I do not believe that my quantitative GRE score accurately represents my true abilities in math. Fast computation has never been one of my strong suits, and for a very long time, that led me to believe I was bad at math. When I have the opportunity to understand a problem and its solution(s), I enjoy math and perform exceptionally well, as evidenced by A grades in real analysis (at MIT), linear algebra (largely proofs-based), proof-writing/foundations of mathematics, vector spaces (taught by an acclaimed mathematician and Abel Prize committee member), and differential equations. I am including this supplement as evidence of my abilities in math and statistics so that my claim is not based solely on transcript figures.

Below are some proofs I have written as part of coursework and for fun. Thank you for your consideration.

Contents

1 Linear Algebra Final Project

The following are two of the Putnam problems I did as part of the final project/problem portfolio for Math 221, Linear Algebra, in undergrad at Emory (taught by Dylanger Pittman). These problems were completed independently and without outside assistance.

1991-A-2. A and B are $n \times n$ matrices with real entries such that $A \neq B$.
If $A^3 = B^3$ and $A^2B = B^2A$, can $A^2 + B^2$ be invertible?

I claim $A^2 + B^2$ cannot be invertible.

Proof: For the sake of contradiction, assume $A^2 + B^2$ has an inverse $(A^2 + B^2)^{-1}$ such that

$$(A^2 + B^2)(A^2 + B^2)^{-1} = I$$

We have

$$A^3 = B^3 \iff A^3 - B^3 = 0$$

$$A^2B = B^2A \iff A^2B - B^2A = 0$$

Therefore,

$$A^3 - B^3 = A^2B - B^2A$$

$$\iff A^3 + B^2A = B^3 + A^2B$$

$$\iff A(A^2 + B^2) = B(A^2 + B^2)$$

Because $A^2 + B^2$ is invertible,

$$A(A^2 + B^2)(A^2 + B^2)^{-1} = B(A^2 + B^2)(A^2 + B^2)^{-1}$$

$$\iff AI = BI$$

$$\iff A = B$$

We know that $A \neq B$, so this is a contradiction. Therefore, $A^2 + B^2$ cannot be invertible. ∎

2014-A-2. Let A be the $n \times n$ matrix whose entry in the i-th row and j-th column is

$$\frac{1}{\min(i, j)}$$

for $1 \leq i, j \leq n$. Compute $\det(A)$.

I claim $\det(A) = \frac{(-1)^{n-1}}{n!(n-1)!}$ and proceed by induction.

Base Case: Let $n = 2$. Then

$$A = \begin{bmatrix} 1 & 1 \\ 1 & \frac{1}{2} \end{bmatrix} \implies \det(A) = \begin{vmatrix} 1 & 1 \\ 1 & \frac{1}{2} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

Using the hypothesis:

$$\det(A) = -\frac{1}{2!} = -\frac{1}{2}$$

It is clear that the hypothesis holds true for the case of $n = 2$.

Induction hypothesis: Assume the hypothesis holds for $n = k$. Then we will have

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \frac{1}{2} & \dots & \frac{1}{2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \frac{1}{2} & \dots & \frac{1}{k} \end{bmatrix}$$

Subtract each row i from the row $i+1$ for all $1 \leq i < k$. We get

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ 0 & -\frac{1}{2} & \dots & -\frac{1}{2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{k} - \frac{1}{k-1} \end{bmatrix}$$

This matrix is an upper triangular matrix whose determinant is equal to $\det(A)$. And

$$\det(A) = \frac{(-1)^{k-1}}{k!(k-1)!}$$

Inductive step: When $n = k + 1 = m$. We will have

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & \frac{1}{2} & \dots & \frac{1}{2} & \frac{1}{2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \frac{1}{2} & \dots & \frac{1}{k} & \frac{1}{k} \\ 1 & \frac{1}{2} & \dots & \frac{1}{k} & \frac{1}{m} \end{bmatrix}$$

Subtract each row i from the row $i+1$ for all $1 \leq i < m$ to get

$$\begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ 0 & -\frac{1}{2} & \dots & -\frac{1}{2} & -\frac{1}{2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{k} - \frac{1}{k-1} & \frac{1}{k} - \frac{1}{k-1} \\ 0 & 0 & \dots & 0 & \frac{1}{m} - \frac{1}{k} \end{bmatrix}$$

Expanding along the mth row (filled with zeros except for in the mth column), we get that

$$\det(A) = \begin{vmatrix} 1 & 1 & \dots & 1 & 1 \\ 0 & -\frac{1}{2} & \dots & -\frac{1}{2} & -\frac{1}{2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{k} - \frac{1}{k-1} & \frac{1}{k} - \frac{1}{k-1} \\ 0 & 0 & \dots & 0 & \frac{1}{m} - \frac{1}{k} \end{vmatrix} = (-1)^{2m} \cdot \left(\frac{1}{m} - \frac{1}{k}\right) \cdot \begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & -\frac{1}{2} & \dots & -\frac{1}{2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{k} - \frac{1}{k-1} \end{vmatrix}$$

Note that

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ 0 & -\frac{1}{2} & \dots & -\frac{1}{2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{k} - \frac{1}{k-1} \end{bmatrix}$$

is the $k \times k$ matrix resulting from subtracting row i from row $i+1$ for all $1 \le i < k$ in the $k \times k$ matrix whose (i,j) entries are each

$$\frac{1}{\min(i,j)}$$

Then from our induction hypothesis, we know that

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & -\frac{1}{2} & \dots & -\frac{1}{2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{k} - \frac{1}{k-1} \end{vmatrix} = \frac{(-1)^{k-1}}{k!(k-1)!}$$

Therefore,

$$\begin{aligned} (-1)^{2m} \cdot \left(\frac{1}{m} - \frac{1}{k}\right) \cdot \begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & -\frac{1}{2} & \dots & -\frac{1}{2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{k} - \frac{1}{k-1} \end{vmatrix} \\ &= \left(\frac{1}{m} - \frac{1}{k}\right) \cdot \frac{(-1)^{k-1}}{k!(k-1)!} \\ &= \frac{k-m}{mk} \cdot \frac{(-1)^{k-1}}{k!(k-1)!} \\ &= \frac{-1}{k(k+1)} \cdot \frac{(-1)^{k-1}}{k!(k-1)!} \\ &= \frac{(-1)^k}{(k+1)k!k(k-1)!} \\ &= \frac{(-1)^k}{(k+1)!k!} \end{aligned}$$

Our hypothesis holds for the $n = k+1$ case. Then by induction, we have

$$\det(A) = \frac{(-1)^{n-1}}{n!(n-1)!}$$

for any $n \times n$ matrix A whose (i,j) entries are each

$$\frac{1}{\min(i,j)}$$

for $1 \le i, j \le n$. ■

2 Foundations of Mathematics

The following are a small selection of proofs I wrote for homework in Math 250, Foundations of Mathematics (an introductory proof-writing class), at Emory, taught by David Zureick-Brown. These problems were completed independently and without outside assistance.

PS2, #3. Show that for all integers a and b, $a^2b^2(a^2-b^2)$ is divisible by 12.

Proof: Suppose $a, b \in \mathbb{Z}$. We can consider $a^2b^2(a^2-b^2)$ in two cases: (i) when at least one of a and b is even and (ii) when both a and b are odd.

(i) Without loss of generality, suppose a is even and b is either even or odd. If a is even, then $\exists\ n \in \mathbb{Z}$ s.t. $a = 2n$. Then $a^2 = 4n^2$, so $4 \mid a^2$. We know that $4 \mid a^2$, and $a^2 \mid a^2b^2(a^2-b^2)$, so by transitivity, we know that $4 \mid a^2b^2(a^2-b^2)$.

(ii) Suppose a and b are both odd. Then $\exists\ n, m \in \mathbb{Z}$ s.t. $a = 2n+1, b = 2m+1$. Note that $a^2 - b^2 = (a-b)(a+b)$, which we can rewrite as $(2n+1-2m-1)(2n+1+2m+1) = 2(n-m)(2)(n+m+1) = 4(n-m)(n+m+1)$. Because 4 is a factor of a^2-b^2, we know that $4 \mid a^2-b^2 \wedge a^2-b^2 \mid a^2b^2(a^2-b^2) \Rightarrow 4 \mid a^2b^2(a^2-b^2)$ by transitivity.

Because $4 \mid a^2b^2(a^2-b^2)$ in both cases, we know that $4 \mid a^2b^2(a^2-b^2)\ \forall\ a, b \in \mathbb{Z}$.

Becase we know that $4 \mid a^2b^2(a^2-b^2)$, we can prove that $12 \mid a^2b^2(a^2-b^2)$ by showing that $3 \mid a^2b^2(a^2-b^2)$ is also true $\forall\ a, b \in \mathbb{Z}$. We can consider $a^2b^2(a^2-b^2)$ in three cases:

(i) When $3 \mid a \vee 3 \mid b$

(ii) When $\exists\ n \in \mathbb{Z}$ s.t. $a = 3n+1 \wedge\ \exists\ m \in \mathbb{Z}$ s.t. $b = 3m+1$, or when $\exists\ n \in \mathbb{Z}$ s.t. $a = 3n+2 \wedge\ \exists\ m \in \mathbb{Z}$ s.t. $b = 3m+2$

(iii) When $\exists\ n \in \mathbb{Z}$ s.t. $a = 3n+2 \wedge\ \exists\ m \in \mathbb{Z}$ s.t. $b = 3m+1$, or when $\exists\ n \in \mathbb{Z}$ s.t. $a = 3n+1 \wedge\ \exists\ m \in \mathbb{Z}$ s.t. $b = 3m+2$.

(i) Without loss of generality, suppose $3 \mid a$. We already know that $a \mid a^2b^2(a^2-b^2)$, so by transitivity, $3 \mid a^2b^2(a^2-b^2)$.

(ii) Suppose $\exists\ n, m \in \mathbb{Z}$ and $c \in \{1, 2\}$ s.t. $a = 3n+c, b = 3m+c$. Then $a - b = 3n + c - 3m - c = 3(n-m)$. We already know that $a - b \mid a^2b^2(a^2-b^2)$ and now that $3 \mid a-b$. By transitivity, $3 \mid a^2b^2(a^2-b^2)$.

(iii) Without loss of generality, suppose $\exists\ n, m \in \mathbb{Z}$ s.t. $a = 3n+1, b = 3m+2$. Then $a + b = 3n+1+3m+2 = 3(n+m+1)$. Knowing that $3 \mid a+b$ and that $a+b \mid a^2b^2(a^2-b^2)$, we know by transitivity that $3 \mid a^2b^2(a^2-b^2)$.

Because $3 \mid a^2b^2(a^2-b^2)$ in all three cases, we know that $3 \mid a^2b^2(a^2-b^2)\ \forall\ a, b \in \mathbb{Z}$. We also already have that $4 \mid a^2b^2(a^2-b^2)\ \forall\ a, b \in \mathbb{Z}$, so we can say that $\exists\ n, m \in \mathbb{Z}$ s.t. $a^2b^2(a^2-b^2) = 4n$ and $a^2b^2(a^2-b^2) = 3m$. Combining these equations, we get that $4a^2b^2(a^2-b^2) - 12m = 3a^2b^2(a^2-b^2) - 12n \Rightarrow a^2b^2(a^2-b^2) = 12(m-n)$. Because $n, m \in \mathbb{Z}, (m-n) \in \mathbb{Z}$ as well. Therefore, we conclude that $a^2b^2(a^2-b^2)$ is divisible by 12 for all $a, b \in \mathbb{Z}$. ■

PS9, #4. Let $f : A \to B$ be a function. Let $W \subseteq A$, and let $X, Y \subseteq B$. Prove or disprove $f(f^{-1}(X)) \subseteq X$.

Proof: Let $a \in f(f^{-1}(X))$. Then $\exists\ b \in f^{-1}(X)$ s.t. $f(b) = a$. Since $b \in f^{-1}(X)$, $f(b) \in X$. Since $f(b) \in X$ and we know $f(b) = a$, it must be true that $f(b) = a \in X$. Thus, $f(f^{-1}(X)) \subseteq X$. ■

> **PS12, #3.** Let $f : A \to B$ and $g : B \to C$. Prove that if f and g are invertible, so is $g \circ f$ and that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Proof: Let f and g be invertible. Then by the theorem, f and g are also bijective. Then f and g are surjective and injective. Since the composition of two injections is itself injective, $g \circ f$ is injective. Since the composition of two surjections is itself surjective, $g \circ f$ is surjective. Since $g \circ f$ is injective and surjective, $g \circ f$ is bijective. Then by the theorem, $g \circ f$ has an inverse. ■

For the inverse of $g \circ f$, $(g \circ f)^{-1}$, $(g \circ f) \circ (g \circ f)^{-1} = \mathrm{id}_C$. Let f^{-1} be the inverse of f and g^{-1} be the inverse of g. Then by associativity, $(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ f^{-1}) \circ g^{-1}$. Since f and f^{-1} are inverses, $f \circ f^{-1} = \mathrm{id}_B$. Then $g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ \mathrm{id}_B \circ g^{-1}$. Since $\mathrm{id}_B \circ g^{-1} = g^{-1}$, $g \circ \mathrm{id}_B \circ g^{-1} = g \circ g^{-1} = \mathrm{id}_C$. ■

3 Abstract Vector Spaces

The following are a small selection of proofs I wrote for homework in Math 321, Abstract Vector Spaces, at Emory, taught by Parimala Raman. These problems were completed independently and without outside assistance.

> **PS9, #4.** Let T be an invertible linear operator. Prove that a scalar λ is an eigenvalue of T if and only if λ^{-1} is an eigenvalue of T^{-1}.

Proof: Let $T : V \to V$ be an isomorphism.

($\implies$) Suppose λ is an eigenvalue of T with a corresponding eigenvector $\mathbf{v}$. Then $T(\mathbf{v}) = \lambda\mathbf{v}$. Let T^{-1} be the inverse of T. Then $T^{-1}(T(\mathbf{v})) = T^{-1}(\lambda\mathbf{v}) \iff Id_V(\mathbf{v}) = \mathbf{v} = \lambda T^{-1}(\mathbf{v})$. Then $T^{-1}(\mathbf{v}) = \lambda^{-1}\mathbf{v}$, so λ^{-1} is an eigenvalue of T^{-1}.

($\impliedby$) Since λ is an eigenvalue of $T \implies \lambda^{-1}$ is an eigenvalue of T^{-1}, λ^{-1} is an eigenvalue of $T^{-1} \implies (\lambda^{-1})^{-1} = \lambda$ is an eigenvalue of $(T^{-1})^{-1} = T$. Thus, λ^{-1} is an eigenvalue of $T^{-1} \implies \lambda$ is an eigenvalue of T. ■

> **PS10, #4.** Let A be an $n \times n$ matrix that is similar to an upper triangular matrix and has the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ with corresponding multiplicities $m_1, m_2, \ldots, m_k$. Prove that $\operatorname{tr}(A) = \sum_{i=1}^{k} m_i\lambda_i$ and that $\det(A) = \prod_{i=1}^{k}(\lambda_i)^{m_i}$.

Proof: Let $A, B, Q \in M_n(F)$, B be upper triangular, and Q invertible such that $B = Q^{-1}AQ$.

Proposition: λ is an eigenvalue of A if and only if λ is an eigenvalue of B.
Proof: Let λ be an eigenvalue of A with a corresponding eigenvector $\mathbf{v}$. Then $A\mathbf{v} = \lambda\mathbf{v} \iff QBQ^{-1}\mathbf{v} = \lambda\mathbf{v} \iff BQ^{-1}\mathbf{v} = \lambda Q^{-1}\mathbf{v}$. Then λ is an eigenvalue of B with a corresponding eigenvector $Q^{-1}\mathbf{v}$. Conversely, let λ be an eigenvalue of B. Then $B\mathbf{v} = \lambda\mathbf{v} \iff Q^{-1}AQ\mathbf{v} = \lambda\mathbf{v} \iff AQ\mathbf{v} = \lambda Q\mathbf{v}$. Then λ is an eigenvalue of A with corresponding eigenvector $Q\mathbf{v}$. Then we can conclude that λ is an eigenvalue of A if and only if λ is an eigenvalue of B. That is, A and B have the same eigenvalues.

Proposition: If $X, Y \in M_n(F)$, then $\operatorname{tr}(XY) = \operatorname{tr}(YX)$.
Proof: $\operatorname{tr}(XY) = \sum_{i=1}^{n}(XY)_{ii} = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} x_{ij}y_{ji}\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} y_{ji}x_{ij}\right) = \sum_{j=1}^{n}(YX)_{jj}$
$= \operatorname{tr}(YX)$.

Proposition: Similar matrices have the same trace.
Proof: $\operatorname{tr}(A) = \operatorname{tr}(QBQ^{-1})$. Since $\operatorname{tr}(XY) = \operatorname{tr}(YX)$, $\operatorname{tr}(QBQ^{-1}) = \operatorname{tr}(QQ^{-1}B) = \operatorname{tr}(B)$.

Since A and B have the same eigenvalues, by §5.2 Exercise 10, B has diagonal entries of $\lambda_1, \lambda_2, \ldots, \lambda_k$, with each λ_i appearing on the main diagonal of B m_i times $(1 \leq i \leq k)$.

Then $\operatorname{tr}(B) = \sum_{i=1}^{k} m_i\lambda_i = \operatorname{tr}(A)$.

Since B has diagonal entries of $\lambda_1, \lambda_2, \ldots, \lambda_k$ with each λ_i appearing m_i times on its main diagonal $(1 \leq i \leq k)$, and since B is upper triangular, $\det(B) = \prod_{i=1}^{k}(\lambda_i)^{m_i}$.

$\det(A) = \det(QBQ^{-1}) = \det(Q)\det(B)\det(Q^{-1}) = \det(Q)(\det(Q))^{-1}\det(B) = \det(B) = \prod_{i=1}^{k}(\lambda_i)^{m_i}$.
So $\det(A) = \prod_{i=1}^{k}(\lambda_i)^{m_i}$. ■

PS10, #5. Let T be a linear operator on a finite-dimensional vector space V. Recall that for any eigenvalue λ of T, λ^{-1} is an eigenvalue of T^{-1}. Prove that the eigenspace of T corresponding to λ is the same as the eigenspace of T^{-1} corresponding to λ^{-1}. Then prove that if T is diagonalizable, then T^{-1} is diagonalizable.

Proof: Let $\dim(V) = n$. Let E_λ represent the eigenspace of T corresponding to λ, and $E_{\lambda^{-1}}$ represent the eigenspace of T^{-1} corresponding to λ^{-1}. Let $\mathbf{v} \in E_\lambda$. Then $T(\mathbf{v}) = \lambda\mathbf{v} \iff T^{-1} \circ T(\mathbf{v}) = T^{-1}(\lambda\mathbf{v}) \iff \mathbf{v} = \lambda T^{-1}(\mathbf{v}) \iff T^{-1}(\mathbf{v}) = \lambda^{-1}\mathbf{v} \iff \mathbf{v} \in E_{\lambda^{-1}}$. Then $E_\lambda \subseteq E_{\lambda^{-1}}$. Conversely, let $\mathbf{v} \in E_{\lambda^{-1}}$. Then $T^{-1}(\mathbf{v}) = \lambda^{-1}\mathbf{v} \iff T \circ T^{-1}(\mathbf{v}) = \lambda^{-1}T(\mathbf{v}) \iff T(\mathbf{v}) = \lambda\mathbf{v} \iff \mathbf{v} \in E_\lambda$. Then $E_{\lambda^{-1}} \subseteq E_\lambda$, so $E_\lambda = E_{\lambda^{-1}}$.

Since T is diagonalizable, a basis β for V comprised of eigenvectors of T yields a diagonal matrix $[T]_\beta$. Since $E_\lambda = E_{\lambda^{-1}}$, β is comprised of eigenvectors of T^{-1}. Then for each ith vector $\mathbf{v}_i$ in β, $T^{-1}(\mathbf{v}_i) = \lambda^{-1}\mathbf{v}_i$, so $[T^{-1}(\mathbf{v}_i)]_\beta = \lambda^{-1}\mathbf{e}_i$, where $\mathbf{e}_i$ is the n-dimensional vector whose ith entry is 1 and all other entries are 0. Then $[T^{-1}]_\beta = (\lambda_1^{-1}\mathbf{e}_1\ \lambda_2^{-1}\mathbf{e}_2\ \ldots\ \lambda_n^{-1}\mathbf{e}_n) = \operatorname{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1})$, a diagonal matrix. Then T^{-1} is diagonalizable. ■

4 Real Analysis

The following are a small selection of proofs I that I have written for homework in 18.100A, Real Analysis, at MIT, taught by Qin Deng. These problems were completed independently and without outside assistance.

> **PS3, #3.** Let $\{x_n\}_{n=1}^{\infty}$ be an increasing sequence. Let $\{x_{n_i}\}_{n=1}^{\infty}$ be a subsequence and suppose $\{x_{n_i}\}$ converges to some $L \in \mathbb{R}$. Prove directly from the definition of limits that $\{x_n\}$ converges to L.

I first claim that $\{x_{n_i}\}$ must be increasing and bounded above by L.

Proof that $\{x_{n_i}\}$ is increasing:

Choose any two $p, q \in \mathbb{N}$ s.t. $p < q$. Then since subsequence indices must be strictly increasing, $n_p < n_q$. Further, since $\{x_n\}$ is increasing, $x_{n_p} \leq x_{n_q}$. Thus, $\{x_{n_i}\}$ is increasing.

Proof that $\{x_{n_i}\}$ is bounded above by L:

For simplicity, let $\{y_n\}_{n=1}^{\infty} := \{x_{m_i}\}_{m=1}^{\infty}$. By the above, we know that $\{y_n\}$ is increasing and that $\lim\limits_{n\to\infty} y_n = L \in \mathbb{R}$. For the sake of contradiction, assume $\{y_n\}$ is not bounded above by L. Then $\exists\ M \in \mathbb{N}$ s.t. $y_M > L$. Then $\forall\ m \in \mathbb{N}_{>M}$, $y_m > L$ since $\{y_n\}$ is increasing. Since $y_m > L$, $|y_m - L| = y_m - L > 0$.

Now let $\varepsilon > 0$. By the definition of limits, $\exists\ N_1 \in \mathbb{N}$ s.t. $\forall\ n \in \mathbb{N}_{>N_1}$, $|y_n - L| < \varepsilon$. Define $N := \max\{N_1, M\}$. Then $\forall\ n \in \mathbb{N}_{>N}$, $|y_n - L| < \varepsilon$ and $y_n - L > 0$.

Choose $\varepsilon = \frac{y_N - L}{2}$ and notice that since $y_N > L$, $y_N - L > 0$ and thus, $\frac{y_N - L}{2} > 0$. Further, $y_n > y_N$, so $y_n - L > y_N - L$, $\frac{y_n - L}{2} < y_n - L$, and $\frac{y_n - L}{2} > \frac{y_N - L}{2}$. Thus, $y_n - L > \frac{y_N - L}{2}$. Since $|y_n - L| = y_n - L$, we have that $|y_n - L| > \frac{y_N - L}{2}$, a contradiction. Thus, $\{y_n\}$ is bounded above by L.

Proof: Let $\varepsilon > 0$. Then $\exists\ N \in \mathbb{N}$ s.t. $\forall\ i \in \mathbb{N} > N$, $|x_{n_i} - L| < \varepsilon$. Since $\{x_{n_i}\}$ is increasing and bounded from above by L, $x_{n_i} \leq L$ for all $n \in \mathbb{N}$, so $|x_{n_i} - L| = L - x_{n_i}$. Let $m \in \mathbb{N}$. Then $\exists\ i \in \mathbb{N}_{>N}$ s.t. $n_i \leq m$. Then $x_{n_i} \leq x_m$. There are two cases: (i) $x_m \leq L$ and (ii) $x_m > L$.

Consider the second case (ii) first: Suppose $x_m > L$. There must be $q \in \mathbb{N}$ s.t. $n_q > m$, and thus that $x_{n_q} > x_m$. Then $x_{n_q} > L$, so $\{x_{n_i}\}$ is not bounded above by L, a contradiction. Therefore, this case is not possible.

Now consider case (i): Suppose $x_{n_i} \leq x_m \leq L$. Then $|x_m - L| = L - x_m$. And since $x_m \geq x_{n_i}$, $-x_m \leq -x_{n_i}$, so $L - x_m \leq L - x_{n_i} < \varepsilon$, and $|x_m - L| < \varepsilon$. Therefore, $\lim\limits_{n\to\infty} x_n = L$. ■

> **PS4, #2.** Let $x \in \mathbb{R}$. Suppose $\{x_n\}_{n=1}^{\infty}$ is a sequence such that every subsequence $\{x_{n_i}\}_{i=1}^{\infty}$ has a subsequence $\{x_{n_{i_j}}\}_{j=1}^{\infty}$ that converges to x. Prove that $\{x_n\}$ is bounded.

Proof: For the sake of contradiction, assume $\{x_n\}$ is not bounded. Then I define a subsequence, $\{x_{n_i}\}_{i=1}^{\infty}$ inductively:

Since $\{x_n\}$ is unbounded, it is clearly not bounded by 1, so when $i = 1$, $\exists\, j \in \mathbb{N}$ s.t. $|x_j| > 1$. Then I define $n_1 = j$ and thus $x_{n_1} = x_j$ so that $|x_{n_1}| > 1$.

For some $i \in \mathbb{N}$, assume $\exists\, k \in \mathbb{N}$ s.t. $|x_k| > i$. Then I define $n_i = k$ and thus $x_{n_i} = x_k$ so that $|x_{n_i}| > i$. Since $\{x_n\}$ is unbounded, every K-tail, including its $(k+1)$-tail, is unbounded as well.* Then since the $(k+1)$-tail of $\{x_n\}$, $\{x_n\}_{n=k+1}^{\infty}$ is unbounded, it is clearly not bounded by $i+1$. Then it follows that $\exists\, m \in \mathbb{N}$ s.t. $m \geq k+1$ and $|x_m| > i+1$. Then I define $n_{i+1} = m$ and thus $x_{n_{i+1}} = x_m$ so that $|x_{n_{i+1}}| > i+1$.

By this definition and the principles of induction, the sequence $\{n_i\}_{i=1}^{\infty}$ is strictly increasing (notice that $m \geq k+1 > k$), so $\{x_{n_i}\}_{i=1}^{\infty}$ is a subsequence of $\{x_n\}$, and further, $|x_{n_i}| > i$ for every $i \in \mathbb{N}$.

Then $\forall\, i \in \mathbb{N}$ where $i > x+1$, $|x_{n_i}| > x+1$. Then for any subsequence of $\{x_{n_i}\}$, $\{x_{n_{i_j}}\}_{j=1}^{\infty}$, $|x_{n_{i_j}}| > x+1$ whenever $j > x+1$ since $n_{i_j} \geq j$. Then for every subsequence $\{x_{n_{i_j}}\}$ of $\{x_{n_i}\}$, $\{x_{n_{i_j}}\}$ does not converge to x, a contradiction.

Therefore, $\{x_n\}$ must be bounded. ∎

Proof that every K-tail of an unbounded sequence is unbounded: Let $\{x_n\}_{n=1}^{\infty}$ be an unbounded sequence. Then $\forall\, b \in \mathbb{R}$, $\exists\, i \in \mathbb{N}$ s.t. $|x_i| > b$. For the sake of contradiction, assume $\exists$ some k-tail such that $\{x_n\}_{n=k}^{\infty}$ is bounded. Then $\exists \sup\{x_n : n \geq k\} =: a$. Now choose $b = \sum_{i=1}^{k-1} |x_i| + a$.

If $n \geq k$, then by the definition of the supremum, $|x_n| \leq b$. If instead $n < k$, then note that since $\{x_n\}$ is unbounded, $\exists\, i \in \mathbb{N}$ s.t. $|x_i| > b$. Notice that $b = |x_1| + |x_2| + \cdots + |x_i| + \cdots + |x_{k-1}|$, so $|x_i| > b \implies |x_1| + |x_2| + \cdots + |x_{i-1}| + |x_{i+1}| + \cdots + |x_{k-1}| < 0$. But $|x_j| \geq 0$ for each $j \in \mathbb{N}$ by absolute value properties and since $x_j \in \mathbb{R}$, so $|x_1| + |x_2| + \cdots + |x_{i-1}| + |x_{i+1}| + \cdots + |x_{k-1}| \geq 0$, a contradiction. Thus, the k-tail of an unbounded sequence $\{x_n\}$ must also be unbounded. ∎

PS5, #1. Let $\sum\limits_{n=1}^{\infty} x_n, \sum\limits_{n=1}^{\infty} y_n$ be two series. Assume that $\exists$ some $N \in \mathbb{N}$ s.t. $y_n \geq x_n$ for any $n > N$. Prove that if $\sum\limits_{n=1}^{\infty} x_n = \infty$, then $\sum\limits_{n=1}^{\infty} y_n = \infty$.

Proof: Let $N \in \mathbb{N}$ s.t. $y_n \geq x_n$ for any $n \in \mathbb{N}, n \geq N$. Since for any $M \in \mathbb{N}$, a series $\sum\limits_{n=1}^{\infty} z_n$ converges iff $\sum\limits_{n=M}^{\infty} z_n$ converges, we know that $\sum\limits_{n=1}^{\infty} x_n, \sum\limits_{n=1}^{\infty} y_n$ converge iff $\sum\limits_{n=N}^{\infty} x_n, \sum\limits_{n=N}^{\infty} y_n$ converge, respectively.

By definition, if $\sum\limits_{n=1}^{\infty} x_n = \infty$, then $\lim\limits_{k\to\infty}\left(\sum\limits_{n=1}^{k} x_n\right) = \infty$ and further, that $\forall\, A \in \mathbb{R}$, $\exists\, M \in \mathbb{N}$ s.t. $\forall k \in \mathbb{N}, k > M$, $\sum\limits_{n=1}^{k} x_n > A$. So let $A \in \mathbb{R}$. Then $\exists N \in \mathbb{R}$ s.t. $\forall k \in \mathbb{N}, k > N$, $\sum\limits_{n=1}^{k} x_n > A + \sum\limits_{n=1}^{N-1} x_n$. Then $\sum\limits_{n=1}^{k} x_n - \sum\limits_{n=1}^{N-1} x_n = \sum\limits_{n=N}^{k} x_n > A$, so we can conclude that if $\sum\limits_{n=1}^{\infty} x_n = \infty$, then $\sum\limits_{n=N}^{\infty} x_n = \infty$. Therefore, it is sufficient to show that if $\sum\limits_{n=N}^{\infty} x_n = \infty$, then $\sum\limits_{n=1}^{\infty} y_n = \infty$.

I first claim that if $\sum\limits_{n=N}^{\infty} x_n = \infty$, then $\sum\limits_{n=N}^{\infty} y_n = \infty$. Let $A \in \mathbb{R}$. Then $\exists M \in \mathbb{N}$ s.t. $\forall m \in \mathbb{N}, m > M$, $\sum\limits_{n=N}^{m} x_n > A$. Since $y_n \geq x_n$ for each $n \in \mathbb{N}$ where $n \geq N$, it follows that $\sum\limits_{n=N}^{m} y_n \geq \sum\limits_{n=N}^{m} x_n > A$, so

$\sum\limits_{n=N}^{m} y_n > A$. Thus, $\sum\limits_{n=N}^{\infty} y_n = \infty$.

Now it remains to show that if $\sum\limits_{n=N}^{\infty} y_n = \infty$, then $\sum\limits_{n=1}^{\infty} y_n = \infty$ as well. Let $A \in \mathbb{R}$. By definition, $\exists\, M \in \mathbb{N}$ s.t. $\forall\, k \in \mathbb{N}, k > M$, $\sum\limits_{n=N}^{k} y_n > A - \sum\limits_{n=1}^{N-1} y_n$. Then clearly, $\sum\limits_{n=N}^{k} y_n + \sum\limits_{n=1}^{N-1} y_n = \sum\limits_{n=1}^{k} y_n > A$. So by definition, $\sum\limits_{n=1}^{\infty} y_n = \infty$.

Thus, if $\sum\limits_{n=1}^{\infty} x_n = \infty$, then $\sum\limits_{n=1}^{\infty} y_n = \infty$. ∎

> **PS6, #3.** Let $S \subseteq \mathbb{R}$, c be a cluster point of S, and $f : S \to \mathbb{R}$ be a function. Suppose that for every sequence $\{x_n\}_{n=1}^{\infty}$ in $S \setminus \{c\}$ such that $\lim\limits_{n\to\infty} x_n = c$, the sequence $\{f(x_n)\}_{n=1}^{\infty}$ is convergent. Show that $f(x)$ converges as $x \to c$.

Proof: Let $\{x_n\}_{n=1}^{\infty}$, $\{y_n\}_{n=1}^{\infty}$ be sequences satisfying the specified conditions. To show that $f(x)$ converges as $x \to c$, it suffices to show that $\{f(x_n)\}_{n=1}^{\infty}$ and $\{f(y_n)\}_{n=1}^{\infty}$ converge to the same limit. For the sake of contradiction, suppose this is not the case. That is, that $\exists\, L_1, L_2 \in \mathbb{R}$ defined by $L_1 := \lim\limits_{n\to\infty} f(x_n)$ and $L_2 := \lim\limits_{n\to\infty} f(y_n)$ such that $L_1 \neq L_2$. By the limit definition for sequences, $\exists\, N_1, N_2 \in \mathbb{N}$ s.t. for every $n, m \in \mathbb{N}$ where $n > N_1, m > N_2$, $|f(x_n) - L_1| < \frac{|L_1-L_2|}{2}$ and $|f(y_m) - L_2| < \frac{|L_1-L_2|}{2}$.

I now construct a new sequence, $\{z_n\}_{n=1}^{\infty}$, where $z_n = x_n$ for every odd $n \in \mathbb{N}$ and $z_n = y_n$ for every even $n \in \mathbb{N}$. Let $\varepsilon > 0$. By definition, $\exists\, M_1, M_2 \in \mathbb{N}$ s.t. $\forall\, n, m \in \mathbb{N}$ where $n > M_1, m > M_2$, $|x_n - c| < \varepsilon$ and $|y_m - c| < \varepsilon$. Then notice that for every $n \in \mathbb{N}$ where $n > \max\{M_1, M_2\}$, $|z_n - c| = |x_n - c|$ if n is odd and $|z_n - c| = |y_n - c|$ if n is even. Since $n > \max\{M_1, M_2\}$, $|x_n - c|, |y_n - c| < \varepsilon$, so we know that $|z_n - c| < \varepsilon$. Then by the limit definition for sequences, the sequence $\{z_n\}_{n=1}^{\infty}$ converges to c. Furthermore, for every $n \in \mathbb{N}$, $x_n, y_n \neq c$, so similarly, for every $n \in \mathbb{N}$, $z_n \neq c$. Then $\{z_n\}$ satisfies the conditions stated in the question.

By assumption, the sequence $\{f(z_n)\}_{n=1}^{\infty}$ must then converge. Then $\exists\, L \in \mathbb{R}$ such that for every $\varepsilon > 0$, $\exists\, N \in \mathbb{N}$ s.t. $\forall\, n \in \mathbb{N}$, $n > N$, $|f(z_n) - L| < \varepsilon$. Consider N_1, N_2 again from above. Then for every $n \in \mathbb{N}$ where $n > \max\{N_1, N_2\}$ and n is odd, $|f(z_n) - L_1| < \frac{|L_1-L_2|}{2}$. Then when n is odd, $f(z_n) \in (L_1 - \frac{|L_1-L_2|}{2}, L_1 + \frac{|L_1-L_2|}{2})$. For every $n \in \mathbb{N}$ where $n > \max\{N_1, N_2\}$ and n is even, $|f(z_n) - L_2| < \frac{|L_1-L_2|}{2}$, so $f(z_n) \in (L_2 - \frac{|L_1-L_2|}{2}, L_2 + \frac{|L_1-L_2|}{2})$. Now considering every $n \in \mathbb{N}$ where $n > \max\{N_1, N_2\}$, the following must both be true:

1. $f(z_n) \in (L_1 - \frac{|L_1-L_2|}{2}, L_1 + \frac{|L_1-L_2|}{2})$,
2. $f(z_n) \in (L_2 - \frac{|L_1-L_2|}{2}, L_2 + \frac{|L_1-L_2|}{2})$.

The above requirements cannot both be true: WLOG, assume $L_1 > L_2$. Then $f(z_n) > L_1 - \frac{L_1-L_2}{2}$ and $f(z_n) < L_2 + \frac{L_1-L_2}{2}$. For this to be possible, $L_1 - \frac{L_1-L_2}{2} < L_2 + \frac{L_1-L_2}{2}$, so $L_1 - L_2 < L_1 - L_2$. This is impossible. Therefore, the sequence $\{f(z_n)\}$ does not converge, a contradiction.

Then the sequences $\{f(x_n)\}_{n=1}^{\infty}$, $\{f(y_n)\}_{n=1}^{\infty}$ converge to the same limit. Therefore, $f(x)$ converges as $x \to c$. ∎

PS7, #1. Let $S \subseteq \mathbb{R}$ and let $c \in \mathbb{R}$. Let $f : S \to \mathbb{R}$ be a function and $L \in \mathbb{R}$. Assume c is a cluster point of both $S \cap (-\infty, c)$ and $S \cap (c, \infty)$. Prove that c is a cluster point of S and that

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

Proof: Let $S \subseteq \mathbb{R}$, $f : S \to \mathbb{R}$ be a function, and $L \in \mathbb{R}$. Assume $c \in \mathbb{R}$ is a cluster point of $S \cap (-\infty, c)$ and $S \cap (c, \infty)$. Let $\varepsilon > 0$. Since c is a cluster point of $S \cap (-\infty, c)$, $\exists\, y \in S \cap (-\infty, c)$ s.t. $|y - c| < \varepsilon$. Since $S \cap (-\infty, c) \subseteq S \setminus \{c\}$, y is also in $S \setminus \{c\}$. Then $\exists\, y \in S \setminus \{c\}$ s.t. $|y - c| < \varepsilon$. Then $\forall\, \varepsilon > 0$, $\exists\, y \in S \setminus \{c\}$ s.t. $|y - c| < \varepsilon$, so by the definition of cluster points, c is a cluster point of S.

I now prove that

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

($\implies$): Suppose $\lim_{x \to c} f(x) = L$. Then by definition, $\forall\, \varepsilon > 0$, $\exists\, \delta > 0$ s.t. $\forall\, x \in S \setminus \{c\}$ where $|x - c| < \delta$, $|f(x) - L| < \varepsilon$. Define the function $f' : S \cap (-\infty, c) \to \mathbb{R}$ by $f'(x) = f(x)$ for all $x \in S \cap (-\infty, c)$. Let $\varepsilon > 0$ and $x \in S \cap (-\infty, c)$. Since $S \cap (-\infty, c) \subseteq S \setminus \{c\}$, $x \in S \setminus \{c\}$, so $\exists\, \delta_1 > 0$ s.t. if $|x - c| < \delta_1$, then $|f(x) - L| < \varepsilon$. Then by definition, $\lim_{x \to c^-} f(x) = L$.

Again, let $\varepsilon > 0$. Let $x \in S \cap (c, \infty)$. Since $S \cap (c, \infty) \subseteq S \setminus \{c\}$, $x \in S \setminus \{c\}$. Then $\exists\, \delta_2 > 0$ s.t. if $|x - c| < \delta_2$, then $|f(x) - L| < \varepsilon$. Then by definition, $\lim_{x \to c^+} f(x) = L$. Thus, $\lim_{x \to c^-} f(x) = \lim_{x \to c^+} f(x) = L$.

($\impliedby$): Suppose $\lim_{x \to c^-} f(x) = \lim_{x \to c^+} f(x)$. Let $\varepsilon > 0$. Since $\lim_{x \to c^-} f(x) = L$, $\exists\, \delta_1 > 0$ s.t. $\forall\, x \in S \cap (-\infty, c)$ where $|x - c| < \delta_1$, $|f(x) - L| < \varepsilon$. Since $\lim_{x \to c^+} f(x) = L$, $\exists \delta_2 > 0$ s.t. $\forall x \in S \cap (c, \infty)$ where $|x - c| < \delta_2$, $|f(x) - L| < \varepsilon$. Now notice that $(S \cap (-\infty, c)) \cup (S \cap (c, \infty)) = S \cap ((-\infty, c) \cup (c, \infty))$, and $(-\infty, c) \cup (c, \infty) = \mathbb{R} \setminus \{c\}$, and further, $S \cap (\mathbb{R} \setminus \{c\}) = S \setminus \{c\}$. Then for any $\varepsilon > 0$, let $x \in S \setminus \{c\}$ s.t. $|x - c| < \min\{\delta_1, \delta_2\}$. If $x < c$, then $x \in S \cap (-\infty, c)$, so since $|x - c| < \delta_1$, we know $|f(x) - L| < \varepsilon$. Similarly, if $x > c$, then $x \in S \cap (c, \infty)$, so since $|x - c| < \delta_2$, we know that once again, $|f(x) - L| < \varepsilon$. Then choosing $\delta = \min\{\delta_1, \delta_2\}$, we see that for any $x \in S \setminus \{c\}$ where $|x - c| < \delta$, $|f(x) - L| < \varepsilon$. Then we can conclude that $\forall\, \varepsilon > 0$, $\exists\, \delta > 0$ s.t. $\forall\, x \in S \setminus \{c\}$ where $|x - c| < \delta$, $|f(x) - L| < \varepsilon$. Then by the definition of limits for functions, $\lim_{x \to c} = L$. ■

PS8, #1. Let $f : (0, 1) \to \mathbb{R}$ be a continuous function so that $\lim_{x \to 0} f(x) = \lim_{x \to 1} f(x) = 0$. Show that f achieves an absolute minimum or an absolute maximum on $(0, 1)$.

Proof: If f achieves an absolute minimum or an absolute maximum on $(0, 1)$, then $\exists\, c \in (0, 1)$ s.t. $\forall\, x \in (0, 1)$, $|f(x)| \le |f(c)|$. Then there are two cases: (i) $\forall\, x \in (0, 1)$, $f(x) = 0$; and (ii) $\exists\, c \in (0, 1)$ s.t. $|f(c)| > 0$.

(i) Sps. that $\forall\, x \in (0, 1)$, $f(x) = 0$. Then f achieves an absolute minimum and maximum on $(0, 1)$. Consider $f(0.5) = 0$. Then $\forall\, x \in (0, 1)$, $f(x) = f(0.5) = 0$, so $|f(x)| \le |f(0.5)| = 0$.

(ii) Sps. instead that $\exists\, c \in (0,1)$ s.t. $|f(c)| > 0$. Since $\lim\limits_{x\to 0} f(x) = 0$, $\exists\, \delta_1 \in (0, \frac{1}{2})$ s.t. $\forall\, x \in (0,1)$ where $|x| = x < \delta_1$, $|f(x)| < |f(c)|$. Similarly, since $\lim\limits_{x\to 1} f(x) = 0$, $\exists\, \delta_2 \in (0, \frac{1}{2})$ s.t. $\forall\, x \in (0,1)$ where $|x-1| = 1-x < \delta_2$, $|f(x)| < |f(c)|$. Consider the closed interval $[\delta_1, 1-\delta_2]$. By EVT, $\exists\, a \in [\delta_1, 1-\delta_2]$ s.t. $\forall\, x \in [\delta_1, 1-\delta_2]$, $|f(x)| \leq |f(a)|$. Then f achieves a minimum or maximum on $(0,1)$ at a if $\forall\, x \in (0,\delta_1) \cup (1-\delta_2, 1)$, $|f(x)| \leq |f(a)|$, which remains to be shown.

By construction, we know that $\forall\, x \in (0,\delta_1) \cup (1-\delta_2, 1)$,[1] $|f(x)| < |f(c)|$ and that $c \in (0,1)$. Then clearly, $c \in (0,1) \setminus ((0,\delta_1) \cup (1-\delta_2, 1)) = [\delta_1, 1-\delta_2]$. Since $\forall\, x \in [\delta_1, 1-\delta_2]$, $|f(x)| \leq |f(a)|$, it follows that $|f(c)| \leq |f(a)|$. Then since $\forall\, x \in (0,\delta_1) \cup (1-\delta_2, 1)$, $|f(x)| < |f(c)| \leq |f(a)|$, $|f(x)| \leq |f(a)|$. Then f achieves a minimum or maximum at $a \in (0,1)$. ■

PS10, #2(a). Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable. Prove that $|f|$ is Riemann integrable.

Proof: From the ε reformulation of integrability, it suffices to show that $\forall\, \varepsilon > 0$, $\exists$ a partition P of $[a,b]$ s.t. $U(P,f) - L(P,f) < \varepsilon$. I begin by showing that for every partition P of $[a,b]$, $U(P,|f|) - L(P,|f|) \leq U(P,f) - L(P,f)$.

Let $P := \{x_0, x_1, \cdots, x_n\}$ partition $[0,1]$ for some $n \in \mathbb{N}$. Then consider the ith subinterval, $S_i := [x_{i-1}, x_i]$ for $i = 1, 2, \cdots, n$. Suppose $\inf(f(S_i)) \geq 0$. Since $\inf(f(S_i)) \geq 0$, and $\forall\, x \in S_i$, $f(x) \geq \inf(f(S_i))$, it must be the case that $\forall\, x \in S_i$, $f(x) \geq 0$ and thus, $|f(x)| = f(x)$. Then $\inf(|f(S_i)|) = \inf(f(S_i))$, and $\sup(|f(S_i)|) = \sup(f(S_i))$. So clearly, $\sup(|f(S_i)|) - \inf(|f(S_i)|) = \sup(f(S_i)) - \inf(f(S_i))$.

If instead $\sup(f(S_i)) \leq 0$, then by definition, for every $x \in S_i$, $f(x) \leq \sup(f(S_i)) \leq 0$, so $|f(x)| = -f(x)$. Then for every $x \in S_i$, $f(x) \geq \inf(f(S_i)) \implies |f(x)| = -f(x) \geq -\inf(f(S_i))$. I claim that then $\sup(|f(S_i)|) = -\inf(f(S_i))$—For the sake of contradiction, suppose this claim is false. Then $\exists\, \varepsilon > 0$ s.t. $\forall\, x \in S_i$, $|f(x)| \leq -\inf(f(S_i))$. Knowing that $|f(x)| = -f(x)$, it then must be true that $-f(x) \leq -\inf(f(S_i)) - \varepsilon$ and then that $f(x) \geq \inf(f(S_i)) + \varepsilon$. But by the ε reformulation of inf, $\exists\, y \in S_i$ s.t. $f(y) \in (\inf(f(S_i)) + \varepsilon, \inf(f(S_i)))$, a contradiction. So $\sup(|f(S_i)|) = -\inf(f(S_i))$. Using similar logic—except that $-f(x) \geq -\sup(f(S_i))$—it is clear that $\inf(|f(S_i)|) = -\sup(f(S_i))$. Then $\sup(|f(S_i)|) - \inf(|f(S_i)|) = -\inf(f(S_i)) + \sup(f(S_i)) = \sup(f(S_i)) - \inf(f(S_i))$.

Finally, suppose that $\inf(f(S_i)) < 0 < \sup(f(S_i))$. Then $\inf(|f(S_i)|) \geq 0$, and $\sup(|f(S_i)|) = \max\{\sup(f(S_i)), -\inf(f(S_i))\}$. If $\sup(|f(S_i)|) = \sup(f(S_i))$, then since $-\inf(f(S_i)) > 0$, $\sup(|f(S_i)|) \leq \sup(f(S_i)) + (-\inf(f(S_i))) = \sup(f(S_i)) - \inf(f(S_i))$. And $\sup(|f(S_i)|) - \inf(|f(S_i)|) \leq \sup(|f(S_i)|)$, so $\sup(|f(S_i)|) - \inf(|f(S_i)|) \leq \sup(f(S_i)) - \inf(f(S_i))$. If instead $\sup(|f(S_i)|) = -\inf(f(S_i))$, then since $\sup(f(S_i)) > 0$, $\sup(|f(S_i)|) \leq -\inf(f(S_i)) + \sup(f(S_i))$. It then follows that $\sup(|f(S_i)|) - \inf(|f(S_i)|) \leq \sup(f(S_i)) - \inf(f(S_i))$.

In all cases, $\sup(|f(S_i)|) - \inf(|f(S_i)|) \leq \sup(f(S_i)) - \inf(f(S_i))$. Now consider $U(P,|f|) - L(P,|f|)$—

$$\begin{aligned} U(P,|f|) - L(P,|f|) &= \sum_{i=1}^{n} (\sup(|f(S_i)|) - \inf(|f(S_i)|))(x_i - x_{i-1}) \\ &\leq \sum_{i=1}^{n} (\sup(f(S_i)) - \inf(f(S_i)))(x_i - x_{i-1}) \\ &= U(P,f) - L(P,f). \end{aligned}$$

[1] If $x \in (1-\delta_2, 1)$, then $1 - \delta_2 < x < 1$, so $1 - x = |x-1| < \delta_2$.

I now show that $|f|$ is integrable via the ε reformulation of integrability. Let $\varepsilon > 0$. Since f is integrable, $\exists$ a partition P s.t. $U(P, f) - L(P, f) < \varepsilon$. As shown above, $U(P, |f|) - L(P, |f|) \leq U(P, f) - L(P, f)$, so it follows that $U(P, |f|) - L(P, |f|) < \varepsilon$ as well. Then $|f|$ is integrable. ■

PS10, #3. Let f, g be functions integrable on $[a, b]$. Prove that $f + g$ is integrable on $[a, b]$ and that

$$\int_a^b (f(x) + g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

Proof: Let f, g be integrable on $[a, b]$. I first demonstrate that for any partition P of $[a, b]$, $U(P, f+g) \leq U(P, f)+U(P, g)$ and that $L(P, f+g) \geq L(P, f+g)+L(P, f+g)$. Let P be a partition. Consider any subinterval of P, S_i. For every $x \in S_i$, $f(x) \leq \sup(f(S_i))$, and $g(x) \leq \sup(g(S_i))$. So $f(x)+g(x) = (f+g)(x) \leq \sup(f(S_i))+\sup(g(S_i))$ and $\forall y \in (f+g)(S_i)$, $y \leq \sup(f(S_i))+\sup(g(S_i))$. Then $\sup((f + g)(S_i)) \leq \sup(f_(S_i)) + \sup(g(S_i))$. Similarly, $f(x) \geq \inf(f(S_i)), g(x) \geq \inf(f(S_i))$, so $(f + g)(x) \geq \inf(f(S_i)) + \inf(g(S_i))$ and thus, $\inf{(f + g)(S_i)} \geq \inf(f(S_i)) + \inf(g(S_i))$. Then $U(P, f+g) \leq U(P, f)+U(P, g)$, and $L(P, f+g) \geq L(P, f)+L(P, g)$. So $U(P, f+g)-L(P, f+g) \leq U(P, f) + U(P, g) - L(P, g) - L(P, g)$.

I now apply the ε reformulation of integrability to show that $f + g$ is integrable on $[a, b]$. Let $\varepsilon > 0$. Since f is integrable on $[a, b]$, $\exists$ a partition P of $[a, b]$ s.t. $U(P, f) - L(P, f) < \frac{\varepsilon}{2}$. Similarly, $\exists$ a partition Q s.t. $U(Q, g) - L(Q, g) < \frac{\varepsilon}{2}$. Let $\tilde{P} = P \cup Q$. $P, Q \subseteq \tilde{P}$, so $\tilde{P}$ is a refinement of P and Q. Then $U(\tilde{P}, f)-L(\tilde{P}, f) < \frac{\varepsilon}{2}$, and $U(\tilde{P}, g)-L(\tilde{P}, g) < \frac{\varepsilon}{2}$. From the above, $U(\tilde{P}, f+g)-L(\tilde{P}, f+g) \leq U(\tilde{P}, f) - L(\tilde{P}, f) + U(\tilde{P}, g) - L(\tilde{P}, g) < \frac{\varepsilon}{2}$. So $f + g$ is integrable on $[a, b]$.

It remains now to show that

$$\int_a^b (f(x) + g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

Since f and g are integrable, I can construct sequences $\{U(P_n, f)\}_{n=1}^{\infty}, \{U(Q_n, g)\}_{n=1}^{\infty}$ such that the sequences converge to $\int_a^b f(x)dx$ and $\int_a^b g(x)dx$, respectively. From these sequences, I construct a third sequence, $\{U(R_n, f) + U(R_n, g)\}_{n=1}^{\infty}$ where for every $i \in \mathbb{N}$, $R_i := P_i \cup Q_i$, a refinement of the original partitions. Then by the additivity of the limits of converging sequences, $\{U(R_n, f) + U(R_n, g)\}$ converges to $\lim_{n\to\infty} U(R_n, f) + \lim_{n\to\infty} U(R_n, g) =$
$\int_a^b f(x)dx + \int_a^b g(x)dx$.

From above, for any partition P, $U(P, f + g) \leq U(P, f) + U(P, g)$, so the sequence
$\{U(R_n, f + g)\}_{n=1}^{\infty} \leq \{U(R_n, f) + U(R_n, g)\}_{n=1}^{\infty}$.

Repeat the process with sequences of partitions J_n, K_n and D_n, where for every $i \in \mathbb{N}$, $D_i := J_i \cup K_i$ and so that $\{L(J_n, f)\}_{n=1}^{\infty}$ and $\{L(K_n, g)\}_{n=1}^{\infty}$ converge to $\int_a^b f(x)dx$, $\int_a^b g(x)dx$, respectively. Then $\{L(D_n, f + g)\}_{n=1}^{\infty} \geq \{L(J_n, f)\} + \{L(K_n, g)\}$.

For any partitions P, Q, $L(P, f + g) \leq \int_a^b (f + g)(x)dx$, and $U(Q, f + g) \geq \int_a^b (f + g)(x)dx$, so for every $n \in \mathbb{N}$, we have

$$L(J_n, f) + L(K_n, g) \leq L(D_n, f + g) \leq \int_a^b (f + g)(x)dx \leq U(R_n, f + g) \leq U(P_n, f) + U(Q_n, g)$$

and thus,

$$L(J_n, f) + L(K_n, g) \leq U(R_n, f+g) \leq U(P_n, f) + U(Q_n, g).$$

Further,

$$\begin{aligned}\lim_{n\to\infty} L(J_n, f) + L(K_n, g) &= \int_a^b f(x)dx + \int_a^b g(x)dx \\ &= \lim_{n\to\infty} U(P_n, f) + U(Q_n, g).\end{aligned}$$

Then by squeeze theorem,

$$\lim_{n\to\infty} U(R_n, f+g) = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

By similar application of the squeeze theorem to the sequence of lower sums, we have that $\lim\limits_{n\to\infty} L(D_n, f+g) = \int_a^b f(x)dx + \int_a^b g(x)dx$. At this point, we can say that

$$\begin{aligned}\overline{\int_a^b}(f+g)(x)dx &\leq \int_a^b f(x)dx + \int_a^b g(x)dx \\ &\leq \underline{\int_a^b}(f+g)(x)dx.\end{aligned}$$

Since $f+g$ is integrable, $\overline{\int_a^b}(f+g)(x)dx = \underline{\int_a^b}(f+g)(x)dx = \int_a^b(f+g)(x)$, so

$$\begin{aligned}0 &\leq \int_a^b f(x)dx + \int_a^b g(x) - \underline{\int_a^b}(f+g)(x)dx \leq 0 \\ \implies & \int_a^b f(x)dx + \int_a^b g(x) - \underline{\int_a^b}(f+g)(x)dx = 0 \\ \implies & \int_a^b f(x)dx + \int_a^b g(x) = \underline{\int_a^b}(f+g)(x)dx = \int_a^b (f+g)(x)dx.\end{aligned}$$

■

I include the following exact transcriptions of select answers from my second midterm in real analysis, including a small error, for the sake of demonstrating abilities under time pressure and without access to any outside resources or assistance (though all homework problems above were also completed entirely independently).

Midterm 2, #4. Prove directly from the $\varepsilon - \delta$ definition of continuity: Let $f : \mathbb{R} \to \mathbb{R}$ be a function which satisfies that $f(x) \neq 0$ for any $x \in \mathbb{R}$. Let $a \in \mathbb{R}$. If f is continuous at a, then $\frac{1}{f}$ is continuous at a.

Let $g : \mathbb{R} \to \mathbb{R}$ be defined by $g(x) := \frac{1}{f(x)} \forall\, x \in \mathbb{R}$. Since $f(x) \neq 0$ for any $x \in \mathbb{R}$, all $x \in \mathbb{R}$ can be in the domain of g (i.e., g is defined validly).

Let $\varepsilon > 0$. Let f be cts at a. Since f is cts at a, $\exists \delta > 0$ s.t. $\forall x \in \mathbb{R}$ where $|x - a| < \delta$, $|f(x) - f(a)| < \min\left\{\frac{1}{2}f(a), \frac{1}{2}\varepsilon f^2(a)\right\}$. Then consider $\left|\frac{1}{f(x)} - \frac{1}{f(a)}\right|$. If $\frac{1}{f}$ is cts at a, then $\exists\, \delta' > 0$ s.t. $\forall\, x \in \mathbb{R}$ w/ $|x - a| < \delta'$, $\left|\frac{1}{f(x)} - \frac{1}{f(a)}\right| < \varepsilon$. $\left|\frac{1}{f(x)} - \frac{1}{f(a)}\right| = \left|\frac{f(x)-f(a)}{f(x)f(a)}\right|$. If $|x - a| < \delta$, then $|f(x) - f(a)| < \frac{1}{2}f(a)$, so $\frac{1}{2}f(a) < f(x) < \frac{3}{2}f(a)$. Then $|f(x)| > \frac{1}{2}|f(a)|$. So $\frac{|f(x)-f(a)|}{|f(x)||f(a)|} < \frac{|f(x)-f(a)|}{\frac{1}{2}f^2(a)} = 2\frac{|f(x)-f(a)|}{f^2(a)}$. Since $|f(x) - f(a)| < \frac{1}{2}\varepsilon f^2(a)$, $2\frac{|f(x)-f(a)|}{f^2(a)} < \frac{2\varepsilon f^2(a)}{2f^2(a)} = \varepsilon$. So $\left|\frac{f(x)-f(a)}{f(x)f(a)}\right| < \varepsilon$. Then $\exists\, \delta'(= \delta)$ s.t. $\forall\, x \in \mathbb{R}$ where $|x - a| < \delta'$, $\left|\frac{1}{f(x)} - \frac{1}{f(a)}\right| < \varepsilon$. By def., $\frac{1}{f}$ is cts at a.

Midterm 2, #5. Let $f : (0, \infty) \to \mathbb{R}$ be a continuous function with the property that

$$f^2(x) = x^2 \text{ for all } x \in (0, \infty).$$

Prove that if $f(1) = 1$, then $f(x) = x$ for all $x \in (0, \infty)$.

For the sake of contradiction, sps. $\exists\, \gamma \in (0, \infty)$ s.t. $f(\gamma) \neq \gamma$. Since $[f(\gamma)]^2 = \gamma^2$, $f(\gamma) = \gamma$ or $-\gamma$. Since $f(\gamma) \neq \gamma$ by assumption, it must be that $f(\gamma) = -\gamma$. Since $\gamma \in (0, \infty)$, $-\gamma < 0$. Then $f(\gamma) < 0 < f(1) = 1$.

Sps. $\gamma > 1$. Then since f is cts on $(0, \infty)$, by IVT, $\exists\, c \in (1, \gamma)$ s.t. $f(c) = 0$. Then $[f(c)]^2 = 0$, and $c^2 > 1 > 0$, a contradiction.

If $\gamma \in (0, 1)$, then similar reasoning has $\exists\, c \in (\gamma, 1)$ s.t. $f(c) = 0$. And then since $c > 0$, $c^2 > 0$, so $c^2 \neq [f(c)]^2 = 0$, again a contradiction.

Thus, $f(x) = x \forall\, x \in (0, \infty)$.

I also include the following exact transcriptions of select answers from my final exam in real analysis, including errors.

Final, #6. Let $f : [a, b] \to \mathbb{R}$ be a decreasing function. Prove that f is Riemann integrable on $[a, b]$.

Let $f : [a, b] \to \mathbb{R}$ be decreasing. Define a partition P_n of $[a, b]$ by $P_n = \{x_0, x_1, \cdots, x_n\}$ where $\forall\, i = 1, 2, \cdots, n$, $x_i - x_{i-1} = \frac{b-a}{n}$ (subintervals of equal width).

Let $\varepsilon > 0$. Choose $n \in \mathbb{N}$ s.t. $n > \frac{(b-a)(f(a)-f(b))}{\varepsilon}$.

Consider any ith subinterval of P_n, $\overline{x}_i = [x_{i-1}, x_i]$. Then since f is dec., $\forall\, x, y \in \overline{x}_i \subseteq [a, b]$, if $x \geq y$, then $f(x) \leq f(y)$. So $\forall x \in \overline{x}_i$, $x_{i-1} \leq x \leq x_i$, so $f(x_{i-1}) \geq f(x) \geq f(x_i)$, and $f(x_{i-1}), f(x_i) \in f(\overline{x}_i)$. So by def., $\sup(f(\overline{x}_i)) = f(x_{i-1})$, $\inf(f(\overline{x}_i)) = f(x_i)$.

So $L(P_n, f) = \sum_{i=1}^{n}(x_{i-1} - x_i) f(x_i)$, $U(P_n, f) = \sum_{i=1}^{n}(x_{i-1} - x_i) f(x_{i-1})$. And $U(P_n, f) - L(P_n, f) = \sum_{i=1}^{n} \frac{b-a}{n} f(x_{i-1}) - \sum_{i=1}^{n} \frac{b-a}{n} f(x_i) = \frac{b-a}{n} \sum_{i=1}^{n}(f(x_{i-1}) - f(x_i)) = \frac{b-a}{n}(f(x_0) - f(x_n))$.

By the definition of a partition, $x_0 = a$, $x_n = b$, so $U(P_n, f) - L(P_n, f) = \frac{b-a}{n}(f(a) - f(b))$.

And $n > \frac{1}{\varepsilon}(b-a)(f(a) - f(b)) \implies \left(\frac{b-a}{n}\right)(f(a) - f(b)) < \varepsilon$.

So $U(P_n, f) - L(P_n, f) < \varepsilon$. By the ε reformulation of Riemann integrability, f is integrable on $[a, b]$.

> **Final, #8.** Let $f : \mathbb{R} \to \mathbb{R}$ be twice differentiable. Suppose that $f''(x) \geq 0$ for all $x \in \mathbb{R}$. Prove that if $f(0) = 0$ and $f(1) = 1$, then $f(x) \geq x$ for all $x \geq 1$.

Let f be a fxn. s.t. the stated conditions hold. By MVT, $\exists\, c \in (0, 1)$ s.t. $f'(c) = \frac{f(1)-f(0)}{1-0} = 1$. And $f''(x) \geq 0\ \forall\, x \in \mathbb{R}$. Let $x, y \in \mathbb{R}$ s.t. $x > y$. Then $\exists\, d \in (x, y)$ s.t. $f''(d) = \frac{f'(y)-f'(x)}{y-x} \geq 0 \implies f'(y) \geq f'(x)$. So f' is (weakly) inc.

Then $\forall\, x > c$, $f'(c) \geq 1$.

Let $x = 1$. Then $f(x) = f(1) = 1$.

Let $x > 1$. Then $\exists\, d \in (1, x)$ s.t. $f'(d) = \frac{f(x)-1}{x-1} \geq 1$. Then

$$\begin{aligned} &f(x) - 1 \geq x - 1 \\ \implies\ &f(x) \geq x. \end{aligned}$$

So $\forall\, x \geq 1$, $f(x) \geq x$.

$\square$

5 Foundational Probability Theory

I completed the following proofs while self-studying basic probability theory as a refresher from my probability and statistics course that I took more than 3 years ago.

De Morgan's Law.
(a) $(A \cap B)^C = A^C \cup B^C$.
(b) $(A \cup B)^C = A^C \cap B^C$.

Proof of (a): ($\subseteq$) Let $\alpha \in (A \cap B)^C$. Then $\alpha \notin A \vee \alpha \notin B$, so $\alpha \in A^C \vee \alpha \in B^C \Rightarrow \alpha \in A^C \cup B^C$. Thus, $(A \cap B)^C \subseteq A^C \cup B^C$.

($\supseteq$) Let $\alpha \in A^C \cup B^C$. Then $\alpha \notin A \vee \alpha \notin B$, so $\alpha \notin A \cap B \Rightarrow \alpha \in (A \cap B)^C$. Thus, $A^C \cup B^C \subseteq (A \cap B)^C$.

We conclude that $(A \cap B)^C = A^C \cup B^C$. ■

Proof of (b): ($\subseteq$) Let $\alpha \in (A \cup B)^C$. Then $\alpha \notin A \cup B$, so $\alpha \notin A \vee \alpha \notin B \Rightarrow \alpha \in A^C \wedge \alpha \in B^C \Rightarrow \alpha \in A^C \cap B^C$. Thus, $(A \cup B)^C \subseteq A^C \cap B^C$.

($\supseteq$) Let $\alpha \in A^C \cap B^C$. Then $\alpha \in A^C \wedge \alpha \in B^C$, so $\alpha \notin A \wedge \alpha \notin B \Rightarrow \alpha \notin A \cup B \Rightarrow$
$\alpha \in (A \cup B)^C$. Thus, $A^C \cap B^C \subseteq (A \cup B)^C$.

We conclude that $(A \cup B)^C = A^C \cap B^C$. ■

Partition Theorem. Let a set of events $\{B_1, B_2, \ldots\}$ be a partition of sample space S. Then for any event A,

$$A = \bigcup_{i=1}^{\infty} (A \cap B_i)$$

Proof: Let $\alpha \in A$. Since $A \subseteq S$, $\alpha \in S$. Then since $\{B_1, B_2, \ldots\}$ is a partition of S, $\alpha \in B_i$ for some $B_i \in \{B_1, B_2, \ldots\}$. Then $\alpha \in \bigcup_{i=1}^{\infty} \{B_1, B_2, \ldots\} \wedge \alpha \in A$, so $\alpha \in \bigcup_{i=1}^{\infty} (A \cap B_i)$. Therefore, $A \subseteq \bigcup_{i=1}^{\infty} (A \cap B_i)$.

Conversely, let $\alpha \in \bigcup_{i=1}^{\infty} (A \cap B_i)$. Then for some $B_i \in \{B_1, B_2, \ldots\}$, $\alpha \in B_i \wedge \alpha \in A$. Thus, $\bigcup_{i=1}^{\infty} (A \cap B_i) \subseteq A$, so we conclude

$$A = \bigcup_{i=1}^{\infty} (A \cap B_i)$$

■

Law of Total Probability. Let $\{B_i\}$ partition S. Then for some event $A \subseteq S$,

$$\Pr(A) = \sum_{j=1}^{\infty} \Pr(A \mid B_j) \cdot \Pr(B_j)$$

Proof: Let $\{B_i\}$ partition S and $A \subseteq S$ be an event. By the partition theorem,

$$A = \bigcup_{j=1}^{\infty} (A \cap B_j).$$

Since $\{(A \cap B_i)\}$ are disjoint, we have the following by the axioms of the probability function:

$$\Pr\left[\bigcup_{j=1}^{\infty}(A \cap B_j)\right] = \sum_{j=1}^{\infty} \Pr(A \cap B_j)$$

Since $\Pr(A \cap B_i) = \Pr(A \mid B_i) \cdot \Pr(B_i)$ for each B_i, we can conclude that

$$\Pr(A) = \sum_{j=1}^{\infty} \Pr(A \mid B_j) \cdot \Pr(B_j)$$

■

Counting Rule. For K sets $S_1, S_2, \ldots, S_K$, each S_i with size n_i, there are $\prod_{j=1}^{K} n_j$ possible sets C that can be created by selecting one element x_i from each set S_i to occupy the ith place of C.

Proof: Proceed by induction. In the base case, let $K = 1$ and denote $n_1 = n$. Sps. $S_1 = \{x_1, x_2, \ldots, x_n\}$. Then C can contain one element of the set S_1 as its first entry. Therefore, $C = \{x_i\}$ for any $i = 1, 2, \ldots, n$. Then C can be any of n distinct sets.

Sps. that for $K = m$ for some $m > 1$, C can be any of $N_m = \prod_{j=1}^{m} n_j$ sets. Consider the case when $K = m + 1$ and a set D that can be created by selecting one element from each set S_i to occupy the ith place of D. Since C represents all of the possible sets that can be created from the first m sets, $D = C \cup \{x_i\}$ for all $x_i \in S_{m+1}$. x_i can take n_{m+1} values. Then for each of the N_m possible values of C, D can take n_{m+1} values. Thus, D can take $N_m \cdot n_{m+1} = \prod_{j=1}^{m+1} n_j$ values. ■

Binomial Theorem. For all $N \in \mathbb{Z}_{\geq 0}$

$$(a+b)^N = \sum_{K=0}^{N} \binom{N}{K} a^K \cdot b^{N-K}$$

Proof: Proceed by induction. In the base case, let $N = 0$. Then $(a+b)^N = 1$. At the same time, $\sum_{K=0}^{0} \binom{0}{K} a^K b^{0-K} = \frac{0!}{0!} a^0 b^0 = 1$.

Assume that for some $m > 0$,

$$(a+b)^m = \sum_{K=0}^{m} \binom{m}{K} a^K \cdot b^{m-K}$$

Then

$$\begin{aligned}
(a+b)^{m+1} &= (a+b) \sum_{K=0}^{m} \binom{m}{K} a^K \cdot b^{m-K} \\
&= \sum_{K=0}^{m} \binom{m}{K} a^{K+1} \cdot b^{m-K} + \sum_{K=0}^{m} \binom{m}{K} a^K \cdot b^{m-K+1} \\
&= \sum_{K=0}^{m} \frac{m!}{k!(m-k)!} a^{K+1} \cdot b^{m-K} + \sum_{K=0}^{m} \frac{m!}{k!(m-k)!} a^K \cdot b^{m-K+1}
\end{aligned}$$

Then for some $0 \leq R \leq m+1$, the coefficient on $a^R b^{m+1-R}$ is

$$\frac{m!}{(R-1)!(m-R+1)!} + \frac{m!}{R!(m-R)!} = \binom{m}{R-1} + \binom{m}{R} = \binom{m+1}{K}$$

by Pacal's Identity.

Then

$$(a+b)^{m+1} = \sum_{K=0}^{m+1} \binom{m+1}{K} a^K \cdot b^{m+1-K}$$

$\blacksquare$