AP Analytical Calculus BC+

Daniel Kim

September 2018 - June 2019

This is a comprehensive collection of class notes, example problems, proofs, and other material for AP Analytical Calculus BC+ taught by Dr. Abramson.

Course Outline

§1 Fields

§1.1 Review

Definition 1.1. A **field** is a set F with two operations, typically called $+$ and $\times$, satisfying:

a) **Closure**: $\forall a, b \in F, a + b \in F$ and $a \times b \in F$.

b) **Commutative**: $\forall a, b \in F, a + b = b + a$ and $ab = ba$.

c) **Associative**: $\forall a, b, c \in F, a + (b + c) = (a + b) + c$ and $a(bc) = (ab)c$.

d) **Distributive**: $\forall a, b, c \in F, a(b + c) = ab + ac$.

e) **Identities**: $\exists 0, 1 \in F, (0 \neq 1)$ s.t. $\forall a \in F, a + 0 = 0 + a = a$ and $a \cdot 1 = 1 \cdot a = a$.

f) **Inverses**: $\forall a \in F, \exists -a \in F$ s.t. $a + -a = 0$.
$\forall a \in F, a \neq 0, \exists a^{-1} \in F$ s.t. $a(a^{-1}) = 1$.

Some examples of fields would be the sets $\mathbb{Q}, \mathbb{R}, \mathbb{C}$. The set $\mathbb{Z}_p$ (all integers mod p, where p is prime) is a field; it is a well known result in number theory that all integers mod p do have a multiplicative inverse. However, the set of integers, $\mathbb{Z}$, is not a field because not every integer has an integer multiplicative inverse.

Here are some exercises to help you refresh the topics from last year. Be careful not to skip any steps in your proofs.

Problem 1.2. Prove $a + (b + c) = c + (a + b)$.

Proof.
$$\begin{aligned} a + (b + c) &= (a + b) + c \\ &= c + (a + b). \end{aligned}$$
□

Problem 1.3. Prove $0 \cdot a = 0$.

Proof.
$$\begin{aligned} 0 + 0 &= 0 \\ a(0 + 0) &= a \cdot 0 \\ a \cdot 0 + a \cdot 0 &= a \cdot 0 \\ -a \cdot 0 + (a \cdot 0 + a \cdot 0) &= -a \cdot 0 + a \cdot 0 \\ (-a \cdot 0 + a \cdot 0) + a \cdot 0 &= -a \cdot 0 + a \cdot 0 \\ 0 + a \cdot 0 &= -a \cdot 0 + a \cdot 0 \\ a \cdot 0 &= -a \cdot 0 + a \cdot 0 \\ a \cdot 0 &= 0 \\ \therefore 0 \cdot a &= 0. \end{aligned}$$
□

Problem 1.4. Prove $-1 \cdot a = -a$.

Proof. We will have to use the result $0 \cdot a = 0$ from Problem 1.3.

$$\begin{aligned}
1 + -1 &= 0 \\
a(1 + -1) &= a \cdot 0 \\
a \cdot 1 + a \cdot -1 &= a \cdot 0 \\
a + a \cdot -1 &= a \cdot 0 \\
a + a \cdot -1 &= 0 \cdot a \\
a + a \cdot -1 &= 0 \\
-a + (a + a \cdot -1) &= -a + 0 \\
(-a + a) + a \cdot -1 &= -a + 0 \\
0 + a \cdot -1 &= -a + 0 \\
a \cdot -1 &= -a + 0 \\
a \cdot -1 &= -a \\
\therefore -1 \cdot a &= -a.
\end{aligned}$$

□

Problem 1.5. Prove $(ab)^{-1} = a^{-1}b^{-1}$, provided $b \neq 0$.

Proof. As $b \neq 0$, b^{-1} exists. Then,

$$\begin{aligned}
(ab) \cdot (ab)^{-1} &= 1 \\
a^{-1}((ab) \cdot (ab)^{-1}) &= a^{-1} \cdot 1 \\
a^{-1}((ab) \cdot (ab)^{-1}) &= a^{-1} \\
(a^{-1}(ab)) \cdot (ab)^{-1} &= a^{-1} \\
((a^{-1}a)b) \cdot (ab)^{-1} &= a^{-1} \\
(1 \cdot b) \cdot (ab)^{-1} &= a^{-1} \\
b \cdot (ab)^{-1} &= a^{-1} \\
b^{-1}(b \cdot (ab)^{-1}) &= b^{-1}a^{-1} \\
(b^{-1}b) \cdot (ab)^{-1} &= b^{-1}a^{-1} \\
1 \cdot (ab)^{-1} &= b^{-1}a^{-1} \\
(ab)^{-1} &= b^{-1}a^{-1} \\
\therefore (ab)^{-1} &= a^{-1}b^{-1}.
\end{aligned}$$

□

Alternative Proof. Here is a slightly faster method. First, note that

$$\begin{aligned}
ab(a^{-1}b^{-1}) &= ((ab)a^{-1})b^{-1} \\
&= (a^{-1}(ab))b^{-1} \\
&= ((a^{-1}a)b)b^{-1} \\
&= (a^{-1}a)(bb^{-1}) \\
&= 1 \cdot 1 \\
&= 1.
\end{aligned}$$

□

We can then conclude:

$$ab(a^{-1}b^{-1}) = 1$$

$$\begin{aligned}
(ab)^{-1}(ab(a^{-1}b^{-1})) &= (ab)^{-1} \cdot 1 \\
(ab)^{-1}(ab(a^{-1}b^{-1})) &= (ab)^{-1} \\
((ab)^{-1}ab)(a^{-1}b^{-1}) &= (ab)^{-1} \\
1 \cdot (a^{-1}b^{-1}) &= (ab)^{-1} \\
a^{-1}b^{-1} &= (ab)^{-1} \\
\therefore (ab)^{-1} &= a^{-1}b^{-1}.
\end{aligned}$$

Now consider the equation $x + b = a$, where a, b are constants. We now define a new operation to represent the solution to this special equation, which is $x = a + -b$.

Definition 1.6. *Subtraction* is defined as $a + -b = a - b$.

Likewise, consider $bx = a$. We define another new operation to represent the solution to this special equation, which is $x = ab^{-1}$.

Definition 1.7. *Division* is defined as $ab^{-1} = \dfrac{a}{b}$, provided $b \neq 0$.

Problem 1.8. Prove $\dfrac{a}{b} + \dfrac{c}{d} = \dfrac{ad + bc}{bd}$, provided $b, d \neq 0$.

Proof. As $b, d \neq 0$, b^{-1} and d^{-1} exist. We will also have to use the result from Problem 1.5 to finish the proof.

$$\begin{aligned}
\frac{a}{b} + \frac{c}{d} &= ab^{-1} + cd^{-1} \\
&= 1 \cdot (ab^{-1} + cd^{-1}) \\
&= 1 \cdot (ab^{-1}) + 1 \cdot (cd^{-1}) \\
&= (ab^{-1}) \cdot 1 + (cd^{-1}) \cdot 1 \\
&= (ab^{-1}) \cdot (dd^{-1}) + (cd^{-1}) \cdot 1 \\
&= a \cdot (b^{-1} \cdot (dd^{-1})) + (cd^{-1}) \cdot 1 \\
&= a \cdot ((dd^{-1}) \cdot b^{-1}) + (cd^{-1}) \cdot 1 \\
&= a \cdot (d(d^{-1}b^{-1})) + (cd^{-1}) \cdot 1 \\
&= (ad)(d^{-1}b^{-1}) + (cd^{-1}) \cdot 1 \\
&= (ad)(d^{-1}b^{-1}) + (cd^{-1}) \cdot (bb^{-1}) \\
&= (ad)(d^{-1}b^{-1}) + c \cdot (d^{-1} \cdot (bb^{-1})) \\
&= (ad)(d^{-1}b^{-1}) + c \cdot ((bb^{-1}) \cdot d^{-1}) \\
&= (ad)(d^{-1}b^{-1}) + c \cdot (b \cdot (b^{-1}d^{-1})) \\
&= (ad)(d^{-1}b^{-1}) + (cb)(b^{-1}d^{-1}) \\
&= (ad)(b^{-1}d^{-1}) + (cb)(b^{-1}d^{-1}) \\
&= (ad)(b^{-1}d^{-1}) + (bc)(b^{-1}d^{-1}) \\
&= (ad + bc)(b^{-1}d^{-1}) \\
&= (ad + bc)(bd)^{-1} \\
&= \frac{ad + bc}{bd}.
\end{aligned}$$

□

The following are additional exercises for practice. You should prove these from scratch.

Exercise 1.9. Prove the **Cancellation Theorems**:

1. $a + b = a + c \to b = c$.
2. $ab = ac \wedge a \neq 0 \to b = c$.

Exercise 1.10. Prove that $(a + c) - (b + c) = a - b$.

Exercise 1.11. Prove that additive and multiplicative identities and inverses are unique.

Exercise 1.12. Prove that $ab = 0 \to (a = 0 \vee b = 0)$.

Exercise 1.13. Prove that $x^2 = y^2 \to x = y \vee x = -y$.

Exercise 1.14. Prove that $(-a)(-b) = ab$. (Hint: try proving $(-a) \cdot b = -(ab)$ first.)

Exercise 1.15. Prove that $-(-a) = a$ and $(a^{-1})^{-1} = a$.

Exercise 1.16. Prove that $\left(\frac{b}{d}\right)^{-1} = \frac{d}{b}$.

Exercise 1.17. Prove that $\frac{a}{b} = \frac{c}{d} \longleftrightarrow ad = bc$ provided $b, d \neq 0$.

Exercise 1.18. Prove that $(-1)(-1) = 1$.

§1.2 Ordered Fields

In order to differentiate the reals from other fields, we add more restrictive axioms of the real numbers:

Definition 1.19. A field F is an **ordered field** if $\exists P \subset F$ with the following properties:

OF 1. (Trichotomy) $\forall a \in F$, exactly one of the following is true:

$$a = 0, \quad a \in P, \text{ or } -a \in P.$$

OF 2. $\forall a, b \in P$, $a + b \in P$.

OF 3. $\forall a, b \in P$, $ab \in P$.

For the purpose of detail, these three properties will be referred to as OF1, OF2, and OF3 respectively.

Note that we have constructed this definition such that P is the set of positive numbers. Furthermore, P is a proper subset of F; $P \neq F$ because $0 \notin P$, by OF1. However, this is still not enough to distinguish the reals from the rationals, another field in which the notion of "positive" is well-defined.

Exercise 1.20. What are some fields that cannot be ordered?

Definition 1.21. $a < b \longleftrightarrow b - a \in P$.

Problem 1.22. Prove that $\forall a, b \in F$, exactly one of the following is true:

$$a = b, \quad a < b, \text{ or } a > b.$$

Proof. If $a, b \in F$, then $a - b \in F$. By OF1, exactly one of the following is true:

$$(a-b) = 0, \quad (a-b) \in P, \text{ or } -(a-b) = (b-a) \in P.$$

However, these rearrange to $a = b$, $a < b$, and $a > b$ respectively, the latter two following from Definition 1.21. □

Problem 1.23. Prove that $\forall a, b, c \in F$, $((a < b) \wedge (b < c)) \to (a < c)$.

Proof. If $a < b$ and $b < c$, then by Definition 1.21, $b - a \in P$ and $c - b \in P$. By OF2, $(b-a) + (c-b) = c - a \in P$, i.e. $a < c$. □

Problem 1.24. Prove that $\forall a, b, c \in F$, $a < b \to a + c < b + c$.

Proof. By Definition 1.21, $a < b \longleftrightarrow b - a \in P$. Note that $b - a = (b-a) + 0 = (b-a) + (c-c) = (b+c) - (a+c) \in P$, so $a + c < b + c$. □

Problem 1.25. Prove that $\forall a, b, c \in F$, if $c \in P$, then $a < b \to ac < bc$.

Proof. We have $b - a \in P$ and $c \in P$, so by OF3, $(b-a)c = bc - ac \in P$, i.e. $ac < bc$. □

Definition 1.26. We define the set of negative numbers, N, as follows: $a \in N \longleftrightarrow -a \in P$.

Problem 1.27. Prove that $a, b \in N \to ab \in P, a + b \in N$.

Proof. By Definition 1.26, $a \in N$ and $b \in N$ imply $-a \in P$ and $-b \in P$. By OF3, $(-a)(-b) = ab \in P$. By OF2, $-a + -b = -(a+b) \in P \longleftrightarrow (a+b) \in N$. □

Problem 1.28. Prove that $a \in N, b \in P \to ab \in N$.

Proof. By Definition 1.26, $a \in N \longleftrightarrow -a \in P$. By OF3, $-a \cdot b = -(ab) \in P \longleftrightarrow ab \in N$. □

Problem 1.29. Prove that $1 \in P$.

Proof. Since $1 \in F$, by OF1, exactly one of the following is true:

$$1 = 0, \quad 1 \in P, \text{ or } -1 \in P.$$

We proceed by casework. Note that $1 = 0$ is clearly false by the field axioms. Now assume that $-1 \in P$ is the only true statement. if $-1 \in P$, then by OF3, $(-1)(-1) = 1 \in P$. We cannot have both $1 \in P$ and $-1 \in P$, so we have a contradiction. Therefore, the only remaining statement is $1 \in P$, which is necessarily true based on the definition. □

Alternative Proof. We can also take advantage of a previous result. Take $a \in P$, and therefore $a = a \cdot 1 \in P$. As $1 \in F$, exactly one of the following is true:

$$1 = 0, \quad 1 \in P, \text{ or } -1 \in P.$$

If $1 \in N$, then $a \cdot 1 \in N$ by Problem 1.28. This is a contradiction. Obviously $1 \neq 0$ by the field axioms. Therefore $-1 \in N \longleftrightarrow 1 \in P$. □

Problem 1.30. Prove that if $a \in P$, $a^{-1} \in P$, and if $a \in N$, $a^{-1} \in N$.

Proof. As $a^{-1} \in F$, by OF1, exactly one of the following is true:

$$a^{-1} = 0, \quad a^{-1} \in P, \text{ or } -a^{-1} \in P.$$

Similar to the proof of the previous problem, we proceed by casework. Assume $a^{-1} = 0$ is true. Then $a \cdot a^{-1} = a \cdot 0$, i.e. $1 = 0$, a contradiction.

Assume $-a^{-1} \in P$ is true. Since $a \in P$, by OF3, $a \cdot -a^{-1} = -aa^{-1} = -1 \in P$, which is clearly false as $1 \in P$, by Problem 1.29.

Therefore, $a^{-1} \in P$ must be the only true statement.

By the same reasoning, we can similarly prove $a \in N \to a^{-1} \in N$. □

Problem 1.31. Prove that $\forall a \in F, a \neq 0 \to a^2 \in P$.

Proof. By OF1, exactly one of $a = 0$, $a \in P$, or $-a \in P$ must be true. Clearly $a = 0$ contradicts the given condition $a \neq 0$. If $a \in P$ is true, then by OF3, $a \cdot a = a^2 \in P$. If $-a \in P$, then OF3 still implies $(-a)(-a) = a^2 \in P$. The latter two cases yield the result $a^2 \in P$ regardless, so we are done. □

Alternative Proof. However, we can use a result proven in a previous problem to yield a much simpler proof. As $a \in F$ and $a \neq 0$, then either $a \in P$ or $a \in N$ is true, by OF1. If $a \in P$, then $a \cdot a = a^2 \in P$ by OF3. Otherwise, if $a \in N$, then by Problem 1.27, $a^2 \in P$. □

Problem 1.32. Prove that $a < b \wedge c < d \to a + c < b + d$.

Proof. As $b - a \in P$ and $d - c \in P$, by OF2, $(b - a) + (d - c) = (b + d) - (a + c) \in P$ i.e. $a + c < b + d$. □

Problem 1.33. Prove that $0 < a < b \wedge 0 < c < d \to ac < bd$.

Proof. Note that $0 < a < b$ implies $0 < a$, $0 < b$, and $a < b$, which are respectively equivalent to $a \in P$, $b \in P$, and $b - a \in P$. Similarly, $c \in P$, $d \in P$, and $d - c \in P$. By OF3,

$$\begin{aligned} d \in P, b - a \in P &\to d(b - a) = bd - ad \in P. \\ a \in P, d - c \in P &\to a(d - c) = ad - ac \in P. \end{aligned}$$

Therefore, since $bd - ad \in P$ and $ad - ac \in P$, by OF2, $(bd - ad) + (ad - ac) = bd - ac \in P$, i.e. $ac < bd$. □

Problem 1.34. Prove that $0 < a < b \to \dfrac{1}{b} < \dfrac{1}{a}$.

Proof. Like the previous problem, the condition $0 < a < b$ implies $a \in P$, $b \in P$, and $b - a \in P$. By OF3, $ab \in P$. By Problem 1.30, $(ab)^{-1} \in P$, i.e. $a^{-1}b^{-1} \in P$. By OF3 on $b - a \in P$ and $a^{-1}b^{-1} \in P$, $(b - a)(a^{-1}b^{-1}) = ba^{-1}b^{-1} - aa^{-1}b^{-1} = a^{-1} - b^{-1} \in P$, i.e. $b^{-1} < a^{-1}$, or $\dfrac{1}{b} < \dfrac{1}{a}$ as desired. □

Problem 1.35. Prove that $\forall n \in \mathbb{N}$, x^n is increasing on P.

Proof. We must prove that $\forall a, b \in P$, $a < b \to a^n < b^n$.

Since $a < b$, $b - a \in P$. Furthermore, by repeated application of OF3, $a \in P$ implies that $a \cdot a \cdots a \in P$ for any arbitrary number of a terms multiplied together, i.e. $a^k \in P$ for some $k \in \mathbb{N}$. The same holds for b. As a sidenote, this can be rigorously justified by a brief induction argument.

Therefore, we have $a^k b^m \in P$ for any $k, m \in \mathbb{N}$. Then,

$$b^{n-1} + b^{n-2}a + b^{n-3}a^2 + \ldots + b^2a^{n-3} + ba^{n-2} + a^{n-1} \in P.$$

By OF3, $(b-a)(b^{n-1} + b^{n-2}a + b^{n-3}a^2 + \ldots + b^2a^{n-3} + ba^{n-2} + a^{n-1}) = b^n - a^n \in P$, i.e. $a^n < b^n$. $\square$

Alternative Proof. We may also solve this problem using induction. We will prove a slightly stronger claim: that $0 < a < b \to 0 < a^n < b^n$ for all $n \in \mathbb{N}$. The base case is when $n = 1$ - that is, $0 < a < b \to 0 < a^1 < b^1$. This is clearly true, so the base case is proved.

Next, assume that for some $k \in \mathbb{N}$, $0 < a < b \to 0 < a^k < b^k$. Then, we can apply Problem 1.33 to get that $a \cdot a^k < b \cdot b^k$, i.e. $a^{k+1} < b^{k+1}$. Moreover, we know that $0 < a \to 0 < a^{k+1}$ as we did in the first proof of this theorem, and so we have that $0 < a^{k+1} < b^{k+1}$ as desired. By induction, we have proven that $0 < a < b \to 0 < a^n < b^n$, and the theorem follows.

$\square$

Problem 1.36. Prove that $\forall n \in \mathbb{N}$, x^{-n} is decreasing on P.

Proof. In similar style as the last problem, we must prove that $\forall a, b \in P$, $a < b \to b^{-n} < a^{-n}$.

By Problem 1.35, $a < b \to a^n < b^n$. By Problem 1.34, $a^n < b^n \to \frac{1}{b^n} < \frac{1}{a^n}$, i.e. $b^{-n} < a^{-n}$, as desired. $\square$

Problem 1.37. Prove that if n is odd and greater than 0, x^n is increasing on the whole ordered field.

Proof. The theorem is equivalent to the statement: if $n > 0$ is odd, then $\forall a, b \in F$, $a < b \to a^n < b^n$. As a, b can be positive or negative, we proceed via casework on the signs of a and b. We already proved the case where both are positive in Problem 1.35. To get the other cases, we need a lemma:

Lemma

If n is odd, then $(-x)^n = -(x^n)$.

Proof. $(-x)^n = (-1 \cdot x)^n = (-1^n) \cdot (x^n) = (-1)^{\left(\frac{n-1}{2} \cdot 2\right)} \cdot (-1) \cdot (x^n) = \left((-1)^2\right)^{\left(\frac{n-1}{2}\right)} \cdot (-1) \cdot (x^n) = 1^{\text{something}} \cdot (-1) \cdot (x^n) = -x^n$. $\square$

Assume $a < b < 0$ so that $-a > -b > 0$. By the lemma, $(-a)^n = -a^n, (-b)^n = -b^n$. Then by Problem 1.35, we have $(-a)^n > (-b)^n \to -a^n > -b^n \to a^n < b^n$.

If $b < 0 < a$, the problem follows from comparing each of a^n and b^n to 0 and using transitivity. $\square$

Exercise 1.38. Prove that $a > 1 \to a^2 > a$.

Exercise 1.39. Prove that $0 < a < 1 \to a^2 < a$.

Exercise 1.40. Prove that $a < b \wedge c < 0 \to bc < ac$.

§2 Limits

§2.1 Review

Consider the $\varepsilon - \delta$ definition of the limit:

Definition 2.1. The limit of $f(x)$ as x approaches a is L, i.e. $\lim_{x\to a} f(x) = L$, iff

$$\forall\varepsilon > 0\ \exists\delta > 0\ \forall x,\ 0 < |x-a| < \delta \to |f(x) - L| < \varepsilon.$$

Exercise 2.2. Prove $\lim_{x\to 3} 7x + 117 = 138$.

Remark. If the definition is true for a specific ε_0, then it is also true for any $\widetilde{\varepsilon_0} > \varepsilon_0$. If a particular δ_0 works, then any $\widetilde{\delta_0} < \delta_0$ will also work.

Convince yourself why this must be true. This piece of observation indicates that there is flexibility in choosing our ε and δ in a proof. The following problem demonstrates this idea:

Problem 2.3. Prove $\lim_{x\to 99} x^2 = 9801$.

Solution. We must prove that

$$\forall\varepsilon > 0\ \exists\delta > 0\ \forall x,\ 0 < |x-99| < \delta \to \left|x^2 - 9801\right| < \varepsilon.$$

We rewrite $0 < |x-99| < \delta$ as $99-\delta < x < 99+\delta$, and $\left|x^2 - 9801\right| < \varepsilon$ as $9801-\varepsilon < x^2 < 9801+\varepsilon$. We want to take the square root of the latter inequality, but this is not possible if $9801-\varepsilon$ is negative. Therefore, we assume $\varepsilon < 9801$ (remember, as long as ε works, any $\widetilde{\varepsilon} > \varepsilon$ will also work!). Then, $9801-\varepsilon$ is definitely positive, so we can take the square root to get $\sqrt{9801-\varepsilon} < x < \sqrt{9801+\varepsilon}$. Note that we are assuming that the square root function is increasing, but we will take it granted for now.

Therefore our goal is to show that

$$99-\delta < x < 99+\delta \to \sqrt{9801-\varepsilon} < x < \sqrt{9801+\varepsilon}.$$

With interval notation, we want to show that the interval $(99-\delta, 99+\delta)$ is contained in the interval $(\sqrt{9801-\varepsilon}, \sqrt{9801+\varepsilon})$. It is sufficient that

$$\begin{aligned}\sqrt{9801-\varepsilon} &< 99-\delta,\\ 99+\delta &< \sqrt{9801+\varepsilon}.\end{aligned}$$

These rearrange to

$$\begin{aligned}\delta &< 99-\sqrt{9801-\varepsilon},\\ \delta &< \sqrt{9801+\varepsilon}-99.\end{aligned}$$

We want both of these inequalities to be true, so we let

$$\delta < \min\left\{99-\sqrt{9801-\varepsilon}, \sqrt{9801+\varepsilon}-99\right\},$$

so the implication $\forall\varepsilon > 0\ \exists\delta > 0\ \forall x,\ 0 < |x-99| < \delta \to \left|x^2 - 9801\right| < \varepsilon$ is true. $\square$

§2.2 Uncle Vanya's

Uncle Vanya's Proof-in-a-Minute is a method of proving the limits of polynomials. We introduce the underlying concept behind Uncle Vanya's:

Theorem 2.4 (The Factor Theorem)
For a polynomial P, $P(a) = 0 \longleftrightarrow (x-a) \mid P(x)$.

The application of this result will be necessary in every use of Uncle Vanya's. We will demonstrate how we use Uncle Vanya's in the general case:

Example 2.5
Using Uncle Vanya's, prove that all polynomials are continuous.

Later we will prove the sum and product laws of limits, from which this fact will follow more easily.

Proof. Consider an arbitrary polynomial $P(x)$. $P(x)$ is continuous iff $\forall a$ in the domain of P, $\lim_{x \to a} P(x) = P(a)$.

Consider the definition:

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x, \ 0 < |x-a| < \delta \to |P(x) - P(a)| < \varepsilon.$$

By the Factor Theorem, the polynomial $P(x) - P(a)$ is divisible by $x - a$, because $P(a) - P(a) = 0$. Therefore, $P(x) - P(a) = (x-a)Q(x)$ where $Q(x)$ is a polynomial. It then suffices to show that

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x, \ 0 < |x-a| < \delta \to |x-a| \, |Q(x)| < \varepsilon.$$

We know that the assumption $|x-a| < \delta$ implies that $|x-a|$ must be a small value. Furthermore, $|Q(x)|$ must be bounded, so $|x-a| \, |Q(x)|$ must be small.

To begin the proof, first assume $\delta < 1$. This step is valid because any $\widetilde{\delta} < \delta$ will also work. The choice of 1 is arbitrary; you could choose any convenient number as you see fit. Now note that $|x-a| < \delta < 1$. By the Triangle Inequality,

$$|x| \leq |x-a| + |a| < 1 + |a| \, .$$

Now let $Q(x) = \sum\limits_{k=0}^{n} b_k x^k$, where the general term b_k represents the constant coefficient of its corresponding term x^k in the polynomial. Then by the generalization of the Triangle Inequality,

$$|Q(x)| = \left| \sum_{k=0}^{n} b_k x^k \right| \leq \sum_{k=0}^{n} |b_k| \, |x|^k < \sum_{k=0}^{n} |b_k| \, (1 + |a|)^k.$$

Note that $\sum\limits_{k=0}^{n} |b_k| \, (1+|a|)^k$ is constant, and that $|Q(x)| < \sum\limits_{k=0}^{n} |b_k| \, (1+|a|)^k$, showing that it is bounded, which is what we desired.

Now we have

$$|x-a| \, |Q(x)| < \delta \cdot \left[\sum_{k=0}^{n} |b_k| \, (1+|a|)^k \right].$$

We want $|x-a|\,|Q(x)|$ to be less than ε. It is sufficient that we have

$$\delta \cdot \left[\sum_{k=0}^{n} |b_k|\,(1+|a|)^k\right] < \varepsilon,$$

which rearranges to

$$\delta < \frac{\varepsilon}{\sum\limits_{k=0}^{n} |b_k|\,(1+|a|)^k}.$$

However, earlier we have assumed that $\delta < 1$. We want to satisfy both of these inequalities, so we let

$$\delta < \min\left\{1, \frac{\varepsilon}{\sum\limits_{k=0}^{n} |b_k|\,(1+|a|)^k}\right\}.$$

Given a ε, we have found a δ such that $0 < |x-a| < \delta \to |P(x)-P(a)| < \varepsilon$, and therefore $\lim\limits_{x\to a} P(x) = P(a)$. □

Example 2.6

Prove that $\lim\limits_{x\to 3} x^4 - 2x^3 - 5x^2 + 11x - 1 = 14$.

Proof. We want to prove that

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x,\ 0 < |x-3| < \delta \to \left|x^4 - 2x^3 - 5x^2 + 11x - 1 - 14\right| < \varepsilon.$$

By synthetic division, we find that $x^4-2x^3-5x^2+11x-1-14 = x^4-2x^3-5x^2+11x-15 = (x-3)(x^3+x^2-2x+5)$.

We start by assuming $\delta < 1$, which implies that $|x-3| < \delta < 1$. Then by the Triangle Inequality, $|x| \le |x-3| + |3| < 1+3 = 4$. By the more generalized Triangle Inequality, we have

$$\begin{aligned}\left|x^3+x^2-2x+5\right| &\le |x|^3 + |x|^2 + |-2x| + |5| \\ &= |x|^3 + |x|^2 + 2\,|x| + 5 \\ &< 4^3 + 4^2 + 2\cdot 4 + 5 = 93.\end{aligned}$$

Therefore,

$$\begin{aligned}\left|x^4 - 2x^3 - 5x^2 + 11x - 1 - 14\right| &= \left|x^4 - 2x^3 - 5x^2 + 11x - 15\right| \\ &= |x-3|\left|x^3+x^2-2x+5\right| \\ &< \delta \cdot 93.\end{aligned}$$

We want this to be less than a given ε. This suggests that we should have $\delta < \dfrac{\varepsilon}{93}$. However, we have initially assumed $\delta < 1$, so it is sufficient to take

$$\delta < \min\left\{1, \frac{\varepsilon}{93}\right\}$$

to satisfy both inequalities. □

Exercise 2.7. Prove Problem 2.3 using Uncle Vanya's.

Exercise 2.8. Prove $\lim\limits_{x\to -2} 2x^4 + 5x^3 - 7x^2 - 11x + 2 = -12$.

Exercise 2.9. Prove $\lim\limits_{x\to -2} x^3 - x^2 - 13x + 11 = 25$.

§2.3 Other Limits

Problem 2.10. Prove $\lim_{x\to 3} \frac{1}{x} = \frac{1}{3}$.

Proof. We want to prove that

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x,\ 0 < |x-3| < \delta \to \left|\frac{1}{x} - \frac{1}{3}\right| < \varepsilon.$$

Assume $\delta < 1$, so $0 < |x-3| < \delta < 1$. Then by the Triangle Inequality, $|3| \leq |3-x| + |x| = |x-3| + |x| < 1 + |x|$, which rearranges to $|x| > 3 - 1 = 2$. Clearly $|x|$ is greater than zero, so we can take the reciprocal of this inequality (by Problem 1.34) to get $\frac{1}{|x|} < \frac{1}{2}$.

Now note that

$$\left|\frac{1}{x} - \frac{1}{3}\right| = \left|\frac{3-x}{3x}\right| = \frac{|3-x|}{|3x|} = \frac{|x-3|}{3\,|x|} = \frac{1}{3} \cdot \frac{1}{|x|} \cdot |x-3| < \frac{1}{3} \cdot \frac{1}{2} \cdot \delta = \frac{\delta}{6}.$$

We want $\left|\frac{1}{x} - \frac{1}{3}\right|$ to be less than a given ε, so we should take $\delta < 6\varepsilon$. However, we have also assumed $\delta < 1$, so it is sufficient to take $\delta < \min\{1, 6\varepsilon\}$ to satisfy both inequalities. □

Alternative Proof. There is an alternative method that is similar to the proof of Problem 2.3. Consider the implication $|x-3| < \delta \to \left|\frac{1}{x} - \frac{1}{3}\right| < \varepsilon$. We break up the absolute value signs to get

$$3 - \delta < x < 3 + \delta \to \frac{1}{3} - \varepsilon < \frac{1}{x} < \frac{1}{3} + \varepsilon.$$

Assume $\varepsilon < \frac{1}{3}$, which allows us to take the reciprocal of the latter inequality to get

$$3 - \delta < x < 3 + \delta \to \frac{1}{\frac{1}{3} + \varepsilon} < x < \frac{1}{\frac{1}{3} - \varepsilon}.$$

We want to show that the interval $(3-\delta, 3+\delta)$ is contained inside $\left(\frac{1}{\frac{1}{3}+\varepsilon}, \frac{1}{\frac{1}{3}-\varepsilon}\right)$. It is sufficient that

$$\frac{1}{\frac{1}{3}+\varepsilon} < 3 - \delta,$$
$$3 + \delta < \frac{1}{\frac{1}{3}-\varepsilon}.$$

These rearrange to

$$\delta < 3 - \frac{1}{\frac{1}{3}+\varepsilon},$$
$$\delta < \frac{1}{\frac{1}{3}-\varepsilon} - 3.$$

Therefore we take

$$\delta < \min\left\{3 - \frac{1}{\frac{1}{3}+\varepsilon}, \frac{1}{\frac{1}{3}-\varepsilon} - 3\right\}$$

to satisfy both inequalities. □

Problem 2.11. Let $f(x) = \begin{cases} 7 & x \leq 4 \\ 5 & x > 4 \end{cases}$. Prove that $\lim_{x \to 4} f(x)$ does not exist.

Proof. We proceed with proof by contradiction. Assume that $\lim_{x \to 4} f(x) = L$. Then,

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x,\ 0 < |x - 4| < \delta \to |f(x) - L| < \varepsilon.$$

Consider the negation of this statement:

$$\exists \varepsilon > 0\ \forall \varepsilon > 0\ \exists x,\ 0 < |x - 4| < \delta \wedge |f(x) - L| \geq \varepsilon.$$

As long as we can find a single value of ε such that the implication is false for any δ, then we will have shown that the limit does not exist.

First, we rewrite the implication to get

$$4 - \delta < x < 4 + \delta \to L - \varepsilon < f(x) < L + \varepsilon.$$

Assume $\varepsilon = \frac{1}{2}$ (in fact, any value of $\varepsilon \in (0, 1)$ can be used to demonstrate a contradiction). Then we have

$$4 - \delta < x < 4 + \delta \to L - \frac{1}{2} < f(x) < L + \frac{1}{2},$$

which can be represented as

$$x \in (4 - \delta, 4 + \delta) \to f(x) \in \left(L - \frac{1}{2}, L + \frac{1}{2}\right).$$

For any $\delta > 0$, there are values of x in the interval $(4 - \delta, 4 + \delta)$ such that $x < 4$ and $x > 4$, which imply $f(x) = 5$ and $f(x) = 7$ respectively. However, the interval $\left(L - \frac{1}{2}, L + \frac{1}{2}\right)$ has a length of 1; it is impossible that 5 and 7 are both contained within this interval, so we have a contradiction. Therefore the limit does not exist. $\square$

Definition 2.12. The **right-hand limit** as $x \to a$ of $f(x)$ is L when

$$\lim_{x \to a^+} f(x) = L \longleftrightarrow \forall \varepsilon > 0\ \exists \delta > 0\ \forall x,\ 0 < x - a < \delta \to |f(x) - L| < \varepsilon.$$

Definition 2.13. The **left-hand limit** as $x \to a$ of $f(x)$ is L when

$$\lim_{x \to a^-} f(x) = L \longleftrightarrow \forall \varepsilon > 0\ \exists \delta > 0\ \forall x,\ 0 < a - x < \delta \to |f(x) - L| < \varepsilon.$$

Exercise 2.14. Let $f(x)$ be the function considered in Problem 2.11. Evaluate and prove $\lim_{x \to 4^+} f(x)$ and $\lim_{x \to 4^-} f(x)$.

Exercise 2.15. Given $\lim_{x \to a^+} f(x)$ and $\lim_{x \to a^-} f(x)$ exist, prove that $\lim_{x \to a^+} f(x) = \lim_{x \to a^-} f(x) = L \longleftrightarrow \lim_{x \to a} f(x) = L$.

Definition 2.16 (Infinite Limits)**.** We introduce new variables N and M to deal with cases when x or $f(x)$ go to ∞ or $-\infty$.

- $\lim_{x \to \infty} f(x) = L \longleftrightarrow \forall \varepsilon > 0\ \exists N\ \forall x,\ x > N \to |f(x) - L| < \varepsilon.$
- $\lim_{x \to -\infty} f(x) = L \longleftrightarrow \forall \varepsilon > 0\ \exists N\ \forall x,\ x < N \to |f(x) - L| < \varepsilon.$

- $\lim\limits_{x \to \infty} f(x) = \infty \longleftrightarrow \forall M\ \exists N\ \forall x,\ x > N \to f(x) > M.$
- $\lim\limits_{x \to -\infty} f(x) = \infty \longleftrightarrow \forall M\ \exists N\ \forall x,\ x < N \to f(x) > M.$
- $\lim\limits_{x \to \infty} f(x) = -\infty \longleftrightarrow \forall M\ \exists N\ \forall x,\ x > N \to f(x) < M.$
- $\lim\limits_{x \to -\infty} f(x) = -\infty \longleftrightarrow \forall M\ \exists N\ \forall x,\ x < N \to f(x) < M.$
- $\lim\limits_{x \to a} f(x) = \infty \longleftrightarrow \forall M\ \exists \delta > 0\ \forall x,\ 0 < |x-a| < \delta \to f(x) > M.$
- $\lim\limits_{x \to a} f(x) = -\infty \longleftrightarrow \forall M\ \exists \delta > 0\ \forall x,\ 0 < |x-a| < \delta \to f(x) < M.$

Remark. When considering the right-hand or left-hand limits as $f(x)$ goes to ∞ or $-\infty$, we can just replace $0 < |x-a| < \delta$ (in the original definition) with $0 < x-a < \delta$ or $0 < a-x < \delta$ respectively.

Problem 2.17. Prove $\lim\limits_{x \to \infty} \dfrac{1}{x} = 0$.

Proof. We want to prove that

$$\forall \varepsilon > 0\ \exists N\ \forall x,\ x > N \to \left|\frac{1}{x} - 0\right| < \varepsilon.$$

Assume $N > 0$ (we have the same freedom in choosing N as we had with δ). Then $x > N > 0 \to \dfrac{1}{x} < \dfrac{1}{N}$. Since $x > 0$ by assumption, $x = |x|$, so we have $\dfrac{1}{|x|} < \dfrac{1}{N}$, i.e. $\left|\dfrac{1}{x} - 0\right| < \dfrac{1}{N}$. We want $\left|\dfrac{1}{x} - 0\right|$ to be less than a given ε, which suggests that $\dfrac{1}{N} < \varepsilon$. Therefore it is sufficient to take $N > \dfrac{1}{\varepsilon}$. □

Problem 2.18. Prove $\lim\limits_{x \to 0^-} \dfrac{1}{x} = -\infty$.

Proof. Consider the corresponding definition

$$\forall M\ \exists \delta > 0\ \forall x,\ 0 < 0 - x < \delta \to \frac{1}{x} < M.$$

We rewrite the former inequality as $0 > x > -\delta$. Since x and $-\delta$ are both negative, we can take the reciprocal to get $\dfrac{1}{x} < -\dfrac{1}{\delta}$. However, we want $\dfrac{1}{x}$ to be less than a given M, so it is sufficient to take $-\dfrac{1}{\delta} < M$, i.e. $\delta < -\dfrac{1}{M}$. □

Theorem 2.19 (Uniqueness of Limits)
If $\lim\limits_{x \to a} f(x) = L$ and $\lim\limits_{x \to a} f(x) = M$, then $L = M$.

Proof. For the sake of contradiction, assume $L \neq M$. Then it follows that $|L - M| > 0$. Now consider

$$\lim_{x \to a} f(x) = L \longleftrightarrow \forall \varepsilon > 0\ \exists \delta_1 > 0\ \forall x,\ 0 < |x-a| < \delta_1 \to |f(x) - L| < \varepsilon,$$

$$\lim_{x\to a} f(x) = M \longleftrightarrow \forall\varepsilon > 0\ \exists\delta_2 > 0\ \forall x,\ 0 < |x-a| < \delta_2 \to |f(x) - M| < \varepsilon.$$

Choose $\delta = \min\{\delta_1, \delta_2\}$, so we have

$$\forall\varepsilon > 0\ \exists\delta > 0\ \forall x,\ 0 < |x-a| < \delta \to |f(x) - L| < \varepsilon \wedge |f(x) - M| < \varepsilon.$$

Now consider $\varepsilon = \dfrac{|L-M|}{2}$ (recall that $|L-M| > 0$, i.e. $\dfrac{|L-M|}{2}$ by assumption). It becomes true that

$$0 < |x-a| < \delta \to |f(x) - L| < \frac{|L-M|}{2} \wedge |f(x) - M| < \frac{|L-M|}{2}.$$

Then note that

$$\begin{aligned}|L-M| &= |L - f(x) + f(x) - M| \\ &< |L - f(x)| + |f(x) - M| \\ &= |f(x) - L| + |f(x) - M| \\ &< \frac{|L-M|}{2} + \frac{|L-M|}{2} = |L-M|.\end{aligned}$$

We have $|L-M| < |L-M|$, a contradiction. Therefore $L = M$. □

Problem 2.20. Prove that if $f(x) \geq 0\ \forall x$ and $\lim_{x\to a} f(x) = L$, then $L \geq 0$.

Proof. Assume for the sake of contradiction that $L < 0$. We are given the definition

$$\lim_{x\to a} f(x) = L \longleftrightarrow \forall\varepsilon > 0\ \exists\delta > 0\ \forall x,\ 0 < |x-a| < \delta \to |f(x) - L| < \varepsilon.$$

As $f(x) \geq 0$ is given, $f(x) > L$, i.e. $f(x) - L > 0$. Therefore $|f(x) - L| = f(x) - L$. Consider $\varepsilon = -L$, which is valid because $-L > 0$. We can also choose $\varepsilon = -\dfrac{L}{2}$, or any other value in terms of L that would be positive. It follows that $\exists\delta > 0\ \forall x,\ 0 < |x-a| < \delta \to f(x) - L < -L$. But $f(x) - L < -L$ rearranges to $f(x) < 0$, which contradicts the given $f(x) \geq 0\ \forall x$. Therefore $L \geq 0$. □

§2.4 Limit Properties

Theorem 2.21 (Limit of the Identity Function)

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

Proof. It is sufficient to let $\delta = \varepsilon$. Then it is obvious that $0 < |x-a| < \delta \longleftrightarrow 0 < |x-a| < \varepsilon \to |x-a| < \varepsilon$. □

Theorem 2.22 (Limit of a Constant)

$\lim_{x\to a} c = c$ for some constant c.

Proof. Any δ is valid, since $\forall\varepsilon > 0$ we have $|c - c| = 0 < \varepsilon$ which is always true. □

Theorem 2.23 (Sum and Product of Limits)
Suppose $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$. Then

a) $\lim_{x\to a} (f(x) + g(x)) = L + M$.

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

Proof. a) We are given $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$, therefore

$$\begin{aligned} \forall\varepsilon > 0\ \exists\delta_1 > 0\ \forall x,\ 0 < |x-a| < \delta_1 \longrightarrow |f(x) - L| < \frac{\varepsilon}{2}, \\ \exists\delta_2 > 0\ \forall x,\ 0 < |x-a| < \delta_2 \longrightarrow |g(x) - M| < \frac{\varepsilon}{2}. \end{aligned}$$

This is allowed because we can always choose a δ_1 or δ_2 that is small enough to make each $|f(x) - L|$ and $|g(x) - M|$ smaller than $\frac{\varepsilon}{2}$.

Let $\delta = \min\{\delta_1, \delta_2\}$. This is because we want to be able to use the same δ for both definitions. Then,

$$\forall\varepsilon > 0\ \exists\delta > 0\ \forall x,\ 0 < |x-a| < \delta \longrightarrow |f(x) - L| < \frac{\varepsilon}{2} \wedge |g(x) - M| < \frac{\varepsilon}{2}.$$

Note that $|f(x)+g(x)-(L+M)| = |(f(x)-L)+(g(x)-M)| \leq |f(x)-L|+|g(x)+M|$ by the Triangle Inequality. Therefore, given $\varepsilon > 0$, we have found a δ such that

$$|x-a| < \delta \longrightarrow |f(x) + g(x) - (L+M)| \leq |f(x) - L| + |g(x) + M| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

i.e. $\lim_{x\to a} (f(x) + g(x)) = L + M$, as desired.

b) Let $\widetilde{\varepsilon}$ represent some quantity in terms of ε. Eventually we will figure out exactly what expression we should set $\widetilde{\varepsilon}$ equal to, by doing the proof 'backwards,' to demonstrate the motivation and intuition behind the choice of $\widetilde{\varepsilon}$. After we find out what $\widetilde{\varepsilon}$ exactly is, we will write the proof forwards.

As we are given $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$,

$$\begin{aligned} \forall\varepsilon > 0\ \exists\delta_1 > 0\ \forall x,\ 0 < |x-a| < \delta_1 \longrightarrow |f(x) - L| < \widetilde{\varepsilon}, \\ \exists\delta_2 > 0\ \forall x,\ 0 < |x-a| < \delta_2 \longrightarrow |g(x) - M| < \widetilde{\varepsilon}. \end{aligned}$$

Like before, let $\delta = \min\{\delta_1, \delta_2\}$. Then we can cover both cases:

$$\forall\varepsilon > 0\ \exists\delta > 0\ \forall x,\ 0 < |x-a| < \delta \longrightarrow |f(x) - L| < \widetilde{\varepsilon} \wedge |g(x) - M| < \widetilde{\varepsilon}.$$

Now we apply the Triangle Inequality based on what we are trying to prove:

$$\begin{aligned} |f(x)g(x) - LM| &= |f(x)g(x) - f(x)M + f(x)M - LM| \\ &= |f(x)(g(x) - M) + M(f(x) - L)| \\ &\leq |f(x)|\,|g(x) - M| + |M|\,|f(x) - L|. \end{aligned}$$

Now we should introduce a bound on $|f(x)|$. Therefore we should assume $\varepsilon < 1$. Then it follows that $|f(x) - L| < 1$. By another application of the Triangle Inequality,

$$|f(x)| \leq |f(x) - L| + |L| < 1 + |L|.$$

We can now conclude that

$$\begin{aligned} |f(x)g(x) - LM| &\leq |f(x)|\,|g(x) - M| + |M|\,|f(x) - L| \\ &< (1 + |L|)\widetilde{\varepsilon} + |M|\,\widetilde{\varepsilon} \\ &= (1 + |L| + |M|)\widetilde{\varepsilon}. \end{aligned}$$

It would be sufficient to let $\widetilde{\varepsilon} = \dfrac{\varepsilon}{1 + |L| + |M|}$, so $|f(x)g(x) - LM| < \widetilde{\varepsilon}(1 + |L| + |M|) < \varepsilon$, as desired.

Now we should write the proof forwards:
As we are given $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$,

$$\begin{aligned} \forall\varepsilon > 0\ \exists\delta_1 > 0\ \forall x,\ 0 < |x - a| < \delta_1 &\longrightarrow |f(x) - L| < \frac{\varepsilon}{1 + |L| + |M|}, \\ \exists\delta_2 > 0\ \forall x,\ 0 < |x - a| < \delta_2 &\longrightarrow |g(x) - M| < \frac{\varepsilon}{1 + |L| + |M|}. \end{aligned}$$

Let $\delta = \min\{\delta_1, \delta_2\}$. Then $\exists\delta > 0$ such that

$$0 < |x - a| < \delta \longrightarrow |f(x) - L| < \frac{\varepsilon}{1 + |L| + |M|} \wedge |g(x) - M| < \frac{\varepsilon}{1 + |L| + |M|}.$$

Assume $\varepsilon < 1$, so that $|f(x) - L| < 1$. Then $|f(x)| \leq |f(x) - L| + |L| < 1 + |L|$. Therefore, we conclude that

$$\begin{aligned} |f(x)g(x) - LM| &= |f(x)g(x) - f(x)M + f(x)M - LM| \\ &= |f(x)(g(x) - M) + M(f(x) - L)| \\ &\leq |f(x)|\,|g(x) - M| + |M|\,|f(x) - L| \\ &< (1 + |L|)\frac{\varepsilon}{1 + |L| + |M|} + |M|\frac{\varepsilon}{1 + |L| + |M|} \\ &= (1 + |L| + |M|)\frac{\varepsilon}{1 + |L| + |M|} \\ &= \varepsilon. \end{aligned}$$

□

Corollary 2.24 (Difference of Limits)
$\lim_{x\to a}(f(x) - g(x)) = \lim_{x\to a} f(x) - \lim_{x\to a} g(x)$.

Proof. It follows from Theorem 2.22 and Theorem 2.23 that $\lim_{x\to a}(f(x) - g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a}(-g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a}(-1)\lim_{x\to a} g(x) = \lim_{x\to a} f(x) - \lim_{x\to a} g(x)$. □

Lemma 2.25
If $0 \leq j(x) \leq k(x)\ \forall x$ and $\lim_{x\to a} k(x) = 0$, then $\lim_{x\to a} j(x) = 0$.

Proof. It follows that $k(x) - j(x) \geq 0 \ \forall x$. By Problem 2.20, $\lim_{x \to a} k(x) - j(x) \geq 0$. But note that $\lim_{x \to a} k(x) - j(x) = \lim_{x \to a} k(x) - \lim_{x \to a} j(x) = -\lim_{x \to a} j(x)$, therefore $\lim_{x \to a} j(x) \leq 0$. However, as we are also given $j(x) \geq 0 \ \forall x$, so by Problem 2.20, $\lim_{x \to a} j(x) \geq 0$. Thus we can only conclude $\lim_{x \to a} j(x) = 0$. □

Theorem 2.26 (Squeeze Theorem)

Suppose $f(x) \leq g(x) \leq h(x) \ \forall x$ and $\lim_{x \to a} f(x) = \lim_{x \to a} h(x) = L$. Then $\lim_{x \to a} g(x) = L$.

Proof. The given inequality rearranges to $0 \leq g(x) - f(x) \leq h(x) - f(x)$. Note that $\lim_{x \to a}(h(x) - f(x)) = \lim_{x \to a} h(x) - \lim_{x \to a} f(x) = L - L = 0$. By Lemma 2.25, $\lim_{x \to a}(g(x) - f(x)) = 0$, i.e. $\lim_{x \to a} g(x) = \lim_{x \to a} f(x) = L$, as desired. □

Alternative Proof. In fact, we could prove this theorem from scratch, without using previous lemmas.

By definition of limits, there exist δ_1, δ_2 such that

$$\begin{aligned} 0 < |x - a| < \delta_1 \to |f(x) - L| < \varepsilon, \\ 0 < |x - a| < \delta_2 \to |h(x) - L| < \varepsilon. \end{aligned}$$

Let $\delta = \min\{\delta_1, \delta_2\}$, it follows that

$$|x - a| < \delta \to L - \varepsilon < f(x), h(x) < L + \varepsilon.$$

Then

$$L - \varepsilon < f(x) \leq g(x) \leq h(x) < L + \varepsilon \implies |g(x) - L| < \varepsilon,$$

so this choice of δ establishes that $\lim_{x \to a} g(x) = L$. □

Theorem 2.27 (Reciprocal of Limits)

Suppose $\lim_{x \to a} g(x) = L$, and $L \neq 0$. Then $\lim_{x \to a} \frac{1}{g(x)} = \frac{1}{L}$.

Proof. Given $\varepsilon > 0$, there exists δ_1 such that $0 < |x - a| < \delta_1 \to |g(x) - L| < \frac{|L|}{2}$, since $\frac{|L|}{2}$ is a number greater than 0. Likewise, there exists δ_2 such that $0 < |x - a| < \delta_2 \to |g(x) - L| < \frac{\varepsilon |L|^2}{2}$. Then take $\delta = \min\{\delta_1, \delta_2\}$, such that

$$0 < |x - a| < \delta \to |g(x) - L| < \frac{|L|}{2} \wedge |g(x) - L| < \frac{\varepsilon |L|^2}{2}.$$

By the Triangle Inequality, $|L| \leq |L - g(x)| + |g(x)| = |g(x) - L| + |g(x)| < \frac{|L|}{2} + |g(x)|$, therefore $|g(x)| > \frac{|L|}{2}$. As both quantities are positive, we can take the reciprocal to get $\frac{1}{|g(x)|} < \frac{2}{|L|}$. Note that we applied the Triangle Inequality on $|L|$ because we seek a lower bound for $|g(x)|$, such that we have an upper bound for $\frac{1}{|g(x)|}$, which is needed to finish the proof.

Therefore, we have found a δ such that

$$\left|\frac{1}{g(x)} - \frac{1}{L}\right| = \left|\frac{L-g(x)}{g(x)L}\right| = \frac{|g(x)-L|}{|g(x)|\,|L|} < \frac{2}{|L|} \cdot \frac{1}{|L|} \cdot \frac{\varepsilon\,|L|^2}{2} = \varepsilon,$$

given $\varepsilon > 0$, and we can conclude that $\lim_{x\to a} \frac{1}{g(x)} = \frac{1}{L}$. $\square$

Problem 2.28. Evalute and justify $\lim_{x\to\infty} \frac{\sin x}{x} = 0$.

Solution. It is true that $-1 \le \sin x \le 1$. We divide by x to get $-\frac{1}{x} \le \frac{\sin x}{x} \le \frac{1}{x}$. Note that $\lim_{x\to\infty} -\frac{1}{x} = \lim_{x\to\infty} \frac{1}{x} = 0$. We can then apply the Squeeze Theorem (we did not prove the case when $x \to \infty$, but the proof for this is virtually identical to the one given) to conclude that $\lim_{x\to\infty} \frac{\sin x}{x} = 0$. $\square$

Corollary 2.29 (Quotient of Limits)

Provided $\lim_{x\to a} g(x) \neq 0$, $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$.

Proof. This follows from application of the reciprocal rule with the product rule. $\square$

Lemma 2.30 (Continuity of Polynomials)

Let $P(x)$ be a polynomial. Then $\lim_{x\to a} P(x) = P(a)$.

Proof. As shown in Example 2.5, we have proven this fact using Uncle Vanya's. However, the proof of this lemma immediately follows from the application of the sum and product rules of limits, which have been proven at this stage.

Let $P(x) = \sum_{k=0}^{n} b_k x^k$. Then,

$$\begin{aligned}\lim_{x\to a} P(x) &= \lim_{x\to a} \sum_{k=0}^{n} b_k x^k \\ &= \sum_{k=0}^{n} \lim_{x\to a} b_k \cdot (\lim_{x\to a} x)^k \\ &= \sum_{k=0}^{n} b_k (a)^k \\ &= P(a).\end{aligned}$$

$\square$

Lemma 2.31 (Continuity of Rational Functions)

Let $P(x)$, $Q(x)$ be polynomials. If $Q(x) \neq 0\ \forall x$, then

$$\lim_{x\to a} \frac{P(x)}{Q(x)} = \frac{P(a)}{Q(a)}.$$

Proof. We combine the results of Corollary 2.29 with Lemma 2.30. $\square$

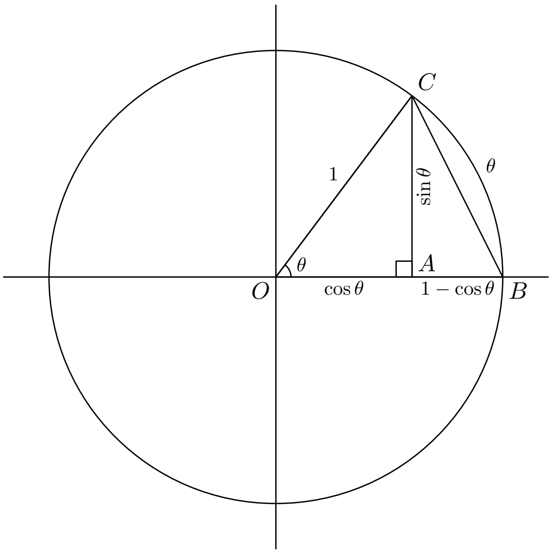
§2.5 Trigonometric Limits

We define sine, cosine via the unit circle.

Example 2.32

Prove that $\lim_{\theta\to 0} \sin\theta = 0$ and $\lim_{\theta\to 0} \cos\theta = 1$.

Proof. Assume $0 < \theta < \frac{\pi}{2}$, where θ is in radians. We only care about θ being close to 0, which is why we restrict θ to be less than $\frac{\pi}{2}$ radians. Consider the unit circle:



We will take for granted that $\widehat{BC} > \overline{BC}$ (we will glance over the rigorous proof for this), and since $\overline{BC}$ is the hypotenuse of $\triangle ABC$, $\overline{BC} > \overline{AC}$ and $\overline{BC} > \overline{AB}$. Therefore, $\widehat{BC} > \overline{AC}$ and $\widehat{BC} > \overline{AB}$. Since θ is in radians, the length of $\widehat{BC}$ is θ. Note that $\overline{AC} = \sin\theta$ and $\overline{AB} = 1 - \cos\theta$. We have the inequalities:

$$\begin{aligned} 0 &< \sin\theta < \theta, \\ 0 &< 1 - \cos\theta < \theta. \end{aligned}$$

As $\lim_{\theta\to 0^+} 0 = \lim_{\theta\to 0^+} \theta = 0$, by Theorem 2.26, $\lim_{\theta\to 0^+} \sin\theta = \lim_{\theta\to 0^+} 1 - \cos\theta = 0$; the latter rearranges to $\lim_{\theta\to 0^+} \cos\theta = 1$.

Lemma 2.33

Assuming both limits exist, $\lim_{x\to a^+} f(x) = \lim_{x\to -a^-} f(-x)$.

Proof. Consider the definition of $\lim_{x \to a^+} f(x) = L$,

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x, \ 0 < x - a < \delta \to |f(x) - L| < \varepsilon.$$

Substitute $x \to -x$. Then,

$$\begin{aligned} 0 < (-x) - a < \delta \to |f(-x) - L| < \varepsilon \\ 0 < -a - x < \delta \to |f(-x) - L| < \varepsilon \end{aligned}$$

which is the definition of $\lim_{x \to -a^-} f(-x) = L$. Therefore the definitions are equivalent, i.e. $\lim_{x \to a^+} f(x) = \lim_{x \to -a^-} f(-x)$. □

By Lemma 2.33, it follows that $\lim_{\theta \to 0^+} \sin\theta = \lim_{\theta \to -0^-} \sin(-\theta) = -\lim_{\theta \to -0^-} \sin\theta = -\lim_{\theta \to 0^-} \sin\theta$. Since $\lim_{\theta \to 0^+} \sin\theta = 0$, $-\lim_{\theta \to 0^-} \sin\theta = 0$, i.e. $\lim_{\theta \to 0^-} \sin\theta = 0$. By Exercise 2.15, we conclude that $\lim_{\theta \to 0} \sin\theta = 0$.

Similarly, we note that $\lim_{\theta \to 0^+} \cos\theta = \lim_{\theta \to -0^-} \cos(-\theta) = \lim_{\theta \to -0^-} \cos\theta = \lim_{\theta \to 0^-} \cos\theta$ by Lemma 2.33. Since $\lim_{\theta \to 0^+} \cos\theta = 1$, we have $\lim_{\theta \to 0^-} \cos\theta = 1$. By Exercise 2.15, we conclude that $\lim_{\theta \to 0} \cos\theta = 1$. □

Lemma 2.34

$\lim_{x \to a} f(x) = L \longleftrightarrow \lim_{h \to 0} f(a+h) = L$.

Proof. The definition of $\lim_{x \to a} f(x) = L$ is

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x, \ 0 < |x - a| < \delta \to |f(x) - L| < \varepsilon.$$

Substitute $x \to a + h$. Then,

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall h, \ 0 < |a + h - a| < \delta \to |f(a+h) - L| < \varepsilon,$$

which simplifies to

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall h, \ 0 < |h - 0| < \delta \to |f(a+h) - L| < \varepsilon,$$

i.e. the definition of $\lim_{h \to 0} f(a+h) = L$. Since the definitions are equivalent, $\lim_{x \to a} f(x) = L \longleftrightarrow \lim_{h \to 0} f(a+h) = L$. □

Theorem 2.35 (Continuity of Trigonometric Functions)

If a is in the domain, then $\lim_{x \to a} \mathrm{Trig}(x) = \mathrm{Trig}(a)$, where $\mathrm{Trig} = \sin, \cos, \tan, \csc, \sec, \cot$.

Proof. We will prove this for $\sin x$ and $\cos x$, and the rest will follow by Theorem 2.27.

We proceed by Lemma 2.34 on $\sin x$ and $\cos x$.

$$\lim_{x \to a} \sin x = \lim_{h \to 0} \sin(a + h)$$

$$\begin{aligned}
&= \lim_{h \to 0} (\sin a \cos h + \cos a \sin h) \\
&= \lim_{h \to 0} \sin a \cdot \lim_{h \to 0} \cos h + \lim_{h \to 0} \cos a \cdot \lim_{h \to 0} \sin h \\
&= \sin a \cdot 1 + \cos a \cdot 0 \\
&= \sin a.
\end{aligned}$$

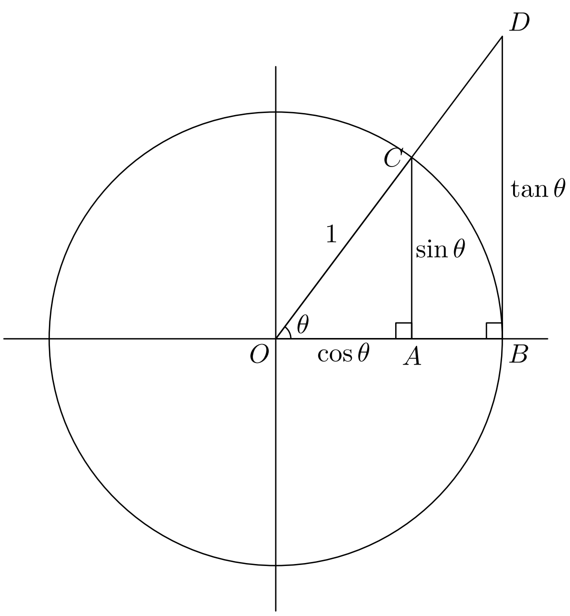
$$\begin{aligned}
\lim_{x \to a} \cos x &= \lim_{h \to 0} \cos(a + h) \\
&= \lim_{h \to 0} (\cos a \cos h - \sin a \sin h) \\
&= \lim_{h \to 0} \cos a \cdot \lim_{h \to 0} \cos h - \lim_{h \to 0} \sin a \cdot \lim_{h \to 0} \sin h \\
&= \cos a \cdot 1 + \sin a \cdot 0 \\
&= \cos a.
\end{aligned}$$

□

Theorem 2.36

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

Proof. Assume $0 < \theta < \frac{\pi}{2}$, where θ is in radians. Consider the following diagram:



It is clear that we have the following inequality:

$$\text{Area of } \triangle OAC < \text{Area of sector } OBC < \text{Area of } \triangle OBD.$$

We find our respective areas:

$$\frac{\sin \theta \cos \theta}{2} < \frac{\theta}{2} < \frac{\tan \theta}{2}.$$

This rearranges to

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

Note that $\lim_{\theta \to 0^+} \frac{1}{\cos\theta} = 1$ and $\lim_{\theta \to 0^+} \cos\theta = 1$, therefore by Theorem 2.26, we have that $\lim_{\theta \to 0^+} \frac{\sin\theta}{\theta} = 1$.

We can then proceed with Lemma 2.33, and eventually conclude that $\lim_{\theta \to 0^-} \frac{\sin\theta}{\theta} = 1$, so therefore $\lim_{\theta \to 0} \frac{\sin\theta}{\theta} = 1$. □

Lemma 2.37

$\lim_{x \to 0} f(x) = L \longleftrightarrow \lim_{x \to 0} f(kx) = L$ provided $k \neq 0$.

Proof. The definition of $\lim_{x \to 0} f(x) = L$ is

$$\forall \varepsilon > 0\ \exists \delta_1 > 0\ \forall x,\ 0 < |x - 0| < \delta_1 \to |f(x) - L| < \varepsilon.$$

Substitute $x \to kx$. Then we have,

$$\forall \varepsilon > 0\ \exists \delta_1 > 0\ \forall x,\ 0 < |kx - 0| < \delta_1 \to |f(kx) - L| < \varepsilon,$$

which rearranges to

$$\forall \varepsilon > 0\ \exists \delta_1 > 0\ \forall x,\ 0 < |k|\,|x| < \delta_1 \to |f(kx) - L| < \varepsilon.$$

Choose $\delta_2 = \frac{\delta_1}{|k|}$. It follows that

$$\forall \varepsilon > 0\ \exists \delta_2 > 0\ \forall x,\ 0 < |x - 0| < \delta_2 \to |f(kx) - L| < \varepsilon,$$

which is just the definition of $\lim_{x \to 0} f(kx) = L$. □

Lemma 2.38

$\lim_{x \to a} f(x) = L \longleftrightarrow \lim_{x \to \frac{a}{k}} f(kx) = L$ provided $k \neq 0$.

The proof is very similar to that of the previous lemma.

Proof. The definition of $\lim_{x \to a} f(x) = L$ is

$$\forall \varepsilon > 0\ \exists \delta_1 > 0\ \forall x,\ 0 < |x - a| < \delta_1 \to |f(x) - L| < \varepsilon.$$

Substitute $x \to kx$. Then we have,

$$\forall \varepsilon > 0\ \exists \delta_1 > 0\ \forall x,\ 0 < |kx - a| < \delta_1 \to |f(kx) - L| < \varepsilon,$$

which rearranges to

$$\forall \varepsilon > 0\ \exists \delta_1 > 0\ \forall x,\ 0 < |k| \left| x - \frac{a}{k} \right| < \delta_1 \to |f(kx) - L| < \varepsilon.$$

Choose $\delta_2 = \frac{\delta_1}{|k|}$. It follows that

$$\forall \varepsilon > 0\ \exists \delta_2 > 0\ \forall x,\ 0 < \left| x - \frac{a}{k} \right| < \delta_2 \to |f(kx) - L| < \varepsilon,$$

which is just the definition of $\lim_{x \to 0} f(kx) = L$. □

Problem 2.39. Find and justify the following limits:

1. $\lim\limits_{x\to 0} \cos 2x$
2. $\lim\limits_{x\to 0} \dfrac{\sin 2x}{2x}$
3. $\lim\limits_{x\to 0} \dfrac{\sin 2x}{3x}$
4. $\lim\limits_{x\to 0} \dfrac{\sin 6x}{\sin 5x}$
5. $\lim\limits_{x\to 0} \dfrac{\tan 3x}{x}$
6. $\lim\limits_{x\to 0} \dfrac{1-\cos x}{x}$

Solution. We apply the results of Example 2.32, Theorem 2.36 with Lemma 2.37 after some clever algebraic manipulations.

1. $\lim\limits_{x\to 0} \cos 2x = \boxed{1}$.
2. $\lim\limits_{x\to 0} \dfrac{\sin 2x}{2x} = \boxed{1}$.
3. $\lim\limits_{x\to 0} \dfrac{\sin 2x}{3x} = \lim\limits_{x\to 0} \dfrac{\sin 2x}{2x} \cdot \dfrac{2}{3} = \boxed{\dfrac{2}{3}}$.
4. $\lim\limits_{x\to 0} \dfrac{\sin 6x}{\sin 5x} = \lim\limits_{x\to 0} \dfrac{\frac{\sin 6x}{x}}{\frac{\sin 5x}{x}} = \lim\limits_{x\to 0} \dfrac{\frac{\sin 6x}{6x}\cdot 6}{\frac{\sin 5x}{5x}\cdot 5} = \dfrac{6}{5}\lim\limits_{x\to 0} \dfrac{\frac{\sin 6x}{6x}}{\frac{\sin 5x}{5x}} = \dfrac{6}{5}\cdot\dfrac{1}{1} = \boxed{\dfrac{6}{5}}$.
5. $\lim\limits_{x\to 0} \dfrac{\tan 3x}{x} = \lim\limits_{x\to 0} \dfrac{\sin 3x}{x\cos 3x} = \lim\limits_{x\to 0} \dfrac{\frac{\sin 3x}{3x}}{x\cdot\frac{\cos 3x}{3x}} = \lim\limits_{x\to 0} \dfrac{\frac{\sin 3x}{3x}}{\frac{\cos 3x}{3}} = \dfrac{\lim\limits_{x\to 0}\frac{\sin 3x}{3x}}{\frac{1}{3}\lim\limits_{x\to 0}\cos 3x} = \dfrac{1}{\frac{1}{3}\cdot 1} =$ $\boxed{3}$.
6. $\lim\limits_{x\to 0} \dfrac{1-\cos x}{x} = \lim\limits_{x\to 0} \dfrac{1-\cos x}{x}\cdot\dfrac{1+\cos x}{1+\cos x} = \lim\limits_{x\to 0} \dfrac{1-\cos^2 x}{x(1+\cos x)} = \lim\limits_{x\to 0} \dfrac{\sin^2 x}{x(1+\cos x)} =$ $\lim\limits_{x\to 0} \dfrac{\sin x}{x} \cdot \lim\limits_{x\to 0} \sin x \cdot \lim\limits_{x\to 0} \dfrac{1}{1+\cos x} = 1\cdot 0\cdot\dfrac{1}{2} = \boxed{0}$.

□

Exercise 2.40. Find and justify $\lim\limits_{x\to 0} \dfrac{\tan^2 x + 2x}{x+x^2}$ and $\lim\limits_{x\to 0} \dfrac{x^2(3+\sin x)}{(x+\sin x)^2}$.

§2.6 Advanced Concepts

Theorem 2.41

Let $P(x)$, $Q(x)$ be polynomials such that $P(x) = \sum_{i=0}^{n} a_i x^i$ and $Q(x) = \sum_{k=0}^{m} b_k x^k$. Then,

$$\lim_{x\to\infty} \frac{P(x)}{Q(x)} = \begin{cases} \pm\infty & \deg P > \deg Q \\ \dfrac{a_n}{b_n} & \deg P = \deg Q \\ 0 & \deg P < \deg Q \end{cases}$$

Proof. If $\deg P > \deg Q$, then $n > m$, and we have

$$\lim_{x\to\infty} \frac{P(x)}{Q(x)} = \lim_{x\to\infty} \frac{a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0}{b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0}.$$

We divide the numerator and denominator by x^n. Then,

$$\lim_{x\to\infty} \frac{a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0}{b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0} = \lim_{x\to\infty} \frac{a_n + a_{n-1}x^{-1} + \ldots + a_1 x^{1-n} + a_0 x^{-n}}{b_m x^{m-n} + b_{m-1}x^{m-n-1} + \ldots + b_1 x^{1-n} + b_0 x^{-n}}.$$

As x goes to ∞, any term of x raised to a negative power goes to 0. Therefore, the denominator goes to 0 and the numerator goes to a_n, so the limit is $\pm\infty$, where the sign would depend on the sign of a_n and whether the other terms approached 0^+ or 0^-.

If $\deg P = \deg Q$, then $n = m$. We will primarily use n. We have

$$\lim_{x\to\infty} \frac{P(x)}{Q(x)} = \lim_{x\to\infty} \frac{a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0}{b_n x^n + b_{n-1}x^{n-1} + \ldots + b_1 x + b_0}.$$

We divide the numerator and denominator by x^n to get

$$\lim_{x\to\infty} \frac{P(x)}{Q(x)} = \lim_{x\to\infty} \frac{a_n + a_{n-1}x^{-1} + \ldots + a_1 x^{1-n} + a_0 x^{-n}}{b_n + b_{n-1}x^{-1} + \ldots + b_1 x^{1-n} + b_0 x^{-n}}.$$

All terms with x raised to a negative power approach 0 as x goes to ∞, so we are left with $\frac{a_n}{b_n}$ as our limit.

Lastly, if $\deg P < \deg Q$, then $n < m$. We have

$$\lim_{x\to\infty} \frac{P(x)}{Q(x)} = \lim_{x\to\infty} \frac{a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0}{b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0}.$$

Similar to the first case, we divide the numerator and denominator by x^m. Then,

$$\lim_{x\to\infty} \frac{a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0}{b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0} = \lim_{x\to\infty} \frac{a_n x^{n-m} + a_{n-1}x^{n-m-1} + \ldots + a_1 x^{1-m} + a_0 x^{-m}}{b_m + b_{m-1}x^{-1} + \ldots + b_1 x^{1-m} + b_0 x^{-m}}.$$

As x goes to ∞, all terms in the numerator go to 0, and all terms except for b_m in the denominator go to 0, so the limit is $\frac{0}{b_m} = 0$. □

Problem 2.42. Evaluate and justify the following limits:

1. $\lim_{x\to 3} \frac{x^2-7x+12}{x^3-27}$
2. $\lim_{x\to\infty} \sqrt{x^2+2x}-x$
3. $\lim_{x\to-\infty} \sqrt{x^2+2x}-x$

Solution. 1. As we cannot directly plug in and compute, we first factor and cancel out like terms: $\lim_{x\to 3} \frac{x^2-7x+12}{x^3-27} = \lim_{x\to 3} \frac{(x-3)(x-4)}{(x-3)(x^2+3x+9)} = \lim_{x\to 3} \frac{x-4}{x^2+3x+9} = \boxed{-\frac{1}{27}}$.

2. The radical motivates us to 'rationalize the numerator': $\lim_{x\to\infty} \sqrt{x^2+2x}-x = \lim_{x\to\infty} (\sqrt{x^2+2x}-x)\cdot \frac{\sqrt{x^2+2x}+x}{\sqrt{x^2+2x}+x} = \lim_{x\to\infty} \frac{2x}{\sqrt{x^2+2x}+x}$. As we are considering x going to ∞, we assume $x>0$ and divide the numerator and denominator by x, which means that we divide the inner content of the radical by x^2, as such: $\lim_{x\to\infty} \frac{2x}{\sqrt{x^2+2x}+x} = \lim_{x\to\infty} \frac{2}{\sqrt{1+\frac{2}{x}}+1} = \frac{2}{2} = \boxed{1}$.

3. As shown in the previous problem, $\lim_{x\to-\infty} \sqrt{x^2+2x}-x = \lim_{x\to-\infty} \frac{2x}{\sqrt{x^2+2x}+x}$. However, as we are considering x going to $-\infty$, we can safely assume $x<0$, so $x=-\sqrt{x^2}$, suggesting that when we divide the numerator and denominator by x, we divide the inner content of the radical by x^2 then making the radical a negative term: $\lim_{x\to-\infty} \frac{2x}{\sqrt{x^2+2x}+x} = \lim_{x\to-\infty} \frac{2}{-\sqrt{1+\frac{2}{x}}+1}$. As x goes to $-\infty$, $\sqrt{1+\frac{2}{x}} \to 1^-$, so $-\sqrt{1+\frac{2}{x}} \to -1^+$, i.e. the denominator $-\sqrt{1+\frac{2}{x}}+1$ approaches 0^+. The numerator stays at 2, so since the overall sign is positive, the limit is $\boxed{\infty}$.

□

Exercise 2.43. Evaluate and justify the following limits:

1. $\lim_{x\to 4} \frac{|x^2-16|}{x-4}$
2. $\lim_{x\to 1} \frac{\sin(3x-3)}{\sin(2x-2)}$
3. $\lim_{x\to-\infty} \sqrt{x^2+4x}-x$
4. $\lim_{x\to\infty} x(\sqrt{x+2}-\sqrt{x})$
5. $\lim_{x\to\infty} \frac{x^3+4x-7}{7x^2-x+1}$

Lemma 2.44

$\lim_{x\to\infty} f(x) = \lim_{y\to 0^+} f\left(\frac{1}{y}\right)$.

Proof. Assume $\lim_{x\to\infty} f(x) = L$, and consider its definition:

$$\forall \varepsilon > 0\ \exists N\ \forall x,\ x > N \to |f(x) - L| < \varepsilon.$$

Assume $N > 0$. Then substitute $x \to \frac{1}{y}$, so we have that $x > 0 \to \frac{1}{y} > 0$, and $y > 0$. The definition becomes

$$\forall \varepsilon > 0\ \exists N\ \forall y,\ \frac{1}{y} > N \to \left|f\left(\frac{1}{y}\right) - L\right| < \varepsilon.$$

As $N > 0$ and $\frac{1}{y} > 0$, we can take the reciprocal of $\frac{1}{y} > N$ to get $y < \frac{1}{N}$. As $\frac{1}{N}$ is some positive number, we let $\delta = \frac{1}{N}$, and it follows that

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall y,\ 0 < y < \delta \to \left|f\left(\frac{1}{y}\right) - L\right| < \varepsilon,$$

which is the definition of $\lim_{y\to 0^+} f\left(\frac{1}{y}\right) = L$, as desired. □

Exercise 2.45. Evaluate and justify $\lim_{x\to\infty} x \sin\left(\frac{1}{x}\right)$. What about $\lim_{x\to\infty} x^2 \sin\left(\frac{1}{x}\right)$?

Example 2.46

Suppose $\forall n, m \in \mathbb{N}$, A_n is a finite subset of $[0, 1]$ and for $m \neq n$, $A_m \cap A_n = \emptyset$. Now define $f : [0, 1] \to \mathbb{R}$ in the following way:

$$f(x) = \begin{cases} \frac{1}{n}, & x \in A_n \\ 0, & x \notin \bigcup_{i=1}^{\infty} A_i \end{cases}$$

Prove that $\lim_{x\to a} f(x) = 0\ \forall a \in [0, 1]$.

Proof. The idea is to exclude all of the elements that belong to A_i with small i in our interval, as those elements will yield large values of f. Formally, given an $\varepsilon > 0$, define

$$A = \bigcup_{i=1}^{\left\lceil \frac{1}{\varepsilon} \right\rceil + 1} A_i$$

to be the set of elements to exclude. Then since A is finite, we can choose an interval I centered at a that doesn't intersect A (except maybe at a itself, which doesn't matter).

Then in this interval,

$$f(x) = \begin{cases} \dfrac{1}{n} < \dfrac{1}{\lceil \frac{1}{\varepsilon} \rceil + 1} < \varepsilon, & x \in A_n \\ 0 < \varepsilon, & x \notin \text{any of } A_i \end{cases}$$

so $0 \leq f(x) < \varepsilon$ for all x in $I \implies \lim\limits_{x \to 0} f(x) = 0$. □

Example 2.47

Consider the following function:

$$f(x) = \begin{cases} 0 & x \in \mathbb{Q} \\ 1 & x \notin \mathbb{Q} \end{cases}$$

Prove $\forall a \in \mathbb{R}$, $\lim\limits_{x \to a} f(x)$ does not exist.

Proof. Assume for the sake of contradiction that $\lim\limits_{x \to a} f(x) = L$. Consider the definition,

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x,\ 0 < |x - a| < \delta \to |f(x) - L| < \varepsilon.$$

Note that $|f(x) - L| < \varepsilon$ rearranges to $f(x) \in (L - \varepsilon, L + \varepsilon)$. Assume that $\varepsilon = \dfrac{1}{4}$. Then $f(x) \in \left(L - \dfrac{1}{4}, L + \dfrac{1}{4}\right)$, an open interval of size $\dfrac{1}{2}$.

For the time being, we will have assume some properties of rational and irrational numbers that we have not proven yet. It is true that there exists a rational and irrational number within the interval $(a - \delta, a + \delta)$, given that $a \in \mathbb{R}$. Then, $f(x)$ takes on values of both 0 when $x \in \mathbb{Q}$ and 1 when $x \notin \mathbb{Q}$, according to the piecewise definition. However, both 0 and 1 cannot be both contained in the open interval $f(x) \in \left(L - \dfrac{1}{4}, L + \dfrac{1}{4}\right)$ of size $\dfrac{1}{2}$, so we have a contradiction, and there does not exist a limit L. □

Example 2.48

Consider the following function:

$$f(x) = \begin{cases} \dfrac{1}{q} & x \in \mathbb{Q} \text{ can be written as } \dfrac{p}{q} \text{ in lowest terms} \\ 1 & x \notin \mathbb{Q} \end{cases}$$

Prove that $\forall a \in \mathbb{R}$, $\lim\limits_{x \to a} f(x) = 0$.

Proof. We make the substitution $x \to x + \lfloor a \rfloor$. Then the limit becomes $\lim\limits_{x \to a - \lfloor a \rfloor} f(x + \lfloor a \rfloor)$. Now, for every natural number i, construct the set A_i as follows:

$$A_i = \left\{ \frac{p}{i} \;\middle|\; 0 \leq p \leq i\ \wedge \frac{p}{i} \text{ is in lowest terms} \right\}$$

It is clear that all of the A_i are disjoint, since each fraction can only be expressed in lowest terms in one way. Define the function

$$g(x) = f(x + \lfloor a \rfloor).$$

Now, for any $x \in [0, 1]$, let's look at how $g(x)$ behaves. If $x \in A_i$ for any i, then we know that x can be written as $\frac{p}{i}$ in lowest terms. Therefore, $x + \lfloor a \rfloor$ would also have denominator i, since $\lfloor a \rfloor$ is an integer. Thus,

$$g(x) = f(x + \lfloor a \rfloor) = \frac{1}{i}.$$

If x is not in any of the A_i, then x is not rational (why?), and so

$$g(x) = f(x + \lfloor a \rfloor) = 0.$$

Now we see that all of the conditions of Example 2.46 are met. Thus, since $a - \lfloor a \rfloor \in [0, 1]$,

$$\lim_{x \to a - \lfloor a \rfloor} f(x + \lfloor a \rfloor) = g(x) = 0,$$

which is equivalent to the desired result. □

§2.7 Continuity

Although we have mentioned the notion of continuity before, we formally define it here:

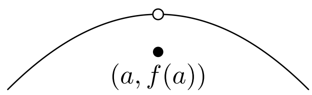
Definition 2.49. $f(x)$ is **continuous** at $x = a$ if $\lim_{x \to a} f(x) = f(a)$.

We have already proven the following results:

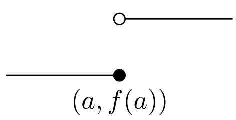
1. Polynomials are continuous everywhere, by Lemma 2.30.
2. Rational functions are continuous where they are defined, by Lemma 2.31.
3. Trigonometric functions are continuous where they are defined, by Theorem 2.35.

We have notions of different kinds of discontinuity:

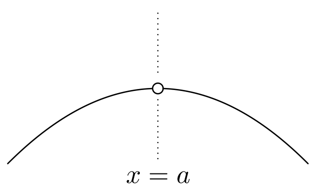
Definition 2.50. When $f(a)$ exists and $\lim_{x \to a} f(x) \neq f(a)$, there is a **removable discontinuity** at $x = a$.



Definition 2.51. When $f(a)$ exists and $\lim_{x \to a} f(x)$ does not exist, there is a **essential discontinuity** at $x = a$.



There can also be a discontinuity when $f(a)$ does not exist at all.



Problem 2.52. Is $f(x) = \dfrac{\sin x}{x}$ continuous? If not, what modifcations can we make to f such that it is continuous?

Solution. Note that $\dfrac{\sin x}{x}$ does not exist at $x = 0$, so we have a discontinuity. We can define the piecewise function

$$f(x) = \begin{cases} \dfrac{\sin x}{x} & x \neq 0 \\ 1 & x = 0 \end{cases}$$

and since $\lim_{x \to 0} \dfrac{\sin x}{x} = 1$, the function is now continuous. $\square$

Theorem 2.53 (Continuity of Composite Functions)
If g is continuous at a and f is continuous at $g(a)$, then $f \circ g$ is continuous at a.

Proof. Given that g is continuous at a, it is true that $\lim_{x \to a} g(x) = g(a)$, i.e.

$$\forall \varepsilon_1 > 0 \; \exists \delta_1 > 0 \; \forall x, \; 0 < |x - a| < \delta_1 \to |g(x) - g(a)| < \varepsilon_1.$$

We are given that f is continuous at $g(a)$, therefore $\lim_{y \to g(a)} f(y) = f(g(a))$, i.e.

$$\forall \varepsilon_2 > 0 \; \exists \delta_2 > 0 \; \forall y, \; 0 < |y - g(a)| < \delta_2 \to |f(y) - f(g(a))| < \varepsilon_2.$$

Let $y = g(x)$, so the definition above can be rewritten as

$$\forall \varepsilon_2 > 0 \; \exists \delta_2 > 0 \; \forall x, \; 0 < |g(x) - g(a)| < \delta_2 \to |f(g(x)) - f(g(a))| < \varepsilon_2.$$

Our objective is to show that $0 < |x - a| < \delta_1 \to |f(g(x) - f(g(a))| < \varepsilon_2$ from the two given definitions. It suffices to let $\varepsilon_1 = \delta_2$, then it follows that $0 < |x - a| < \delta_1 \to |g(x) - g(a)| < \delta_2$. By the rewritten form of the second given definition, $0 < |g(x) - g(a)| < \delta_2 \to |f(g(x)) - f(g(a))| < \varepsilon_2$. Then, by hypothetical syllogism,

$$\forall \varepsilon_2 > 0 \; \exists \delta_1 > 0 \; \forall x, \; 0 < |x - a| < \delta_1 \to |f(g(x)) - f(g(a))| < \varepsilon_2,$$

so $\lim_{x \to a} f(g(x)) = f(g(a))$, i.e. $f \circ g$ is continuous at a, as desired. $\square$

Exercise 2.54. Evaluate $\lim_{x \to \pi} \sin^3(7x)$.

Problem 2.55. Prove $f(x) = \sqrt{x}$ is continuous $\forall x > 0$.

Proof. For a given ε, let $\delta < \varepsilon\sqrt{a}$, assuming $a > 0$. Note that

$$|x - a| < \delta \longrightarrow \left|\sqrt{x} - \sqrt{a}\right| = \left|\sqrt{x} - \sqrt{a}\right| \left|\frac{\sqrt{x} + \sqrt{a}}{\sqrt{x} + \sqrt{a}}\right| = \frac{|x - a|}{\sqrt{x} + \sqrt{a}} \leq \frac{|x - a|}{\sqrt{a}} < \frac{\delta}{\sqrt{a}} < \varepsilon.$$

Therefore, $\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x, \; |x - a| < \delta \longrightarrow |\sqrt{x} - \sqrt{a}| < \varepsilon$, i.e. $\lim_{x \to a} \sqrt{x} = \sqrt{a}$. $\square$

Exercise 2.56. Prove that $f(x) = \sqrt{x}$ is right-continuous at 0, i.e. $\lim_{x \to 0^+} f(x) = f(0)$.

Theorem 2.57

Suppose that f is continuous at $x = a$ and $f(a) > 0$. Then $\exists \delta > 0$ such that $\forall x \in (a - \delta, a + \delta)$, $f(x) > 0$.

Proof. Assuming f is continuous at $x = a$, then $\lim_{x \to a} f(x) = f(a)$, i.e.

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x, \; 0 < |x - a| < \delta \to |f(x) - f(a)| < \varepsilon.$$

Now consider $\varepsilon = \dfrac{f(a)}{2}$, as $f(a) > 0$ implies that $\dfrac{f(a)}{2} > 0$, and ε can be any positive number. Then, by the definition,

$$\exists \delta > 0 \; \forall x, \; 0 < |x - a| < \delta \to |f(x) - f(a)| < \frac{f(a)}{2}.$$

We expand the absolute inequalities to get

$$\exists \delta > 0 \; \forall x, \; a - \delta < x < a + \delta \to \frac{f(a)}{2} < f(x) < \frac{3f(a)}{2},$$

and we rewrite this based on the statement of the theorem, as such:

$$\exists \delta > 0 \; \forall x \in (a - \delta, a + \delta), \; \frac{f(a)}{2} < f(x) < \frac{3f(a)}{2}.$$

We have $f(x) > \dfrac{f(a)}{2}$, which clearly implies that $f(x) > 0$, so we are done. □

Theorem 2.58

Suppose that f is continuous at $x = a$ and $f(a) < 0$. Then $\exists \delta > 0$ such that $\forall x \in (a - \delta, a + \delta)$, $f(x) < 0$.

Proof. The proof is virtually identical to that of Theorem 2.57. □

Example 2.59 (Cauchy's Functional Equation)

Let f be a continuous function on $\mathbb{R}$ for which $f(x + y) = f(x) + f(y) \; \forall x, y$. Prove $f(x) = cx$ for some $c \in \mathbb{R}$.

Proof. Consider $x = 0$. Then $f(0 + y) = f(0) + f(y) \implies f(0) = 0$. Then substitute in $x = -y$ to get $f(-y + y) = f(-y) + f(y) \implies f(-y) = -f(y)$, so f is an odd function.

Now, we can show that

$$f(\underbrace{x + x + \ldots + x}_{m \text{ times}}) = \underbrace{f(x) + f(x) + \ldots + f(x)}_{m \text{ times}}, \; m \in \mathbb{N},$$

using an inductive argument. Thus, $\forall m \in \mathbb{N}, \; f(mx) = mf(x)$.

Similarly,

$$\underbrace{f\left(\frac{m}{n}x + \frac{m}{n}x + \ldots + \frac{m}{n}x\right)}_{n \text{ times}} = \underbrace{f\left(\frac{m}{n}x\right) + f\left(\frac{m}{n}x\right) + \ldots + f\left(\frac{m}{n}x\right)}_{n \text{ times}}, \; m, n \in \mathbb{N},$$

so we have that $f\left(n\cdot\frac{m}{n}x\right) = nf\left(\frac{m}{n}x\right)$, $\forall m,n\in\mathbb{N}$.

Note that $f\left(n\cdot\frac{m}{n}x\right) = f(mx) = mf(x)$, so $mf(x) = nf\left(\frac{m}{n}x\right)$, i.e. $f\left(\frac{m}{n}x\right) = \frac{m}{n}f(x)$, $\forall m,n\in\mathbb{N}$. Using the fact that f is odd, it follows that this identity is true for negative n,m as well, so we conclude that

$$f(qx) = qf(x), \forall q\in\mathbb{Q},$$

which is a substantial result. Letting $x=1$, we get $f(q) = qf(1)$. Since $f(1)$ is constant, let $c = f(1)$ for some $c\in\mathbb{R}$, since f is defined over $\mathbb{R}$. Thus, we have

$$f(q) = cq\ \forall q\in\mathbb{Q}.$$

Now we want to show that this is true on $\mathbb{R}$. Consider $g(x) = f(x) - cx$. If $q\in\mathbb{Q}$, then $g(q) = 0$. Now consider an irrational number, i.e. let $\Omega\in\overline{\mathbb{Q}}$. Since f is continuous on $\mathbb{R}$, we have that g is also continuous on $\mathbb{R}$, so it must be continuous at Ω. Then,

$$\forall\varepsilon>0\ \exists\delta>0\ \forall x,\ |x-\Omega|<\delta\to|g(x)-g(\Omega)|<\varepsilon.$$

We will assume that for any δ, we can find x in the interval $(\Omega-\delta,\Omega+\delta)$ that is rational. A rigorous proof of this assumption will be supplied in the next chapter. Consider such a rational x. Then $g(x) = 0$, so we have $|x-\Omega|<\delta\to|g(\Omega)|<\varepsilon$. But since ε a given positive number, $|g(\Omega)|<\varepsilon$ implies $g(\Omega)=0$, i.e. $f(\Omega)-c\Omega=0$, or $f(\Omega)=c\Omega$.

Thus, $f(x) = cx$ for all rational and irrational x, so $f(x) = cx$ on $\mathbb{R}$. $\square$

§3 Three Hard Theorems

§3.1 Introduction

We will now introduce the following three theorems related to continuous functions. Their proofs will not be given until later.

Theorem 3.1 (Intermediate Value Theorem)
Let $f(x)$ be continuous on $[a,b]$ and suppose $f(a) < 0 < f(b)$. Then $\exists a < c < b$ such that $f(c) = 0$.

Theorem 3.2
Let $f(x)$ be continuous on $[a,b]$. Then $f(x)$ is bounded on $[a,b]$, i.e. $\exists c_1, c_2$ such that $\forall x \in [a,b]$, $c_1 < f(x) < c_2$.

Theorem 3.3
Let $f(x)$ be continuous on $[a,b]$. Then $f(x)$ has a maximum and a minimum on $[a,b]$.

Example 3.4
Show that these theorems fail when $f(x)$ is defined over $\mathbb{Q}$.

Solution. First, consider the function $f(x) = x^2 - 3$ on $[0, 10]$. Note that $f(0) = -3$ and $f(10) = 97$, therefore by Theorem 3.1, there should exist a c such that $f(c) = 0$. However, there does not exist a rational for which $x^2 - 3 = 0$, and we prove this as follows:

Proof. Suppose $\frac{p}{q}$ is a fraction in lowest terms for which $\left(\frac{p}{q}\right)^2 = 3$. We rearrange this to get $p^2 = 3q^2$, which implies that $3 \mid p^2$, or $3 \mid p$. Since p is some multiple of 3, let $p = 3\widetilde{p}$. Then $9\widetilde{p}^2 = 3q^2$, which reduces to $3\widetilde{p}^2 = q^2$, i.e. $3 \mid q^2$, or $3 \mid q$. As $3 \mid p$ and $3 \mid q$, we contradict the assumption that $\frac{p}{q}$ is in lowest terms, therefore there does not exist any rational number whose square is equal to 3. $\square$

Since there does not exist $c \in \mathbb{Q}$ for which $f(c) = 0$, we conclude that Theorem 3.1 fails over the rationals.

Let $g(x) = \frac{1}{x^2 - 2}$. We will show that this function fails both Theorem 3.2 and Theorem 3.3. Note that this function is always continuous over the rationals, because the denominator cannot equal 0 (the proof of this is analogous to what was demonstrated earlier).

We will show that $g(x)$ is unbounded, by claiming that we can always take a rational number whose square is arbitrarily close to 2, such that the fraction $\frac{1}{x^2 - 2}$ can be made arbitrarily large. The proof of this claim is as follows:

Proof. Suppose $\dfrac{m^2}{n^2} < 2$. Then it follows that

$$\begin{aligned} m^2 &< 2n^2 \\ m^2 + (m^2 + 4mn + 2n^2) &< 2n^2 + (m^2 + 4mn + 2n^2) \\ 2m^2 + 4mn + 2n^2 &< m^2 + 4mn + 4n^2 \\ 2(m+n)^2 &< (m+2n)^2 \\ \therefore \frac{(m+2n)^2}{(m+n)^2} &> 2, \end{aligned}$$

i.e. $\left(\dfrac{m}{n}\right)^2 < 2 \to \left(\dfrac{m+2n}{m+n}\right)^2 > 2$.

It is also true that $\left(\dfrac{m}{n}\right)^2 > 2 \to \left(\dfrac{m+2n}{m+n}\right)^2 < 2$, and the proof of this is nearly identical.

Now we compare the distances from $\left(\dfrac{m}{n}\right)^2$ and $\left(\dfrac{m+2n}{m+n}\right)^2$ to 2 respectively. Consider the first implication $\left(\dfrac{m}{n}\right)^2 < 2 \to \left(\dfrac{m+2n}{m+n}\right)^2 > 2$. Then we seek to compare the distances $2 - \left(\dfrac{m}{n}\right)^2$ and $\left(\dfrac{m+2n}{m+n}\right)^2 - 2$. These rearrange to $\dfrac{2n^2 - m^2}{n^2}$ and $\dfrac{2n^2 - m^2}{(m+n)^2}$ respectively, and it is clear that $\dfrac{2n^2 - m^2}{(m+n)^2} < \dfrac{2n^2 - m^2}{n^2}$, therefore the distance from $\left(\dfrac{m+2n}{m+n}\right)^2$ to 2 is less than the distance from $\left(\dfrac{m}{n}\right)^2$ to 2. In other words, $\left(\dfrac{m+2n}{m+n}\right)^2$ is closer to 2 than $\left(\dfrac{m}{n}\right)^2$.

It can be analogously shown that the same result follows for the second implication $\left(\dfrac{m}{n}\right)^2 > 2 \to \left(\dfrac{m+2n}{m+n}\right)^2 < 2$.

Now, let's see what happens when you repeat the above process twice. We go from $\dfrac{m}{n}$ to $\dfrac{m+2n}{m+n}$ to $\dfrac{3m+4n}{2m+3n}$. Note that, since the process changes which side of 2 the square is on, $\left(\dfrac{m}{n}\right)^2$ and $\left(\dfrac{3m+4n}{2m+3n}\right)^2$ are both on the same side of 2. WLOG assume that both of them are less than 2 (this proof is identical if both of them are greater). Now, let's compare $2 - \left(\dfrac{m}{n}\right)^2$ and $2 - \left(\dfrac{3m+4n}{2m+3n}\right)^2$. Note that

$$\begin{aligned} 2 - \left(\frac{m}{n}\right)^2 &= \frac{2n^2 - m^2}{n^2} \\ 2 - \left(\frac{3m+4n}{2m+3n}\right)^2 &= \frac{2n^2 - m^2}{(2m+3n)^2} < \frac{2n^2 - m^2}{(3n)^2} = \frac{1}{9} \cdot \frac{2n^2 - m^2}{n^2}. \end{aligned}$$

In other words, for any $\dfrac{m}{n}$, we have found another rational whose square is nine times closer to 2. We can keep on repeating this process, and get arbitrarily close to 2 (since if we keep on dividing the distance by 9, the distance gets arbitrarily small). Thus, we can get the square of a rational number as close as we want to 2, concluding the proof. $\square$

As $x \in \mathbb{Q}$ can be made arbitrarily close to 2, the overall function can be made arbitrarily large, therefore it is unbounded and cannot have a maximum, so it fails both Theorem 3.2 and Theorem 3.3. $\square$

Corollary 3.5

Any polynomial of odd degree has a root.

Proof. Let n be odd and $P(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$. We assume the polynomial is monic because given a polynomial that is not monic, we can always divide the rest of the coefficients of the polynomial by the leading coefficient.

For this proof, we will assume that x^n with odd n can either be negative or positive.

We want to show that $P(x)$ can be sometimes negative and sometimes positive, then we can apply Theorem 3.1 to conclude that $P(x)$ does have a root.

It suffices to show that for large $|x|$, the function $P(x)$ approaches the behavior of x^n. We factor out x^n to obtain the expression

$$P(x) = x^n\left(1 + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \ldots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}\right).$$

Assume $|x| > 1$. Then,

$$\begin{aligned}\left|\frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \ldots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}\right| &\le \frac{|a_{n-1}|}{|x|} + \frac{|a_{n-2}|}{|x|^2} + \ldots + \frac{|a_1|}{|x|^{n-1}} + \frac{|a_0|}{|x|^n} \\ &\le \frac{|a_{n-1}| + |a_{n-2}| + \ldots + |a_1| + |a_0|}{|x|}.\end{aligned}$$

We want to show that

$$\left|\frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \ldots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}\right| < 1,$$

so we could then conclude that $1 + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \ldots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n} > 0$, i.e. as $|x|$ approaches large values, $P(x)$ becomes a product of x^n and some positive quantity. Then, it would follow that $P(x)$ can be sometimes negative and positive, depending on x^n.

It is sufficient to assume $|x| > |a_{n-1}| + |a_{n-2}| + \ldots + |a_1| + |a_0|$, then

$$\left|\frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \ldots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}\right| \le \frac{|a_{n-1}| + |a_{n-2}| + \ldots + |a_1| + |a_0|}{|x|} < 1.$$

Keeping in mind the idea that we are considering large $|x|$, we have assumed $|x| > \max\{1, |a_{n-1}| + |a_{n-2}| + \ldots + |a_1| + |a_0|\}$, so we would demonstrate that $P(x)$ follows the behavior of x^n for odd n. $\square$

Corollary 3.6

Any polynomial of even degree has an extremum.

Proof. Let n be even and $P(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x_1$. Note that we have omitted the constant term because if we have shown that $P(x)$ has an extremum, then $P(x) + c$ for any constant c will also have an extremum.

Again, we consider the rewritten form

$$P(x) = x^n \left(1 + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \ldots + \frac{a_1}{x^{n-1}}\right).$$

By a similar process used in the proof of Corollary 3.5, we have that $P(x)$ tends toward x^n when $|x|$ is sufficiently large. However, since we are considering even n in this proof, note that $x^n > 0$. Therefore, for sufficiently large $|x|$, $P(x) > 0$.

Let some large number M denote this threshold for being 'sufficiently large,' such that for $|x| > M$, $P(x) > 0$. Then consider $|x| \le M$, i.e. $P(x)$ on the interval $[-M, M]$. As we have omitted the constant term, $P(0) = 0$. By Theorem 3.3, $P(x)$ has a minimum on $[-M, M]$. As $x = 0$ is in this interval, it follows that this minimum will be less than or equal to $P(0) = 0$. However, we know that $P(x) > 0$ for any x not in $[-M, M]$, therefore this minimum is a minimum for P on all of $\mathbb{R}$, so therefore $P(x)$ has an extremum. □

Problem 3.7. Let $f(x)$ be a continuous function with domain and range $[0, 1]$. Prove $\exists 0 \le c \le 1$ for which $f(c) = c$.

Proof. Consider $g(x) = f(x) - x$. Then it is sufficient to prove that $g(c) = 0$. This motivates us to take advantage of Theorem 3.1. First, by the given restrictions on domain and range of f, we know that

$$g(0) = f(0) - 0 \ge 0,$$
$$g(1) = f(1) - 1 \le 0.$$

If $g(0) = 0$, then let $c = 0$. Similarly, if $g(1) = 0$, then let $c = 1$. In either case, we have shown that $\exists 0 \le c \le 1$ such that $g(c) = 0$.

Otherwise, we must have $c \ne 0, 1$, and consider $g(0) > 0$ and $g(1) < 0$, i.e. $g(1) < 0 < g(0)$. By Theorem 3.1, $\exists 1 < c < 0$ (i.e. $0 < c < 1$) such that $g(c) = 0$, and we are done. □

§3.2 Supremum, Infimum

Consider $f(x) = x^2 - 2$ on $U = [1, 2] \cap \mathbb{Q}$. Where lies the fallacy in these assumptions?

As demonstrated in Example 3.4, there does not exist a rational whose square is equal to 2. As $f(1) = -1$ and $f(2) = 2$, Theorem 3.1 fails.

Clearly, there must be something wrong with the rational numbers. What distinguishes the reals from the rationals? We introduce the following concepts related to set theory to explore this discrepancy.

Definition 3.8. Let X be a set of numbers. Then,

- m is an **upper bound** of X if $\forall x \in X,\ x \le m$.
- n is a **lower bound** of X if $\forall x \in X,\ x \ge n$.

Definition 3.9. A set which has an upper bound is called **bounded above**.

A set which has a lower bound is called **bounded below**.

Definition 3.10. Given a set X that is bounded above, its **least upper bound** or its **supremum** is the smallest number that serves as an upper bound of X. It is denoted as $\sup X$.

Given a set X that is bounded below, its **greatest lower bound** or its **infimum** is the greatest number that serves as an lower bound of X. It is denoted as $\inf X$.

Now, we establish the axiom that represents one of the most important properties of the real numbers. The statement may seem underwhelming, but its significance will soon be demonstrated.

Axiom 3.11. Every non-empty set that is bounded above has a supremum.

Theorem 3.12

Every non-empty set that is bounded below has an infimum.

Proof. Let X be a non-empty set that is bounded below. Define $-X = \{-x \mid x \in X\}$. Let m be a lower bound for X. By Definition 3.8, $\forall x \in X,\ x \geq m$. This implies $\forall -x \in -X$, $-x \leq -m$, i.e. by Definition 3.8, $-X$ is bounded above. Then by Axiom 3.11, $-X$ must have a supremum. Let $\sup(-X) = a$. Then by Definition 3.10, $\forall -x \in -X,\ -x \leq a \leq m$, which suggests that $\forall x \in X,\ x \geq -a \geq -m$, so $-a$ is the infimum of X. □

Theorem 3.13

Let X be non-empty and bounded above. Then

$$m = \sup X \longleftrightarrow (m \text{ is an upper bound of } X) \wedge (\forall \varepsilon > 0\ \exists x \in X \text{ s.t. } m - x < \varepsilon)\,.$$

Proof. First, we prove the right direction. Assume for the sake of contradiction that $\exists \varepsilon > 0\ \forall x \in X,\ m - x \geq \varepsilon$. As we are given that $m = \sup X$, we know that $\forall x \in X,\ x \leq m$. However, by rearranging our first assumption, $\forall x \in X,\ x \leq m - \varepsilon$, and $m - \varepsilon$ is less than m. Then $m - \varepsilon$ would be a smaller upper bound, which contradicts that m is the supremum.

For the left direction, assume for the sake of contradiction that m is not the supremum, i.e. there exists some upper bound n such that $n < m$, which rearranges to $m - n > 0$.

We are given that $\forall \varepsilon > 0\ \exists x \in X,\ m - x < \varepsilon$. As $m - n$ is a positive quantity, assume $\varepsilon < m - n$. Then $\exists x \in X,\ m - x < \varepsilon < m - n$, which simplifes to $x > n$. But this contradicts the assumption that n is an upper bound (by Definition 3.8, $\forall x \in X,\ x \leq n$). Therefore, m is the supremum of X. □

Theorem 3.14

Let X be non-empty and bounded below. Then

$$m = \inf X \longleftrightarrow (m \text{ is a lower bound of } X) \wedge (\forall \varepsilon > 0\ \exists x \in X \text{ s.t. } x - m < \varepsilon)\,.$$

Proof. We prove the right direction. Assume for the sake of contradiction that $\exists \varepsilon > 0\ \forall x \in X,\ x - m \geq \varepsilon$, or $x \geq m + \varepsilon$. We are given that $m = \inf X$, so $\forall x \in X,\ x \geq m$. However, $m + \varepsilon$ is a lower bound and $m + \varepsilon > m$. Then, $m + \varepsilon$ is a greater lower bound, which contradicts that m is the infimum.

For the left direction, assume for the sake of contradiction that m is not the infimum, i.e. there exists some lower bound n such that $n > m$, which rearranges to $n - m > 0$.

We are given that $\forall \varepsilon > 0\ \exists x \in X,\ x - m < \varepsilon$. As $n - m$ is a positive quantity, assume $\varepsilon < n - m$ (assuming $\varepsilon = n - m$ works fine as well). Then $\exists x \in X,\ x - m < \varepsilon < n - m$, which simplifes to $x < n$. But this contradicts the assumption that n is a lower bound (by Definition 3.8, $\forall x \in X,\ x \geq n$). Therefore, m is the infimum of X. □

Example 3.15

Let $X = \left\{x \in \mathbb{Q}^+ \;\middle|\; x^2 < 2\right\}$. Prove $\sup X$ exists, and prove $(\sup X)^2 = 2$.

Proof. To demonstrate the existence of a supremum, we can use Axiom 3.11. It is sufficient to show that X is non-empty and bounded above. Clearly, $1 \in X$, so X is non-empty. Then, we claim that 4 is an upper bound: assume for the sake of contradiction that $\forall x \in X,\ x > 4$. Then $x^2 > 16$, which contradicts that $x^2 < 2$ as x is a member of set X. Therefore, $\forall x \in X,\ x \leq 4$, so 4 is an upper bound. Since X is non-empty and bounded above, there exists a supremum, i.e. $\sup X$.

For the second proof, let $z = \sup X$. We want to prove that $z^2 = 2$.

1. Suppose $z^2 < 2$. As shown earlier in the proof of Example 3.4, we can always find a rational number whose square is closer to 2 than z. By this fact, $\exists \widetilde{z} \in \mathbb{Q}^+$ such that $\widetilde{z}^2 < 2$ and $z < \widetilde{z}$. However, the former implies that $\widetilde{z} \in X$ and the latter implies that z is not an upper bound of X, so we arrive at a contradiction.

2. Suppose $z^2 > 2$. By similar reasoning, $\exists \widetilde{z} \in \mathbb{Q}^+$ such that $\widetilde{z}^2 > 2$ and $\widetilde{z} < z$. Then it must follow that $\widetilde{z}$ is a lower upper bound than z, which contradicts the assumption that z is the supremum of X.

Hence, $z^2 = 2$. $\square$

Definition 3.16. We denote $A < B$ if $\forall a \in A,\ \forall b \in B,\ a < b$.

Example 3.17

Let A, B be non-empty sets of numbers. Prove that if $A < B$, then $\sup A \leq \inf B$.

Proof. First, we start by proving that the supremum of A and the infimum of B exist. Note that any element of B is an upper bound of A, so by Axiom 3.11, $\sup A$ exists. Similarly, any element of A is a lower bound of B, so by Theorem 3.12, $\inf B$ exists.

Assume for the sake of contradiction that $\sup A > \inf B$. Then, by Theorem 3.13 and Theorem 3.14,

$$\begin{aligned} \forall \varepsilon > 0,\ \exists a \in A,\ &\sup A - a < \varepsilon, \\ \exists b \in B,\ &b - \inf B < \varepsilon. \end{aligned}$$

Note that $\sup A - \inf B > 0$ by assumption, so consider $\varepsilon = \dfrac{\sup A - \inf B}{2}$. Then,

$$\begin{aligned} \exists a \in A,\ \sup A - a &< \frac{\sup A - \inf B}{2}, \\ \exists b \in B,\ b - \inf B &< \frac{\sup A - \inf B}{2}. \end{aligned}$$

Adding the two inequalities yields $\sup A - \inf B - a + b < \sup A - \inf B$, which simplifies to $a > b$, which contradicts the given assumption that any element in A is less than any element in B. Therefore, $\sup A \leq \inf B$. $\square$

Theorem 3.18

For $A < B$, $\sup A = \inf B \longleftrightarrow \forall \varepsilon > 0\ \exists a \in A,\ b \in B,\ \text{s.t. } b - a < \varepsilon$.

Proof. First, we will prove the right direction. As established in the proof of Example 3.17, $\sup A$ and $\inf B$ exist. Let $z = \sup A = \inf B$. Then, by Theorem 3.13 and Theorem 3.14,

$$\forall \varepsilon > 0,\ \exists a \in A,\ z - a < \frac{\varepsilon}{2},$$
$$\exists b \in B,\ b - z < \frac{\varepsilon}{2}.$$

We can add the inequalities together, and it follows that $\forall \varepsilon > 0,\ \exists a \in A,\ b \in B$ such that $(b - z) + (z - a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$, i.e. $b - a < \varepsilon$, as desired.

For the left direction, it suffices to show that $\sup A$ cannot be strictly less than $\inf B$, as we have already established in Example 3.17 that $\sup A \leq \inf B$. Assume for the sake of contradiction that $\sup A < \inf B$, or $\inf B - \sup A > 0$. Then consider $\varepsilon = \inf B - \sup A$, so $\exists a \in A,\ b \in B$ such that $b - a < \inf B - \sup A$. However, by Definition 3.9, we know that $\forall b \in B$, $b \geq \inf B$ and $\forall a \in A$, $a \leq \sup A$, i.e. $-a \geq -\sup A$, so we add these to get $b - a \geq \inf B - \sup A\ \forall a \in A$, $b \in B$, which contradicts that there exists $a \in A$ and $b \in B$ such that $b - a < \inf B - \sup A$. Therefore, $\sup A = \inf B$. □

Alternative Proof. We mostly proceed with the same steps as the first given proof, but for the left direction, one can use Theorem 3.13 instead of a proof by contradiction.

Note that $\forall b \in B,\ b \geq \inf B$, or $b - a \geq \inf B - a$. However, given $\varepsilon > 0$, we can find $a \in A$ and $b \in B$ such that $b - a < \varepsilon$, so $\inf B - a \leq b - a < \varepsilon$, or $\inf B - a < \varepsilon$, which is just the symbolic definition of the supremum of A by Theorem 3.13. Therefore, $\inf B = \sup A$. □

Problem 3.19. Let A, B be sets. Prove that if $\sup A = \sup B$, then $\forall \varepsilon > 0\ \exists a \in A,\ b \in B$ such that $|a - b| < \varepsilon$.

Proof. Let $M = \sup A = \sup B$. By Theorem 3.13, we have

$$\forall \varepsilon > 0,\ \exists a \in A,\ M - a < \varepsilon \to a > M - \varepsilon,$$
$$\exists b \in B,\ M - b < \varepsilon \to b > M - \varepsilon.$$

As the supremum is an upper bound, we also have

$$\forall x \in A,\ x \leq M,$$
$$\forall y \in B,\ y \leq M.$$

So given $\varepsilon > 0$, there exist $a \in A$, $b \in B$ such that

$$M - \varepsilon < a \leq M,$$
$$M - \varepsilon < b \leq M.$$

We add the inequalities $M - \varepsilon < a$ and $b \leq M$ to get $M - \varepsilon + b < a + M$, i.e. $a > b - \varepsilon$.

Likewise, we add the inequalities $a \leq M$ and $M - \varepsilon < b$ to get $a + M - \varepsilon < M + b$, i.e. $a < b + \varepsilon$.

Combining our results, given $\varepsilon > 0$ we have found $a \in A,\ b \in B$ such that $b - \varepsilon < a < b + \varepsilon$, which rearranges to $|a - b| < \varepsilon$, and we are done. □

Problem 3.20. Let A, B be non-empty sets of numbers, bounded above. Define $A + B = \left\{a + b \;\middle|\; a \in A,\ b \in B\right\}$. Prove $\sup(A + B) = \sup A + \sup B$.

Proof. By Axiom 3.11, $\sup A$ and $\sup B$ exist. It is not hard to show that $\sup(A + B)$ exists as well. It suffices to demonstrate $\sup(A + B) \leq \sup A + \sup B$ and $\sup(A + B) \geq \sup A + \sup B$, and the result would follow.

By Definition 3.8, we have

$$\forall a \in A,\ a \leq \sup A,$$
$$\forall b \in B,\ b \leq \sup B.$$

Then, $\forall a \in A,\ b \in B,\ a + b \leq \sup A + \sup B$, i.e. $\sup A + \sup B$ is an upper bound of the set $A + B$. Therefore, $\sup(A + B) \leq \sup A + \sup B$.

To get the other inequality, assume for the sake of contradiction that $\sup(A + B) < \sup A + \sup B$. Then there is some 'gap' between the two quantities, so let $\varepsilon = \sup A + \sup B - \sup(A + B)$ which is clearly positive by assumption. By Theorem 3.13,

$$\exists a \in A,\ \sup A - a < \frac{\varepsilon}{2},$$
$$\exists b \in B,\ \sup B - b < \frac{\varepsilon}{2}.$$

Then, $\forall \varepsilon > 0\ \exists a \in A,\ b \in B$ such that $(\sup A - a) + (\sup B - b) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$, which rearranges to $a + b > \sup A + \sup B - \varepsilon$. Substituting in $\varepsilon = \sup A + \sup B - \sup(A + B)$, we get $a + b > \sup A + \sup B - (\sup A + \sup B - \sup(A + B))$, or $a + b > \sup(A + B)$, and we arrive at a contradiction. Thus, $\sup(A + B) \geq \sup A + \sup B$.

We then conclude that $\sup(A + B) = \sup A + \sup B$. □

Example 3.21

Prove Theorem 3.1.

Proof. Define $S = \left\{x \in [a, b] \;\middle|\; f(x) < 0\right\}$. Note that S is non-empty because $a \in S$, and it is bounded above by b. Therefore, by Axiom 3.11, S has a supremum. Let $c = \sup S$. As b is an upper bound of S, we know that $c \leq b$. We will show that $f(c) = 0$.

1. Suppose $f(c) < 0$. As f is continuous at c, by Theorem 2.58, $\exists \delta > 0$ such that $\forall x \in (c - \delta, c + \delta),\ f(x) < 0$. Then $f(x) < 0$ for $x = c + \frac{\delta}{2}$, so $c + \frac{\delta}{2} \in S$. Therefore, c is not an upper bound of S, which contradicts that c is the supremum.

2. Suppose $f(c) > 0$. As f is continuous at c, by Theorem 2.57, $\exists \delta > 0$ such that $\forall x \in (c - \delta, c + \delta),\ f(x) > 0$. All x in that interval would not be in S, so consider $x = c - \frac{\delta}{2}$. However, $c - \frac{\delta}{2}$ would be a lower upper bound of S than c, which contradicts that c is the supremum.

As we cannot have $f(c) < 0$ or $f(c) > 0$, we conclude that $f(c) = 0$. □

Example 3.22

Prove Theorem 3.2.

Proof. Define $S = \left\{x \in [a,b] \;\middle|\; f \text{ is bounded on } [a,x]\right\}$. Note that $a \in S$ and b is an upper bound of S, so by Axiom 3.11, S has a supremum. Let $c = \sup S$, and it follows that $c \le b$. We consider each case separately.

Suppose $c < b$. As f is continuous on c, we know that $\lim_{x\to c} f(x) = f(c)$, i.e.

$$\forall\varepsilon > 0\ \exists\delta > 0\ \forall x \in (c-\delta, c+\delta),\ |f(x) - f(c)| < \varepsilon.$$

Consider $\varepsilon = 1$. Then,

$$\exists\delta > 0\ \forall x \in (c-\delta, c+\delta),\ |f(x) - f(c)| < 1.$$

Note that $|f(x) - f(c)| < 1$ rearranges to $f(c) - 1 < f(x) < f(c) + 1$, which implies that f is bounded $\forall x \in (c-\delta, c+\delta)$. Thus, f is bounded on $\left[c - \frac{\delta}{2}, c + \frac{\delta}{2}\right]$. Next, consider the following lemma.

Lemma 3.23
$S = \left\{x \in [a,b] \;\middle|\; x < c\right\}$.

Proof. Consider $\forall\widetilde{c} < c$ and let $\varepsilon = c - \widetilde{c} > 0$. Since $c = \sup S$, by Theorem 3.13, $\exists c_1 \in S$ such that $c - c_1 < \varepsilon$, i.e. $c - c_1 < c - \widetilde{c}$ which rearranges to $c_1 > \widetilde{c}$.

Since $c_1 \in S$, f is bounded on $[a, c_1]$, and since $[a,\widetilde{c}] \subset [a, c_1]$, f is bounded on $[a,\widetilde{c}]$. Therefore, $\widetilde{c} \in S$. □

As $c - \frac{\delta}{2}$ is less than c, by Lemma 3.23, $c - \frac{\delta}{2} \in S$, therefore f is bounded on $\left[a, c - \frac{\delta}{2}\right]$.

As f is bounded on the intervals $\left[a, c - \frac{\delta}{2}\right]$ and $\left[c - \frac{\delta}{2}, c + \frac{\delta}{2}\right]$, we conclude that f is bounded on $\left[a, c + \frac{\delta}{2}\right]$. However, this implies $c + \frac{\delta}{2} \in S$, so c is not an upper bound, which contradicts that $c = \sup S$. Therefore, c cannot be less than b.

We must have $c = b$. Note that f is continuous at b, i.e. $\lim_{x\to b} f(x) = f(b)$, so

$$\forall\varepsilon > 0\ \exists\delta > 0\ \forall x \in (b-\delta, b+\delta),\ |f(x) - f(b)| < \varepsilon.$$

Like before, consider $\varepsilon = 1$, so we have that $\exists\delta > 0$ such that $\forall x \in (b-\delta, b+\delta)$, $f(b) - 1 < f(x) < f(b) + 1$, i.e. $f(x)$ is bounded $\forall x \in (b-\delta, b+\delta)$. Then we have that f is bounded on $\left[b - \frac{\delta}{2}, b\right]$.

By Lemma 3.23, we have that f is bounded on $\left[a, c - \frac{\delta}{2}\right]$. Since $c = b$, it follows that f is bounded on $\left[a, b - \frac{\delta}{2}\right]$. As f is also bounded on $\left[b - \frac{\delta}{2}, b\right]$, we have that f is bounded on $[a,b]$, as desired. □

Definition 3.24. We define $\mathbb{N}$ inductively as follows:

- $1 \in \mathbb{N}$.
- $\forall n \in \mathbb{N}$, if $n \in \mathbb{N}$, then $n + 1 \in \mathbb{N}$.

Problem 3.25. Prove $\mathbb{N}$ is unbounded above.

Proof. Assume for the sake of contradiction that $\mathbb{N}$ is bounded above. As $1 \in \mathbb{N}$, the set is nonempty, so by Axiom 3.11, $\mathbb{N}$ has a supremum. Let $c = \sup \mathbb{N}$. By Definition 3.8, we know that $\forall n \in \mathbb{N}$, $c \geq n$. However, by Definition 3.24, we know that $n+1 \in \mathbb{N}$, so $c \geq n+1$, i.e. $c - 1 \geq n$. However, $c-1$ is an upper bound which is less than c, which contradicts that c is the supremum. Hence, $\mathbb{N}$ has no upper bound. □

Problem 3.26. Prove $\forall \varepsilon > 0 \; \exists n \in \mathbb{N}$ such that $\dfrac{1}{n} < \varepsilon$.

Proof. Assume for the sake of contradiction that

$$\exists \varepsilon > 0 \; \forall n \in \mathbb{N}, \; \frac{1}{n} \geq \varepsilon.$$

However, $\dfrac{1}{n} \geq \varepsilon$ rearranges to $n \leq \dfrac{1}{\varepsilon}$. However, this implies that $\dfrac{1}{\varepsilon}$ is an upper bound of $\mathbb{N}$, which contradicts Problem 3.25. □

Definition 3.27. Define the set of integers as $\mathbb{Z} = \mathbb{N} \cup \{0\} \cup \{-x \mid x \in \mathbb{N}\}$.

Corollary 3.28

$\mathbb{Z}$ is unbounded above and below.

Example 3.29

Prove that the floor function exists, i.e. $\forall x \in \mathbb{R} \; \exists n \in \mathbb{Z}$, $n \leq x$ such that if $\forall m \in \mathbb{Z}$, $m \leq x$, then $m \leq n$.

Proof. Given $x \in \mathbb{R}$, consider the set $S = \{n \in \mathbb{Z} \mid n \leq x\}$. Note that S is bounded above by x. Assume for the sake of contradiction that S is empty. Then there is no $n \in \mathbb{Z}$ such that $n \leq x$, i.e. all integers are greater than x. This implies that $\mathbb{Z}$ is bounded below by x, which contradicts Corollary 3.28. Therefore, S is nonempty, and it follows by Axiom 3.11 that S has a supremum. Let $c = \sup S$.

By Theorem 3.13, $\forall \varepsilon > 0 \; \exists n \in S$ such that $c - n < \varepsilon$. Consider $\varepsilon = 1$. Then $\exists n \in S$ such that $c - n < 1$, i.e. $c < n + 1$. Furthermore, since c is the supremum, we have $n \leq c$. Therefore, we have $n \leq c < n + 1$.

As x is an upper bound, we have $c \leq x$. Furthermore, note that $n + 1 > c$ implies $n + 1 \notin S$, so $n + 1 > x$.

Combining everything we have established, we have the inequality $n \leq c \leq x < n + 1$. Thus, we have found two consecutive integers n and $n+1$ such that $n \leq x < n + 1$, for a given $x \in \mathbb{R}$.

Now consider any integer m. By properties of integers, we know that $m \leq n$ or $m \geq n + 1$. If $m \leq x$, then $m < n + 1$, so it must be true that $m \leq n$. In other words, we have found an n which satisfies the condition $\forall m \in \mathbb{Z}, m \leq x \to m \leq n$, and we conclude the proof. □

Definition 3.30. Let $n = \lfloor x \rfloor$ denote the greatest integer less than or equal to x, as defined in Example 3.29.

Remark. The quantifier $\exists! x$ denotes "there exists a unique x."

Problem 3.31. Prove that if $x \notin \mathbb{Z}$, $\exists! n \in \mathbb{Z}$ such that $n < x < n + 1$.

Proof. Take $n = \lfloor x \rfloor$. The existence has been established in Example 3.29, and it suffices to show that the solution is unique.

Assume that $m \in \mathbb{Z}$ also satisfies $m < x < m+1$. If $m > n$, then $m \geq n+1 > x$, which contradicts the given assumption. Then we have $m \leq n$. However, if $m < n$, then $m+1 \leq n < x$, which also leads to a contradiction. Hence, $m = n$, so the solution is unique. □

Problem 3.32. Prove that if $x \notin \mathbb{Z}$, $\exists! n \in \mathbb{Z}$ such that $x < n < x+1$.

Proof. Since $x+1 \notin \mathbb{Z}$, by Problem 3.31, there exists n such that $n < x+1 < n+1$. This gives $n < x+1$ and if we rearrange the inequality to get $n-1 < x < n$, it follows that $x < n$. Therefore, $x < n < x+1$, as desired. □

Corollary 3.33

Any interval of length at least 1 contains an integer.

Problem 3.34. Prove that between any two real numbers, there exist rational and irrational numbers.

Proof. Consider $a, b \in \mathbb{R}$, WLOG $a < b$. Then, $b - a > 0$. Then by Problem 3.26, $\exists n \in \mathbb{N}$ such that $\frac{1}{n} < b - a$, or $n > \frac{1}{b-a}$. This rearranges to $bn - an > 1$, so we have an interval (an, bn) with length greater than 1. By Corollary 3.33, there exists $m \in \mathbb{Z}$ such that $an < m < bn$. We then divide by n to get $a < \frac{m}{n} < b$, and therefore we have a rational number $\frac{m}{n}$ between a and b.

Now we have demonstrated that there is a rational number between any two real numbers. Let q be a rational number between $\frac{a}{\sqrt{2}}$ and $\frac{b}{\sqrt{2}}$, which are both real numbers.

We have shown that $\sqrt{3}$ is irrational in the proof of Example 3.4, and similarly we can show that $\sqrt{2}$ is irrational as well.

Then $\frac{a}{\sqrt{2}} < q < \frac{b}{\sqrt{2}} \to a < q\sqrt{2} < b$, and $q\sqrt{2}$ is clearly irrational, so we are done. □

Alternatively, we can find explicit formulas for rational and irrational numbers between a and b:

$$x = \lfloor a \rfloor + \left\lceil \frac{a - \lfloor a \rfloor}{\frac{1}{\left\lceil \frac{1}{b-a} \right\rceil}} \right\rceil \cdot \frac{1}{\left\lceil \frac{1}{b-a} \right\rceil}$$

$$y = \lfloor a \rfloor + \left\lceil \frac{a - \lfloor a \rfloor}{\frac{1}{\sqrt{2}\left\lceil \frac{1}{b-a} \right\rceil}} \right\rceil \cdot \frac{1}{\sqrt{2}\left\lceil \frac{1}{b-a} \right\rceil}$$

It can be shown that both x and y are between a and b, and also that x is rational and y is irrational.

Problem 3.35. Define $-A = \{-a \mid a \in A\}$. Given that they exist, prove that $-\sup A = \inf(-A)$.

Proof. We know that $\forall a \in A,\ a \leq \sup A$, i.e. $-a \geq -\sup A$, so $-\sup A$ is a lower bound for $-A$. Then, $\inf(-A) \geq -\sup A$.

Now we want to demonstrate $\inf(-A) \leq -\sup A$. Assume by contradiction that $\inf(-A) > -\sup A$. Then let $\varepsilon = \inf(-A) + \sup A$. By Theorem 3.13, $\exists a \in A$ such that $\sup A - a < \inf(-A) + \sup A$, or $-a < \inf(-A)$, which contradicts the property of the infimum. Therefore, $\inf(-A) \leq -\sup A$, and we conclude that $\inf(-A) = -\sup(A)$. □

Alternative Proof. In fact, an argument using Theorem 3.13 and contradiction is unnecessary. First, we have $\forall a \in A,\ a \leq \sup A$. Then $\forall -a \in -A,\ -a \geq -\sup A$, so $-\sup A$ is a lower bound of $-A$.

Since $\inf(-A)$ is the greatest lower bound of $-A$, we have $\forall -a \in -A,\ -a \geq \inf(-A) \geq -\sup A$. Then multiply the compound inequality by -1 to get $\forall a \in A,\ a \leq -\inf(-A) \leq \sup A$. However, $\sup A$ is the least upper bound, so we must have $-\inf(-A) = \sup A$, i.e. $-\sup A = \inf(-A)$. □

Problem 3.36. Define $2A = \{2a \mid a \in A\}$. Given that they exist, prove that $\inf(2A) = 2\inf A$.

Proof. First, do not fall into the trap of concluding that $A+A = 2A$. Consider $A = \{1, 3\}$. While $A + A = \{2, 4, 6\}$, $2A = \{2, 6\}$, so the two sets are different.

We know $\forall a \in A,\ a \geq \inf A$, i.e. $2a \geq 2\inf A$, so $2\inf A$ is a lower bound of $2A$. Then, $\inf(2A) \geq 2\inf A$.

Like before, it is sufficient to show $\inf(2A) \leq 2\inf A$. Assume for the sake of contradiction that $\inf(2A) > 2\inf A$. Let $\varepsilon = \inf(2A) - 2\inf A$. Then, by Theorem 3.14, $\exists a \in A$ such that $a - \inf A < \dfrac{\varepsilon}{2}$, i.e. $2a - 2\inf A < \varepsilon$, or $2a - 2\inf A < \inf(2A) - 2\inf A$, which simplifies to $2a < \inf(2A)$, a contradiction. Therefore, $\inf(2A) \leq 2\inf A$, so we conclude that $\inf(2A) = 2\inf A$. □

Alternative Proof. In fact, we do not even need to use the ε definition of supremum or infimum. First, note that $\forall a \in A,\ a \geq \inf A$. Then $2a \geq 2\inf A$, so $2\inf A$ is a lower bound of $2A$.

Then $\forall 2a \in 2A, 2a \geq \inf(2A) \geq 2\inf A$, since $\inf(2A)$ is the greatest lower bound of $2A$. Then we divide this compound inequality by 2 to get that $\forall a \in A,\ a \geq \dfrac{\inf(2A)}{2} \geq \inf A$. However, $\inf A$ is the greatest lower bound of A, so we must have $\dfrac{\inf(2A)}{2} = \inf A$, i.e. $\inf(2A) = 2\inf A$, as desired. □

Problem 3.37. Suppose $A \subseteq B$, with A and B non-empty. If B is bounded above and below, what relation exists between their supremums and infimums?

Solution. By Axiom 3.11, B has a supremum and an infimum. It follows that A also has a supremum and an infimum.

We should demonstrate that $\sup A \leq \sup B$ and $\inf A \geq \inf B$. First, we prove the former.

Assume for the sake of contradiction that $\sup A > \sup B$. Let $\varepsilon = \sup A - \sup B$. By Theorem 3.13, $\exists x \in A$ such that $\sup A - x < \sup A - \sup B$, i.e. $x > \sup B$. However, as $A \subseteq B$, if $x \in A$, then $x \in B$. So $\exists x \in B$ such that $x > \sup B$, so we have a contradiction. Therefore, $\sup A \leq \sup B$.

Now we prove the latter. Similarly, assume for the sake of contradiction that $\inf A < \inf B$. Let $\varepsilon = \inf B - \inf A$. By Theorem 3.14, $\exists x \in A$ such that $x - \inf A < \inf B - \inf A$, or $x < \inf B$. As stated before, $x \in A \to x \in B$, so $\exists x \in B,\ x < \inf B$, which is a contradiction. So, $\inf A \geq \inf B$, as desired. □

Problem 3.38. Given a non-empty set A, consider the set $B = \{3 - 2a \mid a \in A\}$. Assuming that they exist, prove $\inf B = 3 - 2\sup A$.

Proof. Let $M = \sup A$. Then $\forall a \in A,\ a \leq M$, i.e. $-2a \geq -2M$ or $3 - 2a \geq 3 - 2M$. Then $\forall b \in B, b \geq 3 - 2M$, so $3 - 2M$ is a lower bound of B.

Let $\widetilde{M} = \inf B$. Then $\forall b \in B, b \geq \widetilde{M} \geq 3-2M$, as $\widetilde{M}$ is the greatest lower bound. Then $b \geq \widetilde{M} \geq 3-2M \to -3+b \geq -3+\widetilde{M} \geq -2M \to \dfrac{-3+b}{-2} \leq \dfrac{-3+\widetilde{M}}{-2} \leq M$. However, we are given that $\forall b \in B,\ \forall a \in A,\ b = 3 - 2a$. Then $\dfrac{-3+b}{-2} = \dfrac{-3+(3-2a)}{-2} = \dfrac{-2a}{-2} = a$, so we have that $\forall a \in A, a \leq \dfrac{-3+\widetilde{M}}{-2} \leq M$. As $M = \sup A$ is the least upper bound, we must have $M = \dfrac{-3+\widetilde{M}}{-2}$. This rearranges to $-2M = -3 + \widetilde{M}$, or $\widetilde{M} = 3 - 2M$, i.e. $\inf B = 3 - 2\sup A$, as desired. $\square$

Example 3.39

Prove Theorem 3.3.

Proof. Given that f is continuous on $[a, b]$, we will show that f has a maximum on $[a, b]$, and the proof regarding its minimum will be analogous.

Consider the set $S = \{f(x) \mid x \in [a, b]\}$. Clearly it is non-empty, and by Theorem 3.2, S must be bounded. Therefore, by Axiom 3.11, it has a supremum. Let $M = \sup S$.

We know that $\forall x \in [a, b],\ f(x) \leq M$. If $\exists x \in [a, b]$ such that $f(x) = M$, then M is the maximum on this interval, so we are done.

Otherwise, assume for the sake of contradiction that $\forall x \in [a, b],\ f(x) < M$. Let $g(x) = \dfrac{1}{M - f(x)}$. We are given that $f(x)$ is continuous on $[a, b]$. By applying various limit rules, it follows that $g(x)$ is also continuous on $[a, b]$, so by Theorem 3.2, $\exists z \in \mathbb{N}$ such that $\forall x \in [a, b],\ \dfrac{1}{M - f(x)} < z$.

Now we apply Theorem 3.13 on S. We have that $\forall \varepsilon > 0,\ \exists x \in [a, b]$ such that $M - f(x) < \varepsilon$. Let $\varepsilon = \dfrac{1}{n}$ for any $n \in \mathbb{N}$. Then $\forall n \in \mathbb{N},\ \exists x \in [a, b]$ such that $M - f(x) < \dfrac{1}{n}$. Note that this rearranges to $(M - f(x))n < 1$, i.e. $\dfrac{1}{M - f(x)} > n$, which contradicts that $\exists z \in \mathbb{N}$, s.t. $\forall x \in [a, b],\ \dfrac{1}{M - f(x)} < z$. Therefore, $f(x) \not< M\ \forall x \in [a, b]$.

Hence, $\exists x \in [a, b]$ such that $f(x) = M$, i.e. $f(x)$ attains the maximum value M on the interval. $\square$

Problem 3.40. Prove or disprove that there exists a function continuous on $[0, 1]$ for which $f\left(\dfrac{1}{n}\right) = n\ \forall n \in \mathbb{N}$.

Solution. Indeed, there does not exist a function. By Theorem 3.2, f is bounded on $[0, 1]$. When we consider $\dfrac{1}{n} \in [0, 1]$, note that $0 < \dfrac{1}{n} \leq 1 \to n \geq 1\ \forall n \in \mathbb{N}$, which is the output of f. Thus, by Problem 3.25, given any upper bound, we can find a natural number that is greater than that upper bound. This contradicts that f is bounded on $[0, 1]$, so there does not exist some function f that satisfies this condition. $\square$

§4 Derivatives

§4.1 Review

Consider a function f. When, at any part, the function is increasing, we have

$$b > a \longrightarrow f(b) > f(a),$$

for two values $x = a, b$ in that part of the function. We rewrite this as

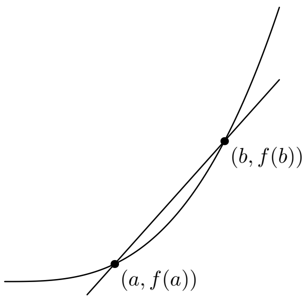
$$b - a > 0 \longrightarrow f(b) - f(a) > 0.$$

When we consider the two points $(a, f(a))$ and $(b, f(b))$, the slope of the line containing both points is also positive:

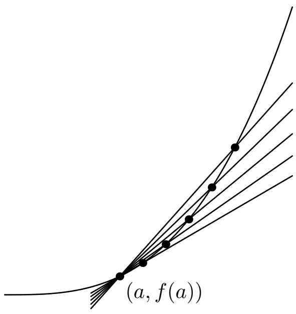
$$b - a > 0 \longrightarrow \frac{f(b) - f(a)}{b - a} > 0.$$

Likewise, we can use analogous reasoning for decreasing parts of a graph as well.

Definition 4.1. We call this line connecting $(a, f(a))$ and $(b, f(b))$ the **secant line**, as shown below:



Now what happens to the slope of the secant line connecting $(a, f(a))$ and $(b, f(b))$ as b is approaching a? Observe the secant line in the figure below:



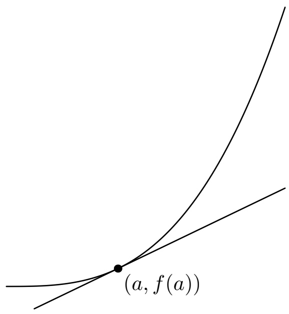
Since we are considering b approaching a, this leads us to use limits to introduce a new, important property of functions:

Definition 4.2. For some constant a, $\lim_{x\to a} \frac{f(x) - f(a)}{x - a} = f'(a)$ is called the **derivative** of $f(x)$ at $x = a$. By Lemma 2.34, note that

$$f'(a) = \lim_{x\to a} \frac{f(x) - f(a)}{x - a} = \lim_{h\to 0} \frac{f(a + h) - f(a)}{h},$$

which is another form of the definition of the derivative.

Definition 4.3. The **tangent line** of $f(x)$ at $x = a$ is the line of slope $f'(a)$ through $(a, f(a))$.



Definition 4.4. If $f'(a)$ exists, f is said to be **differentiable** at $x = a$.

Example 4.5

Let $f(x) = mx + b$ (i.e. an arbitrary linear function with slope m). For any a, prove that $f'(a) = m$.

Proof. We have

$$\begin{aligned} f'(a) &= \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \\ &= \lim_{x \to a} \frac{(mx + b) - (ma + b)}{x - a} \\ &= \lim_{x \to a} \frac{m(x - a)}{x - a} \\ &= \lim_{x \to a} m \\ &= m. \end{aligned}$$

Therefore, for $f(x) = mx + b$, $f'(a) = m$. $\square$

Problem 4.6. Let $f(x) = x^2$. Find $f'(2)$.

Solution. As usual, we apply the definition:

$$\begin{aligned} f'(2) &= \lim_{x \to 2} \frac{f(x) - f(2)}{x - 2} \\ &= \lim_{x \to 2} \frac{x^2 - 4}{x - 2} \\ &= \lim_{x \to 2} \frac{(x + 2)(x - 2)}{x - 2} \\ &= \lim_{x \to 2} x + 2 \\ &= \boxed{4}. \end{aligned}$$ $\square$

Exercise 4.7. Let $f(x) = x^3 - 3x$. Find $f'(2)$.

Exercise 4.8. Let $f(x) = \frac{1}{x}$. Find $f'(-3)$.

Exercise 4.9. Let $f(x) = \dfrac{2x-5}{1-3x}$. Find $f'(4)$.

Exercise 4.10. Let $f(x) = \sqrt{x}$. Find $f'(4)$.

Problem 4.11. What is the equation of the tangent line of $f(x)$ at $x = a$?

Solution. The slope of the tangent line is $f'(a)$. This line must pass through the point $(a, f(a))$, so by the point-slope form of a line, the equation of the tangent line would be $y - f(a) = f'(a)(x - a)$, which rearranges to

$$y = f'(a)(x - a) + f(a).$$

□

Problem 4.12. Find all tangent lines of $y = x^2$ that go through $(7, 1)$.

Solution. Define $f(x) = x^2$. Consider some point $(a, f(a))$ that lies on f and has a tangent line that goes through $(7, 1)$. By Problem 4.11, the general equation of the tangent line would be

$$y = f'(a)(x - a) + f(a).$$

Note that $f(a) = a^2$, and $f'(a) = \lim_{x\to a} \dfrac{f(x) - f(a)}{x - a} = \lim_{x\to a} \dfrac{x^2 - a^2}{x - a} = \lim_{x\to a} \dfrac{(x - a)(x + a)}{x - a} = \lim_{x\to a} x + a = 2a$, so our tangent line is

$$y = 2a(x - a) + a^2.$$

Since our tangent line must also contain $(7, 1)$, we plug those values into the equation:

$$1 = 2a(7 - a) + a^2.$$

This rearranges to $a^2 - 14a + 1 = 0$, and the quadratic formula yields $a = 7 \pm 4\sqrt{3}$, and in fact, both are solutions to the problem. Thus, our tangent lines are

$$y = 2(7 + 4\sqrt{3})(x - (7 + 4\sqrt{3})) + 97 + 56\sqrt{3},$$
$$y = 2(7 - 4\sqrt{3})(x - (7 - 4\sqrt{3})) + 97 - 56\sqrt{3}.$$

□

Example 4.13

Let $f(x) = |x|$. Find $f'(7)$ and $f'(0)$.

Solution. Note that $f'(7) = \lim_{x\to 7} \dfrac{|x| - |7|}{x - 7}$. As we are considering the limit as x goes to 7, we can assume that x is positive because we essentially only care about the values of x that are close to 7. Then,

$$\lim_{x\to 7} \frac{|x| - |7|}{x - 7} = \lim_{x\to 7} \frac{x - 7}{x - 7} = \boxed{1}.$$

For $f'(0)$, we want to find $\lim_{x\to 0} \dfrac{|x|}{x}$. However, evaluating the right-hand and left-hand limits separately gives us 1 and -1 respectively, therefore $\lim_{x\to 0} \dfrac{|x|}{x}$ does not exist, and it follows that there is no tangent line at $x = 0$. In fact, if we graph $f(x) = |x|$, we see that there is a cusp (a pointed end) at $x = 0$, so we can intuitively figure out that there cannot be a tangent line at that point.

□

Definition 4.14. For a function $f(x)$, the **derivative** of $f(x)$ is denoted as $f'(x)$ with respect to all x defined on f. As stated earlier, the limit can be written in two forms:

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

When evaluating derivatives using the limit definition, note that the former involves factoring, while the latter involves expanding, so choose the appropriate definition depending on the function being dealt with.

Definition 4.15. There are two different ways to denote the derivative of f: either $f'(x)$, which is Newton's notation, or $\dfrac{d}{dx}(f(x))$, which is Leibniz's notation. **Leibniz notation** is useful since we don't have to define the function to take its derivative. For example, we can write $\frac{d}{dx}(x^3)$, but not $\left(x^3\right)'$.

If $y = f(x)$, the following all mean the same thing:

$$\frac{d}{dx}(f(x)) = f'(x).$$

$$\frac{d}{dx}(y) = y' = \frac{dy}{dx}.$$

Definition 4.16. To **differentiate** a function is to evaluate the derivative of that function.

Example 4.17

Differentiate the following functions:

1. x^2
2. x^3
3. x^5
4. x^n, $\forall n \in \mathbb{Z}^+$
5. $\dfrac{3x+11}{2x-9}$
6. $\sqrt{x}$
7. $\sin x$

Solution. We may either take the limit of the given function as $z \to x$ or $h \to 0$.

1. Either method works:

$$\lim_{z \to x} \frac{f(z) - f(x)}{z - x} = \lim_{z \to x} \frac{z^2 - x^2}{z - x} = \lim_{z \to x} \frac{(z-x)(z+x)}{z - x} = \lim_{z \to x} z + x = \boxed{2x}.$$

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \to 0} \frac{x^2 + 2xh + h^2 - x^2}{h} = \lim_{h \to 0} 2x + h = \boxed{2x}.$$

2. Factoring seems like a relatively quicker solution:

$$\lim_{z \to x} \frac{f(z) - f(x)}{z - x} = \lim_{z \to x} \frac{z^3 - x^3}{z - x} = \lim_{z \to x} \frac{(z-x)(z^2 + zx + x^2)}{z - x} = \lim_{z \to x} z^2 + zx + x^2 = \boxed{3x^2}.$$

3. Using the helpful factorization technique $a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \ldots + ab^{n-2} + b^{n-1})$, we can quickly deduce

$$\lim_{z\to x} \frac{f(z)-f(x)}{z-x} = \lim_{z\to x} \frac{z^5 - x^5}{z-x} = \lim_{z\to x} z^4 + z^3x + z^2x^2 + zx^3 + x^4 = \boxed{5x^4}.$$

Of course, we also have the option of expanding:

$$\begin{aligned}\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} &= \lim_{h\to 0} \frac{(x+h)^5 - x^5}{h} \\ &= \lim_{h\to 0} \frac{x^5 + 5x^4h + 10x^3h^2 + 10x^2h^3 + 5xh^4 + h^5 - x^5}{h} \\ &= \lim_{h\to 0} 5x^4 + 10x^3h + 10x^2h^2 + 5xh^3 + h^4 = \boxed{5x^4}.\end{aligned}$$

4. From the previous examples, it seems that we have a pattern here. In fact, we can show that $\frac{d}{dx}(x^n) = nx^{x-1}$ using the Binomial Theorem.

$$\begin{aligned}\frac{d}{dx}(x^n) &= \lim_{h\to 0} \frac{(x+h)^n - x^n}{h} \\ &= \lim_{h\to 0} \frac{\left(\binom{n}{0}x^n + \binom{n}{1}x^{n-1}h + \ldots + \binom{n}{n-1}xh^{n-1} + \binom{n}{n}h^n\right) - x^n}{h} \\ &= \lim_{h\to 0} \frac{\binom{n}{1}x^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \ldots + \binom{n}{n-2}x^2h^{n-2} + \binom{n}{n-1}xh^{n-1} + h^n}{h} \\ &= \lim_{h\to 0} \binom{n}{1}x^{n-1} + \binom{n}{2}x^{n-2}h + \ldots + \binom{n}{n-2}x^2h^{n-3} + \binom{n}{n-1}xh^{n-2} + h^{n-1} \\ &= \boxed{nx^{n-1}}.\end{aligned}$$

We can also use the fact $a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \ldots + ab^{n-2} + b^{n-1})$ as mentioned before:

$$\begin{aligned}\frac{d}{dx}(x^n) &= \lim_{z\to x} \frac{z^n - x^n}{z-x} \\ &= \lim_{z\to x} \frac{(z-x)(z^{n-1} + z^{n-2}x + \ldots + zx^{n-2} + x^{n-1})}{z-x} \\ &= \lim_{z\to x} z^{n-1} + z^{n-2}x + \ldots + zx^{n-2} + x^{n-1} \\ &= x^{n-1} + x^{n-2}\cdot x + \ldots + x \cdot x^{n-2} + x^{n-1} \\ &= \boxed{nx^{n-1}}.\end{aligned}$$

5. It is up to your choice which method you prefer. The following uses expanding.

$$\begin{aligned}\frac{d}{dx}\left(\frac{3x+11}{2x-9}\right) &= \lim_{h\to 0} \frac{\frac{3(x+h)+11}{2(x+h)-9} - \frac{3x+11}{2x-9}}{h} \\ &= \lim_{h\to 0} \frac{\frac{3x+3h+11}{2x+2h-9} - \frac{3x+11}{2x-9}}{h} \\ &= \lim_{h\to 0} \frac{(3x+3h+11)(2x-9) - (3x+11)(2x+2h-9)}{h(2x+2h-9)(2x-9)}\end{aligned}$$

$$\begin{aligned}
&= \lim_{h\to 0} \frac{6x^2 - 27x + 6xh - 27h + 22x - 99 - 6x^2 - 6xh + 27x - 22x - 22h + 99}{h(2x+2h-9)(2x-9)} \\
&= \lim_{h\to 0} \frac{-49h}{h(2x+2h-9)(2x-9)} \\
&= \lim_{h\to 0} -\frac{49}{(2x+2h-9)(2x-9)} \\
&= \boxed{-\frac{49}{(2x-9)^2}}.
\end{aligned}$$

6. A common strategy in this situation is to multiply by the radical conjugate.

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

7. This is one of the more nontrivial functions to differentiate using the limit definition.

$$\begin{aligned}
\frac{d}{dx}(\sin x) &= \lim_{h\to 0} \frac{\sin(x+h) - \sin x}{h} \\
&= \lim_{h\to 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\
&= \lim_{h\to 0} \frac{\cos x \sin h - \sin x(1 - \cos h)}{h} \\
&= \lim_{h\to 0} \frac{\cos x \sin h - \sin x \cdot \frac{\sin^2 h}{1+\cos h}}{h} \\
&= \lim_{h\to 0} \frac{\sin h\left(\cos x - \frac{\sin x \sin h}{1+\cos h}\right)}{h} \\
&= \lim_{h\to 0} \frac{\sin h}{h} \cdot \lim_{h\to 0}\left(\cos x - \frac{\sin x \sin h}{1+\cos h}\right) \\
&= \lim_{h\to 0}\left(\cos x - \frac{\sin x \sin h}{1+\cos h}\right) \\
&= \lim_{h\to 0} \cos x - \lim_{h\to 0}\left(\frac{\sin x \sin h}{1+\cos h}\right) \\
&= \cos x - \frac{\lim_{h\to 0}(\sin x \sin h)}{\lim_{h\to 0}(1+\cos h)} \\
&= \cos x - \frac{\sin x \cdot \lim_{h\to 0} \sin h}{1 + \lim_{h\to 0} \cos h}
\end{aligned}$$

$$
\begin{aligned}
&= \cos x - \frac{\sin x \cdot 0}{1+1} \\
&= \boxed{\cos x}.
\end{aligned}
$$

□

Theorem 4.18 (Linearity of Derivatives)

Let $h(x) = f(x) + g(x)$ and $j(x) = k \cdot f(x)$ for some $k \in \mathbb{R}$. Assume that $f(x)$ and $g(x)$ are differentiable.

a) $h'(x) = f'(x) + g'(x)$.

b) $j'(x) = k \cdot f'(x)$.

Proof. a) By Theorem 2.23,

$$
\begin{aligned}
h'(x) &= \lim_{z \to x} \frac{h(z) - h(x)}{x - z} \\
&= \lim_{z \to x} \frac{f(z) + g(z) - (f(x) + g(x))}{x - z} \\
&= \lim_{z \to x} \frac{(f(z) - f(x)) + (g(z) - g(x))}{x - z} \\
&= \lim_{z \to x} \left(\frac{f(z) - f(x)}{x - z} + \frac{g(z) - g(x)}{x - z} \right) \\
&= \lim_{z \to x} \frac{f(z) - f(x)}{x - z} + \lim_{z \to x} \frac{g(z) - g(x)}{x - z} \\
&= f'(x) + g'(x).
\end{aligned}
$$

b) We can always factor out constants from limits, which follows from Theorem 2.22 and Theorem 2.23:

$$
\begin{aligned}
j'(x) &= \lim_{z \to x} \frac{j(z) - j(x)}{z - x} \\
&= \lim_{z \to x} \frac{k \cdot f(z) - k \cdot f(x)}{z - x} \\
&= \lim_{z \to x} \frac{k(f(z) - f(x))}{z - x} \\
&= \lim_{z \to x} k \cdot \frac{f(z) - f(x)}{z - x} \\
&= k \cdot \lim_{z \to x} \frac{f(z) - f(x)}{z - x} \\
&= k \cdot f'(x).
\end{aligned}
$$

□

Example 4.19

Let $g(x) = f(2x)$. If $f'(x) = \alpha(x)$, what is $g'(x)$?

Solution. First, note that

$$
g'(x) = \lim_{z \to x} \frac{g(z) - g(x)}{z - x}
$$

$$
\begin{aligned}
&= \lim_{z \to x} \frac{f(2z) - f(2x)}{z - x} \\
&= 2 \cdot \lim_{z \to x} \frac{f(2z) - f(2x)}{2z - 2x}.
\end{aligned}
$$

As z approaches x, $2z$ approaches $2x$, so this is equivalent to

$$
2 \cdot \lim_{2z \to 2x} \frac{f(2z) - f(2x)}{2z - 2x} = \boxed{2\alpha(2x)}.
$$

□

Theorem 4.20

If $f(x)$ is differentiable at $x = a$, then it is continuous at $x = a$.

Proof. We want to show that if $\lim\limits_{x \to a} \dfrac{f(x) - f(a)}{x - a}$ exists, then $\lim\limits_{x \to a} f(x) = f(a)$, or equivalently, $\lim\limits_{x \to a} f(x) - f(a) = 0$.

We know that $\lim\limits_{x \to a} \dfrac{f(x) - f(a)}{x - a}$ and $\lim\limits_{x \to a}(x - a)$ both exist. We have that

$$
\lim_{x \to a} \frac{f(x) - f(a)}{x - a} \cdot \lim_{x \to a}(x - a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \cdot (x - a) = \lim_{x \to a} f(x) - f(a).
$$

But note that $\lim\limits_{x \to a}(x - a) = 0$, and since the RHS will be 0, the LHS, i.e. $\lim\limits_{x \to a} f(x) - f(a)$, will necessarily be 0. □

Problem 4.21. Define the following functions:

$$
f(x) = \begin{cases} x \sin\left(\dfrac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}
$$

$$
g(x) = \begin{cases} x^2 \sin\left(\dfrac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}
$$

Are these functions differentiable at $x = 0$?

Solution. Consider $f(x)$ at $x = 0$. By the limit definition of a derivative, we have

$$
\begin{aligned}
f'(0) &= \lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} \\
&= \lim_{x \to 0} \frac{x \sin(\frac{1}{x}) - 0}{x - 0} \\
&= \lim_{x \to 0} \sin\left(\frac{1}{x}\right) \\
&= \lim_{x \to 0^+} \sin\left(\frac{1}{x}\right).
\end{aligned}
$$

By Lemma 2.44, $\lim\limits_{x \to 0^+} \sin\left(\dfrac{1}{x}\right) = \lim\limits_{y \to \infty} \sin y$, and this limit does not exist, since $\sin y$ oscillates between -1 and 1. Therefore, $f'(0)$ does not exist, so f is not differentiable at $x = 0$.

Now consider $g(x)$ at $x = 0$. By the limit definition of a derivative, we have

$$\begin{aligned} g'(0) &= \lim_{x \to 0} \frac{g(x) - f(0)}{x - 0} \\ &= \lim_{x \to 0} \frac{x^2 \sin(\frac{1}{x}) - 0}{x - 0} \\ &= \lim_{x \to 0} x \sin\left(\frac{1}{x}\right). \end{aligned}$$

We can split this limit into right-hand and left-hand limits, then apply Squeeze Theorem on both, to determine that $\lim_{x \to 0} x \sin\left(\frac{1}{x}\right) = 0$. Thus, $g'(0) = 0$, so g is differentiable at $x = 0$. $\square$

Theorem 4.22 (Product Rule of Derivatives)
Assuming that the derivatives of $f(x)$ and $g(x)$ exist,

$$\frac{d}{dx}(f(x)g(x)) = f(x)g'(x) + f'(x)g(x).$$

Proof. This proof uses a similar strategy to that of Theorem 2.23.

$$\begin{aligned} \frac{d}{dx}(f(x)g(x)) &= \lim_{z \to x} \frac{f(z)g(z) - f(x)g(x)}{z - x} \\ &= \lim_{z \to x} \frac{f(z)g(z) - f(z)g(x) + f(z)g(x) - f(x)g(x)}{z - x} \\ &= \lim_{z \to x} \frac{f(z)\left(g(z) - g(x)\right) + g(x)\left(f(z) - f(x)\right)}{z - x} \\ &= \lim_{z \to x} \left(\frac{f(z)\left(g(z) - g(x)\right)}{z - x} + \frac{g(x)\left(f(z) - f(x)\right)}{z - x}\right) \\ &= \lim_{z \to x} f(z) \cdot \lim_{z \to x} \frac{g(z) - g(x)}{z - x} + \lim_{z \to x} \frac{f(z) - f(x)}{z - x} \cdot \lim_{z \to x} g(x) \\ &= \lim_{z \to x} f(z) \cdot g'(x) + f'(x)g(x). \end{aligned}$$

As we have assumed that $f(x)$ and $g(x)$ are differentiable, by Theorem 4.20, $\lim_{z \to x} f(z) = f(x)$, therefore

$$\frac{d}{dx}(f(x)g(x)) = f(x)g'(x) + f'(x)g(x).$$ $\square$

Theorem 4.23 (Quotient Rule of Derivatives)
Assuming that the derivatives of $f(x)$ and $g(x)$ exist,

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{g(x)f'(x) - f(x)g'(x)}{g(x)^2}.$$

Proof. First, we expand the fractions, then apply a similar strategy used in the previous proof.

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \lim_{z \to x} \frac{\frac{f(z)}{g(z)} - \frac{f(x)}{g(x)}}{z - x}$$

$$\begin{aligned}
&= \lim_{z\to x} \frac{f(z)g(x) - f(x)g(z)}{g(x)g(z)(z-x)} \\
&= \lim_{z\to x} \frac{f(z)g(x) - f(x)g(x) + f(x)g(x) - f(x)g(z)}{g(x)g(z)(z-x)} \\
&= \lim_{z\to x} \frac{g(x)\left(f(z) - f(x)\right) + f(x)\left(g(x) - g(z)\right)}{g(x)g(z)(z-x)} \\
&= \lim_{z\to x} \frac{1}{g(x)} \cdot \lim_{z\to x} \frac{1}{g(z)} \cdot \left[\lim_{z\to x} \frac{f(z)-f(x)}{z-x} \cdot \lim_{z\to x} g(x) + \lim_{z\to x} f(x) \cdot \lim_{z\to x} \frac{g(x)-g(z)}{z-x}\right].
\end{aligned}$$

By our assumptions and Theorem 4.20, we have

$$\begin{aligned}
\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) &= \frac{1}{g(x)} \cdot \frac{1}{g(x)} \cdot \left[f'(x) \cdot g(x) + f(x) \cdot -g'(x)\right] \\
&= \frac{g(x)f'(x) - f(x)g'(x)}{g(x)^2}.
\end{aligned}$$

□

Problem 4.24. Evaluate the derivatives of $f(x) = x^3 \sin x$ and $f(x) = \dfrac{2x-1}{3x+2}$.

Solution. For the first function, we apply Theorem 4.22.

$$\begin{aligned}
\frac{d}{dx}(x^3 \sin x) &= \frac{d}{dx}(x^3) \cdot \sin x + x^3 \cdot \frac{d}{dx}(\sin x) \\
&= 3x^2 \cdot \sin x + x^3 \cdot \cos x \\
&= \boxed{x^2(3\sin x + x\cos x)}.
\end{aligned}$$

For the latter, apply Theorem 4.23.

$$\begin{aligned}
\frac{d}{dx}\left(\frac{2x-1}{3x+2}\right) &= \frac{(3x+2) \cdot \frac{d}{dx}(2x-1) - (2x-1) \cdot \frac{d}{dx}(3x+2)}{(3x+2)^2} \\
&= \frac{2(3x+2) - 3(2x-1)}{(3x+2)^2} \\
&= \boxed{\frac{7}{(3x+2)^2}}.
\end{aligned}$$

□

Problem 4.25. By Example 4.17, we have established that the derivative of $\sin x$ is $\cos x$. Using Theorem 4.22 and Theorem 4.23, find the derivatives of the rest of the trigonometric functions.

Solution. The proof of $\dfrac{d}{dx}(\cos x)$ is similar to that of $\dfrac{d}{dx}(\sin x)$.

$$\begin{aligned}
\frac{d}{dx}(\cos x) &= \lim_{h\to 0} \frac{\cos(x+h) - \cos x}{h} \\
&= \lim_{h\to 0} \frac{\cos x \cos h - \sin x \sin h - \cos x}{h} \\
&= \lim_{h\to 0} \frac{\cos x(\cos h - 1) - \sin x \sin h}{h} \\
&= \lim_{h\to 0} \frac{\cos x \cdot -\frac{\sin^2 h}{\cos h + 1} - \sin x \sin h}{h} \\
&= \lim_{h\to 0} \frac{-\sin h\left(\frac{\cos x \sin h}{\cos h + 1} + \sin x\right)}{h}
\end{aligned}$$

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

Now that we have found the derivatives of sine and cosine, we can apply Theorem 4.23 to find $\frac{d}{dx}(\tan x)$.

$$
\begin{aligned}
\frac{d}{dx}(\tan x) &= \frac{d}{dx}\left(\frac{\sin x}{\cos x}\right) \\
&= \frac{\cos x \cdot \frac{d}{dx}(\sin x) - \sin x \cdot \frac{d}{dx}(\cos x)}{\cos^2 x} \\
&= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} \\
&= \frac{1}{\cos^2 x} \\
&= \boxed{\sec^2 x}.
\end{aligned}
$$

The proof of $\frac{d}{dx}(\cot x)$ is analogous to that of $\frac{d}{dx}(\tan x)$.

$$
\begin{aligned}
\frac{d}{dx}(\cot x) &= \frac{d}{dx}\left(\frac{\cos x}{\sin x}\right) \\
&= \frac{\sin x \cdot \frac{d}{dx}(\cos x) - \cos x \cdot \frac{d}{dx}(\sin x)}{\sin^2 x} \\
&= \frac{-\sin^2 x - \cos^2 x}{\sin^2 x} \\
&= -\frac{1}{\sin^2 x} \\
&= \boxed{-\csc^2 x}.
\end{aligned}
$$

For $\frac{d}{dx}(\sec x)$, use Theorem 4.23 and note that the derivative of 1 is simply 0.

$$
\begin{aligned}
\frac{d}{dx}(\sec x) &= \frac{d}{dx}\left(\frac{1}{\cos x}\right) \\
&= \frac{\cos x \cdot \frac{d}{dx}(1) - 1 \cdot \frac{d}{dx}(\cos x)}{\cos^2 x} \\
&= \frac{\sin x}{\cos^2 x} \\
&= \boxed{\tan x \sec x}.
\end{aligned}
$$

The proof of $\frac{d}{dx}(\csc x)$ is analogous to that of $\frac{d}{dx}(\sec x)$.

$$
\begin{aligned}
\frac{d}{dx}(\csc x) &= \frac{d}{dx}\left(\frac{1}{\sin x}\right) \\
&= \frac{\sin x \cdot \frac{d}{dx}(1) - 1 \cdot \frac{d}{dx}(\sin x)}{\sin^2 x}
\end{aligned}
$$

$$= -\frac{\cos x}{\sin^2 x}$$
$$= \boxed{-\cot x \csc x}. \qquad \square$$

Problem 4.26. Prove the following results about the derivatives of function transformations.

a) $\dfrac{d}{dx}\left(f(x+k)\right) = f'(x+k)$.

b) $\dfrac{d}{dx}\left(f(kx)\right) = kf'(kx)$.

Proof. The general strategy is to substitute the transformations $x+k$ and kx with convenient variables to facilitate the manipulation of the limit definition of the derivative.

a) Let $\widetilde{z} = z + k$ and $\widetilde{x} = x + k$. Note that $\widetilde{z} - \widetilde{x} = z - x$, and that z goes to x as $\widetilde{z}$ goes to $\widetilde{x}$. Then,

$$\begin{aligned}\frac{d}{dx}\left(f(x+k)\right) &= \lim_{z\to x}\frac{f(z+k)-f(x+k)}{z-x}\\ &= \lim_{\widetilde{z}\to\widetilde{x}}\frac{f(\widetilde{z})-f(\widetilde{x})}{\widetilde{z}-\widetilde{x}}\\ &= f'(\widetilde{x})\\ &= f'(x+k).\end{aligned}$$

b) Let $\widetilde{z} = kz$ and $\widetilde{x} = kx$. Note that z goes to x as $\widetilde{z}$ goes to $\widetilde{x}$. Then,

$$\begin{aligned}\frac{d}{dx}\left(f(kx)\right) &= \lim_{z\to x}\frac{f(kz)-f(kx)}{z-x}\\ &= \lim_{z\to x}\frac{f(kz)-f(kx)}{kz-kx}\cdot k\\ &= \lim_{\widetilde{z}\to\widetilde{x}}\frac{f(\widetilde{z})-f(\widetilde{x})}{\widetilde{z}-\widetilde{x}}\cdot k\\ &= f'(\widetilde{x})\cdot k\\ &= kf'(kx).\end{aligned} \qquad \square$$

Remark. Keep in mind that $f'(x+k)$ and $f'(kx)$ are NOT the same as $\dfrac{d}{dx}\left(f(x+k)\right)$ and $\dfrac{d}{dx}\left(f(kx)\right)$.

If *something* is some function of x, then $f'(\textit{something})$ is the derivative of f evaluated at *something*, while $\dfrac{d}{dx}\left(f(\textit{something})\right) = f'(\textit{something})$ evaluated at x.

For instance, if we let $f(x) = \sin x$, then $f'(2x) = \cos 2x$ while $\dfrac{d}{dx}(f(2x)) = 2\cos 2x$ according to the stated theorem.

Problem 4.27. Using the results of Problem 4.26, quickly evaluate $\dfrac{d}{dx}(\cos x)$.

Solution. We can take advantage of the cofunction identity of sine and cosine:

$$\frac{d}{dx}(\cos x) = \frac{d}{dx}\left(\sin\left(x+\frac{\pi}{2}\right)\right) = \sin'\left(x+\frac{\pi}{2}\right) = \cos\left(x+\frac{\pi}{2}\right) = -\sin x. \qquad \square$$

Problem 4.28. Use the product rule twice to find $\frac{d}{dx}(f(x)g(x)h(x))$.

Solution. Initially suppose that $f(x)g(x)$ is a single function for the first application of Theorem 4.22, then apply the theorem again to break up $f(x)g(x)$.

$$\begin{aligned}\frac{d}{dx}(f(x)g(x)h(x)) &= f(x)g(x)\frac{d}{dx}(h(x)) + \frac{d}{dx}(f(x)g(x))h(x) \\ &= f(x)g(x)h'(x) + (f(x)g'(x) + f'(x)g(x))h(x) \\ &= f(x)g(x)h'(x) + f(x)g'(x)h(x) + f'(x)g(x)h(x).\end{aligned}$$ □

It seems that there is a pattern for the derivative of the product of an arbitrary number of functions. Here is the general statement:

Theorem 4.29 (Generalized Product Rule)

Consider arbitrary functions $f_1, f_2, \ldots, f_n$. Then,

$$\frac{d}{dx}\left(\prod_{k=1}^{n} f_k(x)\right) = \left(\prod_{k=1}^{n} f_i(x)\right)\left(\sum_{i=1}^{n} \frac{f_i'(x)}{f_i(x)}\right).$$

This result can be proved by mathematical induction.

Corollary 4.30 (General Power Rule)

$$\frac{d}{dx}(f(x)^n) = nf(x)^{n-1}f'(x).$$

Proof. We consider a special case of Theorem 4.29 where $f_1 = f_2 = \ldots = f_n$.

$$\begin{aligned}\frac{d}{dx}(f(x)^n) &= \left(\prod_{i=1}^{n} f(x)\right)\left(\sum_{i=1}^{n} \frac{f'(x)}{f(x)}\right) \\ &= f(x)^n\left(\frac{nf'(x)}{f(x)}\right) \\ &= nf(x)^{n-1}f'(x).\end{aligned}$$ □

Problem 4.31. Recall the result of Example 4.17 where it was proven that $\frac{d}{dx}(x^n) = nx^{n-1}$ for all $n \in \mathbb{Z}^+$. Demonstrate that this identity holds for all negative integers n as well.

Proof. Let $n \in \mathbb{Z}^+$. By Theorem 4.23, note that

$$\begin{aligned}\frac{d}{dx}(x^{-n}) &= \frac{d}{dx}\left(\frac{1}{x^n}\right) \\ &= \frac{x^n \cdot \frac{d}{dx}(1) - 1 \cdot \frac{d}{dx}(x^n)}{(x^n)^2} \\ &= \frac{-nx^{n-1}}{x^{2n}} \\ &= -nx^{-n-1}.\end{aligned}$$

Let $m = -n$, such that m is a negative integer. Then we have $\frac{d}{dx}(x^m) = mx^{m-1}$, and we are done. $\square$

Clearly, for $n = 0$, $\frac{d}{dx}(x^0) = \frac{d}{dx}(1) = 0 = 0 \cdot x^{0-1}$. We have now established that $\frac{d}{dx}(x^n) = nx^{n-1}$ for all integers n. However, we can extend this rule to the rational numbers:

Problem 4.32. Let $n \in \mathbb{Z}^+$. Evaluate $\frac{d}{dx}\left(x^{\frac{1}{n}}\right)$.

Proof. We take advantage of the following factorization:

$$\begin{aligned} z - x &= \left(z^{\frac{1}{n}}\right)^n - \left(x^{\frac{1}{n}}\right)^n \\ &= \left(z^{\frac{1}{n}} - x^{\frac{1}{n}}\right)\left(z^{\frac{n-1}{n}} + z^{\frac{n-2}{n}}x^{\frac{1}{n}} + \ldots + z^{\frac{1}{n}}x^{\frac{n-2}{n}} + x^{\frac{n-1}{n}}\right). \end{aligned}$$

Thus, we have

$$\begin{aligned} \frac{d}{dx}\left(x^{\frac{1}{n}}\right) &= \lim_{z \to x} \frac{z^{\frac{1}{n}} - x^{\frac{1}{n}}}{z - x} \\ &= \lim_{z \to x} \frac{z^{\frac{1}{n}} - x^{\frac{1}{n}}}{\left(z^{\frac{1}{n}} - x^{\frac{1}{n}}\right)\left(z^{\frac{n-1}{n}} + z^{\frac{n-2}{n}}x^{\frac{1}{n}} + \ldots + z^{\frac{1}{n}}x^{\frac{n-2}{n}} + x^{\frac{n-1}{n}}\right)} \\ &= \lim_{z \to x} \frac{1}{z^{\frac{n-1}{n}} + z^{\frac{n-2}{n}}x^{\frac{1}{n}} + \ldots + z^{\frac{1}{n}}x^{\frac{n-2}{n}} + x^{\frac{n-1}{n}}} \\ &= \frac{1}{nx^{\frac{n-1}{n}}} \\ &= \frac{1}{n}x^{\frac{1}{n}-1}. \end{aligned}$$

$\square$

Problem 4.33. Use Theorem 4.23 to find $\frac{d}{dx}\left(x^{-\frac{1}{n}}\right)$ for $n \in \mathbb{Z}^+$.

Proof. We take the same approach as we did for Problem 4.31.

$$\begin{aligned} \frac{d}{dx}\left(x^{-\frac{1}{n}}\right) &= \frac{d}{dx}\left(\frac{1}{x^{\frac{1}{n}}}\right) \\ &= \frac{x^{\frac{1}{n}} \cdot \frac{d}{dx}(1) - 1 \cdot \frac{d}{dx}\left(x^{\frac{1}{n}}\right)}{\left(x^{\frac{1}{n}}\right)^2} \\ &= \frac{-\frac{1}{n}x^{\frac{1}{n}-1}}{x^{\frac{2}{n}}} \\ &= -\frac{1}{n}x^{-\frac{1}{n}-1}. \end{aligned}$$

$\square$

We now have shown that $\frac{d}{dx}\left(x^{\frac{1}{n}}\right) = \frac{1}{n}x^{\frac{1}{n}-1}$ for all nonzero integers n. Finally, we deal with the general $\frac{p}{q}$ case for rational numbers, using results that we have discovered so far:

Theorem 4.34 (Power Rule)

$$\forall n \in \mathbb{Q}, \ \frac{d}{dx}\left(x^n\right) = nx^{n-1}.$$

Proof. Using the previous result and Corollary 4.30, we will evaluate $\dfrac{d}{dx}\left(x^{\frac{p}{q}}\right)$ for $p, q \in \mathbb{Z}$ and $q \neq 0$.

$$\begin{aligned}
\frac{d}{dx}\left(x^{\frac{p}{q}}\right) &= \frac{d}{dx}\left(\left(x^{\frac{1}{q}}\right)^p\right) \\
&= p\left(x^{\frac{1}{q}}\right)^{p-1}\left(\frac{1}{q}x^{\frac{1}{q}-1}\right) \\
&= \frac{p}{q}\left(x^{\frac{1}{q}}\right)^{p-1}\left(x^{\frac{1}{q}-1}\right) \\
&= \frac{p}{q}x^{\frac{p}{q}-1}.
\end{aligned}$$

□

Problem 4.35. Let $P(x)$ be the polynomial $\sum_{k=0}^{n} a_k x^k$. Find $P'(x)$.

Proof. As a direct result of Theorem 4.34, we have

$$\begin{aligned}
P'(x) &= \sum_{k=0}^{n} \frac{d}{dx}(a_k x^k) \\
&= \sum_{k=0}^{n} k a_k x^{k-1}.
\end{aligned}$$

□

Now, notice the similarity among the following:

- $\dfrac{d}{dx}\left(f(x+c)\right) = f'(x+c)$.

  When we evaluate $\dfrac{d}{dx}\left(f(x+c)\right)$, we are actually evaluating $\dfrac{d}{dx}\left(f \circ g(x)\right)$, where $g(x) = x + c$.

- $\dfrac{d}{dx}\left(f(kx)\right) = kf'(kx)$.

  When we evaluate $\dfrac{d}{dx}(f(kx))$, we are actually evaluating $\dfrac{d}{dx}\left(f \circ g(x)\right)$, where $g(x) = kx$.

- $\dfrac{d}{dx}\left(f(x)^n\right) = nf(x)^{n-1}f'(x)$.

  When we evaluate $\dfrac{d}{dx}(f(x)^n)$, we are actually evaluating $\dfrac{d}{dx}\left(g \circ f(x)\right)$, where $g(x) = x^n$.

Considering the *composition* of functions will lead us into our next important property of derivatives:

Theorem 4.36 (Chain Rule)

Under appropriate conditions (i.e. functions are well-behaved, continuous; there are no domain or range issues, etc.),

$$\frac{d}{dx}(g(f(x))) = g'(f(x))f'(x).$$

For convenience of explanations, we will refer to f in this statement as the 'inner function' and g as the 'outside function.'

If we let $y = f(x)$ and $z = g(y)$, then it follows from Theorem 4.36 that

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}.$$

The proof of this theorem lies beyond the scope of this course.

Problem 4.37. Use Theorem 4.36 to find $\frac{d}{dx}(f(x^n))$.

Solution. Note that the inner function is x^n, and the outside function is f. Using Theorem 4.36 gives us

$$\frac{d}{dx}(f(x^n)) = f'(x^n) \cdot \frac{d}{dx}(x^n) = f'(x^n) \cdot nx^{n-1}. \qquad \square$$

Exercise 4.38. Use Theorem 4.36 to demonstrate Corollary 4.30.

Problem 4.39. Let f be a function such that $f'(x) = \frac{1}{\sqrt{1-x^2}}$ for $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Evaluate $\frac{d}{dx}(f(\sin x))$.

Solution. By Theorem 4.36, we have $\frac{d}{dx}(f(\sin x)) = f'(\sin x)\frac{d}{dx}(\sin x) = f'(\sin x)\cos x$. To evaluate $f'(\sin x)$, we simply plug in $\sin x$ into the input of f', i.e. $f'(\sin x) = \frac{1}{\sqrt{1-\sin^2 x}}$. Since $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, we have that $\sqrt{1-\sin^2 x} = \cos x$, such that $f'(\sin x) = \frac{1}{\cos x}$. Thus, $\frac{d}{dx}(f(\sin x)) = \frac{1}{\cos x} \cdot \cos x = \boxed{1}$. $\square$

Problem 4.40. Differentiate the following:

1. $\sin^3(x^2)$
2. $\sqrt{1+4\sin x}$
3. $\frac{x^2}{1+x^2}$
4. $\tan(x^3+x^2)$
5. $\sqrt[3]{\cos^4\left(\sin^5\left(x^6\right)\right)}$

Solution. You will need to be familiar with applying Theorem 4.36:

1. After applying Theorem 4.36 twice, we get $\frac{d}{dx}(\sin^3(x^2)) = 3\sin^2(x^2) \cdot \frac{d}{dx}(\sin(x^2)) = 3\sin^2(x^2) \cdot \cos(x^2) \cdot 2x = \boxed{6x\sin^2(x^2)\cos(x^2)}$.

2. Similar to the previous one, $\frac{d}{dx}(\sqrt{1+4\sin x}) = \frac{1}{2}(1+4\sin x)^{-\frac{1}{2}} \cdot \frac{d}{dx}(1+4\sin x) = \frac{1}{2}(1+4\sin x)^{-\frac{1}{2}} \cdot 4\cos x = \boxed{\frac{2\cos x}{\sqrt{1+4\sin x}}}$.

3. We use Theorem 4.23:

$$\begin{aligned}\frac{d}{dx}\left(\frac{x^2}{1+x^2}\right) &= \frac{(1+x^2)\cdot\frac{d}{dx}\left(x^2\right) - x^2\cdot\frac{d}{dx}(1+x^2)}{(1+x^2)^2} \\ &= \frac{2x(1+x^2) - 2x\cdot x^2}{(1+x^2)^2} \\ &= \boxed{\frac{2x}{(1+x^2)^2}}.\end{aligned}$$

4. By Theorem 4.36, $\frac{d}{dx}(\tan(x^3+x^2)) = \sec^2(x^3+x^2)\cdot\frac{d}{dx}(x^3+x^2) = \boxed{\sec^2(x^3+x^2)\cdot(3x^2+2x)}$.

5. Repeatedly apply the Theorem 4.36:

$$\begin{aligned}\frac{d}{dx}\left(\sqrt[3]{\cos^4\left(\sin^5\left(x^6\right)\right)}\right) &= \frac{4}{3}\cos^{\frac{1}{3}}\left(\sin^5\left(x^6\right)\right)\cdot\frac{d}{dx}\left(\cos\left(\sin^5\left(x^6\right)\right)\right) \\ &= \frac{4}{3}\cos^{\frac{1}{3}}\left(\sin^5\left(x^6\right)\right)\cdot - \sin\left(\sin^5\left(x^6\right)\right)\cdot\frac{d}{dx}\left(\sin^5\left(x^6\right)\right) \\ &= \frac{4}{3}\cos^{\frac{1}{3}}\left(\sin^5\left(x^6\right)\right)\cdot - \sin\left(\sin^5\left(x^6\right)\right)\cdot 5\sin^4\left(x^6\right)\cdot\frac{d}{dx}\left(\sin\left(x^6\right)\right) \\ &= \frac{4}{3}\cos^{\frac{1}{3}}\left(\sin^5\left(x^6\right)\right)\cdot - \sin\left(\sin^5\left(x^6\right)\right)\cdot 5\sin^4\left(x^6\right)\cdot\cos(x^6)\cdot 6x^5 \\ &= \boxed{-40\sin\left(\sin^5\left(x^6\right)\right)\sin^4\left(x^6\right)\cos^{\frac{1}{3}}\left(\sin^5\left(x^6\right)\right)\cos\left(x^6\right)x^5}.\end{aligned}$$

Generally, this is as far as problems involving the application of Theorem 4.36 go. □

Definition 4.41. Let $n \in \mathbb{Z}^+$. The **nth derivative** of $f(x)$ is what you get when you take the derivative of the function n times. We write this as $f^{(n)}(x)$ or repeat the $'$ symbol n times, as shown:

$$\begin{aligned}y &= f(x) \\ y' &= f'(x) \\ y'' &= f''(x) \\ y''' &= f'''(x)\end{aligned}$$

and so on.

We could also use roman numerals, i.e. we would represent the fourth derivative of $f(x)$ as $f^{\mathrm{IV}}(x)$.

Using Leibniz Notation, we can represent the second derivative of y as

$$y'' = \frac{d}{dx}\left(y'\right) = \frac{d}{dx}\frac{dy}{dx} = \frac{d^2y}{dx^2}.$$

Problem 4.42. Let $y = x^3 + x^2 + x + 1$. Find y', y'', y''', y^{IV}, and $y^{(100)}$.

Solution.

$$
\begin{aligned}
y' &= 3x^2 + 2x + 1 \\
y'' &= 6x + 2 \\
y''' &= 6 \\
y^{\mathrm{IV}} &= 0 \\
y^{(100)} &= 0
\end{aligned}
$$

□

Exercise 4.43. Let $y = x^n$. Find $y^{(n)}$.

Exercise 4.44. Let $y = \sin(3x)$. Find $y^{(99)}$.

Exercise 4.45. Find $\dfrac{d^{2017}y}{dx^{2017}}(\cos(6x))$.

Exercise 4.46. Find $\dfrac{d^2y}{dx^2}(f(x)g(x))$. What does the result resemble?

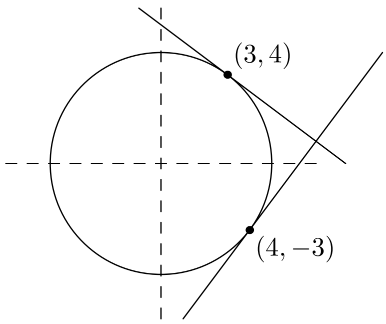
§4.2 Implicit Differentiation

We introduce the concept with an illuminating example.

Example 4.47

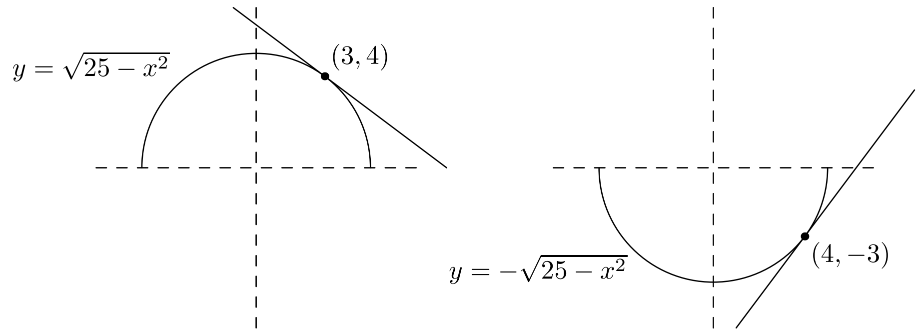
Find the slope of the tangent line at the point $(3, 4)$ of the equation $x^2 + y^2 = 25$. What about $(4, -3)$?

Solution. This is a circle of radius 5 centered at the origin, so consider the following diagram:



The graph of this clearly fails the vertical line test, implying that y cannot be a function of x (there exist multiple y values given one x value). Therefore, we cannot apply our usual tactics.

One approach is to consider cases separately. First, we could solve for y and get two separate equations: $y = \sqrt{25 - x^2}$ and $y = -\sqrt{25 - x^2}$, then evaluate the derivatives separately for $(3, 4)$, which lies in the top half, and $(4, -3)$ which lies in the bottom half, as shown:



For the former, we get $y' = -\frac{x}{\sqrt{25 - x^2}}$, and then we plug in 3 from the point $(3, 4)$ to get the slope $-\frac{3}{4}$. For the latter, we get $y' = \frac{x}{\sqrt{25 - x^2}}$, and then we plug in 4 to get the slope $\frac{4}{3}$.

However, this method was inefficient and time-consuming. It is not even guaranteed that we are able to find an explicit function for y, as the example $x^2 + y^2 = 25$ was conveniently simple. In fact, we can do better.

Note that the relation $x^2 + y^2 = 25$ implies that x and y are *tied* to each other (i.e. share a relation with one another) in some way, but not in a way such that one is an *explicit* function of another. In this situation, we acknowledge that y is an **implicit function** of x, and actually be explicitly written as $y = f(x)$.

We can still find the derivative of an implicit function. As x and y have some sort of relation to each other, we let $y = f(x)$, and the following relation holds:

$$x^2 + f(x)^2 = 25.$$

As this is an identity, we note that their derivatives must also be equal:

$$\frac{d}{dx}\left(x^2 + f(x)^2\right) = \frac{d}{dx}(25).$$

Applying our various rules, we end up with $2x + 2f(x)f'(x) = 0$, i.e.

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

If we plug in the point $(3, 4)$, we have that $x = 3$ and $f(3) = 4$, giving our correct answer $-\frac{3}{4}$. If we plug in the point $(4, -3)$, we have that $x = 4$ and $f(4) = -3$, giving our correct answer $\frac{4}{3}$.

To take this a step further, we don't even need to bother with substituting $y = f(x)$. Leibniz notation allows us to take the derivative of both sides of the given equation and solve for $\frac{dy}{dx}$, while treating y as some function of x. For the particular equation $x^2 + y^2 = 25$, we would do the following work:

$$\begin{aligned}
x^2 + y^2 &= 25 \\
\frac{d}{dx}\left(x^2 + y^2\right) &= \frac{d}{dx}(25) \\
\frac{d}{dx}\left(x^2\right) + \frac{d}{dx}\left(y^2\right) &= 0 \\
2x + \frac{d}{dx}\left(y^2\right) &= 0.
\end{aligned}$$

How would we evaluate $\frac{d}{dx}(y^2)$? Recall that we are treating y as a function of x, therefore we use Theorem 4.36:

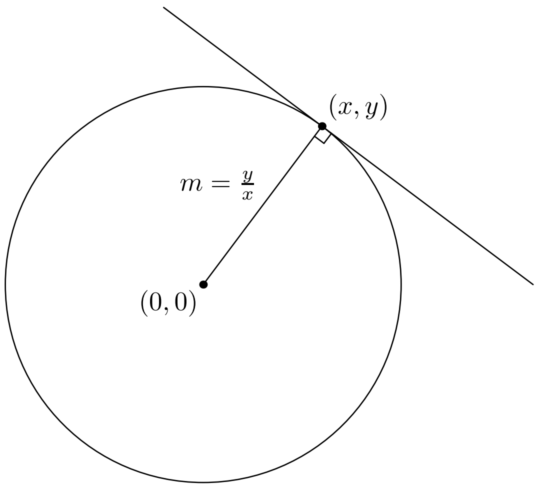
$$2x + 2y\frac{d}{dx}(y) = 0.$$

Nothing that $\frac{d}{dx}(y)$ is just $\frac{dy}{dx}$, we conclude that

$$\begin{aligned} 2x + 2y\frac{dy}{dx} &= 0 \\ \frac{dy}{dx} &= \boxed{-\frac{x}{y}}. \end{aligned}$$

Now, we can simply use this formula by plugging in the appropriate values when given points like $(3,4)$ and $(4,-3)$, and we can confirm that this formula yields the correct answers as well. □

Remark. The result from Example 4.47 is consistent with geometric results, i.e. the tangent line is always perpendicular to the radius at the point of intersection. Consider the following diagram:



As the slope of the line connecting the radius and the point of intersection is $\frac{y}{x}$, the slope of the tangent line must be $-\frac{x}{y}$, since the tangent is perpendicular.

Problem 4.48. Find $\frac{dy}{dx}$ (i.e. y') for $xy^2 + x^2y^3 = 7$.

Solution. As always, we take the derivative of both sides of the equation and then simplify and solve for $\frac{dy}{dx}$. Never forget that we must treat y as a function of x, as $\frac{dy}{dx}$ means that we are evaluating the derivative of y with respect to x. Hence, we must use Theorem 4.36 in the process.

$$\begin{aligned} \frac{dy}{dx}(xy^2 + x^2y^3) &= \frac{dy}{dx}(7) \\ 1 \cdot y^2 + x \cdot 2y \cdot \frac{dy}{dx} + 2x \cdot y^3 + x^2 \cdot 3y^2 \cdot \frac{dy}{dx} &= 0 \\ \frac{dy}{dx} &= \boxed{\frac{-y^2 - 2xy^3}{2xy + 3x^2y^2}}. \end{aligned}$$ □

Problem 4.49. Find $\dfrac{dy}{dx}$ (i.e. y') for $x^2 + xy + y^2 = 12$.

Solution.

$$\begin{aligned} \frac{d}{dx}\left(x^2 + xy + y^2\right) &= \frac{d}{dx}(12) \\ 2x + y + x\frac{dy}{dx} + 2y\frac{dy}{dx} &= 0 \\ \frac{dy}{dx} &= \boxed{-\frac{2x+y}{x+2y}}. \end{aligned}$$

□

Problem 4.50. Find all points on the oblique ellipse from Problem 4.49 when the tangent lines are horizontal and vertical, respectively.

Solution. When the tangent line is horizontal, $-\dfrac{2x+y}{x+2y} = 0$. This simplifies to $y = -2x$, and then we substitute this into the equation $x^2 + xy + y^2 = 12$ to get $x^2 - 2x^2 + 4x^2 = 12$, i.e. $x^2 = 4 \longrightarrow x = \pm 2$, and we get our respective y-values to get the points $\boxed{(2, -4)}$ and $\boxed{(-2, 4)}$.

When the tangent line is vertical, its slope is undefined. The only way $-\dfrac{2x+y}{x+2y}$ can be undefined is when $x + 2y = 0$, i.e. $y = -\dfrac{x}{2}$. We plug this into the first equation to get $x^2 - \dfrac{x^2}{2} + \dfrac{x^2}{4} = 12$, i.e. $x^2 = 16 \longrightarrow x = \pm 4$, and we solve for our respective y-values to get the points $\boxed{(4, -2)}$ and $\boxed{(-4, 2)}$. □

Problem 4.51. Find all horizontal and vertical tangent lines on the oblique ellipse $x^2 - xy + y^2 = 7$.

Solution. First, we find $\dfrac{dy}{dx}$. Note that

$$\begin{aligned} \frac{dy}{dx}(x^2 - xy + y^2) &= \frac{dy}{dx}(7) \\ 2x - \left(y + x\frac{dy}{dx}\right) + 2y\frac{dy}{dx} &= 0 \\ 2x - y - x\frac{dy}{dx} + 2y\frac{dy}{dx} &= 0 \\ \frac{dy}{dx} &= \frac{y-2x}{2y-x}. \end{aligned}$$

The tangent line is horizontal when $-\dfrac{y-2x}{2y-x} = 0$, or $y = 2x$. We substitute this into the equation $x^2 - xy + y^2 = 7$ to get $x^2 - x(2x) + (2x)^2 = 7$, which simplifies to $x = \pm\sqrt{\dfrac{7}{3}}$, so our horizontal tangent lines are $y = 2\sqrt{\dfrac{7}{3}}$ and $y = -2\sqrt{\dfrac{7}{3}}$.

The tangent line is vertical when its slope is undefined. The only way $-\dfrac{y-2x}{2y-x}$ can be undefined is when $2y - x = 0$, i.e. $x = 2y$. We plug this into the initial equation to get $(2y)^2 - (2y)y + y^2 = 7$ and obtain $y = \pm\sqrt{\dfrac{7}{3}}$, so our vertical tangent lines are $x = 2\sqrt{\dfrac{7}{3}}$ and $x = -2\sqrt{\dfrac{7}{3}}$. □

Problem 4.52. Find $\dfrac{dy}{dx}$ (i.e. y') for $x\sin(xy^3)+\sin^2(y)=1$.

Solution. We repeatedly apply Theorem 4.36, so you should be familiar with using it.

$$\begin{aligned}
\frac{d}{dx}\left(x\sin(xy^3)+\sin^2(y)\right) &= \frac{d}{dx}(1)\\
\frac{d}{dx}(x)\cdot\sin(xy^3)+x\cdot\frac{d}{dx}\left(\sin(xy^3)\right)+2\sin(y)\cdot\frac{d}{dx}\left(\sin(y)\right) &= 0\\
\sin(xy^3)+x\cdot\cos(xy^3)\cdot\frac{d}{dx}\left(xy^3\right)+2\sin(y)\cos(y)\frac{dy}{dx} &= 0\\
\sin(xy^3)+x\cdot\cos(xy^3)\left(y^3+x\cdot 3y^2\frac{dy}{dx}\right)+2\sin(y)\cos(y)\frac{dy}{dx} &= 0\\
\sin(xy^3)+xy^3\cos(xy^3)+3x^2y^2\cos(xy^3)\frac{dy}{dx}+2\sin(y)\cos(y)\frac{dy}{dx} &= 0\\
\left(3x^2y^2\cos(xy^3)+2\sin(y)\cos(y)\right)\frac{dy}{dx} &= -\left(\sin(xy^3)+xy^3\cos(xy^3)\right)
\end{aligned}$$

Thus, we have

$$\frac{dy}{dx}=\boxed{-\frac{\sin(xy^3)+xy^3\cos(xy^3)}{3x^2y^2\cos(xy^3)+\sin(2y)}}.$$

□

Problem 4.53. Find $\dfrac{dy}{dx}$ for the equation $\sin(xy)=x^2+y^2$.

Solution.

$$\begin{aligned}
\frac{d}{dx}\left(\sin(xy)\right) &= \frac{d}{dx}\left(x^2+y^2\right)\\
\cos(xy)\frac{d}{dx}(xy) &= 2x+2y\frac{dy}{dx}\\
\cos(xy)\left(y+x\frac{dy}{dx}\right) &= 2x+2y\frac{dy}{dx}\\
y\cos(xy)+x\cos(xy)\frac{dy}{dx} &= 2x+2y\frac{dy}{dx}\\
\frac{dy}{dx} &= \boxed{\frac{y\cos(xy)-2x}{2y-x\cos(xy)}}.
\end{aligned}$$

□

Exercise 4.54. Find $\dfrac{dy}{dx}$ for $x^3+y^3=4$.

Exercise 4.55. Find $\dfrac{dy}{dx}$ for $y=\sin(3x+4y)$.

Exercise 4.56. Find $\dfrac{dy}{dx}$ for $y=x^2y^3+x^3y^2$.

Exercise 4.57. Find $\dfrac{dy}{dx}$ for $\cos^2 x+\cos^2 y=\cos(2x+2y)$.

Exercise 4.58. Find $\dfrac{dy}{dx}$ for $x=\sqrt{x^2+y^2}$.

Problem 4.59. Find $\dfrac{d^2y}{dx^2}$ for $x^2+y^2=25$.

Solution. From Example 4.47, we have found that $\frac{dy}{dx} = -\frac{x}{y}$. Note that $\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right)$, so $\frac{d^2y}{dx^2} = \frac{d}{dx}\left(-\frac{x}{y}\right)$. We then use Theorem 4.23 to get

$$\begin{aligned}\frac{d^2y}{dx^2} &= \frac{y\cdot(-1)-(-x)\frac{dy}{dx}}{y^2}\\ &= \frac{-y-\frac{x^2}{y}}{y^2}\\ &= \frac{-x^2-y^2}{y^3}\\ &= -\frac{x^2+y^2}{y^3}.\end{aligned}$$

We are almost done! At this point, we can simply use the equation we are given, $x^2+y^2=25$, and substitute in the appropriate value, to get $\frac{d^2y}{dx^2} = \boxed{-\frac{25}{y^3}}$. □

Exercise 4.60. Find $\frac{d^2y}{dx^2}$ for $x^2+xy+y^2=1$.

§4.3 Related Rates

Recall the formula

$$\text{rate} = \frac{\text{distance}}{\text{time}}.$$

Depending on the problem, rate can also refer to speed or velocity.

When applying derivatives to problems about rate, we consider all measurements as functions of time, as such:

$$\underbrace{s(\overbrace{t}^{t\,=\,\text{time}})}_{\text{something you can measure}}$$

Definition 4.61. The **average rate of change** of s on the interval $[t_0, t_1]$ is equal to

$$\frac{s(t_1)-s(t_0)}{t_1-t_0} = \frac{\Delta s}{\Delta t}.$$

Note that this is just the slope of the secant line connecting points $(t_0, s(t_0))$ and $(t_1, s(t_1))$.

We then consider the rate of change at the point t_0.

$$\lim_{t_1\to t_0}\frac{s(t_1)-s(t_0)}{t_1-t_0} = s'(t_0).$$

We will appropriately use Δs and Δt to represent the differences $s(t_1)-s(t_0)$ and t_1-t_0, respectively.

Definition 4.62. The **instantaneous rate of change** of s at point t is equal to $s'(t)$.

$$\lim_{\Delta t\to 0}\frac{\Delta s}{\Delta t} = \frac{ds}{dt}.$$

This represents the rate of change at a single point in time, t, rather than taking the average of two points.

We now introduce the concept of related rates problems with a series of examples.

Example 4.63

Consider a spherical balloon which is being blown up. Its volume is changing at a constant rate, K. How does the radius of the balloon change? What about its surface area?

Solution. If the volume of the balloon is changing at a constant rate K, then

$$\frac{dV}{dt} = K.$$

We want to find the change in radius, i.e. $\frac{dr}{dt}$.

However, we keep in mind that V and r are functions of time t (always remember that measurements are functions of time!).

For a sphere, we can use the formula of a balloon, i.e.

$$V = \frac{4}{3}\pi r^3.$$

This is our main *relation* between V and r, hence 'related rates.'

We proceed with implicit differentiation; take the derivative of both sides and then use Theorem 4.36 on V and r as necessary:

$$\begin{aligned}\frac{d}{dt}(V) &= \frac{d}{dt}\left(\frac{4}{3}\pi r^3\right)\\ \frac{dV}{dt} &= 4\pi r^2\frac{dr}{dt}\\ \frac{dr}{dt} &= \boxed{\frac{K}{4\pi r^2}}.\end{aligned}$$

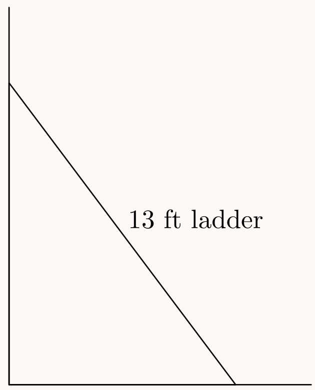
Let S denote the surface area of the balloon. Then the surface area of a sphere is known to be

$$S = 4\pi r^2.$$

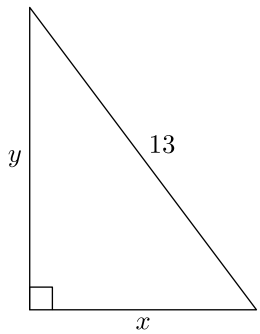
We implicitly differentiate to get $\frac{dS}{dt} = \frac{d}{dt}(4\pi r^2)$,which simplifies to $\frac{dS}{dt} = 8\pi r\frac{dr}{dt}$. We have already found that $\frac{dr}{dt} = \frac{K}{4\pi r^2}$, so we plug this back into the equation to get $\frac{dS}{dt} = \boxed{\frac{2K}{r}}$. □

Example 4.64

How fast is the top of the ladder falling down the wall when the top is at a height of 12 ft. while the bottom of the ladder is being pushed outward 6 inches per minute?



Solution. First, we draw a diagram that simplifies the problem:



For the changing lengths, we label them with variables x and y.

As it states in the problem, the bottom of the ladder is moving outward at 6 inches per minute. We can interpret this as $\frac{dx}{dt} = \frac{1}{2}\frac{\text{ft.}}{\text{min.}}$.

As we want to find the rate of y falling down, we want $\frac{dy}{dt}$ when $y = 12$.

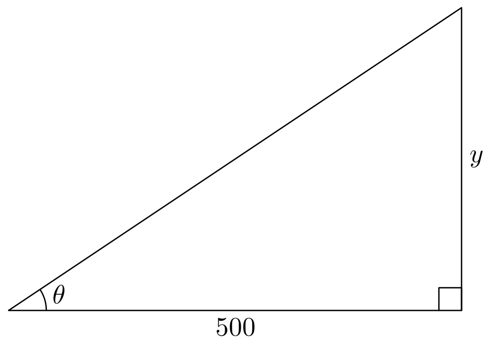
As we have a right triangle, we take advantage of the Pythagorean Theorem: $x^2 + y^2 = 169$. We differentiate both sides to get $\frac{d}{dt}\left(x^2 + y^2\right) = (169)$. Remember, we always treat x and y as functions of time, therefore we use the chain rule to get

$$2x\frac{dx}{dt} + 2y\frac{dy}{dt} = 0.$$

We plug in $\frac{dx}{dt} = \frac{1}{2}$ to get $10 \cdot \frac{1}{2} + 24 \cdot \frac{dy}{dt} = 0$, which yields $\frac{dy}{dt} = -\frac{5}{24}$. It is numerically negative, but based on the word problem, we interpret this as the ladder falling down at $\boxed{\frac{5}{2}\frac{\text{in.}}{\text{min.}}}$. $\square$

Problem 4.65. You are standing 500 ft. away from a tiny rocket. The rocket rises at a rate of 100 feet per second. How fast is its angle of elevation from you changing when the rocket is $500\sqrt{3}$ ft. high?

Solution. Again, we draw a simplified diagram to interpret the problem mathematically:



We are given that y increases 100 feet per second, i.e. $\frac{dy}{dt} = 100$.

We want the change in θ, which is $\frac{d\theta}{dt}$, when $y = 500\sqrt{3}$.

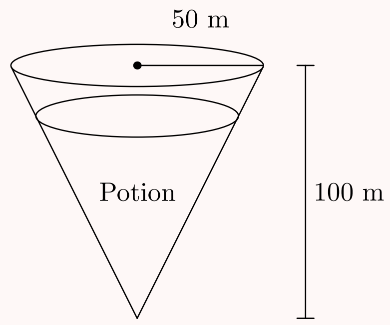
Using this right triangle relationship, we can deduce that $\tan\theta = \frac{y}{500}$. We differentiate this to get

$$\sec^2\theta \frac{d\theta}{dt} = \frac{1}{500}\frac{dy}{dt}.$$

Note that $y = 500\sqrt{3}$ when $\theta = \frac{\pi}{3}$, i.e. $\sec\theta = 2$. Substituting that and $\frac{dy}{dt} = 100$ into the equation, we therefore have $4\frac{d\theta}{dt} = \frac{1}{500} \cdot 100$, from which we solve to get $\frac{d\theta}{dt} = \frac{1}{20}$ radian per second, i.e. the angle of elevation is increasing at $\boxed{\frac{1}{20}\frac{\text{radian}}{\text{second}}}$. □

Example 4.66

Consider a cone-shaped cauldron (with a radius of 50 meters and a height of 100 meters) that holds a potion. The potion leaks out at 2 cubic meters per minute. How fast is the height of the potion changing when the height of the potion in the cauldron is 80 meters?



Solution. Let r and h be the radius and height of the cone with the leaking potion. We are given that $\frac{dV}{dt} = -2$. We want $\frac{dh}{dt}$ when $h = 80$.

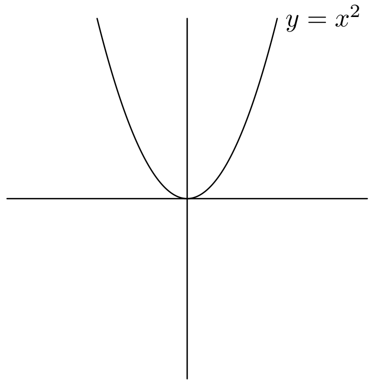
We use the fact that the cone with the leaking potion is similar to the cone of radius 50 meters and height 100 meters. Since the radius to height ratio is $\frac{50}{100} = \frac{1}{2}$, we know that $\frac{r}{h} = \frac{1}{2}$. Furthermore, we know the volume of the cone is $V = \frac{1}{3}\pi r^2 h$.

We rearrange to get $r = \dfrac{h}{2}$ and substitute this into the volume formula to get $V = \dfrac{1}{12}\pi h^3$, then differentiate it to get

$$\frac{dV}{dt} = \frac{1}{4}\pi h^2 \frac{dh}{dt}.$$

We plug in the given $\dfrac{dV}{dt} = -2$ and $h = 80$ to get that $\dfrac{dh}{dt} = -\dfrac{1}{800\pi}$ meters per minute, i.e. the height is decreasing at $\boxed{\dfrac{1}{800\pi} \text{ meters per minute}}$. □

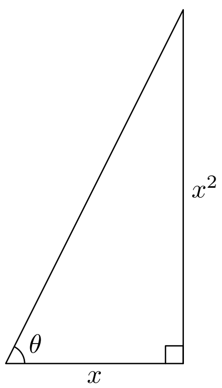
Problem 4.67. Eric is walk along the path of the graph $y = x^2$, as shown:



His x-coordinate increases 10 meters per second. How fast is the angle of inclination from the origin changing when his x-coordinate is 3 meters?

Solution. We let x denote Eric's x-coordinate. We are given that his x-coordinate increases 10 meters per second, which means that $\dfrac{dx}{dt} = 10$. We let θ be the angle of inclination from the origin. We want to find $\dfrac{d\theta}{dt}$ when $x = 3$.

We note that the angle of inclination from the origin is simply the angle between the line containing the point on $y = x^2$ and the origin, and the x-axis. We also have to consider x, the x-coordinate. These relationships motivate us to draw a simplified diagram, as such:



From this right triangle, it becomes apparent that $\tan\theta = x$, and this is our relation. We differentiate both sides with respect to t to get $\sec^2\theta \dfrac{d\theta}{dt} = \dfrac{dx}{dt}$.

How would we find $\sec^2\theta$? We can use the identity $\tan^2\theta + 1 = \sec^2\theta$ (which can be derived from the Pythagorean identity $\sin^2\theta + \cos^2\theta = 1$).

When $x = 3$, then $\tan\theta = 3$, therefore $\sec^2\theta = 3^2 + 1 = 10$, so we have

$$10\frac{d\theta}{dt} = 10.$$

as we are given that $\frac{dx}{dt} = 10$. We solve that $\frac{d\theta}{dt} = 1$, therefore we can conclude that the angle of inclination from the origin increases at a rate of $\boxed{\text{1 radian per second}}$. □

Problem 4.68. The length of a rectangle increases by 10 cm. per hour. Its width decreases by 2 cm. per hour. When the length is 60 cm. and the width is 80 cm., determine how fast each of the following is changing:

a) Perimeter

b) Area

c) Length of diagonal

Solution. Let the length be l and width be w. The problem statement implies that $\frac{dw}{dt} = -2$ and $\frac{dl}{dt} = 10$. Let the perimeter, area, and length of diagonal be P, A, and L respectively. Therefore the problem is asking us to find $\frac{dP}{dt}$, $\frac{dA}{dt}$, and $\frac{dL}{dt}$, when $w = 80$ and $l = 60$.

a) We have our relation $P = 2(l + w)$. We differentiate both sides to get

$$\frac{dP}{dt} = \frac{d}{dt}\left(2(l+w)\right).$$

Applying the chain rule and substituting in the given information gives us

$$\frac{dP}{dt} = 2\left(\frac{dl}{dt} + \frac{dw}{dt}\right) = 2(10-2) = 16.$$

Therefore, the perimeter is increasing at a rate of $\boxed{\text{16 cm. per hour}}$.

b) Similarly, we have that $A = lw$, and differentiating both sides and simplifying yield:

$$\begin{aligned}\frac{dA}{dt} &= \frac{d}{dt}(lw)\\ &= \frac{dl}{dt}w + l\frac{dw}{dt}\\ &= 10\cdot 80 + 60\cdot -2\\ &= 680.\end{aligned}$$

Thus, the area increases at a rate of $\boxed{\text{680 cm.}^2\text{ per hour}}$.

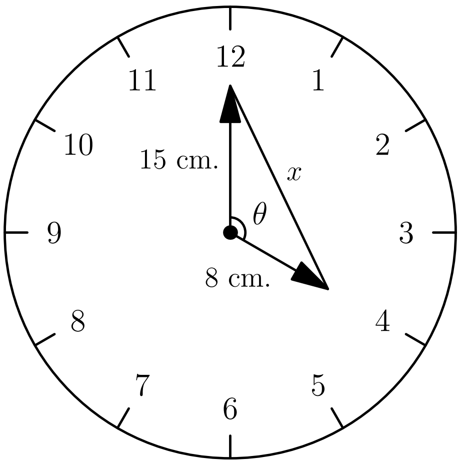
c) We can either use the relation $L = \sqrt{l^2 + w^2}$, or $L^2 = l^2 + w^2$, then differentiate accordingly. I will focus on $L = \sqrt{l^2 + w^2}$:

$$\frac{dL}{dt} = \frac{d}{dt}\left(\sqrt{l^2 + w^2}\right)$$

$$\begin{aligned}
&= \frac{1}{2\sqrt{l^2+w^2}} \cdot \frac{d}{dt}\left(l^2+w^2\right) \\
&= \frac{1}{2\sqrt{l^2+w^2}} \cdot \left(2l\frac{dl}{dt} + 2w\frac{dw}{dt}\right) \\
&= \frac{1}{2\sqrt{60^2+80^2}} \cdot (2 \cdot 60 \cdot 10 + 2 \cdot 80 \cdot -2) \\
&= \frac{1}{200}(1200 - 320) \\
&= \frac{22}{5}.
\end{aligned}$$

The length of the diagonal increases at a rate of $\boxed{\frac{22}{5} \text{ cm. per hour}}$. □

Problem 4.69. Observe the clock below. The minute hand is 15 cm. and the hour hand is 8 cm. How fast is the distance between the tips of the hour hand and minute hand changing at the given time?



Solution. Let the distance between the tips of the hour hand and minute hand be x, and the angle between the two hands be θ. We intend to use the Law of Cosines on the triangle formed by the side lengths (the two hands of the clock) and the third side, which is x. The problem is asking us to find $\frac{dx}{dt}$ when $\theta = \frac{2\pi}{3}$, which signifies the time of 4 o'clock.

Starting off with Law of Cosines, we have:

$$\begin{aligned}
x^2 &= 15^2 + 8^2 - 2(15)(8)\cos\theta \\
&= = 225 + 64 - 240\cos\theta \\
&= 289 - 240\cos\theta.
\end{aligned}$$

Differentiate both sides to get

$$2x\frac{dx}{dt} = 240\sin\theta\frac{d\theta}{dt}.$$

When $\theta = \frac{2\pi}{3}$, $x = \sqrt{289 - 240\cos\left(\frac{2\pi}{3}\right)} = \sqrt{409}$.

Now we approach the main question: what is $\frac{d\theta}{dt}$? Note that the minute hand is changing at 2π radians per hour, and the hour hand is changing at $\frac{\pi}{6}$ radians per hour.

We then consider the context of this problem: as we are approaching 4 o'clock, the quicker minute hand is approaching the slower hour hand, and therefore θ (which is the angle between the two hands) is decreasing, so $\frac{d\theta}{dt} = \frac{\pi}{6} - 2\pi = -\frac{11\pi}{6}$ radians per hour.

If the time was instead 8 o'clock, $\frac{d\theta}{dt}$ would instead be $2\pi - \frac{\pi}{6} = \frac{11\pi}{6}$ radians per hour, as the minute hand would have to keep turning in the clockwise direction until it meets the hour hand (which will have slowly moved two-thirds of the way from the 8 to the 9 markings), and so θ would be increasing in that scenario.

Plugging in our information, we have

$$2\sqrt{409}\frac{dx}{dt} = 240 \cdot \frac{\sqrt{3}}{2} \cdot -\frac{11\pi}{6}.$$

Solving, we get $\frac{dx}{dt} = -\frac{110\pi\sqrt{3}}{\sqrt{409}}$, so our final answer would be that the distance between the tips of the hour hand and the minute hand is changing at a rate of $\boxed{-\frac{110\pi\sqrt{3}}{\sqrt{409}} \text{ cm. per hour}}$. □

§4.4 Significance of the Derivative

Definition 4.70. $f(x)$ has a **local maximum** at $x = a$ if $\exists\delta > 0$ such that $\forall x \in (a-\delta, a+\delta)$, $f(x) \leq f(a)$.

Definition 4.71. $f(x)$ has a **local minimum** at $x = a$ if $\exists\delta > 0$ such that $\forall x \in (a-\delta, a+\delta)$, $f(x) \geq f(a)$.

Definition 4.72. An **extremum** refers to either a local maximum or minimum.

Theorem 4.73

If the function f has a local extremum at $x = a$, then either $f'(a) = 0$ or $f'(a)$ does not exist.

Proof. Without loss of generality, let there be a local maximum at $x = a$. The proof for the local minimum will be analogous.

Suppose $f'(a)$ exists, so we have

$$f'(a) = \lim_{x\to a}\frac{f(x)-f(a)}{x-a}.$$

This implies that $f'(a) = \lim_{x\to a^+}\frac{f(x)-f(a)}{x-a} = \lim_{x\to a^-}\frac{f(x)-f(a)}{x-a}$. We analyze the one-sided limits separately:

1. If $x \to a^+$, then $x - a > 0$ by assumption. As $x = a$ is a local maximum, $f(x) - f(a) \leq 0$ by Definition 4.70, therefore $\frac{f(x)-f(a)}{x-a} \leq 0$, i.e. $\lim_{x\to a^+}\frac{f(x)-f(a)}{x-a} \leq 0$.

2. If $x \to a^-$, then $x - a < 0$ by assumption. Since $f(x) - f(a) \le 0$ by Definition 4.70, we conclude that $\dfrac{f(x) - f(a)}{x - a} \ge 0$, i.e. $\displaystyle\lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} \ge 0$.

We have that $\displaystyle\lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} \le 0$ and $\displaystyle\lim_{x \to a^-} \frac{f(x) - f(a)}{x - a} \ge 0$, and since these one-sided limits must be equal, we necessarily have $\displaystyle\lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a^-} \frac{f(x) - f(a)}{x - a} = 0$. Therefore $f'(a)$ must be 0.

If $f'(a)$ does not exist, then we have a cusp at $(a, f(a))$ which would be the local extremum. □

Definition 4.74. The function f has a **critical point** at $x = a$ if $f'(a) = 0$ or $f'(a)$ does not exist.

Theorem 4.75 (Extreme Value Theorem)
The maximum and minimum values of a continuous function f on $[a, b]$ occur either at the end points or critical points.

Proof. This theorem follows from Theorem 3.3, Theorem 4.73, and Definition 4.74. □

The following problems will demonstrate the usefulness of this theorem regarding minimizing and maximizing functions.

Problem 4.76. Maximize and minimize $f(x) = x^2 - x$ on $[-4, 4]$.

Solution. Note that $f'(x) = 2x - 1$, so the only critical point is $x = \dfrac{1}{2}$. Now, we examine all end points and critical points:

$$\begin{aligned} f(-4) &= 20, \\ f\left(\frac{1}{2}\right) &= -\frac{1}{4}, \\ f(4) &= 12. \end{aligned}$$

Therefore, the maximum and minimum values of f on $[-4, 4]$ are $\boxed{20}$ and $\boxed{-\dfrac{1}{4}}$ respectively. □

Problem 4.77. Maximize and minimize $f(x) = x^3 - 3x + 1$ on $[-4, 4]$.

Solution. We have $f'(x) = 3x^2 - 3$, so the critical points are $x = 1, -1$. Now, we examine all end points and critical points:

$$\begin{aligned} f(-4) &= -51, \\ f(-1) &= -1, \\ f(1) &= 3, \\ f(4) &= 53. \end{aligned}$$

Therefore, the maximum and minimum values are $\boxed{53}$ and $\boxed{-51}$ respectively. □

Problem 4.78. Maximize and minimize $f(x) = \sin 2x$ on $[0, \pi]$.

Solution. Since $f'(x) = 2\cos 2x$, the critical points are $x = \dfrac{\pi}{4}, \dfrac{3\pi}{4}$. Now, we examine all end points and critical points:

$$\begin{aligned} f(0) &= 0, \\ f\left(\frac{\pi}{4}\right) &= 1, \\ f\left(\frac{3\pi}{4}\right) &= -1, \\ f(\pi) &= 0. \end{aligned}$$

Hence, the maximum and minimum values of f on $[0, \pi]$ are $\boxed{1}$ and $\boxed{-1}$ respectively. □

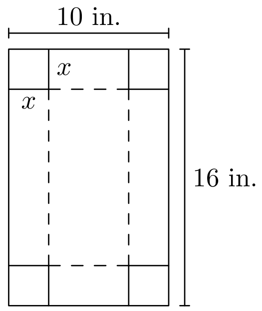
Problem 4.79. Maximize and minimize $f(x) = 4x + \dfrac{3}{x}$ on $\left[\dfrac{1}{2}, 5\right]$.

Solution. After we obtain $f'(x) = 4 - \dfrac{3}{x^2}$, we solve for the critical points: $4 - \dfrac{3}{x^2} = 0 \to 4x^2 = 3$, which gives us $x = \pm\dfrac{\sqrt{3}}{2}$. Since we are considering the interval $\left[\dfrac{1}{2}, 5\right]$, we discard the negative value of x. Now, we examine all end points and critical points:

$$\begin{aligned} f\left(\frac{1}{2}\right) &= 8, \\ f\left(\frac{\sqrt{3}}{2}\right) &= 4\sqrt{3}, \\ f(5) &= \frac{103}{5}. \end{aligned}$$

Thus, the maximum and minimum values of f on $\left[\dfrac{1}{2}, 5\right]$ are $\boxed{\dfrac{103}{5}}$ and $\boxed{4\sqrt{3}}$ respectively. □

Problem 4.80. Squares of side length x will be cut from each corner of a 10×16 in. cardboard such that the remaining will be folded up into an open box. Find the value of x which will maximize the volume of the box.



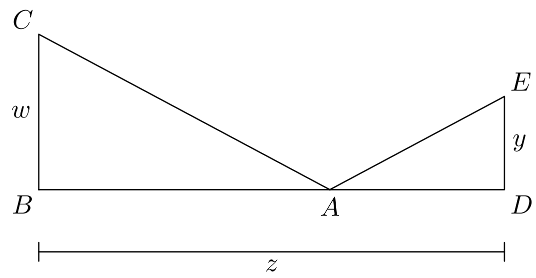
Solution. We are asked to maximize the function $V(x)$ which represents the volume of the resulting box. First, we must consider the constraints of the cardboard. Considering the 10 inch side, the side length of the squares cannot be greater than 5 inches. Therefore we must maximize $V(x)$ over the interval $[0, 5]$. We include the edge cases so we can apply Theorem 4.75.

The height of the box is x, and the dimensions of the base of the box would be $10-2x$ and $16-2x$, since each side is reduced by the 2 squares with side length x cut off from the corners. Therefore, $V(x) = x(16-2x)(10-2x) = 4x^3 - 52x^2 + 160x$. Then $V'(x) = 12x^2 - 104x + 160 = 4(3x-20)(x-2)$, so the critical points are $x = \frac{20}{3}, 2$. However, since $\frac{20}{3}$ is not in the interval $[0,5]$, we ignore this value, so we only check $x = 0, 2, 5$, as shown:

$$\begin{aligned} V(0) &= 0, \\ V(2) &= 144, \\ V(5) &= 0. \end{aligned}$$

Therefore, the maximum volume would be $\boxed{144 \text{ in.}^3}$. □

Problem 4.81. Consider the following diagram:



Suppose $\triangle ABC$ and $\triangle ADE$ are right triangles, with a common vertex at A. Let $BC = w$, $DE = y$, and $BD = z$. Locate A such that the sum of the hypotenuses AC and AE is minimal. Then, demonstrate that regardless of the values of w, y, z, the minimum sum of AC and AE occurs when $\angle BAC \cong \angle DAE$.

Solution. For some $x \in (0, z)$, let $BA = z - x$ and $DA = x$, as this problem would not make sense if x was equal to either of the end points (resulting in one triangle disappearing completely). By the Pythagorean Theorem, we have $AC = \sqrt{w^2 + (z-x)^2}$ and $AE = \sqrt{y^2 + x^2}$. Then consider

$$f(x) = \sqrt{y^2 + x^2} + \sqrt{w^2 + (z-x)^2},$$

which will represent the sum of the hypotenuses based on x. We wish to minimize this function over $(0, z)$. Then

$$f'(x) = \frac{x}{\sqrt{x^2 + y^2}} - \frac{z-x}{\sqrt{(z-x)^2 + w^2}}.$$

We solve for the critical point(s):

$$\begin{aligned} \frac{x}{\sqrt{x^2+y^2}} - \frac{z-x}{\sqrt{(z-x)^2+w^2}} &= 0 \\ \frac{x}{\sqrt{x^2+y^2}} &= \frac{z-x}{\sqrt{(z-x)^2+w^2}} \\ \frac{x^2}{x^2+y^2} &= \frac{(z-x)^2}{(z-x)^2+w^2} \end{aligned}$$

$$\begin{aligned}\frac{x^2+y^2}{x^2} &= \frac{(z-x)^2+w^2}{(z-x)^2}\\ 1+\frac{y^2}{x^2} &= 1+\frac{w^2}{(z-x)^2}\\ \frac{y^2}{x^2} &= \frac{w^2}{(z-x)^2}\\ \frac{y}{x} &= \frac{w}{z-x}.\end{aligned}$$

We then solve to get $x = \dfrac{yz}{w+y}$ as a critical point. After much simplification, we get

$$f\left(\frac{yz}{w+y}\right) = \sqrt{(w+y)^2+z^2},$$

which is our minimum sum of the hypotenuses.

In order to achieve this minimal sum, we must have $\dfrac{y}{x} = \dfrac{w}{z-x}$, as established earlier. This implies that $\dfrac{DE}{DA} = \dfrac{BC}{BA}$, so $\triangle ABC \sim \triangle ADE$. Thus, $\angle BAC \cong \angle DAE$. □

Next, we introduce two very significant results.

Theorem 4.82 (Rolle's Theorem)

If f is continuous on $[a,b]$ and differentiable on (a,b), and $f(a) = f(b)$, then $\exists a < c < b$ such that $f'(c) = 0$.

Proof. By Theorem 3.3, f has a maximum and a minimum on $[a,b]$. By Theorem 4.75, these can occur either at end points or critical points.

If either is at a critical point c where $a < c < b$, then $f'(c) = 0$ since f is differentiable on (a,b). In this case, we are done.

Otherwise, if either is at an end point, f must be constant on $[a,b]$ since $f(a) = f(b)$ and this common value would both be the maximum and the minimum. Thus, we have $f'(c) = 0\ \forall c \in (a,b)$. □

Problem 4.83. Prove $f(x) = x^3 + x + 1$ has exactly one root.

Proof. Note that $f(-1) = -1$ and $f(0) = 1$, so by Theorem 3.1, $\exists -1 < a < 0$ such that $f(a) = 0$. Thus, there exists at least one root.

Assume there is another root b. Then $f(a) = f(b) = 0$. Clearly f is continuous and differentiable over the domain, so by Theorem 4.82, $\exists a < c < b$ such that $f'(c) = 0$. However, $f'(x) = 3x^2 + 1 > 0\ \forall x$, thus $f'(c)$ cannot be 0. We arrive at a contradiction, so a is the one and only root of $f(x)$. □

Theorem 4.84 (Mean Value Theorem)

If f is continuous on $[a,b]$ and differentiable on (a,b), then $\exists a < c < b$ such that

$$f'(c) = \frac{f(b)-f(a)}{b-a}.$$

Proof. Consider the following function:

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Note that $g(a) = f(a)$, and $g(b) = f(b) - (f(b) - f(a)) = f(a)$.

As g is defined by subtracting a linear term from $f(x)$ (so all of its components are continuous and differentiable), g is also continuous on $[a, b]$ and differentiable on (a, b).

Thus, by Theorem 4.82, $\exists a < c < b$ such that $g'(c) = 0$, i.e.

$$f'(c) - \frac{f(b) - f(a)}{b - a} = 0,$$

which rearranges to

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

as desired. □

From this theorem, we have immediate results:

Corollary 4.85

If $f'(c) = 0 \ \forall c$ on a certain interval, then f is constant on that interval.

Proof. Let a, b be any two points on that interval. Then by Theorem 4.84, $\exists a < c < b$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a} = 0.$$

Thus, $f(a) = f(b)$ for any points a, b on that interval, so f is constant on that interval. □

Corollary 4.86

If $f' = g'$ on an interval, then $\exists c \in \mathbb{R}$ such that $f(x) = g(x) + c$.

Proof. Let $h(x) = f(x) - g(x)$. Then $h'(x) = f'(x) - g'(x) = 0$ on an interval. By Corollary 4.85, h is constant on that interval, i.e. $h(x) = c$, so $f(x) - g(x) = c$, i.e. $f(x) = g(x) + c$. □

Definition 4.87. f is **increasing** on an interval if $\forall a, b$ on the interval, $a < b \to f(a) < f(b)$.

Definition 4.88. f is **decreasing** on an interval if $\forall a, b$ on the interval, $a < b \to f(a) > f(b)$.

Corollary 4.89

If $f' > 0$ on an interval, f is increasing on that interval. Similarly, if $f' < 0$ on an interval, f is decreasing on that interval.

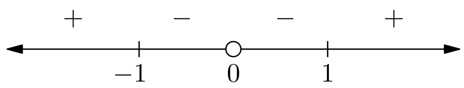
Proof. For all a, b on the interval, let $a < b$ without loss of generality. By Theorem 4.84, $\exists a < c < b$ such that $f'(c) = \dfrac{f(b) - f(a)}{b - a} > 0$, so $f(b) - f(a) > 0$, i.e. $f(a) < f(b)$. Thus, $\forall a, b$ on the interval, $a < b \to f(a) < f(b)$, so f is increasing.

The proof for the latter statement is analogous. □

From Corollary 4.89, we can now determine which parts of a given function are increasing or decreasing, using a method called the **first derivative number line**, which will be demonstrated through the following examples.

Problem 4.90. Determine the intervals on which the function $f(x) = x + \dfrac{1}{x}$ is increasing and decreasing.

Solution. We find the first derivative to be $f'(x) = 1 - \dfrac{1}{x^2}$, which has critical points at -1 and 1. We use the first derivative number line, as shown below. Note that the function is undefined at $x = 0$, so we have an open circle at 0 and must consider the intervals on both sides of the open circle. Keep in mind that $x = 0$ is technically not considered a critical point, but we still include it on the first derivative number line anyway.



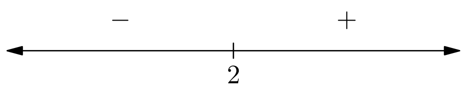
Testing values in between each of the intervals, we find that the function is increasing on $(-\infty, 1]$, decreasing on $[-1, 0)$, decreasing on $(0, 1]$, and increasing on $[1, \infty)$. $\square$

Problem 4.91. Determine the intervals on which the following functions are increasing and decreasing.

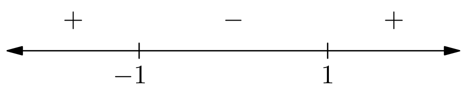
1. $f(x) = 3x - 7$
2. $f(x) = x^2 - 4x + 3$
3. $f(x) = x^3 - 3x$
4. $f(x) = x^3$

Solution. We proceed with the first derivative number line:

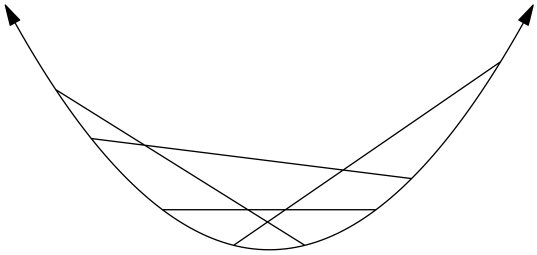
1. Since $f'(x) = 3$, it is always positive, therefore the function is increasing on the entire domain.
2. We get $f'(x) = 2x - 4$, so we get the critical point $x = 2$. Thus, our number line is

   Thus, f is decreasing on $(-\infty, 2]$ and increasing on $[2, \infty)$.
3. We evaluate $f'(x) = 3x^2 - 3$ to get the critical points $x = \pm 1$, so the number line is

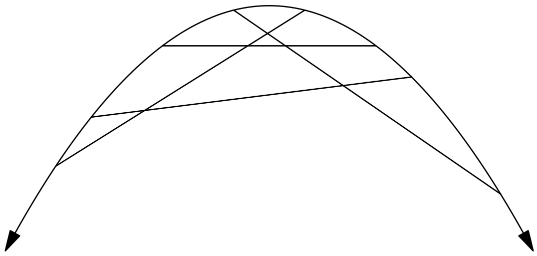
   

   We conclude that f is increasing on $(-\infty, -1]$, decreasing on $[-1, 1]$, and increasing on $[1, \infty)$.
4. We have $f'(x) = 3x^2$, which is clearly always positive. Thus, f is increasing over the entire domain. $\square$



Some secant lines of a convex function

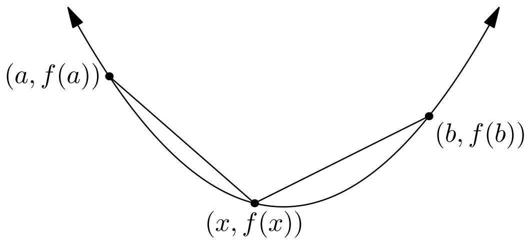
Definition 4.92. $f(x)$ is **convex** on an interval if $\forall a, b$ in the interval, the secant line connecting $(a, f(a))$ to $(b, f(b))$ lies above the graph on that part of the interval.



Some secant lines of a concave function

Definition 4.93. $f(x)$ is **concave** on an interval if $\forall a, b$ in the interval, the secant line connecting $(a, f(a))$ to $(b, f(b))$ lies below the graph on that part of the interval.

Let two arbitrary points in a convex interval be a, b, and WLOG $a < b$. Consider any point $a < x < b$.



By our definition of convexity, $f(x)$ must be less than the corresponding point on the secant line connecting a and b. We can algebraically represent this as

$$\forall a < x < b,\ f(x) < \frac{f(b) - f(a)}{b - a}(x - a) + f(a),$$

where $y = \dfrac{f(b) - f(a)}{b - a}(x - a) + f(a)$ is the equation of the secant line. We can rearrange this inequality as

$$\forall a < x < b,\ \frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(a)}{b - a},$$

which is another way of representing the condition for convexity.

Definition 4.94. For any two points a, b in a convex interval, $\forall a < x < b$, $\dfrac{f(x) - f(a)}{x - a} < \dfrac{f(b) - f(a)}{b - a}$.

Definition 4.95. For any two points a, b in a concave interval, $\forall a < x < b$, $\dfrac{f(x) - f(a)}{x - a} > \dfrac{f(b) - f(a)}{b - a}$.

Theorem 4.96

If $f'' > 0$ over an interval, then f is convex on that interval.

Proof. Let a, b be two points on the interval, and WLOG $a < b$. Let x be some point between a and b.

First, note that f' is increasing by Corollary 4.89. Applying Theorem 4.84 on the intervals (a, x) and (x, b), $\exists c, d$ such that $f'(c) = \dfrac{f(x) - f(a)}{x - a}$ and $f'(d) = \dfrac{f(b) - f(x)}{b - x}$. Since $c < d$ and f' is increasing, we have $f'(c) < f'(d)$, or $\dfrac{f(x) - f(a)}{x - a} < \dfrac{f(b) - f(x)}{b - x}$. Then, note that

$$\begin{aligned}
\frac{f(x) - f(a)}{x - a} &< \frac{f(b) - f(x)}{b - x} \\
(f(x) - f(a))(b - x) &< (f(b) - f(x))(x - a) \\
bf(x) - xf(x) - bf(a) + xf(a) &< xf(b) - af(b) - xf(x) + af(x) \\
bf(x) - bf(a) + xf(a) &< xf(b) - af(b) + af(x) \\
bf(x) - bf(a) + xf(a) + af(a) &< xf(b) - af(b) + af(x) + af(a) \\
bf(x) - bf(a) - af(x) + af(a) &< xf(b) - af(b) - xf(a) + af(a) \\
(b - a)(f(x) - f(a)) &< (x - a)(f(b) - f(a)) \\
\frac{f(x) - f(a)}{x - a} &< \frac{f(b) - f(a)}{b - a}.
\end{aligned}$$

Therefore, the interval between a and b is convex, as desired. □

Exercise 4.97. Is the converse of Theorem 4.96 true? If not, what would be a counterexample?

Theorem 4.98

If $f'' < 0$ over an interval, then f is concave on that interval.

Proof. The proof is analogous to that of the previous theorem. □

Theorem 4.99

If $f'' > 0$ on an interval, its tangent lines lie below the graph.

Proof. By Corollary 4.89, f' is increasing.

Consider a fixed point a on the interval. Then for any arbitrary x on the interval, we have two cases:

- If $x > a$, then by Theorem 4.84, $\exists a < b < x$ such that $f'(b) = \dfrac{f(x) - f(a)}{x - a}$. As f' is increasing, we have $f'(a) < f'(b)$, i.e. $f'(a) < \dfrac{f(x) - f(a)}{x - a}$. This rearranges to $f'(a)(x - a) + f(a) < f(x)$, which is what we wanted.
- If $x < a$, then by Theorem 4.84, $\exists x < b < a$ such that $f'(b) = \dfrac{f(a) - f(x)}{a - x}$. As f' is increasing, we have $f'(b) < f'(a)$, i.e. $\dfrac{f(a) - f(x)}{a - x} < f'(a)$. This rearranges to $f'(a)(x - a) + f(a) < f(x)$, the same result as before. $\square$

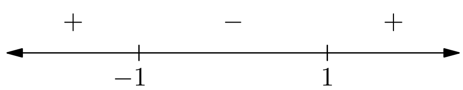
We now have all the tools in calculus to sketch graphs effectively. When given such a task, we consider:

1. The first derivative, which tells us whether f is increasing or decreasing on which intervals.
2. The second derivative, which gives information on the concavity (convex or concave) on which intervals.
3. Any asymptotes; consider values of x that f would be undefined in, the behavior of f as x goes to infinity or negative infinity, or other special values, etc.
4. The function's roots and y-intercept.

Example 4.100

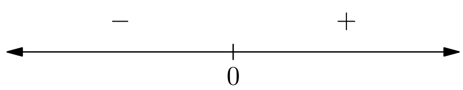
Sketch $f(x) = x^3 - 3x$ and label all relevant points.

Solution. We find that $f'(x) = 3x^2 - 3$, so its critical points are ± 1, and we appropriately set up our first derivative number line and find the signs in each interval:



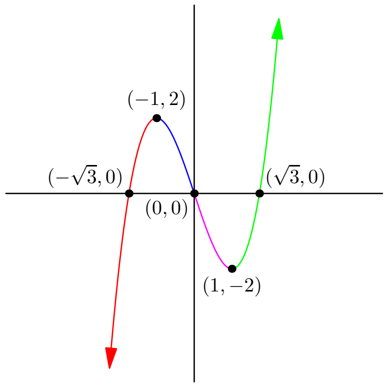
Now we know that f is increasing on $(-\infty, -1]$, decreasing on $[-1, 1]$, and increasing on $[1, \infty)$.

We take the derivative of $f'(x)$ to get that $f''(x) = 6x$. Our only possible point of inflection is 0, so we now set up our second derivative number line and find the signs:



Now we know that f is concave down on $(-\infty, 0]$ and concave up on $[0, \infty)$, and that 0 is a point of inflection.

Lastly, we find the roots, which are $\pm\sqrt{3}$, and it's not hard to see that the graph passes through the origin. It can be noted that since this function is a cubic, there are no asymptotes.



Starting from the left, we begin by sketching a sharply increasing concave curve until we reach -1. This portion is signified as red in the diagram.

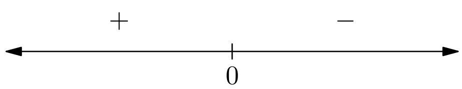
Then, according to the first derivative number line, the function starts decreasing. Keep in mind that the shape is still concave. This portion is represented as blue in the diagram.

Next, the concavity changes at $x = 0$, an inflection point, but the function is still decreasing. This is the pink portion of the graph.

Lastly, we finish the sketch by sharply increasing outwards while maintaining the convex shape. This is the green part of the graph. $\square$

Problem 4.101. Sketch $y = \dfrac{1}{x^2+1}$.

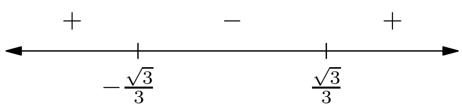
Solution. Its first derivative is $y' = \dfrac{-2x}{(x^2+1)^2}$, so its first derivative number line would be:



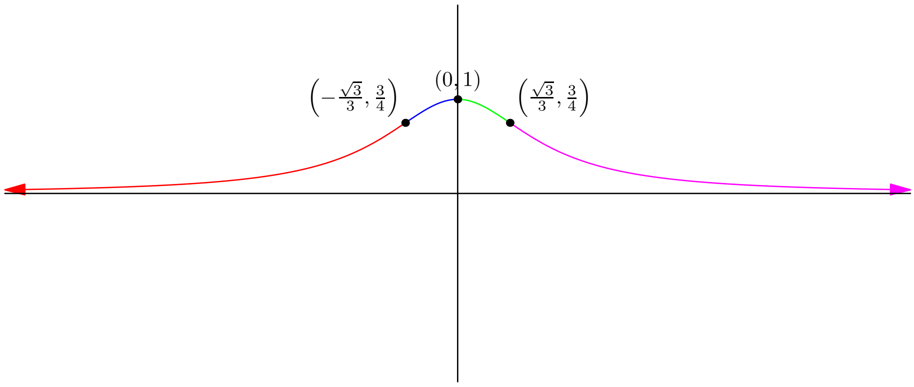
Next, we find our second derivative:

$$\begin{aligned} y'' &= \frac{-2(x^2+1)^2 + 2x \cdot 2(x^2+1) \cdot 2x}{(x^2+1)^4} \\ &= \frac{-2(x^4+2x^2+1) + 8x^2(x^2+1)}{(x^2+1)^4} \\ &= \frac{6x^4+4x^2-2}{(x^2+1)^4} \\ &= \frac{2(x^2+1)(3x^2-1)}{(x^2+1)^4}. \end{aligned}$$

The possible points of inflection are $\pm\dfrac{\sqrt{3}}{3}$, and our second derivative number line would be:



Note that $\displaystyle\lim_{x\to\infty} \frac{1}{x^2+1} = \lim_{x\to-\infty} \frac{1}{x^2+1} = 0$, so $y = 0$ is an asymptote as x goes to negative and positive infinity.



Considering all of this information, we start off our sketch with the red line very close to the asymptote $y = 0$, increasing and convex.

When we hit $x = -\dfrac{\sqrt{3}}{3}$, we have the line (now represented by a blue line in the figure below) become concave down, but still increasing.

When $x = 0$, the graph now is decreasing, so we continue with the line (now green) decreasing.

Lastly, when we reach $x = \dfrac{\sqrt{3}}{3}$, the line (which is now pink) becomes concave up again, but continues to decrease and approach the asymptote $y = 0$ as x goes to infinity. □

Exercise 4.102. Sketch $y = x^2 - \dfrac{1}{x}$.

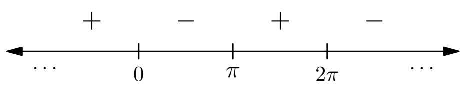
Exercise 4.103. Sketch $y = \dfrac{x}{x^2 + 1}$.

Problem 4.104. Sketch $y = x + \sin x$.

Solution. Note that $y' = 1 + \cos x$. Since $\cos x \in [-1, 1]$, we have $1 + \cos x \geq 0$. Thus, y' is always positive except for critical points $(\ldots - \pi,\ \pi,\ 3\pi,\ \ldots)$, indicating that y is always increasing except at the critical points, at which the tangent lines would be horizontal.

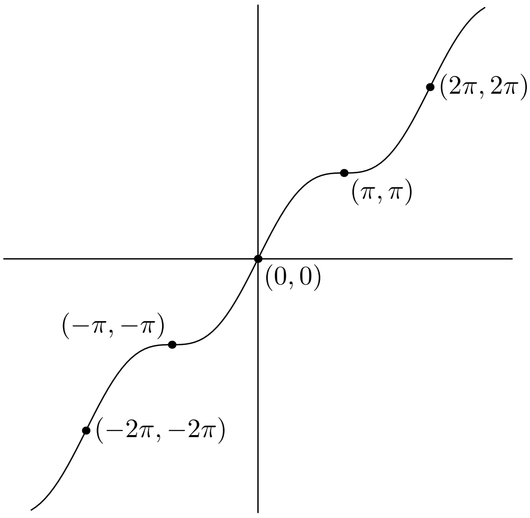
Don't be afraid to deal with infinitely many critical points, because there will probably be a recognizable pattern.

Then, we evaluate the second derivative to be $-\sin x$, so our second derivative number line would be: Notice the pattern that $+$ and $-$ are infinitely alternating. This suggests



that the graph switches concavity at every multiple of π.

After considering this information, the sketch should be similar to this:



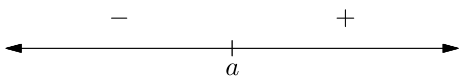
$\square$

Theorem 4.105
If $f'(a) = 0$ and $f''(a) > 0$ then f has a local minimum at $x = a$.

Proof. Note that

$$\begin{aligned} f''(a) &= \lim_{h \to 0} \frac{f'(a+h) - f'(a)}{h} \\ &= \lim_{h \to 0} \frac{f'(a+h)}{h}, \end{aligned}$$

and we are given that $f''(a) > 0$, so $\lim_{h \to 0} \dfrac{f'(a+h)}{h} > 0$. If $h > 0$, then $f'(a+h) > 0$, and if $h < 0$, then $f'(a+h) < 0$. There is no need to consider $h = 0$ since we are taking the limit. Thus, if we consider the first derivative number line, then



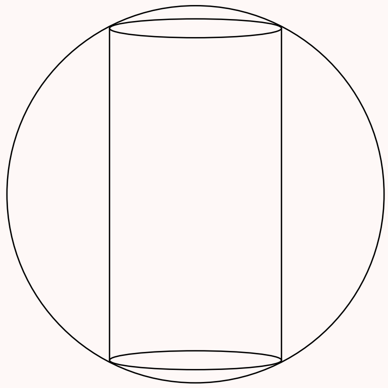
which indicates that $x = a$ is a local minimum. $\square$

Theorem 4.106
If $f'(a) = 0$ and $f''(a) < 0$ then f has a local maximum at $x = a$.

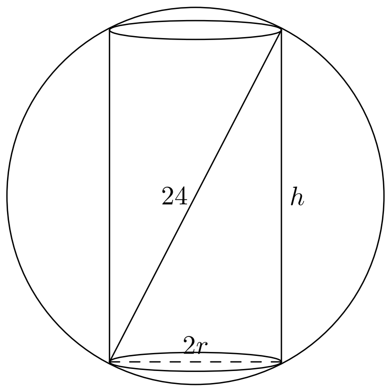
Proof. The proof is nearly identical to that of the previous theorem. $\square$

Example 4.107

If the radius of the sphere is 12, what is the volume of the largest cylinder that can be inscribed in the sphere?



Solution. Let r, h denote the radius and height of the cylinder respectively. Notice that the diagonal of the cylinder is twice the radius of the sphere, and then we can take advantage of a right triangle relationship:



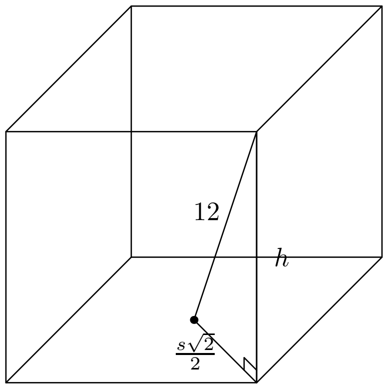
Consider the formula for the volume of a cylinder, $V = \pi r^2 h$. Since there are two variables we cannot find the maximum of this function. However, we have $h^2 + 4r^2 = 576$ from the right triangle, so we solve $r^2 = \dfrac{576 - h^2}{4}$, and substitute this back into the volume formula to get

$$V(h) = \pi \left(\frac{576 - h^2}{4} \right) h = \frac{\pi}{4}(576h - h^3).$$

We differentiate to get $V'(h) = \dfrac{\pi}{4}(576 - 3h^2)$, from which we get the critical point $h = 8\sqrt{3}$. Instead of going through the trouble to set up a first derivative number line, we can plug it into the second derivative. Note that $V''(h) = -\dfrac{3\pi}{2}h$ and thus $V''(8\sqrt{3}) < 0$. Then, by Theorem 4.106, $h = 8\sqrt{3}$ is a local maximum. However, since $8\sqrt{3}$ is the only critical point, it is therefore the global maximum, so the volume is minimized at $V(8\sqrt{3}) = \boxed{768\pi\sqrt{3}}$. $\square$

Problem 4.108. Inside a hemisphere of radius 12 is inscribed a box with a square base. What dimensions will maximize its volume?

Solution. This diagram represents the box with side length s and height h.



Given that it is inscribed in a hemisphere, the distance between the center of the square base (and the base of the hemisphere) and one of the four upper corners of the box must be the radius of the sphere, or 12. By properties of a square, the distance from the center of the base to one of the four lower corners must be half the diagonal of the square, or $\frac{s\sqrt{2}}{2}$.

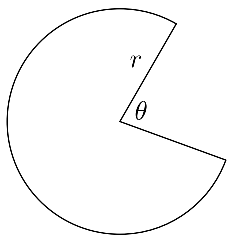
Then, we have a right-triangle relationship and we can use the Pythagorean Theorem to relate all three sides: $h^2 + \left(\frac{s\sqrt{2}}{2}\right)^2 = 144$. This rearranges to $2h^2 + s^2 = 288$.

Given that the formula for the volume of this box is s^2h, we solve the prior equation to get $s^2 = 288 - 2h^2$. Thus, we end up with $V(h) = (288 - 2h^2)h = 288h - 2h^3$.

We differentiate to get $V'(h) = -6h^2 + 288$, and the critical points are $h = \pm 4\sqrt{3}$. However, we discard $h = -4\sqrt{3}$ since length cannot be negative.

We confirm that $h = 4\sqrt{3}$ is a global maximum by substituting it into the second derivative (we have $V''(h) = -12h$, so $V''(4\sqrt{3}) < 0$). Thus, the maximum volume of the box is $V(4\sqrt{3}) = \boxed{768\sqrt{3}}$. □

Problem 4.109. Consider a circle with a sector cut out, at angle θ. Let this resulting figure have a fixed area A. What values of r and θ will minimize the perimeter of this figure?



Solution. We are given that the area A is constant, so the area formula $A = \frac{2\pi - \theta}{2\pi} \cdot \pi r^2 = \frac{2\pi - \theta}{2} r^2$ establishes a relation between r and θ. Since we wish to maximize the perimeter, which is $2r + 2\pi r - r\theta$, we need to find the expression for the perimeter in terms of one variable only, so we could apply our usual optimization techniques.

Note that $A = \frac{2\pi - \theta}{2} r^2$ rearranges to $\theta = 2\pi - \frac{2A}{r^2}$. We substitute this back into the initial expression for the perimeter to get $P(r) = 2r + 2\pi r - r\left(2\pi - \frac{2A}{r^2}\right) =$

$r\left(2+\dfrac{2A}{r^2}\right)=2r+\dfrac{2A}{r}$.

We can then evaluate $P'(r)=2-\dfrac{2A}{r^2}$, from which we get the critical points $r=\pm\sqrt{A}$. We discard the solution $r=-\sqrt{A}$ as length cannot be negative, so we are left with only $r=\sqrt{A}$. We get $P''(r)=\dfrac{4A}{r^3}$, so $P''(\sqrt{A})=\dfrac{4}{\sqrt{A}}>0$, thus we confirm that $\boxed{r=\sqrt{A}}$ and therefore $\boxed{\theta=2\pi-2}$ yield the minimum perimeter. □

Problem 4.110. Show that if the point (a,b) on the parabola $y=x^2$ is the closest point to $(0,c)$ (given $c>0$), then the line connecting (a,b) to $(0,c)$ is perpendicular to the tangent at (a,b).

Solution. The distance from (a,b) to $(0,c)$ is $\sqrt{a^2+(c-b)^2}$. We wish to minimize this distance, but it is sufficient to minimize the expression inside the radical, as this simplifies our work tremendously when differentiating. Note that $b=a^2$, as the point lies on the parabola $y=x^2$.

Consider $D(a)=a^2+(c-a^2)^2$, so $D'(a)=2a+2(c-a^2)(-2a)=2a(1-2c+2a^2)$. The critical points are $a=0,\pm\sqrt{\dfrac{2c-1}{2}}$. However, if $c\le\dfrac{1}{2}$, then $a=0$ would be the only critical point.

Assume $c>\dfrac{1}{2}$. Then $D''(a)=2-4c+12a^2$, so $D''\left(\sqrt{\dfrac{2c-1}{2}}\right)=8c-4>0$, so $a=\sqrt{\dfrac{2c-1}{2}}$ is a local minimum. The same follows for $a=-\sqrt{\dfrac{2c-1}{2}}$.

Thus, the two equally closest points are $\left(\sqrt{\dfrac{2c-1}{2}},\dfrac{2c-1}{2}\right)$ and $\left(-\sqrt{\dfrac{2c-1}{2}},\dfrac{2c-1}{2}\right)$. The slope of the line connecting $(0,c)$ and $\left(\sqrt{\dfrac{2c-1}{2}},\dfrac{2c-1}{2}\right)$ is $\dfrac{\frac{2c-1}{2}-c}{\sqrt{\frac{2c-1}{2}}}=-\dfrac{1}{2\sqrt{\frac{2c-1}{2}}}$.

Note that the derivative of $y=x^2$ is $y'=2x$, so the slope of the tangent line through $\left(\sqrt{\dfrac{2c-1}{2}},\dfrac{2c-1}{2}\right)$ is $2\sqrt{\dfrac{2c-1}{2}}$, and we confirm that these slopes are negative reciprocals of each other, so the lines are perpendicular. The same reasoning can be applied to $\left(-\sqrt{\dfrac{2c-1}{2}},\dfrac{2c-1}{2}\right)$ as well. □

Problem 4.111. Maximize and minimize $a\sin x+b\cos x$ using calculus. Assume $a,b\neq 0$.

Solution. Let $f(x)=a\sin x+b\cos x$. Then $f'(x)=a\cos x-b\sin x$, so x is a critical point only when $a\cos x=b\sin x$, or $\tan x=\dfrac{a}{b}$. This indicates that $\sin x=\dfrac{a}{\sqrt{a^2+b^2}}$ and $\cos x=\dfrac{b}{\sqrt{a^2+b^2}}$, or $\sin x=-\dfrac{a}{\sqrt{a^2+b^2}}$ and $\cos x=-\dfrac{b}{\sqrt{a^2+b^2}}$.

When we substitute the former values into the second derivative, $f''(x)=-a\sin x-b\cos x$, we get $-a\cdot\dfrac{a}{\sqrt{a^2+b^2}}-b\cdot\dfrac{b}{\sqrt{a^2+b^2}}=\dfrac{-(a^2+b^2)}{\sqrt{a^2+b^2}}=-\sqrt{a^2+b^2}<0$, so there is a local maximum at the value of x that satisfies $\sin x=\dfrac{a}{\sqrt{a^2+b^2}}$ and $\cos x=\dfrac{b}{\sqrt{a^2+b^2}}$.

However, notice that $f(x)=-f''(x)$, so we can immediately conclude that $-(-\sqrt{a^2+b^2})=\sqrt{a^2+b^2}$ is the maximum value.

Substituting in the latter values eventually leads to the similar conclusion that $-\sqrt{a^2+b^2}$ is the minimum value. Note that in solving this problem we did not have to explicitly solve for the critical points to be able to plug them into the second derivative. $\square$

§4.5 L'Hôpital's Rule

First, we introduce a precursory theorem to aid us in proving the main formula for this section.

Theorem 4.112 (Cauchy Mean Value Theorem)

Suppose f, g are continuous on $[a, b]$ and differentiable on (a, b). Then $\exists a < x < b$ such that $(f(b)-f(a))g'(x) = (g(b)-g(a))f'(x)$, or

$$\frac{g'(x)}{f'(x)} = \frac{g(b)-g(a)}{f(b)-f(a)}.$$

Proof. Let $h(x) = (f(b)-f(a))g(x) - (g(b)-g(a))f(x)$. Note that $h(a) = (f(b)-f(a))g(a) - (g(b)-g(a))f(a) = f(b)g(a) - f(a)g(b)$, and $h(b) = (f(b)-f(a))g(b) - (g(b)-g(a))f(b) = f(b)g(a) - f(a)g(b)$, so $h(a) = h(b)$. Therefore, by Theorem 4.82, $\exists a < x < b$ such that $h'(x) = 0$, i.e. $(f(b)-f(a))g'(x) - (g(b)-g(a))f'(x) = 0$, which rearranges to our desired result. $\square$

Theorem 4.113 (L'Hôpital's Rule)

Suppose $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$ and that $\lim\limits_{x\to a} \frac{g'(x)}{f'(x)}$ exists. Then $\lim\limits_{x\to a} \frac{g(x)}{f(x)} = \lim\limits_{x\to a} \frac{g'(x)}{f'(x)}$.

Proof. Assuming that f and g are differentiable, note that they must also be continuous, by Theorem 4.20. Thus, $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$ indicates that $f(a) = g(a) = 0$.

For some x near a, by Theorem 4.112, $\exists b$ between a and x such that

$$\frac{g'(b)}{f'(b)} = \frac{g(x)-g(a)}{f(x)-f(a)} = \frac{g(x)}{f(x)}.$$

Then $\lim\limits_{x\to a} \frac{g(x)}{f(x)} = \lim\limits_{x\to a} \frac{g'(b)}{f'(b)}$. However, note that as x approaches a, b will also approach a since b is between x and a. Thus, $\lim\limits_{x\to a} \frac{g(x)}{f(x)} = \lim\limits_{x\to a} \frac{g'(b)}{f'(b)} = \frac{g'(a)}{f'(a)} = \lim\limits_{x\to a} \frac{g'(x)}{f'(x)}$, as desired. $\square$

Problem 4.114. Evaluate $\lim\limits_{x\to 2} \frac{x^2-3x+2}{x^2-5x+6}$ using Theorem 4.113.

Solution. When we plug in $x = 2$, the numerator and denominator both become 0, so we can apply Theorem 4.113 to get $\lim\limits_{x\to 2} \frac{x^2-3x+2}{x^2-5x+6} = \lim\limits_{x\to 2} \frac{2x-3}{2x-5} = \boxed{-1}$. $\square$

Exercise 4.115. Evaluate $\lim\limits_{x\to 0} \frac{\sin x}{x}$ and $\lim\limits_{x\to 0} \frac{\sin 3x}{\sin 5x}$ using Theorem 4.113.

Problem 4.116. Evaluate $\lim_{x\to 0} \frac{1-\cos 3x}{1-\cos 5x}$ using Theorem 4.113.

Solution. First, we have $\lim_{x\to 0} \frac{1-\cos 3x}{1-\cos 5x} = \lim_{x\to 0} \frac{3\sin 3x}{5\sin 5x}$, but this still evaluates to 0 when $x=0$ is plugged in. Therefore, we apply Theorem 4.113 again: $\lim_{x\to 0} \frac{3\sin 3x}{5\sin 5x} = \lim_{x\to 0} \frac{9\cos 3x}{25\cos 5x} = \boxed{\frac{9}{25}}$. □

Problem 4.117. Evaluate $\lim_{x\to 2} \frac{x^2-5x+6}{x^2-4x+4}$ using Theorem 4.113.

Solution. We have $\lim_{x\to 2} \frac{x^2-5x+6}{x^2-4x+4} = \lim_{x\to 2} \frac{2x-5}{2x-4}$, but this limit does not exist. Therefore $\lim_{x\to 2} \frac{x^2-5x+6}{x^2-4x+4}$ does not exist. □

Problem 4.118. Compute $\lim_{x\to 0} \frac{3x-\sin 3x}{x^3}$.

Solution. We repeatedly apply Theorem 4.113:

$$\begin{aligned}
\lim_{x\to 0} \frac{3x-\sin 3x}{x^3} &= \lim_{x\to 0} \frac{3-3\cos 3x}{3x^2} \\
&= \lim_{x\to 0} \frac{1-\cos 3x}{x^2} \\
&= \lim_{x\to 0} \frac{3\sin 3x}{2x} \\
&= \frac{3}{2} \lim_{x\to 0} \frac{\sin 3x}{x} \\
&= \frac{3}{2} \lim_{x\to 0} 3\cos 3x \\
&= \boxed{\frac{9}{2}}.
\end{aligned}$$

□

Remark. Theorem 4.113 extends to $x\to\pm\infty$, $x\to a^+$, and $x\to a^-$.

Theorem 4.119

If $\lim_{x\to a} \frac{f'(x)}{g'(x)} = \pm\infty$, then $\lim_{x\to a} \frac{f(x)}{g(x)} = \pm\infty$. The same holds for right-hand and left-hand limits. In other words, Theorem 4.113 extends to $\pm\frac{\infty}{\infty}$.

Proof. Suppose $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \infty$. Assuming that $\lim_{x\to a} \frac{f(x)}{g(x)}$ exists, we rearrange the fraction,

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{\frac{1}{g(x)}}{\frac{1}{f(x)}},$$

and now it is appropriate to apply Theorem 4.113, as the numerator and denominator now go to 0. Then we differentiate,

$$\lim_{x\to a} \frac{\frac{1}{g(x)}}{\frac{1}{f(x)}} = \lim_{x\to a} \frac{-\frac{g'(x)}{g(x)^2}}{-\frac{f'(x)}{f(x)^2}} = \lim_{x\to a} \frac{f(x)}{g(x)} \cdot \frac{f(x)}{g(x)} \cdot \frac{g'(x)}{f'(x)}.$$

Thus,

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f(x)}{g(x)}\cdot\lim_{x\to a}\frac{f(x)}{g(x)}\cdot\lim_{x\to a}\frac{g'(x)}{f'(x)},$$

so $\lim_{x\to a}\frac{f'(x)}{g'(x)}=\lim_{x\to a}\frac{f(x)}{g(x)}$, which is what we want. □

§4.6 Inverses

We begin the section with a warm-up.

Problem 4.120. Prove that if f is increasing, then f^{-1} is increasing.

Proof. Looking back to Definition 4.87, $\forall a<b\to f(a)<f(b)$. Then the contrapositive of this implication must be true, i.e. $f(b)\le f(a)\to b\le a$. Let $\widetilde{b}=f(b)$ and $\widetilde{a}=f(a)$. It follows that $f^{-1}(\widetilde{b})=b$ and $f^{-1}(\widetilde{a})=a$, so we have $\widetilde{b}\le\widetilde{a}\to f^{-1}(\widetilde{b})\le f^{-1}(\widetilde{a})$. However, note that f and f^{-1} are one-to-one (or else the inverse would not exist), so we can safely get rid of the equal signs to conclude $\widetilde{b}<\widetilde{a}\to f^{-1}(\widetilde{b})<f^{-1}(\widetilde{a})$. Thus, f^{-1} is increasing. □

Theorem 4.121

Suppose f is continuous and increasing on $[a,b]$. Then f^{-1} is continuous on $[f(a),f(b)]$.

The proof, which will involve the $\varepsilon-\delta$ definition of continuity, is left to the reader.

Now it makes sense to consider the derivative of inverses. We now prove the general formula:

Theorem 4.122 (Derivative of the Inverse)

Let f^{-1} denote the inverse function of f. Then,

$$(f^{-1})'(x)=\frac{1}{f'(f^{-1}(x))}.$$

Proof. Assume f is continuous and one-to-one, or else it would not have an inverse. Let $f^{-1}(x)=y$, or $x=f(y)$. Note that

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

When h is small and approaches 0, $f^{-1}(x+h)$ approaches $f^{-1}(x)=y$. So, let $f^{-1}(x+h)=y+\widetilde{h}$, for some small $\widetilde{h}$. This also suggests that $x+h=f(y+\widetilde{h})$. Thus,

$$\begin{aligned}\lim_{h\to 0}\frac{f^{-1}(x+h)-f^{-1}(x)}{h}&=\lim_{h\to 0}\frac{y+\widetilde{h}-y}{h}\\&=\lim_{h\to 0}\frac{\widetilde{h}}{x+h-x}\\&=\lim_{h\to 0}\frac{\widetilde{h}}{f(y+\widetilde{h})-f(y)},\end{aligned}$$

and note that as h goes to 0, $\widetilde{h}$ must also go to 0, as we keep in mind that $f^{-1}(x) = y$. Thus,

$$\lim_{h\to 0} \frac{\widetilde{h}}{f(y+\widetilde{h}) - f(y)} = \lim_{\widetilde{h}\to 0} \frac{\widetilde{h}}{f(y+\widetilde{h}) - f(y)} = \frac{1}{f'(y)} = \frac{1}{f'(f^{-1}(x))},$$

so $(f^{-1})'(x) = \dfrac{1}{f'(f^{-1}(x))}$. □

Exercise 4.123. Use Theorem 4.36 to derive this formula. Why can't this method be a rigorous proof for Theorem 4.122?

Example 4.124

Find $\dfrac{d}{dx}(\sin^{-1}(x))$, $\dfrac{d}{dx}(\cos^{-1}(x))$, and $\dfrac{d}{dx}(\tan^{-1}(x))$.

Solution. We can use three methods:

1. We can directly apply the formula from Theorem 4.122: $\dfrac{d}{dx}(\sin^{-1}(x)) = \dfrac{1}{\cos(\sin^{-1}(x))} = \boxed{\dfrac{1}{\sqrt{1-x^2}}}$.

2. Alternatively, let $y = \cos^{-1}(x)$, so $\cos y = x$. Then, we can implicitly differentiate,

$$\frac{d}{dx}(\cos y) = \frac{d}{dx}(x) \longleftrightarrow -\sin(y)\frac{dy}{dx} = 1,$$

so $\dfrac{dy}{dx} = -\dfrac{1}{\sin y} = -\dfrac{1}{\sin(\cos^{-1}(x))} = \boxed{-\dfrac{1}{\sqrt{1-x^2}}}$.

3. Lastly, consider the identity $x = \tan(\tan^{-1}(x))$. We implicitly differentiate to get

$$\frac{d}{dx}(x) = \frac{d}{dx}(\tan(\tan^{-1}(x))) \longleftrightarrow 1 = \sec^2(\tan^{-1}(x))\frac{d}{dx}(\tan^{-1}(x)),$$

so $\dfrac{d}{dx}(\tan^{-1}(x)) = \dfrac{1}{\sec^2(\tan^{-1}(x))} = \boxed{\dfrac{1}{x^2+1}}$. □

Exercise 4.125. Evaluate the derivatives of the rest of the inverse trigonometric functions.

Exercise 4.126. Use Theorem 4.122 to prove that $\forall n \in \mathbb{N}$, $\dfrac{d}{dx}(x^{\frac{1}{n}}) = \dfrac{1}{n}x^{\frac{1}{n}-1}$. Then, prove the general case: $\forall r \in \mathbb{Q}$, $\dfrac{d}{dx}(x^r) = rx^{r-1}$.

Exercise 4.127. Find $\dfrac{d}{dx}\left(\tan^{-1}\left(x^{\frac{7}{2}}\right)\right)$.

§4.7 Parametric and Polar Equations

Theorem 4.128 (Derivative of Parametric Equations)
Assume $f(t), g(t)$ are continuous, such that

$$x = f(t),$$
$$y = g(t).$$

If $f'(t) \neq 0$, then $\dfrac{dy}{dx} = \dfrac{g'(t)}{f'(t)}$.

Proof. Since the equations $x = f(t)$, $y = g(t)$ do not necessarily define a function (there could be loops or spirals), we need to impose some restrictions. For some t_0, y is a "function" of x locally near t_0.

We can restrict the parametric equation to a certain domain and range such that it is an actual function. In this area, let $y = h(x)$, so $y = g(t) = h(x) = h(f(t))$. Then we implicitly differentiate to get

$$g'(t) = h'(f(t))f'(t) \longleftrightarrow h'(x) = \frac{g'(t)}{f'(t)},$$

which is what we wanted. □

Exercise 4.129. Given the parametric equations $x = \cos\theta$, $y = \sin\theta$ (what is this graph?), find $\dfrac{dy}{dx}$.

Exercise 4.130. Given the parametric equations $x = r(\theta - \sin\theta)$, $y = r(1 - \cos\theta)$ (what is this graph?), find $\dfrac{dy}{dx}$.

Problem 4.131. How can we get the slope of $r = f(\theta)$ at some fixed $\theta = \theta_0$?

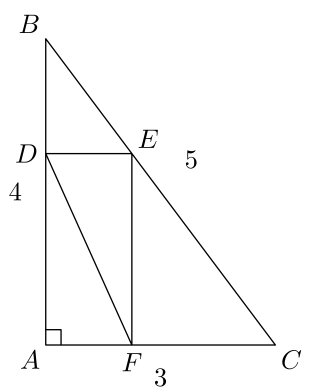
Solution. Note that $x = r\cos\theta$, $y = r\sin\theta$, so substitute $r = f(\theta)$ to get $x = f(\theta)\cos\theta$ and $y = f(\theta)\sin\theta$. Apply Theorem 4.128 to this, and we get

$$\frac{dy}{dx} = \frac{f'(\theta)\sin\theta + f(\theta)\cos\theta}{f'(\theta)\cos\theta - f(\theta)\sin\theta}.$$

Therefore the slope of the tangent line at $\theta = \theta_0$ is $\dfrac{f'(\theta_0)\sin\theta_0 + f(\theta_0)\cos\theta_0}{f'(\theta_0)\cos\theta_0 - f(\theta_0)\sin\theta_0}$. □

§4.8 Review

Problem 4.132. Consider $\triangle ABC$ where $AC = 3$, $AB = 4$, and $BC = 5$. Optimize (minimize or maximize) the perimeter of $\triangle DEF$.



Solution. Let $DE = x$. By AA similarity, $\triangle BDE \sim \triangle BAC$, so $\dfrac{BD}{4} = \dfrac{x}{3}$, which rearranges to $BD = \dfrac{4}{3}x$. Therefore $DA = EF = 4 - \dfrac{4}{3}x$, and $DF = \sqrt{x^2 + \left(4 - \dfrac{4}{3}x\right)^2}$ by the Pythagorean Theorem. We want to optimize the perimeter, so consider the function,

$$f(x) = x + 4 - \frac{4}{3}x + \sqrt{x^2 + \left(4 - \frac{4}{3}x\right)^2}.$$

We differentiate to get

$$f'(x) = -\frac{1}{3} + \frac{\frac{25}{9}x - \frac{16}{3}}{\sqrt{\frac{25}{9}x^2 - \frac{32}{3}x + 16}},$$

and we set this equal to 0 to find the critical points, resulting in the quadratic $25x^2 - 96x + 90 = 0$. Therefore, the critical points are $x = \dfrac{48 - 3\sqrt{6}}{25}$ and $x = \dfrac{48 + 3\sqrt{6}}{25}$, and it is left to the reader to verify which yield local maximums or minimums. □

Exercise 4.133. Find $\dfrac{d}{dx}\left(\tan^{-1}\left(\tan^{-1}(x)\right)\right)$ and $\dfrac{d}{dx}\left(x \sin^{-1} x\right)$.

Problem 4.134. Let $f(x) = x^3 + x + 1$. Find $(f^{-1})'(31)$.

Solution. By Theorem 4.122, $(f^{-1})'(x) = \dfrac{1}{3(f^{-1}(x))^2 + 1}$. To find $(f^{-1})'(31)$, we need to find $f^{-1}(31)$ in order to use the formula. Note that 3 is the only real root that satisfies $x^3 + x + 1 = 31$, so $f^{-1}(31) = 3$, so $(f^{-1})'(31) = \dfrac{1}{3(f^{-1}(31))^2 + 1} = \dfrac{1}{3 \cdot 3^2 + 1} = \boxed{\dfrac{1}{28}}$. □

Problem 4.135. Find $\displaystyle\lim_{x \to \frac{\pi}{2}^-} \left(x - \frac{\pi}{2}\right) \tan x$, $\displaystyle\lim_{x \to 0} \frac{\sin^{-1} x}{\tan^{-1} x}$, and $\displaystyle\lim_{x \to 1} \frac{x^{100} - 1}{x^{99} - 1}$.

Solution. We apply Theorem 4.113.

1. For the first limit, rewrite $\tan x = \dfrac{\sin x}{\cos x}$ to get

$$\lim_{x \to \frac{\pi}{2}^-} \frac{\left(x - \frac{\pi}{2}\right) \sin x}{\cos x},$$

so it is now appropriate to differentiate the numerator and denominator to get

$$\lim_{x \to \frac{\pi}{2}^-} \frac{\sin x + (\cos x)\left(x - \frac{\pi}{2}\right)}{-\sin x} = \boxed{-1}.$$

2. Recalling that $\dfrac{d}{dx}(\sin^{-1}(x)) = \dfrac{1}{\sqrt{1 - x^2}}$ and $\dfrac{d}{dx}(\tan^{-1}(x)) = \dfrac{1}{x^2 + 1}$, we have

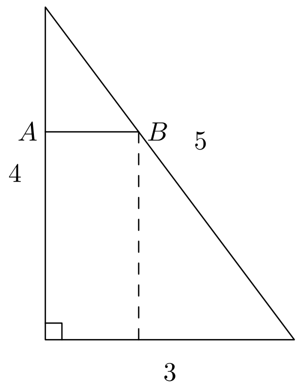
$$\lim_{x \to 0} \frac{\sin^{-1} x}{\tan^{-1} x} = \lim_{x \to 0} \frac{\frac{1}{\sqrt{1 - x^2}}}{\frac{1}{x^2 + 1}} = \lim_{x \to 0} \frac{x^2 + 1}{\sqrt{1 - x^2}} = \boxed{1}.$$

3. Likewise,

$$\lim_{x \to 1} \frac{x^{100} - 1}{x^{99} - 1} = \lim_{x \to 1} \frac{100x^{99}}{99x^{98}} = \boxed{\frac{100}{99}}.$$

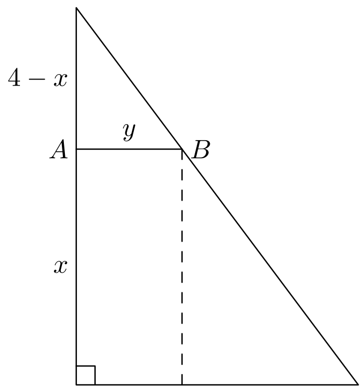
□

Problem 4.136. Consider a $3-4-5$ right triangle as shown.



The horizontal segment $\overline{AB}$ moves downward at a rate of $\frac{1}{3}$ unit per year. How fast is the length AB changing when $AB = 2$ units?

Solution. Let $AB = y$, and the altitude from AB to base of side length 3 be x. We label the triangle as such:



Note that the small right triangle with side lengths $4-x$ and y is similar to the $3-4-5$ triangle by AA similarity. Thus, $\frac{y}{3} = \frac{4-x}{4}$, which rearranges to $4y + 3x = 12$.

Interpreting the problem, we are given $\frac{dx}{dt} = -\frac{1}{3}$, and we want to find $\frac{dy}{dt}$ when $y = 2$. We implicitly differentiate the equation above:

$$\frac{d}{dt}(4y + 3x) = \frac{d}{dt}(12) \longleftrightarrow 4\frac{dy}{dt} + 3\frac{dx}{dt} = 0,$$

and we substitute the given values to get $4\frac{dy}{dt} + 3\left(-\frac{1}{3}\right) = 0$ to get $\frac{dy}{dt} = \boxed{\frac{1}{4}}$ unit per year. Note that we did not even need to use $y = 2$, because the horizontal segment's rate of change is constant, given the relation $4y + 3x = 12$. □

Problem 4.137. A box with a square base and an open top has volume 1000 cubic inches. What dimensions will minimize its surface area?

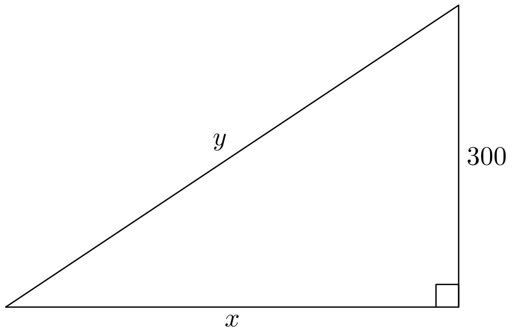
Solution. Let the side length of the square base be x, and therefore the height of the box will be $\frac{1000}{x^2}$. Then the surface area can be represented by the function

$$f(x) = x^2 + 4 \cdot \frac{1000}{x},$$

and we differentiate to get $f'(x) = 2x - \dfrac{4000}{x^2}$, yielding the critical point $x = 10\sqrt[3]{2}$. Since $f''(x) = 2 + \dfrac{8000}{x^3}$, $f''(10\sqrt[3]{2}) > 0$, so $x = 10\sqrt[3]{2}$ yields the minimum surface area. Thus, the dimensions of the box will be $\boxed{10\sqrt[3]{2} \times 10\sqrt[3]{2} \times 5\sqrt[3]{2}}$. □

Problem 4.138. Dan K flies a kite at the height of 300 feet. The wind carries the kite horizontally from Dan K at a rate of 25 feet per second. How fast must he let out the string when the kite is 500 feet away from him?

Solution. We interpret the problem using a right triangle.



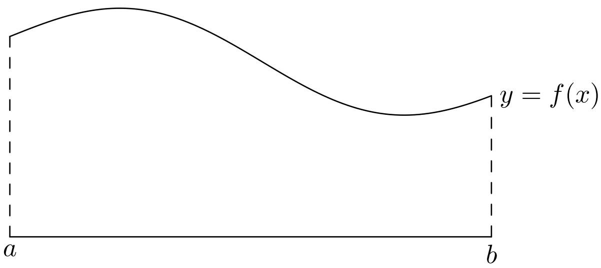
We are given $\dfrac{dx}{dt} = 25$, and we want to find $\dfrac{dy}{dt}$ when $y = 500$. By the Pythagorean Theorem, $300^2 + x^2 = y^2$, so we implicitly differentiate it to get $2x\dfrac{dx}{dt} = 2y\dfrac{dy}{dt}$. Then we substitute in values: $400 \cdot 25 = 500\dfrac{dy}{dt}$, i.e. $\dfrac{dy}{dt} = \boxed{20}$ feet per second. □

§5 Integrals

§5.1 Introduction

We seek to come up with a formal definition for the area between the x-axis and a function. For the concept of area to make sense, we will assume that all functions discussed in this chapter are *bounded*, unless stated otherwise.

Note that the diagrams will assume that the function in discussion will be positive, but you will realize that it is not necessarily positive, according to the way we choose to define concepts in this section.

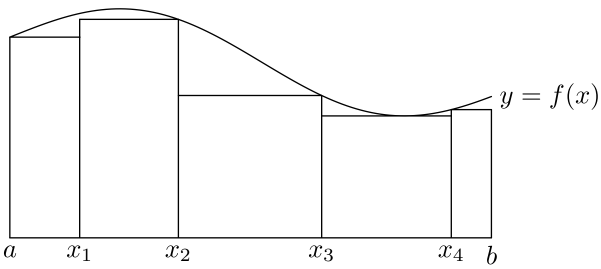


First, we introduce new definitions to formalize the concept of the area under the curve.

Definition 5.1. Given an interval $[a, b]$, a **partition** P is a set of points

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b.$$

By dividing the interval into smaller subintervals, we are essentially chopping up the region beneath $y = f(x)$ into *rectangles* that either leave remaining space in the intended region or cover the entire region plus some extra space. Here is an illustration of the former case:



How would we know where to extend the rectangles to? In this case, for a subinterval $[x_{i-1}, x_i]$, the height of the corresponding rectangle would be the minimum value of $f(x)$ on $[x_{i-1}, x_i]$.

We have assumed $f(x)$ to be bounded, but it is not necessarily continuous (although we have assumed it to be continuous in the diagram above). Therefore, it would not be appropriate to use min or max to define the height of the rectangles, but rather, inf and sup.

Thus, the height of a rectangle on $[x_{i-1}, x_i]$ in this case would be $\inf(\{f(x) \mid x_{i-1} \le x \le x_i\})$. Then, the area of an individual rectangle would be $(x_i - x_{i-1}) \cdot \inf(\{f(x) \mid x_{i-1} \le x \le x_i\})$. This is certainly an unwieldy expression, so we introduce convenient notation for brevity.

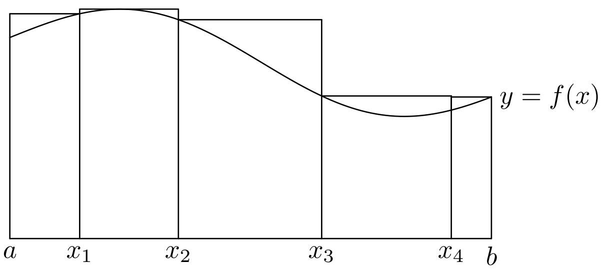
Definition 5.2. On a subinterval $[x_{i-1}, x_i]$, let $m_i = \inf(\{f(x) \mid x_{i-1} \le x \le x_i\})$.

Then the area of a particular rectangle in this diagram would be $m_i(x_i - x_{i-1})$. Note that all of these rectangles are below the graph, leading to our expression for the sum of all the areas of the rectangles:

Definition 5.3. For a function f and partition P of n points, we define the **lower sum**, or the sum of the areas of all rectangles below the graph, as

$$L(f, P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}).$$

Now consider the other case, where the rectangles cover the entire region in addition to some extra space:



Now, the height of a rectangle on $[x_{i-1}, x_i]$ is the maximum value of $f(x)$ on that subinterval. Thus, we establish analogous definitions.

Definition 5.4. On a subinterval $[x_{i-1}, x_i]$, let $M_i = \sup\left(\{f(x) \mid x_{i-1} \le x \le x_i\}\right)$.

Likewise, we consider the sum of all areas of the rectangles, which are above the graph in this case.

Definition 5.5. For a function f and partition P of n points, we define the **upper sum**, or the sum of the areas of all rectangles above the graph, as

$$U(f, P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}).$$

Before moving on to discussing the purpose of defining these sums, we first state and prove some expected results related to these.

The following lemma is a symbolic representation of an intuitive idea: if we chop up the region into more rectangles, the lower sum is greater than before (so it is approaching the actual area from below) and the upper sum is lesser than before (so it is approaching the actual area from above).

Lemma 5.6

Suppose $P \subseteq P'$, where P and P' are partitions. Then $L(f, P) \le L(f, P')$ and $U(f, P) \ge U(f, P')$.

Proof. First, let $[a, b]$ be the interval for the partitions P and P'. We will prove $L(f, P) \le L(f, P')$ by induction, and the proof will be analogous for upper sums.

To start, we will show that the inequality holds for P' containing exactly one more point than P.

Consider $x^\star \in [a,b]$, where $x^\star \notin P$ and let $P' = P \cup \{x^\star\}$. Note that we represent the partition P as

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b.$$

Suppose $x_{k-1} < x^\star < x_k$, such that P' can be represented as

$$a = x_0 < \ldots < x_{k-1} < x^\star < x_k < \ldots < x_n = b.$$

Then we have

$$\begin{aligned} L(f,P) &= \sum_{i=1}^{n} m_i(x_i - x_{i-1}), \\ L(f,P') &= \sum_{i=1}^{k-1} m_i(x_i - x_{i-1}) + m_k^\star(x^\star - x_{k-1}) + m_k^{\star\star}(x_k - x^\star) + \sum_{i=k+1}^{n} m_i(x_i - x_{i-1}), \end{aligned}$$

where we define

$$\begin{aligned} m_k^\star &= \inf\left(\{f(x) \mid x_{k-1} \le x \le x^\star\}\right), \\ m_k^{\star\star} &= \inf\left(\{f(x) \mid x^\star \le x \le x_k\}\right). \end{aligned}$$

It follows that $L(f,P')$ will share the same summands as $L(f,P)$, except that $m_k(x_k - x_{k-1})$ is replaced by $m_k^\star(x^\star - x_{k-1}) + m_k^{\star\star}(x_k - x^\star)$. Thus, to prove $L(f,P) \le L(f,P')$, it suffices to prove that

$$m_k(x_k - x_{k-1}) \le m_k^\star(x^\star - x_{k-1}) + m_k^{\star\star}(x_k - x^\star),$$

which can be rewritten as

$$m_k(x^\star - x_{k-1}) + m_k(x_k - x^\star) \le m_k^\star(x^\star - x_{k-1}) + m_k^{\star\star}(x_k - x^\star).$$

But note that $[x_{k-1}, x^\star] \subset [x_{k-1}, x_k]$ and $[x^\star, x_k] \subset [x_{k-1}, x_k]$, so we conclude that $m_k \le m_k^\star$ and $m_k \le m_k^{\star\star}$, and the inequality is clearly true as a result.

We have shown that $L(f,P) \le L(f,P')$ holds for P' containing exactly one more point than P. However, for a general $P' \supseteq P$, we know that P and P' are finite collections of points, so we can simply inductively add points to P until the partition becomes P' (since the inequality will always hold in this manner). □

Theorem 5.7

For any partitions P and Q, $L(f,P) \le U(f,Q)$.

Proof. Note that $P \subset P \cup Q$, so by Lemma 5.6, $L(f,P) \le L(f,P \cup Q)$. Obviously, $L(f,P\cup Q) \le U(f,P\cup Q)$ since the supremum of $f(x)$ on a subinterval is always greater than or equal to the infimum of $f(x)$ on that same subinterval. Lastly, we apply Lemma 5.6 again to get $U(f,P\cup Q) \le U(f,Q)$. We can now string these inequalities together to conclude that $L(f,P) \le U(f,Q)$. □

This theorem establishes that the set of all lower sums is bounded above by any upper sum, and vice-versa. Thus, by Axiom 3.11, $\sup\left(\{L(f,P) \mid P \text{ is a partition of } [a,b]\}\right)$ and $\inf\left(\{U(f,P) \mid P \text{ is a partition of } [a,b]\}\right)$ exist.

If the area, A, of the region between $y = f(x)$ and the x-axis is to be meaningful, we should have

$$\forall \text{ partitions } P,\ L(f, P) \leq A \leq U(f, P).$$

Furthermore, there should be *only one value* of A for which this is true (it would be unreasonable to have multiple possible values represent the area!).

Note that

$$\forall \text{ partitions } P,\ L(f, P) \leq \sup\left(\{L(f, P)\}\right) \leq \inf\left(\{U(f, P)\}\right) \leq U(f, P),$$

so it would only be reasonable to have $A = \sup\left(\{L(f, P)\}\right) = \inf\left(\{U(f, P)\}\right)$. This leads us into our much-awaited definition:

Definition 5.8. The function f is **integrable** on $[a, b]$ if

$$\sup\left(\{L(f, P) \mid P \text{ is a partition of } [a, b]\}\right) = \inf\left(\{U(f, P) \mid P \text{ is a partition of } [a, b]\}\right).$$

If f is integrable, then we define the **integral** of f on $[a, b]$ to be equal to this common value, and denote it as

$$\int_a^b f.$$

Problem 5.9. Define $f(x) = 1$. Is f integrable on $[0, 1]$? What about

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}?$$

Solution. For the first function, no matter how we choose the partition of points, we will always have $L(f, P) = U(f, P) = 1$. Thus, we can say that the area of the region between f and the x-axis is 1, i.e.

$$\int_0^1 f = 1.$$

This is consistent with our previous knowledge on the area of the unit square, which is 1.

Now consider the second function. For some arbitrary partition P, consider the subinterval $[x_{i-1}, x_i]$. By previous results, we know there exists an irrational number in $[x_{i-1}, x_i]$, so $f(x)$ will be 0 somewhere on that interval. Thus, m_i is automatically 0, since the function can only either be 0 or 1.

Likewise, we know that there is a rational number in $[x_{i-1}, x_i]$, so M_i is automatically 1. Thus, $L(f, P) = 0$ and $U(f, P) = 1$. However, there is no *unique* number between 0 or 1, and it follows that f is not integrable. $\square$

Exercise 5.10. Let $f(x) = c$ for some $c \in \mathbb{R}$. Show that $\displaystyle\int_0^1 f$ exists, and find its value.

Now recall the result of Theorem 3.18. This immediately leads to another condition for integrability:

Theorem 5.11

f is integrable on $[a, b] \longleftrightarrow$

$$\forall \varepsilon > 0\ \exists P, Q \text{ s.t. } U(f, Q) - L(f, P) < \varepsilon.$$

We can now use this result to prove a slightly stronger theorem:

Theorem 5.12

f is integrable on $[a, b] \longleftrightarrow$

$$\forall \varepsilon > 0 \; \exists P \text{ s.t. } U(f, P) - L(f, P) < \varepsilon.$$

Proof. The left direction is simple, by Theorem 5.11 (let $Q = P$). Now we prove the right direction. Given $\varepsilon > 0$, by Theorem 5.11, there exist R and S such that $U(f, S) - L(f, R) < \varepsilon$. Let $P = R \cup S$. We have

$$U(f, P) \leq U(f, S),$$
$$L(f, P) \geq L(f, R).$$

Thus, $U(f, P) - L(f, P) \leq U(f, S) - L(f, R) < \varepsilon$, and we are done. $\square$

We will also prove some lemmas about supremums and infimums that will allow us to calculate integrals more easily.

Lemma 5.13

If A, B are sets with $A \leq B$ and $\widetilde{A} \subseteq A$ and $\widetilde{B} \subseteq B$ satisfy $\sup(\widetilde{A}) = \inf(\widetilde{B})$, then $\sup A, \inf B$ are both equal to this common value.

Proof. Clearly $\sup A \geq \sup \widetilde{A}$ and $\inf B \leq \inf \widetilde{B}$. But as $A \leq B$, we have $\sup A \leq \inf B$. Thus

$$\sup \widetilde{A} \leq \sup A \leq \inf B \leq \inf \widetilde{B},$$

but as the extremes are equal, it follows that everything in between must equal the common value. $\square$

Lemma 5.14

If $\sup(A) = \inf(B)$ and $a < k < b \; \forall a \in A, \; b \in B$, then $\sup(A) = \inf(B) = k$.

Proof. Let $M = \sup(A) = \inf(B)$. Given that $a < k < b \; \forall a \in A, \; b \in B$, we conclude that A is bounded above by k, so $M \leq k$ since M is the least upper bound of A. Furthermore, B is bounded below by k, so $M \geq k$ as M is the greatest lower bound of B. Thus, $M = k$, which concludes the proof. $\square$

We apply these lemmas to calculate the area under a parabola.

Example 5.15

For $a > 0$, prove $\displaystyle\int_0^a x^2 = \frac{a^3}{3}$.

Proof. In order to approximate the area with lower and upper sums, it will be convenient to consider evenly spaced partitions. For $n \in \mathbb{N}$, define P_n such that

$$0 = x_0 < x_1 < x_2 < \ldots < x_n = a,$$

where $x_i = \dfrac{ia}{n}$, as we keep in mind that each subinterval is length $\dfrac{a}{n}$.

We also note that as x^2 is increasing,

$$\sup\{f(x) \mid x_{i-1} \leq x \leq x_i\} = x_i^2,$$
$$\inf\{f(x) \mid x_{i-1} \leq x \leq x_i\} = x_{i-1}^2.$$

If we define $f(x) = x^2$, then by definition,

$$\begin{aligned}
U(f, P_n) &= \sum_{i=1}^{n} x_i^2 (x_i - x_{i-1}) \\
&= \sum_{i=1}^{n} \left(\frac{ia}{n}\right)^2 \cdot \frac{a}{n} \\
&= \frac{a^3}{n^3} \sum_{i=1}^{n} i^2 \\
&= \frac{a^3}{n^3} \cdot \frac{n(n+1)(2n+1)}{6} \\
&= \frac{a^3}{3}\left(1 + \frac{1}{n}\right)\left(1 + \frac{1}{2n}\right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
L(f, P_n) &= \sum_{i=1}^{n} x_{i-1}^2 (x_i - x_{i-1}) \\
&= \sum_{i=1}^{n} \left(\frac{(i-1)a}{n}\right)^2 \cdot \frac{a}{n} \\
&= \frac{a^3}{n^3} \sum_{i=1}^{n} (i-1)^2 \\
&= \frac{a^3}{n^3} \cdot \frac{(n-1) \cdot n \cdot (2n-1)}{6} \\
&= \frac{a^3}{3}\left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{2n}\right).
\end{aligned}$$

First, note that $U(f, P_n) - L(f, P_n) = \frac{a^3}{n^3} \cdot n^2 = \frac{a^3}{n}$ which can be made arbitrarily small by taking a sufficiently large n. Thus, by Theorem 5.12, f is integrable on $[0, a]$.

Moreover, note that $L(f, P_n)$ and $U(f, P_n)$ can be made arbitrarily close to $\frac{a^3}{3}$ for a sufficient $n \in \mathbb{N}$ (the sequence of possible $L(f, P_n)$ is decreasing and the sequence of possible $U(f, P_n)$ is increasing, so the limit is clearly $\frac{a^3}{3}$). Thus, we have

$$L(f, P_n) = \frac{a^3}{3}\left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{2n}\right) < \frac{a^3}{3} < U(f, P_n) = \frac{a^3}{3}\left(1 + \frac{1}{n}\right)\left(1 + \frac{1}{2n}\right).$$

By Lemma 5.13, $\sup(\{L(f, P_n)\}) = \inf(\{U(f, P_n)\}) = \frac{a^3}{3}$, so the integral is equal to this value by the definition of the integral. $\square$

Exercise 5.16. For $a > 0$, prove $\int_0^a x = \frac{a^2}{2}$.

§5.2 Conditions for Integrability

In this section, we will show that nondecreasing and nonincreasing functions as well as continuous functions are always integrable. First, we will introduce a useful definition:

Definition 5.17. The **span** of f on $[a,b]$ is the quantity $\sup(\{f(x) \mid x \in [a,b]\}) - \inf(\{f(x) \mid x \in [a,b]\})$.

This definition is useful in proving integrability because it accounts for the differences between the upper and lower sums in the subintervals. This gives

$$U(f,P) - L(f,P) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \operatorname{span}(f \text{ on } [x_{i-1}, x_i]).$$

First, we begin with a theorem that will assist us in proving main results about integrals:

Theorem 5.18 (Smallest-Span Theorem)

If f is continuous on $[a,b]$, then $\forall \varepsilon > 0$, $\exists$ a partition $P : a = x_0 < x_1 < \ldots < x_n = b$ such that the span of f on each $[x_{i-1}, x_i] < \varepsilon$.

Proof. Given $\varepsilon > 0$, and call a partition *suitable* if the span of f on each subinterval is less than ε. Define $S = \{x \in [a,b] \mid [a,x] \text{ has a suitable partition}\}$. As f is continuous at $x = a$, we have

$$\exists \delta > 0,\ x \in (a-\delta, a+\delta) \to |f(x) - f(a)| < \frac{\varepsilon}{2}.$$

It follows that $a + \frac{\delta}{2} \in S$, and since S is bounded above by b, $\sup S$ exists, and let $c = \sup S$. We must have $c \leq b$.

First assume that $c < b$. By continuity at c, we can choose δ such that the span of f on $[c-\delta, c+\delta] < \varepsilon$. Moreover, as S contains every number in $[a,c)$, we have $c - \delta \in S$ and there exists a suitable partition from $\left[a, c - \frac{\delta}{2}\right]$. Gluing this interval with $\left[c - \frac{\delta}{2}, c + \frac{\delta}{2}\right]$ shows that all points in $\left[a, c + \frac{\delta}{2}\right]$ is also in S, contradicting that c is the supremum.

Thus, we must have $c = b$. Now using continuity at b shows that $b \in S$, so we conclude $S = [a,b]$, which proves the theorem. □

Theorem 5.19

If f is continuous on $[a,b]$, it is integrable on $[a,b]$.

Proof. The theorem easily follows from Theorem 5.18. Pick a partition P such that the span of f on each subinterval is less than $\frac{\varepsilon}{b-a}$. Then,

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \operatorname{span}(f \text{ on } [x_{i-1}, x_i]) \\ &< \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \frac{\varepsilon}{b-a} \end{aligned}$$

$$= (b-a) \cdot \frac{\varepsilon}{b-a} = \varepsilon,$$

proving that f is integrable. □

Theorem 5.20

If f is nondecreasing on $[a, b]$, f is integrable on $[a, b]$.

Proof. We again consider the evenly spaced partition P_n for $n \in \mathbb{N}$ such that

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b,$$

where the length of each subinterval is $\dfrac{b-a}{n}$ and thus $x_i = \dfrac{i(b-a)}{n} + a$.

Given that f is nondecreasing,

$$M_i = \sup\left(\{f(x) \mid x_{i-1} \le x \le x_i\}\right) = f(x_i),$$
$$m_i = \inf\left(\{f(x) \mid x_{i-1} \le x \le x_i\}\right) = f(x_{i-1}).$$

Then,

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{i=1}^{n} M_i(x_i - x_{i-1}) - \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \\ &= \sum_{i=1}^{n} \frac{b-a}{n} \cdot \operatorname{span}(f \text{ on } [x_{i-1}, x_i]) \\ &= \frac{b-a}{n}\left(f(b) - f(a)\right), \end{aligned}$$

and as this quantity can be made arbitrarily small by taking sufficiently large n, we conclude that f is integrable. □

Problem 5.21. Let $f(x) = \begin{cases} 1, & x \neq 1 \\ 2, & x = 1 \end{cases}$. Prove f is integrable on $[0, 2]$ and compute $\displaystyle\int_0^2 f$.

Solution. Let P be a partition with none of the $x_i = 1$, and

$$0 = x_0 < x_1 < x_2 < \ldots < x_n = 2,$$

and suppose $x_{k-1} < 1 < x_k$.

For the lower sum, note that m_i will always be equal to 1, so $\displaystyle\sum_{i=1}^{n} m_i(x_i - x_{i-1}) = \sum_{i=1}^{n}(x_i - x_{i-1})$, which is simply the sum of all the subintervals determined by the partition, i.e. 2, the length of the entire interval. Thus, $L(f, P) = 2$.

For the upper sum, note that M_i will always equal 1 except for the subinterval $[x_{k-1}, x_k]$, which contains 1, indicating that $f(x) = 2$ for some point in this subinterval. Thus, the supremum of $f(x)$ over $[x_{k-1}, x_k]$ is 2. Then, we can express the upper sum as

$$U(f,P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) = 2 + (x_k - x_{k-1}),$$

and so $U(f,P) - L(f,P) = x_k - x_{k-1}$ can be made arbitrarily small since we can always choose a partition such that the distance between x_{k-1} and x_k is less than a given $\varepsilon > 0$. Thus, f is integrable on $[0,2]$.

To calculate the integral, note that the lower sum will always equal 2 no matter what partition is chosen. Any upper sum will be equal to 2 plus some arbitrarily small number, so the infimum of the set of all upper sums is 2, thus the integral is 2. □

§5.3 Properties of Integrals

Using our intuitive notion of area, we will prove some useful facts about the integral. We first introduce some new notation:

Definition 5.22. Define $\displaystyle\int_a^b f(x)\,dx = \int_a^b f$.

The dx here cannot exist independently. Its meaning will become clear as we prove more results about integrals.

The following theorem says that we can shift integrals horizontally without changing the area:

Theorem 5.23

$$\int_a^b f(x)\,dx = \int_{a+c}^{b+c} f(x-c)\,dx.$$

Proof. Given a partition P on $[a,b]$ such that $a = x_0 < x_1 < x_2 < \ldots < x_n = b$, define P_c as the partition of $[a+c, b+c]$ such that $a+c = x_0+c < x_1+c < x_2+c < \ldots < x_n+c = b+c$.

We want to prove that $L(f(x), P) = L(f(x-c), P_c)$, and similarly for the upper sum. Indeed, if we make the substitution $x \to x - c$, each of the sets $\{f(x) \mid x_{i-1} \le x \le x_i\}$ is equivalent to $\{f(x-c) \mid x_{i-1} + c \le x \le x_i + c\}$, so the supremum and infimum of these are equal as well. The lengths of the subintervals are also unchanged (since they are merely shifted over), so both the lower and upper sums are equal. It follows that the integrals must also be equal, as desired. □

Exercise 5.24. Evaluate $\displaystyle\int_{-3}^{4} (x+3)^2\,dx$.

Furthermore, we are able to pull constants from the integral.

Theorem 5.25

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx.$$

Proof. We consider $c > 0$, $c < 0$ separately. Consider a partition P on $[a,b]$ such that $a = x_0 < x_1 < x_2 < \ldots < x_n = b$. Then note that if $A_i = \{f(x) \mid x_{i-1} \le x \le x_i\}$, then $\{cf(x) \mid x_{i-1} \le x \le x_i\} = cA_i$. We know that $\sup(cA) = c\sup(A)$ for any $c > 0$, so we have

$$U(cf, P) = \sum_{i=1}^{n} (x_i - x_{i-1}) \sup(cA_i) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot c \cdot \sup A_i = cU(f,P),$$

so if $B = \{L(f,P) \mid P \text{ partition}\}$, then $\{L(cf,P) \mid P\} = \{cL(f,P) \mid P\} = cB$. Following the same reasoning, we have $\inf(\{L(cf,P)\}) = c\inf(\{L(f,P)\})$. We obtain a similar result for upper sums, so the result follows for this case. For $c < 0$, we do the exact same thing but instead use $\sup(cA) = c\inf(A)$. □

We can also reflect the function over the y-axis.

Theorem 5.26

$$\int_a^b f(x)\,dx = \int_{-b}^{-a} f(-x)\,dx.$$

Proof. Given P, a partition of $[a, b]$, define P' as the partition of $[-b, -a]$, which results from multiplying every point in P by -1.

Since we are considering the transformation $f(x) \to f(-x)$, make the substitution $x \to -x$. It follows that each of the sets $\{f(-x) \mid x_{i-1} \le -x \le x_i\}$ i.e. $\{f(-x) \mid -x_i \le x \le -x_{i-1}\}$ is equivalent to $\{f(x) \mid x_{i-1} \le x \le x_i\}$, so the infimums and supremums of these are equal.

Furthermore, the lengths of the subintervals are also unchanged, since we are merely reflecting the subintervals across the y-axis.

Thus, we can conclude that $L(f(x), P) = L(f(-x), P')$ and $U(f(x), P) = U(f(-x), P')$, and the integrals are the same. □

Theorem 5.27

Let $a < b < c$. Then f is integrable on $[a, c] \longleftrightarrow f$ is integrable on $[a, b]$ and on $[b, c]$.

Proof. Let P, P' be partitions of $[a, b]$ and $[b, c]$ respectively. Then $P \cup P'$ is a partition of $[a, c]$. Thus, $L(f, P) + L(f, P') = L(f, P \cup P')$, as we are merely combining the subintervals. Similarly, $U(f, P) + U(f, P') = U(f, P \cup P')$.

Note that

$$\begin{aligned} U(f, P \cup P') - L(f, P \cup P') &= (U(f, P) + U(f, P')) - (L(f, P) + L(f, P')) \\ &= (U(f, P) - L(f, P)) + (U(f, P') - L(f, P')). \end{aligned}$$

To prove the left direction, apply Theorem 5.12 to get

$$\begin{aligned} \exists P, P' \text{ s.t. } U(f, P) - L(f, P) &< \frac{\varepsilon}{2}, \\ U(f, P') - L(f, P') &< \frac{\varepsilon}{2}, \end{aligned}$$

so thus $U(f, P \cup P') - L(f, P \cup P') < \varepsilon$, i.e. f is integrable on $[a, c]$.

For the right direction, let Q be a partition of $[a, c]$ containing b. Then let $P = Q \cap [a, c]$ and $P' = Q \cap [b, c]$. We are given that $U(f, P \cup P') - L(f, P \cup P') < \varepsilon$ for some $\varepsilon > 0$, by Theorem 5.12.

Clearly $U(f, P) \le U(f, P \cup P')$ and $L(f, P) \le L(f, P \cup P')$. We can string together the following inequalities,

$$\begin{aligned} U(f, P \cup P') - L(f, P \cup P') &\ge U(f, P) - L(f, P), \\ U(f, P \cup P') - L(f, P \cup P') &\ge U(f, P') - L(f, P'), \end{aligned}$$

and we can easily conclude that

$$U(f,P) - L(f,P) < \varepsilon,$$
$$U(f,P') - L(f,P') < \varepsilon. \qquad \square$$

Theorem 5.28

From Theorem 5.27, for $a < b < c$, $\displaystyle\int_a^b f + \int_b^c f = \int_a^c f$.

Proof. In the same style as the previous proof, declare P, P' to be partitions of $[a,b]$ and $[b,c]$ respectively, such that $P \cup P'$ is a partition of $[a,c]$. Moreover, $L(f,P) + L(f,P') = L(f, P \cup P')$.

By the definition of the integral, we have

$$L(f,P) \le \int_a^b f \le U(f,P),$$
$$L(f,P') \le \int_b^c f \le U(f,P').$$

We add these together and use the fact from above:

$$L(f,P) + L(f,P') \le \int_a^b f + \int_b^c f \le U(f,P) + U(f,P'),$$
$$L(f,P\cup P') \le \int_a^b f + \int_b^c f \le U(f,P\cup P').$$

However, recall that $\displaystyle\int_a^c f$ is the unique number between $L(f, P\cup P')$ and $U(f, P \cup P')$.

Thus, we conclude that $\displaystyle\int_a^b f + \int_b^c f = \int_a^c f$, as desired. $\square$

We introduce further definitions to develop the concept of the integral.

Definition 5.29. $\displaystyle\int_a^a f = 0$.

Definition 5.30. If $a < b$, then $\displaystyle\int_b^a f = -\int_a^b f$.

Corollary 5.31

For any choices a, b, c, $\displaystyle\int_a^b f + \int_b^c f = \int_a^c f$.

Proof. We proceed by casework. Consider all six permutations separately:

$$a < b < c,$$
$$a < c < b,$$
$$b < a < c,$$

$$
\begin{aligned}
b &< c < a, \\
c &< a < b, \\
c &< b < a.
\end{aligned}
$$

Each case should be relatively easy to deal with, using the stated definition above. □

We now have results for odd and even functions.

Theorem 5.32

If f is even, $\int_{-a}^{a} f = 2\int_{0}^{a} f$.

Proof. By Theorem 5.28,

$$\int_{-a}^{a} f = \int_{-a}^{0} f + \int_{0}^{a} f.$$

Note that $f(x) = f(-x)$, so by Theorem 5.26, $\int_{-a}^{0} f = \int_{0}^{a} f$. Thus, $\int_{-a}^{a} f = 2\int_{0}^{a} f$. □

Theorem 5.33

If f is odd, $\int_{-a}^{a} f = 0$.

Proof. Using Theorem 5.26 and the fact that $f(-x) = -f(x)$, we have

$$
\begin{aligned}
\int_{-a}^{a} f(x)\,dx &= \int_{-a}^{a} -f(-x)\,dx \\
&= -\int_{-a}^{a} f(-x)\,dx \\
&= -\int_{-a}^{a} f(x)\,dx.
\end{aligned}
$$

Thus $\int_{-a}^{a} f(x)\,dx = 0$, as desired. □

Now, we introduce our main theorem allowing us to take the sum of integrals.

Theorem 5.34

Given that f, g are integrable on $[a, b]$, $\int_{a}^{b} f + \int_{a}^{b} g = \int_{a}^{b} f + g$.

Proof. Let P be a partition of $[a, b]$ such that

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b.$$

On a subinterval $[x_{i-1}, x_i]$, let

$$
\begin{aligned}
m_i &= \inf\left(\{f(x) \mid x_{i-1} \le x \le x_i\}\right), \\
M_i &= \sup\left(\{f(x) \mid x_{i-1} \le x \le x_i\}\right).
\end{aligned}
$$

Similarly, we define m_i', M_i' for $g(x)$ and m_i'', M_i'' for $f(x)+g(x)$.

Note that $m_i + m_i'$ is a lower bound of $f(x)+g(x)$ on the subinterval, and m_i'' is the greatest lower bound of $f(x)+g(x)$ on the subinterval. Thus, we have $m_i'' \geq m_i + m_i'$, and similarly $M_i'' \leq M_i + M_i'$.

As a result,

$$\begin{aligned}
L(f+g,P) &\geq L(f,P)+L(g,P),\\
U(f+g,P) &\leq U(f,P)+U(g,P).
\end{aligned}$$

Thus,

$$U(f+g,P)-L(f+g,P) \leq (U(f,P)-L(f,P))+(U(g,P)-L(g,P)).$$

To show that $f+g$ is integrable on $[a,b]$, we apply Theorem 5.12 here:

$$\begin{aligned}
\exists P \text{ s.t. } U(f,P)-L(f,P) &< \frac{\varepsilon}{2},\\
U(g,P)-L(g,P) &< \frac{\varepsilon}{2}.
\end{aligned}$$

Thus, we have $U(f+g,P)-L(f+g,P)<\varepsilon$.

To prove the main result, we use the same argument in the proof of Theorem 5.28. First, we have

$$\begin{aligned}
L(f,P) \leq \int_a^b f \leq U(f,P),\\
L(g,P) \leq \int_a^b g \leq U(g,P).
\end{aligned}$$

Adding them together,

$$L(f,P)+L(g,P) \leq \int_a^b f + \int_a^b g \leq U(f,P)+U(g,P).$$

Note that $\int_a^b f + \int_a^b g$ must be the unique number between the two pairs of sums.

We have shown that $\int_a^b f+g$ exists, so we also have

$$L(f+g,P) \leq \int_a^b f+g \leq U(f+g,P).$$

Now we can apply all the inequalities discovered earlier to get

$$L(f,P)+L(g,P) \leq L(f+g,P) \leq \int_a^b f+g \leq U(f+g,P) \leq U(f,P)+U(g,P),$$

and by the uniqueness of integrals, we conclude $\int_a^b f + \int_a^b g = \int_a^b f+g$. □

Now we consider a horizontal stretch or shrink by a factor of c.

Theorem 5.35

For any constant $c \neq 0$,

$$c \int_a^b f(x)\,dx = \int_{ca}^{cb} f\left(\frac{x}{c}\right)\,dx.$$

Proof. First, assume $c > 0$. For some partition P on $[a, b]$, define P' as the partition that results from multiplying each of the points in P by c. We want to prove that $c \cdot L(f(x), P) = L\left(f\left(\frac{x}{c}\right), P'\right)$ and a similar result for upper sums.

Given a subinterval $[x_{i-1}, x_i]$, consider the set $\{f(x) \mid x_{i-1} \leq x \leq x_i\}$. Substitute $x \to \frac{x}{c}$, resulting in the set $\left\{f\left(\frac{x}{c}\right) \mid x_{i-1} \leq \frac{x}{c} \leq x_i\right\} = \left\{f\left(\frac{x}{c}\right) \mid cx_{i-1} \leq x \leq cx_i\right\}$. Note that these two sets are the same, so the supremums and infimums are the same.

Thus, the only thing that is ultimately changing is that the widths of the subintervals are multipled by c. It becomes clear that $c \cdot L(f(x), P) = L\left(f\left(\frac{x}{c}\right), P'\right)$ since the m_i terms are equal and c times the subintervals of P yields P' as that is what we defined P' to be. Similarly, $c \cdot U(f(x), P) = U\left(f\left(\frac{x}{c}\right), P'\right)$, so $c \int_a^b f(x)\,dx = \int_{ca}^{cb} f\left(\frac{x}{c}\right)\,dx$ for $c > 0$. □

Problem 5.36. Suppose $\int_0^1 x^n\,dx = c_n$, for $n \in \mathbb{N} \cup \{0\}$. By using an appropriate transformation, find $\int_0^a x^n\,dx$.

Solution. By Theorem 5.35, $a \int_0^1 x^n\,dx = \int_0^a \left(\frac{x}{a}\right)^n\,dx = \int_0^a \frac{x^n}{a^n}\,dx = \frac{1}{a^n}\int_0^a x^n\,dx$. Thus, $ac_n = \frac{1}{a^n}\int_0^a x^n\,dx$, from which we conclude that $\int_0^a x^n\,dx = a^{n+1}c_n$. □

Theorem 5.37

For $n \in \mathbb{N} \cup \{0\}$, $\int_0^1 x^n\,dx = \frac{1}{n+1}$.

Proof. We proceed by strong induction. Like Problem 5.36, let $\int_0^1 x^n\,dx = c_n$, for $n \in \mathbb{N} \cup \{0\}$.

By Problem 5.9, $c_0 = \int_0^1 x^0\,dx = \int_0^1 1\,dx = 1$. Then, our base case is done.

Assume $\int_0^1 x^k\,dx = c_k = \frac{1}{k+1}$, $\forall k \leq n$. We will prove that $c_n = \frac{1}{n+1}$. Consider

$$\int_0^{2a} x^n\,dx.$$

By Problem 5.36, we know that this equals $(2a)^{n+1}c_n = 2^{n+1}a^{n+1}c_n$.

However, we can use the Binomial Theorem to evaluate it in a different way. First, note that $\int_0^{2a} x^n\,dx = \int_{-a}^{a} (x+a)^n\,dx$ by Theorem 5.23. Then, the Binomial Theorem states that

$$\int_{-a}^{a} (x+a)^n \, dx = \int_{-a}^{a} \sum_{k=0}^{n} \binom{n}{k} a^{n-k} x^k \, dx = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} \int_{-a}^{a} x^k \, dx,$$

also using that we can pull out constants from integrals and that the integral of a sum can be broken up into the sums of integrals.

If we consider all odd $k \le n$, then those parts of the summation become 0, since $\int_{-a}^{a} x^k \, dx = 0$ in that case, by Theorem 5.33. Thus, we have

$$\sum_{k=0}^{n} \binom{n}{k} a^{n-k} \int_{-a}^{a} x^k \, dx = \sum_{k \text{ even}}^{n} \binom{n}{k} a^{n-k} \int_{-a}^{a} x^k \, dx = \sum_{k \text{ even}}^{n} \binom{n}{k} a^{n-k} \cdot 2 \cdot \int_{0}^{a} x^k \, dx,$$

with Theorem 5.32 justifying the latter step.

By Problem 5.36, we have $\int_{0}^{a} x^k \, dx = a^{k+1} c_k$. By assumption, $c_k = \frac{1}{k+1}$, so $a^{k+1} c_k = \frac{a^{k+1}}{k+1}$. Then,

$$\sum_{k \text{ even}}^{n} \binom{n}{k} a^{n-k} \cdot 2 \cdot \int_{0}^{a} x^k \, dx = \sum_{k \text{ even}}^{n} \binom{n}{k} a^{n-k} \cdot 2 \cdot \frac{a^{k+1}}{k+1} = 2a^{n+1} \sum_{k \text{ even}}^{n} \binom{n}{k} \frac{1}{k+1}.$$

Now, we note the following combinatorial identity:

$$\begin{aligned}
\binom{n}{k} \frac{1}{k+1} &= \frac{n!}{(n-k)!k!} \cdot \frac{1}{k+1} \\
&= \frac{n!}{(n-k)!(k+1)!} \\
&= \frac{(n+1)!}{(n-k)!(k+1)!} \cdot \frac{1}{n+1} \\
&= \binom{n+1}{k+1} \frac{1}{n+1}.
\end{aligned}$$

Hence,

$$2a^{n+1} \sum_{k \text{ even}}^{n} \binom{n}{k} \frac{1}{k+1} = \frac{2a^{n+1}}{n+1} \sum_{k \text{ even}}^{n} \binom{n+1}{k+1}.$$

By Pascal's Identity,

$$\frac{2a^{n+1}}{n+1} \sum_{k \text{ even}}^{n} \binom{n+1}{k+1} = \frac{2a^{n+1}}{n+1} \sum_{k \text{ even}}^{n} \left[\binom{n}{k} + \binom{n}{k+1} \right],$$

which is just

$$\frac{2a^{n+1}}{n+1} \sum_{k=0}^{n} \binom{n}{k} = \frac{2a^{n+1}}{n+1} 2^n = \frac{2^{n+1} a^{n+1}}{n+1}.$$

Remember that this expression is equal to $\int_{0}^{2a} x^n \, dx$. However, we have also previously mentioned that $\int_{0}^{2a} x^n \, dx = 2^{n+1} a^{n+1} c_n$ as well. Therefore, $2^{n+1} a^{n+1} c_n = \frac{2^{n+1} a^{n+1}}{n+1}$, which simplifies to $c_n = \frac{1}{n+1}$, proving our inductive step. This concludes the proof. $\square$

Corollary 5.38

$$\int_a^b x^n\,dx = \frac{b^{n+1}-a^{n+1}}{n+1}.$$

Exercise 5.39. Using previous results, prove Corollary 5.38. (Hint: use Corollary 5.31).

§5.4 Riemann Integrals

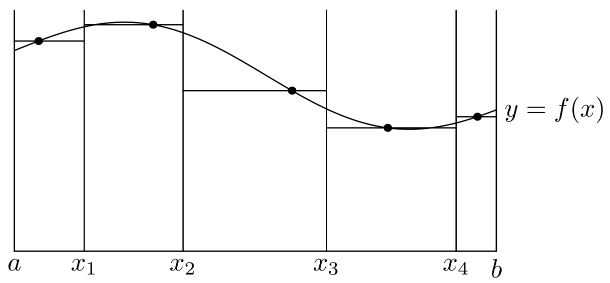
Up to now, we have assumed the integral to be defined using lower sums and upper sums. However, in order to compute some more complicated integrals, we will be looking at integrals through another perspective.

Definition 5.40. Let P be a partition of $[a,b]$ such that $a = x_0 < x_1 < x_2 < \ldots < x_n = b$. On each interval $[x_{i-1}, x_i]$, select a **sample point** $x_i^\star$ on this interval. Then

$$\sum_{i=1}^{n} f(x_i^\star)(x_i - x_{i-1})$$

is called a **Riemann sum**.

Here is an example:



In this graph, the dotted points are the selected sample points for each subinterval of the partition. Unlike lower or upper sums, we are not necessarily selecting the infimum or supremum of a set of the function's values over a subinterval. With Riemann sums, the sample points can be completely arbitrary.

It must follow that

$$L(f,P) \le \sum_{i=1}^{n} f(x_i^\star)(x_i - x_{i-1}) \le U(f,P),$$

for all partitions P, since the sample points are necessarily between the infimums and supremums of the subintervals. However, since we select the heights of the rectangles arbitrarily, we cannot know for sure whether a Riemann sum is less than or greater than the integral.

Ultimately, if we want this definition to be useful, the partitioned subintervals should be made very small in order to properly approximate the integral. This leads us into our next definition:

Definition 5.41. If P is a partition, we define the **mesh** of P as ΔP, which is the maximum of the lengths of the subintervals determined by the partition. Essentially, it is the biggest "hole" of the partition.

The following theorem states that a small mesh for a continuous function will result in a close approximation of the integral.

Theorem 5.42

Let f be continuous on $[a, b]$. Then, $\forall \varepsilon > 0$, $\exists \delta > 0$ such that for any partition $P: \ a = x_0 < x_1 < \ldots < x_n = b$ with $\Delta P < \delta$ and any choice of sample points, then

$$\left| \sum_{i=1}^{n} f(x_i^\star)(x_i - x_{i-1}) - \int_a^b f(x)\,dx \right| < \varepsilon.$$

Proof. Given $\varepsilon > 0$, by Theorem 5.18, let P be a partition on $[a, b]$, where $a = x_0 < x_1 < x_2 < \ldots < x_m = b$ for which the span of each subinterval is less than $\dfrac{\varepsilon}{2(b-a)}$.

Let δ be the length of the shortest interval in P.

Then define $\widetilde{P}: \ a = t_0 < t_1 < t_2 < \ldots < t_n = b$ for which $\Delta\widetilde{P} < \delta$. This ensures that all subintervals of $\widetilde{P}$ are smaller than the subintervals of P.

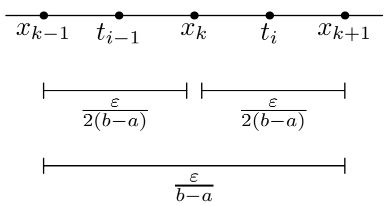
For each pair t_{i-1}, t_i, one of the following is true:

1. $\exists x_{k-1}, \ x_k$ such that $x_{k-1} \le t_{i-1} \le t_i \le x_k$. In other words, a subinterval of $\widetilde{P}$ completely fits inside a subinterval of P.

   Then, since $[t_{i-1}, t_i] \subseteq [x_{i-1}, x_i]$, the span of f on $[t_{i-1}, t_i]$ is also less than $\dfrac{\varepsilon}{2(b-a)}$.

2. $\exists x_{k-1}, \ x_k, \ x_{k+1}$ such that $x_{k-1} < t_{i-1} < x_k < t_i < x_{k+1}$.

   As $[t_{i-1}, t_i] \subset [x_{k-1}, x_{k+1}]$, the span of f on $[t_{i-1}, t_i]$ must be less than $2 \cdot \dfrac{\varepsilon}{2(b-a)} = \dfrac{\varepsilon}{b-a}$. Here is a visual diagram that explains this better:



Thus, the span of each subinterval on $\widetilde{P}$ is less than $\dfrac{\varepsilon}{b-a}$. We have

$$\begin{aligned} U(f, \widetilde{P}) - L(f, \widetilde{P}) &= \sum_{i=1}^{n} (t_i - t_{i-1}) \cdot \operatorname{span}(f \text{ on } [t_{i-1}, t_i]) \\ &< \sum_{i=1}^{n} (t_i - t_{i-1}) \cdot \frac{\varepsilon}{b-a} \\ &= (b-a) \cdot \frac{\varepsilon}{b-a} = \varepsilon. \end{aligned}$$

Since

$$L(f, \widetilde{P}) \le \sum_{i=1}^{n} f(t_i^\star)(t_i - t_{i-1}) \le U(f, \widetilde{P}),$$

and the integral is also between the lower and upper sums, we may conclude that the Riemann sum and the integral can be made arbitrarily close to each other.

Given $\varepsilon > 0$, we have found a $\delta > 0$ such that for any partition $\widetilde{P}$ with $\Delta\widetilde{P} < \delta$, we have $U(f, \widetilde{P}) - L(f, \widetilde{P}) < \varepsilon$ which thereby implies the result that

$$\left| \sum_{i=1}^{n} f(t_i^\star)(t_i - t_{i-1}) - \int_a^b f(x)\,dx \right| < \varepsilon,$$

using the fact that the Riemann sum is always between the lower and upper sums. ☐

Theorem 5.43

Let f be continuous on $[a, b]$. Let $P_n:\ a = x_0 < x_1 < \ldots < x_n = b$ be a partition with equal length subintervals. Choose sample points $x_{i-1} \le x_i^\star \le x_i$ uniformly in some fashion (i.e. all left end points, midpoints, etc.). Then,

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^\star)(x_i - x_{i-1}) = \int_a^b f(x)\, dx.$$

We won't be proving this theorem in this course, but this theorem should rigorously express the notion that making the partitioned subintervals very small will allow us to closely approximate the value of the actual integral.

In fact, this theorem highlights the advantage of using a Riemann integral over computing lower or upper sums; we can use our concept of limits to evaluate integrals that were not possible to do before.

Example 5.44

Evaluate $\displaystyle\int_0^a \sin x\, dx$.

Solution. Consider an evenly-spaced partition $P_n:\ 0 = x_0 < x_1 < \ldots < x_n = a$ and uniformly choose the sample points $x_i^\star$ as the right end-points. This means that for some subinterval $[x_{i-1}, x_i]$, we will choose $x_i^\star = x_i = \dfrac{ia}{n}$. Since each subinterval is length $\dfrac{a}{n}$, our Riemann integral is

$$\sum_{i=1}^{n} \sin\left(\frac{ia}{n}\right) \cdot \frac{a}{n} = \frac{a}{n} \sum_{i=1}^{n} \sin\left(\frac{ia}{n}\right).$$

For convenience, let $k = \dfrac{a}{n}$. Then, focusing on the summation, we have

$$\sum_{i=1}^{n} \sin\left(\frac{ia}{n}\right) = \sin k + \sin 2k + \sin 3k + \ldots + \sin nk.$$

Now, we apply a clever trick: multiply and divide all terms by $\sin\dfrac{k}{2}$:

$$\frac{1}{\sin\frac{k}{2}}\left[\sin k \sin\frac{k}{2} + \sin 2k \sin\frac{k}{2} + \sin 3k \sin\frac{k}{2} + \ldots + \sin nk \sin\frac{k}{2}\right].$$

Then, we apply the product-to-sum formula $\sin a \sin b = \dfrac{1}{2}(\cos(a-b) - \cos(a+b))$ to the terms inside the brackets:

$$\frac{1}{2\sin\frac{k}{2}}\left[\cos\frac{k}{2} - \cos\frac{3k}{2} + \cos\frac{3k}{2} - \cos\frac{5k}{2} + \ldots + \cos\left(nk - \frac{k}{2}\right) - \cos\left(nk + \frac{k}{2}\right)\right].$$

Then we see that the terms inside the brackets telescope, leaving us with:

$$\frac{1}{2\sin\frac{k}{2}}\left[\cos\frac{k}{2} - \cos\left(nk + \frac{k}{2}\right)\right].$$

Rewriting in the context of our initial Riemann integral, we have

$$\frac{a}{n}\sum_{i=1}^{n}\sin\left(\frac{ia}{n}\right) = \frac{a}{n}\left(\frac{1}{2\sin\frac{a}{2n}}\right)\left[\cos\frac{a}{2n} - \cos\left(a+\frac{a}{2n}\right)\right]$$
$$= \frac{\frac{a}{2n}}{\sin\frac{a}{2n}}\left[\cos\frac{a}{2n} - \cos\left(a+\frac{a}{2n}\right)\right].$$

As $\sin x$ is continuous, our partition is defined to be evenly spaced, and our sample points are uniformly the right end-points of their respective subintervals, we can apply Theorem 5.43 to get the desired integral:

$$\int_0^a \sin x\,dx = \lim_{n\to\infty}\sum_{i=1}^{n}\sin\left(\frac{ia}{n}\right)\cdot\frac{a}{n}$$
$$= \lim_{n\to\infty}\frac{\frac{a}{2n}}{\sin\frac{a}{2n}}\left[\cos\frac{a}{2n} - \cos\left(a+\frac{a}{2n}\right)\right].$$

As n goes to ∞, $\frac{a}{2n}$ must go to 0. Then, $\frac{\frac{a}{2n}}{\sin\frac{a}{2n}} \to 1$, $\cos\frac{a}{2n} \to 1$, and $\cos\left(a+\frac{a}{2n}\right) \to \cos a$, leaving us with:

$$\int_0^a \sin x\,dx = 1 - \cos a.$$ □

We can apply Corollary 5.31 to quickly obtain:

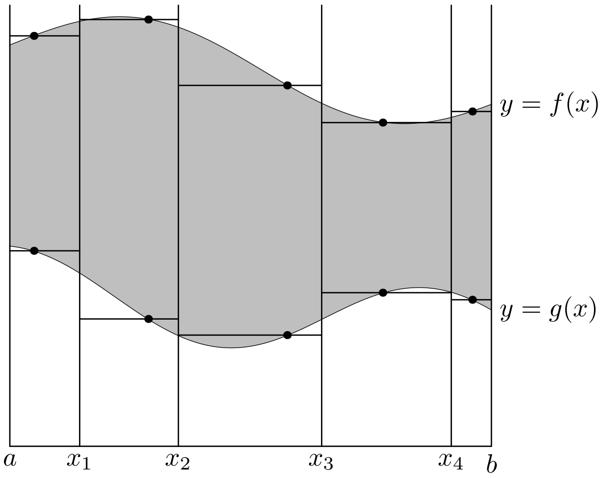
$$\int_a^b \sin x\,dx = \int_0^b \sin x\,dx - \int_0^a \sin x\,dx$$
$$= 1 - \cos b - (1 - \cos a)$$
$$= \cos a - \cos b.$$

Problem 5.45. Using this result, evaluate $\int_a^b \cos x\,dx$.

Solution. The key point is to use the fact that cosine and sine are phase shifts of each other.

$$\int_a^b \cos x\,dx = \int_a^b \sin\left(\frac{\pi}{2} - x\right)\,dx$$
$$= -\int_a^b \sin\left(x - \frac{\pi}{2}\right)\,dx$$
$$= -\int_{a-\frac{\pi}{2}}^{b-\frac{\pi}{2}} \sin x\,dx$$
$$= -\left[\cos\left(a-\frac{\pi}{2}\right) - \cos\left(b-\frac{\pi}{2}\right)\right]$$
$$= -(\sin a - \sin b)$$
$$= \sin b - \sin a.$$ □

Using the Riemann sum, we can find formulas to compute the areas of various shapes - not just the region between a function and the x-axis.



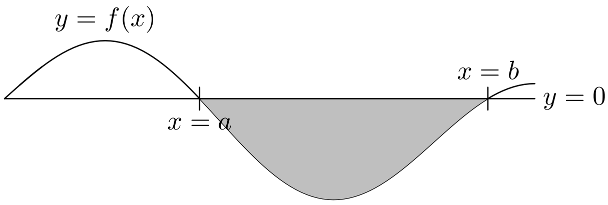
If we choose the same sample points $x_i^\star$ for $f(x)$ and $g(x)$, then the area of the region between $y = f(x)$ and $y = g(x)$ can be approximated by subtracting the area of the rectangles of $g(x)$ from those of $f(x)$, assuming that $f(x)$ is greater than $g(x)$ on this interval. Then, our Riemann sum for this region is

$$\sum_{i=1}^{n}(f(x_i^\star) - g(x_i^\star))(x_i - x_{i-1}),$$

so we can get the exact area of the region by taking the limit:

$$\lim_{n\to\infty}\sum_{i=1}^{n}(f(x_i^\star) - g(x_i^\star))(x_i - x_{i-1}) = \int_a^b (f(x) - g(x))\,dx.$$

We've never really addressed area that was below the x-axis, but this formula enables us to do so. Consider the area of some portion of a graph $y = f(x)$ that is negative, i.e. below the x-axis:

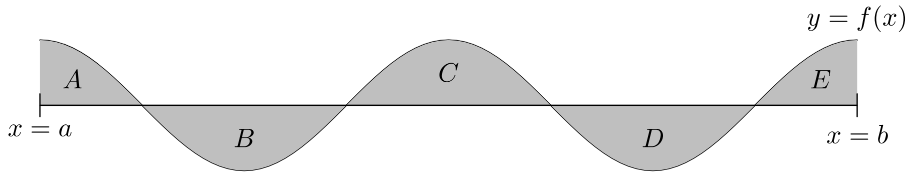


Note that the x-axis can also be written as a function, $y = 0$. Then the shaded area would be

$$\int_a^b (0 - f(x))\,dx = \int_a^b -f(x)\,dx = -\int_a^b f(x)\,dx.$$

Thus, since this area is beneath the x-axis, the numerical answer we get from the integral is also negative. This is a strange concept, since we'd never think of area as a negative quantity. However, in the discussion of integrals, the distinction must be noted.

What if we took the integral of $f(x)$ on some interval where the graph was both above and below the x-axis? Then, we'd have positive and negative areas to deal with, so the integral would be the sum of those quantities, accounting for the signs.

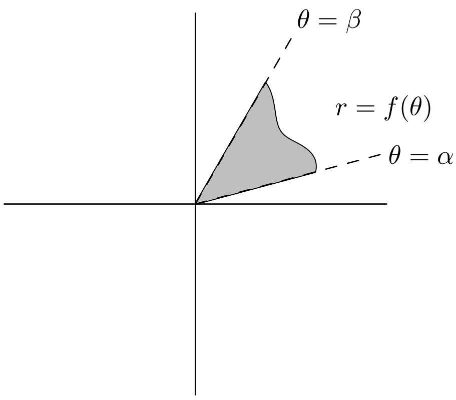


In the diagram above, we'd have

$$\int_a^b f(x)\,dx = A - B + C - D + E.$$

If we wanted to find the total area disregarding sign, then we would have to split up the whole integral into positive and negative integrals (based on which intervals $f(x)$ would be positive or negative in).

Now, what if we wanted to find the area for some part of a polar graph? Unlike regular functions, we must instead consider the area as a 'pizza slice' from the origin to the polar curve.



Here, the area is bounded by $\theta = \alpha$, $\theta = \beta$, and $r = f(\theta)$ from $\theta = \alpha$ to β.

Rather than the usual approach of dividing the region into many rectangles, we will use *sectors* to approximate. First, partition the interval $[\alpha, \beta]$ into subintervals

$$\alpha = \theta_0 < \theta_1 < \ldots < \theta_n = \beta.$$

For each subinterval $[\theta_{i-1}, \theta_i]$, choose a uniform sample point $\theta_i^\star$. Similar to how we set the height of the rectangle to be equal to the function evaluated at the sample point, we will establish the radius of the sector for this subinterval to be $f(\theta_i^\star)$. Then, the area will be

$$\frac{f(\theta_i^\star)^2 \cdot (\theta_i - \theta_{i-1})}{2}.$$

In the same way as we did with rectangles, the principles of the Riemann sum also apply for circular sectors. We evaluate the limit of the Riemann sum to obtain our formula:

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{f(\theta_i^\star)}{2}(\theta_i - \theta_{i-1}) = \int_\alpha^\beta \frac{1}{2} f(\theta)^2\,d\theta.$$

Example 5.46

Find the area of one petal of $r = \cos 2\theta$.

Solution. To find the expression we need to evaluate, we must first find the limits of integration, i.e. α and β in the given formula.

How is one petal formed? As we sketch the graph, notice that we must start from the origin, draw our petal, then end up back at the origin. It is helpful to consider the angles of θ when $r = 0$, since this is when we are at the origin. Solving $\cos 2\theta = 0$ gives us $\theta = \dfrac{\pi}{4} + 2\pi k,\ \dfrac{3\pi}{4} + 2\pi k$ for $k \in \mathbb{Z}$.

Choosing any two consecutive values will suffice (this won't always be the case; always try sketching the function first to see how it is drawn and in what order). Using previous results, we have

$$\begin{aligned}\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \frac{1}{2}\cos^2 2\theta \, d\theta &= \frac{1}{4}\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} 1 + \cos 4\theta \, d\theta \\ &= \frac{1}{4}\left[\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} 1 \, d\theta + \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \cos 4\theta \, d\theta\right] \\ &= \frac{1}{4}\left[\frac{\pi}{2} + \frac{1}{4}\int_{-\pi}^{\pi} \cos\theta \, d\theta\right] \\ &= \frac{1}{4}\left[\frac{\pi}{2} + \frac{1}{4}(\sin\pi - \sin(-\pi))\right] \\ &= \boxed{\frac{\pi}{8}}.\end{aligned}$$

□

Of course, discussion of polar functions also leads us to parametric functions. Let functions f, g be differentiable and f be one-to-one on an interval $[a, b]$.

Partition $a = t_0 < t_1 < \ldots < t_n = b$. Pick sample point $t_{i-1} \le t_i^\star \le t_i$.

The width of this rectangle would be $f(t_i) - f(t_{i-1})$ (since this is the difference in x-coordinates), while the height would be $g(t_i^\star)$. Thus, we have the sum of all the rectangles as

$$\sum_{i=1}^{n} g(t_i^\star)(f(t_i) - f(t_{i-1})).$$

But this is not quite a Riemann sum because of the $f(t_i) - f(t_{i-1})$ term. It would be if we replaced it with $t_i - t_{i-1}$. In fact, we can force this to happen:

$$\sum_{i=1}^{n} g(t_i^\star)(f(t_i) - f(t_{i-1})) = \sum_{i=1}^{n} g(t_i^\star)\frac{f(t_i) - f(t_{i-1})}{t_i - t_{i-1}} \cdot (t_i - t_{i-1}).$$

The $\dfrac{f(t_i) - f(t_{i-1})}{t_i - t_{i-1}}$ term motivates us to apply the Mean Value Theorem: there is a point $t_i^{\star\star}$ such that $f'(t_i^{\star\star}) = \dfrac{f(t_i) - f(t_{i-1})}{t_i - t_{i-1}}$. Then, we have

$$\sum_{i=1}^{n} g(t_i^\star) f'(t_i^{\star\star})(t_i - t_{i-1}).$$

We can take the limit of the mesh of the partitions going to 0, such that $t_i^\star$ and $t_i^{\star\star}$ approach the same number. Thus, we similarly obtain the formula

$$\int_a^b g(t)f'(t)\, dt,$$

as the area under the parametric curve on $[a, b]$, where $x = f(t)$ and $y = g(t)$.

Exercise 5.47. Find the area under the curve $x = \cos\theta$, $y = \sin\theta$ as $\theta = 0$ to π.

Now, what if we wanted to compute the *arc length* of some portion of a parametric function? Although this is not directly related to area, it turns out that the integral will be involved in the result, as a result of another Riemann sum.

Let $x = f(t)$, $y = g(t)$, where $a \le t \le b$. Partition the interval into $a = t_0 < t_1 < \ldots < t_n = b$. Then the curve is partitioned into points $(f(t_i), g(t_i))$ where $0 \le i \le n$. We will approximate the arc length by taking the sum of all the distances between consecutive points on the curve based on the partition. In other words, our approximated length would be

$$\sum_{i=1}^{n} \sqrt{(f(t_i) - f(t_{i-1}))^2 + (g(t_i) - g(t_{i-1}))^2},$$

using the distance formula.

Using the same trick as before, we have

$$\begin{aligned}
&\sum_{i=1}^{n} \sqrt{(f(t_i) - f(t_{i-1}))^2 + (g(t_i) - g(t_{i-1}))^2} \\
&= \sum_{i=1}^{n} \sqrt{\left(\frac{f(t_i) - f(t_{i-1})}{t_i - t_{i-1}}\right)^2 + \left(\frac{g(t_i) - g(t_{i-1})}{t_i - t_{i-1}}\right)^2} \cdot (t_i - t_{i-1}) \\
&= \sum_{i=1}^{n} \sqrt{f'(t_i^\star)^2 + g'(t_i^{\star\star})^2} \cdot (t_i - t_{i-1}),
\end{aligned}$$

using the Mean Value Theorem on two terms. Thus, we can take the limit of this approximation to obtain the exact arc length, represented by the formula

$$\int_a^b \sqrt{f'(t)^2 + g'(t)^2}\, dt.$$

From here, we can easily get the formula for the arc length of a regular function $y = f(x)$. We have the simple parametrization $x = t$, $y = f(t)$, and plugging this into our formula gives

$$\int_a^b \sqrt{1 + f'(x)^2}\, dx.$$

We can also address one last property about a function using Riemann sums: the average value over some interval. What does the average value of a function even mean? There's infinitely many values to consider if we proceeded with the obvious method of adding all values up and then dividing by the number of values.

Instead, we will consider the average value of *sample points* taken on a partition, then take the Riemann sum to infinity as usual, to obtain an answer in terms of an integral.

On an interval $[a, b]$, let $P_n :\ a = x_0 < x_1 < \ldots < x_n = b$ be an equal length partition. Choose sample points $x_{i-1} \le x_i^\star \le x_i$. The average value of the sample points would be

$$\frac{1}{n} \sum_{i=1}^{n} f(x_i^\star).$$

Again, we are encouraged to turn this into a Riemann sum, namely by forcing a $x_i - x_{i-1}$ term inside the summation. Here, we take advantage of the property that the partitions are equal length, i.e. $x_i - x_{i-1} = \dfrac{b-a}{n}$. We could then rewrite $\dfrac{1}{n}$ as $\dfrac{1}{b-a} \cdot \dfrac{b-a}{n}$, resulting in

$$\frac{1}{b-a}\sum_{i=1}^{n} f(x_i^\star) \cdot \frac{b-a}{n} = \frac{1}{b-a}\sum_{i=1}^{n} f(x_i^\star) \cdot (x_i - x_{i-1}).$$

From here, we can easily take the limit as $n \to \infty$ of this Riemann sum to get our formula,

$$\frac{1}{b-a}\int_a^b f(x)\,dx,$$

for the average value of a function $f(x)$ over the interval $[a, b]$.

There is a more intuitive way to think about this formula. Notice that $\displaystyle\int_a^b f(x)\,dx$ is the area of the region, and $b-a$ is the "base" of the region. If we divide the area by the base, then we get the height. Thus, the formula $\displaystyle\frac{1}{b-a}\int_a^b f(x)\,dx$ actually gives us the average "height" of the region.

Problem 5.48. Find the average value of $f(x) = x^2$ on $[0, 3]$.

Solution. We directly apply the formula:

$$\frac{1}{3-0}\int_0^3 x^2\,dx = \frac{1}{3} \cdot \frac{3^3}{3} = \boxed{3}.$$ □

Problem 5.49. Find the average value of $f(x) = \sin x$ on $[0, \pi]$.

Solution. Again, we have

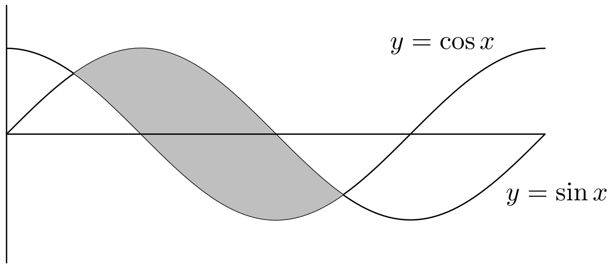
$$\frac{1}{\pi-0}\int_0^\pi \sin x\,dx = \frac{1}{\pi} \cdot (1 - \cos\pi) = \boxed{\frac{2}{\pi}}.$$ □

Toward the end of this section, we have discovered various formulas that result from the Riemann integral. Here is a summary of those formulas:

- The area between two functions $f(x)$ and $g(x)$ on $[a, b]$ is
$$\int_a^b (f(x) - g(x))\,dx.$$
- The area of a polar graph $r = f(\theta)$ from $\theta = \alpha$ to β is
$$\int_\alpha^\beta \frac{1}{2} f(\theta)^2\,d\theta.$$
- Given parametric equations $x = f(t)$, $y = g(t)$, the area under the parametric curve on $[a, b]$ is
$$\int_a^b g(t)f'(t)\,dt.$$
- Given parametric equations $x = f(t)$, $y = g(t)$, the arc length of the parametric curve on $[a, b]$ is
$$\int_a^b \sqrt{f'(t)^2 + g'(t)^2}\,dt.$$

- The arc length of a function $y = f(x)$ on $[a, b]$ is
$$\int_a^b \sqrt{1 + f'(x)^2}\, dx.$$
- The average value of $f(x)$ on $x = a$ to b is
$$\frac{1}{b-a}\int_a^b f(x)\, dx.$$

Problem 5.50. Find the integrals that represent the area and perimeter of the shaded region below:

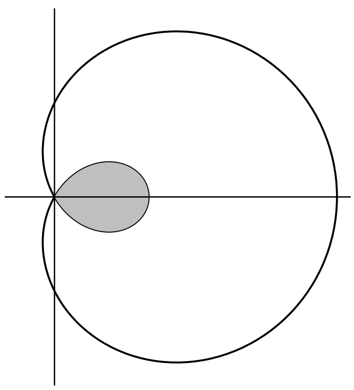


Solution. To find either value, we must find the interval on which this region occurs. We need to find the intersection of the two graphs, i.e. solve $\cos x = \sin x$, within the given diagram. This gives us $x = \dfrac{\pi}{4}, \dfrac{5\pi}{4}$ as our limits of integration.

For the area, note that $\sin x$ is above $\cos x$ for the shaded region, so our integral is
$$\int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \sin x - \cos x\, dx.$$

For the perimeter, we can simply add the arc lengths of $\sin x$ and $\cos x$ respectively on $\left[\dfrac{\pi}{4}, \dfrac{5\pi}{4}\right]$, and we obtain
$$\int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \sqrt{1 + \cos^2 x}\, dx + \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \sqrt{1 + \sin^2 x}\, dx = \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \sqrt{1 + \cos^2 x} + \sqrt{1 + \sin^2 x}\, dx. \quad \square$$

Problem 5.51. Given the graph of $r = 1 + 2\cos\theta$, find the area of the inside loop (shaded below):



Also, how would we get the area of the rest of the enclosed region?

Solution. Again, we need to find the values of θ for which $r = 0$, i.e. when we hit the origin, since that is when the inside loop is made. We solve $1 + 2\cos\theta = 0$ to get $\theta = \frac{2\pi}{3}, \frac{4\pi}{3}$. We can confirm that the inner loop is drawn when we go from $\theta = \frac{2\pi}{3}$ to $\frac{4\pi}{3}$ (this can be done by brief sketching or plugging-in of values). Thus, the area of that region is

$$\int_{\frac{2\pi}{3}}^{\frac{4\pi}{3}} \frac{1}{2}(1 + 2\cos\theta)^2 \, d\theta.$$

For the outside region, we can simply start from $\theta = \frac{4\pi}{3}$ up to the next time we hit the origin, which is $\frac{2\pi}{3} + 2\pi = \frac{8\pi}{3}$. Thus, the area of the outside region would be

$$\int_{\frac{4\pi}{3}}^{\frac{8\pi}{3}} \frac{1}{2}(1 + 2\cos\theta)^2 \, d\theta. \qquad \square$$

§5.5 Fundamental Theorems of Calculus

Finally, we tie together the concept of the derivative and integral with two parts of the Fundamental Theorem of Calculus. We start with a lemma and a preliminary theorem:

Lemma 5.52

If $f(x)$ is integrable on $[a, b]$ and $m \le f(x) \le M$ $\forall x \in [a, b]$, then

$$m(b - a) \le \int_a^b f(x) \, dx \le M(b - a).$$

Proof. If m and M are our bounds on the interval $[a, b]$, then for any partition of $[a, b]$, the infimums on any subinterval of that partition must be greater than or equal to m. Likewise, the supremums of the subintervals must be less than or equal to M.

Then,

$$\begin{aligned} L(f, P) &= \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \\ &\ge \sum_{i=1}^{n} m(x_i - x_{i-1}) \\ &= m \sum_{i=1}^{n} (x_i - x_{i-1}) \\ &= m(b - a), \end{aligned}$$

since the sum of the lengths of all subintervals is simply the length of the entire interval.

Likewise, we can show that $U(f, P) \le M(b - a)$, and since the integral is between the lower and upper sums, we can conclude that it is also between $m(b-a)$ and $M(b-a)$. $\square$

Theorem 5.53

Let f be integrable on $[a, b]$. For $x \in [a, b]$, define $F(x) = \displaystyle\int_a^x f(t)\,dt$. Then F is continuous on $[a, b]$.

Proof. Since f is integrable, it must also be bounded, so let $m \le f(x) \le M$ $\forall x \in [a, b]$.

Recalling back to our limit definition of continuity, we want to prove $\lim\limits_{z \to x} F(z) = F(x)$. This is equivalent to proving $\lim\limits_{h \to 0} F(x+h) = F(x)$, or $\lim\limits_{h \to 0} F(x+h) - F(x)$.

First, note that

$$\begin{aligned} F(x+h) - F(x) &= \int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt \\ &= \int_x^{x+h} f(t)\,dt, \end{aligned}$$

by Corollary 5.31. Now, we proceed by casework on the sign of h.

If $h > 0$, then by Lemma 5.52 on $[x, x+h]$, we have

$$m(x+h-x) = mh \le \int_x^{x+h} f(t)\,dt \le M(x+h-x) = Mh.$$

Then, we apply Squeeze Theorem here: note that $\lim\limits_{h \to 0} mh = \lim\limits_{h \to 0} Mh = 0$, we must have $\lim\limits_{h \to 0} \displaystyle\int_x^{x+h} f(t)\,dt = \lim\limits_{h \to 0} F(x+h) - F(x) = 0$.

If $h < 0$, we apply Lemma 5.52 on $[x+h, x]$ to get

$$m(x-(x+h)) = -mh \le \int_{x+h}^{x} f(t)\,dt \le M(x-(x+h)) = -Mh.$$

Multiplying by -1, we have

$$Mh \le -\int_{x+h}^{x} f(t)\,dt \le mh,$$

and we can rewrite $-\displaystyle\int_{x+h}^{x} f(t)\,dt = \int_x^{x+h} f(t)\,dt$, so we have

$$Mh \le \int_x^{x+h} f(t)\,dt \le mh,$$

from which we can finish by applying the Squeeze Theorem again. □

The proof of the first part of the Fundamental Theorem of Calculus uses a similar argument.

Theorem 5.54 (Fundamental Theorem of Calculus, Part 1)

For constants a, b, define $F(x) = \displaystyle\int_a^x f(t)\,dt$, and suppose that f is continuous on $[a, b]$. Then $F'(x) = f(x)$.

Proof. By the limit definition of the derivative, we consider

$$F'(x) = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} \frac{1}{h} \int_x^{x+h} f(t)\, dt.$$

Assume $h > 0$, and it is left to the reader to prove the analogous case $h < 0$. Define m_h and M_h to be the infimum and supremum of $f(x)$ on $[x, x+h]$ respectively.

By Lemma 5.52, we have

$$hm_h \le \int_x^{x+h} f(t)\, dt \le hM_h,$$

from which we divide by h to get

$$m_h \le \frac{1}{h} \int_x^{x+h} f(t)\, dt \le M_h.$$

Since f is continuous at x, we have $\lim_{h\to 0} f(x+h) = f(x)$. Thus, as $h \to 0$, then m_h and M_h both go to $f(x)$. Therefore, $F'(x) = \lim_{h\to 0} \frac{1}{h} \int_x^{x+h} f(t)\, dt = f(x)$. □

Henceforth, we will abbreviate this theorem as FTC1. Our immediate consequence of FTC1 is that

$$\frac{d}{dx} \int_a^x f(t)\, dt = f(x),$$

for some constant a.

Problem 5.55. Evaluate the following:

1. $\dfrac{d}{dx} \displaystyle\int_2^x \frac{1}{1+t^{20}}\, dt$
2. $\dfrac{d}{dx} \displaystyle\int_x^2 \frac{1}{1+t^{20}}\, dt$
3. $\dfrac{d}{dx} \displaystyle\int_{17}^{x^3} \frac{1}{1+t^{20}}\, dt$
4. $\dfrac{d}{dx} \displaystyle\int_{x^4}^{x^5} \frac{1}{1+t^{20}}\, dt$

Solution.

1. By FTC1, we immediately conclude that $\dfrac{d}{dx} \displaystyle\int_2^x \frac{1}{1+t^{20}}\, dt = \frac{1}{1+x^{20}}$.
2. As $\dfrac{d}{dx} \displaystyle\int_x^2 \frac{1}{1+t^{20}}\, dt = -\frac{d}{dx} \int_2^x \frac{1}{1+t^{20}}\, dt$, we have $\dfrac{d}{dx} \displaystyle\int_x^2 \frac{1}{1+t^{20}}\, dt = -\frac{1}{1+x^{20}}$.
3. Now this is tricky, because we have x^3 as the upper limit of integration. If we define $F(x) = \dfrac{d}{dx} \displaystyle\int_{17}^x \frac{1}{1+t^{20}}\, dt$, then $F'(x) = \dfrac{1}{1+x^{20}}$. Then evaluating the given integral is equivalent to finding $\dfrac{d}{dx}(f(x^3))$. We can then apply the chain rule:

$$\frac{d}{dx} \int_{17}^{x^3} \frac{1}{1+t^{20}}\, dt = \frac{d}{dx}(f(x^3)) = f'(x^3) \cdot 3x^2 = \frac{3x^2}{1+x^{60}}.$$

4. This time, the lower limit of integration is x^4, which is not a constant. Fortunately, we can use Corollary 5.31 to get

$$\frac{d}{dx}\int_{x^4}^{x^5}\frac{1}{1+t^{20}}\,dt = \frac{d}{dx}\int_a^{x^5}\frac{1}{1+t^{20}}\,dt - \frac{d}{dx}\int_a^{x^4}\frac{1}{1+t^{20}}\,dt,$$

for any constant a. Then, we can evaluate each of these terms using the chain rule:

$$\frac{d}{dx}\int_a^{x^5}\frac{1}{1+t^{20}}\,dt - \frac{d}{dx}\int_a^{x^4}\frac{1}{1+t^{20}}\,dt = \frac{5x^4}{1+x^{100}} - \frac{4x^3}{1+x^{80}}.$$ □

Theorem 5.56 (Fundamental Theorem of Calculus, Part 2)
Suppose f is continuous on $[a, b]$ and $\exists$ a function g such that $g'(x) = f(x)\ \forall x$. Then,

$$\int_a^b f(t)\,dt = g(b) - g(a).$$

Proof. Define $F(x) = \displaystyle\int_a^x f(t)\,dt$. By FTC1, $F'(x) = f(x)$.

Given that there exists a function g such that $g'(x) = f(x)$, it must follow that F and g differ by a constant, by Corollary 4.86. This result makes sense since the constant terms disappear when we differentiate. Thus, $\exists c \in \mathbb{R}$ such that $g(x) = F(x) + c\ \forall x$, or $c = g(x) - F(x)$.

If we substitute $x = a$, then we end up with $F(a) = 0$. This indicates that $c = g(a) - F(a) = g(a)$.

Since $F(x) = g(x) - c$, the result $c = g(a)$ tells us that $F(x) = g(x) - g(a)$, or

$$\int_a^x f(t)\,dt = g(x) - g(a),$$

from which considering $x = b$ gives us the desired result. □

The statement of this theorem has tremendous consequences. After not mentioning the derivative at all throughout this chapter, we suddenly have these Fundamental Theorems, which serve as the link between the derivative and the integral, the two main parts of calculus.

We will explore the results in great detail in the next chapter.

§5.6 Logarithmic and Exponential Functions

For now, we will temporarily depart from the main course by discussing logarithmic and exponential functions, with which I assume that you have had prior experience from your prior classes in mathematics. In this chapter, we will define the concept of the logarithm using integrals, which will then allow us to consider taking the limits and derivatives of these two types of functions. Our discussion will also introduce another famous constant that has a large presence in calculus: e.

To address notation, we will set e as our base for $\log x$, not 10 (as other mathematical texts may establish). However, we will not define the logarithm as the reverse of the exponential; instead, this will be a property that will follow from our calculus-based definition of $\log x$.

First, we need to establish a few facts:

Theorem 5.57

Let $a, b > 0$. Then,
$$\int_1^a \frac{1}{x}\,dx + \int_1^b \frac{1}{x}\,dx = \int_1^{ab} \frac{1}{x}\,dx.$$

Proof. We will use Theorem 5.35. Define $f(x) = \frac{1}{x}$. Since $b > 0$, we have $f\left(\frac{x}{b}\right) = \frac{b}{x} = b \cdot \frac{1}{x}$. We can then apply Theorem 5.35:
$$b\int_1^a \frac{1}{x}\,dx = \int_b^{ab} b \cdot \frac{1}{x}\,dx = b\int_b^{ab} \frac{1}{x}\,dx,$$
from which we conclude $\int_1^a \frac{1}{x}\,dx = \int_b^{ab} \frac{1}{x}\,dx$. But by Corollary 5.31,
$$\int_1^b \frac{1}{x}\,dx + \int_b^{ab} \frac{1}{x}\,dx = \int_1^{ab} \frac{1}{x}\,dx.$$
Since we have established $\int_1^a \frac{1}{x}\,dx = \int_b^{ab} \frac{1}{x}\,dx$, we have
$$\int_1^a \frac{1}{x}\,dx + \int_1^b \frac{1}{x}\,dx = \int_1^{ab} \frac{1}{x}\,dx.$$
□

If we define $F(x) = \int_1^x \frac{1}{t}\,dt$ for $x > 0$, then Theorem 5.57 states that $F(a) + F(b) = F(ab)$.

This is the first property of function F which will motivate us to define the logarithmic function based on F.

Furthermore, note that by FTC1, $F'(x) = \frac{1}{x}$, and therefore $F''(x) = -\frac{1}{x^2}$. Then, F is a function whose domain is $(0, \infty)$ satisfying:

- F is increasing since $F' > 0$.
- F is concave as $F'' < 0$.

In the context of the next few problems, we will be using the definition of F above.

Problem 5.58. Prove that $\forall k \in \mathbb{R}^+,\ F(kx) = F(k) + F(x)$.

Proof. By the Chain Rule, $\frac{d}{dx}(F(kx)) = k \cdot \frac{1}{kx} = \frac{1}{x} = \frac{d}{dx}(F(x))$. Because $\frac{d}{dx}(F(kx)) = \frac{d}{dx}(F(x))$, by Corollary 4.86, they differ by a constant, i.e. $\exists c \in \mathbb{R}$ such that $F(kx) = F(x) + c\ \forall x \in \mathbb{R}^+$.

Consider $x = 1$. Plugging this back into our relation above, we have $F(k) = F(1) + c$. But notice that $F(1) = \int_1^1 \frac{1}{t}\,dt = 0$, so we have $c = F(k)$. If we plug this back into the relation, we have
$$F(kx) = F(k) + F(x).$$
□

Exercise 5.59. Prove that $\forall n \in \mathbb{N},\ F(x^n) = nF(x)$.

Problem 5.60. Prove that $\forall r \in \mathbb{Q}^+$, $F(x^r) = rF(x)$.

Proof. Let $r = \dfrac{p}{q}$ for $p \in \mathbb{Z}^+$ and $q \in \mathbb{N}$. Consider adding the term $F\left(x^{\frac{p}{q}}\right)$ to itself q times, i.e.

$$\underbrace{F\left(x^{\frac{p}{q}}\right) + F\left(x^{\frac{p}{q}}\right) + \ldots + F\left(x^{\frac{p}{q}}\right)}_{q \text{ times}}.$$

This is obviously equal to $qF\left(x^{\frac{p}{q}}\right)$, but it is also equal to

$$F\underbrace{\left(x^{\frac{p}{q}} \cdot x^{\frac{p}{q}} \cdots x^{\frac{p}{q}}\right)}_{q \text{ times}} = F\left(x^{\frac{p}{q}+\frac{p}{q}+\ldots+\frac{p}{q}}\right) = F\left(x^p\right) = pF(x),$$

by repeated application of the property $F(a) + F(b) = F(ab)$.

Now we have $qF\left(x^{\frac{p}{q}}\right) = pF(x)$, which rearranges to $F\left(x^{\frac{p}{q}}\right) = \dfrac{p}{q}F(x)$. We substitute $r = \dfrac{p}{q}$ back into this to get $F(x^r) = rF(x)$, as desired. □

Alternative Proof. Since we can apply Theorem 4.34 on the rationals, we have

$$\frac{d}{dx}(F(x^r)) = \frac{1}{x^r} \cdot rx^{r-1} = \frac{r}{x} = \frac{d}{dx}(rF(x)).$$

Since $\dfrac{d}{dx}(F(x^r)) = \dfrac{d}{dx}(rF(x))$, we can use Corollary 4.86 again to show that

$$F(x^r) = rF(x) + \widetilde{c},$$

for some $\widetilde{c} \in \mathbb{R}$.

Again, substituting $x = 1$ gives $F(1) = rF(1) + \widetilde{c}$, and since $F(1) = 0$, we must have $\widetilde{c} = 0$. Thus, we have another property:

$$F(x^r) = rF(x).$$ □

If we look back at all these properties of F that we have established, we can see that they closely match the properties of logarithms that you have studied in a prior class in mathematics. In fact, we will define the logarithm equal to F:

Definition 5.61. $\log x = \displaystyle\int_1^x \frac{1}{t}\,dt$.

Then, Theorem 5.57 and Problem 5.60 tell us:

$$\log x + \log y = \log xy \ \forall x, y \in \mathbb{R}^+,$$

$$\log(x^r) = r \log x \ \forall r \in \mathbb{Q}^+.$$

It will be true that the property $\log(x^r) = r \log x$ holds for all real numbers r, but this realization will only manifest itself later when we define exponential functions.

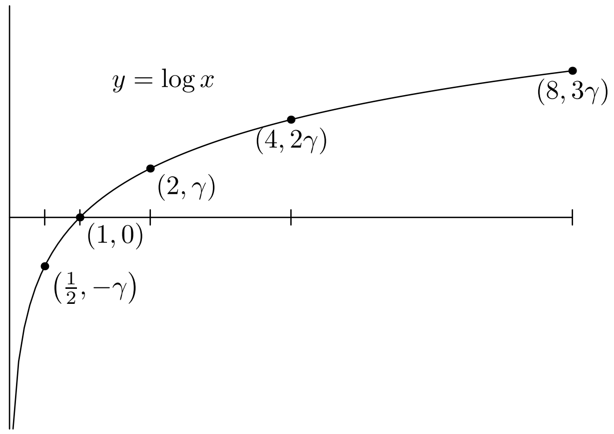
To get a sense of what this function's graph will look like, we need to take some limits:

- Since $\log 2 > 0$, we have $\displaystyle\lim_{r\to\infty} \log(2^r) = \lim_{r\to\infty} r \log 2 = -\infty$.
- Likewise, $\displaystyle\lim_{r\to-\infty} \log(2^r) = \lim_{r\to-\infty} r \log 2 = -\infty$.

As $r \to \infty$, $2^r \to \infty$. Since both of these numbers are going toward infinity, we can replace both of them by x, so we have $\lim\limits_{x\to\infty} \log x = \infty$.

Likewise, as $r \to -\infty$, $2^r \to 0^+$. Then, we have $\lim\limits_{x\to 0^+} \log x = -\infty$.

Now that we know the behavior of $\log x$ in both directions, we can proceed with a sketch. Let $\gamma = \log 2$. This will serve as our reference for the points we will plot. We can also use the observations that $\log x$ is increasing and concave on the domain.



From here, it is clear that the domain of $\log x$ is $(0, \infty)$ while the range is $\mathbb{R}$. Furthermore, $\log x$ is indeed one-to-one.

Definition 5.62. Let e be the number such that $\log e = 1$ is satisfied.

It follows that $\log e^r = r \log e = r \cdot 1 = r$.

Using the one-to-one property of $\log x$, we can consider a related function:

Definition 5.63. Let the function $\exp x$ be the inverse function of $\log x$. In other words,

$$y = \exp x \longleftrightarrow x = \log y.$$

Problem 5.64. Find $\exp'(x)$.

Solution. This leads to one of the most interesting properties of $\exp x$. By Theorem 4.122, we have

$$\exp'(x) = \frac{1}{\log'(\exp x)} = \frac{1}{\frac{1}{\exp x}} = \exp x.$$

The derivative of $\exp x$ is the same as the original function! □

Furthermore, we can find properties of this function based on what we had found for $\log x$. Note that $\log xy = \log x + \log y$, which implies $xy = \exp(\log x + \log y)$. Let $a = \log x$ and $b = \log y$, so we have $x = \exp a$ and $y = \exp b$. Then, it follows that

$$\exp a \exp b = \exp(a + b).$$

If $\log e = 1$, then $\exp 1 = e$. We also have $\log 1 = 0$, so $\exp 0 = 1$.

Lastly, since $\log e^r = r$ for $r \in \mathbb{Q}^+$, we can establish that $\exp r = e^r$.

However, we defined exp to be the inverse of $\log x$, but the range of $\log x$ is $\mathbb{R}$. It must follow that the domain of exp must be $\mathbb{R}$. Then $\exp r = e^r$ must work for all $r \in \mathbb{R}$.

Definition 5.65. We can extend the definition for e^r $\forall r \in \mathbb{R}$ as $\exp r = e^r$.

It then follows that $\log(x^r) = r\log x$ is satisfied $\forall r \in \mathbb{R}$. Thus, by establishing the range of the function $\log x$ as $\mathbb{R}$, we can define $\exp x$ with domain in $\mathbb{R}$ and thereby extend the definition to work for all real numbers.

Our previously discovered property $\exp a \exp b = \exp(a+b)$ then translates to $e^a \cdot e^s = e^{r+s}$, which is consistent with what we already know about exponents.

Furthermore, since we know that $\exp'(x) = \exp x$, we have $\frac{d}{dx}(e^x) = e^x$.

We can summarize the following properties about $\log x$ and $\exp x$ as shown:

- Define $\log x = \int_1^x \frac{1}{t}\,dt$ for $x > 0$, i.e. $x \in (0, \infty)$.
- $\log x + \log y = \log xy$, where $x, y > 0$.
- We can similarly show $\log\left(\frac{x}{y}\right) = \log x - \log y$.
- $\log(x^r) = r\log x$ for all real numbers r.
- For $\log x$, the domain is $(0, \infty)$ and the range is $\mathbb{R}$.
- As $\exp x$ is defined as the inverse of $\log x$, its domain is $\mathbb{R}$ and its range is $(0, \infty)$.
- $\exp n = e^n$ for all real numbers.
- $\exp a \exp b = \exp(a+b)$, or $e^a \cdot e^b = e^{a+b}$ for all real numbers.
- $(\exp r)^s = \exp(rs)$, or $(e^r)^s = e^{rs}$ for all real numbers.
- $\exp'(x) = \frac{d}{dx}(e^x) = e^x$.

Problem 5.66. Find $\frac{d}{dx}(e^{(x^2)})$ and $\frac{d}{dx}\left(\frac{x^3}{e^x}\right)$.

Solution. The first one is a mere application of the chain rule:

$$\frac{d}{dx}(e^{(x^2)}) = e^{(x^2)} \cdot \frac{d}{dx}(x^2) = \boxed{2xe^{(x^2)}}.$$

The second one also includes the quotient rule:

$$\frac{d}{dx}\left(\frac{x^3}{e^x}\right) = \frac{3x^2e^x - x^3e^x}{e^{2x}} = \boxed{\frac{3x^2 - x^3}{e^x}}.$$ □

For some constant $a > 0$, how would we deal with the derivative of a^x? Note that we can rewrite a as $e^{\log a}$, so $a^x = e^{x\log a}$. Since $e^{x\log a}$ is defined for all real numbers, we can extend the definition of a^x to work for all real numbers x.

To confirm that the exponent rules work the same way, we can evaluate

$$a^x a^y = e^{x\log a}e^{y\log a} = e^{(x+y)\log a} = a^{x+y}.$$

To find the derivative, we take advantage of the fact that $a^x = e^{x\log a}$. We have,

$$\frac{d}{dx}(a^x) = \frac{d}{dx}(e^{x\log a}) = e^{x\log a}\frac{d}{dx}(x\log a) = \log a \cdot e^{x\log a} = a^x\log a.$$

Definition 5.67. For $a > 0$, $\log_a x$ is the inverse of a^x.

Again, by Theorem 4.122, we have

$$\frac{d}{dx}(\log_a x) = \frac{1}{a^{\log_a x} \cdot \log a} = \frac{1}{x \log a}.$$

Problem 5.68. Show that Theorem 4.34 works for all real numbers.

Proof. Note that $\frac{d}{dx}(x^n) = \frac{d}{dx}(e^{n \log x})$. We know that $e^{n \log x}$ is defined for all real numbers, so we can safely take the derivative of this keeping that in mind.

$$\frac{d}{dx}(e^{n \log x}) = \frac{d}{dx}(n \log x) \cdot e^{n \log x} = \frac{n}{x} \cdot x^n = nx^{n-1}. \qquad \square$$

Problem 5.69. Find $\frac{d}{dx}(x^x)$.

Solution. Again, we rewrite x^x as $e^{x \log x}$ and then apply the chain rule.

$$\begin{aligned}
\frac{d}{dx}(x^x) &= \frac{d}{dx}(e^{x \log x}) \\
&= e^{x \log x}(\log x + 1) \\
&= x^x(1 + \log x).
\end{aligned} \qquad \square$$

Problem 5.70. Minimize x^x on $(0, \infty)$.

Solution. The critical points satisfy $x^x(1 + \log x) = 0$. Obviously x^x cannot equal 0, but we can solve $1 + \log x = 0$ to get $x = \frac{1}{e}$.

A first derivative number line will tell us that $x^x(1 + \log x)$ is negative on $\left(0, \frac{1}{e}\right]$ and positive on $\left[\frac{1}{e}, \infty\right)$, so we must conclude that $x = \frac{1}{e}$ yields a global minimum. Thus, the minimum value of x^x on $(0, \infty)$ is $\left(\frac{1}{e}\right)^{\frac{1}{e}}$, or $\boxed{\sqrt[e]{\frac{1}{e}}}$. $\square$

Problem 5.71. Maximize $\sqrt[x]{x}$ on $(0, \infty)$.

Solution. We want to find $\frac{d}{dx}(x^{\frac{1}{x}})$ first. We have

$$\begin{aligned}
\frac{d}{dx}(x^{\frac{1}{x}}) &= \frac{d}{dx}(e^{\frac{1}{x} \log x}) \\
&= e^{\frac{1}{x} \log x} \frac{d}{dx}\left(\frac{1}{x} \log x\right) \\
&= x^{\frac{1}{x}} \left[-\frac{1}{x^2} \cdot \log x + \frac{1}{x^2}\right] \\
&= x^{\frac{1}{x}} \left[\frac{1 - \log x}{x^2}\right].
\end{aligned}$$

Solving for critical points, we end up with the only equation $1 - \log x = 0$ which gives us $x = e$. The first derivative number line confirms that this is a global maximum, so our maximum value of $\sqrt[x]{x}$ on $(0, \infty)$ is $e^{\frac{1}{e}}$, or $\boxed{\sqrt[e]{e}}$. $\square$

From our discussion of logarithmic and exponential functions, we can introduce a new, powerful technique to differentiate complicated functions, called **logarithmic differentiation**. The following example should suffice in demonstrating how it is used.

Example 5.72

Find $\frac{d}{dx}\left(\frac{x^3\sin^7 x}{\sqrt{x^5+x+2}}\right)$.

Solution. It would be incredibly painful to use the quotient and chain rules here, so we use logarithmic differentiation.

Let $y = \frac{x^3\sin^7 x}{\sqrt{x^5+x+2}}$. Then take the logarithm of both sides,

$$\log y = \log\left(\frac{x^3\sin^7 x}{\sqrt{x^5+x+2}}\right).$$

Using the facts that $\log x + \log y = \log xy$, $\log(x^r) = r\log x$, and $\log\left(\frac{x}{y}\right) = \log x - \log y$, we can break up the right-hand side:

$$\begin{aligned}\log y &= \log\left(\frac{x^3\sin^7 x}{\sqrt{x^5+x+2}}\right)\\ &= 3\log x + 7\log(\sin x) - \frac{1}{2}\log(x^5+x+2).\end{aligned}$$

Finally, we take the derivative of both sides, yielding:

$$\begin{aligned}\frac{y'}{y} &= \frac{3}{x} + 7\cot x - \frac{1}{2}\cdot\frac{5x^4+1}{x^5+x+2}\\ y' &= y\left[\frac{3}{x} + 7\cot x - \frac{1}{2}\cdot\frac{5x^4+1}{x^5+x+2}\right].\end{aligned}$$

Of course, we cannot just leave the y term as-is, so we substitute $y = \frac{x^3\sin^7 x}{\sqrt{x^5+x+2}}$ back into the expression to get

$$y' = \frac{x^3\sin^7 x}{\sqrt{x^5+x+2}}\left[\frac{3}{x} + 7\cot x - \frac{1}{2}\cdot\frac{5x^4+1}{x^5+x+2}\right],$$

as our answer.

As long as you apply the rules of logarithms and differentiate correctly, you should arrive at the answer much more seamlessly. □

Problem 5.73. Find $\frac{d}{dx}(x^{(10^x)})$.

Solution. Let $y = x^{(10^x)}$. Then $\log y = 10^x \log x$, so we differentiate both sides to get

$$\begin{aligned}\frac{y'}{y} &= 10^x \log 10 \log x + 10^x\cdot\frac{1}{x}\\ y' &= x^{(10^x)}\cdot 10^x\left[\log 10\log x + \frac{1}{x}\right].\end{aligned}$$ □

Exercise 5.74. Solve the previous problem using the method of rewriting $x^{(10^x)}$ as an exponential function.

Problem 5.75. Differentiate $(\sin x)^{\cos x}$.

Solution. Let $y = (\sin x)^{\cos x}$. Then $\log y = \log\left((\sin x)^{\cos x}\right) = (\cos x)\log(\sin x)$. Differentiating both sides yields

$$\frac{y'}{y} = \cos x \cdot \frac{\cos x}{\sin x} + (-\sin x)\log(\sin x),$$

which rearranges to $y' = (\sin x)^{\cos x}\left[\cos x \cot x - \sin x \log(\sin x)\right]$. $\square$

Recall that some indeterminate forms included $\frac{\pm\infty}{\pm\infty}$, $\frac{0}{0}$, and $\infty - \infty$. Following the notion of exponents, we have three more indeterminate forms:

$$0^0,\ \infty^0,\ 1^\infty.$$

We will still have to apply L'Hôpital's Rule, but after some rewriting.

Problem 5.76. Evaluate $\lim\limits_{x\to\infty} x^{\frac{1}{x}}$.

Solution. Rewrite $\lim\limits_{x\to\infty} x^{\frac{1}{x}}$ as $\lim\limits_{x\to\infty} e^{\frac{1}{x}\log x}$. As e^x is continuous, we can take the limit of the exponent itself and then plug back what we get into the main function.

In other words, $\lim\limits_{x\to\infty} e^{\frac{1}{x}\log x} = e^{\left(\lim_{x\to\infty}\frac{1}{x}\log x\right)}$.

It suffices to find $\lim\limits_{x\to\infty}\frac{1}{x}\log x = \lim\limits_{x\to\infty}\frac{\log x}{x}$. But this is an indeterminate form that we can deal with: $\frac{\infty}{\infty}$. We can apply L'Hôpital's Rule to get $\lim\limits_{x\to\infty}\frac{1}{x} = 0$. Thus, our overall limit is $e^0 = \boxed{1}$. $\square$

Problem 5.77. Evaluate $\lim\limits_{x\to 0^+} x^x$.

Solution. Rewrite $\lim\limits_{x\to 0^+} x^x$ as $\lim\limits_{x\to 0^+} e^{x\log x}$.

We can then evaluate $\lim\limits_{x\to 0^+} x\log x = \lim\limits_{x\to 0^+}\frac{\log x}{\frac{1}{x}}$, which becomes $\lim\limits_{x\to 0^+}\frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim\limits_{x\to 0^+} -x = 0$ after L'Hôpital's Rule. Thus, the overall limit is $e^0 = \boxed{1}$. $\square$

Problem 5.78. Evaluate $\lim\limits_{x\to\infty}\left(1+\frac{r}{x}\right)^x$ $\forall r\in\mathbb{R}$.

Solution. Again, we rewrite $\lim\limits_{x\to\infty}\left(1+\frac{r}{x}\right)^x$ as $\lim\limits_{x\to\infty} e^{x\log\left(1+\frac{r}{x}\right)}$. We then have to compute $\lim\limits_{x\to\infty} x\log\left(1+\frac{r}{x}\right)$.

We have

$$\lim_{x\to\infty} x\log\left(1+\frac{r}{x}\right) = \lim_{x\to\infty}\frac{\log\left(1+\frac{r}{x}\right)}{\frac{1}{x}} = \lim_{x\to\infty}\frac{\log(x+r)-\log x}{\frac{1}{x}},$$

from which we apply L'Hôpital's Rule to get

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

Thus, the overall limit is $\boxed{e^r}$. $\square$

From the previous result, we can obtain our alternative definition of e:

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

Problem 5.79. Compute $\lim_{x\to 1} x^{\frac{1}{x-1}}$.

Solution. We have $\lim_{x\to 1} x^{\frac{1}{x-1}} = \lim_{x\to 1} e^{\frac{1}{x-1}\log x}$. We can then compute $\lim_{x\to 1} \frac{1}{x-1}\log x = \lim_{x\to 1} \frac{\log x}{x-1}$, which yields indeterminate form $\frac{0}{0}$. By L'Hôpital's Rule, we have $\lim_{x\to 1} \frac{\log x}{x-1} = \lim_{x\to 1} \frac{1}{x} = 1$.

Thus, the overall limit is $e^1 = \boxed{e}$. ☐

Problem 5.80. There exists a unique constant k such that $\lim_{x\to 0} \frac{e^x - 1 - kx}{x^2}$ exists. Find this value of k, then find the limit.

Solution. We can first note that $\lim_{x\to 0} \frac{e^x - 1 - kx}{x^2}$ gives an indeterminate form $\frac{0}{0}$ regardless of what k is, so we can apply L'Hôpital's Rule to get $\lim_{x\to 0} \frac{e^x - k}{2x}$.

If we plug in $x = 0$, we get $\frac{1-k}{0}$. This suggests that if $k = 1$, we would end up with another indeterminate form $\frac{0}{0}$, from which we'd have to apply L'Hôpital's Rule again.

Thus, assume $k = 1$. Then $\lim_{x\to 0} \frac{e^x - k}{2x} = \lim_{x\to 0} \frac{e^x}{2} = \frac{1}{2}$.

Thus, for the unique constant $k = \boxed{1}$, we have found the limit to be $\boxed{\frac{1}{2}}$.

If k did not equal 1, then $\lim_{x\to 0} \frac{e^x - k}{2x}$ would not exist. ☐

§6 Differential Equations

Definition 6.1. A **differential equation** is an equation involving derivatives.

In this section, we will only go over basic forms of differential equations. The first form is the classic

$$y' = f(x),$$

for which we would have to solve for the function y. In this example, the solution is easily

$$y = \int f(x)\,dx.$$

From here, we can see that there are an infinite number of solutions for y, because of the constant of integration which arises from taking the antiderivative of y'.

Problem 6.2. Solve the given **initial value problem**:

$$y' = 3x^2 + 2x + 1,\ \ y(2) = 4.$$

Solution. First, we have

$$y = \int 3x^2 + 2x + 1\,dx = x^3 + x^2 + x + C,$$

where C is some real number. However, it asks for a function of y such that $y(2) = 4$. Then, we can plug these numbers back into the antiderivative and solve for C:

$$4 = 2^3 + 2^2 + 2 + C \implies C = -10,$$

therefore $y = x^3 + x^2 + x - 10$. □

Problem 6.3. Solve the given initial value problem:

$$y' = 3\sin 2x - 4\cos 3x,\ \ y(0) = 3.$$

Solution. First,

$$y = \int 3\sin 2x - 4\cos 3x\,dx = -\frac{3}{2}\cos 2x - \frac{4}{3}\sin 3x + C,$$

from we solve $C = 3 + \frac{3}{2}\cos 0 + \frac{4}{3}\sin 0 = \frac{9}{2}$. This gives us the solution

$$y = -\frac{3}{2}\cos 2x - \frac{4}{3}\sin 3x + \frac{9}{2}.$$ □

Now, we move on to a different type of differential equation.

Definition 6.4. A differential equation is called **separable** if it can be written in the form

$$\frac{dy}{dx} = f(x) \cdot g(y).$$

Exercise 6.5. Determine which of the following are separable: $\frac{dy}{dx} = xy$, $\frac{dy}{dx} = x + y$, and $\frac{dy}{dx} = e^{x+y}$.

Now we try to solve the general form of a separable differential equation

$$\frac{dy}{dx} = f(x) \cdot g(y).$$

First, define $h(y) = \dfrac{1}{g(y)}$, such that we can rewrite the above as

$$h(y)\frac{dy}{dx} = f(x).$$

Then, we integrate both sides:

$$\int h(y)\frac{dy}{dx}\, dx = \int f(x)\, dx.$$

Here, we end up with a complication: how do we take care of $\dfrac{dy}{dx}\, dx$? First, let's define $H(y)$ such that $H'(y) = h(y)$. Then, if we take $\dfrac{d}{dx}(H(y))$, then using the Chain Rule, we end up with $h(y)\dfrac{dy}{dx}$. Therefore, it makes sense to say that

$$\int h(y)\frac{dy}{dx}\, dx = H(y).$$

It happens that $\dfrac{dy}{dx}$ can be considered a "fraction" with dy in the numerator and dx in the denominator, such that we may conveniently cancel the dx terms in $\dfrac{dy}{dx}\, dx$ to end up with $\int h(y)\, dy$, which is also equal to $H(y)$.

Therefore, if we are at the step

$$h(y)\frac{dy}{dx} = f(x),$$

then we can "multiply both sides" of the equation by dx to get

$$h(y)\, dy = f(x)\, dx,$$

from which integrating both sides would be appropriate.

Problem 6.6. Solve the differential equation $\dfrac{dy}{dx} = e^{x+y}$.

Solution. We rewrite the given as $\dfrac{dy}{dx} = e^x \cdot e^y$. Then, we can separate the right-hand side by moving e^y, which a function in terms of y, to the other side. We end up with

$$e^{-y}\, dy = e^x\, dx,$$

from which integrating both sides yields

$$-e^{-y} + C = e^x + \widetilde{C}.$$

Since $\widetilde{C} - C$ is also some constant, it is unnecessary to write the constant of integration for both sides of the equation. We have

$$e^{-y} = -e^x + \widetilde{\widetilde{C}},$$

and finally we take the natural logarithm of both sides to get

$$y = -\log(\widetilde{\widetilde{C}} - e^x)$$

as our family of solutions. $\square$

Exercise 6.7. Solve $\dfrac{dy}{dx} = \dfrac{x}{y}$.

Example 6.8

Solve $y' = ky$, where k is a constant.

Solution. We rewrite y' as $\dfrac{dy}{dx}$, and rearrange such that only the k term remains on the right-hand side:

$$\frac{1}{y}\,dy = k\,dx.$$

Taking the integral of both sides gives us $\log|y| = kx + C$. Then $|y| = e^{kx+C}$. What should we do about the absolute value? Note that $e^{kx+C} = e^C \cdot e^{kx}$, and e^C also a constant. Let $\widetilde{C} = e^C$, such that

$$|y| = \widetilde{C}e^{kx}.$$

Furthermore, since we have defined $\widetilde{C}$ to denote e^C, and e^C is always positive, it follows that $\widetilde{C} > 0$. Then, we can extend this definition to all real numbers, i.e. let $\widetilde{C} \in \mathbb{R}$. Then, $\widetilde{C}e^{kx}$ can either be positive or negative, which allows us to drop the absolute signs on y. We have

$$y = \widetilde{C}e^{kx},\ \ \widetilde{C} \in \mathbb{R},$$

as our overall family of solutions. Note that $y' = 0$ is also included. □

§7 Sequences and Series

§7.1 Introduction: Estimating Inverse Tangent

Consider the function $\frac{1}{x^2+1}$. While we have used long division to divide a smaller degree polynomial into a higher degree polynomial (in particular, when evaluating integrals of rational functions), we can still apply long division here, with $1+x^2$ as the divisor and 1 as the dividend. Applying the process for a few times gives us

$$\frac{1}{x^2+1} = 1 - x^2 + x^4 - x^6 + x^8 \dots$$

and we can see that this goes on infinitely. We can confirm that $\frac{1}{x^2+1}$ is indeed equal to this infinite series by applying the formula for an infinite geometric series: the first term is 1, the common ratio is $-x^2$, so the sum is $\frac{1}{1-(-x^2)} = \frac{1}{x^2+1}$, as expected.

Rather than having this series continue forever, consider a general last term of this series. For some $n \in \mathbb{Z}^+$, if we continue the long division up to some point, we have $1-x^2+x^4-x^6+x^8-\dots+(-1)^n x^{2n}$ as the quotient and $(-1)^{n+1}x^{2n+2}$ as the remainder. Then,

$$\frac{1}{x^2+1} = 1 - x^2 + x^4 - x^6 + x^8 - \dots + \frac{(-1)^{n+1}x^{2n+2}}{x^2+1}.$$

We have an identity to consider. If we take the definite integral from 0 to $a \in \mathbb{R}$ of both sides, we get

$$\int_0^a \frac{1}{x^2+1}\,dx = \int_0^a 1 - x^2 + x^4 - x^6 + x^8 - \dots + \frac{(-1)^{n+1}x^{2n+2}}{x^2+1}\,dx.$$

This evaluates to

$$\tan^{-1} a = a - \frac{a^3}{3} + \frac{a^5}{5} - \frac{a^7}{7} + \dots + (-1)^n \frac{a^{2n+1}}{2n+1} + (-1)^{n+1}\int_0^a \frac{x^{2n+2}}{x^2+1}\,dx.$$

We *almost* have a clean series for $\tan^{-1} a$, but we have this complicated integral as a remainder. Let

$$R_n = \int_0^a \frac{x^{2n+2}}{x^2+1}\,dx.$$

It is not advisable to directly evaluate this integral, but we can "estimate" its value by comparing the inner function to a similar function, $\frac{1}{x^2+1}$. Note that

$$R_n < \int_0^a x^{2n+2}\,dx = \frac{a^{2n+3}}{2n+3}.$$

At this point, we have to consider what kind of number a is. Clearly, if $a > 1$ or $a < -1$, then $\frac{a^{2n+3}}{2n+3}$ grows unbounded as n gets larger. Then, if we restrict possible values of a to $[-1,1]$, then $\lim_{n\to\infty} \frac{a^{2n+3}}{2n+3} = 0$. Therefore, whenever $|a| \le 1$, $\lim_{n\to\infty} R_n = 0$.

Since

$$\tan^{-1} a = a - \frac{a^3}{3} + \frac{a^5}{5} - \frac{a^7}{7} + \dots + (-1)^n \frac{a^{2n+1}}{2n+1} + (-1)^{n+1}R_n,$$

it follows that as long as $a \in [-1,1]$ and we take sufficiently many terms (i.e. consider larger n), we can make

$$a - \frac{a^3}{3} + \frac{a^5}{5} - \frac{a^7}{7} + \dots + (-1)^n \frac{a^{2n+1}}{2n+1}$$

as close as we want to $\tan^{-1} a$.

For instance, take $a = 1$. We know that $\tan^{-1}(1) = \frac{\pi}{4}$, but we can also utilize the approximation above. Since

$$\tan^{-1}(1) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots + (-1)^n \frac{1}{2n+1},$$

we can conclude that $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots$ can be made as close as we want to $\frac{\pi}{4}$ by taking sufficiently many terms.

For some $n \in \mathbb{N}$, we know that there will be some **error** $(-1)^{n+1} R_n$. However, recall that

$$R_n < \frac{a^{2n+3}}{2n+3},$$

so it follows that

$$\left|(-1)^{n+1} R_n\right| = |\text{Error}| < \left|\frac{a^{2n+3}}{2n+3}\right|.$$

We can prove the usefulness of approximating the error through another example. Recall that $\tan^{-1} \frac{1}{3} + \tan^{-1} \frac{1}{2} = \frac{\pi}{4}$ (one can prove this using the tangent addition formula). Let's evaluate the first two terms of the approximations for each of $\tan^{-1} \frac{1}{3}$ and $\tan^{-1} \frac{1}{2}$:

$$\tan^{-1} \frac{1}{3} = \frac{1}{3} - \frac{1}{81} + \ldots,$$
$$\tan^{-1} \frac{1}{2} = \frac{1}{2} - \frac{1}{24} + \ldots$$

For $\tan^{-1} \frac{1}{3}$, our error will be less than $\frac{1}{3^5 \cdot 5} = \frac{1}{1215}$ at $n = 1$. Likewise, the error for $\tan^{-1} \frac{1}{2}$ will be less than $\frac{1}{160}$. Then, we conclude that $\frac{\pi}{4}$ is within $\frac{1}{1215} + \frac{1}{160} = \frac{55}{7776}$ of $\frac{1}{3} - \frac{1}{81} + \frac{1}{2} - \frac{1}{24} = \frac{505}{648}$. This, along with a calculator, indicates that π is within $\frac{55}{1944} \approx 0.02829$ of $\frac{505}{162} \approx 3.11728$, i.e. $3.08899 \leq \pi \leq 3.14557$, which is consistent with our knowledge that $\pi \approx 3.14159\ldots$ This shows that we are able to estimate the value and its corresponding error up to any $n \in \mathbb{Z}^+$ with certainty.

§7.2 Taylor Polynomials

Consider a polynomial $P(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$. We can first see that $P(0) = a_0$. We can differentiate the polynomial to get

$$P'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1}.$$

Then, we see that $P'(0) = a_1$. Again, we differentiate and get

$$P''(x) = 2a_2 + 6a_3 x + \ldots + n(n-1) a_n x^{n-2}.$$

Now, $P''(0) = 2a_2$, or $a_2 = \frac{P''(0)}{2}$. If we continue this process for the third derivative, we can also see that $a_3 = \frac{P'''(0)}{6}$.

From $a_0 = P(0)$, $a_1 = P'(0)$, $a_2 = \frac{P''(0)}{2}$, and $a_3 = \frac{P'''(0)}{6}$, we can see that the general pattern is

$$a_k = \frac{P^{(k)}(0)}{k!}.$$

Why did we choose 0 in particular? It turns out that we can generalize the result: consider the polynomial

$$P(x) = a_0 + a_1(x - r) + a_2(x - r)^2 + \ldots + a_n(x - r)^n.$$

We say that this polynomial is **centered** at r. Then, it is true that

$$a_k = \frac{P^{(k)}(r)}{k!}.$$

Example 7.1

Write $P(x) = x^3$ as a polynomial centered at 2.

Solution. Since the nth derivative of $P(x)$ is just 0 for $n \geq 4$, we only have to calculate values for the first three. First, note that $a_0 = P(2) = 8$. We differentiate $P(x)$ to get $P'(x) = 3x^2$, from which we obtain $a_1 = P'(2) = 12$. Repeating this process gives us $a_2 = \frac{12}{2!} = 6$ and $a_3 = \frac{6}{3!} = 1$. Now that we have our a_k coefficients, we can write out the entire polynomial:

$$x^3 = 8 + 12(x-2) + 6(x-2)^2 + (x-2)^3.$$ □

Exercise 7.2. Write $2x^4 - x^3 + 5x^2 + 6x - 2$ as a polynomial centered at -1.

Definition 7.3. Let f be a function that can be repeatedly differentiated many times at $x = a$. The nth order **Taylor Polynomial** for $f(x)$ at $x = a$ is

$$P_{n,a,f}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

If the function in discussion is clear from context, I will simply refer to its nth order Taylor Polynomial as $P_{n,a}(x)$.

Problem 7.4. Does the degree of a Taylor Polynomial always equal its order?

Solution. No, because $f^{(k)}(a)$ for some values of k can equal 0. □

Here are the Taylor Polynomials that you should remember (and know how to derive) for well-known functions:

- For $f(x) = e^x$,

$$P_{n,0,f}(x) = \sum_{k=0}^{n} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!}.$$

- For $f(x) = \sin x$,

$$P_{2n+1,0,f}(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}.$$

  Note that $P_{2n+2,0,f}(x) = P_{2n+1,0,f}(x)$.

- For $f(x) = \cos x$,

$$P_{2n,0,f}(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + (-1)^n \frac{x^{2n}}{(2n)!}.$$

- For $f(x) = \log(x+1)$,

$$P_{n,0,f}(x) = \sum_{k=1}^{n} (-1)^{k-1}\frac{x^k}{k} = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots + (-1)^{n-1}\frac{x^n}{n}.$$

Problem 7.5. If $P_{n,a}(x)$ is the nth order Taylor Polynomial for $f(x)$ around $x = a$, what can you say about $P'_{n,a}(x)$?

Solution. Since

$$P_{n,a,f}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k,$$

we have

$$\begin{aligned}
P'_{n,a,f}(x) &= \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!} k(x-a)^{k-1} \\
&= \sum_{k=1}^{n} \frac{f^{(k)}(a)}{(k-1)!}(x-a)^{k-1} \\
&= \sum_{k=0}^{n-1} \frac{f^{(k+1)}(a)}{k!}(x-a)^k \\
&= P_{n-1,a,f'}(x).
\end{aligned}$$

□

Exercise 7.6. Find $P_{n,\pi}(x)$ for $f(x) = \sin x$.

Exercise 7.7. Find $P_{4,\frac{\pi}{4}}(x)$ for $f(x) = \cos x$.

Definition 7.8. We say that a function f is n-times differentiable at $x = a \in \mathbb{R}$ if $f(a), f'(a), f''(a), \ldots, f^{(n)}(a)$ exist.

Theorem 7.9

Let $f(x)$ be n-times differentiable at a for $a \in \mathbb{R}$. Then, $\lim\limits_{x\to a} \dfrac{f(x) - P_{n,a}(x)}{(x-a)^n} = 0.$

Proof. We proceed by induction. For the base case $n = 0$, we just need to show that $\lim\limits_{x\to a} f(x) - P_{0,a}(x) = 0$. However, $P_{0,a}(x) = f(a)$, and $\lim\limits_{x\to a} f(x) - f(a)$ follows since f is necessarily continuous (differentiable implies continuous).

Now, assume that the theorem holds for $n = k$. Then, we can apply L'Hôpital's Rule to get

$$\lim_{x\to a} \frac{f(x) - P_{k+1,a}(x)}{(x-a)^{k+1}} = \lim_{x\to a} \frac{f'(x) - P_{k,a,f'}(x)}{(k+1)(x-a)^k},$$

recalling the result from Problem 7.5. Then, by the inductive hypothesis (where $f'(x)$ is also a function with lots of derivatives),

$$\lim_{x\to a} \frac{f'(x) - P_{k,a,f'}(x)}{(x-a)^k} = 0,$$

so the overall limit is 0, concluding our inductive proof. □

Definition 7.10. $f(x)$ is equal to $g(x)$ up to order n at $x = a$ if

$$\lim_{x\to a} \frac{f(x) - g(x)}{(x-a)^n} = 0.$$

From this definition, Theorem 7.9 states that $f(x)$ is equal to $P_{n,a,f}(x)$ up to order n at $x = a$.

Lemma 7.11

Let $P(x)$, $Q(x)$ be polynomials of degree $\leq n$. If $P(x)$ is equal to $Q(x)$ up to order n at $x = a$, then $P(x) = Q(x)$.

Proof. Given $\lim\limits_{x\to a} \frac{P(x)-Q(x)}{(x-a)^n} = 0$, we want to prove that $P(x) = Q(x)$. Let $R(x) = P(x) - Q(x)$. Then $\deg R \leq n$, and it suffices to prove that if $\lim\limits_{x\to a} \frac{R(x)}{(x-a)^n} = 0$, then $R(x) = 0$.

We can prove this modified assertion using induction. For the base case $n = 0$, note that $\lim\limits_{x\to a} \frac{R(x)}{(x-a)^0} = \lim\limits_{x\to a} R(x) = 0$ by assumption. Since $R(x)$ is a polynomial, it is necessarily continuous, thus $\lim\limits_{x\to a} R(x) = R(a)$, from which we conclude $R(a) = 0$. Since the degree of R is necessarily equal to 0, the polynomial must be constant, so $R(x) = 0$ for all x.

Assume that the statement holds for $n \leq k$. We have established that $R(a) = 0$, and by the Factor Theorem, $(x - a) \mid R(x)$. Let $R(x) = (x - a)\widetilde{R}(x)$.

To prove that the statement is true for $n = k + 1$, assume that $\lim\limits_{x\to a} \frac{R(x)}{(x-a)^{k+1}} = 0$. Observe that $\lim\limits_{x\to a} \frac{R(x)}{(x-a)^{k+1}} = \lim\limits_{x\to a} \frac{(x-a)\widetilde{R}(x)}{(x-a)^{k+1}} = \lim\limits_{x\to a} \frac{\widetilde{R}(x)}{(x-a)^k} = 0$. Given this, the inductive hypothesis allows us to conclude $\widetilde{R}(x) = 0$. Then, $R(x) = (x - a)\widetilde{R}(x) = 0$, finishing our inductive step. □

Lemma 7.12

Let $f(x)$ be n-times differentiable at $a \in \mathbb{R}$. If $f(x)$ is equal to $P(x)$, a polynomial of degree $\leq n$, up to order n at $x = a$, then $P(x) = P_{n,a,f}(x)$.

Proof. This immediately follows from Theorem 7.9 and Lemma 7.11. □

Now, we can use these lemmas to find an appropriate Taylor Polynomial for $\tan^{-1} x$. Recall from the first section that for $n \in \mathbb{Z}^+$,

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots + (-1)^n \frac{x^{2n+1}}{2n+1} + (-1)^{n+1} \int_0^x \frac{t^{2n+2}}{t^2+1}\, dt.$$

Example 7.13

Let $f(x) = \tan^{-1} x$. For $n \in \mathbb{Z}^+$, prove that

$$P_{2n+2,0,f}(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots + (-1)^n \frac{x^{2n+1}}{2n+1}.$$

Proof. By Lemma 7.12, it suffices to prove that $\tan^{-1} x$ is equal to $x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots + (-1)^n \frac{x^{2n+1}}{2n+1}$ up to order $2n+2$ at $x = 0$.

Let $P(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots + (-1)^n \frac{x^{2n+1}}{2n+1}$. Then,

$$\lim_{x \to 0} \frac{\tan^{-1} x - P(x)}{x^{2n+2}} = (-1)^{n+1} \lim_{x \to 0} \frac{\int_0^x \frac{t^{2n+2}}{1+t^2}\, dt}{x^{2n+2}}.$$

Now, note that we have an indeterminate form, $\frac{0}{0}$, so we apply L'Hôpital's Rule and FTC 2 to get

$$(-1)^{n+1} \lim_{x \to 0} \frac{\frac{x^{2n+2}}{1+x^2}}{(2n+2)x^{2n+1}} = \frac{(-1)^{n+1}}{2n+2} \lim_{x \to 0} \frac{x}{x^2+1} = 0,$$

as desired. □

Problem 7.14. What is the 11th derivative of $\tan^{-1} x$ at $x = 0$?

Solution. We've discovered that

$$\tan^{-1} 0 + (\tan^{-1})'(0)x + \frac{(\tan^{-1})''(0)}{2}x^2 + \frac{(\tan^{-1})'''(0)}{3!}x^3 + \ldots + \frac{(\tan^{-1})^{(2n+1)}(0)}{(2n+1)!}x^{2n+1}$$
$$= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots + (-1)^n \frac{x^{2n+1}}{2n+1}.$$

Matching up coefficients, we can see that $\frac{(\tan^{-1})^{(11)}(0)}{11!}x^{11} = -\frac{1}{11}x^{11}$. We can solve this to get $(\tan^{-1})^{(11)}(0) = -\frac{11!}{11} = \boxed{-10!}$.

We can see that in general, the nth derivative of $\tan^{-1} x$ at $x = 0$ is equal to 0 when n is even, and $(-1)^k \cdot (2k)!$ when n is odd and can be written as $2k+1$ for $k \in \mathbb{Z}^+$. □

Now that we've discovered that the function can be approximated to its Taylor Polynomial, there must be some kind of "remainder" since, in most cases, the function and its Taylor Polynomial are not exactly equal.

Definition 7.15. The nth **remainder term** for $f(x)$ of order n at $x = a$ is $R_{n,a}(x)$ where $f(x) = P_{n,a}(x) + R_{n,a}(x)$.

Theorem 7.16

If f is a function for which $P_{n,a}(x)$ exists, then $R_{n,a}(x) = \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n\, dt.$

Proof. We proceed by induction. For base case $n = 0$, note that $\int_a^x \frac{f'(t)}{0!}(x-t)^0\, dt = \int_a^x f'(t)\, dt = f(x) - f(a) = f(x) - P_{0,a}(x)$.

Assume that the statement is true for $n = k$. Then, we can evaluate

$$\int_a^x \frac{f^{(k+2)}(t)}{(k+1)!}(x-t)^{k+1}\, dt$$

using integration by parts. Let $u = \frac{(x-t)^{k+1}}{(k+1)!}$ and $dv = f^{(k+2)}(t)\,dt$, so $du = -\frac{(x-t)^k}{k!}\,dt$ (keep in mind that we are differentiating u with respect to t, not x) and $v = f^{(k+1)}(t)$. Then, we obtain

$$\int_a^x \frac{f^{(k+2)}(t)}{(k+1)!}(x-t)^{k+1}\,dt = \frac{(x-t)^{k+1}}{(k+1)!}f^{(k+1)}(t)\Big|_a^x + \int_a^x f^{(k+1)}(t)\frac{(x-t)^k}{k!}\,dt$$
$$= -\frac{(x-a)^{k+1}}{(k+1)!}f^{(k+1)}(a) + R_{k,a}(x)$$

by our inductive hypothesis. To finish the proof, we first observe that $-\frac{(x-a)^{k+1}}{(k+1)!}f^{(k+1)}(a) = P_{k,a}(x)-P_{k+1,a}(x)$. Then, we see that $f(x) = P_{k,a}(x)+R_{k,a}(x) = P_{k+1,a}(x)+R_{k+1,a}(x)$, or $P_{k,a}(x)-P_{k+1,a}(x) = R_{k+1,a}(x)-R_{k,a}(x)$, thus $-\frac{(x-a)^{k+1}}{(k+1)!}f^{(k+1)}(a) = R_{k+1,a}(x)-R_{k,a}(x)$. Then,

$$\int_a^x \frac{f^{(k+2)}(t)}{(k+1)!}(x-t)^{k+1}\,dt = R_{k+1,a}(x) - R_{k,a}(x) + R_{k,a}(x) = R_{k+1,a}(x),$$

finishing the inductive step. □

As $R_{n,a}(x) = \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n\,dt$ is the "error" between the function and its Taylor Polynomial, we want the bounds on this integral to be made comparatively small.

Theorem 7.17 (Taylor's Theorem)

If f is a function for which $P_{n,a}(x)$ exists, then for some $c \in [a,x]$, $R_{n,a}(x) = \dfrac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$.

Definition 7.18. The form of the remainder presented in Theorem 7.17 is known as the **Lagrange form** of the error.

Proof. By HT3, suppose $m \le f^{(n+1)}(t) \le M\ \forall t \in [a,x]$. Then by Theorem 7.16,

$$\int_a^x \frac{m}{n!}(x-t)^n\,dt \le R_{n,a}(x) \le \int_a^x \frac{M}{n!}(x-t)^n\,dt.$$

Evaluating both sides of the compound inequality yields

$$\frac{m(x-a)^{n+1}}{(n+1)!} \le R_{n,a}(x) \le \frac{M(x-a)^{n+1}}{(n+1)!}.$$

By the Intermediate Value Theorem (note that these functions are differentiable, and therefore continuous), $\exists c \in [a,x]$ such that $f^{(n+1)}(c) \in [m,M]$. Thus, we must have

$$\frac{m(x-a)^{n+1}}{(n+1)!} \le \frac{f^{(n+1)}(c)(x-a)^{n+1}}{(n+1)!} \le \frac{M(x-a)^{n+1}}{(n+1)!},$$

and so $R_{n,a} = \dfrac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$ for some $c \in [a,x]$. □

First, we look at the following problem to aid us in proving the next lemma.

Problem 7.19. Compute $\lim\limits_{n\to\infty} \frac{x^n}{n!}$.

Solution. Let $f(n) = \frac{x^n}{n!}$, and let $N = \lceil x \rceil$. We can write the function in a recursive fashion: $f(n+1) = \frac{x}{n+1}f(n)$. Confirm that this relationship is true. Then, $f(N+1) = \frac{x}{N+1}f(N)$.

For $f(N+2)$, notice that $f(N+2) = \frac{x}{N+2}f(N+1) = \frac{x}{N+2}\cdot\frac{x}{N+1}f(N) < \left(\frac{x}{N+1}\right)^2 f(N)$. We can inductively show that $f(N+k) < \left(\frac{x}{N+1}\right)^k f(N)$ for all $k \in \mathbb{N}$. As $k \to \infty$, noting that $\frac{x}{N+1} < 1$, we can conclude that $\left(\frac{x}{N+1}\right)^k$ goes to 0. Thus, $f(N+k)$ goes to 0. Since $(N+k) \to \infty$ as $k \to \infty$, we have computed the limit as $n \to \infty$ of $f(n)$ to be $\boxed{0}$. □

Lemma 7.20

$\lim\limits_{n\to\infty} R_{n,a}(x) = 0.$

Proof. By Theorem 7.17, note that $R_{n,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$ and $\lim\limits_{n\to\infty} \frac{(x-a)^{n+1}}{(n+1)!} = 0$ by Problem 7.19. Thus, the overall limit goes to 0. □

This lemma establishes that we can make the remainder, or error, as small as we want by taking sufficiently many terms. Let's consider $\sin x$ as an example. Note that

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + R_{2n+2,0}(x).$$

We could also use $R_{2n+1,0}(x)$, since $P_{2n+1,0}(x) = P_{2n+2,0}(x)$. In this example, we will use $R_{2n+2,0}(x)$.

From Theorem 7.17, we have that $R_{2n+2,0}(x) = \frac{\sin^{(2n+3)}(c)}{(2n+3)!}x^{2n+3}$. Now we make one important observation: any derivative of $\sin x$ is bounded by -1 and 1, i.e. $\left|\sin^{(2n+3)}(c)\right| \leq 1$. This indicates that $|R_{2n+2,0}(x)| \leq \frac{|x|^{2n+3}}{(2n+3)!}$. If we wanted to estimate $\sin 1$ up to $n = 1$, for instance, we would have

$$\sin 1 = 1 - \frac{1}{6} + R_{4,0}(x) = \frac{5}{6} + R_{4,0}(x).$$

Then, we know that $|R_{4,0}(x)| \leq \frac{1}{5!} = \frac{1}{120}$, thus we can approximate

$$\frac{5}{6} - \frac{1}{120} \leq \sin 1 \leq \frac{5}{6} + \frac{1}{120},$$

or

$$\frac{99}{120} \leq \sin 1 \leq \frac{101}{120},$$

which can be verified using a calculator.

Say we wanted to estimate the remainder for e^x. We have

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + R_{n,0}(x).$$

By Theorem 7.17, $R_{n,0}(x) = \frac{e^c}{(n+1)!}x^{n+1}$. How do we deal with $c \in [0,x]$? Again, we can derive an appropriate bound by considering the maximum value of e^x on $[0,x]$, which is just e^x as it is exponential. Then, we have $R_{n,0}(x) \leq \frac{e^x}{(n+1)!}x^{n+1}$. Now that we have already determined $e < 3$, we can assert that $R_{n,0}(x) < \frac{3^x}{(n+1)!}x^{n+1}$. While this can be made as small as we want by taking n sufficiently large compared to x, it can be difficult to determine such n.

We can simplify this problem by only considering $0 \le x \le 1$. Then, we have $3^x \le 3$ and $x^{n+1} \le 1$ for all such x, and thus $R_{n,0}(x) < \frac{3}{(n+1)!}$.

Theorem 7.21

e is irrational.

Proof. Assume for the sake of contradiction that e is rational, and let it be equal to $\frac{m}{n}$ in lowest terms. As we have shown that $1 < e < 3$, we know that

$$1+1+\frac{1}{2!}+\frac{1}{3!}+\ldots+\frac{1}{n!}+\frac{1}{(n+1)!} < e < 1+1+\frac{1}{2!}+\frac{1}{3!}+\ldots+\frac{1}{n!}+\frac{3}{(n+1)!}.$$

We multiply everything by $n!$, and let $k = n!\left(1+1+\frac{1}{2!}+\frac{1}{3!}+\ldots+\frac{1}{n!}\right)$. Then, we have

$$k+\frac{1}{n+1} < n!e < k+\frac{3}{n+1}.$$

Subtracting k gives

$$\frac{1}{n+1} < n!e-k < \frac{3}{n+1}.$$

Here, $n!e-k$ is also an integer. It is necessarily true that $n \ge 1$. This implies $\frac{2}{n+1} \le 1$, i.e. $\frac{3}{n+1}-\frac{1}{n+1} \le 1$. Thus, we are trying to fit the integer $n!e-k$ in an interval $\left(\frac{1}{n+1}, \frac{3}{n+1}\right)$ that has length at most 1, which is impossible. Hence, we have a contradiction, and e must be irrational. ☐

Now, we return to Taylor Polynomials. It turns out that we have two important properties. Let $f(x), g(x)$ be functions. Then,

1. $P_{n,a,f+g}(x) = P_{n,a,f}(x) + P_{n,a,g}(x)$.
2. $P_{n,a,fg}(x) = [P_{n,a,f}(x) \cdot P_{n,a,g}(x)]_n$, where $[P]_n$ denotes the **truncation** of P to degree n. In other words, we only take the sum of all terms of P that have degree $\le n$.

Exercise 7.22. Let $f(x) = \sin x$ and $g(x) = \cos x$. Find $P_{n,0,f+g}(x)$.

Problem 7.23. Let $f(x) = e^x \sin x$. Find $P_{4,0,f}(x)$.

Solution. Since $f(x)$ is the product of functions e^x and $\sin x$, we can simply take the truncation of the product of the Taylor Polynomials for e^x and $\sin x$ respectively, up to degree 4. We have

$$\left[\left(1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}\right)\left(x-\frac{x^3}{3!}\right)\right]_4 = x+x^2+\frac{x^3}{2!}+\frac{x^4}{3!}-\frac{x^3}{3!}-\frac{x^4}{3!} = \boxed{x+x^2+\frac{1}{3}x^3},$$

where we ignore all terms with degrees higher than 4. ☐

It is also true that if we want to find Taylor Polynomials for e^{2x} or $e^{(x^3)}$, we can simply substitute x for $2x$ or x^3 into the Taylor Polynomial for e^x to obtain our desired answers.

Problem 7.24. Find the 6th order Taylor Polynomial for $3\sin(x^2) + 2\cos(x^3)$.

Solution. Since the Taylor Polynomial for $\sin x$ is $x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots$, the Taylor Polynomial for $\sin(x^2)$ is $x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \ldots$. Then, the Taylor Polynomial for $3\sin(x^2)$ would be $3x^2 - \frac{3x^6}{3!} + \frac{3x^{10}}{5!} - \ldots$. However, we are only concerned up to the 6th order, so we stop at terms with degrees less than or equal to 6. This gives us $3x^2 - \frac{3x^6}{3!}$ for $3\sin(x^2)$. Similarly, we can find $2 - \frac{2x^6}{2!}$ for $2\cos(x^3)$, and thus our overall 6th order Taylor Polynomial is $\boxed{-\frac{3}{2}x^6 + 3x^2 + 2}$. □

Example 7.25

Approximate $\displaystyle\int_0^1 e^{(x^3)}\,dx$ to within 0.01.

Solution. By Theorem 7.17,

$$e^{(x^3)} = 1 + x^3 + \frac{x^6}{2!} + \frac{x^9}{3!} + \ldots + \frac{x^{3n}}{n!} + \frac{f^{(n+1)}(c)}{(n+1)!}x^{3n+3},$$

for some $c \in [0, x]$. On the interval $[0, 1]$, we see that the maximum value of $f^{(n+1)}(c)$ on $[0, 1]$ is e, and we know that $e < 3$. Thus, we can bound the error, further noting that $x \leq 1$:

$$\frac{f^{(n+1)}(c)}{(n+1)!}x^{3n+3} < \frac{3x^{3n+3}}{(n+1)!} \leq \frac{3}{(n+1)!}.$$

We want this to be less than or equal to 0.01, i.e. $\frac{3}{(n+1)!} \leq 10^{-2}$. The smallest such value of n to satisfy this is 5, so it is sufficient to list out the Taylor Polynomial of $e^{(x^3)}$ to degree 15, i.e.

$$\int_0^1 1 + x^3 + \frac{x^6}{2!} + \frac{x^9}{3!} + \frac{x^{12}}{4!} + \frac{x^{15}}{5!}\,dx = 1 + \frac{1}{4} + \frac{1}{14} + \frac{1}{60} + \frac{1}{312} + \frac{1}{1800}.$$ □

Problem 7.26. Approximate $\sin(0.1)$ to within 0.0001.

Solution. Assuming that the center is $x = 0$, we have $\sin x = P_{2n+1,0}(x) + R_{2n+1,0}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^n\frac{x^{2n+1}}{(2n+1)!} + \frac{f^{(2n+2)}(c)}{(2n+2)!}x^{2n+2}$. Since $\left|f^{(2n+2)}(c)\right| \leq 1$, we have $\left|\frac{f^{(2n+2)}(c)}{(2n+2)!}x^{2n+2}\right| \leq \frac{|x|^{2n+2}}{(2n+2)!}$. The maximum value of this over the interval $[0, 0.1]$ is $\frac{(0.1)^{2n+2}}{(2n+2)!}$, and we want this to be less than or equal to 0.0001, i.e. we want to find the least n such that $\frac{(0.1)^{2n+2}}{(2n+2)!} \leq 0.0001$. This rearranges to $10^{2-2n} \leq (2n+2)!$, and it is easy to see that $n = 1$ works. Thus, we can estimate $\sin(0.1) = 0.1 - \frac{(0.1)^3}{3!} \approx 0.0998333\ldots$, which we can confirm to be within 0.0001 of the exact value $0.0998334166\ldots$ using a calculator. □

Problem 7.27. Approximate $\displaystyle\int_0^{0.1} \sin(x^2)\,dx$ within 0.0001.

Solution. We have

$$\sin(x^2) = x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \ldots + \frac{(-1)^n x^{4n+2}}{(2n+1)!} + \frac{f^{(2n+2)}(c)}{(2n+2)!}x^{4n+4}$$

for some $c \in [0, 0.1]$. Since $\left|f^{(2n+2)}(c)\right| \leq 1$ (any derivative of sin is bounded by -1 and 1), it follows that

$$\left|\frac{f^{(2n+2)}(c)}{(2n+2)!}x^{4n+4}\right| \leq \frac{|x|^{4n+4}}{(2n+2)!} \leq \frac{(0.1)^{4n+4}}{(2n+2)!},$$

and we want this to be less than 0.0001. In fact, $n = 0$ is the least, sufficient n that satisfies this, so we can estimate $\sin(x^2)$ as x^2 and obtain $\int_0^{0.1} x^2\,dx = \frac{1}{3}(0.1)^3 = 3.333\ldots \times 10^{-4}$, which we can confirm to be within 0.0001 of the actual value 3.3333095×10^{-4} using a calculator. $\square$

Exercise 7.28. Estimate $\int_0^{0.1} \frac{\sin x}{x}\,dx$ to within 0.0001.

Exercise 7.29. Estimate $e^{-0.1}$ to the nearest thousandth.

Problem 7.30. Compute $\displaystyle\lim_{x\to 0} \frac{\sin(x^3)}{xe^{(x^2)} - x}$.

Solution. Applying L'Hôpital's Rule would be quite messy here, so we can use an alternative method using Taylor Polynomials.

The Taylor Polynomial for $e^{(x^2)}$ is $1 + x^2 + \frac{x^4}{2!} + \ldots$, and therefore the Taylor Polynomial for $xe^{(x^2)}$ is $x + x^3 + \frac{x^5}{2!} + \ldots$ From here, the Taylor Polynomial for $xe^{(x^2)} - x$ is $x^3 + \frac{x^5}{2!} + \ldots$, and furthermore, the Taylor Polynomial for $\sin(x^3)$ is $x^3 - \frac{x^9}{3!} + \ldots$

After we have found Taylor Polynomials for the numerator and denominator, we can see that

$$\lim_{x\to 0} \frac{\sin(x^3)}{xe^{(x^2)} - x} = \lim_{x\to 0} \frac{x^3 - \frac{x^9}{3!} + \ldots}{x^3 + \frac{x^5}{2!} + \ldots}.$$

We divide the numerator and denominator by x^3 to get

$$\lim_{x\to 0} \frac{1 - \frac{x^6}{3!} + \ldots}{1 + \frac{x^2}{2!} + \ldots} = \boxed{1}.$$

$\square$

Theorem 7.31 (Trapezoid Rule)

Consider $f(x)$ from $x = a$ to b. Partition $a = x_0 < x_1 < \ldots < x_n = b$ into equal length subintervals, each of length Δx. Then the estimated area of the integral from a to b is

$$T_n = \left(\frac{1}{2}f(x_0) + \sum_{i=1}^{n-1} f(x_i) + \frac{1}{2}f(x_n)\right)\Delta x.$$

Consider the function $f(x) = (1+x)^\alpha$, where $\alpha \in \mathbb{R}$. To calculate $P_{n,0}(x)$, we can list out the first few derivatives evaluated at 0:

$$\begin{aligned}
f(0) &= 1 \\
f'(x) &= \alpha(1+x)^{\alpha-1} \\
f'(0) &= \alpha \\
f''(x) &= \alpha(\alpha-1)(1+x)^{\alpha-2} \\
f''(0) &= \alpha(\alpha-1) \\
f'''(x) &= \alpha(\alpha-1)(\alpha-2)(1+x)^{\alpha-3} \\
f'''(0) &= \alpha(\alpha-1)(\alpha-2)
\end{aligned}$$

$$\vdots$$

From here, we can see that the general pattern is that $f^{(k)}(0) = \alpha(\alpha-1)(\alpha-2)\cdots(\alpha-(k-1))$. Then,

$$P_{n,0}(x) = \sum_{k=0}^{n} \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-(k-1))}{k!} x^k.$$

In fact, $\frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-(k-1))}{k!}$ reminds us of $\frac{\alpha!}{k!(\alpha-k)!} = \binom{\alpha}{k}$. We had initially assumed that $\binom{\alpha}{k}$ was only defined for $\alpha, k \in \mathbb{Z}$, but we could actually extend this definition to cover all real α, such that $\binom{\alpha}{k} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-(k-1))}{k!}$. Then, our Taylor Polynomial is then

$$\sum_{k=0}^{n} \binom{\alpha}{k} x^k.$$

Then, how would we compute a term like $\binom{1/2}{2}$? We use a recursive definition:

$$\frac{\binom{\alpha}{k+1}}{\binom{\alpha}{k}} = \frac{\alpha-k}{k+1} \implies \binom{\alpha}{k+1} = \frac{\alpha-k}{k+1}\binom{\alpha}{k}.$$

Then, $\binom{1/2}{2} = \frac{1/2-1}{1+1}\binom{1/2}{1} = -\frac{1}{8}$. For $f(x) = \sqrt{1+x}$, our Taylor Polynomial centered at $x = 0$ would be

$$\begin{aligned}\binom{1/2}{0} + \binom{1/2}{1}x + \binom{1/2}{2}x^2 + \binom{1/2}{3}x^3 + \binom{1/2}{4}x^4 + \dots \\ = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 + \dots\end{aligned}$$

From this, for instance, we are able to estimate $\sqrt{1.1}$ as approximately $1 + 0.05 - 0.00125 = 1.04875$.

Overall, you should be familiar with the following:

§7.3 Analyzing Sequences

Definition 7.32. A **sequence** is a function whose domain is $\mathbb{N}$.

Recall the following fact: A sequence $\{a_n\}$ **converges** to L if $\lim_{n\to\infty} a_n = L$, i.e.

$$\forall \varepsilon > 0\ \exists N \text{ s.t. } \forall n > N,\ |a_n - L| < \varepsilon.$$

For a sequence $\{a_n\}$, let $a_n = f(n)$ for $n \in \mathbb{N}$. Then, $\lim\limits_{x\to\infty} f(x) = \lim\limits_{n\to\infty} a_n$. However, this is not entirely true. Suppose $f(x) = \sin(\pi x)$. We know that $\lim\limits_{x\to\infty} f(x)$ does not exist, since it is an oscillating function, but we can see that a sequence can be generated: $a_1 = 0,\ a_2 = 0,\ a_3 = 0, \ldots$

The issue is that $\lim\limits_{x\to\infty} f(x)$ does not exist, therefore we can fix the assertion as such:

Theorem 7.33

For $a_n = f(n)\ \forall n \in \mathbb{N}$, as long as $\lim\limits_{x\to\infty} f(x)$ exists, $\lim\limits_{x\to\infty} f(x) = \lim\limits_{n\to\infty} a_n$.

Proof. Recall that the definition of $\lim\limits_{x\to\infty} f(x)$ is

$$\forall \varepsilon > 0\ \exists N\ \forall x > N,\ |f(x) - L| < \varepsilon.$$

Simply let $x = n$, and since $a_n = f(n)$, we have

$$\forall \varepsilon > 0\ \exists N\ \forall n > N,\ |a_n - L| < \varepsilon,$$

which is the definition of $\lim\limits_{n\to\infty} a_n$. □

Theorem 7.34

If $\{a_n\}$ is non-decreasing and bounded above, then it converges.

Proof. Let $L = \sup(\{a_n\})$. By definition of supremum, we have

$$\forall \varepsilon > 0,\ \exists N \text{ s.t. } L - a_N < \varepsilon.$$

As $\{a_n\}$ is non-decreasing and bounded above, $\forall n > N,\ a_n \geq a_N$. Then, $L - a_n \leq L - a_N < \varepsilon$, so we have

$$\forall \varepsilon > 0,\ \exists N\ \text{ s.t. } \forall n > N,\ L - a_n < \varepsilon,$$

so it converges. □

Theorem 7.35

If $\{a_n\}$ is non-increasing and bounded below, then it converges.

The proof is nearly identical to that of the previous theorem.

Theorem 7.36

Let $\{a_i\}$ be a bounded sequence. Then there exists a convergent subsequence of $\{a_i\}$.

Proof. Call a_N a **peak point** of the sequence if $a_N \geq a_n$ $\forall n > N$. We proceed by casework.

1. Case 1: There exists an infinite number of peak points. Then, these form a subsequence, i.e.
$$a_{N_1}, a_{N_2}, a_{N_3}, \ldots$$
This is non-increasing and bounded, so by Theorem 7.35, it is a convergent subsequence.

2. Case 2: There are a finite number of peak points. Let a_N be the last peak point. Consider $n_1 > N$. We know a_{n_1} is not a peak point, so $\exists n_2 > n_1$ such that $a_{n_1} < a_{n_2}$. We repeat this process for n_3, n_4, etc. and we end up with a necessarily decreasing subsequence, which must converge by Theorem 7.34. □

Definition 7.37. A sequence $\{a_n\}$ is called a **Cauchy sequence** if
$$\forall \varepsilon > 0\ \exists N\ \forall m, n > N,\ |a_m - a_n| < \varepsilon.$$

Theorem 7.38

A sequence is Cauchy if and only if it converges.

Proof. The left direction is easier. Let L be the limit of the convergent sequence. Then,
$$\forall \varepsilon > 0\ \exists N \text{ s.t. } \forall n > N,\ |a_n - L| < \frac{\varepsilon}{2},$$
$$\forall m > N,\ |a_m - L| < \frac{\varepsilon}{2}.$$

Then, $\forall \varepsilon > 0\ \exists N$ s.t. $\forall m, n > N, |a_m - a_n| \leq |a_m - L| + |L - a_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, by the Triangle Inequality.

For the right direction, first let $\varepsilon = 1$. Then $\exists N$ such that $\forall m > N$, $|a_m - a_n| < 1$, or $-1 + a_N < a_m < 1 + a_N$. Our sequence is bounded by the minimum and maximum of
$$\{a_1, a_2, \ldots, a_{N-1}, -1 + a_N, 1 + a_N\}.$$

By Theorem 7.36, $\{a_n\}$ has a convergent subsequence. Let this subsequence converge to L. Then,
$$\forall \varepsilon > 0\ \exists N_1 \text{ s.t. } \forall n > N_1,\ |a_n - L| < \frac{\varepsilon}{2}.$$

Furthermore, $\exists N_2$ such that $\forall m, n > N_2$, $|a_n - a_m| < \frac{\varepsilon}{2}$, by definition of a Cauchy sequence. Let $N = \max\{N_1, N_2\}$. Then,
$$\forall m > N, |a_m - L| \leq |a_n - a_m| + |L - a_n| < \varepsilon,$$

so $\{a_n\}$ converges. □

Definition 7.39. An **infinite series** is an expression of the form

$$\sum_{n=1}^{\infty} a_n.$$

Definition 7.40. A series **converges** to L if its sequence of partial sums converges, where

$$S_k = \sum_{n=1}^{k} a_n, \text{ and } \lim_{k\to\infty} S_k = L.$$

For example, $S_1 = a_1$, $S_2 = a_1 + a_2$, $S_3 = a_1 + a_2 + a_3$, and so on.

First, we address the first special kind of series:

> **Theorem 7.41** (Convergent Geometric Series)
> Define the geometric sequence $a_n = ar^{n-1}$. Let $S_k = a + ar + ar^2 + \ldots + ar^{k-1} = \dfrac{a(1-r^k)}{1-r}$. Then, $\{a_n\}$ converges to $\dfrac{a}{1-r}$ if and only if $|r| < 1$.

An important note: whether we have $\sum\limits_{n=7}^{\infty} a_n$ or $\sum\limits_{n=10^{100}}^{\infty} a_n$, if one of these converges, then the other must necessarily converge because they differ by a *finite number of terms*. Thus, even though we write $\sum\limits_{n=1}^{\infty} a_n$ in our theorem statements below, it does not matter whether n starts at 1, 7, 1324821, etc.

Then, we will write $\sum a_n$ as shorthand for n that can start at any positive integer (usually 1), going to ∞.

> **Theorem 7.42** (Test for Divergence)
> If $\lim\limits_{n\to\infty} a_n \neq 0$, then $\sum a_n$ diverges.

Proof. We will prove the contrapositive of the statement, i.e. if $\sum a_n$ converges, then $\lim\limits_{n\to\infty} a_n = 0$.

Let $S_n = \sum\limits_{k=1}^{n} a_k$, and suppose the series converges to S, i.e. $\lim\limits_{n\to\infty} S_n = S$.

Note $\lim\limits_{n\to\infty} S_{n-1} = S$ as well (since we are considering n going to ∞ anyway). Then, $\lim\limits_{n\to\infty} S_n - S_{n-1} = 0$, but $S_n - S_{n-1} = a_n$, thus $\lim\limits_{n\to\infty} a_n = 0$. □

Problem 7.43. Using the Test for Divergence, determine whether the sequences $\sum(-1)^n n$ and $\sum \frac{1}{n^2}$ converge or diverge.

Solution. As $\lim\limits_{n\to\infty}(-1)^n n \neq 0$, so the series diverges. For $\sum \frac{1}{n^2}$, we have $\lim\limits_{n\to\infty} \frac{1}{n^2} = 0$. However, from just the Test for Divergence, we cannot determine whether it converges or diverges (avoid committing the inverse or converse error!). □

> **Theorem 7.44** (Comparison Test)
> Suppose $0 \leq a_n \leq b_n\ \forall n$. If $\sum b_n$ converges, then $\sum a_n$ converges.

Proof. Let $\sum\limits_{k=1}^{\infty} a_k = S_n$. By $0 \leq a_n\ \forall n$, $\{S_n\}$ non-decreasing and by $a_n \leq b_n$, it is also bounded above by $\sum\limits_{k=1}^{\infty} b_k$. By Theorem 7.34, we conclude that $\{S_n\}$, and thereby $\{a_n\}$, converge. □

Theorem 7.45 (Integral Test)
Suppose $f(x) \geq 0$ is an eventually decreasing function for which $\lim\limits_{n\to\infty} f(x) = 0$. Define $a_n = f(n)\ \forall n \in \mathbb{N}$. Then, $\sum a_n$ converges if and only if $\displaystyle\int_1^{\infty} f(x)\,dx$ converges.

Proof. Let $S_n = \sum\limits_{k=1}^{n} a_k = \sum\limits_{k=1}^{n} f(k)$. As $f(x) \geq 0$, S_n is monotonically increasing and bounded below. Then, by Theorem 7.35, $\{S_n\}$ converges.

For the right direction, let $t_n = \int_1^n f(x)\,dx$. We know that $\{t_n\}$ is increasing. Then, as S_n consists of the upper sums, $s_n \geq t_n\ \forall n$. Thus, $\{t_n\}$ is bounded above by $\{S_n\}$. By Theorem 7.34, $\{t_n\}$ converges, i.e. $\int_1^n f(x)\,dx$ converges.

The proof for the left direction is nearly identical - we just need to consider the lower sums. □

Exercise 7.46. Does $\sum\limits_{n=1}^{\infty} \frac{n}{e^n}$ converge or diverge?

Theorem 7.47 (p-Series Test)
$\sum\limits_{n=1}^{\infty} \dfrac{1}{n^p}$ converges if and only if $p > 1$.

Proof. Recall that $\int_1^{\infty} \frac{1}{x^p}\,dx$ converges if and only if $p > 1$. The result follows by applying the Integral Test. □

For example, $\sum \frac{1}{n^2}$ converges, but $\sum \frac{1}{n}$ diverges. In fact, $\sum \frac{1}{n}$, or $1 + \frac{1}{2} + \frac{1}{3} + \ldots$, has a special name: the **harmonic series**.

As a brief digression, let's probe further into some special characteristics of the harmonic series. Consider the graph of $y = \frac{1}{x}$ starting from $x = 1$. Since it is decreasing, the right sum will be a lower sum and the left sum will be an upper sum. Up to some $n \in \mathbb{N}$, the right sum is

$$\frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n},$$

and the left sum is

$$1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1}.$$

Clearly the integral is bounded by the lower and upper sums, so we have

$$\frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} < \int_1^n \frac{1}{x}\,dx = \log n < 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1}.$$

Now, let's see how $\log n$ compares to the harmonic series $1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}$. Let

$$S_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} - \log n.$$

From the fact that $\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}<\log n$, we have $1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}-\log n<1$. Furthermore, $\log n<1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n-1}$ tells us that $1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}-\log n>\frac{1}{n}$. Therefore, all S_n terms are positive and less than 1. But does it converge?

Since we know that S_n is bounded, we should figure out whether S_n is increasing or decreasing. Note that

$$S_{n+1}-S_n=\frac{1}{n+1}-\log(n+1)+\log n=\frac{1}{n+1}-\log\left(\frac{n+1}{n}\right).$$

We need to see whether this is positive or negative. In other words, we need to compare $\frac{1}{n+1}$ and $\log\left(\frac{n+1}{n}\right)$. This is equivalent to comparing $e^{\frac{1}{n+1}}$ and $\frac{n+1}{n}$ (from raising both as exponents of e). Let $z=\frac{1}{n+1}$. Then we are comparing e^z and $\frac{1}{1-z}$. Note that as $n\to\infty$, $z\in(0,1)$. Then, we are allowed to compare $(1-z)e^z$ and 1 (from multiplying by $1-z$), which rearranges to $1-z$ vs. e^{-z}. Now, we make use of a lemma:

Lemma

$\forall x$, $e^x\geq 1+x$ with equality only if $x=0$.

Proof. Let $g(x)=e^x-x-1$. Then $g'(x)=e^x-1$, so the critical point is $x=0$. After using a first derivative number line, we see that $x=0$ is the global minimum. Thus, $g(x)$ is always nonnegative, so $e^x-x-1\geq 0$, i.e. $e^x\geq x+1$, as desired. □

Hence, we finally conclude that $e^{-z}\geq 1-z\ \forall z$. Now, we can trace backwards from what we have been doing all along:

$$\begin{aligned}
e^{-z}&\geq 1-z\\
1&\geq(1-z)e^z\\
\frac{1}{1-z}&\geq e^z\quad\text{since } 0<z<1\implies 0<1-z<1\\
\frac{n+1}{n}&\geq e^{\frac{1}{n+1}}\ \text{ where } z=\frac{1}{n+1}\\
\log\left(\frac{n+1}{n}\right)&\geq\frac{1}{n+1}\ \text{ since } \log \text{ is increasing.}
\end{aligned}$$

Finally, we may conclude that $S_{n+1}-S_n=\dfrac{1}{n+1}-\log\left(\dfrac{n+1}{n}\right)\leq 0$, and thus $\{S_n\}$ is nonincreasing. Combining this with the fact that $\{S_n\}$ is bounded below by 0, Theorem 7.35 states that $\{S_n\}$ must converge. It turns out that

$$\lim_{n\to\infty}S_n=\gamma\approx 0.5772156649\ldots$$

where γ is known as **Euler's constant**.

Before we move on to more convergence tests, here are a few review problems:

Problem 7.48. Determine whether $\sum\frac{\log n}{n}$ converges or diverges.

Proof. For $n\geq e$, $0\leq 1\leq\log n$, or $0\leq\frac{1}{n}\leq\frac{\log n}{n}$. We know that $\sum\frac{1}{n}$ diverges, so by the contrapositive of the Comparison Test, $\sum\frac{\log n}{n}$ must also diverge. □

Problem 7.49. Determine whether $\displaystyle\sum_{n=10^{10^{100}}}^{\infty}\frac{1}{n+10^{100}}$ converges or diverges.

Proof. We can ignore n starting at $10^{10^{100}}$. We can guess that this series will probably diverge, because as n goes to ∞, the $+10^{100}$ becomes negligible, leaving us with $\frac{1}{n}$ which is the divergent harmonic series. Indeed, we can simply reindex the summation, as such:

$$\sum_{n=10^{10^{100}}}^{\infty} \frac{1}{n+10^{100}} = \sum_{n=10^{10^{100}}+10^{100}}^{\infty} \frac{1}{n},$$

from which we conclude that it diverges. □

Problem 7.50. Determine whether $\sum \frac{1}{n^2-88}$ converges or diverges.

Proof. Looking at this series, we consider $\sum \frac{1}{n^2}$ as a possible series to compare it to. However, $\frac{1}{n^2-88}$ is always greater than $\frac{1}{n^2}$ for all positive n. We can simply fix this by seeing that for n sufficiently large (in particular, when $n \geq \sqrt{176}$),

$$n^2 - 88 \geq \frac{n^2}{2} \implies \frac{1}{n^2-88} \leq \frac{2}{n^2},$$

and $\sum \frac{2}{n^2}$ is simply 2 times the convergent series $\sum \frac{1}{n^2}$ (by p-Series Test), so by the Comparison Test, $\sum \frac{1}{n^2-88}$ must converge. □

Theorem 7.51 (Limit Comparison Test)

Let $\sum a_n$, $\sum b_n$ be series of non-negative terms such that $\lim\limits_{n\to\infty} \frac{b_n}{a_n} = L > 0$. Then $\sum a_n$ converges if and only if $\sum b_n$ converges (in other words, both series either both diverge or both converge).

If $L = 0$, then the theorem is only true for the right direction: if $\sum a_n$ converges, then $\sum b_n$ converges.

Proof. Take $0 < m < L < M$. Since $L = \lim\limits_{n\to\infty} \frac{b_n}{a_n}$, we see that for sufficiently large n, $\frac{b_n}{a_n}$ must be very close to L, and thereby between m and M. Hence, $\exists N$ such that $\forall n > N$, $m < \frac{b_n}{a_n} < M$. This rearranges to $ma_n < b_n < Ma_n$.

Left direction: If $\sum b_n$ converges, then by Comparison Test, $\sum ma_n$ also converges, i.e. $\sum a_n$ converges.

Right direction: If $\sum a_n$ converges, then $\sum Ma_n$ converges. By Comparison Test, we conclude that $\sum b_n$ converges. □

Problem 7.52. Use Theorem 7.51 to solve Problem 7.50.

Solution. Note that

$$\lim_{n\to\infty} \frac{\frac{1}{n^2}}{\frac{1}{n^2-88}} = \lim_{n\to\infty} \frac{n^2-88}{n^2} = \lim_{n\to\infty} 1 - \frac{88}{n^2} = 1 > 0,$$

and since $\sum \frac{1}{n^2}$ converges, we conclude that $\sum \frac{1}{n^2-88}$ must also converge by the Limit Comparison Test. □

Problem 7.53. Determine whether $\sum \frac{\sqrt[3]{n^2+1}}{\sqrt[4]{n^{17}-n^{16}+204}}$ converges or diverges.

Solution. Again, we have

$$\lim_{n\to\infty} \frac{\frac{\sqrt[3]{n^2}}{\sqrt[4]{n^{17}}}}{\frac{\sqrt[3]{n^2+1}}{\sqrt[4]{n^{17}-n^{16}+204}}} = \lim_{n\to\infty} \frac{\sqrt[4]{n^{17}-n^{16}+204}}{\sqrt[4]{n^{17}}} \cdot \frac{\sqrt[3]{n^2}}{\sqrt[3]{n^2+1}}$$

$$= \lim_{n\to\infty} \sqrt[4]{1-\frac{1}{n}+\frac{204}{n^{17}}} \cdot \sqrt[3]{\frac{1}{1+\frac{1}{n^2}}} = 1 > 0.$$

Note that $\sum \frac{\sqrt[3]{n^2}}{\sqrt[4]{n^{17}}} = \sum \frac{1}{n^{43/12}}$ is a p-Series with $p = \frac{43}{12} > 1$. By the p-Series Test, $\sum \frac{1}{n^{43/12}}$ converges, so $\sum \frac{\sqrt[3]{n^2+1}}{\sqrt[4]{n^{17}-n^{16}+204}}$ converges by the Limit Comparison Test. □

Theorem 7.54 (Ratio Test)

Suppose $\sum\limits_{n=1}^{\infty} a_n$ is a series of non-negative terms and suppose $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} = R$. Then,

a) If $R < 1$, then $\sum\limits_{n=1}^{\infty} a_n$ converges.

b) If $R > 1$, then $\sum\limits_{n=1}^{\infty} a_n$ diverges.

c) If $R = 1$, then the test is inconclusive.

Proof. For $R = 1$, we can consider $\sum \frac{1}{n}$ and $\sum \frac{1}{n^2}$, which diverge and converge respectively. Hence, it is inconclusive.

For $R < 1$, take $R < \widetilde{R} < 1$. Similar to what we did in proving the Limit Comparison Test, we observe that as n grows large, $\frac{a_{n+1}}{a_n}$ gets closer to R (and thereby less than $\widetilde{R}$). Then, $\exists N$ such that $\forall n > N$, $\frac{a_{n+1}}{a_n} < \widetilde{R}$, i.e. $a_{n+1} < \widetilde{R}a_n$. This holds for $\frac{a_{N+1}}{a_N}$, $\frac{a_{N+2}}{a_{N+1}}$, $\frac{a_{N+3}}{a_{N+2}}$, etc. from which we get

$$\begin{aligned} a_{N+1} &< \widetilde{R}a_N \\ a_{N+2} &< \widetilde{R}a_{N+1} < \widetilde{R}^2 a_N \\ a_{N+3} &< \widetilde{R}a_{N+2} < \widetilde{R}^3 a_N \\ &\vdots \end{aligned}$$

We see that $a_{N+k} < \widetilde{R}^k a_N$ for $k \in \mathbb{N}$. As $\widetilde{R} < 1$, $\widetilde{R}^k a_N$ is a convergent geometric series. By the Comparison Test, $\{a_n\}$ will eventually converge after a_N.

We can similarly prove the case for $R > 1$ by letting $\widetilde{R} > R > 1$, and comparing our series (using contrapositive of Comparison Test) with a geometric series with common ratio $\widetilde{R} > 1$, which is thereby divergent. □

Theorem 7.55 (Root Test)

Let $\sum\limits_{n=1}^{\infty} a_n$ be a series of non-negative terms. Suppose $\sum\limits_{n=1}^{\infty} \sqrt[n]{a_n} = L > 0$. Then,

a) If $L < 1$, then $\sum\limits_{n=1}^{\infty} a_n$ converges.

b) If $L > 1$, then $\sum\limits_{n=1}^{\infty} a_n$ diverges.

c) If $L = 1$, then the test is inconclusive.

Proof. The proof is nearly identical to that of the Ratio Test. Consider $L < 1$. Then $L < \widetilde{L} < 1$. Then, $\exists N$ such that $\forall n > N$, $\sqrt[n]{a_n} < \widetilde{L}$, i.e. $a_n < \widetilde{L}^n$. We can easily finish from here. □

Problem 7.56. Let $a > 0$. Does $\sum \frac{a^n}{n!}$ converge or diverge?

Solution. We have

$$\lim_{n\to\infty} \frac{\frac{a^{n+1}}{(n+1)!}}{\frac{a^n}{n!}} = \lim_{n\to\infty} \frac{a}{n+1} = 0 < 1,$$

so $\sum \frac{a^n}{n!}$ converges by the Ratio Test. □

However, notice that $1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \ldots + \frac{a^n}{n!}$ is the nth order Taylor Polynomial for e^a. As $\sum \frac{a^n}{n!}$ converges as $n \to \infty$, we can introduce the notion of a **Taylor Series**:

$$e^a = 1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \ldots$$

As a sidenote, you should be familiar with the following:

$$\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \ \forall x, \\
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots \ \forall x, \\
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots \ \forall x, \\
\tan^{-1} x &= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots \ |x| \le 1, \\
\log(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots \ -1 < x \le 1.
\end{aligned}$$

Problem 7.57. Find all $a \in \mathbb{R}$ for which $\sum \frac{n^{10^{100}}}{a^n}$ converges.

Solution. We have

$$\lim_{n\to\infty} \frac{\frac{(n+1)^{10^{100}}}{a^{n+1}}}{\frac{n^{10^{100}}}{a^n}} = \lim_{n\to\infty} \frac{1}{a}\left(1 + \frac{1}{n}\right)^{10^{100}} = \frac{1}{a}.$$

By the Ratio Test, if $a > 1$, then it converges, and if $a < 1$, then it diverges. For $a = 1$, note that the series simplifies to $\sum n^{10^{100}}$, for which it is obvious that it will diverges (using the Test for Divergence). □

Problem 7.58. Determine whether $\sum \frac{2^n+1}{3^n+1}$ converges or diverges.

Solution. Note that

$$\lim_{n\to\infty} \frac{\frac{2^n}{3^n}}{\frac{2^n+1}{3^n+1}} = \lim_{n\to\infty} \frac{2^n(3^n+1)}{3^n(2^n+1)} = \lim_{n\to\infty} \frac{1+\frac{1}{3^n}}{1+\frac{1}{2^n}} = 1 > 0.$$

Further note that $\sum \frac{2^n}{3^n} = \sum \left(\frac{2}{3}\right)^n$ is a convergent geometric series with common ratio $\frac{2}{3} < 1$. Thus, by Limit Comparison Test, $\sum \frac{2^n+1}{3^n+1}$ converges. □

Problem 7.59. Determine whether $\sum \frac{(1.001)^n}{n^{10^{100}}}$ converges or diverges.

Solution. We have

$$\lim_{n\to\infty} \frac{\frac{(1.001)^{n+1}}{(n+1)^{10^{100}}}}{\frac{(1.001)^n}{n^{10^{100}}}} = \lim_{n\to\infty} 1.001 \left(\frac{1}{1+\frac{1}{n}}\right)^{10^{100}} = 1.001 > 1,$$

so by the Ratio Test, it diverges. □

Problem 7.60. Determine whether $\sum \frac{\log(\log n)}{n}$ converges or diverges.

Solution. For $n > e^e$, $\log(\log n) > 1$. Then, $\frac{1}{n} < \frac{\log(\log n)}{n}$. As $\sum \frac{1}{n}$ is divergent, by Comparison Test, $\sum \frac{\log(\log n)}{n}$ must also diverge. □

Lemma 7.61

Assume sequences $\{a_n\}$, $\{b_n\}$ have positive terms. If $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = 0$, then for sufficiently large n, $a_n < b_n$.

Proof. This lemma should seem rather intuitive (eventually $\{b_n\}$ must "overwhelm" $\{a_n\}$). Let $\varepsilon = 1$. Then,

$$\exists N \text{ s.t. } \forall n > N, \ \left|\frac{a_n}{b_n}\right| < \varepsilon,$$

i.e. $\frac{a_n}{b_n} < 1$, which rearranges to $a_n < b_n$, and we're done. □

Problem 7.62. Prove that for sufficiently large n, $\log n < n^\alpha$ $\forall \alpha > 0$.

Proof. Applying L'Hôpital's Rule on $\lim\limits_{n\to\infty} \frac{\log n}{n^\alpha}$, we get

$$\lim_{n\to\infty} \frac{\log n}{n^\alpha} = \lim_{n\to\infty} \frac{\frac{1}{n}}{\alpha n^{\alpha-1}} = \lim_{n\to\infty} \frac{1}{\alpha n^\alpha} = 0,$$

so by Lemma 7.61, $\log n < n^\alpha$ eventually for large n. □

Problem 7.63. Determine whether $\sum \frac{\log n}{n^2}$ converges or diverges.

Solution. By Problem 7.62, we know that $\log n < n^{1/2}$. Then, $\frac{\log n}{n^2} < \frac{1}{n^{3/2}}$. Note that $\sum \frac{1}{n^{3/2}}$ is a p-Series with $p = \frac{3}{2} > 1$, so by the p-Series Test, it converges. Then, by Comparison Test, $\sum \frac{\log n}{n^2}$ must also converge. □

Definition 7.64. $\sum\limits_{n=1}^{\infty} a_n$ is **absolutely convergent** if $\sum\limits_{n=1}^{\infty} |a_n|$ converges.

Theorem 7.65

If a series is absolutely convergent, then it converges.

If we can show that the absolute value of some series converges, then we will have shown that the original series converges as well.

For example, we can conclude that $a_n = (-1)^{n+1}\frac{1}{n^2}$, which gives the alternating sum $1 - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + \ldots$, converges because $\sum |a_n| = \sum \frac{1}{n^2}$ converges. Similarly, we can assert that

$$1 - \frac{1}{4} - \frac{1}{9} - \frac{1}{16} + \frac{1}{25} - \frac{1}{36} - \frac{1}{49} - \frac{1}{64} + \frac{1}{81} - \ldots$$

also converges by the same argument.

Theorem 7.66

A series is absolutely convergent if and only if its positive terms converge and its negative terms converge.

Now, we introduce another significant test that helps us deal with sequences alternating in sign.

Theorem 7.67 (Leibniz's Test)

Suppose we have an alternating series

$$\sum_{n=1}^{\infty}(-1)^n a_n \text{ or } \sum_{n=1}^{\infty}(-1)^{n+1}a_n.$$

If $\{a_n\}$ is decreasing and $\lim_{n\to\infty} a_n = 0$, then the series converges.

By this test, for instance, we can see that

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \ldots$$

converges even though the absolute value of the sequence is the divergent harmonic series.

Problem 7.68. Determine whether $\sum \frac{\sin n}{n^2}$ converges or diverges.

Solution. If we can show that $\sum \left|\frac{\sin n}{n^2}\right|$ converges, then by Theorem 7.65, we will have shown that it converges.

First, note that $|\sin n| \le 1$, so $\left|\frac{\sin n}{n^2}\right| \le \frac{1}{n^2}$. Since $\sum \frac{1}{n^2}$ is a p-Series with $p = 2 > 1$, it converges by the p-Series Test. By the Comparison Test, $\sum \left|\frac{\sin n}{n^2}\right|$ must converge, and hence, $\sum \frac{\sin n}{n^2}$ converges. □

Problem 7.69. Determine whether $\sum \frac{2^n n!}{n^n}$ converges or diverges.

Solution. We have

$$\lim_{n\to\infty} \frac{\frac{2^{n+1}(n+1)!}{(n+1)^{n+1}}}{\frac{2^n n!}{n^n}} = 2\lim_{n\to\infty}\left(\frac{1}{1+\frac{1}{n}}\right)^n = 2\lim_{n\to\infty}\frac{1}{\left(1+\frac{1}{n}\right)^n} = \frac{2}{e} < 1,$$

so it converges by the Ratio Test. □

Theorem 7.70

If the sequence $\{a_n\}$ follows the conditions of the Leibniz Test (i.e. $\{a_n\}$ is decreasing and $\lim_{n\to\infty} a_n = 0$), then the error in using the first n terms to approximate the actual sum is at most the absolute value of the $(n+1)$st term.

Problem 7.71. Estimate e^{-1} to within 0.001.

Solution. If we list out the first few terms of the Taylor series, we get

$$1 - 1 + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} - \frac{1}{120} + \frac{1}{720} - \frac{1}{5040} + \dots$$

and we can easily verify that it satisfies Leibniz. Then, by Theorem 7.70, the error in using $1-1+\frac{1}{2}-\frac{1}{6}+\frac{1}{24}-\frac{1}{120}+\frac{1}{720}$ to approximate e^{-1} is at most $\frac{1}{5040}$, which is the greatest term that is less than 0.001. Thus, our estimate is $1-1+\frac{1}{2}-\frac{1}{6}+\frac{1}{24}-\frac{1}{120}+\frac{1}{720} = \boxed{\frac{53}{144}}$. $\square$

Definition 7.72. Let $\{f_n(x)\}$ be a sequence of functions and suppose for x in the desired domain, $\lim_{n\to\infty} f_n(x)$ exists. We will call this $f(x)$, i.e. $\lim_{n\to\infty} f_n(x) = f(x)$.

Definition 7.73. Suppose $\{f_n(x)\} \to f(x)$, on some domain D. We say this sequence **converges uniformly** to $f(x)$ if

$$\forall \varepsilon > 0\ \exists N\ \forall x \in D,\ \forall n > N,\ |f_n(x) - f(x)| < \varepsilon.$$

Theorem 7.74

Suppose $\{f_n(x)\} \to f(x)$ uniformly on $[a, b]$.

1. If each f_n is continuous on $[a, b]$, so is f.
2. $\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx$.

Theorem 7.75

If $\{f_n(x)\}$ converges to $f(x)$ uniformly on $[a, b]$, and if each of $f_n(x)$ $\forall n$ are integrable on $[a, b]$, then

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx.$$

Theorem 7.76

Let $\{f_n\}$ uniformly converge to $f(x)$ on $[a, b]$. If f_n is continuous on $[a, b]$ $\forall n$, so is $f(x)$.

Definition 7.77. The series $\sum_{i=1}^{\infty} f_i(x)$ **sums uniformly** to $f(x)$ on $[a, b]$ if the partial sums $\left\{\sum_{i=1}^{n} f_i(x)\right\}$ uniformly converges on $[a, b]$.

Definition 7.78. A **power series** at $x = a$ is an expression of the form $f(x) = \sum_{k=0}^{\infty} c_k(x-a)^k$.

Usually, we will consider $f(x) = \sum_{k=0}^{\infty} c_k x^k$.

Problem 7.79. For which x does $f(x) = \sum_{k=0}^{\infty} |c_k| |x|^k$ converge?

Solution. We have

$$\lim_{k\to\infty} \frac{|c_{k+1}| |x+1|^k}{|c_k| |x|^k} = |x| \lim_{k\to\infty} \left|\frac{c_{k+1}}{c_k}\right|.$$

Let $z = \lim_{k\to\infty} \left|\frac{c_{k+1}}{c_k}\right|$. We know that $z \geq 0$. Then, by the Ratio Test, $f(x) = \sum_{k=0}^{\infty} |c_k| |x|^k$ converges when $z = 0$ or $|x| \cdot z < 1 \implies |x| < \frac{1}{z}$. It will diverge when $|x| \cdot z > 1 \implies |x| > \frac{1}{z}$. □

Theorem 7.80

Given a power series $\sum_{k=0}^{\infty} c_k(x-a)^k$, one of the following is true:

1. It converges absolutely $\forall x$.
2. It converges only at $x = a$.
3. $\exists R > 0$ such that it converges absolutely whenever $|x-a| < R$, or diverges when $|x-a| > R$.

We can use this theorem with the Ratio Test to determine which power series are convergent. Here are a few examples:

1. For $\sum \frac{x^n}{n!}$, we have

$$\lim_{n\to\infty} \left|\frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}}\right| = \lim_{n\to\infty} \frac{|x|}{n+1} = 0 < 1,$$

so it converges absolutely $\forall x$. In other words, our **interval of convergence** is $\mathbb{R}$. The **radius of convergence** is half the length of the interval of convergence. In this case, the radius of convergence would be ∞.

2. For $\sum \frac{x^n}{n}$, we have

$$\lim_{n\to\infty} \left|\frac{\frac{x^{n+1}}{n+1}}{\frac{x^n}{n}}\right| = \lim_{n\to\infty} |x| \cdot \frac{n}{n+1} = |x|.$$

By the Ratio Test, this sum converges absolutely if $|x| < 1$, which means that it converges on the interval $-1 < x < 1$. However, we aren't done yet: we need to test the end points -1 and 1. For $x = -1$, it converges by Leibniz test, and $x = 1$ gives us the well-known divergent harmonic series. Thus, our interval of convergence is $[-1, 1)$. The radius of convergence is $\frac{1-(-1)}{2} = 1$.

3. For $\sum \frac{x^n}{n^2}$, we have

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

   We have nearly the same result as before, but this time, $x = 1$ also implies absolute convergence. Thus, the interval of convergence is $[-1, 1]$. Again, the radius of convergence is 1.

4. For $\sum n!x^n$, we have

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

   Thus, it diverges $\forall x$. Thus, the radius of convergence is 0.

5. For $\sum \frac{2^n(x-3)^n}{4^n+1}$, we have

$$\lim_{n\to\infty}\left|\frac{2^{n+1}(x-3)^{n+1}}{4^{n+1}+1}\cdot\frac{4^n+1}{2^n(x-3)^n}\right| = \frac{1}{2}|x-3|.$$

   We solve $\frac{1}{2}|x-3| < 1$ to get $1 < x < 5$, and we can verify that $x = 1, 5$ yield divergent series (using the Test for Divergence), so the interval of convergence is $(1, 5)$. Then, the radius of convergence is 2.

Theorem 7.81 (Weierstrauss M-Test)

Given a series defined as $\sum_{k=1}^{\infty} f_k(x)$ on A, suppose for each k there exists M_k such that $|f_k(x)| < M_k \ \forall x \in A$. Moreover, suppose $\sum_{k=1}^{\infty} M_k$ converges. Then $\sum_{k=1}^{\infty} f_k(x)$ converges uniformly and absolutely to some $f(x)$.

Theorem 7.82 (Dirichlet's Theorem)

Suppose $\{a_n\}$ and $\{b_n\}$ are series for which a_n is decreasing and going to 0, and the partial sums $b_1, b_1 + b_2, b_1 + b_2 + b_3, \ldots$ are bounded. Then, $\sum a_n b_n$ converges.

Problem 7.83 (Abel's Partial Summation Formula). Let $\{a_n\}$, $\{b_n\}$ be two sequences; let $A_n = \sum_{k=1}^{n} a_k$. Then

$$\sum_{k=1}^{n} a_k b_k = A_n b_{n+1} + \sum_{k=1}^{n} A_k(b_k - b_{k+1}).$$