Mathematical Analysis Notes

(Mostly) Adapted from Rudin

Navid Shamszadeh

March 6, 2021

Contents

1 The Real and Complex Number Systems

In order to study the main concepts in analysis, i.e. convergence, continuity, differentiation, integration, etc., an adequate number system must be defined. Of course later we'll see how many of the topics covered generalize to metric spaces and even topological spaces; however, many motivating examples of such topics come from the real or complex number system, so we will define and explore them in this section first.

To begin, we show that the rational number system does not contain all numbers which we are interested in working with. In particular, the classic example of this is that $\sqrt{2} \notin \mathbb{Q}$. First we define what it means to be a rational number:

Definition 1.1. (Rational Number) A number n is said to be *rational* if $n = \frac{p}{q}$ where $p, q \in \mathbb{Z}$ and $q \neq 0$. If this is the case, we call n a *rational number*. The set of all rational numbers is denoted by $\mathbb{Q}$.

Of course this begs the question what does it mean to be an integer, and what exactly is a fraction, and these questions may be addressed in a later update of these notes; but in brief, integers are constructed through an equivalence relation on the natural numbers $\mathbb{N}$, and rationals are constructed in much the same way (see ring of fractions, a topic explored in abstract algebra). There are different ways in which one can construct the natural numbers $\mathbb{N}$, again, this may be addressed in a later update of these notes. One such way is axiomatically (see the Peano axioms).

Claim 1. The number $\sqrt{2}$ is not rational. Note that we can approximate $\sqrt{2}$ with finite (hence rational) decimal numbers:

$$1, 1.4, 1.41, 1.414, 1.4142, \ldots$$

We say this expansion "tends to $\sqrt{2}$"; however, as we will see, $\sqrt{2}$ cannot be clearly defined as a rational number. Also, as we will see later, the notion of a sequence tending to or approximating a value has a rigorous technical definition. For now, we present an elementary proof of the claim.

Proof. Suppose for the sake of contradiction that $\sqrt{2} \in \mathbb{Q}$. Then there exist $p, q \in \mathbb{Z}$ with $q \neq 0$ such that $\sqrt{2} = \frac{p}{q}$. Note that there are many ways of writing the same number fractionally, e.g. $\frac{1}{2} = \frac{2}{4} = \frac{3}{6}$. In examining these representations we see that there is a way we can write such a fraction so that it is fully reduced, i.e. $\frac{2}{4} = \frac{1}{2}$ by dividing the numerator and denominator by 2, and $\frac{3}{6} = \frac{1}{2}$ by dividing the numerator and denominator by 3. This division leaves the value unchanged since it is being applied to both numerator and denominator. We can assume without loss of generality that $\frac{p}{q}$ is fully reduced, otherwise we can divide p and q by all their common factors before continuing.

Observe that $\sqrt{2} = \frac{p}{q}$ implies $2 = \frac{p^2}{q^2}$, which leads to the equality

$$2q^2 = p^2$$

which in turn shows that p^2 is even. We claim that p is therefore even. Suppose p were odd. Then $p = 2k + 1$ for some integer k, and $p^2 = (2k+1)^2 = 4k^2 + 4k + 1$, which means p^2 is odd as well, but since we know this is not the case, we conclude that p is even.

Since p is even we can write it as $p = 2m$ for some integer m. Then we see that

$$2q^2 = p^2 = 2^2 m^2$$

i.e., $q^2 = 2m^2$, which shows that q^2 and therefore q are even as well. This is a contradiction since now we can write q as $2n$ for some $n \in \mathbb{Z}$, demonstrating that $\frac{p}{q} = \frac{2m}{2n}$ is not fully reduced as previously assumed. We conclude that $\sqrt{2}$ cannot be a rational number. □

2 Basic Topology

Problem 1. Let $\{f_n\}_{n=1}^{\infty}$ be functions such that $f_n : \mathbb{N} \to \mathbb{R}$ for each n. If

$$\sum_{n=1}^{\infty} \left(\sum_{k=1}^{\infty} f_n(k) \right) = \infty \tag{1}$$

then

$$\sum_{k=1}^{\infty}\left(\sum_{n=1}^{\infty} f_n(k)\right) = \infty \tag{2}$$

Proof. Suppose (1) holds true. We divide the proof into two cases. Case 1: Suppose for all $n \in \mathbb{N}$

$$\sum_{k=1}^{\infty} f_n(k) = a_n \in \mathbb{R}^+. \tag{3}$$

That is, for all n, the above series converges to some positive real number. We know this number is positive because the codomain of every f_n is $\mathbb{R}^+$. We can now rewrite (1) as

$$\sum_{n=1}^{\infty} a_n = \infty \tag{4}$$

Since this sum diverges to ∞, by definition its partial sums grow without bound, i.e. for all $M+1 > 0$ there exists $N_0 \in \mathbb{N}$ such that $N \geq N_0$ implies $\sum_{n=1}^{N} a_n > M$. Since (3) holds true for all n, by definition 3.21 in Rudin and by the definition of convergent sequences we know for every $\varepsilon > 0$, in particular letting $\varepsilon := \frac{1}{N}$, there exists $K_n \in \mathbb{N}$ such that $K_N \geq K_n$ implies

$$\left| a_n - \sum_{k=1}^{K_N} f_n(k) \right| < \varepsilon$$

This implies that

$$a_n < \varepsilon + \sum_{k=1}^{K_N} f_n(k)$$

Let $K = \max\{K_1, \ldots, K_N, N\}$ We know K exists because the partial sum inequality $M < \sum_{n=1}^{N} a_n$ ensures we are summing over finitely many f_n, hence there are only finitely many K_n (the exact number is dependent on the value of M). This further implies that

$$M + 1 < \sum_{n=1}^{N} a_n < \sum_{n=1}^{N}\left(\sum_{k=1}^{K} f_n(k)\right) + N\varepsilon \tag{5}$$

Note that we chose $\varepsilon := \frac{1}{N}$, hence $N\varepsilon = 1$ and (5) simplifies to

$$M < \sum_{n=1}^{N} a_n < \sum_{n=1}^{N}\left(\sum_{k=1}^{K} f_n(k)\right) \tag{6}$$

Since finite sums commute, (6) is equivalent to

$$M < \sum_{k=1}^{K}\left(\sum_{n=1}^{N} f_n(k)\right) \tag{7}$$

That is, for any $M > 0$ there exists $N, K \in \mathbb{N}$ so that the above inequality holds, so the sequence of partial sums diverges, hence by definition 3.21 in Rudin

$$\sum_{k=1}^{\infty}\left(\sum_{n=1}^{\infty} f_n(k)\right) = \infty$$

Case 2: Suppose there exists finite $m_1, \ldots, m_l \in \mathbb{N}$ such that

$$\sum_{k=1}^{\infty} f_{m_i}(k) = \infty, \qquad i \in \{1, \ldots, l\} \tag{8}$$

Let $m = \min\{m_1, \ldots, m_l\}$ and $M > 0$. Then by assumption there exists $K_0 \in \mathbb{N}$ such that $K \geq K_0$ implies

$$M < \sum_{k=1}^{K} f_m(k)$$

Since for all $n \in \mathbb{N}$ the codomain of f_n is $\mathbb{R}^+$ we know

$$M < \sum_{k=1}^{K} f_m(k) < \sum_{n=1}^{m-1} \left(\sum_{k=1}^{K} f_n(k) \right) + \sum_{k=1}^{K} f_m(k) = \sum_{n=1}^{m} \left(\sum_{k=1}^{K} f_n(k) \right) \tag{9}$$

Again, since finite sums commute the inequality in (9) can be rewritten as

$$M < \sum_{k=1}^{K} \left(\sum_{n=1}^{m} f_n(k) \right)$$

Since M can be arbitrarily large this shows that the sequence of partial sums grows without bound, hence

$$\sum_{k=1}^{\infty} \left(\sum_{n=1}^{\infty} f_n(k) \right) = \infty$$

This completes the proof. □

Problem 2. $K = \{f \in l^1 : |f(k)| \leq 1/k \ \forall k\}$ is compact in l^1.
Let's first take a slight detour and prove if a subset $A \subset X$ of a metric space X with metric d_X is not bounded then it is not compact, because for some reason this proof isn't in Rudin.

Proof. From Rudin, a subset A is not bounded if for all $M \in \mathbb{R}^+$ and for all $q \in X$ there exists $p \in A$ such that $d_X(p, q) \geq M$. To show that A is not compact we generate an open cover and show there is no finite subcover. Let $r \in A$. We claim $\{B_n(r)\}_{n \in \mathbb{N}}$ is an open cover of A. To show this, let $r' \in A$ be an arbitrary point. Then since X is a metric space we know $d_X(r, r') = D \in \mathbb{R}$. By the Archimedean property we know there exists $N \in \mathbb{N}$ such that $N > D$, that is, $d_X(r, r') < N$, which means $r' \in B_N(r) \in \{B_n(r)\}_{n \in \mathbb{N}}$, hence $r' \in \bigcup_{n \in \mathbb{N}} B_n(r)$ and $A \subset \bigcup_{n \in \mathbb{N}} B_n(r)$. Since we showed $\{B_n(r)\}_{n \in \mathbb{N}}$ is an open cover of A, suppose for contradiction that there exists a finite subcover $\{B_{n_k}(r)\}_{k=1}^{m}$. Let $N = \max\{n_1, \ldots, n_m\}$. Then $\bigcup_{k=1}^{m} B_{n_k}(r) \subset B_N(r)$. To see this, let $b \in \bigcup_{k=1}^{m} B_{n_k}(r)$. Then there exists $i \in \{1, \ldots, m\}$ such that $b \in B_{n_i}(r)$, which means $d_X(r, b) < n_i \leq N$. Hence $d_X(r, b) < N$, so $b \in B_N(r)$. Therefore $B_N(r)$ alone covers A, so for all $a \in A$ $d_X(r, a) < N$, but this contradicts the fact that A is not bounded, since by definition we know there must exist some $a' \in A$ such that $d_X(r, a') \geq N$. □

Since we proved unbounded subsets of a metric space cannot be compact, we now show that $K = \{f \in l^1 : |f(k)| \leq 1/k \ \forall k\}$ is not bounded.

Proof. Suppose for contradiction that K is bounded. Then from Rudin this means there exists $M/2 \in \mathbb{R}^+$ and there exists $q \in l^1$ such that $d_1(f, q) < M/2$ for all $f \in K$. This is equivalent to $K \subset B_{M/2}(q)$. Let $f_0 : \mathbb{N} \to \mathbb{R}$ be such that $m \mapsto 0 \ \forall m \in \mathbb{N}$. Then since $0 < 1/k$ for all $k \in \mathbb{N}$, $f_0 \in K$. Next let $f_n : \mathbb{N} \to \mathbb{R}$ be such that $k \mapsto 1/k$ if $k \leq n \in \mathbb{N}$ and $k \mapsto 0$ otherwise. Then for all $n \in \mathbb{N}$ $f_n \in K$. By assumption, $d_1(f_0, q) < M/2$ and $d_1(f_n, q) < M/2$. By the triangle inequality we see that

$$d(f_0, f_n) \leq d(f_0, q) + d(q, f_n) < 2\frac{M}{2} = M$$

Hence for all $n \in \mathbb{N}$, $f_n \in B_M(f_0)$. Which means

$$d_1(f_n, f_0) = \sum_{m=1}^{\infty} |f_n(m)| < M$$

But from Rudin we know that the harmonic series $\sum_{m=1}^{\infty} 1/m$ diverges to ∞, which means for any $M > 0$ there exists $N_0 \in \mathbb{N}$ such that $N \geq N_0$ implies

$$\sum_{m=1}^{N} \frac{1}{m} > M$$

But this is precisely the same sum as

$$\sum_{m=1}^{\infty} f_N(m)$$

Hence for $N \in \mathbb{N}$, $d_1(f_N, f_0) > M$, a contradiction, which means K is not bounded and therefore is not compact. $\square$

Problem 3. $K = \{f \in l^\infty : |f(k)| \leq 1 \ \forall k\}$ is compact in l^∞.

Proof. We claim K is compact and use sequential compactness to prove. Recall a subset $A \subset X$ of a metric space X is compact if all sequences in A have a subsequence converging to a point in A. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence in K. Then by the definition of K, $\{f_n(k)\}_{n=1}^{\infty} \subset \left[\frac{-1}{k}, \frac{1}{k}\right]$ for all $k \in \mathbb{N}$. Furthermore, each closed interval $\left[\frac{-1}{k}, \frac{1}{k}\right]$ is compact in $\mathbb{R}$, therefore by sequential compactness we know each $\{f_n(k)\}_{n=1}^{\infty}$ has a subsequence converging to some value in the corresponding closed interval, that is, there exists a strictly increasing function $\gamma_1 : \mathbb{N} \to \mathbb{N}$ so that $\{f_{\gamma(n)}(1)\}_{n=1}^{\infty}$ converges to some $a_1 \in [-1, 1]$, there exists a strictly increasing $\gamma_2 : \mathbb{N} \to \mathbb{N}$ so that $\{f_{\gamma_1 \circ \gamma_2(n)}(2)\}_{n=1}^{\infty}$ converges to some $a_2 \in \left[\frac{-1}{2}, \frac{1}{2}\right]$, and in general, for any $k \in \mathbb{N}$ there exists a strictly increasing $\gamma_k : \mathbb{N} \to \mathbb{N}$ so that $\{f_{\gamma_1 \circ \cdots \circ \gamma_k(n)}(k)\}_{n=1}^{\infty}$ converges to some $a_k \in \left[\frac{-1}{k}, \frac{1}{k}\right]$. Let $\beta : \mathbb{N} \to \mathbb{R}$ be defined so that for each $k \in \mathbb{N}$, $k \mapsto a_k$, where a_k is defined as above. By construction, we know that $range(\beta) \subset [-1, 1]$, hence β is bounded and therefore is an element of l^∞. We claim that

$$\lim_{n \to \infty} d_\infty(f_{\gamma_1 \circ \cdots \circ \gamma_n}, \beta) = 0$$

Note that this claims the sequence of functions $\{f_n\}_{n=1}^{\infty}$ has a subsequence which converges to $\beta \in l^\infty$. This claim is equivalent to

$$\lim_{n \to \infty} \sup_{k \in \mathbb{N}} \{\left|f_{\gamma_1 \circ \cdots \circ \gamma_n(n)}(k) - \beta(k)\right|\} = 0$$

Recall that for each $k \in \mathbb{N}$, $\beta(k) = a_k \in [-1/k, 1/k]$ and $\{f_{\gamma_1 \circ \cdots \circ \gamma_k(n)}(k)\}_{n=1}^{\infty}$ converges to a_k. Let $\varepsilon > 0$. So for each k there exists $N_k \in \mathbb{N}$ so that $M_k \geq N_k$ implies

$$\left|f_{\gamma_1 \circ \cdots \circ \gamma_k(M_k)}(k) - \beta(k)\right| < \varepsilon$$

We know by the Archimedean property that no matter how small ε is, there exists $N_0 \in \mathbb{N}$ so that $1/N_0 < \varepsilon$. Therefore we know for all $k \in \{1, 2, \ldots, N_0\}$ there exists N_k so that $M_k \geq N_k$ implies

$$\left|f_{\gamma_1 \circ \cdots \circ \gamma_k(M_k)}(k) - \beta(k)\right| < \varepsilon$$

Since there are now finitely many k, the exact number depending on ε, we can let $M = \max\{N_1, \ldots, N_{N_0}, N_0\}$. Then $M_k \geq M$ implies

$$\sup_{k \in \mathbb{N}} \{\left|f_{\gamma_1 \circ \cdots \circ \gamma_k(M_k)}(k) - \beta(k)\right|\} < \varepsilon$$

Hence

$$\lim_{n \to \infty} \sup_{k \in \mathbb{N}} \{\left|f_{\gamma_1 \circ \cdots \circ \gamma_n(n)}(k) - \beta(k)\right|\} = 0$$

as desired. Since $\{f_n\}_{n=1}^{\infty}$ was chosen arbitrarily, we have proven that K is sequentially compact, therefore also compact, in l^∞. $\square$

Problem 4. If A, B are subsets of $\mathbb{C}$ with A compact, then there exists a point $a \in A$ such that: for all $x \in A$ and $y \in B$, there exists $b \in B$ such that $|a - b| \leq |x - y|$.

Proof. Define the distance between two sets $d(A, B) := \inf\{|p - q| : p \in A,\ q \in B\}$. We have two cases. Case 1: Suppose there exist $a \in A$ and $b \in B$ that satisfy $|a - b| = d(A, B)$. Then by the definition of the greatest lower bound, $|a - b| \leq |x - y|$ for any $x \in A$ and $y \in B$. Case 2: Suppose there is not $a \in A$ and $b \in B$ satisfying $|a - b| = d(A, B)$. Since $|p - q|$ is always nonnegative, $d(A, B)$ is bounded below by 0, hence by the least upper bound property, which according to Rudin is equivalent to the greatest lower bound property, $d(A, B) = \inf\{|p - q| : p \in A, q \in B\} = \delta$ exists. We claim δ is a limit point of $\{|p - q| : p \in A, q \in B\}$. Let $\varepsilon > 0$. Then $\delta \neq \delta + \varepsilon/n$, $(\delta + \varepsilon/n) \in B_\varepsilon(\delta)$, and $\delta + \varepsilon/n \in \{|p - q| : p \in A, q \in B\}$ for sufficiently large $n \in \mathbb{N}$ depending on ε. Hence every open ball centered at δ contains a point not equal to δ in $\{|p-q| : p \in A, q \in B\}$. Since δ is a limit point of $\{|p - q| : p \in A, q \in B\}$, by theorem 3.2 in Rudin we know there exists a sequence $\{r_n\}_{n=1}^\infty$ in $\{|p - q| : p \in A, q \in B\}$ that converges to δ. Since each element of the set is of the form $|p - q|$ we see that the sequence can be written as $\{|p_n - q_n|\}_{n=1}^\infty$, where $p_n \in A$ and $q_n \in B$ for all $n \in \mathbb{N}$. Since A is compact, by theorem 3.6 in Rudin we know there exists a subsequence $\{p_{n_k}\}$ converging to a value $a \in A$, and since $\{r_n\}_{n=1}^\infty$ converges to δ we know every subsequence of $\{r_n\}$ converges to δ, that is, $|a - q_{n_k}|$ must converge to δ. Hence $\{q_{n_k}\}$ must converge to some $b \in B$ satisfying $|a - b| = \delta = \inf\{|p - q| : p \in A, q \in B\}$. This completes the proof. $\square$

3 Limits of Functions and Continuity

Unless otherwise noted, in this section X and Y are metric spaces with respective metrics d_X and d_Y, and $f : X \to Y$.

Definition 3.1. (Rudin) Suppose $E \subset X$ and that p is a limit point of E. We define the limit of $f(x)$ as x (in E) approaches p, and write

$$\lim_{x \to p} f(x) = q \tag{10}$$

if there is a point $q \in Y$ such that for every $\varepsilon > 0$ there exists $\delta > 0$ such that $0 < d_X(x, p) < \delta$ implies $d_Y(f(x), q) < \varepsilon$. Symbolically, this can be written as

$$\lim_{x \to p} f(x) = q \iff \forall \varepsilon > 0\ \exists \delta > 0 \ni 0 < d_X(x, p) < \delta \implies d_Y(f(x), q) < \varepsilon.$$

Intuitively, q being the limit of f at p means no matter how close $f(x)$ is to q (or how small ε is), we can always find a bound $\delta > 0$ for the distance between x and a. Personally, I find that the negation of the definition provides more intuitive insight:

$$\lim_{x \to p} f(x) \neq q \iff \exists \varepsilon > 0\ \forall \delta > 0,\ 0 < d_X(x, p) < \delta \wedge d_Y(f(x), q) \geq \varepsilon$$

Here, $\wedge$ stands for the logical *and* used in propositional logic. This negation basically means the limit of $f(x)$ as x approaches p cannot be q because no matter how close x is to p (or how small δ is), there will always be a *fixed* positive distance greater than or equal to ε between $f(x)$ and q, that is, $f(x)$ does not get arbitrarily close to q as x gets arbitrarily close to p.

Remark. Note that in the definition we require p to be a *limit point*, not necessarily a point in E itself. Recall that limit points of a set E are points that are not necessarily elements of E but are nonetheless arbitrarily close to E. (Every neighborhood of a limit point p intersects E with some element $s \neq q$).

Theorem 3.2. (Rudin) Let $E \subset X$ and p be a limit point of E. Then

$$\lim_{x \to p} f(x) = q \iff \lim_{n \to \infty} f(p_n) = q \tag{11}$$

for every sequence $\{p_n\}_{n=1}^\infty$ in E such that $\forall n \in \mathbb{N}\ p_n \neq p$ and $\lim_{n \to \infty} p_n = p$.

Proof. Suppose $\lim_{x \to p} f(x) = q$ and let $\{p_n\}_{n=1}^\infty$ be defined as above. Then we know for every $\varepsilon > 0$ there exists $\delta > 0$ such that $0 < d_X(x, p) < \delta \implies d_Y(f(x), q) < \varepsilon$. Since $\delta > 0$, we also know there exists $N_0 \in \mathbb{N}$ such

that $n \geq N_0 \implies 0 < d_X(p_n, p) < \delta$. We know $0 < d_X(p_n, p)$ because $p_n \neq p$ for all n. Since $\lim_{x \to p} f(x) = q$ and for $n \geq N_0$, $0 < d_X(p_n, p) < \delta$, we know $d_Y(f(p_n), q) < \varepsilon$. Since ε is arbitrary this proves $\lim_{n \to \infty} f(p_n) = q$.
We use the contrapositive to prove the other direction of implication. Suppose $\lim_{x \to p} f(x) \neq q$. Then there exists $\varepsilon_0 > 0$ such that for any $\delta > 0$ we have $0 < d_X(x, p) < \delta$ and $\varepsilon_0 \leq d_Y(f(x), q)$. Let $\{p_n\}_{n=1}^{\infty}$ be as before. Then we know for any $\varepsilon > 0$, in particular for any $\delta > 0$ as above, there exists $N_0 \in \mathbb{N}$ such that $n \geq N_0$ implies $0 < d_X(p_n, p) < \delta$; however, since $\lim_{x \to p} f(x) \neq q$, this means $\varepsilon_0 \leq d_Y(f(p_n), q)$ for all $n \geq N_0$. Since we must show these inequalities hold for all n, we must now show for all $n \in \{1, \dots, N_0 - 1\}$, there exists $\varepsilon_n > 0$ such that $\varepsilon_n \leq d_Y(f(p_n), q)$. We know for $n \in \{1, \dots, N_0 - 1\}$ there exists $\delta_n > 0$ such that $0 < d_X(p_n, p) < \delta_n$. Hence there exists $\varepsilon_n > 0$ such that $\varepsilon_n \leq d_Y(f(p_n), q)$. Let $\varepsilon = \min\{\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{N_0}\}$. Then for all $n \in \mathbb{N}$, $\varepsilon \leq d_Y(f(p_n), q)$ as desired. □

Definition 3.3. Let $p \in E \subset X$. We say f is continuous at p if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d_X(x, p) < \delta \implies d_Y(f(x), f(p)) < \varepsilon.$$

If f is continuous at every point in E we say that f is continuous on E.

Remark. Note that for f to be continuous at a point p, $f(p)$ must exist. Also note that if we modify our restriction and make p a limit point of E as well as a point in E, we obtain that f is continuous at p if and only if

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

Theorem 3.4. f is continuous at a point $a \in E$ if and only if for all sequences $\{a_n\}_{n=1}^{\infty}$ in E that converge to a,

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

Proof. Let $\{a_n\}_{n=1}^{\infty}$ be as above and suppose $f(a_n) \to f(a)$. Let $\varepsilon > 0$. Our goal is to find a $\delta > 0$ such that $d_X(a, x) < \delta \implies d_Y(f(a), f(x)) < \varepsilon$. Suppose for contradiction that there exists $x \in E$ such that $d_X(a, x) < \delta$ and $d_Y(f(a), f(x) \geq \varepsilon$. □

Problem 1. If f is a continuous mapping from a metric space X into a metric space Y, prove that

$$f\left(\overline{E}\right) \subset \overline{f(E)} \tag{12}$$

Proof. First note that if E is closed then the result immediately follows, so suppose E is not closed. Since $E \subset \overline{E}$, $f(E) \subset f\left(\overline{E}\right)$. Let $y \in f\left(\overline{E}\right)$.

Case 1: $y \in f(E)$. The result immediately follows from the fact that $f(E) \subset \overline{f(E)}$.

Case 2: $y \in f\left(\overline{E}\right) \setminus f(E)$. Then there exists an $x \in \overline{E}$ such that $y = f(x)$. Since $x \in \overline{E}$ we know x is a limit point of E, so there exists a sequence $\{x_n\}_{n=1}^{\infty} \subset E$ such that $\lim_{n \to \infty} x_n = x$. By continuity of f this means $\lim_{n \to \infty} f(x_n) = f(x) = y$. Therefore for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $n \geq N$ implies $|f(x_n) - y| < \varepsilon$. Since $x_n \in E$ for all $n \in \mathbb{N}$, this means $f(x_n) \in f(E)$, which means $f(x_n) \in N_\varepsilon(y)$, so $f(x_n) \in \underline{E} \cap N_\varepsilon(y)$. By choice of y we know $f(x_n) \neq y$, which means y is a limit point of $f(E)$ and therefore $y \in \overline{f(E)}$ as desired. □

Problem 2. Let f be a continuous real function on a metric space X. Prove that $\ker(f)$ is closed.

Proof. It suffices to show $\ker(f)$ contains all of its limit points. Let x be a limit point of $\ker(f)$. Then by definition there exists a sequence $\{x_n\}_{n=1}^{\infty} \subset \ker(f)$ such that $\lim_{n \to \infty} x_n = x$. By continuity this means $\lim_{n \to \infty} f(x_n) = f(x)$, but since $x_n \in \ker(f)$ for all $n \in \mathbb{N}$, we know $\{f(x_n)\}_{n=1}^{\infty}$ is the constant zero sequence. Therefore $f(x_n) \to 0 = f(x)$, which implies $x \in \ker(f)$. □

Problem 3. Let f, g be continuous mappings of a metric space X into a metric space Y, and let E be a dense subset of X. Prove that $f(E)$ is dense in $f(X)$. If $g(p) = f(p)$ for all $p \in E$, prove that $g(p) = f(p)$ for all $p \in X$, i.e., show that a continuous mapping is determined by its values on a dense subset of its domain.

Proof that $f(E)$ is dense in $f(X)$. We know from (4.2) that $f\left(\overline{E}\right) \subset \overline{f(E)}$. Since E is dense in X we know $\overline{E} = X$, i.e., $f(X) \subset \overline{f(E)}$. □

Problem 4. Let $h : [0,1) \longrightarrow \mathbb{R}$ be uniformly continuous on its domain. Then there exists a unique continuous function $g : [0,1] \longrightarrow \mathbb{R}$ such that $g(x) = h(x)$ for all $x \in [0,1)$.

Proof. Consider $\lim_{x\to 1} h(x)$. We first prove this limit exists. Suppose for contradiction that $\lim_{x\to 1} h(x)$ does not exist. Suppose there is some sequence $\{x_n\}_{n=1}^{\infty}$ in the domain of h such that $\lim_{n\to\infty} x_n = 1$ and $\lim_{n\to\infty} h(x_n) = \infty$. Then the sequence $h(x_n)$ is said to diverge to ∞. Since $\{x_n\}_{n=1}^{\infty}$ is convergent and $\mathbb{R}$ is complete, we know that $\{x_n\}$ is a Cauchy sequence. So for all $\varepsilon > 0$ there exists $N_0 \in \mathbb{N}$ such that $n > m \geq N_0 \implies |x_n - x_m| < \varepsilon$. Now since h is uniformly continuous we know for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x, y \in [0,1)$, if $|x - y| < \delta$ then $|h(x) - h(y)| < \varepsilon$. Therefore since $\{x_n\}_{n=1}^{\infty} \subset [0,1)$ is a Cauchy sequence, we know for any $\delta > 0$ there exists $N_0 \in \mathbb{N}$ such that $n > m \geq N_0 \implies |x_n - x_m| < \delta \implies |h(x_n) - h(x_m)| < \varepsilon$. But this means that $h(x_n)$ is a Cauchy sequence as well, hence it must also converge. Therefore

$$\lim_{x\to 1} h(x) \neq \infty$$

Now suppose there exist two sequences $\{a_n\} \to 1$ and $\{b_n\} \to 1$ in $[0,1)$ such that

$$\lim_{x\to 1} h(a_n) = A \neq B = \lim_{x\to 1} h(b_n)$$

We know by theorem 3.3 in Rudin that $\{a_n\} - \{b_n\} \to 0$, that is, for every $\delta > 0$ there exists $N_0 \in \mathbb{N}$ such that $n \geq N_0 \implies |a_n - b_n| < \delta$. Since h is uniformly continuous we know that this implies for any $\varepsilon > 0$, $|h(a_n) - h(b_n)| < \varepsilon$, which means

$$\lim_{n\to\infty} h(a_n) - h(b_n) = 0 \iff \lim_{n\to\infty} h(a_n) = \lim_{n\to\infty} h(b_n)$$

So in fact $\{a_n\}, \{b_n\} \to 1$ and $h(a_n), h(b_n) \to A = B$. Therefore the limit $\lim_{x\to 1} h(x)$ is uniquely defined. Since $h(x)$ is uniformly continuous it is continuous and by theorem 4.2 in Rudin we know that $g(x)$ must be continuous. We know $g(x)$ is unique because if there exists another continuous $f : [0,1] \longrightarrow \mathbb{R}$ such that $f(x) = h(x)$ for all $x \in [0,1)$ then since $\lim_{x\to 1} h(x)$ exists, f is continuous if and only if $f(x) = \lim_{x\to 1} h(x)$. □

Problem 5. Suppose X is a compact metric space and $f : X \longrightarrow \mathbb{R}$ is a not necessarily continuous function for which $f^{-1}([t,\infty))$ is closed for all $t \in \mathbb{R}$. Then f achieves its maximum value on X.

Proof. Let $U_r = \{\left(f^{-1}([r,\infty))\right)^c : r \in R\}$. We claim that $\bigcup_{r\in\mathbb{R}} U_r$ is an open cover of X, that is,

$$X \subset \bigcup_{r\in\mathbb{R}} \left(f^{-1}([r,\infty))\right)^c \tag{13}$$

First note that by assumption each preimage $f^{-1}([r,\infty))$ is closed, hence the complement is open. Specifically, the preimage is defined as

$$f^{-1}([r,\infty)) = \{x \in X : f(x) \in [r,\infty)\}$$

Hence the complement is defined to be

$$\left(f^{-1}([r,\infty))\right)^c = \{x \in X : f(x) \notin [r,\infty)\} \tag{14}$$

Now let $x \in X$ be an arbitrary point. Since $f(x) \in \mathbb{R}$, we know that since $f(x) < f(x) + 1$, $x \notin f^{-1}([f(x) + 1,\infty))$, so $x \in \left(f^{-1}([f(x)+1,\infty))\right)^c$. Since x was arbitrarily chosen this verifies (1). Since X is compact there exists a finite subcover $\{\left(f^{-1}([r_i,\infty))\right)^c : i \in \{1,\dots,n\}\}$. Let $r_M = \max\{r_1,\dots,r_n\}$. We claim for all $i \in \{1,\dots,n\}$ that $\left(f^{-1}([r_i,\infty))\right)^c \subset \left(f^{-1}([r_M,\infty))\right)^c$. Let i be arbitrarily chosen and suppose $x' \in \left(f^{-1}([r_i,\infty))\right)^c$. Then from (2) we know $f(x') \notin [r_i,\infty)$. Since $r_i \leq r_M$, we know $[r_M,\infty) \subset [r_i,\infty)$. Hence

$f(x') \notin [r_M, \infty)$, i.e. $x' \in \left(f^{-1}([r_M, \infty))\right)^c$. Since i was arbitrarily chosen this shows for all $i \in \{1, \ldots, n\}$ that $\left(f^{-1}([r_i, \infty))\right)^c \subset \left(f^{-1}([r_M, \infty))\right)^c$. Therefore by compactness, (1), and the definition of containment we see that $X \subset \left(f^{-1}([r_M, \infty))\right)^c$, in other words, for all $x \in X$, $x \notin f^{-1}([r_M, \infty))$, which means $f(x) \notin [r_M, \infty)$. Hence f is bounded above on X. Since $f(X)$ is bounded above and $f(X) \subset \mathbb{R}$, by the least upper bound property there exists a supremum $M := \sup(f(X))$. We claim that

$$f^{-1}([M, \infty)) \neq \emptyset \tag{15}$$

Consider the family of sets $\{V_n : V_n = f^{-1}([M - \frac{1}{n}, \infty)),\ n \in \mathbb{N}\}$. Then since $[M-1, \infty) \supset [M - \frac{1}{2}, \infty) \supset \cdots$, we see that $V_1 \supset V_2 \supset V_3 \supset \cdots$. Therefore, since each preimage is closed and X is compact, each preimage is compact and by corollary 2.36 in Rudin we know that $\bigcap_{n \in \mathbb{N}} V_n \neq \emptyset$. In still proving (3), we claim that $\bigcap_{n \in \mathbb{N}} V_n = f^{-1}([M, \infty))$. Since $f^{-1}([M, \infty)) \subset V_n$ for all n, we need only prove $\bigcap_{n \in \mathbb{N}} V_n \subset f^{-1}([M, \infty))$. Suppose for contradiction that there exists some $x \in \bigcap_{n \in \mathbb{N}} V_n$ such that $x \notin f^{-1}([M, \infty))$. Then we know $f(x) < M$, so $M - f(x) > 0$ and by the Archimedean property there exists $m \in \mathbb{N}$ such that $M - f(x) > \frac{1}{m}$. But then $f(x) \notin [M - \frac{1}{m}, \infty)$, which means $x \notin f^{-1}(M - \frac{1}{m}, \infty)$. Since $\bigcap_{n \in \mathbb{N}} V_n \subset f^{-1}([M - \frac{1}{m}, \infty))$, $x \notin \bigcap_{n \in \mathbb{N}} V_n$, a contradiction since we assumed $x \in \bigcap_{n \in \mathbb{N}} V_n$. Therefore $\bigcap_{n \in \mathbb{N}} V_n = f^{-1}([M, \infty)) \neq \emptyset$. Since $f^{-1}([M, \infty)) \neq \emptyset$, there exists $x_0 \in X$ such that $f(x_0) \geq M$, but since $M = \sup(f(X))$, we know $f(x_0) \leq M$, hence $f(x_0) = M$ and f therefore achieves its maximum on X. □

Problem 6. Suppose that g is a not necessarily continuous positive real valued function of a real number. If $a < b$ are real numbers; then there is a finite sequence $a = t_0 < t_1 < \cdots < t_n = b$ of real numbers such that in each interval $[t_k, t_{k+1}]$ there is a point x_0 such that $|t_{k+1} - t_k| < g(x_0)$.

Proof. Disclaimer: I worked with Troy and Cody on this problem. Define the family of open sets $\{B_{\frac{g(x_\alpha)}{4}}(x_\alpha) : x_\alpha \in [a, b]\}$. We claim this is an open cover of $[a, b]$. Let $x \in [a, b]$. Then we know $x \in B_{\frac{g(x)}{4}}(x)$, hence $x \in \bigcup_\alpha B_{\frac{g(x_\alpha)}{4}}(x_\alpha)$, which shows us that $[a, b] \subset \bigcup_\alpha B_{\frac{g(x_\alpha)}{4}}(x_\alpha)$, so our claim is verified. Since $[a, b]$ is compact we know there exists a finite subcover, i.e. there is a finite set $\{x_1, \ldots, x_m\} \subset [a, b]$ which we can order so that $x_1 < x_2 < \cdots < x_m$ with $[a, b] \subset \bigcup_{i=1}^m B_{\frac{g(x_i)}{4}}(x_i)$. While this is indeed a finite subcover, we want to make sure there are no open sets contained in other open sets of this family. In other words, for all $x_j \in \{x_1, \ldots, x_m\}$, if for some $x_i \neq x_j$ we have $B_{\frac{g(x_j)}{4}}(x_j) \subset B_{\frac{g(x_i)}{4}}(x_i)$, then we discard $B_{\frac{g(x_j)}{4}}(x_j)$ from our family of sets unless $x_j = a$ or $x_j = b$. In the latter two cases we discard $B_{\frac{g(x_i)}{4}}(x_i)$ and add the open set $B_{\frac{g(x_j)}{4}}(x_j) \cup B_{\frac{g(x_i)}{4}}(x_i)$, extending the neighborhoods containing the boundary points a or b. Once all redundant sets are removed or extended in the case of boundary points, we obtain a new ordered finite set $\{x_1, \ldots, x_n\}$ with $n \leq m$. Let $a = t_0$. For $i \in \{1, \ldots, n\}$ let $t_i = \min(\{x_1, \ldots, x_r\} \setminus \{t_0 \ldots, t_{i-1}\})$. If $b \notin \{x_1, \ldots, x_r\}$ then let $t_{r+1} = b$. Otherwise let $t_r = b$. Notice that for any $k \in \{1, \ldots, n-1\}$ that $B_{\frac{g(t_k)}{4}}(t_k) \cap B_{\frac{g(t_{k+1})}{4}}(t_{k+1}) \neq \emptyset$. Suppose for instance that $B_{\frac{g(t_k)}{4}}(t_k) \cap B_{\frac{g(t_{k+1})}{4}}(t_{k+1}) = \emptyset$. Then there exists some $x \in [a, b]$ such that $t_k + \frac{g(t_k)}{4} \leq x \leq t_{k+1} - \frac{g(t_{k+1})}{4}$. Hence $x \in B_{\frac{g(t_j)}{4}}(t_j)$ for some $k \neq j \neq k+1$, but then $B_{\frac{g(t_k)}{4}}(t_k) \subset B_{\frac{g(t_j)}{4}}(t_j)$ or $B_{\frac{g(t_{k+1})}{4}}(t_{k+1}) \subset B_{\frac{g(t_j)}{4}}(t_j)$, a contradiction. Therefore $B_{\frac{g(t_k)}{4}}(t_k) \cap B_{\frac{g(t_{k+1})}{4}}(t_{k+1}) \neq \emptyset$. We claim for each interval $[t_k, t_{k+1}]$ there is a point in $[t_k, t_{k+1}]$ such that $|t_{k+1} - t_k| < g(x_k)$. By construction, $|t_{k+1} - t_k| \leq \frac{1}{2}\max\{g(t_k), g(t_{k+1})\} < \max\{g(t_k), g(t_{k+1})\}$. Since $t_k, t_{k+1} \in [t_k, t_{k+1}]$, this concludes the proof. □

Problem 7. Let $f : [0, 1] \times [0, 1] \longrightarrow \mathbb{R}$ be continuous and assume for all $x \in [0, 1]$ there is a unique y_x such that $f(x, y_x) = \max\{f(x, y) : y \in [0, 1]\}$. If $g(x) = y_x$ then $g : [0, 1] \longrightarrow [0, 1]$ is continuous.

Proof. Suppose for contradiction that $g(x)$ is not continuous at some point $x_0 \in [0, 1]$. Since $[0, 1]$ is compact and g is not continuous at $x_0 \in [0, 1]$ we know there exists some sequence $\{a_n\}_{n=1}^\infty$ in $[0, 1]$ which has a convergent subsequence $\{a_{n_k}\} \longrightarrow x_0$ such that $g(a_{n_k}) \longrightarrow p \neq g(x_0)$. For the sake of clarity and simplification of notation, since $\{a_{n_k}\}_{n=1}^\infty$ is a convergent subsequence, it is in fact a convergent sequence, so we can write $\{a_{n_k}\}_{n=1}^\infty$ as the sequence $\{x_n\}_{n=1}^\infty \longrightarrow x_0$ with $g(x_n) \longrightarrow p \neq g(x_0)$. By theorem 3.4 in Rudin, $(x_n, g(x_n))$ converges to (x_0, p) if and only if $\{x_n\} \longrightarrow x_0$ and $g(x_n) \longrightarrow p$. Therefore since f is continuous on $[0, 1] \times [0, 1]$, $\{x_n\} \longrightarrow x_0$, and

$g(x_n) \longrightarrow p$, by theorem 4.2 in Rudin $f(x_n, g(x_n)) \longrightarrow f(x_0, p)$. By the assumption of uniqueness, we know $f(x_0, p) \neq f(x_0, g(x_0))$. Therefore $|f(x_0, p) - f(x_0, g(x_0))| > 0$. Let $\varepsilon_0 = |f(x_0, p) - f(x_0, g(x_0))|$. Then by the definition of convergence we know for $\varepsilon_0/2$ there exists $N_0 \in \mathbb{N}$ such that $n \geq N_0$ implies $|f(x_n, g(x_n)) - f(x_0, p)| < \varepsilon_0/2$. By the triangle inequality this implies for all $n \geq N_0$, $|f(x_n, g(x_n)) - f(x_0, g(x_0))| \geq \varepsilon_0/2$. In other words, for sufficiently large n, $f(x_n, g(x_n))$ will always be a fixed positive distance $\varepsilon_0/2$ or larger away from $f(x_0, g(x_0))$. Now since each $x \in [0, 1]$ has a unique y_x such that $f(x, y_y) = \max\{f(x, y) : y \in [0, 1]\}$, we know in particular x_0 has a unique y_{x_0} and each x_n has a unique y_{x_n} and that satisfies this property. □

Problem 8. If $\{a_n\}$ is a **non-increasing** sequence of positive real numbers such that $\sum_n a_n$ converges, then

$$\lim_{n\to\infty} na_n = 0$$

Proof. By the Cauchy criterion we know since $\sum_n a_n$ converges, for all $\varepsilon/2 > 0$ there exists $N_\varepsilon \in \mathbb{N}$ such that $m \geq n \geq N_\varepsilon$ implies

$$\left|\sum_{i=n}^{m} a_i\right| < \varepsilon/2$$

In particular, if $m = 2n$, then we see that

$$\left|\sum_{i=n}^{2n} a_i\right| < \varepsilon/2$$

Since $\{a_n\}$ is non-increasing, we know $a_n \geq a_{n+1} \geq \cdots$ hence,

$$\left|\sum_{i=n}^{2n} a_{2n}\right| \leq \left|\sum_{i=n}^{2n} a_i\right| < \varepsilon/2$$

By the definition of partial sums, the leftmost term is equivalent to $(2n - n + 1)a_{2n} = (n + 1)a_{2n}$. Since our sequence is always positive we need not worry about absolute value. Hence
$na_{2n} < (n + 1)a_{2n} < \varepsilon/2$. Multiplying by 2 gives us $2na_{2n} < \varepsilon$. Therefore for every $\varepsilon > 0$ we know there exists N_ε such $n \geq N_\varepsilon$ implies $2na_{2n} < \varepsilon$. But we also know since $(n + 1)a_{2n} < \varepsilon/2$ that $(2n + 2)a_{2n} < \varepsilon$. Therefore since $\{a_n\}$ is non-increasing and since $2n + 1 < 2n + 2$, we know that

$$(2n + 1)a_{2n+1} \leq (2n + 1)a_{2n} < (2n + 2)a_{2n} < \varepsilon$$

Therefore $(2n+1)a_{2n+1} < \varepsilon$. So we now know for all $\varepsilon > 0$ there exists N_ε such that $n \geq N_\varepsilon$ implies $2na_{2n} < \varepsilon$ and $(2n + 1)a_{2n+1} < \varepsilon$. Therefore for all $\varepsilon > 0$ there exists an N, namely $N := 2N_\varepsilon$, such that $n \geq N$ implies $na_n < \varepsilon$. □

Problem 9. If $f : \mathbb{R} \longrightarrow \mathbb{R}$ is continuous and $E \subset \mathbb{R}$ is closed, then $f(E)$ is closed.

Counterexample. Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be defined so that

$$f(x) = \begin{cases} \frac{1}{x} & \text{if } x \geq 1 \\ 1 & \text{otherwise} \end{cases}$$

We first claim that f is continuous on its domain. Let $x_0 \in \mathbb{R}$. Case 1: $x_0 \in [1, \infty)$. Let $\varepsilon > 0$. Pick

$$\delta < \min\{\varepsilon/2, |x_0 - 1| + \frac{1}{2}\}$$

Then $|x - x_0| < \delta$ implies

$$\left|\frac{1}{x} - \frac{1}{x_0}\right| = \left|\frac{x_0 - x}{x_0 x}\right| = \left|\frac{1}{x_0 x}\right| |x - x_0| < \frac{\delta}{|x x_0|}$$

Since $x > \frac{1}{2}$ and $x_0 > 1$ we know this means

$$\left|\frac{1}{x} - \frac{1}{x_0}\right| \leq 2\frac{\delta}{x_0} < 2\delta \leq \varepsilon$$

Therefore f is continuous on $[1, \infty)$. Case 2: $x_0 \in (-\infty, 1]$. Then $f(x_0) = f(x) = 1$ for all $x \in (-\infty, 1]$. Therefore $|f(x_0) - f(x)| = |1 - 1| = 0 < \varepsilon$ for any $\varepsilon > 0$. Hence we can pick any $\delta > 0$ and $|x - x_0| < \delta$ implies $|f(x) - f(x_0)| < \varepsilon$ in $(-\infty, 1]$. We now claim that $f([1, \infty)) = (0, 1]$. Clearly for all $x \in [1, \infty)$, $\frac{1}{x} \leq 1$, hence 1 is the maximum of the image set. Now if $0 \in f([1, \infty))$, that means there exists $x \in [1, \infty)$ such that $f(x) = 0$, but this is impossible since for all $x \in [1, \infty)$ we know $0 < f(x) = \frac{1}{x}$, but we also know 0 is the infimum of the image set because for any $\varepsilon > 0$, by the Archimedean property there exist $n \in \mathbb{N}$ such that $f(n) = \frac{1}{n} < \varepsilon$. Therefore $f([1, \infty)) = (0, 1]$. Since $[1, \infty)$ contains all of its limit points it is closed, and since 0 is a limit point of $(0, 1]$, we know $(0, 1] = f([1, \infty))$ is not closed, thus we have a counterexample. $\square$

Problem 10. Suppose $f : [0, 1] \longrightarrow (0, 1)$ is a **non-decreasing** function (**NOT** assumed to be continuous). Then there exists $x \in (0, 1)$ such that $f(x) = x$.

Proof. Suppose for contradiction there is no fixed point. Consider the sets $C_1 = \{x \in [0, 1] : f(x) > x\}$ and $C_2 = \{x \in [0, 1] : f(x) < x\}$. By assumption, there are no fixed points, so all points in the domain of f are contained in $C_1 \cup C_2 = [0, 1]$. We also know since the codomain of f is $(0, 1)$ that $f(0) > 0$ and $f(1) < 1$, hence C_1 and C_2 are nonempty. Since $C_1 \subset [0, 1]$ we know $\sup C_1$ exists in $[0, 1]$. Denote $\sup C_1$ by p. We have two cases. Case 1: $p \in C_1$. Then we know $f(p) > p$. Therefore there exists $\varepsilon > 0$ such that $f(p) = p + \varepsilon$. Since p is the supremum of C_1, we know for any elements greater than p, namely $p + \varepsilon$, that $f(p + \varepsilon) < p + \varepsilon$. But since f is non-decreasing we see that $f(p) \leq f(p + \varepsilon)$, hence

$$p + \varepsilon = f(p) \leq f(p + \varepsilon) < p + \varepsilon$$

that is, $p + \varepsilon < p + \varepsilon$, a contradiction. Therefore the only other possibility is case 2: $p \in C_2$. Then we know $f(p) < p$ by definition of C_2. Since $p = \sup C_1$ we know there exists a sequence $\{x_n\} \subset C_1$ such that $\{x_n\} \to p$. Since $\{x_n\} \subset C_1$ we know for all $n \in \mathbb{N}$ that $f(x_n) > x_n$. Furthermore, because $\{x_n\}$ is defined on nonnegative numbers we need not consider the absolute values when considering convergence. Now, if $f(p) < p$ we know there exists $\varepsilon_0 > 0$ such that $\varepsilon_0 = p - f(p)$. Therefore since $\{x_n\}$ converges to p we know there exists $N_0 \in \mathbb{N}$ such that $p - x_{N_0} < \varepsilon_0$. This is equivalent to $p < \varepsilon_0 + x_{N_0}$. Since we picked $\varepsilon_0 = p - f(p)$ we know this is equivalent to $p < x_{N_0} + p - f(p)$, so $f(p) < x_{N_0} + p - p \equiv f(p) < x_{N_0}$. Additionally, since $x_{N_0} \in C_1$ we know $x_{N_0} < f(x_{N_0})$, so $f(p) < x_{N_0} \implies f(p) < f(x_{N_0})$. But then since f is non-decreasing we know $x_{N_0} < p \implies f(x_{N_0}) \leq f(p)$, so we have $f(p) < f(x_{N_0})$ and $f(p) \geq f(x_{N_0})$, a contradiction. Therefore $p \notin C_1 \cup C_2$, but since $p \in [0, 1]$ this can only be true if $f(p) = p$. Hence there indeed exists a fixed point. $\square$

Problem 11. If the function $f : [0, \infty) \longrightarrow \mathbb{R}$ is continuous and $\lim\limits_{x \to \infty} f(x)$ exists as a finite real number, then f is uniformly continuous on its domain.

Proof. Let $\varepsilon > 0$. We know since $\lim_{x\to\infty} f(x) = L \in \mathbb{R}$ that there exists $M > 0$ such that $x \geq M \implies |f(x) - L| < \varepsilon/2$. We first claim that $f|_{[0,M+1]}$ is uniformly continuous. This is immediately true since f is continuous and $[0, M + 1]$ is compact. Therefore there exists a *positive* δ_0 less than 1 such that for any $x, y \in [0, M + 1]$ we have $|x - y| < \delta_0 < 1 \implies |f(x) - f(y)| < \varepsilon$. Next let $x, y \in [M, \infty)$. Then we know from above that $|f(x) - L| < \varepsilon/2$ and $|f(y) - L| < \varepsilon/2$. Therefore by the triangle inequality we see that

$$|f(x) - f(y)| \leq |f(x) - L| + |L - f(y)| < \varepsilon$$

Hence for any $x, y \geq M$ we see that $|f(x) - f(y)| < \varepsilon$. Therefore for any $\delta > 0$, in particular picking $\delta := \delta_0$, as long as $x, y \geq M$ we know $|x - y| < \delta_0 < 1$ and $|f(x) - f(y)| < \varepsilon$, a condition that is *stronger* than uniform continuity, i.e. $|x - y| < \delta_0 \implies |f(x) - f(y)| < \varepsilon$. Since we restricted $\delta_0 < 1$, we know if $x \in [0, M + 1]$ and $y \in [M, \infty)$ that $|x-y| < \delta_0$ implies $x, y \in [M, M+2] \subset [M, \infty)$, hence $|x-y| < \delta_0$ still ensures $|f(x)-f(y)| < \varepsilon$. Therefore f is uniformly continuous on its domain. $\square$

4 Important Definitions (To be organized later)

Definition 4.1. (Algebraic Operations on Functions) Let $f, g : E \longrightarrow \mathbb{C}$ and $x \in E$. We define $(f+g)(x) := f(x) + g(x)$, $(f \cdot g)(x) := f(x) \cdot g(x)$, $(f - g)(x) := f(x) - g(x)$, and for all values of x such that $g(x) \neq 0$ we define $\frac{f}{g}(x) := \frac{f(x)}{g(x)}$.

If $f(x) = c$ for all $x \in E$, where $c \in \mathbb{C}$ is some constant, we say f is a constant function and write $f = c$ or $f \equiv c$. If $f(x) \geq g(x)$ for all $x \in E$ we write $f \geq g$.

Similarly, if $\boldsymbol{f}, \boldsymbol{g} : E \longrightarrow \mathbb{R}^n$ we define

$$(\boldsymbol{f} + \boldsymbol{g})(x) := \boldsymbol{f}(x) + \boldsymbol{g}(x), \qquad (\boldsymbol{f} \cdot \boldsymbol{g})(x) := \boldsymbol{f}(x) \cdot \boldsymbol{g}(x) \tag{16}$$

Where we take $(\cdot) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ as point-wise multiplication. Furthermore, if $\lambda \in \mathbb{R}$ is some constant, we define $(\lambda \boldsymbol{f})(x) := \lambda \boldsymbol{f}(x)$

Remark. These algebraic operations on functions will become very important when we encounter *algebras* of functions and the Stone-Weierstrass theorem.

Definition 4.2. (Continuity) Suppose $(X, d_X), (Y, d_Y)$ are metric spaces, $E \subset X$, $p \in E$, and $f : E \longrightarrow Y$. Then f is said to be *continuous* at p if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $d_X(x, p) < \delta$ implies $d_Y(f(x), f(p)) < \varepsilon$, i.e.

$$\forall \varepsilon > 0 \; \exists \delta > 0 \ni d_X(x, p) < \delta \implies d_Y(f(x), f(p)) < \varepsilon \tag{17}$$

Assuming p is a limit point of E, by definition this means f is continuous at p if and only if

$$\lim_{x \to p} f(x) = f(p) \tag{18}$$

Recall. A norm $||\cdot|| : V \longrightarrow \mathbb{R}$ is a function mapping a vector space V into $\mathbb{R}$ that satisfies some nice stuff including:

1. $||v|| \geq 0$ for all $v \in V$
2. $||v|| = 0 \iff v = 0$
3. $||\lambda v|| = |\lambda| \cdot ||v||$ for any $\lambda \in \mathbb{R}$
4. $||u + v|| \leq ||u|| + ||v||$ for any $u \in V$ (Triangle inequality)

Definition 4.3. (p-norm, $\infty-$norm) Let E^n be an n-dimensional Euclidean space ($\mathbb{R}^n$ or $\mathbb{C}^n$ for our purposes, honestly I don't know if there are any other Euclidean spaces but whatever) and let $p \in \mathbb{Z}^+$ be a positive integer. Then we define the *p-norm* to be the map

$$||\cdot|| : E^n \longrightarrow \mathbb{R}, \qquad (x_1, x_2, \ldots, x_n) \mapsto \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \tag{19}$$

where $|x_i|$ is taken to be the absolute value in $\mathbb{R}$ or the modulus in $\mathbb{C}$.

We define the $\infty-$*norm* to be the map

$$||\cdot||_\infty : E^n \longrightarrow \mathbb{R}, \qquad (x_1, \ldots, x_n) \mapsto \max\{|x_1|, \ldots, |x_n|\} \tag{20}$$

Unless otherwise noted, when working in a Euclidean space E^n we will denote the 2-norm as $||\cdot||$ or just $|\cdot|$.

Remark. The terminology for these norms will become more clear when the ℓ^∞ and ℓ^p sequence spaces are defined. These spaces will provide for many interesting studies of topology and continuity later on.

Remark. (Ignore if you don't know abstract algebra) It is easy to show $\mathbb{C} \cong \mathbb{R}^2$, i.e. $\mathbb{C}$ is ring-*isomorphic* to $\mathbb{R}^2$. One can also show that $\mathbb{C}$ and $\mathbb{R}^2$ are *isometric*, i.e. there exists a bijection that preserves the metric structure of the two spaces. Spaces that admit a norm implicitly admit a metric corresponding to the norm. A consequence of this is that the modulus in $\mathbb{C}$ is the same as the 2-norm in $\mathbb{R}^2$, i.e. for any $z = \alpha + i\beta \in \mathbb{C}$ we know $(\alpha, \beta) \in \mathbb{R}^2$ and

$$|z| = \sqrt{\alpha^2 + \beta^2} \iff ||(\alpha, \beta)||_2 = \sqrt{\alpha^2 + \beta^2} \tag{21}$$

Definition 4.4. (Bounded Function) A mapping $\boldsymbol{f} : E \longrightarrow \mathbb{R}^n$ is said to be *bounded* if there exists a real number $M \in \mathbb{R}$ such that $||\boldsymbol{f}|| \leq M$ for all $x \in E$.

Definition 4.5. (Uniform Continuity) Let $f : X \longrightarrow Y$ be a mapping of metric spaces. We say that f is *uniformly continuous* on X if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $d_Y(f(x), f(y)) < \varepsilon$ for any $x, y \in X$ such that $d_X(x, y) < \delta$. Symbolically this can be interpreted as

$$\forall \varepsilon > 0 \; \exists \delta > 0 \ni \forall x, y \in X \; d_X(x, y) < \delta \implies d_Y(f(x), f(y)) < \varepsilon \tag{22}$$

Remark. Consider the differences between continuity and uniform continuity. Evidently, uniform continuity on a set X implies continuity on X. When defining continuity, we only consider continuity at certain *points* as opposed to entire sets. To say a function is continuous on a set is to say it is continuous at every point in the set. Compare this to uniform continuity where the definition does not mention any points in a set. Rather, uniform continuity is a *global* property of a function. This property is always with regards to a set of points and considers when all the points are of close enough distance. Later we will see that uniform continuity gives us nice properties with respect to differentiation and integration.

Claim 2. The map $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $x \mapsto x^2$ is NOT uniformly continuous on $\mathbb{R}$.

Proof. Suppose for contradiction that f is uniformly continuous on $\mathbb{R}$. Then for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for any $x, y \in \mathbb{R}$, $|x - y| < \delta$ implies $|x^2 - y^2| < \varepsilon$. In particular, pick $\varepsilon = 1$, and pick $x \in \mathbb{R}$ so that $\frac{1}{\delta} + \frac{\delta}{4} < x$. Then by letting $y := x - \frac{\delta}{2}$ we see that $|x - y| = |x - x + \frac{\delta}{2}| = |\frac{\delta}{2}| = \frac{\delta}{2} < \delta$. Therefore we know $|x^2 - y^2| < 1$. But a closer observation yields

$$|x^2 - y^2| = |x + y||x - y| = \left|2x + \frac{\delta}{2}\right| \left|\frac{\delta}{2}\right| = \left|\delta x - \frac{\delta^2}{4}\right| = \delta \left|x - \frac{\delta}{4}\right| < 1 \tag{23}$$

$$\implies \left|x - \frac{\delta}{4}\right| < \frac{1}{\delta} \implies x - \frac{\delta}{4} < \frac{1}{\delta} \implies x < \frac{1}{\delta} + \frac{\delta}{4} \tag{24}$$

a contradiction to our choice of x. Therefore $f(x) = x^2$ is not uniformly continuous on $\mathbb{R}$. $\square$

Recall. The closure of a nonempty set A, denoted by $\overline{A}$, is the set A together with all its limit points. A point p is a limit point of A if for every $\varepsilon > 0$ the epsilon neighborhood $N_\varepsilon(p)$ intersects A at some point other than p.

Definition 4.6. (Connected/Separated Sets) Two subsets A, B of a metric space X are said to be *separated* if both $A \cap \overline{B}$ and $\overline{A} \cap B$ are empty, i.e., if no point of A lies in the closure of B and no point of B lies in the closure of A.

A set $E \subset X$ is said to be *connected* if it is *not* the union of two nonempty separated sets.

4.1 Discontinuities

Definition 4.7. (Left/Right hand limits) Let f be defined on (a, b). Consider any point x such that $a \leq x < b$. We write

$$f(x+) = q \tag{25}$$

if $f(t_n) \to q$ as $n \to \infty$ for all sequences $\{t_n\}_{n=1}^{\infty} \subset (x, b)$ such that $t_n \to x$. This is the right hand limit. The left hand limit $f(x-)$ is defined similarly by considering sequences in (a, x).

Remark. Evidently, $\lim_{t \to x} f(t)$ exists if and only if

$$f(x+) = f(x-) = \lim_{t \to x} f(t) \tag{26}$$

Definition 4.8. (Simple Discontinuity) Let f be defined on (a, b). If f is discontinuous at a point x, and if $f(x+)$ and $f(x-)$ exist, then f is said to have a *discontinuity of the first kind* or more simply, a *simple discontinuity* (no pun intended).

Otherwise, we say f has a discontinuity of the *second kind*.

Example 4.9. Define f as

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{Q}^c \end{cases} \tag{27}$$

Then since $\mathbb{Q}$ is dense in $\mathbb{R}$ we see that $f(x+)$ and $f(x-)$ do not exist for any $x \in \mathbb{R}$ and therefore f has a discontinuity of the second kind at every point x.

Example 4.10. Define f as

$$f(x) = \begin{cases} x & x \in \mathbb{Q} \\ 0 & x \in \mathbb{Q}^c \end{cases} \tag{28}$$

Then f is continuous at 0 and has a discontinuity of the second kind everywhere else.

4.2 Monotonic Functions

Definition 4.11. Let $f : (a,b) \longrightarrow \mathbb{R}$. Then f is said to be *monotonically increasing* on (a,b) if for $x, y \in (a,b)$, $x < y$ implies $f(x) \leq f(y)$. We define monotonically decreasing similarly.

4.3 Infinite Limits and Limits at Infinity

Rather than going by with Rudin's treatment of this topic, we'll just use sequences because they are good and cool.

Definition 4.12. Let $f : E \longrightarrow \mathbb{R} \cup \{-\infty, \infty\}$ be a map from a subset $E \subset \mathbb{R}$ to the extended reals. Here are a bunch of definitions:

$$\lim_{x \to \pm\infty} f(x) = \lim_{n \to \infty} f(a_n) \tag{29}$$

where $\{a_n\}_{n=1}^{\infty} \subset \mathbb{R}$ is any sequence diverging to $\pm\infty$. Likewise,

$$\lim_{x \to p} f(x) = \pm\infty \iff \forall M \in \mathbb{R}\ \exists \delta_M > 0 \ni |x - p| < \delta_M \implies |f(x)| > M \tag{30}$$

where the latter equation can also be interpreted using sequences.

5 Differentiation

Definition 5.1. (Derivative of a real function) Let $f : [a,b] \longrightarrow \mathbb{R}$. Consider the limit

$$\lim_{t \to x} \frac{f(t) - f(x)}{t - x} \tag{31}$$

We define the derivative $f'(x)$ at a point $x \in [a,b]$ as the limit in (18) provided it exists. If $f'(x)$ exists at some point $x \in [a,b]$ then we say f is *differentiable* at x. If f' is defined at every point in $E \subset [a,b]$ then we say f is differentiable on E.

5.1 Mean Value Theorems

Definition 5.2. (Local maximum/minimum) Let f be a real function defined on a metric space X. We say that f has a *local maximum* at a point $p \in X$ if there exists a $\delta > 0$ such that $f(q) \leq f(p)$ for all $q \in X$ with $d_X(p,q) < \delta$. Local minima are defined likewise.

Problem 1. Let f be a real valued function on $[1, \infty)$, satisfying $f(1) = 1$ and $f'(x) = 1/\left(x^2 + f(x)^x\right)$. Then $\lim\limits_{x \to \infty} f(x)$ exists.

Proof. Consider $g : [1, \infty) \longrightarrow \mathbb{R}$, $x \mapsto -x^{-1} + 3$. We claim $\lim\limits_{x\to\infty} \frac{1}{x} = 0$. Let $\{x_n\} \to \infty$ be a divergent sequence in $\mathbb{R}$. Let $M_0 > 0$. We know there exists $M \in \mathbb{N}$ such that $M_0 < M$, so we pick a positive integer M which is greater than our arbitrary $M_0 > 0$. Then there exists $N \in \mathbb{N}$ such that $n \geq N \implies 0 < M < x_n$. Therefore by proposition 1.18 in Rudin we have that $\frac{1}{x_n} = f(x_n) < \frac{1}{M}$. Let $\varepsilon > 0$. Then we know by the Archimedean property we can pick M so that $\frac{1}{M} < \varepsilon$, and we know for $n \geq N$ that $0 < f(x_n) < \frac{1}{M} < \varepsilon$, hence $f(x_n) \to 0$ for any arbitrary sequence diverging to ∞, hence $f(x) \to 0$ as $x \to \infty$, and by theorem 4.4 in Rudin we know $\lim\limits_{x\to\infty} g(x) = \lim\limits_{x\to\infty} \frac{-1}{x} + \lim\limits_{x\to\infty} 3 = 0 + 3 = 3$. Hence by theorem 3.2 in Rudin we know $g(x)$ is bounded. We also know for all $x \in [1, \infty)$

$$g'(x) = \frac{1}{x^2} \geq \frac{1}{x^2 + f(x)^2} = f(x)$$

Notice that since $f'(x) = 1/\left(x^2 + f(x)^x\right)$, $f'(x) > 0$ for all $x \in [1, \infty)$, in particular, for all $x \in (1, \infty)$. Hence by theorem 5.11 in Rudin we know f is monotone increasing. It suffices to show $f(x)$ is bounded, which would imply $\lim\limits_{x\to\infty} f(x)$ exists as a real number. We claim that $f(x) < g(x)$ for all $x \in [1, \infty)$. Suppose for contradiction there exists $x_0 \in [1, \infty)$ such that $f(x_0) \geq g(x_0)$. We see that $g(1) = 2 > 1 = f(1)$, hence $x_0 \in (1, \infty)$. Consider $h(x) := g(x) - f(x)$. Then by theorem 5.3 in Rudin we know $h'(x) = g'(x) - f'(x)$. Now by the mean value theorem we know there exists $c \in (1, x_0)$ such that $h'(c) = \frac{h(x_0) - h(1)}{x_0 - 1}$. Notice that $g(x_0) \leq f(x_0)$ and $f(1) < g(1)$, hence $h(x_0) \leq 0$ and $0 < h(1)$. Since $0 < x_0 - 1$ this means $h'(c) < 0$. By construction of h we know this means $g'(c) < f'(c)$, a contradiction since we showed above that $f'(x) \leq g'(x)$ for all $x \in [1, \infty)$. Therefore for all $x \in [1, \infty)$ we know $f(x) < g(x)$. Since $g(x)$ is bounded this implies $f(x)$ is bounded (above). Since $f(x)$ is monotone increasing this implies $\lim\limits_{x\to\infty} f(x)$ exists. $\square$

Problem 2. Suppose $f : \mathbb{R} \longrightarrow \mathbb{R}$ is continuous, $f(1) > 0$, and $f(x)f(y) = f(x + y)$ for all $x, y \in \mathbb{R}$. Then there is a constant $c \in \mathbb{R}$ such that $f(x) = e^{cx}$.

Proof. We first claim that $f(x) > 0$ for all $x \in \mathbb{R}$. Suppose for contradiction there exists $x_0 \in \mathbb{R}$ such that $f(x_0) = 0$. Let $x \in \mathbb{R}$ be arbitrary. Then there exists $y \in \mathbb{R}$ such that $x_0 + y = x$, but that means $f(x) = f(x_0 + y) = f(x_0)f(y) = 0$. Therefore $f(x) = 0$ for all $x \in \mathbb{R}$, a contradiction to the assumption that $f(1) > 0$. Notice now that $x = \frac{x}{2} + \frac{x}{2}$. Hence $f(x) = f(\frac{x}{2} + \frac{x}{2}) = f(\frac{x}{2})^2 > 0$. Therefore for any $x \in \mathbb{R}$ we know $f(x) > 0$. Notice that $f(x) = f(0 + x) = f(0)f(x)$, hence $f(x) = f(0)f(x) \implies f(0) = 1$. Now let $n \in \mathbb{N}$. Then we know adding summing 1 to itself n times equals n, i.e., $1 + \cdots + 1 = n \implies f(n) = f(1 + \cdots + 1) = f(1)^n$. Hence for any natural number n we know $f(n) = f(1)^n$. For $m \in \mathbb{Z} \setminus \mathbb{N}$ notice that $f(m - m) = f(0) = 1 = f(m)f(-m) = f(m)f(1)^{-m}$, which means $f(m) = f(1)^m$, so for all integers $l \in \mathbb{Z}$ we see that $f(l) = f(1)^l$. Now let $\frac{p}{q} \in \mathbb{Q}$. Recall that $f(x)$ is always positive and $q \neq 0$, hence taking the q^{th} root is a well defined operation. Therefore, notice that by adding $\frac{p}{q}$ to itself q times we obtain

$$\sum_{i=1}^{q} \frac{p}{q} = \frac{pq}{q} = p \implies f\left(\sum_{i=1}^{q} \frac{p}{q}\right) = f(p) = f(1)^p$$

Hence

$$f\left(\frac{p}{q}\right) = f\left(\frac{p}{q}\right)^{q/q} = \prod_{i=1}^{q} f\left(\frac{p}{q}\right)^{1/q} = f\left(\sum_{i=1}^{q} \frac{p}{q}\right)^{1/q} = f(p)^{1/q} = f(1)^{p/q}$$

Therefore for any rational $\frac{p}{q}$ we know $f\left(\frac{p}{q}\right) = f(1)^{p/q}$. Now let $x \in \mathbb{R} \setminus \mathbb{Q}$. We know $\mathbb{Q}$ is dense in $\mathbb{R}$, hence there exists a sequence of rational numbers $\{q_n\}_{n=1}^{\infty} \to x$. Since f is assumed to be continuous we know this means $f(q_n) \to f(x)$. By definition we know for every $n \in \mathbb{N}$, $f(q_n) = f(1)^{q_n}$. Hence by the uniqueness of limits and by continuity of f,

$$\lim_{n\to\infty} f(1)^{q_n} = f(1)^x = \lim_{n\to\infty} f(q_n) = f(x)$$

Hence for any $x \in \mathbb{R}$ we know $f(x) = f(1)^x$. Consider

$$f(x) = e^{\log(f(x))} = e^{\log(f(1)^x)} = e^{x\log(f(1))} \implies f(x) = e^{x\log(f(1))}$$

Well we are given $f(1) > 0$, so $\log f(1) > 0$ is a well defined constant. Let $c := \log f(1)$. Then we see that $f(x) = e^{cx}$.

$\square$

Problem 3. If f is a differentiable function from the reals into the reals, $f'(x) > f(x)$ for all x, and $f(0) = 0$; then $f(x) > 0$ for all $x > 0$.

Proof. We first show that there exists a neighborhood $(0, \delta_0)$ such that $x \in (0, \delta_0) \implies f(x) > 0$. Given $f'(0) > f(0) = 0$, we know by the definition of the derivative that for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|x - 0| = |x| < \delta \implies \left| \frac{f(x) - f(0)}{x - 0} - f'(0) \right| = \left| \frac{f(x)}{x} - f'(0) \right| < \varepsilon$$

Let $\varepsilon = \frac{1}{2} f'(0)$, which we know is positive since $f'(0)$ is positive. Then there exists $\delta_0 > 0$ such that

$$|x| < \delta_0 \implies \frac{-1}{2} f'(0) < \frac{f(x)}{x} - f'(0) < \frac{1}{2} f'(0)$$

adding $f'(0)$ gives,

$$|x| < \delta_0 \implies \frac{1}{2} f'(0) < \frac{f(x)}{x} < \frac{3}{2} f'(0)$$

So our choice of $\varepsilon = \frac{1}{2} f'(0)$ gives us a $\delta_0 > 0$ such that $\frac{1}{2} f'(0) < \frac{f(x)}{x}$ for any x such that $|x| < \delta_0$. In particular, any $x \in (0, \delta_0)$ satisfies $|x| < \delta_0$. Consider arbitrary $x \in (0, \delta_0)$, which we know satisfies $x > 0$. Since $\frac{1}{2} f'(0)$ is positive this means $\frac{f(x)}{x}$ is positive. Therefore since x is positive, $f(x)$ must be positive. Now that we know there exists $\delta_0 > 0$ such that for any $x \in (0, \delta_0)$, $f(x) > 0$, suppose for contradiction there exists $x_0 > 0$ such that $f(x_0) \leq 0$. Consider the set $A = \{\delta > 0 : x \in (0, \delta) \implies f(x) > -\}$. We now know this set is nonempty and is bounded above by x_0, hence $\alpha = \sup A$ exists.

Now consider the set $B = \{x > 0 : f(x) \leq 0\}$. By assumption $x_0 \in B$, so B is nonempty. Furthermore we see that any $y \in (0, \delta_0)$ is a lower bound for B, so $\beta = \inf B$ exists. We claim that $\alpha = \beta$. Suppose not. Then if $\alpha > \beta$ there exists $\varepsilon > 0$ such that $\beta + \varepsilon = \alpha$, i.e., $\beta < \beta + \varepsilon/2 < \alpha$. But then $f(\beta + \varepsilon/2) \leq 0$ since $\beta = \inf B$, and at the same time since $\beta + \varepsilon/2 < \alpha = \sup A$ we know $f(\beta + \varepsilon/2) > 0$, a contradiction. On the other hand if $\alpha < \beta$ we know there exists $\varepsilon > 0$ such that $\alpha + \varepsilon = \beta$, so $\alpha < \alpha + \varepsilon/2 < \beta$. Then $f(\alpha + \varepsilon/2) > 0$ since $\alpha + \varepsilon/2 < \beta = \inf B$. But since $\alpha = \sup A$ this means there is no $\delta \in A$ such that $\alpha + \varepsilon/2 \in (0, \delta)$, which means there is some $\gamma > 0$ such that $\alpha + \gamma \in (\alpha, \alpha + \varepsilon/2)$ with $f(\alpha + \gamma) \leq 0$, but this is impossible since $\alpha + \gamma < \beta = \inf B = \inf\{x > 0 : f(x) \leq 0\}$. Therefore $\alpha = \beta$.

We claim that $\beta \in B$. Suppose for contradiction $\beta \notin B$. Then by definition of B we know $f(\beta) > 0$. If $\beta \notin A$ then there exists some $x \in (0, \beta)$ such that $f(x) \leq 0$, contradicting the fact that $\beta = \inf B$, so we know $\beta \in A$. Since $f(\beta) > 0$ and f is continuous we know for $n \in \mathbb{N}$ that there exists $\delta > 0$ such that

$$|x - \beta| < \delta \implies |f(x) - f(\beta)| < \frac{f(\beta)}{n} \implies f(\beta) - \frac{f(\beta)}{n} < f(x) < f(\beta) + \frac{f(\beta)}{n}$$

Since $f(\beta) - \frac{f(\beta)}{n} > 0$ this means

$$|x - \beta| < \delta \iff \beta - \delta < x < \beta + \delta \implies f(x) > 0$$

In particular, since $\beta \in A$, we now know $x \in (0, \beta + \delta) \implies f(x) > 0$, but this contradicts the fact that $\beta = \alpha = \sup A$. Hence $\beta \in B$. Therefore $f(\beta) \leq 0$. Let $\delta_0 \in A$ and let $a \in (0, \delta_0)$ By the mean value theorem we know there exists $c \in (a, \beta)$ such that $f'(c) = \frac{f(\beta) - f(a)}{\beta - a}$. But since $a \in (0, \delta_0)$ we know $f(a) > 0$, and we also know $f(\beta) \leq 0$ and $\beta > a$, hence $f'(c) = \frac{f(\beta) - f(a)}{\beta - a} \leq 0$, but β being $\inf\{x > 0 : f(x) \leq 0\}$ means $f(c) > 0$ since $c < \beta$, so $f'(c) < f(c)$, a contradiction since by assumption $f(c) < f'(c)$. Therefore $f(x) > 0$ for all $x > 0$. $\square$

Problem 4. Let f be a function of a real variable with f' continuous. If there exist real L, M such that

$$\lim_{x \to \infty} f(x) = M \qquad \lim_{x \to \infty} f'(x) = L$$

then $L = 0$.

Proof. Fix $\varepsilon_0 > 0$, $\{x_n\} \to \infty$, and let $\{y_n\} = \{x_n + \varepsilon_0\}$. Since $\lim_{x\to\infty} f(x) = M$ there exists $R_n > 0$ such that $x_n, y_n \geq R_n \implies |f(x_n) - f(y_n)| < \varepsilon_0$. Since f is differentiable, by the mean value theorem for all $n \in \mathbb{N}$ there exists $c_n \in (x_n, y_n)$ such that

$$|f'(c_n)| = \frac{|f(y_n) - f(x_n)|}{|x_n - y_n|} < \frac{1}{n|x_n - y_n|} = \frac{1}{n} \cdot \frac{1}{\varepsilon_0}$$

Since ε_0 is fixed, we know $\frac{1}{n} \cdot \frac{1}{\varepsilon_0} \to 0$ as $n \to \infty$. Therefore by the squeeze theorem,

$$0 < |f'(c_n)| < \frac{1}{n} \cdot \frac{1}{\varepsilon_0} \; \forall n \in \mathbb{N} \implies \lim n \to \infty |f'(c_n)| \to 0$$

Therefore by continuity of $f'(x)$ we know for all sequences $\{a_n\}$ diverging to ∞,

$$\lim_{n\to\infty} |f'(a_n)| = 0 \implies \lim_{x\to\infty} |f'(x)| = 0 \implies \lim_{x\to\infty} f'(x) = 0.$$

Where the last implication follows from the fact that for any $\varepsilon > 0$ there exists $M > 0$ such that $x > M \implies \big||f'(x)| - 0\big| < \varepsilon \iff |f'(x) - 0| < \varepsilon$. Therefore $L = 0$. $\square$

Problem 5. If $f : \mathbb{R} \longrightarrow \mathbb{R}^n$ and $|f(t)| = 1$ for all $t \in \mathbb{R}$, then the scalar product $\langle f(t), f'(t)\rangle = 0$ for all $t \in \mathbb{R}$.

Proof. Let $t \in \mathbb{R}$ and let $f(t) = (f_1(t), \ldots, f_n(t))$, where $f_i : \mathbb{R} \longrightarrow \mathbb{R}$ for all $i \in \{1, \ldots, n\}$. Recall the Euclidean norm is defined so that $|f(t)| = \sqrt{f_1(t)^2 + \cdots + f_n(t)^2} = 1$. Since $1^2 = 1$ this means $f_1(t)^2 + \cdots + f_n(t)^2 = 1$. By assumption, $f'(t)$ exists, hence we can differentiate the equality with respect to t to obtain

$$2f_1(t)f_1'(t) + \cdots + 2f_n(t)f_n'(t) = 0 \implies f_1(t)f_1'(t) + \cdots + f_n(t)f_n'(t) = 0 \iff \big\langle f(t), f'(t)\big\rangle = 0.$$

$\square$

Problem 6. The sum

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} e^{t/n} \tag{32}$$

defines a continuously differentiable function on all real numbers.

Proof. We first claim that (1) is a convergent series. The alternating series test gives a natural proof of this claim. Fix $t_0 \in \mathbb{R}$ and notice that for all $n \in \mathbb{N}$ we have $\left|\frac{(-1)^n}{n} e^{t_0/n}\right| = \frac{e^{t_0/n}}{n}$, $\frac{1}{n} > \frac{1}{n+1}$, and finally by theorem 8.6 in Rudin we know e^x is a strictly increasing positive function of x, hence

$$\frac{e^{t_0/n}}{n} \geq \frac{e^{t_0/(n+1)}}{n+1} \qquad \forall n \in \mathbb{N}$$

Furthermore, with the risk of being overly pedantic, the alternating series test in theorem 3.43 of Rudin requires a re-indexing of our series. Notice that

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} e^{t_0/n} = \sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{k+1} e^{t_0/(k+1)}$$

With this new indexing, we see that for terms of the form $2m - 1$ where $m \in \mathbb{N}$,

$$\frac{(-1)^{(2m-1)+1}}{(2m-1)+1} e^{\frac{t_0}{(2m-1)+1}} = \frac{1}{2m} e^{t_0/(2m)} > 0$$

And for terms of the form $2m$,

$$\frac{(-1)^{2m+1}}{2m+1} e^{t_0/(2m+1)} = \frac{-1}{2m+1} e^{t_0/(2m+1)} < 0$$

Let $c_k = \frac{(-1)^{k+1}}{k+1} e^{t_0/(k+1)}$. Then we just showed that $c_{2m-1} > 0$ and $c_{2m} < 0$. For the sake of simplicity we revert back to the original indexing in (1). The final requirement for the series to converge is for $\lim_{n\to\infty} c_n = 0$. We know that $\lim_{n\to\infty} \frac{1}{n} = 0$, and we also proved in class that e^x is a continuous function on $\mathbb{R}$. Hence by theorem 3.3 in Rudin and by continuity of e^x we know $\lim_{n\to\infty} e^{t_0/n} = e^0 = 1$. Now let $\varepsilon > 0$. Then we know by the Archimedean property there exists $N \in \mathbb{N}$ such that $n \geq N$ implies $\frac{1}{n} < \varepsilon$. Since $\left|\frac{(-1)^n}{n}\right| = \frac{1}{n} < \varepsilon$, by theorem 3.3 in Rudin we see that $\lim_{n\to\infty} \frac{(-1)^n}{n} e^{t_0/n} = 0$. Therefore by theorem 3.43 (the alternating series test) we know that (1) converges.

Now that we know (1) converges, define $f(t) = (1)$ and define $f_n(t) = \frac{(-1)^n}{n} e^{t/n}$, so $f(t) = \sum_{n=1}^{\infty} f_n(t)$. Let $[a, b] \subset \mathbb{R}$ be an arbitrary closed interval. Define $M_n := \frac{1}{n^2} e^b$. Then by theorem 3.28 and theorem 3.3 in Rudin we know $\sum_{n=1}^{\infty} M_n$ converges. Since e^x is strictly increasing we know for $t \in [a, b]$ that $|f_n'(t)| = \frac{1}{n^2} e^{t/n} \leq \frac{1}{n^2} e^b$. Therefore by theorem 7.10 in Rudin we know $\sum_{n=1}^{\infty} f_n'(t)$ converges uniformly on $[a, b]$. Since $[a, b]$ is an arbitrary closed interval this means the series converges uniformly on $\mathbb{R}$.

Now define the sequences of partial sums $\{s_k\} := \sum_{n=1}^{k} f_n(t)$ and $\{s_k'\} := \sum_{n=1}^{k} f_n'(t)$. Then we know since $\sum_{n=1}^{\infty} f_n'(t)$ converges uniformly that $\{s_k'\}$ converges uniformly. Recall that we proved $f(t_0)$ converges for any fixed $t_0 \in \mathbb{R}$, then by theorem 7.17 in Rudin we know $\{s_k\}$ is uniformly convergent and more importantly

$$f'(t) = \lim_{k\to\infty} s_k'(t)$$

Hence $f'(t)$ exists on $\mathbb{R}$, implying $f(t)$ is continuous on $\mathbb{R}$. By theorem 7.12 in Rudin, since $\{s_k'\}$ is uniformly convergent we know $f'(t)$ is continuous. □

Problem 7. The polynomial

$$P_n(x) = 1 + \frac{x^1}{1!} + \cdots + \frac{x^n}{n!}$$

has exactly 1 zero if n is odd and none if n is even.

Lemma 5.3. If $n \in \mathbb{N}$ is odd and $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ is a polynomial of degree n such that $a_i \in \mathbb{R}^+$, then $p(x_0) \leq 0$ for some $x_0 \in \mathbb{R}$.

Proof of lemma. Suppose for contradiction that $p(x) = a_n x^n + \cdots + a_0 > 0$ for all $x \in \mathbb{R}$. Then we know

$$-(a_{n-1} x^{n-1} + \cdots + a_0) < a_n x^n$$

Since $a_n > 0$ this implies

$$-\frac{a_{n-1}}{a_n} x^{n-1} - \frac{a_{n-2}}{a_n} x^{n-2} \cdots - \frac{a_0}{a_n} < x^n$$

If $|x| > 1$, then $-x^{n-1} < -x - x^2 - \cdots - x^{n-1}$, which means

$$-\frac{a_{n-1}}{a_n} x^{n-1} - \frac{a_{n-2}}{a_n} x^{n-1} - \cdots - \frac{a_0}{a_n} x^{n-1} < -\frac{a_{n-1}}{a_n} x^{n-1} - \frac{a_{n-2}}{a_n} x^{n-2} \cdots - \frac{a_0}{a_n} < x^n$$

Let $c = \max\{\frac{a_j}{a_n}, 1.5 : 0 \leq j \leq n-1\}$, which we know is greater than 1. Then we have that

$$-ncx^{n-1} < -\frac{a_{n-1}}{a_n} x^{n-1} - \frac{a_{n-2}}{a_n} x^{n-1} - \cdots - \frac{a_0}{a_n} x^{n-1} < x^n$$

So $x^n > -ncx^{n-1}$ for $|x| > 1$. Let $x = -nc$. Then $x^n = (-nc)^n = -(nc)^n$, but also $-ncx^{n-1} = -(nc)^n$, so $x^n = -ncx^{n-1}$, but we just showed $x^n > -ncx^{n-1}$, a contradiction, therefore there must exists $x_0 \in \mathbb{R}$ so that $p(x_0) \leq 0$. □

Proof of main problem. We induct on n, separating into cases of even and odd parity. Even base case: $n = 0$, then $P_0(x) = 1$, which has no zeroes. Odd base case: $n = 1$. Then $P_1(x) = 1 + x$, which has 1 real zero at $x = -1$.

Induction Hypothesis: suppose $P_n(x)$ has no real zeroes for even n and only one real zero for odd n. Case 1: n is even. Then $n+1$ is odd, so $P_{n+1}(x)$ is odd. Notice that

$$P'_{n+1}(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} = P_n(x)$$

Since n is even, by hypothesis $P'_{n+1}(x) = P_n(x)$ has no real zeroes, therefore by the intermediate value theorem we know $P_n(x) < 0$ or $P_n(x) > 0$ for all $x \in \mathbb{R}$. Since $P_n(0) = 1 > 0$, we know $P_n(x) > 0$ for all $x \in \mathbb{R}$. Therefore we know $P_{n+1}(x)$ is strictly increasing. Since $n+1$ is odd we apply lemma 0.1 to show there exists at least one $x_0 \in \mathbb{R}$ so that $P_{n+1}(x_0) \leq 0$. Since we know $P_{n+1}(x)$ is strictly increasing this means there is no more than one zero.

Case 2: n is odd, i.e. $n+1$ is even. Then we know $P'_{n+1}(x) = P_n(x)$ has exactly point $x_0 \in \mathbb{R}$ such that $P'_{n+1}(x_0) = 0$, and we also know $P'_{n+1}(x)$ is strictly increasing, so for $x \in (-\infty, x_0)$, $P_n(x) = P'_{n+1}(x) < 0$, which means $P_{n+1}(x)$ is decreasing on $(-\infty, x_0)$, and for $x \in (x_0, \infty)$ we know $P_n(x) = P'_{n+1}(x) > 0$, so $P_{n+1}(x)$ is increasing on $x \in (x_0, \infty)$. Since $P_n(x_0) = P'_{n+1}(x_0) = 0$, we know $P_{n+1}(x_0)$ is a local minimum. Since $n-1$ is even, by hypothesis $P''_{n+1}(x) = P'_n(x) = P_{n-1}(x)$ is always positive, hence $P_{n+1}(x)$ is concave upward, implying $P_{n+1}(x_0)$ is in fact a *global* minimum since by hypothesis there exists no other critical points (i.e. zeroes of $P_n(x)$). Since $P_n(0) = 1$, we know $x_0 \neq 0$, and since $n+1$ is even, $P_{n+1}(x_0) = P_n(x_0) + \frac{1}{(n+1)!}x^{n+1} > 0$, which means the global minimum $P_{n+1}(x_0)$ is positive, hence $P_{n+1}(x) > 0$ for all $x \in \mathbb{R}$. This concludes the induction. □

Problem 8. If f is a twice-differentiable real valued function on the real line such that $|f(x)| \leq 1$ and $|f''(x)| \leq 1$ for all $x \in \mathbb{R}$, then $|f'(x)| < \sqrt{2}$ for all $x \in \mathbb{R}$.

Proof. Suppose for contradiction that $|f'(x_0)| \geq \sqrt{2}$ for some $x_0 \in \mathbb{R}$. Recall the fundamental theorem of calculus: if $f'(x)$ exists on some interval $[a,b] \subset \mathbb{R}$ with $f'(x)$ bounded and integrable, then

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

Since $f(x)$ is twice differentiable on the real line we know $f'(x)$ is integrable. If f' were not bounded then $|f(x)| > 1$ for some $x \in \mathbb{R}$, so $f'(x)$ is bounded. Therefore, for some interval $[0, 2m]$

$$\int_0^{2m} f'(x)dx = f(2m) - f(0) \leq 2$$

□

Problem 9. Let f be a differential real valued function on $[0,1]$ such that $f(0) = 0$ and $0 \leq f'(x) \leq 2f(x)$ for all $x \in [0,1]$. Then $f(x) = 0$ for all $x \in [0,1]$.

Proof. Consider f on the subset $[0, 1/3]$. Since f is differentiable we know by the mean value theorem there exists $c_0 \in (0, 1/3)$ such that $f'(c_0) = \frac{f(1/3)-f(0)}{1/3}$, since $f(0) = 0$, we know this means $\frac{1}{3}f'(c_0) = f(1/3)$, or $f'(c_0) = 3f(1/3)$. Since $0 \leq f'(x)$ for all $x \in [0,1]$ we know $f(x)$ is monotonically increasing, so $0 \leq 3f(c_0) \leq 3f(1/3) = f'(c_0)$, but by assumption we have
$f'(c_0) \leq 2f(c_0)$, so $3f(c_0) \leq f'(c_0) \leq 2f(c_0)$, so $3f(1/3) \leq 2f(1/3)$, which in turn implies $f(1/3) = 0$ since $0 \leq f(x)$ for all $x \in [0,1]$ and f is monotonically increasing. Again, by monotinicity and nonnegativity, since $f(1/3) = 0$ we deduce that $f(x) = 0$ for all $x \in [0, 1/3]$.

Next consider $[1/3, 2/3]$. By the mean value theorem there exists $c_1 \in (1/3, 2/3)$ so that
$f'(c_1) = \frac{f(2/3)-f(1/3)}{1/3}$, which means $f'(c_1) = 3f(2/3)$. Since $f'(x) \leq 2f(x)$ for all x we again see that $f'(c_1) = 3f(2/3) \leq 2f(c_1) \leq 2f(2/3)$, so $3f(2/3) \leq 2f(2/3)$. By monotonicity and nonnegativity of f this again shows that $f'(c_1) = f(c_1) = 0$ and $f(2/3) = 0$.

Finally consider $[2/3, 1]$. The mean value theorem again shows us that there exists $c_2 \in (2/3, 1)$ such that $f'(c_2) = \frac{f(1)-f(2/3)}{1/3} = 3f(1)$. By assumption, $0 \leq f'(c_2) = 3f(1) \leq 2f(c_2) \leq 2f(1)$
$\implies 3f(1) \leq 2f(1)$ which by the same arguments as above means $f(1) = 0$. Since $f(1) = 0$ and f is nonnegative and monotone increasing on $[0,1]$ this means $f(x) = 0$ for all $x \in [0,1]$. □

Problem 10. If f is a functions from the reals to the reals satisfying $2f(x) = f(2x)$ for all x and f is differentiable at 0, then f is linear.

Claim 3. $f(\frac{x}{2^n}) = \frac{1}{2^n}f(x)$ for all $n \in \mathbb{N}$.

Proof of claim. We induct on n. Base case: n = 1. We have $f(x) = f\left(2 \cdot \frac{x}{2}\right) = 2f\left(\frac{x}{2}\right)$. Dividing by 2 yields $\frac{1}{2}f(x) = f\left(\frac{x}{2}\right)$ as desired.

Induction hypothesis: Suppose $f\left(\frac{x}{2^{n-1}}\right) = \frac{1}{2^{n-1}}f(x)$ for some $n-1 \in \mathbb{N}$. Then since $\frac{x}{2^{n-1}} \in \mathbb{R}$ and we proved the base case we know

$$f\left(\frac{x}{2^n}\right) = f\left(\frac{x}{2^{n-1}} \cdot \frac{1}{2}\right) = \frac{1}{2^{n-1}}f(\frac{x}{2}) = \frac{1}{2^n}f(x)$$

This concludes the proof. □

Proof of problem. Notice that since $f(0) = f(2 \cdot 0) = 2f(0)$ then $2f(0) = f(0) \implies f(0) = 0$. Fix $x_0 \neq \mathbb{R}$ and set $y_0 := f(x_0)$. Furthermore, we know $\lim\limits_{n\to\infty} \frac{1}{2^n} = 0$. Then since f is continuous and differentiable at 0 we know

$$f'(0) = \lim_{n\to\infty} \frac{f(\frac{1}{2^n}x_0) - f(0)}{\frac{1}{2^n}x_0 - 0} = \lim_{n\to\infty} \frac{\frac{1}{2^n}f(x_0)}{\frac{1}{2^n}x_0} = \frac{y_0}{x_0}$$

We claim that $f(x) = \frac{y_0}{x_0}x$ for all $x \in \mathbb{R}$. Suppose for contradiction there exists $x_1 \in \mathbb{R}$ such that $f(x_1) \neq \frac{x_0}{y_0}x_1$. Set $y_1 := f(x_1)$. Then since $\frac{y_0}{x_0}x_1 \neq y_1$ we know $\frac{y_1}{x_1} \neq \frac{y_0}{x_0}$. Notice that we can differentiate f again:

$$f'(0) = \lim_{n\to\infty} \frac{f(\frac{1}{2^n}x_1) - f(0)}{\frac{1}{2^n}x_1 - 0} = \lim_{x\to\infty} \frac{\frac{1}{2^n}f(x_1)}{\frac{1}{2^n}x_1} = \frac{y_1}{x_1} \neq \frac{y_0}{x_0}$$

But this is impossible since f is given to be differentiable at 0, hence its derivative (i.e. the limit evaluated above) must be uniquely defined. Therefore $f(x) = \frac{y_0}{x_0}x$ for all $x \in \mathbb{R}$, so $f(x)$ is a linear function. □

Problem 11. Every differentiable function $f : \mathbb{R} \longrightarrow \mathbb{R}$ with the property that $(2h)f'(x) = f(x+h) - f(x-h)$ for all $x, h \in \mathbb{R}$ has the form $f(x) = ax^2 + bx + c$.

Claim 4. For fixed h, $2hf^{(n)}(x)$ exists in the form $f^{(n-1)}(x+h) - f^{(n-1)}(x-h)$ for all $n \in \mathbb{N}$.

Proof of claim. We induct on n. Base case: n = 1. We are given that $(2h)f'(x) = f(x+h) - f(x-h)$, so $f'(x) = \frac{f(x+h)-f(x-h)}{2h}$.

Induction hypothesis: suppose $2hf^{(n-1)}(x) = f^{(n-2)}(x+h) - f^{(n-2)}(x-h)$. Then we know $\frac{d}{dx}\left(f^{(n-2)}(x+h) - f^{(n-2)}(x\right.$
$f^{(n-1)}(x+h) - f^{(n-1)}(x-h) = 2hf^{(n)}(x)$. Therefore $f^{(n)}$ exists for all $n \in \mathbb{N}$. In particular, this shows that f is infinitely differentiable for fixed h and by theorem 5.2 in Rudin we know f is smooth. □

Proof of problem. Fix x and consider $\frac{d}{dh}(2hf'(x)) = 2f'(x) = f'(x+h) + f'(x-h)$. Differentiating with respect to h again leads to $0 = f''(x+h) - f''(x-h)$, and $0 = f'''(x+h) + f'''(x-h) \implies f'''(x+h) = -f'''(x-h)$. Since f is smooth we know f''' is continuous on its domain, hence $\lim\limits_{h\to 0} f'''(x+h) = \lim\limits_{h\to 0} -f'''(x-h) \implies f'''(x) = -f'''(x) \implies f'''(x) = 0$. Therefore derivatives of order greater than or equal to 3 vanish.

Now let $x \in \mathbb{R}$. Let $[-b, b] \in \mathbb{R}$ be some closed interval containing 0 and let $\alpha = 0$ and β be distinct points in $[-b, b]$. Then by Taylor's theorem we know there exists some point $x_0 \in (\alpha, \beta)$ (or $x_0 \in (\beta, \alpha)$, whichever makes sense) such that

$$\sum_{k=0}^{2} \frac{f^{(k)}(\alpha)}{k!}(\beta)^k + \frac{f^{(3)}(x_0)}{3!}(\beta)^3 = \sum_{k=0}^{2} \frac{f^{(k)}(0)}{k!}(\beta)^k + 0 = f(\beta)$$

Since $[-b, b]$ is an arbitrary interval, we can set $\beta := x \neq 0$ and we have an equation for $f(x)$:

$$f(x) = \sum_{k=0}^{2} \frac{f^{(k)}(0)}{k!}(x)^k$$

But this sum is just a two degree polynomial in x, i.e. $f(x) = ax^2 + bx + c$ for $x \neq 0$ and for some constants $a, b, c \in \mathbb{R}$. It suffices to show $f(x)$ exists at $x = 0$. Continuity of f, continuity of polynomials in x, and the existence of $f^{(n)}(0)$ implies

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} \sum_{k=0}^{2} \frac{f^{(k)}(0)}{k!}(x)^k \lim_{x\to 0} ax^2 + bx + c = c$$

Hence $f(x)$ is of the form $ax^2 + bx + c$ for all $x \in \mathbb{R}$ as desired. □

Problem 12. If $\{f_j\}_{j=1}^{\infty}$ and g are twice differentiable functions such that, for each $x \in \mathbb{R}$, $\lim_{j\to\infty} f_j(x) = g(x)$ and $f_j''(x) \geq 0$, then $g''(x) \geq 0$.

Proof. Let $(a, b) \subset \mathbb{R}$ be an arbitrary open interval. By exercise 23 in chapter 4 of Rudin (We are just using this to cite the definition of a convex function), a real valued function f is said to be convex on (a, b) if for all $\lambda \in (0, 1)$ and $x, y \in (a, b)$

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

By exercise 14 in chapter 5 of Rudin, $f''(x) \geq 0$ for all x is equivalent to f being convex. It suffices to prove g is convex. Suppose for contradiction that g is not convex. So there exists $x_0, y_0 \in (a, b)$ and $\lambda_0 \in (0, 1)$ such that

$$g(\lambda_0 x_0 + (1 - \lambda_0)y_0) > \lambda_0 g(x_0) + (1 - \lambda_0)g(y_0)$$

Let $L = g(\lambda_0 x_0 + (1-\lambda_0)y_0) - \lambda_0 g(x_0) - (1-\lambda_0)g(y_0)$, which we know is greater than 0. Since $\{f_j\}_{j=1}^{\infty}$ converges pointwise to g we know for each $x \in \mathbb{R}$ there exists $N_x \in \mathbb{N}$ such that $n \geq N_x \implies |f_n(x) - g(x)| < L/6$. Set $z_0 := \lambda_0 x_0 + (1 - \lambda_0)y_0$ and pick $N = \max\{N_{x_0}, N_{y_0}, N_{z_0}\}$. Then by the triangle inequality, $n \geq N$ implies

$$|g(z_0) - f_n(z_0) - \lambda_0 g(x_0) + \lambda_0 f_n(x_0) - (1 - \lambda_0)g(y_0) + (1 - \lambda_0)f_n(y_0)|$$

$$\begin{aligned} &\leq |g(z_0) - f)n(z_0)| + |\lambda_0 f_n(x_0) - \lambda_0 g(x_0)| + |(1 - \lambda_0)f_n(y_0) - (1 - \lambda_0)g(y_0)| \\ &< \frac{L}{6} + \lambda_0 \frac{L}{6} + (1 - \lambda_0)\frac{L}{6} < L \end{aligned}$$

where the final inequality follows from the fact that $\lambda_0 \in (0, 1)$. Therefore

$$g(z_0) - \lambda_0 g(x_0) - (1 - \lambda_0)g(y_0) - f_n(z_0) + \lambda_0 f_n(x_0) + (1 - \lambda_0)f_n(y_0) < g(z_0) - \lambda_0 g(x_0) - (1 - \lambda_0)g(y_0)$$

$$\implies 0 < f_n(z_0) - \lambda_0 f_n(x_0) - (1 - \lambda_0)f_n(y_0)$$

i.e.,

$$f_n(z_0) = f_n(\lambda_0 x_0 + (1 - \lambda_0)y_0) > \lambda_0 f_n(x_0) + (1 - \lambda_0)f_n(y_0)$$

Which implies f_n is not convex for $n \geq N$, but this contradicts the fact that $f_j''(x) \geq 0$ for all $x \in \mathbb{R}$ and for all $j \in \mathbb{N}$, so g must be convex.

□

6 The Riemann-Stieltjes Integral

6.1 Definition and Existence of the Integral

Definition 6.1. (Partition) Let $[a, b]$ be a given interval. By a *partition* P of $[a, b]$ we mean a finite set of points $x_0, x_1, \ldots, x_n$, where

$$a = x_0 \leq x_1 \leq \cdots \leq x_{n-1} \leq x_n = b.$$

We write

$$\Delta x_i = x_i - x_{i-1} \qquad i \in \{1, \ldots, n\}$$

Definition 6.2. (Upper/Lower Sums) Suppose f is a bounded real function defined on $[a, b]$. Corresponding to each partition P of $[a, b]$ we define

$$U(P, f) = \sum_{k=1}^{n} \sup \{f(x) \ : \ x \in [x_{i-1}, x_i]\} \, \Delta x_i$$

$$L(P, f) = \sum_{k=1}^{n} \inf\{f(x) \ : \ x \in [x_{i-1}, x_i]\}\Delta x_i$$

as the respective upper and lower sums of f on $[a, b]$ corresponding to a partition P.

Definition 6.3. (Upper/Lower/Riemann Integral) We define the upper Riemann integral as

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

and the lower Riemann integral as

$$\underline{\int_a^b} f dx = \sup_P \{L(P, f)\}$$

where the inf and sup are taken over all partitions of the interval $[a, b]$.

If the upper and lower Riemann integrals exist and are equal, we say that f is *Riemann integrable* on $[a, b]$, we write $f \in \mathscr{R}$ (i.e., $\mathscr{R}$ denotes the set of Riemann integrable functions), and we denote the common value as

$$\overline{\int_a^b} f dx = \underline{\int_a^b} f dx = \int_a^b f dx$$

Definition 6.4. (Stieltjes Integral) Let α be a monotonically increasing function on $a[a, b]$. Since $\alpha(a)$ and $\alpha(b)$ are finite, it follows that α is bounded on $[a, b]$. Corresponding to each partition P of $[a, b]$ we write

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) \qquad i \in \{1, \ldots, n\}$$

Since α is monotone increasing we know that $\Delta\alpha_i \geq 0$ for every i. For any real function f which is bounded on $[a, b]$ we put

$$U(P, f, \alpha) = \sum_{k=1}^{n} \sup\{f(x) \ : \ x \in [x_{i-1}, x_i]\}\Delta\alpha_i$$

$$L(P, f, \alpha) = \sum_{k=1}^{n} \inf\{f(x) \ : \ x \in [x_{i-1}, x_i]\}\Delta\alpha_i$$

and define

$$\overline{\int_a^b} f d\alpha = \inf_P \{U(P, f, \alpha)\}$$

$$\underline{\int_a^b} f d\alpha = \sup_P \{L(P, f, \alpha)\}$$

If the upper and lower integrals are equal, we denote their common value by

$$\int_a^b f d\alpha$$

This is the *Riemann-Stieltjes integral* (or simply the *Stieltjes integral*) of f with respect to α over $[a, b]$.

If the integral above exists, i.e., if the upper and lower integrals are equal, then we say that f is integrable with respect to α, in the Riemann sense, and write $f \in \mathscr{R}(\alpha)$.

Remark. By taking $\alpha(x) = x$, we see that the Riemann integral is a special case of the Stieltjes integral. Note that by our definition, $\alpha(x)$ need not be continuous.

Definition 6.5. (Common Refinement/Finer/Coarser) Let P be a partition of $[a, b]$. We say that the partition P^* is a *refinement* of P if $P \subset P^*$. If $P \subsetneq P^*$ we say P^* is a *finer* partition with respect to P or conversely P is a *coarser* partition with respect to P^*.

Given two partitions P_1, P_2, we say that P^* is their *common refinement* if $P^* = P_1 \cup P_2$.

Definition 6.6. (Measure Zero) A set $A \subset \mathbb{R}$ has Lebesgue *measure zero* if for every $\varepsilon > 0$ there exists a countable family of open intervals $\{U_j\}_{j=1}^{\infty}$ such that $A \subset \bigcup_{j=1}^{\infty} U_j$ and $\sum_{j=1}^{\infty} |U_j| < \varepsilon$, where $|U_j|$ denotes the length of the open interval.

Problem 1. If $f : [0, 1] \longrightarrow \mathbb{R}$ and

$$\lim_{x \to c} f(x) = L \in \mathbb{R} \tag{33}$$

exists for all $c \in [a, b] \subset [0, 1]$, then f is Riemann integrable on $[a, b]$.

Claim 5. f is bounded on $[a, b]$.

Proof of claim. Given (1), we know for all $c \in [a, b]$ there exists $\delta_c > 0$ such that $|x - c| < \delta_c$ implies $|f(x) - L| < 1$. By the triangle inequality, we see that

$$|f(x)| = |f(x) - L + L| \leq |f(x) - L| + |L| < 1 + |L| \quad \forall x \in N_{\delta_c}(c)$$

Therefore for all $c \in [a, b]$, $f(x)$ is bounded on $N_{\delta_c(c)} = (\delta_c - c, \delta_c + c)$.

Note that $[a, b] \subset \bigcup_{c \in [a,b]} N_{\delta_c}(c)$. By compactness of $[a, b]$ we know there exists a finite subset $\{c_1, \ldots, c_n\}$ of $[a, b]$ such that

$$[a, b] \subset \bigcup_{i=1}^{n} N_{\delta_{c_i}}(c_i)$$

We showed above that for each $c_i \in \{c_1, \ldots, c_n\}$ f is bounded on $N_{\delta_{c_i}}(c_i)$, i.e. there exists $M_i \in \mathbb{R}$ such that for all $x \in N_{\delta_{c_i}}(c_i)$ $|f(x)| \leq M_i$. Since there are finitely many M_i, let $M = \max\{M_i : 1 \leq i \leq n\}$. Then for all $i \in \{1, \ldots, n\}$ and for all $x \in N_{\delta_{c_i}}(c_i)$, $|f(x)| \leq M$, hence $|f(x)| \leq M$ for all $x \in \bigcup_{i=1}^{n} N_{\delta_{c_i}}(c_i)$. Since $[a, b] \subset \bigcup_{i=1}^{n} N_{\delta_{c_i}}(c_i)$, this implies f is bounded on $[a, b]$. $\square$

Claim 6. $\lim\limits_{x \to c} = L \in \mathbb{R}$ implies for all $\varepsilon/2 > 0$ there exists a neighborhood around c such that for all x, y in the neighborhood, $|f(x) - f(y)| < \varepsilon/2$.

Proof. Note that $\lim\limits_{x \to c} f(x) = L \in \mathbb{R}$ implies for $\varepsilon/4 > 0$ there exists $\delta_c > 0$ such that $|x - \delta_c|$, $|y - \delta_c| < 0$ implies $|f(x) - L|$, $|f(y) - L| < \varepsilon/4$, hence by the triangle inequality,

$$|f(x) - f(y)| = |f(x) - L + L - f(y)| \leq |f(x) - L| + |L - f(y)| = |f(x) - L| + |f(y) - L| < \varepsilon/4 + \varepsilon/4 = \varepsilon/2$$

Therefore for $x, y \in N_{\delta_c}(c) = (c - \delta_c, c + \delta_c)$, we know $|f(x) - f(y)| < \varepsilon/2$. $\square$

Proof of main problem. Let $\varepsilon' > 0$. By claim 2 we know for all $c \in [a, b]$ there exists $\delta_c/2 > 0$ such that $x, y \in N_{\delta_c/2}(c)$ implies $|f(x) - f(y)| < \varepsilon$. Notice that we have an open cover

$$[a, b] \subset \bigcup_{c \in [a,b]} N_{\delta_c/2}(c)$$

which we can make into a finite subcover by compactness:

$$[a, b] \subset \bigcup_{i=1}^{n-1} N_{\delta_{c_i}/2}(c_i)$$

By the total ordering of $\mathbb{R}$, we can rearrange the points $\{c_1, \ldots, c_{n-1}\}$ so that $c_1 \leq c_2 \leq \cdots \leq c_{n-1}$. Notice that $P = \{a = c_0, c_1, \ldots, c_{n-1}, c_n = b\}$ is a partition of $[a, b]$. Notice that for $i \in \{1, \ldots, n-1\}$ we have $\overline{N_{\delta_{c_i}/2}} \subset N_{\delta_{c_i}}(c_i)$. By **Claim 2** we know this implies

$$\sup\left\{f(x) : x \in \overline{N_{\delta_{c_i}/2}(c_i)}\right\} \in N_{\varepsilon/2}(L_i)$$

and

$$\inf\left\{f(x) : x \in \overline{N_{\delta_{c_i}/2}(c_i)}\right\} \in N_{\varepsilon/2}(L_i)$$

where $L_i = \lim_{x \to c_i} f(x)$. Since $[a, b] \subset [0, 1]$ we know $\sum_{i=1}^{n} c_i - c_{i-1} < 1$, therefore

$$\sum_{i=1}^{n}\left(\sup\left\{f(x) : x \in \overline{N_{\delta_{c_i}/2}(c_i)}\right\} - \inf\left\{f(x) : x \in \overline{N_{\delta_{c_i}/2}(c_i)}\right\}\right)(c_i - c_{i-1}) < \varepsilon$$

which we know is equivalent to

$$\sum_{i=1}^{n}\sup\left\{f(x) : x \in \overline{N_{\delta_{c_i}/2}(c_i)}(c_i - c_{i-1})\right\} - \sum_{i=1}^{n}\inf\left\{f(x) : x \in \overline{N_{\delta_{c_i}/2}(c_i)}\right\}(c_i - c_{i-1}) < \varepsilon$$

By theorem 6.6 in Rudin this implies f is Riemann integrable on $[a, b]$.

□

Problem 2. If $f : \mathbb{R} \longrightarrow \mathbb{R}$ is continuous of $[0, \infty)$ and

$$\lim_{x \to \infty}\left(f(x) + \int_0^x f(t)dt\right)$$

exists, then

$$\lim_{x \to \infty} xf(x) = 0$$

Counterexample. Consider $g(x) = \frac{\sin x}{x}$. We first claim that $\lim_{x \to 0} g(x) = 1$. L'hopital's rule gives an immediate proof. Notice that $\sin x$ and x are differentiable on $\mathbb{R}$, $\lim_{x \to 0} \sin x = 0$, and $\lim_{x \to 0} x = 0$. Furthermore, $\frac{d \sin x}{dx} = \cos x$ and $\frac{dx}{dx} = 1$. Therefore we know

$$\lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{\cos x}{1} = 1$$

Now define the function

$$f(x) = \begin{cases} \frac{\sin x}{x} & x \in (-\infty, 0) \cup (0, \infty) \\ 1 & x = 0 \end{cases}$$

Since x and $\sin x$ are well-defined and continuous on $(-\infty, 0) \cup (0, \infty)$ we know $f(x)$ is continuous on $(-\infty, 0) \cup (0, \infty)$. Since $\lim_{x \to 0} f(x) = f(0) = 1$ we know $f(x)$ is continuous at $x = 0$, therefore $f(x)$ is continuous on $[0, \infty)$ as desired.

We next claim that $\lim_{x \to \infty} f(x) = 0$. Notice that for all $x \in \mathbb{R}$, $-1 \leq \sin x \leq 1$, hence for all $x \in \mathbb{R} \setminus \{0\}$, $\frac{-1}{x} \leq \frac{\sin x}{x} \leq \frac{1}{x}$. Therefore by the squeeze lemma,

$$0 = \lim_{x \to \infty} \frac{-1}{x} \leq \lim_{x \to \infty} \frac{\sin x}{x} \leq \lim_{x \to \infty} \frac{1}{x} = 0 \implies \lim_{x \to \infty} \frac{\sin x}{x} = 0$$

In [1], Chen gives a rigorous treatment via real-analytic methods for the evaluation of

$$\lim_{x \to \infty} \int_0^x f(t)dt = \frac{\pi}{2}$$

Therefore we know

$$\lim_{x \to \infty}\left(f(x) + \int_0^x f(t)dt\right) = \frac{\pi}{2}$$

however, we also see that $\lim_{x \to \infty} xf(x) = \lim_{x \to \infty} \sin x$, which we know does not converge. □

Problem 3.

$$\int_0^1 x^{-x}dx = \sum_{n=1}^{\infty} n^{-n} \tag{34}$$

Proof. First note that $x^{-x} = e^{\ln x^{-x}} = e^{-x\ln x}$, which means the evaluating integral in (1) is equivalent to evaluating

$$\int_0^1 e^{-x\ln x}dx = \int_0^1 \sum_{k=0}^{\infty} \frac{(-x)^k(\ln x)^k}{k!}dx \tag{35}$$

We first claim for fixed $k \in \mathbb{N}$ that

$$\lim_{x\to 0} \frac{(-x)^k(\ln x)^k}{k!} = 0 \tag{36}$$

By theorem 8.6 in Rudin we know $\ln(x)$ is differentiable (hence continuous) on $\mathbb{R}^+$. We also know $e^x \to 0$ as $x \to -\infty$ by the same theorem. Therefore we define the sequence $\{e^{-n}\}_{n=1}^{\infty}$ and obtain the limit

$$\lim_{n\to\infty} (\ln e^{-n}) = \lim_{n\to\infty} (-n) = -\infty \tag{37}$$

Continuity of $\ln(x)$ therefore implies $\lim\limits_{x\to 0}(\ln x) = -\infty$. Furthermore, we know $\frac{1}{-x} \to -\infty$ as $x \to 0$. Then since we know $\frac{1}{-x}$ is differentiable on $(0,1)$, $\ln x \neq 0$ for all $x \in (0,1]$ and is differentiable on $(0,1)$, and since

$$\frac{-x(\ln x)}{k!} = \frac{(\ln x)}{\frac{k!}{-x}} \tag{38}$$

we can apply L'Hospital's rule to evaluate

$$\lim_{x\to 0} \frac{-x(\ln x)}{k!} = \frac{1}{k!}\lim_{x\to 0} \frac{\ln x}{\frac{1}{-x}} = \frac{1}{k!}\lim_{x\to 0} \frac{1/x}{1/x^2} = \frac{1}{k!}\lim_{x\to 0} x = 0 \tag{39}$$

Therefore by taking the k products we obtain

$$\lim_{x\to 0} \frac{(-x\ln x)^k}{k!} = \frac{1}{k!}\prod_{i=1}^{k} \lim_{x\to 0} \frac{\ln x}{\frac{1}{-x}} = 0 \tag{40}$$

Now that we have verified (3), define the function

$$f_k(x) = \begin{cases} \frac{(-x^k \ln x)^k}{k!} & x \in (0,1] \\ 0 & x = 0 \end{cases} \tag{41}$$

Where $k \in \mathbb{N}$. Notice that f_k is continuous on $[0,1]$ since we know $(-x)^k$ and $(\ln x)^k$ are continuous on $(0,1]$ and $\lim\limits_{x\to 0} f_k(x) = 0 = f(0)$ as proven above. Since $[0,1]$ is compact, theorems 4.15 and 4.16 in Rudin imply that f_k is bounded and there exists $p \in [0,1]$ such that $|f_k(p)| = |\sup f([0,1])| = M_k$ We claim that $\sum f_k$ is uniformly convergent in $[0,1]$. By theorem 7.10 it suffices to show $\sum M_k$ converges, and we've seen this series before!

$$\sum_{k=0}^{\infty} M_k = \sum_{k=0}^{\infty} \left|\frac{(-p\ln p)^k}{k!}\right| = \sum_{k=0}^{\infty} \frac{\left|(-p\ln p)^k\right|}{k!} = \sum_{k=0}^{\infty} \frac{|(p\ln p)|^k}{k!} = e^{|p\ln p|} = |p|^{|p|} \tag{42}$$

Of course this is only true if $p \neq 0$; however, in the case of $p = 0$ we know that $M_k = 0$, hence $\sum M_k = \sum 0 = 0$, which still implies uniform convergence of $\sum f_k$.

Since $\sum f_k$ is uniformly convergent, by theorem 7.16 in Rudin we know

$$\int_0^1 \sum_{k=0}^{\infty} \frac{(-x\ln x)^k}{k!}dx = \sum_{k=0}^{\infty} \int_0^1 \frac{(-x\ln x)^k}{k!}dx = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\int_0^1 (x\ln x)^k dx \tag{43}$$

Let $-u = \ln x$, so $\lim_{x\to 0} -u = \infty$. Then by theorem 6.19 we know (10) is equivalent to

$$\lim_{a\to\infty}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}\int_a^0 e^{-uk}(-u)^k(-e^{-u})du = \lim_{a\to\infty}\sum_{k=0}^{\infty}\frac{(-1)^{2k}}{k!}\int_a^0 (-1)^{-u}e^{-u(k+1)}(-u)^k du \tag{44}$$

$$= \lim_{a\to\infty}\sum_{k=0}^{\infty}\frac{1}{k!}\int_0^a e^{-u(k+1)}u^k du \tag{45}$$

Let $v = u(k+1)$, so $dv = (k+1)du$. Then (12) becomes

$$\lim_{a\to\infty}\sum_{k=0}^{\infty}\frac{1}{k!}\int_0^a e^{-v}\left(\frac{v}{k+1}\right)^k \frac{1}{k+1}dv = \lim_{a\to\infty}\sum_{k=0}^{\infty}\frac{1}{k!(k+1)^{k+1}}\int_0^a e^{-v}v^k dv$$

Definition 8.17 in Rudin gives a nice function of a similar form to what we currently have above.

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

Setting $t = v$ and $k+1 = x$ yields

$$\sum_{k=0}^{\infty}\frac{\Gamma(k+1)}{k!(k+1)^{k+1}}$$

By theorem 8.18 in Rudin our final expression becomes

$$\sum_{k=0}^{\infty}\frac{1}{(k+1)^{k+1}} = \sum_{k=0}^{\infty}(k+1)^{-(k+1)}$$

or with a re-indexing,

$$\sum_{n=1}^{\infty} n^{-n}$$

as desired. Disclaimer: from the first u substitution and on I heavily relied on Fehlau's (nonrigorous) solution found in [2]. All other work was produced by myself or cited accordingly. □

Problem 4. If $f : [-1,1] \longrightarrow \mathbb{R}$ is continuous and for all non-negative integers n

$$\int_{-1}^{1}\sin^n(x)f(x)dx = 0 \tag{46}$$

then $f \equiv 0$.

Proof. Define the set

$$\mathscr{A} := \left\{\sum_{k=0}^{n} a_k \sin^k x : a_k \in \mathbb{R}\right\} \tag{47}$$

We claim $\mathscr{A}$ forms an *algebra* of functions on $[-1,1]$. Let $\sum_{k=0}^{n} a_k \sin^k x,\ \sum_{k=0}^{m} b_k \sin^k x \in \mathscr{A}$. Then clearly

$$\sum_{k=0}^{n} a_k \sin^k x + \sum_{k=0}^{m} b_k \sin^k x = \sum_{k=0}^{\max\{n,m\}} (a_k + b_k)\sin^k x \in \mathscr{A} \tag{48}$$

where we take the coefficients corresponding to the summation with $\min\{n,m\}$ terms to be zero for all $k > \min\{n,m\}$. Next notice by (15)

$$\left(\sum_{k=0}^{n} a_k \sin^k x\right)\left(\sum_{k=0}^{m} b_k \sin^k x\right) = \sum_{k=0}^{m+n}\left(\sum_{j=0}^{k} a_k \sin^k(x) b_{k-j}\sin^{k-j}(x)\right) \tag{49}$$

$$= \sum_{k=0}^{m+n} \left(\sum_{j=0}^{k} a_k b_{k-j} \sin^{2k-j} x \right) \in \mathscr{A} \tag{50}$$

Let $c \in \mathbb{R}$. Then we see that

$$c \sum_{k=0}^{m} a_k \sin^k x = \sum_{k=0}^{m} c a_k \sin^k x = \sum_{k=0}^{m} c_k \sin^k x \in \mathscr{A} \tag{51}$$

where $c_i = ca_i$ for $i \in \{0, \ldots, n\}$. Therefore $\mathscr{A}$ is an algebra of real functions on the set $[-1, 1]$. Evidently, $a \sin^0 x = a \in \mathscr{A}$, where $a \in \mathbb{R}$. Hence $\mathscr{A}$ vanishes at no point in $[-1, 1]$.

Now, we claim $\mathscr{A}$ separates points on $[-1, 1]$. By definition this claim is equivalent to proving there exists an injective function $\sum_{k=0}^{n} a_k \sin^k x \in \mathscr{A}$. Consider $\sin x \in \mathscr{A}$ and recall Euler's formula

$$e^{ix} = \cos x + i \sin x \tag{52}$$

From (19) we gather that $\sin x = \operatorname{Im}(e^{ix})$. The series expansion of e^{ix} given in Rudin is

$$e^{ix} = \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} \tag{53}$$

Taking the imaginary part yields

$$\operatorname{Im}(e^{ix}) = \operatorname{Im} \left(\sum_{n=0}^{\infty} \frac{(ix)^n}{n!} \right) = \sum_{n=0}^{\infty} \operatorname{Im} \left(\frac{(ix)^n}{n!} \right) \tag{54}$$

Notice that $i^1 = i$, $i^2 = -1$, $i^3 = -i$, and $i^4 = 1$, which implies $i^5 = i$, $i^6 = i^2$, and so on. In general, $i^{2k} = (i^2)^k = \pm 1$ depending on the parity of k, so $i^{2k+1} = \pm i$, i.e. $\operatorname{Im}(i^{2k+1}) = \pm 1$ depending on the parity of k. Therefore $\operatorname{Im}(i^k)$ vanishes for integers of even powers and for integers of odd power $2k + 1$ we see that $\operatorname{Im}(i^{2k+1}) = (-1)^k$, which implies the series expansion in (21) is equivalent to

$$\operatorname{Im}(e^{ix}) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = \sin x \tag{55}$$

Hence proving injectivity of $\sin x$ on $[-1, 1]$ is equivalent to proving injectivity of (22).

We claim if a differentiable function $h : E \to \mathbb{R}$ has positive derivative on $E \subset \mathbb{R}$, then it is injective in E. Suppose for contradiction there exists $x_1 \neq x_2 \in E$ such that $h(x_1) = h(x_2)$. Without loss of generality, suppose $x_1 < x_2$. Then by the mean value theorem there exists $c \in (x_1, x_2)$ such that $h'(c) = \frac{h(x_2)-h(x_1)}{x_2-x_1} = 0$, a contradiction.

Now, by theorem 8.7 in Rudin we know differentiating (22) with respect to x yields

$$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = \cos x \tag{56}$$

and again by Rudin we know $\cos \pi/2 = 0$, $\cos x > 0$ in $[0, \pi/2)$, and $\cos x$ has period 2π, hence $\cos x > 0$ in $[-1, 1] \subset (\pi/2 - 2\pi, \pi/2)$. Therefore $\sin x$ is injective on $[-1, 1]$, which ultimately means $\mathscr{A}$ separates points on $[-1, 1]$.

Since $[-1, 1]$ is compact, by the Stone-Weierstrass theorem there exists a sequence of functions $\{f_n\}_{n=0}^{\infty}$ in $\mathscr{A}$ uniformly converging to f, i.e.

$$\{f_n\}_{n=0}^{\infty} = \left\{ \sum_{k=0}^{m_n} a_{k_n} \sin^k x \right\}_{n=0}^{\infty} \longrightarrow f(x) \tag{57}$$

Therefore we know

$$\int_{-1}^{1}(f(x))^2dx = \lim_{n\to\infty}\int_{-1}^{1}\left(\sum_{k=0}^{m_n} a_{k_n}\sin^k x\right)f(x)dx = \lim_{n\to\infty}\sum_{k=0}^{m_n}\int_{-1}^{1} a_{k_n}\sin^k(x)f(x)dx \tag{58}$$

By assumption, the limit in (25) must therefore vanish, i.e.

$$\lim_{n\to\infty}\sum_{k=0}^{m_n}\int_{-1}^{1} a_{k_n}\sin^k(x)f(x)dx = \lim_{n\to\infty}\sum_{k=0}^{m_n} 0 = 0 \tag{59}$$

which implies

$$\int_{-1}^{1}(f(x))^2\,dx = \lim_{n\to\infty}\sum_{k=0}^{m_n} 0 = 0 \tag{60}$$

We proved in class that for $f \in \mathscr{R}[-1,1]$, since f^2 is always non-negative

$$\int_{-1}^{1}(f(x))^2dx = 0 \implies f \equiv 0 \tag{61}$$

this concludes the proof. $\square$

Problem 5. If f is Riemann integrable on $[0,1]$ and continuous at 0, then

$$\lim_{t\to 0^+}\int_0^1 tx^{t-1}f(x)dx = f(0) \tag{62}$$

Proof. Let $c\in\mathbb{R}$ be some constant. Then notice that

$$\lim_{t\to 0^+}\int_0^1 tx^{t-1}cdx = \lim_{t\to 0^+} ct\int_0^1 x^{t-1}dt \tag{63}$$

Since for $0<t<1$ this integral is improper at $x=0$, we pass to another limit

$$\lim_{t\to 0^+} ct \lim_{b\to 0^+}\int_b^1 x^{t-1}dx = \lim_{t\to 0^+} ct \lim_{b\to 0^+}\left(\frac{1^t}{t}-\frac{b}{t}\right) = \lim_{t\to 0^+} c\frac{t}{t} = c \tag{64}$$

Consider $g(x) := f(x) - f(0)$ and note that $f(0)\in\mathbb{R}$ is just some constant. Since (29) holds for functions of constant values, it suffices to show

$$\lim_{t\to 0^+}\int_0^1 tx^{t-1}g(x)dx = 0 \tag{65}$$

Again, since for values of $t\in(0,1)$ the integral is improper, we must evaluate two limits as before,

$$\lim_{t\to 0^+} t \lim_{b\to 0^+}\int_b^1 x^{t-1}g(x)dx \tag{66}$$

By assumption, f is continuous at $x=0$, which implies g is continuous at $x=0$, therefore for $t\in(0,1)$ there exists $\delta>0$ such that $0<x<\delta$ implies $|g(x)-g(0)| = |g(x)| < t$. Therefore we can write (33) as

$$\lim_{t\to 0^+} t \lim_{b\to 0^+}\left(\int_b^\delta x^{t-1}g(x)dx + \int_\delta^1 x^{t-1}g(x)dx\right) \tag{67}$$

Furthermore, since f is Riemann integrable by assumption we know g is Riemann integrable and therefore bounded. Let $M := \sup g([0,1])$. Then we can create a bound on (34)

$$\lim_{t\to 0^+} t \lim_{b\to 0^+}\left|\int_b^\delta x^{t-1}g(x)dx + \int_\delta^1 x^{t-1}g(x)dx\right| \leq \lim_{t\to 0^+} t \lim_{b\to 0^+}\left(\left|\int_b^\delta x^{t-1}g(x)dx\right| + \int_\delta^1 x^{t-1}Mdx\right) \tag{68}$$

$$= \lim_{t\to 0^+} t \lim_{b\to 0^+}\left(\left|\int_b^\delta x^{t-1}g(x)dx\right| + \frac{M}{t}(1-\delta^t)\right) \tag{69}$$

Now since we know $|g(x)|<t$ for $x\in(0,\delta)$, we know $|x^{t-1}g(x)| = x^{t-1}|g(x)| < tx^{t-1}$, which means

$$\int_b^\delta x^{t-1}g(x)dx < \int_b^\delta tx^{t-1}dx \tag{70}$$

and therefore (36) is less than

$$\lim_{t\to 0^+} t \lim_{b\to 0^+}\left(\int_b^\delta tx^{t-1}dx + \frac{M}{t}(1-\delta^t)\right) = \lim_{t\to 0^+} t\left(\lim_{b\to 0^+}\frac{M}{t}(1-\delta^t) + (\delta^t - b^t)\right) \tag{71}$$

$$= \lim_{t\to 0^+}\lim_{b\to 0^+} M(1-\delta^t) + t(\delta^t - b^t) = \lim_{t\to 0^+} M(1-\delta^t) + t\delta^t = M(1-1) = 0 \tag{72}$$

Therefore by (33) and (35) we have that

$$\lim_{t\to 0^+}\lim_{b\to 0^+}\left|\int_b^1 tx^{t-1}g(x)dx\right| \leq \lim_{t\to 0^+} M(1-\delta^t) + t\delta^t = 0 \tag{73}$$

And since the non-negative integral in (40) is bounded above by a limit converging to 0 we must have

$$\lim_{t\to 0^+}\left|\int_0^1 tx^{t-1}g(x)dx\right| = 0 \tag{74}$$

which implies

$$\lim_{t\to 0^+}\int_0^1 tx^{t-1}g(x)dx = 0 \tag{75}$$

□

Problem 6. If $f:\mathbb{R}\longrightarrow\mathbb{R}$ is a strictly increasing function, $f(0)=0$, $a>0$, and $b>0$, then

$$\int_0^a f + \int_0^b f^{-1} \geq ab \tag{76}$$

Counterexample. Define

$$f(x) = \begin{cases} x & x\in(-\infty,1) \\ x+1 & x\in[1,\infty) \end{cases} \tag{77}$$

Therefore f is strictly increasing, $f(0)=0$, dom $f=\mathbb{R}$, and ran $f=(-\infty,1)\cup[2,\infty)=\mathbb{R}\setminus[1,2)$, which means we define the inverse so that dom $f^{-1}=\mathbb{R}\setminus[1,2)$ and ran $f=\mathbb{R}$.

$$f^{-1}(x) = \begin{cases} x & x\in(-\infty,1) \\ x-1 & x\in[2,\infty) \end{cases} \tag{78}$$

Notice that for $b=2$ the integral

$$\int_0^b f^{-1} \tag{79}$$

is undefined by construction of f. Therefore there is no way of telling whether the inequality in (43) holds true for f and f^{-1}, so the claim is not true. □

7 Sequences and Series of Functions

7.1 Pointwise and Uniform Convergence

Definition 7.1. (Pointwise Convergence) Suppose $\{f_n\}_{n=1}^{\infty}$ is a sequence of functions defined on a set E, and suppose that the sequence of numbers $\{f_n(x)\}_{n=1}^{\infty}$ converges for every $x\in E$. We then define the function f by

$$f(x)=\lim_{n\to\infty} f_n(x), \qquad x\in E$$

and we say f_n *converges pointwise* to f on E.

If $\sum_n f_n(x)$ converges for every $x\in E$ then we define

$$f(x)=\sum_{n=1}^{\infty} f_n(x), \qquad x\in E$$

as the limit of partial sums.

Remark. Pointwise convergence is a fairly *weak* notion of convergence. Say f_n is continuous for every $n\in\mathbb{N}$ and f_n converges pointwise to some f. It is natural to ask whether

$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t)$$

and in general, this is not the case.

Example 7.2. Consider the double sequence

$$s_{m,n} = \frac{m}{m+n}, \qquad m, n \in \mathbb{N}$$

Then for every fixed $n \in \mathbb{N}$ we have $\lim\limits_{m\to\infty} s_{m,n} = 1$, which would imply $\lim\limits_{n\to\infty}\lim\limits_{m\to\infty} s_{m,n} = 1$, but on the other hand if we fix m, then we see that $\lim\limits_{n\to\infty} s_{m,n} = 0$, which would imply $\lim\limits_{m\to\infty}\lim\limits_{n\to\infty} s_{m,n} = 0$, i.e.,

$$\lim_{n\to\infty}\lim_{m\to\infty} s_{m,n} \neq \lim_{m\to\infty}\lim_{n\to\infty} s_{m,n}$$

Example 7.3. Let

$$f_n(x) = \frac{x^2}{(1+x^2)^n}, \qquad x \in \mathbb{R}, n \in \mathbb{N}$$

and consider

$$f(x) = \sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \frac{x^2}{(1+x^2)^n}$$

Notice that $f_n(0) = 0$ for every $n \in \mathbb{N}$, which implies $f(0) = 0$. Now for $x \neq 0$ the above series is a convergent geometric series with sum $1 + x^2$, hence we have

$$f(x) = \begin{cases} 0 & x = 0 \\ 1 + x^2 & x \neq 0 \end{cases}$$

with some shallow thinking it can be seen that this series of continuous functions converges to a discontinuous function.

Definition 7.4. (Uniform Convergence) We say a sequence of functions $\{f_n\}_{n=1}^{\infty}$ *converges uniformly* on the domain E and write $f_n \rightrightarrows f$ if for every $\varepsilon > 0$ there exists an integer $N \in \mathbb{N}$ such that $n \geq N$ implies $|f_n(x) - f(x)| < \varepsilon$ for every $x \in E$. More concisely,

$$f_n \rightrightarrows f \iff \forall \varepsilon > 0 \exists N \in \mathbb{N} \ni n \geq N \wedge x \in E \implies |f_n(x) - f(x)| < \varepsilon$$

We say the series $\sum_n f_n$ converges uniformly to f if the sequence of partial sums is uniformly convergent.

Theorem 7.5. (Cauchy criterion for uniform convergence) The sequence of functions $\{f_n\}$ defined on E converges uniformly on E if and only if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $n, m \geq N$ and $x \in E$ implies $|f_n(x) - f_m(x)| < \varepsilon$.

Proof. (Do this later) for the forward implication us $\varepsilon/2$ and the triangle inequality. For the converse use the fact that $\mathbb{R}$ is complete, fix n, and let $m \to \infty$. □

Theorem 7.6. (Weierstrass M-test) Suppose $\{f_n\}$ is a sequence of functions defined on E, and suppose

$$|f_n(x)| \leq M_n, \qquad x \in E, n \in \mathbb{N}$$

Then $\sum_n f_n$ converges uniformly if $\sum_n M_n$ converges.

Proof. (Do later) use the general triangle inequality and the Cauchy criterion. □

7.2 Uniform Convergence and Continuity

Theorem 7.7. Suppose $f_n \rightrightarrows f$ on E in some metric space and let x be a limit point of E, and suppose that $\lim\limits_{t\to x} f_n(t) = A_n$ for $n \in \mathbb{N}$. Then $\{A_n\}$ converges and $\lim\limits_{t\to x} f(t) = \lim\limits_{n\to\infty} A_n$. In other words,

$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t)$$

Proof. (Do this later) This is actually kind of hard to prove □

Theorem 7.8. If $f_n \rightrightarrows f$ on E and f_n is continuous on E for each $n \in \mathbb{N}$, then f is continuous on E.

Theorem 7.9. Suppose K is compact and

1. $\{f_n\}$ is a sequence of continuous functions on K
2. $\{f_n\}$ converges pointwise to a continuous function f on K
3. $f_n(x) \geq f_{n+1}(x)$ for all $x \in K$ and $n \in \mathbb{N}$

then $f_n \rightrightarrows f$ on K.

Recall. The set $\mathcal{C}(X)$ denotes the set of all continuous bounded functions with domain X, and this set is a metric space with the norm $||f||_\infty = \sup\{|f(x)| : x \in X\}$, which is well defined by the construction of $\mathcal{C}(X)$.

Theorem 7.10. A sequence $\{f_n\}$ converges to f with respect to the metric $||\cdot||_\infty$ of $\mathcal{C}(X)$ if and only if $f_n \rightrightarrows f$ on X.

Theorem 7.11. $\mathcal{C}(X)$ is complete.

Proof. We only prove Cauchy implies convergent, since the converse is trivial at this point. Let $\{f_n\} \subset \mathcal{C}(X)$ be a Cauchy sequence of continuous functions. Fix $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $n, m \geq N$ implies $||f_n - f_m||_\infty < \varepsilon$. By definition this means

$$|f_n(x) - f_m(x)| \leq \sup\{\ |f_n(x) - f_m(x)| : x \in X\} < \varepsilon, \qquad x \in X$$

By the Cauchy criterion this implies f_n converges uniformly to some f, and since f_n is continuous for all $n \in \mathbb{N}$, this means f is continuous. It suffices to show f is bounded. (Do this) $\square$

7.3 Uniform Convergence and Integration

There is really only one important theorem (and a corollary) in this section (we are only considering Riemann integrals for now).

Theorem 7.12. Suppose $f_n \in \mathscr{R}([a,b])$ for all $n \in \mathbb{N}$ and suppose $f_n \rightrightarrows f$ on $[a,b]$. Then $f \in \mathscr{R}([a,b])$, and

$$\int_a^b f dx = \int_a^b \lim_{n\to\infty} f_n dx = \lim_{n\to\infty} \int_a^b f_n dx$$

Theorem 7.13. If $f_n \in \mathscr{R}([a,b])$ for all $n \in \mathbb{N}$, and if $\sum_n f_n \rightrightarrows f$ on $[a,b]$, then

$$\int_a^b f dx = \int_a^b \sum_{n=1}^\infty f_n dx = \sum_{n=1}^\infty \int_a^b f_n dx$$

7.4 Uniform Convergence and Differentiation

There is only one important theorem in this section as well. Note that uniform convergence of f_n says NOTHING about the sequence f_n', so we need a stronger hypothesis to say $f_n \rightrightarrows f \implies f_n' \rightrightarrows f'$.

Theorem 7.14. Suppose $\{f_n\}$ is a sequence of differentiable functions on $[a,b]$, and suppose there exists a point $x_0 \in [a,b]$ such that $\{f_n(x_0)\}$ converges. If $\{f_n'\}$ converges uniformly on $[a,b]$, then $\{f_n\}$ converges uniformly on $[a,b]$ to a function f, and

$$f' = \lim_{n\to\infty} f_n'$$

7.5 Equicontinuous Families of Functions

So there's continuous functions, then there's *uniformly continuous* functions, which are considered even more continuous than regular continuous functions. Now we learn about the ultimate form of continuity, equicontinuous families of functions are basically functions that are *really really* continuous.

The motivation for the proceeding theory building comes from the observation that every bounded complex sequences contains a convergent subsequence (Bolzano-Weierstrass). We would like to extend this notion to sequences of functions. In order to do this, we must first define two types of boundedness.

Definition 7.15. (Pointwise and Uniformly bounded) Let $\{f_n\}$ be a sequence if functions defined on a set E. We say that $\{f_n\}$ is *pointwise bounded* if $\{f_n(x)\}$ is bounded for every $x \in E$, i.e., for every $x \in E$ there exists a finite-valued function φ defined on E such that

$$|f_n(x)| \leq \varphi(x), \qquad x \in E, \ n \in \mathbb{N}$$

On the other hand, we say that $\{f_n\}$ is *uniformly bounded* if there exists a number $M \in \mathbb{R}$ such that

$$|f_n(x)| < M, \qquad \forall x \in E, \forall n \in \mathbb{N}$$

Now, neither of these two notions of boundedness give us anything useful for the time being. Even if $\{f_n\}$ is a uniformly bounded sequence on a compact set K, we need not have a convergent, pointwise nor uniform, subsequence.

Example 7.16. Let $f_n(x) = \sin nx$, $x \in [0, 2\pi]$, $n \in \mathbb{N}$. Suppose there exists a sequence $\{n_k\}$ such that $\{\sin n_k x\}$ converges for every $x \in [0, 2\pi]$. Evidently, we must have

$$\lim_{k\to\infty} (\sin n_k x - \sin n_{k+1} x) = 0$$

which would imply

$$\lim_{k\to\infty} (\sin n_k x - \sin n_{k+1} x)^2 = 0$$

Lebesgue's dominated convergence theorem (which we have yet to define) implies that

$$\lim_{k\to\infty} \int_0^{2\pi} (\sin n_k x - \sin n_{k+1} x)^2 dx = 0$$

however, a (not so) simple calculation shows that

$$\int_0^{2\pi} (\sin n_k x - \sin n_{k+1} x)^2 dx = 2\pi$$

which is a blatant contradiction.

Another point to bring up is that convergent sequences of functions (even uniformly bounded sequences on compact sets) need not have uniformly convergent subsequences.

Example 7.17. Let

$$f_n(x) = \frac{x^2}{x^2 + (1 - nx)^2}, \qquad 0 \leq x \leq 1, n \in \mathbb{N}$$

Then $|f_n(x)| \leq 1$ for every $x \in [0, 1]$ and $n \in \mathbb{N}$, so $\{f_n\}$ is uniformly bounded on $[0, 1]$. Furthermore, $\lim\limits_{n\to\infty} f_n(x) = 0$, but

$$f_n\left(\frac{1}{n}\right) = 1, \qquad \forall n \in \mathbb{N}$$

so no subsequence can converge uniformly on $[0, 1]$ (why?).

What we have established is that continuity and uniform boundedness and compactness and convergence and all those nice things that give nice properties are still not enough to guarantee the existence of a uniformly convergent subsequence of functions.

The missing concept we need is called equicontinuity, and it basically means *really really* continuous.

Definition 7.18. A family $\mathscr{F}$ of complex functions f defined on a set E in a metric space X is said to be *equicontinuous* on E if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$x, y \in E \wedge f \in \mathscr{F} \wedge d(x, y) < \delta \implies |f(x) - f(y)| < \varepsilon$$

i.e.,

$$\forall \varepsilon > 0 \exists \delta > 0 \ni x, y \in E \wedge f \in \ \wedge d(x, y) < \delta \implies |f(x) - f(y)| < \varepsilon$$

So like these guys are *really* continuous. It is easy to show equicontinuity implies uniform continuity (like really easy to show).

Alright now time for an unrelated theorem.

Theorem 7.19 (Unrelated)**.** If $\{f_n\}$ is a pointwise bounded sequence of complex functions on a countable set E, then $\{f_n\}$ has a pointwise convergent subsequence $\{f_{n_k}\}$, i.e., there exists a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in E$.

This is unrelated to equicontinuity, but it is related to our original question of convergent subsequences of functions. The theorem is rather weak. Countability of E is a fairly strong restriction, and the subsequence is only pointwise convergent.

Theorem 7.20. If K is a compact metric space, $f_n \in \mathcal{C}(K)$ for $n \in \mathbb{N}$, and if $\{f_n\}$ converges uniformly on K, then $\{f_n\}$ is equicontinuous on K.

So we have uniform convergence on a compact set implies equicontinuity (the proof of which involves applying the triangle inequality like 3 times or something). Now, time for the main result of this long-winded section:

Theorem 7.21 (Arzela-Ascoli)**.** If K is compact, and if $\{f_n\}_{n=1}^{\infty} \subset \mathcal{C}(K)$ is pointwise bounded and equicontinuous on K, then

1. $\{f_n\}$ is uniformly bounded on K
2. $\{f_n\}$ contains a uniformly convergent subsequence

7.6 The Stone-Weierstrass Theorem

Theorem 7.22. If f is a continuous complex function on $[a, b]$, then there exists a sequence of polynomials P_n such that $P_n \rightrightarrows f$ on $[a, b]$. If f is taken to be real, then P_n may be taken to be real as well.

Problem 1. Let c_0 be the subspace of ℓ^∞ consisting of sequences that converge to 0. Let α be a non-zero multiplicative linear functional on c_0. Then there exists a natural number m such that for $f \in c_0$, $\alpha(f) = f(m)$.

Proof. Suppose for contradiction for every $m \in \mathbb{N}$ there exists $f_m \in c_0$ such that $\alpha(f_m) \neq f_m(m)$. Recall $\ell^\infty = \{\ f : \mathbb{N} \longrightarrow \mathbb{R} \ \big|\ \sup_n |f(n)| < \infty\ \}$, which means $c_0 = \{\ f \in \ell^\infty \big|\ \lim_{n\to\infty} f(n) = 0\ \}$. We first show that c_0 is an *algebra* of real functions. By **Theorem 3.3** in Rudin we know if $f, g \in c_0$ then $\lim_{n\to\infty}(f + g)(n) = \lim_{n\to\infty} f(n) + g(n) = 0$, $\lim_{n\to\infty}(fg)(n) = \lim_{n\to\infty} f(n)g(n) = 0$, and furthermore for $r \in \mathbb{R}$, $\lim_{n\to\infty}(rf)(n) = r \cdot \lim_{n\to\infty} f(n) = r0 = 0$. Therefore c_0 is closed under pointwise addition and multiplication and scalar multiplication, i.e. c_0 is an algebra.

Let $\iota : \mathbb{N} \longrightarrow \mathbb{R}$ be defined so that $n \mapsto 1$ for all $n \in \mathbb{N}$. We extend c_0 to a new algebra, c, so that $\iota \in c$. To do this, we define $c := \{\ \{a_n\}_{n=1}^{\infty} + \{\lambda\}\ :\ \{a_n\}_{n=1}^{\infty} \in c_0,\ \lambda \in \mathbb{R}\}$, where $\{\lambda\}$ denotes the constant sequence of λ. Clearly $\iota \in c$. To see why c is an algebra, we invoke **Theorem 3.3** once more. If $f + \{a\}, g + \{b\} \in c$ where $f, g \in c_0$ and $\{a\}, \{b\}$ are constant sequences, we define addition so that $(f + \{a\}) + (g + \{b\}) = (f + g) + \{a + b\} \in c$.

Since c_0 is an algebra we know $f+g \in c_0$ and we also know adding constant sequences yields another constant sequence. We define multiplication as such: $(f+\{a\})\cdot(g+\{b\}) = fg + f\{b\} + g\{a\} + \{ab\} = h + \{ab\}$ where $h = fg + f\{b\} + g\{a\}$. Note that pointwise multiplication of a sequence by a constant sequence is equivalent to scalar multiplication hence $h \in c_0$. Now if $r \in \mathbb{R}$ then $r(f+\{a\}) = rf + \{ra\}$ where we know $rf \in c_0$ and $\{ra\}$ is still constant, hence c is closed under pointwise addition, multiplication, and scalar multiplication and is therefore an algebra containing ι.

We next extend $\alpha : c_0 \longrightarrow \mathbb{R}$ to $\beta : c \longrightarrow \mathbb{R}$ by defining $\beta(f+\{a\}) = \alpha(f) + a$. We claim under this definition β is a multiplicative linear functional. Notice that

$$\beta\left(f+\{a\}+g+\{b\}\right) = \beta(f+g+\{a+b\}) = \alpha(f+g) + a + b$$

$$= \alpha(f) + a + \alpha(g) + b = \beta(f+\{a\}) + \beta(g+\{b\})$$

furthermore,

$$\beta(fg + \{b\}f + \{a\}g + \{ab\}) = \alpha(fg) + b\alpha(f) + a\alpha(g) + ab$$
$$= (\alpha(f)+a)(\alpha(g)+b) = \beta(f+\{a\})\beta(g+\{b\})$$

and finally $\beta(rf+ra) = r\alpha(f) + ra = r(\alpha(f)+a) = r\beta(f+\{a\})$. Therefore β is a well defined linear functional on c.

Back to the matter at hand, by assumption, for every $m \in \mathbb{N}$ there exists $f_m \in c_0$ such that $\alpha(f_m) \neq f_m(m)$. Define a new sequence $g_m : \mathbb{N} \longrightarrow \mathbb{R}$ by $g_m = f_m - \{\alpha(f_m)\}$. Note that $\alpha(f_m) \in \mathbb{R}$, hence $g_m \in c$. Consider the map $G : \mathbb{N} \longrightarrow \mathbb{R}$ defined by $m \mapsto g_m(m) = f_m(m) - \alpha(f_m)$ and note that $G \in c$. Furthermore G is never zero by construction. Since $G \in c$ we can apply β to obtain

$$\beta(G) = \beta(f_m - \{\alpha(f_m)\}) = \alpha(f_m) - \alpha(f_m) = 0 \tag{80}$$

We have shown on previous homework that if β is a non-zero linear map then $\beta(1) = 1$. Since G is never zero we can take the reciprocal sequence $\frac{1}{G}$ to obtain the following result

$$1 = \beta(1) = \beta\left(G\frac{1}{G}\right) = \beta(G)\beta\left(\frac{1}{G}\right) = 0 \cdot \beta\left(\frac{1}{G}\right) \tag{81}$$

which is a contradiction.

□

Problem 2. If α is a non-zero multiplicative linear functional on $C_{\mathbb{R}}[0,1]$, then there exists $t \in [0,1]$ such that $\alpha(f) = f(t)$.

Proof. Suppose for contradiction that for every $t \in [0,1]$ there exists $f_t \in C_{\mathbb{R}}[0,1]$ such that $\alpha(f_t) \neq f_t(t)$. Define the continuous function $g_t(x) = f_t(x) - \alpha(f_t)$ and note that by construction $g_t(t) \neq 0$, whereas by linearity, $\alpha(g_t) = \alpha(f_t - \alpha(f_t)) = \alpha(f_t) - \alpha\left(\alpha(f_t)\right)$. Note that $\alpha(f_t) \in \mathbb{R}$ is just some scalar constant. Therefore we know $\alpha(f_t) - \alpha(\alpha(f_t)) = \alpha(f_t) - \alpha(f_t)\cdot\alpha(1)$. Therefore, $\alpha(f_t) - \alpha(\alpha(f_t)) = \alpha(f_t) - \alpha(f_t)\cdot\alpha(1) = \alpha(f_t) - \alpha(\alpha(f_t)) = \alpha(f_t) - \alpha(f_t) = 0$, i.e. $g_t \in \mathrm{Ker}(\alpha)$.

We claim there exists a neighborhood U_t around t such that $g_t(x) \neq 0$ for all $x \in U_t$. Since $g_t(t) \neq 0$ we know $0 < |g_t(t)|$ and by continuity of g_t, there exists a $\delta > 0$ such that $|x-t| < \delta$ implies $|g_t(x) - g_t(t)| < |g_t(x)|$. There are two cases for the value of $|g_t(t)|$. Case 1: $|g_t(t)| = g_t(t)$. Then we know $|g_t(x) - g_t(t)| < g_t(t)$ which implies $-g_t(t) < |g_t(x) - g_t(t)| < g_t(t)$, so $-g_t(t) < g_t(x) - g_t(t) < g_t(t)$, i.e. $0 < g_t(x) < 2g_t(t)$. The key takeaway here is that $|x-t| < \delta$ implies $0 < g_t(x)$. Case 2: $|g_t(t)| = -g_t(t)$. Then we know $g_t(t) < g_t(x) - g_t(t) < -g_t(t)$, i.e. $2g_t(t) < g_t(x) < 0$, in particular $g_t(x) < 0$ for $|x-t| < \delta$. Take $U_t = (t-\delta, t+\delta)$. Then by the above arguments we know for $x \in U_t$, $g_t(x) \neq 0$.

Notice how we have an open cover of $[0,1]$:

$$[0,1] \subset \bigcup_{t\in[0,1]} U_t \tag{82}$$

By compactness of $[0,1]$ this means there exists a finite subcover:

$$[0,1] \subset \bigcup_{i=1}^{n} U_{t_i} \tag{83}$$

where $\{t_1, \ldots, t_n\} \subset [0,1]$. Now, define the function $g(x) := \sum_{i=1}^{n} (g_{t_i}(x))^2$ and notice by construction $g(x) \neq 0$ for all $x \in [0,1]$. Recall $\alpha(g_t) = 0$ for all $t \in [0,1]$, therefore by linearity we know

$$\alpha(g) = \alpha\left(\sum_{i=1}^{n} (g_{t_i})^2\right) = \sum_{i=1}^{n} \alpha\left(g_{t_i}^2\right) = \sum_{i=1}^{n} (\alpha(g_{t_i}))^2 = \sum_{i=1}^{n} 0^2 = 0 \tag{84}$$

However, since $g(x) \neq 0$ for all $x \in [0,1]$, we know the reciprocal $\frac{1}{g}$ is a well defined function in $C_{\mathbb{R}}[0,1]$, hence

$$1 = \alpha(1) = \alpha\left(g \cdot \frac{1}{g}\right) = \alpha(g)\alpha\left(\frac{1}{g}\right) = 0 \cdot \alpha\left(\frac{1}{g}\right) \tag{85}$$

which is a contradiction. □

Problem 3. The subalgebra of $C_{\mathbb{R}}[0,1]$ generated by the function $f(t) = \cos t$ is dense.

Proof. By definition, the subalgebra generated by $\cos t$ is of the form

$$\mathscr{A} := \left\{ \sum_{k=0}^{n} a_k (\cos t)^k \,\middle|\, a_k \in \mathbb{R},\ n \in \mathbb{N} \right\} \tag{86}$$

Evidently, $a \cos^0 x = a \in \mathscr{A}$, where $a \in \mathbb{R}$. Hence by taking $a \neq 0$ we see that $\mathscr{A}$ vanishes at no point in $[0,1]$. To apply the Stone-Weierstrass theorem we must first demonstrate that $\mathscr{A}$ separates points in $[0,1]$, which is equivalent to saying $\cos t$ is injective in $[0,1]$. Recall the identity

$$\sin^2 x + \cos^2 x = 1 \tag{87}$$

Suppose for contradiction there exist $x_0, x_1 \in [0,1]$ such that $x_0 \neq x_1$ and $\cos x_0 = \cos x_1$. Then by (8) we know

$$\sin^2 x_0 + \cos^2 x_0 = \sin^2 x_1 + \cos^2 x_0 \implies \sin^2 x_0 = \sin^2 x_1 \tag{88}$$

Recall in previous homework we proved sin is injective on $[-1,1]$. Therefore sin is injective on $[0,1] \subset [-1,1]$. We also know by **Chapter 8.6** in Rudin that $\sin 0 = 0$ and sin is increasing in $(0, \frac{\pi}{2})$, hence sin is non-negative in $[0,1]$, but this means (9) is a contradiction, so cos is indeed injective on $[0,1]$ and therefore $\mathscr{A}$ separates points in $[0,1]$. By the Stone-Weierstrass theorem this implies $\mathscr{A}$ is dense in $C_{\mathbb{R}}[0,1]$ as desired. □

Problem 4. If C is the Cantor function, calculate

$$\int_0^1 C \; dC$$

Claim 7. If f is a monotone non-decreasing continuous function on $[0,1]$, then

$$\int_0^1 f \; df = \frac{1}{2}\left((f(1))^2 - (f(0))^2\right)$$

Proof of claim. **Theorem 6.17** in Rudin tells us that

$$\int_0^1 f \; df = \int_0^1 f f' dx \tag{89}$$

Hence integration by parts yields

$$\int_0^1 ff'df = f^2\Big|_0^1 - \int_0^1 ff'dx \implies 2\int_0^1 ff'dx = f^2(1) - f^2(0) \tag{90}$$

i.e.

$$\int_0^1 fdf = \frac{1}{2}\left(f^2(1) - f^2(0)\right)$$

as desired. □

Corollary 1.

$$\int_0^1 C\ dC = \frac{1}{2}\left(C^2(1) - C^2(0)\right) = \frac{1}{2}$$

8 Functions of Several Variables

At this point in time you should be comfortable with basic linear algebra concepts and definitions (basis, dimension, linear transformation, vector space, linear dependence/independence, etc).

Recall. All of linear algebra.

8.1 Linear Transformations

Definition 8.1. Let X, Y be vector spaces. We denote the set of all linear transformations from X to Y by $\mathcal{L}(X, Y)$. If $Y = X$ then we simply write $\mathcal{L}(X)$.

If $A \in \mathcal{L}(X, Y)$ and $B \in \mathcal{L}(Y, Z)$, then we denote their product BA by function composition:

$$(BA)\mathbf{x} = B(A\mathbf{x}), \quad \mathbf{x} \in X$$

note that $BA \in \mathcal{L}(X, Z)$. Recall this operation need not be commutative ($BA \neq AB$).

For $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ we define the norm $||A||$ of A so that

$$||A|| := \sup\{||A\mathbf{x}||_m : \mathbf{x} \in \mathbb{R}^n, ||\mathbf{x}||_n \leq 1\}$$

note the inequality $||A\mathbf{x}||_m \leq ||A||\ ||x||_n$ holds for all $x \in \mathbb{R}^n$. Also if $\lambda \in \mathbb{R}$ is so that $||A\mathbf{x}||_m \leq \lambda||\mathbf{x}||_n$, then $||A|| \leq \lambda$

Theorem 8.2.

1. If $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, then $||A| < \infty$ and A is a uniformly continuous mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$.

2. If $A, B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and c is a scalar, then

$$||A + B|| \leq ||A|| + ||B||, \quad ||cA|| = |c|\ ||A||$$

Evidently, $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is a metric space with the norm defined above.

3. If $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and $B \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^k)$ then

$$||BA|| \leq ||B||\ ||A||$$

Since $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is a metric space, notions of open sets and continuity make sense.

Theorem 8.3. Let Ω be the set of all invertible linear operators on $\mathbb{R}^n$.

1. If $A \in \Omega$, $B \in \mathcal{L}(\mathbb{R}^n)$, and

$$||B - A|| \cdot ||A^{-1}|| < 1$$

then $B \in \Omega$, i.e., B is invertible.

2. Ω is an open subset of $\mathcal{L}(\mathbb{R}^n)$, and the mapping $A \longrightarrow A^{-1}$ is continuous on Ω (this mapping is also injective onto Ω).

8.2 Matrices

We are pretty familiar with the concept of matrices already. The main takeaway in this section is that every linear transformation from a finite dimensional vector space to another admits a matrix representation. This matrix is determined by picking a basis and applying the transformation to the basis vectors. Note that there are many different bases one can pick for any given finite dimensional vector space, and each one of these bases will give a different matrix representation.

Definition 8.4. (Matrix of a linear transformation) Suppose $\{x_1, \dots, x_n\}$ and $\{y_1, \dots, y_m\}$ are bases for $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, and $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Then we have for every $1 \leq j \leq n$

$$Ax_j = \sum_{i=1}^{m} a_{ij} y_i, \qquad a_{ij} \in \mathbb{R}$$

and we represent A by the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{bmatrix}$$

Remark. Suppose $x = \sum_{j=1}^{n} c_j x_j$. Then by linearity we see that (and you should write this all out by hand to really reinforce your understanding)

$$Ax = A\left(\sum_{j=1}^{n} c_j x_j\right) = \sum_{j=1}^{n} c_j A x_j = \sum_{j=1}^{n} c_j \sum_{i=1}^{m} a_{ij} y_i = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij} c_j y_i$$

What we have shown is that the i^{th} coordinate of Ax with respect to the basis $\{y_1 \dots, y_m\}$ is $\sum_{j=1}^{n} a_{ij} c_j$.

8.3 Differentiation

Definition 8.5. (Differentiable) Let V, W be normed vector spaces. Let $E \subset V$ be open, and suppose $f : E \longrightarrow W$. We say that f is *differentiable* at $x \in E$ if there exists $A \in \mathcal{L}(V, W)$ such that A is bounded and for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $||h||_V < \delta$ implies $||f(x+h) - f(x) - Ah||_W < \varepsilon ||h||_V$, i.e.,

$$\exists A \in \mathcal{L}(V, W) \ni ||A|| < \infty \land \forall \varepsilon > 0 \exists \delta > 0 \ni ||h||_V < \delta \implies ||f(x+h) - f(x) - Ah||_W < \varepsilon ||h||_V$$

Equivalently,

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

Provided A exists, we write $f'(x) = A$. Note that A depends on x, i.e., if f is also differentiable at $y \in E$ with $y \neq x$ then it might be the case that $f'(y) \neq A$.

Theorem 8.6. Let $f : E \longrightarrow W$ be differentiable at $x \in E$. Then the differential $f'(x)$ is unique.

Theorem 8.7 (Chain Rule)**.** Suppose $E \subset \mathbb{R}^n$ is open, $f : E \longrightarrow \mathbb{R}^m$ is differentiable at $x_0 \in E$, g maps an open set containing $f(E)$ to $\mathbb{R}^k$, and g is differentiable at $f(x_0)$. Then the mapping $F : E \longrightarrow \mathbb{R}^k$ defined by

$$F(x_0) = g(f(x_0))$$

is differentiable at x_0 with

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

8.4 Partial Derivatives

Consider a function $f : E \longrightarrow \mathbb{R}^m$ where $E \subset \mathbb{R}^n$ is open. Let $\{e_1, \ldots e_n\}$ and $\{u_1, \ldots u_m\}$ be the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively.

Definition 8.8. The *components* of f are the real functions $f_1, \ldots, f_m$ defined by

$$f(x) = \sum_{i=1}^{m} f_i(x)u_i, \qquad x \in E$$

or equivalently, $f_i(x) = \langle f(x), u_i \rangle$ for $1 \le i \le m$.

Remark. Note that $f(x) \in \mathbb{R}^m$, so intuitively we have $f(x) = f((x_1, \ldots, x_n)) = (f_1(x), f_2(x), \ldots, f_m(x))$.

Definition 8.9. (Partial Derivative) For $x \in E$, $1 \le i \le m$, and $1 \le j \le n$ we define

$$D_j f_i(x) = \lim_{t \to 0} \frac{f_i(x + te_j) - f_i(x)}{t}$$

provided the limit exists. Writing $f_i(x_1, \ldots, x_n)$ in place of $f_i(x)$, we see that $D_j f_i(x)$ is the derivative of f_i with respect to x_j, keeping the other variables fixed. We often use the notation $\frac{\partial f_i}{\partial x_j}$ in place of $D_j f_i$. We call $D_j f_i$ the *partial derivative* of f_i with respect to x_j.

Theorem 8.10. 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 $(D_i f_j)(x)$ exist, and

$$f'(x)e_j = \sum_{i=1}^{m} (D_j f_i)(x)u_i, \qquad 1 \le j \le n$$

A result of this theorem is that for finite dimensional differentiable maps we have

$$f'(x) = \begin{bmatrix} (D_1 f_1)(x) & (D_2 f_1)(x) & \cdots & (D_n f_1)(x) \\ (D_1 f_2)(x) & (D_2 f_2)(x) & \cdots & (D_n f_2)(x) \\ \vdots & & \ddots & \vdots \\ (D_1 f_m)(x) & \cdots & \cdots & (D_n f_m)(x) \end{bmatrix}$$

This guy is called the *Jacobian* matrix of f.

Definition 8.11. A differentiable mapping f of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$ is said to be *continuously differentiable* in E if f' is a continuous mapping of E into $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Remark. So we have a couple of notions of the derivative that we need to sort out. First, for a given $x \in E$, we have $f'(x) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. The actual operator itself, f', is a different map. f' takes as input elements of E and gives back a linear transformation that depends on the chosen point, so $f' : E \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. f is continuously differentiable if and only if the map $f' : E \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is continuous, i.e.,

$$\forall x \in E \; \forall \varepsilon > 0 \exists \delta > 0 \ni ||x - y||_n < \delta \implies ||f'(x) - f'(y)|| < \varepsilon$$

note that the norms are different here. The latter norm is the one we defined for linear maps at the start of this chapter. One should verify that if $f'(x) = 0$ for all $x \in E$ then f is necessarily constant.

8.5 The Contraction Principle

Definition 8.12. Let X be a metric space. If $\varphi : X \longrightarrow X$ and there is a constant $c \in (0, 1)$ such that for every $x, y \in X$ $d(\varphi(x), \varphi(y)) \le cd(x, y)$, then we say φ is a *contraction* of X into X. It's like φ squishes points together.

This next theorem determines the existence of a unique fixed point for every contraction mapping.

Theorem 8.13. If X is a complete metric space and $\varphi : X \longrightarrow X$ is a contraction, then there exists exactly one $x \in X$ such that $\varphi(x) = x$.

8.6 The Inverse Function Theorem

This next theorem states that a continuously differentiable function f is invertible in a neighborhood of any point x at which the linear transformation $f'(x)$ is invertible.

Theorem 8.14. Suppose f is a continuously differentiable mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$ and $f'(a)$ is invertible for some $a \in E$. Then

1. there exist open sets $U, V \in \mathbb{R}^n$ such that $a \in U$ and $f(a) \in V$, f is injective on U, and $f(U) = V$
2. The inverse $g = f^{-1}$ defined on V exists by the above assertion, g is continuously differentiable on V, and

$$g'(y) = [f'(g(y))]^{-1}, \quad y \in V$$

References

[1] Chen, Guo. A Treatment of the Dirichlet Integral via Methods of Real Analysis. University of Chicago, 26 June 2009.

[2] Fehlau, Jens. Sophomore's Dream. YouTube, YouTube, 9 Mar. 2018, www.youtube.com/watch?v=2Fpu9Ylb9_E.