Notes on MAT157: Analysis 1

Unversity of Toronto

David Duan

Last Updated: September 27, 2023 (draft)

Contents

1 Foundations **1**
1 Introduction . . . 2
2 Proving with Induction . . . 5
3 Functions . . . 6
4 Limits . . . 8
5 Supplementary: Countable Sets . . . 12
6 Continuous functions . . . 13
7 Important theorems . . . 15
8 Uniform Continuity . . . 19

2 Derivatives **20**
1 Derivatives . . . 21
2 Significance of the Derivative . . . 25

3 Inverse Functions **35**

4 Integrals **39**

5 Trigonometric Functions **47**

6 Log and Exp functions **53**

7 Integration in elementary terms **57**
3 List of important basic integrals . . . 57

8 Integration in elementary terms **60**
4 Taylor polynomials . . . 60

Chapter 1

Foundations

1 Introduction . . . 2
2 Proving with Induction . . . 5
3 Functions . . . 6
4 Limits . . . 8
5 Supplementary: Countable Sets . . . 12
6 Continuous functions . . . 13
7 Important theorems . . . 15
8 Uniform Continuity . . . 19

Section 1. Introduction

1.1. Definition: The real number is a **complete ordered field.**

1.2. Definition: A **field** is a set F with two binary operations, $+$, and $\cdot$, which satisfies:

P1.	Associative law for addition	$a + (b + c) = (a + b) + c.$
P2.	Existence of an additive identity	$a + 0 = 0 + a = a.$
P3.	Existence of additive inverse	$a + (-a) = (-a) + a = 0.$
P4.	Commutative law for addition	$a + b = b + a.$
P5.	Associative law for multiplication	$a \cdot (b \cdot c) = (a \cdot b) \cdot c.$
P6.	Existence of a multiplicative identity	$a \cdot 1 = 1 \cdot a = a; \quad 1 \neq 0.$
P7.	Existence of multiplicative inverses	$a \cdot a^(- 1) = a^(- 1) \cdot a = 1$ for $a \neq 0.$
P8.	Commutative law for multiplication	$a \cdot b = b \cdot a.$
P9.	Distributive law	$a \cdot (b + c) = a \cdot b + a \cdot c.$

1.3. Remark:

- 0 is unique. $(0 = 0 + 0\prime = 0\prime)$
- $-a$ is unique.
- *For all* $a, b, \quad (-a) \cdot b = -(ab)$
- *For all* $a, b, \quad (-a) \cdot (-b) = a \cdot b$
- *For all* $a, b, \quad a - b = b - a \iff a \cdot (1 + 1) = b \cdot (1 + 1)$

1.4. Definition: Given a field F. F is an **ordered field** if and only if there exist a subset $P \in F$ which is closed under addition and multiplication, and satisfies the Trichotomy Law, i.e.

P10. (Trichotomy law) For every number $a \in F$, one and only one of the following holds:

- $a = 0,$
- $a \in P,$
- $-a \in P$

P11. If $a, b \in P$, then $a + b \in P$.

P12. If $a, b \in P$, then $a \times b \in P$.

1.5. Remark: if $P \in F$ is an ordered field, then $1 \in P$

1.6. Definition: The **absolute value** of a is

$$|a| := \begin{cases} a & \text{if } a >= 0 \\ -a & \text{if } a < 0 \end{cases}$$

1.7. Theorem: *Triangle inequality and reverse triangle inequality:*

$$|x+y| \leq |x| + |y|, \quad ||x| - |y|| \leq |x-y|.$$

Proof. Note $|x| = \max\{x, -x\}$ and $\pm x \leq |x|$, then

$$a + b \leq |a| + |b| \text{ and } -(a+b) \leq |a| + |b|,$$

which gives the first inequality, then we can use it to get the second statement:

$$|x + (-y) + y| \leq |x-y| + |y| \text{ and } |y + (-x) + x| \leq |y-x| + |x|$$

which means

$$|x| - |y| \leq |x-y| \text{ and } |y| - |x| \leq |y-x|$$

which gives the second inequality. □

1.8. Definition: A set A of elements of an ordered field F is **bounded above***(resp. below)* if there exists an upper*(resp. lower)* bound $b \in F$, such that $b \geq a$ *(resp.* $b \leq a$*)* for all $a \in A$. A **least upper bound (supremum)**(*resp. greatest lower bound/**infimum***) b_0 of A is an upper bound of A and if b is any upper*(resp. lower)* bound, $b_0 \leq b$ (*resp.* $b_0 \geq b$).

1.9. Proposition: *The supremum and infimum of a set is unique if it exists. inf* $(A) \leq$ *sup* (A) *if they both exist*

1.10. Definition: F is a complete ordered field if and only if for every nonempty subset of A such that A which is bounded above has a least upper bound.

1.11. Theorem: *A complete ordered field exist and a complete ordered field is unique up to isomorphism*

1.12. Corollary:
1) For every real number x, there is an integer k such that $k > x$
2)For any $\epsilon > 0$, there is an $n > 0$ such that $0 < \frac{1}{n} < \epsilon$
3) Let $x, y \in \mathbb{R}$, if $y - x > 1$, then there is an $k \in \mathbb{Z}$ with $x < k < y$.
4) $x < y \in \mathbb{R}$, then there is a $r \in \mathbb{Q}$ such that $x < r < y$

1.13. Theorem: *There exist an element $x \in \mathbb{R}$ with $x^2 = 2$, i.e., 2 has a square root.*

Proof. Let

$$A = \{x \in \mathbb{R} | x^2 < 2\},$$

clearly A is non-empty and $\frac{3}{2}$ is an upper bound for A. Since $\mathbb{R}$ is a complete ordered field, there exist an $x = \sup(A) \in \mathbb{R}$ that is the least upper bound of A.
Claim: $x^2 = 2$.
Suppose not, then first suppose $x^2 < 2$, and we will show that for some small $\delta > 0$, $(x+\delta)^2 < 2$ which contradicts the fact that $x = \sup(A)$ as $x < x + \delta \in A$.
To find $\delta > 0$ with

$$(x+\delta)^2 \overset{?}{<} 2,$$

is the same as

$$x^2 + 2x\delta + \delta^2 \overset{x}{<} 2$$

$$\delta(2x+\delta) = 2x\delta + \delta^2 \overset{?}{<} 2 - x^2$$

since $x < 3/2$,

$$2x + \delta \leq 3 + \delta,$$

hence we can simply our inequality to

$$\delta(2x+\delta) \leq \delta(3+\delta) \overset{?}{<} 2 - x^2.$$

since we want δ small, lets just take $\delta < 1$, then we get

$$\delta(3+\delta) \leq 4\delta \overset{?}{<} 2 - x^2$$

Solving for delta, we get

$$0 < \delta < \min(1, \frac{1}{4}(2 - x^2)).$$

Note that we have chosen to include the condition $\delta \leq 1$ but it is not needed. To finish the argument, we can work backwards: let δ be defined as above, then

$$\delta(2x+\delta) \leq \delta(3+\delta) \leq 4\delta < 2 - x^2$$

which implies

$$x^2 + 2x\delta + \delta^2 < 2 \implies (x+\delta)^2 < 2.$$

As required, hence $x^2 < 2$ happen. An analogous argument can be made to show that $x^2 > 2$ is impossible. Therefore, $x^2 = 2$.

□

Section 2. Proving with Induction

2.1. Definition (Principal of mathematical induction):
Let P be a predicate such that

(1). $P(1)$ is true.

(2). $P(k) \implies P(k+1)$.

Then $P(k)$ is true for all natural number k.

2.2. The principle of mathematical induction may be formulated in an equivalent way that is better suited in a mathematical discussion.

2.3. Definition: Suppose A is any collection of natural numbers, then

(1). $1 \in A$,

(2). $k \in A \implies k+1 \in A$

then $A = \mathbb{N}$.

2.4. Definition (Complete induction): If A is a set of natural numbers and

(1). $1 \in A$,

(2). $1, ..., k \in A \implies k+1 \in A$,

then $A = \mathbb{N}$.

2.5. Theorem (Well ordering principle): *If A is a nonempty subset of $\mathbb{N}$, then A has a least element.*

2.6. Proposition: *If m is any integer and n is a positive integer, there exist unique q and r such that $m = qn + r$ and $0 \leq r < n$*

Proof. Let $m, q, r \in \mathbb{Z}$, $n \in \mathbb{Z}^+$, $A = \{m - qn | q \in \mathbb{Z}\} \cap \mathbb{N}$, clearly A is nonempty, since if $m >= 0$ then $m \in A$ when $q = 0$, and if $m < 0$, then the element $m - qn$ such that q is the smallest integer such that $m - qn > 0$ is in the set.
By the well ordering principle, there exist a smallest element $m - qn = r \in A$ where $r > 0$.
To prove $r < n$, assume the opposite, so $r \geq n$, then $m - qn \geq n$ which implies $m - n(q+1) \geq 0$. However, this number also satisfies the conditions to be in A and is smaller than r, which is a contradiction as we claimed r to be the smallest element in A. □

2.7. Theorem: *Mathematical induction ⇔ Complete Mathematical induction⇔ Well Ordering*

Proof. **MI ⇒ CMI:** Let $S = \{k \in \mathbb{N} | 1, ..., k \in \mathbb{N} \implies k+1 \in S\}$ and that $1 \in S$. Assume MI, we want to show $S = \mathbb{N}$, that is, CMI is true. Let $A = \{k \in \mathbb{N} | 1, ..., k \in S\}$, then $1 \in A$. Assume $k \in A$, then by definition of S, $k+1 \in S$, and hence $k+1 \in A$, by mathematical induction, we have $A = \mathbb{N}$, hence $S = \mathbb{N}$, as required.
CMI ⇒ WO: Suppose we have a nonempty set $A = \{a \in \mathbb{N}\}$ and $B = \{n \in \mathbb{N} | n \notin A\}$, we want to show $1 \in A$. Suppose not, A does not have a minimal element,this means $1, ..., k \in B$ which implies $1, ..., k \notin A \implies k+1 \notin A$, as otherwise it would the least element of A. Then by strong induction, $\mathbb{N} \in B$ and $A = \emptyset$, which is a contradiction. Thus $1 \in A$.
WO ⇒ MI: Suppose we have a set P such that $1 \in P$ and $n \in P \Rightarrow n+1 \in P$, we want to show $P = \mathbb{N}$. Suppose not, then we have a non-empty set $S = \{n \in \mathbb{N} | n \notin P\}$. By WO, there exist a least element in S which is not 1. Let k be its least element, then $k-1 \notin S$ which implies $k-1 \in P$. But by definition of P, $k-1 \in P \implies k \in P$, which is a contradiction. Thus $P = \mathbb{N}$ and $WO \Rightarrow MI$. □

2.8. Theorem (Fundamental Theorem of Arithmetic): *Every positive integer except 1 can be represented in one way up to isomorphism as a product of one or more primes.*

Proof. Base case: $2 = 2$, $3 = 3$, $4 = 2 \cdot 2$, $5 = 5$, clearly, first few numbers can be factored into primes.
Inductive step: Suppose every number $n \leq k$ can be factored in to product numbers. We consider $k+1$, it is either a prime in which case we are done, or a composite number, thus it can be written as the product of 2 factors, so $k+1 = n_1 n_2$ s.t. $n_1, n_2 \in \mathbb{Z}$ and $2 <= n_1, n_2 < k+1$. By induction hypothesis, n_1 can be written in the form of $p_1 p_2 ... p_k$ and n_2 can be written in the form of $q_1 q_2 ... q_r$. Multiplying them, we get $p_1 q_1 p_2 q_2 ... p_k q_r$. Therefore, since $k+1$ is a product of prime numbers, by strong induction, all $n \in \mathbb{Z}, n > 1$ can be written uniquely as a prime numbers.

□

Section 3. Functions

3.1. Definition: A function $f : A \to B$ is a subset $S \subseteq A \times B$, where

(1). we write $f(a) = b$ if $(a, b) \in S$

(2). $\forall a \in A, \exists (a, b) \in S$

(3). if $(a_1, b_1), (a_2, b_2) \in S$, then $a_1 = a_2 \Rightarrow b_1 = b_2$

3.2. Remark:
Domain: dom $f = \{a \in A | \exists b \in B, (a, b) \in S\}$
Range: ran $f = \{b \in B | \exists (a, b) \in S\}$

3.3. Theorem (Formula for an ellipse): $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

Proof. An ellipse is defined as the set of points whose distance from each of two "focus" adds up to the same value. For convinience, let them be $(-c, 0), (c, 0)$, and the sum of distances to be $2a$ Using the distance formula:

$$\sqrt{(x-(-c))^2 + y^2} + \sqrt{(x-c)^2 + y^2} = 2a$$
$$\sqrt{(x+c)^2 + y^2} = 2a - \sqrt{(x-c)^2 + y^2}$$
$$x^2 + 2cx + c^2 + y^2 = 4a^2 - 4a\sqrt{(x-c)^2 + y^2} + x^2 - 2cx + c^2 + y^2$$
$$4(cx - a^2) = -4a\sqrt{(x-c)^2 + y^2}$$
$$c^2x^2 - 2cxa^2 + a^4 = a^2(x^2 - 2cx + c^2 + y^2)$$
$$(c^2 - a^2)x^2 - a^2y^2 = a^2(c^2 - a^2)$$
$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1$$

we usually let $b = \sqrt{a^2 - c^2}$ so the equation becomes

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

□

3.4. Remark: The **hyperbola** is defined analogously, except we require the *difference* of the two distances to be constant
$\sqrt{(x-(-c))^2 + y^2} - \sqrt{(x-c)^2 + y^2} = \pm 2a, \implies \frac{x^2}{a^2} - \frac{y^2}{a^2 - c^2} = 1$
However, in this case, we must choose $c > a$, so $a^2 < c^2 = 0$, otherwise its a ellipse. So let $b = \sqrt{c^2 - a^2}$, and get $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$

Section 4. Limits

4.1. Definition (Delta-Epsilon): The function f approches the limit l near a means:

$$\forall \epsilon > 0, \exists \delta > 0, \text{ s.t.} \forall x, 0 < |x-a| < \delta \implies |f(x) - l| < \delta$$

4.2. Definition (Uniqueness of Limit): If a function f approaches l near a, and approaches m near a, then $l = m$

Proof. Suppose the limit is not unique, then

$$\forall \epsilon > 0, \exists \delta_1 > 0, \text{ s.t. } \forall x, 0 < |x-a| < \delta_1 \implies |f(x) - l| < \epsilon$$

and

$$\forall \epsilon > 0, \exists \delta_2 > 0, \text{ s.t. } \forall x, 0 < |x-a| < \delta_2 \implies |f(x) - m| < \epsilon$$

We used δ_1 and δ_2 since we can't ensure that the δ which satisfy one definition will work in the other. However, it is easy to conclude that for all $\epsilon > 0$, there will be some $\delta > 0$ that works if we just simply choose $\delta = min(\delta_1, \delta_2)$. Now let $\epsilon = \frac{|l-m|}{2}$, it follows that

$$0 < |x-a| < \delta \implies |f(x) - m| < \frac{|l-m|}{2} \text{ and } |f(x) - l| < \frac{|l-m|}{2}$$

By triangle inequality,

$$|l - m| = |l - f(x) + f(x) - m| <= |f(x) - l| + |f(x) - m| < 2 \cdot \frac{|l-m|}{2} = |l - m|$$

which is a contradiction. □

Intuitively, we can think of $[l-\epsilon, l+\epsilon]$ as the range of possible $f(x)$, such that no matter what ϵ we are given, we can always find a δ such that any x in the interval $[x-\delta, x+\delta]$ gives a $f(x) \in [l-\epsilon, l+\epsilon]$

4.3. Theorem: *If $\lim_{x\to a} f(x) = l$ and $\lim_{x\to a} g(x) = m$, then*

(1). $\lim_{x\to a}(f+g)(x) = l + m.$

(2). $\lim_{x\to a}(f \cdot g)(x) = l \cdot m.$

(3). If $m \neq 0$, $\lim_{x\to a}(\frac{1}{g})(x) = \frac{1}{m}.$

4.4. Lemma (1): *If $|x - x_0| < \frac{\epsilon}{2}$ and $|y - y_0| < \frac{\epsilon}{2}$, then $|(x+y) - (x_0 + y_0)| < \epsilon$*

Proof.

$$|(x+y) - (x_0 + y_0)| = |(x - x_0) + (y - y_0)| \leq |x - x_0| + |y - y_0| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

□

4.5. Lemma (2): *If $|x - x_0| < min(1, \frac{\epsilon}{2(|y_0|+1)})$ and $|y - y_0| < \frac{\epsilon}{2(|x_0|+1)}$, then $|xy - x_0y_0| < \epsilon$*

Proof. $|x| - |x_0| \leq |x - x_0| < 1 \implies |x| < 1 + |x_0|$.

$$\begin{aligned} |xy - x_0y_0| = |xy - xy_0 + xy_0 - x_0y_0| &\leq |x(y - y_0)| + |y_0(x - x_0)| \\ &\leq |x| \cdot |y - y_0| + |y_0| \cdot |x - x_0| \\ &< (1 + |x_0|) \cdot \frac{\epsilon}{2(|x_0| + 1)} + |y_0| \cdot \frac{\epsilon}{2(|y_0| + 1)} \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon \end{aligned}$$

□

4.6. Lemma (3): *If $y_0 \neq 0$ and $|y - y_0| < min(\frac{|y_0|}{2}, \frac{\epsilon|y_0|^2}{2})$, then $y \neq 0$ and $|\frac{1}{y} - \frac{1}{y_0}| < \epsilon$*

Proof. $|y_0| - |y| \leq |y - y_0| < \frac{|y_0|}{2} \implies |y| > \frac{|y_0|}{2}$. Clearly $y \neq 0$ so $\frac{1}{|y|} < \frac{2}{y_0}$.

$$|\frac{1}{y} - \frac{1}{y_0}| = \frac{|y_0 - y|}{|y| \cdot |y_0|} < \frac{2}{|y_0|} \cdot \frac{1}{|y_0|} \cdot \frac{\epsilon|y_0|^2}{2} = \epsilon$$

□

Proof. Now to prove each theorem in 4.3:
(1) The hypothesis states that there are $\delta_1, \delta_2 > 0$ such that for all x,

$$0 < |x - a| < \delta_1 \implies |f(x) - l| < \frac{\epsilon}{2} \text{ and } 0 < |x - a| < \delta_2 \implies |f(x) - l| < \frac{\epsilon}{2}$$

Let $\delta = min(\delta_1, \delta_2)$, so that $0 < |x - a| < \delta$ implies both implication. By lemma 1, this implies $|(f + g)(x) - (l + m)| < \epsilon$
(2) Similarly,

$$0 < |x-a| < \delta_1 \implies |f(x)-l| < min(1, \frac{\epsilon}{2(|m| + 1)}) \text{ and } 0 < |x-a| < \delta_2 \implies |f(x)-l| < \frac{\epsilon}{2|l| + 1}$$

Let $\delta = min(\delta_1, \delta_2)$, thus if $0 < |x - a| < \delta$, both implication stands and by lemma 2,this implies $|(f \cdot g)(x) - l \cdot m| < \epsilon$.
(3) If $\epsilon > 0$ then there exist $\delta > 0$ such that for all x,

$$0 < |x - a| < \delta \implies |g(x) - m| < min(\frac{|m|}{2}, \frac{\epsilon|m|^2}{2})$$

By lemma 3, this implies $g(x) \neq 0$ and $|(\frac{1}{g})(x) - \frac{1}{m}| < \epsilon$ □

4.7. Lemma: *If* $|g(x) - m|$ *is small we want estimates*

$$|g(x)| < 1 + |m| \quad and \quad if\ m \neq 0, \frac{|m|}{2} < |g(x)|$$

Alternative proof. **(1)**Let

$$\epsilon_1 = \frac{\epsilon}{2}, \epsilon_2 = \frac{\epsilon}{2}$$

Given any $\epsilon > 0$,

$$0 < |x - a| < \delta \implies |(f + g)(x) - l - m| < \epsilon$$
$$|f(x) - l + g(x) - m| \leq |f(x) - l| + |g(x) - m| < \epsilon_1 + \epsilon_2 = \epsilon$$

(2) Suppose $\epsilon_1, \epsilon_2 < 1$, then

$$\epsilon_1\epsilon_2 = \frac{1}{2}(\epsilon_1\epsilon_2 + \epsilon_1\epsilon_2) < \frac{1}{2}(\epsilon_1 + \epsilon_2)$$

we have $0 < |x - a| < \delta \implies |(f \cdot g)(x) - l \cdot m| < \epsilon$

$$\begin{aligned}
|f(x)g(x) - l \cdot m| &= |(f(x) - l)g(x) + lg(x) - m| \\
&= |(f(x) - l)(g(x) - m) + m(f(x) - l) + l(g(x) - m)| \\
&\leq |f(x) - l||g(x) - m| + |m||f(x) - l| + |l||g(x) - m| \\
&< \epsilon_1\epsilon_2 + |m|\epsilon_1 + |l|\epsilon_2 \\
&< \frac{1}{2}(\epsilon_1 + \epsilon_2) + |m|\epsilon_1 + |l|\epsilon_2 \\
&= \epsilon_1(\frac{1}{2} + |m|) + \epsilon_2(\frac{1}{2} + |l|) \\
&\overset{?}{<} \epsilon
\end{aligned}$$

Take

$$\delta_1 \text{ s.t. } \epsilon_1 = \frac{\epsilon}{2} \cdot \frac{1}{\frac{1}{2} + |m|}$$

$$\delta_2 \text{ s.t. } \epsilon_2 = \frac{\epsilon}{2} \cdot \frac{1}{\frac{1}{2} + |l|}$$

$$\epsilon_1 = \min(1, \frac{\epsilon}{1 + 2|m|}), \epsilon_2 = \min(1, \frac{\epsilon}{1 + 2|l|})$$

We take $\delta = \min(\delta - 1, \delta_2)$ and we get LHS $< \epsilon$
(3) by (2),

$$|\frac{f(x)}{g(x)} - \frac{l}{m}| \overset{?}{<} \epsilon$$

is the same as solving whether

$$|\frac{1}{g(x)} - \frac{1}{m}| \overset{?}{<} \epsilon, \text{ or } \frac{|g(x) - m|}{|g(x)||m|} \overset{?}{<} \epsilon$$

To solve this we first prove the *above + below lemma*
Proof: Suppose $|g(x)| - |m| < |g(x) - m| < 1$, then above: $|g(x)| < 1 + |m|$
Suppose $|m| - |g(x)| < |g(x) - m| < \frac{|m|}{2}$, then below: $|m| < \frac{|m|}{2} + |g(x)| \implies \frac{|m|}{2} < |g(x)|$
Back to proving (3), now supposed $|g(x) - m| < \frac{|m|}{2} \quad (m \neq 0)$
then

$$\frac{|g(x) - m|}{|g(x)||m|} < \frac{|g(x) - m|}{\frac{|m|}{2}|m|} = \frac{2}{|m|^2}|g(x) - m| \overset{?}{<} \epsilon$$

.
Take δ s.t. $|g(x) - m| < \min(\frac{|m|}{2}, \frac{\epsilon}{2} \cdot |m|^2)$ and we get the desired equation. □

4.8. Intuition: The lemmas (1,2,3) we proved is merely saying that, when x is close to x_0, and y is closed to y_0, then $x + y$ will be closed to $x_0 + y_0$, and xy will be close to $x_0 y_0$, and $\frac{1}{y}$ will be closed to $\frac{1}{y_0}$

4.9. Definition: Sometimes we would like to only speak about the limit of f approaches some a on one side. In that case, we have limit from above and below which are defined as:

$$\lim_{x \to a^+} f(x) = l \text{ if } \forall \epsilon > 0, \exists \delta > 0, \text{ s.t. } \forall x, 0 < x - a < \delta \implies |f(x) - l| < \epsilon$$

$$\lim_{x \to a^-} f(x) = l \text{ if } \forall \epsilon > 0, \exists \delta > 0, \text{ s.t. } \forall x, 0 < a - x < \delta \implies |f(x) - l| < \epsilon$$

4.10. Note: When talking about the limit of $f(x)$ as x approaches ∞, or $\lim_{x \to \infty} f(x)$, we often call this the limit at infinity, and it is defined as:

$$\forall \epsilon > 0, \exists N, \text{ s.t.} \forall x, x > N \implies |f(x) - l| < \epsilon$$

4.3 also works on limits to inifinity if we take $N = \max(N_1, N_2)$

4.11. Note: When we talk about an infinite limit at a point a, we say that the limit of $f(x)$ as x approaches a diverges to inifinity,

$$\forall M > 0, \exists \delta > 0, s.t. \ \forall x, 0 < |x - a| < \delta \implies f(x) > M$$

Section 5. Supplementary: Countable Sets

5.1. Definition: A set A is countable if there is a **surjective** function $f : \mathbb{N} \to A$

5.2. Fact:

(1). A *is finite* $\implies$ *countable*

(2). $A \in N \implies$ *countable*

(3). If $B \to A$ *surjective then* B *countable* $\implies A$ *countable*

(4). $A \subseteq B$ *then* B *countable* $\implies A$ *countable*

(5). $\mathbb{Q}$ *is countable*

(6). $\mathbb{R}$ *is not countable*

(4). Suppose B is countable, let f be the surjective map from $\mathbb{N} \to B$, let $f^{-1} = \{n | f(n) \in A\}$. Since $f : f^{-1}(A) \subset \mathbb{N} \to A$ is surjective we have that A is countable □

(5). Let's write all $q \in \mathbb{Q}$ as a fraction $\frac{a}{b}$ such that $a, b > 0$ and they have no common factor. By the fundamental theorem of arithmetic:

$$a = p_1^{\alpha_1} \cdots p_r^{a_r}, b = q_1^{\beta_1} \cdots q_s^{\beta_s}, \alpha_i, \beta_j > 0$$

We define a function $f : \mathbb{Q}_+ \to N$ which sends $\frac{a}{b} \mapsto p_1^{2\alpha_1} \cdots p_r^{2a_r} \cdot q_1^{2\beta_1+1} \cdots q_s^{2\beta_s+1} \in \mathbb{N}$
Since prime factorization is unique this function is injective and thus a subset of $\mathbb{N}$, therefore $\mathbb{Q}$ is countable. □

5.3. Fact: $\mathbb{R}$ *is uncountable.*

Section 6. Continuous functions

6.1. Definition: The function f is continuous at a if

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

6.2. Definition: If f and g are continuous at a, then

(1). $f + g$ is continuous at a

(2). $f \cdot g$ is continuous at a

(3). If $g(a) \neq 0$, then $\frac{1}{g}$ is continuous at a

Proof. Since f and g are continuous at a, then

$$\lim_{x \to a} f(x) = f(a) \quad \text{and} \quad \lim_{x \to a} g(x) = g(a)$$

(1) By theorem 4.3,

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

(2) By theorem 4.3,

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

(3)

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

□

6.3. Note (continuous limit):

$$\forall \epsilon > 0, \exists \delta > 0, s.t. \forall x, |x - a| < \delta \implies |f(x) - f(a)| < \epsilon$$

6.4. Theorem: *If g is continuous at a, and f is continuous at $g(a)$, then $f \circ g$ is continuous at a.*

Proof. Let $\epsilon > 0$, our goal is to show there exist $\delta' > 0$ such that

$$|x - a| < \delta' \implies |f(g(x)) - f(g(a))| < \epsilon$$

Since g is continuous at a, we know for all ϵ' that there exist δ' such that

$$|x - a| < \delta' \implies |g(x) - g(a)| < \epsilon'$$

Since f is continuous at $g(a)$, we know for all ϵ such that

$$|g(x) - g(a)| < \delta \implies |f(g(x)) - f(g(a))| < \epsilon$$

Since δ is just some positive number, we can take it as ϵ', so then the equation becomes

$$|x-a|<\delta' \implies |g(x)-g(a)|<\epsilon'(=\delta) \implies |f(g(x))-f(g(a))|<\epsilon$$

□

6.5. Definition: We say f is continuous on (a,b) if it continuous at every point in (a,b). However, if f is continuous on $[a,b]$, then it is continuous on every point in (a,b) and $\lim_{a^+} f(x) = f(a)$ and $\lim_{a\to b^-} f(x) = f(b)$

6.6. Lemma: *Suppose f is continuous at a, and $f(a)>0$, then there exist $\delta>0$, such that $f(x)>0$ for all $x \in |x-a|<\delta$. Similarly, if $f(a)<0$, then there exist $\delta>0$, such that $f(x)<0$ for all $x \in |x-a|<\delta$. Similar argument are also correct for one sided limits.*

Proof. Consier $f(a)>0$, since f is continuous at a, there exist a $\delta>0$, such that for all $\epsilon>0$, $|x-a|<\delta \implies |f(x)-f(a)|<\epsilon$ Since $f(a)>0$, take $\epsilon=f(a)$, then

$$|x-a|<\delta \implies |f(x)-f(a)|<f(a)$$

which implies $f(x)>0$. An analoguous argument can be given for $f(a)<0$ by setting $\epsilon=-f(a)$. As well as one sided arguments. □

Section 7. Important theorems

7.1. Theorem: *Suppose f is, continuous on $[a, b]$ and $f(a) < 0 < f(b)$, then there is some x in $[a, b]$ such that $f(x) = 0$.*

Proof. Let $A = \{x \in [a, b] | f < 0 \text{ on } [a, x]\}$. Then $A \neq \emptyset$ since $a \in A$. By 6.6, there exist an interval $[a, a+\delta)$ such that $f(x) < 0 \; \forall x \in [a, a+\delta)$. Similarly, b is an upper bound of A, and since $f(b) > 0$, there exist $\delta > 0$ such that all $x \in b - \delta < x \leq b$ are upper bounds of A. By P13, there exist a least upper bound α of A, or $\sup(A)$, we now wish to show that $f(\alpha) = 0$. Suppose first $f(\alpha) > 0$, then by 6.6, $f(x) > 0$ on $(\alpha - \delta, \alpha + \delta)$ for some $\delta > 0$. However we know that there is x_0 in A in $\alpha - \delta < x_0 < \alpha$, since otherwise α would not be the least upper bound , but then this means that $f(x_0) > 0$ which is impossible, thus $f(\alpha)$ cannot be larger than 0. Suppose $f(\alpha) < 0$, then $f(x) < 0$ on $(\alpha - \delta, \alpha + \delta)$. Now there is some $x_0 \in A$ which satisfies $\alpha - \delta < x_0 < \alpha$, so f is negative on $[x, x_0]$, but if x_1 is a number on the interval $[\alpha, \alpha + \delta)$, then f is negative on the interval $[a, x_1]$ so $x_1 \in A$, which is impossible as well. Therefore, $f(\alpha) = 0$. □

7.2. Proposition: *If f is continuous on $[a, b]$ and $f(a) < c < f(b)$, then there is some x in $[a, b]$ such that $f(x) = c$.*

Proof. Let $g = f - c$, then g is continuous and $g(a) < 0 < g(b)$, by 7.1, there exist $x \in [a, b]$ such that $g(x) = 0$, which means $f(x) = c$. □

7.3. Proposition: *If f is continuous on $[a, b]$ and $f(a) > c > f(b)$, then there is some x in $[a, b]$ such that $f(x) = c$.*

Proof. $-f$ is continuous on $[a, b]$ and $-f(a) < -c < -f(b)$, by 7.2, there exist $x \in [a, b]$ such that $-f(x) = -c$, which means $f(x) = c$. □

7.4. Fact: *If f is continuous at a, then there is a $\delta > 0$, s.t. f is bounded above on $(a-\delta, a+\delta)$. (Also variation on one side limits)*

Proof. Since f continuous at a, take $\epsilon = 1$, then we have

$$\exists \delta > 0, s.t. \; |x - a| < \delta \implies |f(x) - f(a)| < 1 \implies f(x) < f(a) + 1$$

□

7.5. Theorem: *If f is continuous on $[a, b]$, then f is bounded above on $[a, b]$*

Proof. Let $A = \{x \in [a, b] | f \text{ is bounded above on [a,x]}\}$ By 7.4, $A \neq \emptyset$ as $a \in A$, since b is an upperbound of A, there exist a supremum of A, lets call α, we want to show that $\alpha = b$. We know $a < \alpha \leq b$. Suppose $\alpha < b$, then there is a $\delta > 0$ such that f is bounded on $(\alpha - \delta, \alpha + \delta)$. Since α is the least upper bound there exist x_0 in A satisfying $\alpha - \delta < x_0 < \alpha$. So f is bounded on $[a, x_0]$. But there also exists x_1 such that $\alpha < x_1 < \alpha + \delta$, and so f is bounded on $[x_0, x_1]$. Therefore, f is bounded on $[a, x_1]$ which contradicts the fact that α is an upper bound for A, and thus $\alpha = b$ Now we have proved that f is bounded on $[a, x]$ for all $x < b$, we are only left to prove that f is indeed bounded on $[a, b]$. By 7.4, since b is 'continuous' from below, there exist $\delta > 0$ s.t. f is bounded on $(b - \delta, b]$. Take any x in this interval, we know f is bounded on $[a, x]$ and $[x, b]$, hence f is bounded on $[a, b]$ □

7.6. Proposition: *If f is continuous on $[a, b]$, then f is bounded below on $[a, b]$, i.e. there is some number N such that $f(x) \geq N$ for all $x \in [a, b]$*

Proof. The function $-f$ is continuous on $[a, b]$, so by 7.5 there is a number M such that $-f(x) \leq M$ for all $x \in [a, b]$, which means $f(x) \geq -M$ for all $x \in [a, b]$, so we can let $N = -M$. □

7.7. Corollary: *7.5 and 7.6 together shows that a continuous function f on $[a, b]$ is bounded on $[a, b]$, i.e., there is a number N such that $|f(x)| \leq N$ for all $x \in [a, b]$. Suppose we have N_1 such that $f(x) \leq N_1$, and N_2 such that $f(x) \geq N_2$ for all $x \in [a, b]$, we can take $N = \max(|N_1|, |N_2|)$*

7.8. Theorem: *If f is continuous on $[a, b]$, then there is a numbre y in $[a, b]$ such that $f(y) \geq f(x)$ for all x in $[a, b]$*

Proof. Let $B = \{f(x)|x \in [a, b]\}$. $B \neq 0$, and by 7.5, B is bounded above and $sup(B) = \beta$ exists. Since $\beta \geq f(x)$ for $x \in [a, b]$ it suffices to show that $\beta = f(y)$ for some $y \in [a, b]$. Suppose not, let $g(x) = \frac{1}{\beta - f(x)}$. Then g is continuous on $[a, b]$ since the denominator is never 0, and by 7.5, g is bounded on $[a, b]$. However, by the definition of β, we can find x in $[a, b]$ such that $\beta - f(x)$ can be made arbitrary small. That is,

$$\forall \epsilon > 0, \exists x \in [a, b], s.t.\ \beta - f(x) < \epsilon$$

This, in turn, means,

$$\forall \epsilon > 0 \exists x \in [a, b], s.t.\ g(x) > \frac{1}{\epsilon}$$

Which implies g is not bounded on $[a, b]$, contradicting our assumption. □

7.9. Proposition: *If f is continuous on $[a, b]$, then there is some y in $[a, b]$ such that $f(y) \leq f(x)$ for all x in $[a, b]$*

Proof. The function $-f$ is continuous on $[a, b]$, by 7.8, there is some y in $[a, b]$ such that $-f(y) \geq -f(x)$ for all $x \in [a, b]$, which implies that $f(y) \leq f(x)$ for all $x \in [a, b]$ □

7.10. Theorem: *Every positive number has a square root, in other words, if $\alpha > 0$, then there is some number x such that $x^2 = \alpha$*

Proof. Consider $f(x) = x^2$ which is continuous. The statement of "the number α has a square root" simply means $f(x)$ takes on the value α which is an easy consequence of 7.3. There is obviously a number $b > 0$ such that $f(b) > \alpha$, and α always > 0. So we can apply 7.3 to $[0, b]$. □

7.11. Intuition: The same arugment can be used to prove that a positive number has an nth root, for all natural number n, and if n happens to be odd, we can actually prove that every number has an nth root, i.e., if $x^n = a$, then $(-x)^n = -a$

7.12. Theorem: *If n is odd, then any equation*

$$x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$$

has a root.

Proof. Consider the function $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$. The idea is that for a large $|x|$, this function act very much liek $g(x) = x^n$, and since n is odd, $f(x)$ is postive for large positive x, and negative for large negaive x.

$$f(x) = x^n = a_{n-1}x^{n-1} + \cdots + a_0 = x^n(1 + \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n})$$

Note that,

$$|\frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n}| \leq \frac{|a_{n-1}|}{|x|} + \cdots + \frac{|a_0|}{|x^n|}.$$

So lets choose an x satisfying

$$|x| > 1, 2n|a_{n-1}|, ..., 2n|x_0| \tag{$*$}$$

Then $|x^k| > |x|$ and

$$\frac{|a_{n-k}|}{|x^k|} < \frac{|a^{n-k}|}{|x|} < \frac{|a_{n-k}|}{2n|a_{n-k}|} = \frac{1}{2n}$$

So

$$|\frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n}| \leq \frac{1}{2n} + \cdots + \frac{1}{2n} = \frac{1}{2}$$

In other words,

$$-\frac{1}{2} \leq \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n} \leq \frac{1}{2}$$

Which implies

$$\frac{1}{2} \leq 1 + \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n}$$

Therefore, if we choose an $x_1 > 0$ which satisfies $*$,

$$\frac{(x_1)^n}{2} \leq (x_1)^n(1 + \frac{a_{n-1}}{x_1} + \cdots + \frac{a_0}{(x_1)^n}) = f(x_1)$$

so that $f(x_1) > 0$, and if we choose an $x_2 < 0$ which satisfies $*$, then

$$\frac{(x_2)^n}{2} \geq (x_2)^n(1 + \frac{a_{n-1}}{x_2} + \cdots + \frac{a_0}{(x_1)^n}) = f(x_2)$$

so $f(x_2) < 0$.
Now we can apply 7.1 to the interval $[x_2, x_1]$ and conclude that there must be some x in the interval such that $f(x) = 0$ ∎

7.13. Theorem: *If n is even and $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$, then there is a number y such that $f(y) \leq f(x)$ for all x.*

Proof. Similar to last proof, if we have $M = \max(1, 2n|a_{n-1}, ..., 2n|a_0|)$, then for all x with $|x| \geq M$, we have

$$\frac{1}{2} \leq 1 + \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n}$$

Since n is even, $x^n \geq 0$ for all x, so

$$\frac{(x)^n}{2} \leq (x)^n(1 + \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{(x)^n}) = f(x)$$

if $|x| \geq M$. Now consider $f(0)$, let $b > 0$ be a number that such $b^n \geq 2f(0)$ and $b > M$. Then, if $x \geq b$, we have

$$f(x) \geq \frac{x^n}{2} \geq \frac{b^n}{2} \geq f(0)$$

If $x \leq -b$, then

$$f(x) \geq \frac{x^n}{2} \geq \frac{(-b)^n}{2} = \frac{b^n}{2} \geq f(0)$$

Therefore, if $x \geq b$ or $x \leq -b$, then $f(x) \geq f(0)$.
Apply 7.8 on the interval $[-b, b]$, we conclude that there is a number y such that

$$-b \leq x \leq b \implies f(y) \leq f(x)$$

In particular, $f(y) \leq f(0)$ because $0 \in [-b, b]$, thus

$$x \leq -b \text{ or } x \geq b \implies f(x) \geq f(0) \geq f(y)$$

Combining the last two equation we see that $f(y) \leq f(x)$ for all x. □

7.14. Intuition: The idea here is to show first that a minimum $f(y)$ exist on an interval, ex. $[-b, b]$, then we show that all elements not in the interval are also greater than this minimum.

7.15. Theorem: *Consider the equation*

$$x^n + a_{n-1}x^{n-1} + \cdots + a_0 = c$$

and suppose n is even, then there is a number m such that the equation has a solution for $c \geq m$ and has no solution for $c < m$

Proof. Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$. According to the lsat theorem there is a number y such that $f(y) \leq f(x)$ for all x. Let $m = f(y)$, if $c < m$, then the equation obviously has no solution. If $c = m$, then y is a solution. If $c > m$, let $b > y$ and $f(b) > c$. Then we have $f(y) < c < f(b)$, and by IVT, there exist $x \in [y, b]$ such that $f(x) = c$, so x is a solutoin. □

7.16. Theorem: $\mathbb{N}$ *is not bounded above.*

Proof. Suppose $\mathbb{N}$ is bounded above, since $\mathbb{N} \neq \emptyset$, there eixst a least upper bound α of $\mathbb{N}$, then

$$a \geq n \ \forall n \in \mathbb{N}$$

and also

$$\alpha \geq n+1 \ \forall n \in \mathbb{N}$$

Since $n+1 \in \mathbb{N}$ if $n \in \mathbb{N}$, but this means

$$\alpha - 1 \geq n \ \forall n \in \mathbb{N}$$

which means $\alpha - 1$ is also an upper bound for $\mathbb{N}$, contradicting our assumption. □

7.17. Theorem: *For any $\epsilon > 0$, there is a natural number n with $\frac{1}{n} < \epsilon$*

Proof. Suppose not, then for some $\epsilon > 0$, $\frac{1}{n} \geq \epsilon \ \forall n \in \mathbb{N}$. Then $n \leq \frac{1}{\epsilon} \ \forall n \in \mathbb{N}$. But this means that $\frac{1}{\epsilon}$ is an upper bound for $\mathbb{N}$, contradicting out last theorem. □

Section 8. Uniform Continuity

8.1. Definition: The function f is **uniformly continuous on an interval** A if for every $\epsilon > 0$ there is some $\delta > 0$ such that, for all x and y in A,

$$|x-y| < \delta \implies |f(x) - f(y)| < \epsilon$$

8.2. Lemma (Splice Lemma): *Let $a < b < c$ and let f be continuous on the interval $[a, c]$ s.t. $\forall \epsilon > 0$*

$$\exists \delta_1 > 0 \ s.t. \ \forall x, y \in [a, b], |x-y| < \delta_1 \implies |f(x) - f(y)| < \epsilon \quad (1)$$
$$\exists \delta_2 > 0 \ s.t. \ \forall x, y \in [b, c], |x-y| < \delta_2 \implies |f(x) - f(y)| < \epsilon \quad (2)$$

Then there is $\delta > 0$ s.t. $\forall x, y \in [a, c], |x-y| < \delta \implies |f(x) - f(y)| < \epsilon$

Proof. Since f is continuous at b, there is a $\delta_3 > 0$ such that,

$$|x-b| < \delta_3 \implies |f(x) - f(b)| < \frac{\epsilon}{2}$$

It follows that

$$|x-b| < \delta_3 \text{ and } |y-b| < \delta_3 \implies |f(x) - f(y)| < \epsilon$$

Now choose δ to be the minimum of $\delta_1, \delta_2, \delta_3$. Suppose x, y are any two points in $[a, c]$ with $|x-y| < \delta$. If $x, y \in [a, b]$, then $|f(x) - f(y)| > \epsilon$ by (1). If $x, y \in [b, c]$, then $|f(x) - f(y)| < \epsilon$ by (2). Otherwise, $x < b < y$ or $y < b < x$, and in either case, since $|x-y| < \delta$, $|x-b| < \delta$ and $|y-b| < \delta \implies |f(x) - f(y)| < \epsilon$ by (3). □

8.3. Theorem: *If f is continuous on $[a, b]$, then f is uniformly continuous*

Proof. tba □

Chapter 2

Derivatives

1 Derivatives **21**
2 Significance of the Derivative **25**

Section 1. Derivatives

1.1. Definition: The function f is **differentiable at a** if

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \text{ exists}$$

The limit is denoted by $f'(a)$ and is called the **derivative of f at a**.
We define the **tangent line** to the graph of f at $(a, f(a))$ to be the line through $(a, f(a))$ with slope $f'(a)$.

1.2. Theorem: *If f is differentiable and a, then f is continuous at a.*

Proof.

$$\begin{aligned}\lim_{h\to 0} f(a+h)-f(a) &= \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}\cdot h \\ &= \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}\cdot \lim_{h\to 0} h \\ &= f'(a)\cdot 0 \\ &= 0\end{aligned}$$

Therefore f is continuous at a □

1.3. Theorem: *If f is a constant function, $f(x)=c$, then*

$$f'(a) = 0 \quad \forall a$$

Proof.

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

□

1.4. Theorem: *If f is the identity function, $f(x)=x$, then*

$$f'(a) = 1 \quad \forall a$$

Proof.

$$\begin{aligned}f'(a) &= \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \\ &= \lim_{h\to 0} \frac{a+h-a}{h} \\ &= \lim_{h\to 0} \frac{h}{h} = 1\end{aligned}$$

□

1.5. Theorem: *If f and g are differentiable at a, then $f+g$ is also differentiable at a, and*

$$(f+g)'(a) = f'(a) + g'(a)$$

Proof.

$$\begin{aligned}(f+g)'(a) &= \lim_{h\to 0} \frac{(f+g)(a+h)-(f+g)(a)}{h} \\ &= \lim_{h\to 0} \frac{f(a+h)+g(a+h)-[f(a)+g(a)]}{h} \\ &= \lim_{h\to 0}[\frac{f(a+h)-f(a)}{h} + \frac{g(a+h)-g(a)}{h}] \\ &= \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} + \lim_{h\to 0} \frac{g(a+h)-g(a)}{h} \\ &= f'(a) + g'(a)\end{aligned}$$

□

1.6. Theorem: *If f and g are differentiable at a, then*

$$(f\cdot g)'(a) = f'(a)\cdot g(a) + f(a)\cdot g'(a)$$

Proof.

$$\begin{aligned}(f\cdot g)'(a) &= \lim_{h\to 0} \frac{(f\cdot g)(a+h)-(f\cdot g)(a)}{h} \\ &= \lim_{h\to 0} \frac{f(a+h)\cdot g(a+h)-f(a)g(a)}{h} \\ &= \lim_{h\to 0} \frac{f(a+h)[g(a+h)-g(a)]}{h} + \frac{[f(a+h)-f(a)]g(a)}{h} \\ &= \lim_{h\to 0} f(a+h)\cdot \lim_{h\to 0} \frac{g(a+h)-g(a)}{h} + \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \cdot \lim_{h\to 0} g(a) \\ &= f(a)\cdot g'(a) + f'(a)\cdot g(a)\end{aligned}$$

□

1.7. Theorem: *If $g(x) = cf(x)$ and f is differentiable at a, then g is differentiable at a, and*

$$g'(a) = c\cdot f'(a)$$

Proof. Let $h(x) = c$, so $g = h\cdot f$, by last theorem,

$$\begin{aligned}g'(a) &= (h\cdot f)'(a) \\ &= g(a)\cdot f'(a) + h'(a)\cdot f(a) \\ &= c\cdot f'(a) + 0\cdot f(a) \\ &= cf'(a)\end{aligned}$$

□

1.8. Theorem: *If $f(x) = x^n$ for some natural number n, then*

$$f'(a) = na^{n-1} \quad \forall a$$

Proof. We prove by induction on n. $n = 1$ is clearly true as proven above. Now assume that the theorem is true for n, so if $f(x) = x^n$, then $f'(a) = na^{n-1}$. We want to prove this is true for $g(x) = x^{n+1}$. Let $I(x) = x$, then the equation can be written as

$$g(x) = f(x) \cdot I(x)$$

By product rule,

$$\begin{aligned} g'(a) = (f \cdot I)'(a) &= f'(a) \cdot I(a) + f(a) \cdot I'(a) \\ &= na^{n-1} \cdot a + a^n \cdot 1 \\ &= na^n + a^n \cdot 1 \\ &= na^n + a^n \\ &= (n+1)a^n, \quad \forall a \end{aligned}$$

Which is exactly the formula for the case $n + 1$ □

1.9. Note: With this, we can find the derivative of any polyonmial functions, for a polynomial with degree n,

$$f^{(n)}(x) = n!a_n$$

and for $k > n$, $f^{(k)}(x) = 0$

1.10. Theorem: *If g is differentiable at a, and $g(a) \neq 0$, then $\frac{1}{g}$ is differentiable at a, and*

$$(\frac{1}{g})'(a) = \frac{-g'(a)}{[g(a)]^2}$$

Proof. As always, we have

$$\frac{(\frac{1}{g})(a+h) - (\frac{1}{g})(a)}{h}$$

For sufficiently small h, we have to verify that $(\frac{1}{g})(a+h)$ is defined. We know that g is differentiable at a, therefore g is continuous at a. And it follows from 6.6 that there is some $\delta > 0$ such that $g(a+h) \neq 0$ for $|h| < \delta$. So the equation does make sense for small enough h.

$$\begin{aligned} \lim_{h \to 0} \frac{\left(\frac{1}{g}\right)(a+h) - \left(\frac{1}{g}\right)(a)}{h} &= \lim_{h \to 0} \frac{\frac{1}{g(a+h)} - \frac{1}{g(a)}}{h} \\ &= \lim_{h \to 0} \frac{g(a) - g(a+h)}{h[g(a) \cdot g(a+h)]} \\ &= \lim_{h \to 0} \frac{-[g(a+h) - g(a)]}{h} \cdot \frac{1}{g(a)g(a+h)} \\ &= \lim_{h \to 0} \frac{-[g(a+h) - g(a)]}{h} \cdot \lim_{h \to 0} \frac{1}{g(a) \cdot g(a+h)} \end{aligned}$$

$$= -g'(a) \cdot \frac{1}{[g(a)]^2}.$$

□

1.11. Theorem: *If f and g are differentiable at a and $g(a) \neq 0$, then f/g is differentiable at a, and*

$$(\frac{f}{g})'(a) = \frac{g(a) \cdot f'(a) - f(a) \cdot g'(a)}{[g(a)]^2}$$

Proof. Since $\frac{f}{g} = f \cdot (\frac{1}{g})$ we have

$$\begin{aligned}\frac{f}{g}'(a) &= (f \cdot \frac{1}{g})'(a)\\ &= f'(a) \cdot \frac{1}{g}(a) + f(a) \cdot \frac{1}{g}'(a)\\ &= \frac{f'(a)}{g(a)} + \frac{f(a)(-g'(a))}{[g(a)]^2}\\ &= \frac{f'(a) \cdot g(a) - f(a) \cdot g'(a)}{[g(a)]}^2\end{aligned}$$

□

1.12. Theorem: *If g is differentiable at a, and f is differentiable at $g(a)$, then $f \circ g$ is differentiable at a, and*

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a).$$

Proof. Later. □

Section 2. Significance of the Derivative

2.1. Definition: Let f be a function and A a set of numbers contained in the domain of f. A point x in A is a **maximum point** for f on A if

$$f(x) \geq f(y) \text{ for every } y \in A$$

The number $f(x)$ itself is called the **maximum value** of f on A (and we also say that f has its maximum value on A at x)

2.2. Remark: Notice that a function f can have several different maximum points on A, however it can have at most one maximum value. We are typically interested in the case where A is a closed interval $[a, b]$, if f is continuous, then 7.8 guarantees that f does indeed have a maximum value on $[a, b]$

2.3. Theorem: *Let f be any function defined on (a, b). If x is a maximum (or a minimum) point for f on (a, b), and f is differentiable at x, then $f'(x) = 0$.*

Proof. WLOG, consider the case where f has a maximum at x. If h is any number such that $x + h$ is in (a, b), then $f(x) \geq f(x + h)$, and thus $f(x + h) - f(x) \leq 0$. Thus if $h > 0$ we have

$$\frac{f(x+h) - f(x)}{h} \leq 0$$

and consequently

$$\lim_{h \to 0^+} \frac{f(x+h) - f(x)}{h} \leq 0$$

. On the other hand, if $h < 0$, we have

$$\frac{f(x+h) - f(x)}{h} \geq 0 \implies \lim_{h \to 0^+} \frac{f(x+h) - f(x)}{h} \geq 0$$

By our hypothesis, f is differentiable at x so these two limits must be equal each other, and in fact, equal to $f'(x)$. This means that

$$f'(x) \leq 0 \text{ and } f'(x) \geq 0$$

from which it follows that $f'(x) = 0$. □

2.4. Definition: Let f be a function, and A a set of numbers contained in the domain of f. A point x in A is a **local maximum [minimum] point** for f on A if there is some $\delta > 0$ such that x is a maximum [minimum] point for f on $A \cap (x - \delta, x + \delta)$.

2.5. Theorem: *If x is a local maximum or minimum for f on (a, b) and f is differentiable at x, then $f'(x) = 0$*

Proof. Trivial □

2.6. Warning: The converse of theorem 2 is obviously not true.

$$\mathbf{f'(x) = 0 \text{ does not imply that } x \text{ is a local maximum or minimum point of } } f$$

. Consider the simplest example $f(x) = x^3$, in this case $f'(0) = 0$, but f has no local maximum of minimum anywhere.

2.7. Definition: A **critical point** of a function f is a number x such that

$$f'(x) = 0.$$

The number $f(x)$ iteself is called a **critical value** of f.

In order to locate the maximum and minimum of f, we have to consider three kinds of points:

- The critical points of f in $[a, b]$
- The end points a and b.
- Points x in $[a, b]$ such that f is not differentiable at x.

If x is the max/min on $[a, b]$, then it must be in one of the three classes listed above. For if x is not in the second or third group, then x is in (a, b) and f is differentiable at x, and by 2.3, this means that x is in the first group.
If there are many points in these three categories, it may be impossible to find the maximum and minimum of f. But when there are only a few critical points and a few points where f is not differentiable. We can simply find $f(x)$ for each x satisfying $f'(x) = 0$ or where f is not differentiable at x, and of course, $f(a)$ and $f(b)$. The biggest of these will be the maximum value of f, and the smallest will be the minimum.

2.8. Theorem (Rolles Theorem): *If f is continuous on $[a, b]$ and differentiable on (a, b), and $f(a) = f(b)$, then there is a number x in (a, b) such that $f'(x) = 0$.*

Proof. If follows from the continuity on f on $[a, b]$ that f has a maximum and a minimum value on $[a, b]$. Suppose first that the maximum value occurs at a point x in (a, b). Then $f'(x) = 0$ by 2.3, and we are done.
Suppose that the minimum value of f occurs at some point x in (a, b). Then, again, $f'(x) = 0$ by 2.3.
Finally, suppose the maximum and minimum values both occur at the end points. Since $f(a) = f(b)$, the maximum and minimum values of f are equal, so f is a constant function, and for a constant function we can choose any x in (a, b). □

2.9. Theorem (Mean Value Theorem): *if f is continuous on $[a, b]$ and differentiable on (a, b), then there is a number x in (a, b) such that*

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

Proof. Let

$$h(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

Clearly, h is continuous on $[a, b]$ and differentiable on (a, b), and

$$h(a) = f(a). \tag{2.1}$$

$$h(b) = f(b) - \frac{f(b) - f(a)}{b - a}(x - a) \tag{2.2}$$

$$= f(a). \tag{2.3}$$

Consequently, we may apply Rolle's Theorem to h and conclude that there is some x in (a, b) such that

$$0 = h'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}.$$

so then

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

□

2.10. Corollary: *If f is defined on an interval and $f'(x) = 0$ for all x in the interval, then f is constant on the interval.*

Proof. Let a and b be any two points in the interval with $a \neq b$. Then there is some x in (a, b) such that

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

But $f'(x) = 0$ for all x in the interval, so

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

and so $f(a) = f(b)$. Thus the value of f at any two points in the interval is the same, that is, f is constant on the interval. □

2.11. Corollary: *If f and g are defined on the same interval, and $f'(x) = g'(x)$ for all x in the interval, then there is some number c such that $f = g + c$.*

Proof. For all x in the interval we have $(f - g)'(x) = f'(x) - g'(x) = 0$, and by last corollary, there exist a number c such that $f - g = c$ □

2.12. Definition: A function is **increasing** on an interal if $f(a) < f(b)$ whenever a and b are two numbers in the interval with $a < b$. The function f is **decreasing** on an interval if $f(a) > f(b)$ for all a and b in the interval with $a < b$.

2.13. Corollary: *If $f'(x) > 0$ for all x in an interval, then f is increasing on the interval; if $f'(x) < 0$ for all x in the interval, then f is decreasing on the interval.*

Proof. WLOG, consider the case where $f'(x) > 0$. Let a and b be two points in the interval with $a < b$. Then there is some x in (a, b) with

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

But $f'(x) > 0$ for all x in (a, b), so

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

Since $b - a > 0$ it follows that $f(b) > f(a)$. An analogous proof can be given when $f'(x) < 0$ for all x. □

2.14. Theorem: *Suppose $f'(a) = 0$. If $f''(a) > 0$, then f has a local minimum at a; if $f''(a) < 0$, then f has a local maximum at a.*

Proof. By definition,

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

Since $f'(a) = 0$, this can be written

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

Suppose now that $f''(a) > 0$. Then $\frac{f'(a+h)}{h}$ must be positive for sufficiently small h. Therefore, $f'(a+h)$ must be positive for sufficiently small $h > 0$ and $f'(a+h)$ must be negative for sufficiently small $h < 0$. By 2.13, f is increasing in som einterval to the right of a and decreasing in some interval to the left of a. Thus f has a local minimum at a. The proof for the case $f''(a) < 0$ is similar. □

2.15. Theorem: *Suppose $f''(a)$ exists. If f has a local minimum at a, then $f''(a) \geq 0$; if f has a local maximum at a, then $f''(a) \leq 0$.*

Proof. Suppose f has a local minimum at a. If $f''(a) < 0$, then $f''(a) < 0$, then f would also have a local maximum at a, by the last theorem. Then f would be constant in some interval containing a, so that $f''(a) = 0$, a contradiction. Thus we must have $f''(a) \geq 0$. The case of a local maximum is hanled similarly. □

2.16. Remark: Note that 2.15 is only a partial converse of 2.14, that is, the $\geq$ and $\leq$ cannot be replaced by $>$ and $<$.

2.17. Theorem: *Suppose that f is continuous at a, and that $f'(x)$ exists for all x in some interval containing a, except perhaps for $x = a$. Suppose, moreover, that $\lim_{x\to a} f'(x)$ exists. Then $f'(a)$ also exists, and*

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

Proof. By definition,

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

For sufficiently small $h > 0$ the function f will be continuous on $[a, a+h]$ and differentiable on $(a, a+h)$ (a similar assertion holds for sufficiently small $h < 0$). by MVT there is a number α_h in $(a, a+h)$ such that

$$\frac{f(a+h) - f(a)}{h} = f'(a_h).$$

Now α_h approaches a as h approaches 0, because α_h is in $(a, a+h)$; since $\lim_{x\to a} f'(x)$ exists, it follows that

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

□

2.18. Remark: By the theorem above, the graph of f' can never exhibit a removable discontinuity.

2.19. Theorem (The Cauchy MVT): *If f and g are continuous on $[a, b]$ and differentiable on (a, b), then there is a number x in (a, b) such that*

$$[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x).$$

If $g(b) \neq g(a)$, and $g'(x) \neq 0$, this equation can be written as

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x)}{g'(x)}.$$

2.20. Remark: Notice that if $g(x) = x$ for all x ,then $g'(x) = 1$, and we obtain MVT. On the other hands, applying MVT to f and g separately, we find that there are x and y in (a, b) with

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x)}{g'(y)}.$$

however there is no guarantee that the x and y found in this way will be equal.

Proof. Let

$$h(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)].$$

Then h is continuous on $[a, b]$, differentiable on (a, b), and

$$h(a) = f(a)g(b) - g(a)f(b) = h(b).$$

It follows from Rolle's Theorems that $h'(x) = 0$ for some x in (a, b), which means that

$$0 = f'(x)[g(b) - g(a)] - g'(x)[f(b) - f(a)].$$

Rearrange and we get the desired equation. □

2.21. Theorem (L'Hopital's Rule): *Suppose that* $\lim_{x\to a} f(x) = 0$ *and* $\lim_{x\to a} g(x) = 0$ *and suppose also that*

$$\lim_{x\to a} \frac{f'(x)}{g'(x)}$$

exists. Then $\lim_{x\to a} \frac{f(x)}{g(x)}$ *exists, and*

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}$$

Proof. The hypothesis that $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ exists contains two implicit assumptions:

(1). there is an interval $(a - \delta, a + \delta)$ such that $f'(x)$ and $g'(x)$ exist for all x in $(a - \delta, a + \delta)$ except, perhaps, for $x = a$,

(2). in this interval $g'(x) \neq 0$ with the possible exception of $x = a$

Let $f(a) = g(a) = 0$, then f and g are continuous at a. If $a < x < a + \delta$, then MVT and Cauchy MVT apply to f and g on the interval $[a, x]$ (and a similar statement holds for $a - \delta < x < a$). First applying MVT to g, we see that $g(x) \neq 0$, for if $g(x) = 0$ there would be some x_1 in (a, x) with $g'(x_1) = 0$, contradicting (2). Now applying the Cauchy MVT to f and g, we see that there is a number α_x in (a, x) such that

$$[f(x) - 0]g'(\alpha_x) = [g(x) - 0]f'(\alpha_x)$$

or since $g'(\alpha_x) \neq 0$,

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

Now α_x approaches a as x approaches a, because α_x is in (a, x); since we are assuming that

$\lim_{y\to a} \frac{f'(y)}{g'(y)}$ exists, it follows that

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(\alpha_x)}{g'(\alpha_x)} = \lim_{y\to a} \frac{f'(y)}{g'(y)}.$$

Alternate $\delta - \epsilon$ proof:
Since we know $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ exists, let L be our limit. then

$$\forall \epsilon > 0, \exists \delta > 0, s.t. |a - x| < \delta \implies |L - \frac{f'(x)}{g'(x)}| < \epsilon.$$

But $|\alpha_x - a| < |a - x| < \delta$, so for each ϵ we can use the same δ to see that

$$|x - a| < \delta \implies |\frac{f'(\alpha_x)}{g'(\alpha_x)} - L| = |\frac{f(x)}{g(x)} - L| < \epsilon$$

As required. □

2.22. Definition: A function f is **convex** on an interval, if for all a and b in the interval, the line segment joining $(a, f(a))$ and $(b, f(b))$ lies above the graph of f.

2.23. Note: Sometimes, an analytic definition might be more useful. The straight line between $(a, f(a))$ and $(b, f(b))$ is the graph of the function g defined by

$$g(x) = \frac{f(b) - f(a)}{b - a}(x - a) + f(a).$$

For this line to lie above the grpah of f is just

$$\frac{f(b) - f(a)}{b - a}(x - a) + f(a) > f(x)$$

or

$$\frac{f(b) - f(a)}{b - a}(x - a) > f(x) - f(a)$$

or

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

Therefore, we have an equivalent definition of convexity.

2.24. Definition: A function f is **convex** on an interval if for a, x, and b in the interval with $a < x < b$ we have

$$\frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(a)}{b - a}.$$

2.25. Definition: If the inequality in the last definition is replaced by

$$\frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(a)}{b - a}.$$

Then we obtain the definition of a **concave function**

2.26. Remark: All concave functions are precisely the ones of the form $-f$, where f is convex. So for all theorem below about convex function have immediate corollaries about concave functions as well.

2.27. Theorem: *Let f be convex. If f is differentiable at a, then the graph of f lies above the tangent line through $(a, f(a))$, except at $(a, f(a))$ it self. If $a < b$ and f is differentiable at a and b, then $f'(a) < f'(b)$.*

Proof. Let $0 < h_1 < h_2$, then using 2.24 on $a < a + h_1 < a + h_2$, we get

$$\frac{f(a+h_1) - f(a)}{h_1} < \frac{f(a+h_2) - f(a)}{h_2}$$

This inequality shows that the values of $\frac{f(a+h)-f(a)}{h}$ decreasing as $h \to 0^+$. Consequently,

$$f'(a) < \frac{f(a+h) - f(a)}{h} \text{ for } h > 0$$

Which means that for $h > 0$, the secant line through $(a, f(a))$ and $(a + h, f(a + h))$ has a larger slope than the tangent line, which implies that $(a + h, f(a + h))$ lies above the tangent line. An analogous argument can be used for negative h. Let $h_2 < h_1 < 0$, then

$$\frac{f(a+h_1) - f(a)}{h_1} > \frac{f(a+h_2) - f(a)}{h_2}$$

Which shows that the slope of the tangent line through $(a, f(a))$ is greater than

$$\frac{f(a+h) - f(a)}{h} \text{ for } h < 0$$

Therefore, $f(a + h)$ lies above the tangent line for $h < 0$ as well, proving the first part of the theorem.
Now suppose that $a < b$, then, from the last part,

$$f'(a) < \frac{f(a + (b-a)) - f(a)}{b-a} = \frac{f(b) - f(a)}{b-a} \quad \text{since } h = b - a > 0$$

and

$$f'(b) > \frac{f(b + (a-b)) - f(b)}{a-b} = \frac{f(b) - f(a)}{b-a} \quad \text{since } h = a - b > 0.$$

Combing the inequalities gives us $f'(a) < f'(b)$. $\square$

2.28. Note: This theorem has two converse, to make our proofs easier, we will first prove a lemma which says if f' is increasing, then the graph of f lies below any secant line which happens to be horizontal.

2.29. Lemma: *Suppose f is differentiable and f' is increasing. If $a < b$ and $f(a) = f(b)$, then $f(x) < f(a) = f(b)$ for $a < x < b$.*

Proof. Suppose that $f(x) \geq f(a) = f(b)$ for some x in (a, b). Then the maximum of f on $[a, b]$ occurs at some point x_0 in (a, b) with $f(x_0) \geq f(a)$ and $f'(x_0) = 0$. On the other hand, applying the Mean Value Theorem to the interval $[a, x_0]$, we find that there is x_1 with $a < x_1 < x_0$ and

$$f'(x_1) = \frac{f(x_0) - f(a)}{x_0 - a} \geq 0.$$

Contradicting the fact that f' is increasing. □

2.30. Theorem: *If f is differentiable and f' is increasing, then f is convex.*

Proof. Let $a < b$. Define g by

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

It is obvious that g' is also increasing, moreover, $g(a) = g(b) = f(a)$, applying our lemma to g we conclude that

$$g(x) < f(a) \text{ if } a < x < b$$

In other words, if $a < x < b$, then

$$f(x) - \frac{f(b) - f(a)}{b - a}(x - a) < f(a)$$

or

$$\frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(a)}{b - a}.$$

Hence f is convex. □

2.31. Theorem: *If f is differentiable and the graph of f lies above each tangent line except at the point of contact, then f is convex.*

Proof. Let $a < b$. The tangent line at $(a, f(a))$ is the graph of the function

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

and since $(b, f(b))$ lies above the tangent line, we have

$$f(b) > f'(a)(b - a) + f(a).$$

Similarly, since the tangent line at $(b, f(b))$ is the graph of

$$h(x) = f'(b)(x - b) + f(b),$$

and $(a, f(a))$ lies above the tangent line at $(b, f(b))$, we have

$$f(a) > f'(b)(a - b) + f(b).$$

It then follows from the two inequality that $f'(a) < f'(b)$, and by last theorem, f is convex. □

2.32. Theorem: *If f is differentiable on an interval and intersects each of its tangent lines just once, then f is either convex or concave on that interval.*

Proof. This proof is split in two parts:
(1) First we claim that no straight line can intersect the graph of f in three different points. Suppose not, and that some straight line did intersect the graph of f at $(a, f(a)), (b, f(b))$, and $(c, f(c))$, with $a < b < c$. Then we would have

$$\frac{f(b) - f(a)}{b - a} = \frac{f(c) - f(a)}{c - a}. \tag{2.4}$$

Consider the function

$$g(x) = \frac{f(x) - f(a)}{x - a} \text{ for } x \text{ in } [b, c].$$

Equation (2.4) says that $g(b) = g(c)$, so by Rolle's Theorem, there exist a number x in (b, c) where $0 = g'(x)$, and thus

$$0 = (x - a)f'(x) - [f(x) - f(a)]$$

or

$$f'(x) = \frac{f(x) - f(a)}{x - a}.$$

But this says that the tangent line at $(x, f(x))$ passes through $(a, f(a))$, contradicting the hypotheses.
(2) Suppose that $a_0 < b_0 < c_0$ and $a_1 < b_1 < c_1$ are points in the interval. Let

$$x_t = (1 - t)a_0 + ta_1 \tag{2.5}$$
$$y_t = (1 - t)b_0 + tb_1 \quad 0 \leq t \leq 1. \tag{2.6}$$
$$z_t = (1 - t)c_0 + tc_1 \tag{2.7}$$

Then $x_0 = a_0$ and $x_1 = a_1$ and the points x_t all lie between a_0 and a_1, with analogous statements for y_t and z_t. Moreover, $x_t < y_t < z_t$ for $0 \leq t \leq 1$.
Now consider the function

$$g(t) = \frac{f(y_t) - f(x_1)}{y_t - x_t} - \frac{f(z_t) - f(x_t)}{z_t - x_t} \qquad \text{for } 0 \leq t \leq 1.$$

By step 1, $g(t) \neq 0$ for all t in $[0, 1]$. So either $g(t) > 0$ for all t in $[0, 1]$ or $g(t) < 0$ for all t in $[0, 1]$. Thus, either f is convex or concave. □

Chapter 3

Inverse Functions

2.33. Definition: A function f is **one-one** if $f(a) \neq f(b)$ whenever $a \neq b$

2.34. Definition: For any function f, the **inverse** of f, denoted by f^{-1}, is the set of all pairs (a, b) for which the pairs (b, a) is in f.

2.35. Theorem: *f^{-1} is a function if and only if f is one-one*

Proof. Suppose first that f is one-one. Let (a, b) and (a, c) be two pairs in f^{-1}. Then (b, a) and (c, a) are in f, so $a = f(b)$ and $a = f(c)$. Since f is one-one this implies that $b = c$. Thus f^{-1} is a function.
Conversely, suppose that f^{-1} is a function. If $f(b) = f(c)$, then f contains the pairs $(b, f(b))$ and $(c, f(c)) = (c, f(b))$, so $(f(b), b)$ and $(f(b), c)$ are in f^{-1}. Since f^{-1} is a function this implies that $b = c$. Thus f is one-one □

2.36. Remark: Some important things to remember about inverse functions:

- $(f^{-1} \circ f)(a) = a$
- $(f \circ f^{-1})(a) = a$
- f and f^{-1} are reflected through the line $f(x) = x$
- if f is increasing then f^{-1} is also increasing
- if f is decreasing then f^{-1} is also decreasing
- if f is one-to-one, then $codom(f) = dom(f^{-1})$ and $codom(f^{-1}) = dom(f)$

2.37. Theorem: *If f is continuous and one-one on an interval, then f is either increasing or decreasing on that interval.*

Proof. **(1)** If $a < b < c$ are three points in the interval, then either (i) $f(a) < f(b) < f(c)$ or (ii) $f(a) > f(b) > f(c)$. If not, for example, that $f(a) < f(c)$. If we had $f(b) < f(a)$, then by IVT applied to the interval $[b, c]$ would give an x with $b < x < c$ and $f(x) = f(a)$. contradicting

the fact that f is one-one on $[a, c]$. Similarly, $f(b) > f(c)$ would leader to a contradiction, so $f(a) < f(b) < f(c)$.
(2) If $a < b < c < d$ are four points in the interval, then $f(a) < f(b) < f(c) < f(d)$ or $f(a) > f(b) > f(c) > f(d)$ if we just apply (1) to $a < b < c$ and $b < c < d$.
(3) Now take any $a < b$ in the interval, and suppose that $f(a) < f(b)$. Then f is increasing, for if c and d are any two points, we can apply (2) to the collection of $\{a, b, c, d\}$. ☐

2.38. Remark: Suppose f is continuous and one-one on $I = [a, b]$, then $dom(f) = [a, b]$ and $codom(f) = [f(a), f(b)]$ if f is increasing, and $codom(f) = [f(b), f(a)]$ is f is decreasing.
If the domain of f is an open interval, thus having one of the the forms $(a, b), (-\infty, b), (a, \infty)$, or $\mathbb{R}$, then the codomain of f (domain of f^{-1}) will also have one of these forms

2.39. Theorem: *If f is continuous and one-one on an interval, then f^{-1} is also continuous.*

Proof. We know by the last theorem that f is either increasing or decreasing. So WLOG, lets assume that f is increasing, since we can just take care of the other case by just considering $-f$. Lets also assume that our interval is open, since any continuous one-one function on any interval can be extended to one on a larger open interval. To show continuity, we must show that $\lim_{x \to b} f^{-1}(x) = f^{-1}(b)$ for each b in the domain of f^{-1}. Such a number b is of the form $f(a)$ for some a in the domain of f, and $f^{-1}(b) = a$. For all $\epsilon > 0$, we want to find a $\delta > 0$ such that, for all x,

$$\text{if } f(a) - \delta < x < f(a) + \delta, \text{ then } a - \epsilon < f^{-1}(x) < a + \epsilon.$$

Since $a - \epsilon < a < a + \epsilon$, it follows that

$$f(a - \epsilon) < f(a) < f(a + \epsilon)$$

since $a - \epsilon < a < a + \epsilon$, it follows that $f(a - \epsilon) < f(a) < f(a + \epsilon)$; we let δ be the smaller of $f(a + \epsilon) - f(a)$ and $f(a) - f(a - \epsilon)$. Our choice of δ ensures that

$$f(a - \epsilon) \leq f(a) - \delta \quad \text{and} \quad f(a) + \delta \leq f(a + \epsilon).$$

consequently, if

$$f(a) - \delta < x < f(a) + \delta.$$

then

$$f(a - \epsilon) < x < f(a + \epsilon).$$

Since f is increasing, f^{-1} is also increasing, and we obtain

$$f^{-1}(f(a - \epsilon)) < f^{-1}(x) < f^{-1}(f(a + \epsilon))$$

which is

$$a - \epsilon < f^{-1}(x) < a + \epsilon.$$

which is precisely what we want. ☐

2.40. Theorem: *If f is a continuous one-one function defined on an interval and $f'(f^{-1}) = 0$, then f^{-1} is not differentiable at a.*

Proof. We have

$$f(f^{-1}(x)) = x.$$

If f^{-1} were differentiable at a, the Chain Rule would imply that

$$f'(f^{-1}(a)) \cdot (f^{-1})'(a) = 1,$$

hence

$$0 \cdot (f^{-1})'(a) = 1$$

which is impossible. □

2.41. Theorem: *Let f be a continuous one-one function defined on an interval, and suppose that f is differentiable at $f^{-1}(b)$, with derivative $f'(f^{-1}(b)) \neq 0$. Then f^{-1} is differentiable at b, and*

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))}.$$

Proof. Let $b = f(a)$. Then

$$\lim_{h \to 0} \frac{f^{-1}(b+h) - f^{-1}(b)}{h} = \lim_{h \to 0} \frac{f^{-1}(b+h) - a}{h}$$

Every number $b + h$ in the domain of f^{-1} can be written in the form

$$b + h = f(a + k)$$

for a unique k, then

$$\lim_{h \to 0} \frac{f^{-1}(b+h) - a}{h} = \lim_{h \to 0} \frac{f^{-1}(f(a+k)) - a}{f(a+k) - b} = \lim_{h \to 0} \frac{k}{f(a+k) - f(a)} \quad (1).$$

Now since $b+h = f(a+k)$, then $f^{-1}(b+h) = a+k$ or $k = f^{-1}(b+h) - f^{-1}(b)$ By the last theorem, the function f^{-1} is continuous at b. since k is a function of h, this means that k approaches 0 as h approaches 0. Since

$$\lim_{k \to 0} \frac{f(a+k) - f(a)}{k} = f'(a) = f'(f^{-1}(b)) \neq 0,$$

this implies that

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b)).}$$

By plugging in the inverse to (1) □

Epsilon-delta approach. We can also instead use $\epsilon - \delta$, knowing that f is differentiable at $f^{-1}(b) = a$, then

$$\forall \epsilon > 0, \exists \delta_1, s.t., 0 < |k| < \delta_1 \implies |\frac{f(a+k) - f(a)}{k} - f'(a)| < \epsilon$$

and since f^{-1} is also continuous at b by the last theorem, we have

$$\forall \delta_1 > 0, \exists \delta_2 > 0, s.t. 0 < |h| < \delta \implies |f^{-1}(b+h) - f^{-1}(b)| < \delta_1$$

Take $k = f^{-1}(b+h) - f^{-1}(b)$ and connect the two $\epsilon - \delta$,

$$0 < |h| < \delta_2 \implies 0 < |k| < \delta_1 => |\frac{f(a+k) - f(a)}{k} - f'(a)| < \epsilon$$

That is, as h approaches 0, the derivative at a is $f'(a) = f'(f^{-1}(b))$.
This means, let $g = f^{-1}$, then $g'(x) = \frac{1}{f'(g(x))}$ □

Chapter 4

Integrals

2.42. Definition: Let $a < b$, a **partition** of the interval $[a, b]$ is a finite collection of points in $[a, b]$, one of which is a, and one of which is b.

2.43. Definition: Suppose f is bounded on $[a, b]$ and $P = \{t_0, ..., t_n\}$ is a partition of $[a, b]$. Let

$$m_i = inf\{f(x) : t_{i-1} \leq x \leq t_i\}.$$

$$M_i = sup\{f(x) : t_{i-1} \leq x \leq t_i\}.$$

The **lower sum** of f for P, denoted by $L(f, P)$, is defined as

$$L(f, P) = \sum_{i=1}^{n} m_i(t_i - t_{i-1}).$$

The **upper sum** of f for P, denoted by $U(f, P)$, is defined as

$$U(f, P) = \sum_{i=1}^{n} M_i(t_i - t_{i-1}).$$

2.44. Remark: The following lemma is a very important lemma which says, that more points in a partition P result in a better approximation of the region $R(f, a, b)$.

2.45. Lemma: *If Q contains P, then*

$$L(f, P) \leq L(f, Q),$$

$$U(f, P) \geq U(f, Q).$$

Proof. Consider first the special case in which Q contains just one more point than P:

$$P = \{t_0, ..., t_n\}$$
$$Q = \{t_0, ..., t_{k-1}, u, t_k, ..., t_n\}.$$

where $a = t_0 < t_1 < \cdots < t_{k-1} < u < t_k < \cdots < t_n = b$. Let

$$m' = inf\{f(x) : t_{k-1} \leq x \leq u\}, \tag{4.1}$$

$$m'' = inf\{f(x) : u \leq x \leq t_k\}. \tag{4.2}$$

Then

$$L(f, P) = \sum_{i=1}^{n} m_i(t_i - t_{i-1}),$$

$$L(f, Q) = \sum_{i=1}^{k-1} m_i(t_i - t_{i-1}) + m'(n - t_{k-1}) + m''(t_k - u) + \sum_{i=k+1}^{n} m_i(t_i - t_{i-1})$$

To prove that $L(f, P) \leq L(f, Q)$ it therefore suffices to show that

$$m_k(t_k - t_{k-1}) \leq m'(n - t_{k-1}) + m''(t_k - u)$$

Now the set $\{f(x) : t_{k-1} \leq x \leq t_k\}$ contains all the numbers in $\{f(x); t_{k-1} \leq x \leq u\}$, and possibly some smaller ones, so the infimum of the first set is *less than or equal to* the infimum of the second, thus

$$m_k \leq m'.$$

Similarly,

$$m_k \leq m''.$$

Therefore,

$$m_k(t_k - t_{k-1}) = m_k(u - t_{k-1}) + m_k(t_k - u) \leq m'(u - t_{k-1}) + m''(t_k - u).$$

This proves, in the special case, that $L(f, P) \leq L(f, Q)$. An analogous proof can be given to show that $U(f, P) \geq U(f, Q)$. The general case can now be easily deduced. The partition Q can be obtained from P by adding one point at a time, that is, there is a sequence of partitions

$$P = P_1, P_2, ..., P_\alpha = Q$$

such that P_{k+1} contains juts one more point than P_j. Then

$$L(f, P) = L(f, P_1) \leq L(f, P_2) \leq \cdots \leq L(f, P_\alpha) = L(f, Q).$$

and

$$U(f, P) = U(f, P_1) \geq U(f, P_2) \geq \cdots \geq U(f, P_\alpha) = U(f, Q).$$

□

2.46. Theorem: *Let P_1 and P_2 be partitions of $[a, b]$, and let f be a function which is founded on $[a, b]$. Then*

$$L(f, P_1) \leq U(f, P_2).$$

Proof. There is a partition P which contains both P_1 and P_2. According to the lemma,

$$L(f, P_1) \leq L(f, P) \leq U(f, P) \leq U(f, P_2)$$

□

2.47. Remark: It follows from the theorem that any upper sum $U(f, P')$ is an upper bound for the set of all lower sums $L(f, P)$. Consequently, any upper sum $U(f, P')$ is greater than or equal to the least upper bound of all lower sums:

$$\sup\{L(f, P) : P \text{ a partition of } [a, b]\} \leq U(f, P'),$$

for every P'. This, in turn, also means that $\sup\{L(f, P)\}$ is a lower bound for the set of all upper sums of f. Consequently,

$$\sup\{L(f, P)\} \leq \inf\{U(f, P)\}.$$

2.48. Definition: A function f which is bounded on $[a, b]$ is **integrable** on $[a, b]$ if

$$\sup\{L(f, P) : P \text{ a partition of } [a, b]\} = \inf\{U(f, P) : P \text{ a partition of } [a, b]\}$$

In this case, the common number is called the **integral** of f on $[a, b]$ and is denoted by

$$\int_a^b f.$$

2.49. Theorem: *If f is bounded on $[a, b]$, then f is integrable on $[a, b]$ if and only if for every $\epsilon > 0$ there is a partition P of $[a, b]$ such that*

$$U(f, P) - L(f, P) < \epsilon$$

Proof. Suppose first that for every $\epsilon > 0$ there is a partition P with

$$U(f, P) - L(f, P) < \epsilon.$$

Since

$$\inf\{U(f, P')\} \leq U(f, P),$$

$$\sup\{L(f, P')\} \geq L(f, P),$$

it follows that

$$\inf\{U(f, P')\} - \sup\{L(f, P')\} < \epsilon.$$

it follows that

$$\inf\{U(f, P')\} - \sup\{L(f, P')\} < \epsilon.$$

Since this is true for all $\epsilon > 0$, it follows that

$$\sup\{L(f, P')\} = \inf\{U(f, P')\};$$

by definition, then, f is integrable. The proof of the converse assertion is similar: If f is integrable, then

$$\sup\{L(f, P)\} = \inf\{U(f, P)\}$$

This means that for each $\epsilon > 0$ there are partitions P', P'' with

$$U(f, P'') - L(f, P') < \epsilon$$

Let P be a partition which contains both P' and P''. Then, according to the lemma,

$$U(f, P) \leq U(f, P''),$$

$$L(f, P) \geq L(f, P');$$

consequently,

$$U(f, P) - L(f, P) \leq U(f, P'') - L(f, P') < \epsilon$$

□

2.50. Theorem: *If f is continuous on $[a, b]$, then f is integrable on $[a, b]$.*

Proof. Since f is continuous on $[a, b]$, it must be bounded on $[a, b]$. To show that f is integrable on $[a, b]$, we have to use the theorem 2.49 and show that for all $\epsilon > 0$, there exist a partition P of $[a, b]$ such that

$$U(f, P) - L(f, P) < \epsilon$$

As proved in Chapter 8, f is continuous on $[a, b]$ if and only if f is uniformly continuous on $[a, b]$. Thus there is some δ such that for all $x, y \in [a, b]$

$$|x - y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{2(b-a)}$$

Now the trick is to choose a partition $P = \{t_0, ..., t_n\}$ such that each $|t_i - t_{i-1}| < \delta$, then for each i we have

$$|f(x) - f(y)| < \frac{\epsilon}{2(b-a)} \text{ for all} x, y \in [t_{i-1}, t_i]$$

it follows easily that

$$M_i - m_i \leq \frac{\epsilon}{2(b-a)} < \frac{\epsilon}{b-a}.$$

Since this is true for all i, we have

$$\begin{aligned} U(f, P) - L(f, P) &= \sum_{i=1}^{n} (M_i - m_i)(t_i - t_{i-1}) && (4.3) \\ &< \frac{\epsilon}{b-a} \sum_{i=1}^{n} t_i - t_{i-1} && (4.4) \\ &= \frac{\epsilon}{b-a} \cdot b - a && (4.5) \\ &= \epsilon. && (4.6) \end{aligned}$$

as required. □

2.51. Theorem: *Let $a < c < b$. If f is integrable on $[a, b]$, then f is integrable on $[a, c]$ and $[c, b]$. Conversely, if f integrable on $[a, c]$ and on $[c, b]$ then f is integrable on $[a, b]$. Finally, if f is integrable on $[a, b]$, then*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Proof. Suppose f is integrable on $[a, b]$, then for all $\epsilon > 0$, there is a partition $P = \{t_0, ..., t_n\}$ of $[a, b]$ such that

$$U(f, P) - L(f, P) < \epsilon$$

Lets assume that $c = t_j$ for some j. (Otherwise, let Q be the partition which contains $t_0, ..., t_j$ and c, then Q contains P, and $U(f, Q) - L(f, Q) \leq U(f, P) - L(f, P) < \epsilon$.)
Now $P' = \{t_0, ..., t_j\}$ is a partition of $[a, c]$ and $P'' = \{t_j, ..., t_n\}$ is a partition of $[c, b]$. Since

$$L(f, P) = L(f, P') + L(f, P'').$$
$$U(f, P) = U(f, P') + U(f, P'').$$

we have

$$[U(f, P') - L(f, P')] + [U(f, P'') - L(f, P'')] = U(f, P) - L(f, P) < \epsilon.$$

Since each square bracket is nonnegative, they are each less that ϵ. This shows that f is integrable on $[a, c]$ and $[c, b]$, note also that

$$L(f, P') \leq \int_a^c f \leq U(f, P')$$

$$L(f, P'') \leq \int_c^b f \leq U(f, P'')$$

so that

$$L(f, P) \leq \int_a^c f + \int_c^b f \leq U(f, P).$$

Since this is true for any P, this proves that

$$\int_a^c f + \int_c^b f = \int_a^b f.$$

Now to prove the converse, suppose that f integrable on $[a, c]$ and on $[c, b]$. If $\epsilon > 0$, there is a partition P' of $[a, c]$ and a partition P'' of $[c, b]$ such that

$$U(f, P') - L(f, P') < \frac{\epsilon}{2}.$$

$$U(f, P'') - L(f, P'') < \frac{\epsilon}{2}.$$

If P is the partition of $[a, b]$ containing all the points of P' and P'', then

$$L(f, P) = L(f, P') + L(f, P''),$$
$$U(f, P) = U(f, P') + U(f, P'').$$

Consequently,

$$U(f,P) - L(f,P) = [U(f,P') - L(f,P')] + [U(f,P'') - L(f,P'')] < \epsilon.$$

□

2.52. Remark: With this, we can now add the definition

$$\int_a^a f = 0 \quad \text{and} \quad \int_a^b f = -\int_b^a f \text{ if } a > b.$$

2.53. Theorem: *If f and g are integrable on $[a,b]$, then $f+g$ is integrable on $[a,b]$ and*

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g.$$

Proof. Let $P = \{t_0, ..., t_n\}$ be any partition of $[a,b]$. Let

$$\begin{aligned} m_i &= \inf\{(f+g)(x) : t_{i-1} \le x \le t_i\}, \\ m_i' &= \inf\{f(x) : t_{i-1} \le x \le t_i\}, \\ m_i'' &= \inf\{g(x) : t_{i-1} \le x \le t_i\}, \end{aligned}$$

and define M_i, M_i', M_i'' similarly. Then

$$m_i \ge m_i' + m_i'' \quad \text{and} \quad M_i \le M_i' + M_i''$$

Therefore,

$$L(f,P) + L(g,P) \le L(f+g,P).$$

and

$$U(f+g,P) \le U(f,P) + U(g,P).$$

Thus,

$$L(f,P) + L(g,P) \le L(f+g,P) \le U(f+g,P) \le U(f,P) + U(g,P).$$

Since f and g are integrable, there are partitions P', P'' with

$$U(f,P') - L(f,P') < \frac{\epsilon}{2}.$$

$$U(g,P'') - L(g,P'') < \frac{\epsilon}{2}$$

If P contains both P' and P'', then

$$U(f,P) + U(g,P) - [L(f,P) + L(g,P)] < \epsilon,$$

and consequently

$$U(f+g,P) - L(f+g,P) < \epsilon.$$

This proves that $f+g$ is integrable on $[a,b]$. Moreover, (1).

$$\begin{aligned} L(f,P)+L(g,P) &\le L(f+g,P) \\ &\le \int_a^b (f+g) \\ &\le U(f+g,P) \le U(f,P)+U(g,P); \end{aligned}$$

(2).

$$L(f,P)+L(g,P) \le \int_a^b f + \int_a^b g \le U(f,P)+U(g,P).$$

Since $U(f,P)-L(f,P)$ and $U(g,P)-L(g,P)$ can both be made as small as desired, it follows that

$$U(f,P)+U(g,P)-[L(f,P)+L(g,P)]$$

can also be made as small as desired; it therefore follows from (1) and (2) that

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g.$$

□

2.54. Theorem: *If f is integrable on $[a,b]$, then for any number c, the function cf is integrable on $[a,b]$ and*

$$\int_a^b cf = c \cdot \int_a^b f.$$

Proof. Since f is integrable on $[a,b]$, then for all $\epsilon > 0$, there is a partition $P = \{t_0, ..., t_n\}$ such that $U(f,P) - L(f,P) < \epsilon$. Let

$$cm_i = \inf\{cf(x) : t_{i-1} \le x \le t_i\}$$

$$cM_i = \sup\{cf(x) : t_{i-1} \le x \le t_i\}$$

then

$$L(cf,P) = \sum_{i=1}^{n} cm_i(t_i - t_{i-1}) = c\sum_{i=1}^{n} m_i(t_i - t_{i-1}) = cL(f,P)$$

$$U(cf,P) = \sum_{i=1}^{n} cM_i(t_i - t_{i-1}) = c\sum_{i=1}^{n} M_i(t_i - t_{i-1}) = cU(f,P)$$

if $c \ge 0$, then $c(U(f,P) - L(f,P)) < \epsilon$ which implies

$$\begin{aligned} c((U(f,P) - L(f,P)) &< c\epsilon \\ U(cf,P) - L(cf,P) &< c\epsilon \end{aligned}$$

Since we can make ϵ as small as possible, we can also make $c\epsilon$ as small as possible, and thus cf is integrable on $[a, b]$. Notice also that

$$L(cf, P) = cL(f, P) \leq c \int_a^b f \leq cU(f, P) = U(cf, P) \implies \int_a^b cf = c \int_a^b f$$

If $c \leq 0$, we can use the fact that

$$cm_i = \inf\{cf(x) : t_{i-1} \leq x \leq t_i\} = c \sup\{f(x) : t_{i-1} \leq x \leq t_i\}$$

$$cM_i = \sup\{cf(x) : t_{i-1} \leq x \leq t_i\} = c \inf\{f(x) : t_{i-1} \leq x \leq t_i\}$$

Which implies that

$$L(cf, P) = cU(f, P) \text{ and } U(cf, P) = cL(f, P).$$

Therefore,

$$U(cf, P) - L(cf, P) = cL(f, P) - cU(f, P) = -c(U(f, P) - L(f, P)) < -c\epsilon$$

as required. Likewise

$$L(cf, P) = cU(f, P) \leq c \int_a^b f \leq cL(f, P) = U(cf, P) \implies \int_a^b cf = c \int_a^b f$$

□

2.55. Theorem: *Suppose f is integrable on $[a, b]$ and that*

$$m \leq f(x) \leq M \forall x \in [a, b].$$

Then

$$m(b - a) \leq \int_a^b f \leq M(b - a).$$

Proof. It is clear that

$$m(b - a) \leq L(f, P) \text{ and } U(f, P) \leq M(b - a)$$

for every partition P. Since $\int_a^b f = \sup\{L(f, P)\} = \inf\{U(f, P)\}$, the desired inequality follows immediately. □

Chapter 5

Trigonometric Functions

2.56. Definition:

$$\pi = 2 \cdot \int_{-1}^{1} \sqrt{1-x^2}\, dx.$$

We define π as the area of the unit circle, more precisely, it is twice the area of a semicircle.

2.57. Fact: *The area bounded by the unit circle, the horizontal axis, and a half-line from the origin to $(x, \sqrt{1-x}^2)$ is given by*

$$A(x) = \frac{x\sqrt{1-x^2}}{2} + \int_{x}^{1} \sqrt{1-t^2}\, dt,$$

for all $-1 \leq x \leq 1$.

2.58. For $0 \leq x \leq \pi$, we want to define $\cos x$ and $\sin x$ as the coordinates of a point $P = (\cos x, \sin x)$ on the unit circle which determines a sector with area $\frac{\pi}{2}$.

2.59. Definition: If $0 \leq x \leq \pi$, then $\cos x$ is the unique number in $[-1, 1]$ such that

$$A(\cos x) = \frac{x}{2}$$

and

$$\sin x = \sqrt{1-(\cos x)^2}.$$

2.60. Theorem: *If $0 < x < \pi$, then*

$$\cos'(x) = \sin x,$$

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

Proof. IF $B = 2A$, then the definition $A(\cos x) = \frac{x}{2}$ can be written as

$$B(\cos x) = x;$$

which means B is just the inverse of cos. Taking the derivative of A we see

$$A'(x) = -\frac{1}{2\sqrt{1-x^2}},$$

and so

$$B'(x) = -\frac{1}{\sqrt{1-x^2}}.$$

Consequently

$$\begin{aligned}
\cos'(x) &= (B^{-1})'(x) \\
&= \frac{1}{B'(B^{-1}(x))} \\
&= \frac{1}{-\frac{1}{\sqrt{1-[B^{-1}(x)]^2}}} \\
&= -\sqrt{1-(\cos x)^2} \\
&= -\sin x.
\end{aligned}$$

Since

$$\sin x = \sqrt{1-(\cos x)^2}.$$

we also obtain

$$\begin{aligned}
\sin'(x) = \frac{1}{2}\cdot &= \frac{-2\cos x \cdot \cos'(x)}{\sqrt{1-(\cos x)^2}} \\
&= \frac{\cos x \sin x}{\sin x} \\
&= \cos x.
\end{aligned}$$

□

2.61. For values of $\sin x$ and $\cos x$ for x not in $[0, \pi]$, they can be easily defined by a two-step piecing together process:

- If $\pi \le x \le 2\pi$, then

$$\sin x = -\sin(2\pi - x),$$

$$\cos x = \cos(2\pi - x).$$

- If $x = 2\pi k + x'$ for some integer k, and some $x' \in [0, 2\pi]$, then

$$\sin x = \sin x',$$

$$\cos x = \cos x'.$$

2.62. Definition: For $x \neq k\pi + \frac{\pi}{2}$, we define:

$$\sec x = \frac{1}{\cos x} \quad \text{and} \quad \tan x = \frac{\sin x}{\cos x},$$

and for $x \neq k\pi$, we define:

$$\csc x = \frac{1}{\sin x} \quad \text{and} \quad \cot x = \frac{\cos x}{\sin x}.$$

2.63. Theorem: *If $x \neq k\pi + \frac{\pi}{2}$, then*

$$\sec'(x) = \sec x \tan x,$$

$$\tan'(x) = \sec^2 x.$$

If $x \neq k\pi$, then

$$\csc'(x) = -\csc x \cot x,$$

$$\cot'(x) = -\csc^2 x.$$

Proof. Trivial. □

2.64. The inverses of the trigonometric functions can also be easily differentiated, however, we must restrict them to suitable intervals so that it is one-to-one; the largest possible length obtainable is π, and the intervals usually chosen are

$$[-\pi/2, \pi/2] \text{ for sin, } [0, \pi] \text{ for cos, } (-\pi/2, \pi/2) \text{ for tan.}$$

2.65. Definition: The inverse of the function

$$f(x) = \sin x, -\pi/2 \leq x \leq \pi/2$$

is denoted by **arcsin**, whose domain is $[-1, 1]$.

2.66. Definition: The inverse of the function

$$g(x) = \cos x, \quad 0 \leq x \leq \pi$$

is denoted by **arccos**, whose domain is $[-1, 1]$.

2.67. Definition: The inverse of the function

$$h(x) = \tan x, \quad \pi/2 < x < \pi/2$$

is denoted by **arctan**, whose domain is all of $\mathbb{R}$. This function is one of the simplest examples of a bounded differentiable function that is one-to-one on all of $\mathbb{R}$.

2.68. Theorem: *If* $-1 < x < 1$, *then*

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}} \quad \textit{and} \quad \arccos'(x) = \frac{-1}{\sqrt{1-x^2}}.$$

Moreover, for all x *we have*

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

Proof.

$$\begin{aligned}\arcsin'(x) &= (f^{-1})'(x)\\ &= \frac{1}{f'(f^{-1}(x))}\\ &= \frac{1}{\sin'(\arcsin x))}\\ &= \frac{1}{\cos(\arcsin x)}\end{aligned}$$

Now since

$$\sin^2(\arcsin x) + \cos^2(\arcsin x) = 1,$$

we have

$$x^2 + \cos^2(\arcsin x) = 1;$$

therefore,

$$\cos(\arcsin x) = \sqrt{1-x^2}.$$

This proves the first formula. The second formula can be proofed in a similar way. The third formula is proved as follows.

$$\begin{aligned}\arctan'(x) &= (h^{-1})'(x)\\ &= \frac{1}{h'h^{-1}(x)}\\ &= \frac{1}{\tan'(\arctan x)}\\ &= \frac{1}{\sec^(\arctan x)}\end{aligned}$$

Dividing both sides of the identity

$$sin^2 a + cos^2 a = 1$$

by $\cos^2 a$ yields

$$\tan^2 a + 1 = \sec^2 a.$$

It follows that

$$\tan^2(\arctan x) + 1 = \sec^2(\arctan x)$$

or

$$x^2 + 1 = \sec^2(\arctan x)$$

which proves the last formula. □

2.69. Lemma: *Suppose f has a second derivative everywhere and that*

$$f'' + f = 0, f(0) = 0, f'(0) = 0,$$

then $f = 0$.

Proof. Multiplying both sides of the first equation by f' yields

$$f'f'' + ff' = 0.$$

Thus

$$[(f')^2 + f^2]' = 2(f'f'' + ff') = 0.$$

so $(f')^2 + f^2$ is a constant function. From $f(0) = 0$ and $f'(0) = 0$ it follows that the constant is 0; thus

$$[f'(x)]^2 + [f(x)]^2 = 0 \quad \text{for all } x$$

which implies that

$$f(x) = 0 \quad \text{for all } x.$$

□

2.70. Theorem: *If f has a second derivative everywhere and*

$$f'' + f = 0, f(0) = a, f'(0) = b,$$

then

$$f = b \cdot \sin + a \cdot \cos.$$

Proof. Let

$$g(x) = f(x) - b \sin x - a \cos x.$$

Then

$$g'(x) = f'(x) - b \cos x + a \sin x,$$

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

Consequently,

$$g'' + g = 0, g(0) = 0, g'(0) = 0,$$

which shows that

$$0 = g(x) = f(x) - b \sin x - a \cos x \quad \text{for all } x.$$

□

2.71. Theorem: *If x and y are any two numbers, then*

$$\sin(x+y) = \sin x \cos y + \cos x \sin y,$$

$$\cos(x+y) = \cos x \cos y - \sin x \sin y.$$

Proof. For any number y we can define a function f by

$$f(x) = \sin(x+y).$$

Then

$$f'(x) = \cos(x+y)$$

$$f''(x) = -\sin(x+y).$$

Consequently,

$$f'' + f = 0, f(0) = \sin y, f'(0) = \cos y,$$

It follows from theorem 4 that

$$\sin(x+y) = \cos y \sin x + \sin y \cos x, \quad \text{for all } x.$$

Since any number y could have been chosen to begin with, this proves the first formula for all x and y. The second formula is proved similarly. □

Chapter 6

Log and Exp functions

2.72. Definition: If $x > 0$, then

$$\log x = \int_1^x \frac{1}{t}\, dt.$$

2.73. Note: if $x > 1$, then $\log x > 0$, if $0 < x < 1$, then $\log x < 0$. And for all $x \leq 0$, $\log x$ is not defined as $f(t) = 1/t$ is not bounded on $[x, 1]$.

2.74. Theorem: *If $x, y > 0$, then*

$$\log(xy) = \log x + \log y.$$

Proof. Choose any $y > 0$ and let

$$f(x) = \log(xy)$$

Then

$$f'(x) = \log'(xy) \cdot = \frac{1}{xy} \cdot y = \frac{1}{x}$$

Thus $f' = log'$, this means that there is a number c such that

$$f(x) = \log x + c$$

for all $x > 0$, that is

$$\log(xy) = \log x + c$$

for all $x > 0$. Letting $x = 1$, we obtain

$$\log(1 \cdot y) = \log 1 + c = \log c.$$

This is true for all $y > 0$, so the theorem is proved. □

2.75. Corollary: *If n is a natural number and $x > 0$, then*

$$\log(x^n) = n \log x.$$

Proof. Trivial. □

2.76. Corollary: *If $x, y > 0$, then*

$$\log(\frac{x}{y}) = \log x - \log y.$$

Proof. This follows from the equations

$$\log x = \log(\frac{x}{y} \cdot y) = \log(\frac{x}{y}) + \log y.$$

□

2.77. Definition: The "exponential function," **exp**, is defined as $\log^{-1}$.

2.78. Theorem: *For all numbers x,*

$$\exp'(x) = \exp(x).$$

Proof.

$$\exp'(x) = (\log^{-1})'(x) = \frac{1}{\log'(\log^{-1}(x))} = \frac{1}{\frac{1}{\log^{-1}(x)}} = \log^{-1}(x) = \exp(x).$$

□

2.79. Theorem: *If x and y are any two numbers, then*

$$\exp(x + y) = \exp(x) \cdot \exp(y).$$

Proof. Let $x' = \exp(x)$ and $y' = \exp(y)$, so that

$$x = \log x' \quad \text{and} \quad y = \log y'$$

Then

$$x + y = \log x' + \log y' = \log(x'y').$$

This means that

$$\exp(x + y) = x'y' = \exp(x) \cdot \exp(y).$$

□

2.80. Definition:

$$\mathbf{e} = \exp(1).$$

2.81. Definition: For any number x,

$$\mathbf{e^x} = \exp(x).$$

2.82. Definition: If $a > 0$, then, for any real number x,

$$\mathbf{a^x} = e^{x \log a}.$$

2.83. Theorem: *If $a > 0$, then*

$$(a^b)^c = a^{bc}$$

for all b, c;

$$a^1 = a \quad \textit{and} \quad a^{x+y} = a^x \cdot a^y.$$

for all x, y.

Proof. Trivial. □

2.84. Remark: Just as a^x can be expressed in terms of exp, $\log_a$ can be expressed in terms of *log*. If $y = \log_a x$, then $x = a^y = e^{y \log a}$, so $\log x = y \log a$, or $y = \frac{\log x}{\log a}$.

2.85. Theorem: *If f is differentiable and*

$$f'(x) = f(x) \quad \textit{for all } x,$$

then there is a number c such that

$$f(x) = ce^x$$

for all x.

Proof. Let

$$g(x) = \frac{f(x)}{e^x}.$$

Then

$$g'(x) = \frac{e^x f'(x) - f(x)e^x}{(e^x)^2} = 0.$$

Hence, $g(x)$ is constant and so

$$g(x) = \frac{f(x)}{e^x} = c$$

for all x. □

2.86. Theorem: *For any natural number n,*

$$\lim_{x\to\infty} \frac{e^x}{x^n} = \infty.$$

Proof. We prove by induction. When $n = 1$, we have to prove $e^x > x$ for all x, this is equivalent to $x > \log(x)$ for all x.

- If $x < 0$, then $0 < e^x \leq 1$, so $x < e^x$.
- If $0 < x \leq 1$, then $\log(x) \leq 0 < x$.
- If $x > 1$, then $\log x = \int_1^x \frac{1}{t}\, dt$. Suppose we have a partition $\mathcal{P}$ consisting with 1 block with width $x - 1$ and $M = 1$, this means $\log x < U(\frac{1}{t}, \mathcal{P}) = 1(x-1) < x$.

By induction on n, using L'Hopitals' Rule

$$\lim_{x\to\infty} \frac{e^x}{x^n} = \lim_{x\to\infty} \frac{e^x}{nx^{n-1}} = \frac{1}{n} \lim_{x\to\infty} \frac{e^x}{x^{n-1}} = \infty.$$

□

Chapter 7

Integration in elementary terms

Section 3. List of important basic integrals

(1). $\int a\ dx = ax$

(2). $\int a^n\ dx = \frac{x^{n+1}n+1}{,} n \neq = 1$

(3). $\int \frac{1}{x}\ dx = \log x$

(4). $\int e^x\ dx = e^x$

(5). $\int \sin x\ dx = -\cos x$

(6). $\int \cos x\ dx = \sin x$

(7). $\int \sec^2 x\ dx = \tan x$

(8). $\int \sec x\ \tan x\ dx = \sec x$

(9). $\int \frac{dx}{1+x^2} = \arctan x$

(10). $\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x$

3.1. Theorem (Integration by parts): *If f' and g' are continuous, then*

$$\int f(x)g'(x)\ dx = f(x)g(x) - \int f'(x)g(x)\ dx.$$

$$\int_a^b f(x)g'(x)\ dx = f(x)g(x)|_a^b - \int_a^b f'(x)g(x)\ dx.$$

Proof. Trivial. □

3.2. Note: While using integration by parts, there are two tricks one should know. The first is to consider the function g' to be the fact 1, the obvious example of the use of this is when integrating $\int \log x\ dx$. THe second trick is to ues integration by parts to find $\int h$ in terms of $\int h$ again, and then solve for $\int h$. A simple example:

$$\int (1/x) \cdot \log x\ dx = \log x \cdot \log x - \int (1/x) \cdot \log x\ dx$$

which implies that

$$2\int \frac{1}{x}\log x \; dx = (\log x)^2$$

or

$$\int \frac{1}{x}\log x \; dx = \frac{(\log x)^2}{2}$$

However more complicated calculations is often required, usually by repeated applying integration by parts.

3.3. Theorem (Substitution): *If f and g' are continuous, then*

$$\int_{g(a)}^{g(b)} f(u) \; du = \int_a^b (f(g(x)) \cdot g'(x) \; dx$$

Proof. If F is a primitive of f, then the left side is $F(g(b)) - F(g(a))$. On the other hand,

$$(F \circ g)' = (F' \circ g) \cdot g' = (f \circ g) \circ g'.$$

So $F \circ g$ is a primitive of $(f \circ g) \cdot g'$ and the right side is

$$(F \circ g)(b) - (F \circ g)(a) = F(g(b)) - F(g(a)).$$

□

3.4. Theorem (Partial fraction decomposition): *Every polynomial function*

$$q(x) = x^m + b_{m-1}x^{m-1} + \cdots + b_0$$

can be written as a product

$$q(x) = (x - \alpha_1)^{r_1} \cdot \ldots \cdot (x - \alpha_k)^{r_k}(x^2 + \beta_1 x + \gamma_1)^{s_1} \cdot \ldots \cdot (x^2 + \beta_l x + \gamma_l)^{2_k}$$

(where $r_1 + \cdots + r_k + 2(s_1 + \cdots + s_l) = m$).

3.5. Theorem: *If $n < m$ and*

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0,$$

$$\begin{aligned} a(x) &= x^m + b_{m-1}x^{m-1} + \cdots + b_0 \\ &= (x-\alpha_1)^{r_1} \cdot \ldots \cdot (x-\alpha_k)^{r_k} (x^2+\beta_1 x+\gamma_1)^{s_1} \cdot \ldots \cdot (x^2+\beta_l x+\gamma_l)^{2_k} \end{aligned}$$

then $p(x)/q(x)$ can be written in the form

$$\begin{aligned} \frac{p(x)}{q(x)} &= [\frac{a_{1,1}}{(x-\alpha_1)} + \cdots + \frac{a_{1,r_1}}{(x-\alpha_1)^{r_1}}] + \cdots + [\frac{a_{k,1}}{(x-\alpha_k)} + \cdots + \frac{a_{k,r_1}}{(x-\alpha_k)^{r_k}}] \\ &\quad + [\frac{b_{1,1x+c_{1,1}}}{(x^2-\beta_1 x+\gamma_1)} + \cdots + \frac{b_{1,s_1}+c_{1,s_1}}{(x^2+\beta_1 x+\gamma_1)^{s_1}}] + \cdots + [\frac{b_{l,1x+c_{l,1}}}{(x^2-\beta_l x+\gamma_l)} + \cdots + \frac{b_{l,s_l}+c_{l,s_l}}{(x^2+\beta_l x+\gamma_l)^{s_l}}] \end{aligned}$$

Chapter 8

Integration in elementary terms

Section 4. Taylor polynomials

4.1. Definition: Let

$$a_k = \frac{f^{(k)}(a)}{k!}, \quad 0 \le k \le n$$

The Taylor polynomial of degree n for f at a is:

$$P_{n,a}(x) = a_0 + a_1(x-a) + \cdots + a_n(x-a)^n.$$

4.2. The Taylor polynomial has been defined so that

$$P_{n,a}^{(k)}(a) = f^{(k)}(a) \quad 0 \le k \le n,$$

and it is also the only polynomial of degree $\le n$ with this property.

4.3. Theorem: *Suppose that f is a function which is n-times differentiable. Let*

$$a_k = \frac{f^{(k)}(a)}{k!}. \quad 0 \le k \le n,$$

and define

$$P_{n,a}(x) = a_0 + a_1(x-a) + \cdots + a_n(x-a)^n.$$

Then

$$\lim_{x \to a} \frac{f(x) - P_{n,a}(x)}{(x-a)^n} = 0.$$

Proof. Writing out $P_{n,a}(x)$ explicitly, we obtain

$$\frac{f(x) - P_{n,a}(x)}{(x-a)^n} = \frac{f(x) - \sum_{i=0}^{n-1} \frac{f^{(i)}(a)}{i!}(x-a)^i}{(x-a)^n} - \frac{f^{(n)}(a)}{n!}.$$

Lets introduce two new functions

$$Q(x) = \sum_{i=0}^{n-1} \frac{f^{(i)}(a)}{i!}(x-a)^i \quad \text{and} \quad g(x) = (x-a)^n;$$

so now we must prove that

$$\lim_{x\to a} \frac{f(x) - Q(x)}{g(x)} = \frac{f^{(n)}(a)}{n!}.$$

Notice that $Q^{(k)}(a)$ is just $f^{(k)}(a)$ for all $k \leq n-1$, and $g^{(k)}(x) = n!(x-a)^{n-k}/(n-k)!$. Thus we can apply l'Hopitals rule repeatedly, as

$$\lim_{x\to a}[f(x) - Q(x)] = f(a) - Q(a) = 0,$$

$$\vdots$$

$$\lim_{x\to a}[f^{(n-2)}(x) - Q^{(n-2)}(x)] = f^{(n-2)}(a) - Q^{(n-2)}(a) = 0,$$

We can in fact apply l'Hopital's rules $n-1$ times to obtain

$$\lim_{x\to a} \frac{f(x) - Q(x)}{(x-a)^n} = \lim_{x\to a} \frac{f^{(n-1)}(x) - Q^{(n-1)}(x)}{n!(x-a)}$$

Since Q is a polynomial of degree $n-1$, its $(n-1)$st derivative is a constant, in fact, $Q^{^{(n-1)}(x)} = f^{(n-1)}(a)$. Thus

$$\lim_{x\to a} \frac{f(x) - Q(x)}{(x-a)^n} = \lim_{x\to a} \frac{f^{(n-1)}(x) - f^{(n-1)}(a)}{n!(x-a)}$$

Applying L'Hopital one last time gives us that the last limit is $f^{(n)}(a)/n!$, which is what we want. □

4.4. Theorem: *Suppose that*

$$f'(a) = \cdot = f^{(n-1)}(a) = 0,$$

$$f^{(n)}(a) \neq 0$$

(1). if n is even and $f^{(n)}(a) > 0$, then f has a local minimum at a.

(2). If n is even and $f^{(n)}(a) < 0$, then f has a local maximum at a.

(3). If n is odd, then f has neither a local maximum nor a local minimum at Thena.

Proof. Without loss of generality, assume $f(a) = 0$, since neither the hypothese nor the conclusion are affected if f is replaced by $f - f(a)$. Then, since the first $n-1$ derivatives of f at a are 0, the Taylor polynomial $P_{n,a}$ of f is

$$P_{n,a}(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}}{n!}(x-a)^n = \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

By theorem 1,

$$0 = \lim_{x \to a} \frac{f(x) - P_{n,a}(x)}{(x-a)^n} = \lim_{x \to a} [\frac{f(x)}{(x-a)^n} - \frac{f^{(n)}(a)}{n!}].$$

Which means, if x is sufficiently close to a, then

$$\frac{f(x)}{(x-a)^n} \text{ has the same sign as } \frac{f^{(n)}(a)}{n!}.$$

Suppose now that n is even. In this case $(x-a)^n > 0$ for all $x \neq a$, hence $f(x)$ must have the same sign as $f^n(a)/n!$ for x sufficiently cloes to a. If $f^{(n)}(a) > 0$, then

$$f(x) > 0 = f(a)$$

for x close to a, and hence f has a local minimum at a. An analogous proof works for the case $f^{(n)}(a) < 0$.
Now suppose that n is odd, the same argument as before shows that if x is sufficiently close to a, then

$$\frac{f(x)}{(x-a)^n} \quad \text{must always has the same sign}$$

But $(x-a)^n > 0$ for $x > a$ and $(x-a)^n < 0$ for $x < a$. Therefore $f(x)$ has different signs for $x > a$ and $x < a$, which proves that f is neither a local maximum nor a local minimum at a. □

4.5. Definition: Two functions f and g are **equal up to order** n **at** a if

$$\lim_{x \to a} \frac{f(x) - g(x)}{(x-a)^n} = 0.$$

4.6. Theorem: *Let P and Q be two polynomials in $(x-a)$, of degree $\leq n$, and suppose that P and Q are equal up to order n at a. Then $P = Q$.*

Proof. Spivak pg 419. □

4.7. Corollary: *Let f be n-times differentiable at a, and suppose that P is a polynomial in $(x-a)$ of degree $\leq n$, which equals f up to order n at a. Then $P = P_{n,a,f}$.*