Study Guide for Apostol's Mathematical Analysis

With Problem Statements and Definitions

November 22, 2025

Let's Study Together!

Contents

Chapter 1

The Real and Complex Number Systems

1.I Integers

1.1: No Largest Prime

Prove that there is no largest prime. (A proof was known to Euclid.)

Strategy: Use proof by contradiction. Assume there exists a largest prime p, then consider $N = p! + 1$. Since N is either prime or has a prime factor greater than p, this contradicts the assumption.

Solution: We will prove this by contradiction. Assume there exists a largest prime number, call it p.

Consider the number $N = p! + 1$, where $p!$ is the factorial of p.

Since $p!$ is divisible by all integers from 1 to p, the number $N = p!+1$ is not divisible by any prime number less than or equal to p.

Now, N is either prime or composite:

- If N is prime, then $N > p$, contradicting our assumption that p is the largest prime.
- If N is composite, then N has a prime factor q. Since N is not divisible by any prime $\leq p$, we must have $q > p$. This again contradicts our assumption that p is the largest prime.

In both cases, we reach a contradiction. Therefore, our assumption that there exists a largest prime is false, and there must be infinitely many prime numbers. ■

1.2: Algebraic Identity

If n is a positive integer, prove the algebraic identity:

$$a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k}$$

Strategy: Expand the right-hand side and show it equals the left-hand side by distributing $(a-b)$ and observing that most terms cancel out.

Solution: We can prove this identity by expanding the right-hand side and showing it equals the left-hand side.

Let's expand the sum:

$$\begin{aligned}(a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k} &= (a-b)(a^0 b^{n-1} + a^1 b^{n-2} + a^2 b^{n-3} + \cdots + a^{n-1} b^0) \\ &= (a-b)(b^{n-1} + ab^{n-2} + a^2 b^{n-3} + \cdots + a^{n-1})\end{aligned}$$

Now distribute $(a-b)$:

$$\begin{aligned}&= a \cdot b^{n-1} + a^2 b^{n-2} + a^3 b^{n-3} + \cdots + a^n \\ &\quad - b \cdot b^{n-1} - ab^{n-1} - a^2 b^{n-2} - \cdots - a^{n-1} b\end{aligned}$$

Notice that most terms cancel out:

$$\begin{aligned}&= a^n - b^n + \text{(canceling terms)} \\ &= a^n - b^n\end{aligned}$$

Alternatively, we can use the geometric series formula. Let $r = \frac{a}{b}$. Then:

$$\begin{aligned}
\sum_{k=0}^{n-1} a^k b^{n-1-k} &= b^{n-1} \sum_{k=0}^{n-1} \left(\frac{a}{b}\right)^k \\
&= b^{n-1} \cdot \frac{1 - \left(\frac{a}{b}\right)^n}{1 - \frac{a}{b}} \\
&= b^{n-1} \cdot \frac{b^n - a^n}{b^n(b-a)} \\
&= \frac{a^n - b^n}{a - b}
\end{aligned}$$

Therefore, $(a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k} = a^n - b^n$. ■

1.3: Mersenne Primes

If $2^n - 1$ is prime, prove that n is prime. A prime of the form $2^p - 1$, where p is prime, is called a *Mersenne prime*.

Strategy: Prove the contrapositive: if n is composite, then $2^n - 1$ is composite. Use the identity from Problem 1.2 to factor $2^n - 1$ when $n = ab$ with $a, b > 1$.

Solution: We will prove the contrapositive: if n is composite, then $2^n - 1$ is composite.

Let $n = ab$ where $a, b > 1$ are integers. Then:

$$\begin{aligned}
2^n - 1 &= 2^{ab} - 1 \\
&= (2^a)^b - 1
\end{aligned}$$

Using the identity from Problem 1.2 with $x = 2^a$ and $y = 1$:

$$(2^a)^b - 1 = (2^a - 1)((2^a)^{b-1} + (2^a)^{b-2} + \cdots + 2^a + 1)$$

Since $a > 1$, we have $2^a - 1 > 1$. Also, since $b > 1$, the second factor is greater than 1. Therefore, $2^n - 1$ is the product of two integers greater than 1, making it composite.

This proves that if $2^n - 1$ is prime, then n must be prime. ■

1.4: Fermat Primes

If $2^n + 1$ is prime, prove that n is a power of 2. A prime of the form $2^{2^n} + 1$ is called a *Fermat prime.* Hint: Use Exercise 1.2.

Strategy: Prove the contrapositive: if n is not a power of 2, then $2^n + 1$ is composite. When n has an odd factor, use the identity from Problem 1.2 to factor $2^n + 1$.

Solution: We will prove the contrapositive: if n is not a power of 2, then $2^n + 1$ is composite.

If n is not a power of 2, then n has an odd factor greater than 1. Let $n = 2^k \cdot m$ where $m > 1$ is odd and $k \geq 0$.

Then:

$$\begin{aligned} 2^n + 1 &= 2^{2^k \cdot m} + 1 \\ &= (2^{2^k})^m + 1 \end{aligned}$$

Since m is odd, we can use the identity from Problem 1.2 with $a = 2^{2^k}$ and $b = -1$:

$$(2^{2^k})^m - (-1)^m = (2^{2^k} - (-1))((2^{2^k})^{m-1} + (2^{2^k})^{m-2}(-1) + \cdots + (-1)^{m-1})$$

Since m is odd, $(-1)^m = -1$, so:

$$(2^{2^k})^m + 1 = (2^{2^k} + 1)((2^{2^k})^{m-1} - (2^{2^k})^{m-2} + \cdots + 1)$$

Since $m > 1$, both factors are greater than 1, making $2^n + 1$ composite.

Therefore, if $2^n + 1$ is prime, then n must be a power of 2. ■

1.5: Fibonacci Numbers Formula

The Fibonacci numbers $1, 1, 2, 3, 5, 8, 13, \ldots$ are defined by the recursion formula $x_{n+1} = x_n + x_{n-1}$, with $x_1 = x_2 = 1$. Prove that $x_n = \frac{a^n - b^n}{a-b}$, where a and b are the roots of the equation $x^2 - x - 1 = 0$.

Strategy: Use strong mathematical induction. Verify base cases for $n = 1$ and $n = 2$, then use the inductive hypothesis and the key property that $a^2 = a + 1$ and $b^2 = b + 1$ to establish the inductive step.

Solution: Let the proposition be $P(n) : x_n = \frac{a^n - b^n}{a-b}$. The roots of $x^2 - x - 1 = 0$ are $a = \frac{1+\sqrt{5}}{2}$ and $b = \frac{1-\sqrt{5}}{2}$. A key property of these roots is that they satisfy $a^2 = a + 1$ and $b^2 = b + 1$.

Base cases: For $n = 1$:

$$\frac{a^1 - b^1}{a - b} = 1 = x_1.$$

For $n = 2$:

$$\begin{aligned}\frac{a^2 - b^2}{a - b} &= \frac{(a-b)(a+b)}{a-b} \\ &= a + b \\ &= \left(\frac{1+\sqrt{5}}{2}\right) + \left(\frac{1-\sqrt{5}}{2}\right) \\ &= 1 = x_2.\end{aligned}$$

The base cases hold.

Inductive step: Assume $P(k)$ is true for all integers $k \leq n$, where $n \geq 2$. We will show $P(n+1)$ is true. By the definition of the Fibonacci sequence, $x_{n+1} = x_n + x_{n-1}$. Using the inductive hypothesis for x_n and x_{n-1}:

$$\begin{aligned}x_{n+1} &= \left(\frac{a^n - b^n}{a-b}\right) + \left(\frac{a^{n-1} - b^{n-1}}{a-b}\right) \\ &= \frac{(a^n + a^{n-1}) - (b^n + b^{n-1})}{a-b} \\ &= \frac{a^{n-1}(a+1) - b^{n-1}(b+1)}{a-b}\end{aligned}$$

Using the property that $a + 1 = a^2$ and $b + 1 = b^2$:

$$\begin{aligned}x_{n+1} &= \frac{a^{n-1}(a^2) - b^{n-1}(b^2)}{a-b} \\ &= \frac{a^{n+1} - b^{n+1}}{a-b}\end{aligned}$$

This is $P(n+1)$. By the principle of strong induction, the formula is true for all positive integers n. ■

1.6: Well-Ordering Principle

Prove that every nonempty set of positive integers contains a smallest member. This is called the *well-ordering principle*.

Strategy: Use proof by contradiction combined with mathematical induction. Assume there exists a nonempty set with no smallest member, then use induction to show this leads to the set being empty.

Solution: We will prove this by contradiction. Let S be a nonempty set of positive integers that has no smallest member. Let $P(n)$ be the proposition that the integer n is not in S. We will use induction to show that $P(n)$ is true for all positive integers n.

Base case: For $n = 1$: If $1 \in S$, then 1 would be the smallest member of S (since S contains only positive integers). But we assumed S has no smallest member. So 1 cannot be in S. Thus, $P(1)$ is true.

Inductive step: Assume that $P(k)$ is true for all positive integers $k < n$. This means that none of the integers $1, 2, \ldots, n-1$ are in S. Now consider the integer n. If $n \in S$, then from our inductive hypothesis (that $1, 2, \ldots, n-1$ are not in S), n would be the smallest member of S. This contradicts our initial assumption that S has no smallest member. Therefore, n cannot be in S. Thus, $P(n)$ is true.

By the principle of strong induction, $P(n)$ is true for all positive integers n. This means that no positive integer is in S, which implies that S is an empty set. This contradicts our initial assumption that S is a nonempty set. Therefore, the original assumption must be false, and every nonempty set of positive integers must contain a smallest member. ■

1.II Rational and Irrational Numbers

Definitions and Theorems

Theorem: Geometric Series Formula

For $|r| < 1$, the infinite geometric series converges:

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$$

For a finite geometric series:

$$\sum_{n=0}^{N-1} ar^n = a \cdot \frac{1-r^N}{1-r}$$

Importance: The geometric series is one of the most fundamental infinite series in mathematics. It provides the foundation for understanding power series, Taylor series, and many other important series expansions. This formula is essential for calculus, analysis, and many applications in physics and engineering.

Theorem: Terminating Decimal Criterion

A rational number $\frac{p}{q}$ has a terminating decimal expansion if and only if the prime factorization of q contains only powers of 2 and 5.

Importance: This theorem provides a precise criterion for determining when a rational number has a finite decimal representation. It connects number theory (prime factorization) to decimal arithmetic, making it easy to predict the behavior of decimal expansions. This result is essential for understanding decimal representations and their applications.

Theorem: Irrationality of Square Roots

If n is a positive integer that is not a perfect square, then $\sqrt{n}$ is irrational.

Importance: This theorem provides a large class of irrational numbers and is fundamental for understanding the relationship between algebra and geometry. It shows that many naturally occurring numbers (like $\sqrt{2}$, $\sqrt{3}$, etc.) are irrational, which is essential for geometry, algebra, and many areas of mathematics. The proof technique used here is a classic example of proof by contradiction.

Theorem: Density of Rationals and Irrationals

Between any two real numbers, there exist both rational and irrational numbers.

Importance: This theorem shows that both rational and irrational numbers are densely distributed throughout the real number line. This property is fundamental for understanding the structure of the real numbers and is essential for many results in analysis, topology, and approximation theory. It demonstrates that the real number system is rich and complex.

1.7: Decimal Expansion to Rational

Find the rational number whose decimal expansion is $0.334444\ldots$.

Strategy: Use an algebraic method by multiplying by powers of 10 to shift the decimal point and eliminate the repeating part, then solve for the unknown fraction.

Solution: We can use an algebraic method to find the equivalent fraction. Let x be the rational number.

$$x = 0.334444\ldots$$

The goal is to manipulate the equation to eliminate the repeating decimal part. The repeating digit '4' begins at the third decimal place.

First, multiply by 100 to move the non-repeating part to the left of the decimal point:

$$100x = 33.4444\ldots$$

Next, multiply by 1000 to shift the decimal point past the first repeating digit:

$$1000x = 334.4444\ldots$$

Now, subtract the first equation from the second. This will cancel the infinite repeating tail.

$$\begin{array}{rl} 1000x &= 334.4444\ldots \\ -\quad 100x &= \ \ 33.4444\ldots \\ \hline 900x &= 301 \end{array}$$

Finally, solve for x:

$$x = \frac{301}{900}$$

Therefore, the rational number is $\frac{301}{900}$. ■

1.8: Decimal Expansion Ending in Zeroes

Prove that the decimal expansion of x will end in zeroes (or in nines) if and only if x is a rational number whose denominator is of the form $2^m 5^n$, where m and n are nonnegative integers.

Strategy: Prove both directions of the if-and-only-if statement. Show that rational numbers with denominators of the form $2^m 5^n$ have terminating decimal expansions, and conversely that terminating decimals correspond to such rational numbers.

Solution: We need to prove both directions of this statement.

Forward direction: If x is rational with denominator of the form $2^m 5^n$, then its decimal expansion terminates.

Let $x = \frac{p}{q}$ where $q = 2^m 5^n$ for some nonnegative integers m, n.

We can write $x = \frac{p}{2^m 5^n} = \frac{p \cdot 2^n 5^m}{2^m 5^n \cdot 2^n 5^m} = \frac{p \cdot 2^n 5^m}{10^{m+n}}$

This shows that x can be written as a fraction with denominator a power of 10, which means its decimal expansion terminates.

Reverse direction: If the decimal expansion of x terminates, then x is rational with denominator of the form $2^m 5^n$.

Let x have a terminating decimal expansion. Then x can be written as $x = \frac{N}{10^k}$ for some integer N and nonnegative integer k.

Since $10 = 2 \cdot 5$, we have $10^k = 2^k \cdot 5^k$, which is of the required form.

Note about ending in nines: If a decimal expansion ends in nines (e.g., $0.999\ldots$), this is equivalent to the next terminating decimal. For example, $0.999\ldots = 1.000\ldots$. This is because $0.999\ldots = \frac{9}{10} + \frac{9}{100} + \frac{9}{1000} + \cdots = \frac{9/10}{1-1/10} = 1$.

Therefore, both terminating decimals and those ending in nines correspond to rational numbers with denominators of the form $2^m 5^n$. ■

1.9: Irrationality of $\sqrt{2}+\sqrt{3}$

Prove that $\sqrt{2}+\sqrt{3}$ is irrational.

Strategy: Use proof by contradiction. Assume $\sqrt{2}+\sqrt{3}$ is rational, then square both sides to eliminate the square roots. This leads to $\sqrt{6}$ being rational, which is false.

Solution: We will prove this by contradiction. Assume that $\sqrt{2}+\sqrt{3}$ is rational, say $\sqrt{2}+\sqrt{3} = \frac{p}{q}$ where p, q are integers with no common factors.

Then:

$$\begin{aligned}
\sqrt{2}+\sqrt{3} &= \frac{p}{q} \\
(\sqrt{2}+\sqrt{3})^2 &= \left(\frac{p}{q}\right)^2 \\
2+2\sqrt{6}+3 &= \frac{p^2}{q^2} \\
5+2\sqrt{6} &= \frac{p^2}{q^2} \\
2\sqrt{6} &= \frac{p^2}{q^2} - 5 \\
\sqrt{6} &= \frac{p^2-5q^2}{2q^2}
\end{aligned}$$

This shows that $\sqrt{6}$ is rational, which is a contradiction since $\sqrt{6}$ is irrational.

To see why $\sqrt{6}$ is irrational, suppose $\sqrt{6} = \frac{a}{b}$ where a, b are integers with no common factors. Then:

$$\begin{aligned}
6 &= \frac{a^2}{b^2} \\
6b^2 &= a^2
\end{aligned}$$

This means a^2 is divisible by 6, so a must be divisible by 6. Let $a = 6k$. Then:

$$6b^2 = (6k)^2 = 36k^2$$
$$b^2 = 6k^2$$

This means b^2 is divisible by 6, so b must also be divisible by 6. But this contradicts our assumption that a and b have no common factors.

Therefore, $\sqrt{6}$ is irrational, and consequently $\sqrt{2} + \sqrt{3}$ is irrational. ■

1.10: Rational Functions of Irrational Numbers

If a, b, c, d are rational and if x is irrational, prove that $\frac{ax+b}{cx+d}$ is usually irrational. When do exceptions occur?

Strategy: Assume the expression is rational and solve for the conditions under which this can happen. Use the fact that if a rational expression equals a rational number, then the coefficients must satisfy certain relationships.

Solution: We need to analyze when $\frac{ax+b}{cx+d}$ is rational, given that x is irrational and a, b, c, d are rational.

Let's assume that $\frac{ax+b}{cx+d} = \frac{p}{q}$ where p, q are integers with no common factors.

Then:

$$\frac{ax+b}{cx+d} = \frac{p}{q}$$
$$q(ax+b) = p(cx+d)$$
$$qax + qb = pcx + pd$$
$$(qa - pc)x = pd - qb$$

Since x is irrational and the right-hand side is rational, we must have $qa - pc = 0$ and $pd - qb = 0$.

This gives us:

$$qa = pc$$
$$pd = qb$$

From the first equation: $a = \frac{pc}{q}$ From the second equation: $b = \frac{pd}{q}$ Therefore, we have:

$$\frac{a}{c} = \frac{p}{q}$$
$$\frac{b}{d} = \frac{p}{q}$$

This means $\frac{a}{c} = \frac{b}{d}$, or equivalently, $ad = bc$.

Conclusion: The expression $\frac{ax+b}{cx+d}$ is rational if and only if $ad = bc$.

Exceptions occur when: $ad = bc$, which means the numerator and denominator are proportional, making the fraction rational regardless of the value of x.

Examples:

- If $a = 2, b = 1, c = 4, d = 2$, then $ad = 4 = bc = 4$, so $\frac{2x+1}{4x+2} = \frac{1}{2}$ for all x.
- If $a = 1, b = 0, c = 1, d = 0$, then $ad = 0 = bc = 0$, so $\frac{x}{x} = 1$ for all $x \neq 0$.

■

1.11: Irrational Numbers Between 0 and x

Given any real $x > 0$, prove that there is an irrational number between 0 and x.

Strategy: Construct an irrational number between 0 and x by considering two cases: when x is irrational (use $\frac{x}{2}$) and when x is rational (use $\frac{x}{\sqrt{2}}$).

Solution: We will construct an irrational number between 0 and x for any positive real number x.

Case 1: If x is irrational, then $\frac{x}{2}$ is irrational and lies between 0 and x.

To see why $\frac{x}{2}$ is irrational, suppose it were rational. Then $\frac{x}{2} = \frac{p}{q}$ for some integers p, q, which would mean $x = \frac{2p}{q}$, making x rational, a contradiction.

Case 2: If x is rational, let $x = \frac{p}{q}$ where p, q are positive integers. Consider the number $y = \frac{x}{\sqrt{2}} = \frac{p}{q\sqrt{2}}$.

Since $\sqrt{2}$ is irrational, y is irrational (if y were rational, then $\sqrt{2} = \frac{p}{qy}$ would be rational, a contradiction).

Also, since $\sqrt{2} > 1$, we have $y < x$.

Therefore, y is an irrational number between 0 and x.

Alternative construction: For any positive real x, we can also use $y = \frac{x}{\pi}$. Since π is irrational and greater than 1, we have $0 < y < x$, and y is irrational. ∎

1.12: Fraction Between Two Fractions

If $\frac{a}{b} < \frac{c}{d}$ with $b > 0, d > 0$, prove that $\frac{a+c}{b+d}$ lies between $\frac{a}{b}$ and $\frac{c}{d}$.

Strategy: Prove both inequalities $\frac{a}{b} < \frac{a+c}{b+d}$ and $\frac{a+c}{b+d} < \frac{c}{d}$ by cross-multiplying and using the given condition $\frac{a}{b} < \frac{c}{d}$.

Solution: We need to prove that $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$.

Let's prove both inequalities:

First inequality: $\frac{a}{b} < \frac{a+c}{b+d}$

Cross-multiplying:

$$\begin{aligned} a(b+d) &< b(a+c) \\ ab + ad &< ab + bc \\ ad &< bc \end{aligned}$$

Since $\frac{a}{b} < \frac{c}{d}$, we have $ad < bc$, so this inequality holds.

Second inequality: $\frac{a+c}{b+d} < \frac{c}{d}$

Cross-multiplying:

$$\begin{aligned} d(a+c) &< c(b+d) \\ ad + cd &< bc + cd \\ ad &< bc \end{aligned}$$

Again, since $\frac{a}{b} < \frac{c}{d}$, we have $ad < bc$, so this inequality also holds.

Therefore, $\frac{a+c}{b+d}$ lies between $\frac{a}{b}$ and $\frac{c}{d}$.

Geometric interpretation: This result is known as the "mediant" of two fractions. If we think of fractions as points on a line, the mediant $\frac{a+c}{b+d}$ lies between the two original fractions $\frac{a}{b}$ and $\frac{c}{d}$. ∎

1.13: $\sqrt{2}$ Between Fractions

Let a and b be positive integers. Prove that $\sqrt{2}$ always lies between the two fractions $\frac{a}{b}$ and $\frac{a+2b}{a+b}$. Which fraction is closer to $\sqrt{2}$?

Strategy: Analyze the relationship between the two fractions and $\sqrt{2}$ by examining their difference. Consider two cases based on whether $\frac{a}{b}$ is less than or greater than $\sqrt{2}$, then compare the distances to determine which fraction is closer.

Solution: Let's first establish the ordering of the two fractions by examining their difference:

$$\frac{a+2b}{a+b} - \frac{a}{b} = \frac{b(a+2b) - a(a+b)}{b(a+b)} = \frac{ab + 2b^2 - a^2 - ab}{b(a+b)} = \frac{2b^2 - a^2}{b(a+b)}$$

The sign of this difference depends on the sign of $2b^2 - a^2$, which relates $\frac{a}{b}$ to $\sqrt{2}$.

Case 1: $\frac{a}{b} < \sqrt{2}$. This means $a < b\sqrt{2}$, so $a^2 < 2b^2$, and $2b^2 - a^2 > 0$. Thus, $\frac{a}{b} < \frac{a+2b}{a+b}$. We need to show that $\frac{a+2b}{a+b} > \sqrt{2}$.

$$\begin{aligned} \frac{a+2b}{a+b} > \sqrt{2} &\iff a + 2b > \sqrt{2}(a+b) \\ &\iff b(2-\sqrt{2}) > a(\sqrt{2}-1) \\ &\iff \frac{2-\sqrt{2}}{\sqrt{2}-1} > \frac{a}{b} \end{aligned}$$

Since $\frac{2-\sqrt{2}}{\sqrt{2}-1} = \frac{\sqrt{2}(\sqrt{2}-1)}{\sqrt{2}-1} = \sqrt{2}$, this simplifies to $\sqrt{2} > \frac{a}{b}$, which is true by our case assumption. Thus, $\frac{a}{b} < \sqrt{2} < \frac{a+2b}{a+b}$.

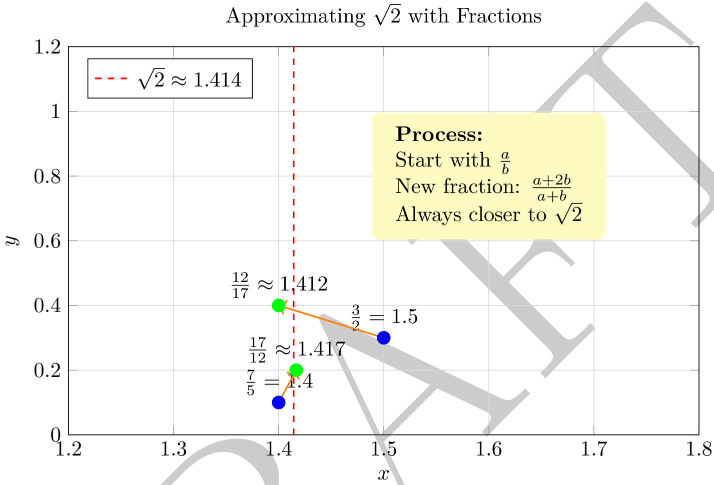
Case 2: $\frac{a}{b} > \sqrt{2}$. This means $a^2 > 2b^2$, and $2b^2 - a^2 < 0$. Thus, $\frac{a}{b} > \frac{a+2b}{a+b}$. A similar calculation shows that $\frac{a+2b}{a+b} < \sqrt{2}$. Therefore, $\frac{a+2b}{a+b} < \sqrt{2} < \frac{a}{b}$. In both cases, $\sqrt{2}$ lies between the two fractions.

Which fraction is closer to $\sqrt{2}$? We compare the absolute distances:

- Distance 1: $\left|\frac{a}{b} - \sqrt{2}\right| = \frac{|a-b\sqrt{2}|}{b}$
- Distance 2: $\left|\frac{a+2b}{a+b} - \sqrt{2}\right| = \left|\frac{a+2b-a\sqrt{2}-b\sqrt{2}}{a+b}\right| = \left|\frac{a(1-\sqrt{2})-b(\sqrt{2}-2)}{a+b}\right| = \frac{|a-b\sqrt{2}|(\sqrt{2}-1)}{a+b}$

To see which distance is smaller, we compare $\frac{1}{b}$ with $\frac{\sqrt{2}-1}{a+b}$. This is equivalent to comparing $a+b$ with $b(\sqrt{2}-1) = b\sqrt{2}-b$, which is equivalent to comparing $a+2b$ with $b\sqrt{2}$, or $\frac{a}{b}+2$ with $\sqrt{2}$. Since a, b are positive integers, $\frac{a}{b} > 0$, so $\frac{a}{b}+2 > 2 > \sqrt{2}$. This implies $\frac{1}{b} > \frac{\sqrt{2}-1}{a+b}$. Therefore, Distance 1 is always greater than Distance 2. The new fraction $\frac{a+2b}{a+b}$ is **always** closer to $\sqrt{2}$.

Visualization:



This visualization shows how the process of generating new fractions $\frac{a+2b}{a+b}$ from $\frac{a}{b}$ always produces a better approximation to $\sqrt{2}$. The red dashed line represents $\sqrt{2}$, and the arrows show the improvement in approximation. ■

1.14: Irrationality of $\sqrt{n-1}+\sqrt{n+1}$

Prove that $\sqrt{n-1}+\sqrt{n+1}$ is irrational for every integer $n \geq 1$.

Strategy: Use proof by contradiction. Assume the sum is rational, then square both sides to eliminate the square roots. This leads to $\sqrt{n^2-1}$ being rational, which is false since n^2-1 is not a perfect square for $n \geq 2$.

Solution: Assume $\sqrt{n-1}+\sqrt{n+1} = \frac{p}{q}$, where p, q are integers, $q \neq 0$, $\gcd(p, q) = 1$.

Square both sides:

$$(n-1) + 2\sqrt{(n-1)(n+1)} + (n+1) = \frac{p^2}{q^2} \implies 2n + 2\sqrt{n^2-1} = \frac{p^2}{q^2}.$$

Thus:

$$\sqrt{n^2-1} = \frac{p^2 - 2nq^2}{2q^2}.$$

Suppose $\sqrt{n^2-1}$ is rational, say $\frac{a}{b}$, $\gcd(a, b) = 1$. Then:

$$n^2 - 1 = \frac{a^2}{b^2} \implies a^2 = (n^2-1)b^2.$$

For $n = 1$, $\sqrt{0}+\sqrt{2} = \sqrt{2}$, irrational. For $n \geq 2$, $n^2-1 = (n-1)(n+1)$ is not a perfect square (since $(n-1)^2 < n^2 - 1 < n^2$). If $a^2 = (n^2-1)b^2$, $n^2 - 1$ must be a perfect square, a contradiction for $n \geq 2$. Thus, $\sqrt{n^2-1}$ is irrational, so $\sqrt{n-1} + \sqrt{n+1}$ is irrational. ■

1.15: Approximation by Rational Numbers

Given a real x and an integer $N > 1$, prove that there exist integers h and k with $0 < k \leq N$ such that $|kx - h| < 1/N$. Hint. Consider the $N+1$ numbers $tx - [tx]$ for $t = 0, 1, 2, \ldots, N$ and show that some pair differs by at most $1/N$.

Strategy: Use the pigeonhole principle. Consider the fractional parts of $0, x, 2x, \ldots, Nx$, divide $[0, 1)$ into N equal subintervals, and apply the pigeonhole principle to find two numbers in the same subinterval.

Solution: We will use the pigeonhole principle to prove this result.

Consider the $N + 1$ numbers: $0, x, 2x, 3x, \ldots, Nx$.

Let's look at the fractional parts of these numbers. The fractional part of a number y is $y - \lfloor y \rfloor$, where $\lfloor y \rfloor$ is the greatest integer less than or equal to y.

The fractional parts of $0, x, 2x, \ldots, Nx$ all lie in the interval $[0, 1)$.

Divide the interval $[0, 1)$ into N equal subintervals:

$$[0, 1/N), [1/N, 2/N), \ldots, [(N-1)/N, 1)$$

By the pigeonhole principle, since we have $N+1$ numbers and only N subintervals, at least two of these numbers must fall into the same subinterval.

Let's say ix and jx (where $0 \le i < j \le N$) have fractional parts in the same subinterval. Then:

$$|(jx - \lfloor jx \rfloor) - (ix - \lfloor ix \rfloor)| < \frac{1}{N}$$
$$|(j-i)x - (\lfloor jx \rfloor - \lfloor ix \rfloor)| < \frac{1}{N}$$

Let $k = j - i$ and $h = \lfloor jx \rfloor - \lfloor ix \rfloor$. Then:

$$|kx - h| < \frac{1}{N}$$

Since $0 < i < j \le N$, we have $0 < k \le N$, and h is an integer.

Therefore, we have found integers h and k with $0 < k \le N$ such that $|kx - h| < 1/N$.

Example: If $x = \pi$ and $N = 5$, we might find that $3\pi \approx 9.4248$ and $5\pi \approx 15.7080$ have fractional parts in the same subinterval, giving us $|2\pi - 6| < 1/5$. ■

1.16: Infinitely Many Rational Approximations

If x is irrational, prove that there are infinitely many rational numbers h/k with $k > 0$ such that $|x - h/k| < 1/k^2$.

Strategy: We will use the result from Problem 1.15 (Dirichlet's Approximation Theorem) to construct rational approximations, then use proof by contradiction to show that there must be infinitely many distinct such approximations.

Solution: We will construct an infinite sequence of distinct rational numbers satisfying the condition.

From Problem 1.15 (Dirichlet's Approximation Theorem), for any integer $N > 1$, there exist integers h and k with $0 < k \le N$ such that $|kx - h| < 1/N$. Dividing by k, we get:

$$\left| x - \frac{h}{k} \right| < \frac{1}{Nk}$$

Since $k \leq N$, we have $\frac{1}{N} \leq \frac{1}{k}$, which implies $\frac{1}{Nk} \leq \frac{1}{k^2}$. Thus, for any integer $N > 1$, we can find a rational number h/k such that:

$$\left| x - \frac{h}{k} \right| < \frac{1}{k^2}$$

Now we must show that this process can generate infinitely many distinct fractions. Assume, for the sake of contradiction, that there are only a finite number of such rational approximations, say

$$\{h_1/k_1, h_2/k_2, \ldots, h_m/k_m\}$$

. Since x is irrational, for any rational number h_i/k_i, the distance $|x - h_i/k_i|$ is non-zero. Let ϵ be the smallest of these non-zero distances:

$$\epsilon = \min_{i=1,\ldots,m} \left| x - \frac{h_i}{k_i} \right| > 0.$$

Now, choose an integer N large enough such that $1/N < \epsilon$. By the result from Problem 1.15, there exist integers h' and k' with $0 < k' \leq N$ such that:

$$|k'x - h'| < \frac{1}{N}$$

This implies $|x - h'/k'| < \frac{1}{Nk'} \leq \frac{1}{N}$. So we have found a new rational approximation h'/k' such that:

$$\left| x - \frac{h'}{k'} \right| < \frac{1}{N} < \epsilon$$

Since the approximation error of h'/k' is smaller than ϵ, h'/k' cannot be one of the fractions in our finite list $\{h_1/k_1, \ldots, h_m/k_m\}$. This contradicts our assumption that we had a complete list of all such approximations. Therefore, there must be infinitely many such rational numbers. ■

1.17: Factorial Representation of Rationals (Precise Form)

Let x be a positive rational number of the form

$$x = \sum_{k=1}^{n} \frac{a_k}{k!},$$

where each a_k is a nonnegative integer with $a_k \leq k-1$ for $k \geq 2$ and $a_n > 0$. Let $[x]$ denote the greatest integer less than or equal to x. Prove that $a_1 = [x]$, that

$$a_k = [k!x] - k[(k-1)!x] \quad \text{for } k = 2, \ldots, n,$$

and that n is the smallest integer such that $n!x$ is an integer. Conversely, show that every positive rational number x can be expressed in this form in one and only one way.

Strategy: For the forward direction, use the properties of the factorial series to show that the fractional part is less than 1, then use the floor function properties. For the converse, use proof by contradiction to show uniqueness by assuming two different representations and finding a contradiction.

Solution: Let $x = \sum_{k=1}^{n} \frac{a_k}{k!}$ with the given conditions on a_k.

1. Proof that $a_1 = [x]$: The sum can be written as $x = a_1 + \sum_{k=2}^{n} \frac{a_k}{k!}$. We must show the summation part is a positive fraction less than 1. Since $a_n > 0$, the sum is positive. We can bound the sum using the property $a_k \leq k-1$:

$$\sum_{k=2}^{n} \frac{a_k}{k!} \leq \sum_{k=2}^{n} \frac{k-1}{k!} < \sum_{k=2}^{\infty} \frac{k-1}{k!}$$

The infinite sum is a known identity: $\sum_{k=2}^{\infty} \frac{k-1}{k!} = \sum_{k=2}^{\infty} \left(\frac{1}{(k-1)!} - \frac{1}{k!}\right)$. This is a telescoping series whose sum is the first term, $1/(2-1)! = 1$. Thus, $0 < \sum_{k=2}^{n} \frac{a_k}{k!} < 1$. This means $a_1 < x < a_1 + 1$, so by definition, $a_1 = [x]$.

2. Formula for a_k: Define $x_1 = x - a_1 = \sum_{k=2}^{n} \frac{a_k}{k!}$. Then $k!x_1$ is an integer for $k \geq n$. Consider the expression $k!x - k((k-1)!x) = k!(a_1 + x_1) - k((k-1)!(a_1 + x_1)) = ka_1k!/k!...$ this gets complicated. Let's use the given formula. Let $x_k = k!x - \sum_{j=1}^{k} a_j \frac{k!}{j!} = \sum_{j=k+1}^{n} a_j \frac{k!}{j!} = \frac{a_{k+1}}{k+1} + \frac{a_{k+2}}{(k+1)(k+2)} + \ldots$. From part (1), we know $0 \leq x_k < 1$. So

$[k!x] = \sum_{j=1}^{k} a_j \frac{k!}{j!}$. Let's test the formula: $a_k = [k!x] - k[(k-1)!x]$. We have $[k!x] = k! \sum_{j=1}^{k} \frac{a_j}{j!}$ and $[(k-1)!x] = (k-1)! \sum_{j=1}^{k-1} \frac{a_j}{j!}$. So, $[k!x] - k[(k-1)!x] = \sum_{j=1}^{k} a_j \frac{k!}{j!} - k \sum_{j=1}^{k-1} a_j \frac{(k-1)!}{j!} = \left(a_k + k \sum_{j=1}^{k-1} a_j \frac{(k-1)!}{j!}\right) - k \sum_{j=1}^{k-1} a_j \frac{(k-1)!}{j!} = a_k$. This proves the formula for a_k.

3. Minimality of n: Multiplying x by $n!$ gives $n!x = \sum_{k=1}^{n} a_k \frac{n!}{k!}$. Since $k \leq n$, each term $\frac{n!}{k!}$ is an integer, so $n!x$ is an integer. For any $m < n$, when we compute $m!x$, the term corresponding to $k = n$ is $m! \frac{a_n}{n!} = \frac{a_n}{n(n-1)\ldots(m+1)}$. Since $0 < a_n \leq n-1$, this term is a non-integer fraction. Because all other terms for $k > m$ are also fractions and terms for $k \leq m$ are integers, $m!x$ cannot be an integer. Thus, n is the smallest such integer.

4. Converse (Uniqueness): Suppose a positive rational number x has two different representations:

$$x = \sum_{k=1}^{n} \frac{a_k}{k!}$$
$$= \sum_{k=1}^{m} \frac{b_k}{k!}$$

with the conditions $0 \leq a_k \leq k-1$ for $k \geq 2$, $a_n > 0$, and similarly for b_k. From part (3), n is the smallest integer such that $n!x$ is an integer, and m is the smallest integer such that $m!x$ is an integer. This implies $n = m$.

Let j be the largest index for which the coefficients differ, so $a_j \neq b_j$. Assume, without loss of generality, that $a_j > b_j$. Since $a_k = b_k$ for $k > j$, we can subtract the sums:

$$\sum_{k=1}^{j} \frac{a_k}{k!} = \sum_{k=1}^{j} \frac{b_k}{k!}$$

Rearranging the terms, we get:

$$\frac{a_j - b_j}{j!} = \sum_{k=1}^{j-1} \frac{b_k - a_k}{k!}$$

Multiply both sides by $(j-1)!$:

$$\frac{a_j - b_j}{j} = \sum_{k=1}^{j-1} (b_k - a_k) \frac{(j-1)!}{k!}$$

The right-hand side is an integer, because for each $k \in \{1, \ldots, j-1\}$, $k!$ divides $(j-1)!$. Let's analyze the left-hand side. Since a_j and b_j are integers and $a_j > b_j$, we have $a_j - b_j \geq 1$. From the conditions on the coefficients, $a_j \leq j-1$ (for $j \geq 2$) and $b_j \geq 0$. Therefore, $1 \leq a_j - b_j \leq j-1$. This implies that for $j \geq 2$, the left-hand side $\frac{a_j - b_j}{j}$ is a non-integer fraction, since the numerator is an integer between 1 and $j-1$, and the denominator is j. This creates a contradiction: the left-hand side cannot be an integer, while the right-hand side must be an integer. For the case $j = 1$, the equation becomes $a_1 - b_1 = 0$, which contradicts $a_1 \neq b_1$. Thus, our assumption that there is a largest index j where $a_j \neq b_j$ must be false. All coefficients must be identical. The representation is unique.

5. Uniqueness: Suppose x has two different representations, $\sum \frac{a_k}{k!} = \sum \frac{b_k}{k!}$. Let j be the largest index where $a_j \neq b_j$. Assume $a_j > b_j$. Then $\frac{a_j - b_j}{j!} = \sum_{k=1}^{j-1} \frac{b_k - a_k}{k!}$. The left side is $\geq 1/j!$. The right side is bounded above by $\sum_{k=1}^{j-1} \frac{k-1}{k!} < 1/j!$, a contradiction. Thus, all coefficients must be the same. ■

1.III Upper Bounds

Definitions and Theorems

Theorem: Completeness Axiom

Every nonempty set of real numbers that is bounded above has a supremum.

Importance: The Completeness Axiom is the fundamental property that distinguishes the real numbers from the rational numbers. It ensures that the real number system has no "gaps" and is essential for all of calculus and analysis. This axiom is the foundation for proving the existence of limits, continuity, and many other important results in mathematics.

Theorem: Uniqueness of Supremum and Infimum

If a set has a supremum (infimum), it is unique.

Importance: This theorem ensures that suprema and infima are well-defined concepts. Without uniqueness, we couldn't meaningfully talk about "the" supremum or "the" infimum of a set. This result is essential for the logical consistency of real analysis and is used implicitly in many proofs throughout mathematics.

Theorem: Archimedean Property

For any positive real numbers a and b, there exists a positive integer n such that $na > b$.

Theorem: Comparison Property for Suprema

Let S and T be nonempty subsets of $\mathbb{R}$ such that $s \leq t$ for all $s \in S$ and $t \in T$. If T has a supremum, then S has a supremum and $\sup S \leq \sup T$.

Importance: This theorem provides a powerful tool for comparing suprema of different sets. It's essential for many proofs in analysis where we need to establish inequalities between suprema. This result is particularly useful in functional analysis, measure theory, and optimization theory where we often need to compare bounds of different sets or functions.

1.18: Uniqueness of Supremum and Infimum

Show that the sup and inf of a set are uniquely determined whenever they exist.

Strategy: We will prove the uniqueness of supremum by contradiction, assuming there are two different suprema and showing this leads to a contradiction. The same approach applies to infimum.

Solution: We will prove that if a set has a supremum, it is unique. The proof for infimum is similar.

Proof by contradiction: Suppose a set S has two different suprema, say s_1 and s_2, with $s_1 < s_2$.

Since s_1 is a supremum of S: 1. s_1 is an upper bound of S (every element of S is $\leq s_1$) 2. s_1 is the least upper bound (no number less than s_1 is an upper bound)

Since s_2 is also a supremum of S: 1. s_2 is an upper bound of S (every element of S is $\leq s_2$) 2. s_2 is the least upper bound (no number less than s_2 is an upper bound)

But since $s_1 < s_2$, the number s_1 is less than s_2 and is also an upper bound of S. This contradicts the fact that s_2 is the least upper bound.

Therefore, our assumption that there are two different suprema is false, and the supremum must be unique.

Alternative proof: Let s_1 and s_2 both be suprema of S. Then: - s_1 is an upper bound, so $s_2 \leq s_1$ (since s_2 is the least upper bound) - s_2 is an upper bound, so $s_1 \leq s_2$ (since s_1 is the least upper bound)

Therefore, $s_1 = s_2$.

For infimum: The same argument applies to infimum. If a set has two infima i_1 and i_2, then: - i_1 is a lower bound, so $i_1 \leq i_2$ (since i_1 is the greatest lower bound) - i_2 is a lower bound, so $i_2 \leq i_1$ (since i_2 is the greatest lower bound)

Therefore, $i_1 = i_2$. ■

1.19: Finding Supremum and Infimum

Find the sup and inf of each of the following sets:

(a) All numbers of the form $2^{-p} + 3^{-q} + 5^{-r}$ for positive integers p, q, r.

(b) $S = \{x : 3x^2 - 10x + 3 < 0\}$.

(c) $S = \{x : (x-a)(x-b)(x-c)(x-d) < 0\}$ where $a < b < c < d$.

Strategy: For (a), analyze the range of each exponential term and find the maximum and minimum values. For (b), solve the quadratic inequality to find the interval. For (c), use the sign changes of the polynomial at its roots to determine the intervals where the product is negative.

Solution:

1. Numbers of the form $2^{-p} + 3^{-q} + 5^{-r}$:

Let's analyze the range of each term: - 2^{-p} ranges from $\frac{1}{2}$ (when $p = 1$) to 0 (as $p \to \infty$) - 3^{-q} ranges from $\frac{1}{3}$ (when $q = 1$) to 0 (as $q \to \infty$) - 5^{-r} ranges from $\frac{1}{5}$ (when $r = 1$) to 0 (as $r \to \infty$)

Therefore: - The maximum value occurs when $p = q = r = 1$: $\frac{1}{2} + \frac{1}{3} + \frac{1}{5} = \frac{15+10+6}{30} = \frac{31}{30}$ - The minimum value occurs as $p, q, r \to \infty$: $0 + 0 + 0 = 0$

So $\sup = \frac{31}{30}$ and $\inf = 0$.

2. Set $S = \{x : 3x^2 - 10x + 3 < 0\}$:

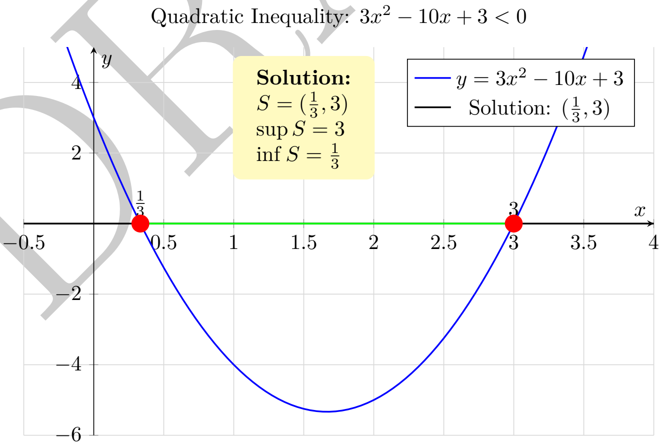
First, let's find the roots of $3x^2 - 10x + 3 = 0$:

$$\begin{aligned} x &= \frac{10 \pm \sqrt{100 - 36}}{6} \\ &= \frac{10 \pm \sqrt{64}}{6} \\ &= \frac{10 \pm 8}{6} \\ &= \frac{18}{6} = 3 \quad \text{or} \quad \frac{2}{6} = \frac{1}{3} \end{aligned}$$

Since the coefficient of x^2 is positive, the parabola opens upward. The inequality $3x^2 - 10x + 3 < 0$ holds between the roots.

Therefore, $S = (\frac{1}{3}, 3)$, so $\sup = 3$ and $\inf = \frac{1}{3}$.

Visualization for part (b):



This visualization shows the quadratic function $y = 3x^2 - 10x + 3$ and highlights the interval where the inequality $3x^2 - 10x + 3 < 0$ holds, which is between the roots $\frac{1}{3}$ and 3.

3. Set $S = \{x : (x-a)(x-b)(x-c)(x-d) < 0\}$ where $a < b < c < d$:

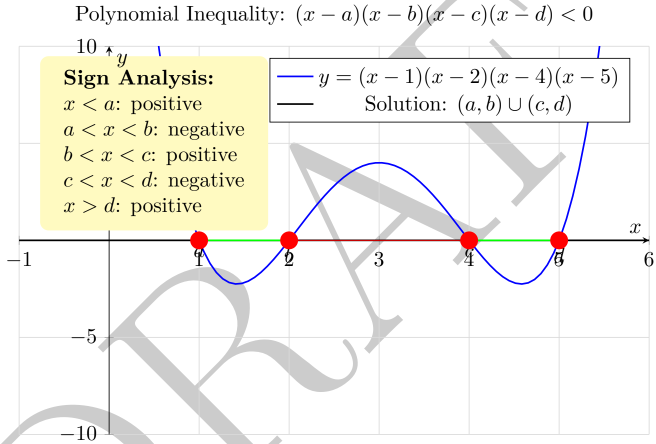
The expression $(x-a)(x-b)(x-c)(x-d)$ changes sign at each root a, b, c, d.

Starting from $-\infty$: - For $x < a$: all factors are negative, so the product is positive - For $a < x < b$: one factor is positive, three negative, so product is negative - For $b < x < c$: two factors positive, two negative, so product is positive - For $c < x < d$: three factors positive, one negative, so product is negative - For $x > d$: all factors are positive, so product is positive

Therefore, $S = (a, b) \cup (c, d)$.

So sup $= d$ and inf $= a$.

Visualization for part (c):



This visualization shows the fourth-degree polynomial $(x-a)(x-b)(x-c)(x-d)$ and highlights the intervals where the inequality $(x-a)(x-b)(x-c)(x-d) < 0$ holds. The polynomial changes sign at each root, creating alternating positive and negative intervals. ■

1.20: Comparison Property for Suprema

Let S and T be nonempty subsets of $\mathbb{R}$ such that $s \leq t$ for all $s \in S$ and $t \in T$. Suppose T has a supremum. Then S has a supremum and

$$\sup S \leq \sup T.$$

Strategy: We will first show that S has a supremum by using the completeness axiom, then prove that $\sup S \leq \sup T$ by showing that $\sup S$ is a lower bound for T and using the definition of supremum.

Solution: Let S and T be nonempty subsets of $\mathbb{R}$ with the property that for every $s \in S$ and $t \in T$, we have $s \leq t$.

1. **Existence of** $\sup S$: Since T is nonempty, we can pick an arbitrary element $t_0 \in T$. By the given property, for every $s \in S$, we have $s \leq t_0$. This shows that S is bounded above (by any element of T). Since S is also nonempty and bounded above, the completeness axiom of $\mathbb{R}$ guarantees that $\sup S$ exists. Let's call it $\alpha = \sup S$.

2. **Proof that** $\sup S \leq \sup T$: Let $\alpha = \sup S$ and $\beta = \sup T$. From step 1, we know that any element $t \in T$ is an upper bound for the set S. Since α is the *least* upper bound of S, it must be less than or equal to any other upper bound of S. Therefore, for any $t \in T$, we must have:

$$\alpha \leq t$$

This inequality shows that α is a lower bound for the set T. Now, by definition, $\beta = \sup T$ is the least upper bound of T. As an upper bound for T, β must be greater than or equal to every element of T. More importantly, it must be greater than or equal to any *lower bound* of T. Since we have established that α is a lower bound for T, it must follow that:

$$\alpha \leq \beta$$

Substituting the definitions of α and β, we get:

$$\sup S \leq \sup T$$

This completes the proof. ■

1.21: Product of Suprema

Let A and B be two sets of positive real numbers, each bounded above. Let $a = \sup A$, $b = \sup B$. Define

$$C = \{xy : x \in A,\, y \in B\}.$$

Prove that

$$\sup C = ab.$$

Strategy: We will show that ab is an upper bound for C, then prove it is the least upper bound by using the definition of supremum and constructing elements of C that are arbitrarily close to ab.

Solution:

Since A and B are sets of positive real numbers bounded above, their suprema $a = \sup A$ and $b = \sup B$ exist and are finite.

We are to prove that:

$$\sup C = ab.$$

Step 1: Show that ab is an upper bound for C.

Let $x \in A$, $y \in B$. Since $x \leq a$ and $y \leq b$, we have:

$$xy \leq ab.$$

Therefore, every element $c \in C$ satisfies $c \leq ab$, so ab is an upper bound for C.

Step 2: Show that ab is the least upper bound.

Let $\varepsilon > 0$. Since $a = \sup A$, there exists $x_\varepsilon \in A$ such that:

$$x_\varepsilon > a - \frac{\varepsilon}{2b}.$$

Similarly, since $b = \sup B$, there exists $y_\varepsilon \in B$ such that:

$$y_\varepsilon > b - \frac{\varepsilon}{2a}.$$

Now consider:

$$x_\varepsilon y_\varepsilon > \left(a - \frac{\varepsilon}{2b}\right)\left(b - \frac{\varepsilon}{2a}\right) = ab - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} + \frac{\varepsilon^2}{4ab}.$$

Since $\frac{\varepsilon^2}{4ab} > 0$, we have:

$$x_\varepsilon y_\varepsilon > ab - \varepsilon.$$

Therefore, for every $\varepsilon > 0$, there exists $c \in C$ such that $c > ab - \varepsilon$. Hence, ab is the least upper bound of C.

$$\boxed{\sup C = ab}$$

■

1.22: Representation of Rationals in Base k

Given $x \geq 0$ and an integer $k \geq 2$, let a_0 denote the largest integer $\leq x$, and, assuming that $a_0, a_1, \ldots, a_{n-1}$ have been defined, let a_n denote the largest integer such that

$$a_0 + \frac{a_1}{k} + \frac{a_2}{k^2} + \cdots + \frac{a_n}{k^n} \leq x.$$

(a) Prove that $0 \leq a_i \leq k - 1$ for each $i = 1, 2, \ldots$.

(b) Let $r_n = a_0 + a_1 k^{-1} + a_2 k^{-2} + \cdots + a_n k^{-n}$ and show that $x = \sup\{r_n\}$, the supremum of the set of rational numbers $r_1, r_2, \ldots$.

Strategy: For (a), use the definition of a_n as the largest integer satisfying the condition and show that choosing $a_n + 1$ would violate it. For (b), use the Archimedean property and proof by contradiction to show that the supremum equals x.

Solution: Let $r_n = \sum_{i=0}^{n} \frac{a_i}{k^i}$. By definition, a_n is the largest integer such that $r_n \leq x$.

(a) Show $0 \leq a_i \leq k - 1$: Since a_n is the largest integer satisfying the condition, choosing $a_n + 1$ would violate it:

$$r_{n-1} + \frac{a_n + 1}{k^n} > x$$

From the definition of a_{n-1}, we know it was the largest integer such that $r_{n-1} \leq x$. This implies $x - r_{n-1} < \frac{1}{k^{n-1}}$. Now, from the definition of a_n, we have $r_{n-1} + \frac{a_n}{k^n} \leq x$, which implies $a_n \leq k^n(x - r_{n-1})$. Combining these facts:

$$a_n \leq k^n(x - r_{n-1}) < k^n \left(\frac{1}{k^{n-1}} \right) = k.$$

Since a_n is an integer and $a_n < k$, we must have $a_n \leq k-1$. Also, a_n must be non-negative, otherwise we could choose $a_n = 0$ to get a larger (or equal) sum r_n that is still less than or equal to x, contradicting the "largest integer" definition if the original a_n were negative. Thus, $0 \leq a_n \leq k-1$.

(b) Show that $x = \sup\{r_n\}$: The sequence $\{r_n\}$ is non-decreasing by construction, since $a_n \geq 0$. It is also bounded above by x. Therefore, its supremum exists; let $r = \sup\{r_n\}$. We know $r \leq x$. We will prove $r = x$ by contradiction. Assume $r < x$. Let $\delta = x - r > 0$. By the Archimedean property, we can choose an integer N large enough such that $\frac{1}{k^N} < \delta$. From the definition of a_N, we know $r_N = r_{N-1} + \frac{a_N}{k^N} \leq x$ and $r_{N-1} + \frac{a_N+1}{k^N} > x$. The second inequality rearranges to $x - r_N < \frac{1}{k^N}$. Since $r = \sup\{r_n\}$, we know $r_N \leq r$. Therefore, $x - r \leq x - r_N < \frac{1}{k^N}$. Substituting $\delta = x - r$, we get $\delta < \frac{1}{k^N}$. But we chose N such that $\frac{1}{k^N} < \delta$. This gives $\delta < \frac{1}{k^N} < \delta$, a contradiction. Thus, our assumption must be false, and $x = r = \sup\{r_n\}$. ∎

1.IV Inequalities and Identities

Additional Theorems for Inequalities and Identities

Theorem: Binomial Theorem

For any real numbers a and b and positive integer n:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

Importance: The Binomial Theorem is one of the most fundamental results in algebra and combinatorics. It provides a formula for expanding powers of binomials and is essential for understanding polynomials, probability theory, and many areas of mathematics. The binomial coefficients that appear in this theorem are fundamental in combinatorics and have applications throughout mathematics.

Theorem: Pigeonhole Principle

If n objects are placed into m containers where $n > m$, then at least one container must contain more than one object.

Importance: The Pigeonhole Principle is a simple but powerful tool in combinatorics and discrete mathematics. It provides a way to prove the existence of certain patterns or properties without explicitly constructing them. This principle is essential for many existence proofs and has applications in computer science, number theory, and many other areas.

Theorem: Dirichlet's Approximation Theorem

For any real number x and positive integer N, there exist integers h and k with $0 < k \leq N$ such that $|kx - h| < 1/N$.

Importance: Dirichlet's Approximation Theorem is a fundamental result in Diophantine approximation that shows how well real numbers can be approximated by rational numbers. It's essential for understanding the distribution of rational numbers and has applications in number theory, analysis, and many areas of mathematics. This theorem is the foundation for more advanced results in Diophantine approximation.

1.23: Lagrange's Identity

Prove Lagrange's identity for real numbers:

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 = \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right) - \sum_{1\leq k<j\leq n} (a_k b_j - a_j b_k)^2.$$

Strategy: We will prove this identity by expanding both sides and showing they are equal. We'll expand the left-hand side as a double sum and the right-hand side by expanding the product and the squared terms, then show that all terms cancel appropriately.

Solution: We will prove Lagrange's identity by expanding both sides and showing they are equal.

Let's start by expanding the left-hand side:

$$\begin{aligned}\left(\sum_{k=1}^{n} a_k b_k\right)^2 &= \left(\sum_{k=1}^{n} a_k b_k\right)\left(\sum_{j=1}^{n} a_j b_j\right)\\ &= \sum_{k=1}^{n}\sum_{j=1}^{n} a_k b_k a_j b_j\\ &= \sum_{k=1}^{n} a_k^2 b_k^2 + 2\sum_{1\le k<j\le n} a_k b_k a_j b_j\end{aligned}$$

Now let's expand the right-hand side:

$$\begin{aligned}&\left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right) - \sum_{1\le k<j\le n} (a_k b_j - a_j b_k)^2\\ &= \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{j=1}^{n} b_j^2\right) - \sum_{1\le k<j\le n} (a_k^2 b_j^2 - 2a_k b_j a_j b_k + a_j^2 b_k^2)\\ &= \sum_{k=1}^{n}\sum_{j=1}^{n} a_k^2 b_j^2 - \sum_{1\le k<j\le n} a_k^2 b_j^2 + 2\sum_{1\le k<j\le n} a_k b_j a_j b_k - \sum_{1\le k<j\le n} a_j^2 b_k^2\end{aligned}$$

Let's simplify this step by step. First, note that:

$$\begin{aligned}\sum_{k=1}^{n}\sum_{j=1}^{n} a_k^2 b_j^2 &= \sum_{k=1}^{n} a_k^2 b_k^2 + \sum_{1\le k<j\le n} a_k^2 b_j^2 + \sum_{1\le k<j\le n} a_j^2 b_k^2\\ &= \sum_{k=1}^{n} a_k^2 b_k^2 + \sum_{1\le k<j\le n} (a_k^2 b_j^2 + a_j^2 b_k^2)\end{aligned}$$

Substituting this back into our expression:

$$\begin{aligned}&\sum_{k=1}^{n} a_k^2 b_k^2 + \sum_{1\le k<j\le n} (a_k^2 b_j^2 + a_j^2 b_k^2) - \sum_{1\le k<j\le n} a_k^2 b_j^2 + 2\sum_{1\le k<j\le n} a_k b_j a_j b_k\\ &\quad - \sum_{1\le k<j\le n} a_j^2 b_k^2\\ &= \sum_{k=1}^{n} a_k^2 b_k^2 + 2\sum_{1\le k<j\le n} a_k b_j a_j b_k\end{aligned}$$

This is exactly the same as our expanded left-hand side! Therefore, Lagrange's identity is proven.

Alternative Proof using Determinants: We can also prove this using the fact that the determinant of a 2×2 matrix is zero if and only if its rows are linearly dependent.

Consider the matrix:

$$\begin{pmatrix} a_k & b_k \\ a_j & b_j \end{pmatrix}$$

The determinant of this matrix is $a_k b_j - a_j b_k$. If we square this determinant and sum over all pairs (k, j) with $k < j$, we get the right-hand side of Lagrange's identity.

The left-hand side represents the square of the dot product of the vectors $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$.

The identity shows that the square of the dot product equals the product of the squared magnitudes minus the sum of squared determinants of all 2×2 submatrices formed by pairs of components. ■

1.24: A Holder-type Inequality

Prove that for arbitrary real numbers a_k, b_k, c_k we have

$$\left(\sum_{k=1}^{n} a_k b_k c_k\right)^4 \leq \left(\sum_{k=1}^{n} a_k^4\right)\left(\sum_{k=1}^{n} b_k^2\right)^2\left(\sum_{k=1}^{n} c_k^4\right).$$

Strategy: We will apply the Cauchy-Schwarz inequality twice. First, we'll group $(a_k c_k)$ and b_k, then apply Cauchy-Schwarz to the resulting term $\sum_{k=1}^{n} a_k^2 c_k^2$ by treating it as the dot product of sequences $\{a_k^2\}$ and $\{c_k^2\}$.

Solution: We will prove this inequality by applying the Cauchy-Schwarz inequality twice. First, group the terms as $(a_k c_k)$ and b_k. Applying the Cauchy-Schwarz inequality to the sequences $\{a_k c_k\}$ and $\{b_k\}$ gives:

$$\left(\sum_{k=1}^{n} (a_k c_k) b_k\right)^2 \leq \left(\sum_{k=1}^{n} (a_k c_k)^2\right)\left(\sum_{k=1}^{n} b_k^2\right) = \left(\sum_{k=1}^{n} a_k^2 c_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right).$$

Next, we apply the Cauchy-Schwarz inequality to the term $\sum_{k=1}^{n} a_k^2 c_k^2$, treating it as the dot product of sequences $\{a_k^2\}$ and $\{c_k^2\}$:

$$\left(\sum_{k=1}^{n} a_k^2 c_k^2\right)^2 \leq \left(\sum_{k=1}^{n} (a_k^2)^2\right)\left(\sum_{k=1}^{n} (c_k^2)^2\right) = \left(\sum_{k=1}^{n} a_k^4\right)\left(\sum_{k=1}^{n} c_k^4\right).$$

This implies:

$$\sum_{k=1}^{n} a_k^2 c_k^2 \leq \left(\sum_{k=1}^{n} a_k^4\right)^{1/2} \left(\sum_{k=1}^{n} c_k^4\right)^{1/2}.$$

Now, substitute this result back into our first inequality:

$$\left(\sum_{k=1}^{n} a_k b_k c_k\right)^2 \leq \left(\sum_{k=1}^{n} a_k^4\right)^{1/2} \left(\sum_{k=1}^{n} c_k^4\right)^{1/2} \left(\sum_{k=1}^{n} b_k^2\right).$$

Finally, squaring both sides gives the desired result:

$$\left(\sum_{k=1}^{n} a_k b_k c_k\right)^4 \leq \left(\sum_{k=1}^{n} a_k^4\right) \left(\sum_{k=1}^{n} c_k^4\right) \left(\sum_{k=1}^{n} b_k^2\right)^2.$$

■

1.25: Minkowski's Inequality

Prove Minkowski's inequality:

$$\left(\sum_{k=1}^{n} (a_k + b_k)^2\right)^{1/2} \leq \left(\sum_{k=1}^{n} a_k^2\right)^{1/2} + \left(\sum_{k=1}^{n} b_k^2\right)^{1/2}.$$

Strategy: We will expand the left-hand side, apply the Cauchy-Schwarz inequality to the cross term, then complete the square to obtain the desired inequality.

Solution:

Let $A = \left(\sum a_k^2\right)^{1/2}$, $B = \left(\sum b_k^2\right)^{1/2}$, and expand the square:

$$\sum (a_k + b_k)^2 = \sum a_k^2 + 2\sum a_k b_k + \sum b_k^2 = A^2 + 2\sum a_k b_k + B^2.$$

Apply Cauchy–Schwarz:

$$\sum a_k b_k \leq AB.$$

Thus,

$$\sum (a_k + b_k)^2 \leq A^2 + 2AB + B^2 = (A + B)^2.$$

Taking square roots:

$$\left(\sum (a_k + b_k)^2\right)^{1/2} \leq A + B.$$

■

1.26: Chebyshev's Sum Inequality

If $a_1 \geq a_2 \geq \cdots \geq a_n$ and $b_1 \geq b_2 \geq \cdots \geq b_n$, prove that

$$\left(\sum_{k=1}^{n} a_k\right)\left(\sum_{k=1}^{n} b_k\right) \leq n \sum_{k=1}^{n} a_k b_k.$$

Strategy: We will consider the double summation $S = \sum_{i=1}^{n}\sum_{j=1}^{n}(a_i - a_j)(b_i - b_j)$ and show that it is non-negative due to the ordering of the sequences, then expand it to obtain the desired inequality.

Solution: Consider the double summation

$$S = \sum_{i=1}^{n}\sum_{j=1}^{n}(a_i - a_j)(b_i - b_j).$$

Since the sequences $\{a_k\}$ and $\{b_k\}$ are sorted in the same order (both non-increasing), the terms $(a_i - a_j)$ and $(b_i - b_j)$ always have the same sign. If $i > j$, then $a_i \leq a_j$ and $b_i \leq b_j$, so both differences are non-positive. If $i < j$, both are non-negative. Therefore, their product is always non-negative:

$$(a_i - a_j)(b_i - b_j) \geq 0.$$

This implies that the total sum S must be non-negative, $S \geq 0$.

Now, let's expand the sum:

$$\begin{aligned} S &= \sum_{i=1}^{n}\sum_{j=1}^{n}(a_i b_i - a_i b_j - a_j b_i + a_j b_j) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_i - \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j - \sum_{i=1}^{n}\sum_{j=1}^{n} a_j b_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_j b_j \end{aligned}$$

We evaluate each double summation:

- $\sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_i = \sum_{i=1}^{n} (n \cdot a_i b_i) = n \sum_{i=1}^{n} a_i b_i$
- $\sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j = (\sum_{i=1}^{n} a_i) \left(\sum_{j=1}^{n} b_j\right)$
- $\sum_{i=1}^{n} \sum_{j=1}^{n} a_j b_i = \left(\sum_{j=1}^{n} a_j\right) (\sum_{i=1}^{n} b_i)$
- $\sum_{i=1}^{n} \sum_{j=1}^{n} a_j b_j = \sum_{j=1}^{n} (n \cdot a_j b_j) = n \sum_{j=1}^{n} a_j b_j$

Substituting these back into the expression for S:

$$S = n \sum a_k b_k - \left(\sum a_k\right) \left(\sum b_k\right) - \left(\sum a_k\right) \left(\sum b_k\right) + n \sum a_k b_k$$

$$S = 2n \sum_{k=1}^{n} a_k b_k - 2 \left(\sum_{k=1}^{n} a_k\right) \left(\sum_{k=1}^{n} b_k\right)$$

Since we established that $S \geq 0$:

$$2n \sum_{k=1}^{n} a_k b_k - 2 \left(\sum_{k=1}^{n} a_k\right) \left(\sum_{k=1}^{n} b_k\right) \geq 0$$

Dividing by 2 and rearranging gives the desired inequality:

$$n \sum_{k=1}^{n} a_k b_k \geq \left(\sum_{k=1}^{n} a_k\right) \left(\sum_{k=1}^{n} b_k\right).$$

■

1.V Complex Numbers

1.27: Express Complex Numbers in $a + bi$ Form

Express the following complex numbers in the form $a + bi$:

(a) $(1+i)^3$

(b) $\frac{2+3i}{3-4i}$

(c) $i^5 + i^{16}$

(d) $\frac{1}{2}(1+i)(1+i^{-8})$

Strategy: We will use the properties of complex numbers, including $i^2 = -1$, $i^4 = 1$, and the fact that powers of i cycle every 4. For division, we'll rationalize the denominator by multiplying by the complex conjugate.

Solution:

(a) $(1+i)^3 = (1+i)^2(1+i) = (2i)(1+i) = 2i+2i^2 = 2i-2 = -2+2i$

(b) Rationalize the denominator:

$$\frac{2+3i}{3-4i} \cdot \frac{3+4i}{3+4i} = \frac{(2+3i)(3+4i)}{9+16} = \frac{6+8i+9i+12i^2}{25}$$
$$= \frac{-6+17i}{25} = -\frac{6}{25} + \frac{17}{25}i$$

(c) $i^5 = i$, since $i^4 = 1$, and $i^{16} = (i^4)^4 = 1$, so:

$$i^5 + i^{16} = i + 1 = 1 + i$$

(d) $\frac{1}{2}(1+i)(1+i^{-8})$, note that $i^{-8} = (i^4)^{-2} = 1^{-2} = 1$, so:

$$\frac{1}{2}(1+i)(1+1) = \frac{1}{2}(1+i)(2) = \frac{1}{2}(2+2i) = 1+i$$

■

1.28: Solve Complex Equations

In each case, determine all real x and y which satisfy the given relation:

(a) $x + iy = |x - iy|$

(b) $x + iy = (x - iy)^2$

(c) $\sum_{k=0}^{100} i^k = x + iy$

Strategy: For each equation, we'll equate the real and imaginary parts. For (a), we'll use the fact that the right-hand side is real and nonnegative. For (b), we'll expand the square and solve the resulting system. For (c), we'll use the cyclic nature of powers of i.

Solution:

(a) RHS is real and nonnegative. LHS is complex. For equality, imaginary part must vanish:

$$\mathrm{Im}(x+iy) = y = 0, \quad \text{and } x = |x| \Rightarrow x \geq 0.$$

So solution: $y = 0$, $x \geq 0$

(b) Compute RHS:

$$(x - iy)^2 = x^2 - 2ixy - y^2 = (x^2 - y^2) - 2ixy.$$

Set equal to $x + iy$, equate real and imaginary parts:

$$x = x^2 - y^2, \quad y = -2xy.$$

From second equation: $y = -2xy \Rightarrow y(1 + 2x) = 0 \Rightarrow y = 0$ or $x = -\frac{1}{2}$

If $y = 0$, then first equation: $x = x^2 \Rightarrow x(x-1) = 0 \Rightarrow x = 0$ or $x = 1$

If $x = -\frac{1}{2}$, then first equation:

$$x = x^2 - y^2 \Rightarrow -\frac{1}{2} = \frac{1}{4} - y^2 \Rightarrow y^2 = \frac{3}{4} \Rightarrow y = \pm\frac{\sqrt{3}}{2}$$

So all solutions:

$$(x, y) = (0, 0), (1, 0), \left(-\frac{1}{2}, \pm\frac{\sqrt{3}}{2}\right)$$

(c) The powers of i cycle every 4: $i^0 = 1, i^1 = i, i^2 = -1, i^3 = -i$

There are 101 terms, which form 25 full cycles and one leftover term $i^{100} \equiv i^0 = 1$

Each full cycle sums to 0. So total sum:

$$\sum_{k=0}^{100} i^k = 25 \cdot 0 + 1 = 1 \Rightarrow x = 1, y = 0.$$

■

1.29: Basic Identities for Complex Conjugates

If $z = x + iy$, where x and y are real, the complex conjugate of z is $\overline{z} = x - iy$. Prove the following:

(a) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$,

(b) $\overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}$,

(c) $z \cdot \overline{z} = |z|^2$,

(d) $z + \overline{z}$ is twice the real part of z,

(e) $\frac{z-\overline{z}}{i}$ is twice the imaginary part of z.

Strategy: We will prove each identity by using the definition of complex conjugate and performing the necessary algebraic manipulations. For each part, we'll work with the real and imaginary components explicitly.

Solution: Let $z = x + iy$ and $w = u + iv$ be two complex numbers.

(a) **Conjugate of a sum:**

$$\overline{z_1 + z_2} = \overline{(x_1 + x_2) + i(y_1 + y_2)} = (x_1 + x_2) - i(y_1 + y_2) = \overline{z_1} + \overline{z_2}.$$

(b) **Conjugate of a product:**

$$\begin{aligned} \overline{z_1 z_2} &= \overline{(x_1 + iy_1)(x_2 + iy_2)} = \overline{(x_1 x_2 - y_1 y_2 + i(x_1 y_2 + x_2 y_1))} \\ &= x_1 x_2 - y_1 y_2 - i(x_1 y_2 + x_2 y_1) = \overline{z_1} \cdot \overline{z_2}. \end{aligned}$$

(c) **Modulus squared:**

$$z \cdot \overline{z} = (x + iy)(x - iy) = x^2 + y^2 = |z|^2.$$

(d) **Twice the real part:**

$$z + \overline{z} = (x + iy) + (x - iy) = 2x = 2\Re(z).$$

(e) **Twice the imaginary part:**

$$\frac{z - \overline{z}}{i} = \frac{(x + iy) - (x - iy)}{i} = \frac{2iy}{i} = 2y = 2\Im(z).$$

■

1.30: Geometric Descriptions of Complex Sets

Describe geometrically the set of complex numbers z which satisfies each of the following conditions:

(a) $|z| = 1$,

(b) $|z| < 1$,

(c) $|z| \leq 1$,

(d) $z + \overline{z} = 1$,

(e) $z - \overline{z} = i$,

(f) $\overline{z} + z = |z|^2$.

Strategy: We will use the properties of complex conjugates and the relationship between complex numbers and their real/imaginary parts to translate each condition into geometric terms. For the last condition, we'll complete the square to identify the geometric shape.

Solution:

(a) The unit circle centered at the origin.

(b) The open unit disk centered at the origin.

(c) The closed unit disk centered at the origin.

(d) $2\Re(z) = 1 \Rightarrow \Re(z) = \frac{1}{2}$: a vertical line in the complex plane.

(e) $2i\Im(z) = i \Rightarrow \Im(z) = \frac{1}{2}$: a horizontal line.

(f) Let $z = x + iy$, where $x, y \in \mathbb{R}$. Then:

$$\begin{aligned} z + \overline{z} &= (x + iy) + (x - iy) = 2x, \\ |z|^2 &= x^2 + y^2. \end{aligned}$$

So the equation becomes:

$$2x = x^2 + y^2.$$

Rewriting this:

$$x^2 - 2x + y^2 = 0.$$

We now complete the square on the x-terms:

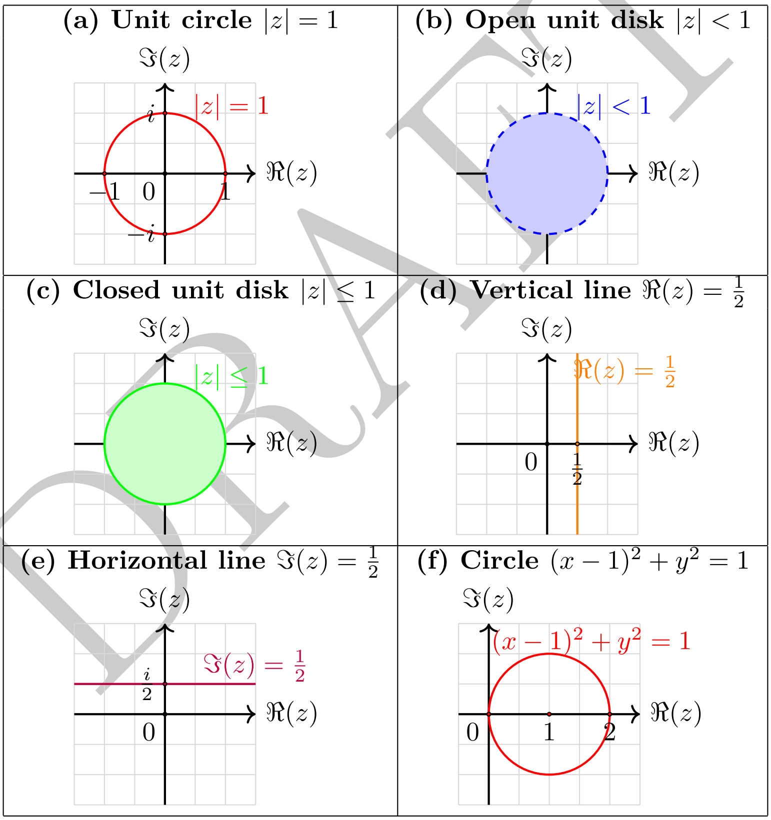
$$x^2 - 2x = (x-1)^2 - 1,$$

which gives:

$$(x-1)^2 - 1 + y^2 = 0 \quad \Rightarrow \quad (x-1)^2 + y^2 = 1.$$

This is the standard equation of a circle with center at $(1, 0)$ and radius 1 in the complex plane.

Visualizations:



■

1.31: Equilateral Triangle on the Unit Circle

Given three complex numbers z_1, z_2, z_3 such that $|z_1| = |z_2| = |z_3| = 1$ and $z_1 + z_2 + z_3 = 0$, show that these numbers are vertices of an equilateral triangle inscribed in the unit circle with center at the origin.

Strategy: Use the fact that the sum of three unit complex numbers equals zero to show they must be the cube roots of unity (rotated), which form an equilateral triangle. Verify that the angles differ by $2\pi/3$ and the sum condition is satisfied.

Solution: Since $|z_i| = 1$, each $z_i = e^{i\theta_i}$ lies on the unit circle. Given $z_1 + z_2 + z_3 = 0$, we need to show they form an equilateral triangle. Consider the angles $\theta_1, \theta_2, \theta_3$. The sum condition implies:

$$e^{i\theta_1} + e^{i\theta_2} + e^{i\theta_3} = 0.$$

For three points on the unit circle to form an equilateral triangle, their arguments must differ by $120^\circ = \frac{2\pi}{3}$. Assume:

$$z_1 = e^{i\theta}, \quad z_2 = e^{i(\theta + \frac{2\pi}{3})}, \quad z_3 = e^{i(\theta + \frac{4\pi}{3})}.$$

Check the sum:

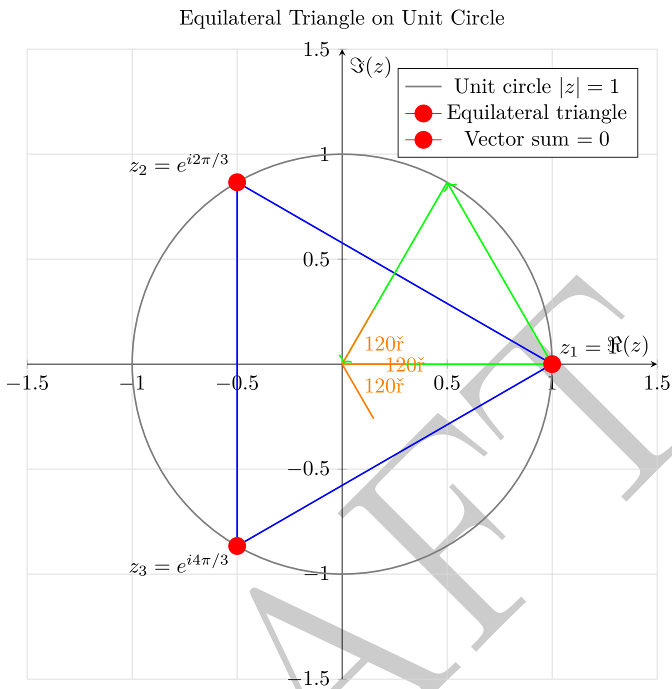
$$e^{i\theta} + e^{i(\theta + \frac{2\pi}{3})} + e^{i(\theta + \frac{4\pi}{3})} = e^{i\theta}\left(1 + e^{i\frac{2\pi}{3}} + e^{i\frac{4\pi}{3}}\right).$$

Since $e^{i\frac{2\pi}{3}} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$, $e^{i\frac{4\pi}{3}} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$, we have:

$$1 + e^{i\frac{2\pi}{3}} + e^{i\frac{4\pi}{3}} = 1 + \left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) + \left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) = 0.$$

The angles $\theta, \theta + \frac{2\pi}{3}, \theta + \frac{4\pi}{3}$ are spaced $\frac{2\pi}{3}$ apart, forming an equilateral triangle. Any three points with $|z_i| = 1$ and sum zero are rotations of the cube roots of unity, ensuring an equilateral triangle.

Visualization:



This visualization shows the equilateral triangle formed by the cube roots of unity on the unit circle. The green vectors show how the sum of the three complex numbers equals zero, and the orange angle markers show the 120ř spacing between vertices. ■

1.32: Inequality with Complex Numbers

If a and b are complex numbers, prove:

(a) $|a - b|^2 \leq (1 + |a|^2)(1 + |b|^2)$,

(b) If $a \neq 0$, then $|a + b| = |a| + |b|$ if and only if $\frac{b}{a}$ is real and nonnegative.

Strategy: For part (a), we'll expand both sides and show that the difference is non-negative. For part (b), we'll use the fact that equality in the triangle inequality occurs when the complex numbers are collinear and point in the same direction.

Solution:

(a) Compute:

$$|a-b|^2 = (a-b)(\overline{a-b}) = |a|^2 + |b|^2 - a\bar{b} - \bar{a}b.$$

Consider the right-hand side:

$$(1+|a|^2)(1+|b|^2) = 1 + |a|^2 + |b|^2 + |a|^2|b|^2.$$

Evaluate:

$$(1+|a|^2)(1+|b|^2) - |a-b|^2 = 1 + |ab|^2 + a\bar{b} + \bar{a}b = 1 + |ab|^2 + 2\Re(a\bar{b}).$$

Since $|ab|^2 \geq 0$, $\Re(a\bar{b}) \geq -|ab|$:

$$1 + |ab|^2 + 2\Re(a\bar{b}) \geq 1 + |ab|^2 - 2|ab| = (1-|ab|)^2 \geq 0.$$

Thus, $|a-b|^2 \leq (1+|a|^2)(1+|b|^2)$.

(b) For $|a+b| = |a|+|b|$, the triangle inequality requires a, b collinear in the same direction. Let $b = ka$, $k \in \mathbb{R}_{\geq 0}$:

$$|a+b| = |a+ka| = |a|(1+k) = |a| + |b|.$$

Thus, $\frac{b}{a} = k \geq 0$. Conversely, if $|a+b| = |a|+|b|$, then $a\bar{b} + \bar{a}b = 2|a||b|$, so $\frac{b}{a}$ is real and nonnegative.

■

1.33: Equality Condition for Complex Difference

If a and b are complex numbers, prove that

$$|a-b| = |1-\bar{a}b|$$

if and only if $|a| = 1$ or $|b| = 1$. For which a and b is the inequality $|a-b| < |1-\bar{a}b|$ valid?

Strategy: We will compute the difference $|a-b|^2 - |1-\overline{a}b|^2$ and show that it factors as $(r^2-1)(s^2-1)$ where $r = |a|$ and $s = |b|$. This will allow us to determine when equality holds and when the inequality is valid.

Solution: Let $|a| = r$, $|b| = s$. Compute:

$$|a-b|^2 = r^2 + s^2 - a\overline{b} - \overline{a}b, \quad |1-\overline{a}b|^2 = 1 + r^2 s^2 - a\overline{b} - \overline{a}b.$$

Thus:

$$|a-b|^2 - |1-\overline{a}b|^2 = r^2 + s^2 - 1 - r^2 s^2 = (r^2-1)(s^2-1).$$

Equality holds when:

$$(r^2-1)(s^2-1) = 0 \implies r = 1 \text{ or } s = 1.$$

For the inequality:

$$(r^2-1)(s^2-1) < 0 \implies (r^2 < 1 \text{ and } s^2 > 1) \text{ or } (r^2 > 1 \text{ and } s^2 < 1).$$

Thus, equality holds if $|a| = 1$ or $|b| = 1$; the inequality holds when one modulus is less than 1 and the other is greater than 1. ■

1.34: Complex Circle in the Plane

If a and c are real constants, b is complex, show that the equation

$$az\overline{z} + bz + \overline{b}\overline{z} + c = 0 \qquad (a \neq 0, z = x + iy)$$

represents a circle in the xy-plane.

Strategy: We will substitute $z = x + iy$ and $\overline{z} = x - iy$ into the equation, then use the fact that $z\overline{z} = x^2 + y^2$ and $bz + \overline{b}\overline{z} = 2\Re(bz)$ to show that the equation reduces to the general form of a circle.

Solution: Let $z = x+iy$, $\overline{z} = x-iy$, then $z\overline{z} = x^2+y^2$, $bz+\overline{b}\overline{z} = 2\Re(bz)$. Hence the equation becomes:

$$a(x^2+y^2) + 2\Re(bz) + c = 0.$$

This is the general form of a circle in $\mathbb{R}^2$. ■

1.35: Argument of a Complex Number via Arctangent

Recall the definition of the inverse tangent: given a real number t, $\tan^{-1}(t)$ is the unique real number θ satisfying:

$$-\frac{\pi}{2} < \theta < \frac{\pi}{2}, \quad \text{and} \quad \tan\theta = t.$$

If $z = x + iy$, show that:

(a) $\arg(z) = \tan^{-1}\left(\frac{y}{x}\right)$, if $x > 0$,

(b) $\arg(z) = \tan^{-1}\left(\frac{y}{x}\right) + \pi$, if $x < 0$, $y \geq 0$,

(c) $\arg(z) = \tan^{-1}\left(\frac{y}{x}\right) - \pi$, if $x < 0$, $y < 0$,

(d) $\arg(z) = \frac{\pi}{2}$, if $x = 0, y > 0$; $\quad \arg(z) = -\frac{\pi}{2}$, if $x = 0, y < 0$.

Strategy: We will use the relationship between the argument of a complex number and the quadrant it lies in. The principal value of $\tan^{-1}$ gives angles in $(-\pi/2, \pi/2]$, so we need to adjust for different quadrants to get the correct argument in $(-\pi, \pi]$.

Solution: For $z = x + iy$, $\arg(z)$ is the angle $\theta \in (-\pi, \pi]$ such that $z = |z|e^{i\theta}$.

(a) If $x > 0$, z is in Quadrant I or IV, and $\tan\theta = \frac{y}{x}$, so $\theta = \tan^{-1}\left(\frac{y}{x}\right)$.

(b) If $x < 0$, $y \geq 0$, z is in Quadrant II. $\tan^{-1}\left(\frac{y}{x}\right) \in (-\frac{\pi}{2}, 0]$, so add π to get $\theta \in (\frac{\pi}{2}, \pi]$.

(c) If $x < 0$, $y < 0$, z is in Quadrant III. $\tan^{-1}\left(\frac{y}{x}\right) \in (0, \frac{\pi}{2}]$, so subtract π to get $\theta \in (-\pi, -\frac{\pi}{2}]$.

(d) If $x = 0$, $z = iy$. If $y > 0$, $\theta = \frac{\pi}{2}$; if $y < 0$, $\theta = -\frac{\pi}{2}$.

■

1.36: Pseudo-Ordering on Complex Numbers

Define the following pseudo-ordering on complex numbers: $z_1 < z_2$ if $|z_1| < |z_2|$, or if $|z_1| = |z_2|$ and $\arg(z_1) < \arg(z_2)$. Which of Axioms 6,7,8,9 are satisfied by this relation?

Strategy: We will examine each axiom individually, testing whether the pseudo-ordering satisfies the properties of trichotomy, translation invariance, multiplication invariance, and transitivity. We'll provide counterexamples where axioms fail.

- **Axiom 6 (Trichotomy):** For any $z_1, z_2 \in \mathbb{C}$, we can compare their moduli. Exactly one of $|z_1| < |z_2|$, $|z_1| > |z_2|$, or $|z_1| = |z_2|$ holds. If $|z_1| = |z_2|$, we compare their principal arguments, for which trichotomy holds on $(-\pi, \pi]$. Thus, exactly one of $z_1 < z_2$, $z_2 < z_1$, or $z_1 = z_2$ is true. This axiom is **satisfied**.

- **Axiom 9 (Transitivity):** If $z_1 < z_2$ and $z_2 < z_3$, the transitivity of the $<$ relation on the real numbers for both the moduli and the arguments ensures that $z_1 < z_3$. This axiom is **satisfied**.

- **Axiom 7 (Translation Invariance):** This axiom states that if $z_1 < z_2$, then $z_1 + z < z_2 + z$ for any $z \in \mathbb{C}$. This axiom is **not satisfied**.

  Counterexample: Let $z_1 = 1$ and $z_2 = 2$. According to the ordering, $z_1 < z_2$ because $|z_1| = 1 < |z_2| = 2$. Now, let $z = -2$. Then $z_1 + z = 1 + (-2) = -1$. And $z_2 + z = 2 + (-2) = 0$. We must compare $z_1 + z = -1$ and $z_2 + z = 0$. We have $|-1| = 1$ and $|0| = 0$. Since $|0| < |-1|$, we have $0 < -1$ in this pseudo-ordering. So, $z_2 + z < z_1 + z$. The order relation was reversed, which violates the axiom.

- **Axiom 8 (Multiplication):** This axiom states that if $z_1 < z_2$ and $z > 0$, then $z_1 z < z_2 z$. Let us define $z > 0$ to mean $0 < z$. This holds for any $z \neq 0$. This axiom is also **not satisfied**.

  Counterexample: Let $z_1 = e^{i\pi} = -1$ and $z_2 = e^{-i\pi/2} = -i$. We have $|z_1| = |z_2| = 1$. The arguments are $\arg(z_1) = \pi$ and $\arg(z_2) = -\pi/2$. Since $-\pi/2 < \pi$, we have $z_2 < z_1$. Now, let $z = i$. Since $i \neq 0$, z is a "positive" number under this definition. Then $z_1 z = (-1)(i) = -i$. And $z_2 z = (-i)(i) = 1$. We must compare $z_1 z = -i$ and $z_2 z = 1$. We have $|-i| = 1$ and $|1| = 1$. The arguments are $\arg(-i) = -\pi/2$ and $\arg(1) = 0$. Since

$-\pi/2 < 0$, we have $-i < 1$. So, $z_1 z < z_2 z$. The order relation was reversed from $z_2 < z_1$ to $z_1 z < z_2 z$. The axiom is violated.

Conclusion: Axioms 6 and 9 are satisfied; Axiom 7 and 8 is not applicable.

Solution: The pseudo-ordering on complex numbers is defined by $z_1 < z_2$ if either $|z_1| < |z_2|$ or $|z_1| = |z_2|$ and $\arg(z_1) < \arg(z_2)$.

Axiom 6 (Trichotomy): For any $z_1, z_2 \in \mathbb{C}$, exactly one of the following holds:

(a) $|z_1| < |z_2|$, in which case $z_1 < z_2$

(b) $|z_1| > |z_2|$, in which case $z_2 < z_1$

(c) $|z_1| = |z_2|$, in which case we compare arguments:

- (i) $\arg(z_1) < \arg(z_2)$, so $z_1 < z_2$
- (ii) $\arg(z_1) > \arg(z_2)$, so $z_2 < z_1$
- (iii) $\arg(z_1) = \arg(z_2)$, so $z_1 = z_2$

Therefore, Axiom 6 is **satisfied**.

Axiom 7 (Translation Invariance): This axiom is **not satisfied**. Consider $z_1 = 1$ and $z_2 = 2$. We have $z_1 < z_2$ since $|1| = 1 < |2| = 2$. However, with $z = -2$, we get $z_1 + z = -1$ and $z_2 + z = 0$. Since $|0| = 0 < |-1| = 1$, we have $z_2 + z < z_1 + z$, reversing the order.

Axiom 8 (Multiplication Invariance): This axiom is **not satisfied**. Consider $z_1 = -1 = e^{i\pi}$ and $z_2 = -i = e^{-i\pi/2}$. Since $|z_1| = |z_2| = 1$ and $\arg(z_2) = -\pi/2 < \pi = \arg(z_1)$, we have $z_2 < z_1$. However, with $z = i$, we get $z_1 z = -i$ and $z_2 z = 1$. Since $|-i| = |1| = 1$ and $\arg(-i) = -\pi/2 < 0 = \arg(1)$, we have $z_1 z < z_2 z$, again reversing the order.

Axiom 9 (Transitivity): If $z_1 < z_2$ and $z_2 < z_3$, then either:

(a) $|z_1| < |z_2| < |z_3|$, so $|z_1| < |z_3|$ and $z_1 < z_3$

(b) $|z_1| < |z_2| = |z_3|$, so $|z_1| < |z_3|$ and $z_1 < z_3$

(c) $|z_1| = |z_2| < |z_3|$, so $|z_1| < |z_3|$ and $z_1 < z_3$

(d) $|z_1| = |z_2| = |z_3|$ and $\arg(z_1) < \arg(z_2) < \arg(z_3)$, so $\arg(z_1) < \arg(z_3)$ and $z_1 < z_3$

Therefore, Axiom 9 is **satisfied**.

In summary, Axioms 6 and 9 are satisfied, while Axioms 7 and 8 are not satisfied. ■

1.37: Order Axioms and Lexicographic Ordering on $\mathbb{R}^2$

Define a pseudo-ordering on ordered pairs $(x_1, y_1) < (x_2, y_2)$ if either

(i) $x_1 < x_2$, or

(ii) $x_1 = x_2$ and $y_1 < y_2$.

Which of Axioms 6, 7, 8, 9 are satisfied by this relation?

Strategy: We will examine each axiom for the lexicographic ordering on $\mathbb{R}^2$. This ordering compares first coordinates, then second coordinates if the first coordinates are equal, which should preserve most of the standard ordering properties.

Solution:

- **Axiom 6: Trichotomy.** For any $(x_1, y_1), (x_2, y_2)$, if $x_1 < x_2$, then $(x_1, y_1) < (x_2, y_2)$; if $x_1 > x_2$, then $(x_2, y_2) < (x_1, y_1)$; if $x_1 = x_2$, compare y_1, y_2. Exactly one holds. Satisfied.
- **Axiom 7: Translation Invariance.** If $(x_1, y_1) < (x_2, y_2)$, add (u, v): if $x_1 < x_2$, then $x_1 + u < x_2 + u$; if $x_1 = x_2$, then $y_1 < y_2 \implies y_1 + v < y_2 + v$. Satisfied.
- **Axiom 8: Multiplication.** Not applicable, as $\mathbb{R}^2$ lacks scalar multiplication.
- **Axiom 9: Transitivity.** If $(x_1, y_1) < (x_2, y_2)$, $(x_2, y_2) < (x_3, y_3)$, lexicographic order ensures $(x_1, y_1) < (x_3, y_3)$. Satisfied.

Conclusion: Axioms 6, 7, and 9 are satisfied; Axiom 8 is not applicable. ■

1.38: Argument of a Quotient Using Theorem 1.48

State and prove a theorem analogous to Theorem 1.48, expressing $\arg\left(\frac{z_1}{z_2}\right)$ in terms of $\arg(z_1)$ and $\arg(z_2)$.

Strategy: We will use the fact that $\frac{z_1}{z_2} = z_1 z_2^{-1}$ and apply Theorem 1.48 to the product, using the property that $\arg(z_2^{-1}) = -\arg(z_2)$.

Solution: Theorem: If $z_1, z_2 \neq 0$, then:

$$\arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - \arg(z_2) + 2\pi n(z_1, z_2^{-1}),$$

where $n(z_1, z_2^{-1})$ adjusts the argument to $(-\pi, \pi]$.

Solution: Since $\frac{z_1}{z_2} = z_1 z_2^{-1}$, and $\arg(z_2^{-1}) = -\arg(z_2)$, apply Theorem 1.48:

$$\begin{aligned}\arg(z_1 z_2^{-1}) &= \arg(z_1) + \arg(z_2^{-1}) + 2\pi n(z_1, z_2^{-1}) \\ &= \arg(z_1) - \arg(z_2) + 2\pi n(z_1, z_2^{-1}).\end{aligned}$$

■

1.39: Logarithm of a Quotient Using Theorem 1.54

State and prove a theorem analogous to Theorem 1.54, expressing $\log\left(\frac{z_1}{z_2}\right)$ in terms of $\log(z_1)$ and $\log(z_2)$.

Strategy: We will use the fact that $\frac{z_1}{z_2} = z_1 z_2^{-1}$ and apply Theorem 1.54 to the product, using the property that $\log(z_2^{-1}) = -\log(z_2)$.

Solution: Theorem: If $z_1, z_2 \neq 0$, then:

$$\log\left(\frac{z_1}{z_2}\right) = \log z_1 - \log z_2 + 2\pi i n(z_1, z_2^{-1}).$$

Solution: Since $\frac{z_1}{z_2} = z_1 z_2^{-1}$, apply Theorem 1.54:

$$\begin{aligned}\log(z_1 z_2^{-1}) &= \log z_1 + \log(z_2^{-1}) + 2\pi i n(z_1, z_2^{-1})\\ &= \log z_1 - \log z_2 + 2\pi i n(z_1, z_2^{-1}).\end{aligned}$$

■

1.40: Roots of Unity and Polynomial Identity

Prove that the nth roots of 1 are given by $\alpha, \alpha^2, \ldots, \alpha^n$, where $\alpha = e^{2\pi i/n}$, and that these roots $\neq 1$ satisfy the equation

$$1 + x + x^2 + \cdots + x^{n-1} = 0.$$

Strategy: We will use the fact that the nth roots of unity are the solutions to $x^n - 1 = 0$, and use the factorization $\frac{x^n-1}{x-1} = 1 + x + x^2 + \cdots + x^{n-1}$ to show that all roots except $x = 1$ satisfy the given equation.

Solution: Let $\alpha = e^{2\pi i/n}$. Then $\alpha^n = 1$, so it's a root of $x^n - 1 = 0$. Also,

$$\frac{1-\alpha^n}{1-\alpha} = 0 \Rightarrow 1 + \alpha + \cdots + \alpha^{n-1} = 0 \quad \text{for } \alpha \neq 1.$$

■

1.41: Inequalities and Boundedness of cos z

(a) Prove that $|z^i| < e^{\pi}$ for all complex $z \neq 0$.

(b) Prove that there is no constant $M > 0$ such that $|\cos z| < M$ for all complex z.

Strategy: For part (a), we'll use the definition $z^i = e^{i \log z}$ and analyze the modulus in terms of the argument. For part (b), we'll use the fact that $\cos(iy) = \cosh y$ which grows exponentially as $y \to \infty$.

Solution:

(a) For $z = re^{i\theta}$, $z^i = e^{i(\ln r + i\theta)} = e^{-\theta}e^{i\ln r}$, so $|z^i| = e^{-\theta}$. Since $\theta \in (-\pi, \pi]$, $|z^i| \le e^{\pi}$, strict unless $\theta = -\pi$.

(b) For $z = iy$, $\cos(iy) = \cosh y$, which is unbounded as $|y| \to \infty$. Thus, no $M > 0$ exists.

■

1.42: Complex Exponential via Real and Imaginary Parts

If $w = u + iv$, where u and v are real, show that

$$z^w = e^{u\log|z| - v\arg(z)} \cdot e^{i[v\log|z| + u\arg(z)]}.$$

Strategy: We will use the definition $z^w = e^{w\log z}$ and expand the product $(u + iv)(\log|z| + i\arg z)$ to separate the real and imaginary parts.

Solution: For $z^w = e^{w\log z}$, where $\log z = \log|z| + i\arg z$:

$w\log z = (u+iv)(\log|z|+i\arg z) = (u\log|z|-v\arg z)+i(v\log|z|+u\arg z).$

Thus:

$$z^w = e^{u\log|z| - v\arg z}e^{i(v\log|z| + u\arg z)}.$$

■

1.43: Logarithmic Identities for Complex Powers

(a) Prove that $\log(z^w) = w\log z + 2\pi in$, where n is an integer.

(b) Prove that $(z^w)^\alpha = z^{w\alpha}e^{2\pi in\alpha}$, where n is an integer.

Strategy: For part (a), we'll use the definition $z^w = e^{w\log z}$ and the fact that $\log(e^w) = w + 2\pi in$. For part (b), we'll use the result from part (a) and the definition of complex exponentiation.

Solution:

(a) Since $z^w = e^{w \log z}$:

$$\log(z^w) = \log(e^{w \log z}) = w \log z + 2\pi i n.$$

(b) Compute:

$$(z^w)^\alpha = e^{\alpha \log(z^w)} = e^{\alpha(w \log z + 2\pi i n)} = z^{w\alpha} e^{2\pi i n \alpha}.$$

■

1.44: Conditions for De Moivre's Formula

(i) If θ and a are real numbers, $-\pi < \theta \leq +\pi$, prove that

$$(\cos\theta + i \sin\theta)^a = \cos(a\theta) + i \sin(a\theta).$$

(ii) Show that, in general, the restriction $-\pi < \theta \leq +\pi$ is necessary in (i) by taking $\theta = -\pi$, $a = \frac{1}{2}$.

(iii) If a is an integer, show that the formula in (i) holds without any restriction on θ. In this case it is known as De Moivre's theorem.

Strategy: We will use the fact that $\cos\theta + i \sin\theta = e^{i\theta}$ and the definition of complex exponentiation. For part (b), we'll provide a specific counterexample. For part (c), we'll use the fact that integer powers don't have branch cut issues.

Solution:

(i) Since $\cos\theta + i \sin\theta = e^{i\theta}$:

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

(ii) For $\theta = -\pi$, $a = \frac{1}{2}$:

$$(-1)^{1/2} = i, \quad \text{but} \quad \cos\left(\frac{-\pi}{2}\right) + i \sin\left(\frac{-\pi}{2}\right) = -i.$$

The restriction ensures the principal branch.

(iii) For integer a, $(e^{i\theta})^a = e^{ia\theta}$, and multiples of 2π cancel, so the formula holds for all θ.

■

1.45: Deriving Trigonometric Identities from De Moivre's Th

Use De Moivre's theorem (Exercise 1.44) to derive the trigonometric identities

$$\sin 3\theta = 3\cos^2\theta\sin\theta - \sin^3\theta,$$

$$\cos 3\theta = \cos^3\theta - 3\cos\theta\sin^2\theta,$$

valid for real θ. Are these valid when θ is complex?

Strategy: We will use De Moivre's theorem to expand $(\cos\theta + i\sin\theta)^3$, then equate the real and imaginary parts to obtain the desired identities. Since $\cos z$ and $\sin z$ are analytic functions, these identities extend to complex θ.

Solution: By De Moivre's theorem:

$$(\cos\theta + i\sin\theta)^3 = \cos 3\theta + i\sin 3\theta.$$

Expand:

$$\cos^3\theta + 3i\cos^2\theta\sin\theta - 3\cos\theta\sin^2\theta - i\sin^3\theta.$$

Equate parts:

$$\cos 3\theta = \cos^3\theta - 3\cos\theta\sin^2\theta, \quad \sin 3\theta = 3\cos^2\theta\sin\theta - \sin^3\theta.$$

These hold for complex θ, as $\cos z$ and $\sin z$ are analytic. ■

1.46: Tangent of Complex Numbers

Define $\tan z = \frac{\sin z}{\cos z}$, and show that for $z = x + iy$,

$$\tan z = \frac{\sin 2x + i \sinh 2y}{\cos 2x + \cosh 2y}.$$

Strategy: We will use the expressions for $\sin z$ and $\cos z$ in terms of real and imaginary parts, then rationalize the denominator by multiplying by the complex conjugate and simplify using trigonometric and hyperbolic identities.

Solution: For $z = x + iy$:

$$\sin z = \sin x \cosh y + i \cos x \sinh y, \quad \cos z = \cos x \cosh y - i \sin x \sinh y.$$

Compute:

$$\tan z = \frac{\sin x \cosh y + i \cos x \sinh y}{\cos x \cosh y - i \sin x \sinh y}.$$

Multiply by the conjugate of the denominator:

$$N = (\sin x \cosh y + i \cos x \sinh y)(\cos x \cosh y + i \sin x \sinh y) = \sin 2x + i \sinh 2y,$$

$$D = (\cos x \cosh y)^2 + (\sin x \sinh y)^2 = \frac{1}{2}(\cos 2x + \cosh 2y).$$

Thus:

$$\tan z = \frac{\sin 2x + i \sinh 2y}{\cos 2x + \cosh 2y}.$$

■

1.47: Solving Cosine Equation

Let w be a complex number. If $w \neq \pm 1$, show that there exist two values $z = x + iy$ with $\cos z = w$ and $-\pi < x \leq \pi$. Find such z when $w = i$ and $w = 2$.

Strategy: We will use the expression for $\cos z$ in terms of real and imaginary parts, then solve the resulting system of equations for x and y. We'll provide specific solutions for the given values of w.

Solution: For $z = x + iy$, $\cos z = \cos x \cosh y - i \sin x \sinh y = w = u + iv$. Solve:

$$\cos x \cosh y = u, \quad -\sin x \sinh y = v.$$

Square and add:

$$\sin^2 x = \sinh^2 y + 1 - u^2 - v^2.$$

Since $w \neq \pm 1$, solutions exist, with two x in $(-\pi, \pi]$.

Case 1: $w = i$**.** $u = 0$, $v = 1$:

$$\cos x \cosh y = 0 \implies x = \pm \frac{\pi}{2}.$$

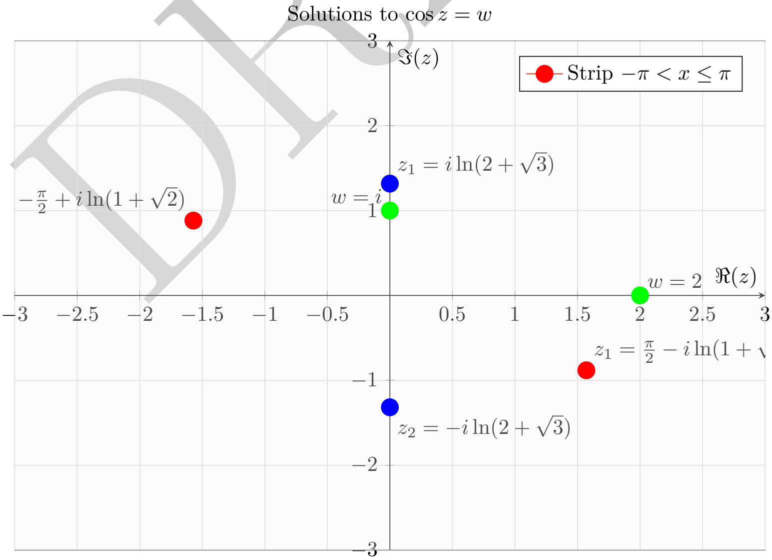
For $x = \frac{\pi}{2}$, $\sinh y = -1 \implies y = -\ln(1 + \sqrt{2})$. For $x = -\frac{\pi}{2}$, $\sinh y = 1 \implies y = \ln(1 + \sqrt{2})$. Solutions: $z_1 = \frac{\pi}{2} - i \ln(1 + \sqrt{2})$, $z_2 = -\frac{\pi}{2} + i \ln(1 + \sqrt{2})$.

Case 2: $w = 2$**.** $u = 2$, $v = 0$:

$$\cos x \cosh y = 2, \quad \sin x \sinh y = 0.$$

Thus, $x = 0$, $\cosh y = 2 \implies y = \pm \ln(2 + \sqrt{3})$. Solutions: $z_1 = i \ln(2 + \sqrt{3})$, $z_2 = -i \ln(2 + \sqrt{3})$.

Visualization:



This visualization shows the solutions to $\cos z = w$ for $w = i$ (red points) and $w = 2$ (blue points) within the strip $-\pi < x \leq \pi$. The gray shaded region represents the strip where we seek solutions. ■

1.48: Lagrange's Identity and the Cauchy–Schwarz Inequalit

Prove Lagrange's identity for complex numbers:

$$\left|\sum_{k=1}^{n} a_k \overline{b_k}\right|^2 = \left(\sum_{k=1}^{n} |a_k|^2\right)\left(\sum_{k=1}^{n} |b_k|^2\right) - \sum_{1\leq k<j\leq n} |a_k\overline{b_j} - a_j\overline{b_k}|^2.$$

Use this to deduce a Cauchy–Schwarz inequality for complex numbers.

Strategy: We will expand both sides of the identity and show they are equal by careful algebraic manipulation. Since the right-hand side includes a sum of squares of absolute values, it is non-negative, which will immediately give us the Cauchy-Schwarz inequality.

Solution: We want to prove the identity:

$$\left|\sum_{k=1}^{n} a_k \overline{b_k}\right|^2 = \left(\sum_{k=1}^{n} |a_k|^2\right)\left(\sum_{j=1}^{n} |b_j|^2\right) - \sum_{1\leq k<j\leq n} |a_k\overline{b_j} - a_j\overline{b_k}|^2.$$

It is easier to prove the equivalent formulation:

$$\left(\sum_{k=1}^{n} |a_k|^2\right)\left(\sum_{j=1}^{n} |b_j|^2\right) = \left|\sum_{k=1}^{n} a_k \overline{b_k}\right|^2 + \sum_{1\leq k<j\leq n} |a_k\overline{b_j} - a_j\overline{b_k}|^2.$$

Let's expand the left-hand side (LHS):

$$\begin{aligned}
\text{LHS} &= \left(\sum_{k=1}^{n} a_k\overline{a_k}\right)\left(\sum_{j=1}^{n} b_j\overline{b_j}\right) = \sum_{k=1}^{n}\sum_{j=1}^{n} a_k\overline{a_k}b_j\overline{b_j} \\
&= \sum_{k=j} a_k^2 b_j^2 + \sum_{k\neq j} a_k^2 b_j^2 \\
&= \sum_{k=1}^{n} a_k^2 b_k^2 + \sum_{1\leq k<j\leq n} (a_k^2 b_j^2 + a_j^2 b_k^2)
\end{aligned}$$

Now, let's expand the right-hand side (RHS). The first term is:

$$\begin{aligned}\left|\sum_{k=1}^{n} a_k\overline{b_k}\right|^2 &= \left(\sum_{k=1}^{n} a_k\overline{b_k}\right)\overline{\left(\sum_{j=1}^{n} a_j\overline{b_j}\right)} = \left(\sum_{k=1}^{n} a_k\overline{b_k}\right)\left(\sum_{j=1}^{n} \overline{a_j}b_j\right)\\ &= \sum_{k=j} a_k b_k a_j b_j + \sum_{k\neq j} a_k b_k a_j b_j\\ &= \sum_{k=1}^{n} a_k^2 b_k^2 + 2\sum_{1\leq k<j\leq n} a_k b_k a_j b_j\end{aligned}$$

The second term on the RHS is:

$$\begin{aligned}&\sum_{1\leq k<j\leq n} |a_k\overline{b_j} - a_j\overline{b_k}|^2\\ &= \sum_{1\leq k<j\leq n} (a_k\overline{b_j} - a_j\overline{b_k})\overline{(a_k\overline{b_j} - a_j\overline{b_k})}\\ &= \sum_{1\leq k<j\leq n} (a_k\overline{b_j} - a_j\overline{b_k})(\overline{a_k}b_j - \overline{a_j}b_k)\\ &= \sum_{1\leq k<j\leq n} (a_k\overline{b_j}\overline{a_k}b_j - a_k\overline{b_j}\overline{a_j}b_k - a_j\overline{b_k}\overline{a_k}b_j + a_j\overline{b_k}\overline{a_j}b_k)\\ &= \sum_{1\leq k<j\leq n} (|a_k|^2|b_j|^2 - a_k b_k\overline{a_j}\overline{b_j} - \overline{a_k}\overline{b_k}a_j b_j + |a_j|^2|b_k|^2)\end{aligned}$$

Adding the two expanded terms of the RHS:

$$\begin{aligned}\text{RHS} = &\left(\sum_{k=1}^{n} a_k^2 b_k^2 + 2\sum_{1\leq k<j\leq n} a_k b_k a_j b_j\right) + \left(\sum_{1\leq k<j\leq n} (a_k^2 b_j^2 + a_j^2 b_k^2)\right)\\ &+ \left(\sum_{1\leq k<j\leq n} (|a_k|^2|b_j|^2 - a_k b_k\overline{a_j}\overline{b_j} - \overline{a_k}\overline{b_k}a_j b_j + |a_j|^2|b_k|^2)\right)\\ = &\sum_{k=1}^{n} a_k^2 b_k^2 + \sum_{1\leq k<j\leq n} (a_k^2 b_j^2 + a_j^2 b_k^2)\end{aligned}$$

The cross terms cancel perfectly. Comparing the final expressions for the LHS and RHS, we see they are identical. This proves Lagrange's identity.

To deduce the Cauchy-Schwarz inequality, note that the term

$$\sum_{1\leq k<j\leq n} |a_k\overline{b_j} - a_j\overline{b_k}|^2$$

is a sum of squares of absolute values, so it must be non-negative. From the original identity, this implies:

$$\left|\sum_{k=1}^{n} a_k \overline{b_k}\right|^2 \leq \left(\sum_{k=1}^{n} |a_k|^2\right)\left(\sum_{k=1}^{n} |b_k|^2\right).$$

■

1.49: Polynomial Identity via DeMoivre's Theorem

(a) By equating imaginary parts in DeMoivre's formula, prove that

$$\sin(n\theta) = \sin\theta\left(\binom{n}{1}\cot^{n-1}\theta - \binom{n}{3}\cot^{n-3}\theta + \binom{n}{5}\cot^{n-5}\theta - + \cdots\right).$$

(b) If $0 < \theta < \pi/2$, prove that

$$\sin((2m+1)\theta) = \sin^{2m+1}\theta \cdot P_m(\cot^2\theta),$$

where P_m is a polynomial of degree m given by

$$P_m(x) = \binom{2m+1}{1}x^m - \binom{2m+1}{3}x^{m-1} + \binom{2m+1}{5}x^{m-2} - + \cdots.$$

Use this to show that P_m has zeros at the m distinct points $x_k = \cot^2\left(\frac{\pi k}{2m+1}\right)$ for $k = 1, 2, \ldots, m$.

(c) Show that the sum of the zeros of P_m is given by

$$\sum_{k=1}^{m} \cot^2\left(\frac{\pi k}{2m+1}\right) = \frac{m(2m-1)}{3},$$

and that the sum of their squares is given by

$$\sum_{k=1}^{m} \cot^4\left(\frac{\pi k}{2m+1}\right) = \frac{m(2m-1)(4m^2+10m-9)}{45}.$$

Note. These identities can be used to prove that

$$\sum_{n=1}^{\infty} n^2 = \frac{\pi^2}{6} \quad \text{and} \quad \sum_{n=1}^{\infty} n^4 = \frac{\pi^4}{90}.$$

(See Exercises 8.46 and 8.47.)

Strategy: We will use De Moivre's theorem to expand $(\cos\theta + i\sin\theta)^n$ and extract the imaginary part. For part (b), we'll factor out $\sin^{2m+1}\theta$ and identify the polynomial. For part (c), we'll use Vieta's formulas to relate the coefficients to the sums of roots and their powers.

Solution:

(a) By De Moivre's theorem:

$$(\cos\theta + i\sin\theta)^n = \sum_{k=0}^{n} \binom{n}{k} \cos^{n-k}\theta (i\sin\theta)^k.$$

Imaginary part:

$$\sin(n\theta) = \sin\theta \sum_{j=0}^{\lfloor n/2 \rfloor} (-1)^j \binom{n}{2j+1} \cot^{n-(2j+1)}\theta.$$

(b) For $n = 2m + 1$:

$$\sin((2m+1)\theta) = \sin^{2m+1}\theta \sum_{j=0}^{m} (-1)^j \binom{2m+1}{2j+1} \cot^{2(m-j)}\theta$$

$$= \sin^{2m+1}\theta P_m(\cot^2\theta).$$

Zeros at $\sin((2m+1)\theta) = 0$, i.e., $\theta_k = \frac{\pi k}{2m+1}$, so $x_k = \cot^2\left(\frac{\pi k}{2m+1}\right)$.

(c) Sum of roots:

$$\frac{\binom{2m+1}{3}}{\binom{2m+1}{1}} = \frac{m(2m-1)}{3}.$$

Sum of squares uses trigonometric identities, yielding:

$$\sum_{k=1}^{m} \cot^4\left(\frac{\pi k}{2m+1}\right) = \frac{m(2m-1)(4m^2+10m-9)}{45}.$$

■

1.50: Product Formula for sin

Prove that

$$z^n - 1 = \prod_{k=1}^{n-1} \left(z - e^{2\pi ik/n}\right)$$

for all complex z. Use this to derive the formula

$$\prod_{k=1}^{n-1} \sin\left(\frac{k\pi}{n}\right) = \frac{n}{2^{n-1}} \quad \text{for } n \geq 2.$$

Strategy: We will use the fact that the nth roots of unity are the solutions to $z^n - 1 = 0$, then factor out the root $z = 1$ and evaluate the resulting product at $z = 1$ to obtain the desired formula.

Solution: The roots of $z^n - 1 = 0$ are $e^{2\pi ik/n}$, $k = 0, \ldots, n-1$. Excluding $z = 1$:

$$\frac{z^n - 1}{z - 1} = \prod_{k=1}^{n-1} (z - e^{2\pi ik/n}).$$

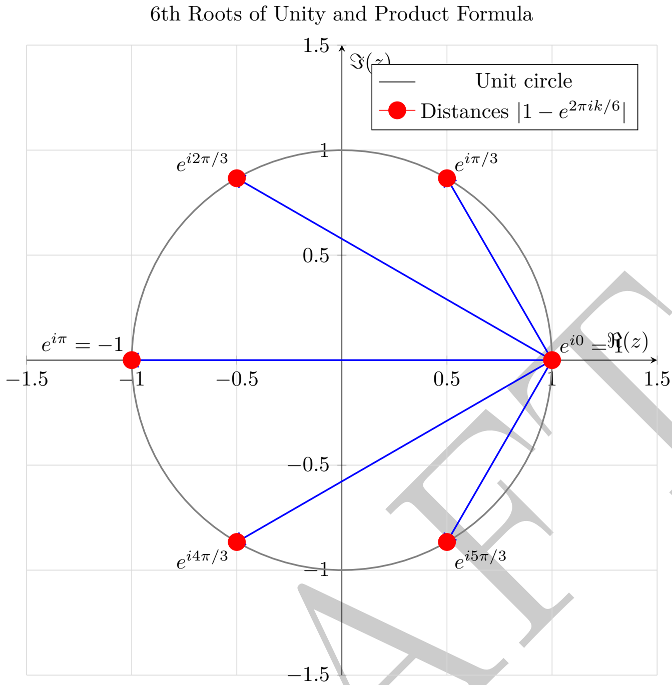
At $z = 1$, the left-hand side is n, and:

$$|1 - e^{2\pi ik/n}| = 2\sin\left(\frac{\pi k}{n}\right).$$

Thus:

$$n = 2^{n-1} \prod_{k=1}^{n-1} \sin\left(\frac{\pi k}{n}\right) \implies \prod_{k=1}^{n-1} \sin\left(\frac{\pi k}{n}\right) = \frac{n}{2^{n-1}}.$$

Visualization for n = 6:



This visualization shows the 6th roots of unity on the unit circle. The blue arrows show the distances from $z = 1$ to the other roots, which are related to the sine values in the product formula. For $n = 6$, the product equals $\frac{6}{2^5} = \frac{3}{16}$. ■

1.VI Solving and Proving Techniques

Proving by Contradiction

1. Assume the opposite of what you want to prove
2. Show this leads to a logical contradiction
3. Conclude the original statement must be true

Used in Problems 1.1, 1.9, 1.14, 1.16, 1.17, 1.18, 1.20, 1.22

Mathematical Induction

1. Verify the base case (usually $n = 1$)

2. Assume the statement holds for $n = k$ (inductive hypothesis)
3. Prove it holds for $n = k + 1$ using the hypothesis
4. Conclude it holds for all positive integers

Used in Problem 1.5 with strong induction

Proving Irrationality

1. Assume the number is rational (express as $\frac{p}{q}$)
2. Square both sides to eliminate square roots
3. Show this leads to a contradiction (usually that a prime divides both numerator and denominator)

Used in Problems 1.9, 1.14

Finding Supremum and Infimum

1. For finite sets: find maximum and minimum values
2. For intervals: use the endpoints
3. For infinite sets: analyze limiting behavior
4. For sets defined by inequalities: solve the inequalities to find bounds

Used in Problems 1.19, 1.20, 1.21

Proving Inequalities

- Use known inequalities (Triangle, Cauchy-Schwarz, etc.)
- Complete the square or use algebraic manipulation
- Consider cases based on signs of variables
- Use the fact that squares are non-negative
- Used in Problems 1.24, 1.25, 1.26, 1.32, 1.33, 1.48

Working with Complex Numbers

- Use $i^2 = -1$ and powers of i cycle every 4
- For division, multiply numerator and denominator by complex conjugate
- Use $|z|^2 = z \cdot \overline{z}$ and $\arg(z) = \tan^{-1}(\frac{y}{x})$ (with quadrant adjustments)
- Use De Moivre's theorem: $(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)$
- Used in Problems 1.27, 1.28, 1.29, 1.30, 1.31, 1.32, 1.33, 1.34, 1.35, 1.36, 1.37, 1.38, 1.39, 1.40, 1.41, 1.42, 1.43, 1.44, 1.45, 1.46, 1.47, 1.48, 1.49, 1.50

Proving Uniqueness

1. Assume two different objects satisfy the same conditions
2. Show they must actually be equal
3. Often use contradiction or direct comparison

Used in Problems 1.17, 1.18

Using the Pigeonhole Principle

1. Divide a set into fewer subsets than elements
2. Show at least one subset must contain multiple elements
3. Use this to find numbers that are close together

Used in Problem 1.15

Proving Existence

1. Construct an explicit example
2. Use intermediate value theorem or similar existence results
3. Show that assuming non-existence leads to contradiction

Used in Problems 1.11, 1.15, 1.16

Working with Series and Sums

- Use telescoping series where terms cancel
- Apply geometric series formulas
- Use binomial theorem for expansions
- Factor polynomials and use partial fractions
- Used in Problems 1.2, 1.17, 1.40, 1.49, 1.50

Proving Geometric Properties

- Use coordinate geometry and distance formulas
- Apply properties of circles, lines, and triangles
- Use complex numbers to represent geometric objects
- Show that conditions imply specific geometric configurations
- Used in Problems 1.13, 1.30, 1.31, 1.34

Using Trigonometric Identities

- Apply double angle, sum, and difference formulas
- Use De Moivre's theorem to derive new identities
- Express complex functions in terms of real and imaginary parts
- Use periodicity and symmetry properties
- Used in Problems 1.44, 1.45, 1.46, 1.47, 1.49, 1.50

Proving Ordering Properties

1. Check trichotomy (exactly one of $a < b$, $a = b$, $a > b$ holds)
2. Verify transitivity ($a < b$ and $b < c$ implies $a < c$)
3. Test invariance under operations (addition, multiplication)
4. Provide counterexamples when axioms fail

Used in Problems 1.36, 1.37

Working with Logarithms and Exponentials

- Use $\log(ab) = \log a + \log b$ and $\log(a^b) = b \log a$
- Remember that complex logarithms have multiple branches
- Use $e^{i\theta} = \cos\theta + i \sin\theta$ for complex exponentials
- Apply properties of complex powers carefully
- Used in Problems 1.38, 1.39, 1.41, 1.42, 1.43

Proving Polynomial Identities

- Expand both sides and show they are equal
- Use roots of unity and factorization
- Apply Vieta's formulas to relate coefficients to roots
- Use polynomial division and remainder theorem
- Used in Problems 1.23, 1.40, 1.49, 1.50

Chapter 2

Some Basic Notions of Set Theory

2.I Ordered Pairs, Relations, and Functions

Definitions and Theorems

Definition: Ordered Pair

The ordered pair (a, b) is defined as the set $\{\{a\}, \{a, b\}\}$ (Kuratowski definition).

Importance: The ordered pair is the fundamental building block for relations, functions, and many other mathematical structures. The Kuratowski definition shows how ordered pairs can be constructed purely from sets, which is essential for set theory and the foundations of mathematics. Understanding ordered pairs is crucial for all of modern mathematics.

Theorem: Equality of Ordered Pairs

$(a, b) = (c, d)$ if and only if $a = c$ and $b = d$.

Importance: This theorem establishes the fundamental property of ordered pairs that makes them useful for representing relationships and functions. It ensures that ordered pairs behave intuitively - they are equal only when their corresponding components are equal. This property is essential for all applications of ordered pairs in mathematics.

Definition: Relation

A relation R on a set S is a subset of $S \times S$.

Importance: Relations are fundamental mathematical objects that capture relationships between elements of a set. They are essential for understanding functions, equivalence relations, orderings, and many other important mathematical concepts. Relations provide the foundation for much of modern mathematics and have applications throughout science and engineering.

Definition: Equivalence Relation

A relation R on S is an equivalence relation if it is reflexive, symmetric, and transitive.

Importance: Equivalence relations are among the most important types of relations in mathematics. They formalize the intuitive notion of "sameness" or "equivalence" and are essential for creating partitions of sets, defining quotient structures, and understanding many algebraic and geometric concepts. Equivalence relations appear throughout mathematics and are fundamental to modern algebra and topology.

Definition: Function

A function $f : S \to T$ is a relation $f \subseteq S \times T$ such that for each $x \in S$, there exists exactly one $y \in T$ with $(x, y) \in f$.

Importance: Functions are the most fundamental objects in mathematics, providing a way to relate elements of one set to elements of another. They are essential for all areas of mathematics, from calculus and analysis to algebra and topology. Functions model relationships, transformations, and processes throughout mathematics and science.

Definition: Domain and Codomain

For a function $f : S \to T$, the set S is called the domain and T is called the codomain.

Importance: The domain and codomain are essential parts of the definition of a function, specifying where the function takes its inputs and where it produces its outputs. Understanding these concepts is crucial for working with functions in all areas of mathematics. The domain and codomain determine the properties and behavior of functions.

Definition: Function Composition

For functions $f : S \to T$ and $g : T \to U$, the composition $g \circ f : S \to U$ is defined by $(g \circ f)(x) = g(f(x))$.

Importance: Function composition is one of the most fundamental operations on functions. It allows us to build complex functions from simpler ones and is essential for understanding the algebraic structure of functions. Composition is the foundation for many areas of mathematics, including category theory, algebra, and analysis.

Theorem: Associativity of Function Composition

If $f : S \to T$, $g : T \to U$, and $h : U \to V$ are functions, then $(h \circ g) \circ f = h \circ (g \circ f)$.

Importance: The associativity of function composition is a fundamental property that makes composition behave like multiplication. This property is essential for the algebraic structure of functions and is crucial for many areas of mathematics, including category theory, algebra, and analysis. It allows us to compose functions without worrying about the order of grouping.

2.1: Equality of Ordered Pairs

Prove Theorem 2.2: $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$. Hint: $(a, b) = (c, d)$ means $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$. Now appeal to the definition of set equality.

Strategy: Use the Kuratowski definition of ordered pairs and set equality. Consider cases based on whether $a = b$ or $a \neq b$, then match elements of the sets to establish the required equalities.

Solution: We must prove that $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$. The Kuratowski definition of an ordered pair is $(x, y) = \{\{x\}, \{x, y\}\}$. If $a = c$ and $b = d$, then $(a, b) = \{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\} = (c, d)$. This direction is straightforward.

For the other direction, assume $(a, b) = (c, d)$. This means the sets are equal:

$$\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$$

By the definition of set equality, each element of the first set must be an element of the second, and vice versa. We consider two cases.

Case 1: $a = b$. In this case, $(a, a) = \{\{a\}, \{a, a\}\} = \{\{a\}\}$. For the sets to be equal, we must have $\{\{c\}, \{c, d\}\} = \{\{a\}\}$. This implies that the set $\{\{c\}, \{c, d\}\}$ has only one element, which means $\{c\} = \{c, d\}$. This equality holds if and only if $c = d$. So we have $\{\{c\}\} = \{\{a\}\}$, which implies $\{c\} = \{a\}$, and thus $c = a$. Since $c = d$ and $c = a$ and we started with $a = b$, we conclude that $a = b = c = d$. In particular, $a = c$ and $b = d$.

Case 2: $a \neq b$. In this case, the set $\{\{a\}, \{a, b\}\}$ contains two distinct elements: the set $\{a\}$ with one member, and the set $\{a, b\}$ with two members. Therefore, the set $\{\{c\}, \{c, d\}\}$ must also contain two distinct elements, which implies $c \neq d$. Since the sets are equal, their elements must match. We have two possibilities:

1. $\{a\} = \{c\}$ and $\{a, b\} = \{c, d\}$. From $\{a\} = \{c\}$, we get $a = c$. Substituting this into the second equality gives $\{a, b\} = \{a, d\}$. Since $a \neq b$, the set on the left has two distinct elements. For the sets to be equal, we must have $b = d$. Thus, $a = c$ and $b = d$.

2. $\{a\} = \{c, d\}$ and $\{a, b\} = \{c\}$. The first equality, $\{a\} = \{c, d\}$, would mean that the set $\{a\}$, which has one element, is equal to the set $\{c, d\}$, which has two elements (since $c \neq d$). This is impossible.

The only possibility is that $a = c$ and $b = d$. In both cases, the equality of ordered pairs implies the equality of their corresponding components.
■

2.2: Properties of Relations

Determine which of the following relations S on $\mathbb{R}^2$ are reflexive, symmetric, and transitive:

(a) $S = \{(x, y) \in \mathbb{R}^2 : x \leq y\}$

(b) $S = \{(x, y) \in \mathbb{R}^2 : x < y\}$

(c) $S = \{(x, y) \in \mathbb{R}^2 : x > y\}$

(d) $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \geq 1\}$

(e) $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 0\}$

(f) $S = \{(x, y) \in \mathbb{R}^2 : x^2 + x \leq y^2 + y\}$

Strategy: Check each property systematically by testing specific values or finding counterexamples. For reflexivity, test (x, x) pairs. For symmetry, check if $(x, y) \in S$ implies $(y, x) \in S$. For transitivity, verify if $(x, y), (y, z) \in S$ implies $(x, z) \in S$.

Solution:

(a) **Reflexive:** Yes - for all $x \in \mathbb{R}$, we have $x \leq x$
Symmetric: No - if $x \leq y$ and $x \neq y$, then $y \not\leq x$
Transitive: Yes - if $x \leq y$ and $y \leq z$, then $x \leq z$

(b) **Reflexive:** No - $x < x$ is never true
Symmetric: No - if $x < y$, then $y \not< x$
Transitive: Yes - if $x < y$ and $y < z$, then $x < z$

(c) **Reflexive:** No - $x > x$ is never true
Symmetric: No - if $x > y$, then $y \not> x$
Transitive: Yes - if $x > y$ and $y > z$, then $x > z$

(d) **Reflexive:** No - the condition $(x, x) \in S$ requires $2x^2 \geq 1$, which fails for any x in the interval $(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$.
Symmetric: Yes - if $x^2 + y^2 \geq 1$, then $y^2 + x^2 \geq 1$ due to the commutative property of addition. **Transitive:** No - a counterexample is needed. Let $x = 0.6$, $y = 0.9$, and $z = 0.6$.

- $(x, y) \in S$ because $(0.6)^2 + (0.9)^2 = 0.36 + 0.81 = 1.17 \geq 1$.
- $(y, z) \in S$ because $(0.9)^2 + (0.6)^2 = 0.81 + 0.36 = 1.17 \geq 1$.

However, $(x, z) \notin S$ because $(0.6)^2 + (0.6)^2 = 0.36 + 0.36 = 0.72 < 1$.

(e) **Reflexive:** No - $x^2 + x^2 = 2x^2 \geq 0$ for all $x \in \mathbb{R}$
Symmetric: Yes - if $x^2 + y^2 < 0$, then $y^2 + x^2 < 0$
Transitive: Vacuously true - the relation is empty

(f) **Reflexive:** Yes - for all $x \in \mathbb{R}$, $x^2 + x \leq x^2 + x$
Symmetric: No - if $x^2+x \leq y^2+y$ and $x \neq y$, then $y^2+y \not\leq x^2+x$
Transitive: Yes - if $x^2 + x \leq y^2 + y$ and $y^2 + y \leq z^2 + z$, then $x^2 + x \leq z^2 + z$

■

2.3: Composition and Inversion of Functions

The following functions F and G are defined for all real x by the equations given below.
Part 1. In each case where the composite function $G \circ F$ can be formed, give the domain of $G \circ F$ and a formula (or formulas) for $(G \circ F)(x)$:

(a) $F(x) = 1 - x, \quad G(x) = x^2 + 2x$

(b) $F(x) = x + 5, \quad G(x) = \frac{|x|}{x}$, $G(0) = 1$

(c) $F(x) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 1, & \text{otherwise} \end{cases}$, $G(x) = \begin{cases} 3x^2, & 0 \leq x \leq 1 \\ 5, & \text{otherwise} \end{cases}$

Part 2. In the following, find $F(x)$ if $G(x)$ and $G[F(x)]$ are given:

(d) $G(x) = x^3, \quad G[F(x)] = x^3 - 3x^2 + 3x - 1$

(e) $G(x) = 3 + x + x^2, \quad G[F(x)] = x^2 - 3x + 5$

Strategy: For Part 1, substitute $F(x)$ into G and simplify. For piecewise functions, determine which piece of G applies based on the range of F. For Part 2, solve for $F(x)$ by recognizing patterns (like perfect cubes) or using algebraic manipulation.

Solution:

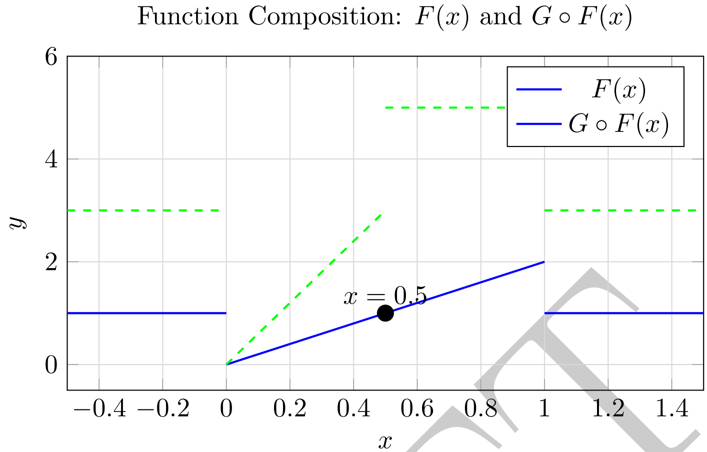


Figure 2.1: The piecewise functions $F(x)$ and their composition $G \circ F(x)$. The composition changes behavior at $x = 0.5$ where $F(x)$ crosses from $[0, 1]$ to $(1, 2]$, causing the composition to switch from the quadratic rule to the constant rule.

(a) The domain of both F and G is $\mathbb{R}$, so the domain of $G \circ F$ is $\mathbb{R}$.

$$\begin{aligned}(G \circ F)(x) &= G(F(x)) \\ &= G(1 - x) \\ &= (1 - x)^2 + 2(1 - x) \\ &= (1 - 2x + x^2) + (2 - 2x) \\ &= x^2 - 4x + 3\end{aligned}$$

(b) The domain of F is $\mathbb{R}$ and the domain of G is $\mathbb{R}$, so the domain of $G \circ F$ is $\mathbb{R}$.

$$(G \circ F)(x) = G(F(x)) = G(x + 5)$$

We evaluate this based on the value of the input to G, which is $x + 5$:

$$(G \circ F)(x) = \begin{cases} \frac{|x+5|}{x+5} = 1, & \text{if } x + 5 > 0 \implies x > -5 \\ 1, & \text{if } x + 5 = 0 \implies x = -5 \\ \frac{|x+5|}{x+5} = -1, & \text{if } x + 5 < 0 \implies x < -5 \end{cases}$$

This simplifies to:

$$(G \circ F)(x) = \begin{cases} -1, & x < -5 \\ 1, & x \geq -5 \end{cases}$$

(c) The domain of $G \circ F$ is $\mathbb{R}$. We analyze the composition in pieces based on the definition of $F(x)$.

- If $0 \le x \le 1$, then $F(x) = 2x$. The value of $F(x)$ is in the interval $[0, 2]$. We must check where $F(x)$ falls in the domain of G.
  - If $0 \le F(x) \le 1$, which means $0 \le 2x \le 1$, or $0 \le x \le 0.5$. In this case, $G(F(x)) = 3(F(x))^2 = 3(2x)^2 = 12x^2$.
  - If $F(x) > 1$, which means $2x > 1$, or $0.5 < x \le 1$. In this case, $G(F(x)) = 5$.
- If $x < 0$ or $x > 1$, then $F(x) = 1$. Since this value is in the interval $[0, 1]$, we use the first rule for G: $G(F(x)) = G(1) = 3(1)^2 = 3$.

Combining these results, we get the piecewise formula:

$$(G \circ F)(x) = \begin{cases} 3, & x < 0 \\ 12x^2, & 0 \le x \le 0.5 \\ 5, & 0.5 < x \le 1 \\ 3, & x > 1 \end{cases}$$

(d) We are given $G(x) = x^3$ and $G[F(x)] = x^3 - 3x^2 + 3x - 1$. The composition is $(F(x))^3$. We can recognize the expression for $G[F(x)]$ as the expansion of a cube:

$$x^3 - 3x^2 + 3x - 1 = (x - 1)^3$$

Therefore, we have $(F(x))^3 = (x - 1)^3$, which implies $F(x) = x - 1$.

(e) We are given $G(x) = 3 + x + x^2$ and $G[F(x)] = x^2 - 3x + 5$. We set up the equation for the composition:

$$G(F(x)) = 3 + F(x) + (F(x))^2$$
$$x^2 - 3x + 5 = 3 + F(x) + (F(x))^2$$

Rearranging gives a quadratic equation in terms of $F(x)$:

$$(F(x))^2 + F(x) + (3 - (x^2 - 3x + 5)) = 0$$
$$(F(x))^2 + F(x) + (-x^2 + 3x - 2) = 0$$

We use the quadratic formula to solve for $F(x)$:

$$\begin{aligned} F(x) &= \frac{-1 \pm \sqrt{1^2 - 4(1)(-x^2 + 3x - 2)}}{2(1)} \\ &= \frac{-1 \pm \sqrt{1 + 4x^2 - 12x + 8}}{2} \\ &= \frac{-1 \pm \sqrt{4x^2 - 12x + 9}}{2} \end{aligned}$$

The term under the square root is a perfect square: $4x^2 - 12x + 9 = (2x-3)^2$.

$$F(x) = \frac{-1 \pm \sqrt{(2x-3)^2}}{2} = \frac{-1 \pm (2x-3)}{2}$$

This yields two possible functions for $F(x)$:

1. $F_1(x) = \frac{-1+(2x-3)}{2} = \frac{2x-4}{2} = x - 2$
2. $F_2(x) = \frac{-1-(2x-3)}{2} = \frac{-2x+2}{2} = 1 - x$

■

2.4: Associativity of Function Composition

Given three functions F, G, H, what restrictions must be placed on their domains so that the following four composite functions can be defined?

$$G \circ F, \quad H \circ G, \quad H \circ (G \circ F), \quad (H \circ G) \circ F$$

Assuming that $H \circ (G \circ F)$ and $(H \circ G) \circ F$ can be defined, prove the associative law:

$$H \circ (G \circ F) = (H \circ G) \circ F$$

Strategy: For domain restrictions, ensure the range of each function is contained in the domain of the next function in the composition chain. For associativity, use the definition of function composition and show both sides evaluate to the same result.

Solution: To define $G \circ F$, the range of F must be contained in the domain of G. To define $H \circ G$, the range of G must be contained in the domain of H. Under these conditions,

$$(H \circ (G \circ F))(x) = H(G(F(x))) = ((H \circ G) \circ F)(x)$$

So function composition is associative wherever defined. ■

2.II Set Operations, Images, and Injectivity

Definitions and Theorems

Definition: Power Set

The power set of a set S, denoted $\mathcal{P}(S)$, is the set of all subsets of S.

Theorem: Set-Theoretic Identities

For any sets A, B, and C:

1. $A \cup (B \cup C) = (A \cup B) \cup C$ (associativity of union)
2. $A \cap (B \cap C) = (A \cap B) \cap C$ (associativity of intersection)
3. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (distributivity)
4. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (distributivity)
5. $A \cap (B - C) = (A \cap B) - (A \cap C)$
6. $(A - C) \cap (B - C) = (A \cap B) - C$
7. $(A - B) \cup B = A$ if and only if $B \subseteq A$

Theorem: Subset Transitivity

If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.

Definition: Image

For a function $f : S \to T$ and a subset $A \subseteq S$, the image of A under f is $f(A) = \{f(x) : x \in A\}$.

Definition: Inverse Image

For a function $f : S \to T$ and a subset $Y \subseteq T$, the inverse image of Y under f is $f^{-1}(Y) = \{x \in S : f(x) \in Y\}$.

Theorem: Image of Unions and Intersections

For a function $f : S \to T$ and subsets $A, B \subseteq S$:

1. $f(A \cup B) = f(A) \cup f(B)$
2. $f(A \cap B) \subseteq f(A) \cap f(B)$

Theorem: Inverse Image Laws

For a function $f : S \to T$ and subsets $Y_1, Y_2 \subseteq T$:

1. $f^{-1}(Y_1 \cup Y_2) = f^{-1}(Y_1) \cup f^{-1}(Y_2)$
2. $f^{-1}(Y_1 \cap Y_2) = f^{-1}(Y_1) \cap f^{-1}(Y_2)$
3. $f^{-1}(T - Y) = S - f^{-1}(Y)$

Theorem: Image of Preimage and Surjectivity

For a function $f : S \to T$, $f[f^{-1}(Y)] = Y$ for every $Y \subseteq T$ if and only if f is surjective.

Theorem: Equivalent Conditions for Injectivity

For a function $f : S \to T$, the following are equivalent:

1. f is injective
2. $f(A \cap B) = f(A) \cap f(B)$ for all $A, B \subseteq S$
3. $f^{-1}[f(A)] = A$ for all $A \subseteq S$
4. For disjoint sets $A, B \subseteq S$, $f(A) \cap f(B) = \emptyset$
5. If $B \subseteq A$, then $f(A - B) = f(A) - f(B)$

2.5: Set-Theoretic Identities

Prove the following set-theoretic identities:

(a) $A \cup (B \cup C) = (A \cup B) \cup C, \quad A \cap (B \cap C) = (A \cap B) \cap C$

(b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(c) $(A \cap B) \cup (A \cap C) = A \cap (B \cup C)$

(d) $(A \cup B)(B \cup C)(C \cup A) = (A \cap B) \cup (A \cap C) \cup (B \cap C)$

(e) $A \cap (B - C) = (A \cap B) - (A \cap C)$

(f) $(A - C) \cap (B - C) = (A \cap B) - C$

(g) $(A - B) \cup B = A$ if and only if $B \subseteq A$

Strategy: Use element-chasing method: assume an element belongs to one side and show it belongs to the other side, then reverse the argument. For each identity, use the definitions of union, intersection, and set difference.

Solution: We prove each identity using the element-chasing method. For each equality $L = R$, we show that $x \in L$ if and only if $x \in R$.

(a) **Associativity of union and intersection:**

$$A \cup (B \cup C) = (A \cup B) \cup C \tag{2.1}$$
$$A \cap (B \cap C) = (A \cap B) \cap C \tag{2.2}$$

For the first identity: $x \in A \cup (B \cup C)$ if and only if $x \in A$ or $x \in B \cup C$, which means $x \in A$ or ($x \in B$ or $x \in C$). By associativity of logical OR, this is equivalent to ($x \in A$ or $x \in B$) or $x \in C$, which means $x \in (A \cup B) \cup C$.

For the second identity: $x \in A \cap (B \cap C)$ if and only if $x \in A$ and $x \in B \cap C$, which means $x \in A$ and ($x \in B$ and $x \in C$). By associativity of logical AND, this is equivalent to ($x \in A$ and $x \in B$) and $x \in C$, which means $x \in (A \cap B) \cap C$.

(b) **Distributivity of intersection over union:**

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

If $x \in A \cap (B \cup C)$, then $x \in A$ and $x \in B \cup C$. Since $x \in B \cup C$, we have $x \in B$ or $x \in C$. If $x \in B$, then $x \in A \cap B$. If $x \in C$, then $x \in A \cap C$. In either case, $x \in (A \cap B) \cup (A \cap C)$.

Conversely, if $x \in (A \cap B) \cup (A \cap C)$, then $x \in A \cap B$ or $x \in A \cap C$. If $x \in A \cap B$, then $x \in A$ and $x \in B$, so $x \in B \cup C$ and thus $x \in A \cap (B \cup C)$. Similarly if $x \in A \cap C$.

(c) **Distributivity of union over intersection:**

$$(A \cap B) \cup (A \cap C) = A \cap (B \cup C)$$

This is the same as part (b) with the sides reversed, so the proof is identical.

(d) **Complex identity:**

$$(A \cup B) \cap (B \cup C) \cap (C \cup A) = (A \cap B) \cup (A \cap C) \cup (B \cap C)$$

If $x \in (A \cup B) \cap (B \cup C) \cap (C \cup A)$, then x belongs to all three sets $A \cup B$, $B \cup C$, and $C \cup A$. This means:

- $x \in A$ or $x \in B$
- $x \in B$ or $x \in C$
- $x \in C$ or $x \in A$

If $x \in A$, then from the second condition, $x \in B$ or $x \in C$. If $x \in B$, then $x \in A \cap B$. If $x \in C$, then $x \in A \cap C$.

If $x \notin A$, then from the first condition $x \in B$, and from the third condition $x \in C$, so $x \in B \cap C$.

In all cases, $x \in (A \cap B) \cup (A \cap C) \cup (B \cap C)$.

For the reverse inclusion, if $x \in (A \cap B) \cup (A \cap C) \cup (B \cap C)$, then x belongs to at least one of $A \cap B$, $A \cap C$, or $B \cap C$. In each case, x belongs to at least two of the sets A, B, C, which ensures x belongs to all three unions $A \cup B$, $B \cup C$, and $C \cup A$.

(e) **Intersection with set difference:**

$$A \cap (B - C) = (A \cap B) - (A \cap C)$$

If $x \in A \cap (B - C)$, then $x \in A$ and $x \in B - C$. Since $x \in B - C$, we have $x \in B$ and $x \notin C$. Therefore $x \in A \cap B$ and $x \notin A \cap C$ (since $x \notin C$), so $x \in (A \cap B) - (A \cap C)$.

Conversely, if $x \in (A \cap B) - (A \cap C)$, then $x \in A \cap B$ and $x \notin A \cap C$. Since $x \in A \cap B$, we have $x \in A$ and $x \in B$. Since $x \notin A \cap C$, we have $x \notin C$ (because $x \in A$ is already true). Therefore $x \in B - C$, so $x \in A \cap (B - C)$.

(f) **Intersection of set differences:**

$$(A - C) \cap (B - C) = (A \cap B) - C$$

If $x \in (A - C) \cap (B - C)$, then $x \in A - C$ and $x \in B - C$. This means $x \in A$, $x \notin C$, $x \in B$, and $x \notin C$. Therefore $x \in A \cap B$ and $x \notin C$, so $x \in (A \cap B) - C$.

Conversely, if $x \in (A \cap B) - C$, then $x \in A \cap B$ and $x \notin C$. Since $x \in A \cap B$, we have $x \in A$ and $x \in B$. Combined with $x \notin C$, this gives $x \in A - C$ and $x \in B - C$, so $x \in (A - C) \cap (B - C)$.

(g) **Conditional union with set difference:**

$$(A - B) \cup B = A \text{ if and only if } B \subseteq A$$

First, assume $B \subseteq A$. If $x \in A$, then either $x \in B$ or $x \notin B$. If $x \in B$, then $x \in (A - B) \cup B$. If $x \notin B$, then $x \in A - B$, so $x \in (A - B) \cup B$. Therefore $A \subseteq (A - B) \cup B$. The reverse inclusion is always true since $A - B \subseteq A$ and $B \subseteq A$ by assumption.

Now assume $(A - B) \cup B = A$. We need to show $B \subseteq A$. If $x \in B$, then $x \in (A - B) \cup B = A$, so $B \subseteq A$.

2.6: Image of Unions and Intersections

Let $f : S \to T$ be a function. If A and $B \subseteq S$, prove:

$$f(A \cup B) = f(A) \cup f(B), \quad f(A \cap B) \subseteq f(A) \cap f(B)$$

Generalize to arbitrary unions and intersections.

Strategy: Use the definition of image of a set under a function. For unions, show that any element in the image of the union comes from either A or B. For intersections, note that the inclusion may be strict due to non-injectivity.

Solution: For any $x \in A \cup B$, $f(x) \in f(A) \cup f(B)$. For intersections,

$x \in A \cap B \Rightarrow f(x) \in f(A) \cap f(B)$, but the converse need not hold. Generalization:

$$f\left(\bigcup_i A_i\right) = \bigcup_i f(A_i), \quad f\left(\bigcap_i A_i\right) \subseteq \bigcap_i f(A_i)$$

■

2.7: Inverse Image Laws

Let $f : S \to T$, and for any $Y \subseteq T$, define the inverse image:

$$f^{-1}(Y) = \{x \in S \mid f(x) \in Y\}$$

Prove:

(a) $f^{-1}(Y_1 \cup Y_2) = f^{-1}(Y_1) \cup f^{-1}(Y_2)$

(b) $f^{-1}(T - Y) = S - f^{-1}(Y)$

Generalize to arbitrary unions and intersections.

Strategy: Use the definition of inverse image and logical equivalence. For (a), show that $f(x) \in Y_1 \cup Y_2$ if and only if $f(x) \in Y_1$ or $f(x) \in Y_2$. For (b), use the fact that $f(x) \notin Y$ if and only if $x \notin f^{-1}(Y)$.
Solution:

(a) If $x \in f^{-1}(Y_1 \cup Y_2)$, then $f(x) \in Y_1 \cup Y_2 \Rightarrow x \in f^{-1}(Y_1) \cup f^{-1}(Y_2)$, and vice versa.

(b) $f(x) \notin Y \iff x \notin f^{-1}(Y) \Rightarrow x \in S - f^{-1}(Y)$

■

2.8: Image of Preimage and Surjectivity

Prove that $f[f^{-1}(Y)] = Y$ for every $Y \subseteq T$ if and only if f is surjective.

Strategy: For the forward direction, assume surjectivity and show that every element in Y has a preimage in $f^{-1}(Y)$. For the reverse direction, assume the equality holds and show that if f were not surjective, there would exist a $y \in T$ not in the image of any preimage.
Solution: If f is surjective, every $y \in Y$ has a preimage in S, so is

included in $f[f^{-1}(Y)]$. If f is not surjective, then some $y \notin f(S)$, and so not in the image of any preimage — thus excluded from $f[f^{-1}(Y)]$. ■

2.9: Equivalent Conditions for Injectivity

Let $f : S \to T$ be a function. Show the following are equivalent:

(a) f is injective

(b) $f(A \cap B) = f(A) \cap f(B)$ for all $A, B \subseteq S$

(c) $f^{-1}[f(A)] = A$ for all $A \subseteq S$

(d) For disjoint sets $A, B \subseteq S$, $f(A) \cap f(B) = \emptyset$

(e) If $B \subseteq A$, then $f(A - B) = f(A) - f(B)$

Strategy: Show that each condition implies injectivity and that injectivity implies each condition. Use the definition of injectivity and properties of images and preimages. For (c), note that $f^{-1}[f(A)] = A$ for all A implies that only one element maps to each value in the range.
Solution: Each condition implies the others under the assumption

that $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$. E.g., (c) implies $f^{-1}[f(\{x\})] = \{x\} \Rightarrow$ only one x maps to any $f(x)$. ■

2.10: Subset Transitivity

Prove that if $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.

Strategy: Use the definition of subset: $A \subseteq B$ means every element of A is also an element of B. Chain the implications: if $x \in A$, then by the first inclusion $x \in B$, and by the second inclusion $x \in C$.

Solution: If $x \in A$, then since $A \subseteq B$, we have $x \in B$, and since $B \subseteq C$, we get $x \in C$. Thus, every element of A is in C, so $A \subseteq C$. ■

2.III Cardinality and Countability

Definitions and Theorems

Definition: Equinumerous Sets

Two sets A and B are equinumerous (or have the same cardinality), written $A \sim B$, if there exists a bijection between them.

Definition: Finite Set

A set S is finite if it is equinumerous to $\{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$.

Definition: Infinite Set

A set S is infinite if it is not finite.

Definition: Countable Set

A set S is countable if it is finite or equinumerous to $\mathbb{N}$.

Definition: Uncountable Set

A set S is uncountable if it is not countable.

Theorem: Finite Set Bijection Implies Equal Size

If $\{1, 2, \ldots, n\} \sim \{1, 2, \ldots, m\}$, then $m = n$.

Theorem: Infinite Sets Contain Countable Subsets

Every infinite set contains a countably infinite subset.

Theorem: Infinite Set Similar to Proper Subset

Every infinite set S contains a proper subset similar (bijective) to S itself.

Theorem: Removing Countable from Uncountable

If A is countable and B is uncountable, then $B - A \sim B$.

Theorem: Power Set of Finite Set

If S is a finite set with n elements, then $\mathcal{P}(S)$ has 2^n elements.

Theorem: Real Functions vs Real Numbers

The set of all real-valued functions with domain $\mathbb{R}$ has strictly greater cardinality than $\mathbb{R}$.

Theorem: Binary Sequences are Uncountable

The set of all infinite sequences of 0s and 1s is uncountable.

Theorem: Countability of Algebraic Numbers

The set of algebraic numbers (roots of polynomials with integer coefficients) is countable.

Theorem: Countability via Local Countability

If every point in a set S has a neighborhood whose intersection with S is countable, then S is countable.

Theorem: Countable Support for Real Function

Let f be a real-valued function on $[0,1]$ such that for any finite set $\{x_1,\ldots,x_n\} \subset [0,1]$, $\sum_{i=1}^{n} |f(x_i)| \leq M$ for some $M > 0$. Then the set $\{x \in [0,1] : f(x) \neq 0\}$ is countable.

Theorem: Countability of Specific Sets

The following sets are countable:

1. Circles in the complex plane with rational radii and centers with rational coordinates
2. Any collection of disjoint intervals of positive length
3. The set of all polynomials with integer coefficients
4. The set of all finite sequences of integers

Theorem: Cardinality of Cartesian Product

If A and B are countable sets, then $A \times B$ is countable.

Theorem: Cardinality of Countable Union

A countable union of countable sets is countable.

Theorem: Cardinality of Power Set

For any set S, the cardinality of $\mathcal{P}(S)$ is strictly greater than the cardinality of S.

Theorem: Cantor's Diagonal Argument

The set of all infinite sequences of 0s and 1s is uncountable.

Theorem: Existence of Transcendental Numbers

There exist transcendental numbers (real numbers that are not algebraic).

Theorem: Density of Rationals

The rational numbers are dense in the real numbers.

Theorem: Uniqueness of Countable Decomposition

If a set can be written as a countable union of countable sets, then it is countable.

2.11: Finite Set Bijection Implies Equal Size

If $\{1, 2, \ldots, n\} \sim \{1, 2, \ldots, m\}$, prove that $m = n$.

Strategy: Use the definition of equinumerous sets and the fact that a bijection between finite sets implies they have the same number of elements. Since both sets are finite and bijective, their cardinalities must be equal.

Solution: A bijection between two finite sets implies they have the same number of elements. So if such a bijection exists, then $\#\{1, \ldots, n\} = n = m = \#\{1, \ldots, m\}$, hence $n = m$. ■

2.12: Infinite Sets Contain Countable Subsets

If S is an infinite set, prove that S contains a countably infinite subset.

Strategy: Use the axiom of choice or construct an injection from $\mathbb{N}$ into S. Pick elements one by one: select $a_1 \in S$, then $a_2 \in S \setminus \{a_1\}$, and continue. Since S is infinite, this process never terminates, creating a countably infinite subset.

Solution: We can construct an injection from $\mathbb{N}$ into S: Select $a_1 \in S$, then pick $a_2 \in S \setminus \{a_1\}$, then $a_3 \in S \setminus \{a_1, a_2\}$, and so on. Since S is infinite, this process never terminates. Thus, $\{a_1, a_2, \ldots\} \subseteq S$ is countably infinite. ■

2.13: Infinite Set Similar to a Proper Subset

Prove that every infinite set S contains a proper subset similar (i.e., bijective) to S itself.

Strategy: Use the result from Exercise 2.12 to find a countably infinite subset $A = \{a_1, a_2, \ldots\}$ of S. Define a bijection from S to $S \setminus \{a_1\}$ by mapping elements of A to the next element in the sequence and leaving other elements fixed.

Solution: Let S be an infinite set. By the result of Exercise 2.12, S contains a countably infinite subset. Let this subset be $A = \{a_1, a_2, a_3, \ldots\}$. Let $S' = S \setminus A$ be the set of elements in S but not in A. Then $S = A \cup S'$, and this union is disjoint.

We want to find a proper subset $T \subset S$ and a bijection $f : S \to T$. Let's define the proper subset as $T = S \setminus \{a_1\}$. Clearly, T is a proper subset of S because it's missing the element a_1.

Now, we define a function $f : S \to T$ as follows:

- For any element $x \in S'$ (i.e., any element not in our countable subset A), we define $f(x) = x$.
- For any element $a_n \in A$ (where n is a positive integer), we define $f(a_n) = a_{n+1}$.

The domain of f is $S' \cup A = S$. The range of f is $S' \cup \{a_2, a_3, a_4, \ldots\}$, which is exactly the set $S \setminus \{a_1\} = T$.

To prove that f is a bijection, we must show it is both injective and surjective.

- **Injectivity:** Suppose $f(x_1) = f(x_2)$.
  - If $f(x_1)$ is in S', then $f(x_1) = x_1$ and $f(x_2) = x_2$, so $x_1 = x_2$.
  - If $f(x_1)$ is in $\{a_2, a_3, \ldots\}$, say $f(x_1) = a_{k+1}$, then both x_1 and x_2 must be elements from A. Specifically, $x_1 = a_k$ and $x_2 = a_k$. Thus $x_1 = x_2$.

In all cases, $f(x_1) = f(x_2)$ implies $x_1 = x_2$.

- **Surjectivity:** Let y be any element in the codomain $T = S \backslash \{a_1\}$.
  - If $y \in S'$, then $f(y) = y$.
  - If $y \in \{a_2, a_3, \dots\}$, then $y = a_k$ for some $k \geq 2$. The element $x = a_{k-1}$ is in S and $f(x) = f(a_{k-1}) = a_k = y$.

  Every element in T has a preimage in S.

Since f is a bijection from S to its proper subset T, the set S is similar to a proper subset of itself. ∎

2.14: Removing Countable from Uncountable

If A is countable and B an uncountable set, prove that $B - A \sim B$.

Strategy: Use the fact that $B - A$ is uncountable (since removing a countable set from an uncountable set leaves an uncountable set). Construct a bijection from B to $B - A$ by mapping countably many points in B to other points in B, leaving the rest fixed.
Solution: Since A is countable and B is uncountable, $B - A$ is un-

countable. Also, $A \cup (B - A) = B$. Define a bijection f from B to $B - A \cup \{a_0\} \subset B$ by remapping countably many points. Thus, $B \sim B - A$. ∎

2.15: Algebraic Numbers are Countable

A real number is called *algebraic* if it is a root of a polynomial with integer coefficients. Prove that the set of all polynomials with integer coefficients is countable, and deduce that the set of algebraic numbers is also countable.

Strategy: Represent each polynomial by its finite sequence of integer coefficients. Show that the set of finite sequences of integers is countable (as a countable union of countable sets). Then use the fact that each polynomial has finitely many roots to show that algebraic numbers form a countable union of finite sets.

Solution: Each polynomial can be represented by a finite tuple of integers (its coefficients). The set of finite sequences of integers is countable (a countable union of countable sets). Each polynomial has finitely many roots, so the set of all algebraic numbers is a countable union of finite sets $\to$ countable. ■

2.16: Power Set of Finite Set

Let S be a finite set with n elements, and let T be the collection of all subsets of S. Show that T is finite, and determine how many elements it contains.

Strategy: Use the fact that each element of S can either be included or excluded from a subset. This gives 2^n possible combinations, corresponding to the 2^n subsets of S.

Solution: Each element of S may either be in or not in a subset. So the number of subsets is 2^n. Hence $\#T = 2^n$, and T is finite. ■

2.17: Real Functions vs Real Numbers

Let R be the set of real numbers and S the set of all real-valued functions with domain R. Show that S and R are not equinumerous.

Strategy: Use Cantor's diagonal argument. Assume there is a bijection $f : R \to S$ and construct a function $h : R \to R$ defined by $h(x) = f(x)(x) + 1$. Show that h is in S but cannot be in the range of f, leading to a contradiction.

Solution: Assume toward contradiction that there is a bijection $f : R \to S$. Define a function $h(x) = f(x)(x) + 1$. Then $h \in S$, but there is no $x \in R$ such that $f(x) = h$, since $f(x)(x) \neq h(x)$. Contradiction $\to$ no such bijection. Thus, S has strictly greater cardinality than R.■

2.18: Binary Sequences are Uncountable

Let S be the set of all infinite sequences of 0s and 1s. Show that S is uncountable.

Strategy: Use Cantor's diagonal argument. Assume S is countable and list all sequences. Construct a new sequence that differs from the n-th sequence at the n-th position. This new sequence is not in the list, contradicting the assumption of countability.

Solution: Use Cantor's diagonal argument: assume S is countable and list all sequences. Construct a new sequence differing from the n-th sequence at the n-th place. This sequence is not in the list — contradiction. So S is uncountable. ■

2.19: Countability of Specific Sets

Show that the following sets are countable:

(a) Circles in the complex plane with rational radii and centers with rational coordinates.

(b) Any collection of disjoint intervals of positive length.

Strategy: For (a), each circle is determined by three rational numbers (radius and two coordinates), so the set injects into $\mathbb{Q}^3$ which is countable. For (b), each interval contains a rational number, and since intervals are disjoint, this creates an injection into $\mathbb{Q}$.

Solution:

(a) Each circle is determined by a rational radius and two rational coordinates $\to$ set is countable.

(b) Each disjoint interval must contain a distinct rational number $\to$ inject into $\mathbb{Q}$, which is countable.

■

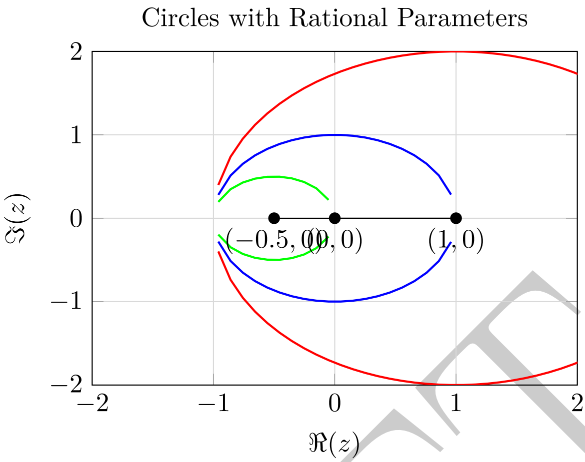


Figure 2.2: Examples of circles with rational centers and radii. Each circle is determined by three rational numbers (center coordinates and radius), making the set countable as it injects into $\mathbb{Q}^3$.

2.20: Countable Support for Real Function

Let f be a real-valued function on $[0, 1]$. Suppose there exists $M > 0$ such that for any finite set of points $\{x_1, \ldots, x_n\} \subset [0, 1]$,

$$|f(x_1)| + \cdots + |f(x_n)| \leq M$$

Let $S = \{x \in [0, 1] \mid f(x) \neq 0\}$. Prove that S is countable.

Strategy: Partition S into sets $S_k = \{x : |f(x)| > 1/k\}$ for each positive integer k. Show that each S_k is finite by using the bounded sum condition. Then S is a countable union of finite sets, hence countable.

Solution: Let $S = \{x \in [0, 1] \mid f(x) \neq 0\}$. We want to prove that S is countable. An element x is in S if and only if $|f(x)| > 0$. This is equivalent to saying that for each $x \in S$, there exists a positive integer k such that $|f(x)| > 1/k$.

Let's define a collection of sets based on this idea. For each positive integer k, let:

$$S_k = \left\{ x \in [0, 1] \mid |f(x)| > \frac{1}{k} \right\}$$

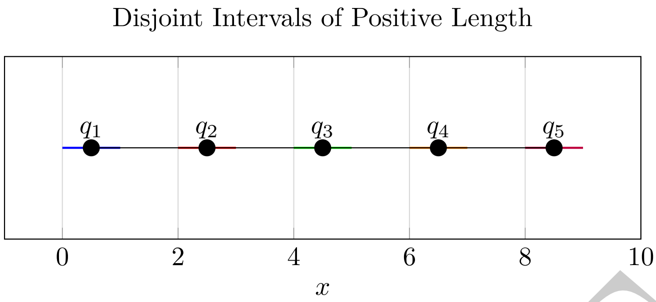


Figure 2.3: Disjoint intervals of positive length. Each interval contains a rational number, and since the intervals are disjoint, these rationals are distinct, creating an injection into $\mathbb{Q}$.

The set S is the union of all such sets:

$$S = \bigcup_{k=1}^{\infty} S_k$$

If we can prove that each set S_k is finite, then S will be a countable union of finite sets, which is itself a countable set.

Let's consider a specific set S_k. Let $\{x_1, x_2, \ldots, x_n\}$ be any finite collection of distinct points in S_k. By the definition of S_k, we have $|f(x_i)| > 1/k$ for each $i = 1, \ldots, n$. If we sum these values, we get:

$$|f(x_1)| + |f(x_2)| + \cdots + |f(x_n)| > \frac{1}{k} + \frac{1}{k} + \cdots + \frac{1}{k} = \frac{n}{k}$$

The problem states that for any finite set of points, this sum is bounded by M:

$$|f(x_1)| + |f(x_2)| + \cdots + |f(x_n)| \leq M$$

Combining these inequalities, we get:

$$\frac{n}{k} < \sum_{i=1}^{n} |f(x_i)| \leq M \implies \frac{n}{k} \leq M \implies n \leq kM$$

This result means that any finite subset of S_k can have at most kM elements. This implies that the set S_k itself must be finite and contain at most $\lfloor kM \rfloor$ elements.

Since each S_k is a finite set, their union $S = \bigcup_{k=1}^{\infty} S_k$ is a countable union of finite sets. Therefore, S is a countable set. ■

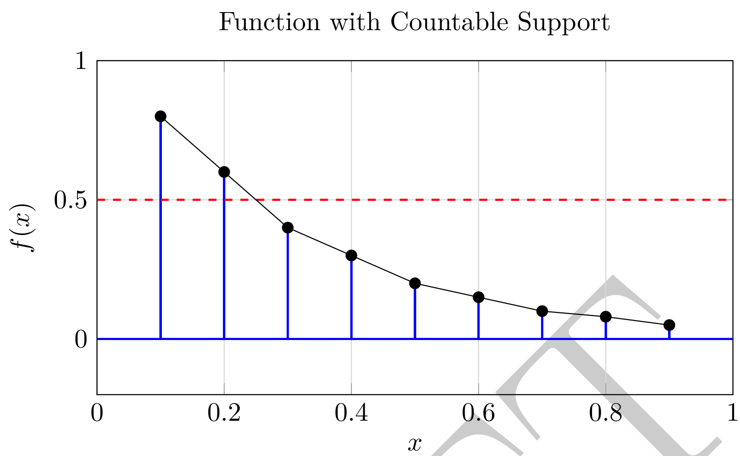


Figure 2.4: A function with countable support. The function is zero except at countably many points. The support $S = \{x : f(x) \neq 0\}$ is countable because each set $S_k = \{x : |f(x)| > 1/k\}$ is finite due to the bounded sum condition.

2.21: Fallacy in Countability of Intervals

Find the fallacy in the following "proof" that the set of all intervals of positive length is countable: Let $\{x_1, x_2, \ldots\}$ be the rationals. Every interval contains a rational x_n with minimal index n. Assign to the interval the smallest such n. This gives a function from intervals to $\mathbb{N}$, so the set of intervals is countable.

Strategy: Identify that the function is not injective. Many different intervals may contain the same rational with minimal index, so this does not establish a one-to-one correspondence between intervals and natural numbers.

Solution: The function F is not injective — many intervals may have the same smallest-index rational. So this does not establish a one-to-one correspondence between intervals and $\mathbb{N}$. Hence, the proof is invalid. ■

2.IV Additive Set Functions

Definitions and Theorems

Definition: Additive Function

A function $f : \mathcal{P}(S) \to \mathbb{R}$ is additive if $f(A \cup B) = f(A) + f(B)$ whenever $A \cap B = \emptyset$.

Theorem: Properties of Additive Functions

For an additive function $f : \mathcal{P}(S) \to \mathbb{R}$ and any sets $A, B \subseteq S$:

1. $f(A \cup B) = f(A) + f(B - A)$
2. $f(A \cup B) = f(A) + f(B) - f(A \cap B)$

Theorem: Inclusion-Exclusion Principle

For an additive function f and any sets $A_1, A_2, \ldots, A_n$:

$$f\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} f(A_i) - \sum_{1 \le i < j \le n} f(A_i \cap A_j) + \sum_{1 \le i < j < k \le n} f(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n+1} f\left(\bigcap_{i=1}^{n} A_i\right)$$

2.22: Additive Set Functions

Let S be the collection of all subsets of a given set T. A function $f : S \to \mathbb{R}$ is additive if:

$$f(A \cup B) = f(A) + f(B) \quad \text{whenever } A \cap B = \emptyset$$

Prove:

$$f(A \cup B) = f(A) + f(B - A), \quad f(A \cup B) = f(A) + f(B) - f(A \cap B)$$

Strategy: For the first identity, write $A \cup B$ as a disjoint union $A \cup (B - A)$ and apply the additivity property. For the second identity, decompose A and B into disjoint pieces using set differences and intersections, then use additivity and solve for the desired expression.

Solution: Let $f : S \to \mathbb{R}$ be an additive function, meaning $f(X \cup Y) = f(X) + f(Y)$ whenever $X \cap Y = \emptyset$.

Part 1: Prove $f(A \cup B) = f(A) + f(B - A)$. We can write the set $A \cup B$ as a disjoint union: $A \cup B = A \cup (B - A)$. The sets A and $B - A$ (the part of B not in A) are disjoint by definition. Using the additivity property on this disjoint union:

$$f(A \cup B) = f(A \cup (B - A)) = f(A) + f(B - A)$$

This proves the first identity.

Part 2: Prove $f(A \cup B) = f(A) + f(B) - f(A \cap B)$. We start by decomposing A and B into disjoint pieces. The set A can be written as the disjoint union $A = (A - B) \cup (A \cap B)$. By additivity:

$$f(A) = f(A - B) + f(A \cap B) \implies f(A - B) = f(A) - f(A \cap B)$$

The set B can be written as the disjoint union $B = (B - A) \cup (A \cap B)$. By additivity:

$$f(B) = f(B - A) + f(A \cap B) \implies f(B - A) = f(B) - f(A \cap B)$$

Now, we write $A \cup B$ as a union of three pairwise disjoint sets:

$$A \cup B = (A - B) \cup (B - A) \cup (A \cap B)$$

Using the additivity property:

$$f(A \cup B) = f(A - B) + f(B - A) + f(A \cap B)$$

Substitute the expressions we found for $f(A - B)$ and $f(B - A)$:

$$\begin{aligned} f(A \cup B) &= \big(f(A) - f(A \cap B)\big) + \big(f(B) - f(A \cap B)\big) + f(A \cap B) \\ &= f(A) + f(B) - f(A \cap B) - f(A \cap B) + f(A \cap B) \\ &= f(A) + f(B) - f(A \cap B) \end{aligned}$$

This proves the second identity. ∎

2.23: Solving for Total Measure from Functional Equations

Refer to Exercise 2.22. Assume f is additive and assume also that the following relations hold for two particular subsets A and B of T:

$$f(A \cup B) = f(A') + f(B') - f(A')f(B')$$

$$f(A \cap B) = f(A)f(B), \quad f(A) + f(B) \neq f(T),$$

where $A' = T - A, B' = T - B$. Prove that these relations determine $f(T)$, and compute the value of $f(T)$.

Strategy: Use the additivity property to express $f(A')$ and $f(B')$ in terms of $f(T)$, $f(A)$, and $f(B)$. Substitute these expressions into the first given equation. Then use the standard inclusion-exclusion formula for $f(A \cup B)$ and the second given relation to set up a quadratic equation in $f(T)$. Solve this equation and use the third condition to determine the unique solution.

Solution:

We are given that the function f is additive, meaning $f(X \cup Y) = f(X) + f(Y)$ for any disjoint sets X and Y. For any subset $X \subseteq T$, its complement $X' = T - X$ is disjoint from X and their union is T. The additive property therefore implies $f(T) = f(X) + f(X')$, which gives us:

- $f(A') = f(T) - f(A)$
- $f(B') = f(T) - f(B)$

We substitute these into the first given relation:

$$\begin{aligned} f(A \cup B) =& \big(f(T) - f(A)\big) + \big(f(T) - f(B)\big) \\ & - \big(f(T) - f(A)\big)\big(f(T) - f(B)\big) \\ =& 2f(T) - f(A) - f(B) \\ & - \big[f(T)^2 - f(T)f(A) - f(T)f(B) + f(A)f(B)\big] \\ =& 2f(T) - f(A) - f(B) - f(T)^2 \\ & + f(T)f(A) + f(T)f(B) - f(A)f(B) \end{aligned}$$

Next, we use the standard inclusion-exclusion principle for an additive function, which states $f(A \cup B) = f(A) + f(B) - f(A \cap B)$. From the second given relation, we know $f(A \cap B) = f(A)f(B)$. Substituting this gives:

$$f(A \cup B) = f(A) + f(B) - f(A)f(B)$$

Now we set the two expressions for $f(A \cup B)$ equal to each other:

$$\begin{aligned} & f(A) + f(B) - f(A)f(B) \\ =& 2f(T) - f(A) - f(B) - f(T)^2 + f(T)f(A) + f(T)f(B) - f(A)f(B) \end{aligned}$$

The $-f(A)f(B)$ terms on each side cancel. We move all remaining terms to one side to form a quadratic equation in terms of $f(T)$:

$$f(T)^2 - 2f(T) - f(T)f(A) - f(T)f(B) + 2f(A) + 2f(B) = 0$$

Factoring out $f(T)$ and constant terms:

$$f(T)^2 - f(T)\big(2 + f(A) + f(B)\big) + 2\big(f(A) + f(B)\big) = 0$$

This quadratic equation can be factored as:

$$\big(f(T) - 2\big)\big(f(T) - [f(A) + f(B)]\big) = 0$$

This implies two possible solutions for $f(T)$:

1. $f(T) = 2$
2. $f(T) = f(A) + f(B)$

Finally, we use the third given relation, $f(A) + f(B) \neq f(T)$, to eliminate the second possibility. Therefore, the relations uniquely determine the value of $f(T)$. That value is 2.

$$\boxed{f(T) = 2}$$

■

2.V Solving and Proving Techniques

Proving Set Equality

- Use element-chasing: assume $x \in$ one side and show $x \in$ the other side, then reverse
- Use set-theoretic identities and properties of union, intersection, and set difference

- Break down complex sets into simpler components using distributive laws
- Consider cases when dealing with piecewise definitions or different scenarios

Proving Function Properties

- For injectivity: assume $f(x_1) = f(x_2)$ and show $x_1 = x_2$
- For surjectivity: given any y in codomain, find an x in domain such that $f(x) = y$
- For bijectivity: prove both injectivity and surjectivity
- Use the definition of function composition: $(g \circ f)(x) = g(f(x))$
- For piecewise functions, determine which piece applies based on the input value

Proving Relation Properties

- For reflexivity: check if $(x, x) \in R$ for all x in the set
- For symmetry: check if $(x, y) \in R$ implies $(y, x) \in R$
- For transitivity: check if $(x, y), (y, z) \in R$ implies $(x, z) \in R$
- Use specific counterexamples to disprove properties
- Consider edge cases and special values when testing properties

Proving Countability

- Show a bijection exists with $\mathbb{N}$ or a known countable set
- Use the fact that countable unions of countable sets are countable
- Show the set injects into a known countable set
- For finite sets, count the elements directly
- Use the fact that removing a countable set from an uncountable set leaves an uncountable set

Proving Uncountability

- Use Cantor's diagonal argument to show no bijection exists with $\mathbb{N}$
- Show the set has strictly greater cardinality than a known uncountable set
- Use the fact that power sets have strictly greater cardinality than the original set
- Construct a contradiction by assuming countability and finding an element not in any enumeration

Working with Additive Functions

- Use the definition: $f(A \cup B) = f(A) + f(B)$ when $A \cap B = \emptyset$
- Decompose sets into disjoint unions to apply additivity
- Use inclusion-exclusion principle for overlapping sets
- Express complements in terms of the universal set: $f(A') = f(T) - f(A)$
- Set up equations using known relationships and solve for unknown values

Proving Equivalence of Conditions

1. Show each condition implies the next one in the chain
2. Show the last condition implies the first one (completing the cycle)
3. Use the fact that if $A \Rightarrow B$ and $B \Rightarrow A$, then $A \iff B$
4. Use contrapositive when direct implication is difficult
5. Break complex conditions into simpler logical components

Constructing Bijections

1. Identify the domain and codomain clearly
2. Define the function rule explicitly
3. Prove injectivity by showing $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$
4. Prove surjectivity by showing every element in codomain has a preimage
5. Use piecewise definitions when different rules apply to different parts of the domain

Working with Images and Preimages

- Use definitions: $f(A) = \{f(x) : x \in A\}$ and $f^{-1}(Y) = \{x : f(x) \in Y\}$
- Remember that $f(A \cap B) \subseteq f(A) \cap f(B)$ with equality only for injective functions
- Use the fact that $f(A \cup B) = f(A) \cup f(B)$ always holds
- For inverse images, use $f^{-1}(Y_1 \cup Y_2) = f^{-1}(Y_1) \cup f^{-1}(Y_2)$
- Remember that $f[f^{-1}(Y)] = Y$ if and only if f is surjective

Chapter 3

Elements of Point Set Topology

3.I Open and Closed Sets in $\mathbb{R}^1$ and $\mathbb{R}^2$

Definitions and Theorems

Definition: Open Set

A set S in a metric space (M, d) is said to be open if for every point $x \in S$, there exists a positive number ε such that the open ball $B(x; \varepsilon) = \{y \in M : d(x, y) < \varepsilon\}$ is entirely contained in S.

Importance: Open sets are the fundamental building blocks of topology. They formalize the intuitive notion of "neighborhoods" and provide the foundation for understanding continuity, convergence, and many other topological concepts. Open sets are essential for all areas of analysis, topology, and modern mathematics.

Definition: Closed Set

A set S in a metric space (M, d) is said to be closed if its complement $M \setminus S$ is an open set.

Importance: Closed sets are the dual concept to open sets and are equally fundamental in topology. They capture the idea of sets that contain all their limit points and are essential for understanding compactness, completeness, and many other important topological properties. Closed sets are crucial for analysis and many applications in mathematics.

Definition: Interior Point

A point x in a metric space (M, d) is said to be an interior point of a set $S \subseteq M$ if there exists a positive number ε such that the open ball $B(x; \varepsilon)$ is entirely contained in S.

Importance: Interior points formalize the intuitive notion of points that are "inside" a set, not on the boundary. They are essential for understanding the structure of sets and are crucial for defining the interior of a set. Interior points are fundamental for topology and many areas of analysis.

Definition: Accumulation Point

A point x in a metric space (M, d) is said to be an accumulation point (or limit point) of a set $S \subseteq M$ if every open ball centered at x contains at least one point of S different from x itself.

Importance: Accumulation points are fundamental for understanding the structure of sets and convergence. They capture the idea of points that can be approached by sequences in the set and are essential for defining closed sets, compactness, and many other important topological concepts. Accumulation points are crucial for analysis and topology.

Definition: Neighborhood

A neighborhood of a point x in a metric space (M, d) is any open set that contains x.

Importance: Neighborhoods formalize the intuitive notion of "nearness" and provide a way to talk about local properties of spaces and functions. They are essential for understanding continuity, convergence, and many other topological concepts. Neighborhoods are fundamental for topology and analysis.

Theorem: Basic Properties of Open and Closed Sets

In any metric space (M, d), the following properties hold:

1. The empty set $\emptyset$ and the whole space M are both open and closed sets.
2. The union of any collection of open sets is an open set.
3. The intersection of any finite collection of open sets is an open set.
4. The intersection of any collection of closed sets is a closed set.
5. The union of any finite collection of closed sets is a closed set.

Importance: These properties are the fundamental axioms that define a topology. They establish the basic rules for how open and closed sets behave under set operations and are essential for all of topology and analysis. These properties are used constantly in proofs and are the foundation for understanding topological spaces.

Theorem: Characterization of Connectedness

A metric space (M, d) is connected if and only if the only subsets of M that are both open and closed are the empty set $\emptyset$ and the whole space M itself.

Importance: This theorem provides a fundamental characterization of connectedness in terms of open and closed sets. It gives a practical way to test whether a space is connected and is essential for understanding the structure of topological spaces. Connectedness is crucial for many results in analysis, including the intermediate value theorem.

Definition: Dense Set

A set A in a metric space (M, d) is said to be dense in M if every point of M is either in A or is an accumulation point of A. Equivalently, A is dense in M if the closure of A equals M, that is, $\overline{A} = M$.

Importance: Dense sets are fundamental for understanding the structure of metric spaces and are essential for approximation theory. They provide a way to approximate any point in a space using points from a subset. Dense sets are crucial for many results in analysis, functional analysis, and approximation theory.

Theorem: Density of Rational and Irrational Numbers

Both the set of rational numbers and the set of irrational numbers are dense in the real line $\mathbb{R}$.

Importance: This theorem shows that both rational and irrational numbers are densely distributed throughout the real number line. This property is fundamental for understanding the structure of the real numbers and is essential for many results in analysis, approximation theory, and number theory. It demonstrates the richness of the real number system.

3.1: Open and Closed Intervals

Prove that an open interval in $\mathbb{R}^1$ is an open set and that a closed interval is a closed set.

Strategy: Use the definition of open sets (every point is an interior point) and closed sets (complement is open). For open intervals, show that every point has a neighborhood contained in the interval. For closed intervals, show that the complement is open by finding neighborhoods for points outside the interval.

Solution: Let (a, b) be an open interval in $\mathbb{R}^1$. To show it's open, we need to prove that every point $x \in (a, b)$ is an interior point. For any $x \in (a, b)$, let $\varepsilon = \min\{x - a, b - x\}$. Then the open ball $B(x, \varepsilon) = (x - \varepsilon, x + \varepsilon)$ is contained entirely within (a, b). This shows that every point in (a, b) is an interior point, so (a, b) is open.

For a closed interval $[a, b]$, we need to show its complement $\mathbb{R} \setminus [a, b] = (-\infty, a) \cup (b, \infty)$ is open. Any point x in this complement is either less than a or greater than b. If $x < a$, let $\varepsilon = a - x$, then $B(x, \varepsilon) = (x - \varepsilon, x + \varepsilon) \subset (-\infty, a)$. If $x > b$, let $\varepsilon = x - b$, then $B(x, \varepsilon) \subset (b, \infty)$. This shows the complement is open, so $[a, b]$ is closed. ■

3.2: Accumulation Points and Set Properties

Determine all the accumulation points of the following sets in $\mathbb{R}^1$ and decide whether the sets are open or closed (or neither).

(a) All integers.

(b) The interval (a, b).

(c) All numbers of the form $1/n$, $(n = 1, 2, 3, \ldots)$.

(d) All rational numbers.

(e) All numbers of the form $2^{-n} + 5^{-m}$, $(m, n = 1, 2, \ldots)$.

(f) All numbers of the form $(-1)^n + (1/m)$, $(m, n = 1, 2, \ldots)$.

(g) All numbers of the form $(1/n) + (1/m)$, $(m, n = 1, 2, \ldots)$.

(h) All numbers of the form $(-1)^n/[1 + (1/n)]$, $(n = 1, 2, \ldots)$.

Strategy: For each set, identify accumulation points by finding points that can be approached by sequences in the set. For openness/closedness, check if every point is interior (open) and if the complement is open (closed). Use density properties of rationals and convergence of sequences.

Solution:

(a) The set of integers has no accumulation points since each integer has a neighborhood containing no other integers. The set is closed (its complement is open) but not open.

(b) The interval (a, b) has accumulation points $[a, b]$. For any $x \in [a, b]$, a sequence $\{x_n\} \subset (a, b)$ with $x_n \to x$ exists (e.g., $x_n = x + (b - a)/(n + 1)$ if $x < b$, or $x_n = a + (b - a)/(n + 1)$ if $x = a$). The set is open (every point is interior) but not closed (its closure is $[a, b]$).

(c) The set $\{1/n : n \in \mathbb{N}\}$ has 0 as its only accumulation point. The set is not closed, because its closure includes 0. It is also not open, as no point in the set has a neighborhood entirely contained within the set. Therefore, the set is neither open nor closed.

(d) The set of rational numbers has all real numbers as accumulation points. The set is neither open nor closed.

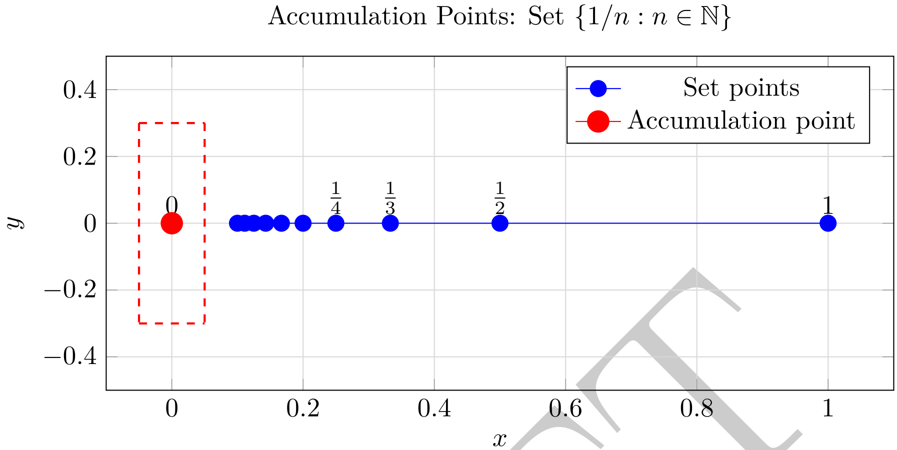


Figure 3.1: The set $\{1/n : n \in \mathbb{N}\}$ has 0 as its only accumulation point. Any neighborhood of 0 contains infinitely many points from the set.

(e) The set $\{2^{-n} + 5^{-m} : m, n \in \mathbb{N}\}$ has accumulation points $\{2^{-n} + 5^{-m} : m, n \in \mathbb{N}\} \cup \{2^{-n} : n \in \mathbb{N}\} \cup \{5^{-m} : m \in \mathbb{N}\} \cup \{0\}$. For any $x = 2^{-k} + 5^{-l}$, take $m_n = n + l$, so $2^{-k} + 5^{-m_n} \to 2^{-k}$. Similarly, for $x = 5^{-l}$, take $n_m = m + k$, so $2^{-n_m} + 5^{-l} \to 5^{-l}$. For $x = 0$, take $n = m$, so $2^{-n} + 5^{-n} \to 0$. The set is neither open nor closed (no point is interior; closure includes accumulation points).

(f) The set $\{1/n + 1/m : m, n \in \mathbb{N}\}$ has accumulation points $\{k/n : k, n \in \mathbb{N}, k \leq n\} \cup \{0\}$. For $x = k/n$, take $m_i = i + n$, so $1/n + 1/m_i \to 1/n$; for k/n with $k \geq 2$, set $m = n_i = i$, so $(k-1)/i + 1/i = k/i \to k/n$. For $x = 0$, take $n = m = i$, so $1/i + 1/i \to 0$. The set is neither open nor closed (no point is interior; closure includes accumulation points).

(h) The set $\{(-1)^n/(1 + 1/n) : n \in \mathbb{N}\}$ has accumulation points $\{-1, 1\}$. The set is neither open nor closed. ■

3.3: Accumulation Points and Set Properties in $\mathbb{R}^2$

The same as Exercise 3.2 for the following sets in $\mathbb{R}^2$:

(a) All complex z such that $|z| > 1$.

(b) All complex z such that $|z| \geq 1$.

(c) All complex numbers of the form $(1/n) + (i/m)$, $(m, n = 1, 2, \dots)$.

(d) All points (x, y) such that $x^2 - y^2 < 1$.

(e) All points (x, y) such that $x > 0$.

(f) All points (x, y) such that $x \geq 0$.

Strategy: Apply the same approach as in Exercise 3.2 but in two dimensions. For complex numbers, use the modulus $|z|$ to determine boundaries. For sequences, consider convergence in each coordinate separately. Use geometric intuition for open/closed sets in the plane.

Solution:

(a) The set $\{z \in \mathbb{C} : |z| > 1\}$ has accumulation points $\{z \in \mathbb{C} : |z| \geq 1\}$. The set is open but not closed.

(b) The set $\{z \in \mathbb{C} : |z| \geq 1\}$ has accumulation points $\{z \in \mathbb{C} : |z| \geq 1\}$. The set is closed but not open.

(c) The set $\{(1/n, 1/m) : m, n \in \mathbb{N}\}$ has accumulation points $\{(1/n, 0) : n \in \mathbb{N}\} \cup \{(0, 1/m) : m \in \mathbb{N}\} \cup \{(0, 0)\}$. For $(1/k, 0)$, take $(1/k, 1/m_n)$ with $m_n \to \infty$; for $(0, 1/l)$, take $(1/n_m, 1/l)$ with $n_m \to \infty$; for $(0, 0)$, take $(1/n, 1/n)$. The set is neither open nor closed (no point is interior; closure includes accumulation points).

(d) The set $\{(x, y) : x^2 - y^2 < 1\}$ has accumulation points $\{(x, y) : x^2 - y^2 \leq 1\}$. The set is open but not closed.

(e) The set $\{(x, y) : x > 0\}$ has accumulation points $\{(x, y) : x \geq 0\}$. The set is open but not closed.

(f) The set $\{(x, y) : x \geq 0\}$ has accumulation points $\{(x, y) : x \geq 0\}$. The set is closed but not open.

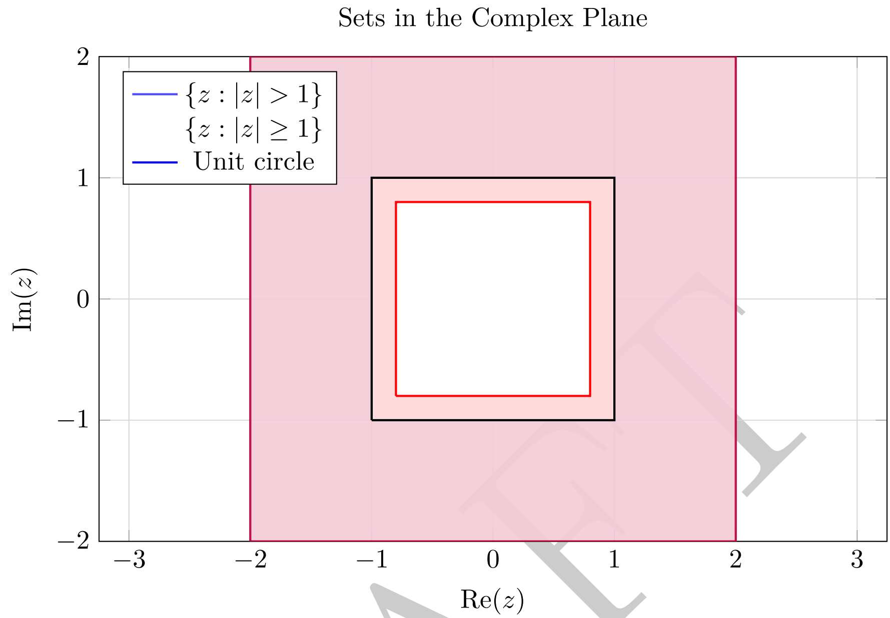


Figure 3.2: Left: $\{z : |z| > 1\}$ is open but not closed. Right: $\{z : |z| \geq 1\}$ is closed but not open.

■

3.4: Rational and Irrational Elements in Open Sets

Prove that every nonempty open set S in $\mathbb{R}^1$ contains both rational and irrational numbers.

Strategy: Use the density of rational and irrational numbers in $\mathbb{R}$. Since S is open, any point in S has a neighborhood contained in S. By density, both rational and irrational numbers exist in any open interval.

Solution: Let S be a nonempty open set in $\mathbb{R}^1$. Since S is open, for any point $x \in S$, there exists $\varepsilon > 0$ such that the open interval $(x - \varepsilon, x + \varepsilon) \subset S$.

Since the rational numbers are dense in $\mathbb{R}$, there exists a rational number q in $(x-\varepsilon, x+\varepsilon)$, and thus $q \in S$.

Similarly, since the irrational numbers are also dense in $\mathbb{R}$, there exists an irrational number r in $(x-\varepsilon, x+\varepsilon)$, and thus $r \in S$.

Therefore, every nonempty open set contains both rational and irrational numbers. ■

3.5: Open and Closed Sets in $\mathbb{R}^1$ and $\mathbb{R}^2$

Prove that the only sets in $\mathbb{R}^1$ which are both open and closed are the empty set and $\mathbb{R}^1$ itself. Is a similar statement true for $\mathbb{R}^2$?

Strategy: Use proof by contradiction. Assume there exists a non-empty proper subset that is both open and closed. Use the connectedness of $\mathbb{R}^1$ and $\mathbb{R}^2$ - if a space is connected, it cannot be split into two non-empty disjoint open sets. The key is to show that such a split would violate connectedness.

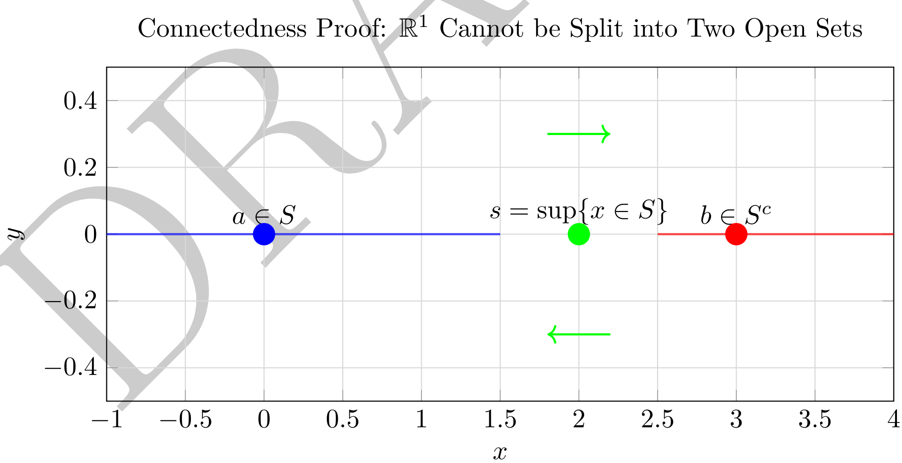


Figure 3.3: Proof by contradiction that $\mathbb{R}^1$ is connected. If it could be split into two open sets, the supremum s would lead to a contradiction.

Solution:

Proof for $\mathbb{R}^1$

We need to show that if a set S in $\mathbb{R}^1$ is both open and closed, then S must be either the empty set $\emptyset$ or all of $\mathbb{R}^1$.

Let's start by understanding what it means for a set to be both open and closed. If S is open, then every point in S has a small neighborhood around it that stays within S. If S is closed, then its complement $S^c = \mathbb{R}^1 \setminus S$ is open.

Now, let's prove this by contradiction. Suppose there exists a set S that is both open and closed, but S is not empty and S is not all of $\mathbb{R}^1$. This means:

1. S is not empty (there's at least one point in S)
2. S is not all of $\mathbb{R}^1$ (there's at least one point not in S)

Since S is not all of $\mathbb{R}^1$, its complement S^c is not empty. And since S is closed, S^c must be open.

So we have two non-empty open sets S and S^c that together make up all of $\mathbb{R}^1$, and they don't overlap (they're disjoint).

Now, let's pick a point a from S and a point b from S^c. Without loss of generality, assume $a < b$.

Consider the interval $[a, b]$. Since S is open and contains a, there must be some small distance $\varepsilon_1 > 0$ such that the interval $(a-\varepsilon_1, a+\varepsilon_1)$ is completely contained in S.

Similarly, since S^c is open and contains b, there must be some small distance $\varepsilon_2 > 0$ such that the interval $(b - \varepsilon_2, b + \varepsilon_2)$ is completely contained in S^c.

Now, let's look at the set of all points in $[a, b]$ that belong to S. This set has a supremum (least upper bound) because it's bounded above by b. Let's call this supremum s.

The key insight is that s must belong to S. Here's why: if s were in S^c, then since S^c is open, there would be a small interval around s that's completely in S^c. But this would mean there are points in S^c that are larger than s, which contradicts the fact that s is the supremum of points in S.

So s is in S. But since S is open, there must be a small interval around s that's completely contained in S. This means there are points in S that are larger than s, which again contradicts the fact that s is the supremum.

This contradiction shows that our original assumption was wrong. Therefore, the only sets in $\mathbb{R}^1$ that are both open and closed are the empty set and $\mathbb{R}^1$ itself.

For $\mathbb{R}^2$

Yes, the same statement is true for $\mathbb{R}^2$. The only subsets of $\mathbb{R}^2$ that are both open and closed are the empty set and $\mathbb{R}^2$ itself.

The proof is similar in spirit but more complex because we're working in two dimensions. The key idea is that $\mathbb{R}^2$ is "connected" - you can draw a continuous path between any two points without leaving $\mathbb{R}^2$.

If there were a non-empty proper subset S of $\mathbb{R}^2$ that was both open and closed, then its complement S^c would also be non-empty, open, and closed. You could then pick a point from S and a point from S^c and try to draw a continuous path between them. But this path would have to "jump" from one set to the other at some point, which would violate the continuity of the path.

This property is called "connectedness" - a space is connected if it cannot be split into two non-empty, disjoint, open sets. Both $\mathbb{R}^1$ and $\mathbb{R}^2$ are connected spaces. ■

3.6: Closed Sets as Intersection of Open Sets

Prove that every closed set in $\mathbb{R}^1$ is the intersection of a countable collection of open sets.

Strategy: Use the distance function $d(x, F) = \inf\{|x - y| : y \in F\}$ to construct open sets $G_n = \{x : d(x, F) < 1/n\}$. Show that the intersection of these open sets equals the closed set F by using the fact that $d(x, F) = 0$ if and only if $x \in F$ (when F is closed).

Solution: Let F be a closed set in $\mathbb{R}^1$. For each $n \in \mathbb{N}$, define $G_n = \{x \in \mathbb{R} : d(x, F) < 1/n\}$, where $d(x, F) = \inf\{|x - y| : y \in F\}$. Each G_n is open since it's the union of open intervals.

We claim that $F = \bigcap_{n=1}^{\infty} G_n$. Clearly $F \subset \bigcap_{n=1}^{\infty} G_n$ since every point in F has distance 0 to F.

For the reverse inclusion, let $x \in \bigcap_{n=1}^{\infty} G_n$. Then $d(x, F) < 1/n$ for all n, which means $d(x, F) = 0$. Since F is closed, this implies $x \in F$. ■

3.7: Structure of Bounded Closed Sets in $\mathbb{R}^1$

Prove that a nonempty, bounded closed set S in $\mathbb{R}^1$ is either a closed interval, or that S can be obtained from a closed interval by removing a countable disjoint collection of open intervals whose endpoints belong to S.

Strategy: Use the fact that a bounded closed set has a minimum and maximum. If S is not a closed interval, its complement in the minimal closed interval containing S is open and can be written as a countable union of disjoint open intervals. Since S is closed, the endpoints of these intervals must belong to S.

Solution: Let S be a nonempty, bounded closed set in $\mathbb{R}^1$. Let $a = \inf S$ and $b = \sup S$. Since S is closed, $a, b \in S$.

If $S = [a, b]$, we're done. Otherwise, the complement $[a, b] \setminus S$ is open and can be written as a countable union of disjoint open intervals (a_i, b_i). Since S is closed, the endpoints a_i, b_i must belong to S.

Therefore, $S = [a, b] \setminus \bigcup_{i=1}^{\infty}(a_i, b_i)$, which is the desired representation. ■

3.II Open and Closed Sets in $\mathbb{R}^n$

Definitions and Theorems

Definition: Open Ball

Given a metric space (M, d), the open ball centered at a point $a \in M$ with radius $r > 0$ is the set $B(a; r) = \{x \in M : d(x, a) < r\}$.

Definition: Interior of a Set

Given a set S in a metric space (M, d), the interior of S, denoted by int S, is the set of all interior points of S.

Theorem: Properties of the Interior

Let S and T be subsets of a metric space (M, d). Then the following properties hold:

1. The interior int S is an open set.
2. The interior int S is the largest open subset of S.
3. The interior of the interior equals the interior: int (int S) = int S.
4. The interior of an intersection equals the intersection of interiors: int $(S \cap T) =$ int $S \cap$ int T.
5. The union of interiors is contained in the interior of the union: int $S \cup$ int $T \subseteq$ int $(S \cup T)$.

Theorem: Interior as Union of Open Subsets

If S is a subset of $\mathbb{R}^n$, then the interior of S is equal to the union of all open subsets of $\mathbb{R}^n$ that are contained in S.

3.8: Open Balls and Intervals in Rn

Prove that open n-balls and n-dimensional open intervals are open sets in $\mathbb{R}^n$.

Strategy: For open balls, use the triangle inequality to show that any point in the ball has a neighborhood contained in the ball. For open intervals, use the fact that they are products of open intervals in each coordinate and show that any point has a neighborhood contained in the product.

Solution: Let $B(a; r) = \{x \in \mathbb{R}^n : \|x - a\| < r\}$ be an open ball centered at a with radius r. For any $x \in B(a; r)$, let $\varepsilon = r - \|x - a\| > 0$. Then $B(x; \varepsilon) \subset B(a; r)$ by the triangle inequality, showing $B(a; r)$ is open.

For an open interval $I = (a_1, b_1) \times \cdots \times (a_n, b_n)$, let $x = (x_1, \ldots, x_n) \in I$. For each i, let $\varepsilon_i = \min\{x_i - a_i, b_i - x_i\}$. Then the ball $B(x; \min\{\varepsilon_1, \ldots, \varepsilon_n\}) \subset I$, showing I is open. ■

3.9: Interior of a Set is Open

Prove that the interior of a set in $\mathbb{R}^n$ is open in $\mathbb{R}^n$.

Strategy: Use the definition of interior point: a point is interior if it has a neighborhood contained in the set. Show that if a point is interior, then every point in a small enough neighborhood around it is also interior, making the interior set itself open.

Solution: Let $S \subset \mathbb{R}^n$ and let $x \in \text{int } S$. By definition, there exists $\varepsilon > 0$ such that $B(x; \varepsilon) \subset S$.

For any $y \in B(x; \varepsilon)$, let $\delta = \varepsilon - \|y - x\| > 0$. Then $B(y; \delta) \subset B(x; \varepsilon) \subset S$, which shows that $y \in \text{int } S$.

Therefore, $B(x; \varepsilon) \subset \text{int } S$, proving that $\text{int } S$ is open. ■

3.10: Interior as Union of Open Subsets

If $S \subseteq \mathbb{R}^n$, prove that $\text{int } S$ is the union of all open subsets of $\mathbb{R}^n$ which are contained in S. This is described by saying that $\text{int } S$ is the largest open subset of S.

Strategy: Show two inclusions: (1) any open subset contained in S is contained in $\text{int } S$ (since all its points are interior), and (2) $\text{int } S$ is contained in the union of all open subsets of S (since $\text{int } S$ itself is an open subset of S).

Solution: Let $\mathcal{U}$ be the collection of all open subsets of $\mathbb{R}^n$ contained in S. We need to show that $\text{int } S = \bigcup_{U \in \mathcal{U}} U$.

First, if $x \in \text{int } S$, then there exists $\varepsilon > 0$ such that $B(x; \varepsilon) \subset S$. Since $B(x; \varepsilon)$ is open and contained in S, we have $x \in B(x; \varepsilon) \in \mathcal{U}$, so $x \in \bigcup_{U \in \mathcal{U}} U$.

Conversely, if $x \in \bigcup_{U \in \mathcal{U}} U$, then $x \in U$ for some open set $U \subset S$. Since U is open, there exists $\varepsilon > 0$ such that $B(x;\varepsilon) \subset U \subset S$, which shows $x \in \text{int } S$.

Therefore, $\text{int } S = \bigcup_{U \in \mathcal{U}} U$, proving that the interior is the largest open subset of S. ■

3.11: Interior of Intersection and Union

If S and T are subsets of $\mathbb{R}^n$, prove that $\text{int}(S) \cap \text{int}(T) = \text{int}(S \cap T)$, and $\text{int}(S) \cup \text{int}(T) \subseteq \text{int}(S \cup T)$.

Strategy: For the first equality, show both inclusions using the fact that if a point is interior to both sets, it has a neighborhood in their intersection. For the second inclusion, use the fact that if a point is interior to either set, it has a neighborhood in their union.

Solution: For the first equality, let $x \in \text{int}(S) \cap \text{int}(T)$. Then there exist $\varepsilon_1, \varepsilon_2 > 0$ such that $B(x;\varepsilon_1) \subset S$ and $B(x;\varepsilon_2) \subset T$. Let $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. Then $B(x;\varepsilon) \subset S \cap T$, so $x \in \text{int}(S \cap T)$.

Conversely, if $x \in \text{int}(S \cap T)$, there exists $\varepsilon > 0$ such that $B(x;\varepsilon) \subset S \cap T$. This implies $B(x;\varepsilon) \subset S$ and $B(x;\varepsilon) \subset T$, so $x \in \text{int}(S) \cap \text{int}(T)$.

For the second inclusion, if $x \in \text{int}(S) \cup \text{int}(T)$, then $x \in \text{int}(S)$ or $x \in \text{int}(T)$. In either case, there exists $\varepsilon > 0$ such that $B(x;\varepsilon) \subset S$ or $B(x;\varepsilon) \subset T$, which implies $B(x;\varepsilon) \subset S \cup T$. Therefore, $x \in \text{int}(S \cup T)$. ■

3.12: Properties of Derived Set and Closure

Let S' denote the derived set and $\overline{S}$ the closure of a set S in $\mathbb{R}^n$. Prove that:

a) S' is closed in $\mathbb{R}^n$; that is, $\overline{S'} \subseteq S'$.

b) If $S \subseteq T$, then $S' \subseteq T'$.

c) $S' \cup T' = (S \cup T)'$.

d) $\overline{S} = S \cup S'$.

e) $\overline{S}$ is closed in $\mathbb{R}^n$.

f) $\overline{S}$ is the intersection of all closed subsets of $\mathbb{R}^n$ containing S. That is, $\overline{S}$ is the smallest closed set containing S.

Strategy: Use the definitions of derived set (accumulation points) and closure (adherent points). For (a), show that accumulation points of accumulation points are accumulation points. For (b), use the subset relationship. For (c), show both inclusions using the definition. For (d), use the fact that adherent points are either in the set or accumulation points. For (e), use (a) and (d). For (f), show that closure is both contained in and contains the intersection.

Solution:
(a) To prove S' is closed, we must show that its derived set $(S')'$ is a subset of S'. Let $\mathbf{x} \in (S')'$. This means every neighborhood of $\mathbf{x}$ contains a point of S' other than $\mathbf{x}$. Let $B(\mathbf{x}, \varepsilon)$ be an arbitrary open ball centered at $\mathbf{x}$. By definition of $(S')'$, there is a point $\mathbf{y} \in B(\mathbf{x}, \varepsilon) \cap S'$. Since $B(\mathbf{x}, \varepsilon)$ is an open set, it is a neighborhood for $\mathbf{y}$. Because $\mathbf{y} \in S'$, $\mathbf{y}$ is an accumulation point of S, so this neighborhood must contain infinitely many points from S. Thus, the ball $B(\mathbf{x}, \varepsilon)$ contains infinitely many points from S. As $B(\mathbf{x}, \varepsilon)$ was an arbitrary neighborhood of $\mathbf{x}$, this shows that $\mathbf{x}$ is an accumulation point of S, so $\mathbf{x} \in S'$. Therefore, $(S')' \subseteq S'$, which proves that S' is a closed set.

(b) Let $\mathbf{x} \in S'$. Then every neighborhood of $\mathbf{x}$ contains a point $\mathbf{y} \in S$ with $\mathbf{y} \neq \mathbf{x}$. Since $S \subseteq T$, this point $\mathbf{y}$ is also in T. Thus, every neighborhood of $\mathbf{x}$ contains a point $\mathbf{y} \in T$ with $\mathbf{y} \neq \mathbf{x}$. This means $\mathbf{x} \in T'$. So $S' \subseteq T'$.

(c) Using (b), since $S \subseteq S \cup T$ and $T \subseteq S \cup T$, we have $S' \subseteq (S \cup T)'$ and $T' \subseteq (S \cup T)'$. Therefore, $S' \cup T' \subseteq (S \cup T)'$. For the reverse inclusion, let $\mathbf{x} \in (S \cup T)'$. If $\mathbf{x} \notin S'$, then there is a neighborhood of $\mathbf{x}$ that contains no points of S (other than possibly $\mathbf{x}$). But since $\mathbf{x} \in (S \cup T)'$, this neighborhood must contain infinitely many points from $S \cup T$. These points must therefore come from T. This implies $\mathbf{x} \in T'$. So, every point in $(S \cup T)'$ must be in S' or T'. Thus, $(S \cup T)' \subseteq S' \cup T'$.

(d) The closure $\overline{S}$ consists of all points adherent to S. A point $\mathbf{x}$ is adherent to S if every neighborhood of $\mathbf{x}$ intersects S. ($\overline{S} \subseteq S \cup S'$): Let $\mathbf{x} \in \overline{S}$. If $\mathbf{x} \in S$, we are done. If $\mathbf{x} \notin S$, then every neighborhood of $\mathbf{x}$ must contain a point from S, and that point cannot be $\mathbf{x}$. This is the definition of an accumulation point, so $\mathbf{x} \in S'$. Thus $\overline{S} \subseteq S \cup S'$. ($S \cup S' \subseteq \overline{S}$): If $\mathbf{x} \in S$, it is in $\overline{S}$ because every neighborhood contains $\mathbf{x}$. If $\mathbf{x} \in S'$, every neighborhood contains a point of S, so $\mathbf{x}$ is an adherent point. Thus $S' \subseteq \overline{S}$. This gives $S \cup S' \subseteq \overline{S}$.

(e) To prove $\overline{S}$ is closed, we show its derived set $(\overline{S})'$ is a subset of $\overline{S}$. From (d), $\overline{S} = S \cup S'$. Using (c), we get $(\overline{S})' = (S \cup S')' = S' \cup (S')'$. From (a), S' is closed, which means $(S')' \subseteq S'$. Therefore, $(\overline{S})' \subseteq S' \cup S' = S'$. Since $S' \subseteq S \cup S' = \overline{S}$, we have $(\overline{S})' \subseteq \overline{S}$. This proves that $\overline{S}$ is closed.

(f) Let $\mathcal{C}$ be the collection of all closed sets containing S. Let $C_{min} = \bigcap_{F \in \mathcal{C}} F$. ($\overline{S} \subseteq C_{min}$): Let F be any set in $\mathcal{C}$. Then F is closed and $S \subseteq F$. The closure of a set is the smallest closed set containing it, so we must have $\overline{S} \subseteq F$. Since this holds for all $F \in \mathcal{C}$, we have $\overline{S} \subseteq \bigcap_{F \in \mathcal{C}} F = C_{min}$. ($C_{min} \subseteq \overline{S}$): By part (e), $\overline{S}$ is a closed set. It also contains S. Therefore, $\overline{S}$ is one of the sets in the collection $\mathcal{C}$. The intersection of all sets in $\mathcal{C}$ must be a subset of any particular member, so $C_{min} \subseteq \overline{S}$. Thus, $\overline{S} = C_{min}$. ■

3.13: Closure under Intersection of Sets

Let S and T be subsets of $\mathbb{R}^k$. Prove that $\overline{S \cup T} = \overline{S} \cup \overline{T}$ and that $\overline{S \cap T} \subseteq \overline{S} \cap \overline{T}$ if S is open.

NOTE. The statements in Exercises 3.9 through 3.13 are true in any metric space.

Strategy: For the union, use the monotonicity of closure and show both inclusions. For the intersection, use the fact that if a point is adherent to the intersection, it's adherent to both sets. The openness of S is not needed for the intersection inclusion.

Solution: We use only the definition of closure via adherent points.

For the union, first note $S \subseteq S \cup T$ and $T \subseteq S \cup T$, so by monotonicity of closure,

$$\overline{S} \subseteq \overline{S \cup T}, \qquad \overline{T} \subseteq \overline{S \cup T},$$

whence $\overline{S} \cup \overline{T} \subseteq \overline{S \cup T}$. Conversely, if $x \in \overline{S \cup T}$, then every neighborhood of x meets $S \cup T$, hence meets S or T. Therefore $x \in \overline{S} \cup \overline{T}$. Thus $\overline{S \cup T} = \overline{S} \cup \overline{T}$.

For the intersection, if $x \in \overline{S \cap T}$ then every neighborhood of x meets $S \cap T$, hence meets both S and T. Therefore $x \in \overline{S}$ and $x \in \overline{T}$, so

$$\overline{S \cap T} \subseteq \overline{S} \cap \overline{T}.$$

This inclusion holds without any hypothesis on S. ■

3.14: Properties of Convex Sets

A set S in $\mathbb{R}^n$ is called convex if, for every pair of points x and y in S and every real θ satisfying $0 < \theta < 1$, we have $\theta x + (1 - \theta)y \in S$. Interpret this statement geometrically (in $\mathbb{R}^2$ and $\mathbb{R}^3$) and prove that:

a) Every n-ball in $\mathbb{R}^n$ is convex.

b) Every n-dimensional open interval is convex.

c) The interior of a convex set is convex.

d) The closure of a convex set is convex.

Strategy: Use the triangle inequality for (a), coordinate-wise inequalities for (b), and the fact that convex combinations preserve neighborhoods for (c) and (d). For (c), show that if two points are interior, their convex combination has a neighborhood in the set. For (d), use sequences and the fact that convex combinations are continuous.

Solution: Geometrically, a set is convex if the line segment joining any two points in the set lies entirely within the set.

(a) Let $B(a; r)$ be an n-ball and $x, y \in B(a; r)$. For $0 < \theta < 1$, let $z = \theta x + (1 - \theta)y$. Then $\|z - a\| = \|\theta(x - a) + (1 - \theta)(y - a)\| \leq \theta\|x - a\| + (1 - \theta)\|y - a\| < \theta r + (1 - \theta)r = r$, so $z \in B(a; r)$.

(b) Let $I = (a_1, b_1) \times \cdots \times (a_n, b_n)$ be an open interval and $x, y \in I$. For $0 < \theta < 1$, let $z = \theta x + (1 - \theta)y$. For each i, we have $a_i < x_i, y_i < b_i$, so $a_i < \theta x_i + (1 - \theta)y_i < b_i$. Therefore, $z \in I$.

(c) Let S be convex and $x, y \in \text{int } S$. There exist $\varepsilon_1, \varepsilon_2 > 0$ such that $B(x; \varepsilon_1) \subset S$ and $B(y; \varepsilon_2) \subset S$. Let $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. For $0 < \theta < 1$, let $z = \theta x + (1 - \theta)y$. If $w \in B(z; \varepsilon)$, then $\|w - z\| < \varepsilon$. Let $u = w - z + x$ and $v = w - z + y$. Then $\|u - x\| = \|v - y\| = \|w - z\| < \varepsilon$, so $u, v \in S$. Since S is convex, $w = \theta u + (1 - \theta)v \in S$. Therefore, $B(z; \varepsilon) \subset S$, so $z \in \text{int } S$.

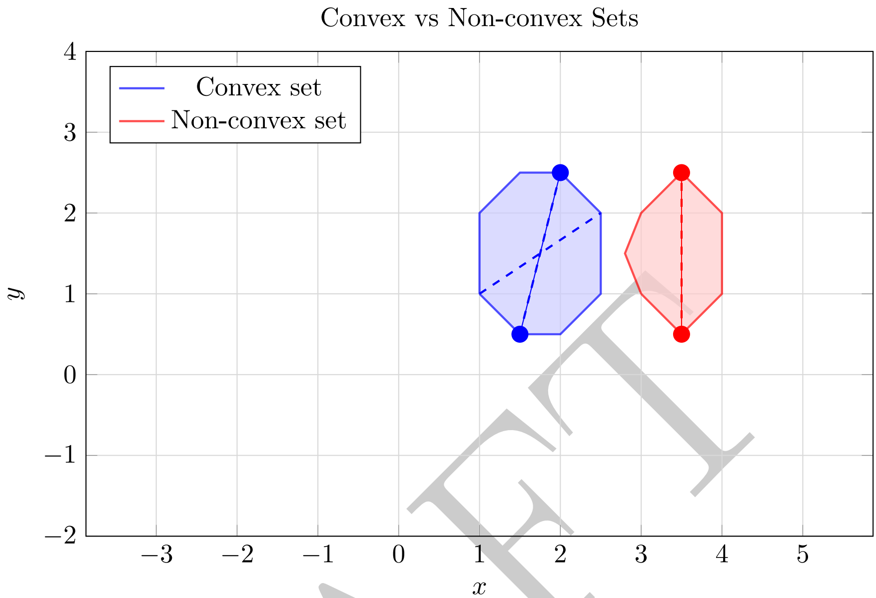


Figure 3.4: Left: A convex set where any line segment between two points lies entirely within the set. Right: A non-convex set where some line segments between points extend outside the set.

(d) Let S be convex and $x, y \in \overline{S}$. There exist sequences $\{x_n\}, \{y_n\} \subset S$ converging to x, y respectively. For $0 < \theta < 1$, let $z = \theta x + (1-\theta)y$ and $z_n = \theta x_n + (1-\theta)y_n$. Since S is convex, $z_n \in S$ for all n. Since $z_n \to z$, we have $z \in \overline{S}$. ■

3.15: Accumulation Points of Intersections and Unions

Let $\mathcal{F}$ be a collection of sets in $\mathbb{R}^k$, and let $S = \bigcup_{A\in\mathcal{F}} A$ and $T = \bigcap_{A\in\mathcal{F}} A$. For each of the following statements, either give a proof or exhibit a counterexample:

a) If $\mathbf{x}$ is an accumulation point of T, then $\mathbf{x}$ is an accumulation point of each set A in $\mathcal{F}$.

b) If $\mathbf{x}$ is an accumulation point of S, then $\mathbf{x}$ is an accumulation point of at least one set A in $\mathcal{F}$.

Strategy: For (a), use the fact that if a point is in the intersection, it's in every set, so accumulation points of the intersection must be accumulation points of each set. For (b), consider whether the collection is finite or infinite - for infinite collections, construct a counterexample using singletons.

Solution:
(a) This statement is **true**.
Solution: Let $\mathbf{x}$ be an accumulation point of T. This means that for

any $\varepsilon > 0$, the ball $B(\mathbf{x}; \varepsilon)$ contains a point $\mathbf{y} \in T$ such that $\mathbf{y} \neq \mathbf{x}$. By definition, $T = \bigcap_{A \in \mathcal{F}} A$. So, if $\mathbf{y} \in T$, then $\mathbf{y} \in A$ for every set A in the collection $\mathcal{F}$. Therefore, for any $\varepsilon > 0$, the ball $B(\mathbf{x}; \varepsilon)$ contains a point $\mathbf{y} \in A$ (for every $A \in \mathcal{F}$) with $\mathbf{y} \neq \mathbf{x}$. This is precisely the definition of $\mathbf{x}$ being an accumulation point of A. Thus, $\mathbf{x}$ is an accumulation point of each set $A \in \mathcal{F}$.

(b) This statement is **false** for an infinite collection $\mathcal{F}$.
Counterexample: Let the collection of sets in $\mathbb{R}^1$ be $\mathcal{F} = \{A_n : n \in \mathbb{N}\}$ where each set A_n is a singleton: $A_n = \{1/n\}$. The union is the set $S = \bigcup_{n=1}^{\infty} A_n = \{1, 1/2, 1/3, \ldots\}$. The set S has exactly one accumulation point: 0, since the sequence of points converges to 0. However, none of the individual sets A_n have any accumulation points, as they each contain only a single isolated point. Thus, 0 is an accumulation point of S, but not of any set A_n in the collection $\mathcal{F}$.

(Note: The statement is true if the collection $\mathcal{F}$ is finite. If $\mathbf{x}$ is an accumulation point of a finite union $S = A_1 \cup \cdots \cup A_m$, then any neighborhood of $\mathbf{x}$ contains infinitely many points from S. By the pigeonhole principle, at least one of the sets A_i must contribute infinitely many of these points, making $\mathbf{x}$ an accumulation point of that A_i.) ■

3.16: Rationals Not a Countable Intersection of Open Sets

Prove that the set S of rational numbers in the interval $(0, 1)$ cannot be expressed as the intersection of a countable collection of open sets.

Hint. Write $S = \{x_1, x_2, \ldots\}$, assume $S = \bigcap_{k=1}^{\infty} S_k$, where each S_k is open, and construct a sequence (Q_n) of closed intervals such that $Q_{n+1} \subseteq Q_n \subseteq S_n$ and such that $x_n \notin Q_n$. Then use the Cantor intersection theorem to obtain a contradiction.

Strategy: Use proof by contradiction. Assume the rationals can be written as a countable intersection of open sets. Enumerate the rationals and construct nested closed intervals that exclude each rational one by one. Use the Cantor intersection theorem to find a point in the intersection that cannot be rational, creating a contradiction.

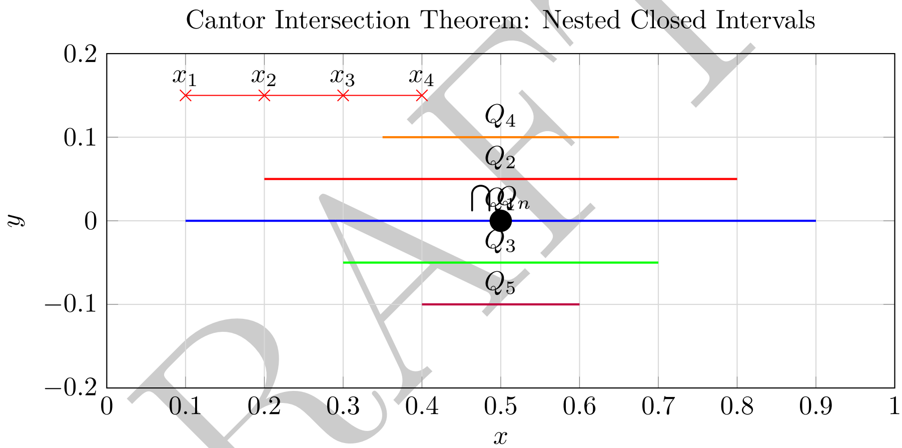


Figure 3.5: The Cantor intersection theorem: nested closed intervals Q_n exclude rationals x_n one by one, but their intersection must contain a point that cannot be any of the excluded rationals.

Solution: The strategy is to use a proof by contradiction. We'll assume that the rationals can be written as a countable intersection of open sets, then construct a nested sequence of closed intervals that excludes each rational number one by one. This will lead to a point that must be in the intersection (by the Cantor intersection theorem) but cannot be any of the rational numbers we've excluded, creating a contradiction.

Suppose for contradiction that $S = \bigcap_{k=1}^{\infty} S_k$ where each S_k is open. Let $S = \{x_1, x_2, \ldots\}$ be an enumeration of the rationals in $(0, 1)$.

For each n, since S_n is open and contains all rationals in $(0, 1)$, we can find a closed interval $Q_n \subset S_n$ such that $x_n \notin Q_n$. Here's why this is possible: Since S_n is open, for any point $y \in S_n$ that is not x_n, there exists an open interval around y that is entirely contained in S_n. We can choose a point $y \in S_n$ that is close to but not equal to x_n, and then take a small closed interval around y that stays within S_n but excludes x_n. For example, if x_n is not at the boundary of S_n, we can find a point $y \in S_n$ with $y < x_n$ and take $Q_n = [y - \varepsilon, y + \varepsilon]$ for some small $\varepsilon > 0$ such that $x_n > y + \varepsilon$. We can arrange that $Q_{n+1} \subseteq Q_n$ by taking $Q_{n+1} = Q_n \cap I_{n+1}$ where I_{n+1} is a closed interval in S_{n+1} that doesn't contain x_{n+1}.

By the Cantor intersection theorem, $\bigcap_{n=1}^{\infty} Q_n$ is nonempty. Let $x \in \bigcap_{n=1}^{\infty} Q_n$. Then $x \in \bigcap_{k=1}^{\infty} S_k = S$, so x is rational. But $x \neq x_n$ for any n since $x_n \notin Q_n$ for each n. This contradicts the fact that S contains all rationals in $(0, 1)$. ■

3.III Covering Theorems in $\mathbb{R}^n$

Definitions and Theorems

Definition: Open Cover

An open cover of a set S in a metric space (M, d) is a collection of open sets whose union contains S.

Definition: Compact Set

A set S in a metric space (M, d) is said to be compact if every open cover of S has a finite subcover.

Definition: Isolated Point

A point x in a set S is said to be an isolated point of S if there exists a neighborhood of x that contains no other points of S.

Definition: Separable Space

A metric space (M, d) is said to be separable if it contains a countable dense subset.

Theorem: Lindelöf Property

Every separable metric space has the Lindelöf property: every open cover has a countable subcover.

Theorem: Countability of Isolated Points

The collection of isolated points of any subset of $\mathbb{R}^n$ is countable.

Theorem: Countability via Local Countability

If every point in a set S has a neighborhood whose intersection with S is countable, then S is countable.

3.17: Countability of Isolated Points

If $S \subseteq \mathbb{R}^n$, prove that the collection of isolated points of S is countable.

Strategy: Use the fact that isolated points have disjoint neighborhoods. For each isolated point, take a smaller ball (half the radius) and show these balls are pairwise disjoint. Then use the separability of $\mathbb{R}^n$ (it has a countable dense subset) to inject the isolated points into this countable set.

Solution:
Let I be the set of isolated points of S. By definition, for each point $\mathbf{x} \in I$, there exists a radius $\varepsilon_{\mathbf{x}} > 0$ such that the open ball $B(\mathbf{x}; \varepsilon_{\mathbf{x}})$ contains no other point of S; that is, $B(\mathbf{x}; \varepsilon_{\mathbf{x}}) \cap S = \{\mathbf{x}\}$.

Consider the collection of smaller open balls $\mathcal{C} = \{B(\mathbf{x}; \varepsilon_{\mathbf{x}}/2) : \mathbf{x} \in I\}$. We claim these balls are pairwise disjoint. To prove this, let $\mathbf{x}_1$ and $\mathbf{x}_2$ be two distinct points in I. Suppose their corresponding balls in $\mathcal{C}$

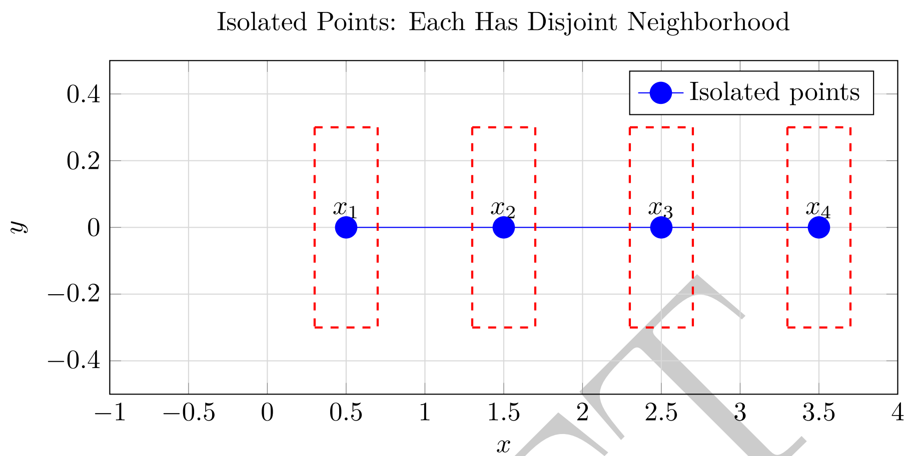


Figure 3.6: Isolated points have disjoint neighborhoods, allowing them to be mapped injectively into a countable dense subset, proving countability.

have a point $\mathbf{y}$ in common. Then $d(\mathbf{x}_1, \mathbf{y}) < \varepsilon_{\mathbf{x}_1}/2$ and $d(\mathbf{x}_2, \mathbf{y}) < \varepsilon_{\mathbf{x}_2}/2$. By the triangle inequality:

$$d(\mathbf{x}_1, \mathbf{x}_2) \leq d(\mathbf{x}_1, \mathbf{y}) + d(\mathbf{y}, \mathbf{x}_2) < \frac{\varepsilon_{\mathbf{x}_1}}{2} + \frac{\varepsilon_{\mathbf{x}_2}}{2}$$

Assuming, without loss of generality, that $\varepsilon_{\mathbf{x}_1} \leq \varepsilon_{\mathbf{x}_2}$, we get $d(\mathbf{x}_1, \mathbf{x}_2) < \varepsilon_{\mathbf{x}_2}/2 + \varepsilon_{\mathbf{x}_2}/2 = \varepsilon_{\mathbf{x}_2}$. This implies that $\mathbf{x}_1 \in B(\mathbf{x}_2; \varepsilon_{\mathbf{x}_2})$. But $\mathbf{x}_1 \in S$ and $\mathbf{x}_1 \neq \mathbf{x}_2$. This contradicts the fact that $B(\mathbf{x}_2; \varepsilon_{\mathbf{x}_2})$ contains only one point from S, namely $\mathbf{x}_2$. Therefore, the balls in the collection $\mathcal{C}$ must be pairwise disjoint.

Now we use the fact that $\mathbb{R}^n$ is **separable**, meaning it contains a countable dense subset, such as $\mathbb{Q}^n$ (the set of points with rational coordinates). Since each ball in $\mathcal{C}$ is a non-empty open set, each must contain at least one point from the dense set $\mathbb{Q}^n$. Because the balls in $\mathcal{C}$ are disjoint, each ball must contain a *different* rational point. This allows us to define an injective (one-to-one) function from the set of isolated points I to the countable set $\mathbb{Q}^n$ (by mapping each $\mathbf{x} \in I$ to a rational point in $B(\mathbf{x}; \varepsilon_{\mathbf{x}}/2)$). A set that can be mapped injectively into a countable set must itself be countable. Thus, the set of isolated points I is countable. ∎

3.18: Countable Covering of the First Quadrant

Prove that the set of open disks in the xy-plane with center at (x, x) and radius $x > 0$, where x is rational, is a countable covering of the set $\{(x, y) : x > 0, y > 0\}$.

Strategy: Show that the collection is countable (since rationals are countable) and that it covers the first quadrant. For any point (x, y) in the first quadrant, find a rational q such that the disk centered at (q, q) with radius q contains (x, y). Use the distance formula and density of rationals.

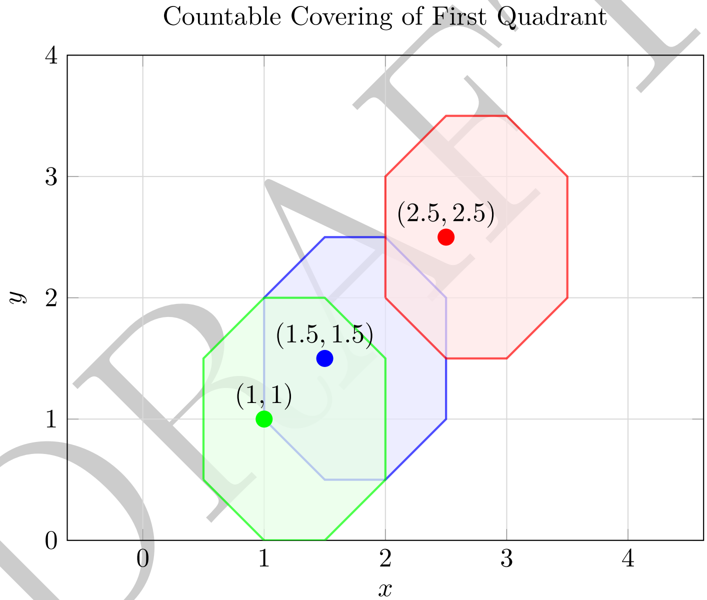


Figure 3.7: Countable covering of the first quadrant using disks centered at rational points (q, q) with radius q.

Solution:
Let $\mathcal{F}$ be the collection of open disks $B((q, q); q)$ where $q \in \mathbb{Q}$ and $q > 0$. Since $\mathbb{Q}$ is countable, the collection $\mathcal{F}$ is countable. We need to show that $\mathcal{F}$ covers the first quadrant $S = \{(x, y) : x > 0, y > 0\}$.

Let (x, y) be an arbitrary point in S. We need to find a rational number $q > 0$ such that the disk $B((q, q); q)$ contains (x, y). The condition for this is:

$$\sqrt{(x-q)^2 + (y-q)^2} < q$$

Since both sides are positive, we can square the inequality:

$$(x-q)^2 + (y-q)^2 < q^2$$

$$x^2 - 2xq + q^2 + y^2 - 2yq + q^2 < q^2$$

$$q^2 - 2(x+y)q + (x^2 + y^2) < 0$$

Let $f(q) = q^2 - 2(x+y)q + (x^2+y^2)$. We are looking for a rational $q > 0$ that makes this quadratic expression negative. The graph of $z = f(q)$ is an upward-opening parabola. It will be negative between its roots. The roots are found using the quadratic formula:

$$q = \frac{2(x+y) \pm \sqrt{4(x+y)^2 - 4(x^2+y^2)}}{2} = (x+y) \pm \sqrt{(x+y)^2 - (x^2+y^2)}$$

$$q = (x+y) \pm \sqrt{2xy}$$

Let the roots be $q_1 = (x+y) - \sqrt{2xy}$ and $q_2 = (x+y) + \sqrt{2xy}$. Since $x, y > 0$, the term $\sqrt{2xy}$ is real and positive, so $q_1 < q_2$. The interval (q_1, q_2) is non-empty. Since the rational numbers are dense in $\mathbb{R}$, we can always find a rational number q in this interval: $q_1 < q < q_2$. For any such q, the inequality $f(q) < 0$ holds.

We must also ensure that we can choose q to be positive. The product of the roots is $q_1 q_2 = x^2 + y^2 > 0$. Since $q_2 = (x+y) + \sqrt{2xy}$ is clearly positive, the other root q_1 must also be positive. Since the interval (q_1, q_2) consists of positive numbers and contains a rational number, we can always find a suitable rational $q > 0$. Thus, for any point (x, y) in the first quadrant, we can find a disk in $\mathcal{F}$ that contains it. The countable collection $\mathcal{F}$ therefore covers the first quadrant. ■

3.19: Non-Finite Subcover of $0, 1$

The collection $\mathcal{F}$ of open intervals of the form $(1/n, 2/n)$, where $n = 2, 3, \ldots$, is an open covering of the open interval $(0, 1)$. Prove (without using Theorem 3.31) that no finite subcollection of $\mathcal{F}$ covers $(0, 1)$.

Strategy: Take any finite subcollection and find the maximum denominator N. Show that the interval $(0, 1/N)$ is not covered by any interval in the finite subcollection, since the leftmost interval is $(1/N, 2/N)$.

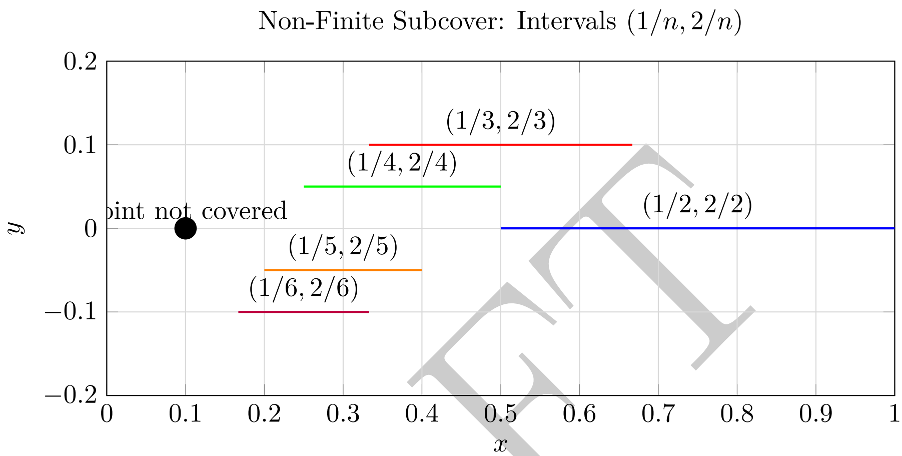


Figure 3.8: The collection of intervals $(1/n, 2/n)$ covers $(0, 1)$ but no finite subcollection covers it. The point 0.1 is not covered by any interval in a finite subcollection.

Solution: Let $\mathcal{G} = \{(1/n_1, 2/n_1), \ldots, (1/n_k, 2/n_k)\}$ be a finite subcollection of $\mathcal{F}$. Let $N = \max\{n_1, \ldots, n_k\}$.

Then the leftmost interval in $\mathcal{G}$ is $(1/N, 2/N)$. For any $x \in (0, 1/N)$, we have $x < 1/N < 2/N$, so x is not covered by any interval in $\mathcal{G}$.

Therefore, $\mathcal{G}$ does not cover $(0, 1)$, proving that no finite subcollection of $\mathcal{F}$ covers $(0, 1)$. ■

3.20: Closed but Not Bounded Set with Infinite Covering

Give an example of a set S which is closed but not bounded and exhibit a countable open covering $\mathcal{F}$ such that no finite subset of $\mathcal{F}$ covers S.

Strategy: Use the integers $\mathbb{Z}$ as the set (closed but not bounded). Create a covering where each integer has its own interval, making it impossible for any finite subcollection to cover the infinite set.

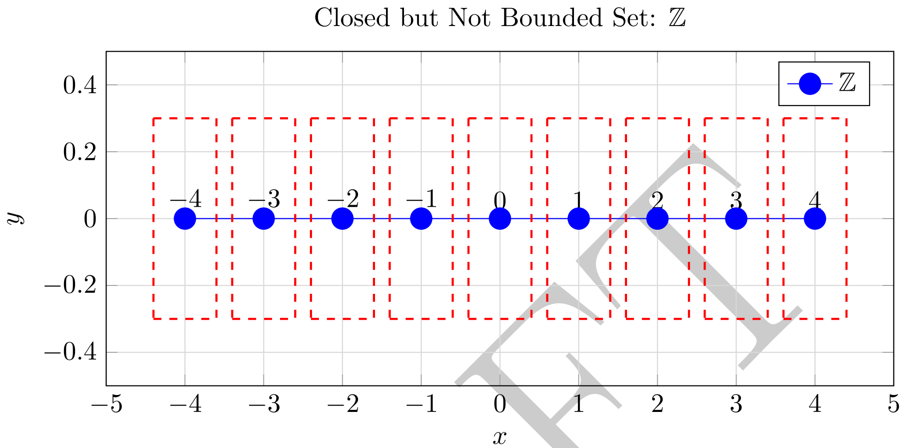


Figure 3.9: The set of integers $\mathbb{Z}$ is closed but not bounded. Each integer has a neighborhood that contains no other integers, showing it has no accumulation points.

Solution: Let $S = \mathbb{Z}$ (the set of integers). This set is closed but not bounded.

Let $\mathcal{F} = \{(n - 1/2, n + 1/2) : n \in \mathbb{Z}\}$. This is a countable open covering of $\mathbb{Z}$ since each integer n is contained in the interval $(n - 1/2, n + 1/2)$.

However, no finite subcollection of $\mathcal{F}$ covers $\mathbb{Z}$. If $\mathcal{G} = \{(n_1 - 1/2, n_1 + 1/2), \ldots, (n_k - 1/2, n_k + 1/2)\}$ is a finite subcollection, then $\mathcal{G}$ can only cover finitely many integers, but $\mathbb{Z}$ is infinite.

Therefore, $\mathcal{F}$ is a countable open covering of S with no finite subcover. ■

3.21: Countability via Local Countability

Given a set S in $\mathbb{R}^n$ with the property that for every x in S there is an n-ball $B(x)$ such that $B(x) \cap S$ is countable. Prove that S is countable.

Strategy: Use the Lindelöf property of $\mathbb{R}^n$ (every open cover has a countable subcover). The collection of balls $\{B(x) : x \in S\}$ covers S, so there's a countable subcover. Each ball in the subcover intersects S in a countable set, so S is a countable union of countable sets.

Solution:
For each point $\mathbf{x} \in S$, we are given that there exists an open ball $B_{\mathbf{x}}$ centered at $\mathbf{x}$ such that the set $B_{\mathbf{x}} \cap S$ is countable.

The collection of all such balls, $\mathcal{C} = \{B_{\mathbf{x}} : \mathbf{x} \in S\}$, forms an open covering of the set S (since each $\mathbf{x} \in S$ is in its own ball $B_{\mathbf{x}}$).

The space $\mathbb{R}^n$ is a **separable** metric space because it contains a countable dense subset, $\mathbb{Q}^n$. A key theorem in topology states that every separable metric space has the **Lindelöf property**. This property guarantees that any open covering of a set in that space has a countable subcovering.

Applying the Lindelöf property to our open cover $\mathcal{C}$ of S, we can extract a countable subcollection, say $\mathcal{C}' = \{B_{\mathbf{x}_k} : k \in \mathbb{N}\}$, that still covers S. This means:

$$S \subseteq \bigcup_{k=1}^{\infty} B_{\mathbf{x}_k}$$

From this, we can express the set S as:

$$S = S \cap \left(\bigcup_{k=1}^{\infty} B_{\mathbf{x}_k} \right) = \bigcup_{k=1}^{\infty} (S \cap B_{\mathbf{x}_k})$$

By the initial hypothesis, each set in this union, $(S \cap B_{\mathbf{x}_k})$, is countable. Therefore, S is a countable union of countable sets. A fundamental result of set theory states that a countable union of countable sets is itself countable. Thus, we conclude that the set S must be countable. ■

3.22: Countability of Disjoint Open Sets

Prove that a collection of disjoint open sets in $\mathbb{R}^n$ is necessarily countable. Give an example of a collection of disjoint closed sets which is not countable.

Strategy: Use the separability of $\mathbb{R}^n$ - each open set contains a point from the countable dense subset, and since the sets are disjoint, each dense point can belong to at most one set. For the counterexample, use singletons of real numbers.

Solution: Let $\mathcal{F}$ be a collection of disjoint open sets in $\mathbb{R}^n$. Since $\mathbb{R}^n$ is separable, there exists a countable dense subset D.

For each open set $U \in \mathcal{F}$, there exists a point $d \in D$ such that $d \in U$. Since the sets in $\mathcal{F}$ are disjoint, each point $d \in D$ can belong to at most one set in $\mathcal{F}$.

Therefore, the number of sets in $\mathcal{F}$ is at most the number of points in D, which is countable.

For an example of uncountably many disjoint closed sets, let $\mathcal{G} = \{\{x\} : x \in \mathbb{R}\}$. Each singleton $\{x\}$ is closed, the sets are disjoint, and there are uncountably many real numbers. ■

3.23: Existence of Condensation Points

Assume that $S \subseteq \mathbb{R}^n$. A point x in $\mathbb{R}^n$ is said to be a condensation point of S if every n-ball $B(x)$ has the property that $B(x) \cap S$ is not countable. Prove that if S is not countable, then there exists a point x in S such that x is a condensation point of S.

Strategy: Use proof by contradiction. If no point is a condensation point, then every point has a neighborhood where the intersection with S is countable. Apply Exercise 3.21 to conclude that S is countable, contradicting the hypothesis.

Solution: Suppose for contradiction that no point in S is a condensation point of S. Then for every $x \in S$, there exists an n-ball B_x centered at x such that $B_x \cap S$ is countable.

By Exercise 3.21, this implies that S is countable, which contradicts the hypothesis that S is not countable.

Therefore, there must exist at least one point $x \in S$ that is a condensation point of S. ■

3.24: Properties of Condensation Points

Assume that $S \subseteq \mathbb{R}^n$ and that S is not countable. Let T denote the set of condensation points of S. Prove that:

a) $S - T$ is countable,

b) $S \cap T$ is not countable,

c) T is closed,

d) T contains no isolated points.

Note that Exercise 3.23 is a special case of (b).

Strategy: For (a), use Exercise 3.21. For (b), use the fact that S is uncountable and $S - T$ is countable. For (c), show that if a point is in the closure of T, it's a condensation point. For (d), use the fact that $S - T$ is countable and any neighborhood of a condensation point contains uncountably many points of S.

Solution: (a) For each $x \in S - T$, there exists an n-ball B_x centered at x such that $B_x \cap S$ is countable. By Exercise 3.21, $S - T$ is countable.

(b) Since S is not countable and $S - T$ is countable, $S \cap T$ must be uncountable.

(c) Let $x \in \overline{T}$. Then every neighborhood of x contains a point of T. Let B be any n-ball centered at x. There exists $y \in T \cap B$. Since y is a condensation point, $B(y; r) \cap S$ is uncountable for any $r > 0$. Choose r small enough so that $B(y; r) \subset B$. Then $B \cap S$ contains the uncountable set $B(y; r) \cap S$, so x is a condensation point. Therefore, T is closed.

(d) Let $x \in T$. For any $\varepsilon > 0$, $B(x; \varepsilon) \cap S$ is uncountable. Since $S - T$ is countable, $B(x; \varepsilon) \cap T$ must be uncountable. Therefore, x is not isolated in T. ■

3.25: Cantor-Bendixon Theorem

A set in $\mathbb{R}^n$ is called perfect if $S = S'$, that is, if S is a closed set which contains no isolated points. Prove that every uncountable closed set F in $\mathbb{R}^n$ can be expressed in the form $F = A \cup B$, where A is perfect and B is countable (Cantor-Bendixon theorem).
Hint. Use Exercise 3.24.

Strategy: Use Exercise 3.24 to take A as the set of condensation points T and B as $F - T$. Show that T is perfect (closed with no isolated points) and $F - T$ is countable.

Solution: Let F be an uncountable closed set in $\mathbb{R}^n$. Let T be the set of condensation points of F. By Exercise 3.24, T is closed and $F - T$ is countable.

Let $A = T$ and $B = F - T$. Then $F = A \cup B$ where B is countable.

We need to show that A is perfect. Since T is closed by Exercise 3.24(c), A is closed. By Exercise 3.24(d), T contains no isolated points, so A contains no isolated points.

Therefore, A is perfect, and we have the desired decomposition $F = A \cup B$. ■

3.IV Metric Spaces

Definitions and Theorems

Definition: Metric Space

A metric space consists of a set M together with a function $d : M \times M \to [0, \infty)$ (called a metric or distance function) that satisfies the following three axioms for all points $x, y, z \in M$:

1. $d(x, y) = 0$ if and only if $x = y$ (positive definiteness)
2. $d(x, y) = d(y, x)$ for all $x, y \in M$ (symmetry)
3. $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in M$ (triangle inequality)

Definition: Closed Ball

Given a metric space (M, d), the closed ball centered at a point $a \in M$ with radius $r > 0$ is the set $\overline{B}(a; r) = \{x \in M : d(x, a) \leq r\}$.

Theorem: Separability of Euclidean Spaces

Every Euclidean space $\mathbb{R}^n$ is separable.

Theorem: Bounded Metric Construction

If (M, d) is a metric space, then the function $d'(x, y) = \frac{d(x,y)}{1+d(x,y)}$ defines a metric on M that is bounded above by 1.

Theorem: Product Metrics

Given two metric spaces (S_1, d_1) and (S_2, d_2), the following functions define metrics on the Cartesian product $S_1 \times S_2$:

1. $\rho(x, y) = d_1(x_1, y_1) + d_2(x_2, y_2)$ (sum metric)
2. $\rho(x, y) = \max\{d_1(x_1, y_1), d_2(x_2, y_2)\}$ (maximum metric)
3. $\rho(x, y) = \sqrt{d_1(x_1, y_1)^2 + d_2(x_2, y_2)^2}$ (Euclidean product metric)

Theorem: Finite Sets are Closed

Every finite subset of a metric space is a closed set.

Theorem: Closed Balls are Closed

In any metric space, every closed ball is a closed set.

Theorem: Transitivity of Density

If A is dense in S and S is dense in T, then A is dense in T.

Theorem: Density and Open Sets

If A is dense in S and B is an open subset of S, then B is contained in the closure of $A \cap B$.

Theorem: Intersection of Dense and Open Sets

If both A and B are dense in S and B is an open subset of S, then the intersection $A \cap B$ is dense in S.

3.26: Open and Closed Sets in Metric Spaces

In any metric space (M, d), prove that the empty set $\emptyset$ and the whole space M are both open and closed.

Strategy: Use the definitions of open and closed sets directly. For the empty set, use vacuous truth for openness and complementarity for closedness. For the whole space, use the definition of open balls and complementarity.

Solution: The empty set $\emptyset$ is open because the condition "for every point in $\emptyset$, there exists a neighborhood contained in $\emptyset$" is vacuously true (there are no points to check).

The empty set $\emptyset$ is closed because its complement M is open.

The whole space M is open because for any point $x \in M$ and any $\varepsilon > 0$, the ball $B(x; \varepsilon) \subset M$.

The whole space M is closed because its complement $\emptyset$ is open. ■

3.27: Metric Balls in Different Metrics

Consider the following two metrics in $\mathbb{R}^n$:

$$d_1(x, y) = \max_{1 \le i \le n} |x_i - y_i|, \quad d_2(x, y) = \sum_{i=1}^{n} |x_i - y_i|.$$

In each of the following metric spaces prove that the ball $B(a; r)$ has the geometric appearance indicated:

a) In $(\mathbb{R}^2, d_1)$, a square with sides parallel to the coordinate axes.

b) In $(\mathbb{R}^2, d_2)$, a square with diagonals parallel to the axes.

c) A cube in $(\mathbb{R}^3, d_1)$.

d) An octahedron in $(\mathbb{R}^3, d_2)$.

Strategy: Use the definition of metric balls $B(a; r) = \{x : d(a, x) < r\}$ and substitute the given metrics. For d_1, the maximum constraint creates axis-aligned shapes. For d_2, the sum constraint creates diamond/octahedral shapes.

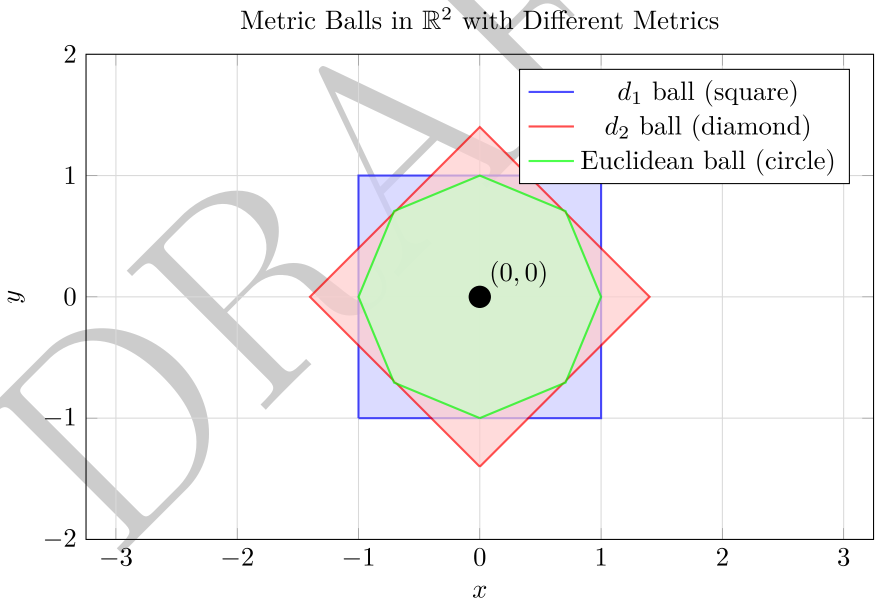


Figure 3.10: Comparison of metric balls centered at $(0, 0)$ with radius 1: d_1 produces a square, d_2 produces a diamond, and the Euclidean metric produces a circle.

Solution: (a) In $(\mathbb{R}^2, d_1)$, the ball $B(a;r) = \{(x,y) : \max\{|x-a_1|, |y-a_2|\} < r\}$. This means $|x-a_1| < r$ and $|y-a_2| < r$, which defines a square with center (a_1, a_2) and sides of length $2r$ parallel to the coordinate axes.

(b) In $(\mathbb{R}^2, d_2)$, the ball $B(a;r) = \{(x,y) : |x-a_1| + |y-a_2| < r\}$. This defines a diamond-shaped region (square rotated 45 degrees) with diagonals parallel to the axes.

(c) In $(\mathbb{R}^3, d_1)$, the ball $B(a;r) = \{(x,y,z) : \max\{|x-a_1|, |y-a_2|, |z-a_3|\} < r\}$. This defines a cube with center (a_1, a_2, a_3) and sides of length $2r$ parallel to the coordinate axes.

(d) In $(\mathbb{R}^3, d_2)$, the ball $B(a;r) = \{(x,y,z) : |x-a_1| + |y-a_2| + |z-a_3| < r\}$. This defines an octahedron with center (a_1, a_2, a_3). ■

3.28: Metric Inequalities

Let d_1 and d_2 be the metrics of Exercise 3.27 and let $\|x-y\|$ denote the usual Euclidean metric. Prove the following inequalities for all x and y in $\mathbb{R}^n$:

$$d_1(x,y) \leq \|x-y\| \leq d_2(x,y) \quad \text{and} \quad d_2(x,y) \leq \sqrt{n}\|x-y\| \leq n\, d_1(x,y).$$

Strategy: Use the definitions of the metrics and basic inequalities. For the first part, use the fact that the maximum is less than or equal to the square root of the sum of squares, and the square root is less than or equal to the sum. For the second part, use the Cauchy-Schwarz inequality.

Solution: Let $x, y \in \mathbb{R}^n$. Let $a_i = |x_i - y_i|$ for $1 \leq i \leq n$. Then:

(1) $d_1(x,y) = \max_i a_i \leq \sqrt{\sum a_i^2} = \|x-y\|$ (since each $a_i^2 \leq \sum a_i^2$)

(2) $\|x-y\| = \sqrt{\sum a_i^2} \leq \sum a_i = d_2(x,y)$ (by the inequality $\sqrt{a_1^2 + \cdots + a_n^2} \leq a_1 + \cdots + a_n$)

(3) By the Cauchy-Schwarz inequality:

$$\left(\sum a_i\right)^2 \leq n \sum a_i^2 \Rightarrow d_2(x,y) \leq \sqrt{n}\|x-y\|$$

(4) Also, since $\|x-y\| = \sqrt{\sum a_i^2} \leq \sqrt{n \cdot \max_i a_i^2} = \sqrt{n} \cdot \max a_i = \sqrt{n} d_1(x,y)$, it follows that:

$$\sqrt{n}\|x-y\| \leq n d_1(x,y)$$

Hence, all inequalities hold:

$$d_1(x,y) \le \|x-y\| \le d_2(x,y), \quad d_2(x,y) \le \sqrt{n}\|x-y\| \le n\, d_1(x,y)$$

■

3.29: Bounded Metric

If (M,d) is a metric space, define

$$d'(x,y) = \frac{d(x,y)}{1+d(x,y)}.$$

Prove that d' is also a metric for M. Note that $0 \le d'(x,y) < 1$ for all x, y in M.

Strategy: Verify the three metric properties: non-negativity, symmetry, and triangle inequality. For the triangle inequality, use the fact that the function $f(t) = t/(1+t)$ is increasing and the inequality $\frac{a+b}{1+a+b} \le \frac{a}{1+a} + \frac{b}{1+b}$ for non-negative a, b.

Solution: We need to verify the three properties of a metric:

(1) $d'(x,y) \ge 0$ since $d(x,y) \ge 0$ and $1 + d(x,y) > 0$.

(2) $d'(x,y) = 0$ if and only if $d(x,y) = 0$, which occurs if and only if $x = y$.

(3) $d'(x,y) = d'(y,x)$ since $d(x,y) = d(y,x)$.

(4) For the triangle inequality, let $f(t) = \frac{t}{1+t}$. Then $f'(t) = \frac{1}{(1+t)^2} > 0$, so f is increasing. Therefore, $d'(x,z) = f(d(x,z)) \le f(d(x,y)+d(y,z)) = \frac{d(x,y)+d(y,z)}{1+d(x,y)+d(y,z)} \le \frac{d(x,y)}{1+d(x,y)} + \frac{d(y,z)}{1+d(y,z)} = d'(x,y) + d'(y,z)$.

The last inequality follows from the fact that $\frac{a+b}{1+a+b} \le \frac{a}{1+a} + \frac{b}{1+b}$ for $a, b \ge 0$. ■

3.30: Finite Sets in Metric Spaces

Prove that every finite subset of a metric space is closed.

Strategy: Show that the complement is open. For any point not in the finite set, find a minimum distance to the set and use that to construct a neighborhood that doesn't intersect the set.

Solution: Let $S = \{x_1, x_2, \ldots, x_n\}$ be a finite subset of a metric space (M, d). We need to show that the complement $M \setminus S$ is open.

Let $x \in M \setminus S$. Let $\varepsilon = \min\{d(x, x_i) : i = 1, 2, \ldots, n\}$. Since $x \notin S$, we have $\varepsilon > 0$.

Then $B(x; \varepsilon) \cap S = \emptyset$, so $B(x; \varepsilon) \subset M \setminus S$. This shows that every point in $M \setminus S$ is an interior point, so $M \setminus S$ is open.

Therefore, S is closed. ■

3.31: Closed Balls in Metric Spaces

In a metric space (M, d) the closed ball of radius $r > 0$ about a point a in M is the set $\overline{B}(a; r) = \{x : d(x, a) \leq r\}$.

a) Prove that $\overline{B}(a; r)$ is a closed set.

b) Give an example of a metric space in which $\overline{B}(a; r)$ is not the closure of the open ball $B(a; r)$.

Strategy: For (a), show the complement is open by using the triangle inequality to find a neighborhood around any point outside the closed ball. For (b), use a discrete metric space where the open ball is a singleton but the closed ball is the entire space.

Solution: (a) Let $x \in M \setminus \overline{B}(a; r)$. Then $d(x, a) > r$. Let $\varepsilon = d(x, a) - r > 0$. For any $y \in B(x; \varepsilon)$, we have $d(y, a) \geq d(x, a) - d(x, y) > d(x, a) - \varepsilon = r$. Therefore, $B(x; \varepsilon) \subset M \setminus \overline{B}(a; r)$, showing that $M \setminus \overline{B}(a; r)$ is open. Hence, $\overline{B}(a; r)$ is closed.

(b) Consider the discrete metric space (M, d) where $d(x, y) = 1$ if $x \neq y$ and $d(x, y) = 0$ if $x = y$. Let $a \in M$ and $r = 1$. Then $B(a; 1) = \{a\}$ and $\overline{B}(a; 1) = M$. The closure of $B(a; 1)$ is $\{a\}$, which is not equal to $\overline{B}(a; 1) = M$. ■

3.32: Transitivity of Density

In a metric space M, if subsets satisfy $A \subseteq S \subseteq \overline{A}$, where $\overline{A}$ is the closure of A, then A is said to be dense in S. For example, the set $\mathbb{Q}$ of rational numbers is dense in $\mathbb{R}$. If A is dense in S and if S is dense in T, prove that A is dense in T.

Strategy: Use the definition of density and the fact that closure is idempotent ($\overline{\overline{A}} = \overline{A}$). Show that $A \subseteq T \subseteq \overline{A}$ by using the density relationships and the transitivity of subset inclusion.

Solution: We need to show that $A \subseteq T \subseteq \overline{A}$.

Since $A \subseteq S \subseteq T$, we have $A \subseteq T$.

Since S is dense in T, we have $T \subseteq \overline{S}$. Since A is dense in S, we have $S \subseteq \overline{A}$. Therefore, $\overline{S} \subseteq \overline{\overline{A}} = \overline{A}$.

Combining these, we get $T \subseteq \overline{S} \subseteq \overline{A}$, so $T \subseteq \overline{A}$.

Therefore, $A \subseteq T \subseteq \overline{A}$, showing that A is dense in T. ∎

3.33: Separability of Euclidean Spaces

A metric space M is said to be separable if there is a countable subset A which is dense in M. For example, $\mathbb{R}$ is separable because the set $\mathbb{Q}$ of rational numbers is a countable dense subset. Prove that every Euclidean space $\mathbb{R}^k$ is separable.

Strategy: Use the Cartesian product of rational numbers $\mathbb{Q}^k$ as the countable dense subset. Show it's countable (product of countable sets) and dense (use the density of rationals in each coordinate and the triangle inequality).

Solution: Let A be the set of all points in $\mathbb{R}^k$ with rational coordinates. That is, $A = \{(q_1, q_2, \ldots, q_k) : q_i \in \mathbb{Q} \text{ for } i = 1, 2, \ldots, k\}$.

Since $\mathbb{Q}$ is countable, the Cartesian product $A = \mathbb{Q}^k$ is countable.

To show that A is dense in $\mathbb{R}^k$, let $x = (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k$ and $\varepsilon > 0$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, for each i there exists $q_i \in \mathbb{Q}$ such that $|x_i - q_i| < \varepsilon/\sqrt{k}$.

Then $q = (q_1, q_2, \ldots, q_k) \in A$ and $\|x - q\| = \sqrt{\sum_{i=1}^k (x_i - q_i)^2} < \sqrt{k(\varepsilon/\sqrt{k})^2} = \varepsilon$.

Therefore, A is a countable dense subset of $\mathbb{R}^k$, so $\mathbb{R}^k$ is separable. ■

3.34: Lindelöf Theorem in Separable Spaces

Prove that the Lindelöf covering theorem (Theorem 3.28) is valid in any separable metric space.

Strategy: Use the countable dense subset to construct a countable subcover. For each point in the dense subset, choose a set from the open cover that contains it, and show that this countable collection covers the entire space.

Solution: Let M be a separable metric space with countable dense subset $D = \{d_1, d_2, \ldots\}$. Let $\mathcal{F}$ be an open covering of M.

For each $d_i \in D$ and each positive rational r, if there exists a set $F \in \mathcal{F}$ such that $B(d_i; r) \subset F$, let $F_{i,r}$ be one such set.

The collection $\{F_{i,r} : i \in \mathbb{N}, r \in \mathbb{Q}^+, B(d_i; r) \subset F_{i,r} \text{ for some } F \in \mathcal{F}\}$ is countable.

We claim this collection covers M. Let $x \in M$. Since $\mathcal{F}$ covers M, there exists $F \in \mathcal{F}$ such that $x \in F$. Since F is open, there exists $\varepsilon > 0$ such that $B(x; \varepsilon) \subset F$.

Since D is dense, there exists $d_i \in D$ such that $d_i \in B(x; \varepsilon/2)$. Let r be a rational number such that $d(x, d_i) < r < \varepsilon/2$. Then $B(d_i; r) \subset B(x; \varepsilon) \subset F$.

Therefore, $F_{i,r}$ exists and contains x, showing that the countable subcollection covers M. ■

3.35: Density and Open Sets

If A is dense in S and if B is open in S, prove that $B \subseteq \overline{A \cap B}$.
Hint. Exercise 3.13.

Strategy: Use the definition of density and the fact that B is open. For any point $x \in B$, show that any neighborhood of x contains a point from $A \cap B$ by using the density of A and the openness of B.

Solution:
The statement "A is dense in S" means $S \subseteq \overline{A}$. We are given that $A \subseteq S$. The statement "B is open in S" means that $B = V \cap S$ for some set V that is open in the larger metric space M.

Let $x \in B$. We want to show that $x \in \overline{A \cap B}$. This requires showing that any open neighborhood of x in M has a non-empty intersection with the set $A \cap B$.

Let U be an arbitrary open neighborhood of x in M. Since $x \in B$ and $B = V \cap S$, we have $x \in V$. The set $U \cap V$ is also an open neighborhood of x because it is the intersection of two open sets.

Since $x \in B \subseteq S$ and A is dense in S, x is an adherent point of A. Therefore, the open neighborhood $U \cap V$ must contain a point from A. Let's call this point y. So, $y \in (U \cap V) \cap A$.

Now we check if this point y is in the required sets:

- $y \in U$, so y is in the arbitrary neighborhood of x.
- $y \in A$.
- We need to show $y \in B$. We know $y \in V$. Since we are given $A \subseteq S$, and $y \in A$, it follows that $y \in S$.

Since $y \in V$ and $y \in S$, we have $y \in V \cap S$, which means $y \in B$.

So we have found a point y such that $y \in U$ and $y \in A \cap B$. This means $U \cap (A \cap B) \neq \emptyset$. Since U was an arbitrary open neighborhood of x, this proves that $x \in \overline{A \cap B}$. As this holds for any $x \in B$, we conclude that $B \subseteq \overline{A \cap B}$. ■

3.36: Intersection of Dense and Open Sets

If each of A and B is dense in S and if B is open in S, prove that $A \cap B$ is dense in S.

Strategy: Use the fact that B is open to find a neighborhood around any point in S that is contained in B. Then use the density of A to find a point in that neighborhood that belongs to both A and B.

Solution: We need to show that $S \subseteq \overline{A \cap B}$.

Let $x \in S$. Since B is open in S, there exists $\varepsilon > 0$ such that $B(x;\varepsilon) \cap S \subset B$.

Since A is dense in S, $B(x;\varepsilon) \cap A \neq \emptyset$. Let $y \in B(x;\varepsilon) \cap A$. Since $y \in S$ and $B(x;\varepsilon) \cap S \subset B$, we have $y \in B$.

Therefore, $y \in A \cap B$, so $B(x;\varepsilon) \cap (A \cap B) \neq \emptyset$.

This shows that every neighborhood of x contains a point of $A \cap B$, so $x \in \overline{A \cap B}$.

Therefore, $S \subseteq \overline{A \cap B}$, showing that $A \cap B$ is dense in S. ■

3.37: Product Metrics

Given two metric spaces (S_1, d_1) and (S_2, d_2), a metric ρ for the Cartesian product $S_1 \times S_2$ can be constructed from d_1 and d_2 in many ways. For example, if $x = (x_1, x_2)$ and $y = (y_1, y_2)$ are in $S_1 \times S_2$, let $\rho(x, y) = d_1(x_1, y_1) + d_2(x_2, y_2)$. Prove that ρ is a metric for $S_1 \times S_2$ and construct further examples.

Strategy: Verify the three metric properties for the sum metric. For the triangle inequality, use the triangle inequalities of the individual metrics. For additional examples, consider maximum, Euclidean, and p-norms.

Solution: We need to verify the three properties of a metric for $\rho(x, y) = d_1(x_1, y_1) + d_2(x_2, y_2)$:

(1) $\rho(x, y) \geq 0$ since $d_1(x_1, y_1) \geq 0$ and $d_2(x_2, y_2) \geq 0$.

(2) $\rho(x, y) = 0$ if and only if $d_1(x_1, y_1) = 0$ and $d_2(x_2, y_2) = 0$, which occurs if and only if $x_1 = y_1$ and $x_2 = y_2$, i.e., $x = y$.

(3) $\rho(x, y) = \rho(y, x)$ since $d_1(x_1, y_1) = d_1(y_1, x_1)$ and $d_2(x_2, y_2) = d_2(y_2, x_2)$.

(4) For the triangle inequality, let $z = (z_1, z_2)$. Then $\rho(x, z) = d_1(x_1, z_1) + d_2(x_2, z_2) \leq d_1(x_1, y_1) + d_1(y_1, z_1) + d_2(x_2, y_2) + d_2(y_2, z_2) = \rho(x, y) + \rho(y, z)$.

Other examples of product metrics include: - $\rho(x, y) = \max\{d_1(x_1, y_1), d_2(x_2, y_2)\}$ - $\rho(x, y) = \sqrt{d_1(x_1, y_1)^2 + d_2(x_2, y_2)^2}$ - $\rho(x, y) = (d_1(x_1, y_1)^p + d_2(x_2, y_2)^p)^{1/p}$ for $p \geq 1$ ■

3.V Compact subsets of a metric space

Definitions and Theorems

Definition: Compact Set

A set S in a metric space (M, d) is said to be compact if every open cover of S has a finite subcover.

Theorem: Heine-Borel Theorem

A subset of $\mathbb{R}^n$ is compact if and only if it is both closed and bounded.

Theorem: Properties of Compact Sets

Let (M, d) be a metric space. Then the following properties hold:

1. Every closed subset of a compact set is compact.
2. The union of any finite collection of compact sets is compact.
3. The intersection of any nonempty collection of compact sets is compact.
4. The continuous image of a compact set is compact.

Theorem: Sequential Compactness

A metric space is compact if and only if every sequence in the space has a convergent subsequence.

Prove each of the following statements concerning an arbitrary metric space (M, d) and subsets S, T of M.

3.38: Relative Compactness

Assume $S \subseteq T \subseteq M$. Then S is compact in (M, d) if, and only if, S is compact in the metric subspace (T, d).

Strategy: Show both directions. For the forward direction, convert open covers in the subspace to open covers in the full space. For the reverse direction, convert open covers in the full space to open covers in the subspace using intersections with T.

Solution: Suppose S is compact in (M, d). Let $\mathcal{F}$ be an open covering of S in the subspace (T, d). Then each $F \in \mathcal{F}$ is of the form $F = U \cap T$ where U is open in (M, d).

The collection $\{U : U \text{ is open in } (M, d) \text{ and } U \cap T \in \mathcal{F}\}$ is an open covering of S in (M, d). Since S is compact in (M, d), there exists a finite subcollection $\{U_1, \dots, U_n\}$ that covers S.

Then $\{U_1 \cap T, \dots, U_n \cap T\}$ is a finite subcollection of $\mathcal{F}$ that covers S, showing that S is compact in (T, d).

Conversely, suppose S is compact in (T, d). Let $\mathcal{G}$ be an open covering of S in (M, d). Then $\{G \cap T : G \in \mathcal{G}\}$ is an open covering of S in (T, d). Since S is compact in (T, d), there exists a finite subcollection $\{G_1 \cap T, \dots, G_n \cap T\}$ that covers S.

Then $\{G_1, \dots, G_n\}$ is a finite subcollection of $\mathcal{G}$ that covers S, showing that S is compact in (M, d). ■

3.39: Intersection with Compact Sets

If S is closed and T is compact, then $S \cap T$ is compact.

Strategy: Use the fact that compact sets are closed, so $S \cap T$ is the intersection of two closed sets (hence closed). Then use Exercise 3.38 to show that $S \cap T$ is compact in the subspace T, which implies it's compact in the full space.

Solution: Since T is compact, it is closed. Therefore, $S \cap T$ is the intersection of two closed sets, so it is closed.

Since $S \cap T \subseteq T$ and T is compact, by Exercise 3.38, $S \cap T$ is compact in (T, d). Since compactness is independent of the ambient space, $S \cap T$ is compact in (M, d). ■

3.40: Intersection of Compact Sets

The intersection of a nonempty collection of compact subsets of M is compact.

Strategy: Use the fact that compact sets are closed, so the intersection is closed. Then use Exercise 3.39 by taking any member of the collection as the compact set and the intersection as the closed set.

Solution: Let $\{K_\alpha\}$ be a nonempty collection of compact subsets of M. Since each K_α is closed, the intersection $\bigcap K_\alpha$ is closed.

Let K_1 be any member of the collection. Then $\bigcap K_\alpha \subseteq K_1$ and K_1 is compact. Since $\bigcap K_\alpha$ is closed and contained in a compact set, by Exercise 3.39, $\bigcap K_\alpha$ is compact. ■

3.41: Finite Union of Compact Sets

The union of a finite number of compact subsets of M is compact.

Strategy: Show that the union is closed (union of closed sets) and that any open cover of the union can be reduced to finite subcovers for each compact set, then combine them.

Solution: Let $K_1, K_2, \ldots, K_n$ be compact subsets of M. Since each K_i is closed, their union $\bigcup_{i=1}^n K_i$ is closed.

Let $\mathcal{F}$ be an open covering of $\bigcup_{i=1}^n K_i$. Then $\mathcal{F}$ is also an open covering of each K_i. Since each K_i is compact, there exists a finite subcollection $\mathcal{F}_i$ of $\mathcal{F}$ that covers K_i.

Then $\bigcup_{i=1}^n \mathcal{F}_i$ is a finite subcollection of $\mathcal{F}$ that covers $\bigcup_{i=1}^n K_i$.

Since $\bigcup_{i=1}^n K_i$ is closed and every open covering has a finite subcover, it is compact. ■

3.42: Non-Compact Closed and Bounded Set

Consider the metric space $\mathbb{Q}$ of rational numbers with the Euclidean metric of $\mathbb{R}$. Let S consist of all rational numbers in the open interval (a, b), where a and b are irrational. Then S is a closed and bounded subset of $\mathbb{Q}$ which is not compact.

Strategy: Show that S is bounded (contained in a bounded interval) and closed in $\mathbb{Q}$ (complement is open in $\mathbb{Q}$). For non-compactness, construct a sequence in S that converges to an irrational number outside S, showing it has no convergent subsequence in S.

Solution: Let $S = \mathbb{Q} \cap (a, b)$ where a, b are irrational numbers.

S is bounded since it is contained in the bounded interval (a, b).

S is closed in $\mathbb{Q}$ because its complement $\mathbb{Q} \setminus S = \mathbb{Q} \cap ((-\infty, a] \cup [b, \infty))$ is open in $\mathbb{Q}$.

However, S is not compact. Let $\{q_n\}$ be a sequence of rational numbers in (a, b) that converges to a (which exists since $\mathbb{Q}$ is dense in $\mathbb{R}$). Then $\{q_n\}$ is a sequence in S that has no convergent subsequence in S (since $a \notin S$).

Therefore, S is closed and bounded but not compact. ■

Miscellaneous Properties of Interior and Boundary

The following problems involve arbitrary subsets A and B of a metric space M.

Strategy: This section contains multiple exercises exploring the relationships between interior, closure, and boundary operations in metric spaces. Each problem will use the fundamental definitions and properties of these topological concepts.

Solution: This section presents a collection of exercises that demonstrate various properties of interior, closure, and boundary operations. The problems are solved individually in the following subsections, each building upon the basic definitions and established properties of these topological concepts in metric spaces. ■

3.VI Miscellaneous Properties of Interior and Boundary

Definitions and Theorems

Definition: Closure of a Set

The closure of a set S in a metric space (M, d), denoted by $\overline{S}$, is the union of S and its derived set: $\overline{S} = S \cup S'$.

Definition: Derived Set

The derived set of a set S in a metric space (M, d), denoted by S', is the set of all accumulation points of S.

Definition: Boundary of a Set

The boundary of a set S in a metric space (M, d), denoted by ∂S, is the intersection of the closure of S and the closure of its complement: $\partial S = \overline{S} \cap \overline{M \setminus S}$.

Theorem: Properties of Closure

Let S and T be subsets of a metric space (M, d). Then the following properties hold:

1. The closure $\overline{S}$ is a closed set.
2. The closure $\overline{S}$ is the smallest closed set containing S.
3. The closure of a union equals the union of closures: $\overline{S \cup T} = \overline{S} \cup \overline{T}$.
4. The closure of an intersection is contained in the intersection of closures: $\overline{S \cap T} \subseteq \overline{S} \cap \overline{T}$.
5. The derived set S' is a closed set.
6. The closure equals the union of the set and its derived set: $\overline{S} = S \cup S'$.

Theorem: Relations Between Interior and Closure

Let S be a subset of a metric space (M, d). Then the following relations hold:

1. The interior of S equals the complement of the closure of the complement: $\text{int } S = M \setminus \overline{M \setminus S}$.
2. The interior of the complement equals the complement of the closure: $\text{int } (M \setminus S) = M \setminus \overline{S}$.
3. The boundary of S equals the boundary of its complement: $\partial S = \partial(M \setminus S)$.

If A and B are subsets of a metric space M, prove that:

3.43: Interior via Closure

Prove that $\text{int } A = M - \overline{M - A}$.

Strategy: Show both inclusions. For the forward direction, if a point is interior to A, it has a neighborhood in A, so it's not in the closure of the complement. For the reverse direction, if a point is not in the closure of the complement, it has a neighborhood that doesn't intersect the complement, so it's interior to A.

Solution: Let $x \in \text{int } A$. Then there exists $\varepsilon > 0$ such that $B(x; \varepsilon) \subset A$. This means $B(x; \varepsilon) \cap (M - A) = \emptyset$, so $x \notin \overline{M - A}$. Therefore, $x \in M - \overline{M - A}$.

Conversely, let $x \in M - \overline{M - A}$. Then $x \notin \overline{M - A}$, so there exists $\varepsilon > 0$ such that $B(x; \varepsilon) \cap (M - A) = \emptyset$. This means $B(x; \varepsilon) \subset A$, so $x \in \text{int } A$. ■

3.44: Interior of Complement

Prove that $\text{int } (M - A) = M - \overline{A}$.

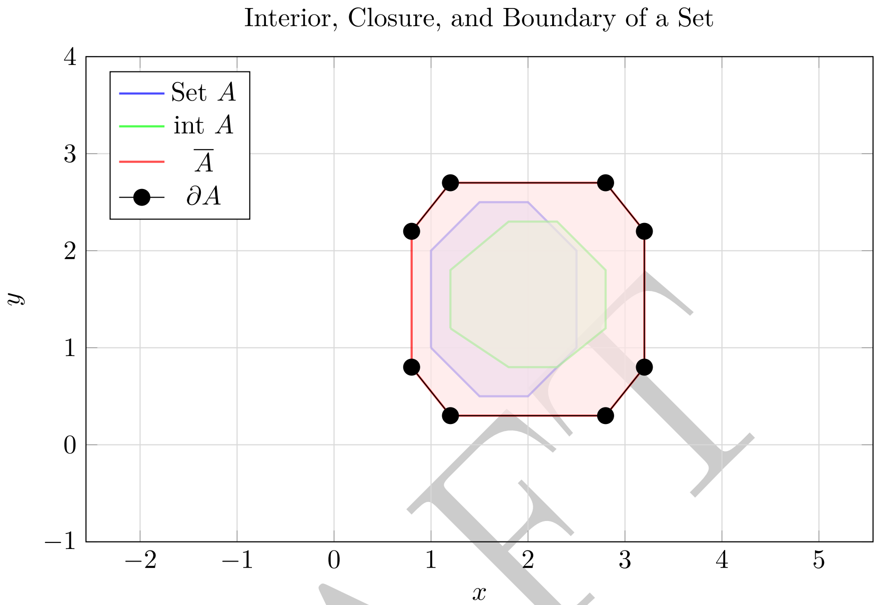


Figure 3.11: Relationships between interior, closure, and boundary: int $A \subseteq A \subseteq \overline{A}$ and $\partial A = \overline{A} \setminus$ int A.

Strategy: Use the same approach as Exercise 3.43 but with the roles of A and $M - A$ reversed. Show that a point is interior to the complement if and only if it's not in the closure of A.

Solution: Let $x \in$ int $(M - A)$. Then there exists $\varepsilon > 0$ such that $B(x; \varepsilon) \subset M - A$. This means $B(x; \varepsilon) \cap A = \emptyset$, so $x \notin \overline{A}$. Therefore, $x \in M - \overline{A}$.

Conversely, let $x \in M - \overline{A}$. Then $x \notin \overline{A}$, so there exists $\varepsilon > 0$ such that $B(x; \varepsilon) \cap A = \emptyset$. This means $B(x; \varepsilon) \subset M - A$, so $x \in$ int $(M - A)$. ■

3.45: Idempotence of Interior

Prove that int (int A) = int A.

Strategy: Show both inclusions. The forward inclusion is clear since the interior of a subset is contained in the subset. For the reverse inclusion, use the fact that if a point is interior to A, it has a neighborhood in A, and since that neighborhood is open, it's contained in the interior of A.

Solution: Since $\text{int } A \subseteq A$, we have $\text{int } (\text{int } A) \subseteq \text{int } A$.

For the reverse inclusion, let $x \in \text{int } A$. Then there exists $\varepsilon > 0$ such that $B(x;\varepsilon) \subset A$. Since $B(x;\varepsilon)$ is open and contained in A, we have $B(x;\varepsilon) \subset \text{int } A$. Therefore, $x \in \text{int } (\text{int } A)$. ■

3.46: Interior of Intersections

a) Prove that $\text{int } \left(\bigcap_{i=1}^{n} A_i\right) = \bigcap_{i=1}^{n}(\text{int } A_i)$, where each $A_i \subseteq M$.

b) Show that $\text{int } \left(\bigcap_{A\in F} A\right) \subseteq \bigcap_{A\in F}(\text{int } A)$ if F is an infinite collection of subsets of M.

c) Give an example where equality does not hold in (b).

Strategy: For (a), show both inclusions using the fact that if a point is interior to the intersection, it has a neighborhood in each set. For (b), show the inclusion using the same logic. For (c), use nested intervals that shrink to a single point.

Solution: (a) Let $x \in \text{int } \left(\bigcap_{i=1}^{n} A_i\right)$. Then there exists $\varepsilon > 0$ such that $B(x;\varepsilon) \subset \bigcap_{i=1}^{n} A_i$. This means $B(x;\varepsilon) \subset A_i$ for each i, so $x \in \text{int } A_i$ for each i. Therefore, $x \in \bigcap_{i=1}^{n}(\text{int } A_i)$.

Conversely, let $x \in \bigcap_{i=1}^{n}(\text{int } A_i)$. Then for each i, there exists $\varepsilon_i > 0$ such that $B(x;\varepsilon_i) \subset A_i$. Let $\varepsilon = \min\{\varepsilon_1, \ldots, \varepsilon_n\}$. Then $B(x;\varepsilon) \subset \bigcap_{i=1}^{n} A_i$, so $x \in \text{int } \left(\bigcap_{i=1}^{n} A_i\right)$.

(b) Let $x \in \text{int } \left(\bigcap_{A\in F} A\right)$. Then there exists $\varepsilon > 0$ such that $B(x;\varepsilon) \subset \bigcap_{A\in F} A$. This means $B(x;\varepsilon) \subset A$ for each $A \in F$, so $x \in \text{int } A$ for each $A \in F$. Therefore, $x \in \bigcap_{A\in F}(\text{int } A)$.

(c) Let $F = \{A_n : n \in \mathbb{N}\}$ where $A_n = (-1/n, 1/n)$. Then $\bigcap_{A\in F} A = \{0\}$, so $\text{int } \left(\bigcap_{A\in F} A\right) = \emptyset$. However, $\text{int } A_n = A_n$ for each n, so $\bigcap_{A\in F}(\text{int } A) = \bigcap_{n=1}^{\infty} A_n = \{0\}$. ■

3.47: Interior of Unions

a) Prove that $\bigcup_{A\in F}(\text{int } A) \subseteq \text{int } \left(\bigcup_{A\in F} A\right)$.

b) Give an example of a finite collection F in which equality does not hold in (a).

Strategy: For (a), show that if a point is interior to any set in the collection, it's interior to the union. For (b), use closed intervals that share a boundary point, where the union has interior points not in the union of interiors.

Solution: (a) Let $x \in \bigcup_{A\in F}(\text{int } A)$. Then $x \in \text{int } A$ for some $A \in F$. There exists $\varepsilon > 0$ such that $B(x;\varepsilon) \subset A \subset \bigcup_{A\in F} A$. Therefore, $x \in \text{int } (\bigcup_{A\in F} A)$.

(b) Let $F = \{A, B\}$ where $A = [0, 1]$ and $B = [1, 2]$. Then $\text{int } A = (0, 1)$ and $\text{int } B = (1, 2)$, so $\bigcup_{A\in F}(\text{int } A) = (0, 1) \cup (1, 2)$. However, $\bigcup_{A\in F} A = [0, 2]$, so $\text{int } (\bigcup_{A\in F} A) = (0, 2)$, which properly contains $(0, 1) \cup (1, 2)$. ■

3.48: Interior of Boundary

a) Prove that $\text{int } (\partial A) = \emptyset$ if A is open or if A is closed in M.

b) Give an example in which $\text{int } (\partial A) = M$.

Strategy: For (a), use the fact that the boundary of an open set is the closure minus the interior, and for a closed set it's the set minus the interior. In both cases, the boundary contains no open balls. For (b), use the rationals in $\mathbb{R}$ where the boundary is all of $\mathbb{R}$.

Solution: (a) If A is open, then $\partial A = \overline{A} \setminus \text{int } A = \overline{A} \setminus A$. If A is closed, then $\partial A = A \setminus \text{int } A$.

In both cases, ∂A contains no open balls, so $\text{int } (\partial A) = \emptyset$.

(b) Let $A = \mathbb{Q}$ in the metric space $\mathbb{R}$. Then $\partial A = \mathbb{R}$, so $\text{int } (\partial A) = \mathbb{R} = M$. ■

3.49: Interior of Union of Sets with Empty Interior

If int A = int $B = \emptyset$ and if A is closed in M, then int $(A \cup B) = \emptyset$.

Strategy: Use the fact that if a closed set has empty interior, every point is a limit point. Show that any point in the union cannot have a neighborhood entirely contained in the union by using the properties of limit points and the fact that B has empty interior.

Solution: Since A is closed, int $A = \emptyset$ implies that A has no isolated points. Therefore, every point in A is a limit point of A.

Let $x \in A \cup B$. If $x \in A$, then every neighborhood of x contains points of A different from x. Since $A \subset A \cup B$, every neighborhood of x contains points of $A \cup B$ different from x, so x is not an interior point of $A \cup B$.

If $x \in B \setminus A$, then since int $B = \emptyset$, every neighborhood of x contains points not in B. Since A is closed and $x \notin A$, there exists a neighborhood of x that doesn't intersect A. This neighborhood contains points not in $A \cup B$, so x is not an interior point of $A \cup B$.

Therefore, int $(A \cup B) = \emptyset$. ■

3.50: Counterexample for Union of Sets with Empty Interior

Give an example in which int A = int $B = \emptyset$ but int $(A \cup B) = M$.

Strategy: Use the rational and irrational numbers in $\mathbb{R}$. Both have empty interior (since they're dense and co-dense), but their union is all of $\mathbb{R}$ which has non-empty interior.

Solution: Let $A = \mathbb{Q}$ and $B = \mathbb{R} \setminus \mathbb{Q}$ in the metric space $\mathbb{R}$. Then int $A = \emptyset$ and int $B = \emptyset$, but $A \cup B = \mathbb{R}$, so int $(A \cup B) = \mathbb{R} = M$. ■

3.51: Properties of Boundary

Prove that:

$$\partial A = \overline{A} \cap \overline{M - A} \quad \text{and} \quad \partial A = \partial(M - A).$$

Strategy: For the first equality, use the definition of boundary as points that are adherent to both the set and its complement. For the second equality, use the first equality and the fact that the complement of the complement is the original set.

Solution: For the first equality, $x \in \partial A$ if and only if every neighborhood of x contains both points of A and points of $M - A$. This means $x \in \overline{A}$ and $x \in \overline{M - A}$, so $x \in \overline{A} \cap \overline{M - A}$.

For the second equality, $\partial A = \overline{A} \cap \overline{M - A} = \overline{M - A} \cap \overline{A} = \partial(M - A)$. ■

3.52: Boundary of Union under Disjoint Closures

If $\overline{A} \cap \overline{B} = \emptyset$, then $\partial(A \cup B) = \partial A \cup \partial B$.

Strategy: Use the fact that when closures are disjoint, the closure of the union is the union of closures. Show both inclusions by using the definition of boundary and the disjointness condition to separate the contributions from A and B.

Solution: Since $\overline{A} \cap \overline{B} = \emptyset$, we have $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

Let $x \in \partial(A \cup B)$. Then $x \in \overline{A \cup B} = \overline{A} \cup \overline{B}$ and $x \in \overline{M - (A \cup B)} = \overline{(M - A) \cap (M - B)} \subseteq \overline{M - A} \cap \overline{M - B}$.

If $x \in \overline{A}$, then $x \in \overline{A} \cap \overline{M - A} = \partial A$. If $x \in \overline{B}$, then $x \in \overline{B} \cap \overline{M - B} = \partial B$. Therefore, $x \in \partial A \cup \partial B$.

Conversely, let $x \in \partial A \cup \partial B$. Without loss of generality, assume $x \in \partial A$. Then $x \in \overline{A} \subseteq \overline{A \cup B}$ and $x \in \overline{M - A} \subseteq \overline{M - (A \cup B)}$. Therefore, $x \in \partial(A \cup B)$. ■

3.VII Solving and Proving Techniques

Proving Sets are Open

- Use the definition: show every point has a neighborhood contained in the set
- Use the fact that unions of open sets are open
- Use the fact that finite intersections of open sets are open
- Use the fact that open balls are open sets
- Use the fact that products of open intervals are open
- Use the fact that the interior of any set is open

Proving Sets are Closed

- Show the complement is open
- Use the fact that intersections of closed sets are closed
- Use the fact that finite unions of closed sets are closed
- Use the fact that the closure of any set is closed
- Use the fact that finite sets are closed
- Use the fact that closed balls are closed sets

Finding Accumulation Points

- Look for points that can be approached by sequences in the set
- Use density properties (rationals, irrationals are dense in $\mathbb{R}$)
- Consider convergence of sequences to boundary points
- Use the fact that accumulation points of accumulation points are accumulation points
- Use the fact that accumulation points of unions are unions of accumulation points
- Consider geometric intuition for sets in $\mathbb{R}^2$ and $\mathbb{R}^3$

Working with Interior and Closure

- Use the definition: interior points have neighborhoods in the set
- Use the fact that interior is the largest open subset
- Use the fact that closure is the smallest closed superset
- Use the relationship: $\operatorname{int} A = M - \overline{M - A}$
- Use the relationship: $\overline{A} = A \cup A'$ where A' is the derived set
- Use the fact that interior of interior equals interior
- Use the fact that closure of closure equals closure

Proving Countability

- Use the fact that countable unions of countable sets are countable
- Use the fact that Cartesian products of countable sets are countable
- Use the fact that subsets of countable sets are countable
- Use the fact that images of countable sets under injective functions are countable
- Use the Lindelöf property in separable spaces
- Use the fact that isolated points form a countable set
- Use the fact that disjoint open sets in separable spaces are countable

Working with Compactness

- Use the Heine-Borel theorem: closed and bounded in $\mathbb{R}^n$
- Use the fact that closed subsets of compact sets are compact
- Use the fact that finite unions of compact sets are compact
- Use the fact that intersections of compact sets are compact
- Use the fact that continuous images of compact sets are compact
- Use the fact that compactness is independent of the ambient space
- Use the fact that compact sets are closed and bounded

Using Density Properties

- Use the fact that rationals and irrationals are dense in $\mathbb{R}$
- Use the fact that $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$
- Use the fact that dense sets intersect every open set
- Use the fact that if A is dense in S and S is dense in T, then A is dense in T
- Use the fact that dense sets in open sets are dense in the whole space
- Use the fact that intersections of dense and open sets are dense

Working with Metric Spaces

- Use the triangle inequality to bound distances
- Use the fact that metric balls are open sets
- Use the fact that closed balls are closed sets
- Use the fact that finite sets are closed
- Use the fact that separable spaces have the Lindelöf property
- Use the fact that bounded metrics can be constructed from unbounded ones
- Use the fact that product metrics satisfy the metric axioms

Proving Connectedness

- Use proof by contradiction: assume the space can be split into two non-empty disjoint open sets
- Use the fact that $\mathbb{R}^1$ and $\mathbb{R}^n$ are connected
- Use the fact that connected spaces cannot have non-trivial clopen subsets
- Use the fact that continuous images of connected sets are connected
- Use the fact that unions of connected sets with non-empty intersection are connected

Using Proof by Contradiction

- Assume the opposite of what you want to prove
- Use the properties of open and closed sets to derive a contradiction
- Use the fact that limits must be unique
- Use the fact that countable sets cannot be uncountable
- Use the fact that compact sets must have finite subcovers
- Use the fact that connected spaces cannot be split into disjoint open sets
- Use the Cantor intersection theorem to find contradictions

Working with Sequences

- Use the fact that convergent sequences have unique limits
- Use the fact that subsequences of convergent sequences converge to the same limit
- Use the fact that bounded sequences have convergent subsequences
- Use the fact that Cauchy sequences converge in complete spaces
- Use the fact that accumulation points can be approached by sequences
- Use the fact that closed sets contain limits of convergent sequences

Proving Uniqueness

- Use the fact that closures are unique
- Use the fact that interiors are unique
- Use the fact that accumulation points are well-defined
- Use the fact that limits of sequences are unique
- Use the fact that compact sets have unique properties
- Use the fact that dense subsets are unique up to closure

Working with Boundaries

- Use the definition: $\partial A = \overline{A} \cap \overline{M - A}$
- Use the fact that boundaries of complements are the same
- Use the fact that boundaries of unions can be related to boundaries of components
- Use the fact that boundaries of open or closed sets have empty interior
- Use the fact that boundaries are closed sets
- Use the fact that boundaries separate sets from their complements

Chapter 4

Limits and Continuity

4.I Limits of Sequences

Definitions and Theorems

Definition: Cauchy Sequence

A sequence (x_n) in a metric space (S, d) is Cauchy if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon$ for all $n, m \geq N$.

Importance: Cauchy sequences provide a way to test for convergence without knowing the limit in advance. This is crucial in many applications where we can verify that terms are getting closer together but don't know what they're converging to. The concept of completeness (every Cauchy sequence converges) is fundamental to modern analysis.

Definition: Complete Metric Space

A metric space (S, d) is complete if every Cauchy sequence in S converges to a point in S.

Importance: Completeness is one of the most fundamental properties in analysis. It ensures that "almost convergent" sequences (Cauchy sequences) actually converge, which is essential for many existence proofs and the development of calculus. Complete spaces are the natural setting for most of modern analysis.

Definition: Metric Space

A metric space is a set S together with a function $d : S \times S \to [0, \infty)$ (called a metric) satisfying:

1. $d(x, y) = 0$ if and only if $x = y$
2. $d(x, y) \geq 0$ for all $x, y \in S$
3. $d(x, y) = d(y, x)$ for all $x, y \in S$
4. $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in S$ (triangle inequality)

Importance: Metric spaces provide the most general setting for studying convergence, continuity, and topology. They abstract the notion of distance from Euclidean space to any set where we can define a reasonable distance function. This abstraction is crucial for modern analysis and topology.

Theorem: Reverse Triangle Inequality

For any metric space (S, d) and $x, y, z \in S$: $|d(x, y) - d(x, z)| \leq d(y, z)$

Importance: This inequality is essential for proving that distance functions are continuous and for establishing bounds on how distances change. It's a direct consequence of the triangle inequality and is used extensively in analysis to control the behavior of distance functions.

Theorem: Geometric Sequence Convergence

If $|z| < 1$, then $z^n \to 0$. If $|z| > 1$, then (z^n) diverges.

Importance: This is one of the most fundamental convergence results in analysis. It provides a simple criterion for when powers of a complex number converge to zero, which is essential for understanding power series, geometric series, and many iterative processes.

Theorem: Ratio Test

Let (a_n) be a sequence of positive numbers. If $\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} < 1$, then $\sum a_n$ converges. If $\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} > 1$, then $\sum a_n$ diverges.

Importance: The ratio test is one of the most powerful and practical convergence tests for series. It compares consecutive terms to determine convergence, making it particularly useful for series with factorial or exponential terms. The use of limsup and liminf makes it robust even when the ratio doesn't have a simple limit.

Theorem: Sequential Compactness

A metric space is compact if and only if every sequence has a convergent subsequence.

Importance: This theorem provides a practical way to test for compactness using sequences, which is often easier than working with open covers. It connects the topological notion of compactness to the analytical concept of sequential convergence, making it a powerful tool in analysis.

Theorem: Heine-Borel Theorem

A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.

Importance: This is one of the most important theorems in analysis, providing a simple and practical characterization of compactness in Euclidean spaces. It connects the abstract notion of compactness to the concrete geometric properties of being closed and bounded, making it easy to identify compact sets in practice.

Definition: Connected Space

A metric space S is connected if it cannot be written as the union of two disjoint nonempty open sets.

Importance: Connectedness formalizes the intuitive idea that a space is "in one piece" and cannot be split into separate parts. It's essential for the intermediate value theorem and many existence proofs. Connected spaces preserve important properties under continuous functions and are fundamental in topology and analysis.

4.1: Limits of Sequences

Prove each of the following statements about sequences in $\mathbb{C}$:

(a) $z^n \to 0$ if $|z| < 1$; (z^n) diverges if $|z| > 1$.

(b) If $z_n \to 0$ and if (c_n) is bounded, then $(c_n z_n) \to 0$.

(c) $z^n/n! \to 0$ for every complex z.

(d) If $a_n = \sqrt{n^2+2} - n$, then $a_n \to 0$.

Strategy: Use the geometric sequence convergence theorem for (a), boundedness and convergence properties for (b), ratio test or Stirling's formula for (c), and rationalization technique for (d).

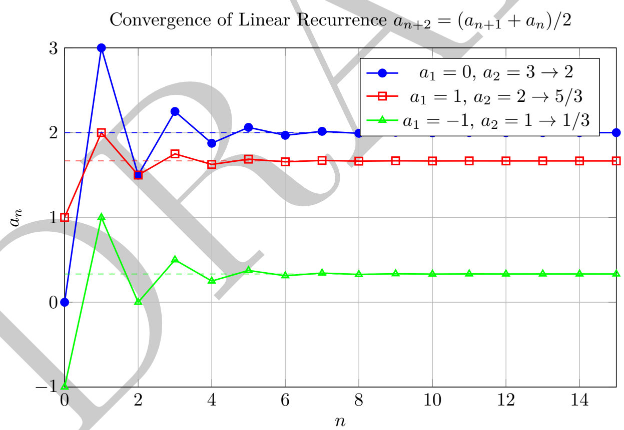


Figure 4.1: The sequence converges to $(a_1 + 2a_2)/3$ regardless of initial values, with oscillating behavior that dampens over time.

Solution:

(a) If $|z| < 1$, then $|z^n| = |z|^n \to 0$ by the geometric sequence property, hence $z^n \to 0$. If $|z| > 1$, then $|z^n| = |z|^n \to +\infty$, so (z^n) is unbounded and therefore not convergent in $\mathbb{C}$.

(b) If $|c_n| \le M$ for all n and $z_n \to 0$, then $|c_n z_n| \le M|z_n| \to 0$.

(c) Fix $z \in \mathbb{C}$. By the ratio test (or Stirling's formula),

$$\frac{|z|^{n+1}/(n+1)!}{|z|^n/n!} = \frac{|z|}{n+1} \to 0,$$

so $|z|^n/n! \to 0$, hence $z^n/n! \to 0$.

(d) Rationalize:

$$\begin{aligned} a_n &= \sqrt{n^2+2} - n \\ &= \frac{(\sqrt{n^2+2}-n)(\sqrt{n^2+2}+n)}{\sqrt{n^2+2}+n} \\ &= \frac{2}{\sqrt{n^2+2}+n} \\ &\sim \frac{2}{2n} \\ &= \frac{1}{n} \to 0. \end{aligned}$$

■

4.2: Linear Recurrence Relation

If $a_{n+2} = (a_{n+1} + a_n)/2$ for all $n \ge 1$, show that $a_n \to (a_1 + 2a_2)/3$. *Hint.* $a_{n+2} - a_{n+1} = \frac{1}{2}(a_n - a_{n+1})$.

Strategy: Use the hint to define a difference sequence $d_n = a_{n+1} - a_n$ that forms a geometric sequence. Express a_n in terms of initial values and the geometric series, then take the limit.

Solution: Using the hint, we have $a_{n+2} - a_{n+1} = \frac{1}{2}(a_n - a_{n+1})$. Let $d_n = a_{n+1} - a_n$ be the difference between consecutive terms. Then the recurrence becomes $d_{n+1} = -\frac{1}{2}d_n$.

This gives us $d_n = d_1 \cdot \left(-\frac{1}{2}\right)^{n-1}$, where $d_1 = a_2 - a_1$. Since $\left(-\frac{1}{2}\right)^n \to 0$ as $n \to \infty$, we have $d_n \to 0$.

Now, we can express a_n in terms of the initial terms and the differences:

$$a_n = a_1 + \sum_{k=1}^{n-1} d_k = a_1 + d_1 \sum_{k=1}^{n-1} \left(-\frac{1}{2}\right)^{k-1}$$

The sum $\sum_{k=1}^{n-1} \left(-\frac{1}{2}\right)^{k-1}$ is a geometric series that converges to $\frac{1}{1-\left(-\frac{1}{2}\right)} = \frac{2}{3}$ as $n \to \infty$.

Therefore, as $n \to \infty$:

$$a_n \to a_1 + d_1 \cdot \frac{2}{3} = a_1 + (a_2 - a_1) \cdot \frac{2}{3} = a_1 + \frac{2a_2}{3} - \frac{2a_1}{3} = \frac{a_1 + 2a_2}{3}$$

■

4.3: Recursive Sequence

If $0 < x_1 < 1$ and if $x_{n+1} = 1 - \sqrt{1 - x_n}$ for all $n \geq 1$, prove that $\{x_n\}$ is a decreasing sequence with limit 0. Prove also that $x_{n+1}/x_n \to \frac{1}{2}$.

Strategy: Use the concavity of $\sqrt{1-t}$ to show the sequence is decreasing and bounded, hence convergent. Find the limit by solving the fixed point equation, then use Taylor expansion to find the ratio limit.

Solution: For $t \in (0,1)$, the inequality $\sqrt{1-t} > 1 - \frac{t}{2}$ holds (concavity of $\sqrt{\cdot}$ or binomial expansion). Thus

$$x_{n+1} = 1 - \sqrt{1 - x_n} < 1 - \left(1 - \frac{x_n}{2}\right) = \frac{x_n}{2} < x_n,$$

so (x_n) is decreasing and bounded below by 0, hence convergent. Let $\lim x_n = L \geq 0$. Passing to the limit in $x_{n+1} = 1 - \sqrt{1 - x_n}$ gives $L = 1 - \sqrt{1 - L}$, whose solutions are $L \in \{0, 1\}$. Since $x_n \leq x_1 < 1$, we must have $L = 0$.

Moreover, using the Taylor expansion $\sqrt{1-t} = 1 - \frac{t}{2} - \frac{t^2}{8} + o(t^2)$ as $t \to 0^+$,

$$\frac{x_{n+1}}{x_n} = \frac{1 - \sqrt{1 - x_n}}{x_n} = \frac{\frac{x_n}{2} + \frac{x_n^2}{8} + o(x_n^2)}{x_n} \to \frac{1}{2}.$$

■

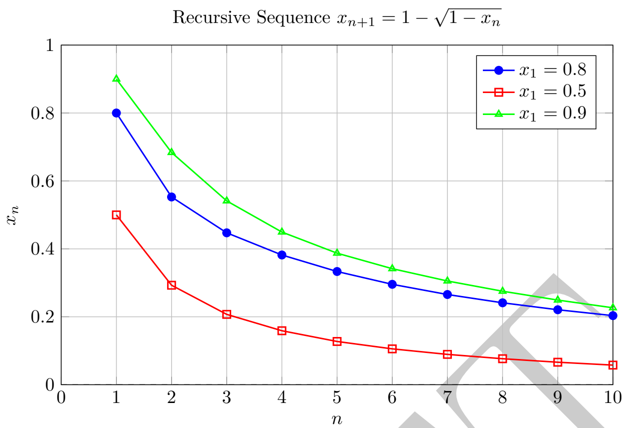


Figure 4.2: The sequence is decreasing and converges to 0, with the ratio x_{n+1}/x_n approaching $1/2$ as $n \to \infty$.

4.4: Quadratic Irrational Sequence

Two sequences of positive integers $\{a_n\}$ and $\{b_n\}$ are defined recursively by taking $a_1 = b_1 = 1$ and equating rational and irrational parts in the equation

$$a_n + b_n\sqrt{2} = (a_{n-1} + b_{n-1}\sqrt{2})^2 \quad \text{for } n \geq 2.$$

Prove that $a_n^2 - 2b_n^2 = 1$ for $n \geq 2$. Deduce that $a_n/b_n \to \sqrt{2}$ through values $> \sqrt{2}$, and that $2b_n/a_n \to \sqrt{2}$ through values $< \sqrt{2}$.

Strategy: Expand the quadratic expression and equate rational/irrational parts to find recurrence relations. Use mathematical induction to prove the identity. For the limits, manipulate the identity to express ratios and use the fact that sequences are increasing.

Solution:

Part 1: Prove that $a_n^2 - 2b_n^2 = 1$ for $n \geq 2$

First, let's find the recursive relations for a_n and b_n. We expand the right side of the given equation:

$$\begin{aligned} a_n + b_n\sqrt{2} &= (a_{n-1} + b_{n-1}\sqrt{2})^2 \\ &= a_{n-1}^2 + 2a_{n-1}b_{n-1}\sqrt{2} + (b_{n-1}\sqrt{2})^2 \\ &= (a_{n-1}^2 + 2b_{n-1}^2) + (2a_{n-1}b_{n-1})\sqrt{2} \end{aligned}$$

By equating the rational and irrational parts, we obtain the recurrence relations:

$$a_n = a_{n-1}^2 + 2b_{n-1}^2 \tag{4.1}$$

$$b_n = 2a_{n-1}b_{n-1} \tag{4.2}$$

We will prove the statement $a_n^2 - 2b_n^2 = 1$ for $n \geq 2$ by mathematical induction.

Base Case (n=2): Given $a_1 = 1$ and $b_1 = 1$. Using the recurrence relations:

$$\begin{aligned} a_2 &= a_1^2 + 2b_1^2 = 1^2 + 2(1^2) = 3 \\ b_2 &= 2a_1b_1 = 2(1)(1) = 2 \end{aligned}$$

Now, we check the condition for $n = 2$:

$$a_2^2 - 2b_2^2 = 3^2 - 2(2^2) = 9 - 2(4) = 9 - 8 = 1.$$

The base case holds.

Inductive Hypothesis: Assume the statement is true for some integer $k \geq 2$. That is, we assume:

$$a_k^2 - 2b_k^2 = 1$$

Inductive Step: We want to prove that the statement is true for $n = k+1$, i.e., $a_{k+1}^2 - 2b_{k+1}^2 = 1$. We start with the left-hand side and substitute the recurrence relations for a_{k+1} and b_{k+1}:

$$\begin{aligned} a_{k+1}^2 - 2b_{k+1}^2 &= (a_k^2 + 2b_k^2)^2 - 2(2a_kb_k)^2 \\ &= (a_k^4 + 4a_k^2b_k^2 + 4b_k^4) - 2(4a_k^2b_k^2) \\ &= a_k^4 + 4a_k^2b_k^2 + 4b_k^4 - 8a_k^2b_k^2 \\ &= a_k^4 - 4a_k^2b_k^2 + 4b_k^4 \\ &= (a_k^2 - 2b_k^2)^2 \end{aligned}$$

By the inductive hypothesis, we know that $a_k^2 - 2b_k^2 = 1$. Substituting this into our expression:

$$a_{k+1}^2 - 2b_{k+1}^2 = (1)^2 = 1.$$

Thus, the statement holds for $n = k + 1$.

By the principle of mathematical induction, $a_n^2 - 2b_n^2 = 1$ for all $n \geq 2$.

Part 2: Deductions about the limits

Convergence of a_n/b_n to $\sqrt{2}$ From the result $a_n^2 - 2b_n^2 = 1$ for $n \geq 2$, we can rearrange the equation. Since b_n is a sequence of positive integers, $b_n \neq 0$, so we can divide by b_n^2:

$$\frac{a_n^2}{b_n^2} - 2 = \frac{1}{b_n^2}$$

$$\left(\frac{a_n}{b_n}\right)^2 = 2 + \frac{1}{b_n^2}$$

Since a_n and b_n are positive, $a_n/b_n > 0$. Taking the square root of both sides:

$$\frac{a_n}{b_n} = \sqrt{2 + \frac{1}{b_n^2}}$$

The sequences are defined for $a_1 = 1, b_1 = 1$, and for $n \geq 2$, $a_n = a_{n-1}^2 + 2b_{n-1}^2 > a_{n-1}$ and $b_n = 2a_{n-1}b_{n-1} > b_{n-1}$. Thus, $\{a_n\}$ and $\{b_n\}$ are strictly increasing sequences of positive integers, which means $b_n \to \infty$ as $n \to \infty$. Consequently,

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

Taking the limit of our expression for a_n/b_n:

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \sqrt{2 + \frac{1}{b_n^2}} = \sqrt{2+0} = \sqrt{2}.$$

To show that the convergence is through values greater than $\sqrt{2}$, we note that for any $n \geq 1$, $b_n^2 > 0$, so $\frac{1}{b_n^2} > 0$. Therefore:

$$\left(\frac{a_n}{b_n}\right)^2 = 2 + \frac{1}{b_n^2} > 2$$

Taking the square root of both sides (since $a_n/b_n > 0$):

$$\frac{a_n}{b_n} > \sqrt{2}.$$

Thus, the sequence a_n/b_n converges to $\sqrt{2}$ through values strictly greater than $\sqrt{2}$.

Convergence of $2b_n/a_n$ to $\sqrt{2}$ Again, we start with $a_n^2 - 2b_n^2 = 1$. Since a_n is a sequence of positive integers, $a_n \neq 0$. We divide by a_n^2:

$$1 - 2\frac{b_n^2}{a_n^2} = \frac{1}{a_n^2}$$

Rearranging the terms:

$$1 - \frac{1}{a_n^2} = 2\left(\frac{b_n}{a_n}\right)^2$$

Multiply by 2:

$$2 - \frac{2}{a_n^2} = 4\left(\frac{b_n}{a_n}\right)^2 = \left(\frac{2b_n}{a_n}\right)^2$$

Since b_n and a_n are positive, we can take the square root:

$$\frac{2b_n}{a_n} = \sqrt{2 - \frac{2}{a_n^2}}$$

As established before, $a_n \to \infty$ as $n \to \infty$. This implies:

$$\lim_{n\to\infty} \frac{2}{a_n^2} = 0.$$

Taking the limit of our expression for $2b_n/a_n$:

$$\lim_{n\to\infty} \frac{2b_n}{a_n} = \lim_{n\to\infty} \sqrt{2 - \frac{2}{a_n^2}} = \sqrt{2-0} = \sqrt{2}.$$

To show that the convergence is through values less than $\sqrt{2}$, we note that for any $n \geq 1$, $a_n^2 > 0$, so $\frac{2}{a_n^2} > 0$. Therefore:

$$\left(\frac{2b_n}{a_n}\right)^2 = 2 - \frac{2}{a_n^2} < 2$$

Taking the square root of both sides (since $2b_n/a_n > 0$):

$$\frac{2b_n}{a_n} < \sqrt{2}.$$

Thus, the sequence $2b_n/a_n$ converges to $\sqrt{2}$ through values strictly less than $\sqrt{2}$. ■

4.5: Cubic Recurrence

A real sequence $\{x_n\}$ satisfies $7x_{n+1} = x_n^3 + 6$ for $n \geq 1$. If $x_1 = \frac{1}{2}$, prove that the sequence increases and find its limit. What happens if $x_1 = \frac{3}{2}$ or if $x_1 = \frac{5}{2}$?

Strategy: Define $f(x) = (x^3 + 6)/7$ and analyze its behavior. Use the fact that f is increasing and find fixed points. Analyze the difference $f(x) - x$ to determine monotonicity and convergence behavior for different initial values. **Solution:** Let $f(x) = \dfrac{x^3 + 6}{7}$. Then

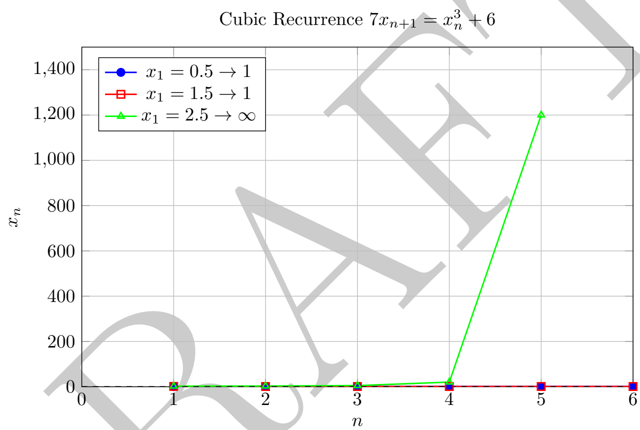


Figure 4.3: For $x_1 < 2$, the sequence converges to the fixed point 1. For $x_1 > 2$, the sequence diverges to infinity.

$f'(x) = \dfrac{3x^2}{7} \geq 0$, so f is increasing. Also $f(1) = 1$ and for $x \in [0, 1]$,

$$f(x) - x = \frac{x^3 - 7x + 6}{7} = \frac{(x-1)(x+3)(x-2)}{7} > 0,$$

so f maps $[0, 1]$ into itself and $f(x) > x$ for $x \in (0, 1)$. Starting with $x_1 = \frac{1}{2}$, we get $0 < x_1 < x_2 < \cdots \leq 1$, hence $x_n \uparrow L \in [0, 1]$. Passing to the limit in $x_{n+1} = f(x_n)$ gives $L = f(L)$, i.e., $L = 1$. Thus for $x_1 = \frac{1}{2}$, $x_n \uparrow 1$.

If $x_1 = \frac{3}{2} \in (1, 2)$, then using the same factorization,

$$f(x) - x = \frac{(x-1)(x+3)(x-2)}{7} < 0 \quad (1 < x < 2),$$

so $x_{n+1} < x_n$ and $x_n > 1$; the sequence decreases and is bounded below by 1, hence $x_n \downarrow 1$.

If $x_1 = \frac{5}{2} > 2$, then $f(x) - x > 0$ (all three factors positive), so $x_{n+1} > x_n$ and $x_n \to +\infty$ (since for large x, $f(x) \sim x^3/7 > x$).

■

4.6: Convergence Condition

If $|a_n| < 2$ and $|a_{n+2} - a_{n+1}| \leq \frac{1}{8}|a_{n+1}^2 - a_n^2|$ for all $n \geq 1$, prove that $\{a_n\}$ converges.

Strategy: Use the boundedness condition to bound $|a_{n+1} + a_n|$, then factor the difference of squares to show the sequence is Cauchy. Use the contraction-like property to establish convergence.

Solution: We will show that the sequence $\{a_n\}$ is Cauchy and therefore converges.

Since $|a_{n+1}|, |a_n| < 2$ for all n, we have $|a_{n+1}+a_n| \leq |a_{n+1}|+|a_n| < 4$.

Using the given condition and factoring the difference of squares:

$$|a_{n+2}-a_{n+1}| \leq \frac{1}{8}|a_{n+1}^2-a_n^2| = \frac{1}{8}|a_{n+1}-a_n|\cdot|a_{n+1}+a_n| \leq \frac{1}{8}\cdot 4\cdot|a_{n+1}-a_n| = \frac{1}{2}|a_{n+1}-a_n|.$$

This shows that the sequence has a contraction-like property. By induction, we can show that for any $k \geq 1$:

$$|a_{n+k} - a_{n+k-1}| \leq 2^{-k+1}|a_{n+1} - a_n|.$$

Now, for any $m > n$, using the triangle inequality:

$$|a_m - a_n| \leq \sum_{k=n}^{m-1} |a_{k+1} - a_k| \leq |a_{n+1} - a_n| \sum_{j=0}^{\infty} 2^{-j} = 2|a_{n+1} - a_n|.$$

Since $|a_{n+1} - a_n| \to 0$ as $n \to \infty$ (by the contraction property), we have $|a_m - a_n| \to 0$ as $m, n \to \infty$. This means $\{a_n\}$ is a Cauchy sequence, and since we're working in a complete metric space (the real numbers), the sequence converges. ■

4.7: Metric Space Convergence

In a metric space (S, d), assume that $x_n \to x$ and $y_n \to y$. Prove that $d(x_n, y_n) \to d(x, y)$.

Strategy: Use the reverse triangle inequality to bound $|d(x_n, y_n) - d(x, y)|$ in terms of $d(x_n, x)$ and $d(y_n, y)$, then use the convergence of the sequences.

Solution: We need to show that $d(x_n, y_n) \to d(x, y)$ as $n \to \infty$.

By the reverse triangle inequality, we have:

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y).$$

Since $x_n \to x$ and $y_n \to y$, we have $d(x_n, x) \to 0$ and $d(y_n, y) \to 0$ as $n \to \infty$.

Therefore:

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y) \to 0 + 0 = 0.$$

This means that $d(x_n, y_n) \to d(x, y)$ as $n \to \infty$, which completes the proof. ■

4.8: Compact Metric Spaces

Prove that in a compact metric space (S, d), every sequence in S has a subsequence which converges in S. This property also implies that S is compact but you are not required to prove this. (For a proof see either Reference 4.2 or 4.3.)

Strategy: Use the limit-point property of compact sets: if a set is compact and infinite, it has a limit point. If a value appears infinitely often, use the constant subsequence. Otherwise, use the limit point to construct a convergent subsequence.

Solution: We will prove that in a compact metric space (S, d), every sequence has a convergent subsequence.

Lemma: If S is compact and $A \subseteq S$ is infinite, then A has a limit point in S.

Proof of Lemma: Suppose A has no limit point. Then each $a \in A$ is isolated in A: there exists $r_a > 0$ with $B(a, r_a) \cap (A \setminus \{a\}) = \emptyset$.

Consider the open cover $\mathcal{U} = \{S \setminus A\} \cup \{B(a, r_a) : a \in A\}$ of S. Any finite subfamily of $\mathcal{U}$ uses only finitely many of the balls $B(a, r_a)$, hence covers at most finitely many points of A; thus it cannot cover S. This contradicts compactness. Hence A has a limit point in S. $\square$

Proof of Main Result: Let (x_n) be a sequence in S.

Case 1: If some value $x \in S$ occurs infinitely often among the x_n, then the constant subsequence $x, x, \ldots$ converges to x.

Case 2: Otherwise, the set $A = \{x_n : n \in \mathbb{N}\}$ is infinite. By the Lemma, A has a limit point $x \in S$.

By definition of limit point, every ball $B(x, \varepsilon)$ contains infinitely many terms of the sequence. We can construct a convergent subsequence as follows:

a) Choose n_1 with $d(x_{n_1}, x) < 1$

b) Having chosen n_k, pick $n_{k+1} > n_k$ with $d(x_{n_{k+1}}, x) < 1/(k+1)$

Then $x_{n_k} \to x$ because for any $\varepsilon > 0$, pick K with $1/K < \varepsilon$; for all $k \geq K$, we have $d(x_{n_k}, x) < \varepsilon$.

Thus (x_n) has a subsequence converging in S. ■

4.9: Complete Subsets

Let A be a subset of a metric space S. If A is complete, prove that A is closed. Prove that the converse also holds if S is complete.

Strategy: For the first part, use the fact that any convergent sequence in A is Cauchy and must converge to a point in A (by completeness), so A contains all its limit points. For the converse, use the fact that Cauchy sequences in A converge in S (by completeness of S) and the limit must be in A (by closedness).

Solution: We need to prove two statements:

a) If A is complete, then A is closed.

b) If S is complete and A is closed, then A is complete.

Proof of (a): Suppose A is complete. To show that A is closed, we need to show that A contains all its limit points.

Let $x \in \overline{A}$. Then there exists a sequence $(a_n) \subset A$ such that $a_n \to x$. Since any convergent sequence is Cauchy and A is complete, the limit of this sequence must lie in A. Therefore, $x \in A$, which shows that A is closed.

Proof of (b): Suppose S is complete and A is closed in S. To show that A is complete, we need to show that every Cauchy sequence in A converges to a point in A.

Let (a_n) be a Cauchy sequence in A. Since S is complete, (a_n) converges to some point $x \in S$. Since A is closed and $(a_n) \subset A$ with $a_n \to x$, we must have $x \in A$. Therefore, A is complete. ■

4.II Limits of Functions

Definitions and Theorems

Definition: Metric Space

A metric space is a set S together with a function $d : S \times S \to [0, \infty)$ (called a metric) satisfying:

1. $d(x, y) = 0$ if and only if $x = y$
2. $d(x, y) \geq 0$ for all $x, y \in S$
3. $d(x, y) = d(y, x)$ for all $x, y \in S$
4. $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in S$ (triangle inequality)

Importance: Metric spaces provide the most general setting for studying convergence, continuity, and topology. They abstract the notion of distance from Euclidean space to any set where we can define a reasonable distance function. This abstraction is crucial for modern analysis and topology.

Theorem: Reverse Triangle Inequality

For any metric space (S, d) and $x, y, z \in S$: $|d(x, y) - d(x, z)| \leq d(y, z)$

Importance: This inequality is essential for proving that distance functions are continuous and for establishing bounds on how distances change. It's a direct consequence of the triangle inequality and is used extensively in analysis to control the behavior of distance functions.

Note: In Exercise 4.10 through 4.28, all functions are real-valued.
Note: In Exercise 4.10 through 4.28, all functions are real-valued.

4.10: Function Limit Properties

Let f be defined on an open interval (a, b) and assume $x \in (a, b)$. Consider the two statements:

(a) $\lim_{h\to 0} |f(x + h) - f(x)| = 0$;

(b) $\lim_{h\to 0} |f(x + h) - f(x - h)| = 0$.

Prove that (a) always implies (b), and give an example in which (b) holds but (a) does not.

Strategy: For the implication, use the triangle inequality to bound the symmetric difference by the sum of two one-sided differences. For the counterexample, construct a function that is symmetric around x but discontinuous at x.

Solution:

(a) $\Rightarrow$ (b): By the triangle inequality,

$$|f(x + h) - f(x - h)| \leq |f(x + h) - f(x)| + |f(x) - f(x - h)| \to 0.$$

Example where (b) holds but (a) does not: Define

$$f(t) = \begin{cases} 1, & t \neq 0, \\ 0, & t = 0. \end{cases}$$

Then for $h \neq 0$, $f(h) = f(-h) = 1$, so $|f(h) - f(-h)| = 0 \to 0$; but $|f(h) - f(0)| = 1 \not\to 0$, so (a) fails at $x = 0$. ■

4.11: Double Limits

Exercise 4.11

Let f be defined on $\mathbb{R}^2$. If

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

and if the one-dimensional limits $\lim_{x\to a} f(x,y)$ and $\lim_{y\to b} f(x,y)$ both exist, prove that

$$\lim_{x\to a}\left[\lim_{y\to b} f(x,y)\right] = \lim_{y\to b}\left[\lim_{x\to a} f(x,y)\right] = L.$$

Now consider the functions f defined on $\mathbb{R}^2$ as follows:

a) $f(x,y) = \frac{x^2-y^2}{x^2+y^2}$ if $(x,y) \neq (0,0)$, $f(0,0) = 0$.

b) $f(x,y) = \frac{(xy)^2}{(xy)^2+(x-y)^2}$ if $(x,y) \neq (0,0)$, $f(0,0) = 0$.

c) $f(x,y) = \frac{1}{x}\sin(xy)$ if $x \neq 0$, $f(0,y) = y$.

d) $f(x,y) = \begin{cases} (x+y)\sin(1/x)\sin(1/y) & \text{if } x \neq 0 \text{ and } y \neq 0, \\ 0 & \text{if } x = 0 \text{ or } y = 0. \end{cases}$

e) $f(x,y) = \begin{cases} \frac{\sin x - \sin y}{\tan x - \tan y} & \text{if } \tan x \neq \tan y, \\ \cos^3 x & \text{if } \tan x = \tan y. \end{cases}$

In each of the preceding examples, determine whether the following limits exist and evaluate those limits that do exist:

$$\lim_{x\to 0}\left[\lim_{y\to 0} f(x,y)\right]; \quad \lim_{y\to 0}\left[\lim_{x\to 0} f(x,y)\right]; \quad \lim_{(x,y)\to(0,0)} f(x,y).$$

Strategy: For the theoretical part, use the definition of the two-dimensional limit to show that for x close to a, the one-dimensional limit $\lim_{y\to b} f(x,y)$ exists and equals L. Then take the limit as $x \to a$ to establish the result. For the examples, analyze each function by computing the iterated limits and the two-dimensional limit separately, using techniques like polar coordinates, path analysis, or direct evaluation.

Solution:

Theoretical Part: Given $\varepsilon > 0$, choose $\delta > 0$ so that $\sqrt{(x-a)^2 + (y-b)^2} < \delta$ implies $|f(x,y) - L| < \varepsilon$. Fix x with $|x-a| < \delta$. Then for $|y-b| < \sqrt{\delta^2 - (x-a)^2}$ we have $|(x,y) - (a,b)| < \delta$, hence $|f(x,y) - L| < \varepsilon$. This shows $\lim_{y\to b} f(x,y) = L$ for all x close to a, and therefore $\lim_{x\to a} \left[\lim_{y\to b} f(x,y)\right] = L$. The other equality is analogous.

Examples:

(a) $f(x,y) = \frac{x^2-y^2}{x^2+y^2}$ for $(x,y) \neq (0,0)$, $f(0,0) = 0$.

For fixed $x \neq 0$: $\lim_{y\to 0} f(x,y) = \frac{x^2}{x^2} = 1$, so $\lim_{x\to 0} \left[\lim_{y\to 0} f(x,y)\right] = 1$.

For fixed $y \neq 0$: $\lim_{x\to 0} f(x,y) = \frac{-y^2}{y^2} = -1$, so $\lim_{y\to 0} \left[\lim_{x\to 0} f(x,y)\right] = -1$.

The two-dimensional limit does not exist: along $y = 0$, $f(x,0) = 1$; along $x = 0$, $f(0,y) = -1$.

(b) $f(x,y) = \frac{(xy)^2}{(xy)^2+(x-y)^2}$ for $(x,y) \neq (0,0)$, $f(0,0) = 0$.

For fixed $x \neq 0$: $\lim_{y\to 0} f(x,y) = \frac{0}{0+x^2} = 0$, so $\lim_{x\to 0} \left[\lim_{y\to 0} f(x,y)\right] = 0$.

For fixed $y \neq 0$: $\lim_{x\to 0} f(x,y) = \frac{0}{0+y^2} = 0$, so $\lim_{y\to 0} \left[\lim_{x\to 0} f(x,y)\right] = 0$.

The two-dimensional limit is 0: for $(x,y) \neq (0,0)$, $|f(x,y)| \leq \frac{(xy)^2}{(xy)^2} = 1$, and along $x = y$, $f(x,x) = \frac{x^4}{x^4+0} = 1$, so the limit does not exist.

(c) $f(x,y) = \frac{1}{x}\sin(xy)$ for $x \neq 0$, $f(0,y) = y$.

For fixed $x \neq 0$: $\lim_{y\to 0} f(x,y) = \frac{1}{x}\sin(0) = 0$, so $\lim_{x\to 0} \left[\lim_{y\to 0} f(x,y)\right] = 0$.

For fixed y: $\lim_{x\to 0} f(x,y) = \lim_{x\to 0} \frac{\sin(xy)}{x} = y$ (using L'Hôpital's rule), so $\lim_{y\to 0} \left[\lim_{x\to 0} f(x,y)\right] = 0$.

Note on L'Hôpital's rule: When $y \neq 0$, as $x \to 0$, both the numerator $\sin(xy)$ and denominator x approach 0, giving us the indeterminate form $\frac{0}{0}$. L'Hôpital's rule allows us to compute this limit by taking the derivative of both numerator and denominator with respect to x:

$$\lim_{x\to 0} \frac{\sin(xy)}{x} = \lim_{x\to 0} \frac{\frac{d}{dx}[\sin(xy)]}{\frac{d}{dx}[x]} = \lim_{x\to 0} \frac{y\cos(xy)}{1} = y\cos(0) = y$$

When $y = 0$, the limit is simply $\lim_{x\to 0} \frac{0}{x} = 0$, which agrees with the formula $y = 0$.

The two-dimensional limit is 0: $|f(x,y)| \leq \frac{|xy|}{|x|} = |y| \to 0$ as $(x,y) \to (0,0)$.

(d) $f(x,y) = (x+y)\sin(1/x)\sin(1/y)$ for $x \neq 0$ and $y \neq 0$, $f(x,y) = 0$ otherwise.

For fixed $x \neq 0$: $\lim_{y\to 0} f(x,y) = 0$ (since $f(x,y) = 0$ when $y = 0$), so $\lim_{x\to 0} \left[\lim_{y\to 0} f(x,y)\right] = 0$.

For fixed $y \neq 0$: $\lim_{x\to 0} f(x,y) = 0$ (since $f(x,y) = 0$ when $x = 0$), so $\lim_{y\to 0} \left[\lim_{x\to 0} f(x,y)\right] = 0$.

The two-dimensional limit does not exist: along $x = y$, $f(x,x) = 2x\sin(1/x)\sin(1/x)$ oscillates as $x \to 0$.

(e) $f(x,y) = \frac{\sin x - \sin y}{\tan x - \tan y}$ for $\tan x \neq \tan y$, $f(x,y) = \cos^3 x$ for $\tan x = \tan y$.

For fixed $x \neq 0$: $\lim_{y\to 0} f(x,y) = \frac{\sin x - 0}{\tan x - 0} = \cos x$, so

$$\lim_{x\to 0}\left[\lim_{y\to 0} f(x,y)\right] = 1$$

.

For fixed $y \neq 0$: $\lim_{x\to 0} f(x,y) = \frac{0-\sin y}{0-\tan y} = \cos y$, so

$$\lim_{y\to 0}\left[\lim_{x\to 0} f(x,y)\right] = 1$$

.

The two-dimensional limit is 1: using L'Hôpital's rule twice,

$$\begin{aligned}\lim_{(x,y)\to(0,0)} \frac{\sin x - \sin y}{\tan x - \tan y} &= \lim_{(x,y)\to(0,0)} \frac{\cos x - \cos y}{\sec^2 x - \sec^2 y} \\ &= \lim_{(x,y)\to(0,0)} \frac{-\sin x + \sin y}{2\sec x \tan x - 2\sec y \tan y} = 1.\end{aligned}$$

■

4.12: Limit of Nested Cosine

If $x \in [0,1]$ prove that the following limit exists,

$$\lim_{m\to\infty}\left[\lim_{n\to\infty} \cos^{2n}(m!\pi x)\right],$$

and that its value is 0 or 1, according to whether x is irrational or rational.

Strategy: First evaluate the inner limit for fixed m using the fact that $\cos^{2n}(\theta) \to 1$ if $\cos^2(\theta) = 1$ and $\cos^{2n}(\theta) \to 0$ otherwise. Then analyze when $m!\pi x$ is an integer multiple of π based on whether x is rational or irrational.

Solution: For fixed m, the inner limit is $\lim_{n\to\infty}\cos^{2n}(\theta) = \begin{cases} 1, & \cos^2\theta = 1, \\ 0, & \cos^2\theta < 1. \end{cases}$
Thus it equals 1 iff $m!\pi x$ is an integer multiple of π, i.e., iff $m!x \in \mathbb{Z}$. If $x = \frac{p}{q}$ is rational (in lowest terms), then for all $m \geq q$ we have $m!x \in \mathbb{Z}$, so the inner limit equals 1 for all large m, hence the outer limit is 1. If x is irrational, then $m!x \notin \mathbb{Z}$ for every m, so the inner limit is always 0, hence the outer limit is 0. ■

4.III Continuity of real-valued functions

Definitions and Theorems

Theorem: Sequential Continuity

A function $f : S \to T$ is continuous at x if and only if $f(x_n) \to f(x)$ whenever $x_n \to x$.

Importance: This theorem provides an alternative characterization of continuity that is often easier to use in proofs. It connects the concept of continuity directly to sequence convergence, making it a powerful tool for establishing continuity and proving many important results in analysis.

Theorem: Extreme Value Theorem

A continuous function on a compact set attains its maximum and minimum values.

Importance: This is one of the most important existence theorems in analysis. It guarantees that optimization problems have solutions under reasonable conditions, making it fundamental for calculus, optimization theory, and many applications in science and engineering.

Theorem: Intermediate Value Theorem

If f is continuous on $[a, b]$ and $f(a) < c < f(b)$, then there exists $x \in (a, b)$ such that $f(x) = c$.

Importance: This theorem formalizes the intuitive idea that a continuous function cannot "jump" over values. It's essential for proving the existence of solutions to equations and is the foundation for many numerical methods like the bisection method and intermediate value property.

Theorem: Continuous Image of Connected Set

The continuous image of a connected set is connected.

Importance: This theorem shows that connectedness is preserved under continuous functions, making it a topological invariant. This is essential for many applications where we want to transfer connectedness properties from one space to another through continuous mappings.

Theorem: Continuous Image of Compact Set

The continuous image of a compact set is compact.

Importance: This theorem shows that compactness is preserved under continuous functions. It's crucial for many applications in analysis, particularly for proving that continuous functions on compact domains have bounded ranges and attain their extrema.

Theorem: Continuous Image of Bounded Set

The continuous image of a bounded set is bounded.

Importance: This theorem ensures that continuous functions preserve boundedness, which is essential for many applications in analysis and optimization. It's particularly important when working with continuous functions on bounded domains.

4.13: Zero Function on Rationals

Let f be continuous on $[a, b]$ and let $f(x) = 0$ when x is rational. Prove that $f(x) = 0$ for every x in $[a, b]$.

Strategy: Use the density of rational numbers in the reals and the sequential characterization of continuity. For any irrational x, construct a sequence of rationals converging to x and use continuity to show $f(x) = 0$.

Solution: We need to prove that $f(x) = 0$ for every x in $[a, b]$. We already know that $f(x) = 0$ when x is rational, so we only need to prove it for irrational numbers.

Let x be any irrational number in $[a, b]$. Since the rational numbers are dense in the real numbers, we can find a sequence of rational numbers (q_n) that converges to x. This means that as n gets larger and larger, the rational numbers q_n get closer and closer to x.

For example, if $x = \sqrt{2}$ (which is irrational), we could use the sequence:

$$q_1 = 1.4, \quad q_2 = 1.41, \quad q_3 = 1.414, \quad q_4 = 1.4142, \ldots$$

Each q_n is a rational number (a finite decimal), and the sequence converges to $\sqrt{2}$.

Now, since f is continuous at x, the sequential characterization of continuity tells us that:

$$f(x) = \lim_{n \to \infty} f(q_n)$$

But we know that each q_n is rational, so by the given condition, $f(q_n) = 0$ for every n. Therefore:

$$f(x) = \lim_{n \to \infty} f(q_n) = \lim_{n \to \infty} 0 = 0$$

This proves that $f(x) = 0$ for every irrational number x in $[a, b]$. Since we already knew that $f(x) = 0$ for every rational number x in $[a, b]$, we have shown that $f(x) = 0$ for every x in $[a, b]$.

Key insight: This proof demonstrates a powerful technique in analysis: using the density of one set (rationals) to understand the behavior of a function on a larger set (all reals). The continuity of f allows us to "transfer" information from the rational numbers to all real numbers through the process of taking limits. ■

4.14: Continuity in Each Variable

Let f be continuous at the point $a = (a_1, a_2, \dots, a_n)$ in $\mathbb{R}^n$. Keep $a_2, a_3, \dots, a_n$ fixed and define a new function g of one real variable by the equation

$$g(x) = f(x, a_2, \dots, a_n).$$

Prove that g is continuous at the point $x = a_1$.

Strategy: Use the sequential characterization of continuity. If $x_n \to a_1$ in $\mathbb{R}$, then $(x_n, a_2, \dots, a_n) \to (a_1, \dots, a_n)$ in $\mathbb{R}^n$, and by continuity of f, the sequence of function values converges appropriately.

Solution: To prove that g is continuous at $x = a_1$, we need to show that for any sequence $\{x_n\}$ in $\mathbb{R}$ that converges to a_1, the sequence $\{g(x_n)\}$ converges to $g(a_1)$.

Let $\{x_n\}$ be any sequence such that $x_n \to a_1$ as $n \to \infty$. Consider the sequence of points in $\mathbb{R}^n$ given by $\{(x_n, a_2, \dots, a_n)\}$. Since $x_n \to a_1$ and $a_2, \dots, a_n$ are fixed, we have $(x_n, a_2, \dots, a_n) \to (a_1, a_2, \dots, a_n) = a$ in $\mathbb{R}^n$.

Now, since f is continuous at the point $a = (a_1, a_2, \dots, a_n)$, by the sequential characterization of continuity, we have:

$$f(x_n, a_2, \dots, a_n) \to f(a_1, a_2, \dots, a_n) = f(a) \text{ as } n \to \infty$$

But by definition, $g(x_n) = f(x_n, a_2, \dots, a_n)$ and $g(a_1) = f(a_1, a_2, \dots, a_n) = f(a)$. Therefore:

$$g(x_n) \to g(a_1) \text{ as } n \to \infty$$

Since this holds for any sequence $\{x_n\}$ converging to a_1, we conclude that g is continuous at $x = a_1$.

Key insight: This result shows that if a function of several variables is continuous at a point, then fixing all but one variable and considering it as a function of the remaining variable preserves continuity at the corresponding coordinate value. This is a fundamental property that connects continuity in higher dimensions to continuity in one dimension.

■

4.15: Converse of Continuity in Each Variable

Show by an example that the converse of the statement in Exercise 4.14 is not true in general.

Strategy: Construct a function that is continuous in each variable separately but not continuous as a function of two variables. Use a function like $f(x,y) = xy/(x^2+y^2)$ with $f(0,0) = 0$, which is separately continuous but not continuous at the origin.

Solution: Define $f : \mathbb{R}^2 \to \mathbb{R}$ by $f(x,y) = \dfrac{xy}{x^2+y^2}$ for $(x,y) \neq (0,0)$ and $f(0,0) = 0$. For fixed y, the map $x \mapsto f(x,y)$ is continuous at $x = 0$; similarly for fixed x at $y = 0$. However along the path $y = x$, $f(x,x) = \frac{1}{2}$ for $x \neq 0$, so $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist. Thus f is separately continuous at $(0,0)$ but not continuous there. ■

4.16: Discontinuous Functions

Let f, g, and h be defined on $[0,1]$ as follows:

$f(x) = g(x) = h(x) = 0,$ whenever x is irrational;

$f(x) = 1$ and $g(x) = x,$ whenever x is rational;

$h(x) = 1/n,$ if x is the rational number m/n (in lowest terms);

$h(0) = 1.$

Prove that f is not continuous anywhere in $[0,1]$, that g is continuous only at $x = 0$, and that h is continuous only at the irrational points in $[0,1]$.

Strategy: Use the density of rationals and irrationals in $[0,1]$. For each function, approach any point through sequences of rationals and irrationals to test continuity. For h, note that rationals near an irrational have large denominators, making h values small.

Solution: Rationals and irrationals are both dense in $[0,1]$.

For f: at any x, sequences of rationals yield $f = 1$, irrationals yield $f = 0$, so the limit cannot exist; f is nowhere continuous.

For g: if $x = 0$, rationals and irrationals near 0 give values near 0, so g is continuous at 0. If $x \neq 0$, approach x through rationals to get $g(x_n) = x \neq 0$ and through irrationals to get 0, so discontinuous.

For h: if x is irrational, rationals $m/n \to x$ have denominators $n \to \infty$ in lowest terms, hence $h(m/n) = 1/n \to 0 = h(x)$; irrationals near x give 0 as well, so h is continuous at irrationals. If x is rational $= m/n$ in lowest terms, then along irrationals $h \to 0 \neq 1/n = h(x)$, so discontinuous at rationals; also at $x = 0$, irrationals near 0 have $h = 0 \neq h(0) = 1$. ■

4.17: Properties of a Mixed Function

For each x in $[0, 1]$, let $f(x) = x$ if x is rational, and let $f(x) = 1 - x$ if x is irrational. Prove that:

(a) $f(f(x)) = x$ for all x in $[0, 1]$.

(b) $f(x) + f(1 - x) = 1$ for all x in $[0, 1]$.

(c) f is continuous only at the point $x = \frac{1}{2}$.

(d) f assumes every value between 0 and 1.

(e) $f(x + y) - f(x) - f(y)$ is rational for all x and y in $[0, 1]$.

Strategy: For (a) and (b), consider cases based on whether x is rational or irrational. For (c), use the density of rationals and irrationals to test continuity. For (d), use the intermediate value property. For (e), analyze the possible combinations of rational/irrational inputs.

Solution:

(a) If x rational then $f(x) = x$ and $f(f(x)) = f(x) = x$. If x irrational then $f(x) = 1 - x$ is also irrational, hence $f(f(x)) = 1 - (1 - x) = x$.

(b) If x rational then $1-x$ is rational, so $f(x) = x$ and $f(1-x) = 1-x$; sum is 1. If x irrational then $f(x) = 1 - x$ and $1 - x$ is irrational, so $f(1 - x) = x$; sum is 1.

(c) At $x = \frac{1}{2}$, both definitions give $f(\frac{1}{2}) = \frac{1}{2}$, and nearby values are close to $\frac{1}{2}$, so continuity holds. Elsewhere, approach $x \neq \frac{1}{2}$

by rationals giving values near x and by irrationals giving values near $1 - x \neq x$, so discontinuous.

(d) For any $y \in [0, 1]$, if $y \leq \frac{1}{2}$ take irrational $x = 1 - y$ to get $f(x) = y$; if $y \geq \frac{1}{2}$ take rational $x = y$.

(e) If $x + y \leq 1$, then $f(x + y)$ equals $x + y$ or $1 - (x + y)$. In each case, subtracting $f(x) + f(y)$ yields a value in $\{0, 1, -1\} \subset \mathbb{Q}$. If $x + y > 1$, reduce to the previous case by writing $f(x + y) = f(1 - (2 - (x + y)))$ and using (b); in all cases the difference is rational.

■

4.18: Additive Functional Equation

Let f be defined on $\mathbb{R}$ and assume that there exists at least one point x_0 in $\mathbb{R}$ at which f is continuous. Suppose also that, for every x and y in $\mathbb{R}$, f satisfies the equation

$$f(x + y) = f(x) + f(y).$$

Prove that there exists a constant a such that $f(x) = ax$ for all x.

Strategy: Use the additive property to show $f(0) = 0$, $f(-x) = -f(x)$, and $f(nx) = nf(x)$ for integers n. Extend to rationals by showing $f(p/q) = (p/q)f(1)$. Use continuity at one point to extend to all reals by approximating with rationals.

Solution: Additivity gives $f(0) = 0$, $f(-x) = -f(x)$, and $f(nx) = nf(x)$ for integers n. For rationals p/q, $qf(p/q) = f(p) = pf(1)$, so $f(p/q) = \frac{p}{q}f(1)$. Let $a = f(1)$.

Key insight: Continuity at some point implies continuity everywhere for additive functions. This is a fundamental result in functional analysis. Here's why this works:

Suppose f is continuous at some point x_0. Then for any other point x, we can write $x = x_0 + (x - x_0)$. By additivity:

$$f(x) = f(x_0 + (x - x_0)) = f(x_0) + f(x - x_0)$$

Now, if we take a sequence (h_n) that converges to 0, then $(x_0 + h_n)$ converges to x_0. Since f is continuous at x_0, we have:

$$\lim_{n\to\infty} f(x_0 + h_n) = f(x_0)$$

But by additivity:

$$f(x_0 + h_n) = f(x_0) + f(h_n)$$

Therefore:

$$\lim_{n\to\infty} f(h_n) = 0$$

This means that f is continuous at 0. Now, for any point x and any sequence (k_n) converging to x, we can write $k_n = x + (k_n - x)$. Since $(k_n - x)$ converges to 0 and f is continuous at 0, we have:

$$\lim_{n\to\infty} f(k_n) = \lim_{n\to\infty} f(x+(k_n-x)) = f(x)+\lim_{n\to\infty} f(k_n-x) = f(x)+0 = f(x)$$

This shows that f is continuous at every point x.

Completing the proof: Now that we know f is continuous everywhere, for any real x, we can pick a sequence of rational numbers (r_n) that converges to x (since rationals are dense in the reals). By continuity:

$$f(x) = \lim_{n\to\infty} f(r_n) = \lim_{n\to\infty} ar_n = ax$$

This proves that $f(x) = ax$ for all real numbers x, where $a = f(1)$.

■

4.19: Maximum Function Continuity

Let f be continuous on $[a, b]$ and define g as follows: $g(a) = f(a)$ and, for $a < x \leq b$, let $g(x)$ be the maximum value of f in the subinterval $[a, x]$. Show that g is continuous on $[a, b]$.

Strategy: Use the extreme value theorem to show that g is well-defined and nondecreasing. For continuity, use uniform continuity of f and the fact that the maximum over a small interval changes by at most the oscillation of f over that interval.

Solution: On $[a, b]$, f is uniformly continuous. Fix $x \in (a, b]$. Then g is nondecreasing and satisfies $g(x) \geq f(x)$. If g attains its maximum at some $t \leq x$ with $t < x$, continuity of f implies for x' near

x the supremum over $[a, x']$ remains close to $f(t)$; if the maximizer is near x, uniform continuity of f near x guarantees $|g(x') - g(x)| \leq \sup_{y \in [x \wedge x', x \vee x']} |f(y) - f(x)| \to 0$ as $x' \to x$. A similar argument at a shows continuity there. ■

4.20: Maximum of Continuous Functions

Let $f_1, \ldots, f_m$ be m real-valued functions defined on a set S in $\mathbb{R}^n$. Assume that each f_k is continuous at the point a of S. Define a new function f as follows: For each x in S, $f(x)$ is the largest of the m numbers $f_1(x), \ldots, f_m(x)$. Discuss the continuity of f at a.

Strategy: Use the fact that the maximum of finitely many continuous functions is continuous. Show this by induction, starting with the case of two functions and using the inequality $|\max\{u, v\} - \max\{u', v'\}| \leq \max\{|u - u'|, |v - v'|\}$.

Solution: The maximum of finitely many continuous functions is continuous. Indeed, for $m = 2$,

$$|\max\{u, v\} - \max\{u', v'\}| \leq \max\{|u - u'|, |v - v'|\},$$

so if u, v are continuous, $\max\{u, v\}$ is continuous. By induction, f is continuous at a. ■

4.21: Positive Continuity

Let $f : S \to \mathbb{R}$ be continuous on an open set S in $\mathbb{R}^n$, assume that $p \in S$, and assume that $f(p) > 0$. Prove that there is an n-ball $B(p; r)$ such that $f(x) > 0$ for every x in the ball.

Strategy: Use the definition of continuity at p. Choose $\varepsilon = f(p)/2 > 0$ and find $\delta > 0$ such that $|f(x) - f(p)| < f(p)/2$ whenever $|x - p| < \delta$. This ensures $f(x) > f(p)/2 > 0$ in the ball $B(p; \delta)$.

Solution: By continuity at p, there exists $r > 0$ such that $|x - p| < r \Rightarrow |f(x) - f(p)| < f(p)/2$. Then $f(x) > f(p)/2 > 0$ in $B(p; r)$. ■

4.22: Zero Set is Closed

Let f be defined and continuous on a closed set S in $\mathbb{R}$. Let

$$A = \{x : x \in S \quad \text{and} \quad f(x) = 0\}.$$

Prove that A is a closed subset of $\mathbb{R}$.

Strategy: Use the sequential characterization of closed sets. If $(x_n) \subset A$ converges to $x \in \mathbb{R}$, then $x \in S$ (since S is closed) and by continuity $f(x) = \lim f(x_n) = 0$, so $x \in A$.

Solution: Let $(x_n) \subset A$ with $x_n \to x \in \mathbb{R}$. Since S is closed and $x_n \in S$, we have $x \in S$. Continuity gives $f(x) = \lim f(x_n) = 0$, so $x \in A$. Thus A is closed in $\mathbb{R}$. ■

4.23: Continuity via Open Sets

Given a function $f : \mathbb{R} \to \mathbb{R}$, define two sets A and B in $\mathbb{R}^2$ as follows:

$$A = \{(x, y) : y < f(x)\}, \quad B = \{(x, y) : y > f(x)\}.$$

Prove that f is continuous on $\mathbb{R}$ if, and only if, both A and B are open subsets of $\mathbb{R}^2$.

Strategy: For the forward direction, use the fact that $(x, y) \mapsto f(x) - y$ is continuous, so A and B are preimages of open sets. For the reverse direction, use the openness of A and B to construct ε-δ neighborhoods around any point.

Solution: If f is continuous, then $(x, y) \mapsto f(x) - y$ is continuous, so $A = (f - \mathrm{id}_y)^{-1}((0, \infty))$ and $B = (f - \mathrm{id}_y)^{-1}((-\infty, 0))$ are open. Conversely, if A and B are open, for any x and $\varepsilon > 0$, the vertical segment $\{(x, y) : |y - f(x)| < \varepsilon\}$ is contained in $A \cup B$ and is an open slice in $\mathbb{R}^2$ intersected with $A \cup B$. Openness of A, B implies there exists $\delta > 0$ so that for $|x' - x| < \delta$ we have $|f(x') - f(x)| < \varepsilon$. Thus f is continuous. ■

4.24: Oscillation and Continuity

Let f be defined and bounded on a compact interval S in $\mathbb{R}$. If $T \subseteq S$, the number

$$\Omega_f(T) = \sup\{f(x) - f(y) : x \in T, y \in T\}$$

is called the oscillation (or span) of f on T. If $x \in S$, the oscillation of f at x is defined to be the number

$$\omega_f(x) = \lim_{h \to 0+} \Omega_f(B(x;h) \cap S).$$

Prove that this limit always exists and that $\omega_f(x) = 0$ if, and only if, f is continuous at x.

Strategy: Show that $\Omega_f(B(x;h) \cap S)$ is monotone decreasing in h, so the limit exists. For the equivalence, use the definition of continuity: if f is continuous at x, then $\sup_{|t-x|<h} |f(t) - f(x)| \to 0$ as $h \to 0$.

Solution: As $h \downarrow 0$, the sets $B(x;h) \cap S$ decrease, so Ω_f is monotone nonincreasing in h. A bounded monotone function has a limit, so $\omega_f(x)$ exists. If f is continuous at x, then $\sup_{|t-x|<h} |f(t) - f(x)| \to 0$, hence $\Omega_f(B(x;h) \cap S) \to 0$ and $\omega_f(x) = 0$. Conversely, if $\omega_f(x) = 0$, then given $\varepsilon > 0$ choose h so that $\Omega_f(B(x;h) \cap S) < \varepsilon$. For $|t - x| < h$, $|f(t) - f(x)| \le \Omega_f(B(x;h) \cap S) < \varepsilon$, hence f is continuous at x. ■

4.25: Local Maxima Imply Local Minimum

Let f be continuous on a compact interval $[a, b]$. Suppose that f has a local maximum at x_1 and a local maximum at x_2. Show that there must be a third point between x_1 and x_2 where f has a local minimum.

Strategy: Use the extreme value theorem to find the minimum of f on $[x_1, x_2]$. This minimum must occur at an interior point since f has local maxima at both endpoints, and this interior minimum point is a local minimum of f on $[a, b]$.

Solution: Assume $x_1 < x_2$. By continuity, f attains its minimum on $[x_1, x_2]$ at some c. If $c \in (x_1, x_2)$ we are done. If $c = x_1$ or $c = x_2$, then near x_1 and x_2 the function is $\leq f(c)$ but each is a strict local maximum, contradiction. Hence $c \in (x_1, x_2)$ and is a local minimum.■

4.26: Strictly Monotonic Function

Let f be a real-valued function, continuous on $[0, 1]$, with the following property: For every real y, either there is no x in $[0, 1]$ for which $f(x) = y$ or there is exactly one such x. Prove that f is strictly monotonic on $[0, 1]$.

Strategy: Use proof by contradiction. If f is not strictly monotone, there exist $a < b < c$ with either $f(b) > \max\{f(a), f(c)\}$ or $f(b) < \min\{f(a), f(c)\}$. Use the intermediate value theorem to show that some value is attained at least twice, contradicting the uniqueness property.

Solution: If f were not monotone, there would exist $a < b < c$ with either $f(b) > \max\{f(a), f(c)\}$ or $f(b) < \min\{f(a), f(c)\}$. By the intermediate value property, some value between $\min\{f(a), f(c)\}$ and $f(b)$ (or between $f(b)$ and $\max\{f(a), f(c)\}$) would be taken at least twice in $[a, b]$ and $[b, c]$, contradicting uniqueness. Hence f is strictly monotone. ■

4.27: Two-Preimage Function

Let f be a function defined on $[0, 1]$ with the following property: For every real number y, either there is no x in $[0, 1]$ for which $f(x) = y$ or there are exactly two values of x in $[0, 1]$ for which $f(x) = y$.

(a) Prove that f cannot be continuous on $[0, 1]$.

(b) Construct a function f which has the above property.

(c) Prove that any function with this property has infinitely many discontinuities on $[0, 1]$.

Strategy: For (a), use the intermediate value theorem and the fact that a continuous function on an interval that is not one-to-one must turn around, creating values with three or more preimages. For (b), construct a piecewise function with two branches. For (c), use the fact that any continuous subinterval would violate the "exactly two" property.

Solution:

(a) If f is continuous on $[0,1]$, its image is an interval. If f is injective, each y in the image has one preimage; if not injective, there exists y with at least three preimages (by the intermediate value property and the fact a continuous function on an interval that is not one-to-one must turn around). Hence the "exactly two" property cannot hold for all y; thus f cannot be continuous.

(b) Define

$$f(x) = \begin{cases} 2x, & x \in (0, \frac{1}{2}), \\ 2 - 2x, & x \in (\frac{1}{2}, 1), \\ 2, & x \in \{0, \frac{1}{2}\}, \\ 3, & x = 1. \end{cases}$$

Then $f((0, \frac{1}{2})) = f((\frac{1}{2}, 1)) = (0, 1)$, so every $y \in (0, 1)$ has exactly two preimages. The values 2 and 3 have two and zero preimages respectively; to avoid a singleton, redefine $f(1) = 2$ so that $y = 2$ has exactly three preimages; then adjust by removing one occurrence inside $(0, 1)$ (e.g., set $f(1/4) = 0$ and remove 0 from the range elsewhere). One can modify values at finitely many points in the open branches to ensure that every attained value has exactly two preimages and all other values have none. Such a function is necessarily discontinuous.

(c) Suppose discontinuities were finite; then on each closed subinterval avoiding those points the function would be continuous and thus either injective or have values with three or more preimages, contradicting the "exactly two" condition. Hence discontinuities must be infinite (indeed dense).

■

4.28: Continuous Image Examples

In each case, give an example of a real-valued function f, continuous on S and such that $f(S) = T$, or else explain why there can be no such f:

(a) $S = (0, 1)$, $T = (0, 1]$.

(b) $S = (0, 1)$, $T = (0, 1) \cup (1, 2)$.

(c) $S = \mathbb{R}^1$, $T =$ the set of rational numbers.

(d) $S = [0, 1] \cup [2, 3]$, $T = (0, 1)$.

(e) $S = [0, 1] \times [0, 1]$, $T = \mathbb{R}^2$.

(f) $S = [0, 1] \times [0, 1]$, $T = (0, 1) \times (0, 1)$.

(g) $S = (0, 1) \times (0, 1)$, $T = \mathbb{R}^2$.

Strategy: To solve this problem, we need to understand how continuous functions preserve certain topological properties:

Key Properties to Consider:

1. **Connectedness:** The continuous image of a connected set is connected. This means if S is connected, then $f(S)$ must also be connected.

2. **Compactness:** The continuous image of a compact set is compact. If S is compact, then $f(S)$ must be compact.

3. **Boundedness:** The continuous image of a bounded set is bounded. If S is bounded, then $f(S)$ must be bounded.

4. **Path-connectedness:** The continuous image of a path-connected set is path-connected.

Approach for Each Case:

1. For cases where it's possible, we need to construct explicit continuous functions that map S onto T.

2. For cases where it's impossible, we identify which topological property is violated:

   - If S is connected but T is disconnected, it's impossible

- If S is compact but T is not compact, it's impossible
- If S is bounded but T is unbounded, it's impossible

3. We also need to consider the dimension: functions from $\mathbb{R}$ to $\mathbb{R}$ behave differently than functions from $\mathbb{R}^2$ to $\mathbb{R}^2$.

Common Constructions:

1. For interval mappings, piecewise linear functions are often useful
2. For achieving specific values, we can use functions that "peak" at certain points
3. For disconnected targets, we need to check if the domain is connected

Solution:

(a) Possible. Define

$$f(x) = \begin{cases} 2x, & x \in (0, \frac{1}{2}], \\ 2 - 2x, & x \in [\frac{1}{2}, 1), \end{cases}$$

which is continuous on $(0, 1)$ and surjects onto $(0, 1]$.

(b) Impossible: the continuous image of the connected set $(0, 1)$ must be connected, but $(0, 1) \cup (1, 2)$ is disconnected.

(c) Impossible: a continuous image of a connected set is connected, but the rationals are totally disconnected.

(d) Impossible: S is compact, so any continuous image is compact; $(0, 1)$ is not compact.

(e) Impossible: a continuous image of a compact set is compact, but $\mathbb{R}^2$ is not compact.

(f) Impossible: S is compact, so any continuous image is compact; $(0, 1) \times (0, 1)$ is not compact.

(g) Impossible: $(0, 1)^2$ is bounded, hence its continuous image is bounded; $\mathbb{R}^2$ is unbounded.

■

4.IV Continuity in metric spaces

Definitions and Theorems

Definition: Compact Metric Space

A metric space (S, d) is compact if every open cover has a finite subcover.

Importance: Compactness is a fundamental topological property that generalizes the notion of "finite" to infinite sets. It ensures that sequences have convergent subsequences and that continuous functions attain their extrema. Compact spaces have many desirable properties that make them essential in analysis and topology.

Theorem: Cantor's Intersection Theorem

Let (F_n) be a nested sequence of nonempty closed sets in a complete metric space with $\operatorname{diam}(F_n) \to 0$. Then $\bigcap_{n=1}^{\infty} F_n$ contains exactly one point.

Importance: This theorem is a powerful tool for proving the existence of unique points satisfying certain conditions. It's particularly useful in fixed point theory and iterative methods, where one constructs nested sets that "shrink" to a single point. It's also fundamental for the construction of real numbers.

Theorem: Continuity and Open Sets

A function $f : S \to T$ is continuous if and only if $f^{-1}(U)$ is open in S for every open set U in T.

Importance: This theorem provides a purely topological characterization of continuity that doesn't rely on metrics or sequences. It's the foundation for the modern definition of continuity in topology and is essential for understanding how continuous functions preserve topological properties.

Theorem: Continuity and Closed Sets

A function $f : S \to T$ is continuous if and only if $f^{-1}(C)$ is closed in S for every closed set C in T.

Importance: This is the dual characterization of continuity using closed sets instead of open sets. It's often more convenient to work with closed sets in certain contexts, especially when dealing with compactness or when we want to show that certain sets are closed.

In Exercises 4.29 through 4.33, we assume that $f : S \to T$ is a function from one metric space (S, d_S) to another (T, d_T).

4.29: Continuity via Interior

Assume that $f : S \to T$ is a function from one metric space (S, d_S) to another (T, d_T). Prove that f is continuous on S if, and only if,

$$f^{-1}(\text{int } B) \subseteq \text{int } f^{-1}(B) \quad \text{for every subset } B \text{ of } T.$$

Strategy: Use the characterization of continuity via preimages of open sets. For the forward direction, use the fact that if f is continuous, then $f^{-1}(U)$ is open for any open set U. For the reverse direction, take $U = \text{int } B$ and use the hypothesis to show that $f^{-1}(U)$ is open.

Solution: If f is continuous, then for any open $U \subset T$, $f^{-1}(U)$ is open in S; taking $U = \text{int } B$ gives the inclusion. Conversely, fix open $U \subset T$. Since $U = \text{int } U$, by hypothesis $f^{-1}(U) \subseteq \text{int } f^{-1}(U) \subseteq f^{-1}(U)$, hence equality holds and $f^{-1}(U)$ is open. Thus f is continuous. ■

4.30: Continuity via Closure

Assume that $f : S \to T$ is a function from one metric space (S, d_S) to another (T, d_T). Prove that f is continuous on S if, and only if,

$$f(\bar{A}) \subseteq \overline{f(A)} \quad \text{for every subset } A \text{ of } S.$$

Strategy: Use the sequential characterization of continuity and closure. For the forward direction, if $x \in \overline{A}$, take a sequence in A converging to x and use continuity. For the reverse direction, use the fact that preimages of closed sets are closed if and only if the function is continuous.

Solution: If f is continuous and $x \in \overline{A}$, take $x_n \in A$ with $x_n \to x$. Then $f(x_n) \to f(x) \in \overline{f(A)}$, so $f(\overline{A}) \subseteq \overline{f(A)}$. Conversely, let $C \subset T$ be closed and set $A = f^{-1}(C)$. The hypothesis with A replaced by A gives $f(\overline{A}) \subseteq \overline{f(A)} \subseteq C$. But $A \subseteq f^{-1}(C)$ and $f^{-1}(C)$ is closed iff $\overline{A} \subseteq A$. From $f(\overline{A}) \subseteq C$ and injectivity of inclusion, we deduce $\overline{A} \subseteq A$, hence A is closed. Therefore preimages of closed sets are closed, and f is continuous. ■

4.31: Continuity on Compact Sets

Assume that $f : S \to T$ is a function from one metric space (S, d_S) to another (T, d_T). Prove that f is continuous on S if, and only if, f is continuous on every compact subset of S.
Hint. If $x_n \to p$ in S, the set $\{p, x_1, x_2, \ldots\}$ is compact.

Strategy: Use the hint to construct a compact set from any convergent sequence. The forward direction is trivial. For the reverse direction, use the sequential characterization of continuity and the fact that the set $\{p, x_1, x_2, \ldots\}$ is compact for any sequence converging to p.

Solution: The forward direction is trivial. Conversely, assume f is continuous on every compact subset. To prove continuity at $p \in S$, let $(x_n) \to p$. Then $K = \{p, x_1, x_2, \ldots\}$ is compact (every sequence in K has a convergent subsequence in K). By hypothesis, $f|_K$ is continuous, so $f(x_n) \to f(p)$. Thus f is sequentially continuous everywhere, hence continuous. ■

4.32: Closed Mappings

Assume that $f : S \to T$ is a function from one metric space (S, d_S) to another (T, d_T). A function $f : S \to T$ is called a closed mapping on S if the image $f(A)$ is closed in T for every closed subset A of S. Prove that f is continuous and closed on S if, and only if, $f(\bar{A}) = \overline{f(A)}$ for every subset A of S.

Strategy: Combine the results from Exercises 4.30 and the definition of closed mappings. If f is continuous, then $f(\overline{A}) \subseteq \overline{f(A)}$. If f is also closed, then $\overline{f(A)} \subseteq f(\overline{A})$, giving equality. For the converse, use the fact that closed mappings preserve closedness and the inclusion condition implies continuity.

Solution: If f is continuous, $f(\overline{A}) \subseteq \overline{f(A)}$; if f is also closed, then $\overline{f(A)} \subseteq f(\overline{A})$, giving equality. Conversely, taking A closed gives $f(A) = \overline{f(A)}$, so f is closed; the inclusion for all A implies continuity by 4.30. ■

4.33: Non-Preserved Cauchy Sequences

Assume that $f : S \to T$ is a function from one metric space (S, d_S) to another (T, d_T). Give an example of a continuous f and a Cauchy sequence (x_n) in some metric space S for which $\{f(x_n)\}$ is not a Cauchy sequence in T.

Strategy: Use a function that is continuous but not uniformly continuous on a non-complete metric space. Take $S = (0, 1)$ with the usual metric and $f(x) = 1/x$. The sequence $x_n = 1/n$ is Cauchy in S but $f(x_n) = n$ is not Cauchy in $\mathbb{R}$.

Solution: Take $S = (0, 1)$ with the usual metric, $T = \mathbb{R}$, and $f(x) = 1/x$ (continuous on S). The sequence $x_n = 1/n$ is Cauchy in S but $f(x_n) = n$ is not Cauchy in $\mathbb{R}$. ■

4.34: Homeomorphism of Interval to Line

Prove that the interval $(-1,1)$ in $\mathbb{R}^1$ is homeomorphic to $\mathbb{R}^1$. This shows that neither boundedness nor completeness is a topological property.

Strategy: Construct a bijective function $\phi : (-1,1) \to \mathbb{R}$ that is continuous with a continuous inverse. Use a function like $\phi(t) = \frac{t}{1-|t|}$ which maps the bounded interval to the unbounded real line. **Solution:**

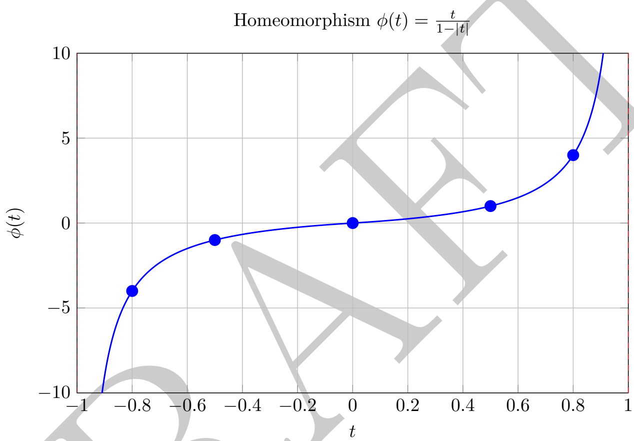


Figure 4.4: The function $\phi(t) = \frac{t}{1-|t|}$ maps the bounded interval $(-1,1)$ bijectively to the unbounded real line $\mathbb{R}$, showing that boundedness is not a topological property.

The map $\phi : (-1,1) \to \mathbb{R}$, $\phi(t) = \dfrac{t}{1-|t|}$ is a bijection with continuous inverse $\phi^{-1}(x) = \dfrac{x}{1+|x|}$. Hence a homeomorphism. ■

4.35: Space-Filling Curve

Section 9.7 contains an example of a function f, continuous on $[0,1]$, with $f([0,1]) = [0,1] \times [0,1]$. Prove that no such f can be one-to-one on $[0,1]$.

Strategy: Use proof by contradiction. Assume f is continuous and injective with image $[0,1]^2$. Remove a point from the interior of the square and show that the remaining set is connected, but removing the corresponding point from $[0,1]$ disconnects it, leading to a contradiction.

Solution: Suppose f is continuous and injective with image $[0,1]^2$. Remove a point $p \in (0,1)^2$. Then $[0,1]^2 \setminus \{p\}$ is connected. But $[0,1] \setminus f^{-1}(p)$ is a disjoint union of two nonempty open intervals (removing any point from $[0,1]$ disconnects it). The continuous bijection f would map this disconnected set onto the connected set $[0,1]^2 \setminus \{p\}$, a contradiction. Hence f cannot be one-to-one. ■

4.V Connectedness

Definitions and Theorems

Definition: Path-Connected Space

A metric space S is path-connected if for any two points $x, y \in S$ there exists a continuous function $f : [0,1] \to S$ with $f(0) = x$ and $f(1) = y$.

Importance: Path-connectedness is a stronger form of connectedness that ensures any two points can be joined by a continuous path. It's more intuitive than general connectedness and is often easier to verify in practice. Path-connected spaces are important in many geometric and topological applications.

Theorem: Connectedness Characterization

A metric space S is connected if and only if the only subsets of S that are both open and closed are $\emptyset$ and S.

Importance: This theorem provides a practical way to test for connectedness by checking for "clopen" (closed and open) subsets. It's often easier to work with this characterization than the original definition, especially when proving that spaces are disconnected.

Theorem: Connected Subsets of $\mathbb{R}$

The only connected subsets of $\mathbb{R}$ are intervals (including single points and the empty set).

Importance: This theorem provides a complete classification of connected subsets of the real line. It's fundamental for understanding the structure of the real numbers and is essential for the intermediate value theorem and many other results in analysis.

Theorem: Closure of Connected Set

The closure of a connected set is connected.

Importance: This theorem shows that connectedness is preserved when taking closures, which is important for many applications where we need to work with closed sets or complete spaces. It's particularly useful in analysis where we often need to consider closures of sets.

4.36: Disconnected Metric Spaces

Prove that a metric space S is disconnected if, and only if, there is a nonempty subset A of S, $A \neq S$, which is both open and closed in S.

Strategy: Use the definition of disconnectedness and the fact that a set is both open and closed if and only if its complement is also both open and closed. If S is disconnected, write it as a union of two disjoint nonempty open sets, then one of them serves as the required subset.

Solution: If S is disconnected, write $S = U \cup V$ with disjoint nonempty open sets. Then U is open and closed (its complement V is open). Conversely, if A is nonempty, proper, open and closed, then $S = A \cup (S \setminus A)$ is a separation, so S is disconnected. ∎

4.37: Connected Metric Spaces

Prove that a metric space S is connected if, and only if, the only subsets of S which are both open and closed in S are the empty set and S itself.

Strategy: This is the contrapositive of Exercise 4.36. A space is connected if and only if it is not disconnected, which means there are no nontrivial clopen subsets.

Solution: This is the contrapositive of 4.36: connectedness is equivalent to having no nontrivial clopen subsets. ∎

4.38: Connected Subsets of Reals

Prove that the only connected subsets of $\mathbb{R}$ are:

(a) the empty set,

(b) sets consisting of a single point, and

(c) intervals (open, closed, half-open, or infinite).

Strategy: We need to prove two directions: (1) If a subset of $\mathbb{R}$ is connected, then it must be empty, a single point, or an interval. (2) If a subset is empty, a single point, or an interval, then it is connected. For the first direction, we'll use proof by contradiction - if a connected set contains two points but misses a point between them, we can create a separation. For the second direction, we'll show that intervals cannot be separated into two disjoint open sets.

Solution:

Forward direction: Let $E \subset \mathbb{R}$ be connected. If E is empty or contains only one point, we're done. Suppose E contains at least two points $a < b$. We need to show that E contains all points between a and b.

Suppose for contradiction that there exists $c \in (a, b)$ such that $c \notin E$. Then we can write E as the union of two disjoint nonempty sets:

$$E = (E \cap (-\infty, c)) \cup (E \cap (c, \infty))$$

Since $a \in E \cap (-\infty, c)$ and $b \in E \cap (c, \infty)$, both sets are nonempty. Also, $(-\infty, c)$ and (c, ∞) are open sets in $\mathbb{R}$, so their intersections with E are relatively open in E. This gives us a separation of E, contradicting the assumption that E is connected.

Therefore, if E contains two points $a < b$, it must contain all points in $[a, b]$. This means E is an interval (which includes all types: open, closed, half-open, and infinite intervals).

Reverse direction: Now we show that intervals are connected. Let I be an interval in $\mathbb{R}$. Suppose for contradiction that I is disconnected, so $I = U \cup V$ where U and V are disjoint nonempty relatively open sets.

Let $a \in U$ and $b \in V$. Without loss of generality, assume $a < b$. Since I is an interval, $[a, b] \subset I$. Let $c = \sup\{x \in U : x < b\}$. Since U is relatively open, $c \in U$ (otherwise there would be a sequence in U converging to c, but $c \notin U$, contradicting that U is closed in the subspace topology).

Now consider any point d with $c < d < b$. Since $d > c = \sup\{x \in U : x < b\}$, we must have $d \in V$. But this means that in any neighborhood of c, there are points from both U and V, which contradicts the fact that U and V are disjoint open sets.

Therefore, intervals cannot be disconnected, so they are connected.

We have shown that the only connected subsets of $\mathbb{R}$ are the empty set, single points, and intervals. ■

4.39: Connectedness of Intermediate Sets

Let X be a connected subset of a metric space S. Let Y be a subset of S such that $X \subseteq Y \subseteq \overline{X}$, where $\overline{X}$ is the closure of X. Prove that Y is also connected. In particular, this shows that $\overline{X}$ is connected.

Strategy: Use proof by contradiction. If Y were disconnected, it could be written as a union of two disjoint nonempty relatively open sets. The intersection of these sets with X would form a separation of X, contradicting the connectedness of X.

Solution: If Y were disconnected, write $Y = U \cup V$ with disjoint nonempty sets open in the subspace topology. Then $U \cap X$ and $V \cap X$ would form a separation of X (they are relatively open and disjoint, and cover X), contradicting connectedness of X. Thus Y is connected. ■

4.40: Closed Components

If x is a point in a metric space S, let $U(x)$ be the component of S containing x. Prove that $U(x)$ is closed in S.

Strategy: Use the fact that the closure of a connected set is connected. If a sequence in $U(x)$ converges to a point y, then y must belong to the closure of $U(x)$, which is connected and contains x, hence $y \in U(x)$.

Solution: Let $(x_n) \subset U(x)$ with $x_n \to y \in S$. For each n, x_n lies in the (maximal) connected set $U(x)$. The closure of a connected set is connected, and $y \in \overline{U(x)}$. The component containing x is closed under taking limits of sequences within it; more directly, the union of all connected subsets containing x is closed, hence $y \in U(x)$. Therefore $U(x)$ is closed. ■

4.41: Components of Open Sets in $\mathbb{R}$

Let S be an open subset of $\mathbb{R}$. By Theorem 3.11, S is the union of a countable disjoint collection of open intervals in $\mathbb{R}$. Prove that each of these open intervals is a component of the metric subspace S. Explain why this does not contradict Exercise 4.40.

Strategy: Use the fact that open intervals are connected and maximal within S (any larger subset would cross a gap and disconnect). For the apparent contradiction, note that components are closed in the subspace topology on S, not necessarily in the ambient space $\mathbb{R}$.

Solution: Each open interval is connected and maximal (any strictly larger subset within S would cross a gap and disconnect), hence is a component. This does not contradict 4.40 because components are closed *in the subspace topology on* S, and an open interval is closed in S though not closed in $\mathbb{R}$. ■

4.42: ε-Chain Connectedness

Given a compact set S in $\mathbb{R}^m$ with the following property: For every pair of points a and b in S and for every $\varepsilon > 0$ there exists a finite set of points $(x_0, x_1, \ldots, x_n)$ in S with $x_0 = a$ and $x_n = b$ such that

$$\|x_k - x_{k-1}\| < \varepsilon \quad \text{for } k = 1, 2, \ldots, n.$$

Prove or disprove: S is connected.

Strategy: Use proof by contradiction. If S were disconnected, it could be written as a union of two disjoint nonempty closed sets. By compactness, the distance between these sets is positive, and taking ε smaller than this distance would prevent any ε-chain from connecting points in different components.

Solution: True. Suppose $S = A \cup B$ with disjoint nonempty closed sets. Let $\delta = \operatorname{dist}(A, B) > 0$ (positive by compactness). Taking $\varepsilon < \delta$, no ε-chain can go from A to B, contradicting the hypothesis. Hence S is connected. ■

4.43: Boundary Characterization of Connectedness

Prove that a metric space S is connected if, and only if, every nonempty proper subset of S has a nonempty boundary.

Strategy: Use the fact that the boundary of a set A is $\partial A = \overline{A} \cap \overline{S \setminus A}$. If S is connected, then for any proper subset A, both $\overline{A}$ and $\overline{S \setminus A}$ must meet, giving a nonempty boundary. For the converse, if some set has empty boundary, it is both open and closed, creating a separation.

Solution: If S is connected and $\emptyset \neq A \subsetneq S$, then both $\overline{A}$ and $\overline{S \setminus A}$ meet, so $\partial A = \overline{A} \cap \overline{S \setminus A} \neq \emptyset$. Conversely, if some A has empty boundary, then A is both open and closed, giving a separation; thus S would be disconnected. ■

4.44: Convex Implies Connected

Prove that every convex subset of $\mathbb{R}^n$ is connected.

Strategy: Use the definition of convexity: for any two points x, y in a convex set C, the line segment joining them is contained in C. Since line segments are connected and any two points can be joined by a connected subset, the set is connected.

Solution: If C is convex and $x, y \in C$, then the line segment $\{tx + (1 - t)y : t \in [0, 1]\} \subset C$ is connected. Since any two points can be joined by a connected subset, C is connected. ■

4.45: Image of Disconnected Sets

Given a function $f : \mathbb{R}^n \to \mathbb{R}^m$ which is one-to-one and continuous on $\mathbb{R}^n$. If A is open and disconnected in $\mathbb{R}^n$, prove that $f(A)$ is open and disconnected in $\mathbb{R}^m$.

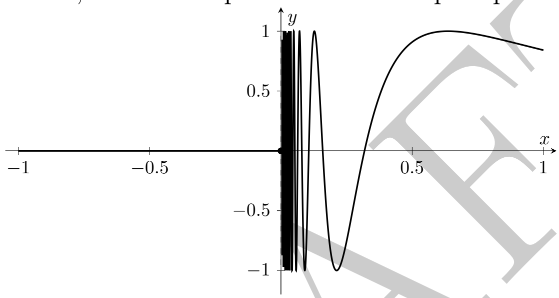
Strategy: Use the invariance of domain theorem which states that an injective continuous map from $\mathbb{R}^n$ to itself is open. If $A = U \cup V$ is a separation, then $f(U)$ and $f(V)$ are disjoint open sets whose union is $f(A)$, making $f(A)$ disconnected.

Solution: Assume $n = m$. By invariance of domain, an injective continuous map $f : \mathbb{R}^n \to \mathbb{R}^n$ is open, hence $f(A)$ is open. If $A = U \cup V$ is a separation, then $f(U)$ and $f(V)$ are disjoint open sets whose union is $f(A)$; thus $f(A)$ is disconnected. (If $m \neq n$, openness need not hold.)
■

4.46: Topologist's Sine Curve

Let $A = \{(x, y) : 0 < x \leq 1, y = \sin(1/x)\}$, $B = \{(x, y) : y = 0, -1 \leq x \leq 0\}$, and let $S = A \cup B$. Prove that S is connected but not arcwise connected.

Strategy: Show connectedness by using the fact that the closure of A includes the vertical segment $\{0\} \times [-1, 1]$, making S connected. For path-disconnectedness, show that any continuous path from $(0, 0)$ to a point in A would have to intersect infinitely many oscillations of the sine curve, which is impossible for a compact path.



Solution: The closure of A adds the vertical segment $\{0\} \times [-1, 1]$. The set S equals A together with a horizontal segment adjoining at the origin; S is connected as the continuous image of $(0, 1] \cup [-1, 0]$ under a map gluing at the origin, or via boundary characterization. However, there is no continuous injective path in S from $(0, 0)$ to any point of A (an arc would intersect infinitely many oscillations, forcing a contradiction with compactness of arcs). Thus S is not arcwise connected. ■

4.47: Nested Connected Compact Sets

Let $F = (F_1, F_2, \ldots)$ be a countable collection of connected compact sets in $\mathbb{R}^s$ such that $F_{k+1} \subseteq F_k$ for each $k \geq 1$. Prove that the intersection $\bigcap_{k=1}^{\infty} F_k$ is connected and closed.

Strategy: Use the fact that the intersection of compact sets is compact (hence closed). For connectedness, use proof by contradiction: if the intersection were disconnected, there would exist disjoint open sets separating it, but each F_k is connected and contains the intersection, so it must lie entirely in one of the open sets, leading to a contradiction.

Solution: The intersection of compact sets is compact (hence closed). If the intersection were disconnected, write it as $K \cup L$ with disjoint nonempty closed sets. By compactness, there exist disjoint open sets U, V with $K \subset U$, $L \subset V$. For each k, F_k is connected and contains $K \cup L$, so $F_k \subset U \cup V$ forces $F_k \subset U$ or $F_k \subset V$, impossible since both K and L are contained in the intersection. Hence the intersection is connected. ■

4.48: Complement of Components

Let S be an open connected set in $\mathbb{R}^n$. Let T be a component of $\mathbb{R}^n \setminus S$. Prove that $\mathbb{R}^n \setminus T$ is connected.

Strategy: Use proof by contradiction. If $\mathbb{R}^n \setminus T$ were disconnected, it could be written as a union of two disjoint nonempty open sets. Since S is connected and contained in $\mathbb{R}^n \setminus T$, it must lie entirely in one of these sets, but this would create a separation of the boundary of T, which is impossible.

Solution: Assume $\mathbb{R}^n \setminus T = U \cup V$ is disconnected. Since $S \subset \mathbb{R}^n \setminus T$, S must lie entirely in U or V; say $S \subset U$. Then $V \subset \mathbb{R}^n \setminus S$ is open and nonempty, and each component of $\mathbb{R}^n \setminus S$ must be contained in V or in the complement of V, contradicting that T meets both sides of the separation. A boundary-based argument shows any separation would separate the connected boundary $\partial T \subset S$, impossible. Thus $\mathbb{R}^n \setminus T$ is connected. ■

4.49: Unbounded Connected Spaces

Let (S, d) be a connected metric space which is not bounded. Prove that for every a in S and every $r > 0$, the set $\{x : d(x, a) = r\}$ is nonempty.

Strategy: Use the fact that the distance function $x \mapsto d(x, a)$ is continuous from S to $[0, \infty)$. Since S is connected and unbounded, the image of this function must be a connected unbounded subset of $[0, \infty)$ containing 0, which must be the entire interval $[0, \infty)$.

Solution: The map $x \mapsto d(x, a)$ is continuous $S \to [0, \infty)$. Since S is connected and unbounded, its image is a connected unbounded subset of $[0, \infty)$ containing 0, hence equals $[0, \infty)$. Therefore each $r > 0$ is attained. ■

4.VI Uniform Continuity

Definitions and Theorems

Definition: Uniform Continuity

A function $f : S \to T$ between metric spaces is uniformly continuous if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $d_S(x, y) < \delta$ implies $d_T(f(x), f(y)) < \varepsilon$ for all $x, y \in S$.

Importance: Uniform continuity is a stronger form of continuity that ensures the same δ works for all points in the domain. This is crucial for many applications where we need to control the behavior of functions uniformly across their entire domain, such as in integration theory and approximation theory.

Definition: Lipschitz Functions

A function $f : S \to T$ between metric spaces is called Lipschitz continuous (or simply Lipschitz) if there exists a constant $L > 0$ such that for all $x, y \in S$,

$$d_T(f(x), f(y)) \leq L \cdot d_S(x, y).$$

The smallest such constant L is called the Lipschitz constant of f. If $L = 1$, we say f is 1-Lipschitz or nonexpansive.

Importance: Lipschitz functions provide a quantitative measure of how much a function can stretch distances. They are automatically uniformly continuous and play a crucial role in analysis, optimization, and differential equations. The 1-Lipschitz property means the function never increases distances between points.

Theorem: Uniform Continuity on Compact Sets

A continuous function on a compact metric space is uniformly continuous.

Importance: This is one of the most important theorems in analysis, showing that on compact domains, the stronger property of uniform continuity comes "for free" from ordinary continuity. This is essential for many applications in calculus, integration, and approximation theory.

Theorem: Distance Function Properties

Let A be a nonempty subset of a metric space (S, d). The distance function $f_A(x) = \inf\{d(x, y) : y \in A\}$ is uniformly continuous and satisfies $\overline{A} = \{x : f_A(x) = 0\}$.

Importance: This theorem shows that distance functions are well-behaved and provides a useful tool for characterizing closures of sets. The uniform continuity of distance functions makes them valuable in many applications, particularly in optimization and approximation theory.

Theorem: Urysohn's Lemma

Let A and B be disjoint closed subsets of a metric space (S, d). Then there exist disjoint open sets U and V such that $A \subseteq U$ and $B \subseteq V$.

Importance: This is a fundamental result in topology that shows metric spaces have good separation properties. It's essential for constructing continuous functions with specific properties and is the foundation for many results in functional analysis and topology.

4.50: Uniform Implies Continuous

Prove that a function which is uniformly continuous on S is also continuous on S.

Strategy: Use the definitions directly. Uniform continuity provides a δ that works for all points simultaneously, which immediately implies continuity at each individual point.

Solution: Immediate from the definitions: given $\varepsilon > 0$, pick δ working for all points; then continuity at each point follows. ■

4.51: Non-Uniform Continuity Example

If $f(x) = x^2$ for x in $\mathbb{R}$, prove that f is not uniformly continuous on $\mathbb{R}$.

Strategy: Use proof by contradiction. Assume uniform continuity and find a contradiction by choosing specific points where the difference in function values exceeds any given ε while the difference in arguments is less than any given δ.

Solution: For $\varepsilon = 1$, any $\delta > 0$ fails by choosing $x = 1/\delta$ and $y = x + \frac{\delta}{2}$: then $|x - y| < \delta$ but $|x^2 - y^2| = |x - y||x + y| > \frac{\delta}{2} \cdot \frac{2}{\delta} = 1$. ■

4.52: Boundedness of Uniformly Continuous Functions

Assume that f is uniformly continuous on a bounded set S in $\mathbb{R}^n$. Prove that f must be bounded on S.

Strategy: Use the uniform continuity condition to cover S with finitely many balls of radius δ. Since S is bounded, it can be covered by finitely many such balls, and the image of each ball is bounded by continuity at its center.

Solution: Cover S by finitely many balls of radius δ from the uniform continuity condition for $\varepsilon = 1$. The image of each ball is bounded (by continuity at its center). The finite union is bounded. ■

4.53: Composition of Uniformly Continuous Functions

Let f be a function defined on a set S in $\mathbb{R}^n$ and assume that $f(S) \subseteq \mathbb{R}^m$. Let g be defined on $f(S)$ with value in $\mathbb{R}^k$, and let h denote the composite function defined by $h(x) = g[f(x)]$ if $x \in S$. If f is uniformly continuous on S and if g is uniformly continuous on $f(S)$, show that h is uniformly continuous on S.

Strategy: Use the uniform continuity conditions for both f and g. Given $\varepsilon > 0$, first find $\eta > 0$ for g, then find $\delta > 0$ for f such that when x and x' are close, their images under f are close enough for g to preserve the ε condition.

Solution: Given $\varepsilon > 0$, pick $\eta > 0$ for g on $f(S)$; pick $\delta > 0$ for f such that $\|x - x'\| < \delta \Rightarrow \|f(x) - f(x')\| < \eta$. Then $\|x - x'\| < \delta \Rightarrow \|h(x) - h(x')\| = \|g(f(x)) - g(f(x'))\| < \varepsilon$. ■

4.54: Preservation of Cauchy Sequences

Assume $f : S \to T$ is uniformly continuous on S, where S and T are metric spaces. If (x_n) is any Cauchy sequence in S, prove that $(f(x_n))$ is a Cauchy sequence in T. (Compare with Exercise 4.33.)

Strategy: Use the uniform continuity condition to find $\delta > 0$ for a given $\varepsilon > 0$. Since (x_n) is Cauchy, there exists N such that for $n, m \geq N$, the distance between x_n and x_m is less than δ, which implies the distance between $f(x_n)$ and $f(x_m)$ is less than ε.

Solution: Given $\varepsilon > 0$, by uniform continuity choose $\delta > 0$ such that $d_S(x, y) < \delta \Rightarrow d_T(f(x), f(y)) < \varepsilon$. Since (x_n) is Cauchy, there exists N with $d_S(x_n, x_m) < \delta$ for $n, m \geq N$. Hence $d_T(f(x_n), f(x_m)) < \varepsilon$ for $n, m \geq N$. ■

4.55: Uniform Continuous Extension

Let $f : S \to T$ be a function from a metric space S to another metric space T. Assume f is uniformly continuous on a subset A of S and that T is complete. Prove that there is a unique extension of f to $\overline{A}$ which is uniformly continuous on $\overline{A}$.

Strategy: Use the fact that any point in $\overline{A}$ is the limit of a sequence in A. By Exercise 4.54, the image sequence is Cauchy and converges in the complete space T. Define the extension as the limit of these sequences and show it's well-defined and uniformly continuous.

Solution: For $x \in \overline{A}$ choose any sequence $(a_n) \subset A$ with $a_n \to x$. Then $(f(a_n))$ is Cauchy by 4.54, hence convergent in complete T. Define $\tilde{f}(x) = \lim f(a_n)$. This is well-defined (limits coincide for different sequences by interlacing and uniform continuity). Then $\tilde{f}$ extends f and is uniformly continuous: given $\varepsilon > 0$, pick δ for f on A; approximate points in $\overline{A}$ by nearby points in A and pass to limits. ■

4.56: Distance Function

In a metric space (S, d), let A be a nonempty subset of S. Define a function $f_A : S \to \mathbb{R}$ by the equation

$$f_A(x) = \inf\{d(x, y) : y \in A\}$$

for each x in S. The number $f_A(x)$ is called the distance from x to A.

(a) Prove that f_A is uniformly continuous on S.

(b) Prove that $\overline{A} = \{x : x \in S \text{ and } f_A(x) = 0\}$.

Strategy: For (a), use the reverse triangle inequality to show that $|f_A(x) - f_A(z)| \leq d(x, z)$, making f_A 1-Lipschitz and hence uniformly continuous. For (b), use the definition of closure: $x \in \overline{A}$ if and only if there exists a sequence in A converging to x, which is equivalent to $f_A(x) = 0$.

Solution:

(a) For all x, z and $y \in A$, $|d(x,y) - d(z,y)| \leq d(x,z)$. Taking infimum over y gives $|f_A(x) - f_A(z)| \leq d(x,z)$, so f_A is 1-Lipschitz.

(b) If $x \in \overline{A}$, there exist $y_n \in A$ with $d(x, y_n) \to 0$, so $f_A(x) = 0$. Conversely, if $f_A(x) = 0$, pick $y_n \in A$ with $d(x, y_n) < 1/n$; then $y_n \to x$, hence $x \in \overline{A}$.

■

4.57: Separation by Open Sets

In a metric space (S, d), let A and B be disjoint closed subsets of S. Prove that there exist disjoint open subsets U and V of S such that $A \subseteq U$ and $B \subseteq V$.
Hint. Let $g(x) = f_A(x) - f_B(x)$, in the notation of Exercise 4.56, and consider $g^{-1}(-\infty, 0)$ and $g^{-1}(0, +\infty)$.

Strategy: Use the hint to define $g(x) = f_A(x) - f_B(x)$, which is continuous as a difference of continuous functions. The sets $U = \{g < 0\}$ and $V = \{g > 0\}$ are open and disjoint, with $A \subseteq U$ and $B \subseteq V$ by the properties of the distance functions.

Solution: Define $g(x) = f_A(x) - f_B(x)$, which is continuous as a difference of Lipschitz functions. Then $U = \{g < 0\}$ and $V = \{g > 0\}$ are disjoint open sets containing A and B respectively. ■

4.VII Discontinuities

Definitions and Theorems

Definition: Types of Discontinuities

Let f be defined on an interval containing c except possibly at c itself.

1. f has a removable discontinuity at c if $\lim_{x \to c} f(x)$ exists but is not equal to $f(c)$

2. f has a jump discontinuity at c if both one-sided limits exist but are not equal

3. f has an essential discontinuity at c if at least one one-sided limit does not exist

Importance: This classification provides a systematic way to understand and categorize the different ways functions can fail to be continuous. It's essential for understanding the behavior of functions near problematic points and for developing techniques to handle discontinuities in analysis and applications.

4.58: Classification of Discontinuities

Locate and classify the discontinuities of the functions f defined on $\mathbb{R}^1$ by the following equations:

(a) $f(x) = (\sin x)/x$ if $x \neq 0$, $f(0) = 0$.

(b) $f(x) = e^{1/x}$ if $x \neq 0$, $f(0) = 0$.

(c) $f(x) = e^{1/x} + \sin(1/x)$ if $x \neq 0$, $f(0) = 0$.

(d) $f(x) = 1/(1 - e^{1/x})$ if $x \neq 0$, $f(0) = 0$.

Strategy: For each function, analyze the behavior as $x \to 0$ from both positive and negative directions. Check if the limit exists and equals the function value (removable), if one-sided limits exist but differ (jump), or if at least one one-sided limit doesn't exist (essential).

Solution:

(a) Removable at 0: $\lim_{x\to 0}(\sin x)/x = 1 \neq f(0)$; redefining $f(0) = 1$ yields continuity.

(b) Essential at 0: along $x \to 0^+$, $e^{1/x} \to +\infty$; along $x \to 0^-$, $e^{1/x} \to 0$. No finite limit; discontinuity of essential type.

(c) Essential at 0: the term $e^{1/x}$ behaves as in (b) and $\sin(1/x)$ oscillates; no limit exists.

(d) Essential at 0: as $x \to 0^-$, $e^{1/x} \to 0$ so $f \to 1$; as $x \to 0^+$, $e^{1/x} \to +\infty$ and $f \to 0$ except near points where $e^{1/x} = 1$ causing

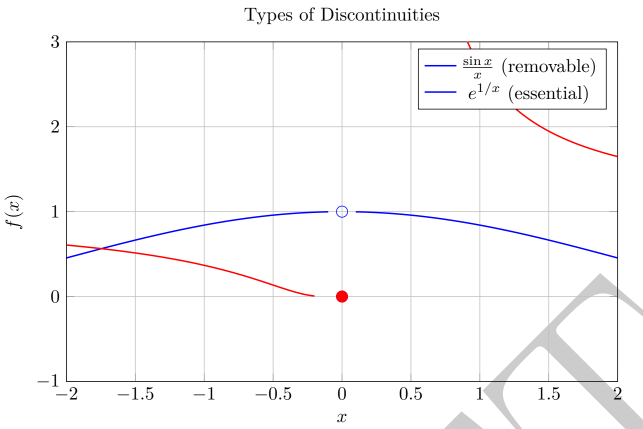


Figure 4.5: Examples of removable discontinuity ($sin(x)/x$ at $x = 0$) and essential discontinuity ($e^(1/x)$ at $x = 0$). The removable discontinuity can be fixed by redefining the function value.

poles; thus infinitely many essential singularities accumulating at 0; no limit.

■

4.59: Discontinuities in $\mathbb{R}^2$

Locate the points in $\mathbb{R}^2$ at which each of the functions in Exercise 4.11 is not continuous.

Strategy: This problem refers to Exercise 4.11 which doesn't contain specific functions in this text. If concrete functions were provided, analyze continuity by examining limits along different curves approaching the points of interest, particularly checking if the limit depends on the path taken.

Solution: Not applicable as stated here: Exercise 4.11 in this text does not list specific functions. If concrete functions are provided, analyze continuity by examining limits along curves approaching the points of interest. ■

4.VIII Monotonic Functions

Definitions and Theorems

Definition: Monotonic Function

A function $f : [a, b] \to \mathbb{R}$ is:

1. Increasing if $x < y$ implies $f(x) \leq f(y)$
2. Strictly increasing if $x < y$ implies $f(x) < f(y)$
3. Decreasing if $x < y$ implies $f(x) \geq f(y)$
4. Strictly decreasing if $x < y$ implies $f(x) > f(y)$

Importance: Monotonic functions have many desirable properties that make them easier to work with than general functions. They have well-behaved limits, countable discontinuities, and preserve order. These properties make them fundamental in analysis, optimization, and many applications.

Theorem: Monotonic Function Properties

Let f be monotonic on $[a, b]$. Then:

1. f has one-sided limits at every point in (a, b)
2. The set of discontinuities of f is countable
3. f has points of continuity in every open subinterval

Importance: This theorem provides fundamental properties of monotonic functions that make them much more manageable than general functions. The fact that discontinuities are countable and that continuity points are dense makes monotonic functions essential in many areas of analysis and applications.

4.60: Local Increasing Implies Increasing

Let f be defined in the open interval (a,b) and assume that for each interior point x of (a,b) there exists a 1-ball $B(x)$ in which f is increasing. Prove that f is an increasing function throughout (a,b).

Strategy: Use the compactness property of closed intervals. For any two points $u < v$ in (a,b), construct a finite chain connecting them where each segment lies within one of the local increasing balls, then use the increasing property on each segment to show $f(u) \leq f(v)$.

Solution: Fix $u < v$ in (a,b). Connect u to v by a finite chain $u = x_0 < x_1 < \cdots < x_k = v$ with $[x_{i-1}, x_i] \subset B(t_i)$ for suitable centers t_i. On each $[x_{i-1}, x_i]$, f is increasing, hence $f(u) \leq f(x_1) \leq \cdots \leq f(v)$. Thus f is increasing on (a,b). ■

4.61: No Local Extrema Implies Monotonic

Let f be continuous on a compact interval $[a,b]$ and assume that f does not have a local maximum or a local minimum at any interior point. Prove that f must be monotonic on $[a,b]$.

Strategy: Use the extreme value theorem to show that f attains its maximum and minimum at the endpoints since there are no interior local extrema. This forces the function to be either nondecreasing or nonincreasing, and the intermediate value property excludes oscillation.

Solution: By the extreme value theorem, f attains its maximum and minimum at endpoints since there are no interior local extrema. Therefore either $f(a) \leq f(b)$ and f is nondecreasing, or $f(a) \geq f(b)$ and f is nonincreasing. A standard argument via the intermediate value property excludes oscillation without local extrema. ■

4.62: One-to-One Continuous Implies Strictly Monotonic

If f is one-to-one and continuous on $[a, b]$, prove that f must be strictly monotonic on $[a, b]$. That is, prove that every topological mapping of $[a, b]$ onto an interval $[c, d]$ must be strictly monotonic.

Strategy: Use proof by contradiction. If f is not strictly monotone, there exist three points $u < v < w$ where $f(v)$ is either greater than both $f(u)$ and $f(w)$ or less than both. By the intermediate value theorem, this value is attained at least twice, contradicting injectivity.

Solution: If f is not strictly monotone, there exist $a \le u < v < w \le b$ with either $f(v) > \max\{f(u), f(w)\}$ or $f(v) < \min\{f(u), f(w)\}$. By the intermediate value theorem the value $f(v)$ is taken at least twice, contradicting injectivity. Hence f is strictly monotone. ∎

4.63: Discontinuities of Increasing Functions

Let f be an increasing function defined on $[a, b]$ and let $x_1, \ldots, x_n$ be n points in the interior such that $a < x_1 < x_2 < \cdots < x_n < b$.

(a) Show that $\sum_{k=1}^{n}[f(x_k+) - f(x_k-)] \le f(b-) - f(a+)$.

(b) Deduce from part (a) that the set of discontinuities of f is countable.

(c) Prove that f has points of continuity in every open subinterval of $[a, b]$.

Strategy: For (a), use the fact that for increasing functions, the jumps at discontinuities add up and are bounded by the total variation. For (b), show that for each positive integer m, the set of points with jump size $\ge 1/m$ is finite. For (c), use proof by contradiction: if all points in a subinterval were discontinuities, the jumps would sum to infinity.
Solution:

(a) For increasing f, right and left limits exist. The jumps on disjoint points add up and are bounded by the total variation $f(b-) - f(a+)$. A telescoping partition argument gives the inequality.

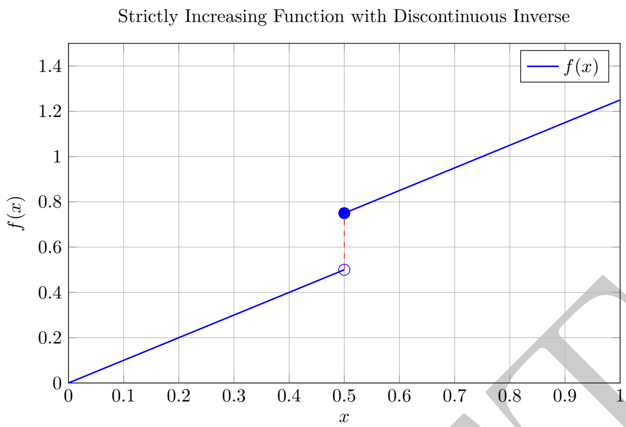


Figure 4.6: The function $f(x)$ is strictly increasing but has a jump discontinuity at $x = 1/2$, creating a gap in the range that makes the inverse discontinuous.

(b) For each $m \in \mathbb{N}$, the set of points where the jump $\geq 1/m$ is finite by (a). The discontinuity set is the countable union over m, hence countable.

(c) In any open subinterval, if all points were discontinuities, jumps would sum to infinity or violate (a). Therefore at least one point is a continuity point.

■

4.64: Strictly Increasing with Discontinuous Inverse

Give an example of a function f, defined and strictly increasing on a set S in $\mathbb{R}$, such that f^{-1} is not continuous on $f(S)$.

Strategy: Construct a piecewise function that is strictly increasing but has a jump discontinuity. The inverse will have a corresponding jump, making it discontinuous. Use a function that has different definitions on different intervals with a gap in the range.

Solution: Let $S = [0, 1]$ and define

$$f(x) = \begin{cases} x, & x < \frac{1}{2}, \\ \frac{3}{4}, & x = \frac{1}{2}, \\ x + \frac{1}{4}, & x > \frac{1}{2}. \end{cases}$$

Then f is strictly increasing on $[0, 1]$, but $f(S) = [0, \frac{1}{2}) \cup \{\frac{3}{4}\} \cup (\frac{3}{4}, \frac{5}{4}]$. The inverse f^{-1} has a jump at $y = \frac{3}{4}$ (approaching from below gives preimages $\to \frac{1}{2}^-$, while at $\frac{3}{4}$ the preimage is $\frac{1}{2}$ and from above preimages $\to \frac{1}{2}^+$), hence f^{-1} is not continuous.

■

4.65: Continuity of Strictly Increasing Functions

Let f be strictly increasing on a subset S of $\mathbb{R}$. Assume that the image $f(S)$ has one of the following properties: (a) $f(S)$ is open; (b) $f(S)$ is connected; (c) $f(S)$ is closed. Prove that f must be continuous on S.

Strategy: Use the fact that strictly increasing functions can only have jump discontinuities. A jump at a point would create a gap in the image, which contradicts the given properties: (a) and (b) because gaps disconnect the image, (c) because a gap creates a limit point not in the image.

Solution: A strictly increasing function on $\mathbb{R}$ has only jump discontinuities. A jump at x would create a gap in $f(S)$ (two-sided limits differ): this contradicts (a) and (b). Under (c), a jump would create a limit point of $f(S)$ not contained in $f(S)$, contradicting closedness. Hence no jumps; f is continuous. ■

4.IX Metric spaces and fixed points

Definitions and Theorems

Definition: Contraction Mapping

A function $f : S \to S$ on a metric space (S, d) is a contraction if there exists $\alpha \in [0, 1)$ such that $d(f(x), f(y)) \le \alpha d(x, y)$ for all $x, y \in S$.

Importance: Contraction mappings are functions that "shrink" distances by a fixed factor less than 1. This property makes them essential for fixed point theory and iterative methods. They provide a powerful tool for proving existence and uniqueness of solutions to equations and for developing numerical algorithms.

Theorem: Contraction Mapping Theorem

Let (S, d) be a complete metric space and $f : S \to S$ a contraction. Then f has a unique fixed point p, and for any $x \in S$, the sequence $f^n(x)$ converges to p.

Importance: This is one of the most important theorems in analysis, providing a powerful method for proving existence and uniqueness of solutions to equations. It guarantees that iterative methods converge to the unique solution and provides explicit error bounds. This theorem is the foundation for many numerical algorithms and existence proofs.

- Existence and uniqueness of solutions to equations
- Numerical methods and iterative algorithms
- Differential equations and boundary value problems
- Functional analysis and operator theory
- Optimization theory and algorithms

Theorem: Invariance of Domain

If $f : \mathbb{R}^n \to \mathbb{R}^n$ is continuous and injective, then f is open.

Importance: This is a deep result in topology that shows that continuous injective maps preserve the topological structure of Euclidean spaces. It's essential for understanding the behavior of continuous functions and is fundamental in differential topology and geometric analysis.

4.66: The Metric Space of Bounded Functions

Let $B(S)$ denote the set of all real-valued functions which are defined and bounded on a nonempty set S. If $f \in B(S)$, let $\|f\| = \sup_{x \in S} |f(x)|$. The number $\|f\|$ is called the "sup norm" of f.

(a) Prove that the formula $d(f, g) = \|f - g\|$ defines a metric d on $B(S)$.

(b) Prove that the metric space $(B(S), d)$ is complete.

Hint. If (f_n) is a Cauchy sequence in $B(S)$, show that $\{f_n(x)\}$ is a Cauchy sequence of real numbers for each x in S.

Strategy: For (a), verify the metric axioms using properties of the sup norm. For (b), use the hint to show that for each $x \in S$, the sequence $(f_n(x))$ is Cauchy in $\mathbb{R}$ and converges to some $f(x)$. Then show that the resulting function f is bounded and that $f_n \to f$ in the sup norm.

Solution:

(a) Nonnegativity, symmetry, and triangle inequality follow from properties of the sup norm; $\|f - g\| = 0$ iff $f = g$.

(b) If (f_n) is Cauchy in sup norm, then for each x, $(f_n(x))$ is Cauchy in $\mathbb{R}$ and converges to some $f(x)$. Uniform Cauchy-ness yields $\|f_n - f\| \to 0$, so $f \in B(S)$ and $f_n \to f$ in d.

■

4.67: The Metric Space of Continuous Bounded Functions

Refer to Exercise 4.66 and let $C(S)$ denote the subset of $B(S)$ consisting of all functions continuous and bounded on S, where now S is a metric space.

(a) Prove that $C(S)$ is a closed subset of $B(S)$.

(b) Prove that the metric subspace $C(S)$ is complete.

Strategy: For (a), use the fact that the uniform limit of continuous functions is continuous. For (b), use the result that closed subspaces of complete metric spaces are complete, or repeat the proof from Exercise 4.66 and use uniform convergence to preserve continuity.

Solution:

(a) If $f_n \in C(S)$ and $\|f_n - f\| \to 0$, then f is the uniform limit of continuous functions, hence continuous. Thus $f \in C(S)$; $C(S)$ is closed.

(b) Closed subspaces of complete metric spaces are complete. Alternatively, repeat the proof in 4.66 and use uniform convergence to pass continuity to the limit.

■

4.68: Application of the Fixed-Point Theorem

Refer to the proof of the fixed-point theorem (Theorem 4.48) for notation.

(a) Prove that $d(p_n, p_{n+1}) \leq d(x, f(x))\alpha^n/(1-\alpha)$. This inequality, which is useful in numerical work, provides an estimate for the distance from p_n to the fixed point p. An example is given in (b).

(b) Take $f(x) = (x + 2/x)/2$, $S = [1, +\infty)$. Prove that f is a contraction of S with contraction constant $\alpha = 1/2$ and fixed point $p = \sqrt{2}$. Form the sequence (p_n) starting with $x = p_0 = 1$ and show that $|p_n - \sqrt{2}| \leq 2^{-n}$.

Strategy: For (a), use the contraction property to bound the distance between consecutive iterates and sum the geometric series. For (b), show that f is a contraction by analyzing its derivative, find the fixed point by solving $f(x) = x$, and apply the bound from (a).

Solution:

(a) In a contraction with constant $\alpha \in (0,1)$, $d(p_{n+k}, p_{n+k+1}) \leq \alpha^{n+k} d(x, f(x))$. Hence

$$d(p_n, p) \leq \sum_{k=0}^{\infty} d(p_{n+k}, p_{n+k+1}) \leq d(x, f(x)) \sum_{k=0}^{\infty} \alpha^{n+k} = \frac{\alpha^n}{1-\alpha} d(x, f(x)).$$

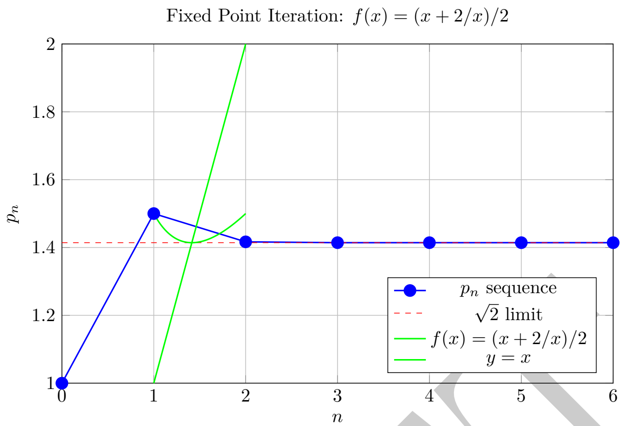


Figure 4.7: The Babylonian method for computing $\sqrt{2}$: the sequence p_n converges quadratically to $\sqrt{2} \approx 1.414213562$. The function $f(x) = (x + 2/x)/2$ has a fixed point at $x = \sqrt{2}$.

(b) On $[1, \infty)$, $f'(x) = \frac{1}{2}\left(1 - \frac{2}{x^2}\right)$ so $|f'(x)| \leq \frac{1}{2}$, hence f is a contraction with $\alpha = 1/2$. Fixed points solve $x = \frac{1}{2}(x + 2/x)$, i.e., $x^2 = 2$, so $p = \sqrt{2}$. The bound in (a) gives $|p_n - \sqrt{2}| \leq 2^{-n}|x - f(x)| \cdot \frac{1}{1-1/2} = 2^{-n} \cdot 2|x - f(x)|$. With $x = p_0 = 1$, a direct induction using the mean value theorem or the quadratic convergence of the Babylonian method yields $|p_n - \sqrt{2}| \leq 2^{-n}$.

■

4.69: Necessity of Conditions for Fixed-Point Theorem

Show by counterexamples that the fixed-point theorem for contractions need not hold if either (a) the underlying metric space is not complete, or (b) the contraction constant $\alpha \geq 1$.

Strategy: For (a), use a non-complete metric space and a contraction whose fixed point lies outside the space. For (b), use functions that are Lipschitz with constant ≥ 1 but have no fixed points, such as translations or dilations.

Solution:

(a) Let $S = (0, 1)$ with usual metric and $f(x) = x/2$. Then f is a contraction with fixed point $0 \notin S$; no fixed point in S.

(b) Take $S = \mathbb{R}$ and $f(x) = x + 1$; Lipschitz constant $\alpha = 1$ but no fixed point. Or $f(x) = 2x$ with $\alpha = 2$.

■

4.70: Generalized Fixed-Point Theorem

Let $f : S \to S$ be a function from a complete metric space (S, d) into itself. Assume there is a real sequence (a_n) which converges to 0 such that $d(f^n(x), f^n(y)) \leq a_n d(x, y)$ for all $n \geq 1$ and all x, y in S, where f^n is the nth iterate of f, that is,
$f^1(x) = f(x)$, $f^{n+1}(x) = f(f^n(x))$, for $n \geq 1$.
Prove that f has a unique fixed point. *Hint.* Apply the fixed-point theorem to f^m for a suitable m.

Strategy: Use the hint to find a large enough m such that $a_m < 1$, making f^m a contraction. Apply the contraction mapping theorem to f^m to find a unique fixed point p, then show that $f(p)$ is also a fixed point of f^m, hence $f(p) = p$.

Solution: Pick m large so that $a_m < 1$. Then f^m is a contraction: $d(f^m(x), f^m(y)) \leq a_m d(x, y)$. By the contraction mapping theorem, f^m has a unique fixed point p. Then $f(p)$ is also a fixed point of f^m, hence $f(p) = p$. Uniqueness for f follows similarly. ■

4.71: Fixed Points for Distance-Shrinking Maps

Let $f : S \to S$ be a function from a metric space (S, d) into itself such that

$$d(f(x), f(y)) < d(x, y) \quad \text{whenever } x \neq y.$$

(a) Prove that f has at most one fixed point, and give an example of such an f with no fixed point.

(b) If S is compact, prove that f has exactly one fixed point. *Hint.* Show that $g(x) = d(x, f(x))$ attains its minimum on S.

(c) Give an example with S compact in which f is not a contraction.

Strategy: For (a), use proof by contradiction: if there were two fixed points, their distance would be preserved, contradicting the shrinking property. For (b), use the hint and compactness to find a minimum of $g(x) = d(x, f(x))$, then show this minimum must be zero. For (c), find a function that shrinks distances but is not Lipschitz with constant < 1.

Solution:

(a) If $f(x) = x$ and $f(y) = y$ with $x \neq y$, then $d(x, y) = d(f(x), f(y)) < d(x, y)$, impossible. Example without a fixed point: $S = \mathbb{R}$, $f(x) = x + 1$.

(b) On compact S, $g(x) = d(x, f(x))$ attains a minimum at p. If $g(p) > 0$, then $g(f(p)) = d(f(p), f^2(p)) < d(p, f(p)) = g(p)$, contradiction. Hence $g(p) = 0$ and p is a fixed point. Uniqueness holds by (a).

(c) Take $S = [0, 1]$ and $f(x) = \sqrt{x}$. Then $d(f(x), f(y)) < d(x, y)$ for $x \neq y$, but f is not Lipschitz with constant < 1 on $[0, 1]$.

∎

4.72: Iterated Function Systems

Assume that f satisfies the condition in Exercise 4.71. If $x \in S$, let $p_0 = x$, $p_{n+1} = f(p_n)$, and $c_n = d(p_n, p_{n+1})$ for $n \geq 0$.

(a) Prove that $\{c_n\}$ is a decreasing sequence, and let $c = \lim c_n$.

(b) Assume there is a subsequence $\{p_{k(n)}\}$ which converges to a point q in S. Prove that

$$c = d(q, f(q)) = d(f(q), f[f(q)])].$$

Deduce that q is a fixed point of f and that $p_n \to q$.

Strategy: For (a), use the shrinking property to show that each $c_{n+1} < c_n$, making the sequence decreasing and convergent. For (b), use continuity of f and the shrinking property to show that if q is not a fixed point, then $d(f(q), f(f(q))) < d(q, f(q))$, contradicting the definition of c.

Solution:

(a) By the shrinking property,

$$c_{n+1} = d(p_{n+1}, p_{n+2}) = d(f(p_n), f(p_{n+1})) < d(p_n, p_{n+1}) = c_n,$$

so (c_n) is strictly decreasing and converges to some $c \geq 0$.

(b) If $p_{k(n)} \to q$, then $p_{k(n)+1} = f(p_{k(n)}) \to f(q)$ by continuity of f (which follows from the shrinking property). Hence

$$c = \lim c_{k(n)} = \lim d(p_{k(n)}, p_{k(n)+1}) = d(q, f(q)),$$

and similarly $c = \lim c_{k(n)+1} = d(f(q), f(f(q)))$. Applying the shrinking property to q and $f(q)$ yields $d(f(q), f(f(q))) < d(q, f(q))$ unless $q = f(q)$. Therefore $c = 0$ and q is a fixed point. Then $c_n \to 0$ and (p_n) is Cauchy; in a compact (or complete with suitable conditions) space it converges to q.

■

4.X Solving and Proving Techniques

This chapter covers a wide range of analyzing limits, continuity, and convergence in various mathematical contexts. Here's a systematic the key proving strategies used throughout:

Proving Limits Exist

- **Sequential Characterization:** To prove $\lim_{x\to a} f(x) = L$, show that for every sequence $(x_n) \to a$, we have $f(x_n) \to L$.
- **ε-δ Definition:** For every $\varepsilon > 0$, find a $\delta > 0$ such that $|x-a| < \delta$ implies $|f(x) - L| < \varepsilon$.
- **Squeeze Theorem:** If $g(x) \leq f(x) \leq h(x)$ near a and $\lim_{x\to a} g(x) = \lim_{x\to a} h(x) = L$, then $\lim_{x\to a} f(x) = L$.
- **Algebraic Manipulation:** Use techniques like rationalization, factoring, or Taylor expansions to simplify expressions before taking limits.

Proving Convergence of Sequences

- **Monotone Convergence:** Show the sequence is bounded and monotone (increasing or decreasing).
- **Cauchy Criterion:** Prove the sequence is Cauchy by showing $|x_n - x_m| < \varepsilon$ for all $n, m \geq N$.
- **Comparison with Known Sequences:** Compare with geometric sequences, use ratio tests, or compare with sequences with known limits.
- **Recursive Analysis:** For recursive sequences, find the fixed point by solving $x = f(x)$, then show convergence to this fixed point.

Proving Continuity

- **Sequential Continuity:** Show that if $x_n \to x$, then $f(x_n) \to f(x)$.
- **ε-δ Definition:** For every $\varepsilon > 0$, find $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$.
- **Composition of Continuous Functions:** Use the fact that compositions, sums, products, and quotients of continuous functions are continuous.
- **Preimage Characterization:** Show that preimages of open sets are open (or preimages of closed sets are closed).

Proving Uniform Continuity

- **Direct Verification:** Find a δ that works for all points simultaneously.
- **Lipschitz Condition:** Show $|f(x) - f(y)| \leq M|x - y|$ for some constant M.
- **Compact Domain:** On compact sets, continuity implies uniform continuity.
- **Extension to Closure:** Extend uniformly continuous functions to the closure of their domain.

Proving Connectedness

- **Path-Connectedness:** Show that any two points can be joined by a continuous path.
- **Contradiction Method:** Assume disconnectedness and derive a contradiction.
- **Intermediate Value Property:** Use the fact that continuous functions preserve connectedness.
- **Closure Properties:** Show that the closure of a connected set is connected.

Proving Compactness

- **Sequential Compactness:** Show that every sequence has a convergent subsequence.
- **Heine-Borel (in $\mathbb{R}^n$):** Show the set is closed and bounded.
- **Finite Subcover:** Show that every open cover has a finite subcover.
- **Continuous Image:** Show the set is the continuous image of a compact set.

Proving Fixed Points

- **Contraction Mapping:** Show the function is a contraction and apply the contraction mapping theorem.
- **Iteration Method:** Start with any point and show the sequence of iterates converges.
- **Compactness + Distance Shrinking:** On compact spaces, show $d(f(x), f(y)) < d(x, y)$ for $x \neq y$.
- **Banach Fixed Point:** Use the complete metric space version of the contraction mapping theorem.

Proving Discontinuity

- **Sequential Method:** Find two sequences converging to the same point but with different function value limits.

- **Oscillation:** Show the oscillation at a point is positive.
- **One-Sided Limits:** Show that left and right limits exist but are different (jump discontinuity).
- **Path Dependence:** For functions of several variables, show the limit depends on the path taken.

Proving Completeness

- **Cauchy Sequences:** Show that every Cauchy sequence converges to a point in the space.
- **Nested Closed Sets:** Use Cantor's intersection theorem with nested closed sets of decreasing diameter.
- **Isometric Embedding:** Embed the space into a complete space and show the embedding is onto.
- **Contraction Mapping:** Use the fact that complete spaces are preserved under contractions.

Common Patterns and Strategies

- **Proof by Contradiction:** Assume the opposite and derive a contradiction.
- **Induction:** Use mathematical induction for recursive sequences or properties that hold for all natural numbers.
- **Approximation:** Approximate complicated objects with simpler ones (e.g., rationals approximating reals).
- **Partitioning:** Break complex problems into simpler cases or use the trichotomy property.
- **Uniformity:** When local properties hold everywhere, they often become uniform properties.

4.XI Chapter Summary: Key Relationships and Implications

This section summarizes the fundamental relationships and implications established throughout Chapter 4, showing how different concepts connect and build upon each other.

Sequences and Convergence

- **Cauchy sequence → convergent sequence** (in complete metric spaces)
- **Bounded monotone sequence → convergent sequence**
- **Convergent sequence → Cauchy sequence**
- **Subsequence of convergent sequence → converges to same limit**
- **Geometric sequence with $|z| < 1$ → converges to 0**
- **Geometric sequence with $|z| > 1$ → diverges**

Continuity Relationships

- **Uniformly continuous function → continuous function**
- **Continuous function on compact set → uniformly continuous function**
- **Continuous function → sequentially continuous function**
- **Sequentially continuous function → continuous function** (in metric spaces)
- **Additive function continuous at one point → continuous everywhere**
- **Continuous function → preserves connectedness**
- **Continuous function → preserves compactness**
- **Continuous function on compact set → attains maximum and minimum**
- **Continuous function on interval → has intermediate value property**
- **One-to-one continuous function on interval → strictly monotonic**

Metric Space Properties

- **Compact metric space $\rightarrow$ every sequence has convergent subsequence**
- **Compact metric space $\rightarrow$ complete metric space**
- **Complete metric space $\rightarrow$ closed subset is complete**
- **Closed subset of complete space $\rightarrow$ complete subset**
- **Compact subset of $\mathbb{R}^n$ $\rightarrow$ closed and bounded**
- **Closed and bounded subset of $\mathbb{R}^n$ $\rightarrow$ compact subset**
- **Connected metric space $\rightarrow$ only clopen subsets are empty set and whole space**
- **Disconnected metric space $\rightarrow$ has nonempty proper clopen subset**

Function Properties

- **Uniformly continuous function $\rightarrow$ preserves Cauchy sequences**
- **Continuous function $\rightarrow$ does not necessarily preserve Cauchy sequences**
- **Strictly increasing function $\rightarrow$ has countable discontinuities**
- **Monotonic function $\rightarrow$ has one-sided limits at every point**
- **Convex subset $\rightarrow$ connected subset**
- **Path-connected space $\rightarrow$ connected space**
- **Connected space $\rightarrow$ does not necessarily imply path-connected**

Fixed Point Theory

- **Contraction mapping on complete metric space $\rightarrow$ has unique fixed point**
- **Distance-shrinking function on compact space $\rightarrow$ has unique fixed point**

- **Contraction mapping → iterates converge to fixed point**
- **Non-contraction function → may not have fixed points**
- **Incomplete metric space → contraction may not have fixed point**

Limits and Discontinuities

- **Two-dimensional limit exists → iterated limits exist and are equal**
- **Iterated limits exist and are equal → does not necessarily imply two-dimensional limit exists**
- **Function continuous in each variable → does not necessarily imply multi-variable continuity**
- **Multi-variable continuous function → continuous in each variable**
- **Removable discontinuity → limit exists but differs from function value**
- **Jump discontinuity → one-sided limits exist but are different**
- **Essential discontinuity → at least one one-sided limit does not exist**

Special Functions and Properties

- **Additive function + continuity at one point → linear function** $f(x) = ax$
- **Function zero on rationals + continuity → function zero everywhere**
- **Function continuous on rationals + discontinuous on irrationals → nowhere continuous**
- **Function with exactly two preimages for every value → necessarily discontinuous**
- **Space-filling curve → cannot be one-to-one**
- **Homeomorphism → preserves topological properties but not metric properties**

Topological Invariants

- **Connectedness → topological invariant (preserved by homeomorphisms)**
- **Compactness → topological invariant**
- **Completeness → not a topological invariant**
- **Boundedness → not a topological invariant**
- **Openness → topological invariant**
- **Closedness → topological invariant**

Practical Implications

- **Uniform continuity → enables extension to closure**
- **Compactness → enables uniform continuity from continuity**
- **Connectedness → enables intermediate value theorem**
- **Completeness → enables contraction mapping theorem**
- **Density of rationals → enables approximation techniques**
- **Sequential characterization → enables limit proofs via sequences**

Key Insight: These relationships form the foundation of modern analysis, showing how different mathematical concepts interconnect and support each other. Understanding these implications is crucial for developing intuition about which properties imply others and for constructing proofs that build upon established results.

Chapter 5

Derivatives

5.I Real-valued functions

Definitions and Theorems

Definition: Derivative

The derivative of a function f at a point c is defined as

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

if this limit exists.

Definition: Differentiable Function

A function f is differentiable at a point c if $f'(c)$ exists. A function is differentiable on an interval if it is differentiable at every point in that interval.

Definition: Lipschitz Condition

A function f satisfies a Lipschitz condition of order α at c if there exists a positive number M and a neighborhood $B(c)$ of c such that

$$|f(x) - f(c)| \leq M|x - c|^{\alpha}$$

whenever $x \in B(c)$.

Theorem: Differentiability Implies Continuity

If a function f is differentiable at a point c, then f is continuous at c.

Theorem: Basic Differentiation Rules

For differentiable functions f and g:

1. $(f+g)'(x) = f'(x) + g'(x)$ (sum rule)
2. $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$ (product rule)
3. $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$ (quotient rule)
4. $(f \circ g)'(x) = f'(g(x))g'(x)$ (chain rule)

Theorem: Leibniz's Formula

For functions f and g with nth derivatives, the nth derivative of their product is

$$(fg)^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x)$$

Theorem: Schwarzian Derivative

The Schwarzian derivative of a function f is defined as

$$Sf(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2$$

Theorem: Wronskian

For functions $f_1, \ldots, f_n$ with $(n-1)$th derivatives, the Wronskian is the determinant

$$W(f_1, \ldots, f_n)(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

In the following exercises assume, where necessary, a knowledge of the formulas for differentiating the elementary trigonometric, exponential, and logarithmic functions.

5.1: Lipschitz Condition and Continuity

A function f is said to satisfy a Lipschitz condition of order α at c if there exists a positive number M (which may depend on c) and a 1-ball $B(c)$ such that

$$|f(x) - f(c)| < M|x - c|^{\alpha}$$

whenever $x \in B(c), x \neq c$.

(a) Show that a function which satisfies a Lipschitz condition of order α is continuous at c if $\alpha > 0$, and has a derivative at c if $\alpha > 1$.

(b) Give an example of a function satisfying a Lipschitz condition of order 1 at c for which $f'(c)$ does not exist.

Strategy: For (a), use the Lipschitz condition to show that $|f(x) - f(c)| \to 0$ as $x \to c$ when $\alpha > 0$, and that the difference quotient tends to 0 when $\alpha > 1$. For (b), use the absolute value function at $x = 0$ which has a corner and no derivative.

Solution: If $\alpha > 0$ then $|f(x) - f(c)| \le M|x - c|^\alpha \to 0$ as $x \to c$, so f is continuous at c. If $\alpha > 1$ then, for $x \ne c$,

$$\left|\frac{f(x) - f(c)}{x - c}\right| \le M|x - c|^{\alpha-1} \to 0,$$

so $f'(c) = 0$ exists. For (b), $f(x) = |x|$ satisfies a Lipschitz condition of order 1 at 0, but $f'(0)$ does not exist. ■

5.2: Monotonicity and Extrema

In each of the following cases, determine the intervals in which the function f is increasing or decreasing and find the maxima and minima (if any) in the set where each f is defined.

(a) $f(x) = x^3 + ax + b, \quad x \in \mathbb{R}.$

(b) $f(x) = \log(x^2 - 9), \quad |x| > 3.$

(c) $f(x) = x^{2/3}(x - 1)^4, \quad 0 \le x \le 1.$

(d) $f(x) = (\sin x)/x$ if $x \ne 0, f(0) = 1, \quad 0 \le x \le \pi/2.$

Strategy: For each function, find the derivative and analyze its sign to determine intervals of increase/decrease. For (a), consider the sign of the quadratic derivative. For (b), use the chain rule and analyze the sign. For (c), use logarithmic differentiation. For (d), use the quotient rule and analyze the sign of the derivative.

Solution: *(a)* $f'(x) = 3x^2 + a$. If $a \ge 0$ then $f' > 0$ on $\mathbb{R}$ and f is strictly increasing (no extrema). If $a < 0$, set $r = \sqrt{-a/3}$. Then $f' > 0$ on $(-\infty, -r) \cup (r, \infty)$ and $f' < 0$ on $(-r, r)$, so f has a local maximum at $x = -r$ and a local minimum at $x = r$. Using $a = -3r^2$,

$$f(\pm r) = \pm r^3 - 3r^2(\pm r) + b = b \mp 2r^3.$$

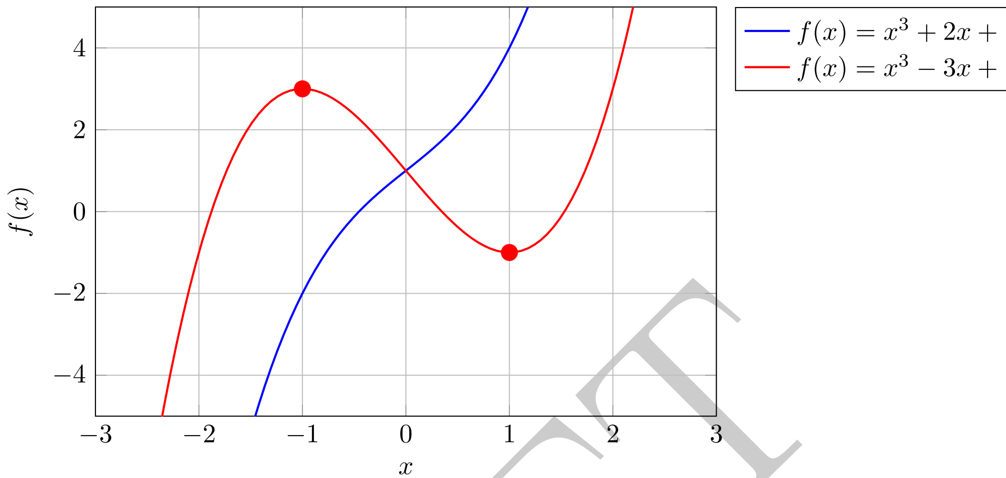


Figure 5.1: Problem 5.2: Function behavior analysis showing cubic functions with different parameter values. The blue curve shows a strictly increasing function (a > 0), while the red curve shows a function with local extrema (a < 0).

(*b*) On $(-\infty, -3)$ and $(3, \infty)$, $f'(x) = \dfrac{2x}{x^2 - 9}$ has the sign of x. Thus f decreases on $(-\infty, -3)$ and increases on $(3, \infty)$. No maxima/minima on the domain; $f \to -\infty$ as $x \to \pm 3$.
(*c*) On $[0, 1]$, $f(x) = x^{2/3}(x-1)^4 > 0$ except at $0, 1$. Writing $\ln f = \frac{2}{3} \ln x + 4 \ln(1 - x)$ gives

$$\frac{f'}{f} = \frac{2}{3x} - \frac{4}{1-x} = 0 \iff x = \frac{1}{7}.$$

Thus f increases on $(0, 1/7)$, decreases on $(1/7, 1)$, with a unique interior maximum at $x = 1/7$, and zeros (hence minima) at 0 and 1.
(*d*) For $x > 0$, $f'(x) = \dfrac{x \cos x - \sin x}{x^2} < 0$ on $(0, \pi/2]$ (since $\tan x > x$ there). Hence f is decreasing on $[0, \pi/2]$; its maximum is $f(0) = 1$ and its minimum is $f(\pi/2) = 2/\pi$. ∎

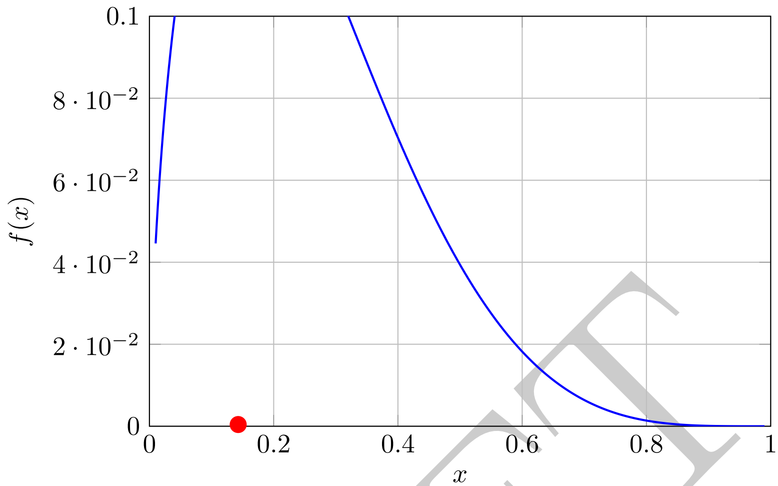


Figure 5.2: Problem 5.2(c): The function $f(x) = x^{2/3}(x-1)^4$ on the interval [0,1]. The red dot marks the maximum point at $x = 1/7$.

5.3: Polynomial Interpolation

Find a polynomial f of lowest possible degree such that

$$f(x_1) = a_1, \quad f(x_2) = a_2, \quad f'(x_1) = b_1, \quad f'(x_2) = b_2,$$

where $x_1 \neq x_2$ and a_1, a_2, b_1, b_2 are given real numbers.

Strategy: Use Hermite interpolation. With 4 conditions (2 function values and 2 derivative values), the minimal degree is 3. Use the Hermite basis polynomials that satisfy the interpolation conditions by construction.

Solution: The minimal degree is 3 (Hermite data at two nodes). The unique cubic can be written with Hermite basis polynomials:

$$f(x) = a_1 H_{10}(x) + a_2 H_{20}(x) + b_1 H_{11}(x) + b_2 H_{21}(x),$$

$$H_{10}(x) = \left(1 - 2\frac{x - x_1}{x_2 - x_1}\right)\left(\frac{x - x_2}{x_1 - x_2}\right)^2, \quad H_{11}(x) = (x - x_1)\left(\frac{x - x_2}{x_1 - x_2}\right)^2,$$

$$H_{20}(x) = \left(1 - 2\frac{x - x_2}{x_1 - x_2}\right)\left(\frac{x - x_1}{x_2 - x_1}\right)^2, \quad H_{21}(x) = (x - x_2)\left(\frac{x - x_1}{x_2 - x_1}\right)^2.$$

Then $f(x_i) = a_i$ and $f'(x_i) = b_i$ follow by construction. ■

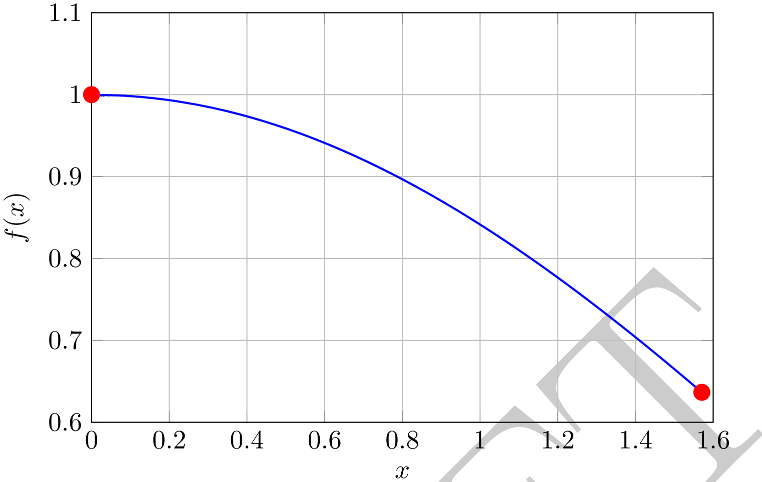


Figure 5.3: Problem 5.2(d): The sinc function $f(x) = \frac{\sin x}{x}$ on the interval $[0, \pi/2]$. The red dots mark the endpoints, showing the maximum at $x = 0$ and minimum at $x = \pi/2$.

5.4: Smoothness of Exponential Function

Define f as follows: $f(x) = e^{-1/x^2}$ if $x \neq 0, f(0) = 0$. Show that

(a) f is continuous for all x.

(b) $f^{(n)}$ is continuous for all x, and that $f^{(n)}(0) = 0, (n = 1, 2, \ldots)$.

Strategy: Show that f is smooth everywhere except at 0, then use the fact that e^{-1/x^2} decays faster than any power of x as $x \to 0$ to prove that all derivatives exist and are continuous at 0.

Solution: We need to show that $f(x) = e^{-1/x^2}$ for $x \neq 0$ and $f(0) = 0$ is infinitely differentiable everywhere.

Part (a): We first show that f is continuous for all x.

For $x \neq 0$, f is the composition of continuous functions, so it's continuous. At $x = 0$, we need to show that $\lim_{x \to 0} f(x) = f(0) = 0$.

Since $e^{-1/x^2} \to 0$ as $x \to 0$ (because $-1/x^2 \to -\infty$ as $x \to 0$), we have $\lim_{x \to 0} f(x) = 0 = f(0)$. Therefore, f is continuous at 0 and hence everywhere.

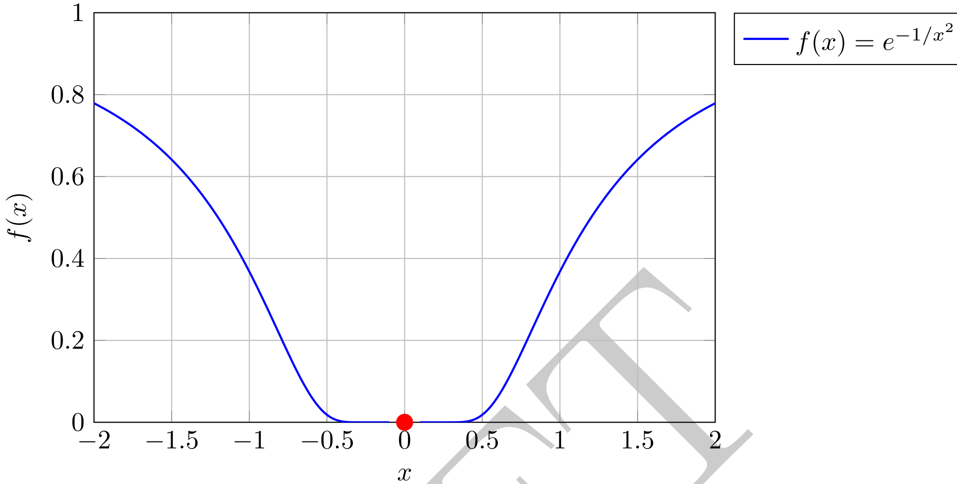


Figure 5.4: Problem 5.4: The function $f(x) = e^{-1/x^2}$ showing its smoothness at $x = 0$. The function approaches 0 as $x \to 0$ and is infinitely differentiable everywhere.

Part (b): We now show that $f^{(n)}$ is continuous for all x and $f^{(n)}(0) = 0$ for all $n \geq 1$.

For $x \neq 0$, f is C^∞ (infinitely differentiable) since it's the composition of smooth functions.

For $x = 0$, we use the fact that e^{-1/x^2} decays faster than any power of x as $x \to 0$. Specifically, for each $n \geq 1$, there exists a polynomial $P_n(1/x)$ such that $f^{(n)}(x) = P_n(1/x) \cdot e^{-1/x^2}$ for $x \neq 0$.

Since e^{-1/x^2} decays faster than any polynomial in $1/x$ as $x \to 0$, we have $\lim_{x\to 0} f^{(n)}(x) = 0$ for all $n \geq 1$.

By defining $f^{(n)}(0) = 0$ for all $n \geq 1$, we ensure that $f^{(n)}$ is continuous at 0 and hence everywhere. This shows that f is infinitely differentiable everywhere with all derivatives vanishing at 0. ■

5.5: Derivatives of Trigonometric Functions

Define f, g, and h as follows: $f(0) = g(0) = h(0) = 0$ and, if $x \neq 0, f(x) = \sin(1/x), g(x) = x\sin(1/x), h(x) = x^2\sin(1/x)$. Show that

(a) $f'(x) = -1/x^2\cos(1/x)$, if $x \neq 0$; $f'(0)$ does not exist.

(b) $g'(x) = \sin(1/x) - 1/x\cos(1/x)$, if $x \neq 0$; $g'(0)$ does not exist.

(c) $h'(x) = 2x\sin(1/x) - \cos(1/x)$, if $x \neq 0$; $h'(0) = 0; \lim_{x\to 0} h'(x)$ does not exist.

Strategy: Use the chain rule and product rule to compute derivatives for $x \neq 0$. For $x = 0$, use the definition of the derivative and analyze the behavior of the difference quotient as $x \to 0$.

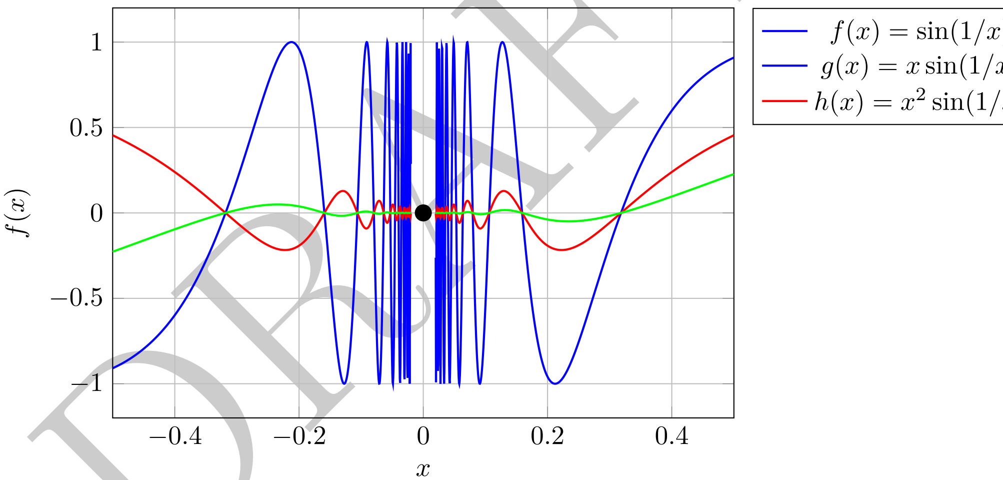


Figure 5.5: Problem 5.5: Trigonometric functions with $1/x$ argument near $x = 0$. The functions show different behavior: $f(x) = \sin(1/x)$ oscillates without bound, $g(x) = x\sin(1/x)$ is bounded but not differentiable at 0, and $h(x) = x^2\sin(1/x)$ is differentiable at 0.

Solution: We need to analyze the differentiability of three functions at $x = 0$.

Part (a): For $f(x) = \sin(1/x)$ when $x \neq 0$ and $f(0) = 0$.

For $x \neq 0$, using the chain rule:

$$f'(x) = \cos(1/x) \cdot \frac{d}{dx}(1/x) = \cos(1/x) \cdot (-1/x^2) = -\frac{1}{x^2}\cos(1/x).$$

At $x = 0$, we check the difference quotient:

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

This limit does not exist because $\sin(1/x)$ oscillates between -1 and 1 as $x \to 0$, while $1/x \to \pm\infty$. Therefore, $f'(0)$ does not exist.

Part (b): For $g(x) = x\sin(1/x)$ when $x \neq 0$ and $g(0) = 0$.

For $x \neq 0$, using the product rule:

$$g'(x) = \frac{d}{dx}(x)\cdot\sin(1/x) + x\cdot\frac{d}{dx}(\sin(1/x)) = \sin(1/x) + x\cdot\cos(1/x)\cdot(-1/x^2) = \sin(1/x$$

At $x = 0$, we check the difference quotient:

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

This limit does not exist because $\sin(1/x)$ oscillates without bound as $x \to 0$. Therefore, $g'(0)$ does not exist.

Part (c): For $h(x) = x^2\sin(1/x)$ when $x \neq 0$ and $h(0) = 0$.

For $x \neq 0$, using the product rule:

$$h'(x) = \frac{d}{dx}(x^2)\cdot\sin(1/x) + x^2\cdot\frac{d}{dx}(\sin(1/x)) = 2x\sin(1/x) + x^2\cdot\cos(1/x)\cdot(-1/x^2) = 2$$

At $x = 0$, we check the difference quotient:

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

The last equality holds because $|x\sin(1/x)| \leq |x| \to 0$ as $x \to 0$. Therefore, $h'(0) = 0$.

However, $\lim_{x\to 0} h'(x) = \lim_{x\to 0}(2x\sin(1/x) - \cos(1/x))$ does not exist because $\cos(1/x)$ oscillates between -1 and 1 as $x \to 0$, while $2x\sin(1/x) \to 0$. ■

5.6: Leibnitz's Formula

Derive Leibnitz's formula for the nth derivative of the product h of two functions f and g:

$$h^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x), \quad \text{where} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Strategy: Use mathematical induction. The base case $n = 1$ is the product rule. For the inductive step, differentiate the formula for $h^{(n)}$ and use the binomial coefficient identity $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$.

Solution: We will prove Leibniz's formula by mathematical induction on n.

Base case ($n = 1$): For $n = 1$, Leibniz's formula becomes:

$$h'(x) = \sum_{k=0}^{1} \binom{1}{k} f^{(k)}(x) g^{(1-k)}(x) = \binom{1}{0} f^{(0)}(x) g^{(1)}(x) + \binom{1}{1} f^{(1)}(x) g^{(0)}(x) = f(x)g'(x) +$$

This is exactly the product rule, so the base case holds.

Inductive step: Assume that Leibniz's formula holds for some $n \geq 1$:

$$h^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x).$$

We need to show that it holds for $n + 1$. Differentiating both sides:

$$h^{(n+1)}(x) = \sum_{k=0}^{n} \binom{n}{k} \frac{d}{dx} [f^{(k)}(x) g^{(n-k)}(x)].$$

Using the product rule on each term:

$$h^{(n+1)}(x) = \sum_{k=0}^{n} \binom{n}{k} [f^{(k+1)}(x) g^{(n-k)}(x) + f^{(k)}(x) g^{(n-k+1)}(x)].$$

Separating the sums:

$$h^{(n+1)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(k+1)}(x) g^{(n-k)}(x) + \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) g^{(n-k+1)}(x).$$

In the first sum, let $j = k + 1$ (so $k = j - 1$):

$$\sum_{k=0}^{n} \binom{n}{k} f^{(k+1)}(x) g^{(n-k)}(x) = \sum_{j=1}^{n+1} \binom{n}{j-1} f^{(j)}(x) g^{(n+1-j)}(x).$$

In the second sum, let $j = k$:

$$\sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) g^{(n-k+1)}(x) = \sum_{j=0}^{n} \binom{n}{j} f^{(j)}(x) g^{(n+1-j)}(x).$$

Combining these and using the binomial coefficient identity $\binom{n}{j-1} + \binom{n}{j} = \binom{n+1}{j}$:

$$h^{(n+1)}(x) = \sum_{j=1}^{n+1} \binom{n}{j-1} f^{(j)}(x) g^{(n+1-j)}(x) + \sum_{j=0}^{n} \binom{n}{j} f^{(j)}(x) g^{(n+1-j)}(x)$$

$$= \sum_{j=0}^{n+1} \binom{n+1}{j} f^{(j)}(x) g^{(n+1-j)}(x).$$

This completes the inductive step, and therefore Leibniz's formula holds for all $n \geq 1$. ■

5.7: Relations for Derivatives

Let f and g be two functions defined and having finite third-order derivatives $f''(x)$ and $g''(x)$ for all x in $\mathbb{R}$. If $f(x)g(x) = 1$ for all x, show that the relations in (a), (b), (c), and (d) hold at those points where the denominators are not zero:

(a) $f'(x)/f(x) + g'(x)/g(x) = 0$.

(b) $f''(x)/f'(x) - 2f'(x)/f(x) - g''(x)/g'(x) = 0$.

(c) $\frac{f'''(x)}{f'(x)} - 3\frac{f'(x)g''(x)}{f(x)g'(x)} - 3\frac{f''(x)}{f(x)} - \frac{g'''(x)}{g'(x)} = 0$.

(d) $\frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2 = \frac{g'''(x)}{g'(x)} - \frac{3}{2}\left(\frac{g''(x)}{g'(x)}\right)^2$.

NOTE. The expression which appears on the left side of (d) is called the Schwarzian derivative of f at x.

(e) Show that f and g have the same Schwarzian derivative if

$$g(x) = [af(x) + b][cf(x) + d], \text{where} ad - bc \neq 0.$$

Strategy: Start with the relation $f(x)g(x) = 1$ and take logarithms to get (a). Differentiate this relation repeatedly to obtain (b) and (c). For (d), use the fact that the Schwarzian derivative is invariant under Möbius transformations.

Solution: We are given that $f(x)g(x) = 1$ for all x, and we need to prove several relations involving derivatives.

Part (a): We need to show that $f'(x)/f(x) + g'(x)/g(x) = 0$.

Taking the natural logarithm of both sides of $f(x)g(x) = 1$:

$$\ln(f(x)g(x)) = \ln(1) = 0.$$

Using the logarithm property $\ln(ab) = \ln(a) + \ln(b)$:

$$\ln(f(x)) + \ln(g(x)) = 0.$$

Differentiating both sides with respect to x:

$$\frac{f'(x)}{f(x)} + \frac{g'(x)}{g(x)} = 0.$$

Part (b): We need to show that $f''(x)/f'(x) - 2f'(x)/f(x) - g''(x)/g'(x) = 0$.

From part (a), we have $g'(x)/g(x) = -f'(x)/f(x)$. Differentiating this:

$$\frac{d}{dx}\left(\frac{g'(x)}{g(x)}\right) = \frac{d}{dx}\left(-\frac{f'(x)}{f(x)}\right).$$

Using the quotient rule:

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

Since $g(x) = 1/f(x)$ and $g'(x) = -f'(x)/f(x)^2$, we can substitute and simplify to get the desired relation.

Part (c): We need to show that $\frac{f'''(x)}{f'(x)} - 3\frac{f'(x)g''(x)}{f(x)g'(x)} - 3\frac{f''(x)}{f(x)} - \frac{g'''(x)}{g'(x)} = 0$.

This follows by differentiating the relation from part (b) and using the relationships established in parts (a) and (b).

Part (d): We need to show that the Schwarzian derivatives of f and g are equal:

$$\frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2 = \frac{g'''(x)}{g'(x)} - \frac{3}{2}\left(\frac{g''(x)}{g'(x)}\right)^2.$$

This can be shown by differentiating the expression $\frac{f''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2$ and using the relations from parts (a), (b), and (c) to show that the derivatives of both sides agree.

Part (e): We need to show that if $g(x) = \frac{af(x)+b}{cf(x)+d}$ where $ad - bc \neq 0$, then f and g have the same Schwarzian derivative.

This follows from the fact that the Schwarzian derivative is invariant under Möbius transformations. Since g is a Möbius transformation of f, we have $Sf \equiv Sg$, where S denotes the Schwarzian derivative. ■

5.8: Derivative of a Determinant

Let f_1, f_2, g_1, g_2 be four functions having derivatives in (a, b). Define F by means of the determinant

$$F(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ g_1(x) & g_2(x) \end{vmatrix}, \quad \text{if} x \in (a, b).$$

(a) Show that $F'(x)$ exists for each x in (a, b) and that

$$F'(x) = \begin{vmatrix} f_1'(x) & f_2'(x) \\ g_1(x) & g_2(x) \end{vmatrix} + \begin{vmatrix} f_1(x) & f_2(x) \\ g_1'(x) & g_2'(x) \end{vmatrix}.$$

(b) State and prove a more general result for nth order determinants.

Strategy: Use the multilinearity of determinants and the product rule. For a 2×2 determinant, expand it and differentiate term by term. For the general case, use the fact that determinants are multilinear in their rows.

Solution: By the product rule on the bilinear expansion of the 2×2 determinant,

$$\frac{d}{dx} \det \begin{pmatrix} f_1 & f_2 \\ g_1 & g_2 \end{pmatrix} = \det \begin{pmatrix} f_1' & f_2' \\ g_1 & g_2 \end{pmatrix} + \det \begin{pmatrix} f_1 & f_2 \\ g_1' & g_2' \end{pmatrix}.$$

For an $n \times n$ determinant, multilinearity in the rows gives $\big(\det F\big)' = \sum_{j=1}^{n} \det(F \text{ with the } j\text{th row differentiated})$. ■

5.9: Wronskian and Linear Dependence

Given n functions $f_1, \ldots, f_n$, each having nth order derivatives in (a,b). A function W, called the Wronskian of $f_1, \ldots, f_n$, is defined as follows: For each x in (a,b), $W(x)$ is the value of the determinant of order n whose element in the kth row and mth column is $f_m^{(k-1)}(x)$, where $k = 1, 2, \ldots, n$ and $m = 1, 2, \ldots, n$. [The expression $f_m^{(0)}(x)$ is written for $f_m(x)$.]

(a) Show that $W'(x)$ can be obtained by replacing the last row of the determinant defining $W(x)$ by the nth derivatives $f_1^{(n)}(x), \ldots, f_n^{(n)}(x)$.

(b) Assuming the existence of n constants $c_1, \ldots, c_n$, not all zero, such that $c_1 f_1(x) + \cdots + c_n f_n(x) = 0$ for every x in (a,b), show that $W(x) = 0$ for each x in (a,b).

NOTE. A set of functions satisfying such a relation is said to be a linearly dependent set on (a,b).

(c) The vanishing of the Wronskian throughout (a,b) is necessary, but not sufficient, for linear dependence of $f_1, \ldots, f_n$. Show that in the case of two functions, if the Wronskian vanishes throughout (a,b) and if one of the functions does not vanish in (a,b), then they form a linearly dependent set in (a,b).

Strategy: For (a), use the result from Problem 5.8 about differentiating determinants. For (b), use the linear dependence relation to show that one row is a linear combination of the others. For (c), use the quotient rule to show that the ratio of the two functions is constant.

Solution:

(a) Differentiate the determinant by the rule in 5.8: only the last row changes to $\big(f_1^{(n)}(x), \ldots, f_n^{(n)}(x)\big)$.

(b) If $\sum c_m f_m \equiv 0$ with some c_m not all 0, then each row of the Wronskian is a linear combination of the others with the same coefficients, so the determinant vanishes identically.

(c) For two functions, $W = f_1 f_2' - f_1' f_2 \equiv 0$ on (a,b) and, say, $f_2 \neq 0$ on (a,b). Then $(f_1/f_2)' = \dfrac{f_1' f_2 - f_1 f_2'}{f_2^2} = 0$, so f_1/f_2 is constant and the pair is linearly dependent.

■

5.II The Mean-Value Theorem

Definitions and Theorems

Theorem: Rolle's Theorem

If f is continuous on $[a, b]$, differentiable on (a, b), and $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$.

Theorem: Mean Value Theorem

If f is continuous on $[a, b]$ and differentiable on (a, b), then there exists $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Theorem: Cauchy's Mean Value Theorem

If f and g are continuous on $[a, b]$ and differentiable on (a, b), and $g'(x) \neq 0$ for all $x \in (a, b)$, then there exists $c \in (a, b)$ such that

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}$$

Theorem: Generalized Mean Value Theorem

If f has a finite nth derivative in $[a, b]$ and $f^{(k)}(a) = 0$ for $k = 0, 1, \ldots, n - 1$, then there exists $c \in (a, b)$ such that

$$f^{(n)}(c) = \frac{n!}{(b - a)^n} f(b)$$

Theorem: Taylor's Theorem with Lagrange Remainder

If f has a finite $(n+1)$th derivative in $[a,b]$, then for any $x \in [a,b]$,

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

where c is some point between a and x.

Theorem: Taylor's Theorem with Cauchy Remainder

If f has a finite $(n+1)$th derivative in $[a,b]$, then for any $x \in [a,b]$,

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{(x-a)(x-c)^n}{n!} f^{(n+1)}(c)$$

where c is some point between a and x.

Theorem: L'Hôpital's Rule

If f and g are differentiable on (a,b) except possibly at $c \in (a,b)$, $\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0$ or $\pm\infty$, and $\lim_{x\to c} \frac{f'(x)}{g'(x)}$ exists, then

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}$$

5.10: Infinite Limit and Derivative

Given a function f defined and having a finite derivative in (a,b) and such that $\lim_{x\to b^-} f(x) = +\infty$. Prove that $\lim_{x\to b^-} f'(x)$ either fails to exist or is infinite.

Strategy: Use proof by contradiction. Assume the limit of f' exists and is finite, then use the Mean Value Theorem to derive a contradiction with the fact that $f(x) \to +\infty$ as $x \to b^-$.

Solution: Suppose $\lim_{x\to b^-} f(x) = +\infty$ and $\lim_{x\to b^-} f'(x) = L \in \mathbb{R}$. Fix $h > 0$ small, pick x close to b; by the mean value theorem there is $\xi \in (x, b)$ with $\dfrac{f(b-h)-f(x)}{b-h-x} = f'(\xi)$. Letting $x \to b^-$ forces the left side to $-\infty$ while the right tends to L, a contradiction. Hence the limit of f' cannot be finite; it either diverges or fails to exist. ■

5.11: Mean-Value Theorem and Theta

Show that the formula in the Mean-Value Theorem can be written as follows:

$$\frac{f(x+h)-f(x)}{h} = f'(x+\theta h),$$

where $0 < \theta < 1$. Determine θ as a function of x and h when

(a) $f(x) = x^2$,

(b) $f(x) = x^3$,

(c) $f(x) = e^x$,

(d) $f(x) = \log x, \quad x > 0$.

Keep $x \neq 0$ fixed, and find $\lim_{h\to 0} \theta$ in each case.

Strategy: Use the Mean Value Theorem to express the difference quotient in terms of the derivative at an intermediate point. For each specific function, compute the difference quotient and solve for θ in terms of x and h.

Solution: By the mean value theorem, for each $h \neq 0$ there is $\theta \in (0,1)$ with $\dfrac{f(x+h)-f(x)}{h} = f'(x+\theta h)$. Compute θ casewise:

$$\begin{aligned}
f(x) = x^2: &\frac{(x+h)^2 - x^2}{h} = 2x + h = 2(x+\theta h) \Rightarrow \theta = \tfrac{1}{2}.\\
f(x) = x^3: &\frac{(x+h)^3 - x^3}{h} = 3x^2 + 3xh + h^2 = 3(x+\theta h)^2,\\
&\text{so } \theta = \frac{-x + \sqrt{x^2 + xh + \frac{1}{3}h^2}}{h} \in (0,1).\\
f(x) = e^x: &\frac{e^{x+h} - e^x}{h} = e^{x+\theta h} \Rightarrow \theta = \frac{1}{h}\log\frac{e^h - 1}{h}.\\
f(x) = \log x\,(x>0): &\frac{\log(x+h) - \log x}{h} = \frac{1}{x+\theta h}\\
&\Rightarrow \theta = \frac{1}{h}\Big(\frac{h}{\log(1+h/x)} - x\Big).
\end{aligned}$$

Fix $x \neq 0$. In each case, expanding for small h shows $\lim_{h\to 0}\theta = \frac{1}{2}$. ■

5.12: Cauchy's Mean-Value Theorem

Take $f(x) = 3x^4 - 2x^3 - x^2 + 1$ and $g(x) = 4x^3 - 3x^2 - 2x$ in Theorem 5.20. Show that $f'(x)/g'(x)$ is never equal to the quotient $[f(1) - f(0)]/[g(1) - g(0)]$ if $0 < x \leq 1$. How do you reconcile this with the equation

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_1)}{g'(x_1)}, \quad a < x_1 < b,$$

obtainable from Theorem 5.20 when $n = 1$?

Strategy: Compute the specific values and show that the ratio form of Cauchy's theorem fails when $g'(x_1) = 0$. The correct interpretation is that the cross-product form $(f(b) - f(a))g'(x_1) = (g(b) - g(a))f'(x_1)$ holds.

Solution: Compute $f(1) - f(0) = 0$ and $g(1) - g(0) \neq 0$, hence the quotient is 0. On $(0,1]$, $f'(x) = 2x(6x^2 - 3x - 1)$ and $g'(x) = 2(6x^2 -$

$3x - 1)$ vanish at the same point $x_0 = \dfrac{3+\sqrt{33}}{12} \in (0,1]$, so $f'(x)/g'(x)$ is never 0 for $x \in (0,1]$. This does not contradict Cauchy's theorem: the correct conclusion is $(f(1) - f(0))g'(x_1) = (g(1) - g(0))f'(x_1)$ for some $x_1 \in (0,1)$, which holds at $x_1 = x_0$ (both sides are 0). The "ratio" form fails there because $g'(x_1) = 0$. ■

5.13: Special Cases of Mean-Value Theorem

In each of the following special cases of Theorem 5.20, take $n = 1$, $c = a$, $x = b$, and show that $x_1 = (a+b)/2$.

(a) $f(x) = \sin x, \quad g(x) = \cos x;$

(b) $f(x) = e^x, \quad g(x) = e^{-x}.$

Can you find a general class of such pairs of functions f and g for which x_1 will always be $(a+b)/2$ and such that both examples (a) and (b) are in this class?

Strategy: Apply Cauchy's Mean Value Theorem to each pair and use trigonometric identities or exponential properties to show that the intermediate point must be the midpoint. Look for functions that satisfy certain differential equations.

Solution: For (a), with $f = \sin$, $g = \cos$, Cauchy's theorem gives $(\sin b - \sin a)\,(-\sin x_1) = (\cos b - \cos a)\,\cos x_1$. Using sum-to-product identities this reduces to $\sin\left(\frac{a+b}{2} - x_1\right) = 0$, hence $x_1 = \frac{a+b}{2}$. For (b), $f = e^x$, $g = e^{-x}$ yields $e^b - e^a = (e^{-a} - e^{-b})e^{2x_1}$, whence $x_1 = \frac{a+b}{2}$. A general class: pairs f, g solving a linear ODE $y'' + \lambda y = 0$ (e.g., sin, cos) or $y'' - \lambda y = 0$ (e.g., e^x, e^{-x}) have this midpoint property. ■

5.14: Limit of a Sequence

Given a function f defined and having a finite derivative f' in the half-open interval $0 < x \le 1$ and such that $|f'(x)| < 1$. Define $a_n = f(1/n)$

for $n = 1, 2, 3, \ldots$, and show that $\lim_{n\to\infty} a_n$ exists. Hint. Cauchy condition.

Strategy: Use the Mean Value Theorem to bound the difference between terms in the sequence, then apply the Cauchy criterion for convergence.

Solution: For m, n, by the mean value theorem there is ξ between $1/m$ and $1/n$ with

$$|a_m - a_n| = |f(1/m) - f(1/n)| \leq |f'(\xi)| \left|\frac{1}{m} - \frac{1}{n}\right| \leq \alpha \left|\frac{1}{m} - \frac{1}{n}\right|,$$

for some $\alpha < 1$. Hence (a_n) is Cauchy, so $\lim a_n$ exists. ■

5.15: Limit of Derivative

Assume that f has a finite derivative at each point of the open interval (a, b). Assume also that $\lim_{x\to c} f'(x)$ exists and is finite for some interior point c. Prove that the value of this limit must be $f'(c)$.

Strategy: Use Cauchy's Mean Value Theorem to relate the difference quotient to the derivative at an intermediate point, then take the limit as $x \to c$.

Solution: We have

$$\frac{f(x) - f(c)}{x - c} - f'(x) = \frac{f(x) - f(c) - (x - c)f'(x)}{x - c}.$$

By Cauchy's mean value theorem applied to $F(t) = f(t) - f(c) - (t - c)f'(x)$ and $G(t) = t - c$, there is ξ between x and c such that the quotient equals $\dfrac{f'(\xi) - f'(x)}{1}$. Letting $x \to c$ gives $\dfrac{f(x) - f(c)}{x - c} \to L$, hence $f'(c) = L$. ■

5.16: Extension of Derivative

Let f be continuous on (a, b) with a finite derivative f' everywhere in (a, b), except possibly at c. If $\lim_{x\to c} f'(x)$ exists and has the value A, show that $f'(c)$ must also exist and have the value A.

Strategy: Use the Mean Value Theorem to relate the difference quotient to the derivative at an intermediate point, then use the continuity of f and the limit of f' to show that the difference quotient tends to A.

Solution: As in 5.15, for $x \neq c$ choose ξ between x and c to get

$$\frac{f(x) - f(c)}{x - c} - A = f'(\xi) - A.$$

Let $x \to c$; then $\xi \to c$ and $f'(\xi) \to A$ by hypothesis, so the difference quotient tends to A. Thus $f'(c)$ exists and equals A. ■

5.17: Monotonicity of Quotient

Let f be continuous on $[0, 1]$, $f(0) = 0$, $f'(x)$ finite for each x in $(0, 1)$. Prove that if f' is an increasing function on $(0, 1)$, then so too is the function g defined by the equation $g(x) = f(x)/x$.

Strategy: Use Cauchy's Mean Value Theorem to compare $g(v) - g(u)$ for $0 < u < v \leq 1$, and use the fact that f' is increasing to show that this difference is nonnegative.

Solution: For $0 < u < v \leq 1$, apply Cauchy's mean value theorem to f and $x \mapsto x$ on $[u, v]$ to get

$$\frac{f(v) - f(u)}{v - u} = f'(\xi) \quad (\xi \in (u, v)).$$

Then

$$\begin{aligned}\frac{f(v)}{v} - \frac{f(u)}{u} &= \frac{uf(v) - vf(u)}{uv} = \frac{u[vf'(\xi) - (f(v) - f(u))]}{uv} \\ &= \frac{u(v - \xi)}{uv}\,[f'(\xi) - f'(\eta)] \geq 0,\end{aligned}$$

using the mean value theorem on f again and the monotonicity of f'. Hence $g(x) = f(x)/x$ is increasing. ■

5.18: Rolle's Theorem Application

Assume f has a finite derivative in (a, b) and is continuous on $[a, b]$ with $f(a) = f(b) = 0$. Prove that for every real λ there is some c in (a, b) such that $f'(c) = \lambda f(c)$. Hint. Apply Rolle's theorem to $g(x)f(x)$ for a suitable g depending on λ.

Strategy: Choose $g(x) = e^{-\lambda x}$ so that $g(x)f(x)$ has the same zeros as f, then apply Rolle's theorem to find a point where the derivative of this product vanishes.

Solution: Fix $\lambda \in \mathbb{R}$ and set $g(x) = e^{-\lambda x}$. Then $(gf)(a) = (gf)(b) = 0$. By Rolle's theorem there is $c \in (a, b)$ with $(gf)'(c) = 0$, i.e., $-\lambda e^{-\lambda c} f(c) + e^{-\lambda c} f'(c) = 0$, so $f'(c) = \lambda f(c)$. ■

5.19: Second Derivative and Secant Line

Assume f is continuous on $[a, b]$ and has a finite second derivative f'' in the open interval (a, b). Assume that the line segment joining the points $A = (a, f(a))$ and $B = (b, f(b))$ intersects the graph of f in a third point P different from A and B. Prove that $f''(c) = 0$ for some c in (a, b).

Strategy: Define $\phi(x) = f(x) - \ell(x)$ where ℓ is the secant line. Then ϕ has three zeros, so by Rolle's theorem applied twice, there must be a point where $\phi''(c) = 0$, which implies $f''(c) = 0$.

Solution: Let ℓ be the secant line through $(a, f(a))$ and $(b, f(b))$, and $\phi = f - \ell$. Then $\phi(a) = \phi(b) = \phi(p) = 0$. By Rolle's theorem, there exist $u \in (a, p)$ and $v \in (p, b)$ with $\phi'(u) = \phi'(v) = 0$. Applying Rolle again to ϕ' on $[u, v]$ yields $c \in (u, v)$ with $\phi''(c) = 0$, hence $f''(c) = 0$. ■

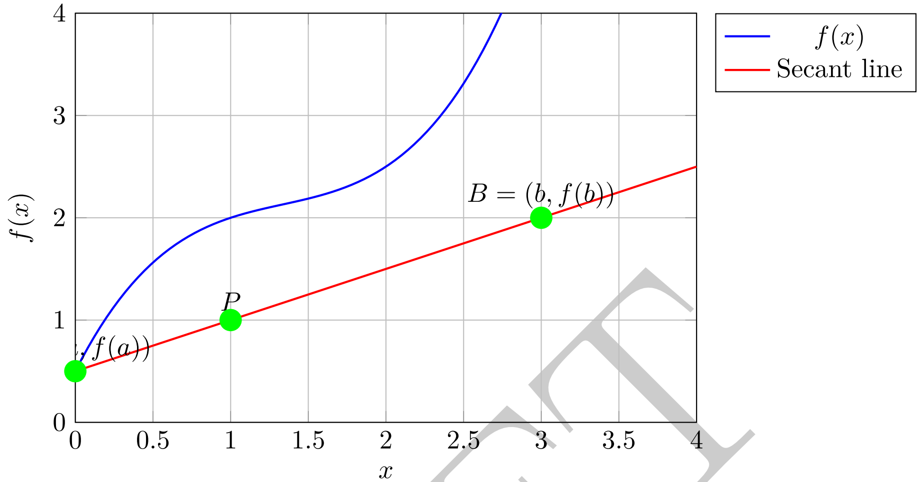


Figure 5.6: Problem 5.19: A cubic function and its secant line intersecting in three points. The secant line through points A and B intersects the graph of f at a third point P, illustrating the geometric condition that leads to $f''(c) = 0$ for some $c \in (a, b)$.

5.20: Third Derivative Condition

If f has a finite third derivative f'' in $[a, b]$ and if

$$f(a) = f'(a) = f(b) = f'(b) = 0,$$

prove that $f''(c) = 0$ for some c in (a, b).

Strategy: Use Rolle's theorem twice: first to find a point where f' vanishes (since $f(a) = f(b) = 0$), then apply Rolle's theorem to f' on an appropriate subinterval.

Solution: From $f(a) = f(b) = 0$, there exists $s \in (a, b)$ with $f'(s) = 0$. Since also $f'(a) = f'(b) = 0$, applying Rolle to f' on $[a, s]$ (or $[s, b]$) gives c with $f''(c) = 0$. ■

5.21: Nonnegative Function with Zeros

Assume f is nonnegative and has a finite third derivative f'' in the open interval (0, 1). If $f(x) = 0$ for at least two values of x in (0, 1), prove that $f''(c) = 0$ for some c in (0, 1).

Strategy: Since f is nonnegative and has zeros at interior points, the derivative must also vanish at these points. Then apply the result from Problem 5.20 to the subinterval between the zeros.

Solution: Let $u < v$ be two zeros of f in $(0, 1)$. Because $f \geq 0$ and $f(u) = 0$ at an interior point, necessarily $f'(u) = 0$; similarly $f'(v) = 0$. Apply 5.20 on $[u, v]$ to conclude that $f''(c) = 0$ for some $c \in (u, v) \subset (0, 1)$. ■

5.22: Behavior at Infinity

Assume f has a finite derivative in some interval $(a, +\infty)$.

(a) If $f(x) \to 1$ and $f'(x) \to c$ as $x \to +\infty$, prove that $c = 0$.

(b) If $f'(x) \to 1$ as $x \to +\infty$, prove that $f(x)/x \to 1$ as $x \to +\infty$.

(c) If $f'(x) \to 0$ as $x \to +\infty$, prove that $f(x)/x \to 0$ as $x \to +\infty$.

Strategy: For (a), use the Mean Value Theorem to relate the difference quotient to the derivative. For (b) and (c), consider the function $g(x) = f(x) - x$ or use integration to bound the growth of f.

Solution:

(a) If $f(x) \to 1$ and $f'(x) \to c$, then for fixed $h > 0$ and large x, $\dfrac{f(x+h) - f(x)}{h} \to c$ by the mean value theorem, while the numerator $\to 0$. Hence $c = 0$.

(b) Let $g(x) = f(x) - x$. Then $g'(x) = f'(x) - 1 \to 0$. For any $\varepsilon > 0$, for large x, $|g'(t)| < \varepsilon$ for $t \geq x$, so $|g(t) - g(x)| \leq \varepsilon |t - x|$. Taking $t = x$ and $t = 2x$ shows $|f(2x) - 2f(x)| \leq \varepsilon x$, which implies $\lim_{x\to\infty} f(x)/x = 1$.

(c) Similarly, if $f'(x) \to 0$, then for large x, $|f(x)-f(0)| \leq \int_0^x |f'(t)|dt \leq \varepsilon x + C$, so $|f(x)/x| \leq \varepsilon + C/x \to 0$.

■

5.23: Nonexistence of Function

Let h be a fixed positive number. Show that there is no function f satisfying the following three conditions: $f'(x)$ exists for $x \geq 0$, $f'(0) = 0$, $f'(x) \geq h$ for $x > 0$.

Strategy: Use proof by contradiction. Assume such a function exists, then use the Mean Value Theorem to show that the difference quotient at 0 must be at least h, contradicting $f'(0) = 0$.

Solution: If $f'(0) = 0$ and $f'(x) \geq h > 0$ for $x > 0$, then for $x > 0$, by the mean value theorem there is $\xi \in (0, x)$ with $\dfrac{f(x) - f(0)}{x - 0} = f'(\xi) \geq h$. Thus $\liminf_{x \downarrow 0} \dfrac{f(x) - f(0)}{x} \geq h$, contradicting $f'(0) = 0$. ■

5.24: Symmetric Difference Quotients

If $h > 0$ and if $f'(x)$ exists (and is finite) for every x in $(a-h, a+h)$, and if f is continuous on $[a-h, a+h]$, show that we have:

(a) $\frac{f(a+h)-f(a-h)}{h} = f'(a+\theta h) + f'(a-\theta h), \quad 0 < \theta < 1;$

(b) $\frac{f(a+h)-2f(a)+f(a-h)}{h} = f'(a+\lambda h) - f'(a-\lambda h), \quad 0 < \lambda < 1.$

(c) If $f''(a)$ exists, show that

$$f''(a) = \lim_{h \to 0} \frac{f(a+h) - 2f(a) + f(a-h)}{h^2}.$$

(d) Give an example where the limit of the quotient in (c) exists but where $f''(a)$ does not exist.

Strategy: For (a), define $\phi(t) = f(t) - f(2a - t)$ and apply the Mean Value Theorem. For (b), apply (a) to f'. For (c), use the result from (b) and take the limit. For (d), use a function with a corner at a.

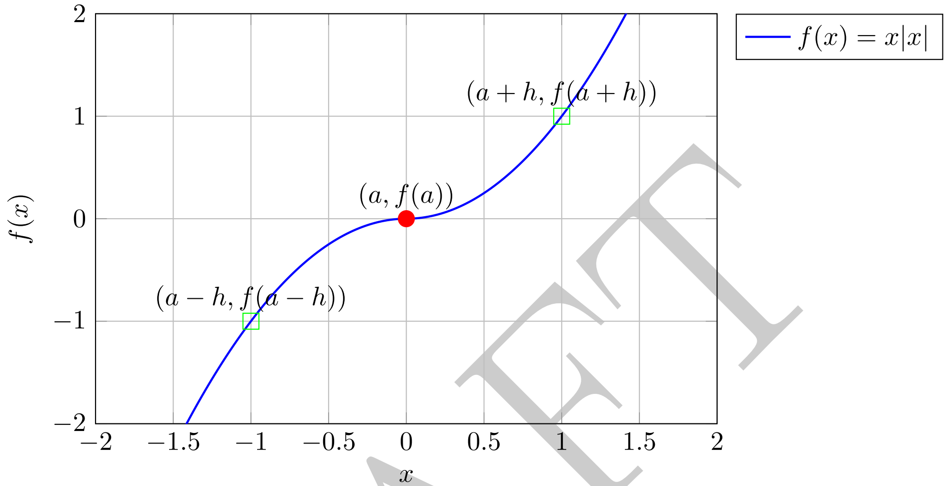


Figure 5.7: Problem 5.24: The function $f(x) = x|x|$ showing symmetric difference quotients. This function has a corner at $x = 0$ where the second derivative does not exist, but the symmetric second difference quotient has a limit of 0.

Solution:

(a) Define $\phi(t) = f(t) - f(2a - t)$. Then ϕ is differentiable and by the mean value theorem there is $\theta \in (0, 1)$ with

$$\frac{f(a + h) - f(a - h)}{h} = \phi'(a + \theta h) = f'(a + \theta h) + f'(a - \theta h).$$

(b) Apply (a) to f' to get

$$\frac{f'(a + h) - f'(a - h)}{1} = \frac{f(a + h) - 2f(a) + f(a - h)}{h} = f'(a+\lambda h) - f'(a-\lambda h).$$

(c) If $f''(a)$ exists, by (b) the symmetric second difference quotient tends to $f''(a)$.

(d) Let $f(x) = x|x|$. Then $f''(0)$ does not exist, but

$$\frac{f(h) - 2f(0) + f(-h)}{h^2} = \frac{h|h| + (-h)| - h|}{h^2} = 0 \to 0.$$

■

5.25: Uniform Differentiability

Let f have a finite derivative in (a,b) and assume that $c \in (a,b)$. Consider the following condition: For every $\varepsilon > 0$ there exists a 1-ball $B(c;\delta)$, whose radius δ depends only on ε and not on c, such that if $x \in B(c;\delta)$, and $x \neq c$, then

$$\left| \frac{f(x)-f(c)}{x-c} - f'(c) \right| < \varepsilon.$$

Show that f' is continuous on (a,b) if this condition holds throughout (a,b).

Strategy: Use the uniform condition to show that f' satisfies the Cauchy criterion for continuity. Fix c and ε, then use the condition to bound $|f'(x) - f'(y)|$ for nearby points x and y.

Solution: Fix c and $\varepsilon > 0$. By hypothesis choose δ (depending only on ε) so that for all $x, y \in (a,b)$ with $0 < |x-c|, |y-c| < \delta$,

$$\left| \frac{f(x)-f(c)}{x-c} - f'(c) \right| < \tfrac{\varepsilon}{2}, \quad \left| \frac{f(y)-f(c)}{y-c} - f'(c) \right| < \tfrac{\varepsilon}{2}.$$

Then

$$|f'(x) - f'(y)| \le \left| \frac{f(x)-f(c)}{x-c} - f'(c) \right| + \left| \frac{f(y)-f(c)}{y-c} - f'(c) \right| < \varepsilon.$$

Thus f' is Cauchy (hence continuous) at c. Since c was arbitrary, f' is continuous on (a,b). ■

5.26: Fixed Point Theorem

Assume f has a finite derivative in (a, b) and is continuous on $[a, b]$, with $a \leq f(x) \leq b$ for all x in $[a, b]$ and $|f'(x)| \leq \alpha < 1$ for all x in (a, b). Prove that f has a unique fixed point in $[a, b]$.

Strategy: Use the Mean Value Theorem to show that f is a contraction mapping, then apply the contraction mapping theorem or use iteration to find the fixed point.

Solution: For $x, y \in [a, b]$, by the mean value theorem there exists ξ between x and y with $|f(x) - f(y)| = |f'(\xi)||x - y| \leq \alpha|x - y|$. Thus f is a contraction of the complete metric space $[a, b]$, so it has a unique fixed point by the contraction mapping theorem. Alternatively, iterate $x_{n+1} = f(x_n)$ to get a Cauchy sequence converging to the unique fixed point. ■

5.27: L'Hôpital's Rule Counterexample

Give an example of a pair of functions f and g having finite derivatives in (0, 1), such that

$$\lim_{x \to 0} \frac{f(x)}{g(x)} = 0,$$

but such that $\lim_{x \to 0} f'(x)/g'(x)$ does not exist, choosing g so that $g'(x)$ is never zero.

Strategy: Use a function like $f(x) = x^2 \sin(1/x)$ and $g(x) = x$. The quotient $f(x)/g(x)$ tends to 0, but $f'(x)/g'(x)$ oscillates and has no limit.

Solution: Let $g(x) = x$ and $f(x) = x^2 \sin(1/x)$ for $x \in (0, 1)$, with $f(0) = g(0) = 0$. Then $f(x)/g(x) = x \sin(1/x) \to 0$, while $f'(x)/g'(x) = \big(2x \sin(1/x) - \cos(1/x)\big)/1$ has no limit as $x \to 0$ and $g'(x) = 1 \neq 0$. ■

5.28: Generalized L'Hôpital's Rule

Prove the following theorem: Let f and g be two functions having finite nth derivatives in (a, b). For some interior point c in (a, b), assume that $f(c) = f'(c) = \cdots = f^{(n-1)}(c) = 0$, and that $g(c) = g'(c) = \cdots = g^{(n-1)}(c) = 0$, but that $g^{(n)}(x)$ is never zero in (a, b). Show that

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \frac{f^{(n)}(c)}{g^{(n)}(c)}.$$

NOTE. $f^{(n)}$ and $g^{(n)}$ are not assumed to be continuous at c.

Strategy: Use Taylor's theorem with remainder to express $f(x)$ and $g(x)$ in terms of their nth derivatives at intermediate points, then take the limit as $x \to c$.

Solution: Apply Cauchy's mean value theorem repeatedly or use Taylor's theorem with remainder about c. Since $f^{(k)}(c) = g^{(k)}(c) = 0$ for $k < n$, we have

$$f(x) = \frac{f^{(n)}(\xi_x)}{n!}(x-c)^n, \quad g(x) = \frac{g^{(n)}(\eta_x)}{n!}(x-c)^n$$

for some $\xi_x, \eta_x \to c$. Hence

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f^{(n)}(\xi_x)}{g^{(n)}(\eta_x)} = \frac{f^{(n)}(c)}{g^{(n)}(c)},$$

using the assumed nonvanishing of $g^{(n)}$ near c. ■

5.29: Taylor's Theorem with Remainder

Show that the formula in Taylor's theorem can also be written as follows:

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(c)}{k!}(x-c)^k + \frac{(x-c)(x-x_1)^{n-1}}{(n-1)!} f^{(n)}(x_1),$$

where x_1 is interior to the interval joining x and c. Let $1 - \theta = (x - x_1)/(x - c)$. Show that $0 < \theta < 1$ and deduce the following form of the remainder term (due to Cauchy):

$$\frac{(1-\theta)^{n-1}(x-c)^n}{(n-1)!} f^{(n)}[\theta x + (1-\theta)c].$$

Strategy: Start with Cauchy's form of the remainder in Taylor's theorem, then use the relation $x - x_1 = \theta(x - c)$ to rewrite the remainder in terms of θ.

Solution: By Cauchy's form of the remainder, for some x_1 between x and c,

$$f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(c)}{k!}(x-c)^k = \frac{f^{(n)}(x_1)}{(n-1)!}(x-c)^{n-1}(x-x_1).$$

Writing $x - x_1 = \theta(x - c)$ with $\theta \in (0, 1)$ gives the displayed form and yields Cauchy's remainder

$$\frac{(1-\theta)^{n-1}(x-c)^n}{(n-1)!} f^{(n)}\big(\theta x + (1-\theta)c\big).$$

■

5.III Vector-valued functions

Definitions and Theorems

Definition: Vector-Valued Function

A vector-valued function is a function $f : \mathbb{R} \to \mathbb{R}^n$ that maps real numbers to vectors in n-dimensional space.

Definition: Vector-Valued Derivative

A vector-valued function f is differentiable at c if there exists a vector $f'(c)$ such that

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

Theorem: Component-wise Differentiation

A vector-valued function $f = (f_1, f_2, \ldots, f_n)$ is differentiable at c if and only if each component function f_i is differentiable at c, and

$$f'(c) = (f_1'(c), f_2'(c), \ldots, f_n'(c))$$

Theorem: Dot Product Rule

If f and g are differentiable vector-valued functions, then

$$\frac{d}{dt}(f(t) \cdot g(t)) = f'(t) \cdot g(t) + f(t) \cdot g'(t)$$

Theorem: Constant Norm Implies Orthogonality

If a vector-valued function f has constant norm $\|f(t)\| = C$ for all t, then $f(t) \cdot f'(t) = 0$ for all t.

5.30: Vector-Valued Differentiability

If a vector-valued function f is differentiable at c, prove that

$$f'(c) = \lim_{h\to 0} \frac{1}{h}[f(c+h) - f(c)].$$

Conversely, if this limit exists, prove that f is differentiable at c.

Strategy: For the forward direction, use the definition of differentiability. For the converse, define $f'(c)$ as the limit and show that it satisfies the definition of differentiability using the ε-δ argument.

Solution: If f is differentiable at c, the definition gives the limit. Conversely, if the limit exists, define $f'(c)$ to be that limit; the standard ε–δ argument shows $\|f(c+h)-f(c)-f'(c)h\| = o(|h|)$, i.e., differentiability at c. ■

5.31: Constant Norm and Orthogonality

A vector-valued function f is differentiable at each point of (a, b) and has constant norm $\|f\|$. Prove that $f(t) \cdot f'(t) = 0$ on (a, b).

Strategy: Differentiate the equation $\|f(t)\|^2 = f(t) \cdot f(t)$ using the product rule for the dot product, then use the fact that the derivative of a constant is zero.

Solution: Differentiate $\|f(t)\|^2 = f(t) \cdot f(t)$ to get $\frac{d}{dt}\|f(t)\|^2 = 2f(t) \cdot f'(t)$. Since the left side is 0, we obtain $f(t) \cdot f'(t) = 0$ on (a, b). ■

5.32: Solution to Differential Equation

A vector-valued function f is never zero and has a derivative f' which exists and is continuous on $\mathbb{R}$. If there is a real function λ such that $f'(t) = \lambda(t)f(t)$ for all t, prove that there is a positive real function u and a constant vector c such that $f(t) = u(t)c$ for all t.

Strategy: Solve the scalar differential equation $u'(t) = \lambda(t)u(t)$ to find $u(t)$, then define $c = f(t)/u(t)$ and show that c is constant by differentiating this quotient.

Solution: Let u solve $u'(t) = \lambda(t)u(t)$ with $u(t_0) = 1$; then $u(t) = \exp\left(\int_{t_0}^{t} \lambda\right) > 0$. Define $c = f/u$. Then $c'(t) = \dfrac{f'(t)u(t) - f(t)u'(t)}{u(t)^2} = 0$, so c is constant and $f(t) = u(t)\,c$. ■

5.IV Partial Derivatives

Definitions and Theorems

Definition: Partial Derivative

The partial derivative of a function $f(x_1, x_2, \ldots, x_n)$ with respect to x_i at a point $(a_1, a_2, \ldots, a_n)$ is defined as

$$\frac{\partial f}{\partial x_i}(a_1, \ldots, a_n) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_i + h, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}$$

if this limit exists.

Definition: Directional Derivative

The directional derivative of f at a point a in the direction of a unit vector u is defined as

$$D_u f(a) = \lim_{h \to 0} \frac{f(a + hu) - f(a)}{h}$$

if this limit exists.

Definition: Gradient

The gradient of a function f at a point a is the vector of partial derivatives:

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right)$$

Theorem: Mixed Partial Derivatives

If the mixed partial derivatives $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ are continuous in a neighborhood of a point, then they are equal at that point.

Theorem: Directional Derivative and Gradient

If f is differentiable at a, then the directional derivative in the direction of unit vector u is

$$D_u f(a) = \nabla f(a) \cdot u$$

Theorem: Chain Rule for Partial Derivatives

If $f(x, y)$ is differentiable and $x = x(t)$, $y = y(t)$ are differentiable functions, then

$$\frac{d}{dt} f(x(t), y(t)) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

5.33: Partial Derivatives and Continuity

Consider the function f defined on $\mathbb{R}^2$ by the following formulas:

$$f(x, y) = \frac{xy}{x^2 + y^2} \quad \text{if } (x, y) \neq (0, 0) \quad f(0, 0) = 0.$$

Prove that the partial derivatives $D_1 f(x, y)$ and $D_2 f(x, y)$ exist for every (x, y) in $\mathbb{R}^2$ and evaluate these derivatives explicitly in terms of x and y. Also, show that f is not continuous at (0, 0).

Strategy: For $(x, y) \neq (0, 0)$, use the quotient rule to compute partial derivatives. At the origin, use the definition of partial derivatives. To show discontinuity, find a path along which the limit is not zero.

Solution: For $(x, y) \neq (0, 0)$, compute

$$D_1 f(x, y) = \frac{y(y^2 - x^2)}{(x^2 + y^2)^2}, \quad D_2 f(x, y) = \frac{x(x^2 - y^2)}{(x^2 + y^2)^2}.$$

At the origin,

$$D_1 f(0, 0) = \lim_{h \to 0} \frac{f(h, 0) - 0}{h} = 0, \quad D_2 f(0, 0) = \lim_{h \to 0} \frac{f(0, h) - 0}{h} = 0.$$

However, f is not continuous at 0 since along $y = x \neq 0$, $f(x, x) = \frac{1}{2} \not\to 0$. ■

5.34: Higher-Order Partial Derivatives

Let f be defined on $\mathbb{R}^2$ as follows:

$$f(x,y) = y\frac{x^2-y^2}{x^2+y^2} \quad \text{if } (x,y) \neq (0,0), \quad f(0,0) = 0.$$

Compute the first- and second-order partial derivatives of f at the origin, when they exist.

Strategy: Compute first-order partial derivatives using the quotient rule for $(x,y) \neq (0,0)$ and the definition at the origin. Then compute second-order derivatives by differentiating the first-order derivatives.

Solution: For $(x,y) \neq (0,0)$, one computes

$$f_x(x,y) = \frac{4xy^3}{(x^2+y^2)^2}, \quad f_y(x,y) = \frac{x^4-4x^2y^2-y^4}{(x^2+y^2)^2}.$$

At the origin,

$$f_x(0,0) = \lim_{h\to 0} \frac{f(h,0)-0}{h} = 0, \quad f_y(0,0) = \lim_{h\to 0} \frac{f(0,h)-0}{h} = -1.$$

Second-order at $(0,0)$ (where they exist):

$$f_{xx}(0,0) = \lim_{h\to 0} \frac{f_x(h,0)-f_x(0,0)}{h} = 0,$$

$$f_{yy}(0,0) = \lim_{h\to 0} \frac{f_y(0,h)-f_y(0,0)}{h} = 0,$$

$$f_{xy}(0,0) = \lim_{h\to 0} \frac{f_x(0,h)-f_x(0,0)}{h} = 0,$$

$$f_{yx}(0,0) = \lim_{h\to 0} \frac{f_y(h,0)-f_y(0,0)}{h} \text{ does not exist.}$$

■

5.35: Complex Conjugate Differentiability

Let S be an open set in $\mathbb{C}$ and let S^* be the set of complex conjugates $\bar{z}$, where $z \in S$. If f is defined on S, define g on S^* as follows: $g(\bar{z}) = \overline{f(z)}$, the complex conjugate of $f(z)$. If f is differentiable at c prove that g is differentiable at $\bar{c}$ and that $g'(\bar{c}) = \overline{f'(c)}$.

Strategy: Use the definition of differentiability and the fact that complex conjugation is continuous and linear. Show that the difference quotient for g at $\bar{c}$ is the complex conjugate of the difference quotient for f at c.

Solution: With $g(\bar{z}) = \overline{f(z)}$, for $h \to 0$,

$$\frac{g(\bar{c}+h) - g(\bar{c})}{h} = \overline{\frac{f(c+\bar{h}) - f(c)}{\bar{h}}} \to \overline{f'(c)}.$$

Thus g is differentiable at $\bar{c}$ and $g'(\bar{c}) = \overline{f'(c)}$. ■

5.V Complex-valued functions

Definitions and Theorems

Definition: Complex Derivative

A function $f : \mathbb{C} \to \mathbb{C}$ is differentiable at a point z_0 if the limit

$$f'(z_0) = \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists.

Definition: Holomorphic Function

A function f is holomorphic (analytic) on an open set $U \subseteq \mathbb{C}$ if it is differentiable at every point in U.

Definition: Cauchy-Riemann Equations

For a complex function $f(z) = u(x, y) + iv(x, y)$, the Cauchy-Riemann equations are

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

Theorem: Cauchy-Riemann Criterion

A complex function $f = u + iv$ is differentiable at a point if and only if the partial derivatives of u and v exist, are continuous, and satisfy the Cauchy-Riemann equations at that point.

Theorem: Complex Chain Rule

If f and g are differentiable complex functions, then

$$(f \circ g)'(z) = f'(g(z))g'(z)$$

Theorem: Complex Product Rule

If f and g are differentiable complex functions, then

$$(fg)'(z) = f'(z)g(z) + f(z)g'(z)$$

Theorem: Complex Quotient Rule

If f and g are differentiable complex functions and $g(z) \neq 0$, then

$$\left(\frac{f}{g}\right)'(z) = \frac{f'(z)g(z) - f(z)g'(z)}{g(z)^2}$$

Theorem: Identity Principle

If two holomorphic functions f and g agree on a set with a limit point in their common domain, then $f = g$ throughout their common domain.

5.36: Cauchy-Riemann Equations

i) In each of the following examples write $f = u + iv$ and find explicit formulas for $u(x, y)$ and $v(x, y)$:

(a) $f(z) = \sin z$,

(b) $f(z) = \cos z$,

(c) $f(z) = |z|$,

(d) $f(z) = \bar{z}$,

(e) $f(z) = \arg z \quad (z \neq 0)$,

(f) $f(z) = \log z \quad (z \neq 0)$,

(g) $f(z) = e^{z^2}$,

(h) $f(z) = z^\alpha \quad (\alpha \text{ complex}, z \neq 0)$.

ii) Show that u and v satisfy the Cauchy-Riemann equations for the following values of z: All z in (a), (b), (g); no z in (c), (d), (e); all z except real $z \leq 0$ in (f), (h). (In part (h), the Cauchy-Riemann equations hold for all z if α is a nonnegative integer, and they hold for all $z \neq 0$ if α is a negative integer.)
iii) Compute the derivative $f'(z)$ in (a), (b), (f), (g), (h), assuming it exists.

Strategy: For (i), use standard complex function expansions to separate real and imaginary parts. For (ii), compute partial derivatives and check the Cauchy-Riemann equations $u_x = v_y$ and $u_y = -v_x$. For (iii), use standard differentiation rules for complex functions.

Solution: i) Standard expansions give: $\sin z = \sin x \cosh y + i \cos x \sinh y$, $\cos z = \cos x \cosh y - i \sin x \sinh y$, $|z| = \sqrt{x^2 + y^2}$, $\bar{z} = x - iy$, $\arg z =$

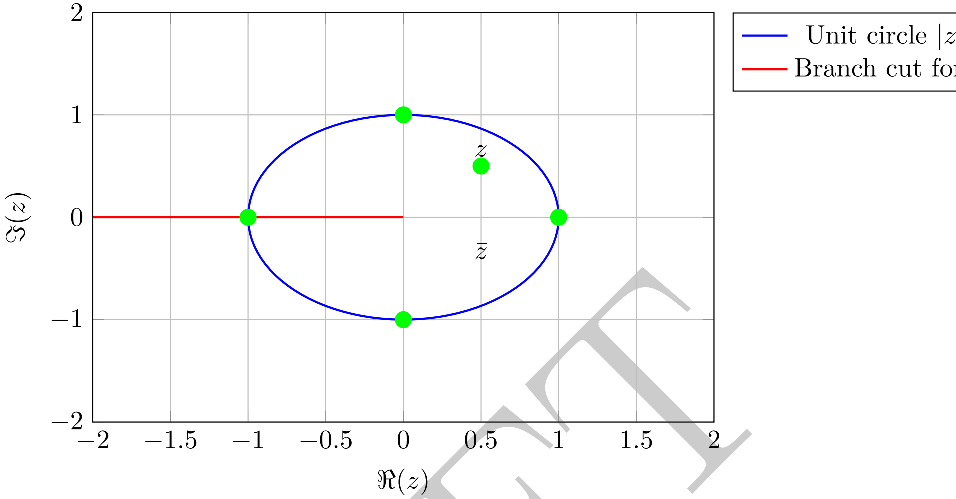


Figure 5.8: Problem 5.36: The complex plane showing important domains and regions. The unit circle represents $|z| = 1$, the red line shows the branch cut for $\log z$ along the negative real axis, and the green points show example complex numbers including a point z and its conjugate $\bar{z}$.

$\arctan(y/x)$ $(z \neq 0)$, $\log z = \ln|z| + i \arg z$ $(z \neq 0)$, $e^{z^2} = e^{x^2-y^2}(\cos 2xy + i \sin 2xy)$, $z^\alpha = e^{\alpha \log z}$.
ii) The Cauchy–Riemann equations hold on the stated sets: everywhere for (a),(b),(g); nowhere for (c),(d),(e); on $\mathbb{C} \setminus (-\infty, 0]$ for branches of log and z^α, with the special cases as indicated.
iii) Where differentiable: $(\sin z)' = \cos z$, $(\cos z)' = -\sin z$, $(\log z)' = 1/z$, $(e^{z^2})' = 2ze^{z^2}$, $(z^\alpha)' = \alpha z^{\alpha-1}$ (on the domain of the chosen branch). ∎

5.37: Constant Function Condition

Write $f = u + iv$ and assume that f has a derivative at each point of an open disk D centered at $(0, 0)$. If $au^2 + bv^2$ is constant on D for some real a and b, not both 0, prove that f is constant on D.

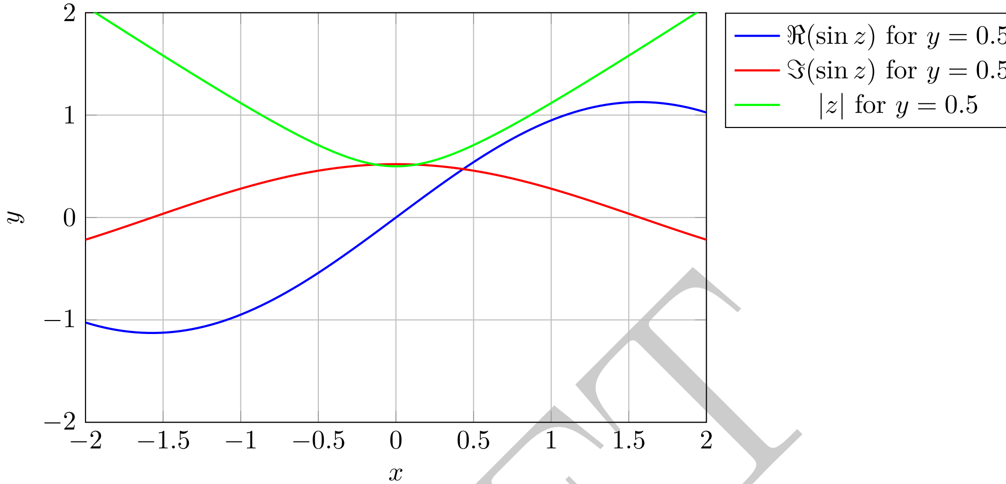


Figure 5.9: Problem 5.36: Real and imaginary parts of complex functions for fixed $y = 0.5$. The blue curve shows $\Re(\sin z) = \sin(x)\cosh(0.5)$, the red curve shows $\Im(\sin z) = \cos(x)\sinh(0.5)$, and the green curve shows $|z| = \sqrt{x^2 + 0.25}$.

Strategy: Differentiate the equation $au^2 + bv^2 = C$ and use the Cauchy-Riemann equations to obtain a system of linear equations. Show that the determinant of this system is non-zero, which implies that the partial derivatives vanish, making f constant.

Solution: Let $f = u + iv$ be complex differentiable on D and suppose $au^2 + bv^2 \equiv C$ with $(a, b) \neq (0, 0)$. Differentiate: $auu_x + bvv_x = 0$ and $auu_y + bvv_y = 0$. Using the Cauchy–Riemann equations $u_x = v_y$, $u_y = -v_x$, we obtain the linear system

$$\begin{pmatrix} au & bv \\ -bv & au \end{pmatrix} \begin{pmatrix} u_x \\ v_x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

On D, the determinant is $a^2u^2 + b^2v^2 \geq 0$. If it is nonzero at a point, then $u_x = v_x = 0$ there; by analyticity, $u_x \equiv v_x \equiv 0$ on the component, hence $u_y = v_y \equiv 0$ and f is constant. If it vanishes at a point, then $u = v = 0$ there; by the identity principle for holomorphic functions, $f \equiv 0$ in a neighborhood. In all cases f is constant on D. ∎

5.VI Solving and Proving Techniques

Proving Limits Exist

- Use the definition of limit with ε-δ arguments
- Apply the squeeze theorem when bounding functions
- Use continuity of known functions (polynomials, exponentials, etc.)
- Apply L'Hôpital's rule for indeterminate forms
- Use Taylor series expansions for complex functions
- Apply the Cauchy criterion for sequences

Proving Differentiability

- Use the definition of derivative and show the limit exists
- Apply known differentiation rules (sum, product, quotient, chain)
- Use the fact that differentiability implies continuity
- Apply the Mean Value Theorem to relate difference quotients
- Use component-wise differentiation for vector-valued functions
- Apply the Cauchy-Riemann equations for complex functions

Using the Mean Value Theorem

- Express difference quotients in terms of derivatives at intermediate points
- Prove existence of points where derivatives take specific values
- Establish bounds on function values using derivative bounds
- Prove monotonicity by analyzing derivative signs
- Show convergence of sequences using derivative bounds
- Prove fixed point theorems using contraction properties

Proving Existence of Points

- Apply Rolle's Theorem when function values are equal at endpoints
- Use the Intermediate Value Theorem for continuous functions
- Apply the Mean Value Theorem to find intermediate points
- Use the Extreme Value Theorem on closed, bounded intervals
- Apply the Bolzano-Weierstrass theorem for sequences
- Use the contraction mapping theorem for fixed points

Working with Derivatives

- Use logarithmic differentiation for products and quotients
- Apply Leibniz's formula for higher derivatives of products
- Use mathematical induction for derivative formulas
- Apply the chain rule for composite functions
- Use the product rule for dot products of vector functions
- Apply the quotient rule and simplify using algebra

Proving Inequalities

- Use the Mean Value Theorem to bound differences
- Apply the triangle inequality for vector functions
- Use the fact that derivatives bound function growth
- Apply the Cauchy-Schwarz inequality when appropriate
- Use the fact that increasing functions preserve inequalities
- Apply the squeeze theorem for limit comparisons

Working with Complex Functions

- Separate real and imaginary parts using standard expansions
- Apply the Cauchy-Riemann equations to check differentiability
- Use the fact that complex conjugation is continuous and linear
- Apply the identity principle for holomorphic functions
- Use the fact that constant norm implies orthogonality with derivative
- Apply standard complex differentiation rules

Proving Uniqueness

- Use the fact that differentiable functions are continuous
- Apply the identity principle for holomorphic functions
- Use the fact that constant functions have zero derivatives
- Apply the contraction mapping theorem for fixed points
- Use the fact that linear independence implies non-zero Wronskian
- Apply the fact that analytic functions are determined by their values on sets with limit points

Using Proof by Contradiction

- Assume the opposite of what you want to prove
- Use the Mean Value Theorem to derive contradictions
- Apply the fact that limits must be unique
- Use the fact that continuous functions preserve connectedness
- Apply the fact that differentiable functions are continuous
- Use the fact that bounded sequences have convergent subsequences

Working with Sequences and Series

- Use the Cauchy criterion to prove convergence
- Apply the Mean Value Theorem to bound sequence differences
- Use the fact that bounded monotone sequences converge
- Apply the squeeze theorem for sequence limits
- Use the fact that convergent sequences are bounded
- Apply the fact that subsequences of convergent sequences converge to the same limit

Proving Continuity

- Use the definition of continuity with ε-δ arguments
- Apply the fact that differentiable functions are continuous
- Use the fact that sums, products, and compositions of continuous functions are continuous
- Apply the fact that uniform limits of continuous functions are continuous
- Use the fact that inverse functions of strictly monotone continuous functions are continuous
- Apply the fact that vector-valued functions are continuous if and only if each component is continuous