suppose f is an even function, i.e., $f(t) = f(-t)$, and set $g(x) = \displaystyle\int_0^x f(t) \, \mathrm{d}t$. Explain why g is odd, i.e., $g(x) = -g(-x)$.

7. Suppose g is a differentiable function and

$$F(x) = \int_a^{g(x)} f(t) \, \mathrm{d}t.$$

Explain why $F'(x) = f(g(x))g'(x)$.

8. Express the limit

$$\lim_{n \to \infty} \frac{1 + 4 + 9 + \cdots + (n-1)^2}{n^3}$$

as an integral of some function over some interval, and find its value.

Bibliography

[Abb15] Stephen Abbott, *Understanding Analysis*, Undergraduate Texts in Mathematics, Springer, New York, NY, 2015. ↑v

[BCGS18] William L. Briggs, Lyle Cochran, Bernard Gillett, and Eric P. Schulz, *Calculus: Early Transcendentals*, Pearson Education, 2018. ↑v

[BO99] Carl M. Bender and Steven A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers I*, Springer, New York, NY, 1999. ↑91

[Cli26] Vaughn Climenhaga, *Calculus I and II*, 2026. ↑v, 91

[GM03] Mariano Giaquinta and Giuseppe Modica, *Mathematical Analysis: Functions of One Variable*, Springer Science & Business Media, 2003. ↑v

[GM10] ———, *Mathematical Analysis: An Introduction to Functions of Several Variables*, Springer Science & Business Media, 2010. ↑v

[LT14] Peter D. Lax and Maria Shea Terrell, *Calculus With Applications*, Undergraduate Texts in Mathematics, Springer, New York, NY, 2014. ↑v, 91

[LT18] ———, *Multivariable Calculus with Applications*, Springer, 2018. ↑v

[Pug15] Charles Chapman Pugh, *Real Mathematical Analysis*, Springer, 2015. ↑v

[Ste22] Ian Stewart, *Galois Theory*, CRC Press, 2022. ↑104

[Ste25] Stefan Steinerberger, *The Unreasonable Elegance of Mathematics*, Springer Nature, 2025. ↑v

[Tao22a] Terence Tao, *Analysis I*, Texts and Readings in Mathematics, vol. 37, Springer Nature, Singapore, 2022. ↑v, 104

[Tao22b] ———, *Analysis II*, Texts and Readings in Mathematics, vol. 38, Springer Nature, Singapore, 2022. ↑v

Index

acceleration, 119, 120
antiderivative, 190
asymptote
    horizontal, 58
    slant (or oblique), 60
    vertical, 60
asymptotic
    rate of convergence, 154
atlas, 116
average value, 187

change of variables, 203
charts, 116
compact, 214
complete, 65
concave down, 141
concave up, 141
continuous, 35, 37
converge
    with order q, 154
critical
    point, 132

decreasing, 17
    strictly, 17
derivative, 68
difference quotient, 22
differentiable, 68
differential, 150
discontinuous, 39
displacement, 119
distance
    signed, 166
dummy variable, 167

e, 201
end behaviors, 57
equation
    differential, 101
even
    function, 143, 191
everywhere
    continuous, 37
exponential function, 201
extreme values, 130

fixed-point
    iteration, 153
fixed-point problem, 152
function
    1-to-1, 7
    decreasing, 137
    exponential, 16
    increasing, 137
    logarithmic, 18
    monotone, 137
    natural exponential , 18
    natural logarithmic, 18
    position, 119
    signum, 25
    Weierstrass monster, 74

Heine–Borel theorem, 213

increasing, 17
    strictly, 17
indeterminate forms, 124
induction
    mathematical, 81
inflection point, 141
integrable, 175, 177
integral
    definite, 164, 176
    indefinite, 190

iterates, 153

Jacques Hadamard, 86

Leibniz rule, 79
linearly
    linearly, 154
logarithmic derivative, 108

maximum
    absolute, 130
    global, 130
    local, 131
maximum value
    absolute, 130
method
    Newton, 155
    Newton–Raphson, 155
minimum
    absolute, 130
    global, 130
    local, 131
minimum value
    absolute, 130
monotone, 17

natural
    logarithm, 199
net displacement, 166
norm
    of a partition, 172

odd
    function, 191
open
    cover, 214
order
    of convergence, 154

partition, 172
    tagged, 172

quadratically
    quadratically, 154

radian, 10
regular
    partition, 173
relative
    maximum, 131
    minimum, 131
relaxation
    method, 155
Riemann sum, 173
    left, 173
    middle, 173
    right, 173
Riemann Surface, 15

simple
    iteration, 153
simple harmonic oscillator, 103
sine function, 202
speed, 120
Substitution Rule, 203
sum
    lower Darboux, 213
    upper Darboux, 213
summation
    index of, 167

tangent line, 160
Taylor series, 158
Theorem
    Vieta's, 44
theorem
    existence, 64
    intermediate value, 64
    mean value, 135
    Rolle, 136
Topologist's Sine Curve, 28
total distance travelled, 166

uniformly
    continuous, 215

velocity, 119, 165
    average, 119
    instantaneous, 119

width
    of a subinterval, 172