Review Notes for the Basic Qualifying Exam

Topics: Analysis & Linear Algebra
Fall 2016 – Spring 2017

Created by
Howard Heaton

Purpose: This document is a compilation of notes generated to prepare for the Basic Qualifying Exam. I have documented some of my solutions so that I may not forget and repeat the frustrations of failing this cursed exam again. Best of luck to anyone using these notes to prepare. Also see Brent Woodhouse's study guide. I do not guarantee accuracy of all the presented solutions.

This document is long and incomplete and I return on occasion to revise solutions and add solutions I have not yet typeset. If the reader finds any typos or corrections to be made, feel free to email me at `heaton@math.ucla.edu` and I will address these and post an updated set of notes to my webpage `math.ucla.edu/~heaton/`.

Contents

1 Introduction

This document provides several previous exam solutions along with some random notes on different topics related to the exam.

2 Analysis

REMARK: The following material is useful for learning about *completeness*:

i) Baby Rudin: pg. 54-55, 150-151

ii) Rosenclicht: pg. 52-53

iii) Kreyszig: pg. 28-39

iv) Tao I: pg: 146-147

v) Tao II: pg. 19-20

vi) Fitzpatrick: 322-323

△

Definition: Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers and $x \in \mathbb{R}$. Then x is a **limit point** of $(a_n)_{n=m}^{\infty}$ if, for every $\varepsilon > 0$ and every $N \geq m$, there exists an $n \geq N$ such that $|a_n - x| \leq \varepsilon$. △

Definition: Let $(a_n)_{n=m}^{\infty}$ be a sequence of real number and $L \in \mathbb{R}$. Then $(a_n)_{n=m}^{\infty}$ **converges** to L if, given any real $\varepsilon > 0$, one can find an $N \geq m$ such that $|a_n - L| \leq \varepsilon \ \forall\ n \geq N$. △

Definition: A sequence $(a_n)_{n=m}^{\infty}$ is a **Cauchy sequence** iff for every $\varepsilon > 0$, there exists an $N \geq 0$ such that $d(a_j, a_k) \leq \varepsilon \ \ \forall\ j, k \geq N$. △

Definition: Let $(a_n)_{n=m}^{\infty}$ and $(b_n)_{n=m}^{\infty}$ be sequences in $\mathbb{R}$. Then $(b_n)_{n=m}^{\infty}$ is a subsequence of $(a_n)_{n=m}^{\infty}$ iff there exists a function $f : \mathbb{N} \to \mathbb{N}$ that is strictly increasing such that $b_n = a_{f(n)} \ \forall\ n \in \mathbb{N}$. △

Definition: Let $X \subseteq \mathbb{R}$, $f : X \to \mathbb{R}$ be a function, and $x_0 \in X$. Then the following statements are equivalent:

i) f is continuous at x_0.

ii) For every sequence $(a_n)_{n=m}^{\infty}$ in X with $\lim_{n\to\infty} a_n = x_0$, we have $\lim_{n\to\infty} f(a_n) = f(x_0)$.

iii) For every $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ for all $x \in X$ with $|x - x_0| < \delta$.

iv) For every $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(x_0)| \leq \varepsilon$ for all $x \in X$ with $|x - x_0| \leq \delta$.

We say that f is **continuous** if f is continuous at every $x_0 \in X$. $\triangle$

Definition: The **Archimedean property** states that whenever $x, \epsilon \in \mathbb{R}$ are positive, there exists $M \in \mathbb{N}$ such that $M\epsilon > x$. $\triangle$

Proposition: $f : X \to Y$ is continuous at x iff whenever $\{x_n\} \to x$, $\{f(x_n)\} \to f(x)$. $\triangle$

Proof:
Let $\{x_n\}$ be a sequence in X that converges to x and $\varepsilon > 0$ be given. Let us also assume f is continuous. Then we must find $N \in \mathbb{N}$ such that for all $n \geq N$, $d(f(x_n), f(x)) < \varepsilon$. By continuity of f, there exists $\delta > 0$ such that whenever $d(x_n, x) < \delta$, $d(f(x_n), f(x)) < \varepsilon$. Since $\{x_n\} \to x$, there is a $N \in \mathbb{N}$ such that $d(x_n, x) < \delta$ for all $n \geq N$. Then whenever $n \geq N$, $d(f(x_n), f(x)) < \varepsilon$.

Conversely, suppose that $\{f(x_n)\} \to f(x)$. We must show that this implies f is continuous at x, i.e., that there is a $\delta > 0$ such that whenever $y \in X$ and $d(x, y) < \delta$, $d(f(x), f(y)) < \varepsilon$. Suppose no such δ exists. Then, letting $\delta = 1/n$ for $n \in \mathbb{N}$, we can identify $x_n \in X$ so that $d(f(x), f(x_n)) > \varepsilon$ while $d(x, x_n) < \delta = 1/n$. This gives a sequence $\{x_n\}$ that converges to x while $\{f(x_n)\}$ does not converge to $f(x)$, which contradicts our initial assumption. Hence f must be continuous. $\square$

Definition: A **homeomorphism** is a bijection $f : X \to Y$ such that f and f^{-1} are continuous. $\triangle$

Proposition: A function $f : X \to Y$ is continuous iff for every open subset $V \subseteq Y$, $f^{-1}(V)$ is open in X. $\triangle$

REMARK: If the topologies X and Y are generated by basic open sets, the above is equivalent to say, for every basic open neighborhood N_y of $f(x_0)$, there is a basic open neighborhood N_x of x_0 such that $f(N_x) \subseteq N_y$. In particular, if X and Y are metric spaces, then f is continuous at x_0 iff $\forall\, \varepsilon > 0$, $\exists\, \delta > 0$ such that $x \in B(x_0, \delta) \Rightarrow f(x_0) \in B(f(x_0), \varepsilon)$. △

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

Proof:
Let $\varepsilon > 0$ be given. We must show there is $\delta > 0$ such that for $x, y \in [a, b]$ we have $|f(x)-f(y)| < \varepsilon$ whenever $|x - y| < \delta$. By the continuity of f, for each $x \in [a, b]$ there exists $\delta_x > 0$ such that $y \in (x-2\delta_x, x+2\delta_x)$ implies $|f(x)-f(y)| < \varepsilon/2$. Also, the collection of $(x-\delta_x, x+\delta_x)$ for $x \in [a, b]$ form an open cover of $[a, b]$. But, $[a, b]$ is closed and bounded and, therefore, compact. This implies there is a finite subcover of $x_1, \ldots, x_n \in [a, b]$ such that

$$[a, b] \subseteq \bigcup_{i=1}^{n} (x_i - \delta_{x_i}, x_i + \delta_{x_i}).$$

Then define $\delta = \min\{\delta_{x_1}, \ldots, \delta_{x_n}\}$ and suppose $x, y \in [a, b]$ such that $|x - y| < \delta$. Since we have a finite cover, there is x_i with $i \in \{1, \ldots, n\}$ such that $|x - x_i| \leq \delta_{x_i}$. This implies

$$|y - x_i| \leq |y - x| + |x - x_i| < \delta + \delta_{x_i} \leq 2\delta_{x_i}.$$

Hence $|f(x_i) - f(x)| < \varepsilon/2$ and $|f(x_i) - f(y)| < \varepsilon/2$. From the triangle inequality, it follows that $|f(x) - f(y)| < \varepsilon$. Hence f is uniformly continuous. □

Intermediate Value Theorem: Let $f : X \to \mathbb{R}$ be continuous and suppose X is connected. Then if f takes on values y_0 and y_1 (with $y_0 < y_1$), then f takes on every value between them, i.e., $\forall\, y \in (y_0, y_1)\ \exists\, x \in X$ such that $f(x) = y$. △

Proof:
Suppose $x_0, x_1 \in X$ such that $f(x_0) = y_0$ and $f(x_1) = y_1$. Then fix any $y \in (y_0, y_1)$, and let $A_0 = \{x \in X \mid f(x) < y\} = f^{-1}(-\infty, y)$ and $A_1 = \{x \in X \mid f(x) > y\} = f^{-1}(y, \infty)$. Both sets are open since f is continuous and since (y, ∞) and $(-\infty, y)$ are open. Both sets are nonempty since $x_0 \in A_0$ and $x_1 \in A_1$. Also, $A_0 \cap A_1 = \emptyset$ since we cannot simultaneously have $f(x) > y$ and $f(x) < y$. Since X is connected, we cannot have $X = A_0 \cup A_1$. Otherwise, X would be disjoint. Thus, there must exists $x \in X - (A_0 \cup A_1)$, which implies there exists $x \in X$ such that $f(x) = y$ and we are done. □

Definition: Let $X \subseteq \mathbb{R}$ and $f : X \to \mathbb{R}$ be a function. Then f is **uniformly continuous** if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(x_0)| \leq \varepsilon$ whenever $x, x_0 \in X$ with $|x - x_0| \leq \delta$. $\triangle$

Note: For uniform continuity one can take a single δ which works for all $x_0 \in X$ while for ordinary continuity each $x_0 \in X$ may use a different δ. Hence a uniformly continuous function is continuous, but not conversely.

Let (X, d) be a compact metric space and (Y, ρ) be a metric space. If $f : X \to Y$ is continuous, then f is uniformly continuous $\triangle$

Proof:
Let $\varepsilon > 0$ be given. By the continuity of f, for each $x \in X$ there exists $\delta_x > 0$ such that $y \in B(x, \delta_x)$ implies $f(y) \in B(f(x), \varepsilon/2)$. The collection of balls $\cup_{x\in X} B(x, \delta_x/2)$ form an open cover for X. Since X is compact, it follows that there is a finite subcover $\cup_{i=1}^n B(x_i, \delta_{x_i}/2)$ of X. Then define $\delta = \min\{\delta_{x_i}/2 \mid 1 \leq i \leq n\}$. Now suppose we have $d(x, y) < \delta$. Since $\cup_{i=1}^n B(x_i, \delta_{x_i}/2)$ covers X, there is an index j such that $x \in B(x_j, \delta_{x_j}/2)$. Then $d(x_j, y) \leq d(x_j, x) + d(x, y) \leq \delta_{x_j}/2 + \delta \leq \delta_{x_j}$. Hence $x, y \in B(x_j, \delta_{x_j})$, which implies

$$\rho(f(x), f(y)) \leq \rho(f(x), f(x_j)) + \rho(f(x_j), f(y)) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Hence f is uniformly continuous. $\square$

Definition: Let X, Y be metric spaces and $f : E \to Y$ be continuous with $E \subseteq X$. If $a \in X$ and $E - \{a\}$ has points arbitrarily close to a, we see the **limit** $\lim\limits_{x\to a, x\in E} f(x)$ exists and is equal to L iff $\forall\ \varepsilon > 0$, $\exists\ \delta > 0$ such that $x \in E - \{a\}$ and $d(x, a) < \delta$ implies $d(f(x), L) < \varepsilon$. $\triangle$

Example: Define $f_n : [0, 1] \to [0, 1]$ by $f(x) = x^n$. Let $f(x) = 0$ if $x \in [0, 1)$ and $f(x) = 1$ if $x = 1$. Then $\{f_n\} \to f$. $\triangle$

Remark: The above example shows that limits of continuous functions need *not* be continuous. Indeed, in the above we see

$$\lim_{x\to 1}\lim_{n\to\infty} f_n(x) = \lim_{x\to 1} f(x) = 0 \neq 1 = \lim_{n\to\infty}\lim_{x\to 1} f_n(x). \tag{1}$$

△

Definition: Let X, Y be metric spaces, $\{f_n : X \to Y\}_{n=1}^{\infty}$ be a sequence of functions, and $f : X \to Y$. Then $\{f_n\}$ **converges pointwise** to f on X if $\forall\ x \in X,\ \lim_{n\to\infty} f_n(x) = f(x)$. That is, $\forall\ x \in X, \varepsilon > 0,\ \exists\ N \in \mathbb{N}$, such that $\forall\ n \geq N$, $|f_n(x) - f(x)| < \varepsilon$. △

Definition: Let X, Y be metric spaces, $\{f_n : X \to Y\}_{n=1}^{\infty}$ be a sequence of functions, and $f : X \to Y$. Then $\{f_n\}$ **converges uniformly** to f on X iff $\forall\ \varepsilon > 0,\ \exists\ N \in \mathbb{N}$ such that $\forall\ x, n \geq N,\ d(f(x), f_n(x)) < \varepsilon$. △

Definition: For metric spaces X and Y, a function $f : X \to Y$ is **Lipschitz** with constant L if $\forall\ x, y \in X$, $d(f(x), f(y)) < Ld(x, y)$. △

Proposition: Lipschitz functions are continuous. △

Proof:
For metric spaces X and Y, define $f : X \to Y$ to be Lipschitz with constant L. Let $\varepsilon > 0$ be given and $x_0 \in X$. For $x \in X$ it follows that $d(x, x_0) < \varepsilon/L$ implies $d(f(x), f(x_0)) \leq Ld(x, x_0) < \varepsilon$. So, at each $x_0 \in X$, f is continuous. Thus, f is continuous and we are done. □

Monotone Convergence Theorem: A monotone sequence converges if and only if it is bounded. Moreover, the bounded monotone sequence $\{a_n\}$ converges to $\sup\{a_n \mid n \in N\}$ if it is monotonically increase and $\inf\{a_n \mid n \in \mathbb{N}\}$ if it is monotonically decreasing. △

Proof:
Suppose $\{a_n\}$ is a convergent sequence. We first show that this sequence is bounded. Let a denote the limit of $\{a_n\}$. Taking $\varepsilon = 1$, it follows from the definition of convergence that this is $N \in \mathbb{N}$ such that $|a_n - a| < 1$ whenever $n \geq N$. Using the triangle inequality, $|a_n| = |(a_n - a) + a| \leq |a_n - a| + a \leq 1 + |a|$ whenever $n \geq N$. Then define $M = \max\{|a_1|, \ldots, |a_{N-1}|, |a| + 1\}$. Then $|a_n| \leq M$ for each $n \in \mathbb{N}$ and so $\{a_n\}$ is bounded.

Now let $\{a_n\}$ be an unbounded monotone sequence. Further suppose that it is monotonically increasing. Then for each positive $M \in \mathbb{R}$ we can find a_n such that $a_n \geq M$. But, since the sequence is increasing, $a_n \geq a_N > M$ for all N, which reveals $x_n \to \infty$ as $n \to \infty$. Similar argument holds for a monotonically decreasing sequence using negative $M \in \mathbb{R}$.

All that remains it to show that a monotone sequence converges to its supremum if it is increasing and its infimum if it is decreasing. First suppose $\{a_n\}$ is a monotonically increasing sequence. Let $a = \sup\{a_n \mid n \in N\}$ and $\varepsilon > 0$ be given. We must find $N \in \mathbb{N}$ such that $|a_n - a| \leq \varepsilon$ for all $n \geq N$. By the definition of a, we have $a_n \leq a < a + \varepsilon$ for all $n \in \mathbb{N}$. Since a is the least upper bound for a, $a - \varepsilon$ is not an upper bound for a, which implies there is $N \in \mathbb{N}$ such that $a - \varepsilon < a_N$. Then $a - \varepsilon < a_N \leq a_n < a + \varepsilon$ for all $n \geq N$. This shows $|a_n - a| < \varepsilon$ whenever $n \geq N$, as desired. The case where $\{a_n\}$ is decreasing follows similarly. $\square$

Heine-Borel Theorem 1: Let (X, d) be a metric space. Then a subset $Y \subseteq X$ is compact iff it is complete and totally bounded. $\triangle$

Heine-Borel Theorem 2: Let $(\mathbb{R}^n, d)$ be a Euclidean space with either the Euclidean metric, the taxicab metric, or the sup norm metric. Let $E \subseteq \mathbb{R}^n$. Then E is compact iff E is closed and bounded. $\triangle$

Bolzano-Weierstrass Theorem in $\mathbb{R}^n$ (aka *Sequential Compactness Theorem*): A bounded subset of $S \subseteq \mathbb{R}^n$ is sequentially compact iff it is closed and bounded. $\triangle$

Proof:
We first show that every sequence $\{x_n\}_{n=1}^{\infty} \subset \mathbb{R}$ has a monotone subsequence. Let us call an integer n a "peak" of the sequence if $m > n$ implies that $x_n > x_m$, i.e., x_n is greater than every subsequent term in the sequence. Now suppose $\{x_n\}_{n=1}^{\infty}$ has infinitely many peaks, $n_1 < n_2 < \cdots < n_j < \cdots$. Then the subsequence corresponding to these peaks $\{x_{n_j}\}$ is monotonically decreasing. Alternatively, suppose $\{x_n\}_{n=1}^{\infty}$ has only finitely many peaks. Let N be the last peak and $n_1 = N + 1$. Then n_1 is not a peak since $n_1 > N$, which implies there exists $n_2 > n_1$ with $x_{n_2} \geq x_{n_1}$. Again $n_2 > N$ is not a peak, and so, by induction, we construct an infinite non-decreasing subsequence $x_{n_1} \leq x_{n_2} \leq x_{n_3} \leq \cdots$, as desired.

Now suppose $\{x_n\}_{n=1}^{\infty}$ is bounded. Then, by the above, there exists a monotone subsequence. Then by the Monotone Convergence Theorem, that subsequence must converge.

(continued on next page)

The case for $\mathbb{R}^n$ follows from the $n = 1$ case through application of a diagonalization argument. Given a bounded sequence in $\mathbb{R}^n$, the sequence of first coordinates is a bounded real sequence and, thus, has a convergent subsequence. Then from this subsequence we can extract a subsubsequence on which the second coordinates converge, and so on, until we have passed from the original subsequence n times, which is still a subsequence of the original sequence. On this final subsequence, each coordinate sequence converges. Hence this subsequence converges. $\square$

The Boundedness Theorem: A continuous function on a closed bounded interval is bounded. $\triangle$

Proof:

By way of contradiction, suppose a continuous function on the closed bounded interval $[a, b] \subset \mathbb{R}$ is not bounded. Then, for each $n \in \mathbb{N}$, there is $x_n \in [a, b]$ such that $f(x_n) > n$. This defines a sequence $\{x_n\}_{n=1}^{\infty}$. Because $[a, b]$ is bounded, the Bolzano-Weierstrass Theorem implies there exists a convergent subsequences $\{x_{n_k}\}_{k=1}^{\infty}$, whose limit shall be denoted by x. Since $[a, b]$ is closed, $x \in [a, b]$. Since f is continuous at x, it follows that $\{f(x_{n_k})\}_{k=1}^{\infty}$ converges to the finite value $f(x)$. But, since $f(x_{n_k}) > n_k \geq k$ for each $k \in \mathbb{N}$, it must follow that $\{f(x_{n_k})\} \to \infty$ as $k \to \infty$ and so this subsequence does not converge to the finite value $f(x)$, a contradiction. Therefore, f is bounded above on $[a, b]$. The proof that f is bounded below follows similarly. $\square$

The Extreme Value Theorem: If $f : [a, b] \to \mathbb{R}$ is continuous, then f must attain a maximum and a minimum. $\triangle$

Proof:

By the boundedness theorem, f is bounded above. By the Dedekind-completeness of $\mathbb{R}$, there is a supremum M of f. We must show there exists $c \in [a, b]$ such that $f(c) = M$. Let $n \in \mathbb{N}$. Then since M is the least upper bound, $M - 1/n$ is not an upper bound for f. Thus, there exists $c_n \in [a, b]$ such that $M - 1/n < f(c_n)$. This defines a sequence $\{c_n\}_{n=1}^{\infty}$. Since M is an upper bound for f, we have $M - 1/n < f(c_n) \leq M$ for all $n \in \mathbb{N}$. Now let $\varepsilon > 0$ be given. Then, by the Archimedean property of $\mathbb{R}$, there is $N \in \mathbb{N}$ such that $1/N \leq \varepsilon$. Thus, $|f(c_n) - M| \leq 1/n \leq \varepsilon$ whenever $n \geq N$ and so $\{f(c_n)\}_{n=1}^{\infty}$ converges to M.

The Bolzano-Weierstrass Theorem implies there is a subsequence $\{c_{n_k}\}_{k=1}^{\infty}$ that converges to some c. Since $[a, b]$ is closed, $c \in [a, b]$. Since f is continuous at c, $\{f(c_{n_k})\}_{k=1}^{\infty}$ converges to $f(c)$. But, $\{f(c_{n_k})\}_{k=1}^{\infty}$ is a subsequence of $\{f(c_n)\}_{n=1}^{\infty}$ that converges to M. Hence $M = f(c)$ and f attains its supremum M at c. Similar proof shows that f attains is infimum. $\square$

Arzela-Ascoli Theorem: Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of functions $f_n : [a,b] \to \mathbb{R}$. If this sequence is uniformly bounded and equicontinuous, then there is a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ that converges uniformly. Conversely, if every subsequence of $\{f_n\}_{n=1}^{\infty}$ itself has a uniformly convergent subsequence, then $\{f_n\}_{n=1}^{\infty}$ is uniformly bounded and equicontinuous. $\triangle$

Proof:

($\Rightarrow$) Let $\{f_n\}$ be a sequence in $\mathcal{F}$. Given $\varepsilon > 0$, we must show there exists $N \in \mathbb{N}$ such that

$$\forall\ t \in [0,1], \quad |f_n(t) - f_m(t)| \leq \varepsilon \ \text{ whenever } n, m \geq N.$$

We proceed using a diagonalization argument. Let $\sigma : \mathbb{N} \to \mathbb{Q} \cap [0,1]$ be an enumeration of the rationals on $[0,1]$. Then $\{f_n(\sigma(1))\}$ is a sequence of rationals. Moreover, because $\mathcal{F}$ is uniformly bounded by some $M > 0$, for each $n \in \mathbb{N}$ we have $f_n(\sigma(1)) \in [-M, M]$. The Bolzano Weierstrass Theorem implies that there is a convergent subsequence $\{f_{n_1(j)}(\sigma(1))\}_{j=1}^{\infty}$. Similarly, we can find a subsequence $\{f_{n_2(j)}\}$ of $\{f_{n_1(j)}\}$ such that $\{f_{n_2(j)}(\sigma(2))\}$ converges. Continuing in an inductive fashion, for each $k \in \mathbb{N}$ we can find a subsequence $n_{k+1}(j)$ of $n_k(j)$ such that $\{f_{n_{k+1}(j)}(\sigma(k))\}$ converges. In fact, by this construction, $\{f_{n_{k+1}(j)}(\sigma(m))\}$ converges for each $m = 1, \ldots, k+1$.

Now define a new sequence $m(j)$ by $m(j) = n_j(j)$. We claim $\{f_{m(j)}(\sigma(k))\}$ converges for each $k \in \mathbb{N}$. Indeed, given $k \in \mathbb{N}$, there are only finitely many terms in the sequence $\{f_{m(j)}(\sigma(k))\}$ that are not in $\{f_{n_k(j)}(\sigma(k))\}$, namely, the $k-1$ terms

$$f_{n_k(1)}(\sigma(k)), \ldots, f_{n_k(k-1)}(\sigma(k)).$$

Thus, in the limit as $k \to \infty$, we have that $\{f_{m(j)}(\sigma(k))\}$ converges to the limit of $\{f_{n_k(j)}(\sigma(k))\}$. Since this k was arbitrarily chosen, this holds for all $k \in \mathbb{N}$. Hence $\{f_{m(j)}(r)\}$ converges and is Cauchy for each $r \in [0,1] \cap \mathbb{Q}$.

Now, because $\mathcal{F}$ is equicontinuous, there is a $\delta > 0$ such that for all $n \in \mathbb{N}$ and $x, y \in [0,1]$,

$$|x - y| \leq \delta \quad \Rightarrow \quad |f_n(x) - f_n(y) \leq \varepsilon/3.$$

The collection of $B(x, \delta/2)$ form an open cover of $[0,1]$. However, $[0,1]$ is closed and bounded. By the Heine-Borel theorem, it follows that $[0,1]$ is compact. Thus, there is a finite subcover of $[0,1]$ by some collection $\cup_{j=1}^{J} B(x_j, \delta/2)$. Since the rationals are dense, there exists $r_1, \ldots, r_J$

with $r_j \in B(x_j, \delta/2)$. And, from the above, for each r_i, there exists N_i such that

$$|f_{m(j)}(r_i) - f_{m(k)}(r_i)| \leq \frac{\varepsilon}{3} \quad \text{whenever } j, k \geq N_i.$$

Let $N = \max\{N_i \mid 1 \leq i \leq J\}$. For each $t \in [0,1]$, it follows that there exists ℓ such that $t \in B(x_\ell, \delta/2)$. Thus,

$$|r_\ell - t| \leq |r_\ell - x_\ell| + |x_\ell + t| \leq \delta/2 + \delta/2 = \delta.$$

For $j, k \geq N$ this implies that for each $t \in [0,1]$

$$\begin{aligned} |f_{m(j)}(t) - f_{m(k)}(t)| &\leq |f_{m(j)}(t) - f_{m(j)}(r_\ell)| + |f_{m(j)}(r_\ell) - f_{m(k)}(r_\ell)| + |f_{m(k)}(r_\ell) - f_{m(k)}(t)| \\ &\leq \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\ &= \varepsilon. \end{aligned}$$

Hence $\{f_{m(j)}\}$ is uniformly Cauchy and every sequence $\{f_n\}$ has a uniformly convergent subsequence.

($\Leftarrow$) I omit the proof of the converse. □

Summation by Parts: Given two sequences $\{a_n\}$ and $\{b_n\}$, define $A_n = \sum_{k=0}^n a_k$ for $n \geq 0$ and put $A_{-1} = 0$. Then, if $0 \leq p \leq q$, we have

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

△

Suppose the partial sums $\sum_n a_n$ form a bounded sequence and that $\{b_n\}$ is a nonnegative monotonically decreasing sequence with $\lim_{n\to\infty} b_n = 0$. Then $\sum_n a_n b_n$ converges. △

Proof:

Choose M such that $A_n = \sum_{k=0}^n a_k$ is bounded above by M for all n. Given $\varepsilon > 0$, there is an integer N such that $b_N \leq (\varepsilon/2M)$. For $N \leq p \leq q$, we have

$$\begin{aligned}\left|\sum_{n=p}^{q} a_n b_n\right| &= \left|\sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p\right| \\ &\leq M\left|\sum_{n=p}^{q}(b_n - b_{n+1}) + b_q + b_p\right| \\ &= 2Mb_p \\ &\leq 2Mb_N \\ &\leq \varepsilon.\end{aligned}$$

Convergence now follows from the Cauchy criterion. We note that the first inequality in the above chain depends on the fact that $b_n - b_{n+1} \geq 0$. □

2.1 Differentiation

Definition: Let $f : X \to \mathbb{R}$ be a function and $x_0 \in X$. Then f attains a **local maximum** at x_0 iff there exists $\delta > 0$ such that the restriction of f to $X \cap (x_0 - \delta, x_0 + \delta)$ attains a maximum at x_0. Similarly, f attains a **local minimum** at x_0 iff there exists $\delta > 0$ such that the restrction of f to $X \cap (x_0 - \delta, x_0 + \delta)$ attains a minimum at x_0. △

Definition: Let $X \subseteq \mathbb{R}$, $x_0 \in X$ be a limit point of X, and $f : X \to \mathbb{R}$ be a function. If the limit

$$\lim_{x \to x_0; x \in X - \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} \tag{2}$$

converges to some $L \in \mathbb{R}$, then we say f is **differentiable at x_0 on X with derivative L,** and write $f'(x_0) = L$. △

Rolle's Theorem: Let $f : [a, b] \to \mathbb{R}$ be a function which is continuous on $[a, b]$ and differentiable on (a, b). Suppose also that $f(a) = f(b)$. Then there exists $x_0 \in (a, b)$ such that $f'(x_0) = 0$. △

Proof:
Since $f : [a, b] \to \mathbb{R}$ is continuous, according to the Extreme Value Theorem, it attains both a minimum and maximum on $[a, b]$. If the maximizers and minimizers occur at the endpoints, then $f : [a, b] \to \mathbb{R}$ is constant and so $f'(x) = 0$ for all $x \in (a, b)$. Otherwise, the function has either a

maximizer or a minimizer at some point $x_0 \in (a,b)$. First suppose that f has a maximizer at x_0. Then, using the definition of the derivative,

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

Then for $x < x_0$ we have

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

and for $x > x_0$

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0.$$

Hence

$$0 \leq \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \leq 0$$

and so $f'(x_0) = 0$. Similar argument applies if x_0 is a minimizer, but with the inequalities reversed. This completes the proof. □

Mean Value Theorem: Let $f : [a,b] \to \mathbb{R}$ be a function which is continuous on $[a,b]$ and differentiable on (a,b). Then there exists $c \in (a,b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{3}$$

△

Proof:
Define $g : [a,b] \to \mathbb{R}$ by $g(x) = f(x) - mx$ for some constant $m \in \mathbb{R}$. Note that g is differentiable since f is differentiable and so also is mx. We seek to apply Rolle's Theorem, which requires $g(b) = g(a)$. Then note

$$g(a) = g(b) \quad \Leftrightarrow \quad f(a) - ma = f(b) - mb \quad \Leftrightarrow \quad m(b-a) = f(b) - f(a) \quad \Leftrightarrow \quad m = \frac{f(b) - f(a)}{b - a}.$$

Then by Rolle's theorem there is $c \in (a,b)$ such that $g'(c) = 0$. Thus, $g'(c) = f'(c) - m = 0$, which implies the desired relation, i.e.,

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

□

2.2 Fixed Point Methods

Definition: For metric spaces X and Y, a function $f : X \to Y$ is a **contraction** if it is Lipschitz with constant $L < 1$. △

Banach Fixed Point Theorem: Let (X, d) be a non-empty complete metric space with the contraction mapping $T : X \to X$. Then T admits a unique fixed-point $x^* \in X$, i.e., $T(x^*) = x^*$. Moreover, for arbitrary $x_0 \in X$, defining $\{x_n\}$ by $x_{n+1} = T(x_n)$ converges to x^*. △

Proof:
First we will show that such an x^* exists. Let $x_0 \in X$. Then define $x_{n+1} = f(x_n)$ and note this implies

$$d(x_{n+2}, x_{n+1}) = d(f(x_{n+1}), f(x_n)) \leq Ld(x_{n+1}, x_n).$$

Thus, by induction, $d(x_{n+1}, x_n) \leq L^n d(x_1, x_0)$. Then for $m > n$ we have

$$\begin{aligned} d(x_m, x_n) &\leq d(x_n, x_{n+1}) + \cdots + d(x_{m-1}, x_m) \\ &\leq \sum_{i=n}^{m-1} L^i d(x_0, x_1) \\ &\leq L^n \cdot d(x_0, x_1) \cdot \frac{1}{1-L}. \end{aligned}$$

However, because $L < 1$, we see $L^n \cdot d(x_0, x_1)/(1-L) \longrightarrow 0$ as $n \longrightarrow \infty$, and so, by the comparison lemma, $\{x_n\}_{n=1}^{\infty}$ is Cauchy. Since X is complete, it follows that $\{x_n\}_{n=1}^{\infty}$ has a limit in X, which we denote by x^*. Then

$$x^* = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right) = f(x^*)$$

where we can bring the limit into the argument of f since f is Lipschitz and, therefore, continuous. Thus, $x^* = f(x^*)$.

All that remains is to show uniqueness of x^*. Suppose also we have $y^* \in X$ such that $f(y^*) = y^*$. If $d(x^*, y^*) > 0$, then, using the fact that $L < 1$,

$$d(x^*, y^*) = d(f(x^*), f(y^*)) \leq Ld(x^*, y^*) < d(x^*, y^*).$$

But, $d(x^*, y^*) < d(x^*, y^*)$ is a contradiction. Hence $d(x^*, y^*) = 0$ and so $x^* = y^*$. Hence the fixed point must be unique. □

2.3 Integration

Definition: Let $a, b \in \mathbb{R}$ with $a < b$. A **partition** of the closed interval $[a, b]$ is a finite sequence of numbers $x_0, \ldots, x_N$ such that $a = x_0 < x_1 < \cdots < x_N = b$. The **width** of the partition is defined to be

$$\text{width } := \max\{x_i - x_{i-1} \mid i = 1, 2, \ldots, N\}.$$

△

Definition: Let $f : [a, b] \to \mathbb{R}$ and $x_0, \ldots, x_N$ be a partition of $[a, b]$. Then a **Riemann sum** for f corresponding to the given partition is given by

$$\sum_{i=1}^{N} f(x_i')(x_i - x_{i-1})$$

where $x_{i-1} \leq x_i' \leq x_i$ for each $i = 1, \ldots, N$. △

Definition: We define the upper and lower Riemann sums, respectively, of f with respect to a partition $P = \{I_1, \ldots, I_n\}$ of a bounded interval $[a, b]$ by

$$U(f; P) = \sum_{k=1}^{n} \left(\sup_{I_k} f \right) |I_k| \quad \text{and} \quad L(f; P) = \sum_{k=1}^{n} \left(\inf_{I_k} f \right) |I_k|. \tag{4}$$

Let Π denote the collection of all partitions of I. We define the upper and lower Riemann integrals of f on I by

$$U(f) = \inf_{P \in \Pi} U(f; P) \quad \text{and} \quad L(f) = \sup_{P \in \Pi} U(f; P). \tag{5}$$

△

Remark: One way to define Riemann integrability is as follows. A bounded function $f : I \to \mathbb{R}$ defined on a bounded interval $[a, b]$ is **Riemann integrable** on $[a, b]$ if its upper integral $U(f)$ and lower integral $L(f)$ are equal. Below is a more concise definition of Riemann integrable △

Definition: Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$. We say that f is **Riemann integrable** on $[a, b]$ if there exists $A \in \mathbb{R}$ such that, for any $\varepsilon > 0$, there is a $\delta > 0$ such that $|S - A| < \varepsilon$ whenever S is a Riemann sum for f corresponding to any partition of $[a, b]$ of width less than δ. In this case, A is called the Riemann integral of f between a and b and is denoted by $\int_a^b f(x)\, \mathrm{d}x$. △

Lemma: A function $f : [a, b] \to \mathbb{R}$ is Riemann integrable on $[a, b]$ iff, given $\varepsilon > 0$, there exists $\delta > 0$ such that $|S_1 - S_2| < \varepsilon$ whenever S_1 and S_2 are Riemann sums for f corresponding to partitions of $[a, b]$ of width less than δ. △

Change of variables theorem: Let U, V be open intervals in $\mathbb{R}$, $\phi : U \to V$ be continuously differentiable, and $f : V \to \mathbb{R}$ be continuous. Then for any $a, b \in U$

$$\int_{\phi(a)}^{\phi(b)} f(v)\, \mathrm{d}v = \int_a^b f(\phi(u))\phi'(u)\, \mathrm{d}u.$$

△

Proof:

Let $F : V \to \mathbb{R}$ be defined by $F(y) = \int_{\phi(a)}^{y} f(v)\, \mathrm{d}v$ for all $y \in V$. Then F is differentiable and $F' = f$. The function $G : U \to \mathbb{R}$ defined by

$$G(x) = \int_{\phi(a)}^{\phi(x)} f(v)\, \mathrm{d}v \tag{6}$$

is the composition $G = F \circ \phi$ of two differentiable functions and, thus, is differentiable itself. By the chain rule, $G'(x) = F'(\phi(x))\phi'(x) = f(\phi(x))\phi'(x)$ for all $x \in U$. Hence we may write

$$G(x) = \int_a^x f(\phi(u))\phi'(u)\mathrm{d}u + c \tag{7}$$

for some $c \in \mathbb{R}$. Equation (6) implies $G(a) = 0$, and so $c = 0$. Then taking $x = b$, we may equate the right hand sides of (6) and (7) to obtain the desired relation.

□

First Fundamental Theorem of Calculus: Let $f : [a, b] \to \mathbb{R}$ be continuous and define $F : [a, b] \to \mathbb{R}$ by

$$F(x) = \int_a^x f(t)\ \mathrm{d}t.$$

Then F is uniformly continuous on $[a, b]$, differentiable on (a, b), and $F'(x) = f(x)$ for all $x \in (a, b)$. $\triangle$

Second Fundamental Theorem of Calculus: Let $f[a, b] \to \mathbb{R}$ and $F[a, b] \to \mathbb{R}$ be such that F is differentiable and $F'(x) = f(x)$ for all $x \in (a, b)$. If f is Riemann integrable on $[a, b]$, then

$$\int_a^b f(x)\ \mathrm{d}x = F(b) - F(a).$$

$\triangle$

Mean Value Theorem for Integrals: Suppose the function $f : [a, b] \to \mathbb{R}$ is continuous. Then there exists $x_0 \in [a, b]$ at which

$$\int_a^b f = f(x_0) \cdot (b - a).$$

$\triangle$

Cauchy Mean Value Theorem: Suppose $f[a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ are continuous and that their restrictions to the open interval (a, b) are differentiable and that $g'(x) \neq 0$ for all $x \in (a, b)$. Then there is a point $x_0 \in (a, b)$ at which

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_0)}{g'(x_0)}.$$

$\triangle$

Archimedes-Riemann Theorem: Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable on $[a, b]$ iff there is a sequence of partitions $\{P_n\}$ of the interval $[a, b]$ such that

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

Moreover, for such a sequence of partitions,

$$\lim_{n\to\infty} L(f, P_n) = \lim_{n\to\infty} U(f, P_n) = \int_a^b f.$$

△

First Fundamental Theorem of Calculus: Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable. Let $F : [a, b] \to \mathbb{R}$ be the function

$$F(x) := \int_{[a,x]} f. \tag{8}$$

Then F is continuous. Also, if $x_0 \in [a, b]$ and f is continuous at x_0, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$. △

Second Fundamental Theorem of Calculus: Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable. If $F : [a, b] \to \mathbb{R}$ is an antiderivative of f, then

$$\int_{[a,b]} f = F(b) - F(a). \tag{9}$$

△

Definition: Let $E \subseteq \mathbb{R}$ and $f : E \to \mathbb{R}$ be a function. If a is an interior point of E, we say that f is **real analytic at** a if there exists an open interval $(a - r, a + r)$ in E for some $r > 0$ such that there exists a power series centered at a which has a radius of convergence greater than or equal to r, and which converges to f on $(a - r, a + r)$. If E is an open set, and f is real analytic at every point a of E, then f is **real analytic on** E**.** △

Sequences and series of functions

Definition: Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from one metric space (X, d_X) to another (Y, d_y), and let $f : X \to Y$ be another function. We say that $(f^{(n)})_{n=1}^{\infty}$ **converges pointwise to** f **on** X if we have

$$\lim_{n\to\infty} f^{(n)}(x) = f(x) \tag{10}$$

for all $x \in X$, i.e.,

$$\lim_{n\to\infty} d_Y(f^{(n)}(x), f(x)) = 0. \tag{11}$$

△

Definition: Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from one metric space (X, d_X) to another (Y, d_Y), and let $f : X \to Y$ be another function. Then $(f^{(n)})_{n=1}^{\infty}$ **converges uniformly to** f **on** X if for every $\varepsilon > 0$ there exists $N > 0$ such that $d_Y(f^{(n)}(x), f(x)) < \varepsilon$ for every $n > N$ and $x \in X$. The function f is the **uniform limit** of the functions $f^{(n)}$. △

Uniform convergence and integration

Leibniz differentiation under the integral sign: Let $f(x,t)$ be a function such that both $f(x,t)$ and its partial derivative $f_x(x,t)$ are continuous in t and x in some region of the (x,t)-plane, including $a(x) \le t \le b(x)$, $x_0 \le x \le x_1$. Also suppose that the functions $a(x)$ and $b(x)$ are both continuous and both have continuous derivatives for $x_0 \le x \le x_1$. Then for $x_0 \le x \le x_1$ we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_{a(x)}^{b(x)} f(x,t)\ \mathrm{d}t\right) = f(x,b(x)) \cdot b'(x) - f(x,a(x)) \cdot a'(x) + \int_{a(x)}^{b(x)} \frac{\partial f}{\partial x}(x,t)\ \mathrm{d}t.$$

$\triangle$

2.4 Metric Space Topology and Analysis

Definition: An order $(L, <)$ is **Dedekind complete** if the following two conditions hold:

i) Every $A \subseteq L$ which is bounded above has a supremum in L, mean a $<$-least element $z \in L$ such that $x \in A \;\Rightarrow\; x \leq z$.

ii) Every $A \subseteq L$ which is bounded below has an infimum in L, meaning a $<$-greatest element $z \in L$ such that $x \in A \;\Rightarrow\; z \leq x$.

$\triangle$

Proposition: $\mathbb{R}$ is Dedekind complete.

Proof:

Let $A \subseteq \mathbb{R}$ be nonempty and bounded above, and let $A' = \{x \mid \exists\, y \in A, x \leq y\}$. Indeed, every upper bound for A is also an upper bound for A', and vice versa. To show that A contains a supremum in $\mathbb{R}$, it then suffices to find the supremum of A'. If A' has a largest element, this element forms the supremum and we are done. Now suppose otherwise. Let $f : \mathbb{N} \to \mathbb{Q}$ be a bijection, and for each $n \in \mathbb{N}$ let $a_n = \max\{f(i) \in A' \mid 1 \leq i \leq n\}$. Then $\{a_n\}_{n=1}^{\infty}$ is nonstrictly increasing. And, the sequence does not have a largest element due to the density of $\mathbb{Q}$ and the fact that A' does not have a largest element. We may choose a subsequence $\{a'_n\}_{n=1}^{\infty}$ of $\{a_n\}_{n=1}^{\infty}$, which is strictly increasing. Using this, and the fact that A' is bounded, we see $[a'_n] \in \mathbb{R}$. And $[a'_n]$ provides an upper bound for A'. Hence we have found our supremum. The proof for the infimum follows similarly. (I am not sure about this proof. Seems to be lacking a couple details.) $\square$

Definition: Any bijection between ordered sets A and B is an **equivalence relation** if it is *symmetric, reflexive,* and *transitive.* $\triangle$

Definition: A **metric space** (X, d) is a space X of objects, together with a metric $d : X \times X \to [0, +\infty)$ for which the following four axioms also hold:

a) For any $x \in X$, $d(x, x) = 0$.

b) (Positivity) For any distinct $x, y \in X$, $d(x, y) > 0$.

c) (Symmetry) For any $x, y \in X$, $d(x, y) = d(y, x)$.

d) (Triangle Inequality) For any $x, y, z \in X$, $d(x, z) \leq d(x, y) + d(y, z)$.

△

Definition: Let (X, d) be a metric space. All points and sets mentioned below are understood to be elements and subsets of X.

i) A **neighborhood** of p is a set consisting of all q such that $d(p, q) < r$, for some $r > 0$. The number r is called the *radius* of this set.

ii) A point p is a **limit point** of the set E if every neighborhood of p contains a point $q \neq p$ such that $q \in E$.

iii) If $p \in E$ and p is not a limit point of E, then p is called an **isolated point** of E.

iv) E is **closed** if every limit point of E is a point of E.

v) A point p is an **interior point** of E if there is a neighborhood N of p such that $N \subset E$.

vi) E is **open** if every point of E is an interior point of E.

vii) The **complement** of E, denoted E^c, is the set of all points $p \in X$ such that $p \notin E$.

viii) E is **perfect** if E is closed and if every point of E is a limit point of E.

ix) E is **bounded** if there is a real number M and a point $q \in X$ such that $d(p, q) < M$ for all $p \in E$.

x) E is **dense** in X if every point of X is a limit point of E, or a point of E (or both).

△

Definition: Let (X, d) be a metric space, $x \in X$, and $r > 0$. We define the **ball** $B_{X,d}(x_0, r)$ in X, centered at x_0, and with radius r, in the metric d, to be the set

$$B_{(X,d)}(x_0, r) := \{x \in X \ \mid \ d(x, x_0) < r\}. \tag{12}$$

△

Every neighborhood is an open set. △

Proof:
Consider a neighborhood $E = B(p, r)$, and let $q \in E$. Then there exists a positive real number h such that $d(p, q) = r - h$. For all points s such that $d(q, s) < h$, we have $d(p, s) \leq d(p, q) + d(q, s) < r - h + h = r$, and so $s \in E$. Thus, q is an interior point of E. □

If p is a limit point of a set E, then every neighborhood of p contains infinitely many points of E. △

Proof:
Suppose there is a neighborhood N of p which contains only a finite number of points of E. Let $q_1, \ldots, q_n$ denote these points of $N \cap E$, which are distinct from p. Then define $r := \min_{1 \leq m \leq n} d(p, q_m)$, which is positive due to the fact that $d(p, q_m) > 0$ whenever q_m and p are distinct. Then the neighborhood $N_r(p)$ contains no point q of E such that $q \neq p$, which implies p is not a limit point of E. This contradicts our hypothesis and so the theorem follows. □

Definition: Let (X, d) be a metric space. Then X is **totally bounded** iff for every $\varepsilon > 0$ there exists a finite set of points $x_1, \ldots, x_n \in X$ such that $X = \cup_{i=1}^n B_\varepsilon(x_i)$. △

The metric space (X, d) is sequentially compact iff it is compact. △

Proof:
First assume X is compact and fix a sequence $\{x_n\}_{n=1}^\infty$ in X. By way of contradiction, suppose $\{x_n\}_{n=1}^\infty$ has no convergent subsequence. So, no $z \in X$ is a limit of a subsequence of $\{x_n\}_{n=1}^\infty$. This implies there exists $r_z > 0$ such that $B(z, r_z) \cap \{x_n \mid n \in \mathbb{N}\} \subseteq \{z\}$. The collection of sets $B(z, r_z)$ for $z \in X$ form an open cover of X. By the compactness of X, there is a finite subcover of the form

$$X \subseteq B(z_1, r_{z_1}) \cup \cdots \cup B(z_k, r_{z_k}).$$

But, then $X \cap \{x_n \mid n \in \mathbb{N}\} \subseteq \{z_1, \ldots, z_k\}$. So, each x_n is contained in $\{z_1, \ldots, z_k\}$. It follows from the pigeonhole principle that there is a z_j with $1 \leq j \leq k$ that shows up infinitely many times in the sequence $\{x_n\}_{n=1}^\infty$. Then there is a subsequence that is constantly z_j. This subsequence then converges to z_j, which contradicts our assumption. The result follows.

Now suppose X is sequentially compact. Let U be an open cover of X and, by way of contradiction, suppose there is no finite subcover of U. Now extract a countable subcover $\{V_1, V_2, \ldots\}$ of U. We construct a sequence as follows. Pick $x_1 \in V_1$. Then for each successive $n \in \mathbb{N}$, pick $x_n \notin V_1 \cup \cdots \cup V_{n-1}$, which is possible since $V_1 \cup \cdots \cup V_{n-1}$ does not cover X. Now, by assumption, X is sequentially compact, which implies there is a subsequence of $\{x_n\}_{n=1}^\infty$ that converges to some $x \in X$. Because $\{V_1, V_2, \ldots\}$ forms a cover for X, there exists V_m such that $x \in V_m$. But, then for each $\varepsilon > 0$ such that $B(x, \varepsilon) \subset V_m$, there is $n \in \mathbb{N}$ such that $x_n \in B(x, \varepsilon)$ and so $x_n \in V_m$. This implies there is an infinite number of terms in U_m, which contradicts the construction of $\{x_n\}_{n=1}^{\infty}$. Thus, the supposition that there is no finite subcover of X was false. Hence sequential compactness implies compactness. $\square$

Definition: Let (X, d) be a metric space, let $E \subseteq X$, and let $x_0 \in X$. Then x_0 is an **interior point** of E if there exists $r > 0$ such that $B(x_0, r) \subseteq E$. We say x_0 is an **exterior point** of E if there exists $r > 0$ such that $B(x_0, r) \cap E = \emptyset$. We say x_0 is a **boundary point** of E if it is neither an interior point nor an exterior point of E. $\triangle$

Definition: Let (X, d) be a metric space, $E \subseteq X$, and $x_0 \in X$. Then x_0 is an **adherent point** of E if for each $r > 0$, the ball $B(x_0, r)$ has a non-empty intersection with E, i.e., $B(x_0, r) \cap E \neq \emptyset$. The set of all adherent points of E is called the **closure** of E and is denoted $\overline{E}$. $\triangle$

Definition: Let (X, d) be a metric space and $E \subseteq X$. Then E is **closed** if it contains all of its boundary points, i.e., $\partial E \subseteq E$. We say E is **open** if it contains none of its boundary points, i.e., $\partial E \cap E = \emptyset$. If E contains some of its boundary points but not others, then it is neither open nor closed. $\triangle$

Theorem:

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

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

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

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

$\triangle$

Let (X, d) be a metric space and $E \subseteq X$.

a) E is open iff $E = \text{int}E$, i.e., for each $x \in E$ there exists $r > 0$ such that $B(x, r) \subseteq E$.

b) E is closed iff E contains all its adherent points, i.e., for every convergent sequence $(a_n)_{n=m}^{\infty}$ in E, the limit $\lim_{n\to} a_n$ of that sequence also lies in E.

c) For any x_0 and $r > 0$, the ball $B(x_0, r)$ is an open set. The closed ball $\{x \in X \mid d(x, x_0) \leq r\}$ is a closed set.

d) If $x_0 \in X$, then the singleton set $\{x_0\}$ is closed.

e) E is open iff $X \backslash E$ is closed.

f) If $E_1, \ldots, E_n$ are open sets in X, then $E_1 \cap \cdots \cap E_n$ is also open. If $F_1, \ldots, F_n$ is a finite collection of closed sets in X, then $F_1 \cup \ldots \cup F_n$ is also closed.

g) If $\{E_\alpha\}_{\alpha \in I}$ is a collection of open sets in X, then $\cup_{\alpha \in I} E_\alpha$ is also open. If $\{F_\alpha\}_{\alpha \in I}$ is a collection of closed sets in X, then $\cap_{\alpha \in I} F_\alpha$ is also closed.

h) $\text{int}(E)$ is the largest open set contained in E. $\bar{E}$ is the smallest closed set which contains E.

△

Definition: Let (X, d) be a metric space and $Y \subseteq X$ and $E \subseteq Y$. Then E is **relatively open with respect to** Y if it is open in the metric space $(Y, d_{Y \times Y})$. We say E is **relatively closed with respect to** Y if it is closed in the metric space $(Y, d_{Y \times Y})$. △

Suppose $Y \subset X$. A subset of E of Y is open relative to Y iff $E = Y \cap G$ for some open subset G of X. △

Definition: By an **open cover** of a set E in a metric space X, we mean a collection $\{G_\alpha\}$ of open subsets of X such that $E \subset \cup_\alpha G_\alpha$. △

Definition: A subset K of a metric space X is said to be **compact** if every open cover of K contains a *finite* subcover. △

If $E \subset \mathbb{R}^n$, then the following properties are equivalent:

i) E is closed and bounded.

ii) E is compact.

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

△

Definition: Two subsets A and 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 E is *not* a union of two nonempty separated sets. $\triangle$

Definition: A space X is **path connected** if for all $x_0, x_1 \in X$, there is a continuous function $f : [0,1] \to X$ with $f(0) = x_0$ and $f(1) = x_1$. $\triangle$

Proposition: Path connectedness implies connectedness. $\triangle$

Proof:
Suppose that a space X is not connected. Then there exists nonempty disjoint open subsets $A_0, A_1 \subseteq X$ such that $X = A_0 \cup A_1$. Then fix $x_0 \in A_0$ and $x_1 \in A_1$. By the assumption of path connectedness, there is a continuous function $f : [0,1] \to X$ with $f(0) = x_1$ and $f(1) = x_1$. Then $f^{-1}(A_0)$ and $f^{-1}(A_1)$ are nonempty since $0 \in f^{-1}(A_0)$ and $1 \in f^{-1}(A_1)$. These sets $f^{-1}(A_0)$ and $f^{-1}(A_1)$ are open by continuity of f, and disjoint by definition of A_0 and A_1, and have union $[0,1]$ since $A_0 \cup A_1 = X$. This contradicts the fact that $[0,1]$ is connected. Thus, we cannot have path connectedness. The hypothesis follows through contraposition. $\square$

A subset E of $\mathbb{R}$ is connected iff it has the following property: If $x, y \in E$ and $x < z < y$, then $z \in E$. $\triangle$

Proposition: The subset $[a,b] \subset \mathbb{R}$ is connected. $\triangle$

Proof:
By way of contradiction, suppose there exist disjoint non-empty open sets A and B such that $A \cup B = [a,b]$. Without loss of generality, suppose $b \in B$. Clearly, $[a,b]$ is bounded. Thus, A is bounded and by the least upper bound principle, we can define $c = \sup(A)$. Since $[a,b]$ is closed, $c \in [a,b]$.

First suppose $c \in A$. Then $c < b$ since $b \in B$ and $A \cap B = \emptyset$. And, since A is open, there exists $\varepsilon > 0$ so that $B(c,\varepsilon) \cap [a,b] \subseteq A$. But, then $c + \min\{\varepsilon, b - \varepsilon\} \in A$, which contradicts the fact that $c = \sup(A)$.

Now suppose $c \in B$. Note then $c \neq a$ since then we'd have $A = \{0\}$, which is closed. Thus, $c \in (a,b]$. Since B is open, there exists $\varepsilon > 0$ such that $B(c,\varepsilon) \cap [a,b] \subseteq B$. But then $c - \min\{\varepsilon, a\}$ is an upper bound for A, again contradicting that $c = \sup A$. Hence $[a,b]$ must be connected. $\square$

Definition: A metric space (X, d) is **complete** iff every Cauchy sequence in (X, d) is in fact convergent in (X, d). △

Let (X, d) be a metric space.

a) Let $(Y, d_{Y \times Y})$ be a subspace of (X, d). If $(Y, d_{Y \times Y})$ is complete, then Y must be closed in X.

b) Suppose (X, d) is complete and $Y \subseteq X$ is closed. Then $(Y, d_{Y \times Y})$ is also complete.

△

Definition: A metric space (X, d) is **sequentially compact** iff every sequence in (X, d) has at least one convergent subsequence. A subset $Y \subseteq X$ is said to be compact if the subspace $(Y, d_{Y \times Y})$ is compact. △

Definition: Let (X, d) be a metric space, and let $Y \subseteq X$. Then Y is **bounded** iff there exists a ball $B(x, r)$ in X which contains Y. △

Definition: Let (X, d) be a compact metric space. Then (X, d) is both complete and bounded. △

Lemma: $C[0, 1]$ is complete. △

Proof:
Let $\{f_n\}$ be a sequence in $C[0, 1]$. We proceed in three parts. First we show the point-wise limit function f exists. Then we show $\{f_n\}$ converges to this limit f in operator norm. Lastly, we verify that f is continuous and, thus, is in $C[0, 1]$.

Let $x \in [0, 1]$ and $\varepsilon > 0$ be given. Then since $\{f_n\}$ is Cauchy, there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$ we have

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\| \leq \varepsilon,$$

which implies that $\{f_n(x)\}$ is Cauchy in $\mathbb{R}$. Since $\mathbb{R}$ is complete, $\lim_{n\to\infty} f_n(x)$ exists. Since this holds for each $x \in [0, 1]$, we define $f : [0, 1] \to \mathbb{R}$ by $f(x) = \lim_{n\to\infty} f_n(x)$ for each $x \in [0, 1]$.

To show $\{f_n\}$ converges to f in operator norm, given $\varepsilon > 0$, we must find $N \in \mathbb{N}$ such that $\|f_n - f\| \leq \varepsilon$ whenever $n \geq N$. Since $\{f_n\}$ is Cauchy, there is $N \in \mathbb{N}$ such that $\|f_n - f_m\| \leq \varepsilon/2$ whenever $n, m \geq N$. So, since the metric of a vector space is continuous, $\lim_{n\to\infty} \|f_n - f_N\| = \|f - f_N\|$ and so $\|f - f_N\| \leq \varepsilon/2$. Thus,

$$\|f - f_n\| \leq \|f_n - f_N\| + \|f - f_N\| \leq \varepsilon$$

whenever $n \geq N$.

Now we show f is continuous. Since each f_n is continuous, for each $x_0 \in [0,1]$ there is $\delta > 0$ such that $|f_n(x) - f_n(x_0)| \leq \varepsilon/3$ whenever $x \in [0,1]$ with $|x - x_0| \leq \delta$. So, Since the f_n converge to f in operator norm, there is an $N \in \mathbb{N}$ such that $\|f_n - f\| \leq \varepsilon/3$ whenever $n \geq N$. Thus, if $|x - x_0| \leq \delta$, then

$$\begin{aligned}
|f(x) - f(x_0)| &\leq |f(x) - f_N(x)| + |f_n(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\
&\leq \|f - f_N\| + |f_n(x) - f_n(x_0)| + \|f_N - f\| \\
&\leq \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\
&= \varepsilon,
\end{aligned}$$

and so f is continuous at x_0. Since x_0 was arbitrary in $[0,1]$, it follows that f is continuous on $[0,1]$. Thence $f \in C[0,1]$ and we are done. □

Let (X,d) be a metric space and $Y \subseteq X$ be compact. Let $(V_i)_{i\in I}$ be a collection of open sets in X that covers Y, i.e.,

$$Y \subseteq \cup_{i\in I} V_i. \tag{13}$$

Then there exists a finite subset $F \subseteq I$ such that

$$Y \subseteq \cup_{i\in F} V_i. \tag{14}$$

△

Definition: A **Baire space** is a topological space with that property that for each countable collection of open dense sets $\{U_n\}_{n=1}^{\infty}$, their intersection $\cap_{n=1}^{\infty} U_n$ is dense. △

Baire Category Theorem: Every complete metric space is a Baire space. Equivalently, a non-empty complete metric space is not the countable union of nowhere-dense closed sets. Equivalently, if a non-empty complete metric space is the countable union of closed sets, then one of these closed sets has non-empty interior.[1] △

Baire Category Theorem: Every locally compact Hausdorff space is a Baire space. △

Stone-Weierstrass Theorem: Suppose $f[a,b] \to \mathbb{R}$ is continuous. Then $\forall\ \varepsilon > 0$, there exists a polynomial $p(x)$ such that $\forall\ x \in [a,b]$, $|f(x) - p(x)| < \varepsilon$, or equivalently, $\|f - p\|_\infty < \varepsilon$. △

Taylor's Theorem: Let $f : [a,b] \to \mathbb{R}$ and suppose $f^{(n-1)}$ is continuous on $[a,b]$ and $f^{(n)}$ exists on (a,b) for some $n \in \mathbb{N}$. Also let $x_0 \in (a,b)$. Then there exists a point ξ between x_0 and $x \in [a,b]$ such that

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n.$$

△

Cauchy Integral Remainder Formula: Let $f : [a,b] \to \mathbb{R}$ and suppose f has n derivatives and $f^{(n)}$ is continuous on $[a,b]$ for some $n \in \mathbb{N}$. Then for each point $x \in [a,b]$

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{1}{n!}\int_{x_0}^{x} f^{(n)}(t)(x - t)^n \ \mathrm{d}t.$$

△

Proof:

Use induction and integration by parts. The base case if the fundamental theorem of calculus. □

[1] See Kreyzig's text on this. It provides solid explanation.

3 Linear Algebra

Definition: The **rank** is the dimension of the image of an operator and the **nullity** is the dimension of the kernel of an operator. △

Rank-Nullity Theorem: Let V be a finite dimensional vector space and $T : V \to V$ be linear. Then

$$\dim V = \dim \mathrm{im} T + \dim \ker T. \tag{15}$$

This is also known as the *Fundamental Theorem of Linear Maps* in Axler's text. △

Proof:
Let $u_1, \ldots, u_m$ be a basis of $\ker T$ so that $\dim \ker T = m$. This can be extended to a basis $u_1, \ldots, u_m, v_1, \ldots, v_n$ of V where $\dim V = m + n$. So, we need only show $\dim \mathrm{im} T = n$, which we do by showing $Tv_1, \ldots, Tv_n$ is a basis of $\mathrm{im} T$.

Let $v \in V$. Then there are unique scalars $a_1, \ldots, a_m, b_1, \ldots, b_n$ such that

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n. \tag{16}$$

Then

$$\begin{aligned} Tv &= T\left(a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n\right) \\ &= a_1 T(u_1) + \cdots + a_m T(u_m) + b_1 T(v_1) + \cdots + b_n T(v_n) \\ &= b_1 T(v_1) + \cdots + b_n T(v_n) \end{aligned} \tag{17}$$

where the final equality holds since $u_1, \ldots, u_m \in \ker T$. This shows $\mathrm{im} T$ is spanned by $Tv_1, \ldots, Tv_n$. All that remains is to show these vectors are linearly independent. Thence suppose there are $c_1, \ldots, c_n$ such that

$$0 = c_1 T(v_1) + \cdots + c_n T(v_n) = T(c_1 v_1 + \cdots c_n v_n) \tag{18}$$

where the second equality holds by linearity of T. This implies $c_1 v_1 + \cdots + c_n v_n \in \ker T$. Since $u_1, \ldots, u_m$ form a basis for $\ker T$, there are scalars $d_1, \ldots, d_m$ such that

$$c_1 v_1 + \cdots c_n v_n = d_1 u_1 + \cdots d_m u_m \quad \Rightarrow \quad 0 = c_1 v_1 + \cdots c_n v_n - d_1 u_1 - \cdots - d_m u_m. \tag{19}$$

Since $v_1, \ldots, v_n, u_1, \ldots, u_m$ forms a basis for V, these vectors are linearly independent and so all the c_i's and d_i's must be zero. Returning to (18), we have that $T(v_1), \ldots, T(v_n)$ are linearly independent, as desired. This completes the proof. □

Lemma: Suppose $T : V \to V$ is linear. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of T and $v_1, \ldots, v_m$ be corresponding eigenvectors. Then $\{v_1, \ldots, v_m\}$ is linearly independent $\triangle$

Proof:
By way of contradiction, suppose otherwise. Let k be the smallest index such that $v_k \in \text{span}(v_1, \ldots, v_{k-1})$. Then there are $a_1, \ldots, a_{k-1}$ not all zero such that

$$v_k = a_1 v_1 + \cdots + a_{k-1} v_{k-1}. \tag{20}$$

Then

$$\begin{aligned} \lambda_k v_k &= T\left(a_1 v_1 + \cdots + a_{k-1} v_{k-1}\right) \\ &= a_1 T(v_1) + \cdots + a_{k-1} T(v_{k-1}) \\ &= a_1 \lambda_1 v_1 + \cdots + a_{k-1} \lambda_{k-1} v_{k-1}. \end{aligned} \tag{21}$$

Multiplying (20) by λ_k and subtracting (21) we obtain

$$0 = (\lambda_k - \lambda_k) v_k = a_1(\lambda_k - \lambda_1) v_1 + \cdots + a_{k-1}(\lambda_k - \lambda_{k-1}) v_{k-1}. \tag{22}$$

By hypothesis, $\lambda_k - \lambda_j \neq 0$ for $j = 1, \ldots, k-1$. Furthermore, by our initial assumption, $\{v_1, \ldots, v_{k-1}\}$ is linearly independent. Hence $a_1 = \cdots = a_{k-1} = 0$. But, this implies $v_k = 0$, which cannot be the case since eigenvectors are, by definition, nonzero. Hence the initial assumption was false and we conclude $\{v_1, \ldots, v_m\}$ is linearly independent. $\square$

Schur's Theorem: Suppose V is a finite-dimensional complex vector space and $T : V \to V$ is a linear operator. Then T has an upper-triangular matrix with respect to some orthonormal basis of V. $\triangle$

Definition: Let V, W be finite dimensional vector spaces and suppose $T : V \to W$ is a linear operator. Then the **adjoint** of T is the function $T^* : W \to V$ such that $\langle Tv, w \rangle = \langle v, T^* w \rangle$ for every $v \in V$ and $w \in W$. $\triangle$

Proposition: Properties of the adjoint:

a) $(S + T)^* = S^* + T^*$ for $S, T \in \mathcal{L}(V, W)$.

b) $(\lambda T)^* = \overline{\lambda} T^*$ for all $\lambda \in \mathbb{F}$ and $T \in \mathcal{L}(V, W)$.

c) $(T^*)^* = T$ for all $T \in \mathcal{L}(V, W)$.

d) $I^* = I$ where I is the identity operator on V.

e) $(ST)^* = T^*S^*$ for all $T \in \mathcal{L}(V, W)$ and $S \in \mathcal{L}(W, U)$.

△

Proposition: More properties of the adjoint. Suppose $T \in \mathcal{L}(V, W)$. Then

a) null $T^* = (\text{range } T)^\perp$,

b) range $T^* = (\text{null } T)^\perp$,

c) null $T = (\text{range } T^*)^\perp$,

d) range $T = (\text{null } T^*)^\perp$.

△

Definition: An linear operator $T : V \to V$ is called **self-adjoint** if $T = T^*$, i.e., if $\langle Tv, w\rangle = \langle v, Tw\rangle$ for all $v, w \in V$. △

Proposition: Every eigenvalue of a self-adjoint operator is real. △

Proof:
Suppose $T : V \to V$ is self-adjoint. Let λ be an eigenvalue of T, and let v be a nonzero vector in V such that $Tv = \lambda v$. Then

$$\lambda\|v\|^2 = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\|v\|^2. \tag{23}$$

Thus, $\lambda = \overline{\lambda}$, which implies that λ is real, as desired. □

Definition: A linear operator $T : V \to V$ on an inner product space is **normal** if it commutes with its adjoint, i.e., if $TT^* = T^*T$. △

Complex Spectral Theorem: Suppose V is a finite dimensional vector space over $\mathbb{C}$ and $T : V \to V$ is a linear operator. Then the following are equivalent:

a) T is normal.

b) V has an orthonormal basis consisting of eigenvectors of T.

c) T has a diagonal matrix with respect to some orthonormal basis of V.

$\triangle$

Real Spectral Theorem: Suppose V is a finite dimensional vector space over $\mathbb{C}$ and $T : V \to V$ is a linear operator. Then the following are equivalent:

a) T is self-adjoint.

b) V has an orthonormal basis consisting of eigenvectors of T.

c) T has a diagonal matrix with respect to some orthonormal basis of V.

$\triangle$

Definition: A linear mapping $T : V \to V$ is called **positive** if T is self-adjoint and $\langle Tv, v\rangle \geq 0$ for all $v \in V$. $\triangle$

Definition: An operator R is called a **square root** of an operator T if $R^2 = T$. $\triangle$

Proposition: Let $T : V \to V$ be a linear mapping. Then the following are equivalent:

a) T is positive.

b) T is self-adjoint and all the eigenvalues of T are nonnegative.

c) T has a positive square root.

d) T has a self-adjoint square root.

e) There exists a linear operator $R : V \to R$ such that $T = R^*R$.

$\triangle$

Definition: A linear operator $S : V \to V$ is called an **isometry** if $\|Sv\| = \|v\|$ for all $v \in V$. If the inner product space is real, then S is also called an **orthogonal** operator. If the inner product space is complex, the S is also called a **unitary** operator. $\triangle$

Cayley-Hamilton Suppose V is complex vector space and $T : V \to V$ is linear. Let q denote the characteristic polynomial of T. Then $q(T) = 0$. △

Proof:
Let $\lambda_1, \ldots, \lambda_p$ be the distinct eigenvalues of T and $d_1, \ldots, d_p$ the dimensions of the corresponding generalized eigenspaces $E_{\lambda_1}, \ldots, E_{\lambda_p}$. We claim the restriction of $(T - \lambda_j I)$ to E_{λ_j} (i.e., $(T - \lambda_j I)|_{E_{\lambda_j}}$) is nilpotent for each $j \in \{1, \ldots, p\}$. This implies $(T - \lambda_j)^{d_j}\big|_{E_{\lambda_j}} = 0$. And, $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$. Thus, to prove $q(T) = 0$, it suffices to show $q(T)|_{E_{\lambda_j}} = 0$ for each j. So, fix any $j \in \{1, \ldots, p\}$. Then

$$q(T) = (T - \lambda_1 I)^{d_1} \cdots (T - \lambda_p I)^{d_p}. \tag{24}$$

Note the operators on the right hand side commute. This implies we can move $(T - \lambda_j I)^{d_j}$ do to be the last term on the right hand side. And, $(T - \lambda_j I)^{d_j}\big|_{E_{\lambda_j}} = 0$. Thence $q(T)|_{E_{\lambda_j}} = 0$, as desired.

All that remains is to verify $(T - \lambda_j I)|_{E_{\lambda_j}}$ is nilpotent. □

3.1 Matrix Exponentiation & Related Topics

Definition: The norm of a matrix A is defined by

$$\|A\| = \sup\{\|Ax\| \ \mid \ x \in \mathbb{R}^n \text{ and } \|x\| = 1\} \tag{25}$$

where $\|x\|$ denote the Euclidean norm in $\mathbb{R}^n$. △

Proposition: For matrices A, B, $\|AB\| \leq \|A\|\|B\|$. △

Proof:
For any $x \in \mathbb{R}^n$ with $\|x\| = 1$ observe that

$$\|A\| \geq \left\| A \frac{Bx}{\|Bx\|} \right\| = \frac{\|ABx\|}{\|Bx\|} \geq \frac{\|ABx\|}{\|B\|}, \tag{26}$$

which implies $\|ABx\| \leq \|A\|\|B\|$. Since x is an arbitrary unit vector in $\mathbb{R}^n$, we must have $\|AB\| \leq \|A\|\|B\|$. □

REMARK: By induction, the above proposition implies $\|A^n\| \leq \|A\|^n$. △

Definition: For a linear operator $T : V \to V$, where V is a finite dimensional inner product space, we define the **exponential** of T to be

$$\exp(T) = \sum_{n=0}^{\infty} \frac{T^n}{n!}.$$

△

Remark: Consider the IVP: $x'(t) = Ax$ and $x(t_0) = x_0$ where A is a square matrix with real or complex scalars. This can be solved using matrix exponentials. Indeed,

$$x(t) = x_0 \exp\left(A(t - t_0)\right).$$

Since $\exp(0) = 1_V$, this choice of $x(t)$ gives the correct initial value. Through expanding out the series, we also see that $x'(t) = Ax(t)$. △

Suppose $A \in M_{n\times n}(\mathbb{C})$. Prove that if $\|A\| < 1$, then $I - A$ is invertible. △

Proof:
We first show $I - A$ is invertible iff 1 is not an eigenvalue of A. We argue by proving the contrapositive of each implication in this claim. First suppose 1 is an eigenvalue of A. Then there is nonzero $v \in \mathbb{C}^n$ such that $(I - A)v = 0$, implying that $I - A$ is not one-to-one and, thus, not invertible. Hence if A is invertible, then 1 is not an eigenvalue of A. Now suppose A is singular. Then $\det(I - A) = 0$, which implies 1 is an eigenvalue of A. Hence if 1 is not an eigenvalue of A, then $I - A$ is invertible.

If 1 is an eigenvalue of A, then there is a unit vector $v \in \mathbb{C}^n$ such that $\|A\| \geq \|Av\| = \|v\| = 1$. Thus, if $\|A\| < 1$, then 1 is not an eigenvalue of A and so $\|A\| < 1$ implies $I - A$ is invertible. □

3.2 Determinants

Definition: The determinant of an $n \times n$ matrix $A = (a_{i,j})$ is given by

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

where sgn denotes the sign function and the sum occurs overall all permutations σ of the set $\{1, 2 \ldots, n\}$, which is the group S_n. △

Proposition: If A is an $n \times n$ matrix, then $\det(-A) = (-1)^n \det(A)$. △

Proof:
Let $B = (b_{ij}) = -A$. Then $b_{ij} = -a_{ij}$ and, using the definition of determinant,

$$\begin{aligned}
\det(-A) = \det(B) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} b_{i,\sigma(i)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} -a_{i,\sigma(i)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(-1)^n \prod_{i=1}^{n} a_{i,\sigma(i)} \\
&= (-1)^n \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \\
&= (-1)^n \det(A).
\end{aligned}$$

□

Proposition: The set of invertible matrices is dense in $M_{n \times n}(\mathbb{C})$. △

Proof:
First note the set of invertible matrices is nonempty since the identity matrix I is itself invertible with $I = I^{-1}$. Let $A \in M_{n \times n}(\mathbb{C})$ be given. If A is invertible, then we are done. So, suppose A is singular. Given $\varepsilon > 0$, we must show there is an invertible matrix within a distance ε of A. Then let $B \in M_{n \times n}(\mathbb{C})$ be invertible and define $L : \mathbb{R} \to M_{n \times n}(\mathbb{C})$ by $L(t) = (1-t)A + tB$. Now define $p : \mathbb{R} \to \mathbb{R}$ by $p(t) = \det(L(t))$ and note that p is a polynomial. Moreover, p is not identically zero since $p(1) = \det(B) \neq 0$. Also, $p(0) = \det(A) = 0$ and since p is a polynomial, it has only

finitely many zeros $t_0, \ldots, t_n$. So, define $d = \min\{|t_i| \mid i = 1, \ldots, n \text{ and } t_i \neq 0\}$. Then $p(t) \neq 0$ for $t \in (-d, d) - \{0\}$. Now define $\delta = \min\{d/2, \varepsilon/\|A - B\|\}$. Then $p(\delta) = \det(L(\delta)) \neq 0$ and so $L(\delta)$ is invertible. Moreover,

$$\|A - L(\delta)\| = \|A - ([1-\delta]A + \delta B)\| = \delta\|A - B\| \leq \varepsilon,$$

as desired. □

Canonical Forms Example: Suppose A is a 12×12 complex matrix with minimal polynomial $(t-\lambda)^6$ where $\lambda \in \mathbb{C}$ is an eigenvalue of A. Suppose also that $\dim(\ker\,(A - \lambda I)) = 4$ and $\dim(\ker\,(A - \lambda I))^2 = 6$. What is the Jordan form of A?

Proof:
Since $\dim(\ker\,(A - \lambda I)) = 4$, we know there are 4 Jordan blocks. Since $\dim(\ker\,(A - \lambda I)^2) - \dim(\ker\,(A - \lambda I)) = 2$, we know there are 2 Jordan blocks of size ≥ 2. Hence 2 of the Jordan blocks are of size 1. From the minimal polynomial, we know one Jordan block is of size 6. The remaining Jordan block must then be of size $12 - 1 - 1 - 6 = 4$, and we are done. □

If J is a matrix in Jordan form, then J^t and J are similar. △

Proof:
Here we provide a proof sketch. Simply note for a Jordan matrix

$$J = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

that

$$\begin{aligned}
&\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & \lambda \\ 0 & 1 & \lambda & 0 \\ 1 & \lambda & 0 & 0 \\ \lambda & 0 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 1 & \lambda & 0 & 0 \\ 0 & 1 & \lambda & 0 \\ 0 & 0 & 1 & \lambda \end{pmatrix} \\
&= J^t.
\end{aligned}$$

This completes the sketch. □

Example: Find an invertible matrix P such that $J = PAP^{-1}$ is in Jordan canonical form where

$$A = \begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix}.$$

Proof:
We first compute the eigenvalues of A to see

$$\begin{aligned}
\chi_A(\lambda) = \det(A - \lambda I) &= (-1)^{3+1}(5)(2\cdot 2 - [4-\lambda]) + (-1)^{3+3}(-\lambda)([-2-\lambda][4-\lambda] + 14) \\
&= 5\lambda - \lambda(\lambda^2 - 2\lambda + 6) \\
&= (-\lambda)(\lambda^2 - 2\lambda + 1) \\
&= (-\lambda)(\lambda - 1)^2.
\end{aligned}$$

So, the eigenvalues are 0 and 1 with multiplicity 1 and 2, respectively. The kernel of $A - 0I$ is found by row reducing the linear system $(A - 0I)x = 0$, i.e.,

$$\left(\begin{array}{ccc|c} -2 & 2 & 1 & 0 \\ -7 & 4 & 2 & 0 \\ 5 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

which implies $\ker(A - 0I) = \text{span}\{(0, 1, -2)\}$. Similarly, from the linear system $(A - 1I)x = 0$ we see

$$\left(\begin{array}{ccc|c} -3 & 2 & 1 & 0 \\ -7 & 3 & 2 & 0 \\ 5 & 0 & 0 & -1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \quad \Rightarrow \quad \begin{cases} x_1 + x_2 = 0 \\ 5x_2 + x_3 = 0. \end{cases}$$

Letting $x_2 = 1$, this gives $(x_1, x_2, x_3) = (1, -1, 5)$, i.e., $\ker(A - 1I) = \text{span}\{(1, -1, 5)\}$. Similarly,

$$\ker(A - 1I)^2 = \ker \begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix}^2 = \ker \begin{pmatrix} 0 & 0 & 0 \\ 10 & -5 & -3 \\ -20 & 10 & 6 \end{pmatrix} = \text{span}\{(1, 2, 0), (1, -1, 5)\}.$$

Thus, letting $P = [v_1\ v_2\ v_3]$ where v_i corresponds to the i-th eigenvector,

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = J = PAP^{-1} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 2 \\ -2 & 5 & 0 \end{pmatrix} \begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 2 \\ -2 & 5 & 0 \end{pmatrix}^{-1}.$$

$\square$

Cauchy-Schwarz inequality: $|\langle u, v\rangle|^2 \leq \langle u, u\rangle \cdot \langle v, v\rangle$, or $|\langle u, v\rangle| \leq \|u\|\|v\|$. $\triangle$

4 Example Solutions

4.1 Old Basic Exam Solutions

2001

F01.01: Let $K \subseteq \mathbb{R}$ be compact and $f : K \to \mathbb{R}$ be continuous. Prove there exists $x_0 \in K$ such that $f(x) \leq f(x_0)$ for all $x \in K$.

Proof:

Suppose f is not bounded above on K. Then, for each $n \in \mathbb{N}$, there exists $x_n \in K$ such that $f(x_n) > n$. This defines a sequence $\{x_n\}$. Because K is compact, it is closed and bounded. So, by the Bolzano-Weierstrss theorem, there exists a convergent subsequence $\{x_{n_k}\}$ of $\{x_n\}$, whose limit we shall denote by x^*. Since K is closed, $x^* \in K$. Because f is continuous, we know $\{f(x_{n_k})\}$ converges to $f(x^*)$. But, $f(x_{n_k}) > n_k \geq k$ for each $k \in \mathbb{N}$, which implies $\{f(x_{n_k})\}$ diverges to ∞, a contradiction. Hence f is bounded above on $[a, b]$.

By the Dedekind-completeness of $\mathbb{R}$, a least upper bound M of f exists. We seek to find $x_0 \in K$ such that $f(x_0) = M$. Let $n \in \mathbb{N}$. Since M is the least upper bound, there must exists $d_n \in [a, b]$ so that $M - 1/n < f(d_n)$. This defines a sequence $\{d_n\}$. Since M is an upper bound for f, it follows that $M - 1/n \leq f(d_n) \leq M$ for each $n \in \mathbb{N}$. Therefore, $\{f(d_n)\} \to M$.

The Bolzano-Weierstrass theorem tells us there exists a subsequence $\{d_{n_k}\}$, which converges to some x_0, and, since K is closed, $x_0 \in K$. Since f is continuous at x_0, the sequence $\{f(d_{n_k})\} \to f(x_0)$. But, $\{f(d_{n_k})\}$ is a subsequence of $\{f(d_n)\}$, which converges to M. Hence $M = f(x_0)$. Therefore, f attains its supremum value at x_0 so that $f(x) \leq f(x_0)\ \forall\ x \in K$. $\square$

F01.02: Let $a_n = (-1)^n/n$ for each $n \in \mathbb{N}$ and let $\alpha \in \mathbb{R}$. Prove there is a bijection $\sigma : \mathbb{N} \to \mathbb{N}$ such that

$$\sum_{n-1}^{\infty} a_{\sigma(n)} = \alpha.$$

Proof:

First, we note by the alternating series test that $\sum_{n-1}^{\infty} a_n$ converges (since $\{1/n\} \to 0$ as $n \to \infty$) while the harmonic series $\sum_{n-1}^{\infty} |a_n|$ diverges. That is, our series is conditionally convergent. For the remaineder of this proof, see F08.5. $\square$

F01.03: Let E be the a set of real numbers and $\{f_n\}$ be a sequence of continuous real-valued functions on E. Prove that if $f_n(x)$ converges to $f(x)$ uniformly on E, then $f(x)$ is continuous on E. (Recall that $f_n(x)$ converges to $f(x)$ uniformly on E means that for $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that whenever $n > N$ and $x \in E$, $|f_n(x) - f(x)| < \varepsilon$.)

Proof:

We first show that if each f_n is continuous at a point $x_0 \in E$, then so must be f. Let $\varepsilon > 0$ be given. To show that f is continuous at x_0, we must find a $\delta > 0$ such that whenever $x \in E$ and $|x - x_0| < \delta$, $|f(x) - f(x_0)| < \varepsilon$. Using the convergence of $\{f_n\}$, there exists $N \in \mathbb{N}$ such that whenever $x \in E$ and $n \geq N$, $|f_n(x) - f(x)| < \varepsilon/3$. Also, by the continuity of f_N, there is a $\delta > 0$ such that $|x - x_0| < \delta$ implies $|f_N(x) - f_N(x_0)| < \varepsilon/3$. Thus, whenever $|x_0 - x| < \delta$,

$$|f(x_0) - f(x)| \leq |f(x_0) - f_N(x_0)| + |f_N(x_0) - f_N(x)| + |f_N(x) - f(x)| < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.$$

Since x_0 was chosen arbitrarily, this holds for all $x \in E$. This implies that f is continuous. □

F01.04: Let S be the set of all sequences $\{x_n\}$ such that for all $n \in \mathbb{N}$, $x_n \in \{0, 1\}$. Prove there does not exists a one-to-one mapping form the set $\mathbb{N}$ onto the set S.

Proof:

By way of contradiction, suppose there exists a bijection $\sigma : N \to S$. Then, for each $n \in \mathbb{N}$, define the sequence $\{x_n\}$ by 1 if $\sigma(n)_n$ is 0 and 1 if $\sigma(n)_n$ is 0. Then $\{x_n\} \in S$. This implies there exists $M \in \mathbb{N}$ such that $\{x_n\} = \sigma(M)$ and so $x_M = \sigma(M)_M$. But, this contradicts the choice of $\{x_n\}$. Hence no such bijection σ can exist. □

F01.05: Suppose that $f : \mathbb{R}^2 \to \mathbb{R}$ is a continuous function such that partial derivatives $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ of f exists everywhere and are continuous everywhere, and $\dfrac{\partial^2 f}{\partial x \partial y}$ and $\dfrac{\partial^2 f}{\partial y \partial x}$ also exist and are continuous everywhere. Prove that

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Proof:

We shall prove the theorem for $z_0 = 0 \in \mathbb{R}^2$. The general case directly follows by replacing $f(z)$

with $f(z+z_0)$. For notational convenience let e_1, e_2 denote the standard basis vectors for $\mathbb{R}^2$ and set $x_1 := x$ and $x_2 := y$. Then define

$$a := \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}\right)(z_0) \quad \text{and} \quad a' := \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right)(z_0).$$

We seek to show $a' = a$. Let $\varepsilon > 0$ be given. Since our double derivatives of f of interest are continuous, there is a $\delta > 0$ such that

$$\left|\frac{\partial^2 f}{\partial x_1 \partial x_2}(z) - a\right| \leq \varepsilon \quad \text{and} \quad \left|\frac{\partial^2 f}{\partial x_2 \partial x_1}(z) - a'\right| \leq \varepsilon$$

whenever $|z| \leq 2\delta$. Now define the quantity

$$X := f(\delta e_1 + \delta e_2) - f(\delta e_1) - f(\delta e_2) + f(0).$$

From the fundamental theorem of calculus, we have

$$f(\delta e_1 + \delta e_2) - f(\delta e_2) = \int_0^\delta \frac{\partial f}{\partial x_1}(x_1 e_1 + \delta e_2)\ \mathrm{d}x_1 \quad \text{and} \quad f(\delta e_1) - f(0) = \int_0^\delta \frac{\partial f}{\partial x_1}(x_1 e_1)\ \mathrm{d}x_1$$

and hence

$$X = \int_0^\delta \left(\frac{\partial f}{\partial x_1}(x_1 e_1 + \delta e_2) - \frac{\partial f}{\partial x_1}(x_1 e_1)\right)\ \mathrm{d}x_1.$$

But by the mean value theorem, for each x_1 we have

$$\frac{\partial f}{\partial x_1}(x_1 e_1 + \delta e_2) - \frac{\partial f}{\partial x_1}(x_1 e_1) = \delta \frac{\partial f}{\partial x_2 \partial x_1}(x_1 e_1 + x_2 e_2)$$

for some $x_2 \in [0, \delta]$. By our construction of δ, we thus have

$$\left|\frac{\partial f}{\partial x_1}(x_1 e_1 + \delta e_2) - \frac{\partial f}{\partial x_1}(x_1 e_1) - \delta a\right| \leq \varepsilon\delta.$$

Integrating from 0 to δ, we thus obtain $|X - \delta^2 a| \leq \varepsilon\delta^2$. We can run the same argument with the subscripts 1 and 2 reversed to obtain $|X - \delta^2 a'| \leq \varepsilon\delta^2$. From the triangle inequality it follows that $|\delta^2 a - \delta^2 a'| \leq 2\varepsilon\delta^2$ and thus $|a - a'| \leq 2\varepsilon$. But this is true for all $\varepsilon > 0$, and a and a' do not depend on ε, and so we must have $a = a'$ as desired. $\square$

F01.07: If V is a real vector space and X is a subspace, let $V^* = \{f : V \to \mathbb{R} \mid f \text{ is linear}\}$ be the dual space of V and $X^0 = \{f \in V^* \mid f(x) = 0\ \forall\ x \in X\}$ be the annihilator of X. Let $T : V \to W$ be a linear transformation on finite dimensional real vector spaces. Recall that the transpose of T is the linear map

$T^t : W^* \to V^*$ defined by $T^t(f) = f \circ T$. Prove the following:

a) $\text{im}(T)^0 = \ker(T^t)$.

b) $\dim(\text{im}(T)) = \dim(\text{im}(T^t))$.

Proof:

a) By definition,

$$\begin{aligned}\text{im}(T)^0 &= \{f \in W^* \ \mid \ f(x) = 0 \ \forall \ x \in \text{im}(T)\}\\ &= \{f \in W^* \ \mid \ f(T(x)) = 0 \ \forall \ x \in V\}\\ &= \{f \in W^* \ \mid \ (f \circ T)(x) = 0 \ \forall \ x \in V\}\\ &= \{f \in W^* \ \mid \ T^t(f)(x) = 0 \ \forall \ x \in V\}\\ &= \{f \in W^* \ \mid \ T^t(f) = 0_{W^*}\}\\ &= \ker(T^t).\end{aligned}$$

b) Let $e_1, \ldots, e_m$ denote a basis of $\text{im}(T)$ so that $m = \dim(\text{im}(T))$. Then this basis can be extended to a basis $e_1, \ldots, e_n$ of W where $m \leq n$. The dual basis of $e_1, \ldots, e_n$ consists of the elements $\phi_1, \ldots, \phi_n$ in W^* defined by

$$\phi_j(e_k) = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases}$$

We claim $\{T^t(\phi_j)\}_{j=1}^m$ forms a basis for $\text{im}(T^t)$. Showing this will reveal the desired relation $\dim(\text{im}(T^t)) = m = \dim(\text{im}(T))$. Indeed, for each $\phi \in W^*$ there are unique $\beta_1, \ldots, \beta_n \in \mathbb{R}$ such that $\phi = \beta_1\phi_1 + \cdots + \beta_n\phi_n$. Similarly, since $T(v) \in \text{im}(T)$, for each $v \in V$ there are unique scalars $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ such that $T(v) = \alpha_1 e_1 + \cdots + \alpha_m e_m$. Thus, using linearity and the definition of each ϕ_j,

$$\begin{aligned} T^t(\phi)(v) = \phi(T(v)) &= \phi(\alpha_1 e_1 + \cdots \alpha_m e_m) \\ &= \sum_{j=1}^m \alpha_j \phi(e_j) \\ &= \sum_{j=1}^m \alpha_j (\beta_1\phi_1 + \cdots + \beta_n\phi_n)(e_j) \\ &= \sum_{j=1}^m \alpha_j \beta_j \phi_j(e_j) \\ &= \sum_{j=1}^m \beta_j \phi_j(T(v)) \\ &= (\beta_1\phi_1 + \cdots + \beta_m\phi_m)(T(v)) \\ &= (\beta_1 T^t(\phi_1) + \cdots + \beta_m T^t(\phi_m))(v) \end{aligned}$$

where we note $\alpha_j\phi_j(e_j) = \phi_j(\alpha_j e_j) = \phi_j(T(v))$. This implies each element $T^t(\phi) \in \text{im}(T^t)$ can be expressed as a linear combination of $T^t(\phi_j)$ for $j = 1, \ldots, m$. Moreover, the choice of β_j was unique and so $\{T^t(\phi_j)\}_{j=1}^m$ is linearly independent. Hence $\{T^t(\phi_j)\}_{j=1}^m$ form a basis for $\text{im}(T^t)$ and we are done.

□

F01.09: Let A be a real symmetric matrix. Prove that there exists an invertible matrix P such that $P^{-1}AP$ is diagonal. [You cannot just quote a theorem, but must prove it from scratch.]

Proof:

Let $T : V \to V$ denote the linear operator associated with the matrix A where $V = \mathbb{R}^n$. Then T is self-adjoint. We claim there exists an orthonormal basis of V consisting of eigenvectors of T, which implies T can be represented by a diagonal matrix relative to an orthonormal basis. Indeed, if $\{v_1, \ldots, v_n\}$ is an orthonormal list of eigenvectors of T, then define $P = [v_1 \ \cdots \ v_n]$ to be the matrix consisting of the column vector of A. Then the ij-th entry of P^TP is

$$(P^TP)_{ij} = \langle v_i, v_j \rangle = \delta_{ij}$$

where δ_{ij} denotes the Kronecker delta. Thus, $P^TP = I$ and so $P^T = P^{-1}$. Moreover, the ij-th entry of the product $P^{-1}AP$ is then

$$(P^TAP)_{ij} = ([v_1 \ \cdots \ v_n]^T[\lambda_1 v_1 \ \cdots \ \lambda_n v_n])_{ij} = \langle v_i, \lambda_j v_j \rangle = \lambda_j \langle v_i, v_j \rangle = \lambda_j \delta_{ij}$$

and so $P^{-1}AP$ is the diagonal matrix $\text{diag}(\lambda_1, \ldots, \lambda_n)$ where λ_i denotes the eigenvalue of v_i for $i = 1, \ldots, n$.

All that remains it to prove there exists an orthonormal basis of V consisting of eigenvectors of T. We proceed by induction on the dimension of V. If $\dim V = 1$, the claim holds trivially. Suppose the claim holds for some dimension $k > 1$, and let $\dim V = k + 1$. Let v denote a nonzero eigenvector of T, $W = \text{span}(v)$, and define $u_{k+1} := v_1/\|v_1\|$, a unit vector in W. Since v is an eigenvector of T, the subspace W of V is invariant under T. And, for $w \in W$ and $z \in W^\perp$ we have

$$\langle Tz, w\rangle = \langle z, Tw\rangle = 0$$

where the first equality holds since T is self-adjoint and the second because $Tw \in U$ and $z \in U^\perp$. Thus, it must follow that $Tz \in W^\perp$ so that $W^\perp$ is invariant under T. Moreover, the restriction $T|_U$ of T to $W^\perp$ is self-adjoint since for $y, z \in W^\perp$ we have

$$\langle (T|_U)y, z\rangle = \langle Ty, z\rangle = \langle y, Tz\rangle = \langle y, (T|_U)z\rangle .$$

Thus, the restriction of T to $W^\perp$ is a symmetric operator. However, $\dim W^\perp = k$ since $\dim W = 1$ and so, by the induction hypothesis, there is an orthonormal basis $\{u_1, \ldots, u_k\}$ of $W^\perp$ consisting of eigenvectors of the restriction of T to W and, thus, of T. And, $\langle u_{k+1}, u_j\rangle = 0$ for $j = 1, \ldots, k$ since $u_j \in W^\perp$. Hence $\{u_1, \ldots, u_{k+1}\}$ is an orthonormal set, consisting of eigenvectors of T, and we have closed the induction, thereby proving the claim. □

F01.10: Let V be a complex vector space and $T : V \to V$ be a linear transformation. Let $v_1, \ldots, v_n$ be non-zero vectors in V, each an eigenvector of a different eigenvalue. Prove that $\{v_1, \ldots, v_n\}$ is linearly independent.

Proof:

We shall proceed by induction on n. Of course, if $n = 1$, and $\alpha_1 v_1 = 0$ with $\alpha_1 \in \mathbb{C}$, then $\alpha_1 = 0$ since $v_1 \neq 0$. Now, suppose the hypothesis holds for some set of k eigenvalues $v_1, \ldots, v_k$. Then suppose there are scalars $\alpha_1, \ldots, \alpha_{k+1} \in \mathbb{C}$ such that a collection of $k + 1$ eigenvectors $v_1, \ldots, v_{k+1}$ satisfies

$$\alpha_1 v_1 + \cdots + \alpha_{k+1} v_{k+1} = 0. \tag{27}$$

Then, using the linearity of T,

$$0 = T(0) = T(\alpha_1 v_1 + \cdots + \alpha_{k+1} v_{k+1}) = \alpha_1 \lambda_1 v_1 + \cdots + \alpha_{k+1} \lambda_{k+1} v_{k+1}. \tag{28}$$

Now, subtracting λ_{k+1} multiplied by (27) from (28) gives

$$0 = \alpha_1(\lambda_{k+1} - \lambda_1)v_1 + \cdots + \alpha_{k+1}(\cancelto{0}{\lambda_{k+1} - \lambda_{k+1}})v_{k+1} = \alpha_1(\lambda_{k+1} - \lambda_1)v_1 + \cdots + \alpha_k(\lambda_{k+1} - \lambda_k)v_k.$$

Since the λ_i's are distinct, it follows that $(\lambda_{k+1} - \lambda_i)v_i$ is a nonzero scalar multiple of an eigenvector of T and, thus, is also an eigenvector of T for $i = 1, \ldots, k$. This implies we have a collection of k non-zero eigenvectors of T with distinct eigenvalues. By the inductive hypothesis, these are linearly independent and so $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$. Returning to (27), this implies

$$0 = 0v_1 + 0v_2 + \cdots + 0v_k + \alpha_{k+1}v_{k+1} = \alpha_{k+1}v_{k+1} \qquad \Rightarrow \qquad \alpha_{k+1} = 0$$

since $v_{k+1} \neq 0$. Thus, each α_i is identically zero for $i = 1, \ldots, k+1$ and we have closed the induction. Hence, such a collection $\{v_1, \ldots, v_n\}$ of non-zero eigenvectors with distinct eigenvalues is linearly independent for each $n \in \mathbb{N}$. $\square$

2002

W02.01: a) State some reasonably general conditions under which this "differentiation under the integral sign" formula is valid:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^b f(x,y)\ \mathrm{d}y = \int_a^b \frac{\partial f}{\partial x}\ \mathrm{d}y.$$

b) Prove that the formula is valid under the conditions you gave in a).

Proof:

a) Let $f : [a,b]\times[c,d] \to \mathbb{R}$. Suppose $\dfrac{\partial f}{\partial x}$ exists on $(a,b)\times[c,d]$ and extends to a continuous function on $[a,b]\times[c,d]$. Then the differentiation under the integral sign formula holds.

b) Define $u(x) := \int_a^b f(x,y)\ \mathrm{d}y$. Then for h with $x+h \in [a,b]$ we have

$$\begin{aligned}\left|\frac{u(x+h)-u(x)}{h} - \int_a^b \frac{\partial f}{\partial x}(x,y)\ \mathrm{d}y\right| &= \left|\int_a^b \left(\frac{f(x+h)-f(x)}{h} - \frac{\partial f}{\partial x}(x,y)\right)\ \mathrm{d}y\right| \\ &\leq \int_a^b \left|\left(\frac{f(x+h)-f(x)}{h} - \frac{\partial f}{\partial x}(x,y)\right)\right|\ \mathrm{d}y.\end{aligned}$$

Let $\varepsilon > 0$ be given. To prove the desired relation, it suffices to find $\delta > 0$ such that the left hand side is less than ε whenever $|h| < \delta$. By the Mean Value Theorem, there exists $c \in (0,1)$ such that

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

Now, since $\dfrac{\partial f}{\partial x}$ is continuous on the compact set $[a,b]\times[c,d]$, it is uniformly continuous. Then we can choose $\delta > 0$ such that $\|(x,y)-(x',y')\| < \delta$ implies

$$\left|\frac{\partial f}{\partial x}(x',y') - \frac{\partial f}{\partial x}(x,y)\right| < \frac{\varepsilon}{b-a}.$$

Using the above, for $|h| < \delta$, it follows that

$$\begin{aligned}\left|\frac{u(x+h)-u(x)}{h} - \int_a^b \frac{\partial f}{\partial x}(x,y)\ \mathrm{d}y\right| &\leq \int_a^b \left|\left(\frac{f(x+h)-f(x)}{h} - \frac{\partial f}{\partial x}(x,y)\right)\right|\ \mathrm{d}y. \\ &= \int_a^b \left|\frac{\partial f}{\partial x}(x+ch,y) - \frac{\partial f}{\partial x}(x,y)\right|\ \mathrm{d}y \\ &< \int_a^b \frac{\varepsilon}{b-a} \\ &= \varepsilon,\end{aligned}$$

and we are done.

$\square$

W02.02: Prove that the unit interval $[0,1]$ is sequentially compact, i.e., that every infinite sequence has a convergent subsequence. [Prove this directly. Do not just quote general theorems like Heini-Borel.]

Proof:

Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in $[0,1]$. We proceed by induction to find a convergent subsequence. Define $I_0 = [0,1]$. Then define $n_0 = 0$ and so, of course, $x_{n_0} \in I_0$. Now, if there are infinitely many terms in the left half of I_0, define $I_1 = [0, 1/2]$. Otherwise, define $I_1 = [1/2, 1]$. Then pick n_1 so that $x_{n_1} \in I_1$. This completes the base case. Now assume that for $k \geq 1$ we have successively found intervals in the fashion I_j and n_j so that $x_{n_j} \in I_j$ and $|I_j| = 2^{-j}$ for $j = 1, \ldots, k$. Now, if there are infinitely many terms in the left side of I_k, define I_{k+1} to be the left hand side of I_k. Otherwise, let I_{k+1} be the right side of I_k. Then pick n_{k+1} so that $x_{n_{k+1}} \in I_{k+1}$. This closed the induction. Let $\varepsilon > 0$ be given. By the Archimedean Property of $\mathbb{R}$, we can pick $N \in \mathbb{N}$ so that $2^{-N} \leq \varepsilon$ (n.b. we could write this more directly using logarithms). Then for all $i, j \geq N$ it follows that

$$|x_{n_i} - x_{n_j}| \leq |I_N| = 2^{-N} \leq \varepsilon,$$

which implies that $\{x_{n_k}\}_{k=1}^{\infty}$ is Cauchy. Because $[0,1]$ is complete, it follows that the Cauchy subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ converges. This completes the proof. $\square$

W02.03: Prove that the open unit ball in $\mathbb{R}^2$, $\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ is connected. [You may assume that intervals in $\mathbb{R}$ are connected. You should not just quote other general results, but give a direct proof.]

Proof:

We begin by first proving two lemmas.

Lemma 1: The image of a connected set under a continuous function is connected.
proof: Let S be a connected set and f be a continuous function. Then suppose that $f(S)$ is not connected, i.e., that there exists nonempty open disjoint subsets $A, B \subset f(S)$ such that $A \cup B = f(S)$. It follows that $f^{-1}(A)$ and $f^{-1}(B)$ are disjoint and non-empty. But, because f is continuous, it further follows that $f^{-1}(A)$ and $f^{-1}(B)$ are open. So, these are nonempty disjoint open subsets of S. In fact, since $f(S) = A \cup B$, it follows that $f^{-1}(A) \cup f^{-1}(B) = S$. But, this contradicts the fact that S is connected. Hence $f(S)$ must be connected.

Lemma 2: If a collection $\{S_\alpha\}_{\alpha \in I}$ consists of connected sets and $\cap_{\alpha \in I} S_\alpha \neq \emptyset$, then $\cup_{\alpha \in I} S_\alpha$ is connected. *proof:* Suppose $\cup_\alpha S_\alpha = A \cup B$ for nonempty disjoint open subsets A, B. Then for each

$\alpha \in I$, $S_\alpha = (S_\alpha \cap A) \cup (S_\alpha \cap B)$. Since $\cap_\alpha S_\alpha \neq \emptyset$, there is $x \in S_\alpha$. This implies either $S_\alpha \cap A$ or $S_\alpha \cap B$ is nonempty. Suppose, without loss of generality, that $x \in S_\alpha \cap A$. Because S_α is connected, it follows that $S_\alpha \cap B = \emptyset$. However, this implies $(\cup_\alpha S_\alpha) \cap B = \emptyset$, which contradicts our initial hypothesis. Hence $\cup_\alpha S_\alpha$ is connected.

For each $\theta \in [0, 2\pi)$ define $f_\theta : \mathbb{R} \to \mathbb{R}^2$ be $f_\theta(t) = (t\cos(\theta), t\sin(\theta))$. Clearly, each f_θ is continuous and $0 \in \cap_{\theta \in [0,2\pi)} f_\theta([0,1))$. Moreover, since $[0,1)$ is connected, by Lemma 1, $f_\theta([0,1))$ is connected. By Lemma 2, $\cup_{\theta\in[0,2\pi)} f_\theta([0,1))$ is connected. However, $\cup_{\theta\in[0,2\pi)} f_\theta([0,1))$ is the unit ball in $\mathbb{R}^2$ and so $\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ is connected. □

W02.04: Prove that the set of irrational numbers $\mathbb{I}$ in $\mathbb{R}$ is not a countable union of closed sets.

Proof:

By way of contradiction, suppose that $\mathbb{I}$ is a countable union of closed sets. Then there is a collection $\{I_k\}_{k=1}^\infty$ such that $\mathbb{I} = \cup_{k=1}^\infty I_k$. This implies that

$$\mathbb{R} = \underbrace{(\cup_{k=1}^\infty I_k)}_{\mathbb{I}} \cup \underbrace{(\cup_{q\in\mathbb{Q}}\{q\})}_{\mathbb{Q}}.$$

This implies $\mathbb{R}$ is the countable union of closed sets. Hence, by the Baire Category Theorem, one of these closed sets has non-empty interior. Since the interior of $\{q\}$ is empty for each $q \in \mathbb{Q}$, there exists $i \in \mathbb{N}$ such that I_i has nonempty interior. Thus, there is $r > 0$ and $x \in I_i$ such that the ball $\mathcal{B}(x,r) \subset I_i$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists $q \in \mathbb{Q}$ such that $q \in \mathcal{B}(x,r)$. This implies $q \in \mathbb{I}$, a contradiction. Hence $\mathbb{I}$ is not the countable union of closed sets. □

W02.05: a) Let $f : U \to \mathbb{R}^k$ be a function on an open set $U \subset \mathbb{R}^n$. Define what it means for f to be differentiable at a point $x \in U$.

b) State carefully the Chain Rule for the composition of differentiable functions of several variables.

c) Prove the Chain Rule you stated in b).

Proof:

a) Let $x \in U$. If there exists a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

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

then f is differentiable at x and we write $f'(x) = A$. If f is differentiable at every $x \in U$, then f is differentiable in U.

b) Suppose $U \subseteq \mathbb{R}^n$ is open and $f : U \to \mathbb{R}^k$ is differentiable at $x_0 \in U$, g maps the open set

$f(U)$ into $\mathbb{R}^m$, and g is differentiable at x_0. Then the composition $F : U \to \mathbb{R}^m$ defined by $F(x) = g(f(x))$ is differentiable at x_0 and $F'(x) = g'(f(x_0))f'(x_0)$.

c) Let $y_0 = f(x_0)$, $A = f'(x_0)$ and $B = g'(y_0)$, and define

$$u(h) = f(x_0 + h) - f(x_0) - A(h) \quad \text{and} \quad v(t) = g(y_0 + t) - g(y_0) - B(t),$$

for all $h \in \mathbb{R}^n$ and $t \in \mathbb{R}^k$ for which $f(x_0 + h)$ and $g(y_0 + t)$ are defined. Now define $\varepsilon(h) = \|u(h)\|/\|h\|$ and $\eta(t) = \|v(t)\|/\|t\|$ and note, by the differentiability of f and g, $\varepsilon(h) \to 0$ as $h \to 0$ and $\eta(t) \to 0$ as $t \to 0$. Given h, let $t = f(x_0 + h) - f(x_0)$. Then

$$\|t\| = \|A(h) + u(h)\| \leq (\|A\| + \|\varepsilon(h)\|)\|h\| \quad \Rightarrow \quad \frac{1}{\|h\|} \leq \frac{\|A\| + \|\varepsilon(h)\|}{\|t\|}$$

and

$$\begin{aligned} \|F(x_0 + h) - F(x_0) - B(A(h))\| &= \|g(y_0 + t) - g(y_0) - B(A(h))\| \\ &= \|B(k - A(h)) + v(t)\| \\ &= \|B(u(h)) + v(t)\| \\ &\leq \|B\|\|u(h)\| + \|v(t)\|. \end{aligned}$$

Since $t = [f(x_0 + h) - f(x_0)] \to 0$ as $h \to 0$, it follows that

$$\begin{aligned} \lim_{h\to 0} \|F(x_0 + h) - F(x_0) - B(A(h))\| &\leq \lim_{h\to 0} \|B\|\frac{\|u(h)\|}{\|h\|} + \frac{\|v(t)\|}{\|h\|} \\ &\leq \lim_{h\to 0} \|B\| \cdot \varepsilon(h) + (\|A\| + \|\varepsilon(h)\|) \cdot \eta(t) \\ &= \|B\| \cdot 0 + (\|A\| + 0) \cdot 0 \\ &= 0 \end{aligned}$$

and we have obtained the desired limit.

□

W02.06: a) State some reasonably general condition on a function $f : \mathbb{R}^2 \to \mathbb{R}$ under which

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

b) Prove the formula under the conditions you stated.

Proof:

a) The formula holds if $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ exists everywhere and are continuous everywhere, and

$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$ and $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$ also exists and are continuous everywhere.

b) See F01.5.

□

W02.09: Let V be a real vector space and $T : V \to V$ be a linear transformation. Let $\lambda_1, \dots, \lambda_m$ be distinct eigenvalues of T. Let $0 \neq v_i$ be an eigenvector of T with eigenvalue λ_i for $1 \leq i \leq m$. Show that $\{v_1, \dots, v_m\}$ is linearly independent.

Proof:

See F01.10. □

W02.10: Let V be a finite dimensional complex inner product space and $f : V \to \mathbb{C}$ a linear functional. Show that there exists a vector $w \in V$ such that $f(v) = \langle v, w\rangle$ for all $v \in V$.

Proof:

Let $e_1, \dots, e_m$ be an orthonormal basis for V. Then the basis of the dual $V^* = \{f : V \to \mathbb{C} \mid f \text{ is linear}\}$ is given by $\phi_1, \dots, \phi_m$ where $\phi_i(e_j) = \delta_{ij}$ with δ_{ij} denoting the Kronecker delta. For a given $f \in V^*$, it follows that there exists unique scalars $a_1, \dots, a_m \in \mathbb{C}$ such that $f = a_1\phi_1 + \cdots + a_m\phi_m$. Now let $v \in V$. Then there similarly exists unique scalars $b_1, \dots, b_m \in \mathbb{C}$ such that $v = b_1e_1 + \cdots b_me_m$. Then, using linearity and the definition of each ϕ_j,

$$f(v) = f\left(\sum_{i=1}^m b_ie_i\right) = \sum_{j=1}^m a_j\phi_j\left(\sum_{i=1}^m b_ie_i\right) = \sum_{j=1}^m\sum_{i=1}^m b_ia_j\phi_j(e_i) = \sum_{j=1}^m\sum_{i=1}^m b_ia_j\delta_{ij} = \sum_{j=1}^m b_ja_j.$$

Define $w := a_1e_1 + \cdots + a_me_m \in V$. Since the e_i are orthonormal, $\langle e_i, e_j\rangle = \delta_{ij}$ and so

$$\langle w, v\rangle = \langle\sum_{j=1}^m b_je_j, \sum_{i=1}^m a_ie_i\rangle = \sum_{j=1}^m\sum_{i=1}^m b_ja_i\langle e_j, e_i\rangle = \sum_{j=1}^m\sum_{i=1}^m b_ja_i\delta_{ij} = \sum_{j=1}^m b_ja_j = f(v),$$

completing the proof. □

W02.11: Let V be a finite dimensional complex inner product space and $T : V \to V$ a linear transformation. Prove that there exists an orthonormal ordered basis for V such that matrix representation A in this basis is upper triangular, i.e., $A_{ij} = 0$ if $i < j$.

Proof:

We proceed by induction. Of course, if $n = 1$, then picking any unit vector in V gives an orthonormal basis and we are done. For the inductive step, presume each k dimensional vector space has an upper triangular representation with respect to a linear operator on it. Then suppose $\dim(V) = k + 1$. Since V is complex, each operator T has an eigenvalue λ. This follows from the fact that $\mathbb{C}$ is algebraically closed. Let E_λ denote the eigenspace corresponding to λ. Then $E_\lambda^\perp = \text{range}(T - \lambda I)$. Because $(T - \lambda I)$ is not one-to-one, it cannot be surjective and so $\dim(E_\lambda^\perp) < \dim(V)$. Furthermore, $E_\lambda^\perp$ is invariant under T since for any $u \in E_\lambda^\perp$ we have $Tu = (T - \lambda I)u + \lambda u$ where $(T - \lambda I)u \in E_\lambda^\perp$, by definition, and $\lambda u \in E_\lambda^\perp$ since $E_\lambda^\perp$ is closed under scalar multiplication. Now the restriction of T to $E_\lambda^\perp$ provides an operator on this subspace and so, by the induction hypothesis, there is an orthonormal basis $e_1, \ldots, e_m$ of $E_\lambda^\perp$ with respect to the restriction of T to U, which has an upper triangular matrix, i.e., for each $j = 1, 2, \ldots, m$ we have $Te_j \in \text{span}\{e_1, \ldots, e_j\}$. Now extend $e_1, \ldots, e_m$ to an orthonormal basis $e_1, \ldots, e_m, v_1, \ldots, v_n$ of V where $k+1 = m+n$. For $k = 1, \ldots, n$ we have $Tv_k = (T - \lambda I)v_k + \lambda v_k$. By definition of $E_\lambda^\perp$, $(T - \lambda I)v_k \in E_\lambda^\perp$ and so this shows $Tv_k \in \text{span}\{e_1, \ldots, e_m, v_1, \ldots, v_k\}$. Then, in turn, implies T has an upper triangular matrix with respect to $e_1, \ldots, e_m, v_1, \ldots, v_k$ and we have closed the induction. □

S02.01: Prove that the closed interval $[0, 1]$ is connected.

Proof:

By way of contradiction, suppose there exist disjoint non-empty open sets A and B relative to $[0, 1]$ such that $A \cup B = [0, 1]$. Without loss of generality, suppose $1 \in B$. Clearly, $[0, 1]$ is bounded. Thus, A is bounded and by the least upper bound principle, we can define $c = \sup(A)$. Since $[0, 1]$ is closed, $c \in [0, 1]$.

First suppose $c \in A$. Then $c < 1$ since $1 \in B$ and $A \cap B = \emptyset$. And, since A is open, there exists $\varepsilon > 0$ so that $B(c, \varepsilon) \cap [0, 1] \subseteq A$. But, then $c + \min\{\varepsilon, 1 - \varepsilon\} \in A$, which contradicts the fact that $c = \sup(A)$.

Now suppose $c \in B$. Note then $c \neq 0$ since then we'd have $A = \{0\}$, which is closed. Thus, $c \in (0, 1]$. Since B is open, there exists $\varepsilon > 0$ such that $B(c, \varepsilon) \cap [0, 1] \subseteq B$. But then $c - \min\{\varepsilon, c\}$ is an upper bound for A, again contradicting that $c = \sup A > 0$. Hence $[0, 1]$ must be connected. □

S02.02: Show that the set $\mathbb{Q}$ of rationals in $\mathbb{R}$ is not expressible as the intersection of a countable collection of open subsets of $\mathbb{R}$.

Proof:

By way of contradiction, suppose that $\mathbb{Q} = \cap_{n=1}^{\infty} U_n$ where $U_n \subset \mathbb{R}$ is dense and open for each

$n \in \mathbb{N}$. Then $\mathbb{Q} \subset U_n$ for each n and, because $\mathbb{Q}$ is dense, so also must be U_n. Since $\mathbb{R}$ is open and dense in $\mathbb{R}$, $\mathbb{R}\backslash\{q\}$ is open and dense in $\mathbb{R}$ for each $q \in \mathbb{Q}$. This set is open since if $B_r(x)$ is a ball in $\mathbb{R}$ about $x \in \mathbb{R}$ of radius $r > 0$, then picking $r_x = \min\{r, \|x - q\|\}$ yields $B_{r_x}(x) \subset \mathbb{R}$. Thus,

$$\emptyset = \mathbb{I} \cap \mathbb{Q} = \left(\bigcap_{q \in \mathbb{Q}} \mathbb{R}\backslash\{q\} \right) \cap \bigcap_{n=1}^{\infty} U_n$$

is a countable intersection of dense open sets. But, the Baire Category Theorem implies that $\emptyset$ must be dense in $\mathbb{R}$, a contradiction.

$\square$

S02.03: Suppose that X is a compact metric space (in the covering sense of the word compact). Prove that every sequence $\{x_n \mid x_n \in X,\ n = 1, 2, 3, \ldots,\}$ has a convergent subsequence. [Prove this directly. Do not just quote a theorem.]

Proof:
Let X be a compact metric space and $\{x_n\}_{n=1}^{\infty} \subset X$. The collection $\cup_{x \in X} B(x, 1)$ of open balls of radius 1 in X forms an open cover of X. Since X is compact, it follows that there is a finite subcover $\cup_{j=1}^{J_1} B(x_j, 1)$ that covers X. Let ℓ_1 denote the index of a ball $B(x_{\ell_1}, 1)$ that contains infinitely many points from the sequence $\{x_n\}_{n=1}^{\infty}$. Existence of such a ball follows from the pigeonhole principle. Let S_1 be the set of all indices in this ball $B(x_{\ell_1}, 1)$. Now form an open covering $\cup_{x \in X} B(x, 1/2)$ of X. Then there is a finite subcover $\cup_{j=1}^{J_2} B(x_j, 1/2)$. Pick an index ℓ_2 for which $B(x_{\ell_1}, 1) \cap B(x_{\ell_2}, 1/2)$ contains infinitely many points from $\{x_n\}_{n=1}^{\infty}$. Then define S_2 to be the set of all indices in $B(x_{\ell_1}, 1) \cap B(x_{\ell_2}, 1/2)$ and note $S_2 \subseteq S_1$. Continuing in an inductive fashion, for each $k \in \mathbb{N}$ we can find an ℓ_k so that $\cap_{i=1}^{k} B(x_{\ell_i}, 1)$ has infinitely many points from $\{x_n\}_{n=1}^{\infty}$ and define S_k to be the set of indices in this intersection so that $S_k \subseteq S_{k-1} \subseteq \cdots \subseteq S_1$.

We define a convergent subsequence as follows. Let $n_1 \in S_1$. Then, for each $k \in \mathbb{N}$ greater than 1, pick $n_k \in S_k$ with $n_k > n_{k-1}$. Note such a choice is possible since, by construction, each S_k is infinite. Let $\varepsilon > 0$ be given. By the Archimedean property of $\mathbb{R}$, there exists $N \in \mathbb{N}$ such that $1/N \leq \varepsilon/2$. For $j, k \geq N$, observe that $x_{n_j}, x_{n_k} \in B(x_{\ell_N}, 1/N)$ and so

$$d(x_{n_j}, x_{n_k}) \leq d(x_{n_j}, x_{\ell_N}) + d(x_{n_k}, x_{\ell_N}) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

which implies the subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ is Cauchy. By the Heine-Borel theorem, X is compact if and only if it is complete and totally bounded. By the definition of completeness, each Cauchy sequence in X converges to a limit in X. Hence $\{x_{n_k}\}_{k=1}^{\infty}$ converges. $\square$

S02.06: Suppose $f : \mathbb{R}^3 \to \mathbb{R}$ is a continuously differentiable function with $\nabla f \neq 0$ at $0 \in \mathbb{R}^3$. Show that there are two other continuously differentiable functions $g : \mathbb{R}^3 \to \mathbb{R}$ and $h : \mathbb{R}^3 \to \mathbb{R}$ such that the function

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

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

Proof:

Since $\nabla f(0) \neq 0$, this can be extended to a basis $\nabla f(0), v_1, v_2$ of $\mathbb{R}^3$. Then define $g : \mathbb{R}^3 \to \mathbb{R}$ and $h : \mathbb{R}^3 \to \mathbb{R}$, respectively, by

$$g(r) = \langle v_1, r\rangle \quad \text{and} \quad h(r) = \langle v_2, r\rangle \quad \forall\ r \in \mathbb{R}^3$$

where $\langle \cdot, \cdot\rangle$ denotes the standard scalar product in $\mathbb{R}^3$. Then g and h are continuously differentiable. Indeed, $\nabla g(r) = v_1$ and $\nabla h(r) = v_2$ for all $r \in \mathbb{R}^3$. Define $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ by $\phi(x, y, z) = (f(x, y, z), g(x, y, z), h(x, y, z))$. Then

$$\phi'(0) = \begin{bmatrix} \nabla f(0) \\ \nabla g(0) \\ \nabla h(0) \end{bmatrix} = \begin{bmatrix} \nabla f(0) \\ v_1 \\ v_2 \end{bmatrix}.$$

Since the rows of $\phi'(0)$ are linearly independent, $\phi'(0)$ is invertible. Moreover, ϕ is continuously differentiable since it is the composition of f, g, and h, which are each continuously differentiable. It follows from the inverse function theorem that there exists open subsets $U, V \subseteq \mathbb{R}^3$ such that $0 \in U$ and $\phi(0) \in V$ where ϕ is one-to-one on U and $\phi(U) = V$. That is, there is a neighborhood U of 0 where ϕ is one-to-one. This completes the proof. $\square$

S02.8: Let V be a finite dimensional real vector space. Let $W \subset V$ be a subspace and $W^0 := \{f : V \to F \text{ linear} \mid f = 0 \text{ on } W\}$. Prove that $\dim(V) = \dim(W) + \dim(W^0)$.

Proof:

Let $e_1, \ldots, e_m$ be an orthonormal basis for W. Then extend this to a basis $e_1, \ldots, e_n$ of V with $n \geq m$ so that $\dim(W) = m$ and $\dim(V) = n$. Then all we must show is that $\dim(W^0) = n - m$. The dual basis of $e_1, \ldots, e_n$ for $V^* = \{f : V \to \mathbb{R} \mid f \text{ is linear}\}$ is given by $\phi_1, \ldots, \phi_n$ where $\phi_i(e_j) = \delta_{ij}$ where δ_{ij} denotes the Kronecker delta. Let $f \in W^* \subset V^*$. Then there are unique scalars $a_1, \ldots, a_n \in \mathbb{R}$ such that $f = a_1\phi_1 + \cdots + a_n\phi_n$. Similarly, let $w \in W$. Then there are unique scalars $b_1, \ldots, b_m \in \mathbb{R}$ such that $w = b_1e_1 + \cdots b_me_m$. Using the linearity of f, it follows

that

$$f(w) = \sum_{j=1}^{n} a_j\phi_j(v) = \sum_{j=1}^{n} a_j\phi_j(b_1e_1 + \cdots b_m e_m) = \sum_{j=1}^{m} a_j\phi_j(b_je_j) = \sum_{j=1}^{m} a_jb_j.$$

where we note

$$\phi_j(b_1e_1 + \cdots b_me_m) = b_1\phi_j(e_1) + \cdots b_m\phi_j(e_m) = 0 + \cdots + 0 + b_j\phi_j(e_j) + 0 + \cdots + 0 = b_j.$$

But, $f = 0$ on W and so $\sum_{j=1}^{m} a_jb_j = 0$. Since this holds for each $w \in W$, we must have $a_j = 0$ for $j = 1, \ldots, m$. Hence each $f \in W^*$ is of the form $f = a_{m+1}\phi_{m+1} + \cdots + a_n\phi_n$, which implies $W^* = \text{span}\{\phi_{m+1}, \ldots, \phi_n\}$ and so $\dim(W^*) = n - m$, as desired. □

S02.10: Let V be a complex inner product space and W a finite dimensional subspace. Let $v \in V$. Prove that there exists a unique vector $v_W \in W$ such that

$$\|v - v_W\| \leq \|v - w\| \tag{29}$$

for all $w \in W$. Deduce that equality holds iff $w = v_W$.

Proof:

Let $e_1, \ldots, e_m$ be an orthonormal basis for W. Extend this to an orthonormal basis $e_1, \ldots, e_m, f_1, \ldots, f_n$ for V. Let $v \in V$ and define v_W to be the projection of v into W, i.e., define

$$v_W := \langle v, e_1\rangle\, e_1 + \cdots + \langle v, e_m\rangle\, e_m$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product. Then $v_W \in W$. Also, $(v - v_W) \in W^\perp$ since $\langle (v - w_W), e_j\rangle = \langle v, e_j\rangle - \langle v, e_j\rangle = 0$ for $j = 1, \ldots, m$. For each $w \in W$ it follows that

$$\|v - v_W\|^2 \leq \|v - v_W\|^2 + \|v_W - w\|^2 = \|(v - v_W) + (v_W - w)\|^2 = \|v - w\|^2, \tag{30}$$

where the first inequality holds because $0 \leq \|v_W - w\|^2$, the following equality holds from the Pythagorean Theorem (which applies because $v - v_W \in W^\perp$ and $v_W - w \in U$), and the last equality holds by simple algebra. Taking square roots gives the desired inequality. Lastly, our inequality above is an equality precisely when $\|v_W - w\| = 0$, which occurs iff $v_W = w$. □

2003

S03.2: Prove that if $a_1, a_2, \ldots$ is a sequence of real number with

$$\sum_{j=1}^{+\infty} |a_j| < \infty, \tag{31}$$

then $\lim\limits_{N\to\infty} \sum_{j=1}^{N} a_j$ exists.

Proof:
The Cauchy Convergence Criterion for series states that the series $\sum_{j=1}^{\infty} a_j$ converges iff for each $\varepsilon > 0$, there is an index N such that

$$|a_{n+1} + \cdots + a_{n+k}| < \varepsilon \quad \text{for all } n \geq N \text{ and } k \in \mathbb{N}. \tag{32}$$

From the triangle inequality, we know

$$\left| \sum_{j=n+1}^{n+k} a_j \right| \leq \sum_{j=n+1}^{n+k} |a_j|. \tag{33}$$

And because $\sum_{j=1}^{+\infty} |a_j|$ converges, it follows from the Cauchy Convergence Criterion that the sequence of partial sums for this series is a Cauchy sequence. So, our above inequality implies the sequence of partial sums for the series $\sum_{j=1}^{\infty} a_j$ is also a Cauchy sequence. Hence, again using the Cauchy Convergence Criterion, $\sum_{j=1}^{\infty} a_j$ converges. □

2004

S04.2: Is $f(x) = \sqrt{x}$ uniformly continuous on $[0, \infty)$? Prove your assertion.

Proof:
We claim $f(x) = \sqrt{x}$ is uniformly continuous on $[0, \infty]$. To verify this, we must show that if $\varepsilon > 0$ is given, then there exists $\delta > 0$ such that $|f(x) - f(y)| \leq \varepsilon$ whenever $x, y \in [0, \infty)$ are such that $|x - y| \leq \delta$. So, let $\varepsilon > 0$ be given. Then define $\delta := \varepsilon^2$ and note that $|\sqrt{x} - \sqrt{y}| \leq |\sqrt{x} + \sqrt{y}|$ since $\sqrt{y}$ is positive. Then

$$|\sqrt{x} - \sqrt{y}|^2 \leq |\sqrt{x} - \sqrt{y}||\sqrt{x} + \sqrt{y}| = |x - y| \leq \delta = \varepsilon^2. \tag{34}$$

Taking the square root, we obtain $|\sqrt{x} - \sqrt{y}| \leq \varepsilon$. Hence we identified $\delta > 0$ such that the desired relation holds, and so $f(x) = \sqrt{x}$ is uniformly continuous. □

F04.4: Suppose (M, ρ) is a metric space, $x, y \in M$, and that $\{x_n\}$ is a sequence in this metric space such that $x_n \to x$. Prove that $\rho(x_n, y) \to \rho(x, y)$.

Proof:

Let $\varepsilon > 0$ be given. We must find $N \in \mathbb{N}$ such that for all $n \geq N$ we have

$$|\rho(x_n, y) - \rho(x, y)| \leq \varepsilon. \tag{35}$$

Let $n \in \mathbb{N}$. Suppose $\rho(x, y) \geq \rho(x_n, y)$. Then, by the triangle inequality,

$$0 \leq \rho(x, y) - \rho(x_n, y) \leq (\rho(x, x_n) + \rho(x_n, y)) - \rho(x_n, y) = \rho(x, x_n). \tag{36}$$

Alternatively, if $\rho(x, y) < \rho(x_n, y)$, then

$$0 < \rho(x_n, y) - \rho(x, y) \leq (\rho(x_n, x) + \rho(x, y)) - \rho(x, y) = \rho(x_n, x). \tag{37}$$

In either case, we have

$$|\rho(x_n, y) - \rho(x, y)| \leq |\rho(x_n, x)|. \tag{38}$$

Since $x_n \to x$, we can pick $N \in \mathbb{N}$ such that $\rho(x_n, x) \leq \varepsilon$ whenever $n \geq N$. Using (38), it follows that $|\rho(x_n, y) - \rho(x, y)| \leq |\rho(x_n, x)| \leq \varepsilon$ whenever $n \geq N$. Hence we have found N such that (35) holds and we are done. $\square$

F04.6: The Bolzano-Weirstrass Theorem in $\mathbb{R}^n$ states that if S is a bounded closed subset of $\mathbb{R}^n$ and $\{x_k\}$ is a sequence which takes values in S, then $\{x_k\}$ has a subsequence which converges to a point in S. Assume this statement is known in the case where $n = 1$, and use it to prove the statement in case $n = 2$.

Proof:

Let $S \subseteq \mathbb{R}^2$ be closed and bounded. We must show that each sequence $\{p_n\}$ in S has a subsequence that converges to a point $p \in S$. Note for each p_n, we can write $(x_n, y_n) = p_n$ for $x_n, y_n \in \mathbb{R}$. And, through the boundedness of S (and therefore each p_n), each x_n and y_n must also be bounded. Assuming the theorem holds for $n = 1$, we know there must exists a subsequence $\{x_{f(n)}\}$ that converges where $f : \mathbb{N} \to \mathbb{N}$ with $f(n) \geq n$ for each $n \in \mathbb{N}$. Also, since $\{y_{f(n)}\}$ is bounded, we can find a convergent subsequence $\{y_{g(f(n))}\}$ of $\{y_{f(n)}\}$ where $g : \mathbb{N} \to \mathbb{N}$ with $g(n) \geq n$ for each $n \in \mathbb{N}$. Of course, since $\{x_{g(f(n))}\}$ is a subsequence of the convergent sequence $\{x_{f(n)}\}$, it also converges. It then follows that each coordinate of the sequence $\{p_{g(f(n))}\}$ converges, which implies that this subsequence of $\{p_n\}$ is itself convergent. Moreover, since S is closed, the limit of $\{p_{g(f(n))}\}$ is in S.

This completes the proof. $\square$

S04.4: Are there infinite compact subsets of $\mathbb{Q}$? Prove your assertion.

Proof:

We claim there are infinite compact subsets of $\mathbb{Q}$. Let $k \in \mathbb{Z}$ and define $S_k = \{k\} \cup \{k + 1/n \mid n = 1, 2, \ldots\}$. Since $k \in \mathbb{Q}$ and $\mathbb{Q}$ is a field, we know $k + 1/n \in \mathbb{Q}$ and so $S_k \subset \mathbb{Q}$. Moreover, S_k is countably infinite by definition. All that remains is to show S_k is compact. Let $U = \{U_i\}$ be an open cover of S_k. Then there exists some U_0 such that $k \in U_0$. Since U_0 is open, there exists $r > 0$ such that $B(k, r) \subseteq U_0$. Furthermore, by the Archimedean property of $\mathbb{R}$, there exists a positive integer N such that $1/N \leq r$. Then for any $n \geq N$ we have $1/n \leq 1/N \leq r$, which implies $k + 1/n \in B(k, r) \subseteq U_0$. For each $i = 1, \ldots, N-1$, there exists some open set $U_i \in U$ containing $k + 1/i$. Then the collection $\{U_0, \ldots, U_{N-1}\}$ is a finite subcover of U that covers S_k. Hence S_k is an infinite compact subset of $\mathbb{Q}$. $\square$

F04.08: Let $A = (a_{ij})$ be a real $n \times n$ symmetric matrix and let $Q(v) = v \cdot Av$ (ordinary dot product) be the associated quadratic form defined for $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$.

a) Show that $\nabla Q(v) = 2Av$ where $\nabla Q(v)$ is the gradient of the function at Q.

b) Let M be the minimum value of $Q(v)$ on the unit sphere $S^n := \{v \in \mathbb{R}^n \ : \ \|v\| = 1\}$ and let $u \in S^n$ be a vector such that $Q(u) = M$. Prove, using Lagrange multipliers, that u is an eigenvector of A with eigenvalue M.

Proof:

a) Observe that

$$Q(v) = v \cdot Av = \sum_{j=1}^{n} v_j (Av)_j = \sum_{j=1}^{n} v_j \left(\sum_{i=1}^{n} A_{ji} v_i \right) = \sum_{i,j=1}^{n} A_{ij} v_i v_j. \tag{39}$$

This implies

$$\begin{aligned} \partial_{x_k} Q(v) &= \partial_{x_k} \sum_{i,j=1}^{n} A_{ij} v_i v_j \\ &= \sum_{i,j=1}^{n} A_{ij} \partial_{x_k} (v_i v_j) \\ &= \sum_{i,j=1}^{n} A_{ij} (\delta_{ik} v_j + v_i \delta_{jk}) \\ &= \sum_{j=1}^{n} A_{kj} v_j + \sum_{i=1}^{n} A_{ik} v_i \\ &= \sum_{j=1}^{n} A_{kj} v_j + \sum_{i=1}^{n} A_{ki} v_i \\ &= 2(Av)_k \end{aligned} \tag{40}$$

where δ_{ij} denotes the Kronecker delta. Thus,

$$\nabla Q(v) = (\partial_{x_1} Q(v), \ldots, \partial_{x_n} Q(v)) = (2(Av)_1, \ldots, 2(Av)_n) = 2Av. \tag{41}$$

b) Define $g(v) = \langle v, v \rangle - 1$ and observe $g(v) = 0$ if and only if $v \in S^n$. Lagrange's theorem for multipliers then states that if Q and g have continuous first derivatives and Q has an extremem point u along the curve $g = 0$, then there is a $\lambda \in \mathbb{R}$ such that $\nabla Q(u) = \lambda \nabla g(u)$. Note, taking I as the identity, the above computations in a) show $\nabla g(v) = 2v$. Thus,

$$\nabla Q(u) = \lambda \nabla g(u) \quad \Leftrightarrow \quad 2Au = 2\lambda u, \tag{42}$$

which implies u is an eigenvalue of A with eigenvalue λ. Then

$$M = Q(u) = u \cdot Au = u \cdot \lambda u = \lambda(u \cdot u) = \lambda. \tag{43}$$

Thus, we conclude u is an eigenvector of A with eigenvalue $M = \lambda$, as desired.

$\square$

F04.09: Let $T : \mathbb{C}^n \to \mathbb{C}^n$ be a linear transformation and P a polynomial such that $P(T) = 0$. Prove that every eigenvalue of T is a root of P.

Proof:
Write $P(x) = a_0 + a_1 x + \cdots + a_m x^m$ for some $a_0, \ldots, a_m \in \mathbb{C}$. Now let $\lambda \in \mathbb{C}$ be an eigenvalue of T. Then there is $v \in \mathbb{C}^n$ with $v \neq 0$ such that $Tv = \lambda v$. So, suppose $T^p v = \lambda^p v$ for some $p \in Z^+$. Then $T^{p+1}v = T^p(Tv) = T^p(\lambda v) = \lambda T^p v = \lambda^{p+1} v$. By induction, this implies $T^k v = \lambda^k v$ for each $k \in \mathbb{Z}^+$. Thus,

$$0 = p(T)v = (a_0 I + a_1 T + \cdots + a_m T^m)v = (a_0 + a_1\lambda + \cdots + a_m \lambda^m)v = P(\lambda)v. \tag{44}$$

where we note $T^0 = I$. Since $v \neq 0$, it follows that $P(\lambda) = 0$. Therefore, we conclude each eigenvalue of T is a root of P. $\square$

F04.10: Let $V = \mathbb{R}^n$ and $T : V \to V$ be a linear transformation. For $\lambda \in \mathbb{C}$, the subspace

$$V(\lambda) = \left\{ v \in V \ : \ (T - \lambda I)^j v = 0 \quad \text{for some } j \geq 1 \right\} \tag{45}$$

is called a generalized eigenspace.

a) Prove there exists a finite number M such that $V(\lambda) = \ker\left((T - \lambda I)^M\right)$.

b) Prove that if $\lambda \neq \mu$, then $V(\lambda) \cap V(\mu) = \{0\}$. Hint: use the following equation by raising both sides to a high power.

$$\frac{T - \lambda I}{\mu - \lambda} + \frac{T - \mu I}{\lambda - u} = I. \tag{46}$$

Solution:

a) We first provide two lemmas:

Lemma 1: $\{0\} = \ker(T^0) \subset \ker(T^1) \subset \ker(T^2) \subset \cdots$.

Proof:

We proceed by induction. Suppose $v \in \ker(T^k)$ for some $k \geq 0$. Then $T^{k+1}(v) = T(T^k v) = T(0) = 0$ where the final equality holds by linearity of T. The result follows by induction. $\square$

Lemma 2: If there is a nonnegative integer m such that $\ker(T^m) = \ker(T^{m+1})$, then $\ker(T^m) = \ker(T^{m+k})$ for all nonnegative integers k.

Proof:

The base case is given. Suppose the statement holds for $k > 1$ and let $v \in \ker(T^{m+k+1})$. This implies

$$0 = T^{m+k+1}(v) = T^{m+1}(T^k v) \quad \Rightarrow \quad T^k v \in \ker(T^{m+1}). \tag{47}$$

But, $\ker(T^{m+1}) = \ker(T^m)$. Hence

$$T^{m+k}(v) = T^m(T^k v) = 0 \quad \Rightarrow \quad v \in \ker(T^{m+k}). \tag{48}$$

This shows $\ker(T^{m+k+1}) \subset \ker(T^{m+k})$. Lemma 1 gives $\ker(T^{m+k}) \subset \ker(T^{m+k+1})$, from which we obtain $\ker(T^{m+k+1}) = \ker(T^{m+k}) = \ker(T^m)$. This closes the induction. $\square$

We now prove a). By way of contradiction, assume the hypothesis is false. Then

$$\ker((T-\lambda I)^0) \subsetneq \ker((T-\lambda I)^1) \subsetneq \cdots \subsetneq \ker((T-\lambda I)^{n+1}) \tag{49}$$

where the inclusions follow from Lemma 1 and they are strict by our assumption and Lemma 2. This implies $\dim(T^{k+1}) \geq \dim(T^k)+1$ for each k, and so $\dim(T^{n+1}) \geq n+1 > n = \dim(V)$. This contradicts the fact that T^{n+1} is a subspace of V. Thus, the claim does hold. This further shows $V(\lambda) = \{v \in V \ : \ (T-\lambda I)^n v = 0\} = \ker((T-\lambda I)^n)$.

b) From a), we have $V(\lambda) = \ker((T-\lambda I)^n)$ and $V(\mu) = \ker((T-\mu I)^n)$. Now let $v \in V(\lambda) \cap V(\mu)$. It suffices to show v is necessarily the zero vector, which we does as follows. Using the given relation, observe

$$v = I^{2n} v = \left(\frac{T-\lambda I}{\mu - \lambda} + \frac{T - \mu I}{\lambda - \mu} \right)^{2n} v \tag{50}$$

But, on the right hand side, each of the terms $(T-\lambda I)$ and $(T-\mu I)$ in the expansion are raised to at least a power of n and $(T-\lambda I)^n v = (T-\mu I)^n v = 0$. Thence

$$v = \left(\frac{T-\lambda I}{\mu - \lambda} + \frac{T - \mu I}{\lambda - \mu} \right)^{2n} v = 0. \tag{51}$$

□

2005

S05.6: Consider the set of $f : [0,1] \to \mathbb{R}$ that obey

$$|f(x) - f(y)| \leq |x - y| \quad \text{and} \quad \int_0^1 f(x)\ \mathrm{d}x = 1. \tag{52}$$

Show that this is a compact subset of $C([0,1])$.

Solution:

The set of f, denoted F, is uniformly bounded. By way of contradiction, suppose otherwise. Then there is $f \in F$ and a $x \in [0,1]$ such that $|f(x)| \geq 3$. This implies for $y \in [0,1]$ that

$$|f(y) - f(x)| \leq |y - x| \leq 1 \quad \Rightarrow \quad |f(y)| \geq |f(x)| - 1 \geq 2 \tag{53}$$

where the right hand side follows by applying the reverse triangle inequality. Assume $f(y) \geq 2$. Then

$$\int_0^1 f(x)\ \mathrm{d}x \geq \int_0^1 2\ \mathrm{d}x = 2 > 1, \tag{54}$$

a contradiction. Similarly, if $f(y) \leq -2$, we get a contradiction. Hence $\|f\|_\infty \leq 3$ for $f \in F$.

To show F is compact, we show it is sequentially compact. Let $\{f_n\} \subset F$ be a sequence. We must show there is a subsequence $\{f_{n(m)}\}_{m=1}^\infty$ of $\{f_n\}_{n=1}^\infty$ that converges to a limit $g \in F$. We first construct the subsequence $\{f_{n(m)}\}_{m=1}^\infty$ by a diagonalization argument. Let $\tau : \mathbb{Z}^+ \to [0,1] \cap \mathbb{Q}$ be an enumeration of the rational numbers in $[0,1]$. Since $\|f_n\|_\infty \leq 3$, $f_n(\tau(1)) \in [-3,3]$ for each $n \in \mathbb{Z}^+$. Because $[-3,3]$ is compact, this sequence has a convergence subsequence, say, $\{f_{n_{1,k}}(\tau(1))\}$. Proceeding inductively, suppose $\{f_{n_{m,k}}(\tau(m))\}_{k=1}^\infty$ converges for $m \in \mathbb{Z}^+$. By the compactness of $[-3,3]$, there is a subsequence $\{f_{n_{m+1,k}}\}_{k=1}^\infty$ of $\{f_{n_{m,k}}\}_{k=1}^\infty$ such that $\{f_{n_{m+1,k}}(\tau(m+1))\}_{k=1}^\infty$ converges. Now define the sequence $n(m) := n_{m,m}$. For each $m \in \mathbb{Z}^+$ there are, at most, $m-1$ terms in the sequence $\{f_{n(m)}(\tau(m))\}_{m=1}^\infty$ not contained in the sequence $\{f_{n_{m,k}}(\tau(m))\}_{k=1}^\infty$, namely, the first $m-1$ terms. Hence

$$\lim_{m\to\infty} f_{n(m)}(\tau(m)) = \lim_{n\to\infty} f_{n_{m,k}}(\tau(m)), \tag{55}$$

and so the limit exists. This implies the sequence $\{f_{n(m)}\}$ converges at each rational $x \in [0,1]$.

Now cover $[0,1]$ by the open balls $B(x, \varepsilon/3)$ with rational $x \in [0,1]$. Since $[0,1]$ is compact and the rationals are dense in $[0,1]$, there is a finite subcover, i.e., there are $x_1, \ldots, x_m \in [0,1] \cap \mathbb{Q}$ such

that

$$[0,1] \subset \bigcup_{i=1}^{m} B(x_i, \varepsilon/3). \tag{56}$$

Let $z \in [0,1]$. Then there is $i \in \{1, \ldots, m\}$ such that $|z - x_i| < \varepsilon/3$. Hence

$$|f(z) - f(x_i)| \leq |z - x_i| < \frac{\varepsilon}{3}. \tag{57}$$

Also, by the convergence of $\{f_{n(m)}(x_i)\}_{m=1}^{\infty}$, there is an N_i such that

$$|f_{n(k)}(x_i) - f_{n(m)}(x_i)| < \frac{\varepsilon}{3} \quad \forall\ k, m \geq N_i. \tag{58}$$

Set $N := \max\{N_1, \ldots, N_m\}$. Then, using the triangle inequality,

$$\begin{aligned} \forall\ k, m \geq N, \quad |f_{n_k}(z) - f_{n_m}(z)| &\leq |f_{n_k}(z) - f_{n_k}(x_i)| + |f_{n_k}(x_i) - f_{n_m}(x_i)| + |f_{n_m}(x_i) - f_{n_m}(z)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon. \end{aligned} \tag{59}$$

Since $f_{n(m)}(z) \in \mathbb{R}$ and $\mathbb{R}$ is complete, the sequence converges. Moreover, it is uniform since this holds for arbitrary $z \in [0,1]$. Define $g : [0,1] \to \mathbb{R}$ as the pointwise limit, i.e., $g(z) := \lim\limits_{m\to\infty} f_{n(m)}(z)$ for $z \in [0,1]$. Since the convergence is uniform, we may write

$$\int_0^1 g(x)\ \mathrm{d}x = \int_0^1 \lim_{m\to\infty} f_{n(m)}(x)\ \mathrm{d}x = \lim_{m\to\infty} \int_0^1 f_{n(m)}(x)\ \mathrm{d}x = \lim_{m\to\infty} 1 = 1. \tag{60}$$

And, for $m \geq N$,

$$\begin{aligned} |g(z) - g(y)| &\leq |g(z) - f_{n(m)}(z)| + |f_{n(m)}(z) - f_{n(m)}(y)| + |f_{n(m)}(y) - g(y)| \\ &\leq \frac{\varepsilon}{3} + |z - y| + \frac{\varepsilon}{3} \\ &\leq |z - y| + \varepsilon. \end{aligned} \tag{61}$$

Hence $|g(z) - g(y)| \leq |z - y| + \varepsilon$. Since none of the terms in this final inequality depend upon ε, we let $\varepsilon \longrightarrow 0$ to obtain $|g(z) - g(y)| \leq |z - y|$, thereby concluding $g \in F$. Thus, F is sequentially compact. This completes the proof. $\square$

S05.12: Let (X, d) be a metric space. Prove that the following are equivalent:

a) There is a countable dense set.

b) There is a countable basis for the topology.

Proof:
First suppose there exists a countable dense subset $E \subseteq X$. Let $\{q_1, q_2 \ldots\}$ be an enumeration of E. We claim that the collection of open balls $U = \{B(q_i, 1/j) \mid (i, j) \in \mathbb{N} \times \mathbb{N}\}$ is countable and forms a basis for the topology. Indeed, U is equinumerous with $\mathbb{N} \times \mathbb{N}$, which is countable since each $(i, j) \in \mathbb{N} \times \mathbb{N}$ can be mapped to $(i + j) \in \mathbb{N}$ while each $k \in \mathbb{N}$ can be mapped to $(k, 1) \in \mathbb{N} \times \mathbb{N}$. To show U forms a basis, we must show that every open set in X can be written as a union of elements of U. So let $V \subseteq X$ be open and $v \in V$. By the definition of open, there exists a neighborhood N of v in V, i.e., $N \subseteq V$. Let r denote the radius of the neighborhood N. Now, by the Archimedean property of $\mathbb{R}$, there exists $j \in \mathbb{N}$ such that $1/j \leq r/2$. And, by the density of E in X, there exists $q_i \in E$ such that v is contained in the open ball about q_i of radius $1/j$, i.e., $v \in B(q_i, 1/j)$. Let Π denote the collection of indices (i, j) of the balls corresponding to each such element of V. Then

$$V \subset \bigcup_{(i,j) \in \Pi} B(q_i, 1/j).$$

And, by the choice of $B(q_i, 1/j)$, for each $v \in V$, we have $B(q_i, 1/j) \subseteq V$ since for any $x \in B(q_i, 1/j)$ we have

$$d(v, x) \leq d(v, q_i) + d(q_i, x) \leq 1/j + 1/j \leq r/2 + r/2 = r. \tag{62}$$

This implies that

$$\bigcup_{(i,j) \in \Pi} B(q_i, 1/j) \subseteq V.$$

Hence U forms a basis for the topology.

We now prove the converse. Suppose $U = \{E_1, E_2, \ldots\}$ is a countable basis for the topology. Now let $x \in X$ and define a neighborhood N of x of some radius $r > 0$. Since N is open, there must exist $E_i \subseteq N$. Since $x \in E_i$, E_i is nonempty, and so we can let q_i denote some element in E_i. Then $q_i \in B(x, r)$. So, for each x and $r > 0$, we can identify a q_i for which $q_i \in B(x, r)$. That is, each $x \in X$ is a limit point of the set $\{q_i\}_{i=1}^{\infty}$. Hence there exists a countable dense set in X. $\square$

F05.1: A real number α is algebraic if there exists a finite set $a_0, \ldots, a_n \in \mathbb{Q}$ not all zero such that α is a root of $a_0 + a_1 x + \cdots + a_n x^n$. Prove that the set of algebraic numbers is countable.

Proof:

The set of polynomials of degree n with rational coefficients is equinumerous with $\mathbb{Q}^{n+1}$. So, the set of roots of such polynomials injects into $\{1, 2, \ldots, n\} \times \mathbb{Q}^{n+1}$ since there are n roots for each degree n polynomial. So, the set of algebraic numbers injects into $\cup_{n=1}^{\infty}\{1, 2, \ldots, n\} \times \mathbb{Q}^{n+1}$. This is a countable union of countable sets where we note $\{1, 2, \ldots, n\} \times \mathbb{Q}^{n+1}$ is countable. This set is then itself countable. So, the set of algebraic numbers injects in $\mathbb{N}$. And, to see there are at least as many algebraic numbers as elements in $\mathbb{N}$, note that each $\alpha \in \mathbb{N}$ is itself algebraic and so $\mathbb{N}$ injects into the set of algebraic numbers. Thus, the set of algebraic numbers is equinumerous with the natural numbers, i.e., is countable. $\square$

F05.3: Let Prove that if $f_j : [0,1] \to \mathbb{R}$ is a sequence of functions which converges uniformly on $[0,1]$ to a (necessarily continuous) function $F : [0,1] \to \mathbb{R}$, then

$$\int_0^1 F^2(x)\ \mathrm{d}x = \lim_{j\to\infty} \int_0^1 f_j^2(x)\ \mathrm{d}x. \tag{63}$$

b) Give an example of a sequence $\{f_j : [0,1] \to \mathbb{R}\}_{j=1}^{\infty}$ of continuous functions which converges to a continuous function $F : [0,1] \to \mathbb{R}$ pointwise and for which

$$\lim_{j\to\infty} \int_0^1 f_j^2(x)\ \mathrm{d}x \text{ exists, but } \lim_{j\to\infty} \int_0^1 f_j^2(x)\ \mathrm{d}x \neq \int_0^1 F^2(x)\ \mathrm{d}x. \tag{64}$$

Proof:

a) Let $\varepsilon > 0$ be given. To prove the claim, we must find $N \in \mathbb{N}$ such that

$$\left| \int_0^1 F^2(x)\ \mathrm{d}x - \int_0^1 f_j^2(x)\ \mathrm{d}x \right| < \varepsilon\ \forall\ n \geq N,\ x \in [0,1]. \tag{65}$$

Since F is continuous and defined on the closed and bounded interval $[0,1]$, it follows from the Extreme value Theorem that F attains a maximum and a minimum on $[0,1]$. So, we can let $M_1 = \max\{|F(x)| : x \in [0,1]\}$. Similarly, each f_j attains a maximum and a minimum on $[0,1]$. So, let $M_2 = \sup\{\max\{|f_j(x)| : x \in [0,1]\} : n \in \mathbb{N}\}$. Using the uniform convergence of $\{f_j\}$ to F, we can pick $N \in \mathbb{N}$ such that $|F(x) - f_j(x)| < \varepsilon/(M_1 + M_2)$ whenever $n \geq N,\ x \in [0,1]$. Then

$$\begin{aligned}
\left| \int_0^1 F^2(x)\ \mathrm{d}x - \int_0^1 f_j^2(x)\ \mathrm{d}x \right| &\leq \int_0^1 \left| F^2(x)\ \mathrm{d}x - \int_0^1 f_j^2(x) \right|\ \mathrm{d}x \\
&= \int_0^1 |F(x) - f_j(x)|\,|F(x) + f_j(x)|\ \mathrm{d}x \\
&\leq \int_0^1 (F(x) - f_j(x)|\,|M_1 + M_2|\ \mathrm{d}x \\
&< \int_0^1 \left(\frac{\varepsilon}{M_1 + M_2} \right) |M_1 + M_2|\ \mathrm{d}x \\
&= \int_0^1 \varepsilon\ \mathrm{d}x \\
&= \varepsilon
\end{aligned} \tag{66}$$

whenever $n \geq N$, and we are done.

b) For each $j \in \mathbb{N}$, define $f_j : [0,1] \to \mathbb{R}$ by

$$f_j(x) = \begin{cases} 0 & \text{if } x \geq 1/j, \\ \sqrt{4j^2 x} & \text{if } 0 \leq x \leq 1/2j, \\ \sqrt{2j - 4j^2(x - 1/2j)} & \text{if } 1/2j < x < 1/j. \end{cases} \tag{67}$$

Then $f_j^2(x)$ forms a tent function of area one, i.e., $\int_0^1 f_j^2(x)\ \mathrm{d}x = 1$ for each $j \in \mathbb{N}$. To show the convergence of $\{f_j\}$, let $\varepsilon > 0$ be given and $x \in [0,1]$. Then by the Archimedean property of $\mathbb{R}$, we can pick $N \in \mathbb{N}$ such that $1/N < x$. It follows that $|f_j(x) - 0| = |0 - 0| = 0 < \varepsilon$ whenever $n \geq N$. Thus, for each $x \in [0,1]$, we have that $f_j(x) \to 0$ as $j \to \infty$. Consequently, $F(x) = 0$ and so $\int_0^1 F^2(x)\ \mathrm{d}x = \int_0^1 0\ \mathrm{d}x = 0 \neq 1 = \int_0^1 f_j^2(x)\ \mathrm{d}x$. (See F08.3 for a more detailed answer.)

$\square$

F05.5: Prove that $\mathbb{R}^2$ is not a (countable) union of sets $S_i,\ i = 1, 2, \ldots$ with each S_i being a subset of some straight line L_i in $\mathbb{R}^2$.

Proof:

Let $\{\ell_n\}_{n=1}^\infty$ be a collection of straight lines in $\mathbb{R}^2$. Then let θ_n be the angle of the line ℓ_n. Since $\{\theta_n\}_{n=1}^\infty$ is countable, we can pick a real number $\theta \in [0, 2\pi]$ that is distinct from all θ_n. Let ℓ be a line at angle θ. Then for all $n \in \mathbb{N}$, $\ell \neq \ell_n$. So, $(\cup_{n=1}^\infty \ell_n) \cap \ell$ is countable. But, ℓ has uncountably many points. Therefore, $\cup_{n=1}^\infty \ell_n$ cannot cover ℓ and $\mathbb{R}^2$ is not a (countable) union of straight lines.

$\square$

F05.8: For a real $n \times n$ matrix A, let $T_A : \mathbb{R}^n \to \mathbb{R}^n$ be the associated linear mapping. Set $\|A\| = \sup\limits_{x \in \mathbb{R}^n,\ \|x\| \leq 1} \|T_A x\|$ using the usual Euclidean norm for x^n.

a) Prove that $\|A + B\| \leq \|A\| + \|B\|$.

b) Use a) to check that the set M of all $n \times n$ matrices is a metric space if the distance function d is defined by $d(A, B) = \|B - A\|$.

c) Prove that M is a complete metric space with this "distance function."

Proof:

a)
$$\begin{aligned}\|A+B\| &= \sup_{x\in\mathbb{R}^n,\ \|x\|\le 1} \|(T_A+T_B)x\| \\ &= \sup_{x\in\mathbb{R}^n,\ \|x\|\le 1} \|T_Ax+T_Bx\| \\ &\le \sup_{x\in\mathbb{R}^n,\ \|x\|\le 1} \|T_Ax\|+\|T_Bx\| \\ &= \sup_{x\in\mathbb{R}^n,\ \|x\|\le 1} \|T_Ax\| + \sup_{x\in\mathbb{R}^n,\ \|x\|\le 1} \|T_Bx\| \\ &= \|A\|+\|B\|.\end{aligned} \tag{68}$$

b) To show that (M,d) forms a metric space we must show that it satisfies the following properties:
   i) For any $A\in M$, $d(A,A)=0$.
   ii) For any distinct $A,B\in M$, $d(A,B)>0$.
   iii) For any $A,B\in M$, $d(A,B)=d(B,A)$.
   iv) For any $A,B,C\in M$, $d(A,B)\le d(A,C)+d(C,B)$.

   Of course, $d(A,A)=\|T_A-T_A\|=\|0\|=0$. ...

c) To show that M is complete, we must show that every Cauchy sequence in (M,d) converges to a matrix in (M,d). So, let $\{A_n\}$ denote a Cauchy sequence in (M,d).

□

2006

W06.1: Show that for each $\varepsilon>0$ there exists a sequence of intervals $\{I_n\}$ with the properties

$$\mathbb{Q}\subset\bigcup_{n=1}^{\infty} I_n \quad\text{and}\quad \sum_{n=1}^{\infty}|I_n|<\varepsilon. \tag{69}$$

Proof:
Because the rationals are countable, there exists a bijection $f:\mathbb{N}\to\mathbb{Q}$. Then for each $i\in\mathbb{N}$, define $I_i=(f(i)-\varepsilon/2^{i+2}, f(i)+\varepsilon/2^{i+2})$. So, $f(i)\in I_i$ and we see

$$\mathbb{Q}\subset\bigcup_{n=1}^{\infty} I_n. \tag{70}$$

Now, observe that $|I_i| = \left|f(i) + \varepsilon/2^{i+2} - \left(f(i) - \varepsilon/2^{i+2}\right)\right| = \varepsilon/2^{i+1}$. This yields that

$$\sum_{n=1}^{\infty} |I_n| = \sum_{n=1}^{\infty} \frac{\varepsilon}{2^{i+1}} = \frac{\varepsilon}{2} \sum_{n=1}^{\infty} \frac{1}{2^i} = \frac{\varepsilon}{2} < \varepsilon. \tag{71}$$

Thus, our choice of $\{I_n\}$ has the desired properties. □

W06.2: Let $\{a_n\}_{n\geq 1}$ be a decreasing sequence of positive numbers such that $\sum_{n=1}^{\infty} a_n = \infty$. Under what condition(s) is the function

$$f(x) = \sum_{n=1}^{\infty} (-1)^n a_n x^n \tag{72}$$

well-defined and left-continuous at $x = 1$? Carefully prove your assertion.

Proof:

We claim that if $\lim\limits_{n\to\infty} a_n = 0$, then f is well-defined and left-continuous at $x = 1$. At $x = 1$, $f(x)$ is well-defined since there the alternating series test asserts the sum converges (since $\lim\limits_{n\to\infty} a_n = 0$). All that remains is to show f is left-continuous at $x = 1$.

Let $\varepsilon > 0$ be given. Then we must show that there exists $\delta > 0$ such that whenever $y < 1$ and $d(1, y) < \delta$, $|f(y) - f(1)| < \varepsilon$. NOT COMPLETE.

□

W06.4: Consider a decreasing sequence of continuous functions $f_n : [0, 1] \to \mathbb{R}$ obeying the uniform bound $|f_n| \leq M$ for some $M \in (0, 1)$. Suppose the point-wise limit $f(x) = \lim\limits_{n\to\infty} f_n(x)$ is continuous on $[0, 1]$. Prove that $f_n \to f$ uniformly on $[0, 1]$. (You may use without proof that $[0, 1]$ is compact as well as sequentially compact.)

Proof:

Let $\varepsilon > 0$ be given. Then we must find an integer N such that $|f_n(x) - f(x)| < \varepsilon$ whenever $n \geq N$ and $x \in [0, 1]$. Since we have a decreasing sequence of functions, $f_{n+1}(x) \leq f_n(x)$ for all $x \in [0, 1]$ and $n \in \mathbb{N}$. Then define $g_n = f_n - f$ and observe that $g_n(x) \geq 0$ for all $x \in [0, 1]$. Now define the sets $E_n = \{x \in [0, 1] \mid g_n(x) < \varepsilon\}$. Then E_n is open since it is the preimage of an open set under the continuous function g_n. Since $\{f_n\}$ is monotonically decreasing, so is $\{g_n\}$ and so $\{E_n\}$ is ascending, i.e., $E_n \subseteq E_{n+1}$ for each $n \in \mathbb{N}$. Because $\{f_n\} \to f$, it follows that $\{g_n\} \to 0$ and so the collection of $\{E_n\}$ form an open cover of $[0, 1]$. By compactness of $[0, 1]$, there is a finite subcover of this collection that covers $[0, 1]$. But, the E_n are ascending and so there is some $N \in \mathbb{N}$ such that E_N covers $[0, 1]$. Then for any $x \in [0, 1]$, if $n \geq N$, $x \in E_N$ and so $|f_n(x) - f(n)| < \varepsilon$, and we are done.

□

W06.6: Let $-\infty < a < b < \infty$. Prove that a continuous function $f : [a, b] \to \mathbb{R}$ attains all values in $[f(a), f(b)]$.

Proof:

First, prove that $[a, b]$ is connected. (See above notes.) Then prove the intermediate value theorem. (See above notes.) □

W06.7: Let V be a complex inner product space and $v, w \in V$. Prove the Cauchy-Schwarz inequality $|\langle v, w\rangle| \leq |v||w|$.

Proof:

We wish to write v as a scalar multiple of w added to some vector $y \in V$ that is orthogonal to w. So, observe that $v = cw + (v - cw)$. The vector cw will be orthogonal to $v - cw$ when $0 = \langle cw, v - cw\rangle = c\langle w, v\rangle - c^2 \langle w, w\rangle = c(\langle w, v\rangle - c\|w\|^2)$. Avoiding the trivial solution of $c = 0$, we see this implies $c = \langle v, w\rangle / \|w\|^2$. Then defining

$$y = v - \frac{\langle v, w\rangle}{\|w\|^2} w$$

yields

$$v = \frac{\langle v, w\rangle}{\|w\|^2} w + y.$$

where w and y are orthogonal. Then, by the Pythagorean Theorem, it follows that

$$\|v\|^2 = \left\| \frac{\langle v, w\rangle}{\|w\|^2} w \right\|^2 + \|y\|^2 \geq \left\| \frac{\langle v, w\rangle}{\|w\|^2} w \right\|^2 = \left(\frac{\langle v, w\rangle}{\|w\|^2} \right)^2 \|w\|^2 = \frac{\langle v, w\rangle^2}{\|w\|^2}.$$

Taking square roots of both sides gives $\|v\| \geq \langle v, w\rangle / \|w\|$, which implies $|\langle v, w\rangle| \leq \|v\|\|w\|$. □

W06.8: Let $T : V \to W$ be a linear transformation of finite dimensional real inner product spaces. Show that there exists a unique linear transformation $T^t : W \to V$ such that

$$\langle T(v), w\rangle_W = \langle v, T^t(w)\rangle_V \quad \forall\, v \in V, w \in W$$

where $\langle \cdot, \cdot\rangle_X$ is the inner product on $X = V$ or W.

Proof:

Let $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$ be orthonormal bases for V and W, respectively. Let A denote the matrix representation of T relative to these basis. By defining $T^t : W \to V$ to have the matrix representation A^t with respect to these basis, we see

$$\langle T(v), w\rangle = \langle Av, w\rangle = (Av)^t w = v^t A^t w = v^t(A^t w) = \langle v, A^t w\rangle = \langle v, T^t(w)\rangle .$$

This show that such a linear transformation T^t exists.

All that remains is to show uniqueness. Suppose $S : W \to V$ also exists and is such that

$$\langle v, T^t(w)\rangle_V = \langle T(v), w\rangle_W = \langle v, S(w)\rangle_V \;\; \forall\; v \in W, w \in W.$$

Then

$$0 = \langle v, T^t(w)\rangle_V - \langle v, S(w)\rangle_V = \langle v, T^t(w) - S(w)\rangle$$

for each $v \in V$. This is true precisely when $T^t(w) - S(w) = 0$. Using linearity, we have $(T^t - S)(w) = 0$ for each $w \in W$. But, this only holds when $T^t - S = 0$, i.e., $T^t = S$. Hence T^t is unique. $\square$

W06.9: Let $A \in M_3(\mathbb{R})$ be invertible and satisfy $A = A^t$ and $\det A = 1$. Prove that A has one as an eigenvalue.

Proof:

In order to show that 1 is an eigenvalue of A, it suffices to show that $\det(A - I) = 0$. Since $A = A^t$, it is self-adjoint and, thus, by the real spectral theorem, A is similar to a diagonal matrix. That is, there exists an orthogonal matrix P and diagonal matrix D such that $A = PDP^{-1}$. Then, using the properties of determinants,

$$\det(A) = \det(PDP^{-1}) = \det(P)\det(D)\det(P^{-1}) = \det(P)\det(D)\frac{1}{\det(P)} = \det(D) = \lambda_1\lambda_2\lambda_3$$

where $\lambda_1, \lambda_2, \lambda_3$ are the eigenvalues of A. Since $1 = \det(A) = \lambda_1\lambda_2\lambda_3$, we have $\lambda_1 = 1/(\lambda_2\lambda_3)$. Thus,

$$\det(A - I) = (\lambda_1 - 1)\,(\lambda_2 - 1)(\lambda_3 - 1) = \left(\frac{1}{\lambda_2\lambda_3} - 1\right)(\lambda_2 - 1)(\lambda_3 - 1).$$

If $\lambda_2 = 1$ or $\lambda_3 = 1$, then we are done. So, suppose this is not the case. Then we must show $1/(\lambda_2\lambda_3) - 1 = 0$, which is equivalent to showing $\lambda_2 = 1/\lambda_3$. Not Complete.
Hence

$$1 = \det(A) = \lambda_1\lambda_2\lambda_3 = \lambda_1\lambda_2(1/\lambda_2) = \lambda_1$$

and so $\lambda_1 = 1$. Hence A must have an eigenvalue of 1.

Suppose $A - I$ is invertible. The inverse of a matrix is unique and, is here given by $\sum_{k=0}^{\infty} A^k$ since

$$(I - A)\sum_{k=0}^{\infty} A^k = (I - A)(I + A + A^2 + \cdots) = I + (A - A) + (A^2 - A^2) + \cdots = I.$$

$\square$

S06.1:

a) Define precisely the notion of Riemann integrability for a function f on $[0,1]$.

b) Suppose that $f_n(x)$ is a sequence of Riemann integrable functions on $[0,1]$ such that $\{f_n(x)\}$ converges uniformly to $f(x)$. Prove that $f(x)$ is Riemann integrable.

Proof:

a) Let $P = \{I_1, \ldots, I_k\}$ be a partition of $[0,1]$. Then the upper sum $U(f,P)$ and lower sum $L(f,P)$ are, respectively, defined by

$$U(f,P) := \sum_{i=1}^{k} \left(\sup_{x \in I_i} f \right) |I_i| \quad \text{and} \quad L(f,P) := \sum_{i=1}^{k} \left(\inf_{x \in I_i} f \right) |I_i|. \tag{73}$$

The function $f(x)$ is Riemann integrable on $[0,1]$ provided

$$\inf_{P \in \Pi} U(f,P) = \sup_{P \in \Pi} L(f,P) \tag{74}$$

where Π denotes the set of all partitions of $[0,1]$.

b) Let $\varepsilon > 0$. Then by the uniform convergence of $\{f_n\}$ there is an $N \in \mathbb{Z}^+$ such that

$$\|f - f_n\|_\infty < \varepsilon \quad \forall\ n \geq N. \tag{75}$$

For any partition $P = \{I_1, \ldots, I_k\}$ this implies

$$\begin{aligned} L(f,P) &:= \sum_{i=1}^{k} \left(\inf_{x \in I_i} f \right) |I_i| \\ &\geq \sum_{i=1}^{k} \left(\inf_{x \in I_i} f_N - \varepsilon \right) |I_i| \\ &= \sum_{i=1}^{k} \left(\inf_{x \in I_i} f_N \right) |I_i| - \varepsilon \sum_{i=1}^{k} |I_i| \\ &= L(f_N, P) - \varepsilon. \end{aligned} \tag{76}$$

Since P was chosen arbitrarily, this holds for each $P \in \Pi$, from which we take the supremum to obtain

$$\sup_{P \in \Pi} L(f,P) \geq \sup_{P \in \Pi} L(f_N, P) - \varepsilon. \tag{77}$$

In similar fashion, we find

$$\inf_{P \in \Pi} U(f,P) \leq \inf_{P \in \Pi} U(f_N, P) + \varepsilon. \tag{78}$$

Then subtracting (77) from (78) we obtain

$$\inf_{P\in\Pi} U(f,P) - \sup_{P\in\Pi} L(f,P) < 2\varepsilon \tag{79}$$

where we have used the fact that f_N satisfies (74). By definition of the upper and lower sums, any partition P gives an upper sum $U(f,P)$ at least as large as $\sup_{P\in\Pi} L(f,P)$. So, the left hand side of (79) must be nonnegative. But, we also see (79) is true for all $\varepsilon > 0$ and none of the terms depend on ε, and so the left hand side must be equal to zero, as desired.

$\square$

S06.4: The point $P = (1,1,1)$ lies on the surface S in $\mathbb{R}^3$ defined by

$$x^2y^3 + x^3z + 2yz^4 = 4. \tag{80}$$

Prove that there exists a differentiable function $f(x,y)$ defined in an open neighborhood N of $(1,1)$ in $\mathbb{R}^2$ such that $f(1,1) = 1$ and $(x,y,f(x,y))$ lies in S for all $(x,y) \in N$.

Proof:
This problem is a routine application of the Implicit Function theorem. First define $F : \mathbb{R}^3 \to \mathbb{R}$ by $F(x,y,z) = x^2y^3 + x^3z + 2yz^4 - 4$. Then

$$\nabla F(x,y,z) = (2xy^3 + 3x^2z, 3x^2y^2 + 2z^4, x^3 + 8yz^3) \quad \Rightarrow \quad \nabla F(1,1,1) = (5,7,9). \tag{81}$$

Since $F_z(1,1,1) \neq 0$, it is invertible and the Implicit Function theorem implies there is a neighborhood $N \subset \mathbb{R}^2$ of $(1,1)$ and an open subset $V \subset \mathbb{R}$ containing 1 and a smooth function $f : N \to V$ such that $f(1,1) = 1$ and for $(x,y) \in N$, $F(x,y,f(x,y))$ is constant (namely, zero). (They probably want me to prove the Implicit Function theorem in this special case... I'll come back.) □

S06.9: Let S be a real $n \times n$ symmetric matrix S, i.e., $S^T = S$.

a) Prove the eigenvalues of S are real.

b) State and prove the Spectral Theorem for S.

Proof:

a) Let $\lambda \in \mathbb{C}$ be an eigenvalue of S and v be the corresponding eigenvector. Let $\langle \cdot, \cdot \rangle$ be the scalar product for $\mathbb{C}^n$. Then

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Av, v \rangle = \langle v, A^* v \rangle = \langle v, Av \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle \tag{82}$$

where $\bar{\lambda}$ denotes the complex conjugate of λ. By hypothesis, $v \neq 0$ and so $\langle v, v \rangle = \|v\|^2 > 0$, which implies $\lambda = \bar{\lambda}$. This implies the imaginary part of λ is zero, i.e., $\lambda \in \mathbb{R}$.

b) The real Spectral Theorem states that since S is symmetric, there exists an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of S and each eigenvalue of S is real. The last part of this was proven in a). We proceed by induction. The case for $n = 1$ holds trivially. Now suppose... $n - 1$ step. Then let $S \in \mathbb{R}^{n \times n}$ be symmetric.

□

2007

S07.7: Let $f : \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function with f'' uniformly bounded, and with a simple root x^* (i.e., $f(x^*) = 0$, $f'(x^*) \neq 0$). Consider the fixed point iteration

$$x_n = F(x_{n-1}) \quad \text{where} \quad F(x) = x - \frac{f(x)}{f'(x)}.$$

Show that if x_0 is sufficiently close to x^*, then there exists a constant C so that for all $n \in \mathbb{N}$, $|x_n - x^*| \leq C|x_{n-1} - x^*|^2$.

Proof:

Since f' is continuous and $f'(x^*) \neq 0$, there is a $\delta > 0$ such that $|f(x) - f(x^*)| \leq |f(x^*)/2|$ whenever $|x - x^*| \leq \delta$. Using the reverse triangle inequality, $|f'(x^*)| - |f'(x)| \leq |f'(x^*) - f'(x)|$ and so for such $|x - x^*| \leq \delta$ we have $|f'(x)| \geq |f'(x^*)|/2 > 0$. Using Taylor's theorem,

$$f(x^*) = f(x) + f'(x)(x^* - x) + f''(\xi)(x^* - x)^2,$$

for some ξ between x and x^*. For $|x - x^*| \leq \delta$, this implies

$$(x^* - x) + \frac{f(x) - f(x^*)}{f'(x)} = -\frac{f''(\xi)(x - x^*)^2}{f'(x^*)}.$$

But, $f(x^*) = 0$ and so

$$|x^* - F(x)| = \left| x^* - \left(x - \frac{f(x)}{f'(x)} \right) \right| = \left| \frac{f''(\xi)(x - x^*)^2}{f'(x^*)} \right| \leq C|x^* - x|^2$$

where C denotes the uniform bound for f'' divided by $|f'(x^*)|$. Let $x_0 \in [x^* - \delta, x^* + \delta]$. Then, using the above, for each $n \in \mathbb{N}$, $|x^* - x_n| = |x^* - F(x_{n-1})| \leq C|x^* - x_{n-1}|^2$. □

S07.8: Suppose the functions f_n are twice continuously differentiable on $[0,1]$ and satisfy

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \forall\, x \in [0,1] \quad \text{and} \quad \|f_n'\| \leq 1 \quad \text{and} \quad \|f_n''\| \leq 1. \tag{83}$$

Prove $f(x)$ is continuously differentiable on $[0,1]$.

Proof:
We first show $f_n \longrightarrow f$ uniformly. It suffices to show $\{f_n\}$ is uniformly Cauchy. Let $\varepsilon > 0$ be given and cover $[0,1]$ by open balls $B(x, \varepsilon/3)$ for rational $x \in [0,1]$. This is possible since the rationals are dense in $[0,1]$. And, since $[0,1]$ is compact this cover has a finite subcover, i.e., ther are $x_1, \ldots, x_p \in [0,1] \cap \mathbb{Q}$ such that

$$[0,1] \subset \bigcup_{i=1}^{m} B(x_i, \varepsilon/3). \tag{84}$$

Using the pointwise convergence of f, for each x_i there is an N_i such that

$$|f_n(x_i) - f_m(x_i)| < \varepsilon/3 \quad \forall\, n, m \geq N_i. \tag{85}$$

Set $N := \max\{N_1, \ldots, N_p\}$. Let $z \in [0,1]$. Then there is a $j \in \{1, \ldots, p\}$ such that $|x_j - z| < \varepsilon/3$. For $n \in \mathbb{Z}^+$, the mean value theorem implies there is c_n between z and x_j such that $f_n'(c_n)\cdot(x_j - z) = f_n(x_j) - f_n(z)$. Using the bound on f_n', we obtain

$$|f_n(x_j) - f_n(z)| = |f_n'(c)||x_j - z| \leq |x_j - z| < \frac{\varepsilon}{3}. \tag{86}$$

This implies

$$\begin{aligned} |f_n(z) - f_m(z)| &\leq |f_n(z) - f_n(x_j)| + |f_n(x_j) - f_m(x_j)| + |f_m(x_j) - f(z)| \\ &< |z - x_j| + \frac{\varepsilon}{3} + |z - x_j| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon \end{aligned} \tag{87}$$

whenever $n, m \geq N$. This holds for every $z \in [0,1]$ and so $\|f_n - f_m\| < \varepsilon$ whenever $n, m \geq N$, i.e., $\{f_n\}$ is uniformly Cauchy. Assume $f_n \longrightarrow g$ pointwise. Then the above process can be similarly applied to show $f_n \longrightarrow g$ uniformly.

We now show f' exists. Using the definition of derivative,

$$\begin{aligned}
\forall\ x \in [0,1], \quad f'(x) &= \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon) - f(x)}{\varepsilon} \\
&= \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{f_n(x+\varepsilon) - f_n(x)}{\varepsilon}] \\
&= \lim_{n \to \infty} \lim_{\varepsilon \to 0} \frac{f_n(x+\varepsilon) - f_n(x)}{\varepsilon}] \\
&= \lim_{n \to \infty} f_n'(x) \\
&= g
\end{aligned} \tag{88}$$

where we are able to interchange the limits with ε and n due to the uniform convergence.

Lastly, we show f' is continuous. Observe for $x, y \in [0,1]$ that $|x-y| < \varepsilon/3$ implies

$$\begin{aligned}
|f'(x) - f'(y)| &\leq |f'(x) - f_N'(x)| + |f_N'(x) - f_N'(y)| + |f_N'(y) - f(y)| \\
&< \frac{\varepsilon}{3} + |x-y| + \frac{\varepsilon}{3} \\
&< \varepsilon,
\end{aligned} \tag{89}$$

and we are done. $\square$

F07.8: Suppose $a_n \geq 0$ and $\sum_{n=1}^{\infty} a_n = \infty$. Does it follow that

$$\sum_{n=1}^{\infty} \frac{a_n}{1+a_n} = \infty? \tag{90}$$

Proof:
We claim that the above series does, in fact, diverge. To verify this, we first show that if $\{a_n/(1+a_n)\}_{n=1}^{\infty} \to 0$, then $\{a_n\}_{n=1}^{\infty} \to 0$. So, suppose $\{a_n/(1+a_n)\}_{n=1}^{\infty} \to 0$ and let $\varepsilon > 0$ be given. Then there exists $N \in \mathbb{N}$ such that $\forall\ n \geq N$,

$$\left| \frac{a_n}{1+a_n} \right| \leq \frac{\varepsilon}{1+\varepsilon}. \tag{91}$$

Rearranging, we see this implies

$$a_n \leq \frac{\varepsilon}{1+\varepsilon}(1+a_n) \quad \Rightarrow \quad a_n \leq \frac{\varepsilon/(1+\varepsilon)}{1-\varepsilon/(1+\varepsilon)} = \varepsilon. \tag{92}$$

Hence $\{a_n\}_{n=1}^{\infty} \to 0$.

Now suppose that $\{a_n\}_{n=1}^{\infty}$ is unbounded. Then $\{a_n\}_{n=1}^{\infty} \nrightarrow 0$ since for each $n \in \mathbb{N}$, we can find $n^* > n$ such that $a_{n^*} \geq \max\{a_1, \ldots, a_n\} + 1$. Thus, by the contrapositive of the above, $\{a_n/(1+a_n)\}_{n=1}^{\infty} \nrightarrow 0$, which implies

$$\sum_{n=1}^{\infty} \frac{a_n}{1+a_n} = \infty. \tag{93}$$

Alternatively, if $\{a_n\}_{n=1}^{\infty}$ is bounded by some $M > 0$, then

$$\sum_{n=1}^{\infty} \frac{a_n}{1+a_n} \geq \sum_{n=1}^{\infty} \frac{a_n}{1+M} = \frac{1}{1+M} \sum_{n=1}^{\infty} a_n = \infty, \tag{94}$$

and we are done. □

2008

S08.1: Let $g \in C([a,b])$, with $g(x) \in [a,b]$ for all $x \in [a,b]$. Prove the following:

a) g has at least one fixed point p in the interval $[a,b]$.

b) If there is a $\gamma < 1$ such that $|g(x) - g(y)| \leq \gamma|x - y|$ for all $x, y \in [a, b]$, then the fixed point p is unique and the iteration $x_{n+1} = g(x_n)$ converges to p for any initial guess $x_0 \in [a, b]$.

Proof:

a) Define $f(x) = g(x) - x$. Then observe that $f(a) \geq 0$ since $g(a) \geq a$ and $f(b) \leq 0$ since $g(b) \leq b$. If $g(a) = a$ or $g(b) = b$, then we are done. So, suppose this is not the case. Then $f(a) > 0$ and $f(b) < 0$. Then by the Intermediate Value Theorem, there is a $p \in (a, b)$ such that $f(p) = 0$, which implies $g(p) - p = 0$ and so $g(p) = p$. Hence there is at least one fixed point of g in $[a, b]$.

b) Suppose that there is a $\gamma < 1$ such that $|g(x) - g(y)| \leq \gamma|x - y|$ for all $x, y \in [a, b]$. Then $|g(x_n) - x_n| \leq \gamma|x_n - x_{n-1}|$ and so, by induction, $|x_{n+1} - x_n| = |g(x_n) - x_n| \leq \gamma^n|x_1 - x_0|$. We claim $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence. Indeed, let $\varepsilon > 0$ be given and note for $m > n$ that

$$\begin{aligned}
|x_m - x_n| &\leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \cdots + |x_{n+1} - x_n| \\
&\leq \sum_{j=0}^{m-n} \gamma^j |x_{n+1} - x_n| \\
&\leq \gamma^n |x_1 - x_0| \sum_{j=0}^{m-n} \gamma^j \\
&\leq \gamma^n |x_1 - x_0| \sum_{j=0}^{\infty} \gamma^j \\
&= \gamma^n \cdot \frac{|x_1 - x_0|}{1 - \gamma}.
\end{aligned}$$

But $\gamma^n \to 0$ as $n \to \infty$ and so we can pick $N \in \mathbb{N}$ such that $\gamma^n \cdot \dfrac{|x_1 - x_0|}{1 - \gamma} \leq \varepsilon$ whenever $n \geq N$, implying that $|x_m - x_n| \leq \varepsilon$ whenever $m > n \geq N$. Hence $\{x_n\}_{n=1}^{\infty}$ is Cauchy and, because $[a, b] \subset \mathbb{R}$ is complete, $\{x_n\}_{n=1}^{\infty}$ converges to some limit $p \in [a, b]$. Then

$$p = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} g(x_n) = g(p)$$

and so $g(p) = p$. All that remains is to verify p is unique. So, suppose there is also a q such that $g(q) = q$. If $q \neq p$, then

$$|q - p| = |g(q) - g(p)| \leq \gamma|q - p| < |q - p|,$$

which implies $|q - p| < |q - p|$, a contradiction. Hence $q = p$.

$\square$

S08.2 Let $\{f_n(x)\}$ be a sequence of continuous functions on the interval $[0,1]$ such that $f_n(x) \geq 0$ for all n and x and such that for all $x \in [0,1]$ $\lim_{n\to\infty} f_n(x) = 0$. Prove or give a counterexample to the assertion:

$$\lim_{n\to\infty} \int_0^1 f_n(x)\ \mathrm{d}x = 0.$$

Proof:
We claim the assertion is false and will prove this by providing a counter example. For each $n \in \mathbb{N}$, define

$$f_n(x) = \begin{cases} n^2 x & \text{if } x \in [0, 1/n], \\ n - n^2(x - 1/n) & \text{if } x \in (1/n, 2/n], \\ 0 & \text{if } x \in (2/n, 1] \end{cases}$$

Of course, f_n is continuous on $[0, 1/n)$, $(1/n, 2/n)$ and $(2/n, 1]$ since f_n is linear over each of these intervals. Thus, to show f_n is continuous, we need only verify that the right and left hand limits agree at $x = 1/n$ and $x = 2/n$. Indeed,

$$\lim_{x\to 1/n-} f(x) = n^2(1/n) = n = n - n^2(0) = n - n^2(1/n - 1/n) = \lim_{x\to 1/n+} f(x),$$

and

$$\lim_{x\to 2/n-} f(x) = n - n^2(2/n - 1/n) = n - n^2(1/n) = n - n = 0 = \lim_{x\to 2/n+} f(x).$$

Hence $\{f_n\}$ is a sequence of continuous functions. Moreover, we claim $\lim_{n\to\infty} f_n(x) = 0$ for each $x \in [0,1]$. Of course, $f_n(0) = 0$ for all $n \in \mathbb{N}$. To show this is true for $x \in (0,1]$, let $\varepsilon > 0$ be given. Then by the Archimedean property of $\mathbb{R}$ we can pick $N \in \mathbb{N}$ such that $1/N \leq x$. Consequently,

$$|f_n(x) - 0| = |0 - 0| \leq \varepsilon \quad \forall\ n \geq N.$$

Now observe that

$$\begin{aligned}
\int_0^1 f_n(x)\ \mathrm{d}x &= \int_0^{1/n} n^2 x\ \mathrm{d}x + \int_{1/n}^{2/n} n - n^2(x - 1/n)\ \mathrm{d}x + \int_{2/n}^{1} 0\ \mathrm{d}x \\
&= \int_0^{1/n} n^2 x\ \mathrm{d}x - \int_{1/n}^{2/n} n^2(x - 1/n)\ \mathrm{d}x + \int_{1/n}^{2/n} n\ \mathrm{d}x \\
&= \int_0^{1/n} n^2 x\ \mathrm{d}x - \int_0^{1/n} n^2 x\ \mathrm{d}x + n(1/n) \\
&= 1.
\end{aligned}$$

Hence $\lim_{n\to\infty} \int_0^1 f_n(x)\ \mathrm{d}x = \lim_{n\to\infty} 1 = 1 \neq 0.$

□

S08.3: Assuming that $f \in C^4[a,b]$ is real, derive a formula for the error of the approximation $E(h)$ when the second derivative is replaced by the finite-difference formula

$$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}$$

and h is the mesh size. Assume that $x, x+h, x-h \in (a,b)$.

Proof:

By Taylor's Theorem, we can expand f about x and evaluate at $x+h$ and $x-h$ to find

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(x)h^2 + \frac{f^{(3)}(x)}{6}h^3 + \frac{f^{(4)}(\xi_1)}{24}h^4$$

and

$$f(x-h) = f(x) - f'(x)h + \frac{1}{2}f''(x)h^2 - \frac{f^{(3)}(x)}{6}h^3 + \frac{f^{(4)}(\xi_2)}{24}h^4$$

where ξ_1 is between x and $x+h$ and ξ_2 is between x and $x-h$. Adding these equations together and then subtracting $2f(x)$ from each side gives

$$f(x+h) - 2f(x) + f(x-h) = f''(x)h^2 + \frac{f^{(4)}(\xi_1) + f^{(4)}(\xi_2)}{24}h^4.$$

Because $f \in C^4$ and $\frac{1}{2}(f^{(4)}(\xi_1) + f^{(4)}(\xi_2))$ is between $f^{(4)}(\xi_1)$ and $f^{(4)}(\xi_2)$, the Intermediate Value Theorem implies there is ξ between ξ_1 and ξ_2 such that $f^{(4)}(\xi) = \frac{1}{2}(f^{(4)}(\xi_1) + f^{(4)}(\xi_2))$. Hence

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} + \frac{f^{(4)}(\xi)}{12}h^2$$

and so

$$E(h) = \frac{f^{(4)}(\xi)}{12}h^2.$$

□

S08.4: Let X be a compact subset of $\mathbb{R}^N$ and let $\{f_n(x)\}$ be a sequence of continuous real functions on X such that $0 \leq f_{n+1}(x) \leq f_n(x)$ and $\lim_{n\to\infty} f_n(x) = 0$ for all $x \in X$. Prove Dini's Theorem that $f_n(x)$ converges to 0 uniformly on X.

Proof:

Let $\varepsilon > 0$ be given. Then we must show there is a $N \in \mathbb{N}$ such that for all $x \in X$, $|f_n(x) - f(x)| \leq \varepsilon$ where $f : X \to \mathbb{R}$ is defined for $x \in X$ by $f(x) := \lim_{n\to\infty} f_n(x) = 0$. To do this, define $E_n = \{x \in [0,1] \mid f_n(x) < \varepsilon\}$. Then E_n is open since $E_n = f_n^{-1}((-\infty, \varepsilon))$ is the preimage of an open set and f_n is continuous. Moreover, because f_n is monotonically decreasing, the the E_n are ascending, i.e., $E_n \subseteq E_{n+1}$ for each $n \in \mathbb{N}$. Since $\{f_n\} \to 0$ as $n \to \infty$, $\{E_n\}_{n=1}^\infty$ forms an open cover of $[0,1]$. Then, by the compactness of $[0,1]$, there is a finite subcover of $[0,1]$, which implies there is a $N \in \mathbb{N}$ such that E_N covers $[0,1]$. Then for all $n \geq N$ and $x \in [0,1]$, we have $x \in E_N$ and so $|f_n(x) - f(x)| = |f_n(x)| \leq \varepsilon$, completing the proof. $\square$

S08.10: Suppose A is an $n \times n$ complex matrix such that A has n distinct eigenvalues. Prove that if B is an $n \times n$ complex matrix such that $AB = BA$, the B is diagonalizable.

Proof:

content... $\square$

S08.12: Let $A \in M_n(\mathbb{R})$ be symmetric and $S := \{x \in \mathbb{R}^n : \|x\| = 1\}$ be the unit sphere in $\mathbb{R}^n$. Let $x \in S$ be such that

$$\langle Ax, x\rangle = \sup_{y\in S} \langle Ay, y\rangle \tag{95}$$

where $\langle z, y\rangle$ is the usual inner product on $\mathbb{R}^n$. (By compactness such x exists.)

a) Prove $\langle x, y\rangle = 0 \;\Rightarrow\; \langle Ax, y\rangle = 0$. Hint: Expand $\langle A(x + \varepsilon y), x + \varepsilon y\rangle$.

b) Use (a) to prove x is an eigenvector for A.

c) Use induction to prove $\mathbb{R}^n$ has an orthonormal basis of eigenvectors for A.

Note: If you use part c) to prove part a) or part b), then your solution should include a proof of c) that does not use part a) or part b).

Solution:

a) Define $f : \mathbb{R}^n \to \mathbb{R}$ by $f(y) = \langle Ay, y\rangle$ and define $g : \mathbb{R}^n \to \mathbb{R}$ by $g(y) = \langle y, y\rangle - 1$. Then $\langle Ax, x\rangle$ is the max of f subject to the constraint $g = 0$. Moreover,

$$\nabla f(y) = 2Ay \quad \text{and} \quad \nabla g(y) = 2y. \tag{96}$$

Lagrange's theorem for multipliers asserts that if the first partials of f and g are continuous and f attains an extremum at a point x subject to $g(x) = 0$, then $\nabla f(x) = \lambda\nabla g(x)$. Indeed, ∇f and ∇g are linear and so the partials of f and g are continuous. Hence

$$2Ax = \nabla f(x) = \lambda g(x) = \lambda 2x \qquad \Rightarrow \qquad Ax = \lambda x, \tag{97}$$

which implies x is an eigenvector of A (n.b. $x \neq 0$ since $\|x\| = 1$). Moreover,

$$0 = \langle x, y\rangle \qquad \Rightarrow \qquad 0 = \lambda \cdot 0 = \lambda \langle x, y\rangle = \langle \lambda x, y\rangle = \langle Ax, y\rangle, \tag{98}$$

as desired.

b) We already proved in a) that x is an eigenvector for A.

c) We now prove the real spectral theorem. If $n = 1$, then x forms an orthonormal basis for $\mathbb{R}$. Now suppose $n > 1$ and that the result holds for $k = 1, \ldots, n-1$. Again let $x \in \mathbb{R}^n$ be the eigenvector of A as above. Set $U = \text{span}(x)$ and observe $Ax = \lambda x \in U$. This implies U is invariant under A. In fact, if $u \in U$ and $w \in U^\perp$, then $0 = \langle u, w\rangle$ and so

$$\langle u, Aw\rangle = \langle A^t u, w\rangle = \langle Au, w\rangle = \langle \lambda u, w\rangle = \lambda \langle u, w\rangle = \lambda \cdot 0 = 0, \tag{99}$$

which implies $Aw \in U^\perp$. Hence $U^\perp$ is invariant under A. Moreover, since $\mathbb{R}^n = U \oplus U^\perp$, we have $\dim U^\perp = \dim C^n - \dim U = n - 1$. Since $U^\perp$ is invariant under A, the inductive hypothesis then implies the restriction of A to $U^\perp$ has an orthonormal basis consisting of eigenvectors of A, say $e_1, \ldots, e_{n-1}$. And, $\langle x, e_i \rangle = 0$ for $i = 1, \ldots, n-1$ since $x \in U$ and $e_i \in U^\perp$. Thus, $x, e_1, \ldots, e_{n-1}$ forms an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors for A. This closes the induction and the result follows.

We lastly verify $\mathbb{R}^n = U \oplus U^\perp$. Suppose $v \in U \cap U^\perp$. Then $\langle v, v \rangle = 0$, which implies $v = 0$ and so $U \cap U^\perp = \{0\}$. Now let $v \in \mathbb{R}^n$. Then

$$v = \underbrace{\langle v, x \rangle x}_{u} + \underbrace{v - \langle v, x \rangle x}_{w} = u + w \tag{100}$$

where $u \in U$ and $w \in U^\perp$. To see that $w \in U^\perp$, simply note that $\langle w, x \rangle = \langle v, x \rangle - \langle v, x \rangle = 0$. This shows $\mathbb{R}^n = U + U^\perp$. Combined with the fact $U \cap U^\perp = \{0\}$, we obtain $\mathbb{R}^n = U \oplus U^\perp$. $\square$

W08.01: Let $g \in C([a,b])$ with $a \leq g(x) \leq b$ for all $x \in [a,b]$. Prove the following:

a) g has at least one fixed point $p \in [a,b]$.

b) If there is $\gamma < 1$ such that $|g(x) - g(y)| \leq \gamma |x - y|$ for all $x, y \in [a,b]$, then the fixed point p is unique, and the iteration $x_{n+1} = g(x_n)$ converges to p for any initial guess $x_0 \in [a,b]$.

Proof:

a) Define $f : [a,b] \to [a,b]$ by $f(x) = g(x) - x$. Then f is continuous since it is the composition of continuous functions. We know $f(a) \geq 0$ since $g(a) \geq a$ and $f(b) \leq 0$ since $g(b) \leq b$. If $f(a) = 0$ or $f(b) = 0$, then we have a fixed point and are done. So, suppose this is not the case. Then $f(a) < 0$ and $f(b) > 0$. Then, by the Intermediate Value Theorem, there exists $x \in (a,b)$ such that $f(x) = 0$. Such an x is a fixed point of g.

b) This is essentially a proof of the Banach Fixed Point theorem. Since $[a,b]$ is complete, it suffices to show that $\{x_n\}$ is Cauchy. First note

$$|x_{n+1} - x_n| = |g(x_n) - g(x_{n-1})| \leq \gamma |x_n - x_{n-1}| \leq \cdots \leq \gamma^n |x_1 - x_0|. \tag{101}$$

This implies for $n > m$

$$\begin{aligned}
|x_n - x_m| &\leq \sum_{j=m}^{n-1} |x_{j+1} - x_j| && \text{Repeatedly apply triangle inequality} \\
&\leq \sum_{j=m}^{n-1} \gamma^j |x_1 - x_0| && \text{Use (101)} \\
&= \gamma^m |x_1 - x_0| \sum_{j=0}^{n-m-1} \gamma^j && \text{Reindex sum} \\
&\leq \gamma^m |x_1 - x_0| \sum_{j=0}^{\infty} \gamma^j && \text{Relate to geometric series} \\
&= \left(\frac{|x_1 - x_0|}{1 - \gamma} \right) \cdot \gamma^m. && \text{Compute limit of series}
\end{aligned} \tag{102}$$

Since $0 < \gamma < 1$, $\gamma^m \longrightarrow 0$ as $m \longrightarrow \infty$. So, for any $\varepsilon > 0$, we can find and N such that

$$|x_n - x_m| \leq \left(\frac{|x_1 - x_0|}{1 - \gamma} \right) \cdot \gamma^m < \varepsilon \quad \text{whenever } n, m \geq N. \tag{103}$$

Thus, the sequence $\{x_n\}$ is Cauchy. Since $[a,b]$ is complete, it follows there is a $x \in [a,b]$ such that $x_n \longrightarrow x$. But, this implies

$$x = \lim_{n \to \infty} x_n = \lim_{n \to \infty} x_{n+1} = \lim_{n \to \infty} g(x_n) = g\left(\lim_{n \to \infty} x_n \right) = g(x) \tag{104}$$

where we have used the continuity of g to pull the limit to the argument of g. Thus, the limit x is a fixed point of g, and this was irrespective of the initial guess x_0.

All that remains is to show that p is unique. By way of contradiction, suppose there is $q \neq p$ such that $g(q) = q$. Then

$$|p - q| = |g(p) - g(q)| \leq \gamma |p - q| < |p - q|, \tag{105}$$

a contradiction. Hence the fixed point is unique.

□

F08.1: For which of the values $a = 0, 1, 2$ is the function $f(t) = t^a$ uniformly continuous on $[0, \infty)$? Prove your assertion.

Proof:

For $a = 0$: We claim $f(t) = t^0 = 1$ is uniformly continuous on $[0, \infty)$. Let $\varepsilon > 0$. Then, for any $x, y \in [0, \infty)$ we see that

$$|f(x) - f(y)| = |x^0 - y^0| = |1 - 1| = 0 < \varepsilon. \tag{106}$$

For $a = 1$: We claim $f(t) = t^1 = t$ is uniformly continuous on $[0, \infty)$. Let $\varepsilon > 0$. Then for any $x, y \in [0, \infty)$ such that $|x - y| < \varepsilon$ we have $|f(x) - f(y)| = |x - y| < \varepsilon$.

For $a = 2$: We claim $f(t) = t^2$ is **not** uniformly continuous on $[0, \infty)$. Let $\varepsilon > 0$ be given. Then for each $\delta > 0$, picking $x = \varepsilon/2\delta$ yields that

$$|f(x + \delta) - f(x)| = |(x + \delta)^2 - x^2| = |2x\delta + \delta^2| > 2x\delta = \varepsilon. \tag{107}$$

Hence there does not exists a $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$ with $x, y \in [0, \infty)$. That is, $f(x)$ is not uniformly continuous. $\square$

F08:3 Give an example of a sequence of continuous real-valued functions f_n on $[0,1]$ such that $f(t) = \lim_{n\to\infty} f_n(t)$ is continuous, but for which $\int_0^1 f_n(t)\ \mathrm{d}t$ does not converge to $\int_0^1 f(t)\ \mathrm{d}t$.

Proof:

For each $n \in \mathbb{N}$, define $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} 0 & \text{if } x \geq 1/n, \\ 4n^2 x & \text{if } 0 \leq x \leq 1/2n, \\ 2n - 4n^2(x - 1/2n) & \text{if } 1/2n < x < 1/n. \end{cases} \tag{108}$$

The function $f_n(x)$ is clearly continuous in each interval $[0, 1/2n)$, $(1/2n, 1/n)$, and $(1/n, 1]$ since there f_n is the sum of a constant function and a multiple of x. Now, to check the two points of question (i.e., $1/2n$ and $1/n$), observe that

$$\lim_{x\to 1/2n^-} 4n^2 x = 4n^2(1/2n) = 2n = 2n - 4n^2(0) = \lim_{x\to 1/2n^+} 2n - 4n^2(x - 1/2n), \tag{109}$$

and

$$\lim_{x\to 1/n^-} 2n - 4n^2(x - 1/2n) = 2n - 4n^2(1/n - 1/2n) = 2n - 4n^2(1/2n) = 0 = \lim_{x\to 1/2n^+} 0. \tag{110}$$

So, f_n is, in fact, continuous. Furthermore, $f_n(x)$ forms a tent function of area one, i.e.,

$$\begin{aligned} \int_0^1 f_n(x)\ \mathrm{d}x &= \int_0^{1/2n} 4n^2 x\ \mathrm{d}x + \int_{1/2n}^{1/n} 2n - 4n^2(x - 1/2n)\ \mathrm{d}x \\ &= \int_0^{1/2n} 4n^2 x\ \mathrm{d}x + \int_{1/2n}^{1/n} 2n\ \mathrm{d}x - \int_0^1 4n^2(x - 1/2n)\ \mathrm{d}x \\ &= \int_0^{1/2n} 4n^2 x\ \mathrm{d}x + \int_0^{1/2n} 2n\ \mathrm{d}x - \int_0^{1/2n} 4n^2 x\ \mathrm{d}x \\ &= \int_0^{1/2n} 2n\ \mathrm{d}x \\ &= 2n(1/2n) \\ &= 1. \end{aligned} \tag{111}$$

To show the convergence of $\{f_n\}$, let $\varepsilon > 0$ be given and $x \in [0,1]$. Then by the Archimedean property of $\mathbb{R}$, we can pick $N \in \mathbb{N}$ such that $1/N < x$. It follows that $|f_n(x) - 0| = |0 - 0| = 0 < \varepsilon$ whenever $n \geq N$. Thus, for each $x \in [0,1]$, we have that $f_n(x) \to 0$ as $n \to \infty$. Consequently, $F(x) = 0$ and so $\int_0^1 F(x)\ \mathrm{d}x = \int_0^1 0\ \mathrm{d}x = 0 \neq 1 = \lim_{n\to\infty} \int_0^1 f_n(x)\ \mathrm{d}x$. $\square$

F08.4:

a) Suppose that K and F are subsets of $\mathbb{R}^2$ with K closed and bounded and F closed. Prove that if $K \cap F = \emptyset$, then $d(F, K) > 0$. Recall that $d(K, F) = \inf\{d(x, y) \mid x \in K, y \in F\}$.

b) Is a) true if K is just closed? Prove your assertion.

Proof:

a) We proceed by proving the contrapositive of the claim. So, suppose $d(F, K) \leq 0$. Since the metric function maps to $[0, \infty)$, it follows that $d(F, K) = 0$. This implies

$$0 = \inf\{d(x, y) \mid x \in K, y \in F\}. \tag{112}$$

That is, for each $\varepsilon > 0$, we can find $x \in K$ and $y \in F$ such that $d(x, y) < \varepsilon$. Thus, there exists sequences $\{x_n,\}$ and $\{y_n\}$ in K and F, respectively, such that $d(x_n, y_n) < 1/n$ for each $n \in \mathbb{N}$. But, because K is closed and bounded, it follows from the sequential compactness theorem that $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$ where $n_k \geq k$. We will denote the limit by x.

We claim that $\{y_n\} \to x$. To show this, let $\varepsilon > 0$ be given. Then, by the convergence of $\{x_n\}$, there exists $N_1 \in \mathbb{N}$ such that $d(x_{n_k}, x) < \varepsilon/2$ whenever $k \geq N_1$. By the Archimedean property of $\mathbb{R}$, there is $N_2 \in \mathbb{N}$ such that $1/N_2 < \varepsilon/2$. Define $N = \max\{N_1, N_2\}$. Then

$$d(y_{n_k}, x) \leq d(y_{n_k}, x_{n_k}) + d(x_{n_k}, x) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon \tag{113}$$

whenever $k \geq N$. Thus, $\{y_{n_k}\} \to x$. But, since F is closed, it must follow that $x \in F$. Similarly, $x \in K$ and so $x \in K \cap F$, which implies $K \cap F \neq \emptyset$.

$\square$

F08.5: Suppose $\sum_{n=1}^{\infty} a_n$ converges, but no absolutely. Show then that for any $a \in \mathbb{R}$, there is a rearrangement of $\sum_{n=1}^{\infty} a_n$ that converges to a.

Proof:

Given a sequence $\{a_n\}$, we introduce the sequences $\{a_n^+\}$ and $\{a_n^-\}$ defined by

$$a_n^+ = \begin{cases} a_n, \text{ if } a_n > 0, \\ 0, \text{if } a_n \leq 0, \end{cases} \qquad \text{and} \quad a_n^- = \begin{cases} -a_n, \text{ if } a_n < 0, \\ 0, \text{ if } a_n \geq 0. \end{cases}$$

Then note $|a_n| = a_n^+ + a_n^-$ and $a_n = a_n^+ - a_n^-$. We claim that $\sum_{n=1}^{\infty} a_n^+ = \sum_{n=1}^{\infty} a_n^- = \infty$. Indeed, since $\sum_{n=1}^{\infty} |a_n| = \sum_{n=1}^{\infty} a_n^+ + \sum_{n=1}^{\infty} a_n^-$, we cannot have $\sum_{n=1}^{\infty} a_n^+$ and $\sum_{n=1}^{\infty} a_n^-$ both be bounded. For this would imply $\sum_{n=1}^{\infty} |a_n|$ is bounded. And, since $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_n^+ - \sum_{n=1}^{\infty} a_n^-$, if either of $\sum_{n=1}^{\infty} a_n^+$ or $\sum_{n=1}^{\infty} a_n^-$ is unbounded, then so must the other. Otherwise, $\sum_{n=1}^{\infty} a_n$ would be unbounded.

We now construct a rearrangement of our series that converges to a. Let $b_1, b_2, \ldots$ be the numbers $a_1^+, a_2^+, \ldots$ in the same order, but with the zeros omitted, and let $c_1, c_2, \ldots$ be the numbers a_1^-, a_2^- in the same order, but with the zeros omitted. Then any series of all the b_n with plus signs and c_n with minus signs will be a rearrangement of $\sum_{n=1}^{\infty} a_n$, which we shall denote by $\sum_{n=1}^{\infty} a_n'$.

First suppose $a \geq 0$. Then let $a_1' = b_1$, $a_2' = b_2$, and so on until the first $n_1 \in \mathbb{N}$ such that

$$b_1 + \cdots + b_{n_1} > a.$$

Then take $a_{n_1+1}' = -c_1$, $a_{n_1+2}' = -c_2$, and so on until the first $n_2 \in \mathbb{N}$ such that

$$b_1 + \cdots + b_{n_1} - c_1 - \cdots - c_{n_2} < a.$$

Then repeat this process, adding b_n's until the sum is greater than a and then subtracting c_n's until the sum is less than a, back and forth, for each natural number. We claim $\sum_{n=1}^{\infty} a_n' = a$. Since $\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} -c_n = \infty$, no matter how many b_n's or c_n's have been used in the sum, those remaining will still add to ∞. So, at each step of the process there will be a sum of b_n that will get the partial sum above a and then c_n's that will get it below a. Let $N \in \mathbb{N}$ so that $n_1 + n_2 + \cdots + n_{k-1} \leq N < n_1 + n_2 + \cdots + n_k$ for some $k \in \mathbb{N}$. Then the difference between the partial sum $\sum_{n=1}^{N} a_n'$ and a is bounded by the finite sum $b_{n_{k-1}+1} + \cdots + b_{n_k}$ when k is odd and $c_{n_{k-1}+1} + \cdots + c_{n_k}$ when k is even. Since the sum of a_n's converges, $\lim_{n\to\infty} a_n = 0$ and so $\lim_{n\to\infty} b_n = \lim_{n\to\infty} c_n = 0$. Thus, given $\varepsilon > 0$, there is a $K \in \mathbb{N}$ such that $\left|\sum_{n=1}^{N} a_n' - a\right| < \varepsilon$ when $N \geq n_1 + n_2 + \cdots + n_{k-1}$ with $k \geq K$. Thus, this rearrangement converges to a. Similarly, if $a < 0$, then just begin with the c_n instead of the b_n.

□

2009

S09.1: Set $a_1 = 0$ and define the sequence $\{a_n\}$ via the recurrence

$$a_{n+1} = \sqrt{6 + a_n} \quad \text{for all } n \geq 1. \tag{114}$$

Show that this sequence converges and determine the limiting value.

Proof:

First, by induction, we show each a_n is bounded above by 3 for $n \in \mathbb{N}$. The base case holds since $a_0 = 0 \leq 3$. Supposing $a_n \leq 3$ for some $n \in \mathbb{N}$, we see

$$a_{n+1} = \sqrt{6 + a_n} \leq \sqrt{6 + 3} = 3 \tag{115}$$

and so we have closed the induction. We claim $a_{n+1} \geq a_n$ for each $n \in \mathbb{N}$. Indeed,

$$a_n^2 - a_n = a_n(a_n - 1) \leq 3(a_n - 1) \leq 3(3 - 1) = 6 \quad \Rightarrow \quad a_n^2 \leq 6 + a_n. \tag{116}$$

This implies $a_n \leq \sqrt{6 + a_n} = a_{n+1}$ and we have closed the induction. Thus, $\{a_n\}_{n=1}^{\infty}$ is monotonically increasing and is bounded above. Then by the monotone convergence theorem, this sequence converges to some limit L. Then, because $f(x) = \sqrt{6 + x}$ is continuous,

$$L = \lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} f(a_n) = f\left(\lim_{n\to\infty} a_n\right) = f(L) = \sqrt{6 + L}, \tag{117}$$

This implies $0 = L^2 - L - 6 = (L - 3)(L + 2)$. Since $a_n \geq 0$, it follows that $L = 3$. □

S09.2: Compute the norm of the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 0 & \sqrt{3} \end{bmatrix}. \tag{118}$$

That is, determine the maximum value of the length of Ax over all unit vectors x.

Solution:

We seek to compute $\|A\|_2$, which is defined by

$$\|A\|_2 := \sup\left\{\|Ax\|_2 \ : \ \|x\| = 1\right\}. \tag{119}$$

Using the spectral radius, we know

$$\|A\|_2 = \sqrt{\lambda_{max}(A^T A)} \tag{120}$$

where $\lambda_{max}(A^T A)$ denotes the maximum eigenvalue of the matrix $A^T A$. We compute the eigenvalues of $A^T A$ as follows. First note

$$A^T A = \begin{bmatrix} 2 & 0 \\ 1 & \sqrt{3} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & \sqrt{3} \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}. \tag{121}$$

So, the characteristic polynomial $\chi(\lambda)$ is given by

$$\chi(\lambda) := \det(A^T A - \lambda I) = (4-\lambda)^2 - 4 = \lambda^2 - 8\lambda + 12 = (\lambda - 6)(\lambda - 2). \tag{122}$$

Hence the eigenvalues are 2 and 6, from which we conclude $\boxed{\|A\|_2 = \sqrt{6}}$. □

S09.3: We wish to find a quadratic polynomial P obeying

$$P(0) = \alpha, \quad P'(0) = \beta, \quad P(1) = \gamma, \quad P'(1) = \delta \tag{123}$$

where $'$ denotes differentiation.

a) Find a minimal system of linear constraints on $(\alpha, \beta, \gamma, \delta)$ such that this is possible.

b) When the constraints are met, what is P? Is it unique? Explain your answer.

Solution:

a) Sine P is quadratic, we can write $P(x) = ax^2 + bx + c$ for some $a, b, c \in \mathbb{R}$. The first constraint gives $P(0) = c = \alpha$, the second $\beta = P'(0) = [2ax + b]_{x=0} = b$, the third $\gamma = P(1) = [ax^2 + bx + c]_{x=1} = a + b + c$, and the fourth $\delta = P'(1) = [2ax + b]_{x=1} = 2a + b$. Expressing this as a linear system we can write

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix}, \tag{124}$$

which can be reduced to

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta - 2\gamma + \beta + 2\alpha \end{bmatrix}. \tag{125}$$

So, a solution exists provided $\delta = 2\gamma - \beta - 2\alpha$. In this case, we obtain

$$\underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}}_{b}. \tag{126}$$

Define A, x, and b to be the underbraced quantities.

b) Yes, the solution P is unique. By simple computation, we find $\det(A) = -1 \neq 0$ and so A is invertible. This implies $x = A^{-1}b$, which gives an explicit expression for each of the coefficients of P. Namely,

$$x = A^{-1}b = \begin{bmatrix} -\alpha - \beta + \gamma \\ \beta \\ \alpha \end{bmatrix} \quad \Rightarrow \quad P(x) = (\gamma - \alpha - \beta)x^2 + \beta x + \alpha \tag{127}$$

where $\delta = 2\gamma - \beta - 2\alpha$.

$\square$

F09.4: Let V be a finite dimensional $\mathbb{R}$-vector space, whose dimension we denote by $\dim(V)$, equipped with an inner product $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$. For a vector space $U \subseteq V$, denote by $U^\perp$ its orthogonal complement, i.e., the set of $v \in V$ such that $\langle v, u\rangle = 0$ for all $u \in U$. Show that $\dim(U) + \dim(U^\perp) = \dim(V)$.

Proof:

First we will show that $V = U \oplus U^\perp$. Let $v \in V$ and $e_1, \ldots, e_m$ be an orthonormal basis of U. Then

$$v = \underbrace{\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m}_{u} + \underbrace{v - \langle v, e_1\rangle e_1 - \cdots - \langle v, e_m\rangle e_m}_{w}. \tag{128}$$

Let u and w be defined as in the above equation. Clearly, $u \in U$. Because $e_1, \ldots, e_m$ is an orthonormal list, for each $j = 1, \ldots, m$ we have $\langle w, e_j\rangle = \langle v, e_j\rangle - \langle v, e_j\rangle = 0$. Thus, w is orthogonal to every vector in $\text{span}\{e_1, \ldots, e_m\}$, which implies $w \in U^\perp$. Thus, we have written $v = u + w$ where $u \in U$ and $w \in U^\perp$. So, $V = U + U^\perp$. Now, suppose $y \in U \cap U^\perp$. Then $\langle y, y\rangle = 0$, which implies $y = 0$. Thus, $U \cap U^\perp \subset \{0\}$ and, since 0 is in every vector space, $U \cap U^\perp = \{0\}$. This implies $V = U \oplus U^\perp$. Now let $f_1, \ldots, f_n$ be a basis for $U^\perp$. Then, because $U \cap U^\perp = \{0\}$ and $U \oplus U^\perp = V$, we see $e_1, \ldots, e_m, f_1, \ldots, f_n$ must form a basis for V. Thus, $\dim(V) = n + m = \dim(U^\perp) + \dim(U)$.

$\square$

F09.12: Let V be an n-dimensional vector space over $\mathbb{C}$ (with $n \geq 2$) with a set of basis vector $e_1, \ldots, e_n$. Let T be a linear transformation of V satisfying $T(e_1) = e_2, \ldots, T(e_{n-1}) = e_n, T(e_n) = e_1$.

i) Show that T has 1 as an eigenvalue and write down an eigenvector with value 1. Show that, up to scaling, it is unique.

ii) Is T diagonalisable? (Hint: Calculate the characteristic polynomial.)

Proof:

i) Let $v = e_1 + \cdots e_n$. Then

$$Tv = T(e_1 + \cdots + e_{n-1} + e_n) = e_2 + \cdots + e_n + e_1 = e_1 + \cdots + e_{n-1} + e_n = v. \tag{129}$$

So, there exists $v \in V$ with eigenvalue 1. Now we must show it is unique up to scaling. Let $w \in V$ be any eigenvector with eigenvalue 1. Then there exists scalars $a_1, \ldots, a_n \in \mathbb{C}$ such that $w = a_1e_1 + \cdots + a_ne_n$. Then

$$T(w) = T(a_1e_1 + \cdots + a_ne_n) = a_1e_2 + \cdots a_ne_1 = a_ne_1 + a_1e_2 + \cdots a_{n-1}e_n. \tag{130}$$

This implies $a_1 = a_2$, $a_2 = a_3$, ..., $a_n = a_1$. That is, $a_1 = a_2 = \cdots = a_n$. Let $\alpha = a_1$. Then $w = \alpha(e_1 + e_2 + \cdots + e_n) = \alpha v$. Hence, up to scaling, the eigenvector v of T with eigenvalue 1 is unique.

ii)

$\square$

2010

S10.1: Let $u_1, \ldots, u_n$ be an orthornormal basis of $\mathbb{R}^n$ and let $y_1, \ldots, y_n$ be a collection of vectors in $\mathbb{R}^n$ satisfying $\sum_i \|y_i\|^2 < 1$. Prove that the vectors $u_1 + y_1, \ldots, u_n + y_n$ are linearly independent.

Proof:
Lemma: Let $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$. If $\|A\| < 1$, then $I - A$ is invertible.
Suppose $I - A$ is singular. Then there exits $v^* \in \mathbb{R}^n$ such that $(I - A)v^* = 0$, which is equivalent to saying $Av^* = Iv^* = v^*$. That is, $I - A$ is singular iff 1 is an eigenvalue of A. If 1 is an eigenvalue of A, then there is a unit vector $v \in \mathbb{C}^n$ such that $\|A\| \geq \|Av\| = \|v\| = 1$. Thus, if $\|A\| < 1$, then 1 is not an eigenvalue of A, which implies $I - A$ is invertible. ■

Define $L : \mathbb{R}^n \to \mathbb{R}^n$ by $L(u_i) = -y_i$. Then the columns of the matrix $I - L$ form the vectors $u_i + y_i$ for $i = 1, \ldots, n$. And $I - L$ is invertible iff these vectors are linearly independent. Then, by the above lemma, it suffices to show $\|A\| < 1$.

Let $x \in \mathbb{C}^n$ be a unit vector. Then there are unique scalars $a_1, \ldots, a_n$ such that $x = a_1u_1 + \cdots + a_nu_n$. Then, using the triangle and Hölder's inequalities,

$$\|Lx\| = \left\| -\sum_{i=1}^n a_i y_i \right\| \leq \sum_{i=1}^n |a_i| \, \|y_i\| \leq \left(\sum_{i=1}^n a_i^2\right)^{1/2} \left(\sum_{i=1}^n \|y_i\|^2\right)^{1/2} \leq \|x\| \cdot \left(\sum_{i=1}^n \|y_i\|^2\right)^{1/2}.$$

So, for all nonzero x we have

$$\frac{\|Lx\|}{\|x\|} \leq \left(\sum_{i=1}^n \|y_i\|^2\right)^{1/2} < 1.$$

To see that the last inequality holds, define $\alpha := \sum_i \|y_i\|^2$. If $\sqrt{\alpha} \geq 1$, then $\alpha = \sqrt{\alpha}\sqrt{\alpha} \geq \sqrt{\alpha} \geq 1$, a contradiction. This implies $\|L\| < 1$ and we are done. □

S10.2: Let A be a $n \times n$ real symmetric matrix and let $\lambda_1 \geq \ldots \geq \lambda_n$ be the eigenvalues of A. Prove that

$$\lambda_k = \max_{U, \dim U = k} \min_{x \in U, \|x\|=1} \langle Ax, x \rangle ,$$

where $\langle \cdot, \cdot \rangle$ denotes the usual scalar product in $\mathbb{R}^n$ and the maximum is taken over all k dimensional subspaces of $\mathbb{R}^n$.

Proof:
First note that since A is symmetric, it follows from the Real Spectral Theorem that there exists an orthogonal matrix P such that $A = P^T \Lambda P$ where Λ is a diagonal matrix, consisting of the eigenvalues of A. Consequently, for any $x \in \mathbb{R}^n$

$$\langle Ax, x \rangle = \langle P^T \Lambda P x, x \rangle = (P^T \Lambda P x)^T x = (\Lambda P x)^T P x = \langle \Lambda P x, P x \rangle .$$

Note

$$\|Px\|^2 = \langle Px, Px \rangle = \langle P^T P x, x \rangle = \langle P^{-1} P x, x \rangle = \langle x, x \rangle = \|x\|^2.$$

Thus, for $x \in U$ with $\|x\| = 1$

$$\langle Ax, x \rangle = \langle \Lambda(Px), Px \rangle = \sum_{i=1}^{n} \lambda_i (Px)_i^2 \geq \lambda_{min} \sum_{i=1}^{n} (Px)_i^2 = \lambda_{min} \|Px\|^2 = \lambda_{min}$$

where λ_{min} denotes the minimum eigenvalue of A with an eigenvector Px with $x \in U$. In each k dimensional subspace U of $\mathbb{R}^n$, A has eigenvectors corresponding to k eigenvalues. From the above, it follows that the subspace U that maximizes the expression is that with the largest eigenvalues; namely, it contains the eigenvectors corresponding to $\lambda_1, \ldots, \lambda_k$. Since λ_k is the smallest of these, it follows that in this subspace the minimum of $\langle Ax, x \rangle$ will be λ_k. Hence the desired equality holds. □

S10.3: Let S and T be two normal transformations in the complex finite dimensional vector space V with a positive definite Hermitian inner product such that $ST = TS$. Prove that S and T have joint basis vectors.

Proof:
Using the commutativity....
Let $v \in E_{\lambda,T}$. Then

$$T(Sv) = S(Tv) = S(\lambda v) = \lambda(Sv) \qquad \Rightarrow \qquad (Sv) \in E_{\lambda,T}$$

and so $S(E_{\lambda,T}) \subset E_{\lambda,T}$. $\square$

S10.4: i) Let $A = (a_{ij})$ be an $n \times n$ real symmetric matrix such that $\sum_{i,j} a_{ij}x_ix_j \leq 0$ for every vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ Prove that if $\text{tr}(A) = 0$, then $A = 0$.
ii) Let T be a linear transformation in the complex finite dimensional vector space V with a positive definite Hermitian inner product. Suppose that $TT^* = 4T - 3I$, where I is the identity transformation. Prove that T is positive definite Hermitian and find all possible eigenvalues of T.

Proof:

i) Recall that $\text{tr}(A) = \lambda_1 + \cdots + \lambda_n$ where λ_i gives the i-th eigenvalue of A for $i = 1, \ldots, n$. Let v_i be an eigenvector of A with eigenvalue λ_i. Then

$$0 \geq \langle Av, v\rangle = \langle \lambda v, v\rangle = \lambda \langle v, v\rangle .$$

Since $v \neq 0$, $\langle v, v\rangle > 0$ and so the above implies $\lambda \leq 0$. Thus, $\text{tr}(A) \leq 0$ with equality holding precisely when $\lambda_i = 0$ for each $i = 1, \ldots, n$. So, if $\text{tr}(A) = 0$, then all of the eigenvalues are zero. And, since A is real symmetric, it is similar to a diagonal matrix, consisting of the eigenvalues of A. This implies A is similar to the zero matrix, i.e., there exists invertible P such that $A = P0P^{-1}$. Then $A = P(0P^{-1}) = P0 = 0$.

ii) Recall that $(TT^*)^* = (T^*)^*T^* = TT^*$. This implies $4T - 3I = (4T - 3I)^* = 4T^* - 3I^* = 4T^* - 3I$. For $v \in V$ we have $\langle (4T - 3I)v, v\rangle = 4\langle Tv, v\rangle - 3\langle v, v\rangle$ and $\langle (4T^* - 3I)v, v\rangle = 4\langle T^*v, v\rangle - 3\langle v, v\rangle$. Equating these, we see $\langle Tv, v\rangle = \langle T^*v, v\rangle$, which implies $T = T^*$ and so T is self adjoint, i.e., T is Hermitian. Then $4T - 3I = TT^* = T^2$ and so $T^2 - 4T + 3I = 0$. Thus, by factoring, we see $(T-3I)(T-I) = 0$ and the eigenvalues of T are contained in $\{1, 3\}$.

All that remains is to show T is positive definite. Since T is normal, it follows from the complex spectral theorem that their exists unitary P and diagonal $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ such that $T = P^*\Lambda P$. This implies for $x \in V$ that

$$\langle Tx, x\rangle = \langle P^*\Lambda Px, x\rangle = \langle \Lambda Px, Px\rangle = \sum_{i=1}^{n} \lambda_i (Px)_i^2 \geq \sum_{i=1}^{n} (Px)_i^2 = \langle Px, Px\rangle = \langle P^*Px, x\rangle .$$

where the inequality follows from the above where it was shown that $\lambda_i \geq 1$ for each i. Thus,

$$\langle Tx, x\rangle = \langle P^*\Lambda Px, x\rangle \geq= \langle P^{-1}Px, x\rangle = \langle x, x\rangle = \|x\|^2 \geq 0.$$

where we use the fact that $P^* = P^{-1}$. Hence for $x \neq 0$ we have $\langle Tx, x\rangle > 0$, i.e., T is positive definite, as desired.

$\square$

S10.5: Let A, B be two $n \times n$ complex matrices which have the same minimal polynomial $M(t)$ and the same characteristic polynomial $P(t) = (t - \lambda_1)^{a_1} \cdots (t - \lambda_k)^{a_k}$ where $\lambda_i \neq \lambda_j$ for $i \neq j$. Prove that if $P(t)/M(t) = (t - \lambda_1) \cdot (t - \lambda_k)$, then these matrices are similar.

Proof:

Using $P(t)/M(t)$, for $i = 1, \ldots, k$ we see that there is a Jordan block of size $a_i - 1$ with the eigenvalue λ_i along the diagonal. This implies there is only one remaining Jordan block for each λ_i and that it is of size 1. So, we can uniquely identify the Jordan form J of A, up the order of the Jordan blocks. So, there is invertible P such that $A = P^{-1}JP$. Similarly, there is invertible Q such that $B = QJQ^{-1}$. Since $J = PAP^{-1}$, it follows that

$$B = Q(PAP^{-1})Q^{-1} = (QP)A(P^{-1}Q^{-1}) = (QP)A(QP)^{-1},$$

which implies that A and B are similar. $\square$

S10.6: Let $A = \begin{bmatrix} 4 & -4 \\ 1 & 0 \end{bmatrix}$.

a) Find the Jordan form J of A and a matrix P such that $P^{-1}AP = J$.

b) Compute A^{100} and J^{100}.

c) Find a formula for a_n when $a_{n+1} = 4a_n - 4a_{n-1}$ and $a_0 = a$ and $a_1 = b$.

Solution:

a) First observe that the characteristic polynomial of A is given by $\chi(\lambda) = \det(\lambda I - A) = \lambda(\lambda - 4) + 4 = (\lambda - 2)^2$. Moreover, the minimal polynomial of A is equal to the characteristic polynomial because

$$A - 2I = \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} \neq 0.$$

This implies the Jordan form of A is a single Jordan block, i.e.,

$$J = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

where $\lambda = 2$ is the single eigenvalue of A and has multiplicity 2. Solving the linear system $(A - 2I)v = 0$, it is clear that $v_1 = (2, 1)$ is an eigenvalue of A. Furthermore, since

$$(A - 2I)^2 = \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

it follows that any vector v_2 not equal to a scalar multiple of v_1 is a generalized eigenvector of A, e.g., $v_2 = (1, 0)$. Then let $P[v_1\ v_2]$ and observe that

$$AP = \begin{bmatrix} 4 & -4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad PJ = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 2 & 1 \end{bmatrix}$$

and so $AP = PJ$. Since v_1 and v_2 are not scalar multiples, they are independent and P is invertible. Hence $A = PJP^{-1}$.

b) Now, observe that

$$J = \lambda I + N = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then, by the binomial theorem,

$$J^n = (\lambda I + N)^n = \sum_{r=0}^{n} \binom{n}{r} \lambda^{n-r} N^r.$$

But, $N^2 = 0$ and so $N^r = 0$ for all $r \geq 2$. Thus,

$$J^n = \sum_{r=0}^{1} \binom{n}{r} \lambda^{n-r} N^r = \binom{n}{0} \lambda^n N^0 + \binom{n}{1} \lambda^{n-1} N = \lambda^n I + n\lambda^{n-1} N = \begin{bmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{bmatrix}.$$

So,

$$J^{100} = \begin{bmatrix} 2^{100} & 100 \cdot 2^{99} \\ 0 & 2^{100} \end{bmatrix}$$

and

$$\begin{aligned} A^{100} &= \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2^{100} & 100 \cdot 2^{99} \\ 0 & 2^{100} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2^{100} & 100 \cdot 2^{99} \\ 0 & 2^{100} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix}. \end{aligned}$$

c) Let $\alpha_n = (a_{n+1}, a_n)$. Then

$$A\alpha_n = \begin{bmatrix} 4 & -4 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 4a_{n+1} - 4a_n \\ a_{n+1} \end{pmatrix} = \alpha_{n+1}.$$

Thus, in general, $\alpha_n = A^n \alpha_0$. But,

$$A^n = PJ^nP^{-1} = \lambda^{n-1} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda & n \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} = \lambda^{n-1} \begin{bmatrix} 2n+\lambda & -4n \\ n & \lambda - 2n \end{bmatrix}.$$

Hence

$$\alpha_n = \lambda^{n-1} \begin{bmatrix} 2n+\lambda & -4n \\ n & \lambda - 2n \end{bmatrix} \begin{pmatrix} a_1 \\ a_0 \end{pmatrix},$$

which implies $a_{n+1} = \lambda^{n-1} \left([2n+\lambda]a_1 - 4na_0 \right)$ and so $\boxed{a_n = \lambda^{n-2} \left([2n-2+\lambda]b - 4(n-1)a \right).}$

□

S10.7: Let $\{f_n\}$ be a sequence of real-valued functions on the line, and assume that there is a $B < \infty$ such that $|f_n(x)| \leq B$ for all n and x. Prove that there is a subsequence $\{f_{n_k}\}$ such that $\lim_{k\to\infty} f_{n_k}(r)$ exists for all rational numbers r.

Proof:

We proceed making use of the standard diagonalization argument to construct the desired subsequence, which is here denoted by $\{f_{m(j)}\}_{j=1}^{\infty}$. First let $\sigma : \mathbb{N} \to \mathbb{Q}$ be a bijection. Since $f_n(\sigma(1)) \in [-B, B]$ for each $n \in \mathbb{N}$ and $[-B, B]$ is closed and bounded, the Bolzano-Weierstrass Theorem implies there exists a subsequence $\{f_{n_1(j)}(\sigma(1))\}_{j=1}^{\infty}$ that converges in $[-B, B]$. Similarly, from the sequence of functions $\{f_{n_1(j)}\}_{j=1}^{\infty}$ we can find a subsequence we can find a subsequence $\{f_{n_2(j)}\}_{j=1}^{\infty}$ such that the limit $\lim_{k\to\infty} f_{n_2(j)}\sigma(2)$ exists. We can continue in this fashion inductively, i.e., for each subsequence $n_k(j)$ we can find a subsubsequence $n_{k+1}(j)$ such that $\lim_{k\to\infty} f_{n_{k+1}(j)}\sigma(k+1)$ exists. By the construction of n_{k+1}, we further have $\lim_{k\to\infty} f_{n_{k+1}(j)}\sigma(i)$ exists for $i = 1, \ldots, k+1$. Using this, define the sequence $\{m(j)\}_{j=1}^{\infty}$ by $m(j) = n_j(j)$.

We claim $\lim_{j\to\infty} f_{m(j)}(q)$ exists for each $q \in \mathbb{Q}$. To show this, let $q \in \mathbb{Q}$ be given. Then there is $k \in \mathbb{N}$ such that $q = \sigma(k)$. Then there are only finitely many terms in the sequence $\{f_{m(j)}(\sigma(k))\}_{j=1}^{\infty}$ that are not also in $\{f_{n_k(j)}(\sigma(k))\}_{j=1}^{\infty}$, namely, the $k-1$ terms $f_{n_k(1)}(\sigma(k)), \ldots, f_{n_k(k-1)}(\sigma(k))$. Hence $\{f_{m(j)}(\sigma(k))\}_{j=1}^{\infty}$ must converge to the limit of $\{f_{n_k(j)}(\sigma(k))\}_{j=1}^{\infty}$. Because q was chosen arbitrarily, this holds for each rational number, completing the proof. $\square$

S10.8: Assume that K is a closed subset of a complete metric space (X, d) with the property that, for any $\varepsilon > 0$, K can be covered by a finite number of sets $B_\varepsilon(x) = \{y \in X \mid d(x,y) < \varepsilon\}$. Prove that K is compact.

Proof:
By definition of the completeness of X, every Cauchy sequence in X converges to a point in X. Since $K \subseteq E$, every Cauchy sequence in K converges to a point in X. Moreover, because K is closed, K contains its limit points and so every Cauchy sequence in K converges in K, i.e., K is complete. Since a K is compact iff it is sequentially compact, we prove that K is sequentially compact. Let $\{x_n\}$ be a sequence in K. It suffices to show this sequence has a convergent subsequence $\{x_{n_k}\}$. By completeness of K, this is accomplished if we show $\{x_n\}$ has a Cauchy subsequence $\{x_{n_k}\}$.

Cover X by finitely many balls of radius 1. (This is possible since X is totally bounded). By the pigeonhole principle, at least one of these balls must have an infinite number of x_i. Call this ball B_1 and let S_1 be the set of integers i for which $x_i \in B_1$. Continuing in an inductive fashion, for each $k \in \mathbb{N}$ with $k > 1$, we define B_k to be the intersection of B_{k-1} and an open ball of radius $1/k$ containing an infinite number elements from the collection $\{x_i\}_{i \in S_{k-1}}$. Let S_k be the collection of indices i of these $x_i \in B_k$. Each of the S_k is infinite. So, we can pick a sequence $\{n_k\}$ with $n_k < n_{k+1}$ for each k. Since the S_k are nested, it follows that whenever $i, j \geq k$, then $n_i, n_j \in S_k$. Hence for $i, j \geq k$, x_{n_i} and x_{n_j} are contained some open ball of radius $1/k$ centered at a point c_k so that

$$d(x_{n_i}, x_{n_j}) \leq d(x_{n_i}, c_k) + d(c_k, x_{n_j}) \leq 2/k.$$

By the Archimedean property of $\mathbb{R}$, for each $\varepsilon > 0$ we can find $N \in \mathbb{N}$ such that $1/N < \varepsilon/2$. Hence $d(x_{n_i}, x_{n_j}) < 1/N < \varepsilon$ whenever $i, j \geq N$. Thus, $\{x_{n_k}\}$ is Cauchy. $\square$

S10.9: Assume $f(x, y, z)$ is real valued and continuously differentiable such that $f(x_0, y_0, z_0) = 0$. If $\vec{\nabla} f(x_0, y_0, z_0) \neq 0$, show that there is a differentiable surface, given parametrically by $(x(s,t), y(s,t), z(s,t))$ with $(x(0,0), y(0,0), z(0,0)) = (x_0, y_0, z_0)$, on which $f = 0$.

Proof:

Since $\nabla f(x_0, y_0, z_0) \neq 0$, at least one of the partial derivatives of f at (x_0, y_0, z_0) is nonzero. Without loss of generality, suppose $f_z(x_0, y_0, z_0) \neq 0$. Then the implicit function theorem implies there is an open subset $U \subset \mathbb{R}^2$ containing (x_0, y_0), an open subset $V \subset \mathbb{R}$ containing z_0, and a differentiable function $g : U \to V$ such that

$$f(x, y, g(x, y)) = f(x_0, y_0, z_0) = 0 \quad \forall (x, y) \in U.$$

For $(x, y) \in U$, define $s := x - x_0$ and $t = y - y_0$ so that $x(s, t)$...

Since

$$\vec{\nabla} f(x, y, z) = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix}$$

and $\vec{\nabla} f(x_0, y_0, z_0) \neq 0$, we may assume, without loss of generality, that $\partial f(x_0, y_0, z_0)/\partial z \neq 0$ and so $(\partial f(x_0, y_0, z_0)/\partial z)^{-1}$ exists. The implicit function theorem implies there is an open subset $U \subset \mathbb{R}^3$ with $(x_0, y_0, z_0) \in U$ and an open subset $W \subset \mathbb{R}^2$ with $(x_0, y_0) \in W$ such that for each $(x, y) \in W$ there is a unique z such that $(x, y, z) \in U$ and $f(x, y, z) = 0$. Due to this uniqueness, we may define $g : W \to \mathbb{R}$ by $z = g(x, y)$, which the implicit function theorem further tells us is differentiable. We have shown by the implicit function theorem that for $(x, y) \in W$ we have $f(x, y, g(x, y)) = 0$. Define $W_0 = \{(x - x_0, y - y_0) \mid (x, y) \in W\}$. We now parametrically define our surface mapping W_0 in $\mathbb{R}^3$ by $x(s, t) = s + x_0$, $y(s, t) = t + y_0$, and $z(s, t) = g(s + x_0, y + y_0)$. As noted above g is differentiable and $s + x_0$ and $y + y_0$ are also and so the composition $z(s, t)$ is differentiable. The functions $x(s, t)$ and $y(s, t)$ are also clearly differentiable. Then $(x(0, 0), y(0, 0), z(0, 0)) = (x_0, y_0, g(x_0, y_0)) = (x_0, y_0, z_0)$, as desired. Furthermore, through this construction we have $f(x(s, t), y(s, t), z(s, t)) = 0$, and we are done. $\square$

S10.10: Let $f(x, y)$ be the function defined by

$$f(x, y) = \frac{xy}{\sqrt{x^2 + y^2}}$$

when $(x, y) \neq (0, 0)$ and $f(0, 0) = 0$.

a) Compute the directional derivatives of $f(x, y)$ at $(0, 0)$ in all directions where they exist.

b) Is $f(x, y)$ differentiable at $(0, 0)$? Prove your answer.

Proof:

a) Let $h = (a, b) \in \mathbb{R}^2$ with $\|(a, b)\| = 1$. Then the directional derivative of f in the direction of h is given by

$$\lim_{t\to 0} \frac{f(th) - \cancelto{0}{f(0)}}{t} = \lim_{t\to 0} \frac{abt^2}{t\sqrt{(at)^2 + (bt)^2}} = \lim_{t\to 0} \frac{abt^2}{t|t|\sqrt{a^2 + b^2}} = \lim_{t\to 0} \frac{t}{|t|} \cdot ab.$$

If $a = 0$ or $b = 0$, then $\lim_{t\to 0} \frac{t}{|t|} \cdot ab = \lim_{t\to 0} \frac{t}{|t|} \cdot 0 = 0$. So, the limit exists along the x and y axes and is equal to zero there. The directional derivative does not exist in other directions because there limit as $t \to 0+$ is nonzero and equals the negative of the limit as $t \to 0-$.

b) We show f is not differentiable at $(0, 0)$ by way of contradiction. Suppose f is differentiable at $(0, 0)$. Then there is a linear map $L : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\lim_{(x,y)\to(0,0)} \frac{|f(x, y) - f(0, 0) - L(x, y)h|}{\|(x, y) - (0, 0)\|} = 0.$$

Since L exists, it is completely determined by the partial derivatives of f at $(0, 0)$. From the limit in a) with either $a = 0$ or $b = 0$ we see these partial derivatives are zero, and, thus, $L = 0$. So,

$$\frac{\left|f(x, y) - \cancelto{0}{f(0, 0)} - \cancelto{0}{L(x, y)h}\right|}{\|(x, y) - (0, 0)\|} = \frac{|f(x, y)|}{\|(x, y)\|} = \frac{xy}{\|(x, y)\|^2} = \frac{xy}{x^2 + y^2}.$$

This implies

$$\lim_{(x,x)\to(0,0)} \frac{\|f(x, y) - f(0, 0) - D(x, y)\|}{\|(x, y) - (0, 0)\|} = \lim_{(x,x)\to(0,0)} \frac{x^2}{2x^2} = \frac{1}{2} \neq 0,$$

a contradiction. $\square$

S10.11: Suppose $\sum_{n=1}^\infty |a_n| < \infty$. Let $\sigma : \mathbb{N} \to \mathbb{N}$ be one-to-one. The series $\sum_{n=1}^\infty a_{\sigma(n)}$ is called a "rearrangement" of $\sum_{n=1}^\infty a_n$. Prove that all rearrangements of $\sum_{n=1}^\infty a_n$ are convergent and have the same sum.

Proof:

Since the series $\sum_{n=1}^\infty |a_n|$ is convergent, it is Cauchy and so, given $\varepsilon > 0$, there is a $N \in \mathbb{N}$ such that $\forall\ \ell > k \geq N$, $\sum_{n=k}^{\ell} |a_n| < \varepsilon$. Thus,

$$\left| \sum_{n=k}^{\ell} a_n \right| \leq \sum_{n=k}^{\ell} |a_n| < \varepsilon, \tag{131}$$

and $\sum_{n=1}^\infty a_n$ converges. Let a denote the limit of this series and $\sigma : \mathbb{N} \to \mathbb{N}$ be a bijection. We seek to show

$$a = \sum_{n=1}^{\infty} a_{\sigma(n)}. \tag{132}$$

This will be done if, given $\varepsilon > 0$, we can find $N \in \mathbb{N}$ such that for all $k \geq N$,

$$\left| \sum_{n=1}^{k} a_{\sigma(n)} - a \right| < \varepsilon. \tag{133}$$

Since $\sum_{n=1}^\infty |a_n|$ is Cauchy, there is $N_1 \in \mathbb{N}$ such that $\sum_{n=N_1+1}^\infty |a_n| < \varepsilon/2$. Now pick $N \geq N_1$ such that $\{1,2,\ldots,N_1\} \subseteq \sigma(\{1,2,\ldots,N\})$. For $k \geq N$ we obtain

$$\begin{aligned}
\left|\sum_{n=1}^{k} a_{\sigma(n)} - a\right| &= \left|\sum_{n=1}^{k} a_{\sigma(n)} - \sum_{n=1}^{N_1} a_n - \sum_{n=N_1+1}^{\infty} a_n\right| \\
&= \left|\sum_{n=1, n\notin\sigma^{-1}(1,2,\ldots,N)}^{k} a_{\sigma(n)} - \sum_{n=N_1+1}^{\infty} a_n\right| \\
&\leq \left|\sum_{n=1, n\notin\sigma^{-1}(1,2,\ldots,N)}^{k} a_{\sigma(n)}\right| + \left|\sum_{n=N_1+1}^{\infty} a_n\right| \\
&\leq \sum_{n=1, \sigma(n)\notin\{1,2,\ldots,N_1\}}^{k} \left|a_{\sigma(n)}\right| + \sum_{n=N_1+1}^{\infty} |a_n| \\
&\leq 2 \sum_{n=N_1+1}^{\infty} |a_n| \\
&< 2 \cdot \varepsilon/2 \\
&= \varepsilon,
\end{aligned} \tag{134}$$

as desired. □

S10.12: Assume that $\{f_n\}$ is a sequence of nonnegative continuous functions on $[0,1]$ such that $\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x = 0$. Is it necessarily true that

a) There is a B such that $f_n(x) \leq B$ for $x \in [0,1]$ for all n?

b) There are points $x_0 \in [0,1]$ such that $\lim_{n\to\infty} f_n(x_0) = 0$?

Proof:

a) Consider the tent function of width $2/n^2$ and height n. For each n, $\int_0^1 f_n(x)\,\mathrm{d}x = 1/n$. To be more rigorous, define $f_1 = 0$ and for $n > 1$ define

$$f_n(x) = \begin{cases} n^3 x & \text{if } x \in [0, 1/n^2] \\ n - n^3(x - 1/n^2) & \text{if } x \in (1/n^2, 2/n) \\ 0 & \text{if } x \in [2/n, 1]. \end{cases}$$

Continuity of f_n is clear along the intervals $[0, 1/n^2)$, $(1/n^2, 2/n^2)$, and $(2/n, 1]$ since there f_n is linear. A simple check at $1/n^2$ and $2/n^2$ shows that the left and right hand limits of f_n agree there, which implies f_n is continuous. So, $f_n(1/n^2) = n$, which grows without bound. Hence the claim of an upper bound B is not necessarily true.

b) Let $\sigma : \mathbb{N} \to \mathbb{Q}$ be a bijection. Then for each $r \in \mathbb{Q}$, define $\tau : \mathbb{Q} \to [0,1]$ by

$$\tau(r) \begin{cases} r & \text{if } r \in [0,1] \\ |r| - \lfloor |r| \rfloor & \text{if } r \notin [0,1]. \end{cases}$$

So, for each $r \in [0,1) \cap \mathbb{Q}$, we can find infinitely many natural numbers that $(\tau \circ \sigma)$ maps to r, which implies for each $k \in \mathbb{N}$ we can find $n \in \mathbb{N}$ with $n > k$ such that $(\tau \circ \sigma)(n) = r$.

For each $n \in \mathbb{N}$, define $g_n : [0,1] \to \mathbb{R}$ for each $x \in [0,1]$ by

$$g_n(x) = \begin{cases} f_n(x + (\tau \circ \sigma)(n) - 1/n^2 + 1) & \text{if } x + a_n - 1/n^2 < 0 \\ f_n(x + (\tau \circ \sigma)(n) - 1/n^2) & \text{if } 0 \leq x + a_n - 1/n^2 \leq 1 \\ f_n(x + (\tau \circ \sigma)(n) - 1/n^2 - 1) & \text{if } x + a_n > 1. \end{cases}$$

Our function g_n is a slide of the tent function f_n in a) around the interval $[0,1]$ so that the peak of the tent occurs at the rational number $(\tau \circ \sigma)(n)$.

Now pick $x_0 \in [0,1]$ and let $n \in \mathbb{N}$ and $\varepsilon > 0$ be given. To show $\lim_{n\to\infty} f_n(x_0) \neq 0$, it suffices to find $k > n$ such that $f_k(x_0) \geq 1/2$. By the density of $\mathbb{Q}$, there exists $r \in [0,1) \cap \mathbb{Q}$ such that

$|x_0 - r| \le 1/2n^2$. And, by construction of $(\tau \circ \sigma)$, there exists $k > n$ such that $r = (\tau \circ \sigma)(k)$. This implies that the peak of g_k occurs at r, i.e., $g_k(r) = n$. Using the definition of g_k, it follows that

$$|g_k(x_0) - 0| \ge |g_k(r + 1/2n^2)| = n/2 \ge 1/2.$$

Thus, there does not exists $N \in \mathbb{N}$ such that $|g_n(x_0) - 0| \le 1/2$ for all $n \ge N$. This shows $\lim_{n\to\infty} g_n(x_0) \ne 0$. Since x_0 was chosen arbitrarily in $[0,1]$, this holds for each $x_0 \in [0,1]$. Hence the claim is false.

$\square$

F10.1: Let F be a closed subset of a metric space X with metric ρ.

a) Show that if $K \subseteq X$ is compact, then $K \cap F = \emptyset$ iff

$$\inf_{x \in K, y \in F} \rho(x, y) > 0.$$

b) Is the statement in a) true if K is only assumed to be closed, rather than compact? Give a proof if it is true, and a counterexample if it is false.

Proof:

a) By way of contradiction, suppose

$$\inf_{x \in K, y \in F} \rho(x, y) = 0.$$

Then for each $n \in \mathbb{N}$, there exists $x_n \in K$ and $y_n \in F$ such that $\rho(x_n, y_n) \leq 1/n$. Now, since K is compact, it is closed and complete. This implies that the sequence $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$ that converges to a limit $x \in K$. Let $\varepsilon > 0$ be given. Then, by the convergence of $\{x_{n_k}\}$, there is an integer $K \in \mathbb{N}$ such that $\rho(x_{n_k}, x) \leq \varepsilon/2$ whenever $k \geq K$. Also, by the Archimedean property of $\mathbb{R}$, there exists $N \in \mathbb{N}$ such that $1/N \leq \varepsilon/2$. Now define $M = \max\{N, K\}$. Then

$$\rho(x, y_{n_k}) \leq \rho(x, x_{n_k}) + \rho(x_{n_k}, y_{n_k}) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$$

whenever $n \geq M$. Hence $\{y_{n_k}\}$ converges to x. Since F is closed, it follows that $x \in F$. Then $x \in K \cap F$, which implies $K \cap F \neq \emptyset$, a contradiction. Thus, if $K \cap F = \emptyset$, then

$$\inf_{x \in K, y \in F} \rho(x, y) > 0.$$

Conversely, now suppose $K \cap F \neq \emptyset$. Then there exists $z \in K \cap F$. Then

$$\inf_{x \in K, y \in F} \rho(x, y) = \rho(z, z) = 0,$$

a contradiction. Hence if $\inf_{x \in K, y \in F} \rho(x, y) > 0$, then $K \cap F = \emptyset$.

b) No, the statement does not hold with these looser conditions. To verify this, we provide a counterexample. Let $X = \mathbb{R}^2$ with the standard metric. Then define $K = \{(n, 0) \mid n \in \mathbb{N}\}$ and $F = \{(n, 1/n) \mid n \in \mathbb{N}\}$. Then both K and F are closed with intersection $K \cap F = \emptyset$. However, for each $n \in \mathbb{N}$, there exists $(n, 0) \in K$ and $(n, 1/n) \in F$ such that

$\rho((n, 0), (n, 1/n)) = 1/n$. Thus,

$$\inf_{x \in K, y \in F} \rho(x, y) = 0.$$

□

F10.2: Suppose f is a bounded function on $[a, b]$.

a) Define "f is Riemann integrable on $[a, b]$".

b) Prove directly from the definition that if f is continuous, then f is Riemann integrable.

Proof:

a) Let $P = \{I_1, \ldots, I_n\}$ be a partition of $[a, b]$. Then lower and upper Riemann sums of f are, respectively, given by

$$L(f, P) = \sum_{k=1}^{n} \left(\inf_{x \in I_k} f(x) \right) |I_k| \quad \text{and} \quad U(f, P) = \sum_{k=1}^{n} \left(\sup_{x \in I_k} f(x) \right) |I_k|.$$

The function f is Riemman integrable on $[a, b]$ if

$$\inf_P U(f, P) = \sup_P L(f, P)$$

where the infimum and supremum are taken over all partitions P of the interval $[a, b]$.

b) To prove the claim, we show that, given any $\varepsilon > 0$,

$$\left| \inf_P U(f, P) - \sup_P L(f, P) \right| \leq \varepsilon.$$

Since $[a, b]$ is closed and bounded, it follows from Bolzano-Weierstrass that it is compact. So, f is uniformly continuous. Then there is a $\delta > 0$ such that, for $x, y \in [a, b]$, if $|x - y| \leq \delta$, then $|f(x) - f(y)| \leq \varepsilon/(b - a)$. By the Archimedean property of $\mathbb{R}$, there is a $N \in \mathbb{N}$ such that $(b - a)/N \leq \delta$. Then define a regular partition $P_N = \{I_k\}_{k=1}^{N}$ of $[a, b]$ so that the width of each subinterval is $(b - a)/N \leq \delta$. Then

$$\left| \sup_{x \in I_k} f(x) - \inf_{y \in I_k} f(y) \right| \leq \frac{\varepsilon}{b - a},$$

which implies

$$\sup_{x \in I_k} f(x) \leq \inf_{y \in I_k} f(y) + \frac{\varepsilon}{b - a}.$$

Consequently,

$$\begin{aligned} U(f,P) &= \sum_{k=1}^{N}\left(\sup_{x\in I_k} f(x)\right)|I_k| \\ &\le \sum_{k=1}^{N}\left(\inf_{x\in I_k} f(x)+\frac{\varepsilon}{b-a}\right)|I_k| \\ &= \sum_{k=1}^{N}\left(\inf_{x\in I_k} f(x)\right)|I_k| + \sum_{k=1}^{N}\frac{\varepsilon}{b-a}\cdot|I_k| \\ &= L(f,P)+\varepsilon. \end{aligned}$$

But, this implies

$$\inf_P U(f,P) \le U(f,P_N) \le L(f,P_N)+\varepsilon \le \sup_P L(f,P)+\varepsilon.$$

But, by definition, $L(f,P)\le U(f,P)$ for each P. Hence

$$\left|\inf_P U(f,P) - \sup_P L(f,P)\right| \le \varepsilon.$$

$\square$

F10.5: Prove or disprove the following two statements: For any two subsets S and S' of a vector space V,

a) $\text{span}(S) \cap \text{span}(S') = \text{span}(S \cap S')$.

b) $\text{span}(S) + \text{span}(S') = \text{span}(S \cup S')$.

Proof:

a) We disprove this claim by a counter example. Let $V = \mathbb{R}$, $S = \{0, 1\}$ and $S' = \{0, 2\}$. Then $\text{span}(S) = \mathbb{R} = \text{span}(S')$. So, $\text{span}(S) \cap \text{span}(S') = \mathbb{R} \cap \mathbb{R} = \mathbb{R}$. However, $\text{span}(S \cap S') = \text{span}(0) = \{0\}$ and so $\text{span}(S) \cap \text{span}(S') \neq \text{span}(S \cap S')$.

b) We prove the truth of this statement. Let $u + v \in \text{span}(S) + \text{span}(S')$. Since $S \subseteq S \cup S'$, $\text{span}(S) \subseteq \text{span}(S \cup S')$ and so $u \in \text{span}(S \cup S')$. Similarly, $v \in \text{span}(S \cup S')$. Since $\text{span}(S \cup S')$ is a vector space, it closed under vector addition and so $u + v \in \text{span}(S \cup S')$, which implies $\text{span}(S) + \text{span}(S') \subseteq \text{span}(S \cup S')$. Now let $w \in \text{span}(S \cup S')$. Then w is a linear combination of elements in $S \cup S'$, which implies we can write $w = w_1 + w_2$ where w_1 is a linear combination of elements in S and w_2 is a linear combination of elements in S'. But, then $w_1 \in \text{span}(S)$ and $w_2 \in \text{span}(S')$. Hence $w \in \text{span}(S) + \text{span}(S')$ and $\text{span}(S \cup S') \subseteq \text{span}(S) + \text{span}(S')$. The desired equality follows.

$\square$

F10.6: Let T be an invertible linear operator on a finite dimensional vector space V over a field F. Prove that there exists a polynomial f over F such that $T^{-1} = f(T)$.

Proof:

Let $f(t) = a_0 + a_1 t + \cdots + a_n t^n$ be the minimal polynomial of T where $a_0, \cdots, a_n \in F$. We claim $a_0 \neq 0$. For, if it did, then we would have $f(t) = t(a_1 + \cdots a_n t^{n-1})$, implying that either $T = 0$ or $a_1 + \cdots + a_n T^{n-1} = 0$, which both contradict the fact that $f(t)$ is the minimal polynomial of T. Hence $a_0 \neq 0$. Since $f(T) = 0$, this implies that, upon subtracting a_0 from each side and then dividing by $-a_0$ we get

$$I = -\frac{1}{a_0}\left(a_1 T + \cdots + a_n T^n\right).$$

Consequently,

$$T^{-1} = -\frac{1}{a_0}\left(a_1 T^{-1} T + \cdots + a_n T^{-1} T^n\right) = -\frac{1}{a_0}\left(a_1 I + \cdots + a_n T^{n-1}\right)$$

and have that T^{-1} is expressible as a polynomial of T, as desired. $\square$

F10.7: Let V and W be inner product spaces over $\mathbb{C}$ such that $\dim(V) \leq \dim(W) < \infty$. Prove that there is a linear transformation $T : V \to W$ satisfying $\langle T(v), T(v')\rangle_W = \langle v, v'\rangle_V$ for all $v, v' \in V$.

Proof:

Let $e_1, \ldots, e_m$ denote an orthonormal basis for V and $f_1, \ldots, f_n$ denote an orthonormal basis for W where $m \leq n$. (Note an orthonormal basis for V and W can be found from any basis of V and W, respectively, using the Gram-Schmidt procedure.) Let $v \in V$. Then there exists $\alpha_1, \ldots, \alpha_m \in \mathbb{C}$ such that $v = \sum_{j=1}^m \alpha_j e_j$. Then define the linear transformation $T : V \to W$ by

$$T(x) = T\left(\sum_{j=1}^m \alpha_j e_j\right) = \sum_{j=1}^m \alpha_j f_j. \tag{135}$$

We claim that this choice of T satisfies the desired relation. To see this, let $v' \in V$. Then there exists $\beta_1, \ldots, \beta_m \in \mathbb{C}$ such that $v' = \sum_{j=1}^m \beta_j e_j$. Since the e_j are orthonormal, $\langle e_j, e_k\rangle$ is 1 if $j = k$ and 0 when $j \neq k$. Hence

$$\langle v, v'\rangle = \langle \sum_{j=1}^m \alpha_j e_j, \sum_{j=1}^m \beta_j e_j\rangle = \sum_{j=1}^m \alpha_j \beta_j. \tag{136}$$

Similarly,

$$\langle T(v), T(v')\rangle = \langle \sum_{j=1}^m \alpha_j f_j, \sum_{j=1}^m \beta_j f_j\rangle = \sum_{j=1}^m \alpha_j \beta_j \tag{137}$$

and so the desired relation holds for this choice of T. $\square$

F10.8: Let W_1 and W_2 be subspaces of a finite dimensional inner product space V. Prove that $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$ where $W^\perp$ is the orthogonal complement of a subspace W of V.

Proof:

First let $w \in W_1^\perp$ and $v \in W_1 \cap W_2$. Then $v \in W_1$, which implies $\langle w, v \rangle = 0$ and so $w \in (W_1 \cap W_2)^\perp$, which implies $W_1^\perp \subseteq (W_1 \cap W_2)^\perp$. Similarly, $W_2^\perp \subseteq (W_1 \cap W_2)^\perp$. Since $(W_1 \cap W_2)^\perp$ is a vector space, it is closed under vector addition, which implies $W_1^\perp + W_2^\perp \subseteq (W_1 \cap W_2)^\perp$.

Now let $v \in (W_1 \cap W_2)^\perp$ and $e_1, \ldots, e_m$ be an orthonormal basis for $W_1^\perp$. Since $W_1^\perp \subseteq (W_1 \cap W_2)^\perp$, we can extend this to an orthonormal basis $e_1, \ldots, e_n$ of $(W_1 \cap W_2)^\perp$ where $n \geq m$. Then define $\hat{v}$ to be the projection of v into $W_1^\perp$, i.e.,

$$\hat{v} := \langle v, e_1 \rangle\, e_1 + \cdots + \langle v, e_m \rangle\, e_m.$$

Of course, $\hat{v} \in W_1^\perp$. Also, $v - \hat{v} \in W_1$ since $\langle v - \hat{v}, e_j \rangle = \langle v, e_j | - \langle v, e_j | = 0$ for each $j = 1, \ldots, m$. We now show $v - \hat{v} \in W_2^\perp$. So, let $y \in W_2$. Because $V = W_1 \oplus W_1^\perp$, there exists unique $y_1 \in W_1$ and $y_2 \in W_1^\perp$ such that $y = y_1 + y_2$. Then $y_1 \in (W_1 \cap W_2)$ and $v - \hat{v} \in (W_1 \cap W_2)^\perp$, which implies $\langle y_1, v - \hat{v} \rangle = 0$. Also, since $y_2 \in W_1^\perp$, $\langle y_2, v - \hat{v} \rangle = 0$. Thus, $\langle y, v - \hat{v} | = \langle y_1, v - \hat{v} | + \langle y_2, v - \hat{v} | = 0$ and so $(v - \hat{v}) \in W_2^\perp$. This shows that $v \in W_1^\perp + W_2^\perp$ and, thus, that $(W_1 \cap W_2)^\perp \subseteq W_1^\perp + W_2^\perp$. The desired equality follows directly. $\square$

F10.9: Consider the following iterative method

$$x_{k+1} = A^{-1}(Bx_k + c)$$

where $c = (1,1)^t$ and A and B are given by

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

a) Assume the iteration converges; to what vector x does the iteration converge?

b) Does this iteration converge for arbitrary initial vectors? Justify your answer.

Proof:

a) We claim that if the iteration converges to some x, then $x = (-1,-1)^t$. Indeed,

$$A^{-1}(Bx + c) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\left[\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right] = \begin{pmatrix} -1 \\ -1 \end{pmatrix} = x.$$

b) Observe that

$$\begin{aligned} x_{k+1} = A^{-1}(Bx_k + c) &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\left[\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} (x_k)_1 \\ (x_k)_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right] \\ &= \frac{1}{2}\begin{pmatrix} 2(x_k)_1 + (x_k)_2 + 1 \\ (x_k)_1 + 2(x_k)_2 + 1 \end{pmatrix} \end{aligned}$$

and so, if the entries of x_k are nonnegative, then so also will be those of x_{k+1}. Thus, if $x_0 = (0,0)$, then for each $k \in \mathbb{N}$ we have

$$\|x_k - x\| \geq \|(0,0) - (-1,-1)\| = \|(1,1)\| = \sqrt{2}.$$

Hence the sequence does not converge for all initial vectors.

□

F10.11: Find the function $g(x)$ which minimizes

$$\int_0^1 |f'(x)|^2 \, \mathrm{d}x$$

among smooth functions $f : [0,1] \to \mathbb{R}$ with $f(0) = f(1)$. Is the optimal solution $g(x)$ unique.

Proof:

Using the Cauchy Schwarz with $f'(x)$ and 1 to get

$$\int_0^1 f'(x)\ \mathrm{d}x \leq \sqrt{\int_0^1 f'(x)^2\ \mathrm{d}x \int_0^1 1^2\ \mathrm{d}x} = \sqrt{\int_0^1 f'(x)^2\ \mathrm{d}x}.$$

However, by the fundamental theorem of calculus, the left hand side of the above is equal to $f(1) - f(0) = 1$. So, squaring both sides gives $1 \leq \int_0^1 |f'(x)|^2\ \mathrm{d}x$. Furthermore, equality in the Cauchy Schwarz relation occurs precisely when f' and the constant function 1 are linearly dependent, i.e., iff f' is a scalar multiple of 1. Given the constraints $f(0) = 0$ and $f(1) = 1$, it follows that f' is constant iff $f'(x) = 1$ iff $f(x) = x$. So, $g(x) = x$ minimizes $\int_0^1 |f'(x)|^2\ \mathrm{d}x$. Moreover, this choice was unique due to the given constraints on $f(0)$ and $f(1)$. $\square$

F10.12: Define $D(t) = \{x^2 + y^2 \leq r^2(t)\} \subset \mathbb{R}^2$, where $r(t) : \mathbb{R} \to \mathbb{R}$ is continuously differentiable. For a given smooth nonnegative function $u(x,t) : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$, express the following quantity in terms of a surface integral:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{D(t)} u(x,t)\ \mathrm{d}x\right) - \int_{D(t)} u_t(x,t)\ \mathrm{d}x.$$

[You may use various theorems in Calculus without proof.]

Proof:

We make use of the Leibniz rule for differentiation under the integral sign. To do this, first we switch to polar coordinates so that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{D(t)} u(x,t)\ \mathrm{d}x\right) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_0^{2\pi}\int_0^{r(t)} u(\rho,\theta,t)\ \rho\mathrm{d}\rho\mathrm{d}\theta \\
&= \int_0^{2\pi}\left[\frac{\mathrm{d}}{\mathrm{d}t}\int_0^{r(t)} u(\rho,\theta,t)\ \rho\mathrm{d}\rho\right]\mathrm{d}\theta \\
&= \int_0^{2\pi}\left[\int_0^{r(t)} u_t(\rho,\theta,t)\ \rho\mathrm{d}\rho + u(r(t),\theta,t)\cdot r(t)\cdot\frac{\mathrm{d}r(t)}{\mathrm{d}t}\right]\mathrm{d}\theta \\
&= \int_0^{2\pi}\int_0^{r(t)} u_t(\rho,\theta,t)\ \rho\mathrm{d}\rho\mathrm{d}\theta + \int_0^{2\pi} u(r(t),\theta,t)\cdot r(t)\cdot\frac{\mathrm{d}r(t)}{\mathrm{d}t}\mathrm{d}\theta \\
&= \int_{D(t)} u_t(x,t)\ \mathrm{d}x + \int_0^{2\pi} u(r(t),\theta,t)\cdot r(t)\cdot\frac{\mathrm{d}r(t)}{\mathrm{d}t}\mathrm{d}\theta
\end{aligned}$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{D(t)} u(x,t)\ \mathrm{d}x\right) - \int_{D(t)} u_t(x,t)\ \mathrm{d}x = \boxed{r(t)\cdot\frac{\mathrm{d}r(t)}{\mathrm{d}t}\cdot\int_0^{2\pi} u(r(t),\theta,t)\ \mathrm{d}\theta.}$$

$\square$

2011

S11.1: Let A be a 3×3 matrix with complex entries. Consider tehs et of A that satisfy $\mathrm{tr}(A) = 4$, $\mathrm{tr}(A^2) = 6$, and $\mathrm{tr}(A^3) = 10$. For each similarity (i.e., conjugacy) class of such matrices, give one member in Jordan normal form. The following identity may be helpful: If $b_1 = a_1 + a_2 + a_3$, $b_2 = a_1^2 + a_2^2 + a_3^2$ and $b_3 = a_1^3 + a_2^3 + a_3^3$, then $6a_1a_2a_3 = b_1^3 + 2b_3 - 3b_1b_2$.

Proof:

Recall that the trace of a matrix is equal to the sum of its eigenvalues. Let $\lambda_1, \lambda_2, \lambda_3$ denote the eigenvalues of A. Further note that λ_i^2 is an eigenvalue of A^2 for each i since, for the corresponding eigenvector v_i of A, we have $A^2 v_i = A(\lambda_i v_i) = \lambda_i(Av_i) = \lambda_i^2 v_i$. Similarly, λ_i^3 gives the eigenvalues of A^3 for $i = 1, 2, 3$. So,

$$4 = \mathrm{tr}(A) = \lambda_1 + \lambda_2 + \lambda_3, \quad 6 = \mathrm{tr}(A^2) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \quad 10 = \mathrm{tr}(A^3) = \lambda_1^3 + \lambda_2^3 + \lambda_3^3.$$

Using the given identity, we also know

$$6\lambda_1\lambda_2\lambda_3 = 4^3 + 2(10) - 3(4)(6) = 12 \quad \Rightarrow \quad \lambda_1\lambda_2\lambda_3 = 2.$$

Let us suppose, without loss of generality, that $\lambda_1 = 1$. Then

$$4 = 1 + \lambda_2 + \lambda_3 = 1 + 2/\lambda_3 + \lambda_3 \quad \Rightarrow \quad \lambda_3^2 - 3\lambda_3 + 2 = 0 \quad \Rightarrow \quad (\lambda_3 - 2)(\lambda_3 - 1) = 0.$$

Again, without loss of generality, suppose $\lambda_3 = 2$. Then $\lambda_2 = 2/2 = 1$. Indeed, by plugging the values $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 2$ into the equations above, we see that this, in fact, the solution. So, depending on the minimal polynomial of A, we could have a Jordan block for the eigenvalue 1 of size 2 or two Jordan blocks of size 1. That is, we have two conjugacy classes, each consisting of matrices similar to one of the following two Jordan matrices:

$$J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{or} \quad J = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

□

S11.5: Let A be an $n \times n$ matrix with real entries, and let b be a $n \times 1$ column vector with real entries. Prove that there exists an $n \times 1$column vector solution x to the equation $Ax = b$ if and only if b is in the orthogonal complement of the kernel of the transpose of A.

Proof:
We begin with the following lemma:

Lemma: Let U be a subspace of an inner product space V. Then $U = (U^{\perp})^{\perp}$.
proof:
Let $u \in U$. Then $\langle u, v\rangle = 0$ for each $v \in U^{\perp}$, which implies u is orthogonal to every element of $U^{\perp}$ and, thus, $u \in (U^{\perp})^{\perp}$. Hence $U \subseteq (U^{\perp})^{\perp}$. Conversely, suppose $z \in (U^{\perp})^{\perp}$. Since we have $V = U \oplus U^{\perp}$, there exists unique $z_1 \in U$ and $z_2 \in U^{\perp}$ such that $z = z_1 + z_2$. Since $U \subseteq (U^{\perp})^{\perp}$, $z_1 \in (U^{\perp})^{\perp}$. Then, by closure of vector addition in a vector space, it follows that $z - z_1 \in (U^{\perp})^{\perp}$. This implies $z_2 = z - z_1 = (U^{\perp}) \cap (U^{\perp})^{\perp}$, which implies $z_2 = 0$. Hence $z = z_1 \in U$ and so $(U^{\perp})^{\perp} \subseteq U$. Therefore, $U = (U^{\perp})^{\perp}$. □

We must show that $b \in \text{im}(A)$ iff $b \in (\ker A^t)^{\perp}$, i.e., show the equality of these vector spaces. Then observe that

$$\begin{aligned} b \in \ker(A^t) &\Leftrightarrow A^t b = 0 \\ &\Leftrightarrow \langle A^t b, v\rangle = 0 \ \ \forall\, v \in \mathbb{R}^n \\ &\Leftrightarrow \langle b, Av\rangle = 0 \ \ \forall\, v \in \mathbb{R}^n \\ &\Leftrightarrow b \in (\text{im}(A))^{\perp}, \end{aligned}$$

where we have used the standard inner product for $\mathbb{R}^n$ so that $\langle A^t b, v\rangle = (A^t b)^t v = (b^t A)v = b^t(Av) = \langle b, Av\rangle$. So, we have shown $\ker(A^t) = (\text{im}(A))^{\perp}$. Now, using the above lemma, it follows that

$$(\ker(A^t))^{\perp} = ((\text{im}(A))^{\perp})^{\perp} = \text{im}(A),$$

completing the proof. □

S11.6: Let V and W be finite dimensional real inner product spaces, and let $A : V \to W$ be a linear transformation. Let $w \in W$. Show that the elements of $v \in V$ for which the norm $\|Av - w\|$ is minimal are exactly the solutions to the equations $A^*Ax = A^*w$.

Proof:
Follow the method shown in several previous problems to show that the minimizer of $\|Av - w\|$ is the projection $\hat{w}$ of w into $\text{im}(A)$ and that $w - \hat{w} \in (\text{im}(A))^\perp$. This implies

$$\begin{aligned}\langle w - \hat{w}, Av\rangle = 0 \ \ \forall\ v \in V &\Leftrightarrow \langle A^*(w - \hat{w}), v\rangle = 0 \ \ \forall\ v \in V \\ &\Leftrightarrow 0 = A^*(w - \hat{w}) = A^*w - A^*\hat{w} \\ &\Leftrightarrow Aw^* = A^*\hat{w} = A^*Ax\end{aligned}$$

where $x \in \text{im}(A)$ such that $Ax = \hat{w}$. This completes the proof. □

S11.7: Prove that there is a real number x such that $x^5 - 3x + 1 = 0$.

Proof:
Define $p : \mathbb{R} \to \mathbb{R}$ by $p(x) = x^5 - 3x + 1$. Since p is a polynomial it is continuous since sums and products of continuous functions are continuous. Then observe that $p(2) = 2^5 - 3(2) + 1 = 32 - 6 + 1 = 27 > 0$ and $p(1) = 1^5 - 3(1) + 1 = 1 - 3 + 1 = -1 < 0$. It follows from the Intermediate Value Theorem that there exists $x_0 \in (1, 2)$ such that $p(x_0) = 0$. □

S11.8: Give examples:

a) A function $f(x)$ on $[0, 1]$ which is not Riemann integrable, for which $|f(x)|$ is Riemann integrable.

b) Continuous functions f_n and f on $[0, 1]$ such that $f_n(t) \to f(t)$ for all $t \in [0, 1]$, but $\int_0^1 f_n(t)\ \mathrm{d}t$ does not converge to $\int_0^1 f(t)\ \mathrm{d}t$.

Solution:

a) Define

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

b) Define f_n to be the tent function of height n, width $2/n$, centered at $1/n$. Then $\lim\limits_{n\to\infty} f_n(x) = 0$ for each $x \in [0, 1]$ and $\lim\limits_{n\to\infty} \int_0^1 f_n(x)\ \mathrm{d}x = \lim\limits_{n\to\infty} 1 = 1 \neq 0 \int_0^1 0\ \mathrm{d}x = \int_0^1 \lim\limits_{n\to\infty} f_n(x)\ \mathrm{d}x$.

□

S11.9: Prove that if $f : [a, b] \to \mathbb{R}$ is continuous and $f(x) \geq 0$, then $\int_a^b f(x) = 0$ implies that $f = 0$.

Proof:

Suppose that $f \neq 0$ and $f(x) \geq 0$. Then there exists $x_0 \in [a, b]$ such that $f(x_0) > 0$. By the continuity of f, there exists $\delta > 0$ such that if $|x_0 - x| \leq \delta$, then $|f(x_0) - f(x)| \leq f(x_0)/2$. This implies $f(x) \geq f(x_0)/2$ whenever $x \in [x_0 - \delta, x_0 + \delta]$. Thus, breaking our integral into parts,

$$\begin{aligned}\int_a^b f(x) &= \int_a^{x_0-\delta} f(x) + \int_{x_0-\delta}^{x_0+\delta} f(x) + \int_{x_0+\delta}^{b} f(x)\\ &\geq \int_{x_0-\delta}^{x_0+\delta} f(x)\\ &\geq \int_{x_0-\delta}^{x_0+\delta} f(x_0)/2\\ &= 2\delta f(x_0)/2\\ &= \delta f(x_0)\\ &> 0.\end{aligned}$$

Thus, $f \neq 0$ implies $\int_a^b f(x) \neq 0$. The desired result follows through contraposition. □

S11.11: Give an example of $X \subseteq \mathbb{R}^2$ which is connected, but not path connected.

Proof:

Take $A = \{(x, y) \in \mathbb{R}^2 \mid (x, y) = (0, 0) \vee (y = \sin(1/x) \wedge x > 0)\}$. The graph of a function is connected and so $A - \{(0, 0)\}$ connected. Then since $(0, 0)$ is a limit point of $A - \{(0, 0)\}$, A is connected. By way of contradiction, suppose A is also pathwise connected. Then, by hypothesis, there is a continuous function $f : [0, 1] \to A$ such that $f(0) = (0, 0)$ and $f(1) = (1/\pi, 0)$. Pick any $\delta > 0$ with $\delta \leq 1/\pi$. Using the Archimedean property of $\mathbb{R}$, we can pick $n \in \mathbb{N}$ so that

$$x_n := \frac{1}{(4n+1)\frac{\pi}{2}} \leq \delta,$$

and note $\sin(1/x_n) = \sin((4n + 1)\pi/2) = 1$. Then

$$|0 - x_n| \leq \delta \quad \text{and} \quad \|f(0) - f(x_n)\| = \sqrt{(x_n - 0)^2 + (\sin(x_n) - 0)^2} = \sqrt{x_n^2 + 1^2} \geq 1.$$

Hence no $\delta > 0$ exists such that $\|f(0) - f(x)\| \leq 1/2$ whenever $x \in [0, \delta)$, i.e., f is not continuous. This contradiction implies that A is not pathwise connected. □

S11.12: Given a metric space M, and a constant $0 < r < 1$, a continuous function $T : M \to M$ is said to be an r-contraction if it is a continuous map and $d(T(x), T(y)) < rd(x, y)$ for all x and y. A well-known fixed piont theorem states that if M is complete and T is an r-contraction, then it must have a unique fixed point (don't prove this). This result is often used to prove the existence of solutions of differential equations with initial conditions.

a) Illustrate this technique for the (trivial) case $f'(t) = f(t)$ and $f(0) = 1$ by letting M be the space of continuous functions $C([0, 1])$ for $c \in (0, 1)$ with the uniform distance $d(f, g) = \sup\{|f(t) - g(t)| \mid t \in [0, c]\}$, and defining $(Tf)(x) = 1 + \int_0^x f(t)\ \mathrm{d}t$. Carefully explain your steps.

b) What approximations do you obtain from the sequence $T(0), T^2(0), T^3(0), \ldots$?

Proof:

a) Confer older problem.

$$\|Tf - Tg\|_\infty = \left[\int_0^x f(t) - g(t)\ \mathrm{d}t\right]_\infty \leq x\|f - g\|_\infty \leq c\|f - g\|_\infty. \tag{138}$$

b) First, $T(0) = 1$, $T^2(0) = T(1) = 1 + x$, $T^3(0) = T^2(1) = T(1 + x) = 1 + x + x^2/2$. In general, we see

$$f(x) = \lim_{n\to\infty} T^n(0) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!} = e^x.$$

□

F11.1: Let (X, d) be a compact metric space and let $f : X \to X$ be a map satisfying

$$d(f(x), d(f(y)) < d(x, y), \quad \forall\ x, y \in X \text{ with } x \neq y.$$

Prove that there is a unique point $x \in X$ so that $f(x) = x$.

Proof:

Let $h = d(f(x), x)$. Then $h : X \to \mathbb{R}$ is continuous since it is the composition of continuous functions. Because X is compact, h achieves a minimum. Let $z \in X$ be such that $h(z)$ is the minimal value of h. If $f(z) \neq z$, then

$$d(f(f(z)), f(z)) < d(f(z), z) = h(z).$$

But, this implies $h(f(z)) < h(z)$, which contradicts our choice of z. Hence there must exists $z \in X$ such that $f(z) = z$.

All that remains is to show uniqueness. So suppose $f(x) = x$ and $f(y) = y$ with $x \neq y$. Then $d(x, y) > 0$ and

$$d(x, y) = d(f(x), f(y)) < d(x, y),$$

which implies $0 < d(x, y) < d(x, y)$, a contradiction. Hence $d(x, y)$ must be zero and so $x = y$. Hence the fixed point must be unique. $\square$

2012

W12.04: For a sequence $\{a_n\}$ of non-negative numbers, let $s_n := \sum_{k=1}^{n} a_k$ and suppose s_n tends to a number $s \in \mathbb{R}$ in Cesaro sense:

$$\lim_{n\to\infty} \frac{s_1 + \cdots + s_n}{n} = s. \tag{139}$$

Show that $\sum_{k=1}^{\infty} a_k$ exists and equals s.

Proof:

The sequence $\{s_n\}_{n=1}^{\infty}$ is monotonically increasing (because each $a_k \geq 0$ for each k). This implies

$$\frac{s_1 + \cdots + s_n}{n} \tag{140}$$

is monotonically increasing because this is simply the arithmetic average of the s_k. Consequently, if $s = 0$, then $a_k = 0$ for each k and the result follows directly. Now suppose $s > 0$. And, by way of contradiction, suppose the limit $\lim\limits_{n\to\infty} s_n$ does not exist. Since the $\{s_n\}$ is monotonically increasing, this implies we can find $N \in \mathbb{Z}^+$ such that $s_N \geq 2s$. For $n > N$ we obtain

$$\begin{aligned} \frac{s_1 + \cdots + s_n}{n} &= \frac{s_1 + \cdots + s_{N-1}}{n} + \frac{s_N + \cdots + s_n}{n} \\ &\geq \frac{s_N + \cdots + s_n}{n} \\ &\geq \frac{(n - N + 1)}{n} \cdot 2s. \end{aligned} \tag{141}$$

But, taking the limit as $n \longrightarrow \infty$, we see

$$s = \lim_{n\to\infty} \frac{s_1 + \cdots + s_n}{n} \geq \lim_{n\to\infty} \frac{(n - N + 1)}{n} \cdot 2s = 2s \cdot \lim_{n\to\infty} 1 - \frac{N+1}{n} = 2s \cdot 1 = 2s, \tag{142}$$

which gives $s = 2s$, a contradiction. Thence the assumption was false and we conclude $\lim\limits_{n\to\infty} s_n$ exists.

Set $\ell := \lim\limits_{n\to\infty} s_n$ and let $\varepsilon > 0$ be given. We shall show $\ell = s$. By the convergence of $\{s_n\}$, there is a $N_1 \in \mathbb{Z}^+$ such that

$$|s_n - \ell| < \frac{\varepsilon}{2} \quad \forall\ n \geq N_1. \tag{143}$$

And, by the Archimedean property of $\mathbb{R}$ we can pick $N_2 \in \mathbb{Z}^+$ greater than N_1 such that

$$\frac{|s_1 - \ell| + \cdots + |s_{N_1} - \ell|}{n} < \frac{\varepsilon}{2} \quad \forall\ n \geq N_2, \tag{144}$$

(n.b. the numerator on the left hand side is a constant). Using the triangle inequality, for $n \geq N_2$

we discover

$$\begin{aligned}\left|\frac{s_1+\cdots+s_n}{n}-\ell\right| &\le \frac{|s_1-\ell|+\cdots+|s_n-\ell|}{n} \\ &= \frac{|s_1-\ell|+\cdots+|s_{N_1}-\ell|}{n}+\frac{|s_{N_1+1}-\ell|+\cdots+|s_n-\ell|}{n} \\ &\le \frac{\varepsilon}{2}+\frac{n-N_1}{n}\cdot\frac{\varepsilon}{2} \\ &\le \varepsilon.\end{aligned} \tag{145}$$

Thence

$$s=\lim_{n\to\infty}\frac{s_1+\cdots+s_n}{n}=\ell, \tag{146}$$

as desired. Consequently, $\lim_{n\to\infty} s_n = s$, and we are done. □

W12.05: Prove there is a unique continuous function $y : [0,1] \to \mathbb{R}$ solving the equation

$$y(x) = e^x + \frac{y(x^2)}{2} \quad \forall\, x \in [0,1]. \tag{147}$$

Proof:
This problem is a direct application of the Banach Fixed Point theorem. Let X be the metric space with set $C[0,1]$ and the sup norm $\|\cdot\|$ for a metric. We claim X is complete (and verify this last). Then define the mapping $T : X \to X$ for $f \in X$ by

$$T(f)(x) = e^x + \frac{y(x^2)}{2} \quad \forall\, x \in [0,1]. \tag{148}$$

Let $f, g \in X$ and define $\phi : [0,1] \to \mathbb{R}$ by $\phi(x) = x^2$. Note $\phi(0) = 0$, $\phi(1) = 1$ and $\phi'(x) = 2x$ for $x \in (0,1)$. Thus, $\phi([0,1]) = [0,1]$, from which we deduce

$$\|f \circ \phi\| = \sup_{x\in[0,1]} |(f \circ \phi)(x)| = \sup_{x\in[0,1]} |f(\phi(x))| = \sup_{x\in[0,1]} |f(x)| = \|f\|. \tag{149}$$

Then

$$\begin{aligned} \|T(f) - T(g)\| &= \left\| \left(e^x + \frac{f(x^2)}{2}\right) - \left(e^x + \frac{g(x^2)}{2}\right) \right\| \\ &= \frac{1}{2}\|(f\circ\phi) - (g \circ \phi)\| \\ &= \frac{1}{2}\|(f-g)\circ\phi\| \\ &= \frac{1}{2}\|f-g\|. \end{aligned} \tag{150}$$

This shows T is Lipschitz with Lipschitz constant $L = 1/2$ (and so T is a contraction). Now let $f_0 \in X$ be arbitrary and define the sequence $\{f_n\}_{n=0}^\infty$ by $f_{n+1} = T(f_n)$ for $n \geq 0$. This gives

$$\|f_{n+1} - f_n\| = \|T(f_n) - T(f_{n-1})\| = L\|f_n - f_{n-1}\| = \cdots = L^n\|f_1 - f_0\|. \tag{151}$$

Consequently, using the triangle inequality for $n > m$ we have

$$\begin{aligned}\|f_n - f_m\| &\leq \|f_n - f_{n-1}\| + \cdots + \|f_{m+1} - f_m\| \\ &= L^n\|f_1 - f_0\| + \cdots + L^m\|f_1 - f_0\| \\ &= \|f_1 - f_0\|L^m \cdot \sum_{k=0}^{n-m} L^k \\ &\leq \|f_1 - f_0\|L^m \cdot \sum_{k=0}^{\infty} L^k \\ &\leq \frac{\|f_1 - f_0\|}{1-L} \cdot L^m.\end{aligned} \tag{152}$$

But, taking the limit as $m \longrightarrow \infty$ on the right hand side,

$$\lim_{m\to\infty} \frac{\|f_1 - f_0\|}{1-L} \cdot L^m = \frac{\|f_1 - f_0\|}{1-L} \cdot \lim_{m\to\infty} L^m = \frac{\|f_1 - f_0\|}{1-L} \cdot 0 = 0. \tag{153}$$

This shows the sequence $\{f_n\}$ is Cauchy. Because X is complete, there is $f \in X$ such that $f_n \longrightarrow f$. Whence

$$f = \lim_{n\to\infty} f_n = \lim_{n\to\infty} f_{n+1} = \lim_{n\to\infty} T(f_n) = T\left(\lim_{n\to\infty} f_n\right) = T(f) \tag{154}$$

where we are able to bring the limit into the argument of T since the mapping is Lipschitz continuous (as shown earlier). Thus, there exists a fixed point of T so that

$$f(x) = T(f)(x) = e^x + \frac{f(x^2)}{2} \quad \forall\, x \in [0,1]. \tag{155}$$

We further claim f is unique. Let $g \in X$ also be a fixed point of T. Then

$$\|f - g\| = \|T(f) - T(g)\| = \frac{1}{2}\|f - g\|, \tag{156}$$

which is only possible if $\|f - g\| = 0$, implying $f = g$. Hence the fixed point f is unique.

All that remains is to show X is complete. Let $\{f_n\} \subset X$ be a Cauchy sequence. Let $\varepsilon > 0$ be given. Then there is $N_1 \in \mathbb{Z}^+$ such that

$$\|f_n - f_m\| < \varepsilon \quad \forall\, n, m \geq N_1. \tag{157}$$

Let $x \in [0,1]$. Then we claim the sequence $\{f_n(x)\} \subset \mathbb{R}$ converges. By definition of the sup norm,

$$|f_n(x) - f_m(x)| \leq \sup_{z\in[0,1]} |f_n(z) - f_m(z)| = \|f_n - f_m\| < \varepsilon \quad \forall\, n, m \geq N_1. \tag{158}$$

So, $\{f_n(x)\}$ is Cauchy. Because $\mathbb{R}$ is complete, $\{f_n(x)\}$ converges. This implies we may define a function f to be the point-wise limit $f(x) := \lim_{n\to\infty} f_n(x)$. And, since the uniform convergence of continuous functions converges to a continuous function, we conclude $f \in C[0,1]$. Thus, X is complete. $\square$

W12.10: Let $A \in M_n(\mathbb{C})$. State and prove under which conditions on A, the following identity holds:

$$\det(e^A) = \exp(\operatorname{tr}(A)). \tag{159}$$

Here the matrix exponentiation is defined via the Taylor series

$$e^A = 1 + A + A^2/2! + A^3/3! + \cdots . \tag{160}$$

You can assume it is known that this sum converges (entry-wise) for all complex matrices A.

Proof:
We claim the identity holds for all $A \in M_n(\mathbb{C})$, which we verify as follows. Let J be the Jordan canonical form of A. Then there is invertible P such that $A = PJP^{-1}$. We claim $A^k = PJ^kP^{-1}$ for each $k \in \mathbb{Z}^+$. This is given for $k = 1$. Now pick any $k \in \mathbb{Z}^+$ and suppose the claim holds. Then

$$A^{k+1} = AA^k = (PJP^{-1})(PJ^kP^{-1}) = PJ(P^{-1}P)J^kP^{-1} = PJIJ^kP^{-1} = PJ^{k+1}P^{-1}, \tag{161}$$

and we have closed the induction. The principle of induction implies the claim. So, for any polynomial $f_n(x) = a_0 + a_1x + \cdots + a_nx^n$ of degree n, we have

$$\begin{aligned} f_n(A) &= a_0I + a_1A + \cdots + a_nA^n \\ &= a_0PP^{-1} + a_1PJP^{-1} + \cdots + a_nPJ^nP^{-1} \\ &= P\left(a_0I + a_1J + \cdots + J^n\right)P^{-1} \\ &= Pf_n(J)P^{-1}. \end{aligned} \tag{162}$$

Thence

$$\begin{aligned} e^A = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{A^k}{k!} &= \lim_{n\to\infty}\sum_{k=0}^{n}P\frac{J^k}{k!}P^{-1} \\ &= \lim_{n\to\infty}P\left(\sum_{k=0}^{n}\frac{J^k}{k!}\right)P^{-1} \\ &= P\left(\lim_{n\to\infty}\sum_{k=0}^{n}\frac{J^k}{k!}\right)P^{-1} \\ &= Pe^JP^{-1} \end{aligned} \tag{163}$$

where we have used the fact P is linear (and thus continuous) to move the limit inside the parentheses.

We now show the diagonal entries of e^J are e^{λ_i} for $i = 1, \ldots, n$ where the λ_i are the eigenvalues of A. Indeed, the diagonal entries of J are $\lambda_1, \ldots, \lambda_n$. We claim $(J^k)_{ii} = (J_{ii})^k$, and show this by induction. The case for $k=1$ is trivial. Suppose $k \in \mathbb{Z}^+$ and $(J^k)_{ii} = (J_{ii})^k$. Then

$$(J^{k+1})_{ii} = (JJ^k)_{ii} = \sum_{j=1}^{n} J_{ij}(J^k)_{ji}. \tag{164}$$

Since J is upper triangular, so also is J^k. This implies $J_{ij} = 0$ for $i > j$ and $(J^k)_{ji} = 0$ for $j > i$. Thence there product is zero whenever $i \neq j$ and so

$$(J^{k+1})_{ii} = J_{ii}(J^k)_{ii} == J_{ii}(J_{ii})^k = (J_{ii})^{k+1}, \tag{165}$$

as desired. For a polynomial f_n as above, this implies

$$(f_n(J))_{ii} = (a_0 I + a_1 J + \cdots + a_n J^n)_{ii} = a_0 I_{ii} + a_1 J_{ii} + \cdots + a_n (J_{ii})^n = f_n(J_{ii}). \tag{166}$$

Thus,

$$(e^J)_{ii} = \left(\lim_{n\to\infty} \sum_{k=0}^{n} \frac{J^k}{k!} \right)_{ii} = \lim_{n\to\infty} \left(\sum_{k=0}^{n} \frac{J^k}{k!} \right)_{ii} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{(J_{ii})^k}{k!} = e^{J_{ii}} = e^{\lambda_i}. \tag{167}$$

Finally, using properties of determinants, we see

$$\det(e^A) = \det(Pe^J P^{-1}) = \det(P)\det(e^J)\det(P^{-1}) = \det(P)\det(e^J)\cdot\frac{1}{\det(P)} = \det(e^J) = \prod_{j=1}^{n} e^{\lambda_i} = e^{\operatorname{tr}(J)}. \tag{168}$$

All that remains is to show $\operatorname{tr}(J) = \operatorname{tr}(A)$. To show this, we verify commutativity of the trace. Let $M, N \in M_n(\mathbb{C})$. Then

$$\operatorname{tr}(MN) = \sum_{i=1}^{n} (MN)_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} M_{ij}N_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{n} M_{ij}N_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{n} N_{ji}M_{ij} = \sum_{j=1}^{n} (NM)_{jj} = \operatorname{tr}(NM). \tag{169}$$

Thus,

$$\operatorname{tr}(A) = \operatorname{tr}(PJP^{-1}) = \operatorname{tr}((JP^{-1})P) = \operatorname{tr}(J), \tag{170}$$

as desired. This completes the proof. □

S12.5: Prove that there is a unique continuous function $y : [0, 1] \to \mathbb{R}$ solving the equation

$$y(x) = e^x + \frac{y(x^2)}{2}, \quad \forall\, x \in [0, 1].$$

Proof:
First note $C[0, 1]$ is complete, which follows from the fact that $C(K)$ is complete for any compact set K. Now define the operator $T : C[0, 1] \to C[0, 1]$ for each $f \in C[0, 1]$ by

$$(Tf)(x) = e^x + \frac{f(x^2)}{2}.$$

We note that e^x is continuous and $f(x^2)/2$ is a composition of continuous functions since x^2 is continuous. And, sums of continuous functions are continuous. So, (Tf) is, in fact, continuous. Then

$$\begin{aligned}
\|Tf - Tg\|_\infty &= \sup\{|Tf(x) - Tg(x)| \ \mid \ x \in [0, 1]\} && \text{Definition of } \|\cdot\|_\infty \\
&= \sup\{|(f(x^2) - g(x^2)/2| \ \mid \ x \in [0, 1]\} && \text{Definition of } T \\
&= \frac{1}{2} \cdot \sup\{|f(x^2) - g(x^2)| \ \mid \ x \in [0, 1]\} && \text{Factor out } 1/2 \\
&= \frac{1}{2} \cdot \sup\{|f(u) - g(u)| \ \mid \ u \in [0, 1]\} && \text{Factor out } 1/2 \\
&= \frac{1}{2} \cdot \|f - g\|_\infty,
\end{aligned}$$

where we let $u = x^2$ and note $u \leq 1 \cdot x \leq 1$ and $u \geq 0 \cdot x = 0$. Hence T is a contraction mapping with Lipschitz constant $L = 1/2$. Then, by the Banach Fixed Point theorem, there exists a unique $y \in C[0, 1]$ such that $y(x) = (Ty)(x) = e^x + y(x^2)/2$. $\square$

S12.9: Prove that if $f(x)$ is a continuous function on $[a, b]$ and $f(x) \geq 0$, then $\int_a^b f(x) = 0$ implies that $f = 0$.

Proof:

Suppose f is continuous on $[a, b]$ and $f(x) \geq 0$. We shall proceed to prove the claim by proving its contrapositive. So, suppose $f \neq 0$. Then there is an $x^* \in [a, b]$ at which $f(x^*) > 0$. By continuity of f, there exists $\delta > 0$ with $\delta \leq \min\{|x^* - a|, |x^* - b\}$ such that, for $x \in [a, b]$, if $|x - x^*| < \delta$, then $|f(x^*) - f(x)| < f(x^*)/2$. This implies $f(x) > 0$ whenever $|x - x^*| < \delta$. Then, by linearity of the integral,

$$\begin{aligned}\int_a^b f(x)\ \mathrm{d}x &= \int_a^{x^*-\delta} f(x) + \int_{x^*-\delta}^{x^*+\delta} f(x)\ \mathrm{d}x + \int_{x^*+\delta}^{b} f(x)\ \mathrm{d}x \\ &= 0 + \int_{x^*-\delta}^{x^*+\delta} f(x)\ \mathrm{d}x + 0 \geq \int_{x^*-\delta}^{x^*+\delta} f(x^*)/2\ \mathrm{d}x \\ &> \delta f(x^*) \\ &> 0.\end{aligned}$$

Hence $\int_a^b f(x)\ \mathrm{d}x \neq 0$. The desired claim then follows. $\square$

S12.10:

F12.1: Let $\{b_n\}_{n=1}^{\infty}$ be a sequence of real numbers with bounded partial sums, i.e., there is $M < \infty$ such that for all N, $|\sum_{n=1}^{N}| \leq M$, and let $\{a_n\}_{n=1}^{\infty}$ be a sequence of positive numbers decreasing to 0. Prove the series $\sum a_n b_n$ converges.

Proof:

For each $N \in \mathbb{N}$, define the N-th partial sum by

$$S_N = \sum_{n=1}^{N} a_n b_n$$

and the sequence $\{B_n\}$ by

$$B_N = \sum_{n=1}^{N} b_n.$$

Let $\varepsilon > 0$ be given. To prove convergence of the partial sums $\{S_N\}$, we must find an integer N^* such that whenever $M, N \geq N_1$, $|S_M - S_N| < \varepsilon$. Then observe that we can rearrange the finite sum S_N to write

$$S_N = a_N B_N + \sum_{n=1}^{N-1} B_n(a_n - a_{n+1}), \tag{171}$$

which implies for $M > N$ that

$$|S_M - S_N| = \left| a_M B_M - a_N B_N + \sum_{n=N}^{M-1} B_n(a_n - a_{n+1}) \right|. \tag{172}$$

Since the $\{a_n\}$ converge to 0, we can pick $N_1 \in \mathbb{N}$ such that $|a_n - 0| < \varepsilon/4M$ whenever $n \geq N_1$. Similarly, since $\{a_n\}$ is Cauchy, we can pick $N_2 \in \mathbb{N}$ such that $|a_j - a_k| < \varepsilon/2M$ whenever $j, k \geq N_2$. Then define $N^* = \max\{N_1, N_2\}$. Using the triangle inequality and bound on B_N, this implies

$$\begin{aligned}
|S_M - S_N| &\leq |a_M B_M - a_N B_N| + \sum_{n=N}^{M-1} |B_n|\,|a_n - a_{n+1}| \\
&\leq a_N \left| \frac{a_M}{a_N} B_M - B_N \right| + \sum_{n=N}^{M-1} M\,|a_n - a_{n+1}| \\
&\leq a_N \left(|B_M| + |B_N|\right) + M\,|a_N - a_{M-1}| \\
&\leq 2M a_N + M|a_N - a_{M-1}| \\
&\leq \varepsilon/2 + \varepsilon/2 \\
&= \varepsilon
\end{aligned} \tag{173}$$

where we have used the fact that $\{a_n\}$ is decreasing to assert $a_M/a_N \leq 1$. Hence the sequence of partial of partial sums $\{S_N\}$ is Cauchy and converges in $\mathbb{R}$. $\square$

F12.5: A subset E of a metric space X is a G_δ set if $E = \cap_{n=1}^\infty G_n$ where each G_n is open in X. Prove that $\mathbb{Q}$ is not a G_δ subset of $\mathbb{R}$.

Proof:

By way of contradiction, suppose $\mathbb{Q} = \cap_{n=1}^\infty G_n$ with each G_n open in X. Note each G_n is dense since $\mathbb{Q} \subseteq G_n$ and $\mathbb{Q}$ is dense. Since $\mathbb{Q}$ is countable, let $\{a_n\}_{n=1}^\infty$ enumerate $\mathbb{Q}$. Then let $D_n = \mathbb{R} - \{a_n\}$. Then D_n is dense open. By the Baire category theorem,

$$\underbrace{\left(\cap_{n=1}^\infty G_n\right)}_{\mathbb{Q}} \bigcap \underbrace{\left(\cap_{n=1}^\infty D_n\right)}_{\mathbb{R}-\mathbb{Q}} \neq \emptyset, \tag{174}$$

which gives our contradiction. $\square$

F12.9: Let A be an $m \times n$ real matrix with $m \geq n$. Let $b \in \mathbb{R}^m$. Let M be the set of vectors $x \in \mathbb{R}^n$ which minimize $\|Ax - b\|$. Show that $M = x_0 + N$ where N is the kernel of A and x_0 is an element of M.

Proof:

Let $\hat{b}$ denote the projection of b into the image of A, i.e., define

$$\hat{b} = \langle b, e_1 \rangle e_1 + \cdots + \langle b, e_k \rangle e_k$$

where $e_1, \ldots, e_k$ with $k \leq m$ denotes an orthonormal basis for im A, the vectors $e_1, \ldots, e_m$ denotes the standard orthonormal basis for $\mathbb{R}^m$, and we use the usual the scalar by product $\langle v_1, v_2 \rangle = v_1^T v_2$ for $v_1, v_2 \in \mathbb{R}^m$. By definition, $\hat{b} \in$ im A. And, $(b - \hat{b}) \in (\text{im } A)^\perp$ since it is orthogonal to every vector in im A. This follows because $\langle b - \hat{b}, e_j \rangle = \langle b, e_j \rangle - \langle b, e_j \rangle = 0$ for each $j = 1, \ldots, k$. Using this, we can apply the Pythagorean Theorem to see that for any $z \in$ im A we have

$$\|b - \hat{b}\|^2 \leq \|b - \hat{b}\|^2 + \|\hat{b} - z\|^2 = \|(b - \hat{b}) + (\hat{b} - z)\|^2 = \|b - z\|^2 \quad \Rightarrow \quad \|b - \hat{b}\| \leq \|b - z\|.$$

Thus, $\hat{b}$ is the closest point in im A to b. Moreover, because $\hat{b} \in$ im A, there exists $x_0 \in \mathbb{R}^n$ such that $Ax_0 = \hat{b}$. Thus, $x_0 \in M$. So, not only have we shown that M is nonempty, but also that for each $x \in M$, $Ax = \hat{b}$. If we also have $y_0 \in M$, then $Ax_0 = \hat{b} = Ay_0$ and so $0 = Ay_0 - Ax_0 = A(y_0 - x_0)$, which implies $(y_0 - x_0) \in N$. That is, $x_0 + (y_0 - x_0) \in x_0 + N$, which implies $M \subseteq x_0 + N$. Now, pick any $z \in N$. Then $A(x_0 + z) = Ax_0 + Az = \hat{b} + 0 = \hat{b}$ and so $(x_0 + z) \in M$. Hence $M = x_0 + N$. $\square$

2013

S13.4: Denote by h_n the n-th harmonic number

$$h_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}.$$

Prove that there is a limit

$$\gamma = \lim_{n\to\infty} (h_n - \ln(n)).$$

Proof:
By Taylor Theorem, for $x \in [0,1]$ there exists $c \in (0,1)$ such that

$$e^x = e^0 + \frac{\mathrm{d}}{\mathrm{d}x}\left[e^x\right]_{x=0} \cdot x + \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left[e^x\right]_{x=c} \cdot \frac{x^2}{2} = 1 + x + \frac{e^c}{2} \cdot x^2.$$

This implies $e^x \geq 1 + x \;\Rightarrow\; x = \ln(e^x) \geq \ln(1+x)$, which follows from the fact $\ln(x)$ is increasing, i.e., $\ln'(x) = 1/x > 0 \;\forall\; x > 0$. So, for each $n \in \mathbb{N}$, $1/n \leq 1$ and so

$$-\ln(1 + 1/n) = \ln(n/(n+1)) = \ln(1 - 1/(n+1)) < -1/(n+1).$$

Now, noting

$$\ln(n) = \sum_{k=1}^{n-1} \ln(k+1) - \ln(k) = \sum_{k=1}^{n-1} \ln(1 + 1/k),$$

we see

$$h_n - \ln(n) = \left(\frac{1}{k} - \sum_{k=1}^{n-1}\frac{1}{k}\right) - \sum_{k=1}^{n-1} \ln\left(1 + \frac{1}{k}\right) = \frac{1}{n} + \sum_{k=1}^{n-1}\frac{1}{k} - \ln\left(1 + \frac{1}{k}\right).$$

Now note $\{h_n - \ln(n)\}_{n=1}^{\infty}$ is increasing since $1/n - \ln(1 + 1/n) \geq 0$ for each $n \in \mathbb{N}$. Moreover, using the above,

$$h_n - \ln(n) \leq \frac{1}{n} + \sum_{k=1}^{n-1}\frac{1}{k} - \frac{1}{k+1} = \frac{1}{n}\sum_{k=1}^{n-1}\frac{1}{k^2+k} < \frac{1}{n}\sum_{k=1}^{n-1}\frac{1}{k^2}.$$

Then, by the integral test, this series converges since

$$\lim_{n\to\infty}\sum_{k=1}^{n-1}\frac{1}{k^2} \leq \lim_{n\to\infty}\int_1^n \frac{\mathrm{d}x}{x^2} = \lim_{n\to\infty} -\frac{1}{x}\bigg|_1^n = \lim_{n\to\infty} 1 - \frac{1}{n} = 1.$$

And, $1/n \to 0$ as $n \to \infty$, which implies $\{h_n - \ln(n)\}_{n=1}^{\infty}$ is also bounded above. Then, by the Monotone Convergence Theorem, $\lim\limits_{n\to\infty} h_n - \ln(n)$ exists. $\square$

S13.11: Define the Fibonacci sequence F_n by $F_0 = 0$, $F_1, = 1$, and recursively by $F_n = F_{n-1} + F_{n-2}$ for $n = 2, 3, 4, \ldots$

a) Show that the limit as $n \to \infty$ of F_n/F_{n-1} exists and find its value.

b) Prove that $F_{2n+1}F_{2n-1} - F_{2n}^2 = 1$ for all $n \geq 1$.

Proof:

a) NOT COMPLETE.

b) We proceed by induction. The base case for $n = 1$ holds since

$$F_{2(1)+1}F_{2(1)-1} - F_{2(1)}^2 = F_3F_1 - F_2^2 = 2 \cdot 1 - 1^2 = 1. \tag{175}$$

For the inductive step, suppose the desired relation holds for case n. Then observe that

$$\begin{aligned} F_{2n+2}F_{2n} - F_{2n+1}^2 &= (F_{2n+1} + F_{2n})F_{2n} - (F_{2n} + F_{2n-1})F_{2n+1} \\ &= F_{2n}^2 - F_{2n+1}F_{2n-1} \\ &= -1. \end{aligned} \tag{176}$$

Similarly,

$$\begin{aligned} (F_{2n+2}F_{2n-1}) - F_{2n+1}F_{2n} &= (F_{2n+1} + F_{2n})F_{2n-1} - (F_{2n} + F_{2n-1})F_{2n} \\ &= F_{2n+1}F_{2n-1} - F_{2n}^2 \\ &= 1. \end{aligned} \tag{177}$$

Using the above, it follows that

$$\begin{aligned} &F_{2(n+1)+1}F_{2(n+1)-1} - F_{2(n+1)}^2 \\ &= F_{2n+3}F_{2n+1} - F_{2n+2}^2 \\ &= (F_{2n+2} + F_{2n+1})(F_{2n} + F_{2n-1}) - (F_{2n+1} + F_{2n})^2 \\ &= \left[F_{2n+2}F_{2n} - F_{2n+1}^2\right] + \left[F_{2n+1}F_{2n-1} - F_{2n}^2\right] + \left[F_{2n+2}F_{2n} - F_{2n+1}^2\right] \\ &= 1 - 1 + 1 \\ &= 1, \end{aligned} \tag{178}$$

and the hypothesis follows from the principle of mathematical induction.

□

F13.5: A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be convex if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \quad \forall\, x, y \in \mathbb{R}^d,\ t \in [0,1]. \tag{179}$$

Assume f is continuously differentiable such that

$$(\nabla f(x) - \nabla f(y)) \cdot (x - y) \geq 0 \quad \forall\, x, y, \in \mathbb{R}^d \tag{180}$$

where ∇f is the gradient of f and $\cdot$ is the inner product on $\mathbb{R}^d$. Prove that f is convex.

Proof:
Fix $x, y \in \mathbb{R}^d$ and define $z : [0,1] \to \mathbb{R}^d$ by $z(t) = tx + (1-t)y$. Then define $g : [0,1] \to \mathbb{R}$ by $g(t) := f(z(t))$ and $h : [0,1] \to \mathbb{R}$ by $h(t) := tf(x) + (1-t)f(y)$. To prove f is convex, we must show $(g-h)(t) \leq 0$ for $t \in [0,1]$. First, by definition of g and h,

$$(g-h)(0) = f(y) - f(y) = 0 \quad \text{and} \quad (g-h)(1) = f(x) - f(x) = 0. \tag{181}$$

We claim $(g-h)'$ is monotonically increasing. Then, by way of contradiction, suppose there is a $t^* \in (0,1)$ such that $(g-h)(t^*) > 0$. The mean value theorem implies there is a $\xi_1 \in (0, t^*)$ such that

$$(g-h)'(\xi_1)(t^* - 0) = (g-h)(t^*) - (g-h)(0) \quad \Rightarrow \quad (g-h)'(\xi_1) = \frac{(g-h)t^*}{t^*} > 0. \tag{182}$$

Similarly, there is a $\xi_2 \in (t^*, 1)$ such that

$$(g-h)'(\xi_2)(1-t^*) = (g-h)(1) - (g-h)(t^*) \quad \Rightarrow \quad (g-h)'(\xi_2) = -\frac{(g-h)(t^*)}{1-t^*} < 0. \tag{183}$$

But, $\xi_1 < \xi_2$ and $g'(\xi_1) > g'(\xi_2)$, contradicting the fact $(g-h)'$ is monotonically increasing. Hence $(g-h)(t) \leq 0$ for all $t \in [0,1]$.

All that remains is to verify $(g-h)'$ is monotonically increasing. Let $t \in (0,1)$. Then

$$g'(t) = \nabla f(z(t)) \cdot z'(t) = \nabla f(z(t)) \cdot (x-y) \quad \text{and} \quad h'(t) = f(x) - f(y). \tag{184}$$

Pick $t_1, t_2 \in (0,1)$ with $t_1 < t_2$. Then

$$(g-h)'(t_2) - (g-h)'(t_1) = [\nabla f(z(t_2)) - \nabla f(z(t_1))] \cdot (x-y) \tag{185}$$

where we note $h'(t_2) = h'(t_1)$. But,

$$z(t_2) - z(t_1) = [t_2 x + (1-t_2)y] - [t_1 x + (1-t_1)y] = (t_2 - t_1)(x - y). \tag{186}$$

Then, applying the hypothesis,

$$\begin{aligned}(t_2 - t_1) \cdot \left[(g-h)'(t_2) - (g-h)'(t_1)\right] &= (t_2 - t_1)\left[\nabla f(z(t_2)) - \nabla f(z(t_1))\right] \cdot (x - y) \\ &= \left[\nabla f(z(t_2)) - \nabla f(z(t_1))\right] \cdot (z(t_2) - z(t_1)) \\ &\geq 0.\end{aligned} \tag{187}$$

Dividing by $(t_2 - t_1)$ and then adding $(g-h)'(t_1)$ to each side, we conclude $(g-h)'(t_2) \geq (g-h)'(t_1)$. That is, $(g-h)'$ is monotonically increasing. This completes the proof. $\square$

F13.01: When $\{a_n\}$ is a sequence of positive real numbers, $a_n > 0$, define $P_n = \prod_{j=1}^{n}(1+a_j)$. Prove that $\lim_{n\to\infty} P_n$ exists and is a non-zero real number if and only if $\sum_{n=1}^{\infty} a_n < \infty$.

Proof:
First assume $\sum_{k=1}^{\infty} a_k < \infty$. Because the sequence is monotonically increasing (recall $a_k > 0$), the monotone convergence theorem implies the sequence of partial sums $\{\sum_{k=1}^{n} a_k\}_{n=1}^{\infty}$ converges to a limit positive $a \in \mathbb{R}$. Let $x > 0$. Then Taylor's theorem implies there is a $\xi_x \in (0, x)$ such that for the function $f(x) = e^x$

$$e^x = f(x) = f(0) + f'(0)x + \frac{f''(\xi_x)}{2} \cdot x^2 = e^0 + e^0 x + \frac{e^{\xi_x}}{2} \cdot x^2, \tag{188}$$

which implies

$$1 + x = e^x - \frac{e^{\xi_x}}{2} \cdot x^2 \le e^x. \tag{189}$$

Then

$$\prod_{k=1}^{n} 1 + a_k \le \prod_{k=1}^{n} e^{a_k} = \exp\left(\sum_{k=1}^{n} a_k\right) \le e^a. \tag{190}$$

But, this implies the sequence of partial products is bounded above. Moreover, because $1 + a_k > 1$ for each k, the sequence $\{\prod_{k=1}^{n} 1 + a_k\}_{n=1}^{\infty}$ is increasing. The monotone convergence theorem therefore implies the sequence converges. Hence $\lim_{n\to\infty} P_n$ exists and is greater than unity.

Conversely, suppose $\lim_{n\to\infty} P_n$ exists and equals, say, $x \in \mathbb{R}$ and note $x > 1$ since $1 + a_k > 1$ for each k. Then

$$\prod_{k=1}^{n} 1 + a_k \le x \quad \Rightarrow \quad \ln(x) \ge \ln\left(\prod_{k=1}^{n} 1 + a_k\right) = \sum_{k=1}^{n} \ln(1 + a_k). \tag{191}$$

where we recall the logarithm function is positive and increasing for arguments greater than unity. Since $\ln(1 + a_k) > 0$ for each k, the sequence of partial sums is monotonically increasing and, by (191), is bounded above. Thence the monotone convergence theorem implies this sum converges. This also yields that this sequence of partial sums is Cauchy, from which it follows that $\ln(1+a_k) \longrightarrow 0$ as $k \longrightarrow \infty$. Hence $a_k \longrightarrow 0$ as $k \longrightarrow \infty$. Consequently,

$$\lim_{k\to\infty} \frac{\ln(1 + a_k)}{a_k} = \lim_{z\to 0} \frac{\ln(1 + z)}{z} = \lim_{z\to 0} \frac{\frac{1}{1+z}}{1} = \frac{1}{1 + 0} = 1 \tag{192}$$

where we have made use of L'Hopital's rule to evaluate the limit. From the direct comparison lemma, we see the sequence of partial sums $\sum_{k=1}^{n} a_k$ converges if and only if the sequence of partial sums $\sum_{k=1}^{n} \ln(1 + a_k)$ converges. Because the latter sequence converges, we conclude the sequence of partial sums $\sum_{k=1}^{n} a_k$ converges. This completes the proof. $\square$

2014

S14.1: a) Find a real matrix A whose minimal polynomial is equal to $\lambda^4 + 1$.
b) Show that the usual real linear map determined by $v \mapsto Av$ has no non-trivial invariant subspace.

Proof:

a) This problem can be solved simply using the companion matrix of $t^4 + 1$, i.e., let

$$A = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Then the characteristic polynomial is

$$\chi(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda & 0 & 0 & -1 \\ 1 & \lambda & 0 & 0 \\ 0 & 1 & \lambda & 0 \\ 0 & 0 & 1 & \lambda \end{vmatrix} = (-1)^{1+1} \cdot \lambda \cdot \lambda^3 + (-1)^{4+1} \cdot (-1) \cdot 1^3 = \lambda^4 + 1.$$

Moreover, the characteristic polynomial factors over $\mathbb{C}$ as $\chi(\lambda) = (\lambda - \sqrt{i})(\lambda + \sqrt{i})(\lambda - i\sqrt{i})(\lambda + i\sqrt{i})$ and, thus, has distinct roots. So, the minimal polynomial is of at least degree 4, but the minimal polynomial divides the characteristic polynomial, which is also of degree 4. Hence these polynomials are equal and so the minimal polynomial of A is $\lambda^4 + 1$.

b) Suppose $v \mapsto Av$ has a non-trivial subspace. Then there is an eigenvector w of A, which then has an associated eigenvalue. But, each of the eigenvalues of A are complex. Hence the linear map $v \mapsto Av$ has no non-trivial invariant subspace. Part b) is supposedly note true.

$\square$

S14.2: Suppose that $S, T \in \text{Hom}(V, V)$ where V is a finite dimensional vector space over $\mathbb{R}$. Let (im S) be the image of S and (ker S) be the kernel of S. Show that

$$\dim(\text{im } S) + \dim(\text{im } T) \leq \dim(\text{im } (S \circ T)) + \dim V.$$

Proof:
Let us use im $S(K) = \{Sk \mid k \in K\}$ where K is a subspace of V. So, im $(S \circ T) =$ im $(S(\text{im } T))$. By the Fundamental Theorem of Linear Maps (FTOLM), it follows that

$$\dim(\text{im } (S(\ker T))) + \dim(\ker S(\ker T)) = \dim(\ker T).$$

and so $\dim(\text{im } (S(\ker T))) \leq \dim(\ker T)$. This implies

$$\dim(\ker T) + \dim(\text{im } (S(\text{im } T))) \geq \dim(\text{im } (S(\ker T))) + \dim(\text{im } (S(\text{im } T))) = \dim(\text{im } S)$$

where we have again used the FTOLM to assert $\dim(\text{im } S(\ker T)) + \dim(\text{im } S(\text{im } T)) = \dim(\text{im } S)$. Again using the FTOLM, $\dim V = \dim(\ker T) + \dim(\text{im } T)$,

$$\begin{aligned} \dim(\text{im } S) + \dim(\text{im } T) &\leq \dim(\text{im } S(\text{im } T)) + \dim(\ker T) + \dim(\text{im } T) \\ &= \dim(\text{im } S(\text{im } T)) + \dim V \\ &= \dim(\text{im } (S \circ T)) + \dim V, \end{aligned}$$

and we are done. □

S14.3: Suppose that $A, B \in M_{n\times n}(\mathbb{C})$ satisfy $AB - BA = A$. Show that A is not invertible.
Proof:
By way of contradiction, suppose A is invertible. Then

$$I = A^{-1}A = A^{-1}(AB - BA) = B - A^{-1}BA \quad \Rightarrow \quad A^{-1}BA = B - I.$$

However, recalling that $\text{tr}(CD) = \text{tr}DC$ for any $C, D \in M_{n\times n}(\mathbb{C})$, we see the above implies

$$\text{tr}(A^{-1}BA) = \text{tr}(A^{-1}AB) = \text{tr}(B) = \sum_{i=1}^{n} b_{ii}$$

while

$$\text{tr}(B - I) = \sum_{i=1}^{n}(b_{ii} - 1) = \left(\sum_{i=1}^{n} b_{ii}\right) - n.$$

But, then

$$\text{tr}(B - I) = \text{tr}(A^{-1}BA) \quad \Leftrightarrow \quad \left(\sum_{i=1}^{n} b_{ii}\right) - n = \sum_{i=1}^{n} b_{ii} \quad \Leftrightarrow \quad n = 0,$$

a contradiction. Hence A cannot be invertible. □

S14.4: Suppose that $A, B \in M_{n\times n}(\mathbb{C})$. Show that the characteristic polynomials of AB and BA are equal. *Hint:* One approach is to first show that it holds when B is invertible.

Proof:

First prove that the set of invertible matrices is dense in $M_{n\times n}(\mathbb{C})$. Let $A, B \in M_{n\times n}(\mathbb{C})$. Then, by the density of invertible matrices in $M_{n\times n}(\mathbb{C})$, there exists a sequence of invertible matrices $\{B_n\}$ that converges to B. And, since the characteristic polynomial χ of a matrix is, in fact, a polynomial, it is continuous. Hence

$$\lim_{n\to\infty} \chi(AB_n) = \chi(\lim_{n\to\infty} AB_n) = \chi(AB).$$

However, for each B_n we have

$$\begin{aligned}
\det(AB_n - \lambda I) &= \det(B_n B_n^{-1}) \det(AB_n - \lambda I) \\
&= \det(B_n) \det(AB_n - \lambda I) \det(B_n^{-1}) \\
&= \det(B_n(AB_n - \lambda I)B_n^{-1}) \\
&= \det(B_n A B_n B_n^{-1} - \lambda B_n I B_n^{-1}) \\
&= \det(B_n A - \lambda I).
\end{aligned}$$

Thus,

$$\lim_{n\to\infty} \chi(AB_n) = \lim_{n\to\infty} \chi(B_n A) = \chi(\lim_{n\to\infty} B_n A) = \chi(BA).$$

Thus, $\chi(AB) = \chi(BA)$ and we are done. $\square$

S14.5: Suppose that V is a finite dimensional real inner product space with inner product $\langle \cdot, \cdot \rangle$, that $L \in \mathrm{Hom}(V, V)$ and that $b \in V$ is fixed. Suppose that $u, v \in V$ both minimize $D(x) = \|L(x) - b\|$. Show that $u - v \in \ker L$.

Proof:

First let $\hat{b}$ denote the projection of b into the subspace $\mathrm{im}(L) \subset V$, i.e., define $\hat{b} = \langle b, e_1 \rangle + \cdots + \langle b, e_m \rangle e_m$ where $e_1, \ldots, e_m$ denotes an orthonormal basis for $\mathrm{im}(L)$. Clearly, $\hat{b} \in \mathrm{im}(L)$ since, by definition, it is expressed as a linear combination of the e_j's. Then $(b - \hat{b}) \in \mathrm{im}(L)^\perp$ since for any $j = 1, \ldots, m$ we have $\langle (b - \hat{b}), e_j \rangle = \langle b, e_j \rangle - \langle b, e_j \rangle = 0$, i.e., $(b - \hat{b})$ is orthogonal to every vector in $\mathrm{im}(L)$. It follows from the Pythagorean Theorem that, for each $z \in \mathrm{im}(L)$,

$$\|b - \hat{b}\|^2 \le \|b - \hat{b}\|^2 + \|\hat{b} - z\|^2 = \|(b - \hat{b}) + (\hat{b} - z)\|^2 = \|b - z\|^2$$

where the Pythagorean Theorem can be applied since $\hat{b} \in \text{im}(L)$ and $(b - \hat{b}) \in \text{im}(L)^{\perp}$. Taking square roots, this implies $\|b - \hat{b}\| \leq \|b - z\|$. Hence $\hat{b}$ is the closest point in $\text{im}(L)$ to b. Moreover, because $\hat{b} \in \text{im}(L)$, there is a $x_0 \in V$ such that $\hat{b} = L(x_0)$. This shows that not only does D have a minimizer x_0, but that $L(x_0) = \hat{b}$.

So, suppose $x, y \in V$ both minimize D. Then $L(x) = \hat{b} = L(y)$ and so $0 = L(x) - L(y) = L(x - y)$, which implies $(x - y) \in \ker(L)$, completing the proof. $\square$

S14.6: Show that if $A \in M_{n\times n}(\mathbb{C})$ is normal, then $A^* = P(A)$ for some polynomial $P(x)$ with complex coefficients. Here A^* is the conjugate transpose of A.

Proof:

Since A is normal, A and A^* are simultaneously diagonalizable, i.e., there is a unitary matrix U so that $D_1 = UAU^*$ and $D_2 = UA^*U^*$ where D_1 and D_2 are diagonal matrices. Then, using Lagrange interpolation, we can construct a polynomial of degree $n-1$ so that $P(\lambda_i) = \overline{\lambda}_i$ for $i = 1, \ldots, n$ where each λ_i is an eigenvalue of A and $\overline{\lambda}_i$ is an eigenvalue of A^*. Since D_1 and D_2 are diagonal matrices, it follows that

$$P(D_1) = \begin{pmatrix} P(\lambda_1) & & \\ & \ddots & \\ & & P(\lambda_n) \end{pmatrix} = \begin{pmatrix} \overline{\lambda}_1 & & \\ & \ddots & \\ & & \overline{\lambda}_n \end{pmatrix} = D_2.$$

There are $a_0, a_1, \ldots, a_{n-1}$ such that $P(t) = a_0 + a_1 t + \cdots + a_{n-1}t^{n-1}$ and so

$$\begin{aligned} P(A) &= a_0 I + a_1 A + \cdots + a_{n-1}A^{n-1} \\ &= a_0 UU^* + a_1 UD_1U^* + \cdots + a_{n-1}UD^{n-1}U^* \\ &= U\left(a_0 I + a_1 D_1 + \cdots + D^{n-1}\right)U^* \\ &= UP(D_1)U^* \end{aligned}$$

where we have used the easily verified relation that $A^n = UD_1^nU^*$. This implies $P(A) = UP(D_1)U^* = UD_2U^* = A^*$, as desired. $\square$

S14.7: Find a doubly infinite sequence $\{a_{n,m} \mid n, m \in \mathbb{Z}\}$ such that for all m, $\sum_n a_{n,m} = 0$ and for all n, $\sum_m a_{n,m} = 0$, with all these series converge absolutely, but such that $\sum_n \sum_m |a_{n,m}| = \infty$.

Proof:

Define the sequence $a_{m,n}$ by

$$a_{n,m} = \begin{cases} \operatorname{sgn}(m \cdot n)\left(1 - \frac{1}{|m|}\right)^{|n|}\left(1 - \frac{1}{|n|}\right)^{|m|} & \text{if } m \cdot n \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Note if $m \in \{-1, 0, 1\}$, then the summation $\sum_n a_{n,m}$ is identically zero. Otherwise, for fixed m, let $r = (1 - 1/|m|)$ so that

$$\begin{aligned} \left|\sum_{n=1}^{\infty} a_{n,m}\right| &= \left|\sum_{n=1}^{\infty} \operatorname{sgn}(m) \cdot \left(1 - \frac{1}{|m|}\right)^{|n|}\left(1 - \frac{1}{|n|}\right)^{|m|}\right| \\ &< \left|\sum_{n=1}^{\infty} \operatorname{sgn}(m) \cdot \left(1 - \frac{1}{|m|}\right)^{|n|}\right| \\ &= \sum_{n=1}^{\infty} r^n \\ &= \frac{r^{-1}}{1 - r} \\ &= \frac{1}{1 - 1/|m|} \cdot \frac{1}{1 - (1 - 1/|m|)} \\ &= \frac{|m|^2}{|m| - 1} \\ &\leq |m| \end{aligned}$$

where we note $(1 - 1/|n|) < 1$ and so $(1 - 1/|n|)^{|m|} < 1$. So, this sequence is bounded above. Moreover, the partial sums are monotonic since the sign of $a_{n,m}$ is constant when m is fixed and the sign of n does not change. Thus, by the monotone convergence theorem, the sum $\sum_{n=1}^{\infty} a_{n,m}$ converges to some limit S_m. Since $a_{n,m} = -a_{-n,m}$, $\sum_{n=-1}^{-\infty} a_{n,m}$ converges to $-S_m$. Moreover, $\sum_{n=-\infty}^{\infty} |a_{n,m}| = 2S_m$ and so the convergence is absolute. Thus, we may rearrange the order of terms in the series without changing the overall limit to see

$$\begin{aligned}
\sum_{n=-\infty}^{\infty} a_{n,m} &= \sum_{n=1}^{\infty} a_{n,m} + \sum_{n=-1}^{-\infty} a_{n,m} \\
&= \sum_{n=1}^{\infty} a_{n,m} - \sum_{n=1}^{\infty} a_{-n,m} \\
&= \sum_{n=1}^{\infty} a_{n,m} + a_{-n,m} \\
&= \sum_{n=1}^{\infty} a_{n,m} - a_{n,m} \\
&= 0.
\end{aligned}$$

Due to the symmetry in the definition of $a_{n,m}$, we see also that $\sum_m a_{n,m} = 0$ for fixed n. Lastly, observe that show that sum exceeets some value proportional to m, and so double sum diverges... $\square$

S14.12: Assume $[0,1] = \cup_{n=1}^{\infty} I_n$ where $I_n = [a_n, b_n] \neq \emptyset$ and $I_n \cap I_m = \emptyset$ whenever $n \neq m$.

a) Let $E = \{a_n \mid n \geq 1\} \cup \{b_n \mid n \geq 1\}$ be the set of endpoints of the intervals above. Prove E is closed.

b) Prove no such family of intervals $\{I_n\}$ can exist.

Proof:

a) We show E is closed by showing its complement E^c is open. By the Baire Category Theorem, if a non-empty complete metric space is the countable union of closed sets, then one of these closed sets has nonempty interior. Since $[0,1]$ is complete and nonempty, it follows that there exists I_j with nonempty interior, i.e., $b_j > a_j$ so that $(a_j, b_j) \neq \emptyset$. So, $E^c = \cup_{n=1}^{\infty}(a_n, b_n) \neq \emptyset$. Let $x \in E^c$. Then there is $k \in \mathbb{N}$ such that $x \in (a_k, b_k)$. Because (a_k, b_k) is open, there exists $r > 0$ such that $(x_r, x+r) \subseteq (a_k, b_k)$. Hence $x \in \text{int}(E^c)$ and so $E^c \subseteq \text{int}(E^c)$. But, by definition of interior, $\text{int}(E^c) \subseteq E^c$. Hence $\text{int}(E^c) = E^c$ and so E^c is open, which implies E is closed.

b) Suppose $x \in E$. Then there exists I_k such that $x \in I_k$, specifically, either $x = a_k$ or $x = b_k$. Then for each $r > 0$, $B(x, r)$ contains points not in I_k since $x \in \text{bd}(I_k)$. So, x is a limit point of some I_j with $j \neq k$. But, because I_j is closed, $x \in I_j$. This contradicts the fact that $I_j \cap I_k = \emptyset$. Hence E must be empty, i.e., $E = \emptyset$. But, then $I_n = \emptyset$ for each $n \in \mathbb{N}$, a contradiction. Thus, there does not exists such a family of $\{I_n\}$.

□

F14.2: Let A, B be two closed subsets of $\mathbb{R}^n$ such that $A \cup B$ and $A \cap B$ are connected. Prove that A is connected.

Proof:

By way of contradiction, suppose A is not connected. Then there are nonempty open subsets $S_1, S_2 \subset \mathbb{R}^n$ with $A \subseteq S_1 \cup S_2$ and $S_1 \cap S_2 = \emptyset$. Then

$$A \cap B \subseteq (S_1 \cup S_2) \cap B = (S_1 \cap B) \cup (S_2 \cap B).$$

By way of contradiction, suppose without loss of generality that $S_2 \cap B = \emptyset$. Then $A - B = S_2$ and so

$$A \cup B = (A - B) \cup B = S_2 \cup B \tag{193}$$

where S_2 and B are disjoint. But, this contradicts the fact $A \cup B$ is connected. Hence $S_1 \cap B$ and $S_2 \cap B$ are nonempty. But, $S_1 \cap B \subset S_1$ and $S_2 \cap B \subset S_2$ and so $(S_1 \cap B) \cap (S_2 \cap B) = \emptyset$. This implies $A \cap B$ is not connected, a contradiction. Thus, the assumption that A is not connected must be false and so we conclude A is connected. $\square$

F14.3: Let $f_n : [0,1] \to \mathbb{R}$ be monotonically increasing continuous functions. Assume that f_n converges pointwise to a continuous function $f : [0,1] \to \mathbb{R}$. Prove that $f_n \longrightarrow f$ uniformly.

Proof:
Let $\varepsilon > 0$ be given. We must show there is an $N \in \mathbb{Z}^+$ such that

$$\|f_n - f\|_\infty < \varepsilon \quad \forall\ n \geq N. \tag{194}$$

Since $[0,1]$ is closed and bounded, Heine-Borel's theorem implies it is compact. This, combined with the fact f is continuous, yields that f is uniformly continuous. Hence there is a $\delta > 0$ such that for $x, y \in [0,1]$,

$$|x - y| < \delta \qquad \Rightarrow \qquad |f(x) - f(y)| < \varepsilon/2. \tag{195}$$

Now cover $[0,1]$ with the collection of open balls $B(x,\delta)$ for $x \in [0,1]$. Since $[0,1]$ is compact, there is a finite subcover, i.e., there are $x_1, \ldots, x_p \in [0,1]$ such that

$$[0,1] \subset \bigcup_{i=1}^{p} B(x_i, \delta). \tag{196}$$

By the pointwise convergence of $f_n \longrightarrow f$, for each $i \in \{1, \ldots, p\}$ there is an N_i such that

$$|f_n(x_i) - f(x_i)| < \frac{\varepsilon}{2} \quad \forall\ n \geq N_i. \tag{197}$$

Set $N := \max\{N_1, \ldots, N_p\}$. We claim this choice of N satisfies (194), which we verify as follows. Let $x \in [0,1]$. Then there is an index i such that $x \in [x_i, x_{i+1}]$. Since f_n is monotonic,

$$f_n(x_i) \leq f_n(x) \leq f_n(x_{i+1}). \tag{198}$$

Then apply (197) to write

$$f(x_i) - \frac{\varepsilon}{2} \leq f_n(x) \leq f(x_{i+1}) + \frac{\varepsilon}{2}. \tag{199}$$

And, by choice of the cover in (196), $|x_i - x| < \delta$ and $|x_{i+1} - x| < \delta$. By the continuity of f in (195), this implies

$$\left[f(x) - \frac{\varepsilon}{2}\right] - \frac{\varepsilon}{2} \leq f_n(x) \leq \left[f(x) + \frac{\varepsilon}{2}\right] + \frac{\varepsilon}{2}, \tag{200}$$

from which it follows

$$|f(x) - f_n(x)| < \varepsilon. \tag{201}$$

Since this argument holds for arbitrary $x \in [0,1]$, it holds for all such x and we conclude (194) holds. This completes the proof. $\square$

F14.4: Let $f_n : [-2, 2] \to [0, 1]$ be a sequence of convex functions. Show that there is a subsequence which converges uniformly on $[-1, 1]$.

Proof:
We claim the restriction $\{g_n\}$ of the sequence $\{f_n\}$ to $[-1, 1]$ is equicontinuous and uniformly bounded. By definition, since f_n maps to $[0, 1]$, $\|f_n\|_\infty \leq 1$ for all $n \in \mathbb{Z}^+$. Let $\varepsilon > 0$ be given. Then $\{g_n\}$ is an equicontinuous family of function provided, given $\varepsilon > 0$, there is a $\delta > 0$ such that

$$|x - y| < \delta \qquad \Rightarrow \qquad |g_n(x) - g_n(y)| < \varepsilon \quad \forall\ x, y \in [-1, 1],\ n \in \mathbb{Z}^+. \tag{202}$$

We first use a diagonalization argument to show there is a subsequence $\{f_{n_k}\}$ that converges pointwise on the rationals in $[-1, 1]$.

Let $\varepsilon > 0$ be given. We must show there is an $N \in \mathbb{Z}^+$ such that

$$|f_{n_k}(x) - f(x)| < \varepsilon \quad \forall\ n \geq N,\ x \in [-1, 1]. \tag{203}$$

□

F14.5: Consider the following sequence:

$$a_1 = \sqrt{2} \quad \text{and} \quad a_{n+1} = \sqrt{2 + a_n} \ \ \forall \ n \geq 1.$$

Prove that this sequence converges and find its limit.

Proof:
We claim that $\{a_n\}_{n=1}^{\infty}$ is monotonically increasing and bounded above. Indeed, $a_1 = \sqrt{2} < 2$ and, if $a_n \leq 2$, then

$$a_{n+1} = \sqrt{2 + a_n} \leq \sqrt{2 + 2} = 2,$$

which closes in the induction. To see that $\{a_n\}_{n=1}^{\infty}$ is increasing, observe that

$$a_n^2 - a_n = a_n(a_n - 1) \leq 2(a_n - 1) \leq 2(2 - 1) = 2.$$

Thus, $a_n^2 \leq 2 + a_n$ and so $a_n \leq \sqrt{2 + a_n} = a_{n+1}$. It follows from the Monotone Convergence Theorem that $\{a_n\}$ converges to some limit L. Then, since the function $f(x) = \sqrt{2 + x}$ is continuous,

$$\lim_{n \to \infty} f(a_n) = f(L) = \sqrt{2 + L}.$$

But,

$$\lim_{n \to \infty} f(a_n) = \lim_{n \to \infty} a_{n+1} = \lim_{n \to \infty} a_n = L.$$

This implies $L = \sqrt{2 + L}$ and so $2 + L = L^2$, which implies $(L - 2)(L + 1) = 0$. Since $a_n \geq \sqrt{2}$ for each $n \in \mathbb{N}$, it follows that the limit is $L = 2$. $\square$

F14.6: Let $f : [0,1] \to \mathbb{R}$ be a C^1 function. Prove that

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \left| f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right) \right| = \int_0^1 |f'(t)| \ \mathrm{d}t. \tag{204}$$

Solution:

Define $g(t) := |f'(t)|$ and observe g is continuous since it is the composition of continuous functions. By definition, g is said to be Riemann integrable on $[0,1]$ provided

$$\sup_{P\in\Pi} L(g,P) = \inf_{P\in\Pi} U(g,P) \tag{205}$$

where Π is the set of all partitions of $[0,1]$ and, for a partition $P = \{I_1, \dots, I_k\} \in \Pi$,

$$L(g,P) := \sum_{i=1}^{k} \left(\inf_{x\in I_i} g\right) |I_i| \quad \text{and} \quad U(g,P) := \sum_{i=1}^{k} \left(\sup_{x\in I_i} g\right) |I_i|. \tag{206}$$

When (205) holds, we write the integral of g as

$$\int_0^1 g(t) \ \mathrm{d}t := \sup_{P\in\Pi} L(g,P) = \inf_{P\in\Pi} U(g,P). \tag{207}$$

We first show (205) holds for our choice of g. It suffices to find a sequence of partitions P_n such that

$$\lim_{n\to\infty} U(g,P_n) = \sup_{P\in\Pi} L(g,P). \tag{208}$$

Let $\varepsilon > 0$ be given. Since g is continuous on $[0,1]$, which is closed and bounded and therefore compact, g is uniformly continuous. So, there is a $\delta > 0$ such that for all $x, y \in [0,1]$,

$$|x-y| \le \delta \quad \Rightarrow \quad |g(x) - g(y)| \le \varepsilon. \tag{209}$$

By the Archimedean property of $\mathbb{R}$, there is a $N \in \mathbb{Z}^+$ such that $1/N < \delta$. For $n \in \mathbb{Z}^+$ set $P_n = \{I_0, \dots, I_{n-1}\} \in \Pi$ where $I_i = [i/n, i+1/n]$. Then for all $n \ge N$ we discover

$$\begin{aligned} U(g,P_n) = \sum_{k=0}^{n-1} \left(\sup_{x\in I_k} g\right) |I_k| &\le \sum_{k=0}^{n-1} \left(\inf_{x\in I_k} g + \varepsilon\right) |I_k| \\ &= \sum_{k=0}^{n-1} \left(\inf_{x\in I_k} g\right) |I_k| + \varepsilon \sum_{k=0}^{n-1} |I_k| \\ &= L(g,P_n) + \varepsilon \\ &\le \sup_{P\in\Pi} L(g,P) + \varepsilon \end{aligned} \tag{210}$$

where the first inequality follows from (209). And, by definition of U, $U(g,P_n) \geq \sup_{P\in\Pi} L(g,P)$ for every $n \in \mathbb{Z}^+$. Hence

$$\left| U(g,P_n) - \sup_{P\in\Pi} L(g,P) \right| \leq \varepsilon \quad \forall\ n \geq N, \tag{211}$$

and so we obtain the limit in (208). A corresponding limit holds for $L(g,P_n)$, which is found by similar argument.

For $k \in \mathbb{Z}$, the mean value theorem implies there exists $\xi_k \in (k/n,(k+1)/n)$ such that

$$\left| f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right) \right| = \left| f'(\xi_k)\cdot\left(\frac{k+1}{n} - \frac{k}{n}\right)\right| = g(\xi_k)\cdot\frac{1}{n}. \tag{212}$$

So,

$$\sum_{k=0}^{n-1} \left| f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right) \right| = \sum_{k=0}^{n-1} g(\xi_k)\cdot\frac{1}{n} = \sum_{k=0}^{n-1} g(\xi_k)\cdot|I_k| \leq \sum_{k=0}^{n-1}\left(\sup_{x\in I_k} g\right)\cdot|I_k| = U(g,P_n). \tag{213}$$

We obtain a corresponding inequality with $L(g,P_n)$. Hence applying (208) and the corresponding limit for $L(g,P_n)$, we see

$$\int_0^1 g(t)\ \mathrm{d}t \leq \lim_{n\to\infty}\sum_{k=0}^{n-1} \left| f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right) \right| \leq \int_0^1 g(t)\ \mathrm{d}t, \tag{214}$$

from which we conclude

$$\lim_{n\to\infty}\sum_{k=0}^{n-1} \left| f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right) \right| = \int_0^1 g(t)\ \mathrm{d}t, \tag{215}$$

as desired.

□

F14.7: Among all the solutions to the system

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 5 & 7 \\ -2 & -1 & 1 & 3 \end{pmatrix} x = \begin{pmatrix} 2 \\ 7 \\ -1 \end{pmatrix}, \tag{216}$$

find the solution with minimal length.

Solution:
Observe this linear system is equivalent to

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 2 & 3 & 5 & 7 & 7 \\ -2 & -1 & 1 & 3 & -1 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 3 & 5 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{217}$$

where we note the third equation of the given linear system is equal to the second equation minus the four times the first equation. Define

$$A := \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 5 \end{pmatrix} \quad \text{and} \quad b := \begin{pmatrix} 2 \\ 5 \end{pmatrix} \tag{218}$$

and observe A has full rank. We seek to solve

$$\min \|x\| \ \text{ such that } \ Ax = b. \tag{219}$$

Since $\|\cdot\|$ is convex, we can equivalently minimize $\|x\|^2/2$ such that $Ax = b$, which we do by applying Lagrange multipliers. The Lagrangian for this problem is given by

$$\mathcal{L}(x, \lambda) = \frac{x^T x}{2} + \lambda^T(Ax - b) = \frac{x^T x}{2} + (A^T\lambda)^T x - \lambda^T b. \tag{220}$$

We seek to minimize $\mathcal{L}$ over x and λ. Due to convexity, it suffices to find a critical point of $\mathcal{L}$. At such a point $(\bar{x}, \bar{\lambda})$,

$$0 = \nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}) = \bar{x} + A^T\bar{\lambda} \quad \Rightarrow \quad \bar{x} = -A^T\bar{\lambda}, \tag{221}$$

and

$$0 = \nabla_\lambda \mathcal{L}(\bar{x}, \bar{\lambda}) = A\bar{x} - b = A(-A^T\bar{\lambda}) - b \quad \Rightarrow \quad \bar{\lambda} = -(AA^T)^{-1}b \tag{222}$$

where the inverse of AA^T is well-defined since A has full rank. Hence, by back substituting, we conclude the solution is given by

$$\boxed{\bar{x} = A^T(AA^T)^{-1}b,} \tag{223}$$

which can be computed explicitly using the definitions of A and b above. □

F14.8: Compute the eigenvalues of the following $n \times n$ matrix:

$$M = \begin{pmatrix} k & 1 & 1 & \cdots & 1 \\ 1 & k & 1 & \cdots & 1 \\ 1 & 1 & k & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & k \end{pmatrix}.$$

Use the eigenvalues to compute $\det(M)$.

Proof:

Note that $M = (k-1)I + \mathbb{1}$ where I denotes the identity matrix and $\mathbb{1}$ is the matrix of all ones. Of course, $(k-1)$ is an eigenvalue of I since I is the identity. Also, for each v observe that

$$\mathbb{1}v = \left(\sum_{i=1}^{n} v_i\right)\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

This gives two possibilities for eigenvectors. Either $\sum_{i=1}^n v_i = 0$ or $\sum_{i=1}^n v_i \neq 0$ and $v_i = c$ for $i = 1, \ldots, n$. If it is the latter, then

$$Mv = ((k-1)I + \mathbb{1})v = (k-1)v + nv = (n+k-1)v$$

and so $(n+k-1)$ is an eigenvalue of M. Note well the dimension of this eigenspace is 1 since it occurs precisely for scalars of the vectors of all ones. In the other case, if $\sum_{i=1}^n v_i = 0$, then

$$Mv = ((k-1)I + \mathbb{1})v = (k-1)v + 0v = (k-1)v,$$

which implies $k-1$ is an eigenvalue of M. Since this is the only other eigenvalue of $\mathbb{1}$, it must follow that it has dimension $n-1$. Since $\det(M)$ is the product of the eigenvalues of M, we find $\boxed{\det(M) = (n+k-1)(k-1)^{n-1}.}$ $\square$

F14.10: What is the largest number of 1's an invertible 0-1 matrix of size $n \times n$ can have? You must show both that this number is possible and that no larger number is possible.

Proof:
The largest number of 1's an invertible 0-1 matrix of size $n \times n$ can have is $n^2 - n$. Define the vector v_i to have a 0 in the i-th position and zeros elsewhere so that v_i contains $(n-1)$ ones. Then we claim the matrix $[v_1 \ v_2 \ \cdots \ v_n]$ is invertible. To verify this, it suffices to show the v_i are independent. Suppose there are scalars $a_1, \ldots, a_n \in \mathbb{R}$ with $a_j \neq 0$ such that

$$0 = \sum_{i=1}^{n} a_i v_i \qquad \Rightarrow \qquad v_j = -\frac{1}{a_j} \cdot \sum_{i=1, i \neq j}^{n} a_i v_i \tag{224}$$

By way of contradiction, suppose there is a 0-1 matrix with more than $n^2 - n$ ones. $\square$

F14.11: Suppose a 4×4 integer matrix has four distinct real eigenvalues $\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4$. Prove that $\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 \in \mathbb{Z}$.

Proof:
Since $\mathbb{Z}$ is a ring closed under multiplication and addition, it follows that M^2 has elements in $\mathbb{Z}$. Let v_i denote the eigenvector corresponding to λ_i for $i = 1, 2, 3, 4$. Then v_i is an eigenvector of M^2 since

$$M^2 v_i = M(Mv_i) = M(\lambda_i v_i) = \lambda_i^2 v_i \quad \forall\ i = 1, 2, 3, 4.$$

Since the trace of a matrix is defined to be the sum of its diagonal elements, it follows that $\text{tr}(M^2) \in \mathbb{Z}$. However, the trace of a matrix is also the sum of its eigenvalues. Hence

$$\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 = \text{tr}(M^2) \in \mathbb{Z},$$

and we are done. $\square$

F14.12: Let $A = (a_{ij})_{1\leq i,j\leq n}$ where $a_{ij} = 1/(i+j-1)$. Prove that A is positive definite.

Proof:
We must show $\langle Ax, x\rangle \geq 0$ for all $x \in V$ where we assume $V = \mathbb{C}^n$. Defining $v = (1, t, \ldots, t^{n-1})$, we discover

$$\int_0^1 (v^T v)_{ij}\ \mathrm{d}t = \int_0^1 t^{i+j-2}\ \mathrm{d}t = \frac{1}{i+j-1} = a_{ij}.$$

This implies, for $x \in \mathbb{C}^n$,

$$\begin{aligned}
\langle Ax, x\rangle &= (Ax)^* x \\
&= x^* A^* x \\
&= x^* A x \\
&= x^* \left(\int_0^1 v^T v\ \mathrm{d}t \right) x \\
&= \int_0^1 x^* v^T v x\ \mathrm{d}t \\
&= \int_0^1 (vx)^* vx\ \mathrm{d}t \\
&= \int_0^1 \|vx\|^2\ \mathrm{d}t.
\end{aligned}$$

The first equality holds by definition of the scalar product in $\mathbb{C}^n$, the second by definition of the conjugate transpose, and the next by the fact A is hermitian. The following equality holds by substitution for each entry in A. Then, using linearity we bring x^* and x inside the integral. Then we use the definition of the norm to substitute and obtain the final equality. And, since $\|vx\|^2 \geq 0$ for all $x \in \mathbb{C}^n$, it follows that $\int_0^1 \|vx\|^2\ \mathrm{d}t \geq 0$. Thus,$\langle Ax, x\rangle \geq 0$, as desired. $\square$

2015

S15.1: Let $f : [0, \infty) \to [0, \infty)$ be continuous with $f(0) = 0$. Show that if

$$f(t) \leq 1 + \frac{1}{10} f(t)^2 \quad \forall\, t \in [0, \infty),$$

then f is uniformly bounded throughout $[0, \infty)$.

Proof:

By way of contradiction, suppose f is not uniformly bounded. Then there exists $b > 0$ such that $f(b) \geq 6$. Then, noting that $f(0) = 0$, it follows from the Intermediate Value Theorem that there exists $x \in [0, b]$ such that $f(x) = 5$ since $5 \in (f(0), f(b))$. However, this implies

$$5 = f(x) \leq 1 + \frac{1}{10} f(x)^2 = 1 + \frac{1}{10} \cdot 25 = 3.5,$$

which is a contradiction. Hence f is uniformly bounded. □

S15.4: Let $f[0, 1] \to \mathbb{R}$ be a function satisfying the intermediate value property, namely, whenever $0 \leq a < b \leq 1$ and y lies between $f(a)$ and $f(b)$, there exists $x \in (a, b)$ such that $f(x) = y$. Assume that for any $y \in \mathbb{R}$, the pre-image $f^{-1}(\{y\})$ is closed. Prove that f is continuous.

Proof:

Let $\varepsilon > 0$ be given. Pick $x_0 \in [0, 1]$. Pick $a_0 = 0$. Since $(f(x_0) + f(a_1))/2$ is contained between $f(x_0)$ and $f(a_1)$, there exists a_2 between x_0 and a_1 such that $f(a_2) = (f(x_0) + f(a_1))/2$. Then

$$|f(x_0) - f(a_2)| = \left| \frac{f(x_0) - f(a_1)}{2} \right|.$$

Continuing to choose a_n in this fashion, through induction it follows that

$$|f(x_0) - f(a_n)| = 2^{-n} \left| f(x_0) - f(a_1) \right|.$$

Note $a_n \leq x_0$ for each $n \in \mathbb{N}$. We can define a similar sequence $\{b_n\}$ with the single change being that $b_0 = 1$. Then $x_0 \in [a_n, b_n]$ for each $n \in \mathbb{N}$.

We claim f is monotonic. By way of contradiction, suppose otherwise. Then there exists $a, b, c \in [0, 1]$ such that c is between a and b, but $f(c)$ is not between $f(a)$ and $f(b)$.

Without loss of generality, suppose $b > a$. Then pick y between $f(a)$ and $f(b)$. Then there exists $c \in (a, b)$ such that $f(c) = y$.

□

S15.5: Let $f : [1, \infty) \to [0, \infty)$ be bounded and monotonically decreasing with $\lim_{x\to\infty} f(x) = 0$. Show that

$$\int_1^{N+1} f(x)\ \mathrm{d}x - \sum_{n=1}^{N} f(n)$$

converges to a finite limit as $N \to \infty$.

Proof:

We first show that the given sequence is bounded. Denote the N-th term of our sequence by a_N, i.e.,

$$a_N = \int_1^{N+1} f(x)\ \mathrm{d}x - \sum_{n=1}^{N} f(n).$$

Indeed, using the definition of Riemann integrable,

$$a_N = \inf_P U(f, P) - \sum_{n=1}^{N} f(n) \leq \sum_{n=1}^{N} f(n) - \sum_{n=1}^{N} f(n) = 0$$

where $U(f, P)$ denotes the upper Riemann sum using partition P of the interval $[1, N + 1]$ and we have used the fact that f is monotonically decreasing to assert $\sum_{n=1}^{N} f(n)$ is an upper Riemann sum. Similarly, $\sum_{n=1}^{N} f(n + 1)$ is a lower Riemann sum for this integral and so

$$a_N = \sup_P L(f, P) - \sum_{n=1}^{N} f(n) \geq \sum_{n=1}^{N} f(n + 1) - \sum_{n=1}^{N} f(n) = f(N + 1) - f(1) \geq -f(1).$$

Thus, $a_N \in [-f(1), 0]$ for each $N \in \mathbb{N}$. Now we claim $\{a_N\}_{N=1}^{\infty}$ is monotonically decreasing. To see this, observe that, by the linearity of the integral,

$$\begin{aligned} a_{N+1} &= \left(\int_N^{N+1} f(x)\ \mathrm{d}x - f(N)\right) + \left(\int_1^{N+1} f(x)\ \mathrm{d}x - \sum_{n=1}^{N} f(n)\right) \\ &= \int_N^{N+1} f(x)\ \mathrm{d}x - f(N) + a_N \\ &\leq a_N \end{aligned}$$

where we use the fact that

$$\int_N^{N+1} f(x)\ \mathrm{d}x \leq \int_N^{N+1} f(N)\ \mathrm{d}x = f(N).$$

Hence $\{a_N\}_{N=1}^{\infty}$ is a bounded monotonically decreasing sequence. By the Monotone Convergence Theorem, $\{a_N\}_{N=1}^{\infty}$ must converge to some limit in $[-f(1), 0]$, and we are done.

Below we present an alternative solution.

Define S_N to be the N-th term of our sequence and note

$$S_N = \int_1^{N+1} f(x)\ \mathrm{d}x - \sum_{n=1}^{N} f(n) = \sum_{n=1}^{N} \int_n^{n+1} f(x)\ \mathrm{d}x - f(n).$$

Since f_n is monotonically decreasing and maps to nonnegative numbers,

$$\left| \int_n^{n+1} f(x)\ \mathrm{d}x - f(n) \right| = f(n) - \int_n^{n+1} f(x)\ \mathrm{d}x \leq f(n) - f(n+1)$$

which implies for $M > N$ we have

$$\begin{aligned} |S_M - S_N| &= \left| \sum_{n=N+1}^{M} \int_n^{n+1} f(x)\ \mathrm{d}x - f(n) \right| \\ &\leq \sum_{n=N+1}^{M} \left| \int_n^{n+1} f(x)\ \mathrm{d}x - f(n) \right| \\ &\leq \sum_{n=N+1}^{M} f(n) - f(n+1) \\ &= f(M) - f(N+1). \end{aligned}$$

But, $\lim_{n\to\infty} f(n) = 0$, which implies $f(n)$ is Cauchy, i.e., given $\varepsilon > 0$, there exist $K \in \mathbb{N}$ such that for all $M, N \geq K$ we have $|f(M) - f(N)| \leq \varepsilon$. Thus, for $M, N \geq K$

$$|S_M - S_N| \leq |f(M) - f(N+1)| \leq \varepsilon,$$

and so $\{S_N\}$ is Cauchy. Since $\mathbb{R}$ is complete, S_N converges to a finite limit and we are done. $\square$

S15.6: Prove that the integral equation

$$f(t) = e^{t^2} + \frac{1}{2}\int_0^1 \cos(s)f(s)\ \mathrm{d}s$$

admits a unique continuous solution $f : [0,1] \to \mathbb{R}$.

Proof:

We proceed by applying the Banach Fixed Point theorem. To do this, we verify $C[0,1]$ is complete and define a contraction $T : C[0,1] \to C[0,1]$. First let $\{f_n\}$ be a Cauchy sequence in $C[0,1]$. For each $x \in [0,1]$, $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{R}$ and since $\mathbb{R}$ is complete its limit is in $\mathbb{R}$. Hence the pointwise limit $\lim\limits_{n\to\infty} f_n(x)$ exists. To see that $f_n \longrightarrow f$ in norm, note that because $\{f_n\}$ is Cauchy, there is $N \in \mathbb{N}$ such that

$$\|f_n - f_m\|_\infty < \varepsilon/3 < \varepsilon \quad \forall\ m, n \geq N. \tag{225}$$

But, since norms are continuous, we have the limit $\lim\limits_{m\to\infty} \|f_n - f_m\| = \|f_n - f\|$ and so

$$\|f_n - f\| < \varepsilon/3 < \varepsilon \quad \forall\ n \geq N, \tag{226}$$

as desired. We now must show that f is continuous. Since f_N is continuous on a compact set, it is uniformly continuous. Thus, there is a $\delta > 0$ such that for $x, y \in [0,1]$,

$$|x - y| < \delta \quad \Rightarrow \quad |f_N(x) - f_N(y)| < \varepsilon/3. \tag{227}$$

Using N as above, this implies that, whenever $|x - y| < \delta$,

$$|f(x) - f(y)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \tag{228}$$

Thus, $f \in C[0,1]$ and so $C[0,1]$ is complete.

(continued on next page)

Now define the operator $T : C[0,1] \to C[0,1]$ for each $f \in C[0,1]$ by

$$(Tf)(t) = e^{t^2} + \frac{1}{2}\int_0^1 \cos(s) f(s)\ \mathrm{d}s,$$

and $T(f)$ is continuous since it is the composition of several continuous functions. For $f, g \in C[0,1]$, observe that

$$\begin{aligned} \|Tf - Tg\|_\infty &= \left\| \frac{1}{2}\int_0^1 \cos(s)[f(s) - g(s)]\ \mathrm{d}s \right\|_\infty \\ &\leq \frac{1}{2}\left\| \int_0^1 \cancelto{1}{\|\cos(s)\|_\infty} \cdot \|f - g\|_\infty\ \mathrm{d}s \right\|_\infty \\ &= \frac{1}{2}\|f - g\|_\infty. \end{aligned}$$

Hence T is a contraction with Lipschitz constant $1/2$. Since T is a contraction and $C[0,1]$ is complete, the Banach Fixed Point theorem implies for any $f_0 \in C[0,1]$ (e.g., $f = 0$), the sequence defined by $f_{n+1} = T(f_n)$ for $n \geq 1$ converges to a unique fixed point f of T, i.e., $\lim_{n\to\infty} f_n = f$ and $T(f) = f$. This means there is a unique $f \in C[0,1]$ such that

$$f(t) = T(f)(t) = e^{t^2} + \frac{1}{2}\int_0^1 \cos(s) f(s)\ \mathrm{d}s,$$

and we are done. $\square$

S15.7: Let $f(x,y,z) = 9x^2 + 6y^2 + 6z^2 + 12xy - 10xz - 2yz$. Does there exists a point (x,y,z) such that $f(x,y,z) < 0$?

Solution:
No, there does not exist such a point. By way of contradiction, suppose there does. Since f is continuous $\square$

S15.12: Let

$$M = \begin{pmatrix} 3 & 5 \\ 1 & -1 \end{pmatrix}.$$

a) Compute $\exp(M)$.

b) Does there exists a real 2×2 matrix A such that $M = \exp(A)$?

Solution:

a) First we diagonalize M. NOT COMPLETE.

b) We begin with the following lemma:

Lemma: Let $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$. 1 is not an eigenvalue of A iff $I - A$ is invertible.

We first show $I - A$ is invertible iff 1 is not an eigenvalue of A. We argue by proving the contrapositive of each implication in this claim. First suppose 1 is an eigenvalue of A. Then there is nonzero $v \in \mathbb{C}^n$ such that $(I - A)v = 0$, implying that $I - A$ is not one-to-one and, thus, not invertible. Hence if A is invertible, then 1 is not an eigenvalue of A. Now suppose A is singular. Then $\det(I - A) = 0$, which implies 1 is an eigenvalue of A. Hence if 1 is not an eigenvalue of A, then $I - A$ is invertible. ■

Now define $Q = I - M$. Then the sought matrix A exists iff $\ln(M) = \ln(I - Q)$ exists. By the above lemma, $I - Q$ is invertible precisely when 1 is not an eigenvalue of Q.

Show eigenvalues are 3 and -3.

□

F15.1 Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of positive numbers such that $a_{n+m} \leq a_n + a_m,\ m, n \geq 1$. Prove that $\lim_{n\to\infty} a_n/n$ exists by showing $\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\geq 1} \frac{a_n}{n}$.

Proof:
Let $\ell := \inf_{n\geq 1} \frac{a_n}{n}$ and $\varepsilon > 0$ be given. We must show there is a $N \in \mathbb{Z}^+$ such that

$$\left|\frac{a_n}{n} - \ell\right| < \varepsilon \quad \text{whenever } n \geq N. \tag{229}$$

Since ℓ is the greatest lower bound of a_n/n, there exists a $K \in \mathbb{Z}^+$ such that

$$\left|\frac{f(K)}{K} - \ell\right| < \frac{\varepsilon}{2}. \tag{230}$$

Then by the Archimedean property of $\mathbb{R}$, there is a $L \in \mathbb{Z}^+$ such that

$$\frac{1}{L} < \frac{MK\varepsilon}{2} \quad \text{where} \quad M := \max_{1\leq r\leq K}\{a_r\}. \tag{231}$$

This implies

$$\frac{a_r}{KL} < \frac{\varepsilon}{2} \quad \text{for } 1 \leq r \leq K. \tag{232}$$

We claim (229) holds if $N := KL$. To see this, let $n \geq N$. By Euclid's division lemma, there are nonnegative $q, r \in \mathbb{Z}$ such that $n = Kq + r$ with $r < K$. And, by choice of N, we know $q \geq L$. Using subaddtivity, we see

$$\frac{a_n}{n} \leq \frac{a_{qK}}{n} + \frac{a_r}{n} = \frac{a_{qK}}{qK+r} + \frac{a_r}{qK+r}. \tag{233}$$

And, from (232), we have

$$\frac{a_r}{qK+r} \leq \frac{a_r}{qK} \leq \frac{a_r}{LK} \leq \frac{\varepsilon}{2}. \tag{234}$$

Applying the lemma below, we know

$$\frac{a_{qK}}{qK+r} \leq \frac{a_{qK}}{qK} \leq \frac{qa_K}{qK} = \frac{a_K}{K} \leq \ell + \frac{\varepsilon}{2}. \tag{235}$$

Combining (233)-(235), we see

$$\frac{a_n}{n} \leq \ell + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \ell + \varepsilon \quad \text{whenever } n \geq N, \tag{236}$$

which implies (229) holds. This completes the proof. □

(Lemma on next page.)

Lemma: If $\{a_n\}_{n=1}^{\infty}$ *is subadditive, then* $a_{mn} \leq m a_n$ *for all* $m, n \geq 1$.

Proof:
We verify this by induction on n. The base case holds trivially. Suppose now this claim holds for some $n \in \mathbb{Z}^+$. Then observe that

$$a_{m(n+1)} \leq a_{mn} + a_m \leq n a_m + a_m = (n+1) a_m, \tag{237}$$

and we have closed the induction. The claim follows by the principle of mathematical induction.

□

F15.2: Let $a, b \in \mathbb{R}$ obey $a < b$. Show that if $g, h : [a, b] \to \mathbb{R}$ are continuous with $h \geq 0$, then there is $c \in [a, b]$ such that
$$\int_a^b g(x)h(x)\ \mathrm{d}x = g(c)\int_a^b h(x)\ \mathrm{d}x.$$

Proof:

Since g is defined on the closed interval $[a, b]$, it follows from the Extreme Value Theorem that g takes on a maximum and minimum value on the interval. Let M and m denote this maximum and minimum, respectively. Then, using the linearity of the integral
$$m\int_a^b h(x)\ \mathrm{d}x = \int_a^b m \cdot h(x)\ \mathrm{d}x \leq \int_a^b g(x)h(x)\ \mathrm{d}x \leq \int_a^b M \cdot h(x)\ \mathrm{d}x = M\int_a^b h(x)\ \mathrm{d}x.$$

If $h = 0$, then $0 = m \cdot 0 \leq \int_a^b g(x)h(x)\ \mathrm{d}x \leq M \cdot 0 = 0$, which implies the desired relation holds for each $c \in [a, b]$. So, suppose this is not the case. Then there exists $x^* \in [a, b]$ such that $h(x^*) > 0$. Because h is continuous, there exists $\delta > 0$ such that $|h(x) - h(x^*)| \leq h(x^*)/2$ whenever $|x - x^*| \leq \delta$, which implies $h(x) \geq h(x^*)/2$ whenever $|x - x^*| \leq \delta$. Hence
$$\int_a^b h(x)\ \mathrm{d}x \geq \int_{x^*-\delta}^{x^*+\delta} h(x)\ \mathrm{d}x \geq \int_{x^*-\delta}^{x^*+\delta} \frac{h(x^*)}{2}\ \mathrm{d}x = 2\delta \cdot \frac{h(x^*)}{2} = \delta h(x^*) > 0.$$

Thus, our above relation can be rewritten as
$$m \leq \frac{\int_a^b g(x)h(x)\ \mathrm{d}x}{\int_a^b h(x)\ \mathrm{d}x} \leq M.$$

By the Intermediate Value Theorem, it follows that there exists $c \in [a, b]$ such that
$$g(c) = \frac{\int_a^b g(x)h(x)\ \mathrm{d}x}{\int_a^b h(x)\ \mathrm{d}x},$$

which implies
$$\int_a^b g(x)h(x)\ \mathrm{d}x = g(c)\int_a^b h(x)\ \mathrm{d}x,$$

and we are done. $\square$

F15.3: Let $\{f_n\}$ be a sequence of continuous functions $f_n : [-1,1] \to [0,1]$ such that for each $x \in [-1,1]$,

1) the sequence of numbers $\{f_n(x)\}_{n=1}^{\infty}$ is non-increasing, and

2) $\lim_{n\to\infty} f_n(x) = 0$.

Define

$$g_n(x) := \sum_{m=1}^{n} (-1)^m f_m(x).$$

Prove that $g_n(x)$ converges to some $g(x) \in \mathbb{R}$ for each $x \in [-1,1]$ and that the function $g[-1,1] \to \mathbb{R}$ thus defined is continuous on $[-1,1]$.

Proof:

Let $x \in [-1,1]$. We first show that $\lim_{n\to\infty} g_n(x)$ exists. Since $\mathbb{R}$ is complete, it suffices to show that the sequence $\{g_n(x)\}$ is Cauchy. Let $\varepsilon > 0$ be given and define $A_0 := 0$ and

$$A_m := \sum_{j=1}^{m} (-1)^m \quad \forall\, m \geq 1, \tag{238}$$

noting this implies $|A_m| \leq 1$ for all $m \geq 0$. Then, for $n, p \in \mathbb{Z}^+$ with $n > p$, we have

$$\begin{aligned}
|g_n(x) - g_p(x)| &= \left| \sum_{m=p+1}^{n} (-1)^m f_m(x) \right| \\
&= \left| \sum_{m=p+1}^{n} (A_m - A_{m-1})\, f_m(x) \right| \\
&= \left| \sum_{m=p+1}^{n} A_m f_m(x) - \sum_{m=p}^{n-1} A_m f_{m+1}(x) \right| \\
&= \left| \sum_{m=p+1}^{n-1} A_m \left[f_m(x) - f_{m+1}(x)\right] + A_n f_n(x) - A_p f_{p+1}(x) \right| \qquad (239) \\
&\leq \left| \sum_{m=p+1}^{n-1} A_m \left[f_m(x) - f_{m+1}(x)\right] \right| + |A_n f_n(x)| + |A_p f_{p+1}(x)| \\
&\leq \left| \sum_{m=p+1}^{n-1} f_m(x) - f_{m+1}(x) \right| + |f_n(x)| + |f_{p+1}(x)| \\
&= |f_{p+1}(x) - f_n(x)| + |f_n(x)| + |f_{p+1}(x)| \\
&\leq 4|f_{p+1}(x)|.
\end{aligned}$$

Since $\{f_n\}$ converges to 0, there is $N \in \mathbb{Z}^+$ such that $|f_n - 0| < \varepsilon/4$ for all $n \geq N$. Consequently, for $n, p \geq N$ with $n \geq p$ we have

$$|g_n(x) - g_p(x)| \leq 4|f_{p+1}| < \varepsilon. \tag{240}$$

Thus, $\{g_n(x)\}$ is Cauchy. Since x was arbitrarily chosen in $[-1, 1]$, we may, thus, define a function $g(x) := \lim_{n\to\infty} g_n(x)$ for each x. This shows pointwise convergence of the sequence of functions $\{g_n\}$ to the limit g.

MUST FINISH AND SHOW CONVERGENCE IS UNIFORM SO WE GET THAT g IS CONTINUOUS. $\square$

F15.4 Let $f_n : [0,\infty) \to \mathbb{R}$ be functions defined recursively by $f_1(x) := 0$ and

$$f_{n+1}(x) := e^{-2x} + \int_0^x f_n(t)e^{-2t}\ \mathrm{d}t, \quad n \geq 1. \tag{241}$$

Show that $f(x) := \lim\limits_{n\to\infty} f_n(x)$ exists for all $x \geq 0$ and identify f explicitly.

Proof:

We first show f exists. Define a mapping $T : C[0,\infty) \to C[0,\infty)$ by

$$T(f) := e^{-2x} + \int_0^x f(t)e^{-2t}\ \mathrm{d}t. \tag{242}$$

Since $C[0,\infty)$ is complete, if we can show that T is a contraction, then the Banach Fixed Point theorem states T has a fixed point and that for any $f_1 \in C[0,\infty)$, the sequence defined by

$$f_{n+1} = T(f_n) \quad \forall\ n \geq 1 \tag{243}$$

will converge to a fixed point of T, implying the limit $\lim\limits_{n\to\infty} f_n$ exists. For $f, g \in C[0,\infty)$ we discover

$$\begin{aligned} \|T(f) - T(g)\|_\infty &= \left\| \int_0^x (f(t) - g(t))e^{-2t}\ \mathrm{d}t \right\|_\infty \\ &\leq \left\| \int_0^x \|f-g\|_\infty e^{-2t}\ \mathrm{d}t \right\|_\infty \\ &= \|f-g\|_\infty - \infty \left| \int_0^x e^{-2t}\ \mathrm{d}t \right| \\ &\leq \|f-g\|_\infty - \infty \left| \int_0^\infty e^{-2t}\ \mathrm{d}t \right| \\ &= \frac{1}{2}\|f-g\|_\infty \end{aligned} \tag{244}$$

Hence T is a contraction. Because the sequence defined in (243) is precisely the sequence defined in the problem statement, we see $f(x) := \lim\limits_{n\to\infty} f_n(x)$ exists for all $x \geq 0$.

We now identify f explicitly. Since $f = T(f)$, we differentiate to find

$$\frac{\mathrm{d}f}{\mathrm{d}x} = -2e^{-2x} + f(x)e^{-2x} = (f(x) - 2)e^{-2x} \qquad \Rightarrow \qquad \frac{\mathrm{d}f}{f-2} = e^{-2x}\ \mathrm{d}x. \tag{245}$$

Integrating gives

$$\int_{f(0)}^{f(x)} \frac{\mathrm{d}f}{f-2} = \int_0^x e^{-2x}\ \mathrm{d}x \qquad \Rightarrow \qquad -[f(x) - 2] = \frac{f(x)-2}{f(0)-2} = \exp\left(-\frac{1}{2}\left[e^{-2x} - 1\right]\right) \tag{246}$$

where we know $f(0) = T(f(0)) = e^{-2\cdot 0} + 0 = 1$. Hence

$$\boxed{f(x) = 2 - \exp\left(-\frac{1}{2}\left[e^{-2x} - 1\right]\right)} \tag{247}$$

$\square$

F15.6: Let $X := \mathbb{R}\backslash\{0\}$. Find a metric ρ on X with the following properties:

i) (X, ρ) is a complete metric space, and

ii) if $\{x_n\}_{n=1}^{\infty} \subset X$ and $x \in X$, then

$$\lim_{n\to\infty} |x_n - x| = 0 \quad \Leftrightarrow \quad x_n \to x \text{ in } (X, \rho).$$

Prove both properties, as well as all of your other assertions, in full detail.

Proof:
Define ρ to be the discrete metric, i.e.,

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

This metric satisfies the properties of a metric. Indeed, ρ equal to zero precisely when $x = y$ and is positive otherwise. Further, ρ is symmetric. All that remains is to verify that the triangle inequality holds, i.e., we need

$$\rho(x, z) \leq \rho(x, y) + \rho(y, z).$$

If $x = z$ then this is trivial. Suppose otherwise. Then, if $x = y$, we have $\rho(x, z) = \rho(y, z) = 0 + \rho(y, z) = \rho(x, y) + \rho(y, z)$ and similarly when $y = z$. If $x \neq y \neq z$, then $1 = \rho(x, z) < 2 = \rho(x, y) + \rho(y, z)$. In each case, the triangle inequality holds.

Now we claim that (X, ρ) is complete. Let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence in X. Then there is a $N \in \mathbb{N}$ such that $|x_m - x_n| < \frac{1}{2}$ whenever $m, n \geq N$. With the discrete metric, this occurs precisely when $x_n = x_m$ for all $n, m \geq N$. That is, if $\{x_n\}_{n=1}^{\infty}$ is Cauchy in the discrete metric, then there exists an integer $N \in \mathbb{N}$ at which $x_n = x_N$ for all $n \geq N$. Hence $\{x_n\} \to x_N$, which is in X.

Now suppose $\{x_n\} \subset X$, $x \in X$ and $\lim\limits_{n\to\infty} |x_n - x| = 0$. Then there exists $N \in \mathbb{N}$ such that $|x_n - x| < \frac{1}{2}$ for all $n \geq N$. This only occurs when $x_n = x \ \forall\ n \geq N$, and so $x_n \to x \in (X, \rho)$. Now suppose $x_n \to x$. Then, in the discrete metric, there is a an integer $N \in \mathbb{N}$ at which $x_n = x$ for all $n \geq N$. But, for such n, $\rho(x_n, x) = |x_n - x| = 0$. Hence $\lim\limits_{n\to\infty} |x_n - x| = 0$ and we are done. $\square$

F15.7: Let A, B be two 4×5 matrices of rank 3, and let $C = A^T B$ (this is a 5×5 matrix). Find all possible values r for the rank of C. To be precise, if the rank r is possible, find an explicit example of such matrices. Then prove that all other values are impossible.

Proof:

Note that $A^T : \mathbb{R}^4 \to \mathbb{R}^5$ has rank 3 and so its nullity must be 1 by the rank-nullity theorem. Similarly, since $B : \mathbb{R}^5 \to \mathbb{R}^4$ and the rank of B is 5, B must have nullity 2. This leaves two cases. First, if $(\ker(A)) \cap (\ker(B)) = \emptyset$, then C must have nullity 3. Alternatively, if $(\ker(A)) \cap (\ker(B)) \neq \emptyset$, then $\ker(A) \subset \ker(B)$ since the nullity of A is 1. Then since B has nullity 2, it follows that C has nullity 2 as well.

NOT COMPLETE. NEED TO ADD EXAMPLES. $\square$

F15.8: Find M^{-2} where

$$M = \begin{pmatrix} 2 & 3 & 2 & 1 \\ 3 & 6 & 4 & 2 \\ 4 & 8 & 6 & 3 \\ 2 & 4 & 3 & 1 \end{pmatrix}. \tag{248}$$

Solution:
We find M^{-1} be using elementary row operations, and then square M^{-1} to obtain our result. Note

$$(M \mid I) \sim \left(\begin{array}{cccc|cccc} 2 & 3 & 2 & 1 & 1 & 0 & 0 & 0 \\ 3 & 6 & 4 & 2 & 0 & 1 & 0 & 0 \\ 4 & 8 & 6 & 3 & 0 & 0 & 1 & 0 \\ 2 & 4 & 3 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & -2 \end{array}\right), \tag{249}$$

which implies

$$M^{-1} = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -2 & 1 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix}. \tag{250}$$

Thus,

$$\boxed{M^{-2} = \begin{pmatrix} 5 & -4 & 1 & 0 \\ -4 & 7 & -3 & -1 \\ 2 & -6 & 4 & -1 \\ 0 & -2 & -1 & 5 \end{pmatrix}}. \tag{251}$$

□

F15.9: Let A be an $n \times n$ real matrix such that $A^T = -A$. Prove that $\det(A) \geq 0$.

Proof:

First observe that $\det(A) = \det(A^T) = \det(-A) = (-1)^n \det(A)$. If n is odd, then $\det(A) = -\det(A)$, which implies $\det(A) = 0$. Now suppose n is even and let x be eigenvector of A with eigenvalue λ. Then

$$\lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle Ax, x \rangle = \langle x, -Ax \rangle = \langle x, -\lambda x \rangle = -\overline{\lambda} \langle x, x \rangle,$$

which implies $\lambda = -\overline{\lambda}$ and so λ must be imaginary. That is, for each λ_j there is an $\alpha_j \in \mathbb{R}$ such that $\lambda_j = i\alpha_j$ where $\lambda_1, \ldots, \lambda_n$ lists the eigenvalues of A. Since $\det(A) = \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n$, if $\lambda_j = 0$ for any $j = 1, \ldots, n$, then $\det(A) = 0$ and we are done. So, suppose this is not the case. Then, since eigenvalues come in conjugate pairs and n is even, $\det(A)$ is the product of pairs of eigenvalues λ_j and λ_k with $\lambda_j = \overline{\lambda_k}$ such that $\lambda_j \lambda_k = (-1) i^2 \alpha_j^2 = \alpha_j^2 > 0$, and so $\det(A) > 0$. This completes the proof. $\square$

F15.11: Let $T : V \to V$ be a linear operator such that $T^6 = 0$ and $T^5 \neq 0$. Suppose $V \cong \mathbb{R}^6$. Prove that there is no linear operator $S : V \to V$ such that $S^2 = T$. Does the answer change if $V \cong \mathbb{R}^{12}$.

Proof:

Let $L : V \to V$ be a linear operator. Suppose k is a nonnegative integer and $v \in \text{null } L^k$. Then $L^{k+1}v = L(L^k v) = L(0) = 0$ and so $v \in \text{null } L^{k+1}$. Through induction it follows that

$$\{0\} = \text{null } L^0 \subseteq \text{null } L^1 \subseteq \text{null } L^2 \subset \cdots \subseteq \text{null } L^k \subseteq \text{null } L^{k+1} \subseteq \cdots .$$

Now if for some $m \in \mathbb{N}$ we have null L^m = null L^{m+1}, then we claim L^m = null L^{m+1} = null $L^{m+2} = \cdots$. To show this, we let $k \in \mathbb{N}$ and verify null L^{m+k} = null L^{m+k+1}. We already know null $L^{m+k} \subseteq$ null L^{m+k+1}. So now suppose $v \in$ null L^{m+k+1}. Then $L^{m+k+1}v = L^{m+1}(T^k v) = 0$. This implies $L^k v \in$ null L^{m+1} = null L^m. Thus, $L^{m+k}v = L^m(L^k v) = 0$, which implies $v \in$ null L^{m+k} and so $L^{m+k+1} \subseteq$ null L^{m+k}, proving L^{m+k+1} = null L^{m+k}.

Let $n = \dim(V) = 6$. Then we claim null L^n = null L^{n+1}, which, by the above, implies null L^n = null L^{n+k} for each $k \in \mathbb{N}$. By way of contradiction, suppose this claim does not hold. Then from the above, we have

$$\{0\} = \text{null } L^0 \subsetneq \text{null } L^1 \subsetneq \cdots \subsetneq \text{null } L^n \subsetneq L^{n+1}.$$

At each of the strict inclusions in the chain above, the dimension increases by at least one and so $\dim(\text{null } T^{n+1}) \geq n + 1$, which contradicts the fact that $n = \dim(V)$ and null $T^{n+1} \subseteq V$.

Now suppose we have $S^2 = T$. Then, by the above, null S^6 = null S^{6+k} for each $k \in \mathbb{N}$. However, this implies null S^{10} = null S^{12}. But, $S^{12} = T^6 = 0$ and so null $S^{12} = V$ and null $S^{10} = V$, which implies $S^{10} = T^5 = 0$, a contradiction. Hence such an S does not exists.

Yes, the answer changes if $V \cong \mathbb{R}^{12}$. For there we needn't have $S^{10} = S^{12}$ since $10 < \dim(V)$ and the above argument does not apply. As an example, suppose $e_1, \ldots, e_{12}$ is the standard orthonormal basis for $\mathbb{R}^{12}$. Then define $T(e_j) = e_{j+2}$ when $j + 2 \leq 12$ and $T(e_j) = 0$ when $j \geq 11$. Then defining $S(e_j) = e_{j+1}$ when $j \leq 11$ and $S(e_{12}) = 0$ gives $S^2 = T$. And, $T^5 = S^{10}(e_1) = e_{11} \neq 0$ while $T^6 = S^{12}(e_i) = 0$ for $i = 1, \ldots, 12$. $\square$

F15.12: Prove that the following $n \times n$ matrix M is positive definite:

$$M = \begin{pmatrix} 2 & 1 & 1 & \cdots & 1 \\ 1 & 3 & 1 & \cdots & 1 \\ 1 & 1 & 4 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & n+1 \end{pmatrix}.$$

Proof:
We must show $\langle Mx, x\rangle \geq 0$ for each $x \in V$ where V is our n dimensional vector space. Let N be the $n \times n$ matrix of all 1's and $\Lambda = \text{diag}(1, 2, \dots, n)$. Then $M = \Lambda + N$. Now observe that for each $x \in V$, Nx is equal to the sum of all the entries of x multiplied by the vector of all 1's, i.e.,

$$Nx = \left(\sum_{j=1}^{n} x_j\right) \underbrace{(1, 1, \dots, 1)}_{n \text{ terms}}.$$

Thus,

$$\langle Nx, x\rangle = \left(\sum_{j=1}^{n} x_j\right) \langle (1, 1, \dots, 1), x\rangle = \left(\sum_{j=1}^{n} x_j\right)\left(\sum_{j=1}^{n} x_j\right) = \left(\sum_{j=1}^{n} x_j\right)^2 \geq 0.$$

For the matrix Λ observe that

$$\langle \Lambda x, x\rangle = \sum_{j=1}^{n} j \cdot x_j \cdot x_j = \sum_{j=1}^{n} j \cdot x_j^2 \geq 0.$$

Hence

$$\langle Mx, x\rangle = \langle (N + \Lambda)x, x\rangle = \langle Nx, x\rangle + \langle \Lambda x, x\rangle = \left(\sum_{j=1}^{n} x_j\right)^2 + \sum_{j=1}^{n} j \cdot x_j^2 \geq 0$$

and so M is positive definite. $\square$

2016

S16.1: For $a < b$ real numbers, let $f[a,b] \times [a,b] \to \mathbb{R}$ be such that

a) for each $y \in [a,b]$, $x \mapsto f(x,y)$ is non-increasing and continuous on $[a,b]$,

b) for each $x \in [a,b]$, $y \mapsto f(x,y)$ is non-decreasing and continuous on $[a,b]$.

Prove that $g(x) := f(x,x)$ is continuous on $[a,b]$.

Proof:

To show that $g(x)$ is continuous, it suffices to show that, given $\varepsilon > 0$, at each $x_0 \in [a,b]$ there exists $\delta > 0$ such that $|g(x) - g(x_0)| \leq \varepsilon$ whenever $|x - x_0| < \delta$. Observe from the triangle inequality that

$$\begin{aligned}|g(x) - g(x_0)| &= |f(x,x) - f(x_0,x_0)| \\ &= |f(x,x) - f(x,x_0) + f(x,x_0) - f(x_0,x_0)| \\ &\leq |f(x,x) - f(x,x_0)| + |f(x,x_0) - f(x_0,x_0)|.\end{aligned} \tag{252}$$

Since the mapping $y \mapsto f(x,y)$ is continuous on a compact set $[a,b] \subset \mathbb{R}$, it is uniformly continuous. So, there is a $\delta_1 > 0$ such that $|f(x,x) - f(x,x_0)| \leq \varepsilon/2$ whenever $|x - x_0| < \delta_1$. Similarly, there exists $\delta_2 > 0$ such that $|f(x,x_0) - f(x_0,x_0)| \leq \varepsilon/2$ whenever $|x - x_0| < \delta_2$. Define $\delta := \min\{\delta_1, \delta_2\}$. Then, whenever $|x - x_0| \leq \delta$,

$$|g(x) - g(x_0)| \leq |f(x,x) - f(x,x_0)| + |f(x,x_0) - f(x_0,x_0)| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon, \tag{253}$$

and we are done. □

S16.2: For $a < b$ real numbers and a function $f : [a, b] \to \mathbb{R}$, do as follows:

a) Define what it means for f to be Riemann integrable on $[a, b]$.

b) Let $\{x_n\}_{n=1}^{\infty} \subset [a, b]$ be a sequence such that $\lim_{n\to\infty} x_n$ exists and suppose that $f : [a, b] \to \mathbb{R}$ is defined by

$$f(x) := \begin{cases} 1 & \text{if } x \notin \{x_n\}_{n=1}^{\infty}, \\ 0 & \text{otherwise.} \end{cases} \tag{254}$$

Using your definition, prove that f is Riemann integrable on $[a, b]$.

Proof:

a) We define the upper and lower Riemann sums, respectively, of f with respect to a partition $P = \{I_1, \ldots, I_n\}$ of $[a, b]$ by

$$U(f; P) := \sum_{i=1}^{k} \left(\sup_{I_i} f \right) |I_i| \quad \text{and} \quad L(f; P) := \sum_{i=1}^{k} \left(\inf_{I_i} f \right) |I_i|. \tag{255}$$

Let Π denote the collection of all partitions of $[a, b]$. Then $f : [a, b] \to \mathbb{R}$ is Riemann integrable on $[a, b]$ if it is bounded and

$$\inf_{P\in\Pi} U(f; P) = \sup_{P\in\Pi} L(f; P). \tag{256}$$

b) We claim that f is Riemann integrable on $[a, b]$. Clearly, f is bounded by one. We claim every upper sum equals $b - a$. Indeed, in each interval I_i of a partition P of $[a, b]$, we can pick $z \in I_i$ such that $z \notin \{x_n\}$ since I_i is uncountable. Thus, for any partition $P = \{I_1, \ldots, I_k\}$ we have

$$U(f; P) = \sum_{i=1}^{k} 1 \cdot |I_i| = b - a. \tag{257}$$

Hence $\inf_{P\in\Pi} U(f; P) = b - a$. Now let $\varepsilon > 0$ be given. Then to show $\sup_{P\in\Pi} L(f; P) = b - a$, it suffices to construct a partition P of $[a, b]$ such that $L(f; P) \geq (b - a) - \varepsilon$. We do this as follows.

Since the sequence $\{x_n\}$ converges to a limit x and $[a, b]$ is closed, we have $x \in [a, b]$. Moreover, by the convergence of $\{x_n\}$ we can find an $N \in \mathbb{Z}^+$ such that, by defining $P_N := [x - \varepsilon/4, x + \varepsilon/4] \cap [a, b]$, we have $x_n \in P_N$ for all $n \geq N$. Now, inductively define

$$P_i := \left(x_i - \frac{\varepsilon}{4 \cdot 2^i}, x_i + \frac{\varepsilon}{4 \cdot 2^i} \right) \cap [a, b] \backslash \{P_N, \ldots, P_{i+1}\} \tag{258}$$

for i going from $N-1$ to 1. Then $\inf_{x\in P_i} f = 0$ for $i = 1, \ldots, N$ and

$$\sum_{i=1}^{N} |P_i| = |P_N| + \sum_{i=1}^{N-1} |P_i| = \frac{\varepsilon}{2} + \frac{\varepsilon}{4}\sum_{i=1}^{N-1} \frac{1}{2^i} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{4}\sum_{i=0}^{\infty} \frac{1}{2^i} = \varepsilon. \tag{259}$$

Now we can extend $\{P_i\}_{i=1}^N$ to a partition $P = \{P_1, \ldots, P_N, K_1, \ldots, K_J\}$ of $[a, b]$. This can be done by defining the K_i to be the intervals in $[a, b]\backslash\{P_1, \ldots, P_N\}$. Then

$$L(f; P) = \sum_{i=1}^{J} \left(\inf_{x\in K_i} f\right) |K_i| = \sum_{i=1}^{J} |K_i| = (b-a) - \varepsilon, \tag{260}$$

as desired. Hence f is Riemann integrable on $[a, b]$.

$\square$

S16.3: Suppose $f : [0,1] \to \mathbb{R}$ is a continuously differentiable function. Show that the limit

$$\lim_{n\to\infty} n \left(\sum_{k=0}^{n} f\left(\frac{k}{n}\right) - n \int_0^1 f(x)\ \mathrm{d}x \right) \tag{261}$$

exists and compute its value.

Proof:

We claim the limit actually does not exist, which we will show with a counterexample. Define $f : [0,1] \to \mathbb{R}$ by $f(x) = x$ for $x \in [0,1]$ and note that f is continuously differentiable. Then

$$\begin{aligned} n \left(\sum_{k=0}^{n} f\left(\frac{k}{n}\right) - n \int_0^1 f(x)\ \mathrm{d}x \right) &= n \left(\sum_{k=0}^{n} \frac{k}{n} - \frac{n}{2} \right) \\ &= n \left(\frac{1}{n} \cdot \frac{n(n+1)}{2} - \frac{n}{2} \right) \\ &= \frac{n^2+n}{2} - \frac{n^2}{2} \\ &= \frac{n}{2}. \end{aligned} \tag{262}$$

Hence

$$\lim_{n\to\infty} n \left(\sum_{k=0}^{n} f\left(\frac{k}{n}\right) - n \int_0^1 f(x)\ \mathrm{d}x \right) = \lim_{n\to\infty} \frac{n}{2} = \infty. \tag{263}$$

□

S16.4: Given continuous functions $\alpha : [0,1] \to \mathbb{R}$ and $\beta : [0,1] \to [0,1)$, define functions $f_n : [0,1] \to \mathbb{R}$ by the recursion

$$f_{n+1}(x) = \alpha(x) + \int_0^x \beta(t) f_n(t)\ \mathrm{d}t \tag{264}$$

with $f_0(x) := 0$ for all $x \in [0,1]$. Prove that, for each $x \in [0,1]$, the limit $f(x) := \lim_{n\to\infty} f_n(x)$ exists and compute its value.

Proof:

Consider the metric space $C[0,1]$ with the sup norm, which we shall denote by X. We proceed by showing X is complete, and then we define a contraction mapping $T : X \to X$ with which we are able to apply the Banach Fixed Point theorem.

First we show completeness of X. Let $\{g_n\} \subset X$ be Cauchy and $\varepsilon > 0$ be given. Then there is a $N \in \mathbb{Z}^+$ such that

$$\|g_n - g_m\| \le \varepsilon/3 \quad \forall\ m, n \ge N. \tag{265}$$

Let $x \in [0,1]$. Then the sequence $\{g_n(x)\}$ is Cauchy since $|g_n(x) - g_m(x)| \le \|g_n - g_m\|\ \forall\ m, n$. But, since $\{g_n(x)\} \subset \mathbb{R}$ and $\mathbb{R}$ is complete, this sequence converges to some limit. Thus, we may define a function $g : [0,1] \to \mathbb{R}$ point-wise by $g(x) := \lim_{n\to\infty} g_n(x)$. Then, taking the limit as $n \to \infty$ and using the fact that norms are continuous,

$$\|g - g_m\| \le \varepsilon/3 \quad \forall\ m \ge N, \tag{266}$$

and so $g_n \longrightarrow g$. Now we verify g is continuous. Let $x \in [0,1]$. Since g_N is continuous, there is a $\delta > 0$ such that for $y \in [0,1]$,

$$|x - y| < \delta \quad \Rightarrow \quad |g_N(x) - g_N(y)| \le \varepsilon/3. \tag{267}$$

Thus, whenever $y \in [0,1]$ and $|x - y| \le \delta$,

$$\begin{aligned} |g(x) - g(y)| &\le |g(x) - g_N(x) + g_N(x) - g_N(y) + g_N(y) - g(y)| \\ &\le |g(x) - g_N(x)| + |g_N(x) - g_N(y)| + |g_N(y) - g(y)| \\ &\le \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon. \end{aligned} \tag{268}$$

Thus, $g \in X$ and so X is complete.

Now define the mapping $T : X \to X$ by

$$T(f)(x) = \alpha(x) + \int_0^x \beta(t) f_n(t) \, \mathrm{d}t. \tag{269}$$

Because $[0,1]$ is closed and bounded, the extreme value theorem implies β attains its maximum value β_* on $[0,1]$. Since $\beta([0,1]) \subset [0,1)$, we see $\beta_* < 1$. Then, for $f, g \in X$, observe we discover

$$\begin{aligned} \|T(f) - T(g)\| &= \left\| \int_0^x \beta(t)[f(t) - g(t)] \, \mathrm{d}t \right\| \\ &\leq \beta_* \left\| \int_0^x f(t) - g(t) \, \mathrm{d}t \right\| \\ &\leq \beta_* \|f - g\| \left| \int_0^x \mathrm{d}t \right| \\ &\leq \beta_* \|f - g\| \\ &< \|f - g\|. \end{aligned} \tag{270}$$

Thus, T is a contraction. And because X is complete, the Banach Fixed Point theorem implies that for any $f_0 \in X$, the sequence $f_{n+1} := T(f_n)$ converges to a fixed point of T. Thence the limit $f(x) := \lim\limits_{n \to \infty} f_n(x)$ exists and satisfies

$$f(x) = T(f)(x) = \alpha(x) + \int_0^x \beta(t) f(t) \, \mathrm{d}t. \tag{271}$$

All that remains is to compute f. We proceed in this by assuming α is differentiable, but this is not a necessary assumption our final answer. Differentiating, we see

$$f'(x) = \alpha'(x) + \beta(x) f(x) \quad \Rightarrow \quad f'(x) - \beta(x) f(x) = \alpha'(x). \tag{272}$$

Including an integrating factor, we obtain

$$\mathrm{d}\left[f(x) \exp\left(-\int_x \beta \right) \right] = \alpha'(x) \exp\left(-\int_x \beta \right) \, \mathrm{d}x \tag{273}$$

which implies

$$f(x) = \alpha(x) + \exp\left(\int_x \beta \right) \cdot \int_x \alpha\beta \exp\left(-\int_t \beta \right) \, \mathrm{d}t \tag{274}$$

where the equality follows by integrating each side (the right hand side by parts) and then multiplying by $\exp\left(\int_x \beta\right)$. $\square$

S16.5: Let $f, g \in \mathbb{R}^2 \to \mathbb{R}$ be continuously differentiable functions such that g attains value zero at at least one point. Suppose that $\nabla g \neq 0$ everywhere on $\mathbb{R}^2$ and assume (x_0, y_0) is a point such that

$$f(x_0, y_0) = \inf \{ f(x,y) : x, y \in \mathbb{R}, g(x,y) = 0 \}. \tag{275}$$

Show there is $\lambda \in \mathbb{R}$ such that $\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$.

Proof:

We proceed by identifying $\lambda \in \mathbb{R}$ such that the desired relation holds. Since $\nabla g(x_0, y_0) \neq 0$, we may take $g_y(x_0, y_0) \neq 0$, without loss of generality. Then the Implicit Function Theorem implies there is a neighborhood $N \subset \mathbb{R}$ containing x and continuously differentiable $\phi : N \to \mathbb{R}$ such that

$$y_0 = \phi(x_0) \quad \text{and} \quad g(x, \phi(x)) = 0 \ \text{ for } x \in N. \tag{276}$$

With this, define $\tau : N \to \mathbb{R}$ by $\tau(x) = f(x, \phi(x))$. Since x_0 is an extremal point of τ, by hypothesis, and N is an open neighborhood, we discover

$$0 = \tau'(x_0) = f_x(x_0, \phi(x_0)) + f_y(x_0, \phi(x_0))\phi'(x_0). \tag{277}$$

But, we also know $0 = g(x_0, y_0) = g(x_0, \phi(x_0))$ and so

$$0 = g_x(x_0, y_0) + g_y(x_0, y_0)\phi'(x_0) \quad \Rightarrow \quad \phi'(x_0) = -\frac{g_x(x_0, y_0)}{g_y(x_0, y_0)} \tag{278}$$

where the division is well-defined since $g_y(x_0, y_0) \neq 0$. Plugging (278) into (277), we see

$$f_x(x_0, y_0) = \frac{g_x(x_0, y_0) f_y(x_0, y_0)}{g_y(x_0, y_0)}. \tag{279}$$

Take $\lambda := f_y(x_0, y_0)/g_y(x_0, y_0)$ so that $f_y(x_0, y_0) = \lambda g_y(x_0, y_0)$. With this choice of λ, (279) becomes $f_x(x_0, y_0) = \lambda g_x(x_0, y_0)$. Thus, we have identified λ yielding $\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$. This completes the proof.[2] $\square$

[2]For a proof in a more general context, see pp. 465-466 of Fitzpatrick [1].

S16.6: A metric ρ in a metric space (X, ρ) is said to be an *ultrametric* if

$$\forall\ x, y, z \in X, \quad \rho(x, y) \leq \max\{\rho(x, z), \rho(y, z)\}. \tag{280}$$

Prove that, in this metric, every open ball $\{y : \rho(x, y) < r\}$ is closed and every closed ball $\{y : \rho(x, y) \leq r\}$ is open.

Proof:
Let B_x^0 be the open ball of radius r centered at x and B_x be the corresponding closed ball. Also suppose $y \in (B_x^0)^c$. We claim $B_y^0 \subset (B_x^0)^c$, from which it follows that $(B_x^0)^c$ is open and, equivalently, that B_x^0 is closed. To see this, suppose, by way of contradiction, there is $z \in B_y^0 \cap B_x^0$. Then

$$\rho(x, y) \leq \max\{\rho(x, z), \rho(y, z)\} < r. \tag{281}$$

But, we have $\rho(x, y) \geq r$ since $y \in (B_x^0)^c$, a contradiction. Thus, $B_y^0 \subset (B_x^0)^c$.

We now show B_x is open. Let $y \in B_x$. We claim $B_y^0 \subset B_x$. Indeed, let $z \in B_y^0$. Then

$$\rho(z, x) \leq \max\{\rho(y, z), \rho(z, y)\} \leq r, \tag{282}$$

and so $z \in B_x$. Hence B_x is open. $\square$

S16.7: An orthogonal $n \times n$ matrix A is called *elementary* if the corresponding linear transformation $L_A : \mathbb{R}^n \to \mathbb{R}^n$ fixes an $(n-2)$-dimensional subspace. Prove that every orthogonal matrix M is a product of at most $(n-1)$ elementary orthogonal matrices.

Proof:
By way of contradiction, suppose there is an orthogonal $n \times n$ matrix M that is the product of n elementary orthogonal matrices..... $\square$

S16.8: Let $A = (a_{ij})$ be a 2×2 real matrix such that

$$a_{11}^2 + a_{12}^2 + a_{21}^2 + a_{22}^2 < \frac{1}{10}. \tag{283}$$

Prove that $(I + A)$ is invertible.

Proof:

To prove $(I + A)$ is invertible, it suffices to show that $\det(I + A) \neq 0$. So, observe that

$$\begin{aligned}
\det(I + A) &= \begin{vmatrix} 1 + a_{11} & a_{12} \\ a_{21} & 1 + a_{22} \end{vmatrix} \\
&= (1 + a_{11})(1 + a_{22}) - a_{12}a_{21} \\
&= 1 + a_{11} + a_{22} + a_{11}a_{22} - a_{12}a_{21} \\
&\geq 1 - |a_{11}| - |a_{22}| - |a_{11}||a_{22}| - |a_{12}||a_{21}| \\
&\geq 1 - \frac{1}{\sqrt{10}} - \frac{1}{\sqrt{10}} - \frac{1}{10} - \frac{1}{10} \\
&= \frac{8 - 2\sqrt{10}}{10} \\
&> 0,
\end{aligned} \tag{284}$$

where we have used the triangle inequality and the fact that $8 - 2\sqrt{10} > 0$ since $8^2 = 64 > 40 = (2\sqrt{10})^2$. Thus, $\det(I + A) \neq 0$. $\square$

S16.9: Let $v_1 = (0,1,x)$, $v_2 = (1,x,1)$, $v_3 = (x,1,0)$. Find all $x \in \mathbb{R}$ for which $\{v_1, v_2, v_3\}$ are linearly independent over $\mathbb{R}$. Similarly, find all $x \in \mathbb{R}$ for which $\{v_1, v_2, v_3\}$ are linearly independent over $\mathbb{Q}$.

Proof:

Let $A = [v_3\ v_2\ v_1]$. Then

$$\det A = \begin{vmatrix} x & 1 & 0 \\ 1 & x & 1 \\ 0 & 1 & x \end{vmatrix} = x \begin{vmatrix} x & 1 \\ 1 & x \end{vmatrix} - 1 \begin{vmatrix} 1 & 0 \\ 1 & x \end{vmatrix} = x(x^2-1) - 1(x-0) = x(x^2-2). \tag{285}$$

We know that the columns of A are independent over $\mathbb{R}$ if $\det(A) \neq 0$. Hence the set of all points for which these columns are linearly independent over $\mathbb{R}$ is given by

$$\{x \in \mathbb{R} \ \mid \ x \neq 0 \wedge x \neq \pm\sqrt{2}\}. \tag{286}$$

Showing linear independence over $\mathbb{Q}$ is slightly more involved. Of course, since $0 \in \mathbb{Q}$, by (285) we know $x \neq 0$. All that remains is to check for when $x \in \{-\sqrt{2}, \sqrt{2}\}$. We claim such choice of x yields that $\{v_1, v_2, v_3\}$ are linearly independent over $\mathbb{Q}$. To see this, we proceed by way of contradiction. So, suppose $x \in \{-\sqrt{2}, \sqrt{2}\}$ and $\{v_1, v_2, v_3\}$ is a linearly dependent set over $\mathbb{Q}$. Then there exists scalars $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{Q}$ not all zero such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0. \tag{287}$$

Since none of the v_i is a scalar multiple of another v_j with $i \neq j$, each α_i is nonzero. We can write

$$v_3 = \frac{\alpha_2}{\alpha_3} v_2 + \frac{\alpha_1}{\alpha_3} v_1. \tag{288}$$

This implies

$$\pm\sqrt{2} = \frac{\alpha_2}{\alpha_3} \cdot 1 + \frac{\alpha_1}{\alpha_3} \cdot 0 = \frac{\alpha_2}{\alpha_3}. \tag{289}$$

But $\alpha_2/\alpha_3 \in \mathbb{Q}$ while $\pm\sqrt{2} \notin \mathbb{Q}$, which gives our desired contradiction. $\square$

S16.10: Let S be a subset of $Mat(3,\mathbb{C})$, the set of 3×3 matrices over $\mathbb{C}$. The set S is called *dense* if every matrix in $Mat(3,\mathbb{C})$ is a limit of a sequence of matrices in S.

a) Prove that the set of matrices with distinct eigenvalues is dense in $Mat(3,\mathbb{C})$.

b) Prove that the sequence of matrices with one Jordan block is not dense in $Mat(3,\mathbb{C})$.

Proof:

a) We prove the theorem for the general case of $Mat(n,\mathbb{C})$, from which the result for $n=3$ follows. Let $A\in Mat(n,\mathbb{C})$ and $\varepsilon>0$. We shall find a sequence $\{A^k\}_{k=1}^\infty$ with limit A. We do this using the norm defined by $\|A\| := \max_{i,j=1,\dots,n}\{|A_{ij}|\}$. We must find $N\in\mathbb{N}$ such that

$$\|A^k - A\| < \varepsilon \quad \forall\ k\geq N. \tag{290}$$

Let J be the Jordan canonical form of A so that there is an invertible matrix P with $A = PJP^{-1}$. Then the eigenvalues $\lambda_1,\dots,\lambda_n$ of A are along the diagonal of J.

We now construct n sequences $\{\delta_j^k\}_{k=1}^\infty$ such that, for fixed k, each $\lambda_j+\delta_j^k$ is distinct and with the property that $\delta_j^k \longrightarrow 0$ as $k\longrightarrow\infty$. First let $\delta_1^k := 1/k$. Then, inductively, define the sequence $\{\delta_j^k\}$ by $\delta_j^k := 1/(m_{jk}k)$ where m_{jk} is defined to be the smallest integer in $\{1,2,\dots,j\}$ such that $\lambda_j+\delta_j^k \neq \lambda_i+\delta_i^k$ for $i=1,\dots,j-1$. Now define the matrix $\Lambda_k\in Mat(n,\mathbb{C})$ by $\Lambda^k := \operatorname{diag}(\delta_1^k,\dots,\delta_n^k)$. We claim $A^k := P(J+\Lambda^k)P^{-1}$ converges to A. Observe that

$$\|A^k - A\| = \|PJP^{-1} - P(J+\Lambda^k)P^{-1}\| = \|P\Lambda^kP^{-1}\| \leq \|P\|\|\Lambda^k\|\|P^{-1}\|. \tag{291}$$

By construction of Λ^k and each δ_j^k, we have $\|\Lambda^k\|\leq 1/k$. And, the Archimedean property of $\mathbb{R}$ implies there is $N\in\mathbb{Z}^+$ such that $1/N \leq \varepsilon/\|P\|\|P^{-1}\|$. Thence

$$\|A^k - A\| \leq \frac{1}{k}\cdot\|P\|\|P^{-1}\| \leq \varepsilon \quad \forall\ k\geq N, \tag{292}$$

as desired.

(continued on next page)

b) Consider the matrix $A := \text{diag}(1, 2, \ldots, n) \in Mat(n, \mathbb{C})$. This is a diagonal matrix with eigenvalues 1,2,$\ldots, n$. By way of contradiction, suppose $\{A^k\} \subset Mat(n, \mathbb{C})$ and $A^k \longrightarrow A$ as $k \longrightarrow \infty$, and A^k has one Jordan block J_λ^k with eigenvalue λ_k. By hypothesis, J_λ^k converges, and so it must converge point-wise, which implies there is a λ such that $\lambda_k \longrightarrow \lambda^*$. And,

$$\begin{aligned} \forall\ k \in \mathbb{Z}^+, \quad \lim_{k\to\infty} \det(A^k - \lambda I) &= \lim_{k\to\infty} \det(P J_\lambda^k P^{-1} - \lambda I) \\ &= \lim_{k\to\infty} \det(J_\lambda^k - \lambda I) \\ &= \lim_{k\to\infty} (\lambda_k - \lambda)^n \\ &= (\lambda^* - \lambda)^n. \end{aligned} \tag{293}$$

And, since the determinant is a polynomial of its entries, it is continuous and

$$\lim_{k\to\infty} \det(A^k - \lambda I) = \det\left(\lim_{k\to\infty} A^k - \lambda I\right) = \det(A - \lambda I) = \prod_{i=1}^{n} (i - \lambda). \tag{294}$$

But, this implies

$$\prod_{i=1}^{n} (i - \lambda) = (\lambda^* - \lambda)^n, \tag{295}$$

a contradiction.

□

S16.12: Let A be a symmetric $n \times n$ real matrix, $n \geq 4$, and let $v_1, \ldots, v_4 \in \mathbb{R}^n$ be nonzero vectors. Suppose $Av_k = (2k-1)v_k$ for $k = 1, \ldots, 4$. Prove that $v_1 + 2v_2$ is orthogonal to $3v_3 + 4v_4$.

Proof:
We must show $\langle v_1 + 2v_2, 3v_3 + 4v_4 \rangle = 0$. Using the linearity of the scalar product in $\mathbb{R}^n$, we see

$$\begin{aligned} \langle v_1 + 2v_2, 3v_3 + 4v_4 \rangle &= \langle v_1, 3v_3 \rangle + \langle v_1, 4v_4 \rangle + \langle 2v_2, 3v_3 \rangle + \langle 2v_2, 4v_4 \rangle \\ &= 3 \langle v_1, v_3 \rangle + 4 \langle v_1, v_4 \rangle + 6 \langle v_2, v_3 \rangle + 8 \langle v_2, v_4 \rangle . \end{aligned} \tag{296}$$

Using the fact that A is symmetric, we see

$$\langle v_1, v_3 \rangle = \langle Av_1, v_3 \rangle = \langle v_1, Av_3 \rangle = \langle v_1, 5v_3 \rangle = 5 \langle v_1, v_3 \rangle , \tag{297}$$

which implies $\langle v_1, v_3 \rangle = 0$. Also,

$$\langle v_2, v_3 \rangle = \frac{1}{3} \langle Av_2, v_3 \rangle = \frac{1}{3} \langle v_2, Av_3 \rangle = \frac{1}{3} \langle v_2, 5v_3 \rangle = \frac{5}{3} \langle v_2, v_3 \rangle , \tag{298}$$

and so $\langle v_2, v_3 \rangle = 0$. In similar fashion, we find $\langle v_1, v_4 \rangle = 0$ and $\langle v_2, v_4 \rangle = 0$. Thus,

$$\langle v_1 + 2v_2, 3v_3 + 4v_4 \rangle = 3 \cdot 0 + 4 \cdot 0 + 6 \cdot 0 + 8 \cdot 0 = 0, \tag{299}$$

as desired. □

F16.10: Find the unique point $(x, y) \in \mathbb{R}^2$ on the curve $x^4 + y^4 = 2$ that is closest to the line $y = x - 100$. Note: Formal calculations alone do not constitute a solution. you must justify rigorously that there is a point that is closest, that it is unique, and that it is the specific point you claim it is.

Proof:
We claim the point $(1, -1)$ is closest to the line $L = \{(x, y) \in \mathbb{R}^2 \ : \ y = x - 100\}$. Let $L_{\bar{x},\bar{y}}$ be the line containing $(\bar{x}, \bar{y})$ that is orthogonal to the line L. Then the distance between $(\bar{x}, \bar{y})$ and the line L is given by the distance between $(\bar{x}, \bar{y})$ and the point in the intersection of L and $L_{\bar{x},\bar{y}}$. Since the line $L_{\bar{x},\bar{y}}$ is orthogonal to $y = x - 100$, it has slope -1. Using the point-slope formula, we find

$$L_{\bar{x},\bar{y}} = \{(x, y) \ : \ y - \bar{y} = -(x - \bar{x})\}. \tag{300}$$

The point (x, y) in the intersection of $L_{\bar{x},\bar{y}}$ and $y = x - 100$ gives

$$-(x - \bar{x}) + \bar{y} = y = x - 100 \quad \Rightarrow \quad x = \frac{\bar{x} + \bar{y}}{2} + 50 \quad \Rightarrow \quad y = \frac{\bar{x} + \bar{y}}{2} - 50. \tag{301}$$

The Euclidean distance between $(\bar{x}, \bar{y})$ is, therefore,

$$\begin{aligned} \sqrt{\left(\bar{x} - \left[\frac{\bar{x} + \bar{y}}{2} + 50\right]\right)^2 + \left(\bar{y} - \left[\frac{\bar{x} + \bar{y}}{2} - 50\right]\right)^2} &= \sqrt{\left[\frac{\bar{x} - \bar{y}}{2} - 50\right]^2 + \left[\frac{\bar{y} - \bar{x}}{2} + 50\right]^2} \\ &= \sqrt{2}\left|\frac{\bar{x} - \bar{y}}{2} + 50\right|. \end{aligned} \tag{302}$$

Define a function $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x, y) = 2\left(\frac{x - y}{2} + 50\right)^2 \tag{303}$$

so that f gives the square of the distance from a point (x, y) to L. Since x^2 is strictly convex, our problem may be stated as

$$\min_{x \in \mathbb{R}^2} f(x) \quad \text{such that} \quad g(x) = 0 \tag{304}$$

where $g(x) = x^4 + y^4 - 2$ is the constraint function. Note the set of points for which $g(x) = 0$ is nonempty since $g(1, -1) = 1^4 + 1^4 - 2 = 1 + 1 - 2 = 0$. And, f is nonnegative (identically zero in L and positive elsewhere) and continuous. So, the problem does admit a solution. Now, Lagrange's theorem states that if f and g continuous differentiable and that an extremum (a, b) of f subject to the constraint $g = 0$ we have $\nabla g(a, b) \neq (0, 0)$, there is $\lambda \in \mathbb{R}$ such that

$$\nabla f(a, b) = \lambda \nabla g(a, b). \tag{305}$$

Indeed,

$$\nabla f(x,y) = 2\left(\frac{x-y}{2} + 50\right)(1,-1) \quad \text{and} \quad \nabla g(x,y) = (4x^3, 4y^3), \tag{306}$$

each of which have continuous partial derivatives. And, $\nabla g(x,y) = 0$ iff $(x,y) = (0,0)$, which is not possible for $g(x,y) = 0$. And, because f is strictly convex, any extremum is a unique minimizer. Thence

$$2(b-a, a-b) = \nabla f(a,b) = \lambda \nabla g(a,b) = 4\lambda(a^3, b^3). \tag{307}$$

This implies

$$4\lambda a^3 = 2(b-a) = -2(a-b) = -(4\lambda b^3). \tag{308}$$

If $\lambda = 0$, then $\nabla f(a,b) = 0$. However, f attains its global minimum (zero) along points in L. And, by the constraint $g(a,b) = 0$, we know $(a,b) \notin L$. So, we take $\lambda \neq 0$ to obtain $a^3 = -b^3$, from which we deduce $a = -b$. Then

$$0 = g(a,b) = a^4 + b^4 - 2 = a^4 + a^4 - 2 \quad \Rightarrow \quad a^4 = 1 \quad \Rightarrow \quad a = 1. \tag{309}$$

Thence we conclude $(1,-1)$ is the unique point in $\mathbb{R}^2$ on the curve $g = 0$ that is closest to L. $\square$

4.2 Other Good Problems

Let $A = (a_{ij})$ be a 2×2 real matrix such that

$$a_{11}^2 + a_{12}^2 + a_{21}^2 + a_{22}^2 < \frac{1}{10}. \tag{310}$$

Prove that $(I - A)$ is invertible.

Proof:

Recall that $(I - A)$ is invertible iff $\det(I - A) \neq 0$. And, $\det(I - A) = 0$ iff 1 is an eigenvalue of A for some unit vector v. In such a case, $\|A\| \geq \|Av\| = \|1v\| = \|v\| \geq 1$. So, it suffices to show $\|A\| < 1$. Let us denote $a_{11} = a$, $a_{21} = c$, $a_{21} = b$, and $a_{22} = c$. Then for (x, y) with $x^2 + y^2 = 1$ we have

$$\begin{aligned}
\left\| \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right\|^2 &= \left\| \begin{pmatrix} ax + cy \\ bx + dy \end{pmatrix} \right\|^2 \\
&= \langle (a, c), (x, y) \rangle^2 + \langle (b, d), (x, y) \rangle^2 \\
&\leq \|(a, c)\|^2 \cancelto{1}{\|(x, y)\|^2} + \|(b, d)\|^2 \cancelto{1}{\|x, y\|^2} \\
&= (a^2 + c^2) + (b^2 + d^2) \\
&< 1/10.
\end{aligned}$$

Thus, $\|A\| \leq 1/\sqrt{10} < 1$ and we are done. $\square$

a) Show that for any $n \times m$ matrix A, the dimension of the span of the rows is equal to the dimension of the span of the columns.

b) Show that the dimensions in a) equal the size of the largest submatrix of A that is square with nonzero determinant.

Proof:

a) Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be the linear map associated with A and $v \in \ker T$. Then for each $w \in V$ $\langle Tv, w \rangle = \langle 0, w \rangle = 0$. But, by definition of the adjoint T^* of T, $\langle Tv, w \rangle = \langle v, T^* w \rangle$ and so $\langle v, T^* w \rangle = 0$ for all w. This is logically equivalent to saying $v \in (\text{im } T^*)^\perp$. Hence $\ker T = (\text{im } T^*)^\perp$. But, $\dim((\text{im } T^*)^\perp) = \dim(\mathbb{R}^n) - \dim(\text{im } T^*)$. Thence $\dim(\text{im } T^*) = \dim(\mathbb{R}^n) - \dim(\ker T)$. This implies

$$\dim(\text{row } A) = \dim(\text{col } A^t) = \dim(\text{im } T^*) = \dim(\mathbb{R}^n) - \dim(\ker T) = \dim(\text{im } T) = \dim(\text{col } A).$$

b) We can row reduce A such that the first $k = \dim(\text{col } A)$ columns are nonzero with the rest of the columns zero, i.e., we can row reduce A to obtain $A \sim B = (v_1\ v_2\ \cdots\ v_k\ 0\ \cdots\ 0)$ for column vectors $v_1, \ldots, v_k$ where the v_i are linearly independent and form a basis for col B. Note that this is the largest number of linearly independent column vectors we can have. Then, using the result of a), the submatrix $B' = (v_1 \ \cdots \ v_k)$ must have k linearly independent rows so that B' can be row reduced to obtain a $m \times k$ matrix $(v_1' \ \cdots \ v_k')$ where the each entry in v_i from $k+1$ to n is zero. Then the submatrix $k \times k$ submatrix $(v_1' \ \cdots \ v_k')$ has linearly independent columns and rows. Note well that, by construction, this matrix is similar to the largest $k \times k$ submatrix of A with linearly independent columns. Since the columns of this submatrix are independent, it is invertible. And, since this is similar to a $k \times k$ submatrix in A, that matrix is also invertible and hence has determinant zero. Since a matrix is invertible iff it has linearly independent columns, this is the largest size of invertible submatrix of A.

$\square$

Suppose $T \in \mathcal{L}(V)$ is self-adjoint, $\lambda \in F$, and $\varepsilon > 0$. Suppose also there is $v \in V$ such that $\|v\| = 1$ and $\|Tv - \lambda v\| < \varepsilon$. Prove T has an eigenvalue λ' such that $|\lambda - \lambda'| < \varepsilon$.[3]

Proof:
Since T is self-adjoint, the spectral theorem implies V has an orthonormal basis of eigenvectors $e_1, \ldots, e_n$ of T with eigenvalues $\lambda_1, \ldots, \lambda_n$. So, there are unique scalars $a_1, \ldots, a_n \in F$ such that $v = a_1 e_1 + \cdots + a_n v_n$ and $1 = \|v\|^2 = |a_1|^2 + \cdots + |a_n|^2$. Then note

$$Tv - \lambda v = (T - \lambda I) \sum_{i=1}^{n} a_i e_i = \sum_{i=1}^{n} a_i (\lambda_i - \lambda) e_i, \tag{311}$$

If T has no eigenvalue λ' with $|\lambda - \lambda'| < \varepsilon$, then, using the fact that the e_i are orthonormal,

$$\|Tv - \lambda v\|^2 = \sum_{i=1}^{n} |a_i|^2 |\lambda_i - \lambda|^2 \geq \sum_{i=1}^{n} |a_i|^2 \varepsilon^2 = \varepsilon^2 \quad \Rightarrow \quad \|Tv - \lambda v\| \geq \varepsilon, \tag{312}$$

which contradicts our hypothesis. Hence T must have an eigenvalue λ' such that $|\lambda - \lambda'| < \varepsilon$. $\square$

[3]This comes from Axler Problem 12 page 224.

References

[1] Patrick M. Fitzpatrick. *Advanced Calculus*, volume II. American Mathematical Society, 2nd edition, 2009.

[2] Erwin Kreyszig. *Introductory Functional Analysis*. John Wiley & Sons (Asia), 1989.

[3] Maxwell Rosenlicht. *Introduction to Analysis*. Dover, 1968.

[4] Walter Rudin. *Real and Complex Analysis*. McGraw Hill Education (India), 3rd edition, 2006.

[5] Walter Rudin. *Principles of Mathematatial Analysis*. McGraw Hill Education (India), 3rd edition, 2013.

[6] Terrence Tao. *Analysis*, volume I. Hindustan Book Agency (India), 3rd edition, 2014.

[7] Terrence Tao. *Analysis*, volume II. Hindustan Book Agency (India), 3rd edition, 2014.