Real Analysis: A Complete Guide

Real Analysis

May 20, 2021

Elijah Thompson,
Physics and Math Honors

Solo Pursuit of Learning

Contents

Part I

Single Variable Analysis

Chapter 1

Topology and Construction of the Real Line

1.1 The Axiom of Completeness

Definition 1.1.1: The Reals

The real number system $\mathbb{R}$ is an **ordered field** *which contains $\mathbb{Q}$ as a subfield, which satisfies the* **axiom of choice**. *In particular, the real numbers is a set $\mathbb{R}$ with two binary operations $+$ and $\cdot$, two distinct elements 0 and 1, and a subset $\mathbb{P}$ of positive numbers satisfying the following 13 postulates:*

1. *Addition is associative:* $\forall a, b, c \in \mathbb{R}, a + (b + c) = (a + b) + c$
2. *The number 0 is an additive identity:* $\forall a \in \mathbb{R}, a + 0 = 0 + a = a$
3. *Additive inverses exist:* $\forall a \in \mathbb{R}; \exists(-a) \in \mathbb{R}$ *s.t.* $a + (-a) = (-a) + a = 0$
4. *Addition is commutative:* $\forall a, b \in \mathbb{R}, a + b = b + a$
5. *Multiplication is associative:* $\forall a, b, c \in \mathbb{R}, a \cdot (b \cdot c) = (a \cdot b) \cdot c$
6. *The number 1 is a multiplicative identity:* $\forall a \in \mathbb{R} a \cdot 1 = 1 \cdot a = a$
7. *Multiplicative inverses exist:* $\forall a \neq 0; \exists a^{-1} \in \mathbb{R}$ *s.t.* $a \cdot a^{-1} = a^{-1} \cdot a = 1$
8. *Multiplication is commutative:* $\forall a, b \in \mathbb{R}, a \cdot b = b \cdot a$
9. *The distributive law:* $\forall a, b, c \in \mathbb{R}, a \cdot (b + c) = a \cdot b + a \cdot c$
10. *The trichotomy of $\mathbb{P}$: for every $a \in \mathbb{R}$, exactly one of the following holds:* $a = 0, a \in \mathbb{P}, (-a) \in \mathbb{P}$
11. *Closure under addition: if $a \in \mathbb{P}$ and $b \in \mathbb{P}$, then $a + b \in \mathbb{P}$*
12. *Closure under multiplication: if $a \in \mathbb{P}$ and $b \in \mathbb{P}$, then $a \cdot b \in \mathbb{P}$*

13. *(to be added)*

From positive postulates we can define the order relations $>, <, \geqslant, \leqslant$ *on* $\mathbb{R}$ *for* $a, b \in \mathbb{R}$ *by*

1. $a > b$ *if* $a - b \in \mathbb{P}$
2. $a < b$ *if* $b > a$
3. $a \geqslant b$ *if* $a > b$ *or* $a = b$
4. $a \leqslant b$ *if* $a < b$ *or* $a = b$

Note in particular $a > 0$ *if and only if* $a \in \mathbb{P}$.

Remark 1.1.1

A few points which follow from the postulates are:

1. *Finite sums such as* $a_1 + a_2 + \ldots + a_n$ *are well defined*
2. *The additive identity is unique (also multiplicative)*
3. *Additive inverses are unique (also multiplicative)*
4. *Subtraction can be defined*
5. $a \cdot b = b \cdot c \iff a = 0 \vee b = c$
6. $a \cdot b = \iff a = 0 \vee b = 0$
7. $a - b = b - a \iff a = b$
8. *A "well behaved" order relation can be defined.*
9. *The "absolute value" function* $a \mapsto |a|$ *can be defined by*

$$|a| = \begin{cases} a, & a \geqslant 0 \\ -a, & a \leqslant 0 \end{cases}$$

and for all $a, b \in \mathbb{R}$, *the triangle inequality*

$$|a + b| \leqslant |a| + |b|$$

holds

Axiom: Axiom of Completeness

Every non-empty subset of the real numbers that is bounded above has a least upper bound.

1.1.1 Upper and Lower Bounds

Definition 1.1.2: Bounds

A set $A \subseteq \mathbb{R}$ is **bounded above** *if there exists a number $b \in \mathbb{R}$ such that $a \leqslant b$ for all $a \in A$. The number b is called an* **upper bound** *for A.*
Similarly, the set A is **bounded below** *if there exists a* **lower bound** *$l \in \mathbb{R}$ satisfying $l \leqslant a$ for every $a \in A$.*

Definition 1.1.3: Least Upper Bound

A real number s is the **least upper bound** *for a set $A \subseteq \mathbb{R}$ if it meets the following two criteria:*

1. *s is an upper bound for A;*

2. *if b is any upper bound for A, then $s \leqslant b$.*

The least upper bound of a set A is also called the **supremum** *of A, and denoted by* $\sup A$.

Definition 1.1.4: Greatest Lower Bound

A real number i is the **Greatest Lower bound** *for a set $A \subseteq \mathbb{R}$ if it meets the following two criteria:*

1. *i is a lower bound for A;*

2. *if b is any lower bound for A, then $b \leqslant i$.*

The greatest lower bound of a set A is also called the **infemum** *of A, and denoted by* $\inf A$.

Remark 1.1.2

From the definitions we assert that the least upper bound and greatest lower bound of a set, if they exist, are unique.

Example

Consider the set

$$A = \left\{ \frac{1}{n} : n \in \mathbb{N} \right\} = \left\{ 1, \frac{1}{2}, \frac{1}{3}, \ldots \right\}$$

The set A is bounded above and below. Moreover, the least upper bound of A is $\sup A = 1$, *which is in A, while* $\inf A = 0$, *which is not in A.*

Definition 1.1.5: Max and Min

A real number a_0 is a **maximum** *of a set A if a_0 is an element of A and $a_0 \geqslant a$ for all $a \in A$. Similarly, a number a_1 is a* **minimum** *of A if $a_1 \in A$ and $a_1 \leqslant a$ for all $a \in A$.*

Example

Consider the open interval $(0, 2)$*, and the closed interval* $[0, 2]$*. Note both sets are bounded above and below, and both have the same infimum and supremum, namely* $\inf = 0$ *and* $\sup = 2$*. However,* $[0, 2]$ *has both a maximum and a minimum, namely its infimum and supremum, while* $(0, 2)$ *has neither.*

Example

Let $A \subseteq \mathbb{R}$ *be a non-empty and bounded above set, and let* $c \in \mathbb{R}$*. Define the set* $c + A$ *by*

$$c + A := \{c + a : a \in A\}$$

Then I claim that $\sup(c + A) = c + \sup A$*.*

Proof.

Let $\alpha = c + \sup A$*. First let us show that* α *is an upper bound of* $c + A$*. Indeed, for* $x \in c + A$ *we can write* $x = c + a$ *for some* $a \in A$ *by definition. Then, by definition we have that* $a \leqslant \sup A$*. Thus, adding c to both sides we obtain*

$$x = c + a \leqslant c + \sup A = \alpha$$

Therefore, as x was arbitrary α *is indeed an upper bound of* $c + A$*. Now, suppose b is an upper bound of* $c + A$*. Then,* $c + a \leqslant b$ *for all* $c + a \in c + A$*, so in particular* $a \leqslant b - c$ *for all* $a \in A$*. Then,* $b - c$ *is an upper bound for A, so as* $\sup A$ *is the least upper bound of A we have that* $\sup A \leqslant b - c$*. Hence, we conclude that* $\alpha = c + \sup A \leqslant b$*. Thus* α *satisfies the axioms of a least upper bound for* $c + A$*, and we conclude that* $\sup(c + A) = c + \sup A$*.*

■

Lemma 1.1.1

Assume $s \in \mathbb{R}$ *is an upper bound for a set* $A \subseteq \mathbb{R}$*. Then* $s = \sup A$ *if and only if, for every choice* $\epsilon > 0$*, there exists an element* $a \in A$ *satisfying* $s - \epsilon < a$*.*

Proof.

Let $s \in \mathbb{R}$ *be an upper bound for a set* $A \subseteq \mathbb{R}$*.*
($\Longrightarrow$ *) First, suppose that* $s = \sup A$*, and choose* $\epsilon \in \mathbb{R}$ *with* $\epsilon > 0$*. Then* $s - \epsilon$ *is not an upper bound of A. Indeed, if* $s - \epsilon$ *was an upper bound then* $s \leqslant s - \epsilon$ *which implies that* $\epsilon \leqslant 0$*, but by assumption* $\epsilon > 0$*. Thus, there must exist* $a \in A$ *such that* $s - \epsilon < a$*, satisfying the implication.*
($\Longleftarrow$ *) Conversely, suppose that for all* $\epsilon > 0$ *there exists* $a \in A$ *such that* $s - \epsilon < a$*. Now, suppose that b is an upper bound of A, and towards a contradiction suppose* $s > b$*. Then* $s - b > 0$*, so there exists* $a \in A$ *such that* $s - (s - b) < a$*. In particular,* $b < a$*. However,* $a \in A$ *and b is an upper bound of A by assumption, so* $b < a$ *is a contradiction. Therefore we conclude that* $s \leqslant b$*, so s is the supremum of A by definition.*

■

Theorem 1.1.2: Archimedean Property for the Reals

For all $x, y > 0$ *in* $\mathbb{R}$*, there exists* $n \in \mathbb{N}$ *such that* $nx > y$*.*

Proof.

Towards a contradiction suppose $nx \leqslant y$ for all $n \in \mathbb{N}$. Then the set $\{nx : n \in \mathbb{N}\}$ is bounded above by y. Thus, by the least upper bound property of $\mathbb{R}$ we have a supremum $\alpha \in \mathbb{R}$. Then for all $n \in \mathbb{N}$ $\alpha \geqslant nx$. In particular, $\alpha \geqslant (n+1)x$ for all $n \in \mathbb{N}$, so $\alpha - x \geqslant nx$ for all $n \in \mathbb{N}$. But this implies that $\alpha - x$ is also an upper bound of the set, which contradicts the fact that α is the least upper bound. Thus, we must have that $nx > y$ for some $n \in \mathbb{N}$, as claimed.

■

Corollary 1.1.3

$\mathbb{N}$ is not bounded above.

Corollary 1.1.4

For any $\epsilon > 0$ there is a natural number n with $1/n < \epsilon$.

Proof.

Consider $0 < 1/\epsilon \in \mathbb{R}$ and $1 \in \mathbb{R}$. Then by the Archimedean Property of $\mathbb{R}$ there exists $n \in \mathbb{N}$ such that $1 \cdot n > 1/\epsilon$. In particular, we have that $n > 0$, so $\epsilon > 1/n$, completing the proof.

■

1.2 Limits

Remark 1.2.1: Motivating Definition

The function f approaches the limit $l \in \mathbb{R}$ near $a \in \mathbb{R}$, if we can make $f(x)$ as "close as we like" to l by requiring that x be "sufficiently close to," but unequal to, a.

Definition 1.2.1: Limit

A real valued function $f : \mathbb{R} \to \mathbb{R}$ **approaches the limit l near a** *if for every $\epsilon > 0$ there is some $\delta > 0$ such that, for all $x \in \mathbb{R}$, if $0 < |x - a| < \delta$, then $|f(x) - l| < \epsilon$.*

Notation

We denote the number l which a function f approaches near $a \in \mathbb{R}$ by $\lim_{x \to a} f(x)$, read **the limit of $f(x)$ as x approaches a**.

Lemma 1.2.1

1. *If*

$$|x - x_0| < \frac{\epsilon}{2} \text{ and } |y - y_0| < \frac{\epsilon}{2}$$

then

$$|(x + y) - (x_0 + y_0)| < \epsilon$$

2. If

$$|x - x_0| < \min\left(1, \frac{\epsilon}{2(|y_0| + 1)}\right) \text{ and } |y - y_0| < \frac{\epsilon}{2(|x_0 + 1)}$$

then

$$|xy - x_0 y_0| < \epsilon$$

3. If $y_0 \neq 0$ and

$$|y - y_0| < \min\left(\frac{|y_0|}{2}, \frac{\epsilon |y_0|^2}{2}\right)$$

then $y \neq 0$ and

$$\left|\frac{1}{y} - \frac{1}{y_0}\right|$$

Proof.

(1) Suppose $|x - x_0| < \frac{\epsilon}{2}$ and $|y - y_0| < \frac{\epsilon}{2}$. Then it follows that

$$\begin{aligned} |(x + y) - (x_0 + y_0)| &= |(x - x_0) + (y - y_0)| \\ &\leqslant |x - x_0| + |y - y_0| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon \end{aligned}$$

as desired.
(2) Next, suppose $|x - x_0| < \min\left(1, \frac{\epsilon}{2(|y_0|+1)}\right)$ and $|y - y_0| < \frac{\epsilon}{2(|x_0+1)}$. Note that as $|x - x_0| < 1$ we have that $|x| - |x_0| < 1$ so $|x| < 1 + |x_0|$. It follows that

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

(3) Suppose $y_0 \neq 0$ and $|y - y_0| < \min\left(\frac{|y_0|}{2}, \frac{\epsilon |y_0|^2}{2}\right)$. Then note that as $|y - y_0| < \frac{|y_0|}{2}$ $-\frac{|y_0|}{2} < y - y_0 < \frac{|y_0|}{2}$. If $y_0 > 0$ we have that $y_0 = |y_0|$ so $\frac{|y_0|}{2} < y < \frac{3|y_0|}{2}$. On the other hand if $y_0 < 0$ then $y_0 = -|y_0|$ so $-\frac{3|y_0|}{2} < y < -\frac{|y_0|}{2}$. In either case we have that $|y| > \frac{|y_0|}{2} > 0$, so $y \neq 0$. Then it follows that

$$\begin{aligned} \left|\frac{1}{y} - \frac{1}{y_0}\right| &= \left|\frac{y - y_0}{y y_0}\right| \\ &< \frac{\epsilon |y_0|^2}{2} \cdot \frac{1}{|y_0|} \cdot \frac{2}{|y_0|} \end{aligned}$$

$$= \epsilon$$

as claimed.

■

Theorem 1.2.2: Limit Laws

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. *Moreover, if* $m \neq 0$*, then* $\lim_{x \to a} \left(\frac{1}{g}\right)(x) = \frac{1}{m}$

Proof.

Suppose that $\lim_{x \to a} f(x) = l$ *and* $\lim_{x \to a} g(x) = m$*. Let* $\epsilon > 0$*.*
(1) Then since $\lim_{x \to a} f(x) = l$ *and* $\lim_{x \to a} g(x) = m$ *there exist* $\delta_1, \delta_2 > 0$ *such that* $|f(x) - l| < \frac{\epsilon}{2}$ *if* $0 < |x - a| < \delta_1$ *and* $|g(x) - m| < \frac{\epsilon}{2}$ *if* $0 < |x - a| < \delta_2$*. Choose* $\delta = \min(\delta_1, \delta_2)$*. Then it follows that for* $0 < |x - a| < \delta$*:*

$$\begin{aligned} |(f + g)(x) - (l + m)| &= |(f(x) - l) + (g(x) - m)| \\ &\leqslant |f(x) - l) + |g(x) - m| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon \end{aligned}$$

Thus, we have that by definition $\lim_{x \to a} (f + g)(x) = l + m$*.*
(2) Now, fix $\epsilon_1 = \min\left(1, \frac{\epsilon}{2(|m|+1)}\right)$ *and* $\epsilon_2 = \frac{\epsilon}{2(|l|+1)}$*. Then since* $\lim_{x \to a} f(x) = l$ *and* $\lim_{x \to a} g(x) = m$ *there exist* $\delta_1, \delta_2 > 0$ *such that if* $0 < |x - a| < \delta_1$ *then* $|f(x) - l| < \epsilon_1$ *and if* $0 < |x - a| < \delta_2$ *then* $|g(x) - m| < \epsilon_2$*. Choose* $\delta = \min(\delta_1, \delta_2)$*. It follows that for* $0 < |x - a| < \delta$*, we have* $|f(x)g(x) - lm| < \epsilon$*, so by definition* $\lim_{x \to a} (f \cdot g)(x) = l \cdot m$*.*
(3) Now, suppose $m \neq 0$*. Fix* $\epsilon_1 = \min\left(\frac{|m|}{2}, \frac{\epsilon |m|^2}{2}\right)$*. Then as* $\lim_{x \to a} g(x) = m$*, there exists* $\delta > 0$ *such that if* $|x - a| < \delta$*,* $|g(x) - m| < \epsilon_1$*. By the previous Lemma we have that if* $|x - a| < \delta$*,* $g(x) \neq 0$ *and* $\left|\frac{1}{g(x)} - \frac{1}{m}\right| < \epsilon$*. Hence, by definition* $\lim_{x \to a} \left(\frac{1}{g}\right)(x) = \frac{1}{m}$*, as desired.*

■

Definition 1.2.2: Limits from Above and Below

The **limit from above** *for a function* f *as* x *goes to* a *is denoted by* $\lim_{x \to a^+} f(x) = l$*, which means for every* $\epsilon > 0$ *there is a* $\delta > 0$ *such that for all* x*,*

$$if \ 0 < x - a < \delta, \ then \ |f(x) - l| < \epsilon$$

where $0 < x - a < \delta$ *is equivalent to* $0 < |x - a| < \delta$ *and* $x > a$*.*

The **limit from below** *for f as x goes to a is denoted by* $\lim\limits_{x \to a^-} f(x) = l$*, and means that for every $\epsilon > 0$ there is a $\delta > 0$ such that, for all x,*

$$if\ 0 < a - x < \delta,\ then\ |f(x) - l| < \epsilon$$

Remark 1.2.2

For a function $f : \mathbb{R} \to \mathbb{R}$, $\lim_{x \to a} f(x)$ exists if and only if $\lim\limits_{x \to a^+} f(x)$ *and* $\lim\limits_{x \to a^-} f(x)$ *both exist and are equal.*

Definition 1.2.3: Limits at Infinity

A **limit at infinity** *is denoted by* $\lim\limits_{x \to \infty} f(x) = l$*, and means that for every $\epsilon > 0$ there is $M \in \mathbb{R}$ such that for all x,*

$$if\ x > M,\ then\ |f(x) - l| < \epsilon$$

A limit at negative infinity is defined analogously, replacing $x > M$ with $x < M$.

Definition 1.2.4

We define $\lim\limits_{x \to a} f(x) = \infty$ *to mean that for all $N \in \mathbb{R}$, there exists $\delta > 0$ such that, for all $x \in \mathbb{R}$, if $0 < |x - a| < \delta$, then $f(x) > N$. (the case for $-\infty$ is defined similarly)*

1.3 Continuous Functions

Definition 1.3.1: Continuity

Let $f : \mathbb{R} \to \mathbb{R}$ be a real function. Then f is said to be **continuous at a point** *a if*

$$\lim_{x \to a} f(x) = f(a) \tag{1.3.1}$$

Theorem 1.3.1

If f and g are continuous at a point a, then

1. *$f + g$ is continuous at a*
2. *$f \cdot g$ is continuous at a*
3. *Moreover, if $g(a) \neq 0$, then $1/g$ is continuous at a.*

Proof.

Suppose f and g are continuous at a point a. Then by the Limit Laws theorem, we have that *as* $\lim_{x\to a} f(x) = f(a)$ *and* $\lim_{x\to a} g(x) = g(a)$,

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

Hence, $f+g$ *is continuous at a. Similarly, again by the Limit Laws theorem, we have that*

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

Thus, $f\cdot g$ *is continuous at a. Finally, if* $g(a)\neq 0$*, the*

$$\lim_{x\to a}(1/g)(x) = 1/g(a) = (1/g)(a)$$

so $1/g$ *is continuous at a.*

■

Theorem 1.3.2

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

Proof.

Let $\epsilon > 0$*. Then by continuity of f there exists* $\delta_1 > 0$ *such that if* $|g(x)-g(a)| < \delta_1$*, then*

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

Then, by the continuity of g, there exists $\delta > 0$ *such that if* $|x-a| < \delta$*, then* $|g(x)-g(a)| < \delta_1$*. Thus, if* $|x-a| < \delta$ *we have that*

$$|(f\circ g)(x) - (f\circ g)(a)| = |f(g(x)) - f(g(a))| < \epsilon$$

proving continuity of $f\circ g$ *at a.*

■

Definition 1.3.2

A function f is called **continuous on** *an open interval* (a,b)*, if f is continuous at x for all* $x\in(a,b)$*.*
A function f is called **continuous on** *a closed interval* $[a,b]$ *if*

1. f *is continuous at x for all* $x\in(a,b)$
2. $\lim_{x\to a^+} f(x) = f(a)$ *and* $\lim_{x\to b^-} f(x) = f(b)$

In general, a function f is **continuous** *if it is continuous at x for all x in its domain.*

Theorem 1.3.3

Suppose f is continuous at a, and $f(a) > 0$*. Then* $f(x) > 0$ *for all x in some interval containing a; more precisely, there is a number* $\delta > 0$ *such that* $f(x) > 0$ *for all x satisfying* $|x-a| < \delta$*. Similarly, if* $f(a) < 0$*, then there is a number* $\delta > 0$ *such that* $f(x) < 0$ *for all x* *satisfying $|x-a|<\delta$.*

Proof.

Suppose f is continuous at a. Let $\epsilon = \frac{f(a)}{2} > 0$. Then by continuity there exists $\delta > 0$ such that if $|x-a| < \delta$, $|f(x) - f(a)| < \epsilon$. Then we have that $-\epsilon < f(x) - f(a) < \epsilon$, so $f(x) > \epsilon > 0$, satisfying the claim. The case for $f(a) < 0$ is proved analogously.

■

1.3.1 Important Theorems and Results on Continuity

Theorem 1.3.4

If f is continuous on a closed interval $[a,b]$ (a compact set) and $f(a) < 0 < f(b)$, then there is some $x \in [a,b]$ such that $f(x) = 0$.

Proof.

Consider an interval $[a,b]$ such that $f(a) < 0 < f(b)$. Define the set

$$a := \{x \in \mathbb{R} : a \leqslant x \leqslant b, \text{and } f \text{ is negative on } [a,x]\}$$

Clearly $A \neq \emptyset$ as $a \in A$. In fact, there exists some $\delta > 0$ such that A contains all points $x \in \mathbb{R}$ satisfying $a \leqslant x < a+\delta$, since f is continuous on $[a,b]$ and $f(a) < 0$. Similarly, b is an upper bound for A and, in fact, there is a $\delta > 0$ such that all points satisfying $b-\delta < x \leqslant b$ are upper bounds for A.
Thus, applying the Least Upper Bound property of $\mathbb{R}$, A has a least upper bound α and $a < \alpha < b$. We wish to show that $f(\alpha) = 0$. First, if $f(\alpha) < 0$, then by a previous result there is a $\delta > 0$ such that $f(x) < 0$ for $\alpha - \delta < x < \alpha + \delta$. In particular, there is some number $x_0 \in A$ satisfying $\alpha - \delta < x < \alpha$ since α is the supremum of A. Thus f is negative on the whole interval $[a, x_0]$. But, if $x_1 \in (\alpha, \alpha+\delta$, then f is also negative on the whole interval $[x_0, x_1]$. Therefore f is negative on the interval $[a, x_1]$ so $x_1 \in A$. But, this contradicts the fact that α is an upper bound for A, so $f(\alpha) < 0$ must be false.
Suppose, on the other hand, that $f(\alpha) > 0$. Then there is a number $\delta > 0$ such that $f(x) > 0$ for all $\alpha - \delta < x < \alpha + \delta$. Now there is some number $x_0 \in A$ such that $\alpha - \delta < x_0 < \alpha$ as α is presumed to be the supremum of A. This means that f is negative on the whole interval $[a, x_0]$, which is impossible since $f(x_0) > 0$. Thus, the assumption $f(\alpha) > 0$ leads to a contradiction, leaving $f(\alpha) = 0$ as the only possible alternative.

■

Lemma 1.3.5

If f is continuous at a, then there is a number $\delta > 0$ such that f is bounded on the interval $(a-\delta, a+\delta)$.

Proof.

Since f is continuous at a we have that $\lim_{x \to a} f(x) = f(a)$. Then, fix $\epsilon = 1$. By continuity it follows that there exists $\delta > 0$ such that for all $x \in (a - \delta, a + \delta)$, $|f(x) - f(a)| < 1$. In particular, we have that $f(x) < f(a) + 1$, so f is bounded above by $f(a) + 1$ on $(a - \delta, a + \delta)$. Moreover, $f(x) > f(a) - 1$, so f is bounded below by $f(a) - 1$ on $(a - \delta, a + \delta)$. Thus f is bounded on the interval $(a - \delta, a + \delta)$ as claimed.

■

Corollary 1.3.6

If $\lim_{x \to a^+} f(x) = f(a)$ then there exists $\delta > 0$ such that f is bounded on the interval $[a, a + \delta)$. Moreover, if $\lim_{x \to b^-} f(x) = f(b)$ then there exists $\delta > 0$ such that f is bounded on the interval $(b - \delta, b]$.

Theorem 1.3.7

If f is continuous on a closed interval $[a, b]$ (a compact set), then f is bounded above on $[a, b]$, that is, there is some number $M \in \mathbb{R}$ such that $f(x) \leqslant M$ for all $x \in [a, b]$ (consequence of the continuous image of a compact set being compact and the Heine-Borel Theorem).

Proof.

Define the set

$$A := \{x \in [a, b] : f \text{ is bounded above on } [a, x]\}$$

Clearly $A \neq \varnothing$ as $a \in A$, and A is bounded above by B, so A has a least upper bound $\alpha \in \mathbb{R}$. We wish to show that $\alpha = b$. Suppose towards a contradiction that $\alpha < b$. Then there exists $\delta > 0$ such that f is bounded on $(\alpha - \delta, \alpha + \delta)$ since f is continuous on $[a, b]$, so in particular f is continuous at α. Since α is the least upper bound of A there is some $x_0 \in A$ satisfying $\alpha - \delta < x_0 < \alpha$. This implies that f is bounded on $[a, x_0]$. But, if x_1 is any number with $\alpha < x_1 < \alpha + \delta$, then f is also bounded on $[x_0, x_1]$. Therefore f is bounded on $[a, x_1]$ so $x_1 \in A$, contradicting the fact that α is an upper bound for A. This contradiction shows that $\alpha = b$. Now, there is a $\delta > 0$ such that f is bounded on $(b - \delta, b]$. There is $x_0 \in A$ such that $b - \delta < x_0 < b$, since $\alpha = b$. Thus f is bounded on $[a, x_0]$, and also on $[x_0, b]$, so f is bounded on $[a, b]$, completing the proof.

■

Theorem 1.3.8

If f is continuous on a closed interval $[a, b]$, then there is some number $y \in [a, b]$ such that $f(y) \geqslant f(x)$ for all $x \in [a, b]$.

Proof.

From the previous theorem we know that f is bounded on $[a, b]$, so the set $\{f(x) : x \in [a, b]\}$ is bounded. This set is obviously non-empty, so it has a least upper bound $\alpha \in \mathbb{R}$. Since $\alpha \geqslant f(x)$ for all $x \in [a, b]$, it suffices to show that $\alpha = f(y)$ for some $y \in [a, b]$. Suppose instead that

$\alpha \neq f(y)$ for all $y \in [a,b]$. Then the function g defined by

$$g(x) = \frac{1}{\alpha - f(x)}, x \in [a,b]$$

is continuous on $[a,b]$ since the denominator is never zero and is the sum of continuous functions. On the other hand, α is the least upper bound of $\{f(x) : x \in [a,b]\}$ so for every $\epsilon > 0$ there exists $x \in [a,b]$ such that $\alpha - \epsilon < f(x)$, so $\alpha - f(x) < \epsilon$. This in turn implies that for every $\epsilon > 0$ there exists $x \in [a,b]$ with $g(x) > 1/\epsilon$. But, this implies that g is not bounded on $[a,b]$, contradicting the previous theorem as g is assumed to be continuous. Hence, g is not continuous, and i particular $\alpha - f(y) = 0$ for some $y \in [a,b]$.

■

Theorem 1.1: Intermediate Value Theorem

If f is continuous on $[a,b]$ and $f(a) < c < f(b)$, then there is some $x \in [a,b]$ such that $f(x) = c$ (continuous image of a connected set is connected).
Moreover, if $f(a) > c > f(b)$, then there is some $x \in [a,b]$ such that $f(x) = c$.

Theorem 1.3.9

If f is continuous on $[a,b]$, then f is bounded below on $[a,b]$, that is, there is some number $M \in \mathbb{R}$ such that $f(x) \geqslant M$ for all $x \in [a,b]$.

Theorem 1.3.10

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

Corollary 1.3.11

For all $\alpha \in \mathbb{P}$, so $\alpha > 0$, there exists $x \in \mathbb{R}$ such that $x^2 = \alpha$.

Proof.

Consider the function $f(x) = x^2$, which is certainly continuous over $\mathbb{R}$. Consider $\alpha \in \mathbb{P}$. Then there exists $b > 0$ such that $f(b) > \alpha$. Indeed, if $\alpha > 1$ we can take $b = \alpha$, and if $\alpha < 1$ we can take $b = 1$. Then, f is defined on the closed interval $[0,b]$ and $f(0) = 0 < \alpha < f(b)$. Therefore, by the Intermediate Value Theorem there exists $c \in [0,b]$ such that $f(c) = \alpha$. In particular, $c^2 = \alpha$.

■

Corollary 1.3.12

If n is odd, then any equation

$$x^n + a_{n-1}x^{n-1} + ... + a_0 = 0 \tag{1.3.2}$$

has a solution, or root.

Proof.

Consider the function $f(x) = x^n + a_{n-1}x^{n-1} + ... + a_0$. *Write*

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

Then note that

$$\left|\frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + ... + \frac{a_0}{x^n}\right| \leqslant \frac{|a_{n-1}|}{|x|} + ... + \frac{|a_0|}{|x^n|}$$

Choose x such that

$$|x| > 1, 2n|a_{n-1}|, ..., 2n|a_0|$$

so $|x^k| > |x|$ *for all* $k > 1$, *and*

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

Thus, we have that

$$\left|\frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + ... + \frac{a_0}{x^n}\right| \leqslant \frac{|a_{n-1}|}{|x|} + ... + \frac{|a_0|}{|x^n|} < \underbrace{\frac{1}{2n} + ... + \frac{1}{2n}}_{n \text{ times}} = \frac{1}{2}$$

In other words,

$$-\frac{1}{2} < \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + ... + \frac{a_0}{x^n} < \frac{1}{2}$$

which implies that

$$\frac{1}{2} < 1 + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + ... + \frac{a_0}{x^n}$$

Choosing $x_1 > 0$ *which satisfies our condition, we have*

$$\frac{x_1^n}{2} \leqslant x_1^n \left(1 + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + ... + \frac{a_0}{x^n}\right) = f(x_1)$$

so that $f(x_1) > 0$. *On the other hand, choosing* $x_2 < 0$ *satisfying our condition,* $x_2^n < 0$ *as n is odd and*

$$\frac{x_2^n}{2} \geqslant x_2^n \left(1 + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + ... + \frac{a_0}{x^n}\right) = f(x_2)$$

so that $f(x_2) < 0$. *Applying the Intermediate Value Theorem to the interval* $[x_2, x_1]$ *we conclude that there exists* $c \in [x_2, x_1]$ *such that* $f(c) = 0$.

■

Theorem 1.3.13

If n is even and $f(x) = x^n + a_{n-1}x^{n-1} + ... + a_0$, *then there is a number y such that* $f(y) \leqslant f(x)$ *for all* $x \in \mathbb{R}$.

Proof.

Choose M such that

$$M = \max(1, 2n|a_{n-1}|, ..., 2n|a_0)$$

Then for all x with $|x| \geqslant M$ we have

$$\frac{1}{2} \leqslant 1 + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + ... + \frac{a_0}{x^n}$$

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

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

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

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

The same holds for $x \leqslant -b$. In particular, if $x \geqslant b$ or $x \leqslant -b$, then $f(x) \geqslant f(0)$. Applying the extreme value theorem to f on the interval $[-b, b]$, we conclude that there is a number $y \in [-b, b]$ such that if $-b \leqslant x \leqslant b$, then $f(y) \leqslant f(x)$. In particular, $f(y) \leqslant f(0)$. Thus, if $x \leqslant -b$ or $x \geqslant b$, then $f(x) \geqslant f(0) \geqslant f(y)$. Combining these results we find that $f(y) \leqslant f(x)$ for all $x \in \mathbb{R}$.

■

Corollary 1.3.14

Consider the equation

$$x^n + a_{n-1}x^{n-1} + ... + a_0 = c \tag{1.3.3}$$

for n even. Then there is a number m such that the equation has a solution for $c \geqslant m$ and has no solution for $c < m$.

Proof.

Let $f(x) = x^n + a_{n-1}x^{n-1} + ... + a_0$. According to our previous theorem there exists $y \in \mathbb{R}$ such that $f(y) \leqslant f(x)$ for all $x \in \mathbb{R}$. Let $m = f(y)$. If $c < m$ then the equation above has no solutiion, since the left hand side has a value $\geqslant m$ always. If $c = m$, then y is a solution of the equation. Finally, for $c > m$, let $b > y$ such that $f(b) > c$. Then the Intermediate Value Theorem applied to the interval $[y, b]$ states that there exists $x \in [y, b]$ such taht $f(x) = c$, so x is a solution of the equation.

■

1.3.2 Uniform Continuity

Definition 1.3.3

A function $f : \mathbb{R} \to \mathbb{R}$ is **uniformly continuous on an interval** I *if for every $\epsilon > 0$ there is some $\delta > 0$ such that, for all $x, y \in I$, if $|x - y| < \delta$ then $|f(x) - f(y)| < \delta$.*

Lemma 1.3.15

Let $a < b < c$ and let f be continuous on the interval $[a, c]$. Let $\epsilon > 0$ and suppose that

1. *if $x, y \in [a, b]$ and $|x - y| < \delta_1$, then $|f(x) - f(y)| < \epsilon$*
2. *if $x, y \in [b, c]$ and $|x - y| < \delta_2$, then $|f(x) - f(y)| < \epsilon$*

Then there is a $\delta > 0$ such that if $x, y \in [a, c]$ and $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$.

Proof.

Fix $\epsilon > 0$. Since f is continuous at b there exists $\delta_3 > 0$ such that if $|x - b| < \delta_3$, then $|f(x) - f(b)| < \frac{\epsilon}{2}$. It follows that if $|x - b| < \delta_3$ and $|y - b| < \delta_3$ then $|f(x) - f(y)| < \epsilon$. Choose $\delta = \min(\delta_1, \delta_2, \delta_3)$. Let $x, y \in [a, c]$ with $|x - y| < \delta$. If x and y are both in $[a, b]$, then $|f(x) - f(y)| < \epsilon$ by assumption. Similarly, if $x, y \in [b, c]$, then again $|f(x) - f(y)| < \epsilon$ by assumption. Finally, without loss of generality suppose $x < b < y$. Since $|x - y| < \delta$ we have that $|x - b| = |b - x| = b - x < y - x = |y - x| < \delta$ and similarly $|y - b| < \delta$. Thus, we have that $|f(x) - f(y)| < \epsilon$, completing the proof.

■

Theorem 1.3.16: Uniform Continuity Theorem

If f is continuous on $[a, b]$, then f is uniformly continuous on $[a, b]$.

Proof.

Consider $\epsilon > 0$. Define the set

$$A(\epsilon) := \{x \in [a, b] : \exists \delta > 0; \forall y, z \in [a, x]; |y - z| < \delta \implies |f(y) - f(z)| < \epsilon\}$$

Then $A(\epsilon) \neq \emptyset$ since $a \in A(\epsilon)$, and $A(\epsilon)$ is bounded above by b, so $A(\epsilon)$ has a least upper bound $\alpha_\epsilon \in \mathbb{R}$. Suppose towards a contradiction that $\alpha < b$. Since f is continuous at α, there is some δ_0 such that if $|y - \alpha| < \delta_0$, then $|f(y) - f(\alpha)| < \epsilon/2$. Consequently, if $|y - \alpha| < \delta_0$ and $|z - \alpha| < \delta_0$, then $|f(y) - f(z)| < \epsilon$. So, f surely satisfies the condition for containment in $A(\epsilon)$ on the interval $[\alpha - \delta_0, \alpha + \delta_0$. On the other hand, since α is the least upper bound of A it is also clear that the condition is satisfies on $[a, \alpha - \delta_0]$, namely $\alpha - \delta_0 \in A$. Then, the Lemma implies that f satisfies the condition on $[a, \alpha + \delta_0]$ since it satisfies it on $[a, \alpha - \delta_0]$ and $[\alpha - \delta_0, \alpha + \delta_0]$. Hence, $\alpha + \delta_0 \in A$, contradicting the fact that α is an upper bound.
To complete the proof we must show that $\alpha = b$ is in A. Since f is continuous at b, there is some $\delta_0 > 0$ such that if $y \in (b - \delta_0, b)$, then $|f(y) - f(b)| < \epsilon/2$. So, for any $x, y \in [b - \delta_0, b]$, $|f(y) - f(x)| < \epsilon$. But, f satisfies the condition for $A(\epsilon)$ on $[a, b - \delta_0]$ since b is the least upper bound of $A(\epsilon)$, so the Lemma implies that f satisfies the condition on $[a, b]$. Therefore,

as $\epsilon > 0$ was arbitrary, we conclude that f is uniformly continuous on $[a, b]$, completing the proof.

■

Chapter 2

Differentiation

2.1 Introduction to Derivatives

Definition 2.1.1: Differentiability

A function $f : \mathbb{R} \to \mathbb{R}$ is said to be **differentiable at** a *if*

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \tag{2.1.1}$$

exists. In this case the limit is denoted by $f'(a)$ *and is called the* **derivative of f at** a*. We also say that f is* **differentiable** *if f is differentiable at a for all a in its domain.*

Definition 2.1.2

We define the **tangent line** *to the graph of f at $(a, f(a))$ to be the line through $(a, f(a))$ with slope $f'(a)$. That is, the tangent line at $(a, f(a))$ is well defined if and only if f is differentiable at a.*

Remark 2.1.1

Given a function f, we denote by f' the function whose domain is the set of all numbers $a \in \mathbb{R}$ such that f is differentiable at a, and whose value at such a number a is

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \tag{2.1.2}$$

The function f' is called the **derivative** *of f.*

Notation

For a given function $f : \mathbb{R} \to \mathbb{R}$, the derivative f' is often denoted by

$$\frac{df(x)}{dx} \tag{2.1.3}$$

and the number $f'(a)$ is denoted by

$$\left.\frac{df(x)}{dx}\right|_{x=a} \tag{2.1.4}$$

Theorem 2.1.1

If f is differentiable at a, then f is continuous at a.

Proof.

Suppose f is differentiable at a point a. Then we have that the limit

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

exists. It follows by Limit Laws that

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

Thus, by Limit Laws the result that $\lim_{h\to 0} f(a+h) - f(a) = 0$ is equivalent to $\lim_{h\to 0} f(a+h) = \lim_{h\to 0} f(a) = f(a)$. Thus, f is continuous at a, replacing $a+h$ with x and $h \to 0$ with $x \to a$.

■

Definition 2.1.3: Higher Order Derivatives

Since the derivative of a function f is also a function, we can take its derivative to obtain the function $(f')' = f''$. In general, we denote the $k+1$-th derivative of f inductively by

$$\begin{aligned}
f^{(1)} &= f' \\
f^{(k+1)} &= (f^{(k)})'
\end{aligned}$$

These are called **higher order derivatives of f**. *We also define $f^{(0)} = f$. In Leibnitzian notation we write*

$$\frac{d^k f(x)}{dx} = f^{(k)} \tag{2.1.5}$$

2.2 Differentiation Results

Theorem 2.2.1

If f is a constant function, $f(x) = c$, then $f'(a) = 0$ for all $a \in \mathbb{R}$.

Proof.

Observe that for $a \in \mathbb{R}$,

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

as desired. ∎

Theorem 2.2.2

If f is the identity function, $f(x) = x$, then $f'(a) = 1$ for all $a \in \mathbb{R}$.

Proof.

Observe that for $a \in \mathbb{R}$,

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

as desired. ∎

Theorem 2.2.3: Linearity

If f and g are differentiable at a, then $f + cg$ is differentiable for all $c \in \mathbb{R}$

Proof.

Observe that

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

as desired. ∎

Theorem 2.2.4: Product Rule

If f and g are differentiable at a, then $f \cdot g$ is also differentiable at a and

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

Proof.

Observe that

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

as claimed, where $\lim_{h \to 0} f(a+h) = f(a)$ since f is differentiable at a, which implies it is also continuous at a.

∎

Theorem 2.2.5: Power Rule

IF $f(x) = x^n$ for some natural number n, then

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

for all a.

Proof.

For the proof we will proceed by induction on n. For $n = 1$ we have shown that $f'(a) = 1 = 1 \cdot a^0$, satisfying the base case. Assume that there exists $k \in \mathbb{N}$ such that if $n = k$, $f'(a) = ka^{k-1}$. Then, for the case of $n = k+1$ we may write $g(x) = x \cdot x^k = I(x) \cdot f(x)$. Hence, by the product rule we have that for all a

$$\begin{aligned}
g'(a) &= (I \cdot f)'(a) \\
&= I'(a) \cdot f(a) + I(a) \cdot f'(a) \\
&= 1 \cdot a^k + a \cdot ka^{k-1} \\
&= (k+1)a^k
\end{aligned}$$

as claimed. Hence, by mathematical induction we conclude that if $f(x) = x^n$ for $n \in \mathbb{N}$, then $f'(a) = na^{n-1}$ for all $a \in \mathbb{R}$.

∎

Theorem 2.2.6: Derivative of a Quotient

If g is differentiable at a, and $g(a) \neq 0$, then $1/g$ is differentiable at a and

$$\left(\frac{1}{g}\right)'(a) = \frac{-g'(a)}{|g(a)|^2}$$

Proof.

Note that since g is differentiable at a it is continuous at a. Moreover, since $g(a) \neq 0$, there exists $\delta > 0$ such that $g(a+h) \neq 0$ for $|h| < \delta$. Therefore, $(1/g)(a+h)$ is well defined for small enough h, and we can write

$$\begin{aligned}
\lim_{h\to 0} \frac{(1/g)(a+h) - (1/g)(a)}{h} &= \lim_{h\to 0} \frac{1/g(a+h) - 1/g(a)}{h} \\
&= \lim_{h\to 0} \frac{g(a) - g(a+h)}{h[g(a)\cdot g(a+h)]} \\
&= \lim_{h\to 0} \frac{-[g(a+h) - g(a)]}{h} \cdot \lim_{h\to 0} \frac{1}{g(a)\cdot g(a+h)} \\
&= -g'(a) \cdot \frac{1}{|g(a)|^2}
\end{aligned}$$

where $\lim_{h\to 0} 1/g(a+h) = 1/g(a)$ by continuity of g.

■

Theorem 2.2.7: Quotient Rule

If f and g are differentiable at a and $g(a) \neq 0$, then f/g is differentiable at a and

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

Proof.

Note that $f/g = f \cdot (1/g)$, so we have

$$\begin{aligned}
(f/g)'(a) &= (f \cdot 1/g)'(a) \\
&= f'(a)\cdot(1/g)(a) + f(a)\cdot(1/g)'(a) && \text{(Product Rule)} \\
&= \frac{f'(a)}{g(a)} - \frac{f(a)g'(a)}{|g(a)|^2} && \text{(Quotient Derivative)} \\
&= \frac{f'(a)g(a) - f(a)g'(a)}{|g(a)|^2}
\end{aligned}$$

as claimed.

■

Theorem 2.2.8: General Product Rule

If $f_1, f_2, \ldots, f_n$ are differentiable at a for some $n \in \mathbb{N}$, then $f_1 \cdot f_2 \cdot \ldots \cdot f_n$ is differentiable at a and

$$(f_1 \cdot \ldots \cdot f_n)'(a) = \sum_{i=1}^{n} f_1(a) \cdot \ldots \cdot f_{i-1}(a) \cdot f_i'(a) \cdot f_{i+1}(a) \cdot \ldots \cdot f_n(a)$$

Proof.

We proceed by induction on n. If $n = 1$ then $f_1'(a) = f_1'(a)$, so the base case holds. Now, suppose the claim is true for some $k \in \mathbb{N}$. Then it follows that if $n = k + 1$

$$\begin{aligned}
(f_1 \cdot \ldots \cdot f_k \cdot f_{k+1})'(a) &= (f_1 \cdot \ldots \cdot f_k)'(a) f_{k+1}(a) + (f_1 \cdot \ldots \cdot f_k)(a) f_{k+1}'(a) && \text{(Product Rule)} \\
&= \left[\sum_{i=1}^{k} f_1(a) \cdot \ldots \cdot f_{i-1}(a) \cdot f_i'(a) \cdot f_{i+1}(a) \cdot \ldots \cdot f_k(a)\right] f_{k+1}(a) \\
&+ f_1(a) \cdot \ldots \cdot f_k(a) \cdot f_{k+1}'(a) && \text{(by Induction Hypothesis)} \\
&= \sum_{i=1}^{k+1} f_1(a) \cdot \ldots \cdot f_{i-1}(a) \cdot f_i'(a) \cdot f_{i+1}(a) \cdot \ldots \cdot f_{k+1}(a)
\end{aligned}$$

as desired. Thus by mathematical induction we conclude that the formula holds for all $n \in \mathbb{N}$. ■

Theorem 2.2.9: Chain Rule

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.

Define a function ϕ as follows:

$$\phi(h) = \begin{cases} \frac{f(g(a+h)) - f(g(a))}{g(a+h) - g(a)}, & \text{if } g(a+h) - g(a) \neq 0 \\ f'(g(a)), & \text{if } g(a+h) - g(a) = 0 \end{cases} \tag{2.2.1}$$

Note that by differentiability of g at a, g is continuous at a as well so as $h \to 0$, $g(a+h) - g(a) \to 0$, so if $g(a+h) - g(a)$ is not zero, then $\phi(h)$ will approach $f'(g(a))$ as h goes to zero. If it is zero then $\phi(h)$ is exactly $f'(g(a))$. Note that as f is differentiable at $g(a)$ we have

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

Thus, if $\epsilon > 0$ there is some number $\delta' > 0$ such that, for all k,

$$(1)\ \text{if } 0 < |k| < \delta', \text{ then } \left| \frac{f(g(a) + k) - f(g(a))}{k} - f'(g(a)) \right| < \epsilon$$

Now, g is differentiable at a, hence continuous at a, so there is $\delta > 0$ such that for all h,

$$(2)\ \text{if } |h| < \delta, \text{ then } |g(a+h) - g(a)| < \delta'$$

Consider now any h with $|h| < \delta$. If $k = g(a+h) - g(a) \neq 0$, then

$$\phi(h) = \frac{f(g(a+h)) - f(g(a))}{g(a+h) - g(a)} = \frac{f(g(a)+k) - f(g(a))}{k}$$

it follows from (2) that $|k| < \delta'$, and hence from (1) that

$$|\phi(h) - f'(g(a))| < \epsilon$$

On the other hand, if $g(a+h) - g(a) = 0$, then $\phi(h) = f'(g(a))$, so it is surely true that

$$|\phi(h) - f'(g9a))| < \epsilon$$

We therefore have proved that

$$\lim_{h\to 0} \phi(h) = f'(g(a))$$

so ϕ is continuous at 0. If $h \neq 0$, then we have

$$\frac{f(g(a+h)) - f(g(a))}{h} = \phi(h) \cdot \frac{g(a+h) - g(a)}{h}$$

even if $g(a+h) - g(a) = 0$. Therefore, we have that

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

by continuity of $\phi(h)$ at 0.

■

2.3 Applications of Derivatives

Definition 2.3.1: Extrema

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

$$f(x) \geqslant f(y) \forall y \in A \tag{2.3.1}$$

The number $f(x)$ is itself called the **maximum value** *of f on A.*
A point $x \in A$ is a **minimum point** *for f on A if*

$$f(x) \leqslant f(y) \forall y \in A \tag{2.3.2}$$

The number $f(x)$ is itself called the **minimum value** *of f on A.*

Theorem 2.3.1

Let f be any function defined on (a,b). If x is an extremum point for f on (a,b), and f is differentiable at x, then $f'(x) = 0$.

Proof.

Consider the case where f has a maximum at x. If h is any number such that $x+h \in (a,b)$, then

$$f(x) \geqslant f(x+h)$$

since f has a maximum on (a,b) at x. This implies that

$$f(x+h) - f(x) \leqslant 0$$

Thus, if $h > 0$ we have that

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

and consequently

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

as otherwise $\frac{f(x+h)-f(x)}{h} > 0$ for some h, contradicting our initial assumptions. Similarly, if $h < 0$ we have

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

so

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

By hypothesis f is differentiable at x, so these two limits must be equal, so in fact $f'(x) \leqslant 0$ and $f'(x) \geqslant 0$. Thus, $f'(x) = 0$.

On the other hand, suppose f has a minimum at x. Then $-f$ has a maximum at x. Indeed, for all $y \in (a,b)$ we have $f(y) \geqslant f(x)$, so $-f(y) \leqslant -f(x)$. Then, from our above argument and the differentiability of f at x, we have $-f'(x) = 0$, which implies that $f'(x) = 0$.

■

Definition 2.3.2: Local Extrema

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 $A \cap (x-\delta, x+\delta)$.*

Definition 2.3.3

A **critical point** *of a function f is a number x such that*

$$f'(x) = 0 \tag{2.3.3}$$

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

Remark 2.3.1

Give a function continuous f, if x is an extrumum of f on $[a,b]$, then one of the following must be satisfied:

1. *x is a critical point of f in $[a,b]$*
2. *$x = a$ or $x = b$ so x is an endpoint of $[a,b]$*
3. *x is a point in $[a,b]$ such that f is not differentiable at x*

Theorem 2.1: Rolle's 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.

It follows from continuity of f on $[a,b]$ that f has a maximum or minimum value on $[a,b]$ (by the Extreme Value Theorem).
Suppose first that the maximum value occurs at a point $x \in (a, b0$. Then $f'(x) = 0$ by Theorem 2.3.1. On the other hand suppose that the minimum value of f occurs at some point x in (a,b). Then, again, $f'(x) = 0$ by Theorem 2.3.1.
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 $x \in (a,b)$ and have $f'(x) = 0$, completing the proof. ■

Theorem 2.2: The 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

$$f'(x) = \frac{f(b) - f(a)}{b - a} \tag{2.3.4}$$

Proof.

Let

$$h(x) = f(x) - \left[\frac{f(b) - f(a)}{b - a}\right](x - a)$$

Evidently, h is continuous on $[a,b]$ and differentiable on (a,b) as it is the sum of correspondingly continuous and differentiable functions. Moreover,

$$\begin{aligned} h(a) &= f(a) \\ h(b) &= f(b) - \left[\frac{f(b) - f(a)}{b - a}\right](b - a) \\ &= f(a) \end{aligned}$$

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

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

so that

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

as desired.

■

Corollary 2.3.2

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.: L

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

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

so $f(b) - f(a) = 0$ and consequently $f(a) = f(b)$. Thus the value of f at any two points in the interval is the same, so f is constant on the interval.

■

Corollary 2.3.3

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

Proof.

For all x in the interval we have $(f - g)'(x) = f'(x) - g'(x) = 0$, so by the previous corollary there is some number c such that $f - g = c$.

■

Definition 2.3.4

A function is **increasing** *on an interval I if $f(a) < f(b)$ whenever $a, b \in I$ with $a < b$. The function f is* **decreasing** *on an interval I if $f(a) > f(b)$ for all $a, b \in I$ with $a < b$.*

Corollary 2.3.4

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.

Consider the case where $f'(x) > 0$. Let $a, b \in I$ with $a < b$. Then by The Mean Value Theorem there exists $x \in (a,b)$ such that

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

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

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

Since $b - a > 0$ we conclude that $f(b) > f(a)$ so f is increasing.
Next, consider the case for $f'(x) < 0$. Then $-f'(x) > 0$ for all $x \in I$, so by the first case we have that for all $a, b \in I$ with $a < b$, $-f(a) < -f(b)$. Multiplying both sides by -1 we have that $f(a) > f(b)$ for all $a, b \in I$ such that $a < b$, so f is decreasing, as desired.

■

Theorem 2.3.5: Second Derivative Test

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 $f'(a) = 0$ by assumption, we can write

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

Suppose now that $f''(a) > 0$. Then there exists $\delta > 0$ such that if $|h| < \delta$ $f'(a+h)/h > 0$. Thus, for $|h| < \delta$, if $h < 0$ we must have $f'(a+h) < 0$ and if $h > 0$ we must have $f'(a+h) > 0$. This means by our previous corollary that f is increasing in the interval $(a, a + \delta)$, and decreasing in $(a - \delta, a)$. Thus, as $f'(a) = 0$, $f(a)$ must be a local minimum.
If $f''(a) < 0$, then $-f''(a) > 0$ so $-f(a)$ is must be a local minimum. That is, there exists $\delta > 0$ such that if $x \in (a - \delta, a + \delta)$, then $-f(x) \geqslant -f(a)$. Hence, it follows that $f(x) \leqslant f(a)$ for all $x \in (a - \delta, a + \delta)$, so $f(a)$ is a local maximum of f.

■

Theorem 2.3.6

Suppose $f''(a)$ exists. If f has a local minimum at a, then $f''(a) \geqslant 0$; if f has a local maximum at a, then $f''(a) \leqslant 0$.

Proof.

Suppose f has a local minimum at a. If $f''(a) < 0$ then by our previous result f would have a local maximum at a. But, this implies that f would be constant in some interval containing a, so that $f''(a) = 0$, which is a contradiction. Thus, we must have that $f''(a) \geqslant 0$.
The case for a local maximum is analogous.

■

Theorem 2.3.7

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*

$$f'(a) = \lim_{x\to a} f'(x) \tag{2.3.5}$$

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 $(a, a+h)$, by assumption (similarly for sufficiently small $h < 0$). By The Mean Value Theorem there is a number $\alpha_h \in (a, a+h)$ such that

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

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

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

For this last equality write $\lim_{x\to a} f'(x) = L \in \mathbb{R}$. Fix $\epsilon > 0$. Then there exists $\delta > 0$ such that for all $x \in (a-\delta, a+\delta)$, $|f'(x) - L| < \epsilon$. It follows that for $|h| < \delta$, if $h > 0$ and $\alpha_h \in (a, a+h) \subset (a-\delta, a+\delta)$ we have $|f'(\alpha_h) - L| < \epsilon$ and if $h < 0$ and $\alpha_h \in (a+h, a) \subset (a-\delta, a+\delta)$, then $|f'(\alpha_h) - L| < \epsilon$. Thus, by definition we have that $\lim_{h\to 0^+} f'(\alpha_h) = \lim_{h\to 0^-} f'(\alpha_h) = L$, so in particular $\lim_{h\to 0} f'(\alpha_h) = L = \lim_{x\to a} f'(x)$, completing the proof.

■

Theorem 2.3: The Cauchy Mean Value Theorem

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

$$[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x) \tag{2.3.6}$$

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 by Rolle's Theorem that $h'(x) = 0$ for some $x \in (a, b)$, which implies that

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

completing the proof.

■

Theorem 2.4: L'Hôpital's Rule

Suppose that

$$\lim_{x\to a} f(x) = 0 \text{ and } \lim_{x\to a} g(x) = 0 \tag{2.3.7}$$

and suppose also that $\lim_{x\to a} f'(x)/g'(x)$ *exists. Then* $\lim_{x\to a} f(x)/g(x)$ *exists, and*

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} \tag{2.3.8}$$

Proof.

The hypothesis that $\lim_{x\to a} f'(x)/g'(x)$ *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,* $x = a$,
2. *in this interval* $g'(x) \neq 0$, *with the possible exception of* $x = a$

If we define $f(a) = g(a) = 0$, *then* f *and* g *are continuous at* a. *If* $x \in (a, a+\delta)$, *then The Mean Value Theorem and The Cauchy Mean Value Theorem apply to* f *and* g *on* $[a, x]$ *(a similar statement holds for* $x \in (a-\delta, a)$*). First, applying the The Mean Value Theorem to* g, *we see that* $g(x) \neq 0$, *for if* $g(x) = 0$ *there would exist* $x_1 \in (a, x)$ *with* $g'(x_1) = 0$, *contradicting 2.. Now, applying The Cauchy Mean Value Theorem to* f *and* g, *we see that there is a number* $\alpha_x \in (a, x)$ *such that*

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

or

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

Now, let $\lim_{y\to a} f'(y)/g'(y) = L \in \mathbb{R}$. *Fix* $\epsilon > 0$. *Then there exists* $\delta' > 0$ *such that if* $y \in (a-\delta', a+\delta')$ *then* $|f'(y)/g'(y) - L| < \epsilon$. *Then, for* $x \in (a, a+\delta)$ *(or* $x \in (a-\delta, a)$*) we have* $(a, x) \subset (a-\delta, a+\delta)$ *(or* $(x, a) \subset (a-\delta, a+\delta)$*). Thus, for* $|x-a| < \delta$ *we have* $\alpha_x \in (a, x) \subset (a-\delta, a+\delta)$ *(or* $\alpha_x \in (x, a) \subset (a-\delta, a+\delta)$*), so* $|f'(\alpha_x)/g'(\alpha_x) - L| < \epsilon$. *Therefore, we conclude that*

$$\lim_{x\to a^+} \frac{f'(\alpha_x)}{g'(\alpha_x)} = L = \lim_{x\to a^-} \frac{f'(\alpha_x)}{g'(\alpha_x)}$$

so in particular

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

completing the proof.

■

2.3.1 Convexity

Definition 2.3.5

A function f is **convex** *on an interval I, if for all $a, b \in I$, the line segment joining $(a, f(a))$ and $(b, f(b))$ lies above the graph of f.*
This is equivalent to stating that for all $x \in (a, b)$,

$$\frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(a)}{b - a} \tag{2.3.9}$$

Definition 2.3.6

A function f is **concave** *on an interval I, if for all $a, b \in I$, the line segment joining $(a, f(a))$ and $(b, f(b))$ lies below the graph of f.*
This is equivalent to stating that for all $x \in (a, b)$,

$$\frac{f(x) - f(a)}{x - a} > \frac{f(b) - f(a)}{b - a} \tag{2.3.10}$$

Theorem 2.3.8

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 $(a, f(a))$ itself. If $a < b$ and f is differentiable at a and b, then $f'(a) < f'(b)$.

Proof.

If $0 < h_1 < h_2$, then $a < a + h_1 < a + h_2$, and applying f's convexity we have that

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

This implies that the values of $[f(a + h) - f(a)]/h$ decrease as $h \to 0^+$. Consequently,

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

In fact, $f'(a)$ is the infimum of these numbers. Similarly, for h negative, if $h_2 < h_1 < 0$, then

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

This shows that the slope of the tangent line is greater that $[f(a + h) - f(a)]/h$ for $h < 0$. In fact, $f'(a)$ is the supremum of all these numbers, so $f(a + h)$ lies above the tangent line if $h < 0$. This satisfies the first part of the theorem. Now, suppose $a < b$. Then we have that

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

since $b - a > 0$ and

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

since $a - b < 0$. Combining these inequalities we obtain $f'(a) < f'(b)$, as desired.

■

Lemma 2.3.9

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 towards a contradiction that $f(x) \geqslant f(a) = f(b)$ for some $x \in (a, b)$. Then the maximum of f on $[a, b]$ occurs at some point $x_0 \in (a, b)$ with $f(x_0) \geqslant f(a)$ and, of course, $f'(x_0) = 0$. 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

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

contradicting the fact that f' is increasing (since $f'(x_0) = 0$ and $x_1 < x_0$).

■

Theorem 2.3.10

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 easy to see that g' is also increasing; moreover, $g(a) = g(b) = f(a)$. Applying the lemma to g we conclude that

$$a < x < b \implies g(x) < f(a)$$

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.

■

Theorem 2.3.11

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 $a < b$. Since the tangent lien 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

$$(1)\ \ f(b) > f'(a)(b - a) + f(a)$$

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

$$(2)\ \ f(a) > f'(b)(a - b) + f(b)$$

It follows from (1) and (2) that $f'(a) < f'(b)$. Then, from our previous theorem we have that f is convex.

■

2.4 Inverse Functions

Definition 2.4.1

For any function f. the **inverse image** *of f, denoted by f^{-1}, is the set of all pairs (a, b) such that $(b, a) \in f$.*

Remark 2.4.1

f^{-1} is a function if and only if f is one-to-one.

Theorem 2.4.1

If f is increasing (decreasing) on an interval I, then f is injective on I so f^{-1} is a function and in fact f^{-1} is increasing (decreasing).

Proof.

Consider the case that f is increasing. Then suppose $a, b \in I$ with $a \neq b$. Without loss of generality suppose $a < b$. Then since f is increasing $f(a) < f(b)$ so in particular $f(a) \neq f(b)$. Therefore, f is injective as claimed, so f^{-1} is a well-defined function on I. Now, consider $a' < b'$ in $f(I) = I'$. Then there exist $x, y \in I$ such that $f(x) = a'$ and $f(y) = b'$, so in particular $f^{-1}(a') = x$ and $f^{-1}(b') = y$. Since f is increasing and $f(x) = a' < b' = f(y)$ we must have that $x < y$. Thus, $f^{-1}(a') = x < y = f^{-1}(b')$, so f^{-1} is increasing as claimed.
Consider the case that f is decreasing. Then $-f$ is increasing so it is injective and $-f^{-1}$ is increasing by the first case. Hence, we have that f^{-1} is decreasing as desired.

■

Theorem 2.4.2

If f is continuous and one-to-one on an interval I, then f is either increasing or decreasing on I.

Proof.

We proceed in three steps:
(1) If $a < b < c$ are three points in I, then I claim either $f(a) < f(b) < f(c)$ or $f(a) > f(b) > f(c)$. Indeed, suppose that $f(a) < f(c)$. If we have $f(b) < f(a)$, then the Intermediate Value Theorem applied to $[b, c]$ gives an $x \in (b, c)$ such that $f(x) = f(a)$, contradicting the fact that f is injective on $[a, c]$. Similarly, if $f(b) > f(c)$ we would find a contradiction, so $f(a) < f(b) < f(c)$. Similar argumentation leads to the result that $f(a) > f(b) > f(c)$ in the second case.
(2) If $a < b < c < d$ are four points in I, then I claim that either $f(a) < f(b) < f(c) < f(d)$ or $f(a) > f(b) > f(c) > f(d)$. Indeed we can apply (1) to $a < b < c$ and then to $b < c < d$.
(3) Take any $a < b$ in I, and suppose $f(a) < f(b)$. Then f is increasing, for if $c, d \in I$ are any two points, we can apply (2) to the collection $\{a, b, c, d\}$ after arranging them in increasing order.

■

Theorem 2.4.3

If f is continuous and one-to-one on an interval, then f^{-1} is also continuous.

Proof.

Since f is continuous and injective on the interval, it is either increasing or decreasing. Consider the case that f is increasing. 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. For any $\epsilon > 0$, we want to find a $\delta > 0$ such that for all x, if $x \in (f(a) - \delta, f(a) + \delta)$, then $|f^{-1}(x) - a| < \epsilon$, as $a = f^{-1}(b) = f^{-1}(f(a))$. Now, since $a - \epsilon < a < a + \epsilon$ we have that $f(a - \epsilon) < f(a) < f(a + \epsilon)$ since f is presumed increasing. Let $\delta = \min(f(a + \epsilon) - f(a), f(a) - f(a - \epsilon))$. Our choice of δ ensures that

$$f(a - \epsilon) \leqslant f(a) - \delta \text{ and } f(a) + \delta \leqslant f(a + \epsilon)$$

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

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

so $a - \epsilon < f^{-1}(x) < a + \epsilon$, which is precisely $|f^{-1}(x) - a| < \epsilon$, as desired.

■

Theorem 2.4.4

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

Proof.

We have $f(f^{-1}(x)) = x$. If f^{-1} were differentiable at a, then 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.

■

Theorem 2.4.5

Let f be a continuous one-to-one function defined on an interval I, 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

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))} \tag{2.4.1}$$

Proof.

Let $b = f(a)$. Then

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

Now, every number $b+h$ in the domain of f^{-1} can be written in the form $b+h = f(a+k)$ for a unique $k(h)$. Then

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

Since $b+h = f(a+k(h))$ we have $f^{-1}(b+h) = a + k(h)$, or $k(h) = f^{-1}(b+h) - f^{-1}(b)$. Now, since f is continuous on I, f^{-1} is also continuous on its domain, and in particular it is continuous at b. This means that $\lim_{h\to 0} k(h) = 0$, so $k(h)$ goes to zero as h goes to 0. Hence, as

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

this implies that f^{-1} is differentiable at b and

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

■

Chapter 3

Integration

3.1 Introduction to Definite Integrals

Definition 3.1.1

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.*

The points in a partition can be numbered $t_0, ..., t_n$ so that

$$a = t_0 < t_1 < ... < t_{n-1} < t_n = b \tag{3.1.1}$$

we shall always assume that such a numbering has been assigned.

Definition 3.1.2

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

$$m_i = \inf\{f(x) : t_{i-1} \leqslant x \leqslant t_i\} \tag{3.1.2}$$
$$M_i = \sup\{f(x) : t_{i-1} \leqslant x \leqslant t_i\} \tag{3.1.3}$$

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

$$L(f,P) := \sum_{i=1}^{n} m_i(t_i - t_{i-1}) \tag{3.1.4}$$

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

$$U(f,P) = \sum_{i=1}^{n} M_i(t_i - t_{i-1}) \tag{3.1.5}$$

Remark 3.1.1

If P is any partition, then

$$L(f, P) \leqslant U(f, P) \tag{3.1.6}$$

because

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

and for each i we have $m_i(t_i - t_{i-1}) \leqslant M_i(t_i - t_{i-1})$.

Lemma 3.1.1

If Q is a partition of $[a, b]$ which contains P, then

$$L(f, P) \leqslant L(f, Q)$$
$$U(f, P) \geqslant U(f, Q)$$

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

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

Let

$$m' = \inf\{f(x) : t_{k-1} \leqslant x \leqslant u\}$$
$$m'' = \inf\{f(x) : u \leqslant x \leqslant t_k\}$$

Then

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

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

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

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

$$m_k \leqslant m'$$

Similarly,

$$m_k \leqslant m''$$

Therefore,

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

This proves, in this special case that $L(f, P) \leqslant L(f, Q)$. *Now, let*

$$M' = \sup\{f(x) : t_{k-1} \leqslant x \leqslant u\}$$
$$M'' = \sup\{f(x) : u \leqslant x \leqslant t_k\}$$

Then

$$U(f, P) = \sum_{i=1}^{n} M_i(t_i - t_{i-1})$$
$$U(f, Q) = \sum_{i=1}^{k-1} M_i(t_i - t_{i-1}) + M'(u - t_{k-1}) + M''(t_k - u) + \sum_{i=k+1}^{n} M_i(t_i - t_{i-1})$$

Hence, to prove that $U(f, Q) \leqslant U(f, P)$ *it suffices to show that*

$$M'(u - t_{k-1}) + M''(t_k - u) \leqslant M_k(t_k - t_{k-1})$$

As before, the set $\{f(x) : t_{k-1} \leqslant x \leqslant t_k\}$ *contains all the numbers in* $\{f(x) : t_{k-1} \leqslant x \leqslant u\}$ *and possibly some larger ones, so the smallest upper bound of the first set is greater than or equal to the smallest upper bound of the second; thus*

$$M_k \geqslant M'$$

Similarly,

$$M_k \geqslant M''$$

Therefore,

$$M_k(t_k - t_{k-1}) = M_k(t_k - u) + M_k(u - t_{k-1}) \geqslant M''(t_k - u) + M'(u - t_{k-1})$$

This proves, in this special case that $U(f, P) \geqslant U(f, Q)$.
The general case can now be deduced quite easily. The partition Q can be obtained from P by adding one point at a time; in otherwords, there is a sequence of partition

$$P = P_1 \subsetneq P_2 \subsetneq P_3 \subsetneq ... \subsetneq P_\alpha = Q$$

such that $P_{j+1} = P_j \cup \{u_{j+1}\}$ *for some* $u_{j+1} \in [a, b] - P_j$. *Then*

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

and

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

completing the proof.

■

Theorem 3.1.2

Let P_1 and P_2 be partitions of $[a,b]$, and let f be a function which is bounded on $[a,b]$. Then

$$L(f,P_1) \leqslant U(f,P_2) \tag{3.1.7}$$

Proof.

There is a partition P which contains both P_1 and P_2 (let $P = P_1 \cup P_2$). According to the lemma

$$L(f,P_1) \leqslant L(f,P) \leqslant U(f,P) \leqslant U(f,P_1)$$

■

Remark 3.1.2

It follows 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:

$$\sup\{L(f,P) : P \subset [a,b]; \exists n \in \mathbb{N}, |P| = n\} \leqslant U(f,P') \tag{3.1.8}$$

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

$$\sup\{L(f,P)\} \leqslant \inf\{U(f,P)\} \tag{3.1.9}$$

It is clear that for all partitions P',

$$L(f,P') \leqslant \sup\{L(f,P\} \leqslant \inf\{U(f,P)\} \leqslant U(f,P') \tag{3.1.10}$$

Definition 3.1.3: Definite Integral

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

$$\int_a^b f \tag{3.1.11}$$

The integral $\int_a^b f$ is also called the **area** *of $R(f,a,b)$ when $f(x) \geqslant 0$ for all $x \in [a,b]$.*

Theorem 3.1.3

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 such a partition P. Since

$$\inf\{U(f,P')\} \leqslant U(f,P)$$
$$\sup\{L(f,P')\} \geqslant L(f,P)$$

it follows that

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

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

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

so by definition, then, f is integrable. Next, if f is integrable then

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

Let M denote the value of this. Then for each $\epsilon > 0$ there exist partitions P' and P'' such that $|U(f,P') - M| < \epsilon/3$ and $|L(f,P'') - M| < \epsilon/2$. Then as $U(f,P') \geqslant L(f,P'')$ from the previous theorem, we have that

$$U(f,P') - L(f,P'') = |U(f,P') - L(f,P'')| \leqslant |U(f,P') - M| + |M - L(f,P'')| < \epsilon$$

Let $P = P' \cup P''$ be a partition. Then, according to the lemma $U(f,P) \leqslant U(f,P')$ and $L(f,P) \geqslant L(f,P'')$ so

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

■

Theorem 3.1.4

If f is continuous on $[a,b]$, then f is integrable on $[a,b]$.

Proof.

Notice, first, that f is bounded on $[a,b]$, because it is continuous on $[a,b]$. To prove that f is integrable on $[a,b]$, we want to use our previous theorem, and show that for every $\epsilon > 0$ there is a partition P of $[a,b]$ such that

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

Now we know, by our result on uniform continuity, that f is uniformly continuous on $[a,b]$. So there is some $\delta > 0$ such that for all $x, y \in [a,b]$, if $|x-y| < \delta$, then $|f(x)-f(y)| < \epsilon/[2(b-a)]$. We choose a partition $P = \{t_0, ..., t_n\}$ such that each $|t_i - t_{i-1}| < \delta$. Then for each i we have

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

for all $x, y \in [t_{i-1}, t_i]$. Then, for the sake of contradiction suppose $M_i - m_i > \frac{\epsilon}{2(b-a)} = \epsilon'$, and let $\delta = \frac{M_i - m_i - \epsilon'}{2}$. Then as $M_i - \delta$ and $m_i + \delta$ are not upper and lower bounds of $\{f(x) : t_{i-1} \leqslant x \leqslant$

$t_i\}$ respectively, there exist $f(u), f(v) \in \{f(x) : t_{i-1} \leqslant x \leqslant t_i\}$ such that $M_i - \delta < f(u) \leqslant M_i$ and $m_i \leqslant f(v) < m_i + \delta$. It follows that

$$\epsilon' = M_i - m_i - 2\delta < f(u) - f(v) \leqslant |f(u) - f(v)| < \epsilon'$$

However, this implies that $\epsilon' < \epsilon'$, a contradiction. Thus, we have that

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

Since this is true for all i, we have that

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

Thus, by our previous theorem f is integrable.

■

Theorem 3.1.5

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

$$\int_a^b f = \int_a^c f + \int_c^b f \tag{3.1.12}$$

Proof.

(1) Suppose f is integrable on $[a,b]$. Then f is bounded on $[a,b]$, so it is bounded on $[a,c]$ and $[c,b]$. Indeed, f being bounded implies that there exists $M \in \mathbb{R}$ such that for all $x \in [a,b]$ $|f(x)| \leqslant M$. Thus, as this applies for all $x \in [a,b]$ and $[a,c],[c,b] \subset [a,b]$, we have that it holds for all $x \in [a,c]$ and all $x \in [c,b]$. Now fix $\epsilon > 0$. Then there exists a partition P of $[a,b]$ such that

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

Without loss of generality suppose $c = t_j$ for some $t_j \in P = \{t_0, t_1, ..., t_n\}$. Then we have partitions $P' = \{t_0, ..., t_j\}$ and $P'' = \{t_j, ..., t_n\}$ for $[a,c]$ and $[c,b]$ respectively. Moreover,

$$\begin{aligned}
U(f,P) &= U(f,P') + U(f,P'') \\
L(f,P) &= L(f,P') + L(f,P'')
\end{aligned}$$

Hence, we have that

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

But $U(f, P') \geqslant L(f, P')$ and $U(f, P'') \geqslant L(f, P'')$, so

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

Therefore, f is integrable on $[a, c]$ and $[c, b]$
(2) Suppose f is integrable on $[a, c]$ and $[c, b]$. Thus, there exists $M_1, M_2 \in \mathbb{R}$ such that for all $x \in [a, c]$ $|f(x)| \leqslant M_1$ and for all $x \in [c, b]$ $|f(x)| \leqslant M_2$. Let $M = \max(M_1, M_2)$. Then for all $x \in [a, b]$ we have $|f(x)| \leqslant M$, so f is bounded on $[a, b]$. Let $\epsilon > 0$. Then there exist partitions P_1, P_2 of $[a, c]$ and $[c, b]$ respectively such that

$$U(f, P_1) - L(f, P_1) < \epsilon/2$$
$$U(f, P_2) - L(f, P_2) < \epsilon/2$$

Let $P = P_1 \cup P_2$, where $P_1 \cap P_2 = \{c\}$. Then we have that

$$U(f, P) - L(f, P) = [U(f, P_1) - L(f, P_1)] + [U(f, P_2) - L(f, P_2)] < \epsilon/2 + \epsilon/2 = \epsilon$$

Therefore, by definition f is integrable on $[a, b]$.
(3) Suppose f is integrable on $[a, b]$, so by the previous results f is integrable on $[a, c]$ and $[c, b]$. Let $\int_a^b f = R$, $\int_a^c f = R_1$, and $\int_c^b f = R_2$. Let P be a partition of $[a, b]$, and without loss of generality suppose $c \in P = \{t_0, ..., t_j = c, ..., t_n\}$. Then let $P_1 = \{t_0, ..., t_j\}$ and $P_2 = \{t_j, ..., t_n\}$ be partitions of $[a, c]$ and $[c, b]$. It then follows that

$$L(f, P_1) \leqslant R_1 \leqslant U(f, P_1)$$
$$L(f, P_2) \leqslant R_2 \leqslant U(f, P_2)$$

Hence, we have that

$$L(f, P) = L(f, P_1) + L(f, P_2) \leqslant R_1 + R_2$$
$$U(f, P) = U(f, P_1) + U(f, P_2) \geqslant R_1 + R_2$$

Thus $L(f, P) \leqslant R_1 + R_2 \leqslant U(f, P)$. Note that this holds for all partitions P, as if P' is a partition, then considering the partition $P'_c = P' \cup \{c\}$ we have that

$$L(f, P') \leqslant L(f, P'_c) \leqslant R_1 + R_2 \leqslant U(f, P'_c) \leqslant U(f, P')$$

Therefore, this holds for all partitions of $[a, b]$, but R is the unique number which does this so we must have that $R = R_1 + R_2$. Thus

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

■

Definition 3.1.4

Using the previous theorem, we defin

$$\int_a^a f := 0 \quad and \quad \int_a^b f := -\int_b^a f,\ for\ a > b \tag{3.1.13}$$

Theorem 3.1.6

If f and g are integrable on $[a,b]$, then $f+g$ is integrable on $[a,b]$ and

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g \tag{3.1.14}$$

Proof.

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

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

and define M_i, M_i', M_i'' similary. Then it follows that

$$m_i \geqslant m_i' + m_i''$$

and

$$M_i \leqslant M_i' + M_i''$$

Therefore, we have that

$$\begin{aligned} L(f+g,P) &\geqslant L(f,P) + L(g,P) \\ U(f+g,P) &\leqslant U(f,P) + U(g,P) \end{aligned}$$

Thus we have that

$$L(f,P) + L(g,P) \leqslant L(f+g,P) \leqslant U(f+g,P) \leqslant U(f,P) + U(g,P)$$

Fix $\epsilon > 0$. Since f and g are integrable there exist partitions P' and P'' of $[a,b]$ such that

$$\begin{aligned} U(f,P') - L(f,P') &< \epsilon/2 \\ U(g,P'') - L(g,P'') &< \epsilon/2 \end{aligned}$$

Let $P_\epsilon = P' \cup P''$. Then we have that

$$U(f+g,P_\epsilon) - L(f+g,P_\epsilon) \leqslant [U(f,P_\epsilon) - L(f,P_\epsilon)] + [U(g,P_\epsilon) - L(g,P_\epsilon)] < \epsilon/2 + \epsilon/2 = \epsilon$$

Therefore $f+g$ is integrable on $[a,b]$. Moreover,

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

and also

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

Then, observe that

$$\begin{aligned}-[U(f,P)+U(g,P)]+(L(f,P)+L(g,P)) &\leqslant \left(\int_a^b f+\int_a^b g\right)-\int_a^b (f+g)\\ &\leqslant U(f,P)+U(g,P)-(L(f,P)+L(g,P))\end{aligned}$$

But, this applies for all partitions P, so in particular from above we have that for $\epsilon>0$ there exists a partition P_ϵ such that

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

Thus, we have for all $\epsilon>0$ that

$$\left|\left(\int_a^b f+\int_a^b g\right)-\int_a^b (f+g)\right|<\epsilon$$

Consequently, we have that

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

■

Theorem 3.1.7

If f is integrable on $[a,b]$, then for any $c\in\mathbb{R}$, the function cf is integrable on $[a,b]$ and

$$\int_a^b cf=c\cdot\int_a^b f \tag{3.1.15}$$

Proof.

(1) Consider $c\geqslant 0$. Let $P=\{t_0,...,t_n\}$ be a partition of $[a,b]$, define

$$\begin{aligned}m_i&=\inf\{(cf)(x):t_{i-1}\leqslant x\leqslant t_i\}\\ m_i'&=\inf\{f(x):t_{i-1}\leqslant x\leqslant t_i\}\end{aligned}$$

and define M_i and M_i' similarly. I claim that $m_i\geqslant cm_i'$ and $M_i\leqslant cM_i'$. Indeed, for all $x\in[t_{i-1},t_i]$ we have that $(cf)(x)\leqslant cM_i'$, so $M_i\leqslant cM_i'$. Similarly, $(cf)(x)\geqslant cm_i'$ so $m_i\geqslant cm_i'$. Then it follows that

$$L(cf,P)\geqslant cL(f,P)$$

and

$$U(cf,P)\leqslant cU(f,P)$$

Thus, we have that

$$cL(f,P)\leqslant L(cf,P)\leqslant U(cf,P)\leqslant cU(f,P)$$

If $c = 0$ then we have that $0 \leqslant L(cf, P) \leqslant U(cf, P) \leqslant 0$, so $L(cf, P) = U(cf, P) = 0$ for all partitions P, and consequently cf is integrable with $\int_a^b cf = 0$. Otherwise, fix $\epsilon > 0$. Then there exists a partition P' such that

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

Then it follows that

$$U(cf, P') - L(cf, P') \leqslant cU(f, P') - cL(f, P') < \epsilon$$

Thus cf is integrable on $[a, b]$. Then we have that

$$cL(f, P) \leqslant L(cf, P) \leqslant \int_a^b cf \leqslant U(cf, P) \leqslant cU(f, P)$$

and also

$$cL(f, P) \leqslant c\int_a^b f \leqslant cU(f, P)$$

Then we have that

$$\left|\int_a^b cf - c\int_a^b f\right| \leqslant cU(f, P) - cL(f, P)$$

where we can make $cU(f, P) - cL(f, P)$ as small as we want by choosing an appropriate partition. Therefore,

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

(2) Consider $c < 0$. Let $P = \{t_0, ..., t_n\}$ be a partition of $[a, b]$, define

$$m_i = \inf\{(cf)(x) : t_{i-1} \leqslant x \leqslant t_i\}$$
$$m_i' = \inf\{f(x) : t_{i-1} \leqslant x \leqslant t_i\}$$

and define M_i and M_i' similarly. I claim that $m_i \geqslant cM_i'$ and $M_i \leqslant cm_i'$. Indeed, for all $x \in [t_{i-1}, t_i]$ we have that $(cf)(x) = cf(x) \leqslant cm_i'$, since $f(x) \geqslant m_i'$ and $c < 0$, so $M_i \leqslant cm_i'$. Similarly, $(cf)(x) \geqslant cM_i'$ so $m_i \geqslant cM_i'$. Then it follows that

$$L(cf, P) \geqslant cU(f, P)$$

and

$$U(cf, P) \leqslant cL(f, P)$$

Thus, we have that

$$cU(f, P) \leqslant L(cf, P) \leqslant U(cf, P) \leqslant cL(f, P)$$

Note that $U(f, P) \geqslant L(f, P)$ so as $c < 0$, $cU(f, P) \leqslant cL(f, P)$. Fix $\epsilon > 0$. Then there exists a partition P' such that

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

Then it follows that

$$U(cf,P') - L(cf,P') \leqslant cL(f,P') - cU(f,P') < \epsilon$$

as $\epsilon/c < L(f,P') - U(f,P')$ *and* $c < 0$. *Thus* cf *is integrable on* $[a,b]$. *Then we have that*

$$cU(f,P) \leqslant L(cf,P) \leqslant \int_a^b cf \leqslant U(cf,P) \leqslant cL(f,P)$$

and also

$$cU(f,P) \leqslant c\int_a^b f \leqslant cL(f,P)$$

Then we have that

$$\left|\int_a^b cf - c\int_a^b f\right| \leqslant cL(f,P) - cU(f,P)$$

where we can make $cL(f,P) - cU(f,P)$ *as small as we want by choosing an appropriate partition. Therefore,*

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

completing the proof.

■

Theorem 3.1.8

Suppose f *is integrable on* $[a,b]$ *and that*

$$m \leqslant f(x) \leqslant M \forall x \in [a,b] \tag{3.1.16}$$

Then

$$m(b-a) \leqslant \int_a^b f \leqslant M(b-a) \tag{3.1.17}$$

Proof.

It is clear that $M \geqslant \sup\{f(x) : x \in [a,b]\}$ *and* $m \leqslant \inf\{f(x) : x \in [a,b]\}$, *so*

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

for every partition P. *Since* $\int_a^b f = \sup\{L(f,P)\} = \inf\{U(f,P)\}$ *we have that*

$$m(b-a) \leqslant \sup\{L(f,P)\} = \int_a^b f = \inf\{U(f,P)\} \leqslant M(b-a)$$

■

Remark 3.1.3

If f *is integrable on* $[a,b]$, *we can define a new function* F *on* $[a,b]$ *by*

$$F(x) = \int_a^x f = \int_a^x f(t)dt \tag{3.1.18}$$

Theorem 3.1.9

If f is integrable on $[a,b]$ and F is defined on $[a,b]$ by

$$F(x) = \int_a^x f$$

then F is continuous on $[a,b]$.

Proof.

Suppose $c \in [a,b]$. Since f is integrable on $[a,b]$ it is, by definition, bounded on $[a,b]$; let M be a number such that

$$|f(x)| \leqslant M, \forall x \in [a,b]$$

If $h > 0$ for $c+h \in [a,b]$, then

$$F(c+h) - F(c) = \int_a^{c+h} f - \int_a^c f = \int_c^{c+h} f$$

Since $-M \leqslant f(x) \leqslant M$ for all $x \in [a,b]$ it follows from the Theorem 3.1.8 that

$$-M \cdot h \leqslant \int_c^{c+h} f \leqslant M \cdot h$$

In other words

$$-M \cdot h \leqslant F(c+h) - F(c) \leqslant M \cdot h$$

If $h < 0$, with $c+h \in [a,b]$ we find the inequality

$$-M \cdot h \geqslant F(c+h) - F(c) \geqslant M \cdot h$$

In either case we have that

$$|F(c+h) - F(c)| \leqslant M \cdot |h|$$

Therefore, if $\epsilon > 0$, we have

$$|F(c+h) - F(c)| < \epsilon$$

provided that $|h| < \epsilon/M$. This proves that

$$\lim_{h \to 0} F(c+h) = F(c)$$

so F is continuous at c.

■

3.2 Reimann Sums

Definition 3.2.1

Suppose $P = \{t_0, ..., t_n\}$ is a partition of $[a, b]$, and for each i choose $x_i \in [t_{i-1}, t_i$. Then we clearly have that

$$L(f, P) \leqslant \sum_{i=1}^{n} f(x_i)(t_i - t_{i-1}) \leqslant U(f, P) \tag{3.2.1}$$

Any sum of the form

$$\sum_{i=1}^{n} f(x_i)(t_i - t_{i-1}) \tag{3.2.2}$$

is called a **Reimann sum** *of f for P.*

Theorem 3.2.1

Suppose that f is integrable on $[a, b]$. Then for every $\epsilon > 0$ there is some $\delta > 0$ such that, if $P = \{t_0, ..., t_n\}$ is any partition of $[a, b]$ with $t_i - t_{i-1} < \delta$ for all $i \in \{1, ..., n\}$, then

$$\left| \sum_{i=1}^{n} f(x_i)(t_i - t_{i-1}) - \int_a^b f(x)dx \right| < \epsilon$$

for any Riemann sum formed by choosing $x_i \in [t_{i-1}, t_i]$.

Proof.

Note that for any partition P both the integral and the Riemann sum lie between $L(f, P)$ and $U(f, P)$. Thus, it suffices to show that for any given $\epsilon > 0$, there exists a $\delta > 0$ such that $U(f, P) - L(f, P) < \epsilon$ for any partition P with $t_i - t_{i-1} < \delta$ for all $i \in \{1, 2, ..., n\}$.
Since f is integrable on $[a, b]$ it is also bounded, so there exists $M \in \mathbb{R}$ such that $|f(x)| \leqslant M$ for all $x \in [a, b]$. First choose some particular partition $P^* = \{u_0, ..., u_K\}$ for which

$$U(f, P^*) - L(f, P^*) < \epsilon/2$$

and then choose a δ such that

$$\delta < \frac{\epsilon}{4MK}$$

For any partition P with $t_i - t_{i-1} < \delta$, we can write the sum

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

as two sums. Let the first consist of those i's forwhcih $[t_{i-1}, t_i] \subseteq [u_{j-1}, u_j]$ for some j. Evidently, this sum is $\leqslant U(f, P^*) - L(f, P^*) < \epsilon/2$. For all other i we will have $t_{i-1} < u_j < t_i$ for some $j \in \{1, ..., K - 1\}$, so there are at most $K - 1$ of them. Consequently, the sum for these terms is $< (K - 1) \cdot 2M \cdot \delta M < \epsilon/2$. Then we have that $U(f, P) - L(f, P) < \epsilon$, completing the proof.

■

3.3 The Fundamental Theorem of Calculus

Theorem 3.1: The Fundamental Theorem of Calculus

Let f be integrable on $[a,b]$, and define F on $[a,b]$ by

$$F(x) = \int_a^x f = \int_a^x f(t)dt \tag{3.3.1}$$

If f is continuous at c in $[a,b]$, then F is differentiable at c, and

$$F'(c) = f(c) \tag{3.3.2}$$

(if $c = a$ or b, then $F'(c)$ is understood to mean the right or left hand derivative of F).

Proof.

First, consider $c \in (a,b)$. By definition,

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

Suppose first that $h > 0$. Then

$$F(c+h) - F(c) = \int_c^{c+h} f$$

Define m_h and M_h as follows:

$$m_h = \inf\{f(x) : c \leqslant x \leqslant c+h\}$$
$$M_h = \sup\{f(x) : c \leqslant x \leqslant c+h\}$$

It follows from Theorem 3.1.8 that

$$m_h \cdot h \leqslant \int_c^{c+h} f \leqslant M_h \cdot h$$

Therefore,

$$m_h \leqslant \frac{F(c+h) - F(c)}{h} \leqslant M_h$$

If $h < 0$, let

$$m_h = \inf\{f(x) : c+h \leqslant x \leqslant c\}$$
$$M_h = \sup\{f(x) : c+h \leqslant x \leqslant c\}$$

It follows from Theorem 3.1.8 that

$$m_h \cdot (-h) \leqslant \int_{c+h}^{c} f \leqslant M_h \cdot (-h)$$

Since

$$F(c+h) - F(c) = \int_c^{c+h} f = -\int_{c+h}^{c} f$$

this yields

$$m_h \cdot h \geqslant F(c+h) - F(c) \geqslant M_h \cdot h$$

Since $h < 0$, we have that

$$m_h \leqslant \frac{F(c+h) - F(c)}{h} \leqslant M_h$$

This inequality is true for any integrable function, continuous or not. Since f is continuous at c, however,

$$\lim_{h\to 0} m_h = \lim_{h\to 0} M_h = f(c)$$

and this proves that

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

Now, if $c = a$ we need only look at when $h > 0$, and in this case we still have

$$m_h \leqslant \frac{F(a+h) - F(a)}{h} \leqslant M_h$$

and from our previous limits,

$$\lim_{h\to 0^+} m_h = \lim_{h\to 0^+} m_h = f(a)$$

thus we have that

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

Similarly, if $c = b$ we need only look at $h < 0$, so we have that

$$\lim_{h\to 0^-} m_h = \lim_{h\to 0^-} m_h = f(b)$$

and

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

completing the proof.

■

Remark 3.3.1

We may consider

$$F(x) = \int_a^x f \tag{3.3.3}$$

when $x < a$. In this case we have that

$$F(x) = -\int_x^a f = -\left(\int_b^a f - \int_b^x f\right) \tag{3.3.4}$$

so for $c \in [a,b]$,

$$F'(c) = -(-f(c)) = f(c) \tag{3.3.5}$$

as before.

Corollary 3.3.1

If f is continuous on $[a,b]$ and $f = g'$ for some function g, then

$$\int_a^b f = g(b) - f(a) \tag{3.3.6}$$

Proof.

Let

$$F(x) = \int_a^x f$$

Then $F' = f = g'$ on $[a,b]$. Consequently there is a number c such that

$$F = g + c$$

Note that

$$0 = F(a) = g(a) + c$$

so $c = -g(a)$; thus

$$F(x) = g(x) - g(a)$$

This is true, in particular, for $x = b$. THus

$$\int_a^b f = F(b) = g(b) - g(a)$$

∎

Remark 3.3.2

It is important to note that this is merely a useful result for certain functiosn f, **not** *a definition.*

Theorem 3.2: Second Fundamental Theorem of Calculus

If f is integrable on $[a,b]$ and $f = g'$ for some function g, then

$$\int_a^b f = g(b) - g(a)$$

Proof.

Let $P = \{t_0, ..., t_n\}$ be any partition of $[a,b]$. By The Mean Value Theorem there is a point $x_i \in [t_{i-1}, t_i]$ such that

$$g(t_i) - g(t_{i-1}) = g'(x_i)(t_i - t_{i-1})$$

$$= f(x_i)(t_i - t_{i-1})$$

If

$$m_i = \inf\{f(x) : t_{i-1} \leqslant x \leqslant t_i\}$$
$$M_i = \sup\{f(x) : t_{i-1} \leqslant x \leqslant t_i\}$$

then clearly

$$m_i(t_i - t_{i-1}) \leqslant f(x_i)(t_i - t_{i-1}) \leqslant M_i(t_i - t_{i-1})$$

that is

$$m_i(t_i - t_{i-1}) \leqslant g(t_i) - g(t_{i-1}) \leqslant M_i(t_i - t_{i-1})$$

Adding these equations for $i \in \{1, 2, ..., n\}$ we obtain

$$L(f, P) \leqslant g(b) - g(a) \leqslant U(f, P)$$

for every partition P. This means that

$$g(b) - g(a) = \int_a^b f$$

■

Remark 3.3.3

If f is any bounded function on $[a, b]$, then

$$\sup\{L(f, P)\} \quad \text{and} \quad \inf\{U(f, P)\} \tag{3.3.7}$$

will both exist. These numbers are called the **lower integral** *of f on $[a, b]$ and the* **upper integral** *of f on $[a, b]$, respectively, and will be denoted by*

$$\mathbf{L}\int_a^b f \quad \text{and} \quad \mathbf{U}\int_a^b f \tag{3.3.8}$$

If $a < c < b$, then

$$\mathbf{L}\int_a^b f = \mathbf{L}\int_a^c f + \mathbf{L}\int_c^b f \quad \text{and} \quad \mathbf{U}\int_a^b f = \mathbf{U}\int_a^c f + \mathbf{U}\int_c^b f \tag{3.3.9}$$

and if $m \leqslant f(x) \leqslant M$ for all $x \in [a, b]$, then

$$m(b-a) \leqslant \mathbf{L}\int_a^b f \leqslant \mathbf{U}\int_a^b f \leqslant M(b-a) \tag{3.3.10}$$

f is integrable precisely when

$$\mathbf{L}\int_a^b f = \mathbf{U}\int_a^b f \tag{3.3.11}$$

Proof.

(Left to the reader)

■

Remark 3.3.4

We shall now demonstrate an alternate proof for the following theorem stated previously.

Theorem 3.3.2

If f is continuous on $[a,b]$, then f is integrable on $[a,b]$.

Proof.

Define function L and U on $[a,b]$ by

$$L(x) = \mathbf{L}\int_a^x f \quad \text{and} \quad U(x) = \mathbf{U}\int_a^x f$$

Let $x \in (a,b)$. If $h > 0$ and

$$\begin{aligned} m_h &= \inf\{f(t) : x \leqslant t \leqslant x+h\} \\ M_h &= \sup\{f(t) : x \leqslant t \leqslant x+h\} \end{aligned}$$

then

$$m_h \cdot h \leqslant \mathbf{L}\int_x^{x+h} f \leqslant \mathbf{U}\int_x^{x+h} f \leqslant M_h \cdot h$$

so

$$m_h \cdot h \leqslant L(x+h) - L(x) \leqslant U(x+h) - U(x) \leqslant M_h \cdot h$$

or

$$m_h \leqslant \frac{L(x+h) - L(x)}{h} \leqslant \frac{U(x+h) - U(x)}{h} \leqslant M_h$$

If $h < 0$ and

$$\begin{aligned} m_h &= \inf\{f(t) : x+h \leqslant t \leqslant x\} \\ M_h &= \sup\{f(t) : x+h \leqslant t \leqslant x\} \end{aligned}$$

one obtains the same inequality, precisely as in the proof of The Fundamental Theorem of Calculus.
Since f is continuous at x, we have

$$\lim_{h\to 0} m_h = \lim_{h\to 0} M_h = f(x)$$

and this proves that

$$L'(x) = U'(x) = f(x), \forall x \in (a,b)$$

THis means that there is a number c such that

$$U(x) = L(x) + c, \forall x \in [a,b]$$

Since $U(a) = L(a) = 0$, the number c must be equal to 0, so

$$U(x) = L(x) \forall x \in [a, b]$$

In particular,

$$\mathbf{U}\int_a^b f = U(b) = L(b) = \mathbf{L}\int_a^b f$$

so f is integrable on $[a, b]$.

■

3.A Trigonometric Functions

Definition 3.A.1

We define the mathematical constant π as the area of the unit circle, or in this case, twice the area of a semi-circle:

$$\pi := 2 \cdot \int_{-1}^{1} \sqrt{1 - x^2} dx \tag{3.A.1}$$

Definition 3.A.2

If $-1 \leqslant x \leqslant 1$, then the area of the sector bounded between the upper unit circle from $[x, 1]$ and the x-axis and radial arm is

$$A(x) := \frac{x\sqrt{1 - x^2}}{2} + \int_x^1 \sqrt{1 - t^2} dt \tag{3.A.2}$$

Remark 3.A.1

For $-1 < x < 1$, A is differentiable at x and

$$\begin{aligned} A'(x) &= \frac{1}{2}\left[\sqrt{1 - x^2} + x \cdot \frac{-2x}{2\sqrt{1 - x^2}}\right] - \sqrt{1 - x^2} \\ &= \frac{1}{2}\frac{1 - x^2 - x^2}{\sqrt{1 - x^2}} - \frac{1 - x^2}{\sqrt{1 - x^2}} \\ &= \frac{1}{2}\frac{-1}{\sqrt{1 - x^2}} \\ &= \frac{-1}{2\sqrt{1 - x^2}} \end{aligned}$$

Note that on $[-1, 1]$, the function A decreases from $A(-1) = \frac{\pi}{2}$ to $A(1) = 0$.

Definition 3.A.3

If $0 \leqslant x \leqslant \pi$, then $\cos x$ is the unique number in $[-1,1]$ such that

$$A(\cos x) = \frac{x}{2} \tag{3.A.3}$$

and

$$\sin x := \sqrt{1-(\cos x)^2} \tag{3.A.4}$$

Note that such a $\cos x$ exists as A is continuous on $[-1,1]$, and $A(-1) = \frac{\pi}{2}$ while $A(1) = 0$. Hence, by Intermediate Value Theorem there exists $y \in [-1,1]$ such that $A(y) = \frac{x}{2}$ for all $x \in [0,\pi]$.

Theorem 3.A.1

If $0 < x < \pi$, then

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

Proof.

If $B = 2A$, then the definition $A(\cos x) = x/2$ can be written

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

in other words, $\cos$ is just the inverse of B. We have already computed taht

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

from which we conclude

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

Consequently we have that

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

Then, since $\sin x = \sqrt{1-(\cos x)^2}$ we also obtain

$$\sin'(x) = \frac{1}{2} \cdot \frac{-2\cos x \cdot \cos'(x)}{\sqrt{1-(\cos x)^2}}$$

$$= \frac{-\cos x \cdot (-\sin x)}{\sin x}$$
$$= \cos x$$

■

Definition 3.A.4

Now, to define sin *and* cos *on* $\mathbb{R}$, *we proceed as follows*

1. *If* $\pi \leqslant x \leqslant 2\pi$, *the*

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

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

$$\sin x = \sin x'$$
$$\cos x = \cos x'$$

Lemma 3.A.2

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*

$$[f'(x)]^2 + [f(x)]^2 = 0 \forall x$$

This implies that

$$f(x) = 0 \forall x$$

■

Theorem 3.A.3

If f has a second derivative everywhere and

$$\begin{aligned} f'' + f &= 0 \\ f(0) &= a \\ f'(0) &= b \end{aligned}$$

then

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

Proof.

Let

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

Then

$$\begin{aligned} g'(x) &= f'(x) - b\cos x + a\sin x \\ g''(x) &= f''(x) + b\sin x + a\cos x \end{aligned}$$

Consequently,

$$\begin{aligned} g'' + g &= 0 \\ g(0) &= 0 \\ g'(0) &= 0 \end{aligned}$$

which shows by the previous lemma that

$$0 = g(x) = f(x) - b\sin x - a\cos x, \forall x$$

■

Theorem 3.A.4

If x and y are any two numbers, then

$$\begin{aligned} \sin(x + y) &= \sin x \cos y + \cos x \sin y \\ \cos(x + y) &= \cos x \cos y - \sin x \sin y \end{aligned}$$

Proof.

For any particular $y \in \mathbb{R}$, we can define a function f by

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

Then $f'(x) = \cos(x + y)$ and $f''(x) = -\sin(x + y)$. Consequently,

$$\begin{aligned} f'' + f &= 0 \\ f(0) &= \sin y \\ f'(0) &= \cos y \end{aligned}$$

It follows from the previous theorem that

$$f = (\cos y) \cdot \sin + (\sin y) \cdot \cos$$

that is

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

Since any number y could have been chosen to begin with, this proves the first formula for x and y.
Similarly, for any $y \in \mathbb{R}$ *define* $f(x) = \cos(x+y)$, *so* $f'(x) = -\sin(x+y)$ *and* $f''(x) = -\cos(x+y)$. *Then* $f'' + f = 0$, $f(0) = \cos y$ *and* $f'(0) = -\sin y$. *Then we have that*

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

proving the second formula.

■

Remark 3.A.2

Since

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

it follows from Second Fundamental Theorem of Calculus that

$$\arcsin x = \arcsin x - \arcsin 0 = \int_0^x \frac{1}{\sqrt{1-t^2}} dt$$

Using this definition of arcsin *we could define* sin *as* $\arcsin^{-1}$, *and the formula for the derivative of an inverse function would show that*

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

which could be defined as $\cos x$.

3.B The Logarithm and Exponential Functions

Definition 3.B.1

If $x > 0$, *then define*

$$\log x := \int_1^x \frac{1}{t} dt \tag{3.B.1}$$

Theorem 3.B.1

If $x.y > 0$, *then*

$$\log(xy) = \log x + \log y \tag{3.B.2}$$

Proof.

Notice first that $\log'(x) = 1/x$*, by The Fundamental Theorem of Calculus. Now, choose a number* $y > 0$ *and let*

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

Then

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

Thus, $f' = \log'$*. This means that there is a number* c *such that* $f(x) = \log(x) + c$ *for all* $x > 0$*, that is,*

$$\log(xy) = \log x + c, \ \forall x > 0$$

The number c *can be evaluated by noting that* $\log(1) = 0$*, so* $\log(1 \cdot y) = c$*. Thus*

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

for all x*. Since this is true for all* $y > 0$*, the theorem is proved.*

■

Corollary 3.B.2

If n *is a natural number and* $x > 0$*, then*

$$\log(x^n) = n \log x \tag{3.B.3}$$

Proof.

We proceed by induction on $n \in \mathbb{N}$*. If* $n = 1$ *we simply have* $\log(x^1) = 1 \cdot \log x$*, so the base case holds. Now suppose inductively that there exists* $k \geqslant 1$ *such that if* $n = k$*,*

$$\log(x^k) = k \log x$$

Then, observe that by the previous theorem

$$\begin{aligned} \log(x^{k+1}) &= \log(x^k x) \\ &= \log(x^k) + \log x \\ &= k \log x + \log x && \text{(by the Induction Hypothesis)} \\ &= (k+1) \log x \end{aligned}$$

as desired. Thus by mathematical induction we conclude that for all $n \geqslant 1$*,* $\log(x^n) = n \log x$*.* ■

Corollary 3.B.3

If $x, y > 0$*, then*

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

Proof.

This result follows from the equation

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

■

Definition 3.B.2

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

Theorem 3.B.4

For all numbers x,

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

Proof.

Observe that

$$\begin{aligned}\exp'(x) &= (\log^{-1})'(x) \\ &= \frac{1}{\log'(\log^{-1}(x))} \\ &= \frac{1}{\frac{1}{\log^{-1}(x)}} \\ &= \log^{-1}(x) = \exp(x)\end{aligned}$$

■

Theorem 3.B.5

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* $x = \log x'$ *and* $y = \log y'$. *Then*

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

This means that

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

■

Definition 3.B.3

We define

$$e := \exp(1) \tag{3.B.4}$$

and this is equivalent to the equation

$$1 = \log e = \int_1^e \frac{1}{t}dt \tag{3.B.5}$$

Then, we note that $\exp(x) = [\exp(1)]^x = e^x$ *for rational x, so we define for any* $x \in \mathbb{R}$,

$$e^x = \exp(x) \tag{3.B.6}$$

Definition 3.B.4

If $a > 0$*, then, for any real number x,*

$$a^x := e^{x \log a} \tag{3.B.7}$$

If $a = e$ *this definition agrees with our previous one.*

Theorem 3.B.6

If $a > 0$*, then*

$$(1)\ \ (a^b)^c = a^{bc},\ \forall a, b \in \mathbb{R}$$

and

$$(2)\ \ a^1 = a \text{ and } a^{x+y} = a^x \cdot a^y,\ \forall x, y \in \mathbb{R}$$

Proof.

First, observe that

$$\begin{aligned}(a^b)^c &= e^{c \log a^b} \\ &= e^{c \log e^{b \log a}} \\ &= e^{cb \log a} \\ &= a^{bc}\end{aligned}$$

Next, observe that

$$a^1 = e^{1 \log a} = e^{\log a} = a$$

and

$$\begin{aligned}a^{x+y} &= e^{(x+y) \log a} \\ &= e^{x \log a + y \log a} \\ &= e^{x \log a} \cdot e^{y \log a} \\ &= a^x \cdot a^y\end{aligned}$$

■

Theorem 3.B.7

If f is differentiable and

$$f'(x) = f(x),\ \forall x \in \mathbb{R}$$

then there is a number c such that

$$f(x) = ce^x,\ \forall x \in \mathbb{R}$$

Proof.

Let $g(x) = f(x)/e^x$, which is possible as $e^x \neq 0$ for all x. Then

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

THerefore, there is a number c such that

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

■

Theorem 3.B.8

For any natural number n,

$$\lim_{x\to\infty} \frac{e^x}{x^n} = \infty \tag{3.B.8}$$

Proof.

Step 1. We claim that $e^x > x$ for all x, and consequently $\lim_{x\to\infty} e^x = \infty$.
For $x \leqslant 0$ this is immediate. Now, it suffices to show $x > \log x$ for all $x > 0$. If $x < 1$ this clearly holds since $\log x < 0$. If $x > 1$, then $x - 1$ is an upper sum for $f(t) = \frac{1}{t}$ on $[1, x]$, so $\log x < x - 1 < x$.
Step 2. We claim $\lim_{x\to\infty} \frac{e^x}{x} = \infty$. First, note that

$$\frac{e^x}{x} = \frac{e^{x/2} \cdot e^{x/2}}{\frac{x}{2} \cdot 2} = \frac{1}{2}\left(\frac{e^{x/2}}{\frac{x}{2}}\right) \cdot e^{x/2}$$

By Step 1. the expression in parentheses is greater than 1, and $\lim_{x\to\infty} e^{x/2} = \infty$; this shows that $\lim_{x\to\infty} e^x/x = \infty$.
Step 3. To prove the main claim note that

$$\frac{e^x}{x^n} = \frac{(e^{x/n})^x}{\left(\frac{x}{n}\right)^n \cdot n^n} = \frac{1}{n^n} \cdot \left(\frac{e^{x/n}}{\frac{x}{n}}\right)^n$$

The expression in parentheses becomes arbitrarily large, by Step 2., so the nth power certainly becomes arbitrarily large.

■

Chapter 4

Sequences and Series

4.1 Approximation by Polynomial Functions

Definition 4.1.1

Given a function f that is n times differentiable in a neighborhood of a point a, let

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

and define

$$P_{n,a}(x) = a_0 + a_1(x-a) + \ldots + a_x(x-a)^n$$

The polynomial $P_{n,a}$ is called the **Taylor polynomial of degree n for f at a.**

Remark 4.1.1

The Taylor polynomial has been defined so that

$$P_{n,a}^{(k)}(a) = f^{(k)}(a) \quad \textit{for} \;\; 0 \leqslant k \leqslant n$$

in fact, it is evidently the only polynomial of degree $\leqslant n$ with this property.

Example

Consider the sin *function. We have*

$$\begin{aligned}
\sin(0) &= 0 \\
\sin'(0) &= \cos 0 = 1 \\
\sin''(0) &= -\sin 0 = 0 \\
\sin'''(0) &= -\cos 0 = -1 \\
\sin^{(4)}(0) &= \sin 0 = 0
\end{aligned}$$

From this point on, the derivatives repeat modulo 4. The coefficients become

$$a_k = \frac{\sin^{(k)}(0)}{k!} = \begin{cases} 0 & \textit{if } \exists l \in \mathbb{N}; k = 2l \\ \frac{(-1)^l}{(2l+1)!} & \textit{if } \exists l \in \mathbb{N}; k = 2l+1 \end{cases}$$

Therefore, the Taylor polynomial $P_{2n+1,0}$ of degree $2n+1$ for sin *at* 0 *is*

$$P_{2n+1,0}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

Via a similar derivation we find that the Taylor polynomial $P_{2n,0}$ of degree $2n$ for cos *at* 0 *is*

$$P_{2n,0} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + (-1)^n \frac{x^{2n}}{(2n)!}$$

Example

Note that for all $k \geqslant 0$, $\exp^{(k)}(0) = \exp(0) = 1$, so the Taylor polynomial of degree n for exp *at* 0 *is*

$$P_{n,0}(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!}$$

For log*, observe that*

$$\begin{aligned} \log'(x) &= \frac{1}{x}, & \log'(1) &= 1 \\ \log''(x) &= -\frac{1}{x^2}, & \log''(1) &= -1 \\ \log'''(x) &= \frac{2}{x^3}, & \log'''(1) &= 2 \end{aligned}$$

in general

$$\log^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{x^k}, \quad \log^{(k)}(1) = (-1)^{k-1}(k-1)!$$

for $k \geqslant 1$, and $\log(1) = 0$. Therefore, the Taylor polynomial of degree n for log *at* 1 *is*

$$P_{n,1}(x) = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \ldots + \frac{(-1)^{n-1}(x-1)^n}{n}$$

If we consider the function $f(x) = \log(1+x)$, then the Taylor polynomial of degree n of f at 0 *is*

$$P_{n,0}(x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + \frac{(-1)^{n-1}x^n}{n}$$

Theorem 4.1.1

Suppose that f is a function and $a \in \mathbb{R}$ such that

$$f'(a), ..., f^{(n)}(a)$$

all exist. Let

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

and define

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

Then

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

Proof.

Weiting out $P_{n,a}(x)$ expliticly we obtain

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

Let us introduce the new functions

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

now we must prove that

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

Note that $Q^{(k)}(a) = f^{(k)}(a)$ for $k \leqslant n-1$, and $g^{(k)}(x) = n! \frac{(x-a)^{n-k}}{(n-k)!}$. By the continuity of $f^{(k)}$ and $Q^{(k)}$ for $k \leqslant n-1$, we have that

$$\lim_{x \to a} \left[f^{(k)}(x) - Q^{(k)}(x) \right] = f^{(k)}(a) - Q^{(k)}(a) = 0$$

and

$$\lim_{x \to a} g^{(k)}(x) = 0$$

Then applying l'Hopital's Rule $n-1$ times, we obtain

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

But the $n-1$st derivative of Q is constant, and in fact $Q^{(n-1)}(x) = f^{(n-1)}(a)$, so

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

by definition of $f^{(n)}(a)$, as desired.

■

Theorem 4.1.2

Suppose that

$$f'(a) = ... = f^{(n-1)}(a) = 0, \quad and \quad f^{(n)}(a) \neq 0$$

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 a.*

Proof.

Let f be as in the hypothesis, and without loss of generality let $f(a) = 0$, as otherwise one can replace f with $f - f(a)$ without affecting the hypothesis. THen, since the first $n-1$ derivatives of f at a are 0, the Taylor polynomial$P_{n,a}$ of f is

$$\begin{aligned} P_{n,a}(x) &= \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i \\ &= \frac{f^{(n)}(a)}{n!}(x-a)^n \end{aligned}$$

Thus, Theorem 4.1.1 states that

$$0 = \lim_{x \to a} \frac{f(x) - P_{n,a}(x)}{(x-a)^n} = \lim_{x \to a} \left[\frac{f(x)}{(x-a)^n} - \frac{f^{(n)}(a)}{n!} \right]$$

Consequently, if x is sufficiently close to a, then $f(x)/(x-a)^n$ has the same sign as $f^{(n)}(a)/n!$. Suppose now that n is even. In this case $(x-a)^n > 0$ for all $x \neq a$. Since $f(x)/(x-a)^n$ has the same sign as $f^{(n)}(a)/n!$ for x sufficeintly close to a, it follows that $f(x)$ itself has the same sign as $f^{(n)}(a)/n!$ for x sufficiently close to a. If $f^{(n)}(a) > 0$, this means that

$$f(x) > 0 = f(a)$$

for x close to a. Consequently, f has a local minimum at a. If $f^{(n)}(a) < 0$, this means that

$$f(x) < 0 = f(a)$$

for x close to a, so f has a local minimum at a.
Conversely, suppose that n is odd. Then if x is sufficiently close to a $f(x)/(x-a)^n$ always has the same sign, since it has the same sign as $f^{(n)}(a)/n!$ which is constant. 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 $x < a$. Hence, f has neither a local maximum nor a local minimum at a.

■

Definition 4.1.2

Two functions f and g are **equal up to order** n **at** a *if*

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

Theorem 4.1.3

Let P and Q be two polynomials in $(x-a)$, of degree $\leqslant n$, and suppose that P and Q are equal up to order n at a. Then $P = Q$.

Proof.

Let $R = P - Q$. Since R is a polynomial of degree $\leqslant n$, it is only necessariy to prove that if $R(x) = b_0 + \ldots + b_n(x-a)^n$ satisfies

$$\lim_{x \to a} \frac{R(x)}{(x-a)^n} = 0$$

then $R = 0$. Now the hypothesis on R surely implies that

$$\lim_{x \to a} \frac{R(x)}{(x-a)^i} = 0 \quad \text{for } 0 \leqslant i \leqslant n$$

For $i = 0$ this condition reads that $\lim_{x \to a} R(x) = 0$. On the other hand $\lim_{x \to a} R(x) = b_0$. Thus $b_0 = 0$. Similarly, we find that

$$\frac{R(x)}{x-a} = b_1 + b_2(x-a) + \ldots + b_n(x-a)^{n-1}$$

and

$$\lim_{x \to a} \frac{R(x)}{x-a} = b_1$$

so $b_1 = 0$. Continuing in this way, by induction we find that

$$b_0 = b_1 = \ldots = b_n = 0$$

so $R(x) = 0$ and $P = Q$.

■

Corollary 4.1.4

Let f be n-times differentiable at a, and suppose that P is a polynomial in $(x-a)$ of degree $\leqslant n$, which equals f up to order n at a. Then $P = P_{n,a,f}$ (The Taylor polynomial of f of degree n at a).

Proof.

Since P and $P_{n,a,f}$ both equal f up to order n at a, using the triangle inequality and an epsilon-delta proof it can be shown that P equals $P_{n,a,f}$ up to order n at a. Consequently, $P = P_{n,a,f}$ by *the preceding Theorem.*

■

Definition 4.1.3

If f is a function for which $P_{n,a}(x)$ exists, we define the **remainder term** *$R_{n,a}(x)$ by*

$$R_{n,a}(x) = f(x) - P_{n,a}(x)$$

If $f^{(n+a)}$ is continuous on $[a, x]$, then

$$R_{n,a}(x) = \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n dt$$

Remark 4.1.2

If m and M are the minimum and maximum of $f^{(n+1)}n!$ on $[a, x]$, then $R_{n,a}(x)$ satisfies

$$m \in_a^x (x-t)^n dt \leqslant R_{n,a}(x) \leqslant M \int_a^x (x-t)^n dt$$

so we can write

$$R_{n,a}(x) = \alpha \cdot \frac{(x-a)^{n+1}}{n+1}$$

for some number α between m and M. Since we have assumed that $f^{(n+1)}$ is continuous, for some $t \in (a, x)$ we can also write

$$R_{n,a}(x) = \frac{f^{(n+1)}(t)}{n!}\frac{(x-a)^{n+1}}{n+1} = \frac{f^{(n+1)}(t)}{(n+1)!}(x-a)^{n+1}$$

This is known as the **Lagrange form of the remainder**.

Lemma 4.1.5

Suppose that the function R is $(n+1)$-times differentiable on $[a, b]$, and

$$R^{(k)}(a) = 0 \quad \text{for } k = 0, 1, 2, ..., n$$

Then for any $x \in (a, b]$ we have

$$\frac{R(x)}{(x-a)^{n+1}} = \frac{R^{(n+1)}(t)}{(n+1)!} \quad \text{for some } t \text{ in } (a, x)$$

Proof.

For $n = 0$, this is simply The Mean Value Theorem, and we will proceed for the remaining n by induction on n. To do this we use The Cauchy Mean Value Theorem to write

$$\frac{R(x)}{(x-a)^{n+2}} = \frac{R'(z)}{(n+2)(z-a)^{n+1}} = \frac{1}{n+2}\frac{R'(z)}{(z-a)^{n+1}} \quad \text{for some } z \text{ in } (a, x),$$

and then apply the induction hypothesis to R' on the interval $[a,z]$ to get

$$\begin{aligned} \frac{R(x)}{(x-a)^{n+2}} &= \frac{1}{n+2}\frac{(R')^{(n+1)}(t)}{(n+1)!} \quad \textit{for some } z \textit{ in } (a,x), \\ &= \frac{R^{(n+2)}(t)}{(n+2)!} \end{aligned}$$

as desired.

■

Theorem 4.1: Taylor's Theorem

Suppose $f', \ldots, f^{(n+1)}$ are defined on $[a,x]$, and that $R_{n,a}(x)$ is defined by

$$R_{n,a}(x) = f(x) - P_{n,a}(x)$$

Then

$$R_{n,a}(x) = \frac{f^{(n+1)}(t)}{(n+1)!}(x-a)^{n+1} \quad \textit{for some } t \textit{ in } (a,x)$$

Proof.

The function $R_{n,a}$ satisfies the conditions of the preceding Lemma by the definition of the Taylor polynomial, so

$$\frac{R_{n,a}(x)}{(x-a)^{n+1}} = \frac{R_{n,a}^{(n+1)}(t)}{(n+1)!}$$

for some t in (a,x). But,

$$R_{n,a}^{(n+1)} = f^{(n+1)}$$

since $R_{n,a} - f$ is a polynomial of degree n. Substituting gives the Lagrange form for the remainder, as desired.

■

Example

Applying Taylor's Theorem to the functions sin, cos, *and* exp, *with $a = 0$, we obtain the following formulas:*

$$\begin{aligned} \sin x &= \sum_{i=0}^{n} \frac{(-1)^i x^{2i+1}}{(2i+1)!} + \frac{\sin^{(2n+2)}(t)}{(2n+2)!}x^{2n+2} \\ \cos x &= \sum_{i=0}^{n} \frac{(-1)^i x^{2i}}{(2i)!} + \frac{\cos^{(2n+1)}(t)}{(2n+1)!}x^{2n+1} \\ \exp x &= \sum_{i=0}^{n} \frac{x^n}{n!} + \frac{\exp t}{(n+1)!}x^{n+1} \end{aligned}$$

(of course, we could go one power higher in the remainder terms for sin *and* cos*)*

4.2 Infinite Sequences

Definition 4.2.1

An **infinite sequence** *of real numbers is a real valued function whose domain is* $\mathbb{N}$.

Definition 4.2.2

A sequence $\{a_n\}$ **converges to** $\ell \in \mathbb{R}$, *denoted by* $\lim_{n\to\infty} a_n = \ell$, *if for every* $\epsilon > 0$ *there exists* $N \in \mathbb{N}$ *such that for all* $n \in \mathbb{N}$, *if* $n \geqslant N$ *then*

$$|a_n - \ell| < \epsilon$$

A sequence $\{a_n\}$ *is said to* **converge** *if such an* ℓ *exists, and to* **diverge** *otherwise.*

Theorem 4.2.1: Limit Laws

If $\lim_{n\to\infty} a_n$ *and* $\lim_{n\to\infty} b_n$ *both exist, then*

$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$$
$$\lim_{n\to\infty} a_n \cdot b_n = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n$$

moreover, if $\lim_{n\to\infty} b_n \neq 0$, *then* $b_n \neq 0$ *for all* n *greater than some* N, *and*

$$\lim_{n\to\infty} a_n/b_n = \lim_{n\to\infty} a_n / \lim_{n\to\infty} b_n$$

Proof.

(To be completed)

■

Theorem 4.2.2

Let f *be a function defined in an open interval containing* c, *except perhaps at* c *itself, with*

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

Suppose that $\{a_n\}$ *is a sequence such that*

1. *each* a_n *is in the domain of* f,
2. *each* $a_n \neq c$,
3. $\lim_{n\to\infty} a_n = c$

Then the sequence $\{f(a_n)\}$ *satisfies*

$$\lim_{n\to\infty} f(a_n) = l$$

Conversely, if this is true for every sequence $\{a_n\}$ *satisfying the above conditions, then* $\lim_{x\to c} f(x) = l$.

Proof.

Suppose first that $\lim_{x \to c} f(x) = l$. *Then for every* $\varepsilon > 0$ *there is a* $\delta > 0$ *such that, for all* x,

$$\text{if } 0 < |x - c| < \delta, \text{ then } |f(x) - l| < \varepsilon$$

If the sequence $\{a_n\}$ *satisfies* $\lim_{n \to \infty} a_n = c$, *then there is a natural number* N *such that for all* $n \in \mathbb{N}$,

$$\text{if } n \geqslant N, \text{ then } 0 < |a_n - c| < \delta$$

By our choice of δ *this implies that*

$$|f(a_n) - l| < \varepsilon$$

showing that $\lim_{n \to \infty} f(a_n) = l$.

Suppose, conversely, that $\lim_{n \to \infty} f(a_n) = l$ *for every sequence* $\{a_n\}$ *satisfying our three conditions. If* $\lim_{x \to c} f(x) = l$ *were not true, there would be some* $\varepsilon > 0$ *such that for all* $\delta > 0$ *there exists* x *such that* $0 < |x - c| < \delta$ *but* $|f(x) - l| \geqslant \varepsilon$. *In particular, for each* n *there would exist* x_n *such that* $0 < |x_n - c| < 1/n$, *but* $|f(x_n) - l| \geqslant \varepsilon$. *Define a sequence* $\{x_n\}$ *using these* x_n. *Then* x_n *is in the domain of* f *for each* n, *and as* $0 < |x_n - c|$ *for each* n, $x_n \neq c$ *for all* n. *Moreover, for all* $\varepsilon' > 0$ *there exists* $N \in \mathbb{N}$ *such that* $\varepsilon' > 1/N > 0$ *(by the Archimedean Property of* $\mathbb{R}$*), so for all* $n \geqslant N$, $0 < |x_n - c| < 1/n < \varepsilon'$. *Consequently,* $\lim_{n \to \infty} x_n = c$, *so the sequences satisfies all of our initial conditions. However, then by assumption* $\lim_{n \to \infty} f(x_n) = l$. *But, by construction, for* ε *we have that for all* $n \in \mathbb{N}$, $0 < |x_n - c| < 1/n$ *but* $|f(x_n) - l| \geqslant \varepsilon$, *so* $f(x_n)$ *does not converge to* l, *contradicting our hypothesis. Thus* $\lim_{x \to c} f(x) = l$ *must be true.*

■

Definition 4.2.3

A sequence $\{a_n\}$ *is* **increasing** *if* $a_{n+1} > a_n$ *for all* n, **nondecreasing** *if* $a_{n+1} \geqslant a_n$ *for all* n, *and* **bounded above** *if there is a number* M *such that* $a_n \leqslant M$ *for all* n. *Similarly, a sequence* $\{a_n\}$ *is* **decreasing** *if* $a_{n+1} < a_n$ *for all* n, **nonincreasing** *if* $a_{n+1} \leqslant a_n$ *for all* n, *and* **bounded below** *if there is a number* m *such that* $a_n \geqslant m$ *for all* n.

Theorem 4.2.3

If $\{a_n\}$ *is nondecreasing and bounded above, then* $\{a_n\}$ *converges.*

Proof.

The set $A := \{a_n : n \in \mathbb{N}\}$ *is, by assumption, bounded above, so* A *has a least upper bound* $\alpha \in \mathbb{R}$. *We claim that* $\lim_{n \to \infty} a_n = \alpha$. *If* $\varepsilon > 0$, *then* $\alpha - \varepsilon$ *is not an upper bound for* A *so there exists* a_N *in* A *such that* $a_N > \alpha - \varepsilon$, *so* $\alpha - a_N < \varepsilon$. *Then for all* $n \geqslant N$, $a_n \geqslant a_N$ *since* $\{a_n\}$ *is nondecreasing so*

$$|\alpha - a_n| = \alpha - a_n \leqslant \alpha - a_N < \varepsilon$$

Consequently, we conclude that $\lim_{n \to \infty} a_n = \alpha$.

■

Definition 4.2.4

*A **subsequence** of a sequence $\{a_n\}$ is a sequence*

$$a_{n_1}, a_{n_2}, a_{n_3}, \ldots$$

where the n_j are natural numbers with $n_1 < n_2 < n_3 < \ldots$.

Lemma 4.2.4

Any sequence $\{a_n\}$ contains a subsequence which is either nondecreasing or nonincreasing.

Proof.

Call a natural number n a "peak point" of a sequence $\{a_n\}$ if $a_m < a_n$ for all $m > n$.

Case 1. The sequence has infinitely many peak points. In this case, if $n_1 < n_2 < n_3 < \ldots$ are the peak points, then $a_{n_1} > a_{n_2} > a_{n_3} > \ldots$, so $\{a_{n_j}\}$ is a nonincreasing subsequence of $\{a_n\}$.

Case 2. THe sequence has only finitely many peak points. In this case, let n_1 be greater than all peak points. Since n_1 is not a peak point, there is some $n_2 > n_1$ such that $a_{n_2} \geqslant a_{n_1}$. Since n_2 is not a peak point, there is some $n_3 > n_2$ such that $a_{n_3} > a_{n_2}$. Suppose there exists $k \geqslant 3$ such that for all $1 \leqslant m < k$, $a_{n_m} \leqslant a_{n_{m+1}}$ and $n_m < n_{m+1}$. Then since n_k is not a peak point, there is some n_{k+1} such that $n_{k+1} > n_k$ and $a_{n_{k+1}} \geqslant a_{n_k}$. Thus, by recursive definition we have constructed a nondecreasing subsequence $\{a_{n_k}\}$ of $\{a_n\}$.

■

Corollary 4.2.5: The Bolzano-Weierstrass Theorem

Every bounded sequence has a convergent subsequence.

Definition 4.2.5

*A sequence $\{a_n\}$ is a **Cauchy sequence** if for every $\varepsilon > 0$ there is a natural number N such that, for all $m, n \in \mathbb{N}$, if $m, n \geqslant N$, then*

$$|a_n - a_m| < \varepsilon$$

(This can be written as $\lim_{m,n\to\infty} |a_m - a_n| = 0$)

Theorem 4.2.6

A sequence $\{a_n\}$ converges if and only if it is a Cauchy sequence.

Proof.

First assume that $\lim_{n\to\infty} a_n = l$ for some $l \in \mathbb{R}$. Then given $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that if $n \geqslant N$, then $|a_n - l| < \varepsilon/2$. Hence, if $m, n \geqslant N$, then

$$|a_n - a_m| \leqslant |a_n - l| + |a_m - l| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

Thus, $\{a_n\}$ is Cauchy.
*Conversely, suppose that $\{a_n\}$ is a Cauchy sequence. I claim that this implies $\{a_n\}$ is bounded. First, for $\varepsilon = 1 > 0$, there exists $N \in \mathbb{N}$ such that if $m, n \geqslant N$, then $|a_m - a_n| < \varepsilon = 1$. In particular, for all $n \geqslant N$, $|a_N - a_n| < 1$. Take $M = \max(a_N + 1, a_{N-1}, ..., a_1)$, and $m = \min(a_N - 1, a_{N-1}, ..., a_1)$. Then for all $k \leqslant N$ we have that $m \leqslant a_k \leqslant M$. On the other hand, for $k \geqslant N$, we have that $a_N - 1 < a_k < a_N + 1$, so $m < a_k < M$. Thus $\{a_n : n \in \mathbb{N}\}$ is bounded. Then, by **??** $\{a_n\}$ has a convergent subsequence $\{a_{n_k}\}$. Let $\lim_{k\to\infty} a_{n_k} = l$, for some $l \in \mathbb{R}$. Then, fix $\varepsilon > 0$. It follows that there exist $K, K' \in \mathbb{N}$ such that for $m, n \geqslant K$ and $j \geqslant K'$, $|a_m - a_n| < \varepsilon/2$, while $|a_{n_j} - l| < \varepsilon/2$. Note that $n_j \geqslant j$, since the sequence $\{n_k\}$ is increasing. Then, for all $i \geqslant \max(K, K')$ we have that*

$$|a_i - l| \leqslant |a_i - a_{n_i}| + |a_{n_i} - l| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

Therefore, by definition $\{a_n\}$ converges to l as well.

■

4.3 Infinite Series

Definition 4.3.1

The sequence $\{a_n\}$ is **summable** *if the sequence $\{s_n\}$ converges, where*

$$s_n = \sum_{i=1}^{n} a_i$$

is the n-th **partial sum**. *In this case, $\lim_{n\to\infty} s_n$ is denoted by*

$$\sum_{n=1}^{\infty} a_n$$

and is called the **sum** *of the sequence $\{a_n\}$.*

Remark 4.3.1

If $\{a_n\}$ and $\{b_n\}$ are summable, then

$$\sum_{n=1}^{\infty}(a_n + b_n) = \sum_{i=1}^{\infty} a_n + \sum_{i=1}^{\infty} b_n$$
$$\sum_{n=1}^{\infty} c \cdot a_n = c \cdot \sum_{i=1}^{\infty} a_n$$

for all $c \in \mathbb{R}$.

Theorem 4.2: The Cauchy Criterion

The sequence $\{a_n\}$ is summable if and only if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $m \geqslant n \in \mathbb{N}$, if $m, n \geqslant N$, then

$$\left|\sum_{i=1}^{m} a_i - \sum_{i=1}^{n} a_i\right| = \left|\sum_{i=n+1}^{m} a_i\right| < \varepsilon$$

Remark 4.3.2

This result is a direct conseqeunce of the fact that a sequence in $\mathbb{R}$ is Cauchy if and only if it converges applied to the sequence of partial sums for $\{a_n\}$.

Theorem 4.3: The Vanishing Condition

If $\{a_n\}$ is summable, then

$$\lim_{n\to\infty} a_n = 0$$

Proof.

If $\lim_{n\to\infty} s_n = l$, then

$$\begin{aligned}\lim_{n\to\infty} a_n &= \lim_{n\to\infty}(s_n - s_{n-1}) = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} \\ &= l - l = 0\end{aligned}$$

■

Example

The **geometric series** *are of the form*

$$\sum_{n=0}^{\infty} r^n = 1 + r + r^2 + r^3 + \dots$$

For $|r| \geqslant 1$, $\lim_{n\to\infty} r^n \neq 0$, so $\{r^n\}$ is not summable. On the other hand, if $|r| < 1$ the sequence is summable. First write

$$s_n = 1 + r + r^2 + ... + r^n, \text{ and } rs_n = r + r^2 + r^3 + ... + r^{n+1}$$

It follows that

$$s_n(1 - r) = 1 - r^{n+1}$$

so

$$s_n = \frac{1 - r^{n+1}}{1 - r}$$

since $r \neq 1$. Finally, it follows that

$$\sum_{n=0}^{\infty} r^n = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{1 - r^{n+1}}{1 - r} = \frac{1}{1 - r}$$

as $|r| < 1$.

Definition 4.3.2

A sequence $\{a_n\}$ *such that* $a_n \geqslant 0$ *for all* $n \in \mathbb{N}$ *is said to be* **nonnegative**.

Theorem 4.4: The Boundedness Criterion

A nonnegative sequence $\{a_n\}$ *is summable if and only if the set of partial sums* s_n *is bounded.*

Proof.

Since $\{a_n\}$ *is nonnegative,* $\{s_n\}$ *is nondecreasing. Hence from our previous results on monotone sequences,* $\{s_n\}$ *converges if and only if* $\{s_n\}$ *is bounded.*

■

Theorem 4.5: The Comparison Test

Suppose that $\{a_n\}$ *and* $\{b_n\}$ *are sequences such that* $0 \leqslant a_n \leqslant b_n$ *for all* $n \in \mathbb{N}$*. Then if* $\sum_{n=1}^{\infty} b_n$ *converges, so does* $\sum_{n=1}^{\infty} a_n$*.*

Proof.

Let s_n *denote the n-th partial sum of* $\{a_n\}$*, and let* t_n *denote the n-th partial sum of* $\{b_n\}$*. Then* $0 \leqslant s_n \leqslant t_n$ *for all* $n \in \mathbb{N}$*. Now* $\{t_n\}$ *converges by assumption, so it is bounded. Hence, there exists* $M \in \mathbb{R}$ *such that* $0 \leqslant s_n \leqslant t_n \leqslant M$ *for all* $n \in \mathbb{N}$*, so* $\{s_n\}$ *is also bounded. Thus by The Boundedness Criterion* $\{a_n\}$ *is summable, so by definition* $\sum_{n=1}^{\infty} a_n$ *converges.*

■

Theorem 4.6: The Limit Comparison Test

If $a_n, b_n > 0$ *for convergent sequences* $\{a_n\}$ *and* $\{b_n\}$*, and* $\lim_{n\to\infty} a_n/b_n = c \neq 0$*, then* $\sum_{n=1}^{\infty} a_n$ *converges if and only if* $\sum_{n=1}^{\infty} b_n$ *converges.*

Proof.

Suppose $\sum_{n=1}^{\infty} b_n$ *converges. Since* $\lim_{n\to\infty} a_n/b_n = c$*, there is some* N *such that*

$$a_n/b_n - c \leqslant c \implies a_n \leqslant 2cb_n, \quad \text{for } n \geqslant N$$

But the sequence $2c \sum_{n=N}^{\infty} b_n$ *certainly converges. Then by The Comparison Test we have that*

$\sum_{n=N}^{\infty} a_n$ *converges, and this implies convergence of the whole series* $\sum_{n=1}^{\infty} a_n$*, which only has finitely many additional terms.*
Note that

$$\lim_{n\to\infty} b_n/a_n = \frac{1}{\lim_{n\to\infty} a_n/b_n} = 1/c \neq 0$$

so the converse follows immediately.

■

Theorem 4.7: The Ratio Test

Let $\{a_n\}$ *be a positive sequence, and suppose that*

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = r$$

for some $r \geqslant 0$*. Then* $\sum_{n=1}^{\infty} a_n$ *converges if* $r < 1$*. On the other hand, if* $r > 1$*, then the terms* a_n *are unbounded, so* $\sum_{n=1}^{\infty} a_n$ *diverges.*

Proof.

Suppose first that $r < 1$*. Choose any number* s *with* $r < s < 1$*. The hypothesis* $\lim_{n\to\infty} a_{n+1}/a_n = r < 1$ *implies that there is some* $N \in \mathbb{N}$ *such that if* $n \geqslant N$*,*

$$a_{n+1}/a_n - r < s - r \implies a_{n+1}/a_n < s$$

This can be written as $a_{n+1} < s a_n$*. Thus,*

$$\begin{aligned} a_{N+1} &< s a_N \\ a_{N+2} &< s a_{N+1} < s^2 a_N \\ &\vdots \\ a_{N+k} &< s^k a_N \end{aligned}$$

Since $\sum_{k=0}^{\infty} a_N s^k = a_N \sum_{k=0}^{\infty} s^k$ *converges, since* $|s| < 1$*, The Comparison Test shows that*

$$\sum_{n=N}^{\infty} a_n = \sum_{k=0}^{\infty} a_{N+k}$$

converges as $a_{N+k} < a_N s^k$ *for all* $k \geqslant 0$*. This implies that* $\sum_{n=0}^{\infty} a_n$ *as a whole converges.*
If $r > 1$*, choose some* $s \in \mathbb{R}$ *such that* $1 < s < r$*. Then there is a number* $N \in \mathbb{N}$ *such that*

$$r - a_{n+1}/a_n < r - s \implies s < a_{n+1}/a_n$$

for all $n \geqslant N$. This implies that

$$a_{N+k} > a_N s^k,$$

for all $k \in \mathbb{N}$, so the terms are unbounded.

■

Theorem 4.8: The Integral Test

Suppose that f is positive and decreasing on $[1, \infty)$, and that $f(n) = a_n$ for all $n \in \mathbb{N}$. Then $\sum_{n=1}^{\infty} a_n$ converges if and only if the limit

$$\int_1^{\infty} f = \lim_{A \to \infty} \int_1^A f$$

exists.

Proof.

The existence of $\lim_{A \to \infty} \int_1^A f$ is equivalent to the convergence of the series

$$\int_1^2 f + \int_2^3 f + \int_3^4 f + \ldots$$

Since f is decreasing, we have

$$f(n+1) < \int_n^{n+1} f < f(n)$$

The first half of this double inequality shows that the series $\sum_{n=1}^{\infty} a_{n+1}$ may be compared to the series $\sum_{n=1}^{\infty} \int_n^{n+1}$, proving that $\sum_{n=1}^{\infty} a_{n+1}$ (and hence $\sum_{n=1}^{\infty} a_n$) converges if $\lim_{A \to \infty} \int_1^A f$ exists.
The second half of the inequality shows that the series $\sum_{n=1}^{\infty} \int_n^{n+1} f$ may be compared to the series $\sum_{n=1}^{\infty} a_n$, proving that $\lim_{A \to \infty} \int_1^A f$ must exist if $\sum_{n=1}^{\infty} a_n$ converges.

■

Corollary 4.3.1

The series $\sum_{n=1}^{\infty} 1/n^p$ converges if and only if $p > 1$.

Proof.

If $p < 0$ then $\lim_{n \to \infty} 1/n^p \neq 0$. If $p > 0$, the convergence of $\sum_{n=1}^{\infty} 1/n^p$ is equivalent, by The

Integral Test, to the existence of

$$\lim_{A\to\infty}\int_1^A \frac{1}{x^p}dx$$

Now, observe that

$$\int_1^A \frac{1}{x^p}dx = \begin{cases} -\frac{1}{p-1}\cdot\frac{1}{A^{p-1}}+\frac{1}{p-1}, & p\neq 1 \\ \log A, & p=1 \end{cases}$$

This shows that $\lim_{A\to\infty}\int_1^A 1/x^p dx$ *exists if* $p>1$*, but not if* $p\leqslant 1$*. Thus,* $\sum_{n=1}^{\infty} 1/n^p$ *converges precisely for* $p>1$.

■

Definition 4.3.3

The series $\sum_{n=1}^{\infty} a_n$ *is* **absolutely convergent** *if the series* $\sum_{n=1}^{\infty}|a_n|$ *is convergent. (In formal language, the sequence* $\{a_n\}$ *is said to be* **absolutely summable** *if the sequence* $\{|a_n|\}$ *is summable.)*

Theorem 4.3.2

Every absolutely convergent series is convergent. Moreover, a series is absolutely convergent if and only if the series formed from its positive terms and the series formed from its negative terms both converge.

Proof.

If $\sum_{n=1}^{\infty}|a_n|$ *converges, then, by The Cauchy Criterion,*

$$\lim_{m,n\to\infty}\sum_{i=n+1}^{m}|a_i| = 0$$

Since

$$\left|\sum_{i=n+1}^{m} a_i\right| \leqslant \sum_{i=n+1}^{m}|a_i|$$

it follows that

$$\lim_{m,n\to\infty}\sum_{i=n+1}^{m} a_i = 0$$

which shows that $\sum_{n=1}^{\infty} a_n$ *converges.*

To prove the second portion of the theorem, let $\{a_n^+\}$ *and* $\{a_n^-\}$ *be sequences defined by*

$$a_n^+ = \begin{cases} a_n, & \text{if } a_n\geqslant 0 \\ 0, & \text{if } a_n\leqslant 0 \end{cases}$$

$$a_n^- = \begin{cases} a_n, & \text{if } a_n\leqslant 0 \\ 0, & \text{if } a_n\geqslant 0 \end{cases}$$

It follows that

$$\sum_{n=1}^{\infty} |a_n| = \sum_{n=1}^{\infty} [a_n^+ - a_n^-] = \sum_{n=1}^{\infty} a_n^+ - \sum_{n=1}^{\infty} a_n^-$$

so if $\sum_{n=1}^{\infty} a_n^+$ *and* $\sum_{n=1}^{\infty} a_n^-$ *are convergent, so is* $\sum_{n=1}^{\infty} |a_n|$.

Conversely, suppose $\sum_{n=1}^{\infty} |a_n|$ *converges. Then by our initial argument* $\sum_{n=1}^{\infty} a_n$ *converges also. Therefore,*

$$\sum_{n=1}^{\infty} a_n^+ = \sum_{n=1}^{\infty} \frac{1}{2}[a_n + |a_n|] = \frac{1}{2}\left(\sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} |a_n|\right)$$

and

$$\sum_{n=1}^{\infty} a_n^- = \sum_{n=1}^{\infty} \frac{1}{2}[a_n - |a_n|] = \frac{1}{2}\left(\sum_{n=1}^{\infty} a_n - \sum_{n=1}^{\infty} |a_n|\right)$$

both converge.

■

Remark 4.3.3

A consequent of this result is that every convergent series with positive terms can be used to obtain infinitely many other convergent series simply by putting minus sings at random.

Definition 4.3.4

A convergent series which is not absolutely convergent is said to be **conditionally convergent**.

Theorem 4.9: Leibniz's Theorem

Suppose that $\{a_n\}$ *is a non-decreasing non-negative sequence such that*

$$\lim_{n\to\infty} a_n = 0$$

Then the series

$$\sum_{n=1}^{\infty} (-1)^{n+1} a_n = a_1 - a_2 + a_3 - a_4 + \ldots$$

converges.

Proof.

First, observe that

1. $s_2 \leqslant s_4 \leqslant s_6 \leqslant \ldots$
2. $s_1 \geqslant s_3 \geqslant s_5 \geqslant \ldots$
3. $s_k \leqslant s_l$ *if* k *is even and* l *is odd.*

Indeed, note that for all n, $a_{2n+1} \geqslant a_{2n+2}$, so $a_{2n+1} - a_{2n+2} \geqslant 0$, so

$$s_{2n+2} = s_{2n} + a_{2n+1} - a_{2n+2} \geqslant s_{2n}$$

and similarly as $a_{2n+2} \geqslant a_{2n+3}$, we have

$$s_{2n+3} = s_{2n+1} - a_{2n+2} + a_{2n+3} \leqslant s_{2n+1}$$

To prove the third inequality, first notice that

$$s_{2n} = s_{2n-1} - a_{2n} \leqslant s_{2n-1}$$

since $a_{2n} \geqslant 0$. Now, if k is even and l is odd, choose n such that $k \leqslant 2n$ and $l \leqslant 2n-1$. Then

$$s_k \leqslant s_{2n} \leqslant s_{2n-1} \leqslant s_l$$

which proves the third inequality.
Now, the sequence $\{s_{2n}\}$ converges, because it is nondecreasing and is bounded above (by s_l for any odd l). Let $\alpha = \sup\{s_{2n}\} = \lim_{n\to\infty} s_{2n}$. Similarly, let $\beta = \inf\{s_{2n+1}\} = \lim_{n\to\infty} s_{2n+1}$. It follows from our third inequality that $\alpha \leqslant \beta$; since

$$s_{2n+1} - s_{2n} = a_{2n+1} \quad \textit{and} \quad \lim_{n\to\infty} a_n = 0$$

it is actually the case that $\alpha = \beta$. This proves that $\alpha = \beta = \lim_{n\to\infty} s_n$.

■

Definition 4.3.5

A sequence $\{a_n\}$ is a **rearrangement** *of a sequence $\{a_n\}$ if each $b_n = a_{f(n)}$ where f is a certain permutation on the natural numbers.*

Theorem 4.3.3

If $\sum_{n=1}^{\infty} a_n$ converges, but does not converge absolutely; then for any number α there is a rearrangement $\{b_n\}$ of $\{a_n\}$ such that $\sum_{n=1}^{\infty} b_n = \alpha$.

Proof.

Let $\sum_{n=1}^{\infty} p_n$ denote the series formed from the positive terms of $\{a_n\}$ and let $\sum_{n=1}^{\infty} q_n$ denote the series of negative terms. It follows from Theorem 4.3.2 that at least one of these series does not converge. As a matter of fact, both must fail to converge, for if one had bounded partial sums, and the other had unbounded partial sums, then the original series $\sum_{n=1}^{\infty} a_n$ would also have unbounded partial sums, contradicting the assumption that it converges.

Let α be any number. Assume, for simplicity, that $\alpha > 0$. Since the series $\sum_{n=1}^{\infty} p_n$ there is a number N such that

$$\sum_{n=1}^{N} p_n > \alpha$$

We will choose N_1 to be the smallest N with this property. This means that

$$\sum_{n=1}^{N_1-1} p_n \leqslant \alpha \quad \text{and} \quad \sum_{n=1}^{N_1} p_n > \alpha$$

Then if $S_1 = \sum_{n=1}^{N_1} p_n$, we have $S_1 - \alpha \leqslant p_{N_1}$. Next, choose the smallest integer M_1 for which

$$T_1 = S_1 + \sum_{n=1}^{M_1} q_n < \alpha$$

As before, we have $\alpha - T_1 \leqslant -q_{M_1}$. We continue this process indefinitely. The sequence

$$p_1, ..., p_{N_1}, q_1, ..., q_{M_1}, q_{N_1+1}, ..., p_{N_2}, ...$$

is a rearrangement of $\{a_n\}$. The partial sums of this rearrangement increase to S_1, then decrease to T_1, then increase to S_2, etc. To complete the proof we note that $|S_k - \alpha|$ and $|T_k - \alpha|$ are less that or equal to p_{N_k} or $-q_{M_k}$, respectively, and that these terms, being numbers of the original sequence $\{a_n\}$, must decrease to 0, since $\sum_{n=1}^{\infty} a_n$ converges.

■

Theorem 4.3.4

If $\sum_{n=1}^{\infty} a_n$ converges absolutely, and $\{b_n\}$ is any rearrangement of $\{a_n\}$, then $\sum_{n=1}^{\infty} b_n$ also converges (absolutely), and in particular

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} b_n$$

Proof.

Denote the partial sums of $\{a_n\}$ by s_n, and the partial sums of $\{b_n\}$ by t_n. Suppose that $\varepsilon > 0$. Since $\sum_{n=1}^{\infty} a_n$ converges, there is some N such that

$$\left|\sum_{n=1}^{\infty} a_n - s_N\right| < \varepsilon$$

Moreover, since $\sum_{n=1}^{\infty} |a_n|$ *converges, we can also choose* N' *so that*

$$\sum_{n=1}^{\infty} |a_n| - (|a_1| + \ldots + |a_{N'}|) < \varepsilon$$

so that

$$\sum_{n=N+1}^{\infty} |a_n| < \varepsilon$$

Choose M *such that each* $a_1, ..., a_N$ *appear among* $b_1, ..., b_M$. *Then whenever* $m > M$, *the difference* $t_m - s)N$ *is the sum of certain* a_i, *where* $i > N$. *Consequently,*

$$|t_m - s_N| \leqslant \sum_{n=N+1}^{\infty} |a_n| < \varepsilon$$

Thus, if $m >$, *then*

$$\begin{aligned}\left|\sum_{n=1}^{\infty} a_n - t_m\right| &= \left|\sum_{n=1}^{\infty} a_n - s_N - (t_m - s_N)\right| \\ &\leqslant \left|\sum_{n=1}^{\infty} a_n - s_N\right| + |(t_m - s_N)| \\ &< \varepsilon + \varepsilon\end{aligned}$$

Since this is true for every $\varepsilon > 0$, *the series* $\sum_{n=1}^{\infty} b_n$ *converges to* $\sum_{n=1}^{\infty} a_n$.

To show that $\sum_{n=1}^{\infty} b_n$ *converges absolutely, note that* $\{|b_n|\}$ *is a rearrangement of* $\{|a_n|\}$; *since* $\sum_{n=1}^{\infty} |a_n|$ *converges absolutely,* $\sum_{n=1}^{\infty} |b_n|$ *converges by the first part of the theorem.*

■

Theorem 4.3.5

If $\sum_{n=1}^{\infty} a_n$ *and* $\sum_{n=1}^{\infty} b_n$ *converge absolutely, and* $\{c_n\}$ *is any sequence containing the products* $a_i b_j$ *for each pair* $(i, j) \in \mathbb{N} \times \mathbb{N}$, *then*

$$\sum_{n=1}^{\infty} c_n = \sum_{n=1}^{\infty} a_n \cdot \sum_{n=1}^{\infty} b_n$$

Proof.

Notice first that the sequence

$$p_L = \sum_{i=1}^{L} |a_i| \cdot \sum_{j=1}^{L} |b_j|$$

converges since $\{a_n\}$ *and* $\{b_n\}$ *are absolutely convergent, and since the limit of a product is the product of the limits. So* $\{p_L\}$ *is a Cauchy sequence, which means that for any* $\varepsilon > 0$*, if there exists N such that for all* $L, L' \geqslant N$*,*

$$\left| \sum_{i=1}^{L'} |a_i| \cdot \sum_{j=1}^{L'} |b_j| - \sum_{i=1}^{L} |a_i| \cdot \sum_{j=1}^{L} |b_j| \right| < \varepsilon/2$$

It follows that

$$\sum_{i \text{ or } j > L} |a_i| \cdot |b_i| \leqslant \varepsilon/2 < \varepsilon \tag{1}$$

Now suppose that N is such that the terms c_n *for* $n \leqslant N$ *include all* $a_i b_j$ *for* $i, j \leqslant L$*. Then the difference*

$$\sum_{n=1}^{N} c_n - \sum_{i=1}^{L} a_i \cdot \sum_{j=1}^{L} b_j$$

consists of terms $a_i b_j$ *with* $i > L$ *of* $j > L$*, so*

$$\left| \sum_{n=1}^{N} c_n - \sum_{i=1}^{L} a_i \cdot \sum_{j=1}^{L} b_j \right| \leqslant \sum_{i \text{ or } j > L} |a_i| \cdot |b_j| < \varepsilon \tag{2}$$

But since the limit of a product is the product of the limits we also have

$$\left| \sum_{i=1}^{\infty} a_i \cdot \sum_{j=1}^{\infty} b_j - \sum_{i=1}^{L} a_i \cdot \sum_{j=1}^{L} b_j \right| < \varepsilon \tag{3}$$

for sufficiently large L. Consequently, if we choose L and then N large enough, we will have

$$\begin{aligned} \left| \sum_{i=1}^{\infty} a_i \cdot \sum_{j=1}^{\infty} b_j - \sum_{i=1}^{N} c_i \right| &\leqslant \left| \sum_{i=1}^{\infty} a_i \cdot \sum_{j=1}^{\infty} b_j - \sum_{i=1}^{L} a_i \cdot \sum_{j=1}^{L} b_j \right| \\ &\quad + \left| \sum_{n=1}^{N} c_n - \sum_{i=1}^{L} a_i \cdot \sum_{j=1}^{L} b_j \right| \\ &< 2\varepsilon \end{aligned}$$

which proves the theorem.

■

4.4 Uniform Convergence and Power Series

Remark 4.4.1

We are now interested in the study of series of functions, or in other words functions of the form

$$f(x) = f_1(x) + f_2(x) + f_3(x) + \dots$$

In such a situation $\{f_n\}$ will be some sequence of functions; for each x we obtain a sequence of numbers $\{f_n(x)\}$, and $f(x)$ is the sum of this sequence. Recall that each sum $f_1(x) + f_2(x) + f_3(x) + \dots$ is, by definition, the limit of the sequence $f_1(x), f_1(x) + f_2(x), f_1(x) + f_2(x) + f_3(x), \dots$. If we define a new sequence of functions $\{s_n\}$ by

$$s_n = f_1 + \dots + f_n$$

then we can express this fact more succinctly by writing

$$f(x) = \lim_{n\to\infty} s_n(x)$$

for some $x \in \mathbb{R}$.

Remark 4.4.2

First let us consider functions of the form

$$f(x) = \lim_{n\to\infty} f_n(x)$$

All this form may seem simple, it is very important to note that **nothing one would hope to be true actually is***. Instead we have a flurry of lovely counter-examples.*

Example: Counter-Example 1

Even if each f_n is continuous, the function f may not be! Indeed, consider the sequence of functions

$$f_n(x) = \begin{cases} x^n, & 0 \leqslant x \leqslant 1 \\ 1, & x \geqslant 1 \end{cases}$$

These functions are all continuous, but the function $f(x) = \lim_{n\to\infty} f_n(x)$ is not continuous; in fact,

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0, & 0 \leqslant x < 1 \\ 1, & x \geqslant 1 \end{cases}$$

Another example of this phenomenon is illustrated by the family of functions

$$f_n(x) = \begin{cases} -1, & x \leqslant -\frac{1}{n} \\ nx, & -\frac{1}{n} \leqslant x \leqslant \frac{1}{n} \\ 1, & \frac{1}{n} \leqslant x \end{cases}$$

In this case, if $x < 0$ $f_n(x)$ is eventually -1, and if $x > 0$, then $f_n(x)$ is eventually 1, while

$f_n(0) = 0$ for all n. Thus,

$$\lim_{n\to\infty} f_n(x) = \begin{cases} -1, & x < 0 \\ 0, & x = 0 \\ 1, & x > 0 \end{cases}$$

so once again $f(x) = \lim_{n\to\infty} f_n(x)$ is not continuous.

Example: Counter-Example 2

It is even possible to produce a sequence of differentiable functions $\{f_n\}$ for which the function $f(x) = \lim_{n\to\infty} f_n(x)$ is not continuous. One such sequence is

$$f_n(x) = \begin{cases} -1, & x \leqslant -\frac{1}{n} \\ \sin\left(\frac{n\pi x}{2}\right), & -\frac{1}{n} \leqslant x \leqslant \frac{1}{n} \\ 1, & \frac{1}{n} \leqslant x \end{cases}$$

These functions are differentiable, but we still have

$$\lim_{n\to\infty} f_n(x) = \begin{cases} -1, & x < 0 \\ 0, & x = 0 \\ 1, & x > 0 \end{cases}$$

Definition 4.4.1

If f is a function defined on some set A, and a sequence of functions $\{f_n\}$, all defined on the same set A, are such that only

$$f(x) = \lim_{n\to\infty} f_n(x)$$

for all $x \in A$. Precisely, $\{f_n\}$ is said to **converge pointwise to f on A** *if for all $\varepsilon > 0$, and for all $x \in A$, there is some N such that if $n \geqslant N$, then $|f(x) - f_n(x)| < \varepsilon$.*

Definition 4.4.2

Let $\{f_n\}$ be a sequence of functions defined on A, and let f be a function which is also defined on A. Then f is called the **uniform limit of $\{f_n\}$ on A** *if for every $\varepsilon > 0$ there is some N such that for all $x \in A$,*

$$\text{if } n > N, \text{ then } |f(x) - f_n(x)| < \varepsilon$$

We also say that $\{f_n\}$ **converges uniformly to f on A**, *or that f_n* **approaches f uniformly on A**.

Remark 4.4.3

Note that uniform convergence implies pointwise convergence, but the converse is not true.

Theorem 4.4.1

Suppose that $\{f_n\}$ is a sequence of functions which are integrable on $[a, b]$, and that $\{f_n\}$

converges uniformly on $[a,b]$ to a function f which is also integrable on $[a,b]$. Then

$$\int_a^b f = \lim_{n\to\infty} \int_a^b f_n$$

Proof.

Let $\varepsilon > 0$. Then since $\{f_n\}$ converges uniformly to f, there exists $N \in \mathbb{N}$ such that for all $x \in [a,b]$, if $n \geqslant N$

$$|f(x)f_n(x)| < \varepsilon$$

Then for all $n \geqslant N$, we have that

$$\begin{aligned}\left|\int_a^b f(x)dx - \int_a^b f_n(x)dx\right| &= \left|\int_a^b [f(x) - f_n(x)]dx\right| \\ &\leqslant \int_a^b |f(x) - f_n(x)|dx \\ &\leqslant \int_a^b \varepsilon dx \\ &= \varepsilon(b-a)\end{aligned}$$

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

$$\int_a^b f = \lim_{n\to\infty} \int_a^b f_n$$

■

Theorem 4.4.2

Suppose that $\{f_n\}$ is a sequence of functions which are continuous on $[a,b]$, and that $\{f_n\}$ converges uniformly on $[a,b]$ to f. Then f is also continuous on $[a,b]$.

Proof.

For each $x \in [a,b]$ we must prove that f is continuous at x. We first deal with $x \in (a,b)$. Fix $\varepsilon > 0$. Since $\{f_n\}$ converges uniformly to f on $[a,b]$, there is some N such that for all $n \geqslant N$ and all $y \in [a,b]$,

$$|f(y) - f_n(y)| < \varepsilon/3$$

In particular, for all h such that $x+h \in [a,b]$, we have

$$\begin{aligned}|f(x) - f_n(x)| &< \varepsilon/3, \\ |f(x+h) - f_n(x+h)| &< \varepsilon/3\end{aligned}$$

Now f_n is continuous, so there is some $\delta > 0$ such that for $|h| < \delta$ we have

$$|f_n(x) - f_n(x+h)| < \varepsilon/3$$

Thus, if $|h| < \delta$, then

$$|f(x+h) - f(x)| = |f(x+h) - f_n(x+h) + f_n(x+h) - f_n(x) + f_n(x) - f(x)|$$

$$
\begin{aligned}
&\leqslant |f(x+h) - f_n(x+h)| + |f_n(x+h) - f_n(x)| + |f_n(x) - f(x)| \\
&< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\
&= \varepsilon
\end{aligned}
$$

This proves that f is continuous at x.

■

Remark 4.4.4

Allow this last two theorems are great successes, differentiability sadly fails. Even if each f_n is differentiable and $\{f_n\}$ converges uniformly to f, it need not be the case that f is differentiable. Moreover, even if f is itself differentiable, it need not be the case that

$$f'(x) = \lim_{n\to\infty} f_n'(x)$$

Example: Counter Example 3

Consider the family of functions

$$f_n(x) = \frac{1}{n}\sin(n^2 x)$$

then $\{f_n\}$ converges uniformly to the function $f(x) = 0$, but

$$f_n'(x) = n\cos(n^2 x)$$

and $\lim_{n\to\infty} n\cos(n^2 x)$ does not even always exist (for example if $x = 0$).

Theorem 4.4.3

Suppose that $\{f_n\}$ is a sequence of functions which are differentiable on $[a, b]$, with integrable derivatives f_n', and that $\{f_n\}$ converges (pointwise) to f. Suppose, moreover, that $\{f_n'\}$ converges uniformly on $[a, b]$ to some continuous function g. Then f is differentiable and

$$f'(x) = \lim_{n\to\infty} f_n'(x)$$

Proof.

Applying Theorem 1 to the interval $[a, x]$, we see that for each x we have

$$
\begin{aligned}
\int_a^x g &= \lim_{n\to\infty} \int_a^x f_n' \\
&= \lim_{n\to\infty} [f_n(x) - f_n(a)] \qquad \text{(by Second Fundamental Theorem of Calculus)} \\
&= f(x) - f(a)
\end{aligned}
$$

Since g is continuous, it follows that $f'(x) = g(x) = \lim_{n\to\infty} f_n'(x)$ for all x in the interval $[a, b]$, by The Fundamental Theorem of Calculus.

■

Definition 4.4.3

The series $\sum_{n=1}^{\infty} f_n$ **converges uniformly** *(more formally, the sequence* $\{f_n\}$ *is* **uniformly summable***)* **to** f **on** A*, if the sequence*

$$f_1, f_1 + f_2, f_1 + f_2 + f_3, \ldots$$

converges uniformly to f *on* A*.*

Corollary 4.4.4

Let $\sum_{n=1}^{\infty} f_n$ *converge uniformly to* f *on* $[a, b]$*.*

1. *If each* f_n *is continuous on* $[a, b]$*, then* f *is continuous on* $[a, b]$*.*
2. *If* f *and each* f_n *is integrable on* $[a, b]$*, then*

$$\int_a^b f = \sum_{n=1}^{\infty} \int_a^b f_n$$

Moreover, if $\sum_{n=1}^{\infty} f_n$ *converges (pointwise) to* f *on* $[a, b]$*, each* f_n *has an integrable derivative* f_n' *and* $\sum_{n=1}^{\infty} f_n'$ *converges uniformly on* $[a, b]$ *to some continuous function, then*

3. $f'(x) = \sum_{n=1}^{\infty} f_n'(x)$ *for all* $x \in [a, b]$

Proof.

Let $\{s_n\}$ *be the sequence of partial sums of the* $\{f_n\}$*. Then since each* f_n *is continuous, so is each* s_n*. Then as* $\{s_n\}$ *converges uniformly to* f *we have by a previous theorem that* f *is also continuous on* $[a, b]$*. Next, since each* f_n *is integrable on* $[a, b]$*, so is each* s_n*. Then as* $\{s_n\}$ *converges uniformly to* f *we have that*

$$\begin{aligned}\int_a^b f &= \lim_{n\to\infty} \int_a^b s_n \\ &= \lim_{n\to\infty} t_n \\ &= \sum_{n=1}^{\infty} \int_a^b f_n\end{aligned}$$

where $\{t_n\}$ *is the sequence such that*

$$t_n = \sum_{i=1}^{n} \int_a^b f_n = \int_a^b s_n$$

Finally, suppose s_n converges (pointwise) to f on $[a,b]$, and each f_n has an integrable derivative f_n'. Then each s_n has an integrable derivative s_n' on $[a,b]$ by the linearity of the derivative and integral operators. Moreover, suppose s_n' converges uniformly on $[a,b]$ to some continuous function g. Then it follows that for all $x \in [a,b]$

$$f'(x) = \lim_{n\to\infty} s_n'(x) = \sum_{n=1}^{\infty} f_n'(x)$$

■

Theorem 4.10: The Weierstrass M-Test

Let $\{f_n\}$ be a sequence of functions defined on A, and suppose that $\{M_n\}$ is a sequence of numbers such that

$$|f_n(x)| \leqslant M_n, \forall x \in A$$

Suppose moreover that $\sum_{n=1}^{\infty} M_n$ converges. Then for each x in A the series $\sum_{n=1}^{\infty} f_n(x)$ converges absolutely, and $\sum_{n=1}^{\infty} f_n$ converges uniformly on A to the function

$$f(x) = \sum_{n=1}^{\infty} f_n(x)$$

Proof.

For each $x \in A$, the series $\sum_{n=1}^{\infty} |f_n(x)|$ converges by The Comparison Test; consequently $\sum_{n=1}^{\infty} f_n(x)$ converges absolutely. Moreover, for all $x \in A$ we have

$$\begin{aligned} \left| f(x) - \sum_{i=1}^{N} f(x) \right| &= \left| \sum_{n=N+1}^{\infty} f_n(x) \right| \\ &\leqslant \sum_{n=N+1}^{\infty} |f_n(x)| \\ &\leqslant \sum_{n=N+1}^{\infty} M_n \end{aligned}$$

Since $\sum_{n=1}^{\infty} M_n$ converges, the number $\sum_{n=N+1}^{\infty} M_n$ can be made as small as desired (by The Cauchy Criterion), by choosing N sufficiently large.

■

Definition 4.4.4

An infinite sum of functions of the form

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n$$

is called a **power series centered at** a*. One especially important family of power series are those of the form*

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

where f *is some infinitely differentiable function at* a*; this series is called the* **Taylor series for** f **at** a*.*

Remark 4.4.5

Given a function f *infinitely differentiable at* a*, we have for* $x \in \mathbb{R}$ *that*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

if and only if the remainder terms satisfy $\lim_{n\to\infty} R_{n,a}(x) = 0$*.*

Theorem 4.4.5

Suppose that the series

$$f(x_0) = \sum_{n=0}^{\infty} a_n x_0^n$$

converges, and let a *be any number with* $0 < a < |x_0|$*. Then on* $[-a, a]$ *the series*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

converges uniformly (and absolutely). Moreover, the same is true for the series

$$g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

Finally, f *is differentiable and*

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

for all x *with* $|x| < |x_0|$*.*

Proof.

First, since $\sum_{n=0}^{\infty} a_n x_0^n$ *converges,* $\lim_{n\to\infty} a_n x_0^n = 0$*. Hence, the sequence* $\{a_n x_0^n\}$ *is surely bounded: there is some number* M *such that*

$$|a_n x_0|^n = |a_n| \cdot |x_0|^n \leqslant M$$

for all n*. Now if* x *is in* $[-a, a]$*, then* $|x| \leqslant |a|$*, so*

$$\begin{aligned} |a_n x^n| &= |a_n| \cdot |x|^n \\ &\leqslant |a_n| \cdot |a|^n \\ &= |a_n| \cdot |x_0|^n \cdot \left|\frac{a}{x_0}\right|^n \\ &\leqslant M \left|\frac{a}{x_0}\right|^n \end{aligned}$$

But $|a/x_0| < 1$*, so the (geometric) series*

$$\sum_{n=0}^{\infty} M \left|\frac{a}{x_0}\right|^n = M \sum_{n=0}^{\infty} \left|\frac{a}{x_0}\right|^n$$

converges. Choosing $M \cdot |a/x_0|^n$ *as the number* M_n *in The Weierstrass M-Test, it follows that* $\sum_{n=0}^{\infty} a_n x^n$ *converges uniformly on* $[-a, a]$*.*

To prove the same assertion for $g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$ *notice that*

$$\begin{aligned} |n a_n x^{n-1}| &= n|a_n| \cdot |x^{n-1}| \\ &\leqslant n|a_n| \cdot |a^{n-1}| \\ &= \frac{|a_n|}{|a|} \cdot |x_0|^n n \left|\frac{a}{x_0}\right|^n \\ &\leqslant \frac{M}{|a|} n \left|\frac{a}{x_0}\right|^n \end{aligned}$$

Since $|a/x_0| < 1$*, the series*

$$\sum_{n=1}^{\infty} \frac{M}{|a|} n \left|\frac{a}{x_0}\right|^n = \frac{M}{|a|} \sum_{n=1}^{\infty} n \left|\frac{a}{x_0}\right|^n$$

converges (by an application of the Ratio Tes). Another appeal to The Weierstrass M-Test proves that $\sum_{n=1}^{\infty} n a_n x^{n-1}$ *converges uniformly on* $[-a, a]$*.*

Finally, our corollary proves, first that g *is continuous, and then that*

$$f'(x) = g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

for all $x \in [-a, a]$*. Since we could have chosen any* a *with* $0 < a < |x_0|$*, this result holds for all* x *with* $|x| < |x_0|$*.*

■

Part II

Higher-Dimesional Analysis

Chapter 5

Metric Spaces

Chapter 6

Higher-Dimensional Differentiation

Chapter 7

Higher-Dimensional Integration

Chapter 8

Manifolds

Chapter 9

Differential Forms

Chapter 10

Integration on Chains

Chapter 11

Integration on Manifolds

Part III

Function Spaces

Chapter 12

Normed Spaces

Chapter 13

Hilbert Spaces

Chapter 14

Banach Spaces

Chapter 15

Differentiation and Integration

Chapter 16

Banach Algebras

Part IV

Measure Theory

Chapter 17

Measures

Chapter 18

L^p Spaces

Chapter 19

Radon Measures

Appendices

.1 Multivariate Calculus - with Applications

.1.1 Vector Functions and Derivatives

Definition .1.1

A **vector function** *of one variable is a function $\vec{f} : J \subseteq \mathbb{R} \to \mathbb{R}^n$ defined by $t \mapsto \vec{f}(t)$, where $\vec{f}(t) \in \mathbb{R}^n$ is unique.*

Definition .1.2

Let $\vec{f} : \mathbb{R} \to \mathbb{R}^n$ be a vector valued function. Then we define the derivative of $\vec{f}$ at t by

$$\frac{d\vec{f}(t)}{dt} = \lim_{\Delta t \to 0} \frac{\vec{f}(t + \Delta t) - \vec{f}(t)}{\Delta t} \tag{.1.1}$$

Remark .1.1: Properties

Let $\vec{f} : J \subseteq \mathbb{R} \to \mathbb{R}^n$ and $\vec{g} : I \subseteq \mathbb{R} \to \mathbb{R}^n$ be vector functions such that

$$\vec{f} = \langle f_1, ..., f_n \rangle \text{ and } \vec{g} = \langle g_1, ..., g_n \rangle \tag{.1.2}$$

The for all $t \in J \cap I$ and all $\lambda : D_\lambda \subseteq \mathbb{R} \to \mathbb{R}$ we have

1. $(\vec{f} + \vec{g})(t) := \vec{f}(t) + \vec{g}(t)$
2. $(\lambda \vec{f})(t) := \lambda(t)\vec{f}(t)$
3. $(\vec{f} \cdot \vec{g})(t) = \vec{f}(t) \cdot \vec{g}(t)$
4. *For $n = 3$,* $(\vec{f} \times \vec{g})(t) = \vec{f}(t) \times \vec{g}(t)$

Recall .1.2

Given $A \in \mathbb{R}^{n\times n}$, the lagrange expansion is

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \det(A_{ij}) \text{ (along } j\text{-th column)} \tag{.1.3}$$

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \det(A_{ij}) \text{ (along } i\text{-th row)} \tag{.1.4}$$

where A_{ij} is the minor matrix of A obtained by deleting the ith ro and jth column of A.

Definition .1.3

Suppose $\vec{f}(t) = \langle f_1(t), ..., f_n(t)\rangle$ *and* $\vec{L} = \langle L_1, ..., L_n\rangle$ *Then*

$$\lim_{t\to t_0} \vec{f}(t) = \vec{L} \implies \lim_{t\to t_0} f_i(t) = L_i, \forall 1 \leqslant i \leqslant n \tag{.1.5}$$

Remark .1.3

The right hand limit $\lim_{t\to t_0^+} \vec{f}(t)$ *and the left hand limit* $\lim_{t\to t_0^-} \vec{f}(t)$ *are defined in the same way.*

Remark .1.4: Limit Rules

If $\lim_{t\to t_0}$ *for* $\vec{f}(t)$, $\vec{g}(t)$, *and* $\lambda(t)$ *exist and* $k \in \mathbb{R}$, *then*

1. $\lim\limits_{t\to t_0}(\vec{f}(t) + \vec{g}(t)) = \lim\limits_{t\to t_0} \vec{f}(t) + \lim\limits_{t\to t_0} \vec{g}(t)$

2. $\lim\limits_{t\to t_0} k\vec{f}(t) = k \lim\limits_{t\to t_0} \vec{f}(t)$

3. $\lim\limits_{t\to t_0} \lambda(t)\vec{f}(t) = (\lim\limits_{t\to t_0} \lambda(t))(\lim\limits_{t\to t_0} \vec{f}(t))$

4. $\lim\limits_{t\to t_0} \vec{f}(t) \cdot \vec{g}(t) = (\lim\limits_{t\to t_0} \vec{f}(t)) \cdot (\lim\limits_{t\to t_0} \vec{g}(t))$

5. $\lim\limits_{t\to t_0} \vec{f}(t) \times \vec{g}(t) = (\lim\limits_{t\to t_0} \vec{f}(t)) \times (\lim\limits_{t\to t_0} \vec{g}(t))$, *for* $n = 3$.

Definition .1.4: Continuity

A vector function $\vec{f}(t) = \langle f_1(t), ..., f_n(t)\rangle$ *is said to be continuous at* $t = t_0$ *if*

$$\lim_{t\to t_0} \vec{f}(t) = \vec{f}(t_0) \tag{.1.6}$$

In other words, each component function is continuous at $t = t_0$.

Definition .1.5: Differentiability

A vector function $\vec{f}(t) = \langle f_1(t), ..., f_n(t)\rangle$ *is said to be differentiable at* $t = t_0$ *if* $\vec{f}(t)$ *is defined at and around t and*

$$\lim_{t\to t_0} \frac{\vec{f}(t) - \vec{f}(t_0)}{t - t_0} \tag{.1.7}$$

exists. We call this limit the derivative of $\vec{f}(t)$ *at* $t = t_0$ *and is denoted by* $\vec{f}'(t_0)$ *if it exists.*

Theorem .1.1

We say that the vector function is differentiable at $t = t_0$ *if and only if its component functions* *are differentiable at $t = t_0$, and*

$$\vec{f}'(t_0) = \langle f_1'(t_0), ..., f_n'(t_0)\rangle \tag{.1.8}$$

Remark .1.5: Differentiation Rules

Let $\vec{f}(t), \vec{g}(t)$ and $\lambda(t)$ be differentiable and $k \in \mathbb{R}$. Then

1. $(\vec{f} + \vec{g})'(t) = \vec{f}'(t) + \vec{g}'(t)$
2. $(k\vec{f})'(t) = k\vec{f}'(t)$
3. $(\lambda\vec{f})'(t) = \lambda'(t)\vec{f}(t) + \lambda(t)\vec{f}'(t)$
4. $(\vec{f} \cdot \vec{g})'(t) = \vec{f}'(t) \cdot \vec{g}(t) + \vec{f}(t) \cdot \vec{g}'(t)$
5. $(\vec{f} \times \vec{g})'(t) = \vec{f}'(t) \times \vec{g}(t) + \vec{f}(t) \times \vec{g}'(t)$ *for* $n = 3$
6. $(\vec{f}(\lambda(t)))' = \vec{f}'(\lambda(t))\lambda'(t)$

Definition .1.6

Let $\vec{f}(t) = \langle f_1(t), ..., f_n(t)\rangle$ be a vector function defined on a closed interval $[a, b]$. We say that $\vec{f}(t)$ is **integrable** *on $[a, b]$ if each $f_i(t)$ is integrable on $[a, b]$. When that is the case we define*

$$\int_a^b \vec{f}(t)dt := \left\langle \int_a^b f_1(t)dt, ..., \int_a^b f_n(t)dt \right\rangle \tag{.1.9}$$

the definite integral of $\vec{f}(t)$ on $[a, b]$

Remark .1.6: Properties

Let $\vec{f}(t), \vec{g}(t)$ be integrable on $[a, b]$ and $k \in \mathbb{R}$. Then

1. $\int\limits_a^b k\vec{f}(t)dt = k\int\limits_a^b \vec{f}(t)dt$
2. $\int\limits_a^b (\vec{f}(t) + \vec{g}(t))dt = \int\limits_a^b \vec{f}(t)dt + \int\limits_a^b \vec{g}(t)dt$
3. $\int\limits_a^b \vec{f}(t)dt = \int\limits_a^c \vec{f}(t)dt + \int\limits_c^b \vec{f}(t)dt$, $a \leqslant c \leqslant b$.
4. $\left|\int\limits_a^b \vec{f}(t)dt\right| \leqslant \int\limits_a^b \left|\vec{f}(t)\right| dt$ *(***The triangle inequality***)*

.1.2 Parametric Curves and Paths

Definition .1.7

Let $\vec{f} : J \subseteq \mathbb{R} \to \mathbb{R}^n$ given by $t \mapsto \langle f_1(t), ..., f_n(t)\rangle$, and $f_j : J \subseteq \mathbb{R} \to \mathbb{R}$ be the jth component function of $\vec{f}$. We define the image

$$\vec{f}(J) = \mathscr{C} \tag{.1.10}$$

and call $\mathscr{C}$ a **curve parametrized** *by $\vec{f}$. If $J = [a, b]$ for $a \leqslant b \in \mathbb{R}$ and $\vec{f}(a) = \vec{f}(b)$, then $\mathscr{C}$ is a* **closed curve**. *If there exists a parametrization $\vec{g} : I \subseteq \mathbb{R} \to \mathbb{R}^n$ of $\mathscr{C}$ such that $\vec{g}$ is injective (except maybe at end points), then $\mathscr{C}$ is said to be a* **non-self intersecting** *curve. If such a curve is closed it is called a* **simple closed curve**.

Remark .1.7

The pair $(\vec{f}(t), J)$ is called a **parametrization** *of the curve $\mathscr{C}$. The triple $(\vec{f}(t), J, \mathscr{C})$ is called a* **path with curve** $\mathscr{C}$.

Definition .1.8

If $\vec{f} : J \subseteq \mathbb{R} \to \mathbb{R}^n$ is a one-to-one function, then the image $\vec{f}(J) = \mathscr{C}$ is an **oriented curve** *and $\vec{f}$ is a* **consistently oriented path** *which covers $\mathscr{C}$.*

Remark .1.8: Ellipse

The parametrization of an ellipse with equation $\frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{b^2} = 1$, $a, b > 0$ are constants, is $\vec{f}(t) = \langle x_0 + a\cos(t), y_0 + b\sin(t)\rangle$, for $t \in [0, 2\pi)$.

Remark .1.9: Line

Let A and B be points in $\mathbb{R}^n$. The parametrization of the line segment connecting A to B is

$$(\vec{f}(t) = \overline{OA} + t(\overline{OB} - \overline{OA}), [0, 1]) \tag{.1.11}$$

Definition .1.9

The tangent vector to a curve $\mathscr{C}$ with parametrization $(\vec{f}, I)$ exists at a point $t = t_0$ if $\vec{f}$ is $\vec{f}'(t)$ exists.

Definition .1.10

Let $\mathscr{C}$ be a parametric curve with parametrization $(\vec{f}(t), [a, b])$. If $\vec{f}(t)$ is differentiable at $t = t_0$, then $\vec{f}(t_0)$ is called a **tangent vector** *to $\mathscr{C}$ at $P_0 = tip(\vec{f}(t_0))$, provided $\vec{f}'(t_0) \neq \vec{0}$. If $\vec{f}(t)$ is differentiable at every $t \in (a, b)$ and $\vec{f}'(t) \neq \vec{0}$, we say that the parametric curve $\mathscr{C}$ is a* **smooth parametric curve**.

Definition .1.11

Let $\mathscr{C}$ be a parametric curve and let $(\vec{f}(t), I)$ be a parametrization. If $P_0 = tip(\vec{f}(t_0))$, for $t_0 \in I$, such that $\vec{f}'(t_0) \neq \vec{0}$. Then

$$\vec{T}(t_0) = \frac{1}{|\vec{f}'(t_0)|}\vec{f}'(t_0) \tag{.1.12}$$

is called the **unit tangent vector** *associated with the parametrization $(\vec{f}(t), I)$. $\vec{T}(t_0)$ always forces the direction in which $\vec{f}(t)$ traces $\mathscr{C}$. The vector*

$$\vec{N}(t_0) = \frac{1}{|\vec{T}'(t_0)|}\vec{T}'(t_0) \tag{.1.13}$$

is perpendicular to $\vec{T}(t_0)$ and is called the **unit principal normal** *to $\mathscr{C}$ at P_0.*

Definition .1.12

A curve $\mathscr{C}$ is called **piecewise smooth** *if it consists of a finite number of smooth parametric curves $\mathscr{C}_1, ..., \mathscr{C}_k$, where the endpoint of $\mathscr{C}_i$ is the starting point of $\mathscr{C}_{i+1}$ for $i = 1, 2, ..., k-1$.*

Definition .1.13

Let $\mathscr{C}$ be a bounded continuous curve specified by a parametrization $\vec{f} : [a, b] \to \mathbb{R}^n$. We consider partitions of $[a, b]$ into n-subintervals by

$$a = t_0 < t_1 < ... < t_n = b \tag{.1.14}$$

So the points $\vec{f}(t_i)$ subdivide $\mathscr{C}$, and using the **chord length** *$|\vec{f}(t_i) - \vec{f}(t_{i-1})|$ we define the sequence of lengths approximating $\mathscr{C}$ by*

$$s_n = \sum_{i=1}^{n} |\vec{f}(t_i) - \vec{f}(t_{i-1})| \tag{.1.15}$$

We say $\mathscr{C}$ is **rectifiable** *if there exists $K \in \mathbb{R}$ such that $s_n \leqslant K$ for all $n \in \mathbb{N}$ and all choices of points. From the completeness axiom of $\mathbb{R}$ there exists a least such K. This K we define as the* **length** *of $\mathscr{C}$ and we denote it by s. Let $\Delta t_i = t_i - t_{i-1}$ and $\Delta\vec{f_i} = \vec{f}(t_i) - \vec{f}(t_{i-1})$ so*

$$s_n = \sum_{i=1}^{n} \left|\frac{\Delta\vec{f_i}}{\Delta t_i}\right| \Delta t_i \tag{.1.16}$$

If $\vec{f}(t)$ has a continuous derivative, then

$$s = \lim_{\substack{n\to\infty \\ sup\Delta t_i \to 0}} s_n = \int_a^b \left|\frac{d\vec{f}}{dt}\right| dt \tag{.1.17}$$

is the **arclength***.*

Definition .1.14

Let $\vec{f} : [a,b] \to \mathbb{R}^n$ be a smooth parammetrization of a curve $\mathscr{C}$. Then the **arclength function of $\vec{f}$** *is a function $s : [a,b] \to \mathbb{R}$ where*

$$s(t) := \int_a^b \left| \frac{d\vec{f}}{dt} \right| dt \tag{.1.18}$$

and the **arclength element** *for $\mathscr{C}$ is given by*

$$ds := \left| \frac{d}{dt}\vec{f} \right| dt \tag{.1.19}$$

Definition .1.15

If $\vec{f} : J \subseteq \mathbb{R} \to \mathbb{R}^n$ parametrizes a curve $\mathscr{C}$ with the parameter being the arclenght along the curve relative to some inital point, then we call this an **arclength** *or* **intrinsic parametrization**. *Such a parametrization traces $\mathscr{C}$ at unit speed*

$$\left| \frac{d\vec{f}(s)}{dt} \right| = 1 \tag{.1.20}$$

.1.3 Functions of Several Variables

Definition .1.16

A function $f : \mathscr{D}(f) \subseteq \mathbb{R}^n \to \mathbb{R}$, where $\mathscr{D}(f)$ is the **domain** *of f, is called a* **scalar field on $\mathscr{D}(f)$**. *The image of f is*

$$\mathrm{Im}(f) := \{x \in \mathbb{R} : \exists \vec{v} \in \mathscr{D}(f), f(\vec{v}) = x\} \tag{.1.21}$$

The natural domain of f is the largest subset of $\mathbb{R}^n$ such that f is well-defined.

Definition .1.17

The **graph** *of a function $f : A \to B$ is the set*

$$\Gamma(f) := \{(x, f(x)) : x \in A\} \tag{.1.22}$$

For

$$f : \prod_{i=1}^{n} X_i \to B \tag{.1.23}$$

we have the graph

$$\Gamma(f) := \{((x_1, ..., x_n), f(x_1, ..., x_n)) : (x_1, ..., x_n) \in \prod_{i=1}^{n} X_i\} \tag{.1.24}$$

Remark .1.10

The grap of a function $f : \mathbb{R}^n \to \mathbb{R}$ can be considered as a surface in $\mathbb{R}^{n+1}$ by a natural embedding.

Definition .1.18

Given a function $f : \mathbb{R}^n \to \mathbb{R}$, a **k-level surface** *of f is a set*

$$S_k := \{\vec{x} \in \mathbb{R}^n : f(\vec{x}) = k\} \tag{.1.25}$$

where k is a fixed constant.

Definition .1.19

Let $\vec{x}_0 \in \mathbb{R}^n$ and $r > 0$ a real number. Then the **open ball of radius** r *centered at $\vec{x}_0$ is defined as*

$$B_r(\vec{x}_0) := \{\vec{x} \in \mathbb{R}^n : ||\vec{x} - \vec{x}_0|| < r\} \tag{.1.26}$$

the **closed ball** *is defined by*

$$\overline{B}_r(\vec{x}_0) := \{\vec{x} \in \mathbb{R}^n : ||\vec{x} - \vec{x}_0|| \leqslant r\} \tag{.1.27}$$

Definition .1.20

A **neighborhood** *of a point $\vec{x}_0 \in \mathbb{R}^n$ is any set $U \subseteq \mathbb{R}^n$ such that there exists $r > 0$ so that $B_r(\vec{x}_0) \subseteq U$.*

Definition .1.21

Let $E \subseteq \mathbb{R}^n$, where we equip $\mathbb{R}^n$ with $\mathscr{T}_{st}$. We say E is **open** *if $E \in \mathscr{T}_{st}$. Equivalently, E is* **open** *if for all $\vec{x} \in E$, E is a* **neighborhood** *of $\vec{x}$. We sau E is* **closed** *if its* **complement** *$E^C = \mathbb{R}^n \backslash E$ is* **open**.

Definition .1.22

A point $\vec{x}_0 \in \mathbb{R}^n$ is called a **boundary point** *of E if for any $r > 0$, $B_r(\vec{x}_0) \cap E \neq \varnothing$ and $B_r(\vec{x}_0) \cap E^C \neq \varnothing$.*

Definition .1.23

The set of all **boundary points** *of a set $E \subseteq \mathbb{R}^n$ is called the* **boundary** *of E.*

Definition .1.24

Let $E \subseteq \mathbb{R}^n$. We say that E is **bounded** *if there exists $R > 0$ such that $||\vec{x}|| \leqslant R$ for all $\vec{x} \in E$. E is* **unbounded** *if for all $R > 0$ there exists $\vec{x}_0 \in E$ such that $||\vec{x}_0|| > R$.*

Definition .1.25

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. We say $\lim_{\vec{x}\to\vec{x}_0} f(\vec{x}) = L$, provided that

1. *Every punctered neighborhood $B_r^*(\vec{x}_0)$ of $\vec{x}_0$ intersects $\mathscr{D}(f)$*

$$B_r^*(\vec{x}_0) \cap \mathscr{D}(f) \neq \varnothing \tag{.1.28}$$

that is, $\vec{x}_0$ is a limit point of $\mathscr{D}(f)$.

2. *For all $\varepsilon > 0$, there exists $\delta > 0$ such that $f(\vec{x}) \in B_\varepsilon(L)$ whenever $\vec{x} \in B_\delta^*(\vec{x}_0) \cap \mathscr{D}(f)$. That is*

$$f(\mathscr{D}(f) \cap B_\delta^*(\vec{x}_0)) \subseteq B_\varepsilon(L) \tag{.1.29}$$

Remark .1.11

As $\mathbb{R}^n$ and $\mathbb{R}$ are metric spaces, they are Hausdorff, so if the limit exists it is unique.

Remark .1.12: Limit Properties

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ and $\vec{x}_0 \in \mathbb{R}^n$ such that $\lim_{\vec{x}\to\vec{x}_0} f(\vec{x}_0) = L$ and $\lim_{\vec{x}\to\vec{x}_0} g(\vec{x}) = M$. Then if $\vec{x}_0$ is not an isolated point of $\mathscr{D}(f) \cap \mathscr{D}(g)$, then

1. $\lim_{\vec{x}\to\vec{x}_0}(f(\vec{x}) \pm g(\vec{x})) = L \pm M$
2. $\lim_{\vec{x}\to\vec{x}_0} f(\vec{x})g(\vec{x}) = LM$
3. $\lim_{\vec{x}\to\vec{x}_0} \frac{f(\vec{x})}{g(\vec{x})} = \frac{L}{M}$ *if $M \neq 0$*
4. *If $F : \mathbb{R} \to X$ is continuous at L, then $\lim_{\vec{x}\to\vec{x}_0} F(f(\vec{x})) = F(L)$.*

Definition .1.26

We say a function $f : \mathbb{R}^n \to \mathbb{R}$ is **continuous at a point** *$\vec{x}_0 \in \mathbb{R}^n$ if*

$$\lim_{\vec{x}\to\vec{x}_0} f(\vec{x}) = f(\vec{x}_0) \tag{.1.30}$$

Remark .1.13

If setting $x = x_0$ and $y = y_0$ in the expression for $f(x,y)$ does not evalutate to a real number, then we can try using polar coordinates: $x = x_0 + r\cos(\theta)$ and $y = y_0 + r\sin(\theta)$. Recall $r = \sqrt{(x-x_0)^2 + (y-y_0)^2}, 0 \leqslant \theta < 2\pi$. As a result, $(x,y) \to (x_0,y_0)$ is equivalent to $r \to 0$, so

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = \lim_{r\to 0} f(x_0 + r\cos(\theta), y_0 + r\sin(\theta)) \tag{.1.31}$$

Theorem .1.2: Squeeze Theorem

Let $f(x,y), g(x,y)$ and $h(x,y)$ be defined in a neighborhood U of (x_0,y_0), except maybe at (x_0,y_0), and such that

$$g(x,y) \leqslant f(x,y) \leqslant h(x,y), \forall (x,y) \in U\backslash\{(x_0,y_0)\} \tag{.1.32}$$

If

$$\lim_{(x,y)\to(x_0,y_0)} g(x) = L, \lim_{(x,y)\to(x_0,y_0)} h(x) = L \tag{.1.33}$$

Then $\lim\limits_{(x,y)\to(x_0,y_0)} f(x) = L.$

Theorem .1.3

If one can find two continuous parametric curves $\mathscr{C}_1$ and $\mathscr{C}_2$ that pass through the point (x_0,y_0) such that

$$\lim_{\substack{(x,y)\to(x_0,y_0)\\(x,y)\in\mathscr{C}_1}} f(x,y) = L_1, \lim_{\substack{(x,y)\to(x_0,y_0)\\(x,y)\in\mathscr{C}_2}} f(x,y) = L_2, \text{ and } L_1 \neq L_2 \tag{.1.34}$$

then $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)$ *does not exist.*

Definition .1.27

Let $f:\mathbb{R}^n \to \mathbb{R}$ be a function.

1. *We say that f is* **continuous** *at $\vec{x}_0 \in \mathbb{R}^n$ if*
   (a) *There exists a neighborhood U of $\vec{x}_0$ such that $f(U)$ is defined*
   (b) $\lim_{\vec{x}\to\vec{x}_0} f(\vec{x}) = f(\vec{x}_0)$
2. *We say that f is continuous on a region D if it s continuous at every point $\vec{x}$ in the region.*

Remark .1.14: Constructing Continuous Functions

Let $f, g:\mathbb{R}^n \to \mathbb{R}$ be continuous at $\vec{x}_0 \in \mathbb{R}^n$, and if $\lambda \in \mathbb{R}$, then

1. *$f \pm g$, $f \cdot g$, and λf are continuous at $\vec{x}_0$*
2. *f/g is continuous at $\vec{x}_0$ provided $g(\vec{x}_0) \neq 0$.*

Suppose $u(t)$ is continuous at $t_0 = f(\vec{x}_0)$. Then $u(f(\vec{x}))$ is continuous at $\vec{x}_0$.

Theorem .1: Extreme Value Theorem

If $f:\mathbb{R}^n \to \mathbb{R}$ is continuous on a closed and bounded region $D \subseteq \mathbb{R}^n$, then there exist $\vec{x}_m, \vec{x}_M \in D$ such that

$$f(\vec{x}_m) \leqslant f(\vec{x}) \leqslant f(\vec{x}_M), \forall \vec{x} \in D \tag{.1.35}$$

$m = f(\vec{x}_m)$ is called the **absolute minimum** *of f on D, while $M = f(\vec{x}_M)$ is called the* **absolute maximum** *of f on D.*

.1.4 Partial Derivatives

Definition .1.28

The **first partial derivatives** *of a function $f : \mathbb{R}^n \to \mathbb{R}$ with respect to the variable x_i, $1 \leqslant i \leqslant n$, is the function*

$$f_i(x_1, ..., x_n) = \lim_{h \to 0} \frac{f(x_1, ..., x_i + h, ..., x_n) - f(x_1, ..., x_i, ..., x_n)}{h} \tag{.1.36}$$

provided the limit exists and f is defined in a neighbrohood of $(x_1, ..., x_n)$.

Notation

We often write

$$\frac{\partial}{\partial x_i} f(x_1, ..., x_n) = f_i(x_1, ..., x_n) = D_i f(x_1, ..., x_n) \tag{.1.37}$$

and

$$\left(\frac{\partial}{\partial x_i} f(\vec{x}) \right) \Big|_{\vec{x}_0} = f_i(\vec{x}_0) = D_i f(\vec{x}_0) \tag{.1.38}$$

Remark .1.15

If a function $f : J \subseteq \mathbb{R}^n \to \mathbb{R}$ has first partial derivatives at $\vec{x}_0$ in a region $D \subseteq \mathbb{R}^n$, then this defines n new functions

$$\frac{\partial f}{\partial x_i}\Big| : \mathbb{R} \to \mathbb{R} \tag{.1.39}$$

where we differentiate f with respect to x_i.

Definition .1.29

Given a function $f : \mathbb{R}^n \to \mathbb{R}$, the **gradient** *of f is the vector function*

$$\nabla f = grad(f) : \mathbb{R}^n \to \mathbb{R}^n \tag{.1.40}$$

such that

$$\nabla f(x_1, ..., x_n) = \left\langle \frac{\partial}{\partial x_1} f, ..., \frac{\partial}{\partial x_n} f \right\rangle \tag{.1.41}$$

where ∇ is the **del operator**

$$\nabla = \left[\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n} \right]^T \tag{.1.42}$$

Definition .1.30

Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined in a neighborhood of a point $\vec{x}_0$ such that its first partial derivatives exist at $\vec{x}_0$. Then by definition, the **linear approximation** *of f at $\vec{x}_0$ is the polynomial of degree 1, $L(\vec{x})$, that matches f at $\vec{x}_0$ and matches its partials at $\vec{x}_0$. In particular, we have that*

$$L(\vec{x}, \vec{x}_0) = f(\vec{x}_0) + \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(\vec{x}_0)(x_i - x_{i,0}) \tag{.1.43}$$

Definition .1.31

Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function defined around $\vec{x}_0 \in \mathbb{R}^n$ with first partials also defined. Let $L(\vec{x})$ be its linear approximation at $\vec{x}_0$. We say that f is **differentiable** *at $\vec{x}_0$ if the limit*

$$\lim_{\vec{x} \to \vec{x}_0} \frac{f(\vec{x}) - L(\vec{x})}{||\vec{x} - \vec{x}_0||} = 0 \tag{.1.44}$$

Definition .1.32

Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a multivariate function defined in a neighborhood of $\vec{x}_0 \in \mathbb{R}^n$ with first partial derivatives also defined. Then the **Jacobian matrix** *of f is defined to be*

$$Df(\vec{x}) = \begin{bmatrix} \partial_1 f_1(\vec{x}) & \partial_2 f_1(\vec{x}) & \dots & \partial_n f_1(\vec{x}) \\ \partial_1 f_2(\vec{x}) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \partial_1 f_m(\vec{x}) & \dots & \dots & \partial_n f_m(\vec{x}) \end{bmatrix} \tag{.1.45}$$

Then we say that f is differentiable at $\vec{x}_0$ if

$$\lim_{\vec{x} \to \vec{x}_0} \frac{||f(\vec{x}) - f(\vec{x}_0) - Df(\vec{x} - \vec{x}_0)||}{||\vec{x} - \vec{x}_0||} = 0 \tag{.1.46}$$

If all the first partial derivatives of f are continuous in a neighborhood of $\vec{x}_0$ then this holds.

Remark .1.16: Properties

For $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$:

1. *If f is differentiable at $\vec{x}_0$, then f is continuous at $\vec{x}_0$*
2. *If f and g are differentiable at $\vec{x}_0$, then $f \pm g, kf, fg$ ($m = 1$) are differentiable at $\vec{x}_0$.*
3. *If the partials of f are continuous in a neighborhood of $\vec{x}_0$, then f is differentiable at $\vec{x}_0$. The converse is not true in general.*

Definition .1.33

Consider $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, we have

$$\partial_i f(\vec{x}_0) = \lim_{t\to 0} \frac{f(\vec{x}_0 + t\hat{e}_i) - f(\vec{x}_0)}{t} \tag{.1.47}$$

so we can generalize this to define the **directional derivative** *in the direction of $\hat{u}$:*

$$\partial_{\hat{u}} f(\vec{x}_0) = \lim_{t\to 0} \frac{f(\vec{x}_0 + t\hat{u}) - f(\vec{x}_0)}{t} \tag{.1.48}$$

Theorem .1.4

If $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\vec{x}_0$, then the directional derivative of f at $\vec{x}_0$ exists in the direction of $\hat{u}$, and is equal to

$$\partial_{\hat{u}} f(\vec{x}_0) = \nabla f(\vec{x}_0) \cdot \hat{u} \tag{.1.49}$$

Theorem .1.5

Notice if $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\vec{x}_0$, and $\hat{u}$ is a unit vector, we have that

$$\partial_{\hat{u}} f(\vec{x}_0) = \nabla f(\vec{x}_0) \cdot \hat{u} = |\nabla f(\vec{x}_0)| \cos(\theta) \tag{.1.50}$$

with θ being the angle between $\nabla f(\vec{x}_0)$ and $\hat{u}$. As a result:

1. *The largest value of $\partial_{\hat{u}} f(\vec{x}_0)$ is equal to $|\nabla f(\vec{x}_0)|$, and occurs when $\hat{u}$ is in the same direction as the gradient*
2. *The smallest value of $\partial_{\hat{u}} f(\vec{x}_0)$ is equal to $-|\nabla f(\vec{x}_0)$ when $\hat{u}$ is in the same direction as $-\nabla f(\vec{x}_0)$*
3. *When $\hat{u}$ is perpendicular to $\nabla f(\vec{x}_0)$, the directional derivative is zero.*

Theorem .1.6: Chain Rule V1

Let $f : D_f \subseteq \mathbb{R}^n \to \mathbb{R}$ and $g : D_g \subseteq \mathbb{R} \to \mathbb{R}$ such that $g(t) = f(\vec{x}(t))$. Then

$$\frac{dg}{dt} = \nabla f(\vec{x}) \cdot \frac{d}{dt}\langle x_1, ..., x_n\rangle = \sum_{i=1}^{n} \partial_i f \frac{dx_i}{dt} \tag{.1.51}$$

Theorem .1.7: Chain Rule V2

Let $f : D_f \subseteq \mathbb{R}^n \to \mathbb{R}$ and $g : D_g \subseteq \mathbb{R}^m \to \mathbb{R}$ such that $g(\vec{x}) = f(y_1(\vec{x}), ..., y_n(\vec{x}))$. Then

$$\partial_i g(\vec{x}) = \nabla f(y_1(\vec{x}), ..., y_n(\vec{x})) \cdot \frac{\partial}{\partial x_i}\vec{y} \tag{.1.52}$$

for $\vec{y} = \langle y_1, ..., y_n\rangle$.

Theorem .1.8: Clairout's Theorem

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ has continous first and second partials on an open ball B_r. Then $f_{ij}(\vec{x}) = f_{ji}(\vec{x})$ for all $\vec{x} \in D$.

.1.5 Implicit Differentiation

Theorem .1.9: Implicit Function Theorem (Two variables)

Consider $F(x, y) = 0$. Let $(x_0, y_0) \in \mathbb{R}^2$ such that $F(x_0, y_0) = 0$, and suppose F's first partials are continuous in a neighborhood of (x_0, y_0). Then

1. *If $F_y(x_0, y_0) \neq 0$, then $F(x, y) = 0$ uniquely defines y as a continuously differentiable function of x in a neighborhood of x_0, and we have that*

$$\frac{dy}{dx} = -\frac{\partial_x F(x, y)}{\partial_y F(x, y)} \tag{.1.53}$$

2. *Similarly for $F_x(x_0, y_0) \neq 0$.*

Theorem .1.10: Implicit Function Theorem (n variables)

Consider $F(\vec{x}) = 0(\star)$, $\vec{x} = (x_1, ..., x_n)$. Let $\vec{a}$ satisfy $F(\vec{a}) = 0$, and suppose $F(\vec{x})$ has continuous first partial derivatives at and in a neighborhood of $\vec{a}$. Let β be one of the variables $\{x_1, ..., x_n\}$, and let $\vec{\alpha}$ be the rest. If $\partial_\beta F(\vec{a}) \neq 0$, then the equation $(\star)$ uniquely defines the variable β as a continuously differentiable function of $\vec{\alpha}$, and for $x_j \neq \beta$, we have

$$\partial_{x_j}\beta(\vec{\alpha}) = -\frac{\partial_{x_j} F(\vec{a})}{\partial_\beta F(\vec{a})} \tag{.1.54}$$

Theorem .2: Implicit Function Theorem (General)

Consider a system of n equations in $n + m$ variables

$$\begin{cases} F_{(1)}(x_1, ..., x_m, y_1, ..., y_n) = 0 \\ \vdots \\ F_{(n)}(x_1, ..., x_m, y_1, ..., y_n) = 0 \end{cases} \tag{.1.55}$$

and a point P_0 which satisfies the system. Suppose each $F_{(i)}$ is differentiable near P_0, so they have continuous first partial derivatives. Finally, suppose

$$\left.\frac{\partial(F_{(1)}, ..., F_{(n)})}{\partial(y_1, ..., y_n)}\right|_{P_0} \neq 0 \tag{.1.56}$$

Then the system defines $y_1, ..., y_n$ uniquely as continuously differentiable functions of

$x_1, ..., x_m$ in some neighborhood of P_0. Moreover,

$$\partial_{x_j} y_i = -\frac{\frac{\partial(F_{(1)},...,F_{(n)})}{\partial(y_1,...,x_j,...,y_n)}}{\frac{\partial(F_{(1)},...,F_{(n)})}{\partial(y_1,...,y_i,...,y_n)}} \tag{.1.57}$$

This formula is a consequence of Cramer's Rule applied to the n linear equations in n unknowns which is the system differentiated with respect to x_j.

.1.6 Differentials

Definition .1.34

Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function defined at and around a point $\vec{a}$. Given $\Delta\vec{x}$, of small magnitude,

$$\Delta f_{\vec{a}}(\Delta\vec{x}) = f(\vec{a} + \Delta\vec{x}) - f(\vec{a}) \tag{.1.58}$$

represents the change in the value of the function associated with the change $\Delta\vec{x}$ in $\vec{x}$ at $\vec{a}$. Then, we approximate this change with the **differential** *at $\vec{a}$ defined by*

$$df_{\vec{a}}(\Delta\vec{x}) = \nabla f(\vec{a}) \cdot \Delta\vec{x} \tag{.1.59}$$

If $\Delta\vec{x}$ is sufficiently small, these changes are approximately equal.

.1.7 Taylor Polynomials

Theorem .3: Taylor's Theorem (One Variable)

et $f(x)$ be a function wit $n+1$ continuous derivatives in the open interval (a, b). Let

$$T_n(x) := \sum_{i=0}^{n} \frac{f^{(i)}(c)(x-c)^i}{i!} \tag{.1.60}$$

be the **degree n Taylor polynomial** *of $f(x)$ centered at $x = c \in (a, b)$. Then, for any $x \in (a, b)$, there exists a number θ between c and x such that*

$$f(x) = T_n(x) + \frac{f^{(n+1)}(\theta)}{(n+1)!}(x-c)^{n+1} \tag{.1.61}$$

Definition .1.35: Two Variable Taylor Polynomial

Let $f(x,y)$ be a smooth function (continuous partial derivatives up to whatever degree needed) in a open set $D \subset \mathbb{R}^2$. The **degree n Taylor polynomial** *of $f(x,y)$ at a point $(a,b) \in D$, is the polynomial $T_n(x,y)$ of degree n that equals $f(x,y)$ and its first n partial derivatives at (a,b). It can be written as*

$$T_n(x,y) := \sum_{i=0}^{n} \frac{\left[(x-a)\partial_x + (y-b)\partial_y\right]^{(i)} f(a,b)}{i!} \tag{.1.62}$$

where $\left[(x-a)\partial_x + (y-b)\partial_y\right]^{(i)}$ is to be expanded as an algebraic expression and the products of ∂_x and ∂_y correspond to composition of operators.

Theorem .4: Taylor's Theorem (Two variables)

et $f(x,y)$ be a function with continuous partial derivatives up to $(n+1)$ in some neighborhood D of $(a,b) \in \mathbb{R}^2$. Let $T_n(x,y)$ be the degree n Taylor polynomial of $f(x,y)$ at (a,b). Then for any $(x,y) \in D$, there exists $(\alpha,\beta) \in D$ such that

$$f(x,y) = T_n(x,y) + \frac{\left[(x-a)\partial_x + (y-b)\partial_y\right]^{(n+1)} f(\alpha,\beta)}{(n+1)!} \tag{.1.63}$$

This is called the **Taylor formula/expansion of order n** *of f at (a,b).*

$$R_n(x,y) := \frac{\left[(x-a)\partial_x + (y-b)\partial_y\right]^{(n+1)} f(\alpha,\beta)}{(n+1)!} = f(x,y) - T_n(x,y) \tag{.1.64}$$

is called the **remainder** *of the expansion.*

Remark .1.17: I

all partial derivatives of order $(n+1)$ are bounded by some constant $M > 0$, then

$$|f(x,y) - T_n(x,y)| \leqslant \frac{M}{(n+1)!} \left[\sum_{j=0}^{n+1} \binom{n+1}{j} |x-a|^{n+1-j}|y-b|^{j}\right] \tag{.1.65}$$

Theorem .5: Taylor's Theorem (General)

et $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a function with continuous partial derivatives of order up to $m+1$ in a neighborhood D of $\vec{a} \in \mathbb{R}^n$. Then for all $\vec{x} \in D$ there exists $\vec{theta} \in D$ such that

$$f(\vec{x}) = T_m(\vec{x}) + R_m(\vec{x}, \vec{\theta}) \tag{.1.66}$$

where

$$T_m(\vec{x}) := \sum_{k=0}^{m} \frac{[(\vec{x}-\vec{a}) \cdot \nabla]^{(k)} f(\vec{a})}{k!} \tag{.1.67}$$

is the degree m Taylor polynomial of f at $\vec{a}$, and

$$R_m(\vec{x}, \vec{\theta}) := \frac{[(\vec{x}-\vec{a}) \cdot \nabla]^{(m+1)} f(\vec{\theta})}{(m+1)!} \tag{.1.68}$$

is the remainder. If all partial derivatives of f are continuous and there exists $r \in \mathbb{R}^+$ such that whenever $||\vec{x}-\vec{a}|| < r$ we have for all $t \in [0,1]$

$$\lim_{m\to\infty} R_m(\vec{x}, \vec{a} + t(\vec{x}-\vec{a})) = 0 \tag{.1.69}$$

Then we can represent $f(\vec{x})$ as the **Taylor series**

$$f(\vec{x}) = \sum_{n=0}^{\infty} \frac{[(\vec{x}-\vec{a}) \cdot \nabla]^{(n)} f(\vec{a})}{n!} \tag{.1.70}$$

.1.8 Local Extrema of Multivariate Functions

Definition .1.36

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of n variables defined in a neighborhood of $\vec{x}_0 \in \mathbb{R}^n$:

1. *We say f has a* **local maximum** *at $\vec{x}_0$ if there exists a neighborhood D of $\vec{x}_0$ for which f is defined and*
$$f(\vec{x}) \leqslant f(\vec{x}_0), \forall \vec{x} \in D \tag{.1.71}$$
2. *We say f has a* **local minimum** *at $\vec{x}_0$ if there exists a neighborhood D of $\vec{x}_0$ for which f is defined and*
$$f(\vec{x}) \geqslant f(\vec{x}_0), \forall \vec{x} \in D \tag{.1.72}$$

↳ *If f has a local maximum or minimum at $\vec{x}_0$, we say f has a* **local extremum** *at $\vec{x}_0$.*

Theorem .1.11: Fermat

If $f : \mathbb{R}^n \to \mathbb{R}$ has a local extremum at $\vec{x}_0$, then one of the following must hold:

1. *$\nabla f(\vec{x}_0) = \vec{0}$ when all first partials of f exist at $\vec{x}_0$*
2. *At least one of the first partials of f are not defined at $\vec{x}_0$*

Definition .1.37

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is defined in a neighborhood of $\vec{x}_0$. If one of the conditions of Fermat's Theorem is satisfied by $\vec{x}_0$, we say $\vec{x}_0$ is a **critical point** *of f.*

Remark .1.18

To determine if f has a local max or min at a critical point $\vec{x}_0$, study the sign of

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) \tag{.1.73}$$

for small $|\vec{h}|$. If it is always positive, f has a local minimum, if it is always negative f has a local maximum, and if it changes sign, f does not have a local extremum and in this case we say f has a **saddle point** *at $\vec{x}_0$.*

Theorem .6: Second Derivative Test

uppose $f(x,y)$ has continuous second partial derivatives in a neighborhood of a critical point (x_0,y_0). Define the **Hessian** *matrix of f at (x_0,y_0) to be*

$$H_f(x_0,y_0) := \begin{bmatrix} f_{xx}(x_0,y_0) & f_{xy}(x_0,y_0) \\ f_{yx}(x_0,y_0) & f_{yy}(x_0,y_0) \end{bmatrix} \tag{.1.74}$$

Let $\delta_1 = f_{xx}(x_0,y_0)$ and $\delta_2 = \det(H_f(x_0,y_0))$. Then

1. *If $\delta_1 > 0$ and $\delta_2 > 0$, then f has a local minimum at (x_0,y_0)*
2. *If $\delta_1 < 0$ and $\delta_2 > 0$, then f has a local maximum at (x_0,y_0)*
3. *If $\delta_2 \neq 0$ but neither case 1 nor case 2 hold, then f has a saddle point at (x_0,y_0).*
4. *If $\delta_2 = 0$ the test is inconclusive.*

Theorem .7: Second Derivative Test (general)

uppose $f : D \subset \mathbb{R}^n \to \mathbb{R}$ has continuous second partial derivatives in a neighborhood of a critical point $\vec{x}_0 \in D$. Define the **Hessian** *matrix of f at $\vec{x}_0$ to be*

$$H_f(\vec{x}_0) := \begin{bmatrix} f_{11}(\vec{x}_0) & f_{12}(\vec{x}_0) & \dots & f_{1n}(\vec{x}_0) \\ f_{21}(\vec{x}_0) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ f_{n1}(\vec{x}_0) & \dots & \dots & f_{nn}(\vec{x}_0) \end{bmatrix} \tag{.1.75}$$

Denote the the kth principal minor of $H_f(\vec{x}_0)$ by

$$\delta_k := \begin{vmatrix} f_{11}(\vec{x}_0) & f_{12}(\vec{x}_0) & \dots & f_{1k}(\vec{x}_0) \\ f_{21}(\vec{x}_0) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ f_{k1}(\vec{x}_0) & \dots & \dots & f_{kk}(\vec{x}_0) \end{vmatrix} \tag{.1.76}$$

. Then

1. *If for all $i \in \{1, 2, ..., n\}$, $\delta_i > 0$, then f has a local minimum at $\vec{x}_0$*
2. *If for all $i \in \{1, 2, ..., n\}$, $\delta_{2i-1} < 0$ and $\delta_{2i} > 0$, then f has a local maximum at $\vec{x}_0$*
3. *If $\delta_n = \det(H_f(\vec{x}_0) \neq 0$ but neither case 1 nor case 2 hold, then f has a saddle point at $\vec{x}_0$.*
4. *If $\delta_n = \det(H_f(\vec{x}_0) = 0$ the test is inconclusive.*

.1.9 Vector Fields

Definition .1.38

A **vector field** *is a vector function $\vec{F} : D \subset \mathbb{R}^n \to \mathbb{R}^n$. In the case of three variables we write*

$$\vec{F}(x, y, z) = \langle P(x, y, z), Q(x, y, z), R(x, y, z) \rangle \tag{.1.77}$$

Remark .1.19

A vector field $\vec{F} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ is said to be of class C^k for $k \in \mathbb{Z}^+$ in D if the first k partial derivatives of the component functions of $\vec{F}$ are continuous in D.

Definition .1.39: Conservative Fields

A vector field $\vec{F} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ is called **conservative** *in a region $E \subseteq D$ if there exists a scalar function $f : D_f \subset \mathbb{R}^n \to \mathbb{R}$ such that*

$$\vec{F}(\vec{x}) = \nabla f(\vec{x}), \forall \vec{x} \in E \tag{.1.78}$$

where f is called a **potential function** *of the vector field $\vec{F}$.*

Definition .1.40

Let $\vec{F} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ be a differentiable vector field. The **divergence** *of $\vec{F}$ is the scalar field*

$$\nabla \cdot \vec{F} = \sum_{i=1}^{n} \partial_i F_i \tag{.1.79}$$

where F_i are the component function of $\vec{F}$.

Definition .1.41

Let $\vec{F} : D \subset \mathbb{R}^3 \to \mathbb{R}^3$ be a differentiable vector field. The **curl** *of $\vec{F}$ is the vector field*

$$\nabla \times \vec{F} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \partial_x & \partial_y & \partial_z \\ P & Q & R \end{vmatrix} \tag{.1.80}$$

where F_i are the component function of $\vec{F}$.

Proposition .1.12: Properties of Divergence

If $\vec{F} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ and $\vec{G} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ vector fields and $f : D \subset \mathbb{R}^n \to \mathbb{R}$ is a scalar field, and $C_1, C_2 \in \mathbb{R}$, then

1. *(Linearity)* $\nabla \cdot (C_1\vec{F} + C_2\vec{G}) = C_1\nabla \cdot \vec{F} + C_2\nabla \cdot \vec{F}$

2. *(Product rule)* $\nabla \cdot (f\vec{F}) = \nabla f \cdot \vec{F} + f\nabla \cdot \vec{F}$

3. *(Laplacian)* $\Delta f = \nabla \cdot \nabla f = \partial^2_{xx} f + \partial^2_{yy} f + \partial^2_{zz} f$

Proposition .1.13: Properties of Divergence

If $\vec{F} : D \subset \mathbb{R}^3 \to \mathbb{R}^3$ and $\vec{G} : D \subset \mathbb{R}^3 \to \mathbb{R}^3$ vector fields and $f : D \subset \mathbb{R}^3 \to \mathbb{R}$ is a scalar field, that are all defined and differentiable in D. Let $C_1, C_2 \in \mathbb{R}$, then

1. *(Linearity)* $\nabla \times (C_1\vec{F} + C_2\vec{G}) = C_1\nabla \times \vec{F} + C_2\nabla \times \vec{F}$

2. *(Product rule)* $\nabla \times (f\vec{F}) = \nabla f \times \vec{F} + f(\nabla \times \vec{F})$

3. *(Conservative Property)* $\nabla \times (\nabla f) = 0$

4. $\nabla \times (\nabla \times \vec{F}) = \nabla(\nabla \cdot \vec{F}) - (\nabla \cdot \nabla)\vec{F}$ *(provided $\vec{F}$ has continuous second partial derivatives, and where $(\nabla \cdot \nabla)\vec{F} = (\Delta P, \Delta Q, \Delta R)$)*

Definition .1.42

Let $E \subseteq \mathbb{R}^n$

1. *We say that E is* **path connected** *if for any two points A and B in E if there exists a continuous function $f : [0, 1] \to E$ such that $f(0) = A$ and $f(1) = B$.*

2. *We say E is* **simply connected** *if E is path connected and any simple closed curve C that completely lies in E can be continuously deformed into a single point without leaving E.*

Theorem .1.14

Let $\vec{F}$ be a class C^1 in an open region $E \subseteq \mathbb{R}^3$. If $\vec{F}$ is conservative in E, then $\nabla \times \vec{F} = \vec{0}$ at every point of E. Moreover, if $\nabla \times \vec{F} = \vec{0}$ at every point of E and E is simply connected, then $\vec{F}$ is conservative in E.

.1.10 Line Integrals

Definition .1.43

Let C be a bounded continuous parametric curve in $\mathbb{R}^n$. Recall that C is a **smooth curve** *if it has a parameterization $\vec{r} : I \subset \mathbb{R} \to \mathbb{R}^n$ such that $\frac{d\vec{r}}{dt}$ is continuous and nonzero in I. We say C is a* **smooth arc** *if it is a smooth curve with finite parameter interval $I = [a, b]$.*

Definition .1.44

Given a smooth curve C with parameterization $\vec{r} : [a, b] \to \mathbb{R}^n$ we have

$$l_C = \int_C ds = \int_a^b \left| \frac{d\vec{r}}{dt} \right| dt \tag{.1.81}$$

In general, for a scalar field $f : \mathbb{R}^n \to \mathbb{R}$, we define the **line integral along C** *to be*

$$\int_C f(\vec{x}) ds = \int_a^b f(\vec{r}(t)) \left| \frac{d\vec{r}}{dt} \right| dt \tag{.1.82}$$

This definition is parameterization independent.

Definition .1.45

If $\vec{F}$ is a continuous vector field and C is an oriented smooth curve, then the **line integral of the tangential component of $\vec{F}$ along C** *is*

$$\int_C \vec{F} \cdot d\vec{r} = \int_C \vec{F} \cdot \hat{T} ds \tag{.1.83}$$

Definition .1.46

If C is a closed curve we also call this line integral the **circulation** *of $\vec{F}$ around C, and we* *denote it by*

$$\oint_C \vec{F} \cdot d\vec{r} \tag{.1.84}$$

Remark .1.20

A line integral over a piecewise smooth path is the sum of the line integrals over the individual smooth arcs

$$\int_{\bigcup_{i=1}^{n} C_i} ds = \sum_{i=1}^{n} \int_{C_i} ds \tag{.1.85}$$

.1.11 Line Integral Theorems

Theorem .8: Fundamental Theorem of Line Integrals

et C be a piecewise smooth parametric curve with initial point A and terminal point B. If $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is a scalar function with continuous first partial derivatives in an open region containing C, then

$$\int_C \nabla f \cdot d\vec{r} = f(B) - f(A) \tag{.1.86}$$

Corollary .1.15

If C is a piecewise smooth closed curve contained in a region D where the vector field $\vec{F} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is conservative, then

$$\oint_C \vec{F} \cdot d\vec{r} \tag{.1.87}$$

Definition .1.47

A vector field $\vec{F}$ is said to be **path-independent** *in a region Ω if for every pair of points A and B in Ω and every pair of piecewise smooth curves C_1 and C_2 with initial point A and terminal point B, we have*

$$\int_{C_1} \vec{F} \cdot d\vec{r} = \int_{C_2} \vec{F} \cdot d\vec{r} \tag{.1.88}$$

Theorem .1.16

Let D be an open connected domain in $\mathbb{R}^n$ and let $\vec{F}$ be a smooth vector field defined on D. Then the following properties are equivalent:

1. *$\vec{F}$ is conservative in D*

2. $\oint_C \vec{F} \cdot d\vec{r} = 0$ *for every piecewise smooth closed curve* $C \subset D$

3. $\vec{F}$ *is path independent in D.*

Theorem .9: Green's Theorem

et R be a closed region in the xy-plane whose boundary ∂R consists of a finite number of piecewise smooth simple closed curves that are positively oriented with respect to R. If $\vec{F}$ is a smooth vector field on R, then

$$\oint_{\partial R} \vec{F} \cdot d\vec{r} = \int\int_R (\nabla \times \vec{F}) \cdot \hat{k} dA \tag{.1.89}$$

In particular $\hat{k}$ is the unit normal field specifying the orientation of R, and ∂R is oriented such that its principal normal field $\vec{N}$ points away from the region and

$$\hat{N} = \hat{T} \times \hat{k} \tag{.1.90}$$

Remark .1.21

You don't need anything past this point yet.

Theorem .10: Plane Divergence Theorem

et R be a closed region in the xy-plane whose boundary ∂R consists of a finite number of piecewise smooth simple closed curves. Let $\vec{N}$ denote the unit outward (from R) normal field on C. If $\vec{F}$ is a smooth vector field on R, then

$$\oint_{\partial R} \vec{F} \cdot \hat{N} ds = \int\int_R \nabla \cdot \vec{F} dA \tag{.1.91}$$

Theorem .11: Stoke's Theorem

et $\mathcal{S}$ be a piecewise smooth oriented surface in 3-space having a unit normal field $\hat{N}$ and boundary C consisting of a finite number of piecewise smooth closed curves with orientation inherited from $\mathcal{S}$. If $\vec{F}$ is a smooth vector field defined on an open set containing $\mathcal{S}$, then

$$\oint_C \vec{F} \cdot d\vec{r} = \int\int_S (\nabla \times \vec{F}) \cdot \hat{N} dS \tag{.1.92}$$

.1.12 Surface Integrals

Definition .1.48

A **parametric surface** *in 3-space is a continuous function $\vec{r} : R \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ for some rectangle R*

$$R = \{(u, v) \in \mathbb{R}^2 : a \leqslant u \leqslant b, c \leqslant v \leqslant d\} \tag{.1.93}$$

in the uv-plane having values in 3-space:

$$\vec{r}(u, v) = \langle x(u, v), y(u, v), z(u, v)\rangle, (u, v) \in R \tag{.1.94}$$

Remark .1.22

If $\vec{r}$ is one-to-one the surface does not intersect itself. In this case $\vec{r}$ maps the boundary of R onto a curve in 3-space called the **boundary of the parametric surface***. A surface with no boundary is called a* **closed surface***.*

Definition .1.49

If a finite number of parametric surfaces are joined pairwise along their boundaries one obtains a **composite surface***, or just a surface thinking geometrically.*

Definition .1.50

A set $S \subseteq \mathbb{R}^3$ is a **smooth surface** *if any point $P \in S$ has a neighborhood N that is the domain of a smooth function $g : N \to \mathbb{R}$ satisfying*

1. $N \cap S = \{Q \in N : g(Q) = 0\}$
2. $\nabla g(Q) \neq \vec{0}$ *if* $Q \in N \cap S$

Remark .1.23

This means the surface has a unique tangent plane at any non-boundary point P.

Definition .1.51

If $\vec{r} : D \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ is a parameterization of a smooth surface S, the **normal vector** *to S at $\vec{r}(u, v)$ is*

$$\vec{n} = \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \tag{.1.95}$$

Definition .1.52

The **area element** *at $\vec{r}(u, v)$ on S is given by*

$$dS = \left|\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right| dudv \tag{.1.96}$$

Then if $f(\vec{r})$ is continuous on $\mathcal{S}$ and the domain of $\vec{r}$ is D in the uv-plane

$$\int\int_{S} f dS = \int\int_{D} f(\vec{r}(u,v)) \left| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right| dudv \tag{.1.97}$$

Definition .1.53

A smooth surface $\mathcal{S}$ in 3-space is said to be **orientable** *if there exists a unit vector field $\hat{N}$ defined on $\mathcal{S}$ that varies* **continuously** *over $\mathcal{S}$, and is everywhere normal to $\mathcal{S}$. Any such vector field $\hat{N}$ induces an orientation on $\mathcal{S}$. The side of $\mathcal{S}$ out of which $\hat{N}$ points is the* **positive side***, and the other side is the* **negative side***. An* **oriented surface** *is a smooth surface with a particular choice of orienting unit normal vector field $\hat{N}$.*

Remark .1.24

An oriented surface $\mathcal{S}$ **induces an orientation** *on any of its boundary curves $\mathcal{C}$; if we stand on the positive side of the surface $\mathcal{S}$ and walk around $\mathcal{C}$ in the direction of its orientation, then $\mathcal{S}$ will be on our left side.*

Definition .1.54

Given any continuous vector field $\vec{F}$, the **flux** *of $\vec{F}$ across the orientable surface $\mathcal{S}$ is the surface integral of the normal component of $\vec{F}$ over $\mathcal{S}$*

$$\int\int_{S} \vec{F} \cdot \hat{N} dS = \int\int_{S} \vec{F} \cdot d\vec{S} \tag{.1.98}$$

and when the surface is closed we write

$$\oiint_{S} \vec{F} \cdot \hat{N} dS = \oiint_{S} \vec{F} \cdot d\vec{S} \tag{.1.99}$$

Remark .1.25

If $\vec{r}(u,v)$ parametrizes $\mathcal{S}$ with domain D, we have normal

$$\vec{n} = \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \tag{.1.100}$$

and $dS = |\vec{n}| dudv$. Hence

$$d\vec{S} = \hat{N} dS = \pm \frac{\vec{n}}{|\vec{n}|} |\vec{n}| dudv = \pm \vec{n} dudv \tag{.1.101}$$

where the sign reflects the orientation of the surface and parameterization.

Theorem .12: Divergence Theorem

et D be a three dimensional domain bounded by piecewise smooth closed surfaces. Suppose its boundary $\mathcal{S}$ is an oriented closed surface with unit normal field $\hat{N}$ pointing out of D. If $\vec{F}$ is a smooth vector field defined on D, then

$$\oiint_S \vec{F} \cdot \hat{N} dS = \int\int\int_D \nabla \cdot \vec{F} dV \tag{.1.102}$$