Analysis Comp Study Guide

The Real and Complex Number Systems and Basic Topology

Theorem 1 (Cauchy-Schwarz Inequality)**.**

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

Proof:
As $\sum_{k=1}^{n}(a_k - tb_k)^2 \geq 0$, as a quadratic in t it has at most one root. So the discriminant must therefore be less than or equal to 0. As $\sum_{k=1}^{n}(a_k - tb_k)^2 = \sum_{k=1}^{n} a_k^2 - 2t\sum_{k=1}^{n} a_k b_k + t^2 \sum_{k=1}^{n} b_k^2$, the discriminant is

$$4\left(\sum_{k=1}^{n} a_k b_k\right)^2 - 4\sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2 \leq 0,$$

or, after rearrangement, we get

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

Note: Equality holds only when $a_k = b_k$.

Theorem 2 (The Archimedean Property)**.**

If $x, y \in \mathbb{R}$ and $x > 0$, then there exists a positive integer n such that $nx > y$.

Proof:
Let A be the set of all nx, where n runs through the positive integers and assume toward contradiction that y is an upper bound for A. A has a least upper bound, so let $\alpha = \sup A$. Since $x > 0$, $\alpha - x < \alpha$, and $\alpha - x$ is not an upper bound of A. Hence $\alpha - x < mx$ for some positive integer m. But then $\alpha < (m+1)x \in A$ which is impossible as α is an upper bound of A. So there exists a positive integer n such that $nx > y$.

Theorem 3 ($\mathbb{Q}$ is dense in $\mathbb{R}$)**.**

If $x, y \in \mathbb{R}$, $x < y$, then there exists a $p \in \mathbb{Q}$ such that $x < p < y$.

Proof:
$y - x > 0$ so, by the Archimedean property, there exists an integer $n \geq 1$ such that $n(y - x) > 1$. Again by the Archimedean property, there exists an integer m such that $m - 1 \leq nx < m$. So combining these two inequalties, we get $nx < m \leq nx + 1 < ny$. As $n > 0$ we may divide and obtain:

$$x < \frac{m}{n} < y,$$

where $\frac{m}{n} \in \mathbb{Q}$.

Theorem 4 (Heine-Borel Theorem)**.**

If a set $E \subseteq \mathbb{R}^k$ has one of the following properties, it has the other two as well:

(a) E is closed and bounded.

(b) E is compact.

(c) Every infinite subset of E has a limit point in E.

Our proof requires two lemmas in order to proceed.

Lemma 1 (Closed subsets of compact sets are compact)**.**

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 $\mathcal{O}$ of K. Since K is compact, there is a finite subcollection $\{\mathcal{U}_n\}$ of $\mathcal{O}$ which covers K and hence F. If F^c is a member of $\{\mathcal{U}_n\}$, we may remove it from $\{\mathcal{U}_n\}$ and still retain an open cover of F. We have thus shown that a finite subcollection of $\{V_\alpha\}$ covers F.

Lemma 2 (If K_α 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)**.**

Fix a member K_1 of $\{K_\alpha\}$ and put $G_\alpha = K_\alpha^c$. Assume that no point of K_1 belongs to every K_α. Then the sets G_α form an open cover of K_1. Since K_1 is compact there exists a subcover $K_1 \subset \{G_1, \ldots, G_n\}$. However, this means that

$$K_1 \bigcap K_2 \bigcap \cdots \bigcap K_n$$

is empty, contradicting our hypothesis.

Lemma 3 (Every k-cell is compact)**.**

Let I be a k-cell, consisting of all points $x = (x_1, \ldots, x_k)$ such that $a_j \leq x_j \leq b_j$ $(1 \leq j \leq k)$, Put

$$\delta = \left\{ \sum_{j=1}^{k} (b_j - a_j)^2 \right\}^2 .$$

Then $|x - y| < \delta$ if $x, y \in I$.
Suppose, toward contradiction, there exists an open cover $\{G_\alpha\}$ of I which contains no finite subcover of I. Put $c_j = (a_j + b_j)/2$. The intervals $[a_j, c_j]$ and $[c_j, b_j]$ then determine $2^k k$-cells Q_i whose union is I. At least one of these, call it I_1 cannot be covered by any finite subcollection of $\{G_\alpha\}$. We can next subdivide I_1 and continue this process. We obtain a sequence $\{I_n\}$ with the following properties:

(i) $I \supset I_1 \supset I_2 \cdots$;

(ii) I_n is not covered by any finite subcollection of $\{G_\alpha\}$;

(iii) if $x \in I_n$ and $y \in I_n$, then $|x - y| < 2^{-n}\delta$.

By (i) and Lemma 2, there is a point x^* which lies in every I_n. For some α, $x^* \in G_\alpha$. Since G_α is open, there exists an $r > 0$ such that $|y - x^*| < r$ implies $y \in G_\alpha$. There exists an n such that $2^{-n}\delta < r$, so by (iii) $I_n \subset G_\alpha$, which contradicts (ii).

Proof of Heine-Borel Theorem:
(**(a)** $\to$ **(b)**) As E is bounded, $E \subset I$, where I is a k-cell. By Lemma 3, I is compact. By Lemma 1, as E is closed and $E \subset I$, E is compact.
(**(b)** $\to$ **(c)**) Let U be an infinite subset of E (if E has only finitely many points the result is trivial). If no point of E is a limit point of U, then each $q \in E$ would have a neighborhood V_q which contains at most one point of U, specifically q if $q \in E$. As E is infinite, no finite subcollection of $\{V_q\}$ can cover U and the same is true of E as $U \subset E$. This contradicts the compactness of E. So U must have a limit point in E.
(**(c)** $\to$ **(a)**) If E is not bounded then E contains points x_n such that $|x_n| > n$, $n = 1, 2, \ldots$. The set S consisting of these points x_n is infinite and clearly has no limit point in $\mathbb{R}^k$ and hence has none in E. So (c) implies E is bounded.
If E is not closed, then there is a point $x_0 \in \mathbb{R}^k$ which is a limit point of E but not a point of E. For $n = 1, 2, 3, \ldots$ there are points $x_n \in E$ such that $|x_n - x_0| < 1/n$. Let S be the set of the points x_n. Then S is infinite (or else we could find a ball around x_0 not containing any points of E), contains x_0 as a limit point and S has no other limit points in $\mathbb{R}^k$. For if $y \in \mathbb{R}^k, y \neq x_0$, then

$$|x_n - y| \geq |x_0 - y| - |x_n - x_0| \geq |x_0 - y| - \frac{1}{n} \geq \frac{1}{2}|x_0 - y|$$

for all but finitely many n. So y is not a limit point of S. Thus S has no limit point in E. By (c) we get then that E must be closed.

Counter Example 1 (If K_α is a collection of closed subsets of a metric space X such that the intersection of every finite subcollection of $\{K_\alpha\}$ is nonempty, then $\bigcap K_\alpha$ is not necessarily nonempty)**.**

Let $K_n = [n, \infty)$. So each K_n is closed and a finite intersection of $K_i s$ is nonempty. However, if we assume there exists some $x \in \bigcap K_n$, then, by the Archimedean property, there exists an integer $N > x$ and so $x \notin K_N$ so $x \notin \bigcap K_n$.

Counter Example 2 (Closed And Bounded Does Not Imply Compact In A General Metric Space)**.**

Let X be the discrete metric with infinitely many points. Then X itself is closed as it is the entire metric space. Similarly, for any $x \in X$, if we look at $B(x, 2)$ this bounds X as for any $y \in X$, $d(x, y) \leq 1$. So X is closed and bounded. However, if we take the set of $\{B(x, 1/2) : x \in X\}$, then this covers X however no finite subcover (in fact, no proper subcover) can cover X. So X is not compact.

Theorem 5 (Bolzano-Weierstrass)**.**

Every bounded infinite subset of $\mathbb{R}^k$ has a limit point in $\mathbb{R}^k$.

Proof:
Being bounded, the set E in question is a subset of a k-cell $I \subset R^k$. By Lemma 3 above, I is compact and so E has a limit point in I by the proof above, we have that E has a limit point in I.

Numerical Sequences and Series

Theorem 6 (Cauchy Sequence Condition)**.**

(a) In any metric space X, every convergent sequence is a Cauchy sequence.

(b) If X is a compact metric space and if $\{p_n\}$ is a Cauchy sequence in X, then $\{p_n\}$ converges to some point in X.

(c) In $\mathbb{R}^k$, every Cauchy sequence converges.

Lemma 4 (If $\overline{E}$ is the closure of a set E in a metric space X, then diam$\overline{E}$ =diam$E$.)**.**

Since $E \subset \overline{E}$, diam$E \leq$ diam$\overline{E}$. Fix $\epsilon > 0$ and choose $p, q \in \overline{E}$. By the definition of $\overline{E}$, there are points $p', q' \in E$ such that $d(p, p') < \epsilon$, $d(q, q') < \epsilon$. Hence

$$d(p, q) \leq d(p, p') + d(p', q') + d(q', q) < 2\epsilon + \text{diam}E.$$

It follows that
diam$\overline{E} \leq 2\epsilon +$ diam$E$. As ϵ was arbitrary, the lemma is proved.

Proof:

(a) If $p_n \to p$ and if $\epsilon > 0$, there is an integer N such that $d(p, p_n) < \epsilon/2$ for all $n \geq N$. Hence

$$d(p_n, p_m) \leq d(p_n, p) + d(p, p_m) < \epsilon$$

for $n, m \geq N$. Thus $\{p_n\}$ is a Cauchy sequence.

(b) Let $\{p_n\}$ be a Cauchy sequence in the compact space X. For $N = 1, 2, 3, \ldots$ let E_N be the set consisting of $p_N, p_{N+1}, \ldots$. Then

$$\lim_{N\to\infty} \text{diam}\overline{E}_N = 0,$$

by Lemma 4. Being a closed subset of the compact space X, each $\overline{E}_N$ is compact by Lemma 1. Also, $E_N \supset E_{N+1}$, so $\overline{E}_N \supset \overline{E}_{N+1}$.

Define $E = \bigcap_{n=1}^{\infty} \overline{E}_n$. By Lemma 2, E is nonempty shows now that there is a unique $p \in X$ which lies in every $\overline{E}_N$.

Let $\epsilon > 0$ be given. As diam$\overline{E}_N \to 0$ there is an integer N_0 such that diam$\overline{E}_N < \epsilon$ for $N_0 \geq N$. Since $p \in \overline{E}_N$, it follows that $d(p, q) < \epsilon$ for every $q \in \overline{E}_N$ hence for every $q \in E_N$. In other words, $d(p, p_n) < \epsilon$ if $n \geq N_0$. This says exactly that $p_n \to p$.

(c) Let $\{x_n\}$ be a sequence in $\mathbb{R}^k$. Define E_N as in **(b)**, with x_i in place of p_i. For some N, diam $E_N < 1$. The range of $\{x_n\}$ is the union of E_N and the finite set $\{x_1, \dots, x_{N-1}\}$. Hence $\{x_n\}$ is bounded. Since every bounded subset of $\mathbb{R}^k$ has compact closure in $\mathbb{R}^k$ (Heine-Borel), **(c)** follows from **(b)**.

Counter Example 3 ($\mathbb{Q}$ has a Cauchy Sequence that does not converge)**.**

Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists a sequence $\{p_n\}$ of rational points in $\mathbb{R}$ converging to $\sqrt{2}$ such that $|p_n - p_{n+1}| < 2^{-n}$. Then in $\mathbb{Q}$, $\{p_n\}$ is a Cauchy sequence since for any $2 > \epsilon > 0$, there exists an N such that $\sum_{k=0}^{N} 2^{-k} > 2 - \epsilon$. So, for $m \geq n > N$,

$$|p_m - p_n| \leq |p_m - p_{m-1}| + \cdots |p_{n+1} - p_n| < \sum_{k=0}^{\infty} 2^{-k} - \sum_{k=0}^{N} 2^{-k} < \epsilon.$$

So our sequence is Cauchy, however $\sqrt{2}$ is not in $\mathbb{Q}$, so this sequence does not converge in $\mathbb{Q}$.

Theorem 7 (Some Special Sequences)**.**

(a) For $p > 0$,

$$\lim_{n\to\infty} \frac{1}{n^p} = 0.$$

Fix $\epsilon > 0$. By the Archimedean property, there is a $n \in \mathbb{Z}^+$ such that $n > (1/\epsilon)^{1/p}$. So

$$\frac{1}{n^p} < \frac{1}{((1/\epsilon)^{1/p})^p} = \epsilon.$$

(b) For $p > 0$,

$$\lim_{n\to\infty} \sqrt[n]{p} = 1.$$

If $p > 1$, let $x_n = \sqrt[n]{p} - 1$. So $x_n > 0$, and $1 + nx_n \leq (1 + x_n)^n = p$, so $0 \leq x_n \leq \frac{p-1}{n}$. So, by the Squeeze Theorem, $x_n \to 0$. If $p = 1$ the result is trivial and if $p < 1$ let $p' = 1/p$ and by above $\sqrt[n]{p'} \to 1$, so $\sqrt[n]{p} \to 1$.

(c) We may assume $n \geq 2$,

$$\lim_{n\to\infty} \sqrt[n]{n} = 1.$$

Let $x_n = \sqrt[n]{n} - 1$. So $x_n \geq 0$. By the Binomial Theorem, $n = (1 + x_n)^n \geq \dfrac{n(n-1)}{2} x_n^2$. So $0 \leq x_n \leq \sqrt{\frac{2}{n-1}}$. So again by the Squeeze Theorem, $x_n \to 0$.

(d) If $p > 0$, $\alpha \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0.$$

Let $k \in \mathbb{Z}^+$ such that $k > \alpha$. For $n > 2k$,

$$(1+p)^n > \binom{n}{k} p^k = \frac{n(n-1)\dots(n-k-1)}{k!} p^k > \frac{n^k p^k}{2^k k!}.$$

So $0 < \dfrac{n^\alpha}{(1+p)^n} < \dfrac{2^k k!}{p^k} n^{\alpha-k}$. As $\alpha - k < 0$ by **(a)**, $n^{\alpha-k} \to 0$.

(e) If $|x| < 1$, then

$$\lim_{n\to\infty} x^n = 0.$$

As $|x| < 1$, $|x| = \dfrac{1}{p+1}$, $p > 0$. By **(d)**, with $\alpha = 0$, we get

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

Theorem 8 (Cauchy Criterion For Sums)**.**

$\sum_{n=1}^{\infty} a_n$ converges if and only if for every $\epsilon > 0$ there is an integer N such that $|\sum_{k=n}^{m} a_k| < \epsilon$ for $m \geq n \geq N$.

Proof:
For a proof of this, treat the partial sums as a sequence and defer to the Cauchy Sequence Conditions.

Theorem 9 (Geometric Series)**.**

For $x \neq 1$ $x > 0$,

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

Proof:
Let $S_n = \sum_{k=0}^{n} x^k$. So $xS_n = S_n - 1 + x^{n+1}$. Rearranging, we get

$$S_n = \frac{1 - x^{n+1}}{1 - x}.$$

Corollary 1 (Inifinite Geometric Series)**.**

For $|x| < 1$,

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

Proof:
If we take the limit

$$\lim_{n\to\infty} S_n = \frac{1}{1 - x}$$

when $|x| < 1$.

Theorem 10 (Convergence Tests for Series)**.**

Given the series $\sum_{n=0}^{\infty} a_n$.

(a) Root Test: Let $\alpha = \limsup_{n\to\infty} \sqrt[n]{|an|}$.
If $\alpha < 1$, the series converges.
If $\alpha > 1$, the series diverges.
If $\alpha = 1$, the result is inconclusive.

Proof:
If $\alpha < 1$, we can choose β such that $\alpha < \beta < 1$ and an integer N such that

$$\sqrt[n]{|a_n|} < \beta$$

for $n \geq N$. That is for $n \geq N$,

$$|a_n| < \beta^n.$$

As $0 < \beta < 1$, $\sum \beta^n$ converges. Convergence of $\sum a_n$ follows from the comparison test.
Note: This actually implies absolute convergence of a_n.
If $\alpha > 1$, then there is a sequence n_k such that

$$\sqrt[n_k]{|a_{n_k}|} \to \alpha.$$

Hence $|a_n| > 1$ for infinitely many values of n, so $a_n \not\to 0$. So $\sum a_n$ diverges.
Consider the series

$$\sum_{n=1}^{\infty} \frac{1}{n}, \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

For both of these series, $\alpha = 1$, however the first diverges and the second converges.

(b) Ratio Test: Let $\alpha(n) = \left|\frac{a_{n+1}}{a_n}\right|$.
If $\limsup_{n\to\infty} \alpha(n) < 1$, the series converges.
If $\alpha(n) \geq 1$ for $n \geq N$, for some $N \in \mathbb{Z}^+$, the series diverges.

Proof:
If $\limsup_{n\to\infty} \alpha(n) < 1$, we can find $\beta < 1$ and an integer N such that

$$\left|\frac{a_{n+1}}{a_n}\right| < \beta$$

for $n \geq N$. In particular, $|a_{N+1}| < \beta|a_n|, \ldots, |a_{N+p}| < \beta^p|a_n|$. That is, $|a_n| < |a_N|\beta^{-N}\beta^n$ for all $n \geq N$ and our series is convergent by the comparison test since $\sum \beta^n$ converges.
If $|a_{n+1}| \geq |a_n|$ for $n \geq N$, then $a_n \not\to 0$, so our series diverges.

(c) Integral Test:

If $N \in \mathbb{Z}$ and f is a monotone decreasing function on $[N, \infty)$, then $\sum_{n=N}^{\infty} f(n)$ converges if and only if $\int_N^{\infty} f(x)dx$ converges to a finite value.

Proof:
Since f is a monotone decreasing Riemann-integrable function, we have $f(x) \leq f(n)$ for $x \in [n, \infty)$ and $f(n) \leq f(x)$ for $x \in [N, n]$. So, for every $n \geq N$,

$$\int_n^{n+1} f(x)dx \leq \int_n^{n+1} f(n)dx = f(n) = \int_{n-1}^{n} f(n)dx \leq \int_{n-1}^{n} f(x)dx.$$

Since the lower estimate is also valid for $f(N)$, we get by summation over all n from N to some larger integer M,

$$\int_N^{M+1} f(x)dx \leq \sum_{n=N}^{M} f(n) \leq f(N) + \int_N^{M} f(x)dx.$$

Letting M tend to infinity, we get the desired result.

Theorem 11 (Power Series)**.**

Given the series $\sum_{n=0}^{\infty} c_n z^n$. Let $\alpha = \limsup_{n\to\infty} \sqrt[n]{c_n}$, $R = \frac{1}{\alpha}$.

Then $\sum_{n=0}^{\infty} c_n z^n$ converges for $|z| < R$ and diverges if $|z| > R$.

Proof:
Proof follows directly from the Ratio Test.

Theorem 12 (Summation by Parts)**.**

Given two sequences $\{a_n\}$,$\{b_n\}$, put

$$A_n = \sum_{k=0}^{n} a_k$$

if $n \geq 0$; put $A_{-1} = 0$. Then, if $0 \leq p \leq q$, we have

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p.$$

Proof:

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q} (A_n - A_{n-1}) b_n = \sum_{n=p}^{q} A_n b_n - \sum_{n=p-1}^{q-1} A_n b_{n+1} = \sum_{n=p}^{q} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p.$$

Continuity

Definition 1.

Let X, Y be metric spaces $E \subset X$, $p \in E$, $f : E \to Y$. f is continuous at p if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$d_Y(f(x), f(p)) < \epsilon$$

for all points $x \in E$ for which $d_X(x, p) < \delta$.

Counter Example 4 (A Nowhere Continuous Function)**.**

Let $f(x) = \begin{cases} 1 & \text{if} \quad x \in \mathbb{Q} \\ 0 & \text{if} \quad x \in \mathbb{Q}^c \end{cases}$

- Case 1: $x \in \mathbb{Q}$.
  Let $0 < \epsilon < 1$. For any $\delta > 0$, as $\mathbb{Q}^c$ is dense in $\mathbb{R}$, there exists an irrational point y such that $|x - y| < \delta$. However, $|f(x) - f(y)| = 1 > \epsilon$, so f is not continuous at x.

- Case 2: $x \in \mathbb{Q}^c$.
  Let $0 < \epsilon < 1$. For any $\delta > 0$, as $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists a rational point y such that $|x - y| < \delta$. However, $|f(x) - f(y)| = 1 > \epsilon$, so f is not continuous at x.
  So f is discontinuous at all points in $\mathbb{R}$.

Theorem 13 (Equivalent Definitions of Continuity)**.**

(a) A mapping f of a metric space X into a metric space Y is continuous on X if and only if $f^{-1}(V)$ is open in X for every open set V in Y

(b) For the definition of continuity, if we assume also that p is a limit point of E, then f is continuous at p if and only if $\lim_{x \to p} f(x) = f(p)$.

Proof:
(a) Suppose first that f is continuous on X and V is an open set in Y. Suppose $p \in X$ and $f(p) \in V$. Since V is open, there exists $\epsilon > 0$ such that $y \in V$ if $d_Y(f(p), y) < \epsilon$; and since f is continuous at p there exists $\delta > 0$ such that $d_Y(f(x), f(p)) < \epsilon$ if $d_X(x, p) < \delta$. Thus $x \in f^{-1}(V)$ as soon as $d_X(x, p) < \delta$.
Conversely, suppose $f^{-1}(V)$ is open in X for every open set V in Y. Fix $p \in X$ and $\epsilon > 0$, let V be the set of all $y \in Y$ such that $d_Y(y, f(p)) < \epsilon$. Then V is open; hence $f^{-1}(V)$ is open; hence there exists $\delta > 0$ such that $x \in f^{-1}(V)$ as soon as $d_X(p, x) < \delta$. But if $x \in f^{-1}(V)$, then $f(x) \in V$, so that $d_Y(f(x), f(p)) < \epsilon$. So f is continuous.
(b) Proof follows directly from the definition of limit and continuity.

Theorem 14 (A Continuous Mapping of a Compact Set Is Compact)**.**

Suppose f is a continuous mapping of a compact metric space X into a metric space Y. Then $f(X)$ is compact.

Proof:

Let $\{V_\alpha\}$ be an open cover of $f(X)$. Since f is continuous, each set $f^{-1}(V_\alpha)$ is open. Since X is compact, there are finitely many indices, $\alpha_1, \dots, \alpha_n$ such that

$$X \subset f^{-1}(V_{\alpha_1}) \bigcup \cdots \bigcup f^{-1}(V_{\alpha_n}).$$

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

$$f(X) \subset V_{\alpha_1} \bigcup \cdots \bigcup V_{\alpha_n}.$$

Corollary 2 (Extreme Value Theorem)**.**

Suppose f is a continuous real function on a compact metric space X, and

$$M = \sup_{p \in X} f(p), m = \inf_{p \in X} f(p).$$

Then there exist points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$.

Proof:
By the Theorem above, as f is compact on X, $f(X)$ is compact and thus closed and bounded. So $f(X)$ contains $\sup f(x)$ and $\inf f(X)$ as these exist and are limit points of $f(X)$.

Theorem 15 (Continuity Preserves Connectedness)**.**

If f is a continuous mapping of a metric space X into a metric space Y, and if E is a connected subset of X, then $f(E)$ is connected.

Proof:
Assume, to the contrary, that $f(E) = A \bigcup B$, where A and B are nonempty separated subsets of Y. Put $G = E \bigcap f^{-1}(A)$, $H = E \bigcap f^{-1}(B)$. Then $E = G \bigcup H$ and neither G nor H is empty.
Since $A \subset \overline{A}$, we have $G \subset f^{-1}(\overline{A})$; the latter set is closed, since f is continuous; hence $\overline{G} \subset f^{-1}(\overline{A})$. It follows that $f(\overline{G}) \subset \overline{A}$. Since $f(H) = B$ and $\overline{A} \bigcap B$ is empty, we conclude that $\overline{G} \bigcap H$ is empty. Similarly, $G \bigcap \overline{H}$ is empty. Thus we have reached a contradiction.

Corollary 3 (Intermediate Value Theorem)**.**

Let f be a continuous real function on the interval $[a, b]$. If $f(a) < f(b)$ and if c is a number such that $f(a) < c < f(b)$, then there exists a point $x \in (a, b)$ such that $f(x) = c$.

Proof:
$[a, b]$ is connected; hence, by the Theorem above $f([a, b])$ is connected subset of $\mathbb{R}$ and the assertion follows by the definition of connectedness in $\mathbb{R}$.

Definition 2 (Uniform Continuity)**.**

Let f be a mapping of a metric space X into a metric space Y. We say that f is uniformly continuous on X if for every $\epsilon > 0$ there exists $\delta > 0$ such that $d_Y(f(x), f(y)) < \epsilon$ for all $x, y \in X$ such that $d_X(x, y) < \delta$.

Counter Example 5 (A Not Uniformly Continuous Function)**.**

Let $f(x) = x^2$. As f is a polynomial, it is continuous on $\mathbb{R}$. Let $\epsilon = 1$ and let $\delta > 0$. By the Archimedean property, there exists an integer n such that $n\delta > 1$. Let $x = n$, $y = n + \delta/2$. So $|x - y| < \delta$. $|f(x) - f(y)| = (n + \delta/2 - n)(n + \delta/2 + n) = n\delta + \delta^2/4 > 1 = \epsilon$. So f cannot be uniformly continuous on $\mathbb{R}$.

Theorem 16 (A Continuous Function Is Uniformly Continuous on a Compact Set)**.**

Let f be a continuous mapping on a compact metric space X into a metric space Y. Then f is uniformly continuous on X.

Proof:

Let $\epsilon > 0$ be given. Since f is continuous, we can associate to each point $p \in X$ a positive number $\phi(p)$ such that

$$q \in X, d_X(p,q) < \phi(p) \text{ implies } d_Y(f(p), f(q)) < \frac{\epsilon}{2}.$$

Let $J(p)$ be the set of all $q \in X$ for which

$$d_X(p,q) < \frac{1}{2}\phi(p).$$

Since $p \in J(p)$, the collection of all sets $J(p)$ is an open cover of X; and since X is compact, there is a finite set of points $p_1, \ldots, p_n$ in X, such that

$$X \subset J(p_1) \bigcap \cdots \bigcap J(p_n).$$

Let $\delta = \frac{1}{2}\min[\phi(p_1), \ldots, \phi(p_n)]$. So $\delta > 0$.
Now let q and p be points of X, such that $d_X(p,q) < \delta$. There is an integer m, $1 \leq m \leq n$ such that $p \in J(p_m)$; hence

$$d_X(p, p_m) < \frac{1}{2}\phi(p_m),$$

and we also have

$$d_x(q, p_m) \leq d_x(p,q) + d_x(p, p_m) < \delta + \frac{1}{2}\phi(p_m) \leq \phi(p_m).$$

Finally, we have

$$d_Y(f(p), f(q)) \leq d_Y(f(p), f(p_m)) + d_Y(f(q), f(p_m)) < \epsilon.$$

Theorem 17 (A Convex Function on an Open Set Is Continuous)**.**

If f is a convex function on an open set $U \in \mathbb{R}$, then f is continuous on U.
Note: A function is convex on $[a,b]$ if for any $x, y \in [a,b]$ and any $t \in [0,1]$, $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$.

Proof:
Let $b \in U$. As U is open there exist $a, c \in U$ such that $a < b < c$.

Counter Example 6 (A Convex Function on a Closed Set Need Not Be Continuous)**.**

If we look at the closed set $[0,1]$ and let $f(x) = 0$ for $x \in [0,1)$ and $f(1) = 1$, then f is convex, but f is not continuous.

Differentiation

Definition 3.

Let f be defined on $[a,b]$. For any $x \in [a,b]$ form the quotient

$$\phi(t) = \frac{f(t) - f(x)}{t - x}$$

and define

$$f'(x) = \lim_{t \to x} \phi(t).$$

Then f' is defined to be the derivative of f at x provided the limit exists.

Theorem 18 (Differentiability Implies Continuity)**.**

Let f be defined on $[a,b]$. If f is differentiable at a point $x \in [a,b]$, then f is continuous at x.

Proof:
As $t \to x$, we have, $f(t) - f(x) = \dfrac{f(t) - f(x)}{t - x} \cdot (t - x) \to f'(x)0 = 0$.

Counter Example 7 (A Continuous Function Need Not Be Differentiable)**.**

Let $f(x) = |x|$ on containing a neighborhood of 0. This function is continuous as $||x| - |y|| < |x - y|$ for all $x, y \in \mathbb{R}$. However, we have

$$\lim_{t \to 0+} \frac{|t|}{t} = \lim_{t \to 0^+} \frac{t}{t} = 1$$

and

$$\lim_{t \to 0-} \frac{|t|}{t} = \lim_{t \to 0^+} \frac{-t}{t} = -1$$

so our two sided limits do not agree, so our function is not differentiable at 0.

Counter Example 8 (The Derivative of a Function Need Not Be Continuous)**.**

Let $f(x) = \begin{cases} x^2 \sin(1/x) & \text{if} \quad x \neq 0; \\ 0 & \text{if} \quad x > 0 \end{cases}$, f is continuous for all positive x and, by the Squeeze Theorem, f is continuous at 0 as well. $f'(x)$ is defined on all nonnegative real numbers as, for $x \neq 0$ $f'(x) = 2x\sin(1/x) - \cos(1/x)$ and for $x = 0$,

$$f'(0) = \lim_{x \to 0^+} \frac{f(x) - 0}{x - 0} = \lim_{x \to 0^+} x \sin(1/x) = 0$$

by the Squeeze Theorem. So $f'(x)$ is differentiable, but is discontinuous at $x = 0$.

Counter Example 9 (A Nowhere Differentiable Continuous Function)**.**

Let $\phi(x) = |x|$ for $-1 < x \leq 1$ and extend this definition to all real numbers by letting $\phi(x + 2) = \phi(x)$. Then, for all s and t,

$$|\phi(s) - \phi(t)| \leq |s - t|$$

since if we let $s = s' + 2k$ where $s' \in (-1, 1]$ and k is an integer. Similarly define $t = t' + 2l$. So $|phi(s) - \phi(t)| = |s' - t'| < 2$, and $|s - t| = |(s' - t') + 2(k - l)|$. So we have in particular that ϕ is continuous on $\mathbb{R}$. Define

$$f(x) \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \phi(4^n x).$$

Since $0 \leq \phi \leq 1$, by the Weierstrass M-test, this series converges uniformly on $\mathbb{R}$. Thus, since partial sum is continuous, f is continuous on $\mathbb{R}$.
Fix a real number x and a positive integer m. Put

$$\delta_m = \pm \frac{1}{2} \cdot 4^{-m}$$

where the sign is so chosen that no integer lies between $4^m x$ and $4^m(x + \delta_m)$. This can be done since $4^m |d_m| = \frac{1}{2}$. Define

$$\gamma_n = \frac{\phi(4^n(x + \delta_m)) - \phi(4^n x)}{\delta_m}.$$

When $n > m$, $4^n \delta_m$ is an even integer so that $\gamma_n = 0$. When $0 \leq n \leq m$, we have that $|\gamma_n| \leq 4^n$, by the fact that for all s, t, $|\phi(s) - \phi(t)| \leq |s - t|$.
Since $|\gamma_m| = 4^m$, we conclude that

$$\left| \frac{f(x + \delta_m) - f(x)}{\delta_m} \right| = \left| \sum_{n=0}^{m} \left(\frac{3}{4}\right)^n \gamma_n \right| \geq 3^m - \sum n = 0^{m-1} 3^n = \frac{1}{2}(3^m + 1).$$

As $m \to \infty$, $\delta_m \to 0$, it follows that f is not differentiable at x.

Theorem 19 (Fermat's Theorem)**.**

Let f be defined on $[a, b]$; if f has a local maximum at a point $x \in (a, b)$ and if $f'(x)$ exists, then $f'(x) = 0$.

Proof:

Choose δ such that for all points $y \in [a,b]$ such that $|x-y| < \delta$ implies $f(y) \le f(x)$. So $a < x-\delta < x < x+\delta < b$. If $x-\delta < t < x$, then

$$\frac{f(t)-f(x)}{t-x} \ge 0.$$

Letting $t \to x$, we see that $f'(x) \ge 0$. Similarly, for $x < t < x+\delta$,

$$\frac{f(t)-f(x)}{t-x} \le 0.$$

Once again letting $t \to x$, we get $f'(x) \le 0$. So $f'(x) = 0$.

Theorem 20 (Generalized Mean Value Theorem)**.**

If f and g are continuous real functions on $[a,b]$ which are differentiable in (a,b), then there is a point $x \in (a,b)$ at which

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

Proof:
Put

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

for $a \le t \le b$. Then h is continuous on $[a,b]$, h is differentiable in (a,b) and $h(a) = f(b)g(a) - f(a)g(b) = h(b)$. It suffices to show that $h'(t) = 0$ somewhere in (a,b). If h is constant, this holds for all $x \in (a,b)$. If $h(t) > h(a)$ for some $t \in (a,b)$, let x be a point on $[a,b]$ at which h attains its maximum, which exists by the Extreme Value Theorem. So $x \in (a,b)$ and Fermat's Theorem shows that $h'(x) = 0$. If $h(t) < h(a)$, we may look at $-h(t)$ and show $-h'(x) = 0$ for some $x \in (a,b)$.

Corollary 4 (Mean Value Theorem)**.**

If f is continuous on $[a,b]$, differentiable on (a,b), then there exists a $c \in (a,b)$ such that

$$f'(c)(b-a) = f(b) - f(a).$$

Proof:
If we set $g(x) = x$ in the above theorem we get the desired result.

Theorem 21 (L'Hôpital's Rule)**.**

Suppose f and g are real and differentiable in (a,b) and $g'(x) \ne 0$ for all $x \in (a,b)$, where $-\infty \le a < b \le \infty$. Suppose

$$\frac{f'(x)}{g'(x)} \to A \text{ as } x \to a.$$

If $f(x) \to 0$ and $g(x) \to 0$ as $x \to a$ or if $g(x) \to \infty$ as $x \to a$, then

$$\frac{f(x)}{g(x)} \to A \text{ as } x \to a.$$

Proof:
We first consider the case in which $-\infty \le A < +\infty$. Choose a real number q such that $A < q$, and then choose r such that $A < r < q$. There is a point $c \in (a,b)$ such that $a < x < c$ implies

$$\frac{f'(x)}{g'(x)} < r.$$

If $a < x < y < c$, then the Mean Value Theorem furnishes a point $r \in (x,y)$ such that

$$\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(t)}{g'(t)} < r.$$

If we assume that $f(x), g(x) \to 0$, then letting $x \to a$ in the above equation yields

$$\frac{f(y)}{g(y)} \leq r < q$$

If we assume $g(x) \to \infty$, keeping y fixed in the above situation, we can choose a pint $c_1 \in (a, y)$ such that $g(x) > g(y)$ and $g(x) > 0$ if $a < x < c_1$. Multiplying the equating by $[g(x) - g(y)]/g(x)$, we obtain

$$\frac{f(x)}{g(x)} < r - r\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)}.$$

If we let $x \to a$, then as $g(x) \to \infty$, there is a point $c_2 \in (a, c_1)$ such that

$$\frac{f(x)}{g(x)} < q.$$

Summing two of the equations above shows, we get

$$\frac{f(y)}{g(y)} + \frac{f(x)}{g(x)} < 2q.$$

So, for any q subject only to $A < q$ there is a point c_2 such that $f(x)/g(x) < q$ if $a < x < c_2$. Similarly, if $-\infty < A \leq \infty$ we can obtain the same result. The result holds also if $g(x) \to -\infty$ by replacing $g(x)$ with $-g(x)$.

Theorem 22 (Taylor's Theorem)**.**

Suppose f is a real function on $[a, b]$, n is a positive integer, $f^{(n-1)}$ is continuous on $[a, b]$, $f^{(n)}(t)$ exists for every $t \in (a, b)$. Let α, β be distinct points of $[a, b]$, and define

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t - \alpha)^k.$$

Then there exists a point x between α and β such that

$$f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta - \alpha)^n.$$

Proof:
Let M be the number defined by

$$f(\beta) = P(\beta) + M(\beta - \alpha)^n$$

and put

$$g(t) = f(t) - P(t) - M(t - \alpha)^n$$

for $a \leq t \leq b$. We have to show that $n!M = f^{(n)}(x)$ for some x between α and β. By the definition of $P(t)$ and $g(t)$, we have

$$g^{(n)}(t) = f^{(n)}(t) - n!M$$

for $a < t < b$. Hence the proof will be complete if we can show that $g^{(n)}(x) = 0$ for some x between α and β. Since $P^{(k)}(\alpha) = f^{(k)}(\alpha)$ for $k = 0, 1, \ldots, n-1$ we have

$$g(\alpha) = g'(\alpha) = \cdots = g^{(n-1)} = 0$$

Our choice of M shows that $g(\beta) = 0$, so that $g'(x_1) = 0$ for some x_1 between α and β, by the mean value theorem. Since $g'(\alpha) = 0$, we conclude similarly that $g''(x_2) = 0$ for some x_2 between α and x_1. After n steps we arrive at the conclusion that $g^{(n)}(x_n) = 0$ for some x_n between α and x_{n-1}, that is, between α and β.

Theorem 23 (Convexity and $f''(x) \geq 0$ Are Equivalent)**.**

A twice-differentiable function f is convex on $[a, b]$ if and only if $f''(x) \geq 0$ for all $x \in (a, b)$.

Proof:
Assume first that $f''(x) \geq 0$.
For any $x_1, x_2 \in (a, b)$, $x_1 < x_2$ and $\lambda \in (0, 1)$.

$f(\lambda x_1+(1-\lambda)x_2)\lambda f(x_1)-(1-\lambda)f(x_2) = \lambda[f(\lambda x_1+(1-\lambda)x_2)-f(x_1)]+(1-\lambda)[f(\lambda x_1+(1-\lambda)x_2)-f(x_2)]$.

By the Mean Value Theorem,

$$= \lambda f'(\xi_1)((\lambda - 1)x_1 + (1 - \lambda)x_2) + (1 - \lambda)f'(\xi_2)(\lambda x_1 - \lambda x_2)$$

where

$$\xi_1 \in (x_1, \lambda x_1 + (1 - \lambda)x_2),$$
$$\xi_2 \in (\lambda x_1 + (1 - \lambda)x_2, x_2).$$

Rearranging the above equation, we get

$$= \lambda(1 - \lambda)f'(\xi_1)(x_2 - x_1) + \lambda(1 - \lambda)f'(\xi_2)(x_2 - x_1) = \lambda(1 - \lambda)(x_2 - x_1)[f'(\xi_1) - f'(\xi_2)].$$

Utilizing the Mean Value Theorem again, we get that this is equal to

$$\lambda(1 - \lambda)(x_2 - x_1)f''(\xi)(\xi_1 - \xi_2) \leq 0$$

for $\xi \in (\xi_1, \xi_2)$ since $\xi_1 < \xi_2$ and we have assumed that $f''(\xi) \geq 0$. So f is convex.
Conversely, assume f is convex. For any $x \in (a, b)$, there exists an $h \neq 0 \in \mathbb{R}$ small enough so that $x \pm h \in (a, b)$.
Since f is convex,

$$f(x) = f(\frac{1}{2}(x + h) + \frac{1}{2}(x - h)) \leq \frac{1}{2}\left(f(x + h) + f(x - h)\right).$$

So

$$\frac{f(x + h) + f(x - h) - 2f(x)}{h^2} \geq 0$$

This result is independent of our choice of h, so we may make h as small as we like. Thus we can take the limit as $h \to 0$ and the result will still hold

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

is an indeterminate form as f is continuous, so L'Hôpital's Rule applies. So

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

Adding and subtracting $\frac{f'(x)}{h}$, we get

$$0 \leq \lim_{h\to 0} \frac{1}{2}\left(\frac{f'(x + h) - f'(x)}{h} + \frac{f'(x) - f'(x - h)}{h}\right)$$

$$= \frac{1}{2}\left(f''(x) + f''(x)\right) = f''(x)$$

So $f''(x) \geq 0$ for all $x \in (a, b)$.

Integration

Definition 4.

Let $[a,b]$ be a given interval. Suppose f is a bounded real function defined on $[a,b]$. Corresponding to each partition P of $[a,b]$, we put
$M_i = \sup f(x), (x_{i-1} \leq x \leq x_i)$,
$m_i = \inf f(x), (x_{i-1} \leq x \leq x_i)$,
$U(P,f) = \sum_{i=1}^{n} M_i \Delta x_i$,
$L(P,f) = \sum_{i=1}^{n} m_i \Delta x_i$, and finally

$$\overline{\int_a^b} f dx = \inf U(P,f),$$

$$\underline{\int_a^b} f dx = \sup U(P,f),$$

where the inf and sup are taken over all partitions P of $[a,b]$. If the upper and lower integrals are equal, we say that f is Riemann-integrable on [a,b] and denote the common value by

$$\int_a^b f(x)dx.$$

Theorem 24 (Equivalent Definition of Integrability)**.**

f is Riemann-integrable on $[a,b]$ if and only if for every $\epsilon > 0$ there eixsts a partition P such that

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

Lemma 5 (If P^* is a refinement of P, then $L(P,f) \leq L(P^*,f)$, $U(P^*,f) \leq U(P,f)$)**.**

It suffices to prove only the first claim as the proof for the second is identical. Suppose first that P^* contains just one more point than P. Let this extra point be x^* and suppose $x_{i-1} < x^* < x_i$. Put
$w_1 = \inf_{x \in [x_{i-1}, x^*]} f(x)$,
$w_2 = \inf_{x \in [x^*, x_i]} f(x)$.
It is clear that $w_1 \geq m_i$ and $w_2 \geq m_i$ where m_i is defined as above. Hence, $L(P^*,f) - L(P,f) =$

$$w_1[x^* - x_{i-1}] + w_2[x_i - x^*] - m_i[x_i - x_{i-1}]$$

$$= (w_1 - m_i)(x^* - x_{i-1}) + (w_2 - m_i)(x_i - x^*) \geq 0.$$

We may repeat this reasoning for all points in P^* that are not in P.

Lemma 6 ($L(P,f) \leq \underline{\int} f dx \leq \overline{\int} f dx \leq U(P,f)$)**.**

It suffices to show that $\underline{\int} f dx \leq \overline{\int} f dx$ as the outer inequalities hold by definition of sup and inf. Let P^* be the common refinement of two partitions P_1 and P_2. By Lemma 5,

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

Hence $L(P_1,f) \leq U(P_2,f)$. If P_2 is fixed and the sup is taken over all P_1, we obtain

$$\underline{\int} f dx \leq U(P_2,f).$$

If we now take the inf over all P_2 we get the desired result.

Proof:
For every P, by Lemma 6 we have

$$L(P,f) \leq \underline{\int} f dx \leq \overline{\int} f dx \leq U(P,f).$$

Thus, if we assume, for any $\epsilon > 0$, there is a partition P such that

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

then we have

$$\overline{\int} f dx - \underline{\int} f dx < \epsilon.$$

As this holds for any ϵ, we have

$$\overline{\int} f dx = \underline{\int} f dx.$$

Conversely, suppose f is Riemann-integrable and let $\epsilon > 0$ be given. Then there exist partitions P_1 and P_2 such that

$$U(P_2, f) - \int f dx < \frac{\epsilon}{2},$$

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

Choosing P to be the common refinement of P_1 and P_2, we get, by Lemma 5,

$$U(P, f) \leq U(P_2, f) < \int f dx + \frac{\epsilon}{2} < L(P_1, f) + \epsilon \leq L(P, f) + \epsilon.$$

After rearrangement, we get

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

Theorem 25 (Conditions For Integrability)**.**

(a) If f is continuous on $[a, b]$ then f is Riemann-integrable on $[a, b]$.

(b) If f is monotonic on $[a, b]$, then f is Riemann-integrable.

(c) Suppose f is Riemann-integrable on $[a, b]$, $m \leq f(x) \leq M$, ϕ is continuous on $[m, M]$, and $h(x) = \phi(f(x))$ on $[a, b]$. Then h is Riemann-integrable on $[a, b]$.

(z) If f has countably many discontinuities, then f is Riemann-integrable.

Proof:
(a) Let $\epsilon > 0$ be given. Choose $\eta > 0$ so that $\eta < \frac{\epsilon}{b - a}$. Since f is uniformly continuous on $[a, b]$, there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \eta$$

for $x, y \in [a, b]$, $|x - y| < \delta$.
If P is any partition of $[a, b]$ such that $\Delta x_i < \delta$ for all i, then the above implies that $M_i - m_i \leq \eta$ and therefore

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

$$\leq \eta \sum_{i=1}^{n} \Delta x_i = \eta(b - a) < \epsilon.$$

So, by the Theorem above, f is Riemann-integrable.
(b) Let $\epsilon > 0$ be given. Choose an integer n such that $n > \frac{(b - a)(f(b) - f(a))}{\epsilon}$. Choose a partition such that $\Delta x_i = \frac{b - a}{n}$.
We suppose that f is monotonically increasing (the proof is analogous in the other case). Then $M_i = f(x_i)$, $m_i = f(x_{i-1})$ for $i = 1, \ldots, n$, so that

$$U(P, f) - L(P, f) = \frac{b - a}{n} \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})]$$

$$= \frac{b-a}{n}(f(b) - f(a)) < \epsilon$$

by our choice of n. So f is Riemann-integrable.
(c) Fix $\epsilon > 0$. Since ϕ is uniformly continuous on $[m, M]$, there exists $\delta > 0$ such that $\delta < \epsilon$ and $|\phi(s) - \phi(t)| < \epsilon$ if $|s - t| \leq \delta$ and $s, t \in [m, M]$.
Since f is Riemann-integrable, there is a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ such that

$$U(P, f) - L(P, f) < \delta^2.$$

Let M_i, m_i have the same meaning as in the definition of integral and let M_i^*, m_i^* be the analogous numbers for h. Divide the numbers $1, \ldots, n$ into two classes: $i \in A$ if $M_i - m_i < \delta$, $i \in B$ if $M_i - m_i \leq \delta$.
For $i \in A$, our choice of δ show that $M_i^* m_i^* \leq \epsilon$.
For $i \in B$, $M_i^* - m_i^* \leq 2K$ where $K = \sup |\phi(t)|$, $m \leq t \leq M$. By our choice of partition, we have

$$\delta \sum_{i \in B} \Delta x_i \leq \sum_{i \in B} (M_i - m_i) \Delta x_i < \delta^2$$

so that $\sum_{i \in B} \Delta x_i < \delta$. It follows that

$$U(P, h) - L(P, h) = \sum_{i \in A} (M_i^* - m_i^*) \Delta x_i + \sum_{i \in B} (M_i^* - m_i^*) \Delta x_i$$

$$\leq \epsilon[b - a] + 2K\delta < \epsilon[b - a + 2K].$$

Since ϵ was arbitrary, we have that h is Riemann-integrable.

Theorem 26 (Fundamental Theorem of Calculus, Part I)**.**

Let f be Riemann-integrable on $[a, b]$. For $a \leq x \leq b$, put

$$F(x) = \int_a^x f(t) dt.$$

Then F is continuous on $[a, b]$; furthermore, if f is continuous at a point x_0 of $[a, b]$, then F is differentiable at x_0, and

$$F'(x_0) = f(x_0).$$

Proof:
Since f is Riemann-integrable, f is bounded. Suppose $|f(t)| \leq M$ for $a \leq t \leq b$. If $a \leq x < y \leq b$, then

$$|F(y) - F(x)| = |\int_x^y f(t) dt| \leq M(y - x).$$

Given $\epsilon > 0$, we see that $|F(y) - F(x)| < \epsilon$, when $y - x < \frac{\epsilon}{M}$.
Now suppose f is continuous at x_0. Given $\epsilon > 0$ choose $\delta > 0$ such that $|f(t) - f(x_0)| < \epsilon$ if $|t - x_0| < \delta$, and $a \leq t \leq b$. Hence, if

$$x_0 - \delta < s \leq x_0 \leq t < x_0 + \delta$$

and

$$a \leq s < t \leq b,$$

we have that

$$\left| \frac{F(t) - F(s)}{t - s} - f(x_0) \right| = \left| \frac{1}{t - s} \int_s^t f(u) - f(x_0) du \right| < \epsilon.$$

It follows that $F'(x_0) = f(x_0)$.

Theorem 27 (Fundamental Theorem of Calculus, Part II)**.**

If f is Riemann-integrable on $[a,b]$ and if there is a differentiable function F on $[a,b]$ such that $F' = f$, then

$$\int_a^b f(x)dx = F(b) - F(a).$$

Proof:
Let $\epsilon > 0$ be given. Choose a partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$ so that $U(P,f) - L(P,f) < \epsilon$. The Mean Value Theorem furnishes points $t_i \in [x_{i-1}, x_i]$ such that

$$F(x_i) - F(x_{i-1}) = f(t_i)\Delta x_i$$

for $i = 1, \ldots, n$. Thus

$$\sum_{i=1}^{n} f(t_i)\Delta x_i = F(b) - F(a).$$

It now follows that

$$\left| F(b) - F(a) - \int_a^b f(x)dx \right| < \epsilon.$$

Since this holds for every $\epsilon > 0$, the proof is complete.

Corollary 5 (Integration by Parts)**.**

Suppose F and G are differentiable functions on $[a,b]$, $F' = f$¡ $G' = g$, both of which are Riemann-integrable, 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:
Put $H(x) = F(x)G(x)$ and apply the Fundamental Theorem of Calculus Part II to H and its derivative.

Theorem 28 (Arc Length)**.**

If γ' is continuous on $[a,b]$, then the arc length of γ is given by

$$\Lambda(\gamma) = \int_a^b |\gamma'(t)|dt.$$

Theorem 29 (Hölder's Inequality)**.**

Let $f(x)$ and $g(x)$ be two integrable functions on $[a,b]$, both bounded. Then

$$\left| \int_a^b f(x)g(x)dx \right| \leq \left(\int_a^b |f(x)|^p dx \right)^{1/p} \left(\int_a^b |f(x)|^q dx \right)^{1/q}$$

where $1 \leq p, q < \infty$ and $1/p + 1/q = 1$.

Lemma 7 (Young's Inequality)**.**

If $u, v \geq 0$, p, q as above, then $uv \leq \dfrac{u^p}{p} + \dfrac{v^q}{q}$.
If u or v are 0, then the inequality hold trivially. If $u, v > 0$, there exist x, y such that $u = e^{x/p}$ and $v = e^{y/q}$. As the second derivative of e^t is e^t and $e^t > 0$, for all t, e^t is convex. So if we let $t = 1/p$ and $1 - t = 1/q$, then we get $uv = e^{x/p}e^{y/q} = e^{tx}e^{(1-t)y} = e^{tx+(1-t)y} \leq te^x + (1-t)e^y = \dfrac{u^p}{p} + \dfrac{v^q}{q}$.

Lemma 8.

If f and g are integrable p and q as above, both are nonnegative $\int_a^b f^p dx = 1 = \int_a^b g^q dx$, then

$$\int_a^b f(x)g(x)dx \leq 1.$$

f^p and g^q are integrable by a composition of a continuous and integrable functions and fg is integrable. Using the inequality found in the lemma above, we get that $fg \leq f^p/p + g^q/q$. Integrating both sides (which is allowed as the inequality holds for all values), we get

$$\int_a^b f(x)g(x)dx \leq 1/p \int_a^b f(x)^p dx + 1/q \int_a^b g(x)^q dx$$

$$= 1/p + 1/q = 1.$$

Proof:

Let f and g be two integrable functions on $[a, b]$. By the Cauchy-Schwarz Inequality,

$$\left| \int_a^b f(x)g(x)dx \right|$$

$$\leq \int_a^b |f(x)||g(x)|dx.$$

Let $F = \int_a^b |f|^p dx$, $G = \int_a^b |g|^q dx$. We can use the lemma above on the integrable functions $|f|/(F^{1/p})$ and $|g|/(G^{1/q})$. Thus we get $\left| \int_a^b f(x)g(x)dx \right| \leq F^{1/p}G^{1/q} = \left(\left| \int_a^b |f(x)|^p dx \right| \right)^{1/p} \left(\left| \int_a^b |g(x)|^q dx \right| \right)^{1/q}$. Note if F or G are equal to zero then the result holds trivially.

Multivariable Differentiation

Definition 5.

Suppose E is an open set in $\mathbb{R}^n$, f maps E into $\mathbb{R}^m$, and $x \in E$. If there exists a linear transformation A of $\mathbb{R}^n$ into $\mathbb{R}^m$ such that

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

then we say that f is differentiable at x and we write

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

Theorem 30 (Multivariable Derivative Is Unique)**.**

Suppose E and f are as in the definition above and

$$\lim_{|h| \to 0} \frac{|f(x+h) - f(x) - Ah|}{|h|} = 0$$

holds for $A = A_1$ and $A = A_2$. Then $A_1 = A_2$.

Proof:
If $B = A_1 - A_2$, then the inequality

$$|Bh| \leq |f(x+h) - f(x) - A_1 h| + |f(x+h) - f(x) - A_2 h|$$

shows that $\frac{|Bh|}{|h|} \to 0$ as $|h| \to 0$. For a fixed $h \neq 0$, it follows that $\frac{|B(th)|}{|th|} \to 0$ as $t \to 0$. The linearity of B shows that this is independent of t, so $Bh = 0$ for all h. So $B = 0$.

Theorem 31 (f Is Differentiable Implies Partial Derivatives Exist)**.**

Suppose f maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, and f is differentiable at a point $x \in E$. Then the partial derivatives $(D_j f_i)(x)$ exist, and

$$f'(x)e_j = \sum_{i=1}^{m} (D_j f_i)(x) u_i$$

for $1 \leq j \leq m$.

Proof:
Fix j. Since f is differentiable at x,

$$f(x + te_j) - f(x) = f'(x)(te_j) + r(te_j)$$

where $|r(te_j)|/t \to 0$ as $t \to 0$. The linearity of $f'(x)$ shows therefore that

$$\lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t} = f'(x)e_j.$$

If we now represent f in terms of components, then the above becomes

$$\lim_{t \to 0} \sum_{i=1}^{m} \frac{f_i(x + te_j) - f_i(x)}{t} u_i = f'(x)e_j.$$

It follows that each quotient in this sum has a limit, as $t \to 0$, so each $(D_j f_i)(x)$ exists.

Counter Example 10 (Existence of Partial Derivatives Do Not Imply Differentiability)**.**

Let

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

$f_x(x, y) = \dfrac{y(x^2 - y^2)}{(x^2 + y^2)^2}$ for $(x, y) \neq (0, 0)$ and if we fix $y = 0$ we get $f(x, 0) = 0$, so $f_x(x, y)$ is defined for all values. We get similar values for f_y. However, f is not differentiable since...

Theorem 32 (Mean Value Estimate)**.**

Suppose f maps a convex open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, f is differentiable in E and there is a real number M such that

$$||f'(x)|| \leq M$$

for every $x \in E$. Then

$$|f(b) - f(a)| \leq M|b - a|$$

for all $a, b \in E$.

Lemma 9 (Mean Value Theorem for Vector Valued Functions)**.**

Suppose f is a continuous mapping of $[a, b]$ into $\mathbb{R}^k$ and f is differentiable in (a, b). Then there exists $x \in (a, b)$ such that

$$|f(b) - f(a)| \leq (b - a)|f'(x)|.$$

Put $z = f(b) - f(a)$ and define

$$\phi(t) = z \cdot f(t).$$

Then ϕ is a real-valued continuous function on $[a, b]$ which is differentiable in (a, b). So by the Mean Value Theorem,

$$\phi(b) - \phi(a) = (b - a)\phi'(x) = (b - a)z \cdot f'(x)$$

for some $x \in (a, b)$. On the other hand,

$$\phi(b) - \phi(a) = z \cdot f(b) - z \cdot f(a) = z \cdot z = |z|^2.$$

The Cauchy-Schwarz Inequality yields

$$|z|^2 = (b-a)|z \cdot f'(x)| \leq (b-a)|z||f'(x)|.$$

Hence $|z| \leq (b-a)|f'(x)|$.

Proof:
Fix $a, b \in E$. Define

$$\gamma(t) = (1-t)a + tb$$

for all $t \in \mathbb{R}$ such that $\gamma(t) \in E$. Since E is convex $\gamma(t) \in E$ if $0 \leq t \leq 1$. Put

$$g(t) = f(\gamma(t)).$$

Then, by the chain rule, we have

$$g'(t) = f'(\gamma(t))\gamma'(t) = f'(\gamma(t))(b-a),$$

so that

$$|g'(t)| \leq ||f'(\gamma(t))|| \cdot |b-a| \leq M|b-a|$$

for all $t \in [0,1]$. Thus, by the above Lemma,

$$|g(1) - g(0)| \leq M|b-a|.$$

But $g(0) = f(a)$ and $g(1) = f(b)$.

Theorem 33 (Equivalence of Continuous Differentiability and Continuity of Partial Derivatives)**.**

Suppose f maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$. Then f is continuously differentiable on E if and only if the partial derivatives $D_j f_i$ exist and are continuous on E for $1 \leq i \leq m$, $1 \leq j \leq n$.

Proof:
Assume first that f is continuously differentiable on E. By above, we get

$$(D_j f_i)(x) = (f'(x)e_j) \cdot u_i$$

for all i, j and all $x \in E$. Hence

$$(D_j f_i)(y) - (D_j f_i)(x) = \{[f'(y) - f'(x)]e_j\} \cdot u_i$$

and since $|u_i| = |e_j| = 1$, it follows that

$$|(D_j f_i)(y) - (D_j f_i)(x)| \leq |[f'(y) - f'(x)]e_j| \leq ||f'(y) - f'(x)||.$$

Hence $D_j f_i$ is continuous as f' is continuous.
For the converse, it suffices to consider the case where $m = 1$. Fix $x \in E$ and $\epsilon > 0$. Since E is open, there is an open ball $S \subset E$ with center at x and radius r, and the continuity of the functions $D_j f$ shows that r can be chosen so that

$$|(D_j f)(y) - (D_j f)(x)| < \frac{\epsilon}{n}$$

for $y \in S$ and $1 \leq j \leq n$.
Suppose $h = \sum_{j=1}^{n} h_j e_j$, $|h| < r$, put $v_0 = 0$ and $v_k = h_1 e_1 + \cdots + h_k e_k$ for $1 \leq n$. Then

$$f(x+h) - f(x) = \sum_{j=1}^{n} [f(x+v_j) - f(x+v_{j-1})].$$

Since $|v_k| < r$ for $1 \leq k \leq n$ and since S is convex, the segments with end points $x + v_{j-1}$ and $x + v_j$ lie in S. Since $v_j = v_{j-1} + h_j e_j$, the Mean Value Theorem shows that the jth summand in the above equality is equal to

$$h_j (D_j f)(x + v_{j-1} + \theta_j h_j e_j)$$

for some $\theta_j \in (0,1)$, and this differs from $h_j(D_jf)(x)$ by less than $|h_j|\epsilon/n$ by our choice of r. By the above equality, it follows that

$$\left| f(x+h) - f(x) - \sum_{j=1}^{n} h_j(D_jf)(x) \right| \leq \frac{1}{n}\sum_{j=1}^{n} |h_j|\epsilon \leq |h|\epsilon$$

for all h such that $|h| < r$. Dividing by $|h|$ gives the desired result.

Theorem 34 (f Is Convex If and Only If Every Critical Point Is A Global Minimizer)**.**

If f is twice differentiable and convex on a convex domain D, then every critical point of f in D must be a global minimizer.

Lemma 10 (f Is Convex If and Only If $H(f)$ Is Positive Semidefinite)**.**

f is convex in D (D as above) if and only if $H(f)$ is positive semidefinite, where $H(f)$ is the Hessian of f and a matrix is positive semidefinite if all eigenvalues are greater than or equal to 0.
For any fixed $x, y \in D$, let $g_{xy}(t) = f(x + t(y-x))$, where t is defined such that $x + t(x-y) \in D$. So $g_{xy}(t)$ is convex since f is. So, by an above theorem, $g''_{xy}(t) \geq 0$. So

$$0 \leq (g'_{xy}(t)) = \left(\sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x + t(y-x))(y_k - x_k) \right)'$$

$$= \sum_{k=1}^{n}\sum_{l=1}^{n} \frac{\partial^2 f}{\partial x_l \partial x_k}(x + t(y-x))(y_k - x_k)(y_l - x_l).$$

So f is convex if and only if $\displaystyle\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) w_i w_j \geq 0$, for $w = (w_1, \ldots, w_n) \in \mathbb{R}^n$, $x \in D$.

Proof:
Let x_0 be a critical point of f. So $\dfrac{\partial f}{\partial x_j}(x_0) = 0$ for $j = 1, 2, \ldots, n$. For any $x \in D$, by Taylor's Theorem,

$$f(x) = f(x_0) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0)(x_j - x_{0j}) + \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\zeta)(x_i - x_{0i})(x_j - x_{0j})$$

for ζ picked appropriately. However, since x_0 is a critical point and by the lemma above, we have

$$f(x) \geq f(x_0) + 0 + 0 = f(x_)).$$

So x_0 is a global minimizer.

Definition 6 (Contraction)**.**

Let X be a metric space, with metric d. If ϕ maps X into X and if there is a number $c < 1$ such that

$$d(\phi(x), \phi(y)) \leq cd(x,y)$$

for all $x, y \in X$, then ϕ is said to be a contraction of X into X.

Theorem 35 (The Contraction Principle)**.**

If X is a complete metric space, and if ϕ is a contraction of X into X, then there exists one and only one $x \in X$ such $\phi(x) = x$.

Proof:
Pick $x_0 \in X$ arbitrarily, and define $\{x_n\}$ recursively, by setting

$$x_{n+1} = \phi(x_n).$$

As ϕ is a contraction, there is a $c < 1$ satisfying the definition. For $n \geq 1$ we then have

$$d(x_{n+1}, x_n) = d(\phi(x_n), \phi(x_{n-1})) \leq cd(x_n, x_{n-1}).$$

By induction, we get

$$d(x_{n+1}, x_n) \leq c^n d(x_1, x_0).$$

If $n < m$, it follows that

$$d(x_n, x_m) \leq \sum_{i=n+1}^{m} d(x_i, x_{i-1}) \leq (c^n + \cdots + c^{m-1}) d(x_1, x_0)$$

$$\leq [(1-c)^{-1} d(x_1, x_0)] c^n$$

as this is a geometric series. Since $c < 1$, as $n \to \infty$, this goes to 0, so $\{x_n\}$ is a Cauchy Sequence. Since X is complete, $\lim_{n\to\infty} x_n = x$ for some $x \in X$.
Since ϕ is a contraction, ϕ is (uniformly) continuous on X. Hence

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

Counter Example 11 (If a contraction has $c = 1$, the above conclusion need not hold)**.**

Let $f(x) = x + \dfrac{1}{1+e^x}$. So $|f(x) - f(y)| = \left| x - y + \dfrac{e^y - e^x}{(1+e^x)(1+e^y)} \right| < |x - y|$. However, $f(x) \neq x$ as $e^x > 0$ for all $x \in \mathbb{R}$.

Theorem 36 (Inverse Function Theorem)**.**

Suppose f is a continuously differentiable mapping on E of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$, $f'(a)$ is invertible for some $a \in E$, and $b = f(a)$. Then

(a) there exist open sets U and V in $\mathbb{R}^n$ such that $a \in U$, $b \in V$, f is one-to-to on U, and $f(U) = V$;

(b) if g is the inverse of f, which exists by **(a)**, defined in V by

$$g(f(x)) = x, (x \in U),$$

then g is continuously differentiable on V and $g'(y) = \{f'(g(y))\}^{-1}$ for $y \in V$.

Proof:
(a) Put $f'(a) = A$, and choose λ so that

$$2\lambda ||A^{-1}|| = 1.$$

Since f' is continuous at a, there is an open ball $U \subset E$, with center at a, such that

$$||f'(x) - A|| < \lambda, (x \in U).$$

We associate to each $y \in \mathbb{R}^n$ a function ϕ, defined by

$$\phi(x) = x + A^{-1}(y - f(x)), (x \in E).$$

(Note that $f(x) = y$ if and only if x is a fixed point of ϕ).
Since $\phi'(x) = I - A^{-1}f'(x) = A^{-1}(A - f'(x))$, by the above, we have

$$||\phi'(x)|| < \frac{1}{2}, (x \in U).$$

Hence

$$|\phi(x_1) - \phi(x_2)| \leq \frac{1}{2}|x_1 - x_2|, (x_1, x_2 \in U),$$

by the Mean Value Estimate. It follows that ϕ has at most one fixed point in U. So f is one-to-one in U.
Next, put $V = f(U)$ and pick $y_0 \in V$. Then $y_0 = f(x_0)$ for some $x_0 \in U$. Let B be an open ball with center

x_0 and radius $r > 0$, so small that its closure $\overline{B}$ lies in U.
Fix y, $|y - y_0| < \lambda r$. With ϕ as above,

$$|\phi(x_0) - x_0| = |A^{-1}(y - y_0)| < ||A^{-1}||\lambda r = \frac{r}{2}.$$

If $x \in \overline{B}$, it follows by the contraction that

$$|\phi(x) - x_0| \leq |\phi(x) - \phi(x_0)| + |\phi(x_0) - x_0| < \frac{1}{2}|x - x_0| + \frac{r}{2} \leq r;$$

hence $\phi(x) \in B$.
Thus ϕ is a contraction of $\overline{B}$ into $\overline{B}$. Being a closed, bounded subset of $\mathbb{R}^n$, $\overline{B}$ is complete. The Contraction Principle implies that ϕ has a fixed point $x \in \overline{B}$. For this x, $f(x) = y$. Thus $y \in f(\overline{B}) \subset f(U) = V$. So V is open in $\mathbb{R}^n$.

Corollary 6 (Open Mapping Theorem)**.**

If f is a continuously differentiable mapping on an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$ and if $f'(x)$ is invertible for every $x \in E$, then $f(W)$ is an open subset of $\mathbb{R}^n$ for every open set $W \subset E$.

Proof:
Let W be an open set in $\mathbb{R}^n$. For every point $a \in W$, since f' is invertible at a, there is an open ball $U(a) \subset W$ such that, for $x, y \in U(a)$, $||f'(y) - f'(x)|| < \dfrac{1}{2||(f'(a))^{-1}||}$. Note that $W = \bigcup_{x \in W} U(x)$. Thus, by the Inverse Function Theorem, $V(x) = f(U(x))$ is open. As any union of open sets is open, we have that $f(W) = f(\bigcup U(x)) = \bigcup V(x)$ is open in $\mathbb{R}^n$.

Theorem 37 (Implicit Function Theorem)**.**

Let f be a continuously differentiable mapping of an open set $E \subset \mathbb{R}^{n+m}$ into $\mathbb{R}^n$, such that $f(a, b) = 0$ for some point $(a, b) \in E$.
Put $A = f'(a, b)$ and assume that A_x is invertible.
Then there exist open sets $U \subset \mathbb{R}^{n+m}$ and $W \subset \mathbb{R}^m$, with $(a, b) \in U$ and $b \in W$, having the following property:
To every $y \in W$ corresponds a unique x such that

$$(x, y) \in U \text{ and } f(x, y) = 0.$$

If this x is defined to be $g(y)$, then g is a continuously differentiable mapping of W into $\mathbb{R}^n$, $g(b) = a$,

$$f(g(y), y) = 0, y \in W,$$

and

$$g'(b) = -(A_x)^{-1}A_y.$$

Theorem 38 (Condition For Changing Order of Derivatives)**.**

Suppose f is defined in an open set $E \subset \mathbb{R}^2$, suppose that $D_1 f$, $D_{21} f$ and $D_2 f$ exist at every point of E, and $D_{21} f$ is continuous at some point $(a, b) \in E$.
Then $D_{12} f$ exists at (a, b) and

$$(D_{12} f)(a, b) = (D_{21} f)(a, b).$$

Counter Example 12 (A Function Whose Derivatives Are Not Interchangable)**.**

Let

$$f(x, y) = \begin{cases} \frac{xy(y^2 - x^2)}{x^2 + y^2} & \text{if}(x, y) \neq (0, 0) \\ 0 & \text{if}(x, y) = (0, 0) \end{cases}$$

So

$$f_x(x, y) = \begin{cases} \frac{(3x^2 - y^3)(x^2 + y^2) - 2x(x^3 y - xy^3)}{(x^2 + y^2)^2} & \text{if}(x, y) \neq (0, 0) \\ 0 & \text{if}(x, y) = (0, 0) \end{cases}$$

$$f_y(x,y) = \begin{cases} \frac{(-3xy^2+x^3)(x^2+y^2)-2y(x^3y-xy^3)}{(x^2+y^2)^2} & \text{if}(x,y) \neq (0,0) \\ 0 & \text{if}(x,y) = (0,0) \end{cases}$$

$$f_{xy}(0,0) = \lim_{h\to 0} = \frac{f_x(0,h) - f_x(0,0)}{h} = \lim_{h\to 0} \frac{-h-0}{h} = -1,$$

$$f_{yx}(0,0) = \lim_{h\to 0} = \frac{f_y(h,0) - f_y(0,0)}{h} = \lim_{h\to 0} \frac{h-0}{h} = 1.$$

So $f_{xy}(0,0) \neq f_{yx}(0,0)$.

Integration of Differential Forms

Theorem 39 (Stokes' Theorem)**.**

Let D be a bounded domain in $\mathbb{R}^n$ with piecewise-$\mathcal{C}^1$ boundary. Let f be a differentiable $(n-1)$-form on $\overline{D}$. Then

$$\int_{\partial D} f = \int_D df,$$

$f = \sum_{j=1}^{n} f_i dx_{(j)}$, $dx_{(j)} = dx_1 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_n$.

Theorem 40 (Green's Theorem)**.**

Let E be an open subset of $\mathbb{R}^2$, $\alpha, \beta \in \mathcal{C}^1(E)$ and D be a closed subset of E with positively oriented boundary ∂D. Then

$$\int_{\partial D} \alpha dx + \beta dy = \int_D \frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} dA.$$

Proof:
Let $\lambda = \alpha dx + \beta dy$. Then

$$d\lambda = d(\alpha dx + \beta dy) = \frac{\partial \alpha}{\partial y} dy \wedge dx + \frac{\partial \beta}{\partial x} dx \wedge dy = \left(\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} \right) dA.$$

So, by Stokes' Theorem

$$\int_{\partial D} \alpha dx + \beta dy = \int_{\partial D} \lambda = \int_D d\lambda = \int_D \frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} dA.$$

Theorem 41 (Divergence Theorem)**.**

Let D be a bounded domain in $\mathbb{R}^n$ with piecewise $\mathcal{C}^1$ boundary. If $f = (f_1, \ldots, f_n) : \overline{D} \to \mathbb{R}^n$ wiht $f_j \in \mathcal{C}^1(D) \cap \mathcal{C}(\overline{D})$ for $1 \leq j \leq n$. Then

$$\int_D \nabla \cdot f dx = \int_{\partial D} f \cdot \nu d\sigma(x),$$

where ν is the unit outer normal vector of ∂D at x, and ν is given by $\dfrac{\nabla f}{\sqrt{\nabla f \cdot \nabla f}}$.

Sequences and Series of Functions

Definition 7.

A sequence $\{f_n\}_{n=1}^{\infty}$ of functions on E is <u>pointwise convergent</u> if, for any fixed $x \in E$, $f_n(x) \to f(x)$ as $n \to \infty$.
A sequence $\{f_n\}$ of functions on E is <u>uniformly convergent</u> to a function $f(x)$ if for every $\epsilon > 0$ there is an integer N such that $n \geq N$ implies

$$|f_n(x) - f(x)| < \epsilon$$

for all $x \in E$.

Counter Example 13 (Pointwise Convergence Does Not Imply Uniform Convergent)**.**

Let $\epsilon > 0$ be given and let $f_n(x) = \dfrac{x}{n}$. For any fixed $x \in \mathbb{R}$, if we choose an integer n such that $|x|/\epsilon < n$, then $|f_n(x) - 0| < \epsilon$, so f_n converges pointwise to $f(x) = 0$. Assumet toward contradiction that $f_n \to f$ uniformly and let $\epsilon < 1$. So there exists an integer N such that for all $n \geq N$, $f_n(x) - f(x)| < \epsilon$ for all $x \in \mathbb{R}$. Fix $n = N$ and choose $x = N$. We have

$$|f_N(N) - 0| = \left|\frac{N}{N}\right| = 1 \not< \epsilon.$$

So f_n does not converge to f uniformly.
Note: Uniform Convergence does imply pointwise convergence.

Theorem 42 (Cauchy Criterion)**.**

The sequence of functions $\{f_n\}$ on E converges uniformly if and only if for every $\epsilon > 0$ there exists an integer N such that for $m, n \geq N$, $x \in E$ imply

$$|f_n(x) - f_m(x)| < \epsilon.$$

Proof:
Suppose $\{f_n\}$ converges uniformly on E and let f be the limit function. Then there is an integer N such that $n \geq N$, $x \in E$ implies

$$|f_n(x) - f(x)| < \frac{\epsilon}{2},$$

so that

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \epsilon$$

if $n, m \geq N$, $x \in E$.
Conversely, suppose the Cauchy criterion holds. By the Cauchy Sequence Condition, the sequence $\{f_n(x)\}$ converges, for every x, to a limit which me way call $f(x)$. Thus the sequence $\{f_n(x)\}$ converges on E, to f. Let $\epsilon > 0$ be given, and choose N such that

$$|f_n(x) - f_m(x)| < \epsilon$$

for $n, m \geq N$, $x \in E$. Fix n, and let $m \to \infty$. Since $f_m(x) \to f(x)$ as $m \to \infty$, this gives

$$|f_n(x) - f(x)| < \epsilon$$

for every $n \geq N$ and every $x \in E$.

Theorem 43 (Weierstrass M-Test)**.**

Suppose $\{f_n(x)\}$ is a sequence of functions defined on E and suppose

$$|f_n(x)| \leq M_n, (x \in E, n = 1, 2, 3, \dots).$$

Then $\displaystyle\sum_{n=1}^{\infty} f_n(x)$ converges uniformly on E if $\displaystyle\sum_{n=1}^{\infty} M_n$ converges.

Proof:
If $\displaystyle\sum_{n=1}^{\infty} M_n$ converges, then, for arbitrary $\epsilon > 0$,

$$\left|\sum_{i=n}^{m} f_i(x)\right| \leq \sum_{i=n}^{m} M_i < \epsilon, (x \in E),$$

provided m and n are large enough. Uniform convergence follows by the Cauchy Criterion for sums.

Theorem 44 (Uniform Convergence and Continuity)**.**

Suppose $\{f_n\}$ is a sequence of continuous functions and $f_n \to f$ uniformly. Then f is continuous.

Proof:
Let $\epsilon > 0$ be fixed. Let $x_0 \in E$. If x_0 is an isolated point of E, then f is trivially continuous at x_0. We may suppose then that x_0 is a limit point of E. Since we have uniform convergence, there is an integer N such that for all $n \geq N$,

$$|f_n(x) - f(x)| < \frac{\epsilon}{3}$$

for all $x \in E$.
Since f_N is continuous at x_0, there is a neighborhood $B(x_0)$ such that for all $x \in B(x_0) \bigcap E$,

$$|f_N(x_0) - f_N(x)| < \frac{\epsilon}{3}.$$

So, for any $x \in B(x_0) \bigcap E$,

$$|f(x_0) - f(x)| \leq |f(x_0) - f_N(x_0)| + |f_N(x_0) - f_N(x)| + |f_N(x) - f(x)| < \epsilon.$$

Counter Example 14 (Sequence of Continuous Functions that Converge to a Discontinuous Function)**.**

Let

$$f(x) = \begin{cases} 0 & \text{if} \quad x < -\frac{1}{n} \\ nx+1 & \text{if} \quad -\frac{1}{n} \leq x < 0 \\ 1 & \text{if} \quad x = 0 \\ nx+1 & \text{if} \quad 0 < x \leq \frac{1}{n} \\ 0 & \text{if} \quad \frac{1}{n} < x \end{cases}$$

Each $f_n(x)$ is continuous for all n. For $x \neq 0$, $f_n(x) \to 0$ as there exists some $N > \dfrac{1}{|x|}$. For $x = 0$, $f_n(x) = 1 \to 1$. So $f_n(x) \to f(x)$ where $f(x)$ is not continuous at 0.

Theorem 45 (Partial Converse to Uniform Convergence and Continuity)**.**

Suppose K is compact and

(a) $\{f_n\}$ is a sequence of continuous functions on K;

(b) $\{f_n\}$ converges pointwise to a continuous function on K;

(c) $f_n(x) \geq f_{n+1}(x)$ for all $x \in K$, $n = 1, 2, 3, \ldots$.

Then $f_n \to f$ uniformly on K.

Proof:
Put $g_n = f_n - f$. Then g_n is continuous, $g_n \to 0$ pointwise, and $g_n \geq g_{n+1}$. It suffices to show that $g_n \to 0$ uniformly on K.
Let $\epsilon > 0$ be given. Let K_n be the set of all $x \in K$ with $g_n(x) \geq \epsilon$. Since g_n is continuous, K_n is closed, hence compact by Heine-Borel because K_n is bounded by the fact that $g_n(K)$ is compact. Since $g_n \geq g_{n+1}$, we have $K_n \supset K_{n+1}$. Fix $x \in K$. Since $g_n(x) \to 0$, we see that $x \notin K_n$ if n is sufficiently large. This $x \notin \bigcap K_n$. In other words, $\bigcap K_n$ is empty. Hence K_N is empty for some N. It follows that $0 \leq g_n(x) < \epsilon$ for all $x \in K$ and for all $n \geq N$.

Counter Example 15.

Note that all of the above conditions are necessary for this result to hold. For instance, if we omit compactness and look at $f_n(x) = \dfrac{1}{nx+1}$ on $(0,1)$, then f_n is continuous, converges pointwise to 0 and is monotonic. However, f_n does not converge uniformly since for any $n \in \mathbb{Z}^+$, $1/n \in (0,1)$ and $f_n(1/n) = 1/2 \neq 0$.
If we instead omit the monotonicity of f and look at $f_n(x) = nx(1-x^2)^n$ on $[0,1]$, we have K is compact, f_n is continuous, $f_n \to 0$ pointwise by L'Hôpital's Rule, but f_n does not converge uniformly to 0, since if we pick $\epsilon < e^{-1}$, and let $x = 1/n$, $f_n(x) = (1+\dfrac{1}{n})^n(1-\dfrac{1}{n})^n > e^{-1}$ as $n \to \infty$. So f_n cannot converge uniformly.

Theorem 46 (Integrability and Uniform Convergence).

Suppose f_n is Riemann-integrable on $[a, b]$, for $n = 1, 2, 3, \ldots.$ and suppose $f_n \to f$ uniformly on $[a, b]$, then f is Riemann-integrable and

$$\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx.$$

Proof:
Let $\epsilon_n = \sup |f_n(x) - f(x)|$, the supremum being taken over $a \leq x \leq b$. Then

$$f_n - \epsilon_n \leq f \leq f_n + \epsilon_n,$$

so that the upper and lower integrals of f satisfy

$$\int_a^b f_n - \epsilon_n dx \leq \underline{\int} f dx \leq \overline{\int} f dx \leq \int_a^b f_n + \epsilon_n dx.$$

Hence

$$0 \leq \overline{\int} f dx - \underline{\int} f dx \leq 2\epsilon_n[b - a].$$

Since $\epsilon_n \to 0$ as $n \to \infty$, the upper and lower integrals are equal. Thus f is Riemann-integrable.

Counter Example 16.

If $f_n \to f$ pointwise on $[a, b]$, $\displaystyle\int_a^b f_n(x)dx$ does not necessarily converge to $\int_a^b f(x)dx$.
Let $f_n(x) = nx(1 - x^2)^n$ on $[0, 1]$. As has been shown, $f_n \to 0$. We have

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

So $\displaystyle\int_0^1 nx(1 - x^2)^n dx \to \frac{1}{2} \neq \int_0^1 f(x)dx = 0.$

Counter Example 17.

If $\displaystyle\int_a^b f_n(x)dx \to \int_a^b f(x)dx$, it does not guarantee that $f_n \to f$ at all (let alone uniformly).
Let $f_n(x) = \sin(x)$, $f(x) = 0$.

$$\int_0^{2\pi} \sin(x)dx = 0 \int_0^{2\pi} f(x)dx.$$

However, $f_n(x) = \sin(x) \not\to f(x) = 0$.

Theorem 47 (Differentiability and Uniform Convergence).

Suppose $\{f_n\}$ is a sequence of functions, differentiable on $[a, b]$ and such that $\{f_n(x_0)\}$ converges for some point x_0 on $[a, b]$. If $\{f'_n\}$ converges uniformly on $[a, b]$, then $\{f_n\}$ converges uniformly on $[a, b]$, to a function f, and

$$f'(x) = \lim_{n\to\infty} f'_n(x), (a \leq x \leq b).$$

Proof:
Let $\epsilon > 0$ be given. Choose N such that $n, m \geq N$ implies

$$|f_n(x_0) - f_m(x_0)| < \frac{\epsilon}{2}$$

and

$$|f'_n(t) - f'_m(t)| < \frac{\epsilon}{2(b - a)}, (a \leq t \leq b).$$

If we apply the Mean Value Theorem to the function $f_n - f_m$, the above inequality furnishes

$$|(f_n(x) - f_m(x)) - f_n(t) + f_m(t)| \leq \frac{|x-t|\epsilon}{2(b-a)} < \frac{\epsilon}{2}$$

for any x and t on $[a,b]$, if $n, m \geq N$. The inequality

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)|$$

implies, by above inequalities, that

$$|f_n(x) - f_m(x)| < \epsilon, (a \leq x \leq b, n, m \geq N),$$

so that $\{f_n\}$ converges uniformly on $[a,b]$. Let

$$f(x) = \lim_{n\to\infty} f_n(x), (a \leq x \leq b).$$

Let us now fix a point x on $[a,b]$ and define

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t-x}, \phi(t) = \frac{f(t) - f(x)}{t-x}$$

for $a \leq t \leq b, t \neq x$. Then

$$\lim_{t\to x} \phi_n(t) = f_n'(x).$$

From what we obtained by using the Mean Value Theorem, we get

$$|\phi_n(t) - \phi_m(t)| < \frac{\epsilon}{2(b-a)}, n, m \geq N,$$

so that $\{\phi_n\}$ converges uniformly, for $t \neq x$. Since $\{f_n\}$ converges to f, we conclude by the definition of ϕ_n and ϕ that

$$\lim_{n\to\infty} \phi_n(t) = \phi(t)$$

uniformly for $a \leq t \leq b$, $t \neq x$.
Since $\{\phi_n\}$ is a set of continuous functions that converge uniformly to ϕ, we have $\lim_{t\to x} \phi(t) = \lim_{n\to\infty} f_n'(x)$. So, specifically, $f'(x) = \lim_{n\to\infty} f_n'(x)$ for $a \leq x \leq b$.

Definition 8.

A family $\mathcal{F}$ of functions f defined on E in a metric space (X,d) is said to be equicontinuous on E if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

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

whenever $d(x,y) < \delta$, $x, y, \in E$ and $f \in \mathcal{F}$.

Theorem 48 (Convergent Subsequence on a Countable Set)**.**

If $\{f_n\}$ is a pointwise bounded sequence of complex functions on a countable set E, then $\{f_n\}$ has a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in E$.

Proof:
Let $\{x_i\}$ $i = 1, 2, 3, \ldots$, be the points of E arranged in a sequence. Since $\{f_n(x_1)\}$ is bounded, by Bolzano Weierstrass there exists a subsequence, which we shall denote by $\{f_{1,k}\}$, such that $\{f_{1,k}(x_1)\}$ converges as $k \to \infty$.
Let us now consider sequences $S_1, S_2, \cdots$, which we represent by the array:

$$S_1 : f_{1,1}, f_{1,2}, f_{1,3}, \cdots$$

$$S_2 : f_{2,1}, f_{2,2}, f_{2,3}, \cdots$$

$$S_3 : f_{3,1}, f_{3,2}, f_{3,3}, \cdots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

and which have the following properties:

(a) S_n is a subsequence of S_{n-1} for $n = 2, 3, 4, \ldots$;

(b) $\{f_{n,k}(x_n)\}$ converges as $k \to \infty$ (the boundedness of $\{f_n(x_k)\}$ makes it possible to choose S_n in this way);

(c) The order in which the functions appear is the same in each sequence, i.e., if one function precedes another in S_1, they are in the same relation in every S_n, until one or the other is deleted. Hence, when going from one row in the above array to the next below, functions may move to the left but never to the right.
We now consider the sequence $S : f_{1,1}, f_{2,2}, f_{3,3}, \ldots$.
By (c), the sequence S (except possibly the first $n-1$ terms) is a subsequence of S_n for $n = 1, 2, 3, \ldots$.
Hence (b) implies that $\{f_{n,n}(x_i)\}$ converges as $n \to \infty$ for every $x_i \in E$.

Theorem 49 (Arzelà-Ascoli)**.**

Let K be a compact set. If $\mathcal{F}$ is a family of continuous functions on K, the following are equivalent:

(a) $\{f_n\}$ is pointwise bounded and equicontinuous.

(b) $\{f_n\}$ has a uniformly convergent subsequence.

Proof:
(**(b)** $\to$ **(a)**) Let $\epsilon > 0$ be given and let E be a countable dense subset of K. By the above Theorem, $\{f_n\}$ has a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in E$.
Put $f_{n_k} = g_k$ and pick $\delta > 0$ such that $|f_n(x) - f_n(y)| < \epsilon/3$ for all n provided that $d(x, y) < \delta$. Let $V(x, \delta)$ be the set of all $y \in K$ with $d(x, y) < \delta$. Since E is dense in K, and K is compact, there are finitely many points $x_1, \ldots, x_m$ in E such that

$$K \subset V(x_1, \delta) \bigcup \cdots \bigcup V(x, \delta).$$

Since $\{g_k(x)\}$ converges for every $x \in E$, there is an integer N such that

$$|g_k(x_s) - g_l(x_s)| < \epsilon/3$$

whenever $k, l \geq N$, $1 \leq s \leq m$.
If $x \in K$, then by above $x \in V(x_s, \delta)$ for some s, so that

$$|g_k(x) - g_k(x_s)| < \epsilon/3$$

for every k. If $k, l \geq N$, it follows that

$$|g_k(x) - g_l(x)| \leq |g_k(x) - g_k(x_s)| + |g_k(x_s) - g_l(x_s)| + |g_l(x_s) - g_l(x)| < \epsilon.$$

Theorem 50 (Stone Weierstrass Theorem)**.**

If f is a continuous complex function on $[a, b]$, there exists a sequence of polynomials P_n such that

$$\lim_{n \to \infty} P_n(x) = f(x)$$

uniformly on $[a, b]$. If f is real, P_n may be taken real.

Some Special Functions

Theorem 51 (Uniqueness of the Gamma Function)**.**

If f is a positive function on $(0, \infty)$ such that

(a) $f(x + 1) = x f(x)$,

(b) $f(1) = 1$,

(c) $\log f$ is convex,

then $f(x) = \Gamma(x)$, where

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.$$

Theorem 52 (Trigonometric Polynomial Approximation)**.**

If f is continuous (with period 2π) and if $\epsilon > 0$, then there is a trigonometric polynomial P such that

$$|P(x) - f(x)| < \epsilon$$

for all real x.
Note: A trigonometric polynomial is a finite sum of the form

$$P(x) = a_0 + \sum_{n=1}^{N} a_n \cos(nx) + b_n \sin(nx)$$

$a_0, \ldots, a_N, b_1, \ldots, b_n \in \mathbb{C}$.

Definition 9.

If f is integrable on $[a, b]$, the Fourier Series of f with respect to $\{\psi_k\}_{k=1}^\infty$ is $f \sim \sum_{n=1}^{\infty} c_n\psi_n$, where c_n is given by $\int_a^b f(x)\overline{\psi_n(x)}dx$.
Note: $\{\psi_k\}$ is an orthonormal set in $L^2[a, b]$ such that

$$\int_a^b \psi_i(x)\overline{\psi_j(x)} = \delta_{ij}.$$

Some examples of orthonormal sets are:

$$\left\{\frac{1}{\sqrt{2\pi}}e^{inx} : n \in \mathbb{Z}\right\}$$

$$\left\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos(nx), \frac{1}{\sqrt{\pi}}\sin(mx) : n, m = 1, 2, 3. \ldots\right\}$$

Theorem 53 (Fourier Series Is the Best Approximation)**.**

Let $\{\psi_n\}$ be orthonormal on $[a, b]$. Then

$$\int_a^b |f(x) - \sum_{k=1}^{n} c_k\psi_k(x)|^2dx \le \int_a^b |f(x) - \sum_{k=1}^{n} \gamma_k\psi_k(x)|^2dx$$

for any $\{\gamma_i\}$ set of numbers. In particular, $\sum_{n=1}^{\infty} |c_n|^2 \le \int_a^b |f(x)|^2dx$.

Proof:
Let $t_n = \sum_{k=1}^{n} \gamma_k\psi_k(x)$, $s_n = \sum_{k=1}^{n} c_k\psi_k(x)$.

$$\int_a^b f(x)\overline{t_n}dx = \int_a^b f(x)\sum_{k=1}^{n} \overline{\gamma}_k\overline{\psi}_k(x) = \sum_{m=1}^{n} c_k\overline{\gamma}_k$$

by the definition of $\{c_k\}$. Also,

$$\int_a^b |t_n|^2dx = \int_a^b \sum_{k=1}^{n} \gamma_k\psi_k(x)\sum_{k=1}^{n} \overline{\gamma}_k\overline{\psi}_k(x)dx = \sum_{k=1}^{n} |\gamma_k|^2$$

since $\{\psi_k\}$ is orthonormal. So

$$\int_a^b |f(x)-t_n|^2 dx = \int_a^b |f(x)|^2 - \int_a^b f(x)\bar{t}_n dx - \int_a^b \overline{f}(x) t_n dx + \int_a^b |t_n|^2 dx.$$

$$= \int_a^b |f(x)|^2 dx - \sum_{k=1}^n c_k \overline{\gamma}_k - \sum_{k=1}^n \overline{c}_k \gamma_k + \sum_{k=1}^n \gamma_k \overline{\gamma}_k$$

$$= \int_a^b |f(x)|^2 - \sum_{k=1}^n |c_k|^2 + \sum_{k=1}^n |\gamma_k - c_k|^2,$$

which is minimized if and only if $\gamma_k = c_k$.
Putting $\gamma_k = c_k$ in this calculation, we get

$$\sum_{k=1}^n |c_k|^2 \le \int_a^b |f(x)|^2 dx,$$

since $\int_a^b |f(x)-t_n|^2 dx \ge 0$.
If we take the limit as $n \to \infty$, we get that

$$\sum_{n=1}^\infty |c_n|^2 \le \int_a^b |f(x)|^2 dx.$$

Theorem 54 (Parseval's Theorem)**.**

Suppsose f and g are Riemann-integrable functions with period 2π and

$$f(x) \sim \sum_{-\infty}^{\infty} c_n e^{inx}, g(x) \sim \sum_{-\infty}^{\infty} \gamma_n e^{inx}.$$

Then

(a)

$$\lim_{N\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - s_N(f;x)|^2 dx = 0,$$

where $s_N(f;x) = \sum_{-N}^{N} c_n e^{inx}$;

(b)

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)} dx = \sum_{-\infty}^{\infty} c_n \overline{\gamma}_n;$$

(c)

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx = \sum_{-\infty}^{\infty} |c_n|^2.$$