Real Analysis Notes

Contents

1. Sequences

Definition (convergence): A sequence (a_n) converges to $a \in \mathbb{R}$ if for all $\varepsilon > 0$, there exists some N such that $n > N \Rightarrow |a_n - a| < \varepsilon$.

Definition (divergence): A sequence can either diverge to positive infinity (for all $M > 0$, there exists an N such that $n > N \Rightarrow a_n > M$), negative infinity (for all $M < 0$, there exists an N such that $n > N \Rightarrow a_n < M$), or neither, in which case the limit does not exist.

Proposition: If a sequence converges, then the limit is unique.

Proof: Suppose $a_n \to x, y$, where $x \neq y$. We know that $|a_n - x|, |a_n - y| < \frac{\varepsilon}{2}$ for arbitrarily large n. Thus we have

$$|x - y| \leq |x - a_n| + |a_n - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

However, this holds for every $\varepsilon > 0$, which implies $x = y$, a contradiction. ■

Proposition: Convergent sequences are bounded.

Proof: Note that eventually $|a_n - a| < 1$, where $\lim_{n\to\infty} a_n = a$. Thus $1 - a < a_n < 1 + a$. Now just take the max and min of the finitely many terms that occur before this happens to get bounds on a_n. ■

Proposition:
a) $(c \cdot a_n) \to c \cdot a$
b) $(a_n + b_n) \to a + b$
c) $(a_n - b_n) \to a - b$
d) $(a_n \cdot b_n) \to a \cdot b$
e) $\left(\frac{a_n}{b_n}\right) \to \frac{a}{b}$

Proof:
a) Suppose $\varepsilon > 0$. Then there exists N such that for all $n \geq N$, we have

$$|a_n - a| < \frac{\varepsilon}{|c|} \Rightarrow |c \cdot a_n - c \cdot a| < \varepsilon.$$

Thus $\lim_{n\to\infty} c \cdot a_n = c \cdot a$.
b) Suppose $\varepsilon > 0$. Then there exists N_1, N_2 such that for all $n_1 \geq N_1, n_2 \geq N_2$, we have

$$\left|a_{n_1} - a\right|, \left|b_{n_2} - b\right| < \frac{\varepsilon}{2}.$$

Set $N = \max\{N_1, N_2\}$. Then, for all $n \geq N$, we have

$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus $\lim_{n\to\infty}(a_n + b_n) = a + b$.
c) Negate (b_n) and use the last two bullets.
d) Since (a_n) converges, we have $|a_n| \leq C$ for some C for all n. There exists some N_1 such that for all $n \geq N_1$, we have $|a_n - a| < \frac{\varepsilon}{2|b|+1}$ (note that $2|b| + 1 > 0$). Similarly, there exists some N_2 such that for all $n \geq N_2$, we have $|b_n - b| < \frac{\varepsilon}{2C+1}$ (note that $2C + 1 > 0$). Let $N = \max\{N_1, N_2\}$. Then, for all $n \geq N$, we have

$$|a_n b_n - ab| = |a_n b_n - a_n b + a_n b - ab| \leq |a_n||b_n - b| + |a_n - a||b| < C \cdot \frac{\varepsilon}{2C + 1} + |b| \cdot \frac{\varepsilon}{2|b| + 1} < \varepsilon.$$

Thus $\lim_{n\to\infty} a_n b_n = ab$.

e) Reciprocate (b_n) (assuming only finitely many terms are 0), and apply the last bullet.

■

Proposition: Suppose (a_n) and (b_n) convergent series and $a_n \leq b_n$ for all n. Then

$$\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n.$$

Proof: Let $a_n \to A$ and $b_n \to B$, and suppose for the sake of contradiction that $A > B$. Then for sufficiently large n we have

$$|a_n - A| < \frac{A-B}{2} \quad \text{and} \quad |b_n - B| < \frac{A-B}{2}.$$

Expanding the absolute values yields

$$\frac{3B-A}{2} < b_n < \frac{A+B}{2} < a_n < \frac{3A-B}{2},$$

which is a contradiction. ■

Theorem (squeeze theorem): Suppose $a_n \leq x_n \leq b_n$ for arbitrarily large n and $a_n, b_n \to L$. Then $x_n \to L$.

Proof: We have

$$L - \varepsilon < a_n \leq x_n \leq b_n < L + \varepsilon$$

for arbitrarily large n, which implies $|x_n - L| < \varepsilon$. ■

Theorem (monotone convergence theorem): A monotone sequence converges if and only if it is bounded. Further, if the sequence is increasing and bounded, then it converges to the supremum of the set of elements of the sequence. If it's decreasing and bounded, then it converges to the infinum of the set of the elements of the sequence. If a monotone sequence diverges, then it diverges to ∞ or $-\infty$, depending on if it's increasing or decreasing.

Proof: If the sequence converges, then clearly it's bounded. Now suppose the sequence is monotone increasing and bounded. Let (a_n) be the sequence and let $S = \{a_n \mid n \geq 1\}$. Since the sequence is bounded, S is bounded, so $\sup(S)$ exists. We claim that $\lim_{n\to\infty} a_n = \sup(S)$. By definition of supremum, for all $\varepsilon > 0$, there exists some N such that $\sup(S) - \varepsilon \leq a_N \leq \sup(S)$. Since the sequence is increasing, this implies $\sup(S) - \varepsilon \leq a_N \leq a_n \leq \sup(S)$ for all $n \geq N$. This implies that $|\sup(S) - a_n| < \varepsilon$ for all $n \geq N$, which means a_n converges to $\sup(S)$ as desired. Negating the sequence proves the infinum case.

The divergence part of the theorem just means that the sequence doesn't bounce around, which is obvious from monotonicity. ■

Proposition: Suppose $S \subseteq \mathbb{R}$ is bounded above. Then there exists a sequence (a_n) where $a_n \in S$ for each n and

$$\lim_{n\to\infty} a_n = \sup(S).$$

Similarly, if S is bounded below, then there exists a sequence (b_n) where $b_n \in S$ for each n and

$$\lim_{n\to\infty} b_n = \inf(S).$$

Proof: We prove the infinum case, as the supremum case follows upon negation.

Note by definition, for each $n \geq 1$, there exist some $x \in S$ such that $\inf(S) \leq x \leq \inf(S) + \frac{1}{n}$. Let such an x be a_n. Then we have

$$\inf(S) \leq a_n \leq \inf(S) + \frac{1}{n}.$$

Note that both the left and the right converge to $\inf(S)$, and thus by the squeeze theorem, (a_n) must also converge to $\inf(S)$. $\blacksquare$

Proposition: A sequence converges to a if and only if every subsequence converges to a.

Proof: Since the oringinal sequence is a subsequence, if all subsequences converge, then so does the original.

Now suppose the original sequence $(a_n) \to a$, and consider some arbitrary subsequence $\left(a_{n_k}\right)$. For all $\varepsilon > 0$, there exists N such that for all $n \geq N$, we have $|a_n - a| < \varepsilon$. Now note that for all $k \geq K$ for some K, we have $n_k \geq N$. Thus, for all $k \geq K$, we have $\left|a_{n_k} - a\right| < \varepsilon$, which means $\left(a_{n_k}\right) \to a$. $\blacksquare$

Proposition: If a monotone sequence (a_n) has a convergent subsequence, then (a_n) converges to the same limit.

Proof: Suppose the sequence is monotone increasing (decreasing is proved the exact same). Then clearly the subsequence is increasing as well. We know by the monotone convergence theorem that

$$\lim_{n\to\infty} a_{n_k} = \sup\left(\left\{a_{n_k} : k \geq 1\right\}\right).$$

Then since $k \leq n_k$ for all k (n_k is a subsequence of $\mathbb{N}$), we have

$$a_k \leq a_{n_k} \leq \sup\left(\left\{a_{n_k} : k \geq 1\right\}\right).$$

Thus (a_n) is bounded, so by monotone convergence, it converges. Thus, since every subsequence converges to the main series' limit, $(a_n) \to \sup\left(\left\{a_{n_k} : k \geq 1\right\}\right)$. $\blacksquare$

Lemma: Every sequence has a monotone subsequence.

Proof: Let (a_n) be the sequence. Define a peak to be an element of the sequence that's bigger than every later element. First suppose the sequence has finitely many peaks. To start the subsequence, pick the next element after the last peak. Then, since there are no more peaks, there must be an element bigger than the one chosen. We can keep doing this and get an increasing subsequence.

Now suppose there are infinitely many peaks. Then each peak must be less than the previous one by definition, so the subsequence of peaks is monotone decreasing. ■

Theorem (Bolzano-Weierstrass theorem): Every bounded sequence has a convergent subsequence.

Proof: Every sequence has a monotone subsequence, and since the original sequence is bounded, this subsequence is bounded. Thus it converges by monotone convergence. ■

Definition (Cauchy): A sequence (a_n) is *Cauchy* if for all $\varepsilon > 0$ there exists some N such that $|a_m - a_n| < \varepsilon$ for all $m, n \geq N$.

Proposition: Every Cauchy sequence is bounded.

Proof: There exists N such that for all $m, n \geq N$, we have

$$|a_m - a_n| < 1.$$

Thus, for all $m \geq N$, we have

$$|a_m - a_N| < 1.$$

This bounds a_m with $m \geq N$ between $a_N - 1$ and $a_N + 1$. Then, simply take the maximum and minimum of all the previous terms to see that the sequence is indeed bounded. ■

Theorem: A sequence converges if and only if it is Cauchy.

Proof: First suppose $(a_n) \to a$. Then, for all $\varepsilon > 0$, there exists N such that for all $n \geq N$, we have $|a_n - a| < \frac{\varepsilon}{2} \Rightarrow -\frac{\varepsilon}{2} < a_n - a < \frac{\varepsilon}{2}$. For any $m \geq N$, we also have $|a_m - a| < \frac{\varepsilon}{2} \Rightarrow -\frac{\varepsilon}{2} < a - a_m < \frac{\varepsilon}{2}$. Adding the two yields

$$-\varepsilon < a_n - a_m < \varepsilon \Rightarrow |a_n - a_m| < \varepsilon$$

for all $n, m \geq N$. Thus (a_n) is Cauchy.

Now suppose (a_n) is Cauchy. Thus, (a_n) is bounded, and so by Bolzano Weierstrass, there is some convergent subsequence. Let this subsequence be $\left(a_{n_k}\right) \to a$. Thus for all $\varepsilon > 0$, there exists K such that for all $n_k \geq K$, we have

$$\left|a_{n_k} - a\right| < \frac{\varepsilon}{2}.$$

Since (a_n) is Cauchy, for all $\varepsilon > 0$, there exists M such that for all $m, n_k \geq M$, we have

$$\left|a_m - a_{n_k}\right| < \frac{\varepsilon}{2}.$$

Let $N = \max\{K, M\}$. Let $m, n_k \geq N$. Then both inequalities are true. Adding the two and using the triangle inequality yields

$$|a_m - a| \leq \left|a_m - a_{n_k}\right| + \left|a_{n_k} - a\right| < \varepsilon.$$

Thus $(a_n) \to a$. ∎

Definition (limsup and liminf): Let (x_n) be a sequence. Then define

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty}\left(\sup_{m\geq n} x_m\right)$$

and

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty}\left(\inf_{m\geq n} x_m\right).$$

Remark: Most of the time we will write $\limsup x_n$ to signify the limit superior, and similarly for the limit inferior.

Proposition: Let $\limsup_{n\to\infty} x_n = L$ and $\liminf_{n\to\infty} x_n = M$. Then, for all $\varepsilon > 0$, there exists N_1 such that for all $n \geq N_1$ we have

$$L + \varepsilon > x_n.$$

Similarly, there exists some N_2 such that for all $n \geq N_2$ we have

$$x_n > M - \varepsilon.$$

Proof: We prove the infinum case, as the superior case follows similarly.

We proceed by contradiction. Suppose there exists some $\varepsilon > 0$ such that for all N, there exists some $n \geq N$ such that $x_n \leq M - \varepsilon$. Thus $\inf_{m\geq n} x_m \leq M - \varepsilon$. Thus we have

$$\varepsilon \leq M - \inf_{m\geq n} x_m = \left|M - \inf_{m\geq n} x_m\right|.$$

However, this is a contradiction, since $\liminf x_n = M$. ∎

Proposition: Let $L = \limsup_{n\to\infty} x_n$ and $M = \liminf_{n\to\infty} x_n$. Then, for all $\varepsilon > 0$, there exist infinitely many N such that

$$L \geq x_N \geq L - \varepsilon$$

and

$$M \leq x_N \leq M + \varepsilon.$$

Proof: We do the supremum case, as the infinum case follows similarly.

Let $\varepsilon > 0$. Suppose for the sake of contradiction that there are only finitely many N such that $L \geq x_N \geq L - \varepsilon$. Let N' be the last of these. Then, for all $n > N'$, we have

$$L - \varepsilon > x_n.$$

This implies that for all $n > N'$, we have

$$\sup_{m\geq n} x_m \leq L - \varepsilon < L \Rightarrow \sup_{m\geq n} x_n - L \leq -\varepsilon < 0 \Rightarrow \left|\sup_{m\geq n} x_n - L\right| \geq \varepsilon.$$

However, this contradicts the fact that $L = \limsup_{n\to\infty} x_n$. Thus we have a contradiction. ■

Proposition: Suppose (x_n) is a bounded sequence. Then there is a subsequence that converges to $\limsup_{n\to\infty} x_n$ and a subsequence that converges to $\liminf_{n\to\infty} x_n$.

Proof: We prove the supremum case, as the infinum case follows similarly. Let $\limsup_{n\to\infty} x_n = L \in \mathbb{R}$, which exists because (x_n) is bounded.. Let N_1 be the smallest integer such that $L - 1 \leq a_{N_1} \leq L$. Then let N_2 be the smallest integer greater than N_1 such that $L - \frac{1}{2} \leq a_{N_2} \leq L$. We know this must exist since by the previous proposition, there are infinitely such N_2 that satisfy the inequality. We can then inductively build the sequence, taking the smallest interger N_k greater than N_{k-1} such that $L - \frac{1}{k} \leq a_{N_k} \leq L$. Then by the squeeze theorem we have that $\left(x_{N_k}\right)$ converges to L, as desired.

Remark: This also proves Bolzano-Weierstrass.

■

Proposition: A sequence converges if and only if $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n$.

Proof: Note that

$$\inf_{m\geq n} x_m \leq x_n \leq \sup_{m\geq n} x_m$$

by definition. Thus, by the squeeze theorem, we can conclude (x_n) converges to $\limsup x_n = \liminf x_n$.

Now suppose $L = \limsup x_n \neq \liminf x_n = M$. By the previous proposition, there are two subsequences that converge to L and M. Since they converge to different numbers, we must have that $\lim_{n\to\infty} x_n$ does not exist. ■

1.1. Interesting Problems

Problem: Suppose (a_n) is a sequence and $f : \mathbb{N} \to \mathbb{N}$ is a bijection. Prove the following:
a) if (a_n) diverges to ∞, then $\left(a_{f(n)}\right)$ diverges to ∞.
b) if (a_n) converges to L, then $\left(a_{f(n)}\right)$ converges to L.

Solution:
a) We have that for every M, there exists N such that $\forall n \geq N$, we have $a_n > M$. Since f is a bijection, there exists some N' such that for all $n' \geq N'$, we have $f(n') \geq N$ (this is because eventually every number less than N, it will be an output of some input to f). Thus $\left(a_{f(n)}\right)$ dose diverge to infinity.
b) Basically the same as before, except we have the convergence condition.

Problem: Suppose (a_n) is a sequence for which $a_n \to a$. Define

$$b_n = \frac{a_1 + \dots + a_n}{n}.$$

Prove that $b_n \to a$.

Solution: Suppose $\forall n \geq N$, we have $|a_n - a| < \frac{\varepsilon}{2}$ for some $\varepsilon > 0$. Let $M = \max\{|a_k - a| : k < N\}$. For all $n \geq \frac{2M(N-1)}{\varepsilon}$, we have

$$\begin{aligned}
\left|\frac{(a_1 - a) + \dots + (a_n - a)}{n}\right| &\leq \frac{1}{n}(|a_1 - a| + \dots + |a_n - a|) \\
&< \frac{1}{n}\left(M(N-1) + \frac{\varepsilon}{2}(n - N)\right) \\
&= \frac{M(N-1)}{n} + \frac{\varepsilon}{2}\left(1 - \frac{N}{n}\right) \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Thus $b_n \to a$.

Problem: Let a_1, a_2 be real numbrers, and define

$$a_n := \frac{a_{n-1} + a_{n-2}}{2}.$$

Does (a_n) converge?

Solution: Using the characteristic equation, we have

$$a_n = \frac{2a_2 + a_1}{3} + \frac{4}{3}(a_2 - a_1)\left(-\frac{1}{2}\right)^n.$$

Letting $n \to \infty$ yields $a_n \to \frac{2a_2+a_1}{3}$.

2. Series

Definition (series convergence): A series converges if the sequence of its partial sums converges.

Proposition: Suppose $\sum_{i=1}^{\infty} a_i = A$ and $\sum_{i=1}^{\infty} b_i = B$.
a) $\sum_{i=1}^{\infty} (a_i + b_i) = A + B$.
b) For any $c \in \mathbb{R}$, we have $\sum_{i=1}^{\infty} c \cdot a_i = c \cdot A$.

Proof:
a) Let (s_n) be the sequence of partial sums for (a_n), and define (t_n) similarly. We have $\sum_{i=1}^{n} (a_i + b_i) = s_n + t_n$. Thus the sequence of partial sums for the sum of the series is $(s_n + t_n)$. Then limit laws imply that the partial sums converge to $A + B$.
b) Follows from the same argument as the previous bullet.

■

Proposition (divergence test): If $a_k \nrightarrow 0$, then $\sum_{i=1}^{\infty} a_i$ diverges.

Proof: We prove the contrapositive. If the sum converges, then the sequence of partial sums is Cauchy. Thus, if $\varepsilon > 0$, there exists N such that $\forall n \geq m \geq N$, we have

$$\left|a_m + a_{m+1} + \cdots + a_n\right| < \varepsilon.$$

Letting $n = m$ yields

$$\left|a_n\right| < \varepsilon,$$

which implies $a_n \to 0$. ■

2.1. Riemann Rearrangement Theorem

A whole section for this because why not.

Lemma: Let $\sum_{n=1}^{\infty} a_n = A$ and $\sum_{n=1}^{\infty} b_n = B$. Define an interlacing of the two sequences as the combination of the two sequences such that if a_k appears in the new sequence, the next term in this sequence that appears from (a_n) is a_{k+1}, and similarly for B. Then the sum of this interlacing series converges to $A + B$.

Proof: Let (c_n) be the interlacing. Pick N_1 such that $\forall n \geq N_1$, we have

$$\left|\sum_{k=1}^{n} a_k - A\right| < \frac{\varepsilon}{2}.$$

Define N_2 similarly for b_n. Define M_1 such that $a_{N_1} = c_{M_1}$, define M_2 similarly for b_n, and let $M = \max\{M_1, M_2\}$. Thus we have

$$\left|\sum_{k=1}^{M} c_k\right| = \left|\sum_{k=1}^{n_1} a_k - A + \sum_{k=1}^{n_2} b_k - B\right| < \left|\sum_{k=1}^{n_1} a_k - A\right| + \left|\sum_{k=1}^{n_2} b_k - B\right|.$$

Since $n_1 \geq N_1$ and $n_2 \geq N_2$ (because of our choice of M), we have

$$\left|\sum_{k=1}^{n_1} a_k - A\right| + \left|\sum_{k=1}^{n_2} b_k - B\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus the interlaced sequence does converge to $A + B$. ■

Lemma: Suppose $\sum_{n=1}^{\infty} a_i$ is a conditionally convergent sequence. Let a_n^+ be nth positive term in the series, and define a_n^- similarly. Then we have

$$\sum_{i=1}^{\infty} a_n^+ = \infty \quad \text{and} + \sum_{i=1}^{\infty} a_n^- = -\infty.$$

Proof: If the two series were to converge to real numbers, we could take the absolute value of the negative series, and then by the previous lemma, for any interlacing, we'll get a convergent series. Since the absolute value of our initial series is an interlacing of the two, that would imply the series is absolutely convergent, which is a contradiction.

Now suppose the positive series diverges and the negative series converges to $-L$ with $L > 0$ (the opposite case is shown to be impossible similarly). For all M, there exists N such that $\forall n \geq N$, we have

$$\sum_{i=1}^{n} a_i^+ > M + L.$$

Pick N' such that a_N^+ shows up in $(a_n)_{1 \leq n \leq N'}$. Then $\forall n \geq N'$, we have

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n_1} a_i^+ + \sum_{i=1}^{n_2} a_i^-.$$

Here we must have $n_1 \geq N$. Thus we have

$$\sum_{i=1}^{n_1} a_i^+ + \sum_{i=1}^{n_2} a_i^- > (M + L) - L = M.$$

Thus the initial series diverges, which is a contradiction. ■

Theorem (Riemann rearrangement theorem): Suppose $\sum_{i=1}^{\infty} a_i$ is a conditionally convergent series. We can find a bijection $f : \mathbb{N} \to \mathbb{N}$ such that for any $\alpha, \beta \in \mathbb{R} \cup \{-\infty, \infty\}$ with $\alpha \leq \beta$, we have

$$\limsup_{n\to\infty}\left(\sum_{i=1}^{n} a_{f(i)}\right) = \beta \quad \text{and} \quad \liminf_{n\to\infty}\left(\sum_{i=1}^{n} a_{f(i)}\right) = \alpha.$$

Proof: Let a_n^+ be the nth positive term in the series, and define a_n^- similarly. By the previous lemma, we have $\sum_{i=1}^{\infty} a_n^+ = \infty$ and $\sum_{i=1}^{\infty} a_n^- = -\infty$. Note that $(a_n) \to \infty$ since the series converges, and since (a_n^+) and (a_n^-) are subsequences, they both must also converge to 0.

We break off into cases:
a) $-\infty < \alpha \leq \beta < \infty$
b) $\beta = \infty$ and α is finite or $\alpha = -\infty$ and β is finite.
c) $\beta = \infty, \alpha = -\infty$

Part a)

Without loss of generality, supose $\beta \geq 0$ (if it wasn't, we'd just the process of creating the rearrangement with negative terms). Let P_1 be the smallest number such that

$$\sum_{i=1}^{P_1} a_n^+ \geq \beta.$$

Let N_1 be the smallest number such that

$$\sum_{i=1}^{P_1} a_n^+ + \sum_{i=1}^{N_1} a_n^- \leq \alpha.$$

Now inductively define P_k as the smallest number such that

$$\sum_{i=1}^{P_1} a_n^+ + \sum_{i=1}^{N_1} a_n^- + \cdots + \sum_{i=P_{k-1}+1}^{P_k} a_i^+ \geq \beta$$

and define N_k to be the smallest number such that

$$\sum_{i=1}^{P_1} a_n^+ + \sum_{i=1}^{N_1} a_n^- + \cdots + \sum_{i=P_{k-1}+1}^{P_k} a_i^+ + \sum_{i=N_{k-1}+1}^{N_k} a_n^- \leq \alpha.$$

This is alawys possible, since we know the positive series and the negative series both diverge to infinities, so starting from any point through the series and adding further terms will still diverge to infinity.

Let (b_n) be the partial sums of the rearranged series, where the rearranged series is

$$a_1^+, a_2^+, \ldots, a_{P_1}^+, a_1^-, a_2^-, \ldots, a_{N_1}^-, a_{P_1+1}^+, \ldots$$

We prove the limsup of the series converges to β. The liminf follows similarly.

Pick $\varepsilon > 0$. There exists some M such that $\forall n \geq M$, we have $|a_n^+| < \varepsilon$. Thus this holds $\forall P_k \geq M$. By construction, we have $b_{P_1+N_1+\cdots+N_{k-1}+P_k-1} \leq \beta \leq b_{P_1+N_1+\cdots+N_{k-1}+P_k}$. Thus we must have $\beta \leq b_{P_1+N_1+\cdots+N_{k-1}+P_k} \leq \beta + \varepsilon$ work any $P_k \geq M$. Note that again by construction, the supremum of a tail of the partial sums sequence will be $b_{P_1+N_1+\cdots+N_{k-1}+P_k}$ for some k. Thus, we have for any $m \geq P_k$ (where $P_k \geq M$ for some working k), we have

$$\left| \sup_{n \geq m} b_n - \beta \right| < \varepsilon.$$

Thus we have $\limsup_{n \to \infty} b_n = \beta$. ■

2.2. Interesting Problems

3. The Topology of $\mathbb{R}$

Definition (open set): A set $U \subseteq \mathbb{R}$ is *open* if for every $x \in U$, there is a number $\delta > 0$ such that

$$(x - \delta, x + \delta) \subseteq U.$$

This is called the δ neighborhood of x, and is denoted $V_\delta(x)$.

Proposition:
a) If $\{U_\alpha\}$ is a collection of open sets, then $\bigcup_\alpha U_\alpha$ is also an open set.
b) If $\{U_\alpha\}$ is a finite collection of open sets, then $\bigcap_\alpha U_\alpha$ is an open set.

Proof:
a) Consider some $x \in \bigcup_\alpha U_\alpha$. Then $x \in U_i$ for some i. Thus for some δ, $V_\delta(x) \subseteq U_i$. Thus $V_\delta(x) \subseteq \bigcup_\alpha U_\alpha$, so $\bigcup_\alpha U_\alpha$ is open.
b) Consider some $x \in \bigcap_\alpha U_\alpha$. For each U_i in $\{U_\alpha\}$, there exists δ_i such that $V_{\delta_i}(x) \subseteq U_i$. Let $\delta = \min\{\delta_\alpha\}$. Then $V_\delta(x) \subseteq V_{\delta_i}(x) \subseteq U_i$. Thus $V_\delta(x) \subseteq \bigcap_\alpha U_\alpha$.

■

Theorem: Every open set is a countable union of disjoint open intervals.

Proof: Let A be an open set. For $x \in A$, let $I_x = (\alpha, \beta)$, where $\alpha = \inf\{a : (a, x) \subseteq A\}$ and $\beta = \sup\{b : (x, b) \subseteq A\}$. For any x, y, we must have $I_x = I_y$ or $I_x \cap I_y = \emptyset$, because if they overlap but aren't equal, then you could extend one of them, contradicting us choosing the largest possible interval.

We claim that these intervals make up A. Note that for every $x \in A$ we have $x \in I_x \subseteq A$, so the union of all the intervals is A. Further, because $\mathbb{Q}$ is dense if $\mathbb{R}$, every open interval contains a rational number, so there cannot be more intervals than rationals. Thus the number of intervals is countable. ■

Definition (closed set): A set $A \subseteq \mathbb{R}$ is *closed* if A^c is open.

Definition (limit point): A point x is a limit point of a set A if there is a sequence of points $a_1, a_2 \ldots$ from $A \setminus \{x\}$ such that $a_n \to x$.

Theorem: A set is closed if and only if it contains all its limit points.

Proof: First we show that if a set is closed, then it contains all its limit points. We proceed by contradiction. Let x be a limit point not in A. Then we have the following:
- There exists a sequence (a_n) with each term in A such that $\lim_{n \to \infty} a_n = x$.
- $x \in A^c$, which is an open set, so there exists δ such that $V_\delta(x) \subseteq A^c$.

Since the sequence converges to x, we must have $|a_n - x| < \delta \Rightarrow x - \delta < a_n < x + \delta$ for all $n \geq N$ for some N. Thus implies $a_n \in V_\delta(x)$ for all $n \geq N$. However, this is impossible, since $a_n \in A$, while $V_\delta(x) \subseteq A^c$. Thus we have a contradiction.

Now we prove the other by contrapositive, that is we prove that if a set is not closed, then it doesn't contain all its limit points. Suppose A is not closed. Then A^c is not open. Thus, there exists some $x \in A^c$ such that every δ neighborhood of x contains some element not in A^c, which is equivalent to it containing an element in A. Let a_n be an element in A that is contained in the $\frac{1}{n}$ neighborhood of x. We claim $\lim_{n \to \infty} a_n = x$, which proves the claim.

Let $\varepsilon > 0$, and pick some integer k such that $\frac{1}{k} < \varepsilon$. Then we have $|a_k - x| < \frac{1}{k} < \varepsilon$ by definition. Note that this implies $|a_n - x| < \frac{1}{k} < \varepsilon$ for all $n \geq k$, since if a_n is in a $\frac{1}{n}$ neighborhood of x, then it's also in a $\frac{1}{k}$ neighborhood of x, which is further in an ε neighborhood of x. Thus a_n converges to x. ■

Proposition:
- If $\{U_\alpha\}$ is a finite collection of closed sets, then $\bigcup_\alpha U_\alpha$ is also a closed set.
- If $\{U_\alpha\}$ is a collection of closed sets, then $\bigcap_\alpha U_\alpha$ is also a clsoed set.

Proof: Follows from the union/intersection proposition of open sets and De Morgan's laws. ■

Definition (cover): Let A be a set. The collection of sets $\{U_\alpha\}$ are a *cover* of A if

$$A \subseteq \bigcup_\alpha U_\alpha .$$

If each U_α is open, then $\{U_\alpha\}$ is an *open cover* of A. If a finite subset of $\{U_\alpha\}$ is a cover of A, then that subset is a finite subcover of A.

Definition (compact): A set A is *compact* if every open cover of A contains a finite subcover of A.

Theorem (Heine-Borel theorem): A set $S \subseteq \mathbb{R}$ is compact if and only if S is closed and bounded.

Proof: Suppose S is compact. Then for every open cover of S, there exists a finite subcover. Let $I_n = (-n, n)$. Clearly the set $\{I_n\}_{n\geq 1}$ is an open cover of S. Thus, there exists a sequence $n_1, ..., n_k$ such that $\left\{I_{n_1}, ..., I_{n_k}\right\}$ is an open cover of S. WLOG $n_1 < n_2 < \cdots < n_k$. We have

$$S \subseteq \bigcup_{j=1}^{k} I_{n_j} = I_{n_k}.$$

Since I_{n_k} is bounded, then clearly S is bounded.

Next we show that S is closed by contradiction. Suppose S is compact and doesn't contain all its limit points. That is, there exists a sequence (a_n) contained in S that converges to some point x not in S. Let $I_n = \left(-\infty, x - \frac{1}{n}\right) \cup \left(x + \frac{1}{n}, \infty\right)$. Clearly

$$S \subseteq \bigcup_{k=1}^{\infty} I_k = \mathbb{R} / \{x\}.$$

Suppose $\left\{I_{n_1}, ..., I_{n_k}\right\}$ is a finite subcover, where the indexes are increasing. Then $\cup_{j=1}^{k} I_{n_j} = I_{n_k}$. Thus

$$S \subseteq I_{n_k} = \left(-\infty, x - \frac{1}{n_k}\right) \cup \left(x + \frac{1}{n_k}, \infty\right),$$

which implies

$$S \cap \left(x - \frac{1}{n_k}, x + \frac{1}{n_k}\right) = \emptyset.$$

However this is a contradiction, since the equation implies the sequence cannot converge to x without containing elements outside of S, contradiction.

Now we prove the other direction. Suppose S is closed and bounded. For every $x \in \mathbb{R}$, define $S_x = S \cap (-\infty, x]$. Suppose $\mathcal{F}$ is an open cover of S, and let

$$B = \{x : \mathcal{F} \text{ contains a finite subcover of } S_x\}.$$

Note that since S is bounded, $M = \sup(S)$ and $L = \inf(S)$ exist. We also know that there exist sequences that converge to both, so they're limit points. Thus, since S is closed, we have $L, M \in S$.

We want to show that $M \in B$, since $S_M = S$. Note that we already have $L \in B$, since the cover $\mathcal{F}$ must contain some set that covers L, just take that set as the subcover.

Assume for the sake of contradiction that $M \notin B$. Then clearly we can't have $x \in B$ for any $x \geq M$, since otherwise we would get a finite subcover for $S \cap (-\infty, x]$ as well. Thus $x < M$ for all $x \in B$, which implies B is bounded from above. Since B is nonempty, we can then let $T = \sup(B)$. Note also that B contains infinitely elements, since for all $x \in B$, any number less than x is also in b, and if $\sup(B) = L$, then we can show by a similar argument as for the first case (next paragraph) that this is impossible. Since B has infinitely many elements, this implies that for all $\varepsilon > 0$, there's some element $b \in B$ such that $T - \varepsilon < b < T$.

We have two cases:

Case 1: $T \in S$

Since S is compact, some finite subset of $\mathcal{F}$ covers S, which contains some open set that contains T, call it U. Pick $\delta = \min\{\delta_1, \delta_2\}$, where $T + \delta_1 < M$ and $V_{\delta_2}(x) \subseteq U$. Thus we have $\left(T - \delta, T + \frac{\delta}{2}\right] \subseteq U$.

Note that $T - \delta \in B$, since if not, then we can't have $T \in B$ via the same argument we made to show that $x < M$ for all $x \in B$. Thus, there exists some finite subcover F of $\mathcal{F}$ that covers $S_{T-\delta}$. However, note that this implies $F \cup \left(T - \delta, T + \frac{\delta}{2}\right)$ covers $S_{T+\frac{\delta}{2}}$, which contradicts $T = \sup(B)$. Thus we have a contradiction in this case.

Case 2: $T \notin S$

Since $T \notin S$ and S is closed, $T \in S^c$, which is an open set. Pick δ so that $V_\delta(T) \subseteq S^c$. Thus $\left[T - \frac{\delta}{2}, T + \frac{\delta}{2}\right] \cap S = \emptyset$, which implies $S \cap \left(-\infty, T - \frac{\delta}{2}\right] = S \cap \left(-\infty, T + \frac{\delta}{2}\right]$.

Note we showed that for all $\varepsilon > 0$, there exists $b \in B$ such that $T - \varepsilon < b < T$. Thus, picking $\varepsilon = \frac{\delta}{2}$, we have that $T - \frac{\delta}{2} < a \in B$, which again by an argument made earlier implies that $T - \frac{\delta}{2} \in B$. Thus there's some finite subcover of $S_{T-\frac{\delta}{2}} = S \cap \left(-\infty, T - \frac{\delta}{2}\right]$. However, we also showed that $S_{T-\frac{\delta}{2}} = S_{T+\frac{\delta}{2}}$, so this same subcover works for this set. This implies that $T + \frac{\delta}{2} \in B$, which contradicts $T = \sup(B)$, so we again have a contradiction ■

Theorem (Heine-Borel expanded): Suppose $A \subseteq \mathbb{R}$. The following are equivalent:
a) A is compact.
b) A is closed and bounded.
c) If (a_n) is a sequence of numbers in A, then there is a subsequence $\left(a_{n_k}\right)$ that converges to a point in A.

Proof: The equivalence of a) and b) was the last theorem. Suppose A is closed and bounded. Then any sequence coming from A is bounded, and so has a convergent subsequence by Bolzano-Weierstrass. The limit of this subsequence is clearly a limit point of A, and since A is closed, it must be contained in A.

If A is not closed, then there's some limit point of A not in A. Let (a_n) be a sequence that converges to this limit point. Then every subsequence must also converge to that limit point, which again is not in A.

If A is not bounded, then we can create an unbounded sequence. Just let a_k be some element of A that is greater than k, which must exist since A is unbounded. Clearly every subsequence of (a_n) also diverges. This establishes the equivalence of b) and c). ■

3.1. Interesting Problems

Problem: Construct a set whose set of limit points is $\mathbb{Z}$.

Solution: Let $A_k = \left\{k + \frac{1}{2}, k + \frac{1}{3}, k + \frac{1}{4}, \ldots\right\}$ for all $k \in \mathbb{Z}$. We claim

$$A = \bigcup_{k=-\infty}^{\infty} A_k$$

has $\mathbb{Z}$ as its set of limit points. First note that for each $k \in \mathbb{Z}$, the sequence $a_n = k + \frac{1}{n}$ for $n \geq 2$ converges to k. Thus $\mathbb{Z}$ is a subset of the set of limit points of A.

Now consider some non-integer α. Note that $\{\alpha\} \in \left(\frac{1}{n+1}, \frac{1}{n}\right]$ for some integer n. Then the closest α gets to any element of A is $\min\left\{\left|\{\alpha\} - \frac{1}{n}\right|, \left|\{\alpha\} - \frac{1}{n+1}\right|\right\}$, where each option corresponds to $\lfloor\alpha\rfloor + \frac{1}{n}$ and $\lfloor\alpha\rfloor\frac{1}{n+1}$ respectively. Thus, we can't get arbitrarily close to any non integer α (choose $\varepsilon = \min\left\{\left|\{\alpha\} - \frac{1}{n}\right|, \left|\{\alpha\} - \frac{1}{n+1}\right|\right\}$), so α is not a limit point.

Problem: Prove that the set of limit points of a set is closed.

Solution: We show the complement is open. Let A be our set, and let L be the set of limit points of A. Consider some $x \in L^c$. Since x is not a limit point, this implies that for every sequence $(a_n) \in A \setminus \{x\}$, there exists $\varepsilon > 0$ such that for all N, there exists some $n \geq N$ such that $|a_n - x| \geq \varepsilon$. Note that this inequality implies $\left|a_n - \left(x + \frac{\varepsilon}{2}\right)\right| \geq \frac{\varepsilon}{2}$ and $\left|a_n - \left(x - \frac{\varepsilon}{2}\right)\right| \geq \frac{\varepsilon}{2}$ (we get these by splitting the inequalities into cases based on if the inside if the absolute value is positive or negative). Thus, no sequence in $A \setminus \{x\}$ converges to $x + \frac{\varepsilon}{2}$ or $x - \frac{\varepsilon}{2}$. To show that these are not limit points, we need to show that the sequences can come from $A \setminus \{x \pm \frac{\varepsilon}{2}\}$. However, this isn't an issue. Since any sequence doesn't converge to those values, the sequences can only contain finitely many terms that are $x \pm \frac{\varepsilon}{2}$, so removing those won't affect the convergence. Similarly, adding in any amount of terms equal to x to the sequences won't change the convergence.

Thus we've showed $x \pm \frac{\varepsilon}{2} \in L^c$. In fact, for any $\delta < \frac{\varepsilon}{2}$, we can show that $x \pm \delta \in L^c$ by a similar method to what we did above. Thus we have $V_{\frac{\varepsilon}{2}}(x) \in L^c$ for any $x \in L^c$, which implies L^c is open, which means L is closed, as desired.

Definition (interior, exterior, boundary):

The *interior* of a set A, denoted $\mathrm{Int}(A)$, is the set of points x such that there is an open neighborhood of x that is a subset of A.

The *exterior* of a set A, denoted $\mathrm{Ext}(A)$, is the set of points x such that there is an open of x that is a subset is a subset of A^c.

The *boundary* of set A, denoted ∂A, is the set of points X such that every neighborhood of x contains points in A and A^c.

Problem: Prove that for any set A, we have $\mathbb{R} = \mathrm{Int}(A) \cup \partial A \cup \mathrm{Ext}(A)$, and is a disjoint union.

Solution: It's clear that $\mathrm{Int}(A) \cap \mathrm{Ext}(A) = \emptyset$. Similarly, $\mathrm{Int}(A) \cap \partial = \emptyset$ and $\mathrm{Ext}(A) \cap \partial A = \emptyset$. Now consider some $x \in \mathbb{R}$, and suppose it's not in $\mathrm{Int}(A)$ or $\mathrm{Int}(B)$. Then that implies that every open neighborhood of x contains points in both A and A^c, which means $x \in \partial A$. We can similarly show that if x is not in two of the sets, then it must be in the other one. Thus, we have $\mathbb{R} = \mathrm{Int}(A) \cup \partial A \cup \mathrm{Ext}(A)$.

4. Continuity

4.1. Functional Limits

Definition (functional limit): Let $f : A \to \mathbb{R}$ and let c be a limit point of A. Then

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

if for all $\varepsilon > 0$ there exists some $\delta > 0$ such that for every $x \in A$ for which $0 < |x - c| < \delta$, we have

$$|f(x) - L| < \varepsilon.$$

For the one sided limit $\lim_{x \to c^+}$, the condition on x is relaxed to $c < x < c + \delta$. Similarly, for the one sided limit $\lim_{x \to c^-}$, the condition on x is relaxed to $c - \delta < x < c$.

Proposition: A limit $\lim_{x \to c} f(x)$ can converge to at most one value.

Proof: Suppose $\varepsilon > 0$. Then there exists δ_1 such that when $0 < |x - c| < \delta_1$, we have $|f(x) - L_1| < \frac{\varepsilon}{2}$. There also exists δ_2 such that $|L_2 - f(x)| < \frac{\varepsilon}{2}$. Let $\delta = \min\{\delta_1, \delta_2\}$. Then we have, for all $0 < |x - c| < \delta$, we have

$$|L_2 - L_1| = |(L_2 - f(x)) + (f(x) - L_1)| \leq |L_2 - f(x)| + |f(x) - L_1| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Since this holds for all $\varepsilon > 0$, this implies $L_2 - L_1 = 0$, as desired. ■

Proposition: $\lim_{x \to x} f(x) = L$ if and only if $\lim_{x \to c^+} f(x) = \lim_{x \to c^-} f(x) = L$.

Proof: Suppose $\lim_{x \to c^+} f(x) = \lim_{x \to c^-} f(x) = L$. Thus, for all $\varepsilon > 0$, there exist $\delta_1, \delta_2 > 0$ such that

$$|f(x) - L| < \varepsilon \quad \text{when} \quad c < x < c + \delta_1$$

and

$$|f(x) - L| < \varepsilon \quad \text{when} \quad c - \delta_2 < x < c.$$

Let $\delta = \min\{\delta_1, \delta_2\}$. Then we have

$$|f(x) - L| < \varepsilon \quad \text{when} \quad 0 < |x - c| < \delta,$$

which implies $\lim_{x \to c} f(x) = L$.

Now suppose $\lim_{x \to c} f(x) = L$. Then for all $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(x) - L| < \varepsilon \quad \text{when} \quad 0 < |x - c| < \delta.$$

This implies

$$|f(x) - L| < \varepsilon \quad \text{when} \quad c < x < c + \delta$$

and

$$|f(x) - L| < \varepsilon \quad \text{when} \quad c - \delta < x < c,$$

which implies that both one-sided limits are equal to L. ■

Theorem: Assume $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$, and c is a limit point of A. Then $\lim_{x \to c} = L$ if and only if, for every sequence a_n from A for which $a_n \neq c$ and $a_n \to c$, we have $f(a_n) \to L$.

Proof: We assume that $a_n, x \neq c$.

First suppose $\lim_{x \to c} f(x) = L$. Then for all $\varepsilon > 0$, there exists δ such that

$$0 < |x - c| < \delta \Rightarrow |f(x) - L| < \varepsilon$$

for $x \in A$. Let (a_n) be an arbitrary sequence in A converging to c. Then, there exists N such that for all $n \geq N$, we have $|a_n - c| < \delta$. This implies $|f(a_n) - L| < \varepsilon$ for all $n \geq N$, which shows that $f(a_n) \to L$.

Now supose $\lim_{x \to c} f(x) \neq L$. That is, there exists $\varepsilon > 0$ such that for all $\delta > 0$ there exists $x \in A$ such that $0 < |x - c| < \delta \Rightarrow |f(x) - L| \geq \varepsilon$. In particular, setting $\delta_n = \frac{1}{n}$, there always exists x_n within $0 < |x - c| < \delta_n$ such that $|f(x) - L| \geq \varepsilon$. Clearly $x_n \to c$, while $f(x_n) \not\to L$, so we're done. ■

Proposition (limit laws): Let f and g be functions from $A \subseteq \mathbb{R}$ to $\mathbb{R}$ and let c be a limit point of A. Assume $\lim_{x \to c} f(x) = L$ and $\lim_{x \to c} g(x) = M$. Then
- $\lim_{x \to c} [k \cdot f(x)] = k \cdot L$
- $\lim_{x \to c} [f(x) + g(x)] = L + M$
- $\lim_{x \to c} [f(x) \cdot g(x)] = L \cdot M$
- $\lim_{x \to c} \frac{f(x)}{g(x)} = \frac{L}{M}$ if $M \neq 0$ and $g(x) \neq 0$ for any $x \in A$.

Proof: The limits holds for all sequences converging to c, and these laws apply to sequences, so in turn these laws hold for limits. ■

Theorem (squeeze theorem): Let f, g, h be functions from $A \subseteq \mathbb{R}$ to $\mathbb{R}$, let c be the limit point of A, suppose

$$f(x) \leq g(x) \leq h(x)$$

for all $x \in A$, and suppose

$$\lim_{x \to c} f(x) = L = \lim_{x \to c} h(x).$$

Then

$$\lim_{x \to c} g(x) = L.$$

Proof: Same reasoning as last proposition. ■

4.2. Continuity

Definition (continuity): A function $f : A \to \mathbb{R}$ is *continuous at a point* $c \in A$ if for all $\varepsilon > 0$, there exists some $\delta > 0$ such that for all $x \in A$ where $|x - c| < \delta$ we have

$$|f(x) - f(c)| < \varepsilon.$$

If f is continuous at every point in its domain, then f is called *continuous.*

Remark: Note that if $c \in A$ is not a limit point of A, then it's automatically continuous, since we can pick δ so that $|x - c| < \delta$ contains no values in $A \setminus \{x\}$. Thus the condition $|f(x) - f(c)| < \varepsilon$ is vacuosly true.

Proposition: Let $f : A \to \mathbb{R}$ and $c \in A$. Then the follwing are equivalent:

a) f is continuous at c.
b) For all $\varepsilon > 0$ there exists some $\delta > 0$ such that for all $x \in A$ where $|x - c| < \delta$ we have $|f(x) - f(c)| < \varepsilon$.
c) For any ε-neighborhood of $f(c)$, denoted $V_\varepsilon(f(c))$, there exists some δ neighborhood of c, denoted $V_\delta(c)$, with the property that for any $x \in A$ for which $x \in V_\delta(c)$, we have $f(x) \in V_\varepsilon(f(c))$.
d) For all sequences $(a_n) \in A$ converging to c, we have $f(a_n) \to f(c)$.
e) $\lim_{x \to c} f(x) = f(c)$ if c is a limit point of A.

Proof: a) is equivalent to b) by definition. b) is equivalent to c), just rephrased in term of neighborhoods. The proof that a) is equivalent to d) is basically identical to the proof of sequences converging to c converge to $\lim_{x \to} f(x)$ under f. d) is equivalent to e) using that same theorem. ■

Proposition (continuity laws): Let $f : A \to \mathbb{R}$ and $g : A \to \mathbb{R}$ be continuous at c, and let $c \in A$. Then the follwing are true:

a) $k \cdot f(x)$ is continuous at c for all $k \in \mathbb{R}$.
b) $f(x) + g(x)$ is continuous at c.
c) $f(x) \cdot g(x)$ is continuous at c.
d) $\frac{f(x)}{g(x)}$ is continuous at c, provided $g(x) \neq 0$ for all $x \in A$.

Proof: By the previous proposition, we can rephrase these as limits, and then we apply our limit laws. ■

Problem (continuous compositions): Suppose $A, B \subseteq \mathbb{R}$, $g : A \to B$ and $f : B \to \mathbb{R}$. If g is continuous at $c \in A$, and f is continuous at $g(c) \in B$, then $f \circ g : A \to \mathbb{R}$ is continuous at c.

Proof: Consider an arbitrary sequence (a_n) from A converging to c. Then by continuity we have that $g(a_n) \to g(c)$. Note that $(g(a_n))$ is a sequence in B converging to $f(c)$, so again by continuity we have $f(g(a_n)) \to f(g(c))$. Since (a_n) was arbitrary, this holds for any sequence converging to c. Thus, $f \circ g$ is continuous at c. ■

4.3. Topological Continuity

Definition (pre-image): Let $X, Y \subseteq \mathbb{R}$ and $f : X \to Y$. For $B \subseteq Y$, define the *pre-image* (or *inverse*)

$$f^{-1}(B) = \{x \in X : f(x) \in B\}.$$

Example: Define $f : [2, 10] \to \mathbb{R}$ as $f(x) = 5x$. Letting $B = (1, 20)$, we have

$$f^{-1}(B) = [2, 4) = \left(\frac{1}{5}, 4\right) \cup [2, 10].$$

Theorem: Let $f : X \to \mathbb{R}$. Then f is continuous if and only if for every open set B, we have $f^{-1}(B) = A \cap X$ for some open set A.

Proof: Let f be continuous and let B be an open set in $\mathbb{R}$. We want to show that for any $x_0 \in f^{-1}(B)$, there exists δ such that if $x \in X \cap V_\delta(x)$, then $x \in f^{-1}(B)$. To that end, suppose $x_0 \in f^{-1}(B)$. This implies $f(x_0) \in B$. Thus, there exists ε such that $V_\varepsilon(f(x_0)) \subseteq B$ (since B is open). By continuity, this implies that there exists δ such that if $x \in X$ and $x \in V_\delta(x_0)$, then $f(x) \in V_\varepsilon(f(x_0))$. In particular, this implies that $x \in f^{-1}(V_\varepsilon(f(x_0))) \subseteq f^{-1}(B)$, as desired.

Now suppose for every open set B, $f^{-1}(B) = A \cap X$ for some open set A. Pick $x_0 \in X$ and $\varepsilon > 0$. Note that $V_\varepsilon(f(x_0))$ is open, so we have that $f^{-1}(V_\varepsilon(f(x_0))) = A \cap X$ for some open set

A. Clearly $x_0 \in A \cap X = f^{-1}(V_\varepsilon(f(x_0)))$, and since A is open, there exists some δ such that $V_\delta(x_0) \subseteq A$. In particular, if $x \in X \cap V_\delta(x_0)$, then

$$x \in X \cap V_\delta(x_0) \subseteq X \cap A = f^{-1}(V_\varepsilon(f(x_0))) \Rightarrow f(x) \in V_\varepsilon(f(x_0)).$$

Thus f is continuous, as desired. ■

Remark: Note that this condition guarantees continuity at every point in the domain.

4.4. The Extreme Value Theorem

Proposition: Suppose $f : X \to \mathbb{R}$ is continuous. If $A \subseteq X$ is compact, then $f(A)$ is compact.

Proof: Suppose $\{U_\alpha\}$ is an open cover of $f(A)$. Consider $\{f^{-1}(U_\alpha)\}$. We have $f^{-1}(U_\alpha) = X \cap V_\alpha$ for some open set V_α by the previous proposition. We show that $\{V_\alpha\}$ is an open cover of A. Consider $x_0 \in A$. Then $f(x_0) \in f(A)$, which means $f(x_0) \in U_i$ for some i, which implies $x_0 \in f^{-1}(U_i) = X \cap V_i \Rightarrow x_0 \in V_i$. Thus $\{V_\alpha\}$ is indeed an open cover of A.

Since A is compact, there exists some finite subcover $\{V_1, V_2, ..., V_k\}$. We claim that $\{U_1, U_2, ..., U_k\}$ is a finite subcover of $f(A)$, where V_i corresponds to U_i through $f^{-1}(U_i) = X \cap V_i$. Consider $y_0 \in f(A)$. Then $f(x_0) = y_0$ for some x_0. Thus $x_0 \in V_i$ for some i, since $\{V_\alpha\}$ is an finite subcover of A. However, $x_0 \in X$, which implies $x_0 \in X \cap V_i = f^{-1}(U_i)$. This implies $y_0 = f(x_0) \in U_i$, as desired. ■

Corollary: A continuous function on a compact set is bounded.

Proof: Suppose $f : A \to \mathbb{R}$ is continuous and A is compact. By the previous proposition, $f(A)$ is compact, which means $f(A)$ is bounded. ■

Theorem (extreme value theorem): A continuous function on a compact set attains a maximum and a minimum.

Proof: Suppose $f : A \to \mathbb{R}$ is continuous and A is compact, and let $M = \sup\{f(x) : x \in A\}$ and $L = \inf\{f(x) : x \in A\}$. These exist since $f(A)$ is bounded by the corollary. Note that since $f(A)$ is compact, it's closed and thus contains all its limit points. Since M and L are limit points of $f(A)$, they must both be in $f(A)$. Thus, there exists x_1 and x_2 such that $f(x_1) = M$, $f(x_2) = L$, as desired. ■

4.5. The Intermediate Value Theorem

Lemma: If f is continuous and $f(c) > 0$, then there exists a δ such that $x \in (c - \delta, c + \delta) \Rightarrow f(x) > 0$. Likewise, if f is continuous and $f(c) < 0$, then there exists a δ such that $x \in (c - \delta, c + \delta) \Rightarrow f(x) < 0$.

Proof: Without loss of generality, suppose $f(c) > 0$. Let $\varepsilon = \frac{f(c)}{2}$. Note that by continuity, there exists δ such that

$$|x - c| < \delta \Rightarrow |f(x) - f(c)| < \frac{f(c)}{2}.$$

Unraveling the second inequality yields

$$0 < \frac{f(c)}{2} < f(x) < \frac{3f(c)}{2}$$

for all $x \in (c - \delta, c + \delta)$, as desired. ■

Proposition: If f is continuous on $[a, b]$ and $f(a)$ and $f(b)$ have different signs, then there is some $c \in (a, b)$ for which $f(c) = 0$.

Proof: Without loss of generality, assume $f(a) > 0 > f(b)$. Let

$$A = \{t : f(x) > 0, \forall x \in [a, t]\}.$$

Note that $a \in A$ and $b \notin A$. Thus A is nonempty and bounded above, so let $c = \sup(A)$. If $f(c) = 0$, then we're done.

Otherwise for the sake contradiction, assume $f(c) > 0$. Then by the previous proposition, we know that there exists δ such that $x \in (c - \delta, c + \delta) \Rightarrow f(x) > 0$, which also implies $x \in \left(c - \delta, c + \frac{\delta}{2}\right] \Rightarrow f(x) > 0$. Note that $c - \delta \in A$, since otherwise it would be an upper bound on A. But this implies that $f(x) > 0$ for all $x \in \left[a, c + \frac{\delta}{2}\right]$, which implies $c + \frac{\delta}{2} \in A$, which implies c is not an upper bound of A.

We can similarly show that $f(c) < 0$ implies a contradiction. ■

Theorem (intermediate value theorem): If f is continuous on $[a, b]$ and α is any number between $f(a)$ and $f(b)$, then there exists $c \in (a, b)$ such that $f(c) = \alpha$.

Proof: If $f(a) = f(b)$, then there's nothing to show, so suppose without loss of generality that $f(a) < \alpha < f(b)$. Now let $g(x) = f(x) - \alpha$. Clearly g is continuous on $[a, b]$, and note that $g(a) = f(a) - \alpha < 0$ and $g(b) = f(b) - \alpha > 0$. Thus by the previous proposition, there exists some $c \in (a, b)$ such that $g(c) = f(c) - \alpha = 0 \Rightarrow f(c) = \alpha$, as desired. ■

4.6. Uniform Continuity

Definition (uniform continuity): Let $f : A \to \mathbb{R}$. f is *uniformly continuous* if for all $\varepsilon > 0$ there exists some $\delta > 0$ such that for all $x, y \in A$, $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$.

Proposition (sequential formulation): A function $f : X \to \mathbb{R}$ is uniformly continuous if and only if for every pair of sequences $(x_n), (y_n) \in X$ such that if $\lim_{n\to\infty}(x_n - y_n) = 0$, then $\lim_{n\to\infty}(f(x_n) - f(y_n)) = 0$.

Proof: First suppose f is uniformly continuous. Let $\varepsilon > 0$. Then there exists δ such that $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. Consider sequences $(x_n), (y_n)$ such that $\lim_{n\to\infty}(x_n - y_n) = 0$. This implies that there exists some N such that for all $n \geq N$, we have $|x_n - y_n| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. Thus $\lim_{n\to\infty}(f(x_n) - f(y_n)) = 0$.

Now suppose f is not uniformly continuous. Then there exists $\varepsilon > 0$ such that for all $\delta > 0$, there exists $x, y \in X$ such that $|x - y| < \delta \Rightarrow |f(x) - f(y)| \geq \varepsilon$. Thus, for $\delta_n = \frac{1}{n}$, there exists $x_n, y_n \in A$ such that $|x_n - y_n| < \delta_n$, which implies $|f(x_n) - f(y_n)| \geq \varepsilon$. Note that $x_n - y_n < \frac{1}{n}$ converges to 0 via the squeeze theorem. However, $|f(x_n) - f(y_n)| \geq \varepsilon$ for all n, which implies $\lim_{n\to\infty}(f(x_n) - f(y_n)) \neq 0$, as desired. ∎

Proposition: If $f : A \to \mathbb{R}$ is continuous and A is compact, then f is uniformly continuous on A.

Proof: Let $\varepsilon > 0$. For each $c \in A$, let $\delta_c > 0$ be the number such that $|x - c| < \delta_c \Rightarrow |f(x) - f(c)| < \frac{\varepsilon}{2}$. Note that $\left\{\left(c - \frac{\delta_c}{2}, c + \frac{\delta_c}{2}\right)\right\}$ over all c forms an open cover of A. Since A is compact, there exists a finite subcover of these open sets,

$$\left\{\left(c_1 - \frac{\delta_{c_1}}{2}, c_1 + \frac{\delta_{c_1}}{2}\right), \dots, \left(c_n - \frac{\delta_{c_n}}{2}, c_n + \frac{\delta_{c_n}}{2}\right)\right\}.$$

Let δ_{c_k} be the minimum over all δ_{c_i}.

Suppose $x, y \in A$ such that $|x - y| < \frac{\delta_{c_k}}{2}$. We have $x \in \left(c_i - \frac{\delta_{c_i}}{2}, c_i + \frac{\delta_{c_i}}{2}\right)$ for some c_i (since the intervals are a finite subcover). Then by the triangle inequality, we have

$$|y - c_i| \leq |y - x| + |x - c_i| < \frac{\delta_{c_k}}{2} + \frac{\delta_{c_i}}{2} \leq \delta_{c_i}.$$

Thus we have $|x - c_i| < \delta_{c_i}$ and $|y - c_i| < \delta_{c_i}$. This implies that $|f(x) - f(c_i)| < \frac{\varepsilon}{2}$ and $|f(y) - f(c_i)| < \frac{\varepsilon}{2}$. Then by the triangle inequality we have

$$|f(x) - f(y)| \leq |f(x) - f(c_i)| + |f(c_i) - f(y)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus f is uniformly continuous. ∎

4.7. Interesting Problems

Problem: Let $f : \mathbb{R} \to \mathbb{R}$. Prove that $\lim_{x\to 0^+} f\left(\frac{1}{x}\right) = \lim_{x\to\infty} f(x)$ if one of converges to L.

Solution: We show that if $\lim_{x\to\infty} f(x) = L$, then we also have $\lim_{x\to 0^+} f\left(\frac{1}{x}\right) = L$. Going the other way is similar.

The hypothesis implies that for all $\varepsilon > 0$, there exists N_ε such that $x > N_\varepsilon \Rightarrow |f(x) - L| < \varepsilon$. Let $\delta = \frac{1}{N_\varepsilon}$. Note that $0 < x < \delta = \frac{1}{N_\varepsilon} \Rightarrow \frac{1}{x} > N_\varepsilon \Rightarrow \left|f\left(\frac{1}{x}\right) - L\right| < \varepsilon$. This implies that $\lim_{x\to 0^+} f\left(\frac{1}{x}\right) = L$, as desired.

Problem: Let $f : [0, 1] \to \mathbb{R}$ be continuous with $f(0) = f(1)$. Show that there exist $x, y \in [0, 1]$ which are a distance $\frac{1}{2}$ apart for which $f(x) = f(y)$.

Solution: Define $g : \left[0, \frac{1}{2}\right] \to \mathbb{R}$ as $g(x) = f\left(x + \frac{1}{2}\right) - f(x)$. We need to prove that $g(c) = 0$ for some $c \in \left[0, \frac{1}{2}\right]$. Clearly g is continuous. Note that $g(0) = f\left(\frac{1}{2}\right) - f(0)$ and $g\left(\frac{1}{2}\right) = f(1) - f\left(\frac{1}{2}\right)$. Adding the equations yields $g(0) + g\left(\frac{1}{2}\right) = f(1) - f(0) = 0 \Rightarrow g(0) = -g\left(\frac{1}{2}\right)$. If $g(0) = 0$, then we're done. Otherwise, $g(0)$ and $g\left(\frac{1}{2}\right)$ have different signs, and by the IVT, $f(c) = 0$ for some $c \in \left[0, \frac{1}{2}\right]$.

Problem: Let S be a dense subset of $\mathbb{R}$, and assume that f and g are continuous functions on $\mathbb{R}$. Prove that if $f(x) = g(x)$ for all $x \in S$, then $f(x) = g(x)$ for all $x \in \mathbb{R}$.

Solution: Consider $x_0 \notin S$. Since S is dense in $\mathbb{R}$, there exists $a_n \in S$ such that $x_0 - \frac{1}{n} < a_n < x_0$. Thus $(a_n) \to x_0$, and note that $f(a_n) = g(a_n)$ for all n. Thus the limits of these functions are the same, and since both are continuous, we have $f(x_0) = g(x_0)$, as desired.

Remark: This shows that if a solution to the Cauchy functional is given to be continuous, it must be linear, since on $\mathbb{Q}$ the function must be linear.

Problem: Prove that if $f : \mathbb{R} \to \mathbb{R}$ is continuous and A is connected, then $f(A)$ is connected.

Solution: We prove the contrapositive. Suppose $f(A)$ is not connected. Thus there exist open sets U, V such that $U \cap V = \emptyset$, they both intersect $f(A)$, and $(U \cap f(A)) \cup (V \cap f(A)) = A$.

Now consider $U' = f^{-1}(U)$ and $V' = f^{-1}(V)$. Note that both are open (since f is continuous), and that $U' \cap V' = \emptyset$, since otherwise this would imply $U \cap V \neq \emptyset$. Now suppose $y_0 \in U \cap f(A)$. Then $f(x_0) = y_0$ for some $x_0 \in A$. Note that x_0 will also be in U'. This $U' \cap A \neq 0$, and similarly, $V' \cap A \neq 0$.

Now we show $(U' \cap A) \cup (V' \cap A) = A$. Suppose $x_0 \in A$. Then $f(x_0) \in f(A)$, which implies $f(x_0)$ is in either U or V, WLOG U. Then $x_0 \in U'$, which implies $x_0 \in U' \cap A \Rightarrow x_0 \in (U' \cap A) \cup (V' \cap A)$. Thus $A \subseteq (U' \cap A) \cup (V' \cap A)$.

Now suppose $x_0 \in (U' \cap A) \cup (V' \cap A)$. WLOG x_0 comes from the first term (the two terms are disjoin by $U' \cap V' = \emptyset$). Then clearly $x_0 \in A$, which implies $(U' \cap A) \cup (V' \cap A) \subseteq A$, so we're done.

Problem: Suppose $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are uniformly continuous.
a) Prove $f + g$ is uniformly continuous.
b) If f and g are bounded, prove that fg is uniformly continuous.
c) Prove that $g \circ f$ is uniformly continuous.

Solution:

a) Let $\varepsilon > 0$. Then there exists δ_1, δ_2 such that $|x - y| < \delta_1 \Rightarrow |f(x) - f(y)| < \frac{\varepsilon}{2}$ and $|x - y| < \delta_2 \Rightarrow |g(x) - g(y)| < \frac{\varepsilon}{2}$. Let $\delta = \min\{\delta_1, \delta_2\}$. If $|x - y| < \delta$, then we have

$$|x - y| < \delta \Rightarrow |f(x) + g(x) - f(y) - g(y)| \leq |f(x) - f(y)| + |g(x) - g(y)| < \varepsilon.$$

Thus $f + g$ is uniformly continuous.

b) Let $\varepsilon > 0$. Let $M = \max\{M_1, M_2\}$, where M_1 bounds f and M_2 bounds g. There exists δ_1, δ_2 such that $|x - y| < \delta_1 \Rightarrow |f(x) - f(y)| < \frac{\varepsilon}{2M}$ and $|x - y| < \delta_2 \Rightarrow |g(x) - g(y)| < \frac{\varepsilon}{2M}$. Let $\delta = \min\{\delta_1, \delta_2\}$. Then we have

$$\begin{aligned} |x - y| < \delta \Rightarrow & |f(x)g(x) - f(y)g(y)| \\ & \leq |f(x)g(x) - f(x)g(y)| + |f(x)g(y) - f(y)g(y)| \\ & \leq M(|f(x) - f(y)| + |g(x) - g(y)|) < \varepsilon. \end{aligned}$$

Thus fg is uniformly continuous.

c) Let $\varepsilon > 0$. Then there exists δ such that $|x - y| < \delta \Rightarrow |g(x) - g(y)| < \varepsilon$. There also exists δ' such that $|x - y| < \delta' \Rightarrow |f(x) - f(y)| < \delta \Rightarrow |g(f(x)) - g(f(y))| < \varepsilon$. This $g \circ f$ is uniformly continuous.

5. Differentiation

Definition (derivative): Let A be an open set (this will often be an interval), $f : A \to \mathbb{R}$, and $c \in A$. We say f is *differnetiable* at c is

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

exists. If C is the collection of points at which f is differentiable, then the *derivative* of f is a function $f' : C \to \mathbb{R}$ where

$$f^c = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}.$$

Remark: This definition is equivalent to

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

Remark: I won't be super picky about the kind of set that functions are defined on in this chapter for basic derivative results. I'll just declare that they're differentiable at some point, or take for granted that sequences exist that converge to limit points, since most of the time the set that functions are defined on in practice are intervals.

Proposition: Suppose $f : A \to \mathbb{R}$ is differentiable ar $c \in A$. Then f is continuous at c.

Proof: We have

$$\begin{aligned} \lim_{x \to c}[f(x) - f(c)] &= \lim_{x \to x}\left[\frac{f(x) - f(c)}{x - c} \cdot (x - c)\right] \\ &= \left(\lim_{x \to c} \frac{f(x) - f(c)}{x - c}\right)\left(\lim_{x \to c} x - c\right) \\ &= f'(c) \cdot 0 = 0 \\ &\Rightarrow \lim_{x \to c} f(x) = f(c). \end{aligned}$$

■

Proposition (derivative rules): Let $f, g : A \to \mathbb{R}$ be differentiable at $c \in A$. Then we have the following:

a) $(f + g)'(c) = f'(c) + g'(c)$
b) $(kf)'(c) = kf'(c)$
c) $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$
d) $\left(\frac{f}{g}\right)'(c) = \frac{f'(c)g(c)-f(c)g'(c)}{g(c)^2}$

Proof:

a)

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

b)

$$\lim_{x\to c} \frac{(kf)(x) - (kf)(c)}{x-c} = k \cdot \lim_{x\to c} \frac{f(x)-f(c)}{x-c} = k \cdot f'(c)$$

c)

$$\begin{aligned}\lim_{x\to c} \frac{f(x)g(x) - f(c)g(c)}{x-c} &= \lim_{x\to c} \frac{f(x)g(x) - f(x)g(c) + f(x)g(c) - f(c)g(c)}{x-c} \\ &= \lim_{x\to c} f(x) \cdot \frac{g(x)-g(c)}{x-c} + \lim_{x\to c} g(c) \cdot \frac{f(x)-f(c)}{x-c} \\ &= f(c)g'(c) + f'(c)g(c)\end{aligned}$$

d) The quotient rule follows much easier using the chain rule and product rule, which we prove next. ∎

Proposition (chain rule): Let $g : A \to B$ and $f : B \to \mathbb{R}$. If g is differentiable at $c \in A$ and f is differentiable at $g(c) \in B$, then

$$(f \circ g)'(c) = f'(g(c))g'(c).$$

Proof: Consider the following function:

$$Q(y) = \begin{cases} \frac{f(y)-f(g(c))}{y-g(c)} \text{ if } y \neq g(c) \\ f'(g(c)) \text{ if } y = g(c) \end{cases}$$

This function takes place of the difference quotient in the limit and ensure that the quotient doesn't have divide by 0 problems (that's what the second case is for).

Note that Q is continuous at $g(c)$ since f is differentiable at $g(c)$ (and approaching Q from above or below $g(c)$ will alawys result in case 1).

Next we show that

$$\frac{f(g(x)) - f(g(c))}{x - c} = Q(g(x)) \cdot \frac{g(x) - g(c)}{x - c}.$$

If $g(x) \neq g(c)$, then we plug in the quotient in case 1. Then the denominator of $Q(g(x))$ cancels with $g(x) - g(c)$ and we're done. If $g(x) = g(c)$, then we want to show that

$$\frac{f(g(x)) - f(g(c))}{x - c} = f'(g(c)) \cdot \frac{g(x) - g(c)}{x - c}.$$

Then applying $g(x) = g(c)$ yields 0 on both sides.

Now we have

$$\begin{aligned}(f \circ g)'(c) &= \lim_{x \to c} \frac{f(g(x)) - f(g(c))}{x - c} \\ &= \lim_{x \to c} Q(g(x)) \cdot \frac{g(x) - g(c)}{x - c} \\ &= f'(g(c))g'(c).\end{aligned}$$

■

5.1. Min and max

Definition (local min/max): Let $f : A \to \mathbb{R}$. Then f has a *local maximum* at $c \in A$ if there exists some $\delta > 0$ such that for all $x \in A$ for which $|x - c| < \delta$, we have

$$f(x) \leq f(c).$$

Similarly, f has a *local minimum* at $c \in A$ if there exists some $\delta > 0$ such that for all $x \in A$ for which $|x - c| < \delta$, we have

$$f(x) \geq f(c).$$

Proposition: Let A be an open set and suppose $f : A \to \mathbb{R}$ is differentiable at $c \in A$. If f has a local max or min at c, then $f'(c) = 0$.

Proof: WLOG the c is a local max, and suppose f on $V_\delta(c)$ is at most $f(c)$. Then pick a sequence (ℓ_n) with $c - \delta < \ell_n < c$ that converges to c and a sequence (r_n) with $c < r_n < c + \delta$ that converges to c. Then we have

$$\frac{f(\ell_n) - f(c)}{\ell_n - x} \geq 0 \quad \text{and} \quad \frac{f(r_n) - f(c)}{r_n - c} \leq 0$$

for all n. Since the sequences converge to c, and f is continuous at c (since it's differentiable at c), both quotients converge to $f'(c)$. Note however that the inequalities on the quotients imply that $f'(c) \geq 0$ and $f'(c) \leq 0$. Thus, $f'(c) = 0$. ■

Theorem (Darboux's theorem): Suppose $f : [a, b] \to \mathbb{R}$ is differentiable. If α is between $f'(a)$ and $f'(b)$, then there exists $c \in (a, b)$ where $f'(c) = \alpha$.

Proof: WLOG $f'(b) < \alpha < f'(a)$. Let $g(x) = f(x) - \alpha x$. Then g is differentiable on $[a, b]$ with $g'(x) = f'(x) - \alpha$. Note also that $g'(a) = f'(\alpha) - \alpha > 0$ and $g'(b) = f'(b) - \alpha < 0$. Since $[a, b]$ is compact, by the extreme value theorem, g attains a maximum on $[a, b]$. We need to show that the max does not occur at a or b.

Suppose the max occurred at a. Then $\frac{g(x)-g(a)}{x-a} \leq 0$ for all $x \in (a, b]$. Thus $g'(a) \leq 0$, but this is a contradiction. We can do basically the same thing for b.

Thus g attains its max at $c \in (a, b)$. It's clear the max is also a local max, so by the previous proposition, we have that $g'(c) = 0 \Rightarrow f'(c) = \alpha$. ■

Remark: This is really, really insane. Essentially what this means is that the derivative of a differentiable function won't have jump or removable discontinuities or asymptotes, but instead will oscillate infinitely into a point of discontinuity i.e. $x^2 \sin\left(\frac{1}{x}\right)$ at 0, where the function at 0 is defined to be 0. Despite being differentiable at 0 with derivative 0, the derivative is not continous at 0.

5.2. The Mean Value Theorem

Theorem (Rolle's theorem): Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there exists $c \in (a, b)$ where $f'(c) = 0$.

Proof: By the extreme value theorem, f hits a max at $c_1 \in [a, b]$ and a min at $c_2 \in [a, b]$. If either of these are in (a, b), then we're done by the local min/max proposition. Otherwise, c_1 and c_2 are endpoints. WLOG $c_1 = f(a)$ and $c_2 = f(b)$. Then $f(a) \geq f(x) \geq f(b)$ for all $x \in [a, b]$. However, since $f(a) = f(b)$, this implies $f(x)$ is constant, and thus $f'(x) = 0$. ■

Theorem (mean value theorem): Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists some $c \in (a, b)$ where

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

Proof: Let

$$L(x) = \left(\frac{f(b) - f(a)}{b - a}\right)(x - a) + f(a),$$

and define $g(x) = f(x) - L(x)$. Then g is continuous on $[a, b]$ and differentiable on (a, b). Note that $g(a) = g(b)$, so by Rolle's theorem, we have $g'(c) = 0$ for some $c \in (a, b)$. Thus

$$g'(x) = f'(x) - L'(x) = f'(x) - \left(\frac{f(b) - f(a)}{b - a}\right) \Rightarrow 0 = f'(c) - \left(\frac{f(b) - f(a)}{b - a}\right).$$

■

Corollary: Let I be an interval and $f : I \to \mathbb{R}$ be differentiable. If $f'(x) = 0$ for all $x \in I$, then f is constant on I.

Proof: Pick $x, y \in I$ with $x < y$. Since f is differentiable on I, it's also differentiable on $[x, y]$. Thus, by the mean value theorem we have

$$f'(c) = \frac{f(x) - f(y)}{x - y}$$

for some $c \in (x, y)$. By assumption, $f'(c) = 0$, so we have

$$0 = \frac{f(x) - f(y)}{x - y} \Rightarrow f(x) = f(y).$$

■

Corollary: Let I be an interval and $f, g : I \to \mathbb{R}$ be differentiable. If $f'(x) = g'(x)$ for all $x \in I$, then $f(x) = g(x) + C$ for some C.

Proof: Let $h(x) = f(x) - g(x)$. Then $h'(x) = f'(x) - g'(x) = 0$, and so by the previous corollary, we have that h is constant on I, which gives the desired result. ■

Corollary: Let I be an interval and $f : I \to \mathbb{R}$ be differentiable.
a) f is monotone increasing if and only if $f'(x) \geq 0$ for all $x \in I$.
b) f is monotone decreasing if and only if $f'(x) \leq 0$ for all $x \in I$.

Proof: We only prove a), since b) is extremely similar.

First supose f is monotone increasing on I. Then for any $x, c \in I$, we have

$$\frac{f(x) - f(c)}{x - c} \geq 0.$$

Thus the limit of the left side as x approaches c, which is $f'(c)$, which be nonnegative. This holds for all $c \in I$.

Now suppose $f'(x) \geq 0$ for all $x \in I$. Pick $x, y \in I$ with $x < y$. By the mean value theorem, we have

$$\frac{f(y) - f(x)}{y - x} = f'(c) \geq 0$$

for some $c \in (x, y)$. Since the denominator of the quotient is positive, the numerator must be nonnegative, which implies $f(x) \leq f(y)$, as desired. ■

Theorem (Cauchy mean value theorem): If f and g are continuous on $[a, b]$ and differentiable on (a, b), then there exists $c \in (a, b)$ such that

$$[f(b) - f(a)] \cdot g'(c) = [g(b) - g(a)] \cdot f'(c).$$

Proof: If $g(b) = g(a)$, then by Rolle's there's c such that $g'(c) = 0$, so the equation holds.

If $g(b) \neq g(a)$, then define

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

Clearly this is differentiable and continuous. Note that $h(a) = h(b)$, so by Rolle's there is c such that $h'(c) = 0$. Thus

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

as desired. ■

5.3. L'Hôpital's Rule

Proposition (L'Hôpital's rule): Suppose I is an open interval containing a point a, and $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ are differentiable on I, except possibly at a. Suppose also $g'(x) \neq 0$ on I. Then, if

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

then

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

provided that $\lim_{x \to a} \frac{f'(x)}{g'(x)}$ exists.

Proof:

■

Proposition (L'Hôpital's rule): Suppose I is an open interval containing a point a, and $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ are differentiable on I, except possibly at a. Suppose also $g'(x) \neq 0$ on I. Then, if

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

then

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

provided that $\lim_{x \to a^+} \frac{f'(x)}{g'(x)}$ exists. We can have the limits approach from the left

Proof: First a lemma.

Lemma: Suppose $\lim_{x \to c} f(x) = \infty$ and $\lim_{x \to c} g(x) = \infty$, and suppose $r, s \in \mathbb{R}$. If $\frac{f(x)-r}{g(x)-s}$ is bounded on some open interval containing c, then

$$\lim_{x \to c} \left[\frac{f(x) - r}{g(x) - s} - \frac{f(x)}{g(x)} \right] = 0.$$

Proof of Lemma:

We can rewrite the inside of the limit as

$$\frac{1}{g(x)} \left(r - s \cdot \frac{f(x) - r}{g(x) - s} \right).$$

Pick $\varepsilon > 0$. Suppose $\frac{f(x)-r}{g(x)-s}$ is bounded on (a, b) with $c \in (a, b)$, and is bounded by M on (a, b). Note that $\lim_{x \to c} g(x) = \infty \Rightarrow \lim_{x \to c} \frac{1}{g(x)} = 0$. Pick δ such that $V_\delta(c) \subseteq (a, b)$ and $\left| \frac{1}{g(x)} \right| < \frac{\varepsilon}{|r - s \cdot M|}$ for all $x \in V_\delta(c)$ (which we can do because the limit approaches 0). Then we have

$$\left| \frac{1}{g(x)} \left(r - s \cdot \frac{f(x) - r}{g(x) - s} \right) \right| < \left| \frac{\varepsilon}{|r - s \cdot M|} (r - s \cdot M) \right| = \varepsilon.$$

for all $0 < |x - c| < \delta$. This works for any $\varepsilon > 0$, so the limit is indeed 0.

We prove the case when the limits approach from the right, since limits approaching from the left is analogous.

Let $\lim_{x \to a^+} \frac{f'(x)}{g'(x)} = L$, and pick $\varepsilon > 0$. By assumption, there exists δ_1 such that

$$a < x < a + \delta_1 \Rightarrow L - \frac{\varepsilon}{2} < \frac{f'(x)}{g'(x)} < L + \frac{\varepsilon}{2}.$$

Now pick $a < x_1 < x_2 < a + \delta_1$. Note that f, g are continous on $[x_1, x_2]$ and differentiable on (x_1, x_2). Thus by Cauchy's mean value theorem, we have

$$\frac{f(x_2) - f(x_1)}{g(x_2) - g(x_1)} = \frac{f'(c)}{g'(c)}$$

for some $c \in (x_1, x_2)$ (note that $g(x_2) - g(x_1)$ can't be 0, since otherwise the regular mean value theorem would imply that $g'(x) = 0$ for some x). Thus, for any $x_1, x_2 \in (a, a + \delta_1)$, we have

$$L - \frac{\varepsilon}{2} < \frac{f(x_2) - f(x_1)}{g(x_2) - g(x_1)} < L + \frac{\varepsilon}{2}.$$

By the lemma, there exists δ_2 such that for all $a < x_2 < a + \delta_2$, we have

$$\frac{f(x_2) - f(x_1)}{g(x_2) - g(x_1)} - \frac{\varepsilon}{2} < \frac{f(x_2)}{g(x_2)} < \frac{f(x_2) - f(x_1)}{g(x_2) - g(x_1)} + \frac{\varepsilon}{2}.$$

Pick $\delta = \min\{\delta_1, \delta_2\}$. Then, for all $a < x < a + \delta$, we have

$$L - \varepsilon < \frac{f(x)}{g(x)} < L + \varepsilon.$$

ε was arbitrary, so we do indeed have $\lim_{x \to a^+} \frac{f(x)}{g(x)} = L$. ■

5.4. Interesting Problems

Problem: Suppose $f : I \to \mathbb{R}$ is differentiable on an interval I. Prove that if f' is bounded, then f is uniformly continuous.

Solution: Consider the difference quotient $\frac{f(x)-f(y)}{x-y}$. Since f is differentiable on I, it's continuous on I, so we can apply the mean value theorem. Thus, for any $x, y \in I$, there exists $c \in I$ such that $\frac{f(x)-f(y)}{x-y} = f'(c)$. Since f' is bounded, the difference quotient must also be bounded, which means

$$\left| \frac{f(x) - f(y)}{x - y} \right| \leq M$$

for some M.

Pick some $\varepsilon > 0$, and let $\delta = \frac{\varepsilon}{M}$. The bound on the difference quotient implies

$$|f(x) - f(y)| \leq M|x - y| < M \cdot \delta = \varepsilon,$$

which implies f is uniformly continuous.

Remark: This solution also shows that f is Lipschitz.

Remark: The converse is not true. Consider $x \sin\left(\frac{1}{x}\right)$ on $[-1, 1]$ with it being defined to be 0 at $x = 0$. The function is continous on $[-1, 1]$ and so is uniformly continous on $[-1, 1]$. However, its derivative is $\sin\left(\frac{1}{x}\right) - \frac{1}{x} \cos\left(\frac{1}{x}\right)$ is unbounded.

Problem: Let I be an interval and $f : I \to \mathbb{R}$ be differentiable. Show that f is Lipschitz on I if and only if f' is bounded on I.

Solution: Suppose f is Lipschitz with constant M. Then

$$\left|\frac{f(x) - f(y)}{x - y}\right| \leq M$$

for all $x, y \in I$. Fix $y = c$ and consider the limit as $x \to c$ of the difference quotient. Clearly it exists since f is differentiable, and since every possible value of the difference quotient is bounded, the derivative at c must be bounded as well. This works for all $c \in I$, so f' is bounded on I.

Now suppose f' is bounded on I. Then, by the mean value theorem, for every $x, y \in I$, we have

$$\left|\frac{f(x) - f(y)}{x - y}\right| = |f'(c)| \leq M,$$

where M is the bound on f'. Thus f is Lipschitz with constant M.

Remark: A nice consequence of this is that if f' is continous on a closed interval, then by the extreme value theorem it's bounded, so f is also Lipschitz.

Problem: Suppose f and g are differentiable functions with $f(a) = g(a)$ and $f'(x) < g'(x)$ for all $x > a$. Prove that $f(b) < g(b)$ for any $b > a$.

Solution: Consider $h = g - f$. Then $h' = g' - f' > 0$ and $h(a) = 0$. Then by the mean value theorem, for any $b > a$, there exists $c \in (a, b)$ such that

$$\frac{h(b) - h(a)}{b - a} = h'(c) > 0 \Rightarrow h(b) > 0 \Rightarrow g(b) > f(b),$$

as desired.

Problem: Assume that $f(0) = 0$ and $f'(x)$ is increasing. Prove that $g(x) = \frac{f(x)}{x}$ is an increasing function on $(0, \infty)$.

Solution: Note that

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

Thus we just need to prove $f'(x)x > f(x) \Rightarrow f'(x) > \frac{f(x)}{x}$. Note that the right side is the mean value theorem on $[0, x]$. Thus $\frac{f(x)}{x} = f'(c)$ for some $c < x$, which means $f'(x) > f'(c)$, which is true. Thus g' is greater than 0, which means $\frac{f(x)}{x}$ is increasing, as desired.

6. Integration

6.1. Darboux Integral

Definition (partition): A *partition* of $[a, b]$ is a finite set

$$P = \{x_0, x_1, ..., x_n\}$$

such that $a = x_0$, $b = x_n$, and $x_0 < x_1 < \cdots < x_n$.

We also denote for a subinterval $[x_i, x_{i+1}]$ that

- $m_i = \inf\{f(x) : x \in [x_i, x_{i+1}]\}$
- $M_i = \sup\{f(x) : x \in [x_i, x_{i+1}]\}$

Definition (upper/lower sums): Consider a function $f : [a, b] \to \mathbb{R}$, and consider a partition $P = \{x_0, x_2, ..., x_n\}$ of $[a, b]$. Define the *upper sum* as

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

and the *lower sum* as

$$L(f, P) = \sum_{i=1}^{n} m_i (x_i - x_{i+1}).$$

Definition (refinement): Consider a partition of P of $[a, b]$. A partition Q of $[a, b]$ is called a *refinement* of P if $P \subseteq Q$.

Proposition: Consider a function $f : [a, b] \to \mathbb{R}$ and a partition $P = \{x_0, ..., x_n\}$ of $[a, b]$. If Q is a refinment of P, then

$$L(f, P) \leq L(f, Q) \quad \text{and} \quad U(f, P) \geq U(f, Q).$$

Proof: We prove the lower sum case, as the upper sum case is similar. We have that

$$L(f, P) = m_1(x_1 - x_0) + \cdots + m_n(x_n - x_{n-1}).$$

Since Q is a refinment of P, there are $x_{\frac{1}{n'}}, x_{\frac{2}{n'}}, ..., x_{\frac{n'-1}{n'}} \in Q$ such that

$$x_0 < x_{\frac{1}{n'}} < \cdots < x_{\frac{n'-1}{n'}} < x_1.$$

It could happen that there are no elements between x_0 and x_1, but if that's the case, then the contribution of the interval $[x_0, x_1]$ into the lower sum is the same for both P and Q, so it doesn't change the inequality.

Note that every element in $\left[x_{\frac{i}{n'}}, x_{\frac{i+1}{n'}}\right]$ is by definition at least m_1, which implies $m_{\frac{i}{n'}} \geq m_1$. Thus we have

$$\sum_{i=1}^{n'} m_{\frac{i}{n'}}\left(x_{\frac{i}{n'}} - x_{\frac{i-1}{n'}}\right) \geq \sum_{i=1}^{n'} m_1\left(x_{\frac{i}{n'}} - x_{\frac{i-1}{n'}}\right) = m_1(x_1 - x_0).$$

This holds for all of the terms in $L(f, P)$, which implies $L(f, P) \leq L(f, Q)$, as desired. ■

Proposition: Let $f : [a, b] \to \mathbb{R}$. If P_1 and P_2 are any partitions of $[a, b]$, then

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

Proof: First note that for any partition P, we have

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

Now let $Q = P_1 \cup P_2$, which is clearly a refinement of both of them. Thus, by the previous proposition we have

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

as desired. ■

6.2. Integrability

Definition (upper/lower integral): Let $f : [a, b] \to \mathbb{R}$ be a bounded function and let $\mathcal{P}$ denote the set of all partitions of $[a, b]$. The *upper integral* of f is defined to be

$$U(f) = \inf\{U(f, P) : P \in \mathcal{P}\},$$

and the *lower integral* of f is defined to be

$$L(f) = \sup\{L(f, P) : P \in \mathcal{P}\}.$$

Lemma (intgral bound): Let $f : [a, b] \to \mathbb{R}$ be a bounded function with $m \leq f(x) \leq M$ for all $x \in [a, b]$. Then

$$m(b - a) \leq L(f) \leq U(f) \leq M(b - a).$$

Proof: The middle inequality follows from the last proposition. Let $P_0 = \{a, b\}$ be a partition of $[a, b]$. Then

$$\begin{aligned} L(f) &= \sup\{L(f,P) : P \in \mathcal{P}\} \\ &\geq L(f, P_0) \\ &\geq m(b-a). \end{aligned}$$

Note that we assume m is the infinum of f over $[a,b]$, and if it wasn't, then we just have one more inequality in the chain. The upper inequality holds similarly. ■

Definition (integrable): A bounded function $f : [a,b] \to \mathbb{R}$ is *integrable* if $L(f) = U(f)$, and we write

$$\int_a^b f(x)\,\mathrm{d}x = L(f) = U(f).$$

Example: Let $f : [0,1] \to \mathbb{R}$ be the Dirichlet function

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

Then f is not integrable. Let P be any partition of $[0,1]$. Note that every subinterval will contain a rational and irrational, since both sets are dense in $\mathbb{R}$. Thus $L(f,P) = 0$ and $U(f,P) = 1$, regardless of what P is. Thus $L(f) \neq U(f)$, and so f is not integrable.

Proposition: Assume that a bounded function $f : [a,b] \to \mathbb{R}$ is integrable and nonnegative on $[a,b]$. Then $\int_a^b f(x)\,\mathrm{d}x \geq 0$.

Proof: By the integral bound, we have $0 \cdot (b-a) \leq L(f) = \int_a^b f(x)\,\mathrm{d}x$, where the equality comes from f being integrable. ■

Proposition: Let $f : [a,b] \to \mathbb{R}$ be bounded. Then f is integrable if and only if, for all $\varepsilon > 0$ there exists a partition P_ε of $[a,b]$ where

$$U(f, P_\varepsilon) - L(f, P_\varepsilon) < \varepsilon.$$

Remark: This is easily motivated by looking at the definitions of integrability. To be integrable, we require

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

which implies that the elements of each set get arbitrarily close.

Proof: First suppose the condition holds for all $\varepsilon > 0$. We have $L(f,P_\varepsilon) \leq L(f)$ and $U(f) \leq U(f, P_\varepsilon)$. Thus

$$|U(f) - L(f)| \leq U(f) - L(f) \leq U(f, P_\varepsilon) - L(f, P_\varepsilon) < \varepsilon.$$

Since this holds for all $\varepsilon > 0$, we have $U(f) - L(f) = 0$, and so f is integrable.

Now suppose f is integrable, which means $U(f) = L(f) = I$. Let P_1 be a partition such that $I - \frac{\varepsilon}{2} < L(f, P_1)$, which exists since I is a supremum. Similarly, there exists P_2 such that $U(f, P_2) < I + \frac{\varepsilon}{2}$. Let $P_\varepsilon = P_1 \cup P_2$ be a refinement. We have

$$L(f, P_\varepsilon) \geq L(f, P_1) > I - \frac{\varepsilon}{2} \quad \text{and} \quad U(f, P_\varepsilon) \leq U(f, P_2) < I + \frac{\varepsilon}{2}.$$

Subtracting the first inequality from the second yields

$$U(f, P_\varepsilon) - L(f, P_\varepsilon) < \varepsilon,$$

as desired. ∎

Corollary: If $f : [a, b] \to \mathbb{R}$ is integrable, then there exists a sequence of partitions (P_n) of $[a, b]$ such that

$$\lim_{n \to \infty} [U(f, P_n) - L(f, P_n)] = 0.$$

Proof: Let P_n be a partition such that $U(f, P_n) - L(f, P_n) < \frac{1}{n}$, which exists by the previous proposition. Since $U(f, P_n) \geq L(f, P_n)$, the sequence is bounded below by 0, and so the squeeze theorem implies the sequence converges to 0. ∎

Remark: This section contains way more examples to order to motivate and help me understand the big theorem at the end.

Proposition: If $f : [a, b] \to \mathbb{R}$ is continuous, then f is integrable.

Remark: Intuitively we can expect this to hold, since on an arbitrarily small subinterval, we can make $M_i - m_i$ arbitrarily small, and then previous results will give us the desired conclusion.

Proof: Since $[a, b]$ is compact, f is bounded, so we can quote integral results. Compactness also gives uniform continuity.

Pick $\varepsilon > 0$. By unform continuity, there exists δ such that $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \frac{\varepsilon}{b-a}$. Pick n such that $\frac{b-a}{n} < \delta$, and let $x_i = a + i \cdot \frac{b-a}{n}$. Then $P_\varepsilon = \{x_0, ..., x_n\}$ is a partition of $[a, b]$.

Note that on the subinterval $[x_i, x_{i+1}]$, f achieves a min and max by the extreme value theorem, m_i and M_i. Then since $x_{i+1} - x_i = \frac{b-a}{n} < \delta$, the range of f on the subinterval is contained within an interval of length $\frac{\varepsilon}{b-a}$. Thus, $|M_i - m_i| < \frac{\varepsilon}{b-a}$. This holds for any subinterval.

Now we have

$$\begin{aligned}U(f,P_\varepsilon)-L(f,P_\varepsilon)&=\sum_{i=1}^{n}(M_i-m_i)(x_i-x_{i-1})\\&<\sum_{i=1}^{n}\frac{\varepsilon}{b-a}(x_i-x_{i-1})\\&=\frac{\varepsilon}{b-a}\cdot(b-a)=\varepsilon.\end{aligned}$$

This holds for all $\varepsilon>0$, so f is indeed integrable. ■

Lemma: Let $f:[a,b]\to\mathbb{R}$ and $a<c<b$. Then f is integrable on $[a,b]$ if and only if f is integrable on both $[a,c]$ and on $[c,b]$.

Proof: First assume f is integrable on $[a,c]$ and $[c,b]$. Then there exists P_ε^1 and P_ε^2 such that

$$U(f,P_\varepsilon^1)-L(f,P_\varepsilon^1)<\frac{\varepsilon}{2}\quad\text{and}\quad U(f,P_\varepsilon^2)-L(f,P_\varepsilon^2)<\frac{\varepsilon}{2}.$$

Let $P_\varepsilon=P_\varepsilon^1\cup P_\varepsilon^2$, and note that it's a partition of $[a,b]$. Note that since the partitions are disjoint except for c, we have $U(f,P_\varepsilon)=U(f,P_\varepsilon^1)+U(f,P_\varepsilon^2)$, and similarly for L. Thus,

$$U(f,P_\varepsilon)-L(f,P_\varepsilon)=U(f,P_\varepsilon^1)-L(f,P_\varepsilon^1)+U(f,P_\varepsilon^2)-L(f,P_\varepsilon^2)<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

Thus f is integrable over $[a,b]$.

Now suppose f is integrable over $[a,b]$. Let P be a partition such that

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

Without loss of generality, $c\in P$, since otherwise we can add it P and both sums get refined, with the difference becoming smaller. Let $P'=P\cap[a,c]$. Then we can write $P=\{x_0,x_1,...,x_T,x_{T+1},...,x_N\}$, where $P'=\{x_0,x_1,...,x_T\}$. Thus

$$\begin{aligned}U(f,P')-L(f,P')&=\sum_{i=1}^{T}(M_i-m_i)(x_i-x_{i-1})\\&\le\sum_{i=1}^{N}(M_i-m_i)(x_i-x_{i-1})\\&=U(f,P)-L(f,P)<\varepsilon.\end{aligned}$$

Thus f is integrable over $[a,c]$, and a similar approach shows that f is integrable over $[c,b]$. ■

6.2.1. Lebesgue's Integrability Criterion

This section focuses on the integrability of functions with discontinuities. We first give a few examples of functions with discontinuities that are integrable, and then dive into weeds of Lebesgue's intergrability criterion.

Example *(one discontinuity)*: Let $f:[0,2]\to\mathbb{R}$, defined by

$$f(x)=\begin{cases}1 & \text{if } x\neq 1\\ 0 & \text{if } x=1.\end{cases}$$

Then f is integrable. Note that the upper sum is always 2, and the lower sum is $2 - \varepsilon$, where ε is the length of the subinterval that contains 1. We can make that subinterval arbitrarily small, so the upper and lower sums get arbitrarily close, meaning the function is integrable.

In general, for any function with one discontinuity, simply make the interval containing the point of discontinuity arbitrarily small.

Example *(finite number of discontinuities)*: We can split the function into separate intervals, each containing a single discontinuity. We know that on these intervals, f is integrable, and by the lemma, f on the overall interval is integrable.

Example *(countable number of discontinuities)*: Since the function's domain is compact, the discontinuities will intuitively cluster around points in the domain. Since partitions must be finite, we can pick a small enough interval around the cluster points that contain countably many of them. Then for the finitely many left discontinuities, we can also pick arbitrarily small intervals.

Example *(discountable number of discontinuities)*: The function that's 1 on the Cantor Set and 0 otherwise is actually integrable. This is especially strange since the Cantor Set is totally disconnected.

Definition (measure zero): A set A has *measure zero* if for all $\varepsilon > 0$ there exists a countable collection $I_1, I_2, I_3, \ldots$ of intervals such that

$$A \subseteq \bigcup_{k=1}^{\infty} I_k \quad \text{and} \quad \sum_{k=1}^{\infty} \mathscr{L}(I_k) < \varepsilon,$$

where $\mathscr{L}(I)$ denotes the length of the interval I.

Proposition: If a countable collection of sets $S_1, S_2, \ldots$ each has measure 0, then the union of the sets has measure 0.

Proof: For S_i, we can find intervals that cover S_i whose length is less than $\frac{\varepsilon}{2^i}$. Since the union of countably many countable sets is countable, and since that sum of the lenghts of the intervals is $\sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon$, the union of the sets does indeed have measure 0. ■

Definition (oscillation on a set): Let f be a function defined on A. The oscillation of f on A is

$$\Omega_f(A) = \sup_{x,y \in A} |f(x) - f(y)|.$$

Remark: Note that if $B \subseteq A$, then $\Omega_f(B) \leq \Omega_f(A)$. This means for the above definition, we can replace the inf with a limit $r \to 0^+$.

Definition (oscillation at a point): Let f be a function on A and $c \in A$. Then the oscillation of f at c is

$$\omega_f(c) = \inf_{r>0} \Omega_f(A \cap (c-r, c+r)).$$

Proposition: Suppose f is defined on A and $c \in A$. Then f is continuous at c if and only if $\omega_f(c) = 0$.

Proof: Suppose f is continuous at c. Then for all $\varepsilon > 0$, there exists $\delta(\varepsilon)$ such that if $x \in A$ and $|x - c| < \delta(\varepsilon)$, then $|f(x) - f(c)| < \frac{\varepsilon}{2}$. Then by the triangle inequality, $|f(x) - f(y)| \leq |f(x) - f(c)| + |f(c) - f(y)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$ for $x, y \in A \cap V_{\delta(\varepsilon)}(c)$. Thus, $0 \leq \Omega_f\left(A \cap V_{\delta(\varepsilon)}(c)\right) \leq \varepsilon$. Then, taking the limit at $\varepsilon \to 0$, by the squeeze theorem we have $\lim_{\varepsilon \to 0} \Omega_f\left(A \cap V_{\delta(\varepsilon)}(c)\right) = 0$, which implies $\omega_f(c) = \inf_{r>0} \Omega_f(A \cap V_r(c)) = 0$.

Now suppose $\omega_f(c) = 0$, and let $\varepsilon > 0$. Then, there exists δ such that $0 \leq \Omega_f(A \cap V_\delta(c)) \leq \frac{\varepsilon}{2}$. Thus $\sup_{x,y \in A \cap V_\delta(c)} |f(x) - f(y)| \leq \frac{\varepsilon}{2} \Rightarrow |f(x) - f(y)| < \varepsilon$ for all $x \in A \cap V_\delta(c)$, which means f is continuous at c. ■

Proposition: Let f be a function with domain $[a, b]$. Then for any $s > 0$, the set

$$A_s = \left\{x \in [a,b] : \omega_f(x) \geq s\right\}$$

is compact.

Proof: Note that clearly A_s is bounded, so we just need to show that it's closed. We do this by showing A_s^c is open relative to $[a, b]$.

Let $x_0 \in A_s^c$. Then $\omega_f(x) = t < s$. This means

$$t = \lim_{r \to 0^+} \sup_{x,y \in V_r(x_0) \cap [a,b]} |f(x) - f(y)|.$$

Then, letting $\varepsilon = \frac{s-t}{2}$, there exists δ such that $0 < r \leq \delta$ implies

$$\left| \sup_{x,y \in V_r(x_0) \cap [a,b]} |f(x) - f(y)| - t \right| < \varepsilon \Rightarrow \sup_{x,y \in V_r(x_0) \cap [a,b]} |f(x) - f(y)| < t + \varepsilon = \frac{t+s}{2}.$$

Pick $y_0 \in V_{\frac{\delta}{2}}(x_0)$. Then $V_{\frac{\delta}{2}}(y_0) \subset V_\delta(x_0)$, and so for all $0 < r' < \frac{\delta}{2}$, we have

$$\begin{aligned} \sup_{x,y \in V_{r'}(y_0) \cap [a,b]} |f(x) - f(y)| &\leq \sup_{x,y \in V_{\frac{\delta}{2}}(y_0) \cap [a,b]} |f(x) - f(y)| \\ &\leq \sup_{x,y \in V_\delta(x_0) \cap [a,b]} |f(x) - f(y)| < \frac{t+s}{2}. \end{aligned}$$

Thus,

$$\lim_{r \to 0^+} \sup_{x,y \in V_r(y_0) \cap [a,b]} |f(x) - f(y)| \leq \frac{t+s}{2} < s.$$

This means $V_{\frac{\delta}{2}}(x) \subset A_s^c$, and so A_s^c is open. ■

Theorem (Lebesgue's Integrability Criterion): A bounded function f on $[a, b]$ is integrable if and only if the set of discontinuities D has measure zero.

Proof: Suppose f is integrable. Let D_k be the set of points such that $\omega_f(x) \geq \frac{1}{2^k}$. Let P_k be a partition $\{x_0, ..., x_n\}$ such that

$$U(f, P_k) - L(f, P_k) < \frac{\varepsilon}{4^k}.$$

Suppose $x \in D_k \cap (x_{j-1}, x_j)$. Then there exists δ such that $V_\delta(x) \subseteq (x_{j-1}, x_j)$. Then we have

$$\frac{1}{2^k} \leq \omega_f(x) \leq \Omega_f(V_\delta(x)) \leq M_k - m_k,$$

where these all follow by definition. Let T be the set of j such that $D_k \cap (x_{j-1}, x_j) \neq \emptyset$. Then we have

$$\frac{1}{2^k} \sum_{j \in T} (x_j - x_{j-1}) \leq \sum_{j=1}^{n} (M_j - m_j)(x_j - x_{j-1}) = U(f, P_k) - L(f, P_k) \leq \frac{\varepsilon}{4^k}.$$

Note that $D_k \subseteq \bigcup_{j \in T} (x_{j-1}, x_j) \cup \bigcup_{j=0}^{n} \{x_j\}$. Note that the length of those intervals totaled is $\sum_{j \in T} (x_j - x_{j-1}) \leq \frac{\varepsilon}{2^k}$. Thus D_k is contained in a union of intervals that can get arbitrarily small, which implies D_k has measure 0. Then the collection of D_k is countable and each one has measure zero, then the union of all of them, which is equal to D, has measure zero.

Now suppose the set of discontinuities D has measure 0. ■