Contents

Chapter 1: The Real and Complex Number Systems

1. If $r \in \mathbb{Q}, r \neq 0$, and x is irrational, prove that $r + x$ and rx are irrational.

   Solution: Suppose that $r + x \in \mathbb{Q}$. Since $\mathbb{Q}$ is a field, $-r \in \mathbb{Q}$ and we have that $-r + r + x = x \in \mathbb{Q}$, a contradiction. Therefore $r + x \in \mathbb{R}\backslash\mathbb{Q}$, i.e. $r + x$ is irrational. Similarly, suppose that $rx \in \mathbb{Q}$. Since $\mathbb{Q}$ is a field, $\frac{1}{r} \in \mathbb{Q}$ and we have that $\frac{1}{r} \cdot rx = x \in \mathbb{Q}$, a contradiction. Therefore $rx \in \mathbb{R}\backslash\mathbb{Q}$, i.e. rx is irrational. $\square$

2. Prove that there is no rational number whose square is 12.

   Solution: Suppose there exist $m, n \in \mathbb{Z}$ such that neither are divisible by 3 (i.e. we assume their ratio is in simplified form) and

$$\frac{m}{n} = \sqrt{12} \implies m^2 = 12n^2 = 2^2 \cdot 3n^2.$$

   Therefore, $3 \mid m^2 \implies 3 \mid m$ because 3 is prime. Therefore, $3^2 \mid m^2 \implies 3^2 \mid 12n^2 \implies 3 \mid n^2 \implies 3 \mid n$, again because 3 is prime. This contradicts that m and n were chosen such that their ratio is in simplified form because both are divisible by 3.

   We can investigate this further, discovering that $\mathbb{Q}$ lacks the greatest/least upper bound property, as is done analogously in Example 1.1 on page 2 of the text. Let $A = \{p \in \mathbb{Q} \mid p > 0, p^2 < 12\}$ and $B = \{p \in \mathbb{Q} \mid p > 0, p^2 > 12\}$. Define

$$q = p - \frac{p^2 - 12}{p + 12} = \frac{p(p + 12) - p^2 + 12}{p + 12} = \frac{12(p + 1)}{p + 12}. \tag{2.1}$$

   Then, we have that

$$q^2 - 12 = \frac{144(p + 1)^2 - 12(p + 12)^2}{(p + 12)^2} = \frac{132(p^2 - 12)}{(p + 12)^2}. \tag{2.2}$$

   Now, if $p \in A$ then $p^2 - 12 < 0$ so by Equation (2.1) $q > p$ and by Equation (2.2) $q^2 < 12$ so $q \in A$. Therefore, A contains no largest number. Similarly, if $p \in B$ then $p^2 - 12 > 0$ so by Equation (2.1) $0 < q < p$ and by Equation (2.2) $q^2 > 12$ so $q \in B$. Therefore, B contains no smallest number. The elements of B are precisely the upper bounds of A, therefore, since B has no least element, A has no least upper bound. Similarly, the elements of A are the lower bounds of B and since A has no greatest element, B has no greatest upper bound. $\square$

3. Prove the following using the axioms of multiplication in a field.

   (a) If $x \neq 0$ and $xy = xz$ then $y = z$.

   (b) If $x \neq 0$ and $xy = x$ then $y = 1$.

   (c) If $x \neq 0$ and $xy = 1$ then $y = 1/x$.

   (d) If $x \neq 0$ the $1/(1/x) = x$.

Solution: To show part (a) we simply use associativity of multiplication and the existence of multiplicative inverses.

$$y = 1y = \left(\frac{1}{x} \cdot x\right)y = \frac{1}{x}(xy) = \frac{1}{x}(xz) = \left(\frac{1}{x} \cdot x\right)z = 1z = z.$$

The statement of (b) and (c) follow directly from (a) if we let $z = 1$ and $z = 1/x$, respectively. Part (d) follows from (c) if we replace x with $1/x$ and y with x. $\square$

4. Let E be a nonempty subset of an ordered set; suppose α is a lower bound of E and β is an upper bound of E. Prove that $\alpha \leq \beta$.

   Solution: Since E is nonempty, let $x \in E$. Then, by definition of upper and lower bounds, we have that $\alpha \leq x \leq \beta$. $\square$

5. Let A be a nonempty set of real numbers which is bounded below. Let $-A$ be the set of all numbers $-x$, where $x \in A$. Prove that $\inf A = -\sup(-A)$.

   Solution: Let $\beta = \sup(-A)$, which exists because A is a nonempty subset of the real numbers which is bounded below, hence $-A$ is a nonempty subset of the real numbers which is bounded above. Thus,

   $$\forall\, x \in A, -x \leq \beta \implies x \geq -\beta$$

   so $-\beta$ is a lower bound for A. Now, suppose there exists γ such that $\gamma > -\beta$ and $\forall\, x \in A, x \geq \gamma$. This implies that $-\gamma < \beta$ and $-x \leq -\gamma$, $\forall\, x \in A$. But then, $-\gamma$ is a smaller upper bound (than β) for $-A$, contradicting that $\beta = \sup(-A)$. Therefore no such γ exists and so $-\sup(-A) = -\beta = \inf(A)$. The reversing of inequalities throughout follows from Proposition 1.18 on page 8 of the text. $\square$

6. Fix $b > 1$.

   (a) If m, n, p, q are integers, $n > 0, q > 0$, and $r = m/n = p/q$, prove that

   $$(b^m)^{1/n} = (b^p)^{1/q}.$$

   Hence it makes sense to define $b^r = (b^m)^{1/n}$.

   (b) Prove that $b^{r+s} = b^r b^s$ if r and s are rational.

   (c) If x is real, define $B(x)$ to be the set of all numbers b^t, where t is rational and $t \leq x$. Prove that

   $$b^r = \sup B(r)$$

   where r is rational. Hence, it makes sense to define

   $$b^x = \sup B(x)$$

   for every real x.

   (d) Prove that $b^{x+y} = b^x b^y$ for every real x and y.

   Solution:

   (a) Frst observe that $(b^{n_1})^{n_2} = b^{n_1 n_2}$ for $n_1, n_2 \in \mathbb{Z}$ by direct expansion of either term. Let $y^n = b^m$ so that $y = (b^m)^{1/n}$. Then, $y^{nq} = b^{mq} = b^{np} \implies (y^q)^n = (b^p)^n$, because $m/n = p/q \implies mq = np$. By Theorem 1.12 on page 10, since n^{th} roots are unique, we have $y^q = b^p$. Then, $(b^m)^{1/n} = y = (b^p)^{1/q}$.

(b) Let $r, s \in \mathbb{Q}$ and write $r = m/n$ and $s = p/q$ for $m, n, p, q \in \mathbb{Z}$. Then, $r+s = m/n+p/q = (mq + np)/nq$. Note, for $n_1, n_2 \in \mathbb{Z}$ we have $b^{n_1}b^{n_2} = b^{n_1+n_2}$ by direct expansion. Then, by (a), $b^{r+s} = (b^{mq+np})^{1/nq} = (b^{mq}b^{np})^{1/nq} = (b^{mq})^{1/nq}(b^{np})^{1/nq}$, where the last equality follows from the Corollary to Theorem 1.12 on page 11 of the text. Again using the result of (a) we can simplify this last expression and obtain $(b^{mq})^{1/nq}(b^{np})^{1/nq} = (b^m)^{1/n}(b^p)^{1/q} = b^r b^s$, showing (b).

(c) Let $t, r \in \mathbb{Q}$ and write $t = p/q, r = m/n$, where $p, q, m, n \in \mathbb{Z}$. Then, $t < r \implies p/q < m/n \implies pn < mq$. Thus, since $b > 1$, we have that $b^{pn} < b^{mq} \implies b^p < (b^{mq})^{1/n} \implies (b^p)^{1/q} < ((b^{mq})^{1/n})^{1/q}$. Using the result of part (a), we can simplify the last expression to obtain $b^t = (b^p)^{1/q} < (b^m)^{1/n} = b^r$, showing that

$$b^r > \{b^t \mid t < r\} \implies b^r = \sup \{b^t \mid t \leq r\} = B(r)$$

and proving part (c).

(d) We first observe that $b^r b^t = b^{r+t}$ by part (b) and so

$$b^x b^y = \sup_{r\leq x, t\leq y} b^r b^t = \sup_{r\leq x, t\leq y} b^{r+t} \leq \sup_{r+t\leq x+y} b^{r+t} = b^{x+y}.$$

Here, $\sup_{r\leq x,t\leq y} b^{r+t} \leq \sup_{r+t\leq x+y} b^{r+t}$ because we are taking the sup over a more restricted set on the right since $r \leq x, t \leq y \implies r + t \leq x + y$ but $r + t \leq x + y$ doesn't imply $r \leq x$ and $t \leq y$. But, we in fact have equality because if $\sup_{r\leq x,t\leq y} b^{r+t} < \sup_{r+t\leq x+y} b^{r+t}$ then there exist $r', t' \in \mathbb{Q}$ such that $r' + t' \leq x + y$ and

$$\sup_{\rho\leq x,\tau\leq y} b^{\rho+\tau} < b^{r'+t'} \leq \sup_{r+t\leq x+y} b^{r+t}. \tag{6.1}$$

If $r' + t' = x + y$ we get a contradiction to Equation (6.1) by choosing $\rho = x, \tau = y$ (this follows from Exercise 1 above). Otherwise, $r' + t' < x + y$ and then we can choose $\rho < x, \tau < y$ such that $r' + t' < \rho + \tau < x + y$, because $\mathbb{Q}$ is dense in $\mathbb{R}$. But then, $b^{\rho+\tau} > b^{r'+t'}$, also contradicting Equation (6.1). Therefore we have equality and part (d) is shown. □

7. Omitted.

8. Prove that no order can be defined in the complex field that turns it into an ordered field. *Hint:* -1 is a square.

   Solution: Suppose there is an ordered defined on the complex field that turns it into an ordered field. Then, we get a contradiction to Proposition 1.18(d) on page 8 because $i \neq 0$ yet $i^2 = -1 < 0$, where $-1 < 0$ is forced by part (a) of the same Proposition. □

9. Suppose $z = a + bi, w = c + di$. Define $z < w$ if $a < c$, and also if $a = c$ but $b < d$. Prove that this turns the set of all complex numbers into an ordered set. (This type of order relation is called a *dictionary order*, or *lexicographic order*, for obvious reasons.) Does this ordered set have the least-upper-bound property?

   Solution: Let $z = a + bi, w = c + di \in \mathbb{C}$. Then, since $\mathbb{R}$ is an ordered set, either $a < b$, $a = b$, or $a > b$. If $a < b$, then $z < w$. If $a > b$ then $z > w$. If $a = b$ we check the relationship between b and d and similarly will conclude that either $z < w$ if $b < d$, $z > w$ if $b > d$, or $z = w$ if $b = d$. To check transitivity, let $u = e + fi \in \mathbb{C}$ and suppose that $z < w$ and $w < u$. Thus $a < c$ or $a = c$ and $b < d$. Also, $c < e$ or $c = e$ and $d < f$. If $a < c$ and $c < e$ then $a < e$ because $\mathbb{R}$ is an ordered set and in this case $z < u$. If $a < c$ and $c = e$ with $d < f$ then $a < c = e$ so $z < u$. If $a = c$ and $b < d$ with $c < e$ then $a = c < e$ and $z < u$. If $a = c$ and $b < d$ with $c = e$ and $d < f$ then $a = e$ but $b < d < f$ so $b < f$ because $\mathbb{R}$ is an ordered set and therefore $z < u$. Thus, $\mathbb{C}$ is an ordered set under this order.

This ordered set does have the least-upper-bound property. To see this, let $E \subseteq \mathbb{C}$ and define $\alpha = \sup\{a \mid a+bi \in E\}$ and $\beta = \sup\{b \mid a+bi \in E\}$. Then, it is clear $\alpha+\beta i \in \mathbb{C}$ because $\mathbb{R}$ has the least-upper bound property and furthermore $\forall\, a+bi \in E, a+bi \leq \alpha+\beta i$, (because $a \leq \alpha$ and $b \leq \beta$) so $\alpha+\beta i$ is an upper bound for E. Now, suppose there exists $\gamma+\delta i \in \mathbb{C}$ such that $\gamma+\delta i < \alpha+\beta i$ and $\gamma+\delta i$ is an upper bound for E. If $\gamma < \alpha$, then we have $a \leq \gamma < \alpha \;\forall\, a$ such that $a+bi \in E$, contradicting the definition of α. An identical contradiction with the definition of β occurs if $\gamma = \alpha$ and $\delta < \beta$. Therefore, $\alpha+\beta i = \sup(E) \in \mathbb{C}$ and this ordered set does indeed have the least-upper-bound property. □

10. Suppose $z = a+bi, w = u+vi$, and

$$a = \left(\frac{|w|+u}{2}\right)^{1/2}, \qquad b = \left(\frac{|w|-u}{2}\right)^{1/2}.$$

Prove that $z^2 = w$ if $v \geq 0$ and that $(\bar{z})^2 = w$ if $v \leq 0$. Conclude that every complex number (with one exception!) has two complex square roots.

Solution: First, suppose that $v \geq 0$. Then

$$\begin{aligned} z^2 = (a+bi)(a+bi) = (a^2-b^2)+2abi &= \left(\frac{(u^2+v^2)^{1/2}+u}{2}\right) - \left(\frac{(u^2+v^2)^{1/2}-u}{2}\right) + 2abi \\ &= u + 2\left(\frac{(u^2+v^2)-u^2}{4}\right)^{1/2} i \\ &= u+vi \\ &= w \end{aligned}$$

where we needed to use that $v \geq 0$ because $(v^2)^{1/2} = |v|$. If $v \leq 0$ then

$$\begin{aligned} (\bar{z})^2 = (a-bi)(a-bi) = (a^2-b^2)-2abi = u-2abi &= u - 2\left(\frac{(u^2+v^2)-u^2}{4}\right)^{1/2} i \\ &= u - |v|i \\ &= w. \end{aligned}$$

Then it is clear each complex number (except 0 of course!) has two complex roots because we can take either *both* positive or *both* negative square roots defining a and b above since if we choose $-a$ and $-b$, the minus signs vanish in $w = (a^2-b^2) \pm 2abi$. It is clear we can't have more than 2 complex roots because if we mixed taking the positive square root defining a and the negative square root defining b, or vice versa, the minus signs would not cancel in the $\pm 2abi$ term. □

11. If z is a complex number, prove that there exists an $r \geq 0$ and a complex number w with $|w| = 1$ such that $z = rw$. Are w and r always uniquely determined by z?

Solution: Simply let $w = z/|z|$ and $r = |z|$. w is not always uniquely determined since if $z = 0$ we can choose any $w \in \mathbb{C}$ such that $|w| = 1$ and $r = 0$. r is always uniquely determined since $|z| = |rw| = |r||w| = |r| = r$ because $r \geq 0$ by hypothesis. If $z \neq 0$ then w is uniquely determined because it must lie on the radius which z lies on, where it intersects the unit circle. To see this, let $z = a + bi$ and $w = c + di$. Then, $a + bi = rc + rdi$ and therefore $c = a/r$ and $d = b/r$. Hence, since r is uniquely determined, so is w provided $z \neq 0$. □

12. If $z_1, \ldots, z_n$ are complex, prove that

$$|z_1 + \cdots + z_n| \leq |z_1| + \cdots + |z_n|.$$

Solution: We proceed by induction on n. For $n = 1$ there is nothing to show and the case $n = 2$ is the triangle inequality which is given by Theorem 1.33 on page 14 of the text. Suppose the result holds for $n - 1$. Then

$$|z_1 + \cdots + z_n| = |z_1 + (z_2 + \cdots + z_n)| \leq |z_1| + |z_2 + \cdots + z_n| \leq |z_1| + |z_2| + \cdots + |z_n|$$

where the last inequality follows by the inductive hypothesis applied to the term $|z_2 + \cdots + z_n|$. □

13. If x, y are complex, prove that

$$\big||x| - |y|\big| \leq |x - y|.$$

Solution: Let $x, y \in \mathbb{C}$ and let $z = y - x$. Then, by the triangle inequality we have that

$$|x + z| \leq |x| + |z| \implies |x + z| - |x| \leq |z|,$$

hence substituting in the expression for z we obtain that

$$|y| - |x| \leq |y - x|. \tag{13.1}$$

Now, set $z = x - y$ and proceeding analogously we find that

$$|y + z| \leq |y| + |z| \implies |y + z| - |y| \leq |z|,$$

hence substituting in the expression for z we obtain that

$$|x| - |y| \leq |x - y| = |y - x|,$$

thus combined with Equation (13.1) above we see that

$$\big||x| - |y|\big| \leq |x - y|,$$

as desired.

Alternate Solution: Let $x = a + bi$ and $y = c + di$ with $a, b, c, d \in \mathbb{R}$. By the triangle inequality (Theorem 1.33(e) on page 14 of the text) we have that

$$|x + y| \leq |x| + |y| \iff |x + y|^2 \leq |x|^2 + |y|^2 + 2|x||y|.$$

Substituting our expressions for x and y and noting that by Definition 1.32 on page 14 of the text we have that $|x| = (a^2 + b^2)^{1/2}$ we find

$$\begin{aligned} a^2 + c^2 + 2ac + b^2 + d^2 + 2bd &= (a + c)^2 + (b + d)^2 \\ &= |(a + c) + (b + d)i|^2 \\ &= |x + y|^2 \\ &\leq |x|^2 + |y|^2 + 2|x||y| \\ &= a^2 + b^2 + c^2 + d^2 + 2(a^2 + b^2)^{1/2}(c^2 + d^2)^{1/2} \end{aligned}$$

which immediately implies that

$$ac + bd \le (a^2 + b^2)^{1/2}(c^2 + d^2)^{1/2}. \tag{13.2}$$

Therefore,

$$\begin{aligned}
\big||x| - |y|\big|^2 = \big|(a^2+b^2)^{1/2} - (c^2+d^2)^{1/2}\big|^2 &= \big[(a^2+b^2)^{1/2} - (c^2+d^2)^{1/2}\big]^2 \\
&= a^2 + b^2 + c^2 + d^2 - 2\big((a^2+b^2)^{1/2}(c^2+d^2)^{1/2}\big) \\
&\le a^2 + b^2 + c^2 + d^2 - 2(ac + bd) \\
&= a^2 + c^2 - 2ac + b^2 + d^2 - 2bd \\
&= (a-c)^2 + (b-d)^2 \\
&= |(a-c) + (b-d)i|^2 \\
&= |(a+bi) - (c+di)|^2 \\
&= |x - y|^2
\end{aligned}$$

hence we have

$$\big||x| - |y|\big| \le |x - y|.$$

Since the triangle inequality was the fundamental trick in both solutions, we might suspect that there is a geometric interpretation. There is and it is easy to visualize if we treat x and y as vectors in $\mathbb{R}^2$. Then, this merely states that the magnitude of the vector given by their difference is greater than or equal to the difference in their magnitudes. This makes sense if we consider x and y to point in opposite directions (i.e. $y = -\alpha x$ where $\alpha > 0$). Then the magnitude of the vector given by their difference is the sum of their magnitudes which is greater than or equal to the difference of their magnitudes because magnitudes are positive or zero. When x and y point in the same direction, we get equality, otherwise we have inequality. When they point in opposite directions we get "maximal inequality". □

14. If z is a complex number such that $|z| = 1$, that is, such that $z\overline{z} = 1$, compute

$$|1 + z|^2 + |1 - z|^2.$$

Solution: Write $z = a + bi$ with $a, b \in \mathbb{R}$. Then

$$\begin{aligned}
|1+z|^2 + |1-z|^2 = |(1+a) + bi|^2 + |(1-a) + bi|^2 &= (1+a)^2 + b^2 + (1-a)^2 + b^2 \\
&= 1 + a^2 + 2a + b^2 + 1 + a^2 - 2a + b^2 \\
&= 2 + 2a^2 + 2b^2 \\
&= 2(a^2 + b^2) + 2 \\
&= 2|z|^2 + 2 \\
&= 4. \ \square
\end{aligned}$$

15. Under what conditions does equality hold in the Schwarz inequality?

Solution: If $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ are complex numbers then the Schwarz inequality states that

$$\left|\sum_{i=1}^n a_i \overline{b_i}\right|^2 \le \sum_{i=1}^n |a_i|^2 \sum_{i=1}^n |b_j|^2.$$

We claim that equality holds if there exists $\lambda \in \mathbb{C}$ such that $b_i = \lambda a_i \; \forall \; i = 1, \ldots, n$. Supposing this condition holds we have

$$\left|\sum_{i=1}^{n} a_i \overline{b_i}\right|^2 = \left|\sum_{i=1}^{n} a_i \overline{a_i}\overline{\lambda}\right|^2 = |\lambda|^2 \left|\sum_{i=1}^{n} |a_i|^2\right|^2 = |\lambda|^2 \left(\sum_{i=1}^{n} |a_i|^2\right)\left(\sum_{i=1}^{n} |a_i|^2\right)$$
$$= \left(\sum_{i=1}^{n} |a_i|^2\right)\left(\sum_{i=1}^{n} |\lambda|^2 |a_i|^2\right)$$
$$= \left(\sum_{i=1}^{n} |a_i|^2\right)\left(\sum_{i=1}^{n} |\lambda a_i|^2\right)$$
$$= \sum_{i=1}^{n} |a_i|^2 \sum_{i=1}^{n} |b_i|^2. \; \square$$

16. Suppose $k \geq 3, \mathbf{x}, \mathbf{y} \in \mathbb{R}^k$, $|\mathbf{x} - \mathbf{y}| = d > 0$, and $r > 0$. Prove:

   (a) If $2r > d$, there are infinitely many $\mathbf{z} \in \mathbb{R}^k$ such that
   $$|\mathbf{z} - \mathbf{x}| = |\mathbf{z} - \mathbf{y}| = r.$$
   (b) If $2r = d$, there is exactly one such $\mathbf{z}$.
   (c) If $2r < d$, there are no such $\mathbf{z}$.

   How must these statements be modified if k is 2 or 1?

   Solution:

   (a) If $\mathbf{z} \in \mathbb{R}^k$ such that $|\mathbf{z} - \mathbf{x}| = |\mathbf{z} - \mathbf{y}| = r$ then $\mathbf{z} \in \partial B(\mathbf{x}, r) \cap \partial B(\mathbf{y}, r)$, where $\partial B(\mathbf{x}, r) = \{\mathbf{z} \in \mathbb{R}^k \mid |\mathbf{z} - \mathbf{x}| = r\}$ is the boundary of the ball of radius r centered at $\mathbf{x}$. If $2r > d \implies r > d/2$ then the two boundaries intersect in a surface (of dimension $k - 2$) and thus there are infinitely many points in this intersection which satisfy the given equation because $k \geq 3$.
   (b) If $r = d/2$ then the two boundaries intersect in a single point, the midpoint of the segment joining $\mathbf{x}$ and $\mathbf{y}$.
   (c) If $r < d/2$ then the boundaries never intersect and so no $\mathbf{z} \in \mathbb{R}^k$ can be on both boundaries simultaneously.

   If $k = 1$ then there are no $z \in \mathbb{R}$ satisfying both equations simultaneously unless $r = d/2$. This is because there are only two points that are a distance r from x for a given $x \in \mathbb{R}$. Without loss of generality we can assume $x < y$. If $r = d/2$, we see that $z = x + r$ is the only point that is a distance r from both x and y. If $r > d/2$ or $r < d/2$ then the boundaries (set's of two points) never intersect. This is because the points of distance r from x are $x + r$ and $x - r$. But, $y - (x + r) = y - x - r = d - r < d - d/2 = d/2 < r$ so $x + r$ is not of distance r from y. Similarly $y - (x - r) = y - x + r = d + r > r$ so $x - r$ is also not of distance r from y. Analogous results hold when $r < d/2$ with inequalities reversed.

   If $k = 2$ then when $r > d/2$ instead of infinitely many points in the intersection of the boundaries, there are just two since in this case the boundaries of the ball's around $\mathbf{x}$ and $\mathbf{y}$ are circles and two different circles can maximally intersect in only 2 points. The cases $r = d/2$ and $r < d/2$ give the same results as they did for $k \geq 3$. $\square$

17. Prove that
$$|\mathbf{x}+\mathbf{y}|^2 + |\mathbf{x}-\mathbf{y}|^2 = 2|\mathbf{x}|^2 + 2|\mathbf{y}|^2$$
if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^k$. Interpret this geometrically, as a statement about parallelograms.

Solution: If we write $\mathbf{x} = (x_1, \dots, x_k)$ and $\mathbf{y} = (y_1, \dots, y_k)$ then expanding the left hand side gives

$$\begin{aligned}
|\mathbf{x}+\mathbf{y}|^2 + |\mathbf{x}-\mathbf{y}|^2 &= \sum_{i=1}^{k}(x_i + y_i)^2 + \sum_{i=1}^{k}(x_i - y_i)^2 \\
&= \sum_{i=1}^{k}(x_i + y_i)^2 + (x_i - y_i)^2 \\
&= \sum_{i=1}^{k}(x_i^2 + 2x_iy_i + y_i^2 + x_i^2 - 2x_iy_i + y_i^2) \\
&= \sum_{i=1}^{k}(2x_i^2 + 2y_i^2) \\
&= 2\sum_{i=1}^{k} x_i^2 + 2\sum_{i=1}^{k} y_i^2 \\
&= 2|\mathbf{x}|^2 + 2|\mathbf{y}|^2.
\end{aligned}$$

If we consider the vectors $\mathbf{x}$ and $\mathbf{y}$ as representing two sides of a parallelogram then $\mathbf{x}-\mathbf{y}$ and $\mathbf{x}+\mathbf{y}$ represent the diagonals. The equation above is then seen to be nothing more then the generalized Pythagorean Theorem. Considering the case $k = 2$ the equation is merely the sum of the ordinary Pythagorean Theorem for the two diagonals. Since for both diagonals the sides of the triangle for which they are the hypotenuse are $\mathbf{x}$ and $\mathbf{y}$ we have that

$$\begin{aligned}
|\mathbf{x}+\mathbf{y}|^2 &= |\mathbf{x}|^2 + |\mathbf{y}|^2 \\
|\mathbf{x}-\mathbf{y}|^2 &= |\mathbf{x}|^2 + |\mathbf{y}|^2
\end{aligned}$$

hence our original equation is just the sum of these two. □

18. If $k \geq 2$ and $\mathbf{x} \in \mathbb{R}^k$, prove that there exists $\mathbf{y} \in \mathbb{R}^k$ such that $\mathbf{y} \neq 0$ but $\mathbf{x} \cdot \mathbf{y} = 0$. Is this also true if $k = 1$?

Solution: We can see right away that the statement must fail for $k = 1$ because in this case the dot product is nothing more than multiplication of real numbers and since the reals are a field $xy = 0 \implies x = 0$ or $y = 0$, i.e. fields have no zero divisors. This means there is no notion of perpendicularity in a 1-dimensional space as we'd expect since perpendicularity implies linear independence and hence extra dimension.

For $k \geq 2$ let $\mathbf{x} = (x_1, \dots, x_k)$ and without loss of generality suppose $\mathbf{x} \neq 0$ because otherwise any non-zero $\mathbf{y} \in \mathbb{R}^k$ will suffice. We consider two cases based on the parity of k. If k is even then we let $\mathbf{y} = (x_2, -x_1, x_4, -x_3, \dots, x_{j+1}, -x_j, \dots, x_k, -x_{k-1})$ where x_{j+1} is in the j^{th} slot. We then have

$$\mathbf{x} \cdot \mathbf{y} = x_1x_2 - x_2x_1 + \cdots + x_jx_{j+1} - x_{j+1}x_j + \cdots + x_{k-1}x_k - x_kx_{k-1} = 0$$

and $\mathbf{y} \neq 0$ because $\mathbf{x} \neq 0$ by hypothesis. Now, suppose that k is odd. Since $\mathbf{x} \neq 0$ by hypothesis we have that $x_j \neq 0$ for some $1 \leq j \leq k$. Supposing that x_j is not the only non-zero entry, we repeat the same argument and choose $\mathbf{y} = (x_2, -x_1, \ldots, x_{j+1}, 0, -x_{j-1}, \ldots, x_k, -x_{k-1})$ where the 0 is in the j^{th} slot and $\mathbf{y} \neq 0$ because we have supposed that x_j is not the only non-zero entry. Then we have

$$\mathbf{x} \cdot \mathbf{y} = x_1x_2 - x_2x_1 + \cdots + x_{j-1}x_{j+1} + 0 - x_{j+1}x_{j-1} + \cdots + x_{k-1}x_k - x_kx_{k-1} = 0.$$

If x_j is the only non-zero entry then we simply choose $\mathbf{y} = (1, \ldots, 1, 0, 1, \ldots, 1)$ where the 0 is in the j^{th} slot. $\square$

Chapter 2: Basic Topology

1. Prove that the empty set is a subset of every set.

   Solution: Let A be a set and let $\emptyset$ be the empty set. To show $\emptyset \subset A$ we need to show that $\forall\, p \in \emptyset, p \in A$. Since there are no $p \in \emptyset$, this statement is vacuously true and therefore $\emptyset \subset A$. $\square$

2. A complex number z is said to be *algebraic* if there are integers $a_0, \ldots, a_n$ not all zero such that
$$a_0z^n + a_1z^{n-1} + \cdots + a_{n-1}z + a_n = 0.$$
   Prove that the set of all algebraic numbers is countable.

   Solution: Let
$$P_n = \left\{ \sum_{i=0}^{n} a_i t^i \;\middle|\; a_i \in \mathbb{Z},\ i = 0, \ldots, n \right\}$$
   so P_n is the set of all degree n or smaller polynomials with integer coefficients. It is clear that
$$P_n \sim \mathbb{Z}^{n+1} \text{ given by}$$
$$\sum_{i=0}^{n} a_i t^i \mapsto (a_0, \ldots, a_n)$$
   where $\sim$ means bijective as sets, so that in particular, P_n is countable for each n. Now, let
$$R_n = \{a \in \mathbb{C} \mid p(a) = 0 \text{ for some } p \in P_n\}$$
   so that R_n is the set of all complex roots of all degree n or smaller polynomials with integer coefficients. Since P_n is countable we have that $P_n \sim \mathbb{N}$ so that in particular to each polynomial in P_n we can associate a unique natural number. Since every degree n or smaller polynomial has at most n complex roots we see that
$$|R_n| \leq \left| \bigcup_{m=0}^{\infty} \{1_m, \ldots, n_m\} \right|$$
   where $m \in \mathbb{N}$ is the index for the elements of the countable set P_n and the set $\{1_m, \ldots, n_m\}$ is a set of n elements for each m representing the maximum number of roots of any degree n or smaller polynomial. Since the set above on the right is a countable union of countable sets it is also countable by Theorem 2.12 on page 20 of the text and so we have that R_n is countable. But since the set of all complex roots of all polynomials with integer coefficients is simply $\cup_{n=0}^{\infty} R_n$, it is also countable by the same Theorem. $\square$

3. Prove that there exist real numbers which are not algebraic.

   Solution: We see that the set of all *real* algebraic numbers is at most countable because it is a subset of the set of all complex algebraic numbers which is countable by Exercise 3 above. But, since $\mathbb{R}$ is uncountable (see Theorem 2.14 on page 30 of the text) there must exist real numbers which are not algebraic. $\square$

4. Is the set of all irrational real numbers countable?

   Solution: We have by the Corollary to Theorem 2.13 on pages 29-30 of the text that $\mathbb{Q}$ is countable. Since $\mathbb{R}$ is uncountable (see Theorem 2.14 on page 30 of the text) we must have that the irrationals are uncountable because the union of two countable sets is countable by Theorem 2.12 on page 29 of the text. $\square$

5. Construct a bounded set of real numbers with exactly three limit points.

   Solution: Consider the set $S = \{1/n \mid n \in \mathbb{N}\} \cup \{1 + 1/n \mid n \in \mathbb{N}\} \cup \{2 + 1/n \mid n \in \mathbb{N}\}$. Then S is bounded because $d(p, 0) < 3 \; \forall \; p \in S$. It is also clear that $0, 1,$ and 2 are limit points of S. No other points are limit points of S since around any other real number we can choose a small enough neighborhood such that it excludes all the points of S, except possibly itself, since we can always choose a radius smaller then $\frac{1}{n} - \frac{1}{n+1}$ for any fixed n. $\square$

6. Let E' be the set of all limit points of a set E. Prove that E' is closed. Prove that E and $\overline{E}$ have the same limit points. (Recall that $\overline{E} = E \cup E'$). Do E and E' always have the same limit points?

   Solution: Let p be a limit point of E'. Thus, for each $n \in \mathbb{N}$ there exists $p_n \in E'$ such that $p_n \neq p$ and $d(p, p_n) < \frac{1}{2n}$. Since each $p_n \in E'$ we have that for each p_n there exists a point $s_n \in E$ such that $s_n \neq p_n$ and $d(p_n, s_n) < \min\{\frac{1}{2n}, d(p_n, p)\}$. Since $d(p_n, s_n) < d(p_n, p)$ we have that $s_n \neq p$. By the triangle inequality we then have that

   $$d(p, s_n) \leq d(p, p_n) + d(p_n, s_n) < \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n}.$$

   Thus, every neighborhood of p contains a point in E not equal to p itself so p is a limit point of E and hence $p \in E'$. Therefore, E' is closed.

   Now, let p be a limit point of E. Then any neighborhood of p contains a point $q \in E$ such that $q \neq p$. Since $E \subset \overline{E}$ we have that $q \in \overline{E}$ so that in particular p is also a limit point of $\overline{E}$. Now, suppose that p is a limit point of $\overline{E}$. Then, for each $n \in \mathbb{N}$ there exists a point $p_n \in \overline{E}$ such that $p_n \neq p$ and $d(p, p_n) < \frac{1}{2n}$. If $p_n \notin E$ then $p_n \in E'$ because $p_n \in \overline{E}$ and we replace p_n with an element $q_n \in E$ such that $q_n \neq p_n$ and $d(p_n, q_n) < \min\{\frac{1}{2n}, d(p_n, p)\}$. Then $q_n \neq p$ because $d(p_n, q_n) < d(p_n, p)$ and $d(p, q_n) \leq d(p, p_n) + d(p_n, q_n) < \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n}$. Therefore, p is also a limit point of E and so E and $\overline{E}$ have the same limit points.

   We can see that E and E' do not always have the same limit points by the simple example $E = (0, 1)$. Here, the limit points of E are 0 and 1 so that $E' = \{0, 1\}$. But, E' itself has no limit points by the Corollary to Theorem 2.20 on pages 32-33 of the text because it is a finite point set. We can see this directly because if $s \neq 0, 1$ then any neighborhood of s of radius $r < \min\{d(s, 0), d(s, 1)\}$ contains no elements of E' and any neighborhood of 0 or 1 of radius less than 1 contains no other points of E' except 0 and 1, respectively, hence E' has no limit points. $\square$

7. Let $A_1, A_2, \ldots$ be subsets of a metric space.

(a) If $B_n = \bigcup_{i=1}^n A_i$, prove that $\overline{B_n} = \bigcup_{i=1}^n \overline{A_i}$, for $n = 1, 2, \ldots$

(b) If $B = \bigcup_{i=1}^\infty A_i$, prove that $\overline{B} \supset \bigcup_{i=1}^\infty \overline{A_i}$.

Show, by an example, that the inclusion in (b) can be proper.

Solution:

(a) We use the notation B' for the set of limit points of B. For (a), suppose that $p \in \overline{B_n}$. Then, either $p \in B_n$ or $p \in B_n'$ (or both). If $p \in B_n = \bigcup_{i=1}^n A_i \implies p \in A_i$ for some i hence $p \in \overline{A_i} \implies p \in \bigcup_{i=1}^n \overline{A_i}$. So, suppose that $p \in B_n'$ so that p is a limit point of B_n. Now, suppose that $p \notin \overline{A_i}\ \forall\, i = 1, \ldots, n$, i.e. that p is not a limit point for any of the A_i's. This means that for each $i\ \exists\ \varepsilon_i > 0$ such that $N(p, \varepsilon_i) \cap A_i \subset \{p\}$. Here $N(p, \varepsilon_i)$ is the ball around p of radius ε_i. That is, there is some non-zero neighborhood of p which intersects A_i in at most p itself. Letting $0 < \varepsilon < \min_i\{\varepsilon_i\}$ we see that $N(p, \varepsilon) \cap A_i \subset \{p\}\ \forall\, i = 1, \ldots, n \implies N(p, \varepsilon) \cap \cup_{i=1}^n A_i = B_n \subset \{p\}$ contradicting that p is a limit point of B_n. Therefore, we must have that $p \in \overline{A_i}$ for at least one i hence $p \in \cup_{i=1}^n \overline{A_i}$ and we obtain the inclusion $\overline{B_n} \subseteq \cup_{i=1}^n \overline{A_i}$. Now, let $p \in \cup_{i=1}^n \overline{A_i}$. Then we have that $p \in \overline{A_i}$ for some i. If $p \in A_i \subset B_n$ then $p \in \overline{B_n}$. If $p \in A_i'$ then every neighborhood of p contains a point $q \in A_i$ such that $q \neq p$. Since $A_i \subset B_n$, we have that $q \in B_n$ also and hence $p \in \overline{B}_n$ showing the other inclusion $\cup_{i=1}^n \overline{A_i} \subseteq \overline{B_n}$ and completing part (a).

(b) The proof is identical to the second inclusion just shown in (a).

As an example, if we let $A_i = \{1/i\}$ and $B = \cup_{i=1}^\infty A_i$ we see that $0 \in B'$ yet $0 \notin A_i'$ for any i because for each i we can choose $0 < \varepsilon < 1/i$ and then $N(0, \varepsilon) \cap A_i = \emptyset$. Hence the inclusion in part (b) is proper. □

8. Is every point of every open set $E \subset \mathbb{R}^2$ a limit point of E? Answer the same question for closed sets in $\mathbb{R}^2$.

Solution: Let $p \in E$ and let $\delta > 0$ be such that $B(p, \delta) \subset E$ so that in particular there exists $q \in B(p, \delta)$ such that $q \neq p$ and of course $q \in E$ because $B(p, \delta) \subset E$. For example, if $p = (p_1, p_2)$ we can choose $q = (p_1 + \delta/2, p_2)$. Now, let $\varepsilon > 0$ be given. If $\varepsilon = \delta$ then by construction $B(p, \varepsilon)$ contains a point of E distinct from p. If $\varepsilon < \delta$ then $B(p, \varepsilon) \subset B(p, \delta) \subset E$ and so $p \neq q = (p_1 + \varepsilon/2, p_2) \in B(p, \varepsilon) \subset E$ so that $B(p, \varepsilon)$ contains a point of E distinct from p. If $\varepsilon > \delta$ then $B(p, \delta) \subset B(p, \varepsilon)$ and so choosing $q = (p_1 + \delta/2, p_2) \in B(p, \delta) \subset B(p, \varepsilon)$ shows that $B(p, \varepsilon)$ contains a point in E distinct from p. Thus every neighborhood of p contains a point of E distinct from p hence p is a limit point of E and so every point of E is a limit point of E. The same is not true of closed sets. Consider $A = \{(n, 0) \mid n \in \mathbb{Z}\} \subset \mathbb{R}^2$. Then A is closed because A contains no limit points at all. This is because any neighborhood of any point in A with radius less than 1 contains no other points of A. Yet, since A is not empty, it contains points which are not limit points. □

9. Let E° denote the set of all interior points of a set E. [See Definition 2.18(e) on page 32 of the text; E° is called the *interior* of E.]

(a) Prove that E° is always open.

(b) Prove that E is open if and only if $E = E^\circ$.

(c) If $G \subset E$ and G is open, prove that $G \subset E^\circ$.

(d) Prove that the complement of E° is the closure of the complement of E.

(e) Do E and $\overline{E}$ always have the same interiors?

(f) Do E and E° always have the same closures?

Solution:

(a) Let $p \in E^\circ$. Then by definition there exists $\delta > 0$ such that $N(p,\delta) \subset E$. By Theorem 2.19 on page 32 of the text we know that every neighborhood is open and thus if $q \in N(p,\delta)$ there exists $\varepsilon > 0$ such that $N(q,\varepsilon) \subset N(p,\delta) \subset E$. Therefore $q \in E^\circ$ and so $N(p,\delta) \subset E^\circ$ and hence E° is open.

(b) If $E = E^\circ$ then E is open by part (a). Now, suppose that E is open. If $p \in E$ then there exists $\delta > 0$ such that $N(p,\delta) \subset E$ and therefore $p \in E^\circ$ by definition. Thus we have the inclusion $E \subset E^\circ$. By definition, interior points of E are elements of E because any neighborhood of a point contains that point. Therefore $E^\circ \subset E$ and we obtain $E = E^\circ$.

(c) Now, let $p \in G \subset E$. If G is open, there exists $\delta > 0$ such that $N(p,\delta) \subset G \subset E$. Thus, $p \in E^\circ$ by definition and we obtain $G \subset E^\circ$, showing part (c).

(d) The result follows from

$$\begin{aligned} p \in (E^\circ)^c &\iff \forall\, \varepsilon > 0, N(p,\varepsilon) \cap E^c \neq \emptyset \\ &\iff p \in E^c \text{ or } \forall\, \varepsilon > 0,\ \exists\, q \in N(p,\varepsilon) \cap E^c \text{ such that } q \neq p \\ &\iff p \in E^c \text{ or } p \in (E^c)' \\ &\iff p \in \overline{E^c}. \end{aligned}$$

(e) Consider $\mathbb{Q}$, which has no interior points because any neighborhood of a rational must contain irrationals because irrationals are dense in $\mathbb{R}$, hence any such neighborhood cannot be contained entirely in $\mathbb{Q}$. Thus $\mathbb{Q}^\circ = \emptyset$. But, since $\overline{\mathbb{Q}} = \mathbb{R}$ (because the limit points of $\mathbb{Q}$ are precisely irrationals) we see that $(\overline{\mathbb{Q}})^\circ = \mathbb{R}^\circ = \mathbb{R}$ (because $\mathbb{R}$ is open in itself) hence $\mathbb{Q}^\circ \neq (\overline{\mathbb{Q}})^\circ$.

(f) The same example of $\mathbb{Q}$ shows that (f) is also false. $\overline{\mathbb{Q}} = \mathbb{R} \neq \emptyset = \overline{\emptyset} = \overline{\mathbb{Q}^\circ}$. $\square$

10. Let X be an infinite set. For $p, q \in X$, define

$$d(p,q) = \begin{cases} 1 & \text{if } p \neq q \\ 0 & \text{if } p = q \end{cases}.$$

Prove that this is a metric. Which subsets of the resulting metric space are open? Which are closed? Which are compact?

Solution: To show it is a metric, only the triangle inequality is not obvious from the definition. So, let $p,q,r \in X$. Then if $p = q$ we have that $d(p,q) = 0 \leq d(p,r) + d(r,q)$ because $d(p,r), d(r,q) \geq 0$. If $p \neq q$ we have that $d(p,q) = 1 \leq d(p,r) + d(r,q)$ because if $r = q$ we have that $r \neq p$ and then $d(p,r) = 1$ and $d(r,q) = 0$ so the inequality holds. If $r = p$ then we similarly get that $r \neq q$ so $d(p,r) = 0$ and $d(r,q) = 1$ so the inequality still holds. If we finally have that $r \neq p, q$ then both $d(p,r) = 1 = d(p,q)$ and the inequality holds. Thus, this is a metric. First, note that $N(p,r) = X$ if $r \geq 1$ for any $p \in X$ and $N(p,r) = \{p\}$ if $r < 1$ for any $p \in X$. Now, let $\emptyset \neq A \subset X$ and let $p \in A$. Then $N(p, 0.5) = \{p\} \subset A$ thus every point in A has a neighborhood entirely contained in A, thus A is open and so every subset of X is open. Since a set is closed if and only if its complement is open by Theorem 2.23 on page 34 of the text, we see that every subset of X is also closed. It is clear from the definition that any finite subset of X will be compact. But, if $A \subset X$ is infinite we see that it cannot be compact since by Theorem 2.37 on page 38 of the text an infinite subset of a compact set has a limit point in the compact set, yet from this metric we see that there are no limit points for any sets. This is because any neighborhood of any point of radius less then 1 contains no other points. Hence, the only compact subsets of X are finite ones. □

11. For $x \in \mathbb{R}$ and $y \in \mathbb{R}$, define

$$\begin{aligned} d_1(x,y) &= (x-y)^2 \\ d_2(x,y) &= \sqrt{|x-y|} \\ d_3(x,y) &= |x^2-y^2| \\ d_4(x,y) &= |x-2y| \\ d_5(x,y) &= \frac{|x-y|}{1+|x-y|}. \end{aligned}$$

Determine, for each of these, whether it is a metric or not.

Solution: For each of these we must determine whether $d_i(x,y) > 0$ for $x \neq y$, $d_i(x,x) = 0 \; \forall \, x \in \mathbb{R}$, $d_i(x,y) = d_i(y,x)$, and $d_i(x,y) \leq d(x,z) + d(z,y)$.

$d_1(x,y)$ is not a metric because it does not satisfy the triangle inequality since $d_1(0,2) = 4, d_1(0,1) = 1$, and $d_1(1,2) = 1$ thus $4 = d_1(0,2) > d_1(0,1) + d_1(1,2) = 1 + 1 = 2$.

$d_2(x,y)$ is a metric from the following: $d_2(x,y) > 0$ if $x \neq y$ and $d_2(x,x) = 0$ for all $x \in \mathbb{R}$. Also, $d_2(x,y) = \sqrt{|x-y|} = \sqrt{|y-x|} = d_2(y,x)$. Finally, since $|x-y| \leq |x-z| + |z-y|$ we have that $\sqrt{|x-y|} \leq \sqrt{|x-z| + |z-y|} \leq \sqrt{|x-z|} + \sqrt{|y-z|}$. The last inequality follows since $a + b \leq a + b + 2\sqrt{ab} = (\sqrt{a} + \sqrt{b})^2 \implies \sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$.

$d_3(x,y)$ is not a metric since $d(1,-1) = 0$.

$d_4(x,y)$ is not a metric since $d_4(1,1) = 1 \neq 0$.

To show that $d_5(x,y)$ is a metric we prove the more general result that whenever $d(x,y)$ is a metric then

$$d'(x,y) = \frac{d(x,y)}{1+d(x,y)}$$

is also be a metric. Since $d(x,y) = |x-y|$ is the ordinary metric on $\mathbb{R}$, it will follow that $d_5(x,y)$ is a metric. Because $d(x,y)$ is a metric we have that $d'(x,y) > 0$ if $x \neq y$, $d'(x,x) = 0$ for all $x \in \mathbb{R}$, and $d'(x,y) = d'(y,x)$. Now, let $p = d(x,y), q = d(x,z), r = d(z,y)$. Since $d(x,y)$ is a metric we have that $p, q, r \geq 0$ and

$$\begin{aligned} p \leq q + r &\implies p \leq q + r + 2qr + pqr \\ &\implies p + pq + pr + pqr \leq (q + pq + qr + pqr) + (r + pr + qr + pqr) \\ &\implies p(1+q)(1+r) \leq q(1+r)(1+p) + r(1+q)(1+p) \\ &\implies \frac{p}{1+p} \leq \frac{q}{1+q} + \frac{r}{1+r} \\ &\implies d'(x,y) \leq d'(x,z) + d'(z,y). \end{aligned}$$

Thus, $d'(x,y)$ is a metric. □

12. Let $K \subset \mathbb{R}$ consist of 0 and the numbers $1/n$, for $n = 1, 2, 3, \ldots$ Prove that K is compact directly from the definition (without using the Heine-Borel theorem).

Solution: Let $\{U_\alpha\}$ be an open cover of K. Let $0 \in U_{\alpha_0}$. Since U_{α_0} is open there exists $\delta > 0$ such that $B(0,\delta) \subset U_{\alpha_0}$ hence we can choose any n such that $1/n < \delta$ and then

$1/n \in B(0,\delta) \subset U_{\alpha_0}$. Thus, we also have that $1/m \in B(0,\delta) \subset U_{\alpha_0}$ for all $m > n$. Then, let $1/i \in U_{\alpha_i}$ for $i = n-1, n-2, \ldots, 1$. Since there are only finitely many i's we have that $\{U_{\alpha_j}\}_{j=1}^{n-1} \cup U_{\alpha_0}$ is a finite subcover. $\square$

13. Construct a compact set of real numbers whose limit points form a countable set.

    Solution: Consider the set

$$A_m = \{m + 1/n \mid n = 1, 2, \ldots\}.$$

Then A_m is compact for all $m \in \mathbb{N}$ by Exercise 12 above (it is merely a translation of the set in that exercise) and the set of limit points of A_m is $\{m\}$. Therefore, if we let

$$A = \bigcup_{i=1}^{\infty} A_i$$

A is a countable union of countable sets so it is countable by Theorem 2.12 on page 29 of the text. The set of limit points of A is then precisely $\mathbb{N}$ so is countable also. $\square$

14. Give an example of an open cover of the segment $(0, 1)$ which has no finite subcover.

    Solution: Consider the open cover $\{(1/n, 1)\}_{n=2}^{\infty}$. This covers $(0, 1)$, because if $r \in (0, 1)$, let $n \geq 2$ be such that $1/n < r$. Then $r \in (1/n, 1)$. Now, consider any finite collection in this open cover $\{(1/i, 1)\}_{i\in I}$ where $|I| < \infty$. Let $n \geq 2$ be such that $n > \max_{i\in I}\{i\}$. Then we have that $1/n \in (0, 1)$ yet $1/n < 1/i \;\; \forall\, i \in I$ and thus $1/n \notin \cup_{i\in I}(1/i, 1)$ hence $\{(1/i, 1)\}_{i\in I}$ is not an open cover of $(0, 1)$. Therefore, no finite subcover of this cover exists. $\square$

15. Show that Theorem 2.36 and its Corollary become false (in $\mathbb{R}$, for example) if the word "compact" is replaced by "closed" or by "bounded".

    Solution: Theorem 2.36, on page 38 of the text, states that If $\{K_\alpha\}$ is a collection of compact subsets of a metric space X such that the intersection of every finite subcollection of $\{K_\alpha\}$ is nonempty, then $\bigcap K_\alpha$ is nonempty. Its Corollary states that if $\{K_n\}$ is a sequence of nonempty compact sets such that $K_n \supset K_{n+1} (n = 1, 2, \ldots)$, then $\bigcap_{n=1}^{\infty} K_n$ is nonempty.

    Consider the sets $K_n = (0, 1/n)$. Then each K_n is bounded and for any finite subcollection $\{K_i\}_{i\in I}$ where $|I| < \infty$ let $n = \min_{i\in I}\{i\}$. Then $\bigcap_{i\in I} K_i = (0, 1/n) \neq \emptyset$. Yet we have that $\bigcap_{n=1}^{\infty} K_n = \emptyset$ since if $0 < r \in \mathbb{R}$ we can choose $n \in \mathbb{N}$ such that $1/n < r$. But then $r \notin (0, 1/n) \implies r \notin \bigcap_{n=1}^{\infty} K_n$.

    Now, let $K_n = \{m \in \mathbb{N} \mid m \geq n\}$. Then each K_n is closed since it has no limit points (because it is discrete), so it vacuously contains them. Then, if $\{K_i\}_{i\in I}$ is a finite subcollection where $|I| < \infty$, let $m \in \mathbb{N}$ such that $m > \max_{i\in I}\{i\}$. Then, $m \in K_i \;\; \forall\, i \in I$ and so $m \in \bigcap_{i\in I} K_i$. But, we have that $\bigcap_{n=1}^{\infty} K_n = \emptyset$ because if $m \in \mathbb{N}$, let $i \in \mathbb{N}$ be such that $i > m$. Then $m \notin K_i$ and so $m \notin \bigcap_{n=1}^{\infty} K_n$. $\square$

16. Omitted.

17. Omitted.

18. Omitted.

19. Omitted.

20. Omitted.

21. Omitted.

22. A metric space is called *separable* if it contains a countable dense subset. Show that $\mathbb{R}^k$ is separable. *Hint:* Consider the set of points which have only rational coordinate.

    Solution: Let $\mathbb{Q}^k = \{(x_1, \ldots, x_k) \in \mathbb{R}^k \mid x_i \in \mathbb{Q} \;\; \forall\, i = 1, \ldots, k\}$. If $(x_1, \ldots, x_k) \in \mathbb{R}^k$ let $\varepsilon > 0$ be given. Then, since $\mathbb{Q}$ is dense in $\mathbb{R}$, choose $p_1, \ldots, p_n \in \mathbb{Q}$ such that $|x_i - p_i| < \frac{1}{\sqrt{k}}\varepsilon$. Then, writing $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{Q}^k$ and $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^k$ we see that

    $$d(\mathbf{p}, \mathbf{x}) = \left(\sum_{i=1}^{k} |x_i - p_i|^2 \right)^{1/2} < \left(\sum_{i=1}^{k} \frac{1}{k}\varepsilon^2 \right)^{1/2} = \varepsilon.$$

    Therefore, $\mathbb{Q}^k$ is dense in $\mathbb{R}^k$ and $\mathbb{Q}^k$ is countable because it can be realized as the disjoint union of $\mathbb{Q}$, k times. □

23. A collection $\{V_\alpha\}$ of open subsets of X is said to be a *base* if the following is true: For every $x \in X$ and every open set $G \subset X$ such that $x \in G$, we have $x \in V_\alpha \subset G$ for some α. In other words, every open set in X is the union of a subcollection of $\{V_\alpha\}$. Prove that every separable metric space has a *countable* base. *Hint:* Take all neighborhoods with rational radius and center in some countable dense subset of X.

    Solution: Since X is separable, by definition it has a countable dense subset $S \subset X$. Let $S = \{p_1, p_2, \ldots\}$ and enumerate $\mathbb{Q} = \{q_1, q_2, \ldots\}$. Let $V_{ij} = N(p_i, q_j)$. Then, $\{V_{ij}\}_{i,j=1}^{\infty}$ is a countable collection of open sets in X by Theorem 2.12 on page 29 of the text. Let $x \in X$ and let $G \subset X$ be open such that $x \in G$. Since G is open, there exists $\delta > 0$ such that $N(x, \delta) \subset G$. Let $p_i \in S$ such that $d(p_i, x) = \varepsilon < \delta/2$ and then choose $q_j \in \mathbb{Q}$ such that $\varepsilon < q_j < \delta/2$, both of which are possible because S is dense in X and $\mathbb{Q}$ is dense in $\mathbb{R}$. Then, consider $V_{ij} = N(p_i, q_j)$. Since $d(x, p_i) = \varepsilon < q_j$ we see that $x \in V_{ij}$. Now, let $a \in V_{ij}$. Then, $d(a, x) \le d(a, p_i) + d(p_i, x) < q_j + \delta/2 < \delta$. Therefore $a \in N(x, \delta) \subset G$ and so we have that $x \in V_{ij} \subset N(x, \delta) \subset G$ showing that $\{V_{ij}\}_{i,j=1}^{\infty}$ is a countable base. □

Chapter 3: Numerical Sequences and Series

1. Prove that convergence of $\{s_n\}$ implies convergence of $\{|s_n|\}$. Is the converse true?

   Solution: Suppose that $s_n \to s$. Let $\varepsilon > 0$ be given and let N be such that

   $$|s_n - s| < \varepsilon \quad \forall\, n \ge N.$$

   Then, by Exercise #13 in Chapter 1 above, we have that

   $$\big||s_n| - |s|\big| \le |s_n - s| < \varepsilon \quad \forall\, n \ge N$$

   Thus $|s_n| \to |s|$. The converse is false because if we let $s_n = (-1)^n$ then $|s_n|$ converges, yet s_n does not. □

2. Omitted.

3. If $s_1 = \sqrt{2}$ and

   $$s_{n+1} = \sqrt{2 + \sqrt{s_n}}$$

   prove that $\{s_n\}$ converges, and that $s_n < 2$ for all $n \ge 1$.

Solution: $s_1 = \sqrt{2} < 2$. Suppose inductively that $s_n < 2$. Then,

$$s_{n+1} = \sqrt{2 + \sqrt{s_n}} < \sqrt{2 + \sqrt{2}} < 2$$

because $2 + \sqrt{2} < 2 + 2 = 4 \implies \sqrt{2 + \sqrt{2}} < \sqrt{4} = 2$. Thus we see that $s_n < 2$ for all $n \geq 1$. We also have that

$$s_2 = \sqrt{2 + \sqrt{\sqrt{2}}} > \sqrt{2} = s_1.$$

Suppose inductively that $s_n > s_{n-1}$. Then

$$s_{n+1} = \sqrt{2 + \sqrt{s_n}} > \sqrt{2 + \sqrt{s_{n-1}}} = s_n$$

and therefore we have that $s_n > s_{n-1}$ for all $n \geq 1$ by induction. Thus the sequence is increasing and since $\cdots > s_n > s_{n-1} > \cdots > s_1 > 0$ we have an increasing sequence of nonnegative terms which is bounded above. Hence by Theorem 3.24 on page 60 of the text $\{s_n\}$ converges. $\square$

4. Find the upper and lower limits of the sequence $\{s_n\}$ defined by

$$s_1 = 0, \quad s_{2m} = \frac{s_{2m-1}}{2}, \quad s_{2m+1} = \frac{1}{2} + s_{2m}.$$

Solution: If we write out the first few terms we obtain

$$\begin{aligned} s_1 &= 0 \\ s_2 &= 0 \\ s_3 &= \frac{1}{2} \\ s_4 &= \frac{1}{4} \\ s_5 &= \frac{1}{2} + \frac{1}{4} \\ s_6 &= \frac{1}{4} + \frac{1}{8} \\ s_7 &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \\ s_8 &= \frac{1}{4} + \frac{1}{8} + \frac{1}{16} \end{aligned}$$

thus we are lead to write

$$\begin{aligned} s_{2m+1} &= \sum_{k=1}^{m} \left(\frac{1}{2}\right)^k \qquad m \geq 1 \\ s_{2m} &= \sum_{k=2}^{m} \left(\frac{1}{2}\right)^k \qquad m \geq 2. \end{aligned} \tag{4.1}$$

To prove this we proceed by induction on m. For $m = 1$ we have that $s_3 = 1/2 = \sum_{k=1}^{1}(1/2)^k$. For $m = 2$ we have that $s_5 = 1/2 + 1/4 = \sum_{k=1}^{2}(1/2)^k$ and $s_4 = 1/4 = \sum_{k=2}^{2}(1/2)^k$. Now, suppose inductively that Equation (4.1) holds for $m \geq 2$. Then,

$$s_{2(m+1)+1} = \frac{1}{2} + s_{2(m+1)} = \frac{1}{2} + \frac{s_{2m+1}}{2} = \frac{1}{2} + \sum_{k=1}^{m} \left(\frac{1}{2}\right)^{k+1} = \sum_{k=1}^{m+1} \left(\frac{1}{2}\right)^{k}$$
$$s_{2(m+1)} = \frac{s_{2m+1}}{2} = \frac{1}{2}\sum_{k=1}^{m} \left(\frac{1}{2}\right)^{k} = \sum_{k=1}^{m} \left(\frac{1}{2}\right)^{k+1} = \sum_{k=2}^{m+1} \left(\frac{1}{2}\right)^{k}$$

and so Equation (4.1) holds. We thus see that there are only two subsequential limits

$$\lim_{m\to\infty} s_{2m} = \sum_{k=2}^{\infty} \left(\frac{1}{2}\right)^{k} = \frac{1}{2} \tag{4.2}$$
$$\lim_{m\to\infty} s_{2m+1} = \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^{k} = 1 \tag{4.3}$$

where the two limiting values are given by Theorem 3.26 on page 61 of the text (geometric series). These are the only possible limits because if we have a subsequence $\{s_{n_k}\}$ then either $\{n_k\}$ has infinitely many evens, infinitely many odds, or infinitely many of both. If there are infinitely many of both even and odd numbers, then $\{s_{n_k}\}$ wouldn't be Cauchy sequence because for all $N \geq 1$ we could find $n_{k_1}, n_{k_2} \geq N$ such that n_{k_1} was even and n_{k_2} was odd and then we would have that $|s_{n_{k_1}} - s_{n_{k_2}}| \geq 1/2$. If there are only finitely many odds in $\{n_k\}$ then we would have that $\lim_{k\to\infty} s_{n_k} = \lim_{m\to\infty} s_{2m}$. Specifically, letting $N > \max_k\{n_k \text{ odd }\}$ we see that $|s_{n_k} - s_{n_j}| = \sum_{k=n_j}^{n_k}(1/2)^k$ for all $n_k \geq n_j \geq N$ precisely because both n_j and n_k are even. Since $\sum_{k=2}^{\infty}(1/2)^k$ converges, it is a Cauchy sequence by Theorem 3.11 on page 53 of the text and so $\sum_{k=n_j}^{n_k}(1/2)^k$ can be made arbitrarily small for large enough n_j and n_k. If there are only finitely many evens in $\{n_k\}$ then we would have that $\lim_{k\to\infty} s_{n_k} = \lim_{m\to\infty} s_{2m+1}$. Specifically, letting $N > \max_k\{n_k \text{ even }\}$ we see that $|s_{n_k} - s_{n_j}| = \sum_{k=n_j}^{n_k}(1/2)^k$ for all $n_k \geq n_j \geq N$ precisely because both n_j and n_k are odd. Since $\sum_{k=1}^{\infty}(1/2)^k$ converges, it is a Cauchy sequence by Theorem 3.11 on page 53 of the text and so $\sum_{k=n_j}^{n_k}(1/2)^k$ can be made arbitrarily small for large enough n_j and n_k. Since Equations (4.2) and (4.3) are the only two subsequential limits we find that the upper and lower limits are

$$\liminf_{m\to\infty} s_m = \frac{1}{2}$$
$$\limsup_{m\to\infty} s_m = 1.$$

□

5. For any two real sequences $\{a_n\}, \{b_n\}$, prove that

$$\limsup_{n\to\infty}(a_n + b_n) \leq \limsup_{n\to\infty}(a_n) + \limsup_{n\to\infty}(b_n),$$

provided the sum on the right is not of the form $\infty - \infty$.

Solution: If either $\limsup$ on the right side above is infinite, we are done, so we can assume that both are finite. Assume also that the left side is finite (hence our discussion here is incomplete). We make some general remarks about this statement, but we will prove it by actually using an equivalent definition of $\limsup$ which is nonetheless easier to work with than Rudin's. First, let $a = \limsup_{n\to\infty}(a_n), b = \limsup_{n\to\infty}(b_n)$, and $c = \limsup_{n\to\infty}(a_n + b_n)$. Then, let $a_{n_k} + b_{n_k} \to \gamma$ be a convergent subsequence of the sequence $\{a_n + b_n\}$. If the component subsequences $\{a_{n_k}\}$ and $\{b_{n_k}\}$ both converge, say to α and β, respectively, then by Theorem 3.3 on page 49 of the text, we have that

$$\gamma = \lim_{k\to\infty}(a_{n_k} + b_{n_k}) = \lim_{k\to\infty} a_{n_k} + \lim_{k\to\infty} b_{n_k} = \alpha + \beta \leq a + b.$$

where the last inequality follows by the definition of a and b as $\limsup$'s. Thus, we need only consider subsequences $\{a_{n_k}+b_{n_k}\}$ where one, or both, of the the component subsequences *do not* converge. But, we can actually do better than this because we can show that if one of the component subsequences converges, then both must. To see this, without loss of generality, let $a_{n_k} + b_{n_k} \to \gamma$ and suppose that $b_{n_k} \to \beta$. Then, by the triangle inequality,

$$\begin{aligned}|a_{n_k} - (\gamma - \beta)| &\leq |a_{n_k} - (\gamma - b_{n_k})| + |(\gamma - b_{n_k}) - (\gamma - \beta)| \\ &= |(a_{n_k} + b_{n_k}) - \gamma| + |(\gamma - b_{n_k}) - (\gamma - \beta)|.\end{aligned}$$

Since $b_{n_k} \to \beta$ we see that $\gamma - b_{n_k} \to \gamma - \beta$ because γ is a constant and so both terms on the right above can be made arbitrarily small for large enough k. This shows that $a_{n_k} \to \gamma - \beta$ and therefore if one of the component subsequences of the convergent subsequence $\{a_{n_k} + b_{n_k}$ converges, both must. Hence, we are reduced to considering only those convergent subsequences $\{a_{n_k}+b_{n_k}\}$ in which *both* component subsequences *do not* converge. That is, sequences such as $a_{n_k} = (-1)^k$ and $b_{n_k} = (-1)^{k+1}$ so that neither converges individually, but their sum does. Now, we would be done if we could show that $\gamma = \lim_{k\to\infty}(a_{n_k}+b_{n_k}) = \limsup_{k\to\infty}(a_{n_k}+b_{n_k}) \leq \limsup_{k\to\infty}(a_{n_k})+\limsup_{k\to\infty}(b_{n_k}) \leq a+b$ which is the statement of the problem, but restricted to the subsequence $\{a_{n_k} + b_{n_k}\}$. Thus, we could say repeat the above argument and for every convergent sub-subsequence $a_{n_{k_j}} + b_{n_{k_j}}$ such that neither component sub-subsequence converges, repeat the argument again. Eventually, it will end because eventually some sub-$\cdots$-subsequence will have only convergent component sequences, since we are assuming that a, b, and c are finite. But, this is hard to make rigorous and so we will now prove the result by resorting to a different, but equivalent, definition of the $\limsup$.

Consider a sequence $\{a_n\}$ and let $A_k = \sup\{a_m \mid m \geq k\}$ and define $A = \lim_{k\to\infty} A_k$ (allowing it to be $\pm\infty$). Now, choose n_1 such that $|a_{n_1} - A_1| < 1$, which is possible by the definition of A_1 as the $\sup$, so $n_1 \geq 1$. Now, choose a_{n_k} inductively such that $|a_{n_k} - A_k| < 1/k$ and such that $a_{n_k} > a_{n_{k-1}}$. That is, by definition of A_k, we have that there exists K such that $|a_m - A_k| < 1/k$ for all $m \geq K$. So, simply require that $n_k > \max\{K, n_{k-1}\}$. Then, this defines a subsequence of $\{a_n\}$ which by construction converges to A, thus we have that $\limsup_{n\to\infty}(a_n) \geq A$. Now, by Theorem 3.17(a) on page 56 of the text we see that $\limsup_{n\to\infty}(a_n)$ is actually a subsequential limit of the sequence $\{a_n\}$ hence there exists some subsequence $a_{n_k} \to \limsup_{n\to\infty}(a_n)$. Since by construction we have that $a_{n_k} \leq A_{n_k}$ taking the limit as $k \to \infty$ of both sides gives $\limsup_{n\to\infty}(a_n) \leq A$ thus $A = \limsup_{n\to\infty}(a_n)$ and our A is an equivalent formulation of the $\limsup$ to Rudin's. The proof of the above inequality is now nearly trivial since if $a_{n_k} + b_{n+k}$ is any subsequence of $\{a_n + b_n\}$ whose limit exists then we of course have that $a_{n_k} + b_{n_k} \leq \sup_{m\geq k}\{a_{n_m}\} + \sup_{m\geq k}\{b_{n_m}\} = A_{n_k} + B_{n_k}$. Taking the limit of both sides gives $\lim_{k\to\infty}(a_{n_k}+b_{n_k}) \leq A+B$, where B is defined analogously to A above. Since $\{a_{n_k}+b_{n_k}\}$ was any subsequence with a limit we see that $\limsup_{n\to\infty}(a_n+b_n) \leq A + B = \limsup_{n\to\infty}(a_n) + \limsup_{n\to\infty}(b_n)$. $\square$

6. Omitted.

7. Prove that the convergence of $\sum_{n=1}^{\infty} a_n$ implies the convergence of

$$\sum_{n=1}^{\infty} \frac{\sqrt{a_n}}{n},$$

if $a_n \geq 0$.

Solution: By the Cauchy-Schwarz inequality (Theorem 1.35 on page 15 of the text), we have that

$$\sum_{j=1}^{n} a_j \overline{b_j} \leq \left(\sum_{j=1}^{n} |a_j|^2\right)\left(\sum_{j=1}^{n} |b_j|^2\right).$$

Thus, for each $n = 1, 2, \ldots$ we have

$$\sum_{j=1}^{n} \sqrt{a_n} \cdot \frac{1}{n} \leq \left(\sum_{j=1}^{n} a_n\right)\left(\sum_{j=1}^{n} \frac{1}{n^2}\right) \leq \left(\sum_{j=1}^{\infty} a_n\right)\left(\sum_{j=1}^{\infty} \frac{1}{n^2}\right) < \infty$$

where the second inequality follows because all terms are positive and the last inequality follows because both sequences on the right are convergent. Thus the partial sums, $s_n = \sum_{i=1}^{n} \sqrt{a_n} \cdot \frac{1}{n}$, form a bounded sequence. Since all the terms are nonnegative, by Theorem 3.24 on page 60 of the text the series $\sum_{i=1}^{\infty} \sqrt{a_n} \cdot \frac{1}{n}$ converges. □

8. If $\sum_{n=1}^{\infty} a_n$ converges, and if $\{b_n\}_{n=1}^{\infty}$ is monotonic and bounded, prove that $\sum_{n=1}^{\infty} a_n b_n$ converges.

Solution: Since $\sum_{n=1}^{\infty} a_n$ converges, its partial sums form a bounded sequence by definition. By Theorem 3.14 on page 55 of the text, we see that $\{b_n\}_{n=1}^{\infty}$ converges because it is monotonic and bounded, so let $b_n \to b$. Since $\{b_n\}_{n=1}^{\infty}$ is monotonic, it is either increasing or decreasing. If it is increasing, define $c_n = b - b_n$ and in this case the sequence $\{c_n\}_{n=1}^{\infty}$ is decreasing and we have that $c_1 \geq c_2 \geq \cdots$ and also $c_n \to 0$ because $b_n \to b$. On the other hand, if $\{b_n\}_{n=1}^{\infty}$ is decreasing then let $c_n = b_n - b$ and in this case $\{c_n\}_{n=1}^{\infty}$ is still a decreasing sequence so $c_1 \geq c_2 \geq \cdots$ and we still have that $c_n \to 0$. Thus, in either case, by Theorem 3.42 on page 70 of the text, we have that the sequence $\sum_{n=1}^{\infty} a_n c_n$ converges. First, suppose that $\{b_n\}_{n=1}^{\infty}$ was increasing and so $c_n = b - b_n$. Then, the partial sums

$$C_m = \sum_{n=1}^{m} a_n c_n = \sum_{n=1}^{m} (a_n b - a_n b_n) = b \sum_{n=1}^{m} a_n - \sum_{n=1}^{m} a_n b_n$$

are a convergent sequence, say $C_m \to c$. But then

$$\sum_{n=1}^{m} a_n b_n \to b \sum_{n=1}^{\infty} a_n - c,$$

hence $\sum_{n=1}^{\infty} a_n b_n$ converges. In the other case we would obtain

$$\sum_{n=1}^{m} a_n b_n \to c' + b \sum_{n=1}^{\infty} a_n$$

where here $c' = \sum_{n=1}^{\infty} a_n c_n$ and $c_n = b_n - b$ and so $\sum_{n=1}^{\infty} a_n b_n$ still converges. □

9. Omitted.

10. Omitted.

11. Suppose that $a_n > 0$ and $s_n = a_1 + \cdots + a_n$, and $\sum_{n=1}^{\infty} a_n$ diverges.

    (a) Prove that $\sum_{n=}^{\infty} \frac{a_n}{1+a_n}$ diverges.

    (b) Prove that
    $$\frac{a_{N+1}}{s_{N+1}} + \cdots + \frac{a_{N+k}}{s_{N+k}} \geq 1 - \frac{s_N}{s_{N+k}}$$
    and deduce that $\sum_{n=1}^{\infty} \frac{a_n}{s_n}$ diverges.

    (c) Prove that
    $$\frac{a_n}{s_n^2} \geq \frac{1}{s_{n-1}} - \frac{1}{s_n}$$
    and deduce that $\sum_{n=1}^{\infty} \frac{a_n}{s_n^2}$ converges.

    (d) What can be said about
    $$\sum_{n=1}^{\infty} \frac{a_n}{1+na_n} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{a_n}{1+n^2 a_n}?$$

Solution:

(a) Suppose that $\sum_{n=1}^{\infty} \frac{a_n}{1+a_n}$ converges. Observe that
$$\frac{a_n}{1+a_n} \to 0 \iff \frac{1}{\frac{1}{a_n}+1} \to 0 \iff \frac{1}{a_n} \to \infty \iff a_n \to 0.$$
Since we are supposing that $\sum_{n=1}^{\infty} \frac{a_n}{1+a_n}$ converges we have by Theorem 3.23 on page 60 0f the text that $\frac{a_n}{1+a_n} \to 0 \implies a_n \to 0$ by the above. Hence, there exists N_1 such that $a_n < 1$ for all $n \geq N_1$. Now, let $\varepsilon > 0$ be given. Then there exists N_2 such that
$$\frac{a_m}{1+a_m} + \cdots + \frac{a_n}{1+a_n} < \frac{\varepsilon}{2} \quad \forall\, n, m \geq N_2.$$
Letting $N = \max\{N_1, N_2\}$ we have that
$$\frac{\varepsilon}{2} > \frac{a_m}{1+a_m} + \cdots + \frac{a_n}{1+a_n} > \frac{a_m}{1+1} + \cdots + \frac{a_n}{1+1} = \frac{a_m}{2} + \cdots + \frac{a_n}{2}$$
for all $n, m \geq N$ thus
$$\varepsilon > a_m + \cdots + a_n \quad \forall\, n, m \geq N \implies \sum_{n=1}^{\infty} a_n \text{ converges},$$
a contradiction. Thus $\sum_{n=1}^{\infty} \frac{a_n}{1+a_n}$ diverges.

(b) Because $a_n > 0$ for all n we have that $s_{n+1} > s_n$ for all n. Thus,
$$\frac{a_{N+1}}{s_{N+1}} + \cdots + \frac{a_{N+k}}{s_{N+k}} > \frac{a_{N+1} + \cdots + a_{N+k}}{s_{N+k}} = \frac{s_{N+k} - s_N}{s_{N+k}} = 1 - \frac{s_N}{s_{N+k}}.$$

Suppose now that $\sum_{n=1}^{\infty} \frac{a_n}{s_n}$ converges and let $\varepsilon > 0$ be given. Then, there exists an N such that for all $n, m \geq N$ we have that

$$\varepsilon > \frac{a_m}{s_m} + \cdots + \frac{a_n}{s_n}.$$

In particular, for $m = N + 1$ and $n = N + k$ we obtain that

$$\varepsilon > \frac{a_{N+1}}{s_{N+1}} + \cdots + \frac{a_{N+k}}{s_{N+k}} > 1 - \frac{s_N}{s_{N+k}}.$$

Now, since $\sum_{n=1}^{\infty} a_n$ diverges and $a_n > 0$ so the s_n are increasing, we see that $s_{N+k} \to \infty$ as $k \to \infty$. Thus we can choose k large enough such that $s_{N+k} > 2s_N \implies \frac{s_N}{s_{N+k}} < \frac{1}{2}$ because N is fixed. But then, we would obtain

$$\varepsilon > 1 - \frac{s_N}{s_{N+k}} > 1 - \frac{1}{2} = \frac{1}{2},$$

contradicting that ε can be chosen arbitrary, i.e. just choose $\varepsilon = \frac{1}{2}$ at the beginning. Thus, $\sum_{n=1}^{\infty} \frac{a_n}{s_n}$ diverges.

(c) As in part (b), since $a_n > 0$ for all n we have that $s_n > s_{n-1}$ for all n. Thus

$$\frac{1}{s_{n-1}} - \frac{1}{s_n} = \frac{s_n - s_{n-1}}{s_n s_{n-1}} > \frac{s_n - s_{n-1}}{s_n^2} = \frac{a_n}{s_n^2} \quad \forall\, n \geq 2.$$

Then since $s_n > 0$ for all n,

$$\sum_{n=2}^{k} \frac{a_n}{s_n^2} < \sum_{n=2}^{k} \left(\frac{1}{s_{n-1}} - \frac{1}{s_n} \right) = \frac{1}{s_1} - \frac{1}{s_k} < \frac{1}{s_1} = \frac{1}{a_1}, \tag{11.1}$$

thus

$$\sum_{n=1}^{k} \frac{a_n}{s_n^2} < \frac{2}{a_1}$$

so the partial sums of $\sum_{n=1}^{\infty} \frac{a_n}{s_n^2}$ form a bounded sequence and since $a_n > 0$ for all n, we have that $\sum_{n=1}^{\infty} \frac{a_n}{s_n^2}$ converges by Theorem 3.24 on page 60 of the text. We could also take the limit of both sides of Equation (11.1) as $k \to \infty$ and note that $\frac{1}{s_k} \to 0$ as $k \to \infty$ because $\sum_{n=1}^{\infty} a_n$ diverges and $a_n > 0$ for all n so $s_n \to \infty$.

(d) $\sum_{n=1}^{\infty} \frac{a_n}{1+na_n}$ may either diverge or converge. If we let $a_n = \frac{1}{n}$ then

$$\sum_{n=1}^{\infty} \frac{a_n}{1 + na_n} = \sum_{n=1}^{\infty} \frac{1/n}{2} = \frac{1}{2} \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

On the other hand, if we let $a_n = \frac{1}{n \log(n)^p}$ where $p > 1$ and $n \geq 2$ then

$$\begin{aligned}
\frac{a_n}{1 + na_n} = \frac{1}{n \log(n)^p} \cdot \frac{1}{1 + n\left(\frac{1}{n \log(n)^p}\right)} &= \frac{1}{n \log(n)^p (1 + \log(n)^{-p})} \\
&= \frac{1}{n \log(n)^p + n} \\
&< \frac{1}{n \log(n)^p}
\end{aligned}$$

thus $\sum_{n=1}^{\infty} \frac{a_n}{1+na_n}$ converges by Theorem 3.29 on page 62 of the text for this choice of a_n. On the other hand,

$$\sum_{n=1}^{\infty} \frac{a_n}{1+n^2 a_n} = \sum_{n=1}^{\infty} \frac{1}{1/a_n + n^2} \leq \sum_{n=1}^{\infty} \frac{1}{n^2}$$

where the inequality follows because $a_n > 0$ for all n. Thus, $\sum_{n=1}^{\infty} \frac{a_n}{1+n^2 a_n}$ converges. □

Chapter 4: Continuity

1. Suppose f is a real function defined on $\mathbb{R}$ which satisfies

$$\lim_{h\to 0}[f(x+h) - f(x-h)] = 0$$

for every $x \in \mathbb{R}$. Does this imply f is continuous?

Solution: No, this does not imply f is continuous because this statement merely says that $\lim_{t\to x^+} f(t) = \lim_{t\to x^-} f(t)$, that is, the right and left handed limits of f are equal, but it says nothing about whether these actually equal $f(x)$ itself. That is, by Theorem 4.6 on page 86, f is continuous at $x \iff \lim_{t\to x} f(t) = f(x)$ and by the comment in Definition 4.25 on page 94 we see that $\lim_{t\to x} f(t)$ exists $\iff \lim_{t\to x^+} f(t) = \lim_{t\to x^-} f(t) = \lim_{t\to x} f(t)$. So, if we define

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

then at any $x \in \mathbb{Z}$ we have that $\lim_{h\to 0} f(x+h) = \lim_{t\to x^+} f(t) = 0 = \lim_{t\to x^-} f(t) = \lim_{h\to 0} f(x-h)$ but of course $\lim_{t\to x^+} f(t) = \lim_{t\to x^-} f(t) = 0 \neq 1 = f(x)$. That is, $\lim_{t\to x} f(t) = 0 \neq 1 = f(x)$ so f is not continuous at any $x \in \mathbb{Z}$. □

2. If f is a continuous mapping of a metric space X into a metric space Y, prove that $f(\overline{E}) \subset \overline{f(E)}$. Show by an example that $f(\overline{E})$ can be a proper subset of $\overline{f(E)}$.

Solution: If E is empty then the inclusion holds trivially, so suppose that it is not empty. Let $x \in E$ so $f(x) \in f(E) \subset \overline{f(E)}$. Thus, $x \in f^{-1}(\overline{f(E)})$ and since x was arbitrary we see that $E \subset f^{-1}(\overline{f(E)})$. But, $\overline{f(E)}$ is closed in Y (Theorem 2.27(a) on page 35 of the text) and since f is continuous we have that $f^{-1}(\overline{f(E)})$ is then closed in X by the Corollary to Theorem 4.8 on page 87 of the text. But, this implies that $\overline{E} \subset f^{-1}(\overline{f(E)})$ because $\overline{E}$ is the smallest closed set containing E so it is a subset of any closed set containing E, i.e. this follows from Theorem 2.27(c) on page 35 of the text. But, this is precisely the statement that $f(\overline{E}) \subset \overline{f(E)})$.

To see that this inclusion can be proper, let $E = \mathbb{Z}$ and define $f : \mathbb{Z} \to \mathbb{R}$ by $f(n) = 1/n$. Then, f is continuous (in fact, it is uniformly continuous if we take $\delta < 1$ because then $d(n,m) < 1 \implies n = m$ when $n, m \in \mathbb{Z}$ and thus $d(f(n), f(m)) = 0 < \varepsilon \; \forall \, \varepsilon > 0$). See the comment at the end of the proof of Theorem 4.20, on page 92 of the text. But, since $\mathbb{Z}$ is closed, we have that $f(\overline{\mathbb{Z}}) = f(\mathbb{Z}) = \{1/n \mid n \in \mathbb{Z}\} \subsetneq \overline{f(\mathbb{Z})} = \overline{\{1/n \mid n \in \mathbb{Z}\}} = \{1/n \mid n \in \mathbb{Z}\} \cup \{0\}$, hence the inclusion is proper. □

Alternate Solution: If $f(\overline{E})$ is empty then the conclusion holds trivially, so suppose that it is not and let $y \in f(\overline{E})$. Thus, there exists a point $p \in \overline{E}$ such that $f(p) = y$. If $p \in E$ then we have that $y = f(p) \in f(E) \subset \overline{f(E)}$. So, suppose that $p \in E'$, so p is a limit point of E. Since f is continuous at p, for all $\varepsilon > 0$ there exists a $\delta > 0$ such that $d(x,p) < \delta \implies d(f(x), f(p)) < \varepsilon$. But, since p is a limit point of E we see that there always exists $x \in N(p, \delta)$ for some $x \in E$ and for any $\delta > 0$. Thus, for such an x, we have that $d(f(x), f(p)) < \varepsilon$ which implies that

$f(p)$ is a limit point of $f(E)$ since we can always find points in $f(E)$ arbitrarily close to it. Thus $y = f(p) \in \overline{f(E)}$ and so $f(\overline{E}) \subset \overline{f(E)}$. $\square$

3. Let f be a continuous real function on a metric space X. Let $Z(f)$ (the *zero set* of f) be the set of all $p \in X$ at which $f(p) = 0$. Prove that $Z(f)$ is closed.

   Solution: Let $p_n \to p$ in $Z(f)$ so that $p_n \in Z(f)$ for all $n = 1, 2, \ldots$ Of course $p \in X$ and since f is continuous on X we see that $\lim_{q \to p} f(q) = f(p)$ by Theorem 4.6 on page 86 of the text. Then, by Theorem 4.2 on page 84 of the text, we see that $0 = \lim_{n \to \infty} f(p_n) = f(p)$ (where $p_n \neq p \; \forall \, n$ because otherwise we are done) hence $p \in Z(f)$ and so $Z(f)$ is closed.

   Another approach is to show that $Z(f)^c = X \backslash Z(f)$ is open. To this end, let $x \in E = Z(f)^c$ so that $f(x) \neq 0$. Without loss of generality, suppose that $f(x) > 0$. Since f is continuous on X we see that for $\varepsilon = f(x)/2 > 0$ there exists a $\delta > 0$ such that $d(f(x), f(y)) = |f(x) - f(y)| < \varepsilon = f(x)/2$ whenever $d(x, y) < \delta$. But, this implies that $f(y) \neq 0$ because otherwise $|f(x) - f(y)| = f(x) > \varepsilon$, a contradiction. This means that $N(x, \delta) \subset E$ so that in particular, E is open. Thus, $Z(f) = E^c$ is closed. If we had that $f(x) < 0$ we would simply let $\varepsilon = -f(x)/2 > 0$. $\square$

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

   Solution: First, let $y \in f(X)$, let $y = f(x)$ and let $\varepsilon > 0$. Then, because f is continuous, there exists a $\delta > 0$ such that $d(f(x), f(p)) < \varepsilon$ whenever $d(x, p) < \delta$. But, since E is dense in X there exists a $p \in E$ such that $d(x, p) < \delta$, which gives that $d(y, f(p)) = d(f(x), f(p)) < \varepsilon$ hence $f(E)$ is dense in $f(X)$ because we can find an element of $f(E)$ arbitrarily close to any element of $f(X)$.

   Now, let $\varepsilon > 0$ be given and suppose that $g(p) = f(p)$ for all $p \in E$. Let $q \in X$. Because f is continuous at q, there exists $\delta_1 > 0$ such that $d(f(p), f(q)) < \varepsilon/2$ whenever $d(p, q) < \delta_1$. Since g is also continuous at q let $\delta_2 > 0$ be such that $d(g(p), g(q)) < \varepsilon/2$ whenever $d(p, q) < \delta_2$ and let $\delta = \min\{\delta_1, \delta_2\}$. Since E is dense in X we can find $p \in E$ such that $d(p, q) < \delta$. Then,

$$
\begin{aligned}
d(g(q), f(q)) &\leq d(g(q), g(p)) + d(g(p), f(p)) + d(f(p), f(q)) \\
&= d(g(q), g(p)) + d(f(p), f(q)) \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon,
\end{aligned}
$$

   where $p \in E$ so $g(p) = f(p)$. Thus, since ε was arbitrary we must have that $g(q) = f(q)$ hence they agree on all of X.

   Another way to see this is that if $q \in X \backslash E$ then we can find a sequence $p_n \to q$ such that $p_n \in E$ for all n and $p_n \neq q$ for all n because E is dense in X, i.e. every element of X is a point or a limit point of E, so q must be a limit point of E. But then because g and f are both continuous we can interchange limits (by Theorem 4.6 on page 86 and Theorem 4.2 on page 84 of the text) so we obtain,

$$
g(q) = g(\lim_{n \to \infty} p_n) = \lim_{n \to \infty} g(p_n) = \lim_{n \to \infty} f(p_n) = f(\lim_{n \to \infty} p_n) = f(q)
$$

   because $g(p_n) = f(p_n) \; \forall \, n$ since they agree on E. $\square$

5. If f is a real continuous function defined on a closed set $E \subset \mathbb{R}$, prove that there exist continuous real functions g on $\mathbb{R}$ such that $g(x) = f(x)$ for all $x \in E$. (Such functions g are called *continuous extensions* of f from E to $\mathbb{R}$.) Show that the result becomes false if the world "closed" is omitted. Extend the result to vector-valued functions. *Hint:* Let the graph of g be a straight line on each of the segments which constitute the complement of E (compare Exercise 29, Chapter 2). The result remains true if $\mathbb{R}$ is replaced by any metric space, but the proof is not so simple.

   Solution: By Exercise 20 in Chapter we see that every open set of $\mathbb{R}$ is an *at most* countable union of disjoint open sets. Hence, since E is closed, we can write

$$E^c = \bigcup_{i=1}^{n} (a_i, b_i)$$

   where $a_i < b_i < a_{i+1} < b_{i+1}$ and we allow n to be ∞. Then, on each (a_i, b_i) define

$$g(x) = f(a_i) + (x - a_i)\frac{f(b_i) - f(a_i)}{b_i - a_i} \tag{5.1}$$

   for $x \in (a_i, b_i)$. That is, the graph of g is nothing more than the straight line connecting the points $f(a_i)$ and $f(b_i)$, where $a_i, b_i \in E$ because $a_i, b_i \notin E^c$. We let $g(x) = f(x)$ for $x \in E$. Then it is clear that g is continuous on the interior of E and since it is a linear function on E^c it is continuous there also. Hence, we need only check that it is continuous at the boundary of E, namely the points $\{a_i, b_i\}_{i=1}^n$. But, from Equation (5.1) we see that $\lim_{t \to a_i} g(t) = f(a_i) = g(a_i)$ because $a_i \in E$, hence g is continuous at a_i by Theorem 4.6 on page 86 of the text. Similarly $\lim_{t \to b_i} g(t) = f(b_i) = g(b_i)$ because $b_i \in E$ hence g is a continuous extension of f.

   The fact that E is closed is essential because if we consider $f(x) = 1/x$ on $E = (0, 1)$, for example, then f is continuous on E yet there is no continuous extension g of f to $\mathbb{R}$ such that $g = f$ on E yet $g(0) \in \mathbb{R}$ since as $x \to 0$ we have that $f(x) \to \infty$ hence we would also have that $g(x) \to \infty$ as $x \to 0$.

   The vector-valued case is simply a natural extension of the single variable case since if f is a continuos vector valued function on $\mathbb{R}$ we can write $f = (f_1, \ldots, f_n)$ and so f_i is continuous for each $i = 1, \ldots, n$ because f is continuous. Hence, by the above, we can extend each f_i to a continuous function g_i on $\mathbb{R}$ and letting $g = (g_1, \ldots, g_n)$ we see that g is then a continuous extension of f to $\mathbb{R}^n$. □

6. Omitted.

7. Omitted.

8. Let f be a real uniformly continuous function on the bounded set E in $\mathbb{R}$. Prove that f is bounded on E. Show that the conclusion is false if boundedness of E is omitted from the hypothesis.

   Solution: Since f is uniformly continuous, there exists a single $\delta > 0$ such that whenever $|x - y| < \delta$ we have that $|f(x) - f(y)| < 1$ for all $x, y \in E$. Since E is bounded there exists $M > 0$ such that $-M < x < M$ for all $x \in E$. Now, note that if $N \geq 2M/\delta$ is an integer, then N neighborhoods of size δ put side by side cover at least all but a finite number of points in $(-M, M)$ (missing the points that lie between the boundaries of two adjacent neighborhoods). Thus, we can cover $(-M, M)$ with finitely many neighborhoods of radius δ, and so we must be able to do the same with E but centering the neighborhoods at elements of E. Hence, some finite subcollection of $\{B(x,\delta) \mid x \in E\}$ must cover E, say $E \subseteq \{B(x_i,\delta)\}_{i=1}^n$. Now, let $x \in E$. Then, for some $1 \le i \le n$ we have that $x \in B(x_i,\delta)$ hence $|f(x) - f(x_i)| < 1$ and thus $-1 < f(x) - f(x_i) < 1$. If we let $M' = \max_i\{f(x_i)\}$ then we see that $-1 < f(x) - M' < 1 \implies |f(x)| < M' + 1$. Hence, f is bounded.

To see that boundedness of E is crucial, consider $f(x) = x$ defined on $\mathbb{R}$. Then f is uniformly continuous because if $\varepsilon > 0$ is given simply let $\delta = \varepsilon$. Then if $|x - y| < \delta$ we have that $|f(x) - f(y)| = |x - y| < \delta = \varepsilon$. But, f is clearly unbounded. $\square$

9. Omitted.

10. Omitted.

11. Suppose f is a uniformly continuous mapping of a metric space X into a metric space Y and prove that $\{f(x_n)\}_{n=1}^\infty$ is a Cauchy sequence in Y for every Cauchy sequence $\{x_n\}_{n=1}^\infty$ in X. Use this to prove the following theorem (Exercise 13): If $E \subset X$ is a dense subset of a metric space X and f is a uniformly continuous function with values in a *complete* metric space Y then f has a continuous extension from E to X. By Exercise 4 then, this extension will be unique because if we have two such extensions they agree on a dense subset of X hence they agree on all of X.

Solution: Let $\varepsilon > 0$ be given and let $\delta > 0$ be such that $d(f(x), f(y)) < \varepsilon$ whenever $d(x,y) < \delta$. Let $\{x_n\}_{n=1}^\infty$ be a Cauchy sequence in X. Thus, there exist N such that $d(x_n, x_m) < \delta$ for all $n, m \ge N$. But this implies that $d(f(x_n), f(x_m)) < \varepsilon$ for all $n, m \ge N$, hence $\{f(x_n)\}_{n=1}^\infty$ is a Cauchy sequence.

Now, let $x \in X$ and let $\{p_n\}_{n=1}^\infty$ be a sequence in E such that $p_n \to x$, which exists because E is dense in X so every element of X is a limit point of E or a point of E, or both (see Definition 2.18(j) on page 32 of the text). If $x \in E$ then let $p_n = x$ for all n. Since $p_n \to x$ we have that $\{p_n\}_{n=1}^\infty$ is a Cauchy sequence by Theorem 3.11(a) on page 53 of the text. By the first result we showed above we then have that $\{f(p_n)\}_{n=1}^\infty$ is Cauchy in Y, hence converges because Y is complete by hypothesis. Therefore, for $x \in X$ we can define $g(x) = \lim_{n\to\infty} f(p_n)$ where $p_n \to x$ and $p_n \in E$ for all n. We have to verify that this definition is well-defined, i.e. it does not depend on the choice of the sequence $\{p_n\}_{n=1}^\infty$ converging to x. So, suppose that $s_n \to x \leftarrow p_n$ both converge to x. Let $\varepsilon > 0$ be given and choose $\delta > 0$ such that $d(f(x), f(y)) < \varepsilon$ whenever $d(x,y) < \delta$. Let N_1 be such that $d(p_n, x) < \delta/2$ for all $n \ge N_1$ and similarly choose N_2 such that $d(s_n, x) < \delta/2$ for all $n \ge N_2$. Letting $N = \max\{N_1, N_2\}$ we see that

$$d(p_n, s_n) \le d(p_n, x) + d(x, s_n) < \frac{\delta}{2} + \frac{\delta}{2} = \delta \; \forall \, n \ge N \implies d(f(p_n), f(s_n)) < \varepsilon \; \forall \, n \ge N.$$

Since ε was arbitrary this means that

$$\lim_{n\to\infty} d(f(p_n), f(s_n)) = 0. \tag{11.1}$$

We now show the following general result: if $x_n \to x$ and $y_n \to y$ in a metric space X then $d(x_n, y_n) \to d(x,y)$. Let $\varepsilon > 0$ be given and choose N such that $d(x_n, x) < \varepsilon/2$ for all $n \ge N$. Then, for $y \in X$, we have that

$$d(x,y) < d(x,x_n) + d(x_n,y) \implies d(x,y) - d(x_n,y) < d(x,x_n) < \varepsilon. \tag{11.2}$$

Reversing the roles of x_n and x we obtain

$$d(x_n, y) < d(x_n, x) + d(x, y) \implies d(x_n, y) - d(x, y) < d(x_n, x) < \varepsilon. \tag{11.3}$$

Thus, combining Equations (11.2) and (11.3) we see that $|d(x,y) - d(x_n, y)| < \varepsilon$ and since ε was arbitrary this shows that $d(x_n, y) \to d(x, y)$ for all $y \in X$. In particular, it holds for $y_j \in \{y_n\}_{n=1}^{\infty}$ hence we obtain that $\lim_{n\to\infty} d(x_n, y_n) = \lim_{m\to\infty}(\lim_{n\to\infty} d(x_n, y_m)) = \lim_{m\to\infty} d(x, y_m) = d(x, y)$. Applying this to our result in Equation (11.1) we have that

$$d\big(\lim_{n\to\infty} f(p_n), \lim_{n\to\infty} f(s_n)\big) = \lim_{n\to\infty} d(f(p_n), f(s_n)) = 0.$$

Thus, $\lim_{n\to\infty} f(p_n) = \lim_{n\to\infty} f(s_n)$ and so defining $g(x) = \lim_{n\to\infty} f(p_n)$ is well-defined since it does not depend on the particular choice of sequence converging to x. Observe that if $x \in E$ then letting $p_n = x$ for all n we see that $g(x) = \lim_{n\to\infty} f(p_n) = f(x)$ hence $g = f$ on E. Now, we have only to verify that g is continuous. To this end, let $x \in X$ and let $\varepsilon > 0$ be given. Let $\delta > 0$ be such that $d(x, y) < \delta \implies d(f(x), f(y)) < \varepsilon/3$, for all $x, y \in X$. Let $y \in X$ such that $d(x, y) < \delta/3$ and let $p_n \to x$ and $s_n \to y$ where $\{p_n\}_{n=1}^{\infty}$ and $\{s_n\}_{n=1}^{\infty}$ are sequences in E. Choose N_1 such that $d(p_n, x) < \delta/3$ and N_2 such that $d(s_n, y) < \delta/3$. Let $N = \max\{N_1, N_2\}$. Then,

$$d(p_n, s_n) \le d(p_n, x) + d(x, y) + d(y, s_n) < \frac{\delta}{3} + \frac{\delta}{3} + \frac{\delta}{3} = \delta \quad \forall\, n \ge N. \tag{11.4}$$

Now, since $g(x) = \lim_{n\to\infty} f(p_n)$ and $g(y) = \lim_{n\to\infty} f(s_n)$ let M_1 and M_2 be such that $d(g(x), f(p_n)) < \varepsilon/3$ for all $n \ge M_1$ and $d(g(y), f(s_n)) < \varepsilon/3$ for all $n \ge M_2$. Let $M = \max\{M_1, M_2\}$ and set $K = \max\{M, N\}$. Then, fixing $n \ge K$ we have that

$$d(g(x), g(y)) \le d(g(x), f(p_n)) + d(f(p_n), f(s_n)) + d(f(s_n), g(y)) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

where $d(f(p_n), f(s_n)) < \varepsilon/3$ because of Equation (11.4) which holds since $n \ge K \ge N$. Thus, g is continuous on X and is the desired extension of f. □

12. A uniformly continuous function of a uniformly continuous function is uniformly continuous. State this more precisely and prove it.

    Solution: The statement can be rephrased as follows. Let $f : X \to Y$ and $g : Y \to Z$ be two uniformly continuous functions such that the image of f is in the domain of g. Then, the composition $(g \circ f) : X \to Z$ is uniformly continuous. To see this, let $\varepsilon > 0$ be given. Then, there exists $\delta_1 > 0$ such that $d(g(y_1), g(y_2)) < \varepsilon$ whenever $d(y_1, y_2) < \delta_1$. Similarly, there exists $\delta_2 > 0$ such that $d(f(x_1), f(x_2)) < \delta_1$ whenever $d(x_1, x_2) < \delta_2$. But, this means that if we choose $x_1, x_2 \in X$ such that $d(x_1, x_2) < \delta_2$ then we have that $d(g(f(x_1)), g(f(x_2))) < \varepsilon$ because $d(f(x_1), f_(x_2)) < \delta_1$. Hence, $(g \circ f)$ is uniformly continuous. □

13. Omitted.

14. Let $I = [0, 1]$ be the closed unit interval. Suppose f is a continuous mapping of I into I. Prove that $f(x) = x$ for at least one $x \in I$.

    Solution: Let $g(x) = f(x) - x$ be defined on I. If $g(1) = 0$ or $g(0) = 0$ then we are finished so suppose this doesn't happen. Then, since f maps into I we see that $f(1) \in [0, 1)$ since $f(1) \neq 1$ by hypothesis, and therefore $g(1) = f(1) - 1 < 0$. Similarly, $g(0) = f(0) - 0 > 0$ because $f(0) \in (0, 1]$ since f maps into I and $f(0) \neq 0$ by hypothesis. But, g is continuous on I by Theorem 4.4(a) on page 85 of the text, and by Theorem 4.6 on page 86 of the text. Hence, by the Intermediate Value Theorem (Theorem 4.23 on page 93 of the text), there exists a $c \in (0, 1)$ such that $g(c) = 0$. But this says precisely that $f(c) = c$, hence f has a fixed point. □

15. Call a mapping of X into Y *open* if $f(V)$ is an open set in Y whenever V is an open set in X. Prove that every continuous open mapping of $\mathbb{R}$ into $\mathbb{R}$ is monotonic.

    Solution: Suppose that we have a continuous open mapping $f : \mathbb{R} \to \mathbb{R}$ which is not monotonic. This means that there exists $x_1 < x_2 < x_3$ such that $f(x_2) > f(x_1)$ and $f(x_2) > f(x_3)$, or $f(x_2) < f(x_1)$ and $f(x_2) < f(x_3)$. More precisely, there exists two points $x_0 < x_1$ such that $f(x_0) < f(x_1)$, and two other points $x_2 < x_3$ such that $f(x_2) > f(x_3)$. But, since f is continuous, by the Intermediate Value Theorem (Theorem 4.23 on page 93 of the text) f achieves all values in $[f(x_0), f(x_3)]$, and so we simply choose x_2 and then we have that $f(x_2) > f(x_1)$ and $f(x_2) > f(x_3)$.

    First, consider the case $f(x_2) > f(x_1)$ and $f(x_2) > f(x_3)$. Since f is continuous on $\mathbb{R}$, for

$$\varepsilon = \frac{f(x_2) - f(x_1)}{2} > 0$$

    there exists $\delta_1 > 0$ such that $|f(x) - f(x_1)| < \varepsilon$ whenever $|x - x_1| < \delta_1$. In particular,

$$f(x) < \frac{f(x_1) + f(x_2)}{2} < f(x_2), \qquad x_1 < x < x_1 + \delta_1 \tag{15.1}$$

    Note that $\delta_1 < x_2 - x_1$ since otherwise we could choose x_2 itself, but then the above inequality would fail. That is, $f(x_2) - f(x_1) \not< (f(x_2) - f(x_1))/2$. So, let $y_1 \in (x_1, x_1 + \delta_1)$. Similarly, for

$$\varepsilon = \frac{f(x_2) - f(x_3)}{2} > 0$$

    there exists a $\delta_2 > 0$ such that $|f(x) - f(x_3)| < \varepsilon$ whenever $|x - x_3| < \delta_2$. In particular,

$$f(x) < \frac{f(x_2) + f(x_3)}{2} < f(x_2), \qquad x_3 - \delta_2 < x < x_3. \tag{15.2}$$

    Again, note that $\delta_2 < x_3 - x_2$ as before since otherwise we could choose x_2 itself and contradict the above inequality. That is, $f(x_2) - f(x_3) \not< (f(x_2) - f(x_3))/2$. Choose $y_2 \in (x_3 - \delta_2, x_3)$. Then, because $\delta_1 < x_2 - x_1$ and $\delta_2 < x_3 - x_2$ we have that $y_2 > y_1$. Since f is continuous on the compact set $[y_1, y_2]$ by Theorem 4.16 on page 89 of the text, f achieves a maximum value at some $p \in [y_1, y_2]$. Now, observe that

$$\sup_{x \in (x_1, x_3)} f(x) \leq \sup_{x \in [y_1, y_2]} f(x)$$

    because of Equations (15.1) and (15.2). That is, $f(x) < f(x_2)$ for all $x \in (x_1, x_1 + \delta_1)$ and for all $x \in (x_3 - \delta_2, x_3)$ hence the $\sup$ must occur on $(x_1 + \delta_1, x_3 - \delta_2) \subset [y_1, y_2]$ because $x_2 \in (x_1, x_3)$. But, since $[y_1, y_2] \subset (x_1, x_3)$ we have that other inclusion

$$\sup_{x \in (x_1, x_3)} f(x) \geq \sup_{x \in [y_1, y_2]} f(x)$$

    hence

$$\sup_{x \in (x_1, x_3)} f(x) = \sup_{x \in [y_1, y_2]} f(x) = f(p), \tag{15.3}$$

since f achieves its maximum on the compact set $[y_1, y_2]$. Now, $f(p) \in f((x_1, x_3))$ hence $f(p)$ must be an interior point because (x_1, x_3) is open and f is an open mapping. Thus, there exists some $\varepsilon > 0$ such that $f(p) + \varepsilon \in f((x_1, x_3))$. But, since $f(p) + \varepsilon > f(p)$ this contradicts Equation (15.3) above which shows that $f(p)$ is the maximum value of f on (x_1, x_3). Hence f must be monotonic. $\square$

16. Omitted.

17. Omitted.

18. Every rational x can be written in the form $x = m/n$, where $n > 0$, and m and n are integers without any common divisors. When $x = 0$, we take $n = 1$. Consider the function f defined on $\mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{if } x \text{ is irrational} \\ \frac{1}{n} & \text{if } x = \frac{m}{n} \end{cases} .$$

Prove that f is continuous at every irrational point, and that f has a simple discontinuity at every rational point.

Solution: Let x be irrational. Then, we see that for any other irrational y we have that $|f(x) - f(y)| = 0$ hence to show continuity at x we need only consider the difference $|f(x) - f(r)|$ for r a rational. Let $\varepsilon > 0$ be given and choose n such that $1/n < \varepsilon$. For each $1 \le j \le n$, consider a neighborhood of x of radius $1/j$. Then, there are at most two rational numbers in lowest terms having j as a denominator in this neighborhood, because the size of the neighborhood is $2/j$ and x is irrational, and the distance between any two adjacent rationals with j in the denominator, i.e. between an m/j and $(m+1)/j$, is $1/j$. That is, if there were three rationals having j as a denominator in this neighborhood, then that would force x to be the middle one, contradicting that x was irrational. Hence we can write

$$\frac{m_j}{j} < x < \frac{m_j + 1}{j},$$

for some m_j so there are no rationals in lowest terms having a denominator j between m_j/j and $(m_j + 1)/j$ because this gap is of size $1/j$ and rationals with denominators j are evenly spaced, i.e. the distance between any such two is at least $1/j$. Then, let

$$I_j = \left(\frac{m_j}{j}, \frac{m_j + 1}{j} \right),$$

and consider $I = \cap_{j=1}^{n} I_j$. Then, I contains no rationals with denominator j for all $1 \le j \le n$ by construction (because each I_j excludes all rationals with denominator j) yet still contains x because $x \in I_j$ for all $1 \le j \le n$ and is open because it is a finite intersection of open sets. Now, let $\delta > 0$ be such that $B(x, \delta) \subset I$. Then, for all rational $r \in B(x, \delta) \subset I$ we see that r has a denominator which is larger than n when written in lowest terms, say $r = p/q$ with $q > n$, for otherwise if $q \le n$ then $q = j$ for some $1 \le j \le n$, a contradiction that I contains no rationals with any such j in the denominator. Therefore,

$$|f(x) - f(r)| = \frac{1}{q} < \frac{1}{n} < \varepsilon,$$

and since r was arbitrary in $B(x, \delta), \varepsilon$ was arbitrary, and x was arbitrary, we see that f is therefore continuous at all irrational x.

Another way to say this is, pick m_j such that

$$\delta_j = \left| x - \frac{m_j}{j} \right| \quad \text{is smallest.}$$

We know it is greater than 0 because x is irrational, and there exists a smallest one because rationals with denominator j do not get arbitrarily close but are always separated by a distance of at least $1/j$ hence there exists a closest rational with denominator $1/j$ to x. Then, if we let $0 < \delta < \min\{\delta_1, \ldots, \delta_n\}$ we see that $B(x, \delta)$ has the same property as the analogous neighborhood constructed above.

Now, suppose that $r = p/q \in \mathbb{R}$ is rational. Then, since the irrationals are dense in $\mathbb{R}$ any neighborhood around r will contain an irrational x. But then, $|f(r) - f(x)| = 1/q$ for any such rational and we see that no matter how small a neighborhood we choose around r there will be points in it such that $|f(r) - f(x)|$ cannot be made arbitrarily small, hence f is discontinuous at every rational point. To see that this is a simple discontinuity, i.e. that $f(x^+)$ and $f(x^-)$ exist, we observe that $\lim_{x \to y} f(x) = 0$ for *any* point $y \in \mathbb{R}$, in particular a rational point, because as discussed above, approaching any number arbitrarily close with rationals implies their denominators in lowest terms are getting arbitrarily large, and if we approach with irrationals then f is 0 when evaluated at them. Since the value $f(r)$ at a rational point is never zero, this shows the discontinuity is simple because both limits exist, they just don't equal the value of the function at the point. In particular, this also shows that f is continuous at irrational points because then there, the value of the limits does equal the value of the function. □

19. Suppose f is a real function with domain $\mathbb{R}$ which has the intermediate value property: If $f(a) < c < f(b)$, then $f(x) = c$ for some $a < x < b$. Suppose also for every rational r, that the set of all x with $f(x) = r$ is closed. Prove that f is continuous.

    Hint: If $x_n \to x_0$ but $f(x_n) > r > f(x_0)$ for some r and for all $n \geq 1$, then $f(t_n) = r$ for some t_n between x_0 and x_n; thus $t_n \to x_0$. Find a contradiction.

    Solution: If f is not continuous at some $x_0 \in \mathbb{R}$ then that means that $f(x_n) \nrightarrow f(x_0)$ for some sequence $x_n \to x_0$, by Theorem 4.6 on page 86 with Theorem 4.2 on page 84 of the text. Since f is not constant (otherwise it is continuous), we can assume without loss of generality that $f(x_n) > r > f(x_0)$ for some rational r for all $n \geq 1$. That is, we rename our sequence to only include terms such that $f(x_n) > f(x_0)$ and if there are only finitely many such terms then we choose r such that $f(x_0) > r > f(x_n)$. This is possible because $\mathbb{Q}$ is dense in $\mathbb{R}$, and if we only include terms where $f(x_n) \neq f(x_0)$ (there are infinitely many such terms otherwise we'd have convergence). Since f has the intermediate value property we see that there exists t_n between x_n and x_0 such that $f(t_n) = r$ for all $n \geq 1$. But, since $x_n \to x_0$ we must have that $t_n \to x_0$ because $d(t_n, x_0) \leq d(x_n, x_0)$, which goes to 0. Since $S = \{x \mid f(x) = r\}$ is closed by hypothesis and $t_n \in S$ for all $n \geq 1$ by construction we see that $x_0 \in S$ so that $f(x_0) = r$, a contradiction to the choice of r. Therefore, f is continuous at every $x \in \mathbb{R}$. □

20. If E is a nonempty subset of a metric space X, define the distance from $x \in X$ to E by

    $$\rho_E(x) = \inf_{z \in E} d(x, z)$$

    (a) Prove that $\rho_E(x) = 0$ if and only if $x \in \overline{E}$.

(b) Prove that $\rho_E(x)$ is a uniformly continuous function on X, by showing that

$$|\rho_E(x) - \rho_E(y)| \leq d(x, y)$$

for all $x, y \in X$.
Hint: $\rho_E(x) \leq d(x,z) \leq d(x,y) + d(y,z)$, so that $\rho_E(x) \leq d(x,y) + \rho_E(y)$.

Solution:

(a) First, suppose that $\rho_E(x) = 0$. This means that for all $n \geq 1$ there exists $z_n \in E$ such that $d(x, z_n) < 1/n$ by definition of $\inf$ hence $z_n \to x$ and so x is a limit point of E thus $x \in \overline{E}$. On the other hand, suppose that $x \in \overline{E}$ so there exists a sequence $z_n \to x$ with $\{z_n\}_{n=1}^{\infty} \subset E$. Then, for all $\varepsilon > 0$ there exists an N such that $d(x, z_n) < \varepsilon$ for all $n \geq N$, hence the distance from x to some point of E can be made arbitrarily small, i.e. $\rho_E(x) = \inf_{z\in E} d(x, z) = 0$. Otherwise, if $\rho_E(x) = \delta > 0$ simply choose $0 < \varepsilon < \delta$. Then, we can find n such that $d(x, z_n) < \varepsilon < \delta = \inf_{z\in E} d(x, z)$, contradicting the definition of the $\inf$.

(b) By the triangle inequality for $x \in X$ we have that for all $z \in E$ and for all $y \in X$

$$\rho_E(x) = \inf_{z'\in E} d(x, z') \leq d(x, z) \leq d(x, y) + d(y, z).$$

Since z was arbitrary we must have that

$$\rho_E(x) \leq d(x, y) + \inf_{z\in E} d(y, z) = d(x, y) + \rho_E(y).$$

Hence

$$\rho_E(x) - \rho_E(y) \leq d(x, y).$$

Reversing the roles of x and y above we obtain $\rho_E(y) - \rho_E(x) \leq d(x, y)$ hence we have

$$|\rho_E(x) - \rho_E(y)| \leq d(x, y).$$

This is precisely uniform continuity because if $\varepsilon > 0$ is given and $x, y \in X$ let $\delta = \varepsilon$. Then, $d(x, y) < \delta = \varepsilon \implies |\rho_E(x) - \rho_E(y)| \leq d(x, y) < \delta = \varepsilon$, hence $\rho_E(x)$ is uniformly continuous on X. $\square$

21. Suppose K and F are disjoint sets in a metric space X such that K is compact and F is closed. Prove that there exists $\delta > 0$ such that $d(p, q) > \delta$ if $p \in K$ and $q \in F$. Show that the conclusion may fail for two disjoint closed sets if neither is compact.

Solution: We show that $\inf\{d(p, q) \mid p \in K, q \in F\} > 0$ and hence is greater than some positive δ. Since $K \cap F = \emptyset$ we have that $d(x, y) > 0$ for all $x \in K, y \in F$ because $d(x, y) = 0 \iff x = y$. Therefore

$$0 \leq \inf\{d(x, y) \mid x \in K, y \in F\}.$$

So, suppose that

$$0 = \inf\{d(x, y) \mid x \in K, y \in F\}.$$

This means that for each $n \in \mathbb{N}$, there exists $x_n \in K, y_n \in F$ such that $d(x_n, y_n) < 1/n$ by definition of the $\inf$. Then, $\{x_n\} \subset K$ is a sequence and since K is compact, there exists a convergent subsequence $\{x_{n_k}\} \to x_0 \in K$. Then, $\{y_{n_k}\} \subset F$ is a sequence. Let $\varepsilon > 0$ be given. Then, there exists a K_1 such that

$$d(x_{n_k}, y_{n_k}) < \frac{\varepsilon}{2}, \quad k \geq K_1$$

and since $x_{n_k} \to x_0$, there exists an K_2 such that

$$d(x_{n_k}, x_0) < \frac{\varepsilon}{2}, \quad k \geq K_2.$$

If we let $K = \max\{K_1, K_2\}$ then

$$d(y_{n_k}, x_0) \leq d(y_{n_k}, x_{n_k}) + d(x_{n_k}, x_0) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad \forall\, k \geq K.$$

Thus, $y_{n_k} \to x_0$ also. Since F is closed it contains its limit points and therefore $x_0 \in F$, contradicting that $K \cap F = \emptyset$. Therefore, we must have that $0 < \inf\{d(x,y) \mid x \in K, y \in F\}$.

Now, consider the sets $A = \mathbb{Z}^+ \backslash \{1,2\} = \{n \in \mathbb{Z} \mid n \geq 3\}$ and $B = \{n + 1/n \mid n \in \mathbb{Z}^+\}$. Both are closed and $A \cap B = \emptyset$. But, for all $\varepsilon > 0$ if we choose $n > 1/\varepsilon$ then we see that $|n + 1/n - n| = 1/n < \varepsilon$ hence the conclusion fails because we can find points in A and B that are arbitrarily close. □

Alternate Solution: We can see this result in another way by using the results of Exercise 20 above. First, by Exercise 20(a) we see that $\rho_F(x) = 0 \iff x \in \overline{F} = F$ because F is closed, hence we have that $\rho_F(k) > 0$ for all $k \in K$ because $K \cap F = \emptyset$. By Exercise 20(b) $\rho_F(k)$ is continuous on K and so by Theorem 4.16 on page 89 of the text, since K is compact, $\rho_F(k)$ achieves its minimum on K, namely there exists $k \in K$ such that $\rho_F(k) = m > 0$, where we have the inequality because $k \in K$ and we saw above that ρ_F was positive on K. Hence, for $p \in K$ and $q \in F$ we have that $0 < m/2 < \rho_F(p) = \inf_{q' \in F} d(p, q') \leq d(p,q)$. □

Chapter 5: Differentiation

1. Let f be defined for all real x, and suppose that

$$|f(x) - f(y)| \leq (x-y)^2$$

for all real x and y. Prove that f is a constant.

Solution: We can rewrite this condition as

$$|f(x) - f(y)| \leq |x-y||x-y|,$$

and therefore we can divide by $|x-y|$ since it is positive and preserve the inequality to obtain

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

Thus the limit as $x \to y$ of the left hand side exists (for a given $\varepsilon > 0$ choose $\delta = \varepsilon$) and taking this limit gives

$$|f'(y)| \leq 0 \implies f'(y) = 0.$$

Since y was arbitrary we see that f is constant by Theorem 5.11 on page 108 of the text. □

2. Suppose that $f'(x) > 0$ on (a, b). Prove that f is strictly increasing in (a, b), and let g be its inverse function. Prove that g is differentiable, and that

$$g'(f(x)) = \frac{1}{f'(x)} \quad (a < x < b).$$

Solution: Let $a < x < y < b$. Then, f is continuous on $[x, y]$ (because it is differentiable on $(a, b) \supsetneq [x, y]$) and differentiable on (x, y). Thus, by The Mean Value Theorem (Theorem 5.10 on page 108 of the text), there exists $c \in (x, y)$ such that

$$f(y) - f(x) = (y - x)f'(c) > 0,$$

where the inequality follows because $f'(c) > 0$ by hypothesis and $y - x > 0$ by construction. Since f is strictly increasing on (a, b), it is injective and therefore the inverse function g is well-defined. Let $a < x < y < b$. Then,

$$g'(f(x)) = \lim_{f(y) \to f(x)} \frac{g(f(y)) - g(f(x))}{f(y) - f(x)} = \lim_{y \to x} \frac{y - x}{f(y) - f(x)} = \frac{1}{f'(x)},$$

where the second equality follows since f is injective (because it is strictly increasing) thus $f(y) \to f(x) \iff y \to x$. $\square$

3. Suppose g is a real function on $\mathbb{R}$, with bounded derivative (say $|g'| \leq M$). Fix $\varepsilon > 0$, and define $f(x) = x + \varepsilon g(x)$. Prove that f is one-to-one if ε is small enough. (A set of admissible values of ε can be determined which depends only on M.)

Solution: Let $x < y$. By the Mean Value Theorem (Theorem 5.10 on page 108 of the text), there exists a $x < c < y$ such that

$$g(x) - g(y) = g'(c)(x - y).$$

Thus, we obtain

$$\begin{aligned} f(x) - f(y) &= x - y + \varepsilon(g(x) - g(y)) \\ &= x - y + \varepsilon g'(c)(x - y) \\ &= (x - y)(1 + \varepsilon g'(c)). \end{aligned} \tag{3.1}$$

Since $|g'| \leq M$ we have that $-M \leq g'(x) \leq M$ for all $x \in \mathbb{R}$. Thus, $1 - \varepsilon M \leq 1 + \varepsilon g'(x) \leq 1 + \varepsilon M$. If we choose $\varepsilon \leq \frac{1}{2M}$ then this inequality becomes

$$\frac{1}{2} = 1 - \frac{1}{2} \leq 1 - \varepsilon M \leq 1 + \varepsilon g'(x) \leq 1 + \varepsilon M \leq 1 + \frac{1}{2} \implies 1 + \varepsilon g'(x) > 0, \quad \forall\, x \in \mathbb{R}. \tag{3.2}$$

In particular, this holds for $x = c$ and thus from Equation (3.1) we see that $f(x) - f(y) < 0$ since $x - y < 0$ by construction and we see in Equation (3.2) that $1 + \varepsilon g'(c) > 0$. That is, $f(x) \neq f(y)$ so f is one-to-one. $\square$

4. If

$$C_0 + \frac{C_1}{2} + \cdots + \frac{C_{n-1}}{n} + \frac{C_n}{n+1} = 0,$$

where $C_0, \ldots, C_n$ are real constants, prove that the equation

$$C_0 + C_1 x + \cdots + C_{n-1}x^{n-1} + C_n x^n = 0$$

has at least one real root between 0 and 1.

Solution: Consider the polynomial defined by

$$p(x) = C_0 x + \frac{C_1}{2}x^2 + \cdots + \frac{C_n}{n+1}x^{n+1}, \quad x \in [0,1].$$

Then, $p(1) = C_0 + \frac{C_1}{2} + \cdots + \frac{C_{n-1}}{n} + \frac{C_n}{n+1} = 0 = p(0)$ and by The Mean Value Theorem (Theorem 5.10 on page 108 of the text) we see that

$$C_0 + C_1 c + \cdots + C_{n-1}c^{n-1} + C_n c^n = p'(c) = \frac{p(1)-p(0)}{1-0} = 0, \quad c \in (0,1),$$

hence c is a root of $C_0 + C_1 x + \cdots + C_{n-1}x^{n-1} + C_n x^n$. $\square$

5. Suppose f is defined and differentiable for every $x > 0$, and $f'(x) \to 0$ as $x \to \infty$. Put $g(x) = f(x+1) - f(x)$. Prove that $g(x) \to 0$ as $x \to \infty$.

   Solution: Let $\varepsilon > 0$ be given and let N be such that $|f'(x)| < \varepsilon$ for all $x \geq N$. Then, for such x, by The Mean Value Theorem (Theorem 5.10 on page 108 of the text) we have that

$$|g(x)| = |f(x+1) - f(x)| = \left|\frac{f(x+1)-f(x)}{x+1-x}\right| = |f'(c)|, \quad c \in (x, x+1).$$

   Thus $|g(x)| = |f'(c)| < \varepsilon$ since $c > x \geq N$. Since ε was arbitrary we have that $g(x) \to 0$ as $x \to \infty$. $\square$

6. Suppose

   (a) f is continuous for $x \geq 0$,

   (b) $f'(x)$ exists for $x > 0$,

   (c) $f(0) = 0$,

   (d) f' is monotonically increasing.

   Put

$$g(x) = \frac{f(x)}{x} \quad x > 0$$

   and prove that g is monotonically increasing.

   Solution: Let $0 < x < y$. By The Mean Value Theorem (Theorem 5.10 on page 108 of the text) we see that there exist $c_1 \in (0,x)$ and $c_2 \in (0,y)$ such that

$$\frac{f(x)}{x} = \frac{f(x)-f(0)}{x-0} = f'(c_1), \quad \frac{f(y)}{y} = \frac{f(y)-f(0)}{y-0} = f'(c_2).$$

   Therefore,

$$g(y) - g(x) = \frac{f(y)}{y} - \frac{f(x)}{x} = f'(c_2) - f'(c_1) \geq 0,$$

   where the inequality follows since $c_2 \geq c_1$ (because $y > x$) and f' is monotonically increasing. Therefore, g is monotonically increasing. $\square$

7. Suppose $f'(x), g'(x)$ exist, $g'(x) \neq 0$, and $f(x) = g(x) = 0$. Prove that

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

(This holds also for complex functions.)

Solution: Simply consider

$$\frac{f'(x)}{g'(x)} = \frac{\lim_{t\to x} \frac{f(t)-f(x)}{t-x}}{\lim_{t\to x} \frac{g(t)-g(x)}{t-x}} = \lim_{t\to x} \frac{f(t)}{g(t)}.$$

Note, we cannot trivially use L'Hospital's Rule because we don't know that $\lim_{t\to x} \frac{f'(t)}{g'(t)} = \frac{f'(x)}{g'(x)}$, that is, we don't know that the derivatives are continuous at x, only that they exist at x. We can see that the same proof works for complex functions if we write $f(x) = f_1(x) = if_2(x)$ where f_1, f_2 are real-valued functions. □

8. Suppose f' is continuous on $[a, b]$ and $\varepsilon > 0$. Prove that there exists $\delta > 0$ such that

$$\left| \frac{f(t) - f(x)}{t - x} - f'(x) \right| < \varepsilon$$

whenever $0 < |t - x| < \delta, a \leq x \leq b, a \leq t \leq b$. (This could be expressed by saying that f is uniformly differentiable on $[a, b]$ if f' is continuous on $[a, b]$.) Does this hold for vector-valued functions too?

Solution: Since f' is continuous on the compact set $[a, b]$ it is uniformly continuous on $[a, b]$ and therefore there exists $\delta > 0$ such that $|f'(t) - f'(x)| < \varepsilon$ for all $0 < |t - x| < \delta$, $a \leq t \leq b, a \leq x \leq b$. By The Mean Value Theorem (Theorem 5.10 on page 108 of the text) there exists a c between t and x such that $f'(c) = \frac{f(t)-f(x)}{t-x}$. Since c is between t and x we have that $0 < |c - x| < \delta$ because $|c - x| < |t - x|$. Therefore $\left|\frac{f(t)-f(x)}{t-x} - f'(x)\right| = |f'(c) - f'(x)| < \varepsilon$.

This does not hold for vector-valued functions. We can use $f(x) = (\cos(x), \sin(x))$ as a counter example. □

9. Let f be a continuous real function on $\mathbb{R}$, of which it is known that $f'(x)$ exists for all $x \neq 0$ and that $f'(x) \to 3$ as $x \to 0$. Does it follow that $f'(0)$ exists?

Solution: Yes, $f'(0)$ must exist. This follows from Corollary to Theorem 5.12 on page 109 of the text. For suppose that $f'(0)$ did not exist. Since $f'(x) \to 3$ as $x \to 0$ this means that f' has a discontinuity of the first kind (i.e. a *simple discontinuity*) at $x = 0$ because the left and right hand limits exist yet f' is not continuous at 0 because it is not defined there. But this contradicts the Corollary. We also give a direct proof of the following more general statement: Suppose that f is continuous on an open interval I containing x_0, and let f' be defined on I except possibly at x_0 and $f'(x) \to L$ as $x \to x_0$. Then $f'(x_0) = L$. To see this, let $g(x) = x$. Then we have that $\lim_{h\to 0} g(h) = 0$ and $\lim_{h\to 0} f(x_0 + h) - f(x_0) = 0$. Therefore, we can apply L'Hospital's Rule (Theorem 5.13 on page 109 of the text) and obtain that

$$\begin{aligned} f'(x_0) &= \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} \\ &= \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{g(h)} \\ &= \lim_{h\to 0} \frac{(f(x_0+h)-f(x_0))'}{g'(h)} \\ &= \lim_{h\to 0} f'(x_0+h) \\ &= L, \end{aligned}$$

where the last equality follows because by hypothesis $f'(x) \to L$ as $x \to x_0$ so we have that $\lim_{h\to 0} f'(x_0+h) = L$. □

10. Suppose f and g are complex differentiable functions on $(0,1)$, $f(x) \to 0, g(x) \to 0, f'(x) \to A, g'(x) \to B$ as $x \to 0$, where A and B are complex numbers and $B \neq 0$. Prove that

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = \frac{A}{B}.$$

Compare with Example 5.18 on page 112 of the text. *Hint:*

$$\frac{f(x)}{g(x)} = \left(\frac{f(x)}{x} - A\right) \cdot \frac{x}{g(x)} + A \cdot \frac{x}{g(x)}.$$

Apply Theorem 5.13 on page 109 of the text to the real and imaginary parts of $f(x)/x$ and $g(x)/x$.

Solution: Write $f(x) = f_1(x) + if_2(x)$ where f_1, f_2 are real-valued functions. Then, we have that

$$\frac{df}{dx} = \frac{d}{dx}(f_1(x) + if_2(x)) = \frac{df_1(x)}{dx} + i\frac{df_2(x)}{dx}. \tag{10.1}$$

Now, since f_1 and f_2 are real-valued, we can apply L'Hospital's Rule (Theorem 5.13 on page 109 of the text) to the functions $\frac{f_1(x)}{x}$ and $\frac{f_2(x)}{x}$, since $f_1(x), f_2(x), x \to 0$ as $x \to 0$. This shows that

$$\lim_{x\to 0} \frac{f_1(x)}{x} = \lim_{x\to 0} f_1'(x)$$

$$\lim_{x\to 0} \frac{f_2(x)}{x} = \lim_{x\to 0} f_2'(x).$$

Therefore, we obtain that

$$\lim_{x\to 0} \frac{f(x)}{x} = \lim_{x\to 0} \left(\frac{f_1(x)}{x} + i\frac{f_2(x)}{x}\right) = \lim_{x\to 0}(f_1'(x) + if_2'(x)) = \lim_{x\to 0} f'(x) = A,$$

where the last equality follows from Equation (10.1). An analogous construction shows that $\lim_{x\to 0} g(x)/x = g'(x) = B$ thus $\lim_{x\to 0} x/g(x) = 1/B$. Then, we have that

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = \lim_{x\to 0} \left(\frac{f(x)}{x} - A\right) \cdot \frac{x}{g(x)} + A \cdot \frac{x}{g(x)} = (A - A) \cdot \frac{1}{B} + A \cdot \frac{1}{B} = \frac{A}{B}.$$

In Example 5.18, we have that $g'(x) \to \infty$ as $x \to 0$, not a complex number. □

11. Suppose f is defined in a neighborhood of x, and suppose that $f''(x)$ exists. Show that

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

Show by an example that the limit may exist even if $f''(x)$ does not. *Hint:* Use Theorem 5.13 on page 109 of the text.

Solution: Applying L'Hospital's Rule (Theorem 5.13 on page 109 of the text), with respect to h, to the functions $F(h)=f(x+h)+f(x-h)-2f(x)$ and $G(h)=h^2$ we see that

$$\lim_{h\to 0}\frac{f(x+h)+f(x-h)-2f(x)}{h^2}=\lim_{h\to 0}\frac{F(h)}{G(h)}=\lim_{h\to 0}\frac{F'(h)}{G'(h)}=\lim_{h\to 0}\frac{f'(x+h)-f'(x-h)}{2h}, \tag{11.1}$$

where we have a minus sign in the numerator of the last expression from the chain rule: $(f(x-h))'=-f'(x-h)$. Now, we have that

$$\begin{aligned} f''(x) &= \frac{1}{2}(f''(x)+f''(x)) \\ &= \frac{1}{2}\left(\lim_{h\to 0}\frac{f'(x+h)-f'(x)}{h}+\lim_{h\to 0}\frac{f'(x-h)-f'(x)}{-h}\right) \\ &= \frac{1}{2}\lim_{h\to 0}\frac{f'(x+h)-f'(x-h)}{h} \\ &= \lim_{h\to 0}\frac{f'(x+h)-f'(x-h)}{2h} \\ &= \lim_{h\to 0}\frac{f(x+h)+f(x-h)-2f(x)}{h^2}, \end{aligned}$$

where the last equality follows by Equation (11.1). To see that the limit may exist even when $f''(x)$ does not, consider the function $f(x)=x|x|$. Then $f''(0)$ does not exist yet

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

□

12. Omitted.

13. Omitted.

14. Let f be a differentiable real function defined on (a,b). Prove that f is convex if and only if f' is monotonically increasing. Assume next that $f''(x)$ exists for all $x\in(a,b)$, and prove that f is convex if and only if $f''(x)\geq 0$ for all $x\in(a,b)$.

Solution: First, suppose that f is convex so that $f(\lambda x+(1-\lambda)y)\leq\lambda f(x)+(1-\lambda)f(y)$ for all $0<\lambda<1$ and $a<x<b, a<y<b$. This means that $f(\lambda(x-y)+y)\leq\lambda(f(x)-f(y))+f(y)$. If we choose $a<y<x<b$ so that $x-y>0$ then subtracting $f(y)$ from both sides and dividing by $\lambda(x-y)$ gives

$$\frac{f(\lambda(x-y)+y)-f(y)}{\lambda(x-y)}\leq\frac{f(x)-f(y)}{x-y},$$

where the inequality is preserved because $\lambda(x-y)>0$. Taking the limit as $y\to x$ of both sides then shows that

$$f'(y)\leq f'(x),\quad a<y<x<b,$$

where we get $f'(y)$ on the left because $y \to x \implies \lambda(x-y) \to 0$. Hence, $f'(x)$ is monotonically increasing on (a,b).

Now, suppose that f' is monotonically increasing on (a,b). Let $x, y \in (a,b)$ and $0 < \lambda < 1$. Without loss of generality assume $x < y$. Set $x_0 = y - \lambda(y-x)$ which means that

$$\begin{aligned} \lambda &= \frac{y-x_0}{y-x}, \\ 1-\lambda &= \frac{y-x}{y-x} - \frac{y-x_0}{y-x} = \frac{y-x-y+x_0}{y-x} = \frac{x_0-x}{y-x}. \end{aligned} \tag{14.1}$$

This shows that

$$\begin{aligned} \lambda x + (1-\lambda)y &= x\left(\frac{y-x_0}{y-x}\right) + y\left(\frac{x_0-x}{y-x}\right) \\ &= \frac{xy - xx_0 + yx_0 - yx}{y-x} \\ &= \frac{yx_0 - xx_0}{y-x} \\ &= x_0. \end{aligned} \tag{14.2}$$

Write $F(x,y,\lambda) = f(\lambda x + (1-\lambda)y) - (\lambda f(x) + (1-\lambda)f(y))$ for notational purposes (so $F(x,y,\lambda)$ is the convexity relation). Then, using Equations (14.1) and (14.2) we have that

$$\begin{aligned} F(x,y,\lambda) &= f(x_0) - \left(\frac{y-x_0}{y-x}\right)f(x) - \left(\frac{x_0-x}{y-x}\right)f(y) \\ &= \left(\frac{y-x_0}{y-x}\right)(f(x_0)-f(x)) + \left(\frac{x_0-x}{y-x}\right)(f(x_0)-f(y)), \end{aligned} \tag{14.3}$$

where the second equality follows because $\frac{y-x_0}{y-x} + \frac{x_0-x}{y-x} = 1$. Now, two applications of the Mean Value Theorem yields

$$\begin{aligned} f(x_0) - f(x) &= f'(\eta)(x_0 - x) \quad \eta \in (x, x_0) \\ f(x_0) - f(y) &= f'(\xi)(x_0 - y) \quad \xi \in (x_0, y). \end{aligned} \tag{14.4}$$

Substituting this into Equation (14.3) we obtain

$$\begin{aligned} F(x,y,\lambda) &= \left(\frac{y-x_0}{y-x}\right)(x_0-x)f'(\eta) + \left(\frac{x_0-x}{y-x}\right)(x_0-y)f'(\xi) \\ &= \left(\frac{y-x_0}{y-x}\right)(x_0-x)f'(\eta) - \left(\frac{x_0-x}{y-x}\right)(y-x_0)f'(\xi) \\ &= \left[\left(\frac{y-x_0}{y-x}\right)(x_0-x)\right](f'(\eta)-f'(\xi)) \\ &\leq 0, \end{aligned} \tag{14.5}$$

where the inequality follows because $y-x_0, y-x, x_0-x > 0$ but from Equation (14.4) we see that $\eta < \xi \implies f'(\eta) \leq f'(\xi)$ (because f' is monotonically increasing by hypothesis) $\implies f'(\eta) - f'(\xi) \leq 0$. But, Equation (14.5) is precisely

$$f(\lambda x + (1-\lambda)y) - (\lambda f(x) + (1-\lambda)f(y)) \leq 0,$$

by the definition of $F(x, y, \lambda)$, hence f is convex. By Theorem 5.11(a) on page 108 we see that $f''(x) \geq 0 \iff f'(x)$ is monotonically increasing because $f''(x) = (f'(x))'$, hence f is convex if and only if $f''(x) \geq 0$ by the above result. $\square$

Chapter 6: The Riemann-Stieltjes Integral

Before we present the solutions to selected exercises in this chapter, we prove one property of the integral which was not shown in the text, namely Theorem 6.12(c) which states that if $f \in \mathscr{R}(\alpha)$ on $[a, b]$ and $a < c < b$ then $f \in \mathscr{R}(\alpha)$ on $[a, c]$ and on $[c, b]$ and

$$\int_a^b f d\alpha = \int_a^c f d\alpha + \int_c^b f d\alpha.$$

First, since $f \in \mathscr{R}(\alpha)$ on $[a, b]$ we can choose a partition $P = \{a = x_0 \leq x_1 \leq \cdots \leq x_n = b\}$ of $[a, b]$ such that

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

Now, let $P^* = P \cup \{c\}$, a (possible) refinement of P, and let $x_j = c$ for some $1 \leq j \leq n - 1$. Then, by Theorem 6.4 on page 123 of the text we see that $L(P, f, \alpha) \leq L(P^*, f, \alpha)$ and $U(P^*, f, \alpha) \leq U(P, f, \alpha)$. If we let $P_1 = \{a = x_0 \leq x_1 \leq \cdots \leq x_j = c\}$ then,

$$\begin{aligned}
\varepsilon &> U(P, f, \alpha) - L(P, f, \alpha) \\
&= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\
&\geq \sum_{i=1}^{j} (M_i - m_i)\Delta\alpha_i \\
&= U(P_1, f, \alpha) - L(P_1, f, \alpha),
\end{aligned}$$

where the second inequality follows since $M_i - m_i \geq 0$ for all i. Therefore, $f \in \mathscr{R}(\alpha)$ on $[a, c]$ since P_1 is a partition of $[a, c]$. Modifying the above procedure accordingly by letting $P_2 = \{c = x_j \leq x_{j+1} \leq \cdots \leq x_n = b\}$ and keeping $\sum_{i=j+1}^{n} (M_i - m_i)\Delta\alpha_i$ instead, we similarly see that $f \in \mathscr{R}(\alpha)$ on $[c, b]$ also. Now, let P be *any* partition of $[a, b]$ and $P^* = P \cup \{c\}$, with P_1 and P_2 analogous definitions as above. Then, we have that

$$\begin{aligned}
U(P, f, \alpha) &\geq U(P^*, f, \alpha) \\
&= \sum_{i=1}^{n} M_i \Delta\alpha_i \\
&= \sum_{i=1}^{j} M_i \Delta\alpha_i + \sum_{i=j+1}^{n} M_i \Delta\alpha_i \\
&= U(P_1, f, \alpha) + U(P_2, f, \alpha) \\
&\geq \int_a^c f d\alpha + \int_c^b f d\alpha.
\end{aligned}$$

Since P was arbitrary, we see that

$$\int_a^b f d\alpha = \inf_P U(P, f, \alpha) \geq \int_a^c f d\alpha + \int_c^b f d\alpha.$$

Now, if we use $L(P, f, \alpha)$ instead we find that

$$\begin{aligned}
L(P, f, \alpha) &\le L(P^*, f, \alpha) \\
&= \sum_{i=1}^{n} m_i \Delta\alpha_i \\
&= \sum_{i=1}^{j} m_i \Delta\alpha_i + \sum_{i=j+1}^{n} m_i \Delta\alpha_i \\
&= L(P_1, f, \alpha) + L(P_2, f, \alpha) \\
&\le \int_a^c f d\alpha + \int_c^b f d\alpha.
\end{aligned}$$

Since P was arbitrary, we see that

$$\int_a^b f d\alpha = \sup_P L(P, f, \alpha) \le \int_a^c f d\alpha + \int_c^b f d\alpha,$$

which together with the result above shows that $\int_a^b f d\alpha = \int_a^c f d\alpha + \int_c^b f d\alpha$ as desired. □

1. Suppose α increases on $[a, b]$, $a \le x_0 \le b$, α is continuous at x_0, $f(x_0) = 1$, and $f(x) = 0$ if $x \ne x_0$. Prove that $f \in \mathscr{R}(\alpha)$ and that $\int_a^b f d\alpha = 0$.

   Solution: By Theorem 6.10 on page 126 of the text we see that $f \in \mathscr{R}(\alpha)$ since α is continuous at the only point where f is discontinuous. Now, we see that

   $$\begin{aligned}
   U(P, f, \alpha) &= \sum_{i=1}^{n} \sup_{[x_{i-1}, x_i]} f(x)\, \Delta\alpha_i = \Delta\alpha_i, \quad x_0 \in [x_{i-1}, x_i], \\
   L(P, f, \alpha) &= \sum_{i=1}^{n} \inf_{[x_{i-1}, x_i]} f(x)\, \Delta\alpha_i = 0,
   \end{aligned}$$

   where $L(P, f, \alpha) = 0$ because for any partition we have that x_0 is contained in some interval of positive length hence the inf of f on that interval will be 0 since it will contain points not equal to x_0. Since P was arbitrary this implies that $\int_a^b f d\alpha = \sup_P L(P, f, \alpha) = 0$, because f was already shown to be Riemann integrable. We also have that that $U(P, f, \alpha) = \alpha_i + \alpha_{i+1}$ if $x_0 = x_i$. So, we see that $\inf_P U(P, f, \alpha) = 0$ since taking finer and finer partitions means that $\Delta\alpha_i \to 0$. Since this is equal to $\sup_P L(P, f, \alpha)$ we again see that $\int_a^b f d\alpha = 0$. □

2. Suppose $f \ge 0$, f is continuous on $[a, b]$, and $\int_a^b f(x)dx = 0$. Prove that $f(x) = 0$ for all $x \in [a, b]$. (Compare this with Exercise 1).

   Solution: Suppose there exists $x_0 \in [a, b]$ such that $f(x_0) = L > 0$ (because $f \ge 0$). Letting $0 < \varepsilon < L$ we see that there exists $\delta > 0$ such that $f(x) \in (L - \varepsilon, L + \varepsilon)$ whenever $x \in (x_0 - \delta, x_0 + \delta)$ by the continuity of f. We then have that

   $$\int_a^b f(x)dx = \int_a^{x_0 - \delta} f(x)dx + \int_{x_0 - \delta}^{x_0 + \delta} f(x)dx + \int_{x_0 + \delta}^{b} f(x)dx > 0,$$

   where the inequality follows since on $(x_0 - \delta, x_0 + \delta)$, $f(x) > 0$ since $L - \varepsilon > 0$ by the choice of ε. This contradicts that $\int_a^b f(x)dx = 0$, hence no such x_0 can exist. Exercise 1 shows that the continuity of f is necessary for this result since in Exercise 1 we have a function whose integral is zero, but $f(x) \ne 0$ for some $x \in [a, b]$. □

3. Omitted.

4. If $f(x) = 0$ for all irrational x and $f(x) = 1$ for all rational x, prove that $f \notin \mathscr{R}$ on $[a, b]$ for any $a < b$.

   Solution: Let $a < b$ and let P be a partition of $[a, b]$. Because rational and irrational numbers are both dense in $\mathbb{R}$ we see that any interval of positive length will contain both types of numbers hence $[x_{i-1}, x_i]$ contains both types of numbers for at least one i, because P is a partition and $a < b$ so there is at least one interval of positive length with endpoints elements of P. Hence,

$$U(P, f) = \sum_{i=1}^{n} \sup_{[x_{i-1}, x_i]} f(x)\, \Delta x_i \geq 1,$$
$$L(P, f) = \sum_{i=1}^{n} \inf_{[x_{i-1}, x_i]} f(x)\, \Delta x_i = 0.$$

   Therefore, $\sup L(P, f) = 0 < 1 \leq \inf U(P, f)$ therefore $f \notin \mathscr{R}$ by definition of $\mathscr{R}$. □

5. Suppose f is a bounded real function on $[a, b]$, and $f^2 \in \mathscr{R}$ on $[a, b]$. Does it follow that $f \in \mathscr{R}$? Does the answer change if we assume $f^3 \in \mathscr{R}$?

   Solution: Defining $f(x) = -1$ for irrational x and $f(x) = 1$ for rational x shows that $f \notin \mathscr{R}$ on $[a, b]$ by Exercise 2 above, yet $f^2(x) = 1$ for all $x \in [a, b]$ hence $f^2 \in \mathscr{R}$ on $[a, b]$. Therefore, we cannot conclude that $f \in \mathscr{R}$ if $f^2 \in \mathscr{R}$.

   By Theorem 6.11 on page 127 of the text we see that if $f \in \mathscr{R}$ on $[a, b]$ and $m \leq f \leq M$, if ϕ is continuous on $[m, M]$ and $h(x) = \phi(f(x))$ on $[a, b]$ then $h \in \mathscr{R}$ on $[a, b]$. Thus, suppose that $f^3 \in \mathscr{R}$ on $[a, b]$. Since f is bounded we have that $m \leq f \leq M$ for some m and M. The function $\phi(x) = x^{1/3}$ is continuous on $[m, M]$ because it is continuous on all of $\mathbb{R}$. Therefore, $h(x) = \phi(f^3(x)) = f(x) \in \mathscr{R}$ on $[a, b]$ as desired. This fails if we only assume $f^2 \in \mathscr{R}$ precisely because $x^{1/2}$ is not defined for negative x. If $f > 0$ and bounded, then $f^2 \in \mathscr{R} \implies f \in \mathscr{R}$ because $x^{1/2}$ is continuous on $[0, M]$ for all M. □

6. Omitted.

7. Suppose that f is a real function on $(0, 1]$ and $f \in \mathscr{R}$ on $[c, 1]$ for every $c > 0$. Define

$$\int_0^1 f(x)dx = \lim_{c \to 0} \int_c^1 f(x)dx$$

   if this limit exists (and is finite).

   (a) If $f \in \mathscr{R}$ on $[0, 1]$, show that this definition of the integral agrees with the old one.

   (b) Construct a function f such that the above limit exists, although it fails to exist with $|f|$ in place of f.

   Solution:

   (a) If $f \in \mathscr{R}$ on $[0, 1]$ then f is bounded on $[0, 1]$ because the Riemann integral is defined only for bounded functions. That is, if f were not bounded then the upper-Riemann sums would be infinite, hence f would not be integrable. Thus, we can write $|f(x)| \leq M$ for all $x \in [0,1]$. Also, (the old) $\int_0^1 f(x)dx$ exists hence we can form the difference

$$\left|\int_0^1 f(x)dx - \int_c^1 f(x)dx\right| = \left|\int_0^c f(x)dx\right| \leq \int_0^c |f(x)|dx \leq cM \to 0 \text{ as } c \to 0.$$

So, the old integral is the limit of $\int_c^1 f(x)dx$ as $c \to 0$.

(b) Let

$$f(x) = n(-1)^n, \quad \frac{1}{(n+1)} < x \leq \frac{1}{n}.$$

Then this function works. □

8. Suppose that $f \in \mathscr{R}$ on $[a,b]$ for every $b > a$ where a is fixed. Define

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

if this limit exists (and is finite). In that case, we say that the integral on the left *converges*. If it also converges after f has been replaced by $|f|$, it is said to converge *absolutely*. Assume that $f(x) \geq 0$ and that f decreases monotonically on $[1,\infty)$. Prove that

$$\int_1^\infty f(x)dx$$

converges if and only if

$$\sum_{n=1}^\infty f(n)$$

converges. (This is the so-called "integral test" for convergence of series).

Solution: First, suppose that $\int_1^\infty f(x)dx$ converges. Thus, we can write

$$\int_1^\infty f(x)dx = C.$$

Since $f \geq 0$ on $[1,\infty)$ we have that $\int_1^\infty f(x)dx = \int_1^N f(x)dx + \int_N^\infty f(x)dx \geq \int_1^N f(x)$. Now, let $N \geq 1$ and let $P' = \{x_0 = 1, \ldots, x_N = N\}$ be a partition of $[1,N]$. Then,

$$\sum_{i=2}^N f(i) = L(P', f) \leq \sup_P L(P,f) = \int_1^N f(x)dx \leq \int_1^\infty f(x)dx = C, \tag{8.1}$$

where the first equality follows because f is decreasing so $\inf_{[n,n+1]} f(x) = f(n+1)$. Thus, $\sum_{i=1}^N f(i)$ is bounded (because $f(1)$ is finite). Since $f \geq 0$ on $[1,\infty)$ we see that $S_N = \sum_{i=1}^N f(i)$ is an increasing sequence, which is bounded above by Equation (8.1), hence it converges.

Conversely, suppose that $\sum_{i=1}^\infty f(i)$ converges. Then, because f is decreasing on $[1,\infty)$ we have that

$$f(x) \leq f(n) \quad \forall\, x \in [n, n+1].$$

Therefore, taking integrals preserves the inequalities, giving

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

Since $f \geq 0$ we arrive at

$$\int_1^N f(x)dx = \sum_{n=1}^{N-1} \int_n^{n+1} f(x)dx \leq \sum_{n=1}^{N-1} f(n) \leq \sum_{n=1}^{\infty} f(n) < \infty.$$

Hence, the sequence $a_N = \int_1^N f(x)dx$ is bounded above. Since $f \geq 0$ this sequence is also increasing $\big(a_{N+1} = \int_1^{N+1} f(x)dx = \int_1^N f(x)dx + \int_N^{N+1} f(x)dx \geq \int_1^N f(x)dx = a_N\big)$, hence it converges. □

9. Omitted.

10. Let p and q be real numbers such that

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{10.1}$$

Prove the following statements.

(a) If $u \geq 0$ and $v \geq 0$ then

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q}.$$

Equality holds if and only if $u^p = v^q$.

(b) If $f \in \mathscr{R}(\alpha), g \in \mathscr{R}(\alpha), f \geq 0, g \geq 0$, and

$$\int_a^b f^p d\alpha = 1 = \int_a^b g^q d\alpha,$$

then

$$\int_a^b fg \, d\alpha \leq 1.$$

(c) If f and g are complex functions in $\mathscr{R}(\alpha)$ then

$$\left|\int_a^b fg \, d\alpha\right| \leq \left(\int_a^b |f|^p d\alpha\right)^{1/p} \left(\int_a^b |g|^q d\alpha\right)^{1/q}.$$

This *Hölder's inequality*. When $p = q = 2$ it is usually called the Schwarz inequality. (Note that Theorem 1.35 is a very special case of this.)

We omit part (d).

Solution:

(a) If $u = 0$ or $v = 0$ then the inequality holds trivially, hence we can assume that both are positive. From Equation (10.1) we can write

$$\frac{q+p}{qp} = 1 \implies q + p = pq \implies p = \frac{q}{q-1}, \tag{10.2}$$

hence we have that $1 < p, q < \infty$ because both are positive. Define $a = u^p/v^q$. By Equation (10.2) above we have that $q/p = q - 1$ hence

$$a^{1/p} = \left(\frac{u^p}{v^q}\right)^{1/p} = \frac{u}{v^{q/p}} = \frac{u}{v^{q-1}} = uv^{1-q}.$$

Therefore, if we show that

$$a^{1/p} \leq \frac{1}{p}a + \frac{1}{q}, \quad 0 < a < \infty, \tag{10.3}$$

then this implies that $uv^{1-q} \leq (u^p/(v^q)(1/p) + 1/q$ which after multiplying through by v^q gives $uv \leq u^p/p + v^q/q$, our desired inequality. Hence, we need only show Equation (10.3) holds. Consider the continuous function

$$f(a) = \frac{1}{p}a + \frac{1}{q} - a^{1/p},$$

which is positive at $a = 0$ and tends to ∞ as $a \to \infty$. It is continuous because $x^{1/p}$ is continuous for $p > 1$ and $x > 0$. Therefore, if it has a minimum on $(0, \infty)$ it must occur where its derivative vanishes. Calculating we have that

$$f'(a) = \frac{1}{p} - \frac{1}{p}a^{1/p-1}.$$

Finding the zero of this, we wish to solve $1/p = (1/p)a^{1/p-1}$ which implies that $a^{1/p-1} = 1$ hence $a = 1$. Since $f(1) = 1/p + 1/q - 1 = 0$ and f is positive at $a = 0$ and tends to infinity as $a \to \infty$, we see that $f(1)$ is in fact the minimum of f for $0 < a < \infty$ hence we find that

$$0 = f(1) \leq f(a) = \frac{1}{p}a + \frac{1}{q} - a^{1/p}, \quad 0 < a < \infty,$$

which is precisely Equation (10.3), hence the desired inequality also holds.

(b) If $f \geq 0$ and $g \geq 0$ then $f^p \in \mathscr{R}(\alpha)$ and $g^q \in \mathscr{R}(\alpha)$ by Theorem 6.11 on page 127 of the text. We also have that $fg \in \mathscr{R}(\alpha)$ by Theorem 6.13(a) on page 129 of the text. By part (a) we have that

$$f(x)g(x) \leq \frac{f(x)^p}{p} + \frac{g(x)^q}{q}, \quad a \leq x \leq b, \tag{10.4}$$

hence we find that

$$\int_a^b fg\, d\alpha \leq \frac{1}{p}\int_a^b f^p d\alpha + \frac{1}{q}\int_a^b g^q d\alpha = \frac{1}{p} + \frac{1}{q} = 1.$$

(c) If f and g are complex-valued in $\mathscr{R}(\alpha)$ then $|f|$ and $|g|$ are nonnegative elements of $\mathscr{R}(\alpha)$ and as in part (b), $fg \in \mathscr{R}(\alpha)$ as well. Moreover, by Theorem 6.13(b) on page 129 of the text we have that

$$\left|\int_a^b fg\, d\alpha\right| \leq \int_a^b |f||g|\, d\alpha.$$

Now, define

$$I = \int_a^b |f|^p d\alpha, \quad J = \int_a^b |g|^q d\alpha.$$

If $I \neq 0$ and $J \neq 0$ then let $c = I^{1/p}$ and $d = J^{1/q}$. Then, we have that

$$\int_a^b \left(\frac{|f|}{c}\right)^p d\alpha = \frac{1}{c^p}\int_a^b |f|^p d\alpha = \frac{1}{c^p} I = 1,$$

and similarly for g

$$\int_a^b \left(\frac{|g|}{d}\right)^q d\alpha = \frac{1}{d^q}\int_a^b |g|^q d\alpha = \frac{1}{d^q} J = 1.$$

Hence, we can apply the results of part (b) to the two functions $|f|/c$ and $|g|/d$ to obtain

$$\int_a^b \frac{|f|}{c}\frac{|g|}{d} d\alpha \leq 1,$$

which implies that

$$\int_a^b |f||g| d\alpha \leq cd = I^{1/p}J^{1/q} = \left(\int_a^b |f|^p d\alpha\right)^{1/p}\left(\int_a^b |g|^q d\alpha\right)^{1/q}.$$

If one of I or J vanish, assume without loss of generality that I vanishes. Then, if $c > 0$ is a constant, applying Equation (10.4) to the two functions $|f|$ and $c|g|$ we obtain that

$$\int_a^b |f|(c|g|) d\alpha \leq \frac{1}{p}\int_a^b |f|^p d\alpha + \frac{1}{q}\int_a^b c^q |g|^q d\alpha = c^q \frac{1}{q}\int_a^b |g|^q d\alpha,$$

since $I = 0$ by assumption. Thus, since $q > 1$ (this comes from part (a) where we saw that $1 < p, q < \infty$),

$$\left|\int_a^b fg\, d\alpha\right| \leq \int_a^b |f||g| d\alpha \leq c^{q-1}\frac{1}{q}\int_a^b |g|^q d\alpha \to 0 \text{ as } c \to 0,$$

thus the inequality still holds. An analogous argument if $J = 0$ proves the result. Note that if both I and J vanish the result holds trivially. $\square$

11. Let α be a fixed increasing function on $[a, b]$. For $u \in \mathscr{R}(\alpha)$, define

$$||u||_2 = \left(\int_a^b |u|^2 d\alpha\right)^{1/2}.$$

Suppose that $f, g, h \in \mathscr{R}(\alpha)$. Prove the triangle inequality,

$$||f - h||_2 \leq ||f - g||_2 + ||g - h||_2$$

as a consequence of the Schwarz inequality of Exercise 10 above, as in the proof of Theorem 1.37.

Solution: Letting $p = 2 = q$ in Exercise 10 above we obtain

$$\int_a^b |uv| d\alpha \leq \left(\int_a^b |u|^2 d\alpha\right)^{1/2}\left(\int_a^b |v|^2 d\alpha\right)^{1/2}, \tag{11.1}$$

because $|uv| \geq 0$ so we don't need the absolute values on the outside of the term on the left from Exercise 10. Then, we have that

$$\begin{aligned}\int_a^b |u+v|^2 d\alpha &\leq \int_a^b |u|^2 d\alpha + 2\int_a^b |uv|\; d\alpha + \int_a^b |v|^2 d\alpha \\ &\leq \int_a^b |u|^2 d\alpha + 2\Big(\int_a^b |u|^2 d\alpha\Big)^{1/2}\Big(\int_a^b |v|^2 d\alpha\Big)^{1/2} + \int_a^b |v|^2 d\alpha \\ &= \bigg(\Big(\int_a^b |u|^2 d\alpha\Big)^{1/2} + \Big(\int_a^b |v|^2 d\alpha\Big)^{1/2}\bigg)^2,\end{aligned}$$

appling Equation (11.1) to $\int_a^b |uv|\; d\alpha$. This means that

$$\begin{aligned}||u+v||_2 &= \Big(\int_a^b |u+v|^2 d\alpha\Big)^{1/2} \\ &\leq \Big(\int_a^b |u|^2 d\alpha\Big)^{1/2} + \Big(\int_a^b |v|^2 d\alpha\Big)^{1/2} \\ &= ||u||_2 + ||v||_2.\end{aligned}$$

Letting $u = f - g$ and $v = g - h$ then gives the desired inequality. □

12. Omitted.

13. Omitted.

14. Omitted.

15. Suppose f is a real, continuously differentiable function on $[a, b]$, $f(a) = f(b) = 0$, and

$$\int_a^b f^2(x)dx = 1.$$

Prove that

$$\int_a^b xf(x)f'(x)dx = -\frac{1}{2}$$

and that

$$\int_a^b [f'(x)]^2 dx \cdot \int_a^b x^2 f^2(x)dx \geq \frac{1}{4}.$$

Solution: First, if we let $F(x) = xf^2(x)$ then $F'(x) = 2xf(x)f'(x) + f^2(x)$. Since f is continuously differentiable on $[a, b]$ so is F hence we can apply the Fundamental Theorem of Calculus, Theorem 6.21 on page 134 of the text, and obtain

$$\int_a^b [2xf(x)f'(x) + f^2(x)]dx = \int_a^b F'(x)dx = F(b) - F(a) = bf(b)^2 - af(a)^2 = 0.$$

Therefore,

$$2\int_a^b xf(x)f'(x)dx = -\int_a^b f^2(x)dx = -1 \implies \int_a^b xf(x)f'(x)dx = -\frac{1}{2}.$$

We can see the same result by using Integration by Parts, Theorem 6.22 on page 134 of the text, on the functions $F(x) = xf(x)$ and $G(x) = f(x)$ since we'd obtain

$$\int_a^b (xf(x))f'(x)dx = -\int_a^b (xf'(x) + f(x))f(x)dx = -\int_a^b xf(x)f'(x)dx - \int_a^b f(x)^2 dx.$$

Finally, by Exercise 10(c) (with $p = 2$) we have that

$$\frac{1}{4} = \left(\int_a^b xf(x)f'(x)dx\right)^2 \leq \int_a^b (xf(x))^2 dx \cdot \int_a^b [f'(x)]^2 dx.$$

□

Chapter 7: Sequences and Series of Functions

1. Prove that every uniformly convergent sequence of bounded functions is uniformly bounded.

   Solution: Suppose that $f_n \to f$ uniformly on some set E and that for each n there exists an M_n such that $|f_n(x)| \leq M_n$ for all $x \in E$. Let N be such that

   $$|f_n(x) - f_m(x)| < 1, \quad \forall\, n, m \geq N, \ \forall\, x \in E.$$

   Thus, for $n > N$ we have that

   $$|f_n(x)| \leq |f_n(x) - f_N(x)| + |f_N(x)| < 1 + M_N, \quad \forall\, x \in E.$$

   Letting $M = \max\{M_1, \dots, M_N\}$ we also have that $|f_n(x)| \leq M$ for all $1 \leq n \leq N$ and for all $x \in E$. Therefore, combining these two results we see that $|f_n(x)| < \max\{1 + M_N, M\}$ for all $n \geq 1$ and for all $x \in E$, hence $\{f_n\}$ is uniformly bounded. □

2. If $\{f_n\}$ and $\{g_n\}$ converge uniformly on a set E, prove that $\{f_n + g_n\}$ converges uniformly on E. If, in addition, $\{f_n\}$ and $\{g_n\}$ are sequences of bounded functions, prove that $\{f_n g_n\}$ converges uniformly on E.

   Solution: Let $f_n \to f$ uniformly and $g_n \to g$ uniformly. Let $\varepsilon > 0$ be given. Let N_1 be such that $|f_n(x) - f(x)| < \varepsilon/2$ for all $n \geq N_1$ and for all $x \in E$ and let N_2 be such that $|g_n(x) - g(x)| < \varepsilon/2$ for all $n \geq N_2$ and for all $x \in E$. Let $N = \max\{N_1, N_2\}$. Then, for all $x \in E$ and for all $n \geq N$ we have that

   $$|f_n(x) + g_n(x) - f(x) - g(x)| \leq |f_n(x) - f(x)| + |g_n(x) - g(x)| < \varepsilon,$$

   hence $f_n + g_n \to f + g$ uniformly on E.

   Now, suppose that for each n there exist M_n and L_n such that $|f_n(x)| \leq M_n$ for all $x \in E$ and $|g_n(x)| \leq L_n$ for all $x \in E$. By Exercise 2 above, we therefore have that $\{f_n\}$ and $\{g_n\}$ are both uniformly bounded so that there exist M and L such that $|f_n(x)| \leq M$ for all $n \geq 1$ and for all $x \in E$ and $|g_n(x)| \leq L$ for all $n \geq 1$ and for all $x \in E$. Let N be such that

   $$|f_n(x) - f(x)| < 1, \quad \forall\, n \geq N, \ \forall\, x \in E.$$

   Then, we have that

   $$|f(x)| \leq |f(x) - f_N(x)| + |f_N(x)| < 1 + M,$$

hence $f(x)$ is bounded. Now, let $\varepsilon > 0$ be given and choose K_1 such that

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2L}, \quad \forall\, n \geq K_1,\, \forall\, x \in E.$$

Similarly, choose K_2 such that

$$|g_n(x) - g(x)| < \frac{\varepsilon}{2(1+M)}, \quad \forall\, n \geq K_2,\, \forall x \in E.$$

Let $K = \max\{K_1, K_2\}$. Then, for all $n \geq K$ and for all $x \in E$ we have that

$$\begin{aligned} |f_n(x)g_n(x) - f(x)g(x)| &\leq |f_n(x)g_n(x) - f(x)g_n(x)| + |f(x)g_n(x) - f(x)g(x)| \\ &= |g_n(x)||f_n(x) - f(x)| + |f(x)||g_n(x) - g(x)| \\ &< L \cdot \frac{\varepsilon}{2L} + (1+M) \cdot \frac{\varepsilon}{2(1+M)} \\ &= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon, \end{aligned}$$

hence $f_n g_n \to fg$ uniformly on E. Notice, we needed that both $\{f_n\}$ and $\{g_n\}$ were bounded because we needed that $f(x)$ was bounded and that $g_n(x)$ was bounded. □

3. Omitted.

4. Omitted.

5. Omitted.

6. Prove that the series

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

converges uniformly in every bounded interval, but does not converge absolutely for any value of x.

Solution: We can write the partial sums of this series as a sum of two series:

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

Let $x \in [a, b]$. Then $x^2 \leq M$ for some M and so

$$\left| (-1)^n \frac{x^2}{n^2} \right| \leq \frac{M}{n^2}, \quad \forall\, x \in [a, b].$$

Since $\sum_{n=1}^{\infty} M/n^2$ converges, we see by Theorem 7.10 on page 148 of the text that $\sum_{n=1}^{\infty} (-1)^n x^2/n^2$ converges uniformly on $[a, b]$. That is,

$$\left| \sum_{n=p}^{m} (-1)^n \frac{x^2}{n^2} \right| \leq \sum_{n=p}^{m} \frac{M}{n^2},$$

and the series on the right can be made arbitrarily small for all $x \in [a, b]$ because it converges. Hence, by the triangle inequality

$$\left|\sum_{n=p}^{m}(-1)^n \frac{x^2+n}{n^2}\right| \leq \left|\sum_{n=p}^{m}(-1)^n \frac{x^2}{n^2}\right| + \left|\sum_{n=p}^{m}(-1)^n \frac{1}{n}\right|,$$

and so the term on the left can be made arbitrarily small for all $x \in [a, b]$ because each of the two terms on the right can be, since $\sum_{n=1}^{\infty}(-1)^n 1/n$ converges. It is clear the series does not converge absolutely for any value of x because if we write

$$\sum_{n=p}^{m} \frac{x^2+n}{n^2} = \sum_{n=p}^{m} \frac{x^2}{n^2} + \sum_{n=p}^{m} \frac{1}{n},$$

we can see that the second term on the right cannot be made arbitrarily small no matter how large p and m are because the series $\sum_{n=1}^{\infty} 1/n$ diverges, and this of course holds for all x because that particular series has no dependence on x. Hence also the sum on the left cannot be made arbitrarily small since the first term on the right is always positive. □

7. For $n = 1, 2, \ldots$ and x real, put

$$f_n(x) = \frac{x}{1+nx^2}.$$

Show that $\{f_n\}$ converges uniformly to a function f, and that the equation

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

is correct if $x \neq 0$, but false if $x = 0$.

Solution: First we show that $f_n \to 0$ uniformly. For each n, we have that

$$f_n'(x) = \frac{1}{1+nx^2} - \frac{2nx^2}{(1+nx^2)^2} = \frac{1-nx^2}{(1+nx^2)^2}. \tag{7.1}$$

To find the critical points we set $f_n'(x) = 0$ hence we wish to solve $1 - nx^2 = 0$ which means that $x = \pm 1/\sqrt{n}$. Observe that $f_n(-1/\sqrt{n}) < f_n(1/\sqrt{n})$ because of the numerator. Note also that for every n, $\lim_{x\to\pm\infty} f_n(x) = 0$ (because we have an x in the numerator but an x^2 in the denominator). Thus, we know that f_n achieves its maximum at $1/\sqrt{n}$. That is, $f_n(x)$ is bounded and goes to 0 at $\pm\infty$ thus its maximum is achieved and occurs at one of the calculated critical points. In particular, we have that

$$||f_n|| = \sup_{x\in\mathbb{R}} f_n(x) = f_n(1/\sqrt{n}) = \frac{1/\sqrt{n}}{1+n(1/n)} = \frac{1}{2\sqrt{n}} \to 0 \text{ as } n \to \infty, \quad \forall\, x \in \mathbb{R},$$

thus $f_n \to 0$ uniformly. Using Equation (7.1) above we have that $f_n'(0) = 1$ for all $n \geq 1$ and if $x \neq 0$ then $\lim_{n\to\infty} f_n'(x) = 0$ (because we have an n in the numerator but an n^2 in the denominator). So the equality $\lim_{n\to\infty} f_n'(x) = f'(x)$ does indeed hold for $x \neq 0$ but fails at $x = 0$ (f being the 0 function here). □

8. If

$$I(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases},$$

and if $\{x_n\}_{n=1}^{\infty}$ is a sequence of distinct points in (a, b), and if $\sum_{n=1}^{\infty} |c_n|$ converges, prove that the series

$$f(x) = \sum_{n=1}^{\infty} c_n I(x - x_n), \quad a \leq x \leq b$$

converges uniformly, and that f is continuous at every $x \neq x_n$.

Solution: Write $f_n(x) = c_n I(x - x_n)$. Then, we wish to show that $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly. We have that for all $x \in [a, b]$

$$|f_n(x)| = |c_n||I(x - x_n)| \leq |c_n|,$$

and by hypothesis $\sum_{n=1}^{\infty} |c_n|$ converges, therefore $\sum_{n=1}^{\infty} f_n(x) = f(x)$ converges uniformly by the Weierstrass M-Test, Theorem 7.10 on page 148 of the text. Now, let $g_m(x) = \sum_{n=1}^{m} c_n I(x - x_n)$. Since $x \neq x_n$ for all $n \geq 1$ let $0 < \delta < \min\{|x - x_1|, \ldots, |x - x_m|\}$. Then, we see that $B(x, \delta) \cap \{x_n\}_{n=1}^{m} = \emptyset$ and therefore if $|t - x| < \delta$ we have that

$$I(x - x_n) = I(t - x_n) \quad \forall\, n = 1, \ldots, m. \tag{8.1}$$

This is because if $x - x_n > 0$ then

$$t - x_n = t - x + x - x_n > -\delta + x - x_n > 0$$

by the definition of δ and because $x - x_n > 0$. Similarly, if $x - x_n < 0$ then

$$t - x_n = t - x + x - x_n < \delta + x - x_n < 0$$

again by the definition of δ and because $x - x_n < 0$. By the definition of I then, Equation (8.1) holds. Hence, for all $\varepsilon > 0$ we see that if $|t - x| < \delta$ then by Equation (8.1) we have that

$$|g_m(x) - g_m(t)| = \left| \sum_{n=1}^{m} c_n I(x - x_n) - \sum_{n=1}^{m} c_n I(t - x_n) \right| = 0 < \varepsilon,$$

therefore $g_m(x)$ is continuous at all x such that $x \neq x_n$ for all $m \geq 1$. But then, as we saw above, f is the uniform limit of these functions, hence is continuous at the same points. That is, choose m large enough such that $|f(x) - g_m(x)| < \varepsilon/2$ for all $x \in [a, b]$. Then, for $|x - t| < \delta$ we have that

$$\begin{aligned} |f(x) - f(t)| &= |f(x) - g_m(x) + g_m(x) - g_m(t) + g_m(t) - f(t)| \\ &\leq |f(x) - g_m(x)| + |g_m(x) - g_m(t)| + |g_m(t) - f(t)| \\ &< \frac{\varepsilon}{2} + 0 + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

□

9. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of continuous functions which converges uniformly to a function f on a set E. Prove that

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

for every sequence of points $x_n \in E$ such that $x_n \to x$, and $x \in E$. Is the converse of this true?

Solution: Let $\varepsilon > 0$ be given. Since $f_n \to f$ uniformly f is also continuous by Theorem 7.11 on page 149 of the text. Specifically, for a large enough (fixed) n we have the inequality

$$|f(t) - f(x)| \leq |f(t) - f_n(t)| + |f_n(t) - f_n(x)| + |f_n(x) - f(x)|,$$

and all three terms on the right can be made small for t close to x, since f_n is continuous, and because $f_n \to f$ uniformly. Now, let N_1 be such that $|f(x) - f_n(x)| < \varepsilon/2$ for all $n \geq N_1$ and for all $x \in E$, i.e. write $||f - f_n||_\infty = \sup_{x \in E} |f(x) - f_n(x)| < \varepsilon/2$ for all $n \geq N_1$. Now, let N_2 be such that $|f(x) - f(x_n)| < \varepsilon/2$ for all $n \geq N_2$, which is possible because f is continuous hence $f(x_n) \to f(x)$ as $n \to \infty$ (see Theorem 4.2 on page 84 of the text). Now, let $N = \max\{N_1, N_2\}$. Then, for each $n \geq N$ (i.e. fix an $n \geq N$) we have by the triangle inequality that

$$\begin{aligned}
|f(x) - f_n(x_n)| &\leq |f(x) - f(x_n)| + |f(x_n) - f_n(x_n)| \\
&\leq |f(x) - f(x_n)| + ||f - f_n||_\infty \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon,
\end{aligned}$$

where we obtain the second inequality with the sup norm because we have *fixed* an $n \geq N$, hence x_n isn't changing. But, since n was arbitrary, this means that $|f(x) - f_n(x_n)| < \varepsilon$ for all $n \geq N$, i.e. $f_n(x_n) \to f(x)$ as $n \to \infty$. Note, we needed that $x \in E$ because it was continuity of f at x that allows us to write $f(x_n) \to f(x)$ as $n \to \infty$ and f must be defined at x to be continuous there.

□

10. Omitted.

11. Omitted.

12. Omitted.

13. Omitted.

14. Omitted.

15. Suppose f is a real continuous function on $\mathbb{R}$ and $f_n(t) = f(nt)$ for $n = 1, 2, \ldots$ and suppose also that $\{f_n\}_{n=1}^{\infty}$ is equicontinuous on $[0, 1]$. What conclusion can you draw about f?

    Solution: We must have that f is constant on $[0, \infty)$. This follows from the following. Let $x, y \in [0, \infty)$ such that $x \neq y$ and suppose that $f(x) \neq f(y)$. Let $\varepsilon < |f(x) - f(y)|$ and let $\delta > 0$ be such that $|f(nt) - f(ns)| = |f_n(t) - f_n(s)| < \varepsilon$ for all $t, s \in [0, 1]$ such that $|t - s| < \delta$ and for all $n \geq 1$. There exists an N_1 such that

    $$\frac{|x - y|}{N_1} < \delta.$$

    Furthermore, since both x and y are greater than or equal to 0 we see that there exist N_2 and N_3 such that

    $$\frac{x}{N_2} \in [0, 1] \quad \text{and} \quad \frac{y}{N_3} \in [0, 1].$$

If we choose $N > \max\{N_1, N_2, N_3\}$ then all three of these conditions hold simultaneously. But then we have that

$$|f(x) - f(y)| = \left| f\left(N\frac{x}{N}\right) - f\left(N\frac{y}{N}\right)\right| = \left| f_N\left(\frac{x}{N}\right) - f_N\left(\frac{y}{N}\right)\right| < \varepsilon,$$

because $|x/N - y/N| = |x - y|/N < \delta$ and $x/N, y/N \in [0,1]$. But, this contradicts the choice of ε to be smaller than $|f(x) - f(y)|$, hence we must have that $f(x) = f(y)$. Since $x, y \in [0,\infty)$ were arbitrary this shows that f is constant on $[0,\infty)$. Note that we cannot use the equicontinuity to conclude anything about the behavior of $f(x)$ when $x < 0$ since $t \in [0,1]$ and $n \geq 1$ hence $nt \geq 0$ for all n and t where the equicontinuity is known. If, though, we had that the f_n's were equicontinuous on $[-1,1]$ then we would have found that f was constant on all of $\mathbb{R}$ by analogous arguments. □

16. Suppose $\{f_n\}_{n=1}^{\infty}$ is an equicontinuous sequence of functions on a compact set K and $\{f_n\}_{n=1}^{\infty}$ converges pointwise on K. Prove that $\{f_n\}_{n=1}^{\infty}$ converges uniformly on K.

Solution: Let $\varepsilon > 0$ be given and let $\delta > 0$ be such that $|f_n(x) - f_n(y)| < \varepsilon/3$ whenever $|x - y| < \delta$ for all $n \geq 1$ and for all $x, y \in K$, which is possible by the equicontinuity of the sequence $\{f_n\}_{n=1}^{\infty}$. Since K is compact and $\{B(x,\delta)\}_{x\in K}$ is an open cover of K there exists a finite subcover, $\{B(p_i,\delta)\}_{i=1}^{k}$. Since $f_n(x) \to f(x)$ for every $x \in K$ we have that $\{f_n(x)\}_{n=1}^{\infty}$ is a Cauchy sequence for every $x \in K$ so, specifically, if $1 \leq i \leq k$ then there exists an N_i such that

$$|f_n(p_i) - f_m(p_i)| < \frac{\varepsilon}{3}, \quad \forall\, n, m \geq N_i,$$

hence if we let $N = \max_i\{N_i\}$ then we see that

$$|f_n(p_i) - f_m(p_i)| < \frac{\varepsilon}{3}, \quad \forall\, n, m \geq N, \ \forall\, 1 \leq i \leq k. \tag{16.1}$$

Now, let $x \in K$ be arbitrary and let $n, m \geq N$. Since $\{B(p_i,\delta)\}_{i=1}^{k}$ cover K we see that $x \in B(p_j,\delta)$ for some $1 \leq j \leq k$. Thus, we have that

$$\begin{aligned} |f_n(x) - f_m(x)| &\leq |f_n(x) - f_n(p_j)| + |f_n(p_j) - f_m(p_j)| + |f_m(p_j) - f_m(x)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon. \end{aligned}$$

Here, we have that $|f_n(x) - f_n(p_j)| < \varepsilon/3$ because $|x - p_j| < \delta$ by construction of the open cover $\{B(p_i,\delta)\}_{i=1}^{k}$ hence the inequality follows from equicontinuity, and identically for the last term, $|f_m(p_j) - f_m(x)| < \varepsilon/3$ also by equicontinuity. The inequality $|f_n(p_j) - f_m(p_j)| < \varepsilon/3$ follows from Equation (16.1) and hence our choice of N *does not* depend on x because Equation (16.1) holds for all $1 \leq i \leq k$ and *every* $x \in K$ is in some $B(p_j,\delta)$ because they cover K. Hence $\{f_n(x)\}_{n=1}^{\infty}$ is uniformly Cauchy and therefore converges uniformly by Theorem 7.8 on page 147 of the text. Note, we could *not* have done the analog of applying the triangle inequality to $|f_n(x) - f(x)|$ instead (where $f(x)$ is defined as the pointwise limit of $f_n(x)$ for each x) because this would have produced

$$|f_n(x) - f(x)| \leq |f_n(x) - f_n(p_j)| + |f_n(p_j) - f(p_j)| + |f(p_j) - f(x)|.$$

Here, we can make the first term on the right small by equicontinuity and we can make the middle term small by choosing N larger than the finitely many N_i's which make each of

$|f_n(p_i) - f(p_i)|$ small, which can be done for each $1 \le i \le k$ because $f_n(p_i) \to f(p_i)$ for each such i. But, we have no control over the behavior of the last term $|f(p_j) - f(x)|$ since we do not know that f is continuous (because we don't yet know that $f_n \to f$ uniformly). □

17. Omitted.

18. Omitted.

19. Omitted.

20. If f is continuous on $[0, 1]$ and if

$$\int_0^1 f(x)x^n dx = 0, \quad n = 0, 1, 2 \ldots,$$

prove that $f(x) = 0$ on $[0, 1]$. *Hint:* The integral of the product of f with any polynomial is zero. Use the Weierstrass theorem to show that $\int_0^1 f^2(x)dx = 0$.

Solution: Let $P(x)$ be any polynomial. Then, we can write

$$P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Hence, we have, by linearity of the integral, that

$$\begin{aligned}
\int_0^1 f(x)P(x)dx &= \int_0^1 f(x)(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0)dx \\
&= a_n \int_0^1 f(x)x^n dx + \cdots + a_1 \int_0^1 f(x)x\, dx + a_0 \int_0^1 f(x)x^0 dx \\
&= 0.
\end{aligned}$$

By The Weierstrass Theorem (Theorem 7.26 on page 159 of the text) we see that there exists a sequence of polynomials P_n such that $P_n(x) \to f(x)$ uniformly on $[0, 1]$. By the above, we have that $\int_0^1 f(x)P_n(x)dx = 0$ for all $n \ge 1$. By Theorem 7.16 on page 151 of the text, because this convergence is uniform we can interchange limits and integration and obtain

$$0 = \lim_{n\to\infty} \int_0^1 f(x)P_n(x)dx = \int_0^1 \lim_{n\to\infty} f(x)P_n(x)dx = \int_0^1 f^2(x)dx.$$

But, $f^2(x) \ge 0$ for all $x \in [0, 1]$ and since f^2 is continuous (because f is and so is the function x^2) we must have that $f^2(x) = 0$ for all $x \in [0, 1]$, because otherwise it would be nonzero in some neighborhood within $[0, 1]$ by continuity, hence the integral would be positive because it would be positive over that neighborhood and greater than or equal to zero over the rest of $[0, 1]$. But, $f^2 = 0$ on $[0, 1] \implies f = 0$ on $[0, 1]$. □