Real Analysis and Multivariable Calculus: Graduate Level Problems and Solutions

Igor Yanovsky
Disclaimer: This handbook is intended to assist graduate students with qualifying examination preparation. Please be aware, however, that the handbook might contain, and almost certainly contains, typos as well as incorrect or inaccurate solutions. I cannot be made responsible for any inaccuracies contained in this handbook.
Contents

1 Countability 5

2 Unions, Intersections, and Topology of Sets 7

3 Sequences and Series 9

4 Notes 13

4.1 Least Upper Bound Property 13

5 Completeness 14

6 Compactness 16

7 Continuity 17

7.1 Continuity and Compactness 18

8 Sequences and Series of Functions 19

8.1 Pointwise and Uniform Convergence 19

8.2 Normed Vector Spaces 19

8.3 Equicontinuity 21

8.3.1 Arzela-Ascoli Theorem 21

9 Connectedness 21

9.1 Relative Topology 21

9.2 Connectedness 21

9.3 Path Connectedness 23

10 Baire Category Theorem 24

11 Integration 26

11.1 Riemann Integral 26

11.2 Existence of Riemann Integral 27

11.3 Fundamental Theorem of Calculus 27

12 Differentiation 30

12.1 \( \mathbb{R} \to \mathbb{R} \) 30

12.1.1 The Derivative of a Real Function 30

12.1.2 Rolle’s Theorem 30

12.1.3 Mean Value Theorem 30

12.2 \( \mathbb{R} \to \mathbb{R}^m \) 31

12.3 \( \mathbb{R}^n \to \mathbb{R}^m \) 31

12.3.1 Chain Rule 34

12.3.2 Mean Value Theorem 35

12.3.3 \( \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) \) 36

12.4 Taylor’s Theorem 37

12.5 Lagrange Multipliers 40
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.1</td>
<td>Contraction Mappings</td>
<td>41</td>
</tr>
<tr>
<td>13.2</td>
<td>Inverse Function Theorem</td>
<td>41</td>
</tr>
<tr>
<td>13.3</td>
<td>Implicit Function Theorem</td>
<td>44</td>
</tr>
<tr>
<td>13.4</td>
<td>Differentiation Under Integral Sign</td>
<td>46</td>
</tr>
</tbody>
</table>
1 Countability

The number of elements in $S$ is the **cardinality** of $S$.

$S$ and $T$ have the **same cardinality** ($S \cong T$) if there exists a bijection $f : S \to T$.

- $\text{card } S \leq \text{card } T$ if $\exists$ injective$^1 f : S \to T$.
- $\text{card } S \geq \text{card } T$ if $\exists$ surjective$^2 f : S \to T$.

$S$ is **countable** if $S$ is finite, or $S \cong \mathbb{N}$.

**Theorem.** $S, T \neq \emptyset$. $\exists$ injection $f : S \to T \iff \exists$ surjection $g : T \to S$.

**Theorem.** $\mathbb{Q}$ is **countable**.

**Proof.** Need to show that there is a bijection $f : \mathbb{N} \to \mathbb{Q}$.

Since $\mathbb{N} \subseteq \mathbb{Q}$, $\text{card } \mathbb{N} \leq \text{card } \mathbb{Q}$, and therefore, $\exists f$ that is injective.

To show $\text{card } \mathbb{N} \geq \text{card } \mathbb{Q}$, construct the following map. The set of all rational numbers can be displayed in a grid with rows $i = 1, 2, 3, \ldots$ and columns $j = 0, -1, 1, -2, 2, -3, 3, \ldots$.

Each $a_{ij}$, the $ij$th entry in a table, would be represented as $\frac{i}{j}$. Starting from $a_{11}$, and assigning it $n = 1$, move from each subsequent row diagonally left-down, updating $n$. This would give a map $g : \mathbb{N} \to \mathbb{Q}$, which will count all fractions, some of them more than once. Therefore, $\text{card } \mathbb{N} \geq \text{card } \mathbb{Q}$, and so $g$ is surjective. Thus, $\text{card } \mathbb{N} = \text{card } \mathbb{Q}$, and $\mathbb{Q}$ is countable.

**Theorem.** $\mathbb{R}$ is **not** countable.

**Proof.** It is enough to prove that $[0, 1) \subset \mathbb{R}$ is not countable. Suppose that the set of all real numbers between 0 and 1 is countable. Then we can list the decimal representations of these numbers (use the infinite expansions) as follows:

- $a_1 = 0.a_{11}a_{12}a_{13} \ldots a_{1n} \ldots$
- $a_2 = 0.a_{21}a_{22}a_{23} \ldots a_{2n} \ldots$
- $a_3 = 0.a_{31}a_{32}a_{33} \ldots a_{3n} \ldots$

and so on. We derive a contradiction by showing there is a number $x$ between 0 and 1 that is not on the list. For each positive integer $j$, we will choose $j$th digit after the decimal to be different than $a_{jj}$:

$x = 0.x_1x_2x_3 \ldots x_n \ldots$, where $x_j = 1$ if $a_{jj} \neq 1$, and $x_j = 2$ if $a_{jj} = 1$.

For each integer $j$, $x$ differs in the $j$th position from the $j$th number on the list, and therefore cannot be that number. Therefore, $x$ cannot be on the list. This means the list as we chose is not a bijection, and so the set of all real numbers is uncountable.

(Need to worry about not allowing 9 tails in decimal expansion: $0.399\ldots = 0.400\ldots$).

---

$^1$injective $= 1$-1: $f(s_1) = f(s_2) \Rightarrow s_1 = s_2$.

$^2$surjective $= \text{onto}$: $\forall t \in T, \exists s \in S, \text{ s.t. } f(s) = t$. 
Problem (F’01, #4). The set of all sequences whose elements are the digits 0 and 1 is not countable.

Let $S$ be the set of all binary sequences. We want to show that there does not exist a one-to-one mapping from the set $\mathbb{N}$ onto the set $S$.

Proof. 1) Let $A$ be a countable subset of $S$, and let $A$ consist of the sequences $s_1, s_2, \ldots$. We construct the sequence $s$ as follows. If the $n$th digit in $s_n$ is 1, let the $n$th digit of $s$ be 0, and vice versa. Then the sequence $s$ differs from every member of $A$ in at least one place; thus $s \notin A$. However, $s \in S$, so that $A$ is a proper\footnote{A is a proper subset of $B$ if every element of $A$ is an element of $B$, and there is an element of $B$ which is not in $A$.} subset of $S$.

Thus, every countable subset of $S$ is a proper subset of $S$, and therefore, $S$ is not countable. 

Proof. 2) Suppose there exists a $f : \mathbb{N} \to S$ that is injective. We can always exhibit an injective map $f : \mathbb{N} \to S$ by always picking a different sequence from the set of sequences that are already listed. (One way to do that is to choose a binary representation for each $n \in \mathbb{N}$).

Suppose $f : \mathbb{N} \to S$ is surjective. Then, all sequences in $S$ could be listed as $s_1, s_2, \ldots$. We construct the sequence $s$ as follows. If the $n$th digit in $s_n$ is 1, let the $n$th digit of $s$ be 0, and vice versa. Then the sequence $s$ differs from every member of the list. Therefore, $s$ is not on the list, and our assumption about $f$ being surjective is false. Thus, there does not exist $f : \mathbb{N} \to S$ surjective.

Theorem. $\text{card } (A) < \text{card } (P(A))$\footnote{Power set of a set $S$ is the set whose elements are all possible subsets of $S$, i.e., $S = \{1, 2\}, \ P(S) = 2^S = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}. \ |P(A)| = 2^{|A|}$, if $|A| = n$.}.

Proof. $\text{card } (A) \leq \text{card } (P(A))$, since $A$ can be injectively mapped to the set of one-element sets of $A$, which is a subset of $P(A)$.

We need to show there is no onto map between $A$ and $P(A)$. So we would like to find a thing in $P(A)$ which is not reached by $f$. In other words, we want to describe a subset of $A$ which cannot be of the form $f(a)$ for any $a \in A$.

Suppose $|A| = |P(A)|$. Then there is a 1-1 correspondence $f : A \to P(A)$. We obtain a contradiction to the fact that $f$ is onto by exhibiting a subset $X$ of $A$ such that $X \neq f(a)$ for any $a \in A$.

For every $a \in A$, either $a \in f(a)$, or $a \notin f(a)$. Let $X = \{a \in A : a \notin f(a)\}$.

Consider $a \in A$. If $a \notin f(a)$, then $a \notin X$, so $f(a) \neq X$.

If $a \notin f(a)$, then $a \in X$, so $f(a) \neq X$. Therefore, $X \neq f(a)$, $\forall a \in A$, a contradiction. Therefore, $\text{card } (A) < \text{card } (P(A))$.

Theorem. Suppose that $f : [0, 1] \to \mathbb{R}$ is an increasing function. Show that $f$ can have at most a countable number of discontinuities.

Proof. Let $E = \{x \in [0, 1] : f \text{ is discontinuous at } x\}$. Given any $x \in E$, we know that $\lim_{t \to x^-} f(t) < \lim_{t \to x^+} f(t)$ and, using this fact, we choose $r(x) \in \mathbb{Q}$ such that $\lim_{t \to x^-} f(t) < r(x) < \lim_{t \to x^+} f(t)$.

⇒ We have defined a $1 - 1$ function $r : E \to \mathbb{Q}$. 

$\Box$
2 Unions, Intersections, and Topology of Sets

Theorem. Let $E_{\alpha}$ be a collection of sets. Then

$$(\bigcup_{\alpha} E_{\alpha})^c = \bigcap_{\alpha} (E_{\alpha}^c).$$

Proof. Let $A = (\bigcup_{\alpha} E_{\alpha})^c$ and $B = (\bigcap_{\alpha} E_{\alpha})$. If $x \in A$, then $x \notin \bigcup_{\alpha} E_{\alpha}$, hence $x \notin E_{\alpha}$ for any $\alpha$, so that $x \in \bigcap_{\alpha} E^c_{\alpha}$. Thus $A \subseteq B$.

Conversely, if $x \in B$, then $x \in E^c_{\alpha}$ for every $\alpha$, hence $x \notin \bigcup_{\alpha} E_{\alpha}$, so that $x \in (\bigcup_{\alpha} E_{\alpha})^c$. Thus $B \subseteq A$. \hfill $\square$

Theorem. 

a) For any collection $G_{\alpha}$ of open sets, $\bigcup_{\alpha} G_{\alpha}$ is open.

b) For any collection $F_{\alpha}$ of closed sets, $\bigcap_{\alpha} F_{\alpha}$ is closed.

c) For any finite collection $G_1, \ldots, G_n$ of open sets, $\bigcap_{i=1}^n G_i$ is open.

d) For any finite collection $F_1, \ldots, F_n$ of closed sets, $\bigcup_{i=1}^n F_i$ is closed.

Proof. a) Put $G = \bigcup_{\alpha} G_{\alpha}$. If $x \in G$, then $x \in G_{\alpha}$ for some $\alpha$. Since $x$ is an interior point of $G_{\alpha}$, $x$ is also an interior point of $G$, and $G$ is open.

b) By theorem above,

$$(\bigcup_{\alpha} F_{\alpha})^c = \bigcap_{\alpha} (F_{\alpha}^c) \quad \Rightarrow \quad (\bigcap_{\alpha} F_{\alpha})^c = \bigcup_{\alpha} (F_{\alpha}^c),$$

and $F_{\alpha}^c$ is open. Hence a) implies that the right equation of (2.1) is open so that $\bigcap_{\alpha} F_{\alpha}$ is closed.

c) Put $H = \bigcap_{i=1}^n G_i$. For any $x \in H$, there exists neighborhoods $N_{r_i}(x)$ of $x$, such that $N_{r_i}(x) \subseteq G_i$ $(i = 1, \ldots, n)$. Put $r = \min(r_1, \ldots, r_n)$. Then $N_{r}(x) \subseteq G_i$ for $i = 1, \ldots, n$, so that $N_{r}(x) \subseteq H$, and $H$ is open.

d) By taking complements, d) follows from c): $(\bigcup_{i=1}^n F_i)^c = \bigcap_{i=1}^n (F_i^c)$. \hfill $\square$

Theorem. $S \subseteq M \subseteq X$. $S$ open relative to $M$ $\Leftrightarrow$ $\exists$ open $U \subseteq X$ such that $S = U \cap M$.

Proof. $\Rightarrow$ $S$ open relative to $M$. To each $x \in S$, $\exists r_x$, $|x - y| < r_x$, $y \in M \Rightarrow y \in S$. Define

$$U = \bigcup_{x \in S} N_{r_x}(x) \quad \Rightarrow \quad U \subseteq X \text{ open.}$$

It is clear that $S \subseteq U \cap M$. By our choice of $N_{r_x}(x)$, we have $N_{r_x}(x) \cap M \subseteq S$, $\forall x \in S$ so that $U \cap M \subseteq S$ $\Rightarrow$ $S = U \cap M$.

$\Leftarrow$ If $U$ is open in $X$ and $S = U \cap M$, every $x \in S$ has a neighborhood $N_{r_x}(x) \subseteq U$. Then $N_{r_x}(x) \cap M \subseteq S$ so that $S$ is open relative to $M$. \hfill $\square$
Theorem. $K \subset Y \subset X$. $K$ compact relative to $X$ $\iff$ $K$ compact relative to $Y$.

Proof. $\Rightarrow$ Suppose $K$ is compact relative to $X$, and let $\{V_\alpha\}$ be a collection of sets, open relative to $Y$, such that $K \subset \bigcup_\alpha V_\alpha$. By the above theorem, $\exists U_\alpha$ open relative to $X$, $\forall \alpha$; and since $K$ is compact relative to $X$, we have

$$K \subset U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}$$

Since $K \subset Y$ $\Rightarrow$ $K \subset V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}$. $\oplus$

$\Leftarrow$ Suppose $K$ is compact relative to $Y$. Let $\{U_\alpha\}$ be a collection of open subsets of $X$ which covers $K$, and put $V_\alpha = Y \bigcap U_\alpha$. Then $\oplus$ will hold for some choice of $\alpha_1, \ldots, \alpha_n$; and since $V_\alpha \subset U_\alpha$, $\oplus$ implies $\ominus$.

Theorem. Compact subsets of metric spaces are closed.

Proof. Let $K \subset X$ be a compact subset. We prove $K^c$ is open.

Suppose $x \in X$, $x \not\in K$. If $y \in K$, let $V_y$ and $W_y$ be neighborhoods of $x$ and $y$, respectively, of radius less than $\frac{1}{2}d(x,y)$. Since $K$ is compact, $K \subset W_{y_1} \cup \cdots \cup W_{y_n} = W$.

If $V = V_{y_1} \cap \cdots \cap V_{y_n}$, then $V$ is a neighborhood of $x$ which does not intersect $W$. Hence $V \subset K^c$, so that $x$ is an interior point of $K^c$.

Theorem. Closed subsets of compact sets are compact.

Proof. Suppose $F \subset K \subset X$, $F$ is closed (relative to $X$), and $K$ is compact. Let $\{V_\alpha\}$ be an open cover of $F$. If $F^c$ is adjoined to $\{V_\alpha\}$, we obtain an open cover $\Omega$ of $K$.

Since $K$ is compact, there is finite subcollection $\Phi$ of $\Omega$ which covers $K$, and hence $F$. If $F^c$ is a member of $\Phi$, we may remove it from $\Phi$ and still retain an open cover of $F$.

We have thus shown that a finite subcollection of $\{V_\alpha\}$ covers $F$.

Corollary. If $F$ is closed and $K$ is compact, then $F \cap K$ is compact.

Proof. $K$ is closed (by a theorem above), and thus $F \cap K$ is closed. Since $F \cap K \subset K$, the above theorem shows that $F \cap K$ is compact.

Theorem. If $\{K_\alpha\}$ is a collection of compact subsets of a metric space $X$ such that the intersection of every finite subcollection of $\{K_\alpha\}$ is nonempty, then $\bigcap K_\alpha$ is nonempty.

Proof. Fix a member $K_1$ of $\{K_\alpha\}$. Assume that no point of $K_1$ belongs to every $K_\alpha$. Then the sets $K_\alpha^c$ form an open cover of $K_1$; and since $K_1$ is compact, $K_1 \subset K_{\alpha_1}^c \cup \cdots \cup K_{\alpha_n}^c$. But this means that $K_1 \cap K_{\alpha_1} \cap \cdots \cap K_{\alpha_n}$ is empty. $\Rightarrow$ contradiction.
3 Sequences and Series

A sequence \( \{ p_n \} \) converges to \( p \in X \) if:
\[
\forall \epsilon > 0, \exists N: \forall n \geq N \quad |p_n - p| < \epsilon \quad \iff \quad \lim_{n \to \infty} p_n = p.
\]
A sequence \( \{ p_n \} \) is Cauchy if:
\[
\forall \epsilon > 0, \exists N: \forall n, m \geq N \quad |p_n - p_m| < \epsilon.
\]

Cauchy criterion: A sequence converges in \( \mathbb{R}^k \) \( \iff \) it is a Cauchy sequence.
Cauchy criterion: \( \sum a_n \) converges \( \iff \forall \epsilon > 0, \exists N: |\sum_{k=n}^{m} a_k| \leq \epsilon \) if \( m \geq n \geq N \).

The series \( \sum a_n \) is said to converge absolutely if \( \sum |a_n| \) converges.

Problem (S’03, #2). If \( a_1, a_2, a_3, \ldots \) is a sequence of real numbers with \( \sum_{j=1}^{\infty} |a_j| < \infty \), then \( \lim_{N \to \infty} \sum_{j=1}^{N} a_j \) exists. ( \( \sum a_n \) converges absolutely \( \Rightarrow \) \( \sum a_n \) converges).

Proof. If \( s_n = \sum_{j=1}^{n} a_j \) is a partial sum, then for \( m \leq n \) we have
\[
|s_n - s_m| = |\sum_{j=m+1}^{n} a_j| \leq \sum_{j=m+1}^{n} |a_j|
\]
Since \( \sum |a_j| \) converges, given \( \epsilon > 0, \exists N, \) s.t. \( m, n \geq N \) (\( m \leq n \)), then \( \sum_{j=m}^{n} |a_j| < \epsilon \).
Thus \( \{s_n\} \) is a Cauchy sequence in \( \mathbb{R} \), and converges. \( \Rightarrow \lim_{N \to \infty} \sum_{j=1}^{N} a_j \) exists. \( \square \)

Problem (F’01, #2). Let \( \mathbb{N} \) denote the positive integers, let \( a_n = (-1)^{n+1} \frac{1}{n} \), and let \( \alpha \) be any real number. Prove there is a one-to-one and onto mapping \( \sigma: \mathbb{N} \to \mathbb{N} \) such that
\[
\sum_{n=1}^{\infty} a_{\sigma(n)} = \alpha.
\]

Proof. \( a_n = (-1)^{n+1} \frac{1}{n} \) \( \Rightarrow \) \( a_1 = -1, a_2 = \frac{1}{2}, a_3 = -\frac{1}{3}, a_4 = \frac{1}{4}, \ldots \)
\( a_{2n} > 0, \quad a_{2n-1} < 0, \quad n = 1, 2, \ldots \)
\( \sum (\text{positive terms}) = \sum_{n=1}^{\infty} a_{2n} = \sum_{n=1}^{\infty} \frac{1}{2n} = \frac{1}{2} \sum_{n=1}^{\infty} \frac{1}{n} \) diverges, and \( \lim_{n \to \infty} \frac{1}{2n} = 0. \)
\( \sum (\text{negative terms}) = \sum_{n=1}^{\infty} a_{2n-1} = -\sum_{n=1}^{\infty} \frac{1}{2n-1} \) diverges by comparison with \( -\sum_{n=1}^{\infty} \frac{1}{2n} \) and \( \lim_{n \to \infty} \frac{1}{2n-1} = 0. \)
Claim: \( \forall \alpha \in \mathbb{R}, \) there is a one-to-one and onto mapping \( \sigma: \mathbb{N} \to \mathbb{N} \) such that \( \sum_{n=1}^{\infty} a_{\sigma(n)} = \alpha, \) where \( \sigma(n) \) is the rearrangement of indices of the original series.
Given \( \alpha, \) choose positive terms in sequential order until their sum exceeds \( \alpha. \) At this switch point, choose negative terms until their sum is less than \( \alpha. \) Repeat the process.
Note: This process never stops because no matter how many positive and negative terms are taken, there are still infinitely many both positive and negative terms left; the sum of positive terms is \( \infty, \) the sum of negative terms is \( -\infty. \)
Let the sum of terms at the \( N \)th step be denoted by \( S_N, \)
\( S_N = \sum_{n=1}^{N} a_{\sigma(n)}. \) At switch point, \( |\alpha - S_N| \) is bounded by the size of the term added:
\[
|\alpha - S_N| \to 0, \quad N \to \infty
\]
All terms \( \{a_n\} \) will eventually be added to the sum (\( \sigma: \mathbb{N} \to \mathbb{N} \) is surjective (onto)) at different steps (\( \sigma: \mathbb{N} \to \mathbb{N} \) is injective (1-1)). \( \square \)
Root Test. Given $\sum a_n$, put $\alpha = \limsup_{n \to \infty} \sqrt[n]{|a_n|}$. Then
a) if $\alpha < 1$, $\sum a_n$ converges;
b) if $\alpha > 1$, $\sum a_n$ diverges;
c) if $\alpha = 1$, the test gives no information.

Ratio Test. The series $\sum a_n$ converges if $\limsup_{n \to \infty} \frac{|a_{n+1}|}{a_n} < 1$,
diverges if $\frac{|a_{n+1}|}{a_n} \geq 1$ for all $n \geq n_0$, where $n_0$ is some fixed integer.

Alternating Series. Suppose
(a) $|a_1| \geq |a_2| \geq \ldots$ ;
(b) $a_{2m-1} \geq 0$, $a_{2m} \leq 0$, $m = 1, 2, 3, \ldots$ ;
(c) $\lim_{n \to \infty} a_n = 0$.
Then $\sum a_n$ converges.

Geometric Series. $|x| < 1$ :

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}.$$  

Proof : 

\begin{align*}
S_n &= 1 + x + x^2 + \cdots + x^n \quad \text{①} \\
xS_n &= x + x^2 + x^3 + \cdots + x^{n+1} \quad \text{②} \\
\text{①} - \text{②} &= (1-x)S_n = 1 - x^{n+1} \\
\Rightarrow S_n &= \sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1-x} 
\end{align*}
Problem (F’02, #4). By integrating the series
\[
\frac{1}{1 + x^2} = 1 - x^2 + x^4 - x^6 + x^8 \ldots
\]
prove that
\[
\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} \ldots
\]

Justify carefully all the steps (especially taking the limit as \(x \to 1\) from below).

Proof. Geometric Series:
\[
\frac{1}{1 + x^2} = 1 - x^2 + x^4 - x^6 + x^8 \ldots ; \quad |x| < 1.
\]
\[
\int \frac{1}{1 + x^2} \, dx = \tan^{-1}(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} \ldots
\]
\[
\int_0^1 \frac{1}{1 + x^2} \, dx = \lim_{\epsilon \to 0} \left[ \int_0^{1-\epsilon} \frac{s}{1 + x^2} \right] \, dx = \lim_{\epsilon \to 0} \int_0^{1-\epsilon} S_N(x) \, dx
\]
converges uniformly for \(|x| \leq 1 - \epsilon\)

Above calculations show \(\tan^{-1} x = \sum_{n=1}^{\infty} (-1)^n x^{2n+1}, \quad |x| < 1\).

Alternating series test \(\Rightarrow\) right side converges.
\[
\frac{\pi}{4} = \sum_{n=1}^{\infty} \frac{(-1)^n}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} \ldots
\]

Logarithm. $|x| < 1$:

$$
\frac{1}{1 + x} = 1 - x + x^2 - x^3 + \cdots = \sum_{n=0}^{\infty} (-1)^n x^n
$$

$$
\log(1 + x) = \int_0^x \frac{dt}{1 + t} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}
$$

$\log(1 + x)$ is not valid for $|x| > 1$, or $x < -1$. Claim: If $x = 1$, the series converges to $\log 2$. Proof: $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}$ is uniformly convergent for $x \in [0, 1]$, since the sum of any number of consecutive terms starting with the $n^{th}$ has absolute value at most $\frac{x^n}{n} \leq \frac{1}{n}$, since for $0 < x < 1$ we have alternating series.

Binomial Series. $|x| < 1$:

$$(1 + x)^\alpha = 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n = 1 + \alpha x + \left(\frac{\alpha}{2}\right) x^2 + \cdots + \left(\frac{\alpha}{n}\right) x^n + \cdots
$$

$$
= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!} x^3 + \cdots
$$

Problem (F’03, #3). The sequence $a_1, a_2, \ldots$ with $a_n = \left(1 + \frac{1}{n}\right)^n$ converges as $n \to \infty$.

Proof. By Binomial Series Theorem with $\alpha = n$, $x = \frac{1}{n}$, we get:

$$
\left(1 + \frac{1}{n}\right)^n = 1 + n \cdot \frac{1}{n} + \frac{n(n-1)}{2!} \cdot \frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!} \cdot \frac{1}{n^3} + \cdots + \frac{n(n-1)(n-2)\cdots1}{n!} \cdot \frac{1}{n^n}
$$

$$
\leq 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} = \sum_{k=0}^{n} \frac{1}{k!} \to e, \ \text{as} \ n \to \infty.
$$
4 Notes

4.1 Least Upper Bound Property

An ordered set \( S \) is said to have the least upper bound\(^5\) property if:

\[ E \subset S, \ E \text{ is not empty, and } E \text{ is bounded above, then } \sup E \text{ exists in } S. \]

Completeness axiom: If \( E \) is a nonempty subset of \( \mathbb{R} \) that is bounded above, then \( E \) has a least upper bound.

Problem (F’02, #2). Show why the Least Upper Bound Property (every set bounded above has a least upper bound) implies the Cauchy Completeness Property (every Cauchy sequence has a limit) of the real numbers.

Proof. Suppose \( \{x_n\} \) Cauchy. The problem is to show that \( \{x_n\} \) converges.

We first show that \( \{x_n\} \) is bounded. Fix \( \epsilon > 0 \) and let \( N \) be such that \( |x_n - x_m| < \epsilon \) if \( n, m > N \). Then for any fixed \( n > N \), the entire sequence is contained in the closed ball of center \( x_n \) and radius \( \max\{d(x_n, x_1), d(x_n, x_2), \ldots, d(x_n, x_N), \epsilon\} \). Thus \( \{x_n\} \) is bounded.

Define \( z_n = \sup\{x_k\}_{k \geq n} \). Since \( \{x_n\} \) is bounded, each \( z_n \) is a finite real number and is bounded above in absolute value by \( M \). If \( m > n \), then \( z_m \) is obtained by taking the sup of a smaller set than is \( z_n \); hence \( \{z_n\} \) is decreasing. By the greatest lower bound property, \( Z = \{z_n|n \in \mathbb{N}\} \) has an infimum. Let \( x = \inf Z \). We claim that \( x_n \to x \).\(^6\)

For each \( \epsilon > 0 \) there is a corresponding integer \( N \) such that \( x \leq z_N \leq x + \epsilon \). Since \( \{x_n\} \) is Cauchy, by taking a larger \( N \) if necessary, we know that \( k \geq N \implies x_k \in [x_N - \epsilon, x_N + \epsilon] \).

It follows that \( z_N \in [x_N - \epsilon, x_N + \epsilon] \). Hence for \( k \geq N \),

\[ |x_k - x| \leq |x_k - x_N| + |x_N - z_N| + |z_N - x| \leq \epsilon + \epsilon + \epsilon = 3\epsilon. \]

---

\(^5\)least upper bound of \( E \equiv \sup E \).

\(^6\)Idea: Since \( \{x_n\} \) is Cauchy, the terms of this sequence would approach one another. \( \{z_n\} \) also approaches \( \{x_n\} \). Since \( z_n \to x \), \( \{z_n\} \) approaches \( x \). It follows that \( \{x_n\} \) approaches \( x \).
5 Completeness

A metric space \(X\) is complete if every Cauchy sequence of elements of \(X\) converges to an element of \(X\).

**Lemma.** A convergent sequence is a Cauchy sequence.

**Proof.** \(x_n \to x\) means \(\forall \epsilon > 0, \exists N, \text{ such that } \forall n \geq N, |x - x_n| < \epsilon\). Hence
\[
|x_n - x_m| \leq |x_n - x| + |x - x_m| \leq 2\epsilon
\]
when \(n, m \geq N\). Thus, \(x_n\) is a Cauchy sequence.

**Lemma.** If \(x_n\) is Cauchy, then \(x_n\) is bounded.

**Proof.** Let \(\epsilon > 0\). Since \(x_n\) is Cauchy \(\Rightarrow\) choose \(k > 0\) so large that \(|x_n - x_m| < \epsilon\) whenever \(n, m \geq k\). Since \(x_n \to x\) \(\Rightarrow\) choose \(l > 0\) so large (i.e. \(N_l\) large) that \(|x_n - x| < \epsilon\) whenever \(j \geq l\). Set \(N = \max(k, N_l)\). If \(m, n_j > N\), then
\[
|x_m - x| \leq |x_m - x_n| + |x_n - x| < \epsilon + \epsilon = 2\epsilon.
\]

**Theorem.** \([a, b]\) is complete.

**Proof.** Let \(x_n\) be a Cauchy sequence in \([a, b]\). Let \(x_{n_k}\) be a monotone subsequence. Since \(a \leq x_{n_k} \leq b\), \(x_{n_k}\) converges (by the Least Upper Bound). \(\Rightarrow\) \(x_{n_k} \to c\). Since \([a, b]\) is closed, \(c \in [a, b]\). \(\Rightarrow\) Any \(x_n\) that is Cauchy in \([a, b]\), converges in \([a, b]\). \(\Rightarrow\) \([a, b]\) is complete.

The above theorem is a specific case of the following Lemma:

**Theorem.** Let \(x_n\) be a Cauchy sequence in a compact metric space \(X\). Then \(x_n\) converges to some point of \(X\).

**Proof.** Since \(X\) is (sequentially) compact, then for any sequence \(x_n \in X\), there is a subsequence \(x_{n_k} \to c, c \in X\). Using the above theorem \((x_n\) Cauchy and \(x_{n_k} \to c \Rightarrow x_n \to c\), we see that \(x_n \to c \in X\).

**Theorem.** \(\mathbb{R}\) is complete.

**Proof.** Let \(x_n\) be a Cauchy sequence in \(\mathbb{R}\). \(x_n\) is bounded (by the Lemma above). \(\Rightarrow\) \(\{x_n\} \subseteq [a, b]\), and see above.

A direct consequence of the above theorem is the following: In \(\mathbb{R}^k\), every Cauchy sequence converges.

**Theorem.** \([0, 1]\) is not complete.
**Bolzano-Weierstrass.** Every bounded, infinite subset \( S \subset \mathbb{R} \) has a limit point.

**Proof.** If \( I_0 \) is a closed interval containing \( S \), denote by \( I_1 \) one of the closed half-intervals of \( I_0 \) that contains infinitely many points of \( S \). Continuing in this way, we define a nested sequence of intervals \( \{I_n\} \), each of which contains infinitely many points of \( S \). If \( c = \bigcap_{n=1}^{\infty} I_n \), then it is clear that \( c \) is a limit point of \( S \). \( \square \)

**Lemma.** Every bounded sequence of \( \mathbb{R} \) has a convergent subsequence.

**Proof.** If a sequence \( x_n \) contains only finitely many distinct points, the conclusion is trivial and obvious. Otherwise we are dealing with a bounded infinite set, to which the Bolzano-Weierstrass theorem applies, giving us a limit point \( x \). If, for each integer \( k \geq 1 \), \( x_{n_k} \) is a point of the sequence such that \( |x_{n_k} - x| \leq 1/k \), then it is clear that \( x_{n_k} \) is a convergent subsequence. \( \square \)

**Theorem.** A closed subspace \( Y \) of a complete metric space \( X \) is complete.

**Proof.** Let \( y_n \) be a Cauchy sequence in \( Y \). Then \( y_n \) is also a Cauchy sequence in \( X \). Since \( X \) is complete, \( \exists x \in X \) such that \( y_n \rightarrow x \) in \( X \). Since \( Y \) is closed, \( x \in Y \). Consequently, \( y_n \rightarrow x \) in \( Y \). \( \square \)

**Theorem.** A complete subspace \( Y \) of a metric space \( X \) is closed in \( X \).

**Proof.** Suppose \( x \in X \) is a limit point of \( Y \). \( \exists y_n \) in \( Y \) that converges to \( x \). \( y_n \) is a Cauchy sequence in \( X \); hence it is also a Cauchy sequence in \( Y \). Since \( Y \) is complete, \( y_n \rightarrow y \in Y \). Since limits of sequences are unique, \( y = x \) and \( x \) belongs to \( Y \). Hence \( Y \) is closed. \( \square \)

**Theorem.** \( V \) is a normed space. If we have the following implication

\[
\sum_{n=1}^{\infty} \|v_n\| < \infty \quad \Rightarrow \quad \sum_{n=1}^{\infty} v_n < \infty,
\]

then \( V \) is complete (\( V \) is a Banach space).

**Proof.** Say \( v_n \) is Cauchy. Choose \( n_k \) such that \( m, n \geq n_k \) \( \Rightarrow \) \( \|v_m - v_n\| < 2^{-k} \). We may assume that \( n_1 < n_2 < \ldots \)

\[
\|v_{n_{k+1}} - v_{n_k}\| < 2^{-k} \quad \Rightarrow \quad \sum_{k=1}^{\infty} \|v_{n_{k+1}} - v_{n_k}\| < \infty.
\]

It follows that \( \sum_{n=1}^{\infty} (v_{n_{k+1}} - v_{n_k}) < \infty \).

\[
s - s_K = \sum_{k=1}^{K} (v_{n_{k+1}} - v_{n_k}) = v_{n_{K+1}} - v_{n_1} \quad \Rightarrow \quad v_{n_{k+1}} \text{ converges} \quad \Rightarrow \quad v_n \text{ converges}.
\]
6 Compactness

$M$ is (sequentially) **compact** if for any sequence $x_n \in M$, there is a subsequence $x_{n_k} \to c$, $c \in M$.

$M$ is (topologically) **compact** if any open cover of $M$, $M \subseteq \bigcup G_\alpha$, $G_\alpha$ open, contains a finite subcover.

**Problem (W’02, #2).** $[a, b]$ is **compact**.

proof. Let $x_n$ be a sequence in $[a, b]$. Let $x_{n_k}$ be a monotone subsequence $\Rightarrow a \leq x_{n_k} \leq b \Rightarrow x_{n_k} \to c$. Since $[a, b]$ is closed, $c \in [a, b]$. $\Rightarrow [a, b]$ is (sequentially) compact. □

**Lemma.** If $M$ is compact, every open cover of $M$ has a countable subcover.

**Theorem.** If $M$ is sequentially compact, then it is topologically compact.

proof. Say that $M \subseteq G_1 \bigcup G_2 \bigcup \cdots$ has a countable subcover. Need to show that there is a finite subcover, i.e. $M \subseteq \bigcup_{k=1}^n G_k$ for some $n$.

Suppose that fails for every $n$; then for every $n = 1, 2, \ldots$, there would exist

$$x_n \in M \setminus \bigcup_{k=1}^n G_k.$$ 

That sequence would have a convergent subsequence $\{x_{n_k}\}$. Let $x$ be its limit, $x_{n_k} \to x$. Then $x$ would be contained in $G_m$ for some $m$, and thus $x_{n_k} \in G_m$ for all $n_k$ sufficiently large, which is impossible for $n_k > m$ (since $x_{n_k} \in M \setminus G_1 \bigcup \cdots \bigcup G_{n_k}$).

We have reached a contradiction. So there must be a finite subcovering. □

**Problem (S’02, #3).** If $M$ is topologically compact, then it is sequentially compact.

proof. Let $x_n \in M$ and $E$ be the range of $\{x_n\}$. If $E$ is finite, then there is $x \in E$ and a sequence $\{n_i\}$, with $n_1 < n_2 < \ldots$, such that

$$x_{n_1} = x_{n_2} = \ldots = x$$

The subsequence $\{x_{n_i}\}$ converges to $x$.

If $E$ is infinite, $E$ has a limit point $x$ in $M$ (as an infinite subset of a compact set). Every neighborhood of $x$ contains infinitely many points of $M$. For each $k$, $B_{\frac{1}{k}}(x)$ contains infinitely many $x_n$’s. Select one and call it $x_{n_k}$, such that, $n_k > n_{k-1} > \ldots$. We have a subsequence $\{x_{n_k}\}$ so that $d(x, x_{n_k}) < \frac{1}{k} \to 0. \Rightarrow x_{n_k} \to x$. □

**Problem (F’02, #1).** Let $K$ be a compact subset and $F$ be a closed subset in the metric space $X$. Suppose $K \cap F = \phi$. Prove that

$$0 < \inf\{d(x, y) : x \in K, y \in F\}.$$ 

proof. Given $x \in K, x \notin F$, $d_x = d(x, F) > 0$. Then, the ball centered at that $x$ with radius $d_x/2$, i.e. $B_{\frac{d_x}{2}}(x)$, satisfies $B_{\frac{d_x}{2}}(x) \cap F = \phi$. Since $x$ was taken arbitrary, this is true $\forall x \in K, x \notin F$.

$K \subseteq \bigcup_{x \in K} B_{\frac{d_x}{2}}(x)$. Since $K$ is compact, $\exists x_1, \ldots, x_n \in K$, $n < \infty$, such that $K \subseteq \bigcup_{k=1}^n B_{\frac{d_x}{2}}(x_k)$, and $B_{\frac{d_x}{2}}(x_k) \cap F = \phi$. Since $\min_k\{d_{x_k}\} > 0$, we have

$$0 < \inf\{d(x, y) : x \in K, y \in F\}. \square$$
7 Continuity

Limits of Functions: \( \lim_{x \to p} f(x) = q \) if:

\( \forall \epsilon > 0, \exists \delta \text{ such that } \forall x \in E \ 0 < |x - p| < \delta \implies |f(x) - q| < \epsilon. \)

A function \( f \) is continuous at \( p \): \( \lim_{x \to p} f(x) = f(p) \) if:

\( \forall \epsilon > 0, \exists \delta \text{ such that } \forall x \in E \ |x - p| < \delta \implies |f(x) - f(p)| < \epsilon. \)

Negation: \( f \) is not continuous at \( p \) if:

\( \exists \epsilon > 0, \forall \delta \text{ such that } \exists x \in E \ |x - p| < \delta \implies |f(x) - f(p)| \geq \epsilon. \)

A function \( f \) is uniformly continuous on \( X \) if:

\( \forall \epsilon > 0, \exists \delta > 0 \text{ such that } \forall x, z \in X \ |x - z| < \delta \implies |f(x) - f(z)| < \epsilon. \)

Negation: \( f \) is not uniformly continuous on \( X \) if:

\( \exists \epsilon > 0, \forall \delta > 0 \text{ such that } \exists x, z \in X \ |x - z| < \delta \implies |f(x) - f(z)| \geq \epsilon. \)

Examples: \( f(x) = \frac{1}{x} \) on \((0, 1]\) and \( f(x) = x^2 \) on \([1, \infty)\) are not uniformly continuous.

**Theorem.** \( f : X \to Y \) is continuous \( \iff \) \( f^{-1}(V) \) is open in \( X \) for every open set \( V \) in \( Y \).

**Proof.** 

Suppose \( f \) is continuous on \( X \). Let \( V \) be an open set in \( Y \). We have to show that every point of \( f^{-1}(V) \) is an interior point of \( f^{-1}(V) \). Let \( x \in f^{-1}(V) \).

Choose \( \epsilon > 0 \) such that \( B_{\epsilon}(f(x)) \subset V \). Since \( f \) is continuous, \( \exists \delta > 0 \text{ such that } f(B_{\delta}(x)) \subset B_{\epsilon}(f(x)) \subset V \). Hence, \( B_{\delta}(x) \subset f^{-1}(V) \). Since \( f^{-1}(V) \) contains an open ball about each of its points, \( f^{-1}(V) \) is open.

\( \iff \) Suppose \( f^{-1}(V) \) is open in \( X \) for every open set \( V \) in \( Y \). Let \( x \in X \) and let \( \epsilon > 0 \). Then \( f^{-1}(B_{\epsilon}(f(x))) \) is open in \( X \). Hence, \( \exists \delta > 0 \text{ such that } B_{\delta}(x) \subset f^{-1}(B_{\epsilon}(f(x))). \)

Applying \( f \), we obtain \( f(B_{\delta}(x)) \subset B_{\epsilon}(f(x)) \), and so \( f \) is continuous. \( \square \)

**Problem (S'02, #4; S'03, #1).** A function \( f : (0, 1) \to \mathbb{R} \) is the restriction to \((0, 1)\) of a continuous function \( F : [0, 1] \to \mathbb{R} \) \( \iff f \) is uniformly continuous on \((0, 1)\).

**Proof.** \( \Leftarrow \) We show that if \( f : (0, 1) \to \mathbb{R} \) is uniformly continuous, then there is a continuous \( F : [0, 1] \to \mathbb{R} \) with \( F(x) = f(x) \) for all \( x \in (0, 1) \).

Let \( x_n \) be a sequence in \((0, 1)\) converging to 0. Since \( f \) is uniformly continuous, given \( \epsilon > 0, \exists \delta, \text{ s.t. } |x - y| < \delta \implies |f(x) - f(y)| < \epsilon. \) Therefore, we have

\[ |f(x_n) - f(x_m)| < \epsilon \]

for \( n, m \) large enough. \( f(x_n) \) is a Cauchy sequence, so it converges to some \( \xi \). Define \( F(0) = \lim_{n \to \infty} f(x_n) = \xi \). We want to show that this limit is well defined. Let \( y_n \) be another sequence, s.t. \( y_n \to 0 \), so \( f(y_n) \) is Cauchy by the same argument. Since the sequence \((f(x_1), f(y_1), f(x_2), f(y_2), \ldots)\) is Cauchy by still the same argument, and that there is a subsequence \( f(x_n) \to \xi \), then the entire sequence converges to \( \xi \). Thus, \( F(0) = \lim_{n \to \infty} f(x_n) = \xi \) is well defined.

By the same set of arguments, \( F(1) = \eta \). The function \( F : [0, 1] \to \mathbb{R} \) given by

\[
F(x) = \begin{cases} 
  f(x) & \text{for } x \in (0, 1), \\
  \xi & \text{for } x = 0, \\
  \eta & \text{for } x = 1.
\end{cases}
\]

is the unique continuous extension of \( f \) to \([0, 1]\).

\( \Rightarrow \) \( F : [0, 1] \to \mathbb{R} \) is continuous, and \([0, 1]\) is compact. Therefore, \( F \) is uniformly continuous on \([0, 1]\). Thus, \( f = F|_{(0,1)} : (0, 1) \to \mathbb{R} \) is uniformly continuous. \( \square \)

---

\({}^7\) Gamelin, Green, p. 26; Edwards, p. 51.

\({}^8\) \( f \) continuous at \( x \) \( \Rightarrow \forall \epsilon > 0, \exists \delta \text{ such that } z \in B_{\delta}(x) \implies f(z) \in B_{\epsilon}(f(x)). \) or \( f(B_{\delta}(x)) \subset B_{\epsilon}(f(x)). \)
7.1 Continuity and Compactness

**Theorem.** Let \( f : X \to Y \) be continuous, where \( X \) is compact. Then \( f(X) \) is compact.

**Proof.** 1) Let \( \{V_\alpha\} \) be an open cover of \( f(X) \), \( (V_\alpha \subset Y) \). Since \( f \) is continuous, \( f^{-1}(V_\alpha) \) is open. Since \( X \) is compact, there are finitely many \( \alpha_1, \ldots, \alpha_n \), such that

\[
X \subset f^{-1}(V_{\alpha_1}) \cup \ldots \cup f^{-1}(V_{\alpha_n}).
\]

Since \( f(f^{-1}(E)) \subset E \) for every \( E \subset Y \), then

\[
f(X) \subset V_{\alpha_1} \cup \ldots \cup V_{\alpha_n}.
\]

\[\Rightarrow f(X) \text{ is compact}.\]

**Proof.** 2) Let \( \{y_n\} \) be a sequence in the image of \( f \). Thus we can find \( x_n \in X \), such that \( y_n = f(x_n) \). Since \( X \) is compact the sequence \( \{x_n\} \) has a convergent subsequence \( \{x_{n_k}\} \) with limit \( s \in X \). Since \( f \) is continuous,

\[
\lim_{k \to \infty} y_{n_k} = \lim_{k \to \infty} f(x_{n_k}) = f(s)
\]

Hence, the given sequence \( \{y_n\} \) has a convergent subsequence which converges in \( f(X) \).

\[\Rightarrow f(X) \text{ is compact}.\]

**Problem (F’01, #1).** Let \( K \subset \mathbb{R} \) be compact and \( f(x) \) continuous on \( K \). Then \( f \) has a maximum on \( K \) (i.e. there exists \( x_0 \in K \), such that \( f(x) \leq f(x_0) \) for all \( x \in K \)).

**Proof.** By theorem above, the image \( f(K) \) is closed and bounded. Let \( b \) be its least upper bound. Then \( b \) is adherent to \( f(K) \). Since \( f(K) \) is closed \( \Rightarrow b \in f(K) \), that is \( \exists x_0 \in K \), such that \( b = f(x_0) \), and thus \( f(x_0) \geq f(x) \), \( \forall x \in K \).

**Theorem.** Let \( f : [a, b] \to \mathbb{R} \) be continuous function. Then:

1) \( f \) is bounded;
2) \( f \) assumes its max and min values;
3) \( f(a) < p < f(b) \Rightarrow \exists x : f(x) = p \).

**Proof.** 1) \( f \) is continuous \( \Rightarrow f([a, b]) \) is compact \( \Rightarrow f([a, b]) \) is closed and bounded.
2) \( \phi \neq f([a, b]) \leq M \). Let \( M_0 = \sup f([a, b]) \Rightarrow M_0 \in \text{closure}(f([a, b])) = f([a, b]) \Rightarrow \exists x_0 \in [a, b] : f(x_0) = M_0 \).
3) \( [a, b] \) is connected \( \Rightarrow f([a, b]) \) is connected \( \Rightarrow f([a, b]) \) is an interval.

**Theorem.** \( f : X \to Y \) continuous and \( X \) compact. Then \( f \) is uniformly continuous.

**Proof.** Suppose that \( f \) is not uniformly continuous. Then there exist \( \epsilon > 0 \) and (setting \( \delta = 1/k \) in the definition) points \( x_k, z_k \in X \) such that \( |x_k - z_k| < 1/k \) while \( |f(x_k) - f(z_k)| \geq \epsilon \). Passing to a subsequence, we can assume that \( x_k \to x \in X \). Since \( |x_k - z_k| \to 0 \), we also obtain \( z_k \to x \). Since \( f \) is continuous, \( f(x_k) \to f(x) \) and \( f(z_k) \to f(x) \), so that \( |f(x_k) - f(z_k)| \leq |f(x_k) - f(x)| + |f(x) - f(z_k)| \to 0 \), a contradiction.

---

\(^9\)Gamelin, Green, p. 26-27; Rudin, p. 91.
\(^{10}\)See the technique of negation in the beginning of the section.
\(^{11}\)Since \( \{x_k\} \) is Cauchy, and the convergent subsequence can be constructed, \( x_k \to x \).
8 Sequences and Series of Functions

8.1 Pointwise and Uniform Convergence

Suppose \( \{f_n\} \), \( n = 1, 2, 3, \ldots \), is a sequence of functions defined on a set \( E \), and suppose that the sequence of numbers \( \{f_n(x)\} \) converges \( \forall x \in E \). Define a function \( f \) by

\[
f(x) = \lim_{n \to \infty} f_n(x), \quad (x \in E).
\]

\( \{f_n\} \) converges pointwise to \( f \) on \( E \).

A sequence of functions \( \{f_n\} \) converges uniformly on \( E \) to \( f \) if \( \forall \epsilon > 0, \exists N, \) such that \( \forall n \geq N, \)

\[
|f_n(x) - f(x)| \leq \epsilon, \quad \forall x \in E.
\]

**Example:** Consider \( f(x) = x^n \) on \([0, 1]\) \( \Rightarrow \) convergent, but not uniformly convergent on \([0, 1]\).

**Problem (F’01, #3).** If \( \{f_n\} \) is a sequence of continuous functions on \( E \), and if \( f_n \to f \) uniformly on \( E \), then \( f \) is continuous on \( E \).

**Proof.** Fix \( \epsilon > 0 \). Since \( f_n \to f \) uniformly, choose \( N, \) s.t. \( \forall n \geq N \)

\[
|f_n(x) - f(x)| < \frac{\epsilon}{3}, \quad \forall x \in E.
\]

Since \( f_n \) is continuous at \( p \), choose \( \delta, \) s.t. \( x \in E, \ |x - p| < \delta \) then

\[
|f_n(x) - f_n(p)| < \frac{\epsilon}{3}, \quad n \geq N
\]

Thus, \( |f(x) - f(p)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(p)| + |f_n(p) - f(p)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \)

Hence, given any \( \epsilon, \exists \delta, \) s.t. \( |x - p| < \delta \Rightarrow |f(x) - f(p)| < \epsilon, \) and \( f \) is continuous at \( p. \) \[ \square \]

8.2 Normed Vector Spaces

**Problem (F’03, #7).** \( C^0[a, b] \) with the metric \( d(f, g) \equiv \sup_{x \in [0, 1]} |f(x) - g(x)| = \|f - g\|_{\infty} \) is complete.

**Proof.** Let \( \{\varphi_n\}_1^\infty \) be a Cauchy sequence of elements of \( C^0[a, b] \). Given \( \epsilon > 0 \), choose \( N \) such that \( m, n \geq N \) \( \Rightarrow \)

\[
||\varphi_m - \varphi_n||_{\infty} < \frac{\epsilon}{2}
\]

Then, in particular, \( ||\varphi_m(x) - \varphi_n(x)||_{\infty} < \frac{\epsilon}{2} \) for each \( x \in [a, b] \). Therefore \( \{\varphi_n(x)\}_1^\infty \)

is a Cauchy sequence of real numbers, and hence converges to some real number \( \varphi(x) \).

It remains to show that the sequence of functions \( \{\varphi_n\} \) converges uniformly to \( \varphi \); if so, it will imply that \( \varphi \) is continuous on \([a, b]\), i.e. \( \varphi \in C^0[a, b] \).

**Claim:** For \( n \geq N, \) (\( N \) same as above, \( n \) fixed), \( |\varphi(x) - \varphi_n(x)| < \epsilon \) for all \( x \in [a, b] \).

To see this, choose \( m \geq N \) sufficiently large (depending on \( x \)) s.t. \( |\varphi(x) - \varphi_m(x)| < \epsilon/2. \)

\[ \Rightarrow |\varphi(x) - \varphi_n(x)| \leq |\varphi(x) - \varphi_m(x)| + |\varphi_m(x) - \varphi_n(x)| < \epsilon/2 + ||\varphi_m - \varphi_n||_{\infty} < \epsilon/2 + \epsilon/2 = \epsilon. \]

Since \( x \in [a, b] \) was arbitrary, it follows that \( ||\varphi_n - \varphi||_{\infty} < \epsilon \) as desired. \[ \square \]
Theorem. Let \( \{f_n\} \in C^1[a, b] \), \( f_n \to f \) pointwise, \( f'_n \to g \) uniformly. Then \( f_n \to f \) uniformly, and \( f \) is differentiable, with \( f' = g \).

Proof. By the Fundamental Theorem of Calculus, we have
\[
f_n(x) = f_n(a) + \int_a^x f'_n \quad \forall n, \forall x \in [a, b].
\]
From this and the Uniform Convergence and Integration theorem, we obtain
\[
f(x) = \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} f_n(a) + \lim_{n \to \infty} \int_a^x f'_n,
\]
\[
f(x) = f(a) + \int_a^x g.
\]
Another application of the Fundamental Theorem yields \( f' = g \) as desired.

To see that convergence of \( f_n \to f \) is uniform, note that
\[
|f_n(x) - f(x)| = \left| \int_a^x f'_n - \int_a^x g \right| + |f_n(a) - f(a)| \leq \int_a^x |f'_n - g| + |f_n(a) - f(a)|
\]
\[
\leq (b-a)||f_n - g||_{\infty} + |f_n(a) - f(a)|.
\]
The uniform convergence of \( f_n \) therefore follows from that of \( f'_n \).

Theorem. \( C^1[a, b] \), with the \( C^1 \)-norm defined by
\[
||\varphi|| = \max_{x \in [a, b]} |\varphi(x)| + \max_{x \in [a, b]} |\varphi'(x)|,
\]
is complete.

Proof. Let \( \{\varphi_n\} \) be a Cauchy sequence of elements of \( C^1[a, b] \). Since
\[
\max_{x \in [a, b]} |\varphi_m(x) - \varphi_n(x)| \leq ||\varphi_m - \varphi_n|| \quad (C^1 \text{- norm}),
\]
\( \varphi_n \) is a uniformly Cauchy sequence of continuous functions. Thus \( \varphi_n \to \varphi \in C^0[a, b] \) uniformly. Similarly, since
\[
\max_{x \in [a, b]} |\varphi'_m(x) - \varphi'_n(x)| \leq ||\varphi'_m - \varphi'_n|| \quad (C^1 \text{- norm}),
\]
\( \varphi'_n \to \psi \in C^0[a, b] \) uniformly. By the above theorem, \( \varphi \) is differentiable with \( \varphi' = \psi \), so \( \varphi \in C^1[a, b] \). Since
\[
\max_{x \in [a, b]} |\varphi_n(x) - \varphi(x)| = ||\varphi_n - \varphi||_{\infty} \quad \text{(sup norm)}, \quad \text{and}
\]
\[
\max_{x \in [a, b]} |\varphi'_n(x) - \varphi'(x)| = ||\varphi'_n - \varphi'||_{\infty} \quad \text{(sup norm)}
\]
the uniform convergence of \( \varphi_n \) and \( \varphi'_n \) implies that \( \varphi_n \to \varphi \) with respect to the \( C^1 \)-norm of \( C^1[a, b] \). Thus every Cauchy sequence in \( C^1[a, b] \) converges.
8.3 Equicontinuity

A family $F$ of functions $f$ defined on a set $E \subset X$ is **equicontinuous** on $E$ if:
\[ \forall \epsilon > 0, \exists \delta, \text{ such that } |x - y| < \delta, \ x, y \in E, \ f \in F \Rightarrow |f(x) - f(y)| < \epsilon. \]

8.3.1 Arzela-Ascoli Theorem

Suppose $\{f_n(x)\}^\infty_{n=1}$ is uniformly bounded and equicontinuous sequence of functions defined on a compact subset $K$ of $X$. Then $\{f_n\}$ is precompact, i.e. the closure of $\{f_n\}$ is compact, i.e. $\{f_n\}$ contains a uniformly convergent subsequence, i.e. $\{f_n\}$ contains a subsequence $\{f_{n_k}\}$ that converges uniformly on $K$ to a function $f \in X$.

9 Connectedness

9.1 Relative Topology

Define the neighborhood of a point in $\mathbb{R}^n$ as $N_\epsilon(x) = \{y : |x - y| < \epsilon\}$. Consider the subset of $\mathbb{R}^n$, $M \subseteq \mathbb{R}^n$. If all we are interested in are just points in $M$, it would be more natural to define a neighborhood of a point $x \in M$ as $N_{M,\epsilon} = \{y \in M : |x - y| < \epsilon\}$. Thus, the relative neighborhood is just a restriction of the neighborhood in $\mathbb{R}^n$ to $M$. Relative interior points and relative boundary points of a set, as well as a relative open set and a relative closed set, can be defined accordingly.

**Alternative definitions:**

$S \subseteq M(\subseteq \mathbb{R}^n)$ is open relative to $M$ if there is an open set $U$ in $\mathbb{R}^n$ such that $S = U \cap M$. $S \subseteq M(\subseteq \mathbb{R}^n)$ is closed relative to $M$ if there is a closed set $V$ in $\mathbb{R}^n$ such that $S = V \cap M$.

**Example:** A set $S = [1, 4)$ is open relative to $M = [1, 10] \subseteq \mathbb{R}$ since for the open set $U = (0, 4)$ in $\mathbb{R}$, we have $S = U \cap M$.

**Example:** A set $S = [1, 3]$ is open relative to $M = [1, 3] \cup [4, 6] \subseteq \mathbb{R}$ since for the open set $U = (0, 4)$ in $\mathbb{R}$, we have $S = U \cap M$.

9.2 Connectedness

$X$ is connected if it cannot be expressed as a disjoint union of two nonempty subsets that are both open and closed, i.e. $M$ is connected if $M = A \cup B$, such that $A, B$ open and $A \cap B = \emptyset$, then $A$ or $B$ is empty; or, $M$ is connected if $M = A \cup B$, $A, B$ open $\Rightarrow A = \emptyset$ or $B = \emptyset$.

Fact: $X$ is connected $\iff X$ and $\emptyset$ are the only subsets which are clopen.

$X$ is disconnected if there are closed and open subsets $A$ and $B$ of $X$ such that $A \cup B = X$, $A \cap B = \emptyset$, $A \neq \emptyset$, $B \neq \emptyset$.

Another way of phrasing: $X$ is disconnected if there is a closed and open $U \subset X$, such that $U \neq \emptyset$ and $U \neq X$. If there is such a $U$, then the complement $V = X \setminus U$ of $U$ is also both closed and open and $X$ is the disjoint union of the nonempty sets $U$ and $V$.

A subset of a space is a connected subset if it is connected in the relative topology.

\[^{12}A \cap B = \emptyset, \text{ then } A \text{ and } B \text{ are disjoint.}\]
Problem (S’02, #1). The closed interval \([a, b]\) is connected.

Proof. Let \([a, b] = G \cup H\), s.t. \(G \cap H = \emptyset\). Let \(b \in H\). Then claim: \(G = \emptyset\). If not, let \(c = \sup G\). Since \(G\) is closed, \(c \in G\). Since \(G\) is open\(^\text{13}\), \(B_i(c) \subseteq G\), i.e. \([c, c + \epsilon) \subseteq G\). That contradicts \(c = \sup G\). Thus \(G = \emptyset\).

Note: Since \([a, b]\) and \((a, b)\) can be expressed as the union of an increasing sequence of compact intervals, these are also connected.

Theorem. Let \(S_\alpha \subseteq M\), \(S_\alpha\) connected. Suppose \(\bigcap S_\alpha \neq \emptyset\). Then \(\bigcup S_\alpha\) is connected.

Proof. Let \(S = \bigcup S_\alpha = G \cup H\), \(G, H\) open in \(\bigcup S_\alpha\).

Choose \(x_0 \in \bigcap S_\alpha\). \(S_\alpha = (S_\alpha \cap G) \cup (S_\alpha \cap H)\). Assume \(x_0 \in G\). Since \(S_\alpha\) is connected and \(x_0 \in S_\alpha \cap G\), we get \(S_\alpha \cap G = \emptyset\), \(\forall \alpha\). Therefore, \((\bigcup S_\alpha) \cap H = \emptyset\). Since \(H \subseteq \bigcup S_\alpha\) \(\Rightarrow H = \emptyset\). Therefore, \(S\) is connected.

Corollary. \(\mathbb{R}\) is connected.

Proof. Let \(\mathbb{R} = \bigcup_{n \in \mathbb{N}} [-n, n]\), \(0 \in \bigcap [-n, n]\). Therefore, \(\mathbb{R}\) is the union of connected subsets. By the theorem above, \(\mathbb{R}\) is connected.

Theorem. Let \(f : M \to N\) is continuous and \(M\) is connected. Then \(f(M)\) is connected.

Proof. Let \(f(M) = G \cup H\), \(G, H \neq \emptyset\). \(G, H\) open. Then \(M = f^{-1}(G) \cup f^{-1}(H)\), where \(f^{-1}(G)\) and \(f^{-1}(H)\) are both open and nonempty. Contradicts connectedness of \(M\).

Theorem. \(a, b \in I\), and \(a < c < b\), then \(c \in I\), i.e. \(I\) is an interval \(\Leftrightarrow I \subseteq \mathbb{R}\) is connected.

Proof. \(\Rightarrow\) Assume \(I\) an interval.

\(S = [a, b] \cap I; S = (a, b), a \geq -\infty\); \(S = [a, b), b \leq \infty\); \(S = (a, b), a \geq -\infty, b \leq \infty\).

\(\Leftrightarrow\) Say \(I\) is not an interval. \(\exists a < c < b\), \(a, b \in I\), \(c \notin I\). \(I = ((-\infty, c) \cap I) \cup ((c, \infty) \cap I)\), i.e. \(I\) is not connected.

Problem (W’02, #3). The open unit ball in \(\mathbb{R}^2\), \{\((x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}\) is connected.

Proof. Let \(f_\theta(t) = t(\cos \theta, \sin \theta), -1 < t < 1\). We have \(f_\theta : (-1, 1) \to (t \cos \theta, t \sin \theta)\).

Since \(f_\theta\) is continuous and \((-1, 1)\) is connected, \(f_\theta((-1, 1))\) is connected. The unit ball can be expressed as \{\((x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\} = \bigcup_{0 \leq \theta < \pi} f_\theta(t)\). We know that the origin is contained in the intersection of \(f_\theta\)'s. Therefore, \(\bigcup_{0 \leq \theta < \pi} f_\theta(t)\) is connected by the theorem above.

\(^{13}\)If \(M = G \cup H\), \(G \cap H = \emptyset\) and \(G, H\) are open, then \(G\) is closed and open, since \(G = H^c\).

\(M = [0, 1] \cup [2, 3]\) is not connected because if \(G = [0, 1], H = [2, 3]\), \(M = G \cup H\), \(G, H\) are clopen in \(M\), \(G \cap H = \emptyset\), and \(G, H \neq \emptyset\).
9.3 Path Connectedness

A path in $X$ from $x_0$ to $x_1$ is a continuous function $\gamma : [0, 1] \to X$, such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$.

$X$ is path-connected if, for every pair of points $x_0$ and $x_1$ in $X$, there is a path $\gamma$ from $x_0$ to $x_1$.

**Theorem.** A path-connected space is connected.

**Proof.** Fix $x_0 \in X$. For each $x \in X$, let $\gamma_x : [0, 1] \to X$ be a path from $x_0$ to $x$. By theorems above, i.e.

1. $X$ connected and $f : X \to Y$ continuous $\Rightarrow f(X)$ connected;
2. any interval in $\mathbb{R}$ is connected;

each $\gamma_x([0, 1])$ is a connected subset of $X$. Each $\gamma_x([0, 1])$ contains $x_0$ and $X = \bigcup \{\gamma_x([0, 1]) : x \in X\}$, so that the theorem above shows that $X$ is connected.

**Theorem.** An open subset of $\mathbb{R}^n$ is connected $\iff$ it is path-connected.

**Problem.** Any subinterval of $\mathbb{R}$ (closed, open, or semiopen) is path-connected.

**Proof.** If $a$, $b$ belong to an interval (of any kind), then $\gamma(t) = (1 - t)a + tb$, $0 \leq t \leq 1$, defines a path from $a$ to $b$ in the interval.

**Problem.** If $X$ is path-connected and $f : X \to Y$ is a map, then $f(X)$ is path-connected.

**Proof.** If $p = f(x)$ and $q = f(y)$, and $\gamma$ is a path in $X$ from $p$ to $q$, then $f \circ \gamma$ is a path in $f(X)$ from $x$ to $y$. 

□
10 Baire Category Theorem

A subset $T \subset X$ is dense in $X$ if $T = X$, i.e. every point of $X$ is a limit point of $T$, or a point of $T$, or both.

A subset $Y \subset X$ is nowhere dense if $Y$ has no interior points, i.e. $\text{int}(Y) = \emptyset$.

$Y$ is nowhere dense $\iff X \setminus Y$ is a dense open subset of $X$.

The interior of $E$ is the largest open set in $E$, i.e. the set of all interior points of $E$.

$\mathbb{Q} = \mathbb{R}$, $(\mathbb{R} \setminus \mathbb{Q}) = \mathbb{R}$.

$[0, 1] \cap \mathbb{Q}$ is not closed; $(0, 1) \cap \mathbb{Q}$ is not open in $\mathbb{R}$.

**Baire’s Category Theorem.** Let $\{U_n\}_{n=1}^{\infty}$ be a sequence of dense open subsets of a complete metric space $X$. Then $\bigcap_{n=1}^{\infty} U_n$ is also dense in $X$. (Any countable intersection of dense open sets in a complete metric space is dense.)

**Corollary.** Let $\{E_n\}_{n=1}^{\infty}$ be a sequence of nowhere dense subsets of a complete metric space $X$. Then $\bigcup_{n=1}^{\infty} E_n$ has empty interior. (In a complete metric space, no nonempty open subset can be expressed as a union of countable collection of nowhere dense sets.)

**Proof.** We apply the Baire Category Theorem to the dense open sets $U_n = X \setminus E_n$.

Then, $\bigcap_{n=1}^{\infty} U_n$ is dense in $X$.

$$\bigcap_{n=1}^{\infty} U_n = \bigcap_{n=1}^{\infty} (X \setminus E_n) = X \setminus \bigcup_{n=1}^{\infty} E_n$$

Therefore, $X \setminus \bigcup_{n=1}^{\infty} E_n$ is dense, $\implies \bigcup_{n=1}^{\infty} E_n$ is nowhere dense $\implies \bigcup_{n=1}^{\infty} E_n$ has empty interior.

A subset of $X$ is of the first category (i.e. $\mathbb{Q}$) if it is the countable union of nowhere dense subsets. A subset (i.e. $\mathbb{I}, \mathbb{R}$) that is not of the first category is said to be of the second category.

$S \subset \mathbb{R}$ is $F_\sigma$ set if $S = \bigcup_{n=1}^{\infty} F_n$, $F_n$ closed.

$S \subset \mathbb{R}$ is $G_\delta$ set if $S = \bigcap_{n=1}^{\infty} G_n$, $G_n$ open.

**Problem (W’02, #4).** The set of irrational numbers $\mathbb{I}$ in $\mathbb{R}$ is not the countable union of closed sets (not an $F_\sigma$ set).

**Proof.** Suppose $\mathbb{I} = \bigcup F_n$, where each $F_n$ is closed.

$$\mathbb{R} = \bigcup_{n \in \mathbb{N}} F_n \cup \{q\}$$

Thus $\mathbb{R}$ can be expressed as the countable union of closed sets. By (corollary to) the Baire Category Theorem, since $\mathbb{R}$ is a nonempty open subset, one of these closed sets has a nonempty interior. It cannot be one of $q$’s, and since any nonempty interval contains rational numbers, it cannot be one of $F_n$’s. Contradiction.

**Problem (S’02, #2; F’02, #3).** The set $\mathbb{Q}$ of rational numbers is not the countable intersection of open sets of $\mathbb{R}$ (not a $G_\delta$ set).

Show that there is a subset of $\mathbb{R}$ which is not the countable intersection of open subsets.

**Proof.** We take complements in the preceding theorem. Suppose $\mathbb{Q} = \bigcap G_n$, where each $G_n$ open. Then,

$$\mathbb{I} = \mathbb{R} \setminus \mathbb{Q} = \mathbb{R} \setminus \bigcap_{n \in \mathbb{N}} G_n = \bigcup_{n \in \mathbb{N}} (\mathbb{R} \setminus G_n),$$

and $\mathbb{R} \setminus G_n$ is closed. Thus, $\mathbb{I}$ is a $F_\sigma$ set, which contradicts the previous theorem.
Problem (S’03, #3). Find $S \subset \mathbb{R}$ such that both (i) and (ii) hold for $S$:

(i) $S$ is not the countable union of closed sets (not $F_\sigma$);

(ii) $S$ is not the countable intersection of open sets (not $G_\delta$).

Proof. Let $A \subset [0,1]$ not $F_\sigma$, $B \subset [2,3]$ not $G_\delta \Rightarrow A \cup B$ is neither $F_\sigma$ nor $G_\delta$.

If $A \cup B$ is $F_\sigma$, say $A \cup B = \bigcup \{F_n\}$ closed $\Rightarrow$

$$A = A \cup B \bigcap \{0,1\} = \bigcup \{F_n \bigcap \{0,1\}\} \equiv F_\sigma \text{ set } \Rightarrow \text{contradiction.}$$

If $A \cup B$ is $G_\delta$, say $A \cup B = \bigcap \{G_n\}$ open $\Rightarrow$

$$B = A \cup B \bigcap \left(\frac{3}{2},\frac{7}{2}\right) = \bigcap \{G_n \bigcap \left(\frac{3}{2},\frac{7}{2}\right)\} \equiv G_\delta \text{ set } \Rightarrow \text{contradiction.}$$

\[\square\]

Problem. $Q$ is not open, is not closed, but is the countable union of closed sets ($F_\sigma$ set).

Proof. Since any neighborhood $(q-\epsilon,q+\epsilon)$ of a rational $q$ contains irrationals, $Q$ has no inner points. $\Rightarrow Q$ is not open. Since every irrational number $i$ is the limit of a sequence of rationals $\Rightarrow Q$ is not closed. Since every one-point-set $\{x\} \subset \mathbb{R}$ is closed and $Q$ is countable, say $(q_n)$ is a sequence of all rational numbers, we find that $Q = \bigcup_{n \in \mathbb{N}} \{q_n\}$ is the countable union of closed sets, i.e. $Q$ is $F_\sigma$.

\[\square\]

Problem. The set of isolated points of a countable complete metric space $X$ forms a dense subset of $X$.

Proof. For each point $x \in X$ that is not an isolated point of $X$, define $U_x = X \setminus \{x\}$. Each such $U_x$ is open and dense in $X$, and the intersection of the $U_x$’s consists precisely of the isolated points of $X$. By the Baire Category Theorem, the intersection of the $U_x$’s is dense in $X$.

\[\square\]

Problem. Suppose that $F$ is a subset of the first category in a metric space $X$ and $E$ is a subset of $F$. Prove that $E$ is of the first category in $X$. Show by an example that $E$ may not be of the first category in the metric space $F$.

Proof. If $F = \bigcup F_n$, where each $F_n$ is nowhere dense, then $E = \bigcup (E \cap F_n)$, and each $E \cap F_n$ is nowhere dense. For example, note the $\mathbb{R}$ is of first category in $\mathbb{R}^2$, but $\mathbb{R}$ is not of first category in itself.

\[\square\]

Problem. Any countable union of sets of the first category in $X$ is again of the first category in $X$.

Proof. A countable union of countable unions is a countable union.
Problem.  a) If \( a, b \in \mathbb{R} \) satisfy \( a < b \), then there exists a rational number \( q \in (a, b) \).

b) The set \( \mathbb{Q} \) of rational numbers is dense in \( \mathbb{R} \).

Problem. The set of irrational numbers is dense in \( \mathbb{R} \).

Proof. If \( i \) is any irrational number, and if \( q \) is rational, then \( q + i/n \) is irrational, and \( q + i/n \to q \). \( \square \)

Problem. Regard the rational numbers \( \mathbb{Q} \) as a subspace of \( \mathbb{R} \). Does the metric space \( \mathbb{Q} \) have any isolated points?

Proof. The rationals have no isolated points. This does not contradict the above, i.e. "The set of isolated points of countable complete metric space \( X \) forms dense subset of \( X \)" because the rationals are not complete. \( \square \)

Problem. Every open subset of \( \mathbb{R} \) is a union of disjoint open intervals (finite, semi-infinite, or infinite).

Proof. For each \( x \in U \), let \( I_x \) be the union of all open intervals containing \( x \) that are contained in \( U \). Show that each \( I_x \) is an open interval (possibly infinite or semi-infinite), any two \( I_x \)’s either coincide or are disjoint, and the union of the \( I_x \)’s is \( U \). \( \square \)

11 Integration

11.1 Riemann Integral

Let \([a, b]\) be a given interval. A partition \( P \) of \([a, b]\) is a finite set of points \( x_0, \ldots, x_n \):

\[
a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b, \quad \triangle x_i = x_i - x_{i-1}.
\]

A Riemann sum for \( f \) corresponding to the given partition is

\[
S(P, f) = \sum_{i=1}^{n} f(x'_i) \triangle x_i, \quad x_{i-1} \leq x'_i \leq x_i.
\]

Definition: \( f \) is Riemann integrable on \([a, b]\), if \( \exists A \in \mathbb{R} \) such that:

\( \forall \epsilon > 0, \exists \delta > 0 \) such that whenever \( S \) is a Riemann sum for \( f \) corresponding to any partition of \([a, b]\) with \( \max(\triangle x_i) < \delta \) \( \Rightarrow \ |S - A| < \epsilon \).

In this case \( A \) is called the Riemann integral of \( f \) between \( a \) and \( b \) and is denoted as \( \int_{a}^{b} f \, dx \).

Alternative Definition: \( f \) is Riemann integrable on \([a, b]\) if:

\( \forall \epsilon > 0, \exists P \) such that \( U(P, f) - L(P, f) < \epsilon \).

If \( f \) is bounded, there exist \( m \) and \( M \), such that \( m \leq f(x) \leq M, \ a \leq x \leq b \). Hence, for every \( P \),

\[
m(b - a) \leq S(P, f) \leq M(b - a),
\]

so that \( S(P, f) \) is bounded. This shows that Riemann sums are defined for every bounded function \( f \).
11.2 Existence of Riemann Integral

**Theorem.** $f$ is integrable on $[a, b] \iff \forall \epsilon > 0, \exists \delta > 0$ such that $S_1(P, f), S_2(P, f), P$ with $\max(\Delta x_i) < \delta$, then $|S_1 - S_2| < \epsilon$.

**Proof.** Suppose $f$ is integrable on $[a, b]$. For any $\epsilon > 0$, $\exists \delta > 0$ such that $|S - \int_a^b f(x)dx| < \epsilon/2$. For such $S_1$ and $S_2$,

$$|S_1 - S_2| = \left| \left( S_1 - \int_a^b f(x)dx \right) - \left( S_2 - \int_a^b f(x)dx \right) \right| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

\[ \iff \forall \epsilon > 0, \exists \delta > 0 \text{ such that } S_1(P, f), S_2(P, F), P \text{ with } \max(\Delta x_i) < \delta \Rightarrow |S_1 - S_2| < \epsilon. \]

For $n = 1, 2, \ldots$, choose $S^{(n)}(P, f), P$ with $\max(\Delta x_i) < 1/n$. Then, $\forall \epsilon > 0, \exists N > 0 (\delta = 1/N)$, such that $|S^{(n)} - S^{(m)}| < \epsilon$, $n, m \geq N \Rightarrow S^{(n)}$ is a Cauchy sequence of real numbers $\Rightarrow S^{(n)}$ converges to some $A \in \mathbb{R} \Rightarrow |S^{(N)} - A| < \epsilon, 1/N < \delta$. Thus for any $S(P, f), P$ with $\max(\Delta x_i) < \delta$, we have $|S - A| \leq |S - S^{(N)}| + |S^{(N)} - A| < 2\epsilon$. \hfill \(\Box\)

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

**Proof.** Since $f$ is uniformly continuous on $[a, b]$, $\forall \epsilon > 0, \exists \delta > 0$ such that $\forall x, z \in [a, b], |x - z| < \delta \Rightarrow |f(x) - f(z)| < \epsilon \odot$. If $P$ is any partition of $[a, b]$ with $\max(\Delta x_i) < \delta$, then $\odot$ implies that $M_i - m_i \leq \epsilon$, $i = 1, \ldots, n$, and therefore

$$U(P, f) - L(P, f) = \sum_{i=1}^{n} (M_i - m_i)\Delta x_i \leq \epsilon \sum_{i=1}^{n} \Delta x_i = \epsilon(b - a).$$

Thus, $\forall \epsilon > 0, \exists P$ such that $|U(P, f) - L(P, f)| < C\epsilon \Rightarrow f$ is integrable. \hfill \(\Box\)

11.3 Fundamental Theorem of Calculus

**Theorem.** If $f : [a, b] \to \mathbb{R}$ is continuous, and $F : [a, b] \to \mathbb{R}$ is defined by

$$F(x) = \int_{a}^{x} f(t)dt,$$

then $F$ is differentiable and $F' = f$.

**Proof.** Since $f$ is continuous, $F(x) = \int_{a}^{x} f(t)dt$ is defined for all $x \in [a, b]$. We have to show that for any fixed $x_0 \in [a, b]$

$$\lim_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0} = f(x_0).$$

For any $x \in [a, b], x \neq x_0$, we have

$$\left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| = \left| \int_{a}^{x} f(t)dt - \int_{a}^{x_0} f(t)dt\right| = \left| \int_{x_0}^{x} f(t)dt \right| = \left| \frac{\int_{x_0}^{x} f(t)dt}{x - x_0} \right| \leq \frac{\int_{x_0}^{x} |f(t) - f(x_0)|dt}{|x - x_0|} \Rightarrow \odot$$

Since $f$ is continuous at $x_0$, given $\epsilon > 0, \exists \delta$, such that $x \in [a, b], |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon$. Then for any $t$ between $x$ and $x_0$, we have $|f(t) - f(x_0)| < \epsilon$.

$$\Rightarrow \odot < \frac{\int_{x_0}^{x} \epsilon dt}{|x - x_0|} = \epsilon.$$  

Thus $F'(x_0) = f(x_0)$. \hfill \(\Box\)
Corollary. If $f : [a, b] \to \mathbb{R}$ is continuous and $F' = f$ on $[a, b]$, then
\[ \int_a^b f(t)dt = F(b) - F(a). \]

Proof. Since $\frac{d}{dx}(\int_a^x f(t)dt - F(x)) = f(x) - f(x) = 0$, $\int_a^x f(t)dt - F(x)$ is constant. Thus $\int_a^b f(t)dt = F(x) + c$, for some $c \in \mathbb{R}$. In particular, $0 = \int_a^a f(t)dt = F(a) + c$, so that $c = -F(a)$. Therefore, $\int_a^b f(t)dt = F(x) - F(a)$. Hence, $\int_a^b f(t)dt = F(b) - F(a)$. $\square$

Integration by Parts. Suppose $f$ and $g$ are differentiable functions on $[a, b]$, $f', g' \in \mathcal{R}$. Then
\[ \int_a^b f(x)g'(x)dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)dx. \]

Proof. Let $h(x) = f(x)g(x)$ and apply the Fundamental Theorem of Calculus to $h$ and its derivative.
\[ \int_a^b h'(x)dx = h(b) - h(a) \]
\[ \int_a^b (f'(x)g(x) + f(x)g'(x))dx = f(b)g(b) - f(a)g(a). \]

Note that $h' \in \mathcal{R}$. $\square$

Mean Value Theorem for Integrals. If $f : [a, b] \to \mathbb{R}$ is continuous, then $\int_a^b f(x)dx = f(c)(b - a)$ for some $c \in [a, b]$.

Proof. Since $f$ is continuous, by the Fundamental Theorem of Calculus, there is a function $F : [a, b] \to \mathbb{R}$ such that $F'(x) = f(x)$ for $x \in (a, b)$, and $\int_a^b f(x)dx = F(b) - F(a)$.

By the Mean Value Theorem for Differentiation, $\exists c \in (a, b)$ such that $F(b) - F(a) = F'(c)(b - a)$. Thus,
\[ \int_a^b f(x)dx = F(b) - F(a) = F'(c)(b - a) = f(c)(b - a). \]

Thus, $\exists c \in (a, b)$ such that $\int_a^b f(x)dx = f(c)(b - a)$. $\square$

Generalized Mean Value Theorem for Integrals. If $f, g : [a, b] \to \mathbb{R}$ are continuous and $g(x) > 0$ for all $x \in [a, b]$, then there exists $c \in [a, b]$ such that
\[ \int_a^b f(x)g(x)dx = f(c) \int_a^b g(x)dx. \]

Proof. Since $g(x) > 0$ for all $x \in [a, b]$ and since $g$ is continuous, $\int_a^b g(x) > 0$.

Suppose $f(x)$ attains its maximum $M$ at $x_2$ and minimum $m$ at $x_1$.

Then $m = f(x_1) \leq f(x) \leq f(x_2) = M$ for $x \in [a, b]$, and
\[ m \int_a^b g(x)dx \leq \int_a^b f(x)g(x)dx \leq M \int_a^b g(x)dx, \]
and hence
\[ m \leq \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} \leq M. \]

Since $f$ is continuous on compact $[a, b]$, $\exists c$ such that $f(c) = \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx}$. $\square$
Uniform Convergence and Integration. Let \( \{f_n\} \) be a sequence of continuous functions on \([a, b]\) and \(f_n \to f\) uniformly on \([a, b]\). Then
\[
\int_a^b \lim_{n \to \infty} f_n(x) \, dx = \lim_{n \to \infty} \int_a^b f_n(x) \, dx.
\]

Proof. Let \( f = \lim_{n \to \infty} f_n \). Since each \( f_n \) is continuous and \( f_n \to f \) uniformly \( \Rightarrow f \) is continuous. In particular, \( f \) is integrable on \([a, b]\). By the definition of uniform convergence, \( \forall \epsilon > 0, \exists N > 0 \) such that \( n > N \), \( |f(x) - f_n(x)| < \epsilon/(b-a) \), \( \forall x \in [a, b] \).

Thus,
\[
-\frac{\epsilon}{b-a} \leq f(x) - f_n(x) \leq \frac{\epsilon}{b-a}, \quad \forall x \in [a, b] \quad \Rightarrow
\]
\[
\Rightarrow -\epsilon \leq \int_a^b (f(x) - f_n(x)) \, dx \leq \epsilon
\]
or
\[
\left| \int_a^b f(x) \, dx - \int_a^b f_n(x) \, dx \right| \leq \epsilon.
\]
The last inequality holds for all \( n > N \), and therefore, \( \lim_{n \to \infty} \int_a^b f_n(x) \, dx = \int_a^b f(x) \, dx \).

Uniform Convergence and Differentiation. Let \( \{f_n\} \) be a sequence of differentiable functions on \([a, b]\) such that \( \{f_n(x_0)\} \) converges for some \( x_0 \in [a, b] \). If \( f'_n \to f' \) uniformly, then \( f_n \to f \) uniformly on \([a, b]\), and
\[
\left( \lim_{n \to \infty} f_n \right)' = \lim_{n \to \infty} f'_n
\]

Proof. See the section on “Sequences and Series of Functions: Normed Vector Spaces” where the weaker statement is proved, i.e. \( \{f_n\} \in C^1, f_n \to f \) pointwise on \([a, b]\). \( \square \)
12 Differentiation

12.1 $\mathbb{R} \to \mathbb{R}$

12.1.1 The Derivative of a Real Function

Let $f : (a, b) \to \mathbb{R}$, $x_0 \in (a, b)$. $f$ is differentiable at $x_0$ if

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists. $f'$ is the derivative of $f$.

**Theorem.** $f : (a, b) \to \mathbb{R}$. $f$ is differentiable at $x_0$ $\Rightarrow$ $f$ is continuous at $x_0$.

**Proof.**

$$\lim_{x \to x_0} (f(x) - f(x_0)) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \cdot \lim_{x \to x_0} (x - x_0) = f'(x_0) \cdot 0 = 0.$$

Since $\lim_{x \to x_0} f(x) = f(x_0)$, $f$ is continuous at $x_0$. $\square$

**Lemma.** If $f'(c) > 0$, then $f$ is locally strictly increasing at $c$, i.e., $\exists \delta > 0$ such that:

$c - \delta < x < c \Rightarrow f(x) < f(c)$,

$c < x < c + \delta \Rightarrow f(c) < f(x)$.

**Proof.** $\lim_{x \to c} \frac{f(x) - f(c)}{x - c} > 0 \Rightarrow \exists \delta > 0$: $\frac{f(x) - f(c)}{x - c} > 0$ whenever $0 < |x - c| < \delta$.

Thus since $f(x) - f(c) = \frac{f(x) - f(c)}{x - c} (x - c)$, $(x \neq c)$, we have

$c - \delta < x < c \Rightarrow f(x) - f(c) < 0$,

$c < x < c + \delta \Rightarrow f(x) - f(c) > 0$. $\square$

**Corollary.** If $f$ has a max (or a min) at $c \in (a, b)$, i.e. $f(x) \leq f(c)$ (or $f(x) \geq f(c)$) for all $x$, then $f'(c) = 0$.

**Proof.** Say $f'(c) \neq 0$. Say $f'(c) > 0$. Then $x > c \Rightarrow f(x) \leq f(c)$ (since $f$ has a max at $c$), contradicting the lemma above. Proofs of other conditions are similar. $\square$

12.1.2 Rolle’s Theorem

**Theorem.** Let $f$ be continuous on $[a, b]$, differentiable on $(a, b)$, and $f(a) = f(b)$. Then $\exists c \in (a, b): f'(c) = 0$.

**Proof.** Let $M = \sup\{f(x) : x \in [a, b]\}$, $m = \inf\{f(x) : x \in [a, b]\}$.

Then $m \leq f(a) = f(b) \leq M$. If $M = m \Rightarrow f(x) = f(a)$, $\forall x \Rightarrow f'(c) = 0$, $\forall c \in (a, b)$.

Say $f(a) = f(b) < M \oplus$. Then choose $c \in [a, b]: f(c) = M$. From $\oplus$, $c \in (a, b)$. We have from corollary, $f'(c) = 0$.

Similarly, say $m < f(a) = f(b) \ominus$. Then choose $c \in [a, b]: f(c) = m$. From $\ominus$, $c \in (a, b)$. We have from corollary, $f'(c) = 0$. $\square$

12.1.3 Mean Value Theorem

**Theorem.** Let $f$ be continuous on $[a, b]$, differentiable on $(a, b)$. Then $\exists c \in (a, b):$

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$
Proof. Let \( g(x) = f(x) - \frac{f(b) - f(a)}{b-a} (x-a) \). Then \( g(a) = g(b) = f(a) \). By Rolle’s Theorem, \( \exists c \in (a,b) \), such that \( 0 = g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a} \).

**Corollary.** Let \( f \) be continuous on \([a, b]\), differentiable on \((a, b)\).

a) \( f'(x) = 0 \), \( \forall x \in (a, b) \) \( \Rightarrow \) \( f \) is constant.
b) \( f'(x) > 0 \), \( \forall x \in (a, b) \) \( \Rightarrow \) \( f \) is strictly increasing.

Proof. b) \( a \leq x_1 < x_2 \leq b \). By mean value theorem, \( f(x_2) - f(x_1) = f'(c)(x_2 - x_1) \) for some \( c \in (x_1, x_2) \). Therefore, \( f(x_2) - f(x_1) > 0 \) for all such \( x_1, x_2 \). Proof of a) is similar.

### 12.2 \( \mathbb{R} \to \mathbb{R}^n \)

**Proof.** Let \( f : \mathbb{R} \to \mathbb{R}^m \) be differentiable at \( c \in \mathbb{R} \) if there exists a linear map \( L : \mathbb{R} \to \mathbb{R}^m \) such that

\[
\lim_{h \to 0} \frac{\|f(c + h) - f(c) - L(h)\|}{\|h\|} = 0.
\]

in which case \( L \) is defined by \( L = df_c = f'(c) = \begin{bmatrix} f'_1(c) \\ \vdots \\ f'_m(c) \end{bmatrix} \).

The linear mapping \( df_c : \mathbb{R} \to \mathbb{R}^m \) is called the differential of \( f \) at \( c \). The matrix of the linear mapping \( f'(c) \) is the derivative. The differential is the linear mapping whose matrix is the derivative.

### 12.3 \( \mathbb{R}^n \to \mathbb{R}^m \)

Let \( U \subseteq \mathbb{R}^n \) open, \( c \in U \). \( f : U \to \mathbb{R}^m \) is differentiable at \( c \) if there exists a linear map \( L : \mathbb{R}^n \to \mathbb{R}^m \) such that

\[
\lim_{h \to 0} \frac{\|f(c + h) - f(c) - L(h)\|}{\|h\|} = 0.
\]

in which case \( L \) is defined by

\[
L = df_c = f'(c) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}_{x=c}
\]

Let \( v \in \mathbb{R}^n \), \( c \in G \subseteq \mathbb{R}^n \). The directional derivative with respect to \( v \) of \( f \) at \( c \) is

\[
D_v f(c) = \lim_{t \to 0} \frac{f(c + tv) - f(c)}{t}
\]

provided that the limit exists. In particular, the partial derivatives of \( f \) at \( c \) are

\[
\frac{\partial f}{\partial x_j}(c) = D_{e_j} f(c) = \lim_{t \to 0} \frac{f(c + te_j) - f(c)}{t} = \lim_{t \to 0} \frac{f(c_1, \ldots, c_j + t, \ldots, c_n) - f(c_1, \ldots, c_j, \ldots, c_n)}{t}
\]
Theorem. Let $f$ be differentiable at $c$ (i.e., $df_c : \mathbb{R}^n \to \mathbb{R}^m$ is defined). Then:

a)\textsuperscript{14} The directional derivative $D_v f(c)$ exists $\forall v \in \mathbb{R}^n$, and $D_v f(c) = df_c(v) = f'(c)v$.

b)\textsuperscript{15} The matrix $f'(c)$ of $df_c$ is $df_c = f'(c) = \left[ \frac{\partial f_i}{\partial x_j}(c) \right]_{m \times n}$.

Proof. a) Since $f$ is differentiable, and letting $h = tv$,

$$
\lim_{t \to 0} \frac{||f(c + tv) - f(c) - df_c(tv)||}{||tv||} = 0,
$$

$$
\frac{1}{||v||} \left[ \lim_{t \to 0} \left| \frac{f(c + tv) - f(c)}{t} - df_c(v) \right| \right] = 0,
$$

$$
\lim_{t \to 0} \frac{f(c + tv) - f(c)}{t} = df_c(v).
$$

i.e. $D_v f(c)$ exists, and equals to $df_c(v)$.

Proof. b) $df_c = \left[ \begin{array}{c} \frac{\partial f_1}{\partial x_1}(c) \\ \vdots \\ \frac{\partial f_n}{\partial x_n}(c) \end{array} \right] = \left[ \begin{array}{c} D_{e_1} f(c) \\ \vdots \\ D_{e_n} f(c) \end{array} \right] = \left[ \begin{array}{c} \frac{\partial f_1}{\partial x_1}(c) \\ \vdots \\ \frac{\partial f_n}{\partial x_n}(c) \end{array} \right]_{m \times n}$.

$f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be \textbf{continuously differentiable} at $c$ if the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist on $B_{\epsilon}(c)$, and are continuous at $c$.

\textsuperscript{14}Edwards, Theorem 2.1

\textsuperscript{15}Edwards, Theorem 2.4
Theorem\textsuperscript{16}. If \( \frac{\partial f_i}{\partial x_j} \) exist and are continuous on \( G \), then at each \( c \in G \),

\[
df_c = f'(c) = \left[ \frac{\partial f_i}{\partial x_j}(c) \right].
\]

(i.e. \( \frac{\partial f_i}{\partial x_j} \) continuous \( \Rightarrow f \) differentiable).

Proof. Since \( f : \mathbb{R}^n \to \mathbb{R}^m \) is differentiable at \( c \) iff each of its component functions \( f_1, \ldots, f_m \) is, we may assume \( m = 1 \); \( f : U \to \mathbb{R} \). Given \( h = (h_1, \ldots, h_n) \), let \( h_0 = (0, \ldots, 0), h_j = (h_1, \ldots, h_j, 0, \ldots, 0), j = 1, \ldots, n \). We have

\[
f(c + h) - f(c) = \sum_{j=1}^n [f(c + h_j) - f(c + h_{j-1})].
\]

\[
f(c + h_j) - f(c + h_{j-1}) = f(c_1 + h_1, \ldots, c_{j-1} + h_{j-1}, c_j + h_j, c_{j+1}, \ldots, c_n)
- f(c_1 + h_1, \ldots, c_{j-1} + h_{j-1}, c_j, c_{j+1}, \ldots, c_n)
= \frac{\partial f}{\partial x_j}(c_1 + h_1, \ldots, c_{j-1} + h_{j-1}, c_j + t, c_{j+1}, \ldots, c_n) \cdot h_j
\]

for some \( 0 \leq t \leq h_j \), by mean-value theorem. Thus

\[
f(c + h_j) - f(c + h_{j-1}) = \frac{\partial f}{\partial x_j}(d_j) \cdot h_j,
\]

for some \( d_j, \|d_j - c\| \leq h\).

\[
\Rightarrow f(c + h) - f(c) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(d_j) \cdot h_j
\]

Also considering \( f'(c) \cdot h = \left[ \frac{\partial f}{\partial x_1}(c) \right] h_1 \cdots \left[ \frac{\partial f}{\partial x_n}(c) \right] h_n \), we have

\[
\lim_{h \to 0} \frac{|f(c + h) - f(c) - f'(c) \cdot h|}{\|h\|} = \lim_{h \to 0} \frac{\left| \sum_{j=1}^n \left[ \frac{\partial f}{\partial x_j}(d_j) - \frac{\partial f}{\partial x_j}(c) \right] h_j \right|}{\|h\|}
\]

\[
\leq \lim_{h \to 0} \sum_{j=1}^n \left| \frac{\partial f}{\partial x_j}(d_j) - \frac{\partial f}{\partial x_j}(c) \right| = 0,
\]

since each \( d_j \to c \) as \( h \to 0 \), and each \( \frac{\partial f}{\partial x_j} \) is continuous at \( c \). \( \square \)

\textsuperscript{16}Edwards, Theorem 2.5
12.3.1 Chain Rule

**Theorem.** $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$. If the mappings $F : U \to \mathbb{R}^m$ and $G : V \to \mathbb{R}^k$ are differentiable at $a \in U$ and $F(a) \in V$ respectively, then their composition $H = G \circ F$ is differentiable at $a$, and

$$dH_a = dG_{F(a)} \circ dF_a \quad \text{(composition of linear mappings)}$$

In terms of derivatives

$$H'(a) = G'(F(a)) \cdot F'(a) \quad \text{(matrix multiplication)}$$

The differential of the composition is the composition of the differentials; the derivative of the composition is the product of the derivatives.

**Proof.** We must show that

$$\lim_{h \to 0} \frac{H(a + h) - H(a) - dG_{F(a)} \circ dF_a(h)}{||h||} = 0.$$  

Define

$$\varphi(h) = \frac{F(a + h) - F(a) - dF_a(h)}{||h||} \quad \Rightarrow \quad F(a + h) - F(a) = dF_a(h) + ||h|| \varphi(h) \quad \text{and} \quad (12.1)$$

$$\psi(k) = \frac{G(F(a) + k) - G(F(a)) - dG_{F(a)}(k)}{||k||} \quad \Rightarrow \quad G(F(a) + k) - G(F(a)) = dG_{F(a)}(k) + ||k|| \psi(k) \quad (12.2)$$

The fact that $F$ and $G$ are differentiable at $a$ and $F(a)$, respectively, implies that

$$\lim_{h \to 0} \varphi(h) = \lim_{k \to 0} \psi(k) = 0.$$  

Then

$$H(a + h) - H(a) = G(F(a + h)) - G(F(a)) = G(F(a) + (F(a) + h) - F(a))) - G(F(a))$$

$$= dG_{F(a)}(F(a) + h) - F(a)) + ||F(a + h) - F(a)|| \cdot \psi(F(a + h) - F(a))$$

$$= (12.1) = dG_{F(a)}(dF_a(h) + ||h|| \varphi(h)) + ||dF_a(h) + ||h|| \varphi(h)|| \cdot \varphi(h)(F(a) - F(a))$$

$$= dG_{F(a)} \circ dF_a(h) + ||h||dG_{F(a)}(\varphi(h)) + ||h||||dF_a(h)\bigg\| \varphi(h) \bigg\| \cdot \varphi(F(a + h) - F(a))$$

$$\Rightarrow \quad \frac{H(a + h) - H(a) - dG_{F(a)} \circ dF_a(h)}{||h||} = dG_{F(a)}(\varphi(h)) + \bigg\|dF_a(h) \bigg\| \varphi(h) \bigg\| \cdot \varphi(F(a + h) - F(a))$$

$dG_{F(a)}$ is linear $\Rightarrow$ continuous, and $\lim_{h \to 0} \varphi(h) = 0 \Rightarrow \lim_{h \to 0} dG_{F(a)}(\varphi(h)) = 0$.

Since $F$ is continuous at $a$ and $\lim_{k \to 0} \psi(k) = 0 \Rightarrow \lim_{h \to 0} \psi(F(a + h) - F(a)) = 0$.

$dF_a$ is linear $\Rightarrow$ continuous $\Rightarrow$ bounded $\Rightarrow \exists M, ||dF_a(h)|| \leq M||h||$.

Therefore, the limit of the entire expression above $\to 0 \Rightarrow dH_a = dG_{F(a)} \circ dF_a$. \(\square\)
**Theorem.** Let \( U \subseteq \mathbb{R}^n \) be open and connected. \( F : U \rightarrow \mathbb{R}^m \). \( F'(x) = 0 \) for all \( x \in U \) \( \iff \) \( F \) is constant.

**Proof.** Since \( F \) is constant \( \iff \) each of its component functions is constant, and the matrix \( F'(x) = 0 \) \( \iff \) each of its rows is \( 0 \), we may assume \( F = f : U \rightarrow \mathbb{R} \).

Suppose \( f'(x) = \nabla f(x) = 0 \), \( \forall x \in U \).

Given \( a \) and \( b \in U \), let \( \gamma : \mathbb{R} \rightarrow U \) be a differentiable mapping with \( \gamma(0) = a \), \( \gamma(1) = b \).

If \( g = f \circ \gamma : \mathbb{R} \rightarrow \mathbb{R} \), then

\[
g'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t) = 0
\]

\( \forall t \in \mathbb{R} \) \( \Rightarrow \) \( g \) is constant on \( [0,1] \), so \( \quad f(a) = f(\gamma(0)) = g(0) = g(1) = f(\gamma(1)) = f(b) \). \( \square \)

**Example 1.** \( w = w(x,y) \), \( x = x(r,\theta) \), \( y = y(r,\theta) \)

\[
\begin{bmatrix}
\frac{\partial w}{\partial r} & \frac{\partial w}{\partial \theta}
\end{bmatrix} = \begin{bmatrix}
\frac{\partial w}{\partial x} & \frac{\partial w}{\partial y}
\end{bmatrix} \begin{bmatrix}
\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta}
\end{bmatrix} \Rightarrow \begin{bmatrix}
\frac{\partial w}{\partial r} & \frac{\partial w}{\partial \theta}
\end{bmatrix} = \left[ \begin{bmatrix}
\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta}
\end{bmatrix} \begin{bmatrix}
\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta}
\end{bmatrix} \right]^{-1}
\]

**Example 2.** \( f(x,y,z) = 0 \), \( \frac{\partial f}{\partial z} \neq 0 \) \( \Rightarrow \) \( z = h(x,y) \), \( f(x,y,h(x,y)) = 0 \).

For example, can solve for \( \frac{\partial z}{\partial x} \):

\( 0 = \frac{\partial}{\partial x} f(x,y,h(x,y)) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial x} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial x} \)

\( \Rightarrow \frac{\partial z}{\partial x} = -\frac{\partial f}{\partial z} \frac{\partial z}{\partial x} \).

Similarly, can solve for \( \frac{\partial y}{\partial x} \) and \( \frac{\partial x}{\partial y} \), from \( \frac{\partial}{\partial x} f(x,y(x,z),z) \) and \( \frac{\partial}{\partial y} f(x(y,z),y,z) \), respectively, and show \( \frac{\partial x}{\partial y} \frac{\partial y}{\partial z} = -1 \).

### 12.3.2 Mean Value Theorem

**Theorem.** Let \( f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \) be differentiable, and \( a \) and \( b \in U \), such that \( [a,b] \subseteq U \). Then \( \exists c \in (a,b) \), such that

\[
f(b) - f(a) = \nabla f(c) \cdot (b - a).
\]

**Proof.** Define \( \gamma : [0,1] \rightarrow [a,b] \) as \( \gamma(t) = a + t(b - a) = (1-t)a + tb \), \( t \in [0,1] \).

Then \( \gamma'(t) = b - a \). Let \( g = f \circ \gamma : \mathbb{R} \rightarrow \mathbb{R} \)

Since \( g : [0,1] \rightarrow \mathbb{R} \), then by single-variable MVT, \( \exists \xi \in [0,1] \), such that \( g(1) - g(0) = g'(\xi) \). If \( c = \gamma(\xi) \), then

\[
f(b) - f(a) = f(\gamma(1)) - f(\gamma(0)) = g(1) - g(0) = g'(\xi) =
\]

\( = \text{Chain Rule} \) \( \nabla f(\gamma(\xi)) \cdot \gamma'(\xi) = \nabla f(c) \cdot (b - a) \). \( \square \)
Theorem. Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$. If $f$, $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ exist and are continuous on $U$ and $\frac{\partial^2 f}{\partial x \partial y}$, $\frac{\partial^2 f}{\partial y \partial x}$ exist on $U$ and are continuous at $a$, then

$$\frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right)$$
onumber on $U$.

Proof. Consider $S_h(x, y) = f(x + h, y + h) - f(x + h, y) - f(x, y + h) + f(x, y)$. Let $g(x, y) = f(x + h, y) - f(x, y)$. Then,

$$S_h(x, y) = f(x + h, y + h) - f(x + h, y) - f(x, y + h) + f(x, y) = MVT = h \frac{\partial g}{\partial y}(x, y + \beta h) = h \left[ \frac{\partial f}{\partial y}(x + h, y + \beta h) - \frac{\partial f}{\partial y}(x, y + \beta h) \right] = MVT = h^2 \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right)(x + \alpha h, y + \beta h)$$

where $0 < \alpha, \beta < 1$.

Let $r(x, y) = f(x, y + h) - f(x, y)$. Then,

$$S_h(x, y) = f(x + h, y + h) - f(x + h, y) - f(x, y + h) + f(x, y) = MVT = h \frac{\partial r}{\partial x}(x + \alpha' h, y) = h \left[ \frac{\partial f}{\partial x}(x + \alpha' h, y + h) - \frac{\partial f}{\partial x}(x + \alpha' h, y) \right] = MVT = h^2 \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right)(x + \alpha' h, y + \beta' h)$$

where $0 < \alpha', \beta' < 1$.

For each small enough $h > 0$, $\frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right)(x + \alpha h, y + \beta h) = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right)(x + \alpha' h, y + \beta' h)$. Since the mixed partial derivatives are continuous at $a = (x, y)$, let $h \to 0 \Rightarrow \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right)(x, y) = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right)(x, y)$. \qed
12.4 Taylor’s Theorem

**n-th degree Taylor polynomial of f at a;** \( (h = x - a) \)

\[ P_n(h) = f(a) + f'(a)h + \cdots + \frac{f^{(n)}(a)}{n!}h^n \]

**Mean Value Theorem \((\mathbb{R} \to \mathbb{R}, \text{Revisited})\).** \( f : [a, b] \to \mathbb{R} \). Suppose that \( f' \) exists on \([a, b]\). \( h = b - a \). Then \( \exists \xi \) between \( a \) and \( b \) such that

\[ R_0(h) = f'(\xi)h \]

\[ f(a + h) = f(a) + f'(\xi)h = P_0(h) + R_0(h) \]

**Proof.** Need to show: \( R_0(h) = f'(\xi)h \). For \( t \in [0, h] \), define

\[ R_0(t) = f(a + t) - P_0(t) = f(a + t) - f(a) \]

So, \( R'_0(t) = f'(a + t) \).

Define \( \varphi : [0, h] \to \mathbb{R} \) by

\[ \varphi(t) = R_0(t) - \frac{R_0(h)}{h}t \implies \varphi(0) = \varphi(h) = 0 \]

\[ \implies \text{By Rolle’s theorem, } \exists c \in (0, h) \text{ such that} \]

\[ 0 = \varphi'(c) = R'_0(c) - \frac{R_0(h)}{h} = f'(a + c) - \frac{R_0(h)}{h} \]

\[ \implies \text{For } \xi = a + c, \ R_0(h) = f'(\xi)h. \]

**Taylor’s Theorem \((\mathbb{R} \to \mathbb{R})\).** \( f : [a, b] \to \mathbb{R} \). Suppose that \( f^{(n+1)} \) exists on \([a, b]\).

\( h = b - a \). Then \( \exists \xi \) between \( a \) and \( b \) such that

\[ R_n(h) = \frac{f^{(n+1)}(\xi)}{(n + 1)!}h^{n+1} \]

\[ f(a + h) = f(a) + f'(a)h + \cdots + \frac{f^{(n)}(a)}{n!}h^n + \frac{f^{(n+1)}(\xi)}{(n + 1)!}h^{n+1} = P_n(h) + R_n(h). \]

**Proof.** Need to show: \( R_n(h) = \frac{f^{(n+1)}(\xi)}{(n + 1)!}h^{n+1} \). For \( t \in [0, h] \), define \( R_n(t) = f(a + t) - P_n(t) \), and note that

\[ R_n(0) = R'_n(0) = \cdots = R^{(n)}_n(0) = 0 \] \hspace{1cm} (12.3)

since the first \( n \) derivatives of \( P_n(t) \) at 0 agree with those of \( f \) at \( a \). Also,

\[ R^{(n+1)}_n(t) = f^{(n+1)}(a + t) \] \hspace{1cm} (12.4)

since \( P^{(n+1)}_n(t) \equiv 0 \) because \( P_n(t) \) is a polynomial of degree \( n \).

Define \( \varphi : [0, h] \to \mathbb{R} \) by

\[ \varphi(t) = R_n(t) - \frac{R_n(h)}{h^{n+1}}f^{n+1} \implies \varphi(0) = \varphi(h) = 0 \]

\[ \implies \text{By Rolle’s theorem, } \exists c_1 \in (0, h) \text{ such that } \varphi'(c_1) = 0. \]

It follows from (12.3) and (12.4) that

\[ \varphi(0) = \varphi'(0) = \cdots = \varphi^{(n)}(0) = 0 \] \hspace{1cm} (12.5)
\[ \varphi^{(n+1)}(t) = f^{(n+1)}(a + t) - \frac{R_n(h)}{h^{n+1}}(n+1)! \]  

(12.6)

Therefore, we can apply Rolle’s theorem to \( \varphi' \) on \([0, c_1] \) to obtain \( c_2 \in (0, c_1) \) such that \( \varphi''(c_2) = 0 \).

By (12.5), \( \varphi'' \) satisfies the hypothesis of Rolle’s theorem on \([0, c_2] \), so we can continue in this way. After \( n + 1 \) applications of Rolle’s theorem, we obtain \( c_{n+1} \in (0, h) \) such that \( \varphi^{(n+1)}(c_{n+1}) = 0 \). From (12.6) we obtain \( R_n(h) = \frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+1} \) where \( \xi = a + c_{n+1} \).

**Problem (F’03, #5).** Assume \( f : \mathbb{R}^2 \to \mathbb{R} \) is a function such that all partial derivatives of order 3 exist and are continuous. Write down (explicitly in terms of partial derivatives of \( f \)) a quadratic polynomial \( P(x, y) \) in \( x \) and \( y \) such that

\[ |f(x, y) - P(x, y)| \leq C(x^2 + y^2)^{\frac{3}{2}} \]

for all \((x, y)\) in some small neighborhood of \((0, 0)\), where \( C \) is a number that may depend on \( f \) but not on \( x \) and \( y \). Then prove the above estimate.

**Proof.** Taylor expand \( f(x, y) \) around \((0, 0)\):

\[

t(x, y) = f(0, 0) + x f_x(0, 0) + y f_y(0, 0) + \frac{1}{2} \left[ x^2 f_{xx}(0, 0) + 2xy f_{xy}(0, 0) + y^2 f_{yy}(0, 0) \right] \\
+ \frac{1}{3!} \left[ x^3 f_{xxx}(0, 0) + 3x^2 y f_{xxy}(0, 0) + 3xy^2 f_{xyy}(0, 0) + y^3 f_{yyy}(0, 0) \right] + O(x^4)
\]

\[
|f(x, y) - P_2(x, y)| = \left| \frac{1}{3!} \left[ x^3 f_{xxx}(\xi, \eta) + 3x^2 y f_{xxy}(\xi, \eta) + 3xy^2 f_{xyy}(\xi, \eta) + y^3 f_{yyy}(\xi, \eta) \right] \right|
\]

Note that \(|x^3|, |x^2y|, |xy^2|, |y^3| \leq (x^2 + y^2)^{\frac{3}{2}}\). Also, since 3rd order partial derivatives are continuous, \( \exists C_1 \in \mathbb{R} \) s.t. \( \max \{ f_{xxx}, f_{xxy}, f_{xyy}, f_{yyy} \} < C_1 \) in some nbd of \((0, 0)\).

Thus,

\[
|f(x, y) - P_2(x, y)| \leq \frac{1}{3!} (x^2 + y^2)^{\frac{3}{2}} C_1 \leq C(x^2 + y^2)^{\frac{3}{2}}.
\]

\( \square \)
Problem (F'02, #5). Suppose \( f : \mathbb{R}^2 \rightarrow \mathbb{R} \) has partial derivatives at every point bounded by \( A > 0 \).

(a) Show that there is an \( M \) such that

\[
|f((x, y)) - f((x_0, y_0))| \leq M((x - x_0)^2 + (y - y_0)^2)^{\frac{1}{2}} \quad \ominus
\]

(b) What is the smallest value of \( M \) (in terms of \( A \)) for which this always works?

(c) Give an example where that value of \( M \) makes the inequality an equality.

Proof. (a) Since \( |\frac{\partial f}{\partial x}| \leq A, |\frac{\partial f}{\partial y}| \leq A \), by the Mean Value Theorem,

\[
|f(x, y) - f(x_0, y_0)| \leq A|x - x_0| \\
|f(x, y) - f(x, y_0)| \leq A|y - y_0| \\
|f(x, y) - f(x_0, y_0)| \leq |f(x, y) - f(x_0, y)| + |f(x_0, y) - f(x_0, y_0)| \\
\leq A(|x - x_0| + |y - y_0|) \leq A\sqrt{2}(|x - x_0|^2 + |y - y_0|^2)^{\frac{1}{2}}
\]

(b) This always works for \( M = A\sqrt{2} \).

(c) If \( |x - x_0| = |y - y_0| \), then we have an equality in \( \ominus \), since then

\[
A(|x - x_0| + |y - y_0|) \leq A\sqrt{2}(|x - x_0|^2 + |y - y_0|^2)^{\frac{1}{2}}
\]

\[
\Rightarrow \quad 2A|x - x_0| \leq A\sqrt{2}(2|x - x_0|^2)^{\frac{1}{2}}
\]

\[
\Rightarrow \quad 2A|x - x_0| \leq 2A|x - x_0|
\]

\[\blacksquare\]

Problem (F'03, #2). Let \( f : \mathbb{R} \rightarrow \mathbb{R} \) be infinitely differentiable function. Assume that \( \forall x \in [0, 1], \exists m > 0 \), such that \( f^{(m)}(x) \neq 0 \).

Prove that \( \exists M \) such that the following stronger statement holds:

\( \forall x \in [0, 1], \exists m > 0, \ m \leq M \) such that \( f^{(m)}(x) \neq 0 \).

Proof. There are uncountably many \( x_\alpha \)'s in \( [0, 1] \), and for each \( x_\alpha \in [0, 1], \exists m_\alpha > 0 \) such that \( f^{(m_\alpha)}(x) \neq 0 \) for \([x_\alpha - \epsilon_\alpha, x_\alpha + \epsilon_\alpha]\), for some \( \epsilon_\alpha > 0 \) (since \( f^{(m_\alpha)} \) is continuous). Let \( \epsilon = \min_\alpha(\epsilon_\alpha) \). Partition \([0, 1]\) into \( n = 1/\epsilon \) subintervals (since \( \epsilon > 0 \), \( n < \infty \)):

\[
0 < \epsilon = x_0 < x_1 < \ldots < x_n = 1 - \epsilon < 1,
\]

such that \( x_0 = \epsilon, \ x_i = x_0 + i\epsilon, \ i = 1, \ldots, n \). Thus \([0, 1]\) is covered by finitely many overlapping intervals \([x_i - \epsilon, x_i + \epsilon]\). For each \( x_i, \ i = 1, \ldots, n, \ \exists m_i > 0 \) such that \( f^{(m_i)}(x_i) \neq 0 \) on \([x_i - \epsilon, x_i + \epsilon]\). Take \( M = \max_{0 \leq i \leq n}(m_i) \). Thus, \( \forall x \in [0, 1], \exists m > 0, \ m \leq M \) such that \( f^{(m)}(x) \neq 0 \).
Problem (S’03, #4). Consider the following equation for a function $F(x, y)$ on $\mathbb{R}^2$:
\[
\frac{\partial^2 F}{\partial x^2} = \frac{\partial^2 F}{\partial y^2} \quad \mathcal{Q}
\]
(a) Show that if a function $F$ has the form $F(x, y) = f(x+y) + g(x-y)$ where $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are twice differentiable, then $F$ satisfies the equation $\mathcal{Q}$.
(b) Show that if $F(x, y) = ax^2 + bxy + cy^2$, $a, b, c \in \mathbb{R}$, satisfies $\mathcal{Q}$ then $F(x, y) = f(x+y) + g(x-y)$ for some polynomials $f$ and $g$ in one variable.

Proof. (a) Let $\xi(x, y) = x + y$, $\eta(x, y) = x - y$, so $F(x, y) = f(\xi(x, y)) + g(\eta(x, y))$. By Chain Rule,
\[
\frac{\partial F}{\partial x} = \frac{df}{d\xi} \cdot \frac{\partial \xi}{\partial x} + \frac{dg}{d\eta} \cdot \frac{\partial \eta}{\partial x} = \frac{df}{d\xi} + \frac{dg}{d\eta}, \quad \frac{\partial^2 F}{\partial x^2} = \frac{d^2 f}{d\xi^2} + \frac{d^2 g}{d\eta^2},
\]
and similarly $\frac{\partial F}{\partial y} = \frac{df}{d\xi} - \frac{dg}{d\eta}$ and $\frac{\partial^2 F}{\partial x^2} = \frac{d^2 f}{d\xi^2} + \frac{d^2 g}{d\eta^2}$, and thus, $\frac{\partial^2 F}{\partial x^2} = \frac{\partial^2 F}{\partial y^2}$.

(b) Suppose $F(x, y) = ax^2 + bxy + cy^2$, $a, b, c \in \mathbb{R}$ satisfies $\mathcal{Q}$, then
\[
\frac{\partial F}{\partial x} = 2ax + by \Rightarrow \frac{\partial^2 F}{\partial x^2} = 2a, \quad \frac{\partial F}{\partial y} = 2cy + bx \Rightarrow \frac{\partial^2 F}{\partial y^2} = 2c \Rightarrow a = c.
\]

$F(x, y) = ax^2 + bxy + ay^2 = a(x^2 + y^2) + bxy = a\left(\frac{(x+y)^2 + (x-y)^2}{2} + b\left(\frac{(x+y)^2 - (x-y)^2}{4}\right)ight)$. \hfill \Box

12.5 Lagrange Multipliers

Theorem. Let $f$ and $g$ be $C^1$ on $\mathbb{R}^2$. Suppose that $f$ attains its maximum or minimum value on the zero set $S$ of $g$ at the point $p$ where $\nabla g(p) \neq 0$. Then
\[
\nabla f(p) = \lambda \nabla g(p)
\]
for some number $\lambda$.

Problem (S’03, #5). Consider the function $F(x, y) = ax^2 + 2bxy + cy^2$ on the set $A = \{(x, y) : x^2 + y^2 = 1\}$.
(a) Show that $F$ has a maximum and minimum on $A$.
(b) Use Lagrange multipliers to show that if the maximum of $F$ on $A$ occurs at a point $(x_0, y_0)$, then the vector $(x_0, y_0)$ is an eigenvector of the matrix $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

Proof. (a) Since $F$ is continuous and the circle is closed and bounded, $F$ attains both a maximum and minimum values on the unit circle $g(x, y) = x^2 + y^2 - 1 = 0$.

(b) Applying the above theorem, we obtain the three equations (for $x, y, \lambda$)
\[
2ax + 2by = 2\lambda x, \quad 2bx + 2cy = 2\lambda y, \quad x^2 + y^2 = 1.
\]

$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$, and $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ is an eigenvector of this matrix. \hfill \Box
13 Successive Approximations and Implicit Functions

13.1 Contraction Mappings

$C \subseteq \mathbb{R}^n$. The mapping $\varphi : C \to C$ is a contraction mapping with contraction constant $k < 1$ if

$$|\varphi(x) - \varphi(y)| \leq k|x - y| \quad \forall x, y \in C.$$

Contraction Lemma. If $M$ is complete and $\varphi : M \to M$ is a contraction mapping with $k < 1$, then $\varphi$ has a unique fixed point $x^*$. 

Proof. $|x_{n+1} - x_n| = |\varphi(x_n) - \varphi(x_{n-1})| \leq k|x_n - x_{n-1}| \leq k^n|x_1 - x_0|$. If $m > n > 0$, then

$$|x_m - x_n| \leq |x_m - x_{m-1}| + \cdots + |x_{n+1} - x_n| \leq (k^{m-1} + \cdots + k^n)|x_1 - x_0| \leq \frac{k^n}{1-k}|x_1 - x_0|.$$

Thus the sequence $\{x_n\}$ is a Cauchy sequence, and therefore converges to a point $x^*$. Letting $m \to \infty$, we get

$$|x^* - x_n| \leq \frac{k^n}{1-k}|x_1 - x_0|.$$

Since $\varphi$ is contraction, it is continuous. Therefore,

$$\varphi(x^*) = \lim_{n \to \infty} \varphi(x_n) = \lim_{n \to \infty} x_{n+1} = x^*.$$

If $x^{**}$ were another fixed point of $\varphi$, we would have $|x^* - x^{**}| = |\varphi(x^*) - \varphi(x^{**})| \leq k|x^* - x^{**}|$. Since $k < 1$, it follows that $x^* = x^{**}$, so $x^*$ is the unique fixed point of $\varphi$. \[\square\]

13.2 Inverse Function Theorem

Lemma. \(^{17}\) Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is $C^1$, $0 \in W$, $f(0) = 0$, $f'(0) = I$. Suppose $||f'(x) - I|| \leq \epsilon$, $\forall x \in B_r$. Then

$$B_{(1-\epsilon)r} \subseteq f(B_r) \subseteq B_{(1+\epsilon)r}. \quad (13.1)$$

If $U = B_r \cap f^{-1}(B_{(1-\epsilon)r})$, then $f : U \to B_{(1-\epsilon)r}$ is bijection, and the inverse mapping $g : V \to U$ is differentiable at $0$.

Proof. $||f'(x) - I|| \leq \epsilon < 1$, $\forall x \in B_r$. Apply MVT\(^{18}\) to $g(x) = f(x) - x$. $a, b \in B_r$, then

$$||f(b) - f(a)|| = ||g(b) - g(a)|| \leq ||g'(\xi)|| ||b - a|| = ||f'(\xi) - I|| ||b - a|| \leq \epsilon ||b - a|| \quad (13.2)$$

$$||b - a|| \leq ||f(b) - f(a)|| \leq (1 + \epsilon) ||b - a|| \quad (13.3)$$

The left-hand inequality shows that $f$ is 1-1 on $B_r$. The right-hand inequality (with $a = 0$) shows $f(B_r) \subseteq B_{(1+\epsilon)r}$.

\(^{17}\)Problem F’01, # 6.

\(^{18}\) If $U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is $C^1$ then $||f(b) - f(a)|| \leq || b - a || \max_{x \in L} || f'(x) ||$. 

To show \( B_{(1-\epsilon)r} \subseteq f(B_r) \), we use contraction mapping theorem. Given \( y \in B_{(1-\epsilon)r} \), define \( \varphi : \mathbb{R}^n \to \mathbb{R}^n \) by
\[
\varphi(x) = x - f(x) + y.
\]

We want to show that \( \varphi \) is a contraction mapping of \( B_r \); its unique fixed point will then be the desired point \( x \in B_r \) such that \( f(x) = y \).

To see that \( \varphi \) maps \( B_r \) into itself:
\[
||\varphi(x)|| = ||f(x) - x + y|| \leq (by (13.2) with \( a = 0 \)) \leq \epsilon ||x|| + (1-\epsilon)r \leq \epsilon r + (1-\epsilon)r = r,
\]
so if \( x \in B_r \), then \( \varphi(x) \in B_r \). Thus, \( \varphi(B_r) \subseteq B_r \).

To see \( \varphi : B_r \to B_r \) is a contraction mapping, note that
\[
||\varphi(y) - \varphi(z)|| = ||f(y) - f(z) - (y - z)|| \leq \epsilon ||y - z||
\]

Thus, \( \varphi \) has a unique fixed point \( x^* \), \( \varphi(x^*) = x^* \), such that \( f(x^*) = y \). From the statement of theorem, \( U = B_r \bigcap f^{-1}(B_{(1-\epsilon)r}) \). \( f \) is a bijection of \( U \) onto \( B_{(1-\epsilon)r} \).

It remains to show that \( g : V \to U \) is differentiable at \( 0 \), where \( g(0) = 0 \). Need to show
\[
\lim_{|h| \to 0} \frac{|g(h) - g(0) - h|}{|h|} = \lim_{|h| \to 0} \frac{|g(h) - h|}{|h|} = 0.
\]

This will prove that \( g'(0) = I \). Applying (13.2) with \( a = 0 \), \( b = g(0) \), \( h = f(b) \), we get
\[
||g(h) - h|| \leq \epsilon ||b|| \leq \text{by (13.3)} \leq \frac{\epsilon}{1-\epsilon} ||f(b)|| = \frac{\epsilon}{1-\epsilon} ||h||
\]

Therefore, \( \lim_{|h| \to 0} \frac{|g(h) - h|}{|h|} = 0 \), with \( g'(0) = I \). \( \square \)

**Theorem.** 19 Suppose \( f : W \subseteq \mathbb{R}^n \to \mathbb{R}^n \) is \( C^1 \), \( a \in W \), \( b = f(a) \), and the matrix \( f'(a) \) is nonsingular. 20 Then there exist \(^{21}\) open sets \( U \subseteq W \) of \( a \) and \( V \) of \( b \), such that \( f \) maps \( U \) bijectively onto \( V \). ( \( \exists \) 1-1 \( C^1 \) mapping \( g : V \to W \) such that
\[
\begin{align*}
g(f(x)) &= x \ 	ext{ for } x \in U, \\
g(f(y)) &= y \ 	ext{ for } y \in V.
\end{align*}
\]

Also, for all \( y \in V \) (\( y = f(x) \)), \( g = f^{-1} \) satisfies \( g'(y) = g'(f(x)) = f'(x)^{-1} \).

**Proof.** Fix \( a \in U \) and let \( b = f(a) \). Put \( T = f'(a) \), a matrix / linear map. Define
\[
\tilde{f}(h) = T^{-1}(f(a + h) - b).
\]

Note \( \tilde{f}(0) = T^{-1} \cdot (0) = 0 \);
\( \tilde{f}'(0) = T^{-1}f'(a) = T^{-1}T = I \). Thus, by previous Lemma, \( \exists U_0 \) open, \( 0 \in U_0 \), such that \( \tilde{f} : U_0 \to V_0 \) is a bijection, \( 0 \in V_0 \).

\( \tilde{f} \) maps \( U_0 \) bijectively onto \( V_0 \) containing \( 0 \). Lets express \( f \) in terms of \( \tilde{f} \).
\[
f(a + h) - f(a) = T\tilde{f}(h).
\]

\( f(x) - f(a) = T\tilde{f}(x - a) \).

---

19 \( f \) is locally one-to-one in \( E \) \( \equiv \) each point \( x \in E \) has a neighborhood in which \( f \) is 1-1.

20 Jacobian of \( f \) is \( | \det f'(a) | = | \frac{\partial f}{\partial x}(a) | \neq 0 \).

21 i.e. A \( C^1 \) map \( f : W \to V \) is locally invertible at \( a \) \( \equiv \) there exist open sets \( U \subseteq W \) of \( a \) and \( V \) of \( b = f(a) \), and a \( C^1 \) map \( g : V \to U \) such that \( f \) and \( g \) are inverse to each other.
\[
f(x) = T \tilde{f}(x - a) + f(a),
\]
\[
f : U_0 + a \to T \tilde{f}(V_0) + f(a), \quad \text{bijection.}
\]

Let’s compute \( f^{-1} \):

Let \( y = T \tilde{f}(x - a) + f(a) \)

\[
T \tilde{f}(x - a) = y - f(a)
\]

\[
\tilde{f}(x - a) = T^{-1}(y - f(a))
\]

\[
x - a = \tilde{f}^{-1}(T^{-1}(y - f(a))) = \tilde{g}(T^{-1}(y - f(a)))
\]

\[
f^{-1}(y) = x = \tilde{g}(T^{-1}(y - f(a))) + a
\]

\[
(f^{-1})'(y) = (\tilde{g})'(T^{-1}(y - f(a))) \cdot T^{-1}
\]

\[
(f^{-1})'(b) = (\tilde{g})'(0) \cdot T^{-1} = T^{-1} = f'(a)^{-1}.
\]

\[\square\]
Problem (S’02, #7; W’02, #7; F’03, #6). Suppose $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is $C^1$ and that the Jacobian matrix of $F$ is everywhere nonsingular. Suppose that $F(0) = 0$ and that $||F(x,y)|| \geq 1$ for all $(x,y)$ with $||(x,y)|| = 1$. Denote $U = \{(x,y) : ||(x,y)|| < 1\}$. Prove that $F(U) \supset U$.

Hint: Show that $F(U) \cap U$ is both open and closed in $U$.

Proof. Since $U$ is connected, clopeness of $F(U) \cap U$ in $U$ implies that either $F(U) \cap U = U$ or $F(U) \cap U = \phi$. Since there exists a point, namely $0$ such that it is inside both $U$ and $F(U)$, $F(U) \cap U$ cannot be empty, and thus clopeness of $F(U) \cap U$ in $U$ would imply that $F(U) \cap U = U$ (which would mean $U \subseteq F(U)$).

1) Show $F(U) \cap U$ is open in $U$.

$F(U)$ is open in $\mathbb{R}^2$. Say $y_0 \in F(U)$, $y_0 = F(x_0)$, $F'(x_0)$ invertible. By inverse function thm, $F$ maps open set $U_0$ onto open set $V_0$; $x \in U_0 \Rightarrow y_0 = F(x_0) \in V_0$. $y_0 \in V_0 \subseteq F(U) \Rightarrow F(U) \cap U$ is open in $U$.

2) Show $F(U) \cap U$ is closed in $U$.

Say $x_n \in F(U) \cap U$, $x_n \rightarrow x^* \in U$. $x_n = F(y_n)$, $y_n \in U \subseteq U$.

There is a subsequence $y_{n_k} \rightarrow y \in \overline{U}$.

Since $F$ is continuous, $F(y_{n_k}) \rightarrow F(y) = x^*$.

$||y|| = 1 \Rightarrow ||F(y)|| \geq 1 \Rightarrow F(y) = x^* \notin U$. Contradiction.

13.3 Implicit Function Theorem

$(x,y) = (x_1, \ldots, x_m, y_1, \ldots, y_n)$

Theorem. Suppose $G : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$ is $C^1$. $G(a, b) = 0$ for some point $(a, b)$. Partial derivative matrix $\frac{\partial G}{\partial y}(a, b)$ is invertible.

Then there exist open sets $U$ of $a$ in $\mathbb{R}^m$ and $W$ of $(a, b)$ in $\mathbb{R}^{m+n}$ and a $C^1$ mapping $h : U \rightarrow \mathbb{R}^n$, such that $y = h(x)$ solves the equation $G(x, y) = 0$ in $W$.

Example. $f(x, y, z) = 0, \frac{\partial f}{\partial z} \neq 0 \Rightarrow$ can solve for $z = h(x, y)$, $f(x, y, h(x, y)) \equiv 0$.

Example. $m = n = 1$. $G(x, y) = 0, y = h(x) \Rightarrow G(x, h(x)) = 0 \Rightarrow \frac{d}{dx} G(x, h(x)) = 0 \Rightarrow \frac{\partial G}{\partial x} + \frac{\partial G}{\partial h} h'(x) = 0 \Rightarrow h'(x) = -\frac{\partial G}{\partial x}/\frac{\partial G}{\partial h} \neq 0$.

Say $G(x, y) = x^2 + y^2 - 1, x^2 + y^2 - 1 = 0 \Rightarrow y = \pm \sqrt{1-x^2}$, problem at $(1,0)$.

$\frac{\partial G}{\partial h}$ cannot be equal to 0.

Problem (S’02, #6). Suppose $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ is $C^1$ with $\nabla f \neq 0$ at $0$. Show that there are two other $C^1$ functions $g, h : \mathbb{R}^3 \rightarrow \mathbb{R}$, such that the function

$(x, y, z) \rightarrow (f(x, y, z), g(x, y, z), h(x, y, z))$

from $\mathbb{R}^3$ to $\mathbb{R}^3$ is one-to-one on some neighborhood of $0$.

Proof. $\nabla f \neq 0 \Rightarrow$ one of $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$ is not 0. $F = (f, g, h)$.

$$\nabla F = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{bmatrix}$$

Need to produce functions $g, h$ such that the matrix above is invertible.

If $\frac{\partial f}{\partial x}(0) \neq 0$, let $g(x, y, z) = z, h(x, y, z) = y$. Then

$$\begin{bmatrix} \frac{\partial f}{\partial x} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad F = F(f(x, y, z), z, y).$$
Similarly, we can find a set of functions $g, h$ by choosing a matrix in each of the other two cases, i.e. when $\frac{\partial f}{\partial y}(0) \neq 0$ and $\frac{\partial f}{\partial z}(0) \neq 0$.

**Problem (F’02, #6).** Suppose $F : \mathbb{R}^3 \to \mathbb{R}^2$ is $C^1$. Suppose for some $v_0 \in \mathbb{R}^3$ and $x_0 \in \mathbb{R}^2$ that $F(v_0) = x_0$ and $F'(v_0) : \mathbb{R}^3 \to \mathbb{R}^2$ is onto. Show that there is a $C^1$ function $\gamma : (-\epsilon, \epsilon) \to \mathbb{R}^3$ for some $\epsilon > 0$, such that

(i) $\gamma'(0) \neq 0 \in \mathbb{R}^3$, and

(ii) $F(\gamma(t)) = x_0$ for all $t \in (-\epsilon, \epsilon)$.

**Proof.** Since $F'(v_0)$ is onto, the matrix $F'(v_0)$ has rank 2. So, 2 of the 3 columns of $F'(v_0)$ are linearly independent.

$$F'(v_0) = \begin{bmatrix}
\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\
\frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3}
\end{bmatrix}_{v_0}.$$ 

Assume the last two columns are linearly independent.

Consider the function $G(x_1, x_2, x_3) = F(x_1, x_2, x_3) - x_0$. Write $(x_1, x_2, x_3) = (s_1, s_2)$ where $s_1 = x_1$, $s_2 = (x_2, x_3)$. Write $v_0 = (u_1, u_2)$. Then $G(u_1, u_2) = 0$ and $\frac{\partial G}{\partial s_2}(v_0)$ is invertible. By Implicit Function Theorem, $\exists \epsilon > 0$ and $h \in C^1$, such that $h : (u_1 - \epsilon, u_1 + \epsilon) \to \mathbb{R}^2$ and $G(s_1, h(s_1)) = 0$, $\forall s_1 \in (u_1 - \epsilon, u_1 + \epsilon)$.

$\Rightarrow$ $F(s_1, h(s_1)) = x_0$, $\forall s_1 \in (u_1 - \epsilon, u_1 + \epsilon)$.

Define $\gamma(t) : (-\epsilon, \epsilon) \to \mathbb{R}^3$ by $\gamma(t) = (u_1 + t, h(u_1 + t))$. Then $\gamma(t)$ is a differentiable curve satisfying

(i) $\gamma'(t) = (1, h'(u_1 + t)) \neq 0$,

(ii) $F(\gamma(t)) = x_0$ for all $t \in (-\epsilon, \epsilon)$. 

$\square$
13.4 Differentiation Under Integral Sign

Problem (W’02, #1). \( f : [a, b] \times [c, d] \to \mathbb{R} \) \((x, y) \mapsto f(x, y))\). Suppose \( \frac{\partial f}{\partial y} \) exists on \([a, b] \times (c, d)\) and extends to a continuous function on \([a, b] \times [c, d]\). Let

\[
F(y) = \int_a^b f(x, y) \, dx.
\]

Then \( F \) is differentiable in \([a, b]\) and

\[
\frac{d}{dy} F(y) = \int_a^b \frac{\partial f}{\partial y} (x, y) \, dx.
\]

Proof.

\[
\left| \frac{F(y+h) - F(y)}{h} - \int_a^b \frac{\partial f}{\partial y} (x, y) \, dx \right| = \left| \int_a^b \left[ \frac{f(x, y+h) - f(x, y)}{h} - \frac{\partial f}{\partial y} (x, y) \right] \, dx \right|
\]

\[
\leq \int_a^b \left| \frac{f(x, y+h) - f(x, y)}{h} - \frac{\partial f}{\partial y} (x, y) \right| \, dx \leq \Rightarrow \quad \star
\]

By MVT, \( \exists c, \ 0 < c < 1 \), such that

\[
\frac{f(x, y+h) - f(x, y)}{h} = \frac{\partial f}{\partial y} (x, y + ch).
\]

\(|(x, y + ch) - (x, y)| = ch \leq h.
\]

Since \( \frac{\partial f}{\partial y} \) is continuous on \([a, b] \times [c, d]\) \( \Rightarrow \) \( \frac{\partial f}{\partial y} \) is uniformly continuous on \([a, b] \times [c, d]\).

Choose \( \delta \) such that \(|(x, y) - (x', y')| \leq \delta \) \( \Rightarrow \)

\[
\left| \frac{\partial f}{\partial y} (x', y') - \frac{\partial f}{\partial y} (x, y) \right| \leq \frac{\epsilon}{b-a}
\]

\[
\Rightarrow \quad \star \quad = \int_a^b \left| \frac{\partial f}{\partial y} (x, y + ch) - \frac{\partial f}{\partial y} (x, y) \right| \, dx \leq \int_a^b \frac{\epsilon}{b-a} \, dx = (b-a) \frac{\epsilon}{b-a} = \epsilon.
\]

\( \square \)