Solutions to Principles of Mathematical Analysis (Walter Rudin)

Jason Rosendale
jason.rosendale@gmail.com

December 19, 2011

This work was done as an undergraduate student: if you really don't understand something in one of these proofs, it is very possible that it doesn't make sense because it's wrong. Any questions or corrections can be directed to jason.rosendale@gmail.com.

Exercise 1.1a

Let r be a nonzero rational number. We're asked to show that $x \notin \mathbb{Q}$ implies that $(r + x) \notin \mathbb{Q}$. Proof of the contrapositive:

$\mapsto r + x$ is rational $\qquad$ assumed

$\rightarrow (\exists p \in \mathbb{Q})(r + x = p) \qquad$ definition of rational

$\rightarrow (\exists p, q \in \mathbb{Q})(q + x = p) \qquad$ we're told that r is rational

$\rightarrow (\exists p, q \in \mathbb{Q})(x = -q + p) \qquad$ existence of additive inverses in $\mathbb{Q}$

Because p and q are members of the closed additive group of $\mathbb{Q}$, we know that their sum is a member of $\mathbb{Q}$.

$\rightarrow (\exists u \in \mathbb{Q})(x = u)$

$\rightarrow x$ is rational $\qquad$ definition of rational

By assuming that $r + x$ is rational, we prove that x must be rational. By contrapositive, then, if x is irrational then $r + x$ is irrational, which is what we were asked to prove.

Exercise 1.1b

Let r be a nonzero rational number. We're asked to show that $x \notin \mathbb{Q}$ implies that $rx \notin \mathbb{Q}$. Proof of the contrapositive:

$\mapsto rx$ is rational $\qquad$ assumed

$\rightarrow (\exists p \in \mathbb{Q})(rx = p) \qquad$ definition of rational

$\rightarrow (\exists p \in \mathbb{Q})(x = r^{-1}p) \qquad$ existence of multiplicative inverses in $\mathbb{Q}$

(Note that we can assume that r^{-1} exists only because we are told that r is nonzero.) Because r^{-1} and p are members of the closed multiplicative group of $\mathbb{Q}$, we know that their product is also a member of $\mathbb{Q}$.

$\rightarrow (\exists u \in \mathbb{Q})(x = u)$

$\rightarrow x$ is rational $\qquad$ definition of rational

By assuming that rx is rational, we prove that x must be rational. By contrapositive, then, if x is irrational then rx is irrational, which is what we were asked to prove.

Exercise 1.2

Proof by contradiction. If $\sqrt{12}$ were rational, then we could write it as a reduced-form fraction in the form of p/q where p and q are nonzero integers with no common divisors.

$\mapsto \frac{p}{q} = \sqrt{12}$ assumed

$\rightarrow (\frac{p^2}{q^2} = 12)$

$\rightarrow (p^2 = 12q^2)$

It's clear that $3|12q^2$, which means that $3|p^2$. By some theorem I can't remember (possibly the definition of 'prime' itself), if a is a prime number and $a|mn$, then $a|m \vee a|n$. Therefore, since $3|pp$ and 3 is prime,

$\rightarrow 3|p$

$\rightarrow 9|p^2$

$\rightarrow (\exists m \in \mathbb{N})(p^2 = 9m)$ definition of divisibility

$\rightarrow (\exists m \in \mathbb{N})(12q^2 = 9m)$ substitution from $p^2 = 12q^2$

$\rightarrow (\exists m \in \mathbb{N})(4q^2 = 3m)$ divide both sides by 3

$\rightarrow (3|4q^2)$ definition of divisibility

From the same property of prime divisors that we used previously, we know that $3|4 \vee 3|q^2$: it clearly doesn't divide 4, so it must be the case that $3|q^2$. But if $3|qq$, then $3|q \vee 3|q$. Therefore:

$\rightarrow (3|q)$

And this establishes a contradiction. We began by assuming that p and q had no common divisors, but we have shown that $3|p$ and $3|q$. So our assumption must be wrong: there is no reduced-form rational number such that $\frac{p}{q} = \sqrt{12}$.

Exercise 1.3 a

If $x \neq 0$ and $xy = xz$, then

$$y = 1y = (x^{-1}x)y = x^{-1}(xy) = x^{-1}(xz) = (x^{-1}x)z = 1z = z$$

Exercise 1.3 b

If $x \neq 0$ and $xy = x$, then

$$y = 1y = (x^{-1}x)y = x^{-1}(xy) = x^{-1}x = 1$$

Exercise 1.3 c

If $x \neq 0$ and $xy = 1$, then

$$y = 1y = (x^{-1}x)y = x^{-1}(xy) = x^{-1}1 = x^{-1} = 1/x$$

Exercise 1.3 d

If $x \neq 0$, then the fact that $x^{-1}x = 1$ means that x is the inverse of x^{-1}: that is, $x = (x^{-1})^{-1} = 1/(1/x)$.

Exercise 1.4

We are told that E is nonempty, so there exists some $e \in E$. By the definition of lower bound, $(\forall x \in E)(\alpha \leq x)$: so $\alpha \leq e$. By the definition of upper bound, $(\forall x \in E)(x \leq \beta)$: so $e \leq \beta$. Together, these two inequalities tell us that $\alpha \leq e \leq \beta$. We're told that S is ordered, so by the transitivity of order relations this implies $\alpha \leq \beta$.

Exercise 1.5

We're told that A is bounded below. The field of real numbers has the greatest lower bound property, so we're guaranteed to have a greatest lower bound for A. Let β be this greatest lower bound. To prove that $-\beta$ is the least upper bound of $-A$, we must first show that it's an upper bound. Let $-x$ be an arbitrary element in $-A$:

$$\mapsto -x \in -A \quad \text{assumed}$$
$$\to x \in A \quad \text{definition of membership in } -A$$
$$\to \beta \leq x \quad \beta = \inf(A)$$
$$\to -\beta \geq -x \quad \text{consequence of 1.18(a)}$$

We began with an arbitrary choice of $-x$, so this proves that $(\forall -x \in -A)(-\beta \geq -x)$, which by definition tells us that $-\beta$ is an upper bound for $-A$. To show that $-\beta$ is the least such upper bound for $-A$, we choose some arbitrary element less than $-\beta$:

$$\mapsto \alpha < -\beta \quad \text{assumed}$$
$$\to -\alpha > \beta \quad \text{consequence of 1.18(a)}$$

Remember that β is the greatest lower bound of A. If $-\alpha$ is larger than $\inf(A)$, there must be some element of A that is smaller than $-\alpha$.

$$\to (\exists x \in A)(x < -\alpha) \quad \text{(see above)}$$
$$\to (\exists -x \in -A)(-x > \alpha) \quad \text{consequence of 1.18(a)}$$
$$\to !(\forall -x \in -A)(-x \leq \alpha)$$
$$\to \alpha \text{ is not an upper bound of } -A \quad \text{definition of upper bound}$$

This proves that any element less than $-\beta$ is not an upper bound of $-A$. Together with the earlier proof that $-\beta$ is an upper bound of $-A$, this proves that $-\beta$ is the least upper bound of $-A$.

Exercise 1.6a

The difficult part of this proof is deciding which, if any, of the familiar properties of exponents are considered axioms and which properties we need to prove. It seems impossible to make any progress on this proof unless we can assume that $(b^m)^n = b^{mn}$. On the other hand, it seems clear that we can't simply assume that $(b^m)^{\frac{1}{n}} = b^{m/n}$: this would make the proof trivial (and is essentially assuming what we're trying to prove).

As I understand this problem, we have defined x^n in such a way that it is trivial to prove that $(x^a)^b = x^{ab}$ when a and b are integers. And we've declared in theorem 1.21 that, *by definition*, the symbol $x^{\frac{1}{n}}$ is the element such that $(x^n)^{\frac{1}{n}} = x$. But we haven't defined exactly what it might mean to combine an integer power like n and some arbitrary inverse like $1/r$. We are asked to prove that these two elements do, in fact, combine in the way we would expect them to: $(x^n)^{1/r} = x^{n/r}$.

Unless otherwise noted, every step of the following proof is justified by theorem 1.21.

$\mapsto b^m = b^m$ assumed

$\to ((b^m)^{\frac{1}{n}})^n = b^m$ definition of $x^{\frac{1}{n}}$

$\to ((b^m)^{\frac{1}{n}})^{nq} = b^{mq}$

We were told that $\frac{m}{n} = \frac{p}{q}$ which, by the definition of the equality of rational numbers, means that $mq = np$. Therefore:

$\to ((b^m)^{\frac{1}{n}})^{nq} = b^{np}$

$\to ((b^m)^{\frac{1}{n}})^{qn} = b^{pn}$ commutativity of multiplication

From theorem 12.1, we can take the n root of each side to get:

$\to ((b^m)^{\frac{1}{n}})^q = b^p$

From theorem 12.1, we can take the q root of each side to get:

$\to (b^m)^{\frac{1}{n}} = (b^p)^{\frac{1}{q}}$

Exercise 1.6b

As in the last proof, we assume that $b^{r+s} = b^r b^s$ when r and s are integers and try to prove that the operation works in a similar way when r and s are rational numbers. Let $r = \frac{m}{n}$ and let $s = \frac{p}{q}$ where $m, n, p, q \in \mathbb{Z}$ and $n, q \neq 0$.

$\mapsto b^{r+s} = b^{\frac{m}{n} + \frac{p}{q}}$

$\to b^{r+s} = b^{\frac{mq+pn}{nq}}$ definition of addition for rationals

$\to b^{r+s} = (b^{mq+pn})^{\frac{1}{nq}}$ from part a

$\to b^{r+s} = (b^{mq} b^{pn})^{\frac{1}{nq}}$ legal because mq and pn are integers

$\to b^{r+s} = (b^{mq})^{\frac{1}{nq}} (b^{pn})^{\frac{1}{nq}}$ corollary of 1.21

$\to b^{r+s} = (b^{\frac{mq}{nq}})(b^{\frac{pn}{nq}})$ from part a

$\to b^{r+s} = (b^{\frac{m}{n}})(b^{\frac{p}{q}})$

$\to b^{r+s} = (b^r)(b^s)$

Exercise 1.6c

We're given that $b > 1$. Let r be a rational number. Proof by contradiction that b^r is an upper bound of $B(r)$:

$\mapsto b^r$ is not an upper bound of $B(r)$ hypothesis of contradiction

$\to (\exists x \in B(r))(x > b^r)$ formalization of the hypothesis

By the definition of membership in $B(r)$, $x = b^t$ where t is rational and $t \leq r$.

$\to (\exists t \in \mathbb{Q})(b^t > b^r \wedge t \leq r)$

It can be shown that $b^{-t} > 0$ (see theorem S1, below) so we can multiply this term against both sides of the inequality.

$\to (\exists t \in \mathbb{Q})(b^t b^{-t} > b^r b^{-t} \wedge t \leq r)$ theorem S2

$\to (\exists t \in \mathbb{Q})(b^{t-t} > b^{r-t} \wedge t \leq r)$ from part b

$\to (\exists t \in \mathbb{Q})(1 > b^{r-t} \wedge r - t \geq 0)$

$\to (\exists t \in \mathbb{Q})(1^{-(r-t)} > b \wedge r - t \geq 0)$

$\to 1 > b$

And this establishes our contradiction, since we were given that $b > 1$. Our initial assumption must have been incorrect: b^r is, in fact, an upper bound of $B(r)$. We must still prove that it is the least upper bound of $B(r)$, though. To do so, let α represent an arbitrary rational number such that $b^\alpha < b^r$. From this, we need to

prove that $\alpha < r$.

$$\mapsto b^\alpha < b^r \qquad \text{hypothesis of contradiction}$$
$$\to b^\alpha b^{-r} < b^r b^{-r} \qquad \text{theorem S2}$$
$$\to b^{\alpha-r} < b^{r-r} \qquad \text{from part b}$$
$$\to b^{\alpha-r} < 1 \qquad \text{from part b}$$

Exercise 1.7 a

Proof by induction. Let $S = \{n : b^n - 1 \geq n(b-1)\}$. We can easily verify that $1 \in S$. Now, assume that $k \in S$:

$$\mapsto k \in S \qquad \text{hypothesis of induction}$$
$$\to b^k - 1 \geq k(b-1) \qquad \text{definition of membership in } S$$
$$\to bb^k - 1 \geq k(b-1) \qquad \text{we're told that } b > 1.$$
$$\to b^{k+1} - b \geq k(b-1)$$
$$\to b^{k+1} \geq k(b-1) + b$$
$$\to b^{k+1} - 1 \geq k(b-1) + b - 1$$
$$\to b^{k+1} - 1 \geq (k+1)(b-1)$$
$$\to k+1 \in S \qquad \text{definition of membership in } S$$

By induction, this proves that $(\forall n \in \mathbb{N})(b^n - 1 \geq n(b-1))$.

Alternatively, we could prove this using the same identity that Rudin used in the proof of 1.21. From the distributive property we can verify that $b^n - a^n = (b-a)(b^{n-1}a^0 + b^{n-2}a^1 + \ldots + b^0 a^{n-1})$. So when $a = 1$, this becomes $b^n - 1 = (b-1)(b^{n-1} + b^{n-2} + \ldots + b^0)$. And since $b > 1$, each term in the b^{n-k} series is greater than 1, so $b^n - 1 \geq (b-1)(1^{n-1} + 1^{n-2} + \ldots + 1^0) = (b-1)n$.

Exercise 1.7 b

$$\mapsto n(b^{\frac{1}{n}} - 1) = n(b^{\frac{1}{n}} - 1)$$
$$\to n(b^{\frac{1}{n}} - 1) = \underbrace{(1 + 1 + \ldots + 1)}_{n \text{ times}}(b^{\frac{1}{n}} - 1)$$
$$\to n(b^{\frac{1}{n}} - 1) = (1^{n-1} + 1^{n-2} + \ldots + 1^0)(b^{\frac{1}{n}} - 1)$$

It can be shown that $b^k > 1$ when $b > 1, k > 0$ (see theorem S4). Replacing 1 with $b^{\frac{1}{n}}$ gives us the inequality:

$$\to n(b^{\frac{1}{n}} - 1) \leq ((b^{\frac{1}{n}})^{n-1} + (b^{\frac{1}{n}})^{n-2} + \ldots + (b^{\frac{1}{n}})^0)(b^{\frac{1}{n}} - 1)$$

Now we can use the algebraic identity $b^n - a^n = (b^{n-1}a^0 + b^{n-2}a^1 + \ldots + b^0 a^{n-1})(b - a)$:

$$\to n(b^{\frac{1}{n}} - 1) \leq ((b^{\frac{1}{n}})^n - 1)$$
$$\to n(b^{\frac{1}{n}} - 1) \leq (b - 1)$$

Exercise 1.7 c

$$\mapsto n > (b-1)/(t-1) \qquad \text{assumed}$$
$$\to n(t-1) > (b-1) \qquad \text{this holds because } n, t, \text{ and } b \text{ are greater than 1}$$
$$\to n(t-1) > (b-1) \geq n(b^{\frac{1}{n}} - 1) \qquad \text{from part b}$$
$$\to n(t-1) > n(b^{\frac{1}{n}} - 1) \qquad \text{transitivity of order relations}$$
$$\to (t-1) > (b^{\frac{1}{n}} - 1) \qquad n > 0 \to n^{-1} > 0 \text{ would be a trivial proof}$$
$$\to t > b^{\frac{1}{n}}$$

Exercise 1.7 d

We're told that $b^w < y$, which means that $1 < yb^{-w}$. Using the substitution $yb^{-w} = t$ with part (c), we're lead directly to the conclusion that we can select n such that $yb^{-w} > b^{\frac{1}{n}}$. From this we get $y > b^{w+\frac{1}{n}}$, which is what we were asked to prove. As a corollary, the fact that $b^{\frac{1}{n}} > 1$ means that $b^{w+\frac{1}{n}} > b^w$.

Exercise 1.7 e

We're told that $b^w > y$, which means that $b^w y^{-1} > 1$. Using the substitution $b^w y^{-1} = t$ with part (c), we're lead directly to the conclusion that we can select n such that $b^w y^{-1} > b^{\frac{1}{n}}$. Multiplying both sides by y gives us $b^w > b^{\frac{1}{n}} y$. Multiplying this by $b^{\frac{-1}{n}}$ gives us $b^{w-\frac{1}{n}} > y$, which is what we were asked to prove. As a corollary, the fact that $b^{\frac{1}{n}} > 1 > 0$ means that, upon taking the reciprocals, we have $b^{\frac{-1}{n}} < 1$ and therefore $b^{w-\frac{1}{n}} < b^w$.

Exercise 1.7 f

We'll prove that $b^x = y$ by showing that the assumptions $b^x > y$ and $b^x < y$ lead to contradictions.

If $b^x > y$, then from part (e) we can choose n such that $b^x > b^{x-\frac{1}{n}} > y$. From this we see that $x - \frac{1}{n}$ is an upper bound of A that is smaller than x. This is a contradiction, since we've assumed that $x = \sup(A)$.

If $b^x < y$, then from part (d) we can choose n such that $y > b^{x+\frac{1}{n}} > b^x$. From this we see that x is not an upper bound of A. This is a contradiction, since we've assumed that x is the least upper bound of A.

Having ruled out these two possibilities, the trichotomy property of ordered fields forces us to conclude that $b^x = y$.

Exercise 1.7 g

Assume that there are two elements such that $b^w = y$ and $b^x = y$. Then by the transitivity of equality relations, $b^w = b^y$, although this seems suspiciously simple.

Exercise 1.8

In any ordered set, all elements of the set must be comparable (the trichotomy rule, definition 1.5). We will show by contradiction that $(0,1)$ is not comparable to $(0,0)$ in any potential ordered field containing $\mathbb{C}$. First, we assume that $(0,1) > (0,0)$:

$\mapsto (0,1) > (0,0)$ hypothesis of contradiction

$\rightarrow (0,1)(0,1) > (0,0)$ definition 1.17(ii) of ordered fields

> We assumed here that $(0,0)$ can take the role of 0 in definition 1.17 of an ordered field. This is a safe assumption because the uniqueness property of the additive identity shows us immediately that $(0,0) + (a,b) = (a,b) \rightarrow (0,0) = 0$.

$\rightarrow (-1,0) > (0,0)$ definition of complex multiplication

$\rightarrow (-1,0)(0,1) > (0,0)$ definition 1.17(ii) of ordered fields, since we initially assumed $(0,1) > 0$

$\rightarrow (0,-1) > (0,0)$ definition of complex multiplication

> It might seem that we have established our contradiction as soon as we concluded that $(-1,0) > 0$ or $(0,-1) > 0$. However, we're trying to show that the complex field cannot be an ordered field under *any* ordered relation, even a bizarre one in which $-1 > -i > 0$. However, we've shown that $(0,1)$ and $(0,-1)$ are both greater than zero. Therefore:

$\rightarrow (0,-1) + (0,1) > (0,0) + (0,1)$ definition 1.17(i) of ordered fields

$\rightarrow (0,0) > (0,1)$ definition of complex multiplication

This conclusion is in contradiction of trichotomy, since we initially assumed that $(0,0) < (0,1)$. Next, we assume that $(0,1) < (0,0)$:

$\mapsto (0,0) > (0,1)$	hypothesis of contradiction
$\rightarrow (0,0) + (0,-1) > (0,1) + (0,-1)$	definition 1.17(i) of ordered fields
$\rightarrow (0,-1) > (0,0)$	definition of complex addition
$\rightarrow (0,-1)(0,-1) > (0,0)$	definition 1.17(ii) of ordered fields
$\rightarrow (-1,0) > (0,0)$	definition of complex multiplication
$\rightarrow (-1,0)(0,-1) > (0,0)$	definitino 1.17(ii) of ordered fields, since we've established $(0,-1) > (0,0)$
$\rightarrow (0,1) > (0,0)$	definition of complex multiplication

Once again trichotomy has been violated.

Proof by contradiction that $(0,1) \neq (0,0)$: if we assume that $(0,1) = (0,0)$ we're led to the conclusion that $(a,b) = (0,0)$ for every complex number, since $(a,b) = a(0,1)^4 + b(0,1) = a(0,0) + b(0,0) = (0,0)$. By the transitivity of equivalence relations, this would mean that every element is equal to every other. And this is in contradiction of definition 1.12 of a field, where we're told that there are at least two distinct elements: the additive identity ('0') and the multiplicative identity ('1').

Exercise 1.9a

To prove that this relation turns $\mathbb{C}$ into an ordered set, we need to show that it satisfies the two requirements in definition 1.5. Proof of transitivity:

$\mapsto (a,b) < (c,d) \wedge (c,d) < (e,f)$	assumption
$\rightarrow [a < c \vee (a = c \wedge b < d)] \wedge [c < e \vee (c = e \wedge d < f)]$	definition of this order relation
$\rightarrow (a < c \wedge c < e) \vee (a < c \wedge c = e \wedge d < f)$	
$\quad \vee (a = c \wedge b < d \wedge c < e) \vee (a = c \wedge b < d \wedge c = e \wedge d < f)$	distributivity of logical operators
$\rightarrow (a < e) \vee (a < e \wedge d < f) \vee (a < e \wedge b < d) \vee (a = e \wedge b < f)$	transitivity of order relation on $\mathbb{R}$

Although we're falling back on the the transitivity of an order relation, we are not assuming what we're trying to prove. We're trying to prove the transitivity of the dictionary order relation on $\mathbb{C}$, and this relation is defined in terms of the standard order relation on $\mathbb{R}$. This last step is using the transitivity of this standard order relation on $\mathbb{R}$ and is not assuming that transitivity holds for the dictionary order relation.

$\rightarrow (a < e) \vee (a < e) \vee (a < e) \vee (a = e \wedge b < f)$	$p \wedge q \rightarrow p$
$\rightarrow a < e \vee (a = e \wedge b < f)$	$p \vee p \rightarrow p$
$\rightarrow (a,b) < (e,f)$	definition of this order relation

To prove that the trichotomy property holds for the dictionary relation on $\mathbb{Q}$, we rely on the trichotomy property of the underlying standard order relation on $\mathbb{R}$. Let (a,b) and (c,d) be two elements in $\mathbb{C}$. From the standard order relation, we know that

$\mapsto (a,b) \in \mathbb{C} \wedge (c,d) \in \mathbb{C}$	assumed
$\rightarrow a,b,c,d \in \mathbb{R}$	definition of a complex number
$\rightarrow (a < c) \vee (a > c) \vee (a = c)$	trichotomy of the order relation on $\mathbb{R}$
$\rightarrow (a < c) \vee (a > c) \vee (a = c \wedge (b < d \vee b > d \vee b = d))$	trichotomy of the order relation on $\mathbb{R}$
$\rightarrow (a < c) \vee (a > c) \vee (a = c \wedge b < d) \vee (a = c \wedge b > d)$	
$\quad \vee (a = c \wedge b = d)$	distributivity of the logical operators
$\rightarrow (a,b) < (c,d) \vee (a,b) > (c,d) \vee (a,b) < (c,d) \vee (a,b) > (c,d)$	
$\quad \vee (a,b) = (c,d)$	definition of the dictionary order relation
$\rightarrow (a,b) < (c,d) \vee (a,b) > (c,d) \vee (a,b) = (c,d)$	

And this is the definition of the trichotomy law, so we have proven that the dictionary order turns the

complex numbers into an ordered set.

Exercise 1.9b

$\mathbb{C}$ does not have the least upper bound property under the dictionary order. Let $E = \{(0, a) : a \in \mathbb{R}\}$. This subset is just the imaginary axis in the complex plane. This subset clearly has an upper bound, since $(x, 0) > (0, a)$ for any $x > 0$. But it does not have a least upper bound: for any proposed upper bound (x, y) with $x > 0$, we see that

$$(x, y) < (\frac{x}{2}, y) < (0, a)$$

So that $(\frac{x}{2}, y)$ is an upper bound less than our proposed least upper bound, which is a contradiction.

Exercise 1.10

This is just straightforward algebra, and is too tedious to write out.

Exercise 1.11

If we choose $w = \frac{z}{|z|}$ and choose $r = |z|$, then we can easily verify that $|w| = 1$ and that $rw = |z|\frac{z}{|z|} = z$.

Exercise 1.12

Set $a_i \sqrt{z_i}$ and $b_i = \sqrt{\bar{z}_i}$ and use the Cauchy-Schwarz inequality (theorem 1.35). This gives us

$$\left| \sum_{j=1}^{n} \sqrt{z_j} \sqrt{\bar{z}_j} \right|^2 \leq \sum_{j=1}^{n} |\sqrt{z_j}|^2 \sum_{j=1}^{n} |\sqrt{\bar{z}_j}|^2$$

which is equivalent to

$$|z_1 + z_2 + \ldots + z_n|^2 \leq (|z_1| + |z_2| + \ldots + |z_n|)^2$$

Taking the square root of each side shows that

$$|z_1 + z_2 + \ldots + z_n| \leq |z_1| + |z_2| + \ldots + |z_n|$$

which is what we were asked to prove.

Exercise 1.13

$$
\begin{aligned}
|x &- y|^2 \\
&= (x - y)(\overline{x - y}) \\
&= (x - y)(\bar{x} - \bar{y}) && \text{Theorem 1.31(a)} \\
&= x\bar{x} - x\bar{y} - y\bar{x} + y\bar{y} \\
&= x\bar{x} - (x\bar{y} + y\bar{x}) + y\bar{y} \\
&= |x|^2 - 2Re(x\bar{y}) + |y|^2 && \text{Theorem 1.31(c), definition 1.32} \\
&\geq |x|^2 - 2|Re(x\bar{y})| + |y|^2 && x \leq |x|, \text{ so } -|x| \geq |x|. \\
&\geq |x|^2 - 2|x\bar{y}| + |y|^2 && \text{Theorem 1.33(d)} \\
&= |x|^2 - 2|x||\bar{y}| + |y|^2 && \text{Theorem 1.33(c)} \\
&= |x|^2 - 2|x||y| + |y|^2 && \text{Theorem 1.33(b)} \\
&= (|x| - |y|)(|x| - |y|) \\
&= (|x| - |y|)(|\bar{x}| - |\bar{y}|) && \text{Theorem 1.33(b)} \\
&= (|x| - |y|)(\overline{|x| - |y|}) && \text{Theorem 1.31(a)} \\
&= ||x| - |y||^2
\end{aligned}
$$

This chain of inequalities shows us that $||x| - |y||^2 \leq |x - y|^2$. Taking the square root of both sides results in the claim we wanted to prove.

Exercise 1.14

$$|1 + z|^2 + |1 - z|^2$$
$$= (1 + z)(\overline{1 + z}) + (1 - z)(\overline{1 - z})$$
$$= (1 + z)(\bar{1} + \bar{z}) + (1 - z)(\bar{1} - \bar{z}) \qquad \text{Theorem 1.31(a)}$$
$$= (1 + z)(1 + \bar{z}) + (1 - z)(1 - \bar{z}) \qquad \text{The conjugate of } 1 = 1 + 0i \text{ is just } 1 - 0i = 1.$$
$$= (1 + \bar{z} + z + z\bar{z}) + (1 - \bar{z} - z + z\bar{z})$$
$$= (2 + 2z\bar{z})$$
$$= (2 + 2) \qquad \text{We are told that } z\bar{z} = 1$$
$$= 4$$

Exercise 1.15

Using the logic and the notation from Rudin's proof of theorem 1.35, we see that equality holds in the Schwarz inequality when $AB = |C|^2$. This occurs when have $B(AB - |C|^2) = 0$, and from the given chain of equalities we see that this occurs when $\sum |Ba_j - Cb_j|^2 = 0$. For this to occur we must have $Ba_j = Cb_j$ for all j, which occurs only when $B = 0$ or when

$$a_j = \frac{C}{B} b_j \quad \text{ for all } j$$

That is, each a_j must be a constant multiple of b_j.

Exercise 1.16

We know that $|z - x|^2 = |z - y|^2 = r^2$. Expanding these terms out, we have

$$|z - x|^2 = (z - x) \cdot (z - x) = |z|^2 - 2z \cdot x + |x|^2$$

$$|z - y|^2 = (z - y) \cdot (z - y) = |z|^2 - 2z \cdot y + |y|^2$$

For these to be equal, we must have

$$-2z \cdot x + |x|^2 = -2z \cdot y + |y|^2$$

which happens when

$$z \cdot (x - y) = \frac{1}{2}[|x|^2 - |y|^2] \tag{1}$$

We also want $|z - x| = r$, which occurs when $z = x + r\hat{u}$ where $|\hat{u}| = 1$. Using this representation of z, the requirement that $r^2 = |z - y|^2$ becomes

$$r^2 = |z - y|^2 = |x + r\hat{u} - y|^2 = |(x - y) + r\hat{u}^2| = |x - y|^2 + 2r\hat{u} \cdot (x - y) + |r\hat{u}|^2 = d^2 + 2r\hat{u} \cdot (x - y) + r^2$$

Rearranging some terms, this becomes

$$\hat{u} \cdot (y - x) = \frac{d^2}{2r} = \frac{d}{2r}|y - x| \tag{2}$$

A quick, convincing, and informal proof would be to appeal to the relationship $a \cdot b = |a||b|\cos(\theta)$ where θ is the angle between the two vectors; the previous equation then becomes

$$|\hat{u}||y - x|\cos(\theta) = \frac{d}{2r}|y - x|$$

Dividing by $|y - x|$ and remembering that $\hat{u} = 1$, this becomes

$$\cos(\theta) = \frac{d}{2r}$$

where θ is the angle between the fixed vector $(y - x)$ and the variable vector $\hat{u}$. It's easy to see that this equation will hold for exactly one $\hat{u}$ when $d = 2r$; it will hold for no $\hat{u}$ when $d > 2r$; it will hold for two values of $\hat{u}$ when $d < 2r$ and $n = 2$; and it will hold for infinitely many values of $\hat{u}$ when $d < 2r$ and $n > 2$. Each value of $\hat{u}$ corresponds with a unique value of z. More formal proofs follow.

part (a)

When $d < 2r$, equation (2) is satisfied for all $\hat{u}$ for which

$$\hat{u} \cdot (y - x) = \frac{d}{2r} |y - x| < |y - x|$$

By the definition of the dot product, this is equivalent to

$$u_1(y_1 + x_1) + u_2(y_2 + x_2) + \ldots + u_k(y_k + x_k) = \frac{d}{2r}|y - x| \tag{3}$$

The only other requirement for the values of u_i is that

$$\sqrt{u_1^2 + u_2^2 + \ldots + u_k^2} = 1 \tag{4}$$

This gives us a system of two equations with k variables. As long as $k \geq 3$ we have more variables than equations and therefore the system will have infinitely many solutions.

part (b)

Evaluating d^2, we have:

$$
\begin{aligned}
d^2 &= |x - y|^2 \\
&= |(x - z) + (z - y)|^2 \\
&= [(x - z) + (z - y)] \cdot [(x - z) + (z - y)] && \text{definition of inner product } \cdot \\
&= (x - z) \cdot (x - z) + 2(x - z) \cdot (z - y) + (z - y) \cdot (z - y) && \text{inner products are distributive} \\
&= |x - z|^2 + 2(x - z) \cdot (z - y) + |z - y|^2 && \text{definition of inner product } \cdot
\end{aligned}
$$

Evaluating $(2r)^2$, we have:

$$
\begin{aligned}
(2r)^2 &= (r + r)^2 = (|z - x| + |z - y|)^2 \\
&= |z - x|^2 + 2|z - x||z - y| + |z - y|^2 && \text{commutativity of multiplication}
\end{aligned}
$$

If $2r = d$ then $d^2 = (2r)^2$ and therefore, by the above evaluations, we have

$$|x - z|^2 + 2(x - z) \cdot (z - y) + |z - y|^2 = |z - x|^2 + 2|z - x||z - y| + |z - y|^2$$

which occurs if and only if

$$2(x - z) \cdot (z - y) = 2|x - z||z - y|$$

From exercise 14 we saw that this equality held only if $(x - z) = c(z - y)$ for some constant c; we know that $|x - z| = |z - y|$ so $c = \pm 1$; we know $x \neq y$ so $c = 1$. Therefore we have $x - z = z - y$, from which we have

$$z = \frac{x + y}{2}$$

and there is clearly only one such z that satisfies this equation.

part (c)

If $2r < d$ then we have

$$|x - y| > |x - z| + |z - y|$$

which is equivalent to

$$|(x - z) + (z - y)| > |x - z| + |z - y|$$

which violates the triangle inequality (1.37e) and is therefore false for all z.

Exercise 1.17

First, we need to prove that $a \cdot (b + c) = a \cdot b + a \cdot c$ and that $(a + b) \cdot c = a \cdot c + b \cdot c$: that is, we need to prove that the distributive property holds between the inner product operation and addition.

$$
\begin{aligned}
a \cdot (b + c) &= \sum a_i(b_i + c_i) && \text{definition 1.36 of inner product} \\
&= \sum (a_i b_i + a_i c_i) && \text{distributive property of } \mathbb{R} \\
&= \sum a_i b_i + \sum a_i c_i && \text{associative property of } \mathbb{R} \\
&= a \cdot b + a \cdot c && \text{definition 1.36 of inner product}
\end{aligned}
$$

$$
\begin{aligned}
(a + b) \cdot c &= \sum (a_i + b_i)c_i && \text{definition 1.36 of inner product} \\
&= \sum (a_i c_i + b_i c_i) && \text{distributive property of } \mathbb{R} \\
&= \sum a_i c_i + \sum b_i c_i && \text{associative property of } \mathbb{R} \\
&= a \cdot c + b \cdot c && \text{definition 1.36 of inner product}
\end{aligned}
$$

The rest of the proof follows directly:

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

Exercise 1.18

If $x = 0$, then $x \cdot y = 0$ for any $y \in \mathbb{R}^k$. If $x \neq 0$, then at least one of the elements $x_1, \ldots, x_k$ must be nonzero: let this element be represented by x_a. Let x_b represent any other element of x. Choose y such that:

$$
y_i = \begin{cases} \dfrac{x_b}{x_a} & i = a \\ -1 & i = b \\ 0 & \text{otherwise} \end{cases}
$$

We can now see that $x \cdot y = x_a \frac{x_b}{x_a} + x_b(-1) = x_b - x_b = 0$. We began with an arbitrary vector x and demonstrated a method of construction for y such that $x \cdot y = 0$: therefore, we can always find a nonzero y such that $x \cdot y = 0$. This is not true in $\mathbb{R}^1$ because the nonzero elements of $\mathbb{R}$ are closed with respect to multiplication.

Exercise 1.19

We need to determine the circumstances under which $|x - a| = 2|x - b|$ and $|x - c| = r$. To do this, we need to manipulate these equalities until they have a common term that we can use to compare them.

$$
\begin{aligned}
|x - a| &= 2|x - b| \\
|x - a|^2 &= 4|x - b|^2 \\
|x|^2 - 2x \cdot a + |a|^2 &= 4|x|^2 - 8x \cdot b + 4|b|^2 \\
3|x|^2 &= |a|^2 - 2x \cdot a + 8x \cdot b - 4|b|^2
\end{aligned}
$$

$$
\begin{aligned}
|x - c| &= r \\
|x - c|^2 &= r^2 \\
|x|^2 - 2x \cdot c + |c|^2 &= r^2 \\
|x|^2 &= r^2 + 2x \cdot c - |c|^2 \\
3|x|^2 &= 3r^2 + 6x \cdot c - 3|c|^2
\end{aligned}
$$

Combining these last two equalities together, we have

$$\mapsto |a|^2 - 2x \cdot a + 8x \cdot b - 4|b|^2 = 3r^2 + 6x \cdot c - 3|c|^2$$
$$\rightarrow |a|^2 - 2x \cdot a + 8x \cdot b - 4|b|^2 - 3r^2 - 6x \cdot c + 3|c|^2 = 0$$
$$\rightarrow |a|^2 - 4|b|^2 + 3|c|^2 - 2x \cdot (a - 4b - 3c) - 3r^2 = 0$$

We can zero out the dot product in this equation by letting $3c = 4b - a$. Of course, this also determines a specific value of c. This substitution gives us the new equality:

$$\rightarrow |a|^2 - 4|b|^2 + 3\left|\frac{4b}{3} - \frac{a}{3}\right|^2 - 3r^2 = 0$$
$$\rightarrow |a|^2 - 4|b|^2 + \frac{3 \cdot 16}{9}|b|^2 - \frac{3 \cdot 8}{9}a \cdot b + \frac{3}{9}|a|^2 - 3r^2 = 0$$
$$\rightarrow 3|a|^2 - 12|b|^2 + 16|b|^2 - 8a \cdot b + |a|^2 - 9r^2 = 0$$
$$\rightarrow 4|a|^2 4|b|^2 - 8a \cdot b - 9r^2 = 0$$
$$\rightarrow 4|a - b|^2 = 9r^2$$
$$\rightarrow 2|a - b| = 3r$$

By choosing $3c = 4b - a$ and $3r = 2|a - b|$, we guarantee that $|x - a| = 2|x - b|$ iff $|x - c| = r$.

Exercise 1.20

We're trying to show that R has the least upper bound property. The elements of R are certain subsets of $\mathbb{Q}$, and $<$ is defined to be "is a proper subset of". To say that an element $\alpha \in R$ has a least upper bound, then, is to say that the subset α has some "smallest" superset β such that $\alpha \subset \beta$. We're asked to omit property III, which told us that each $\alpha \in R$ has no largest element. With this restriction lifted, we have a new definition of "cut" that includes cuts such as $(-\infty, 1]$ and $(-\infty, \sqrt{2}]$.

To prove that each subset of R with an upper bound must have a least upper bound, we will follow Step 3 in the book almost exactly. We will define two subsets $A \subset R$ and $\gamma \subset R$, show that the subset γ is also an *element* of R, and then show that the subset/element γ is the least upper bound of A.

Let A be a nonempty subset of R with an upper bound of β (Note: A is a subset of R, *not* a subset of $\mathbb{R}$. The elements of A are cuts, which are subsets of $\mathbb{Q}$). Define γ to be the union of all cuts in A. This means that γ contains every element from every cut in A: γ consists of elements from Q.

proof that γ is a cut:

The proof of criterion (I) has two parts. (i) γ is the union of elements of A, and we were told that A is nonempty. Therefore γ is nonempty. (ii) We are told that β is an upper bound for A, so $x \in A \rightarrow x \subset \beta$ (remember that $<$ is defined as $\subset$). But γ is just the union of cut elements in A, so γ is the union of proper subsets of β. This means that γ itself is a proper subset of β. This shows that $\gamma \subset \beta \subseteq \mathbb{Q}$, so $\gamma \neq \mathbb{Q}$. So criterion (I) in the definition of "cut" has been met.

To prove part (II), pick some arbitrary rational $p \in \gamma$. γ is the set of cuts in A, so $p \in \alpha$ for some cut $\alpha \in A$. Choose a second rational q such that $q < p$: by the definition of cut, the fact that $p \in \alpha$ and $q < p$ means that $q \in \alpha$ and therefore $q \in \gamma$. And we're being asked to disregard part (III), so this is sufficient to prove that γ is a cut.

proof that γ is the least upper bound of A

(i) Choose any arbitrary cut $\alpha \in A$. We've defined γ as the union of all cuts in A, so it's clear that every rational number in α is also contained in γ: that is, $\alpha \subseteq \gamma$. And by the definition of $<$ for this set, this tells us that $\alpha \leq \gamma$ for every $\alpha \in A$.

(ii) Suppose that $\delta < \gamma$. By the definition of $<$ for this set, this means that $\delta \subset \gamma$. This is a proper subset, so there must be some rational s such that $s \notin \delta$ but $s \in \gamma$. In order for $s \in \gamma$, it must be the case that $s \in \alpha$ for some cut $\alpha \in A$. We can now show that $\delta \subset \alpha$. For every $r \in \delta$, it's also true that $s \notin \delta$. By (II), this means

that $r < s$ (using standard rational order). And since $s \in \alpha, (II)$ also shows that $r \in \alpha$. This shows that $\delta \subset \alpha$, which means $\delta < \alpha \in A$ (using the subset order on R), and therefore δ is not an upper bound of A.

Together, these two facts show that γ is the least upper bound of A.

definition of addition

Following the example in the book, we define the addition of two cuts $\alpha + \beta$ to be the set of all sums $r + s$ where $r \in \alpha$ and $s \in \beta$. The book defined 0^* to be the set of rational numbers < 0, but by omitting requirement (III) we are forced to redefine our zero. Let $0^\#$ be the set of all rational numbers ≤ 0.

The original definition had to omit 0 from 0^* in order to keep R, the set of cuts, closed under addition (otherwise we'd have $0^* + 0^* = (-\infty, 0]$ which is not a cut because of criterion (III)). Our new zero $0^\#$ must include 0 as an element because our newly defined cuts can have a greatest element. The set 0^* can no longer function as our zero since $(-\infty, x] + 0^* = (-\infty, x)$; these two cuts are not equal $((-\infty, x) < (-\infty, x])$, so 0^* has not functioned as the additive identity.

field axioms A1,A2, and A3

The proofs from the book for closure, commutativity, and associativity are directly applicable to our new definition of cut.

field axiom A4: existence of an additive identity

Let α be a cut in R. Let r and s be rationals such that $r \in \alpha$ and $s \in 0^\#$. Then $r + s \leq r$, which means that $r + s \in \alpha$. This shows us that $\alpha + 0^\# \subseteq \alpha$. Now let p and r be rationals such that $p \in \alpha, r \in \alpha$, and $p \leq r$. This means that $p - r \leq 0$, so $p - r \in 0^\#$. Therefore $r + (p - r) \in \alpha + 0^\#$, which means that $p \in \alpha + 0^\#$. This shows us that $\alpha \subseteq \alpha + 0^\#$. Having shown that $\alpha + 0^\# \subseteq \alpha$ and that $\alpha \subseteq \alpha + 0^\#$, we conclude that $\alpha = \alpha + 0^\#$ for any $\alpha \in R$.

field axiom A5: existence of additive inverses

The book constructed the definition of inverse so that the inverse of $(-\infty, x)$ would be $(-\infty, -x)$. This works under the original definition of "cut", since $(-\infty, x) + (-\infty, -x) = (-\infty, 0) = 0^*$. It is tempting to just define the inverse for our new definition of cut so that the inverse of $(-\infty, x]$ is $(\infty, -x]$. This overlooks the fact that both $(-\infty, x]$ and $(-\infty, x)$ are cuts under our new definition: both of these elements must have additive inverses.

To show that A5 is not satisfied, we need to demonstrate that at least one element of R has no additive inverse. Let $\alpha = (-\infty, r)$ for some rational r. Assume that there was some cut β such that $\alpha + \beta = 0^\# = (-\infty, 0]$: we will show that this assumption leads to a contradiction.

(i) Assume that β does not contain any elements greater than $-r$. Let p and q be arbitrary rationals such that $p \in \alpha$ and $q \in \beta$. From our definitions of α and β, we know that $p < r$ and $q \leq -r$. Combining these two inequalities tells us that $p + q < -r + r$, or $p + q < 0$. But p and q were arbitrary members of α and β, so this shows that $0 \notin \alpha + \beta$. So it cannot be the case that $\alpha + \beta = 0^\#$.

(ii) Assume that β contains at least one element q that is greater than $-r$. Let s represent the difference $q - (-r)$: then $q = -r + s$ (note that s must be a positive rational). Let p be an arbitrary rational such that $p \in \alpha$. By the definition of α, we know that $p < r$. And from property (II) of cuts, we know that $r - s < p < r$. Adding $q = -r + s$ to each element in this equality gives us $0 < p + q < s$. And $p + q$ is an element of $\alpha + \beta$, so we see that $\alpha + \beta$ contains some positive rational: it cannot be the case that $\alpha + \beta = 0^\#$.

Whether or not β contains an element greater than $-r$, we find ourselves in contradiction with the initial assumption that β might be the inverse of α. Therefore there can be no inverse for α. In conclusion, we see that omitting (III) forces us to redefine the additive identity, and this new definition results in the existence of elements in R that have no inverse.

Exercise 2.1

This is immediately justified by noting that the definition of subset, $x \in \varnothing \to x \in A$, is satisfied for any set A because of the false antecedent. A more formal proof:

$\mapsto \neg(\exists x \in \varnothing)$ definition of an empty set

$\to (\forall x)(x \notin \varnothing)$ negation of the $\exists$ quantifier

$\to (\forall x)(x \notin A \to x \notin \varnothing)$ Hilbert's PL1

    This previous step is justified by the argument $p \to (q \to p)$: if something is true (like $x \notin \varnothing$), then everything implies it.

$\to (\forall x)(x \in \varnothing \to x \in A)$ contrapositive

$\to \varnothing \subseteq A$ definition of subset

Exercise 2.2

The set of integers is countable, so by theorem 2.13 the set of all $(k+1)$-tuples $(a_0, a_1, \ldots, a_k)$ with $a_0 \neq 0$ is also countable. Let this set be represented by $\mathbb{Z}^k$. For each $a \in \mathbb{Z}^k$ consider the polynomial $a_0 z^k + a_1 z^{k-1} + \ldots + a_k = 0$. From the fundamental theorem of algebra, we know that there are exactly k complex roots for this polynomial.

We now have a series of nested sets that encompass every possible root for every possible polynomial with integer coefficients. More specifically, we have a countable number of $\mathbb{Z}^k$s, each containing a countable number of $(k+1)$-tuples, each of which corresponds with k roots of a k-degree polynomial. So our set of complex roots (call it R) is a countable union of countable unions of finite sets. This only tells us that R is at most countable: it is either countable or finite.

To show that R is not finite, consider the roots for 2-tuples in $\mathbb{Z}^1$. Each 2-tuple of the form $(-1, n)$ corresponds with the polynomial $-z + n = 0$ whose solution is $z = n$. There is clearly a unique solution for each $n \in \mathbb{Z}$, so R is an infinite set. Because R is also at most countable, this proves that R is countable.

I did not use the hint provided in the text, which either means that this proof is invalid or that there is an alternate (simpler?) proof.

Exercise 2.3

The set of real numbers is uncountable, so if every real number were algebraic then the set of real algebraic numbers would be uncountable. However, exercise 2.2 showed that the algebraic complex numbers were countable. The real numbers are a subset of the complex numbers, so the set of algebraic real numbers is at most countable.

Exercise 2.4

The rational real numbers are countable. If the irrational real numbers were countable, then the union of rational and irrational reals would be countable. But this union is just $\mathbb{R}$, which is not countable. So the irrational reals are not countable.

Exercise 2.x: The Cantor Set

The idea behind the Cantor set is that each term in the series $E_1 \supset E_2 \ldots \supset E_n$ removes the middle third of each remaining line segment, so that you get a series of increasingly smaller segments that look like:

We can see that E_0 has one segment of length 1, E_1 has two segments of length $1/3$, and E_m has 2^m segments of length $1/3^m$. Notice that for any possible segment (α, β), we can choose m to be large enough so that the maximum segment length is less than the length of (α, β) and therefore $(\alpha, \beta) \notin E_m$. And if the segment is not in E_m for some m, it's not in the union P. This is the reasoning behind Rudin's statement that "P contains no segment".

To justify the claim that no point in P is an isolated point, we need to show that for every point $p_1 \in P$ and any arbitrarily small radius r, we can find some second point p_2 such that $d(p_1, p_2) < r$. P is defined to be $\bigcap E_i$, so if $p_1 \in P$ then it is a member of every E_i. Choose some element E_m where m is so large that the segments in E_m are all shorter than r. We are assured that this is possible by the Archimedean property of the rationals, since we need only find m such that $3^{-m} < r$.

So we've found an E_m with a line segment containing p_1. Let p_2 represent one of the endpoints of this line. The length of this segment is less than r, so $d(p_1, p_2) < r$. In each subsequent term in the series $\{E_i\}$, the endpoint p_2 will never be "cut off": it will always be the endpoint of a segment from which the middle third is lost. So $p_2 \in P$. And since this was true for an arbitrary point $p_1 \in P$ and an arbitrarily small r, this shows that every neighborhood of every point in P is a limit point and not an isolated point.

Exercise 2.5

Let E be a subset of $(0, 1]$ consisting only all rational numbers of the form $\frac{1}{m}, m \in \mathbb{N}$. No point in E is a limit point, but the limit points of E need not be members of E: we will demonstrate that the point 0 is a limit point of E.

(i) Proof that no point in E is a limit point: let p_m be an arbitrary member of E. Then $p_m = \frac{1}{m}$ for some $m \in \mathbb{N}$. The closest point to p_m is therefore $p_{m+1} = \frac{1}{m+1}$. So we just choose r such that $r = \frac{1}{2} d(p_m, p_{m+1})$ and we see that there are no other points in this neighborhood of p_m, which shows that p_m is an isolated point.

(ii) Proof that 0 is a limit point: choose an arbitrarily small radius r. From the Archimedian property, we can find some $m \in \mathbb{N}$ such that $\frac{1}{m} < r$. This $p_m = \frac{1}{m}$ is a member of E, and $d(0, p_m) = \frac{1}{m} < r$. This shows that 0 is a limit point.

(iii) Proof that no other points are limit points: let x be some point that is neither zero nor a member of E. This point must be either less than zero, greater than 1, or it must lie between two sequential points $p_m, p_{m+1} \in E$. Let the d_x represent the smallest distance from the set $\{d(x, 0), d(x, 1), d(x, p_m), d(x, p_{m+1})\}$. Choose r such that $r = \frac{1}{2} d_x$. There can be no points of E in the neighborhood $N_r(x)$. If there were, then this would indicate the existence of a point of E less than zero, greater than 1, or between $\frac{1}{m}$ and $\frac{1}{m+1}$. None of these can exist in E as we defined it, so no points other than 0 are limit points.

Now let F be the subset of $(2, 3]$ containing rational numbers of the form $2 + \frac{1}{m}$ and let G be the subset of $(4, 5]$ consisting of rational numbers of the form $4 + \frac{1}{m}$. Just as E was shown to have a limit point of only zero, these sets can be shown to have limit points of only 2 and 4. Therefore the union $E \cup F \cup G$ is bounded and has three limit points $\{0, 2, 4\}$.

Exercise 2.6

(i) We're asked to prove that E' is closed. This is equivalent to proving that every limit point of E' is a point of E. And by the definition of E', this is equivalent to proving that every limit point of E' is a limit point of E.

To prove this, let x represent an arbitrary limit point of E'. Choose any arbitrarily small r and let $s = \frac{r}{2}, t = \frac{r}{4}$[1]. Since x is a limit point of E', we can find a point $y \in E'$ in the neighborhood $N_s(x)$. And y, by virtue of being in E', is a limit point of E: so we can find a point $z \in E$ in the neighborhood $N_t(y)$.

The distance $d(x, y)$ is less than s and $d(y, z)$ is less than t, so definition 2.15 of metric spaces assures us that $d(x, z) \leq d(x, y) + d(y, z) < s + t < r$, so $d(x, z) < r$ and therefore z is in the neighborhood $N_r(x)$ for any

---

[1] t was chosen to be less than r to guarantee $x \neq z$.

arbitrarily small r. So the point x is a limit point for E. But x was an arbitrarily chosen limit point of E', so we have proven that every limit point of E' is a limit point of E, which is what we were asked to prove. The following image is helpful for imagining these points in $\mathbb{R}^2$.
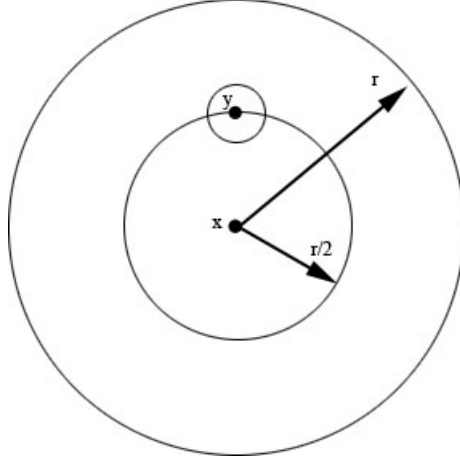
(ii) We can use a similar technique to show that every limit point of $\bar{E}$ is a limit point E. Let x represent an arbitrary limit point of $\bar{E}$ (which we won't assume to be a member of $\bar{E}$). Choose any arbitrarily small r and let $s = \frac{r}{2}, t = \frac{r}{4}$. Since x is a limit point of $\bar{E}$, theorem 2.20 tells us that the neighborhood $N_s(x)$ contains infinitely many points of $\bar{E}$. Because $\bar{E}$ is defined to be $E' \cup E$, each of these infinitely many points in $N_s(x)$ is either a member of E' or a member of E.

(ii.a) First, assume that there exists at least one point y in $N_s(x)$ such that $y \in E'$. By definition, this means that y is a limit point of E, so there is at least one point $z \in E$ in the neighborhood $N_t(y)$. The distance $d(x, y)$ is less than s and $d(y, z)$ is less than t, so definition 2.15 of metric spaces assures us that $d(x, z) \leq d(x, y) + d(y, z) < s + t < r$, so $d(x, z) < r$ and therefore we can find some $z \in E$ in the neighborhood $N_r(x)$ for any arbitrarily small r. So x is a limit point for E.

(ii.b) Next, assume that none of the points in $N_s(x)$ are members of E. This means that all of the infinitely many points of $E' \cup E$ in $N_s(x)$ are members of E. From this, we see that every neighborhood of x contains elements of E. For any neighborhood $N_t(x)$ with $t < s$, the fact that x is a limit point of $\bar{E}$ and that $N_t(x) \subset N_s(x)$ means that $N_t(x)$ contains infinitely many points in E. But this means that x is a limit point of E. For any neighborhood $N_t(x)$ with $t > s$, the fact that $N_s(x) \subset N_t(x)$ means that $N_t(x)$ contains infinitely many points in E. And since every neighborhood of x contains an element of E, x is a limit point for E.

The second part of the proof is to show that every limit point of E is a limit point of $\bar{E}$ and is relatively trivial. Every element of E is also an element of $E \cup E' = \bar{E}$. If x is a limit point of E, then every neighborhood of x contains an element of E. Therefore every neighborhood of x contains an element of $\bar{E}$. So x is a limit point of $\bar{E}$.

We've shown that every limit point of $\bar{E}$ is a limit point of E and vice-versa, which is what we were asked to prove.

(iii) We're asked if E and E' always have the same limit points. The answer is "no", and a counterexample can be found in the previous question. The limit points of E in exercise 2.5 were $E' = \{0, 2, 4\}$. And E' clearly has no limit points whatsoever.

Exercise 2.7a

We'll prove the equality of these sets by showing that each is a subset of the other.

(A) Assume that $x \in \bar{B}_n$. Then, $x \in B_n$ or x is a limit point of B_n.

(A1) If $x \in B_n$, then $x \in A_k$ for some $A_k \in \bigcup A_i$. And if $x \in A_k$, then $x \in \bar{A}_k$. And this means that $x \in \bigcup \bar{A}_i$.

(A2) If x is a limit point of B_n, then it must be a limit point at least one specific $A_k \in \bigcup A_i$. Proof by contradiction: assume that x is not a limit point for any $A_k \in \bigcup A_i$. Then for each i, there is some neighborhood $N_{r_i}(x)$ that contains no elements of A_i. Let $s = \min(r_i)$ (which exists because i is finite). Then $N_s(x)$ contains

no elements from any A_i since $N_s(x) \subseteq N_{r_i}(x)$ for each i. And if this neighborhood contains no points for any A_i, then it contains no points of $B_n = \bigcup A_i$. And this is a contradiction, since x is a limit point of B_n. Therefore, by contradiction, x must be a limit point of at least one specific $A_k \in \bigcup A_i$. So $x \in \bar{A}_k$, which means that $x \in \bigcup \bar{A}_i$.

In either of these two cases, $x \in \bigcup \bar{A}_i$. This proves that $x \in \bar{B}_n \to x \in \bigcup \bar{A}_i$.

(B) Assume that $x \in \bigcup \bar{A}_i$. Then $x \in \bar{A}_k$ for some k. And this means that either $x \in A_k$ or x is a limit point of A_k.

(B1) If $x \in A_k$, then $x \in \bigcup A_i$, which means that $x \in B_n$ and therefore $x \in \bar{B}_N$.

(B2) If x is a limit point of A_k, then every neighborhood of x contain an element of A_k. But every element of A_k is an element of $\bigcup A_i = B_n$, so this means that every neighborhood of x contains an element of B_n. By definition, this means that x is a limit point of B_n, so that $x \in \bar{B}_n$.

In either of these two cases, $x \in \bar{B}_n$. This proves that $x \in \bigcup \bar{A}_i \to x \in \bar{B}_n$. And we've already shown that $x \in \bar{B}_n \to x \in \bigcup \bar{A}_i$, so we have proven that $\bar{B}_n = \bigcup \bar{A}_i$.

Exercise 2.7b

Nothing in part B of the above proof required that i be finite, so it's still the case that $\bigcup_{i=1}^{\infty} \bar{A}_i \subset \bar{B}$. But part A of the proof assumed that we could choose a least element from the set of size i, which we can't do for an infinite set. So we can't assume that $B \subseteq \bigcup \bar{A}_i$. And it's a good thing that we can't assume this, because it's false.

Consider the set from exercise 2.5, the set E consisting of all rational numbers of the form $\frac{1}{m}, m \in \mathbb{N}$. We'll construct this set by defining $A_i = \frac{1}{i}$ and defining $B = \bigcup_{i=1}^{\infty} A_i$. We saw in exercise 2.5 that $0 \in \bar{B}$. But $0 \notin \bar{A}_i$ for any i, so $0 \notin \bigcup_{i=1}^{\infty} \bar{A}_i$. This shows us that $\bar{B} \neq \bigcup_{i=1}^{\infty} \bar{A}_i$ for these sets. And we've already shown that $\bigcup \bar{A}_i \subseteq B$, so for this particular choice of sets we see that $\bigcup \bar{A}_i \subset B$.

Exercise 2.8a

Every point in an open set E in $\mathbb{R}^2$ is a limit point of E. Proof: let p be an arbitrary point in E. We can say two things about the neighborhoods of p. First, since we're dealing with $\mathbb{R}^2$, every neighborhood of p contains infinitely many points. Second, p is an interior point (since E is open) and so there is some r such that $N_r(p) \subset E$. Together, these two facts tell us that $N_r(p)$ contains infinitely many points of E.

We now show that this guarantees that every neighborhood of p contains a point of E. Choose any s such that $r < s$. Because the neighborhood $N_r(p)$ contains some $q \in E$ and $N_r(p) \subset N_s(p)$, we know that $q \in N_s(p)$. Now choose s such that $s < r$. $N_s(p)$ contains infinitely many points, and $N_s(p) \subset N_r(p) \subset E$, so $N_s(p)$ contains infinitely many points of E. We've shown that whether $r < s$ or $r > s$ for any arbitrary r, $N_s(p)$ contains a point of E. Thus p is a limit point of E.

Exercise 2.8b

Consider a closed set consisting of a single point, such as $E = \{(0,0)\}$. This point is clearly not a limit point of E.

Exercise 2.9a

To prove that E° is open, we will show that every point in E° must be an interior point of E°.

Let x be an arbitrary point in E°. By the definition of E°, we know that x is an interior point of E. This means that we can choose some r such that $N_r(x) \subset E$: i.e., every point of $N_r(x)$ is a point of E. To show that x is an interior point of E°, though, we will need to show that every point in the neighborhood $N_r(x)$ is also a

point of $E°$.

Let y be an arbitrary point in $N_r(x)$. Clearly, $0 < d(x, y) < r$. Choose a radius s such that $d(x, y) + s = r$ and let z be an arbitrary point in $N_s(y)$. Every point in $N_s(y)$ is a point in E, since $d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + s = r$ implies $d(x, z) < r$, and we've established that every point in $N_r(x)$ is a member of E. And since every point in $N_s(y)$ is a point in E, by definition we know that y is an interior point of E. But y was an arbitrary point in $N_r(x)$, so we know that every point in $N_r(x)$ is an interior point of E. And this means that x is itself an interior point of the set of interior points of E: that is, x is an interior point of $E°$. And since x was an arbitrary point in $E°$, this means that every point in $E°$ is an interior point of $E°$. And this proves that $E°$ is open.
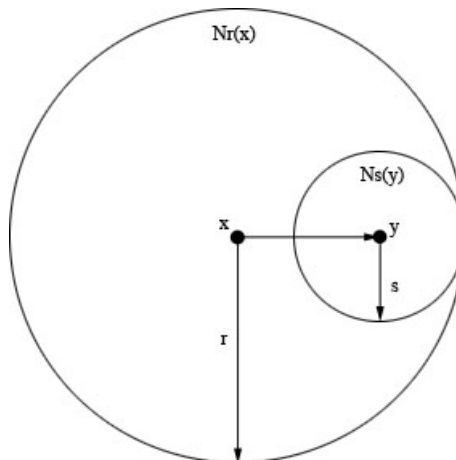
Exercise 2.9b

From exercise 2.9(a), we know that $E°$ is always open: so E is open if $E = E°$. And if E is open, then every point of E is an interior point and therefore $E° = E$: so $E = E°$ if E is open. Together, these two statements show that E is open if and only if $E = E°$.

Exercise 2.9c

Let p be an arbitrary element of G. Because G is open, p is an interior point of G. This means that for every point $p \in G$ there is some neighborhood $N_r(p)$ such that $N_r(p) \subset G \subset E$. This chain of subsets tells us that every point in G is an interior point not only of G, but also of E. And if every point in G is an interior point of E, then this shows us that $G \subset E°$.

Exercise 2.9d

Proof that $\widetilde{E°} \subset \overline{\widetilde{E}}$: let x be a member of $\widetilde{E°}$, the complement of $E°$. From the definition of $E°$ we know that x is not an interior point of E. From the definition of "interior point", this means that every neighborhood of x contains some element y in the complement of E. For each neighborhood, It must be true that either $x = y$ or $x \neq y$. If $x = y$ for one or more of these neighborhoods, then x is in $\widetilde{E}$ (since y is in $\widetilde{E}$ and $x = y$). If $x \neq y$ for every neighborhood, then by definition we know that x is a limit point of $\widetilde{E}$. So we conclude that either $x \in \widetilde{E}$ or x is a limit point for $\widetilde{E}$: that is, x is a member of $\overline{\widetilde{E}}$, the closure of $\widetilde{E}$.

Proof that $\overline{\widetilde{E}} \subset \widetilde{E°}$: this proof is only a very slight modification of the previous one. Let x be a member of $\overline{\widetilde{E}}$, the closure of the complement of E. From the definition of closure, either x is a limit point for $\widetilde{E}$ or x itself is a member of $\widetilde{E}$. In either case, every neighborhood of x contains some point of $\widetilde{E}$. By definition of "interior point", this means that x is not an interior point of E and therefore $x \in \widetilde{E°}$, the complement of $E°$.

Together, these two proofs show us that $\widetilde{E°} = \overline{\widetilde{E}}$.

Exercise 2.9e

No, E and $\bar{E}$ do not always have the same interiors. Consider the set $E = (-1, 0) \cup (0, 1)$ in $\mathbb{R}^1$. The point 0 is not a member of the interior of E, but it is a limit point of E. So the point 0 is an interior point of $\bar{E} = (-1, 1)$.

Exercise 2.9f

No, E and E° do not always have the same closures. Let E be a subset of $(0, 1]$ consisting only all rational numbers of the form $\frac{1}{m}, m \in \mathbb{N}$. As we saw in exercise 2.5, the closure of E is $E \cup \{0\}$. But every neighborhood of every point in E contains a point not in E, so E has no interior points. So $E^\circ = \varnothing$. And the empty set is already closed, so the closure of E° is $\varnothing$. Clearly $E \cup \{0\} \neq \varnothing$, so E and E° do not always have the same closures.

Exercise 2.10

Which sets are closed? Let E be an arbitrary subset of X and let p be an arbitrary point in E. The distance between any two distinct points is always 1, so by choosing $r = .5$ we guarantee that the neighborhood $N_r(p)$ contains only p itself. Because this is true for any point in E, this shows us that E contains no limit points. It's then vacuously true that all of the (nonexistant) limit points of E are points of E: we conclude that E is closed. But our choice of E was arbitrary, so this shows that every subset of X is closed.

Which sets are open? Let E be an arbitrary set in X. We've shown that every set is closed, so it must be the case that $\widetilde{E}$ is closed: and from theorem 2.23, this means that E is open. But our choice of E was arbitrary, so this shows that every subset of X is open.

Which sets are compact? Under this metric, subsets of X are compact if and only if they are finite. Proof: let E be an arbitrary subset of X. This set is either finite or infinite.

If E is finite with cardinality k, then we can always generate a finite subcover for any open cover. For each $e_i \in E$, we select some g_i in our (arbitrary) open cover $\{G_\alpha\}$ such that $e_i \in g_i$ (note that E is finite, so we don't need to use the axiom of choice). This gives us a collection of sets $\bigcup_{i=1}^{k} g_i$ such that $E \subseteq \bigcup_{i=1}^{k} g_i \subseteq \{G_\alpha\}$. By definition, then, $\bigcup_{i=1}^{k} g_i$ is a finite subcover of E. Our choices for E, k, and $\{G_\alpha\}$ were arbitrary, so this shows that any finite subset of X is compact.

If E is infinite, then we can always generate an open cover that has no finite subcover. As we saw previously, every subset of X is open: this includes subsets that consist of a single element. So we can create an infinite open cover of E by letting each set in $\{G_\alpha\}$ be a single element of E. Any subcover of E will need to contain one element of $\{G_\alpha\}$ for each element of E: and since E is infinite, this means that any subcover must be infinite. Therefore there is no finite subcover for this particular open cover, and E is not compact.

Exercise 2.11

d_1 is not a metric: let $x = -1, y = 0, z = 1$. Then $d_1(x, y) + d_1(y, z) = 2$ while $d_1(x, z) = 4$. And this tells us that $d_1(x, z) > d_1(x, y) + d_1(y, z)$ which violates definition 2.15(c) of metric spaces.

d_2 is a metric. It's trivial to verify that d_2 has the properties 2.15(a) and 2.15(b). To show that it has the property 2.15(c):

$\mapsto \lvert x - z \rvert \leq \lvert x - y \rvert + \lvert y - z \rvert$	true by the triangle inequality, theorem 1.37(f)
$\to \lvert x - z \rvert \leq \lvert x - y \rvert + \lvert y - z \rvert + 2\sqrt{\lvert x - y \rvert \lvert y - z \rvert}$	this additional term is > 0
$\to \sqrt{\lvert x - z \rvert}^2 \leq \left(\sqrt{\lvert x - y \rvert} + \sqrt{\lvert y - z \rvert} \right)^2$	
$\to \sqrt{\lvert x - z \rvert} \leq \sqrt{\lvert x - y \rvert} + \sqrt{\lvert y - z \rvert}$	legal because all terms are > 0
$\to d_2(x, z) \leq d_2(x, y) + d_2(y, z)$	

d_3 is not a metric. $d_3(1, -1) = 0$, which violates definition 2.15(a) of metric spaces.

d_4 is not a metric. $d_4(2, 1) = 0$, which violates definition 2.15(a) of metric spaces.

d_5 is a metric. It's trivial to verify that d_2 has the properties 2.15(a) and 2.15(b). To show that it has the property 2.15(c):

$\mapsto |x - z| \leq |x - y| + |y - z|$

This first step is always true because of the triangle inequality (theorem 1.37(f)).

$\rightarrow |x - z| \leq |x - y| + 2|x - y||y - z| + |y - z| + |x - z||x - y||x - z|$

Adding positive terms to the right side does not change the sign of the equality.

$\rightarrow |x - z| + |x - z||x - y| + |x - z||y - z| + |x - z||x - y||y - z|$
$\qquad \leq |x - y| + 2|x - y||y - z| + |y - z| + |x - z||x - y| + 2|x - z||x - y||y - z| + |x - z||y - z|$

It's hard to verify from this ugly mess, but we've just added the same terms to each side.

$\rightarrow (|x - z|)(1 + |x - y| + |y - z| + |x - y||y - z|) \leq (1 + |x - z|)(|x - y| + 2|x - y||y - z| + |y - z|)$

It's still hard to verify, but we've just factored out terms from each side.

$\rightarrow \dfrac{|x - z|}{1 + |x - z|} \leq \dfrac{|x - y| + 2|x - y||y - z| + |y - z|}{1 + |x - y| + |y - z| + |x - y||y - z|}$

And now we've divided each side by two of these factored terms.

$\rightarrow \dfrac{|x - z|}{1 + |x - z|} \leq \dfrac{|x - y|}{1 + |x - y|} + \dfrac{|y - z|}{1 + |y - z|}$
$\rightarrow d_2(x, z) \leq d_2(x, y) + d_2(y, z)$

The logic behind this seemingly arbitrary series of algebraic steps becomes clear if one starts at the end and works back to the beginning (which is, of course, how I initially derived the proof).

Exercise 2.12

Let $\{G_\alpha\}$ be any open cover of K. At least one open set in $\{G_\alpha\}$ must contain the point 0: let G_0 be this set, open in $\mathbb{R}$, that contains 0 as an interior point. As an interior point of G_0, there is some neighborhood $N_r(0) \subset G_0$. This neighborhood contains every $\frac{1}{n} \in K$ such that $\frac{1}{n} < r$: equivalently, this neighborhood contains one element for every natural number n such that $n > \frac{1}{r}$. Not only does this mean that G_0 contains an infinite number of elements of K, it means that it contains all but a finite number ($\frac{1}{r}$, to be exact) of elements of K.

There are infinitely many points contained in Nr(0)
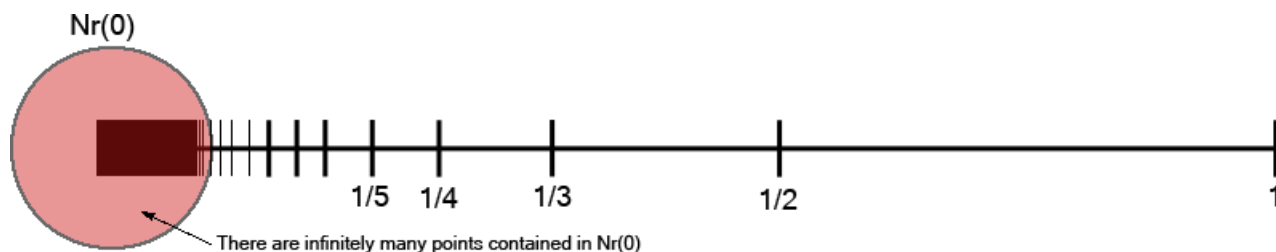
Let G_i represent an element of $\{G_\alpha\}$ containing $\frac{1}{i}$. We can now see that $K \subset G_0 \cup \bigcup_{i=1}^{1/r} G_i$, which is a finite subcover of $\{G_\alpha\}$.

Exercise 2.13

For $i \geq 2$, let A_i represent the set of all points of the form $\frac{1}{i} + \frac{1}{n}$ that are contained in the open interval $\left(\frac{1}{i}, \frac{1}{i-1}\right)$. Each A_i can be shown to have a single limit point of $\frac{1}{i}$ (see exercise 2.5).

Now consider the union of sets $S = \bigcup_{i=2}^{\infty} A_i$. The same reasoning used in exercise 2.5 shows that the set of limit points for S is just the collection of limit points from A_i: that is, the rationals of the form $\frac{1}{i}$ for $i \geq 2$. This shows us that S has one limit point for each natural number greater than 1, which is a countable number.

To show that S is compact, we return to the reasoning found in exercise 2.12. Let $\{G_\alpha\}$ be any open cover of S. For each $i \geq 2$, Let G_i represent the element of $\{G_\alpha\}$ that contains $\frac{1}{i}$. As we saw in exercise 2.12, G_i contains all but a finite number of elements from the interval $\left(\frac{1}{i}, \frac{1}{i-1}\right)$. If we let G_{ij} represent an element of $\{G_\alpha\}$ containing $\frac{1}{i} + \frac{1}{j}$, we see that all of the points of S in the interval $\left(\frac{1}{i}, \frac{1}{i-1}\right)$ can be covered by the finite union of sets $G_i \cup \bigcup_{i=1}^{r_i}$.

Now, let G_0 represent the element of $\{G_\alpha\}$ that contains 0 as an interior point. As an interior point, there is some neighborhood $N_r(0) \subset G_0$. This neighborhood contains every limit point of the form $\frac{1}{n} \in \mathbb{R}$ such that $\frac{1}{n} < r$. As in 2.12, this means that G_0 contains all but a finite number of the limit points of S. More than that, though, it means that G_0 contains all but a finite number of the intervals of the form $(\frac{1}{i}, \frac{1}{i-1})$.

And from these sets we can construct our finite subcover. Our subcover contains G_0. For each of the finitely many intervals not covered by G_0, we include the finite union of sets $G_i \cup \bigcup_{i=1}^{r_i}$. This is a finite collection of a finite number of elements from $\{G_\alpha\}$, and each element of S is included in one of these elements, so we have constructed a finite subcover for an arbitrary cover $\{G_\alpha\}$. This proves that S is compact.

Exercise 2.14

Let G_n represent the interval $\left(\frac{1}{n}, 1\right)$. The union $\{G_\alpha\} = \bigcup_{i=1}^{\infty} G_n$ is a cover for the interval $(0, 1)$ (Proof: for any $x \in (0, 1)$, there is some $n > 1/x$ and therefore some $\frac{1}{n} < x$. So $x \in G_{n+1}$). But there is no finite subcover for $(0, 1)$. Let H be a finite subcover of $\{G_\alpha\}$, and consider the set $\{\frac{1}{i+1} : G_i \in H\}$. This set is finite, so there is a least element. And this least element is in the interval $(0, 1)$ but is not a member of any $G_i \in H$. So our assumption that a finite subcover exists must be false.

Exercise 2.15

For each $i \in \mathbb{N}$, define A_i to be:

$$A_i = \left\{ p \in \mathbb{Q} : \sqrt{2 - \frac{1}{i}} \leq p \leq \sqrt{2 + \frac{1}{i}} \right\}$$

Because the endpoints are irrational, each of these intervals A_i are both bounded and closed (see exercise 2.16 for the proofs). From theorem 1.20, we also know that each interval is nonempty.

The union of any finite collection of these intervals will be nonempty, since the union $\bigcap_{i \in \nabla} A_i$ is equal to A_k, where k is the largest index in ∇. But the infinite union $\bigcap_{i=1}^{\infty} A_i$ is equal to $\{p \in \mathbb{Q} : \sqrt{2} \leq p \leq \sqrt{2}\}$, which is obviously empty.

This proves that the collection $\{A_i : i \in \mathbb{N}\}$ is a counterexample to theorem 2.36 if the word "compact" is replaced by either "closed" or "bounded". This same collection works as a counterexample to the corollary of 2.36, since $A_{n+1} \subset A_n$.

Exercise 2.16

lemma 1: For any real number $x \notin \mathbb{Q}$, the intervals $[x, \infty)$ and (x, ∞) are open in $\mathbb{Q}$. Proof: let x be a real number such that $x \notin \mathbb{Q}$. Let p be a rational number in the interval $[x, \infty)$ or (x, ∞). It cannot be the case that $p = x$ (since p is rational while x is not), so p is in the interval (x, ∞) (i.e, $[x, \infty)$ is identical to (x, ∞) in $\mathbb{Q}$). Because $x < p$, we can rewrite this interval as $(p - d(p, x), \infty)$. Choose $r = d(x, p)$. Then we see that $(p - r, p + r) \subseteq (p - d(p, x), \infty) = (x, \infty)$. This means that $N_r(p)$ is an interior point of (x, ∞): but p was an arbitrary point chosen from $[x, \infty)$, so every point of this interval is an interior point. This proves that $[x, \infty)$ and (x, ∞) are open in $\mathbb{Q}$.

lemma 2: For any real number $x \notin \mathbb{Q}$, the intervals $(-\infty, x]$ and $(-\infty, x)$ are open in $\mathbb{Q}$. Proof: the proof is nearly identical to that of lemma 1.

lemma 3: All of these open intervals $([x, \infty), (x, \infty), (-\infty, x]$, and $(-\infty, x))$ are also closed. Proof: Choose any of these four open sets. Note that its complement is also one of these four open sets. Since its complement is open, the set must be closed.

lemma 4: Every interval of the form (x, y) with $x, y \notin \mathbb{Q}$ is both open and closed in $\mathbb{Q}$. Proof: choose an arbitrary interval $E = (x, y)$. It's complement is $\widetilde{E} = (-\infty, x] \cup [y, \infty)$. From lemma 3, we see that $\widetilde{E}$ is a finite union of closed sets, so $\widetilde{E}$ is closed (theorem 2.24) and E is open (theorem 2.23). But $\widetilde{E}$ is also a finite union of open sets (lemmas 1 and 2), so $\widetilde{E}$ is open (theorem 2.24) and E is closed (theorem 2.23). This proves that (x, y) is both open and closed. The proofs for $[x, y], (x, y]$, and $[x, y)$ are identical to this one.

E is bounded: If $|p| \geq \pm\sqrt{3}$, then $p^2 > 3$ and $p \notin E$. So E is bounded by the interval $(-\sqrt{3}, \sqrt{3})$.

E is open and closed in $\mathbb{Q}$: E is the set of all rational numbers p such that $2 < p^2 < 3$, which means that E is the set of rational numbers in $(-\sqrt{3}, -\sqrt{2}) \cup (\sqrt{2}, \sqrt{3})$. We know that $\pm\sqrt{2}$ and $\pm\sqrt{3}$ are not rational, so by lemma 4 we know that E is both a union of closed sets and also a union of open sets. So by theorem 2.24 we know that E is both open and closed.

E is not compact: Drawing from the example in exercise 2.14, define the interval A_i to be:

$$A_i = \begin{cases} (-\sqrt{3}, -\sqrt{2}) & \text{if } i = 1 \\ \left(\sqrt{2 + \frac{1}{i}}, \sqrt{3}\right) & \text{if } i > 1 \end{cases}$$

Because $\sqrt{2 + \frac{1}{i}}$ is irrational for every $i \in \mathbb{N}$, we know by lemma 4 that each A_i is an open set. Define an open cover of E to be

$$G = \{A_i : i \in \mathbb{N}\}$$

This has no finite subcover for the same reason that the set in exercise 2.14 has no infinite subcover: given any finite collection of elements of G, we can always find an element sufficiently close to $\sqrt{2}$ that is not contained in any of those elements of G.

Exercise 2.17

E is not countable: This is proven directly by theorem 2.14.

E is not dense: Let $x = .660...$ and let $r = .01$. Then $N_r(x) = (.65, .67)$, and every real number in this interval contains a number other than 4 or 7. Therefore $N_r(x)$ contains no points of E. This shows that x is an element of $[0, 1]$ that is neither a point in E nor a limit point of E, which proves that E is not dense in $[0, 1]$.

E is compact: By theorem 2.41, we can prove that E is compact by showing that it is bounded and closed. $E \subset [0, 1]$, so E is bounded. To show that E is closed, we need to show that every limit point of E is a member of E.

Proof by contrapositive that every limit point of E is a member of E: let x be an element of $[0, 1]$ such that $x \notin E$. We will show that x is not a limit point. By the definition of membership in E, the fact that $x \notin E$ means that

$$x = \sum_{i=1}^{\infty} \frac{a_i}{10^i}$$

where some $a_i \notin \{4, 7\}$. Let k represent some index such that $a_k \notin \{4, 7\}$. Choose $r = 10^{-(k+1)}$. Then the neighborhood $N_r(x)$ does not contain any points in E:

i) If $a_{k+1} = 0$, then the $k + 1$st decimal place will be either $9, 0$, or 1 for every element of $(x - r, x + r)$ and so no element of $N_r(x)$ is a member of E.

ii) If $a_{k+1} = 9$, then the $k + 1$st decimal place will be either $8, 9,$ or 0 for every element of $(x - r, x + r)$ and so no element of $N_r(x)$ is a member of E.

iii) If a_{k+1} is neither 0 nor 9, then the kth decimal place will be be a_k for every element of $(x - r, x + r)$ and so no element of $N_r(x)$ is a member of E.

This proves that there is some neighborhood $N_r(x)$ that contains no points in E. This proves that x is not a limit point of E. But x was an arbitrary point such that $x \notin E$, so this proves that $(x \notin E \to x$ is not a limit point of $E)$. By contrapositive, this proves that every limit point of E is a member of E. And this, by definition, means that E is closed.

E **is perfect:** We have already shown that E is closed in the course of proving that E is compact. To prove that E is perfect, we need to show that every point in E is a limit point of E.

Let x be an arbitrary point in E. By the definition of membership in E, we know that

$$x = \sum_{i=1}^{\infty} \frac{a_i}{10^i}$$

where each a_i is either 4 or 7. Define a second number, x_k, such that

$$x_k = \sum_{i=1}^{\infty} \frac{b_i}{10^i}, \quad \text{where} \begin{cases} b_i = a_i & \text{if } i \neq k \\ b_k = 4 & \text{if } a_k = 7 \\ b_k = 7 & \text{if } a_k = 4 \end{cases}$$

From this definition, we see that x_k differs from x only in the kth decimal place, and $d(x, x_k) = 3 \times 10^{-k}$.

We can use this information to show that x is a limit point of E. Let $N_r(x)$ be any arbitrary neighborhood of x. From the archimedian principle, we can find some integer k such that $k > \log_1 0\frac{3}{r}$. And this is algebraically equivalent to finding some integer k such that $3 \times 10^{-k} < r$. This means that we can find some $x_k \in E$ (as defined above) in $N_r(x)$. And this was an arbitrary neighborhood of an arbitrary point in E, so we have proven that every neighborhood of every point in E contains a second point in E: by definition, this means that every point in E is a limit point.

Exercise 2.18

Section 2.44 describes the Cantor set. The Cantor set is a nonempty perfect subset of $\mathbb{R}^1$. Each point in the Cantor set is an endpoint of some segment of the form $\left[\frac{n}{3^k}, \frac{n+1}{3^k}\right]$ with $n, k \in \mathbb{N}$, so each point in the Cantor set is rational.

Let E be the set $\{x + \sqrt{2} : x$ is in the Cantor set$\}$ (that is, we're shifting every element of the Cantor set $\sqrt{2}$ units to the right). Each element in E is irrational (exercise 1.1). The Cantor set was bounded by $[0, 1]$ so E is clearly bounded by $[\sqrt{2}, 1 + \sqrt{2}]$. The proof that E is perfect is identical to the proof that the Cantor set is perfect (given in the book in section 2.44 and also in this document after exercise 2.4).

Exercise 2.19 a

We're told that A and B are disjoint, so $A \cap B = \varnothing$. And if A and B are closed, then by theorem 2.27 we know that $A = \bar{A}$ and $B = \bar{B}$. So we conclude that $A \cap B = \bar{A} \cap B = A \cap \bar{B} = \varnothing$. By definition, this means that A and B are separated.

Exercise 2.19 b

Let A and B be disjoint open sets. Assume that $A \cap \bar{B}$ is not empty. This assumption leads to a contradiction:

$\mapsto A \cap \overline{B}$ is not empty	hypothesis to be contradicted
$\rightarrow (\exists x)(x \in A \cap \overline{B})$	definition of non-emptiness
$\rightarrow (\exists x)(x \in A \cap (B \cup B'))$	definition of closure
$\rightarrow (\exists x)(x \in (A \cap B) \cup (A \cap B'))$	distributivity (section 2.11)
$\rightarrow (\exists x)(x \in \varnothing \cup (A \cap B'))$	A and B are disjoint
$\rightarrow (\exists x)(x \in A \cap B')$	
$\rightarrow (\exists x)(x \in A \wedge x \in B')$	definition of set intersection
$\rightarrow (\exists x)(x$ is an interior point of $A \wedge x \in B')$	A is an open set
$\rightarrow (\exists x)(x$ is an interior point of $A \wedge x$ is a limit point of $B)$	definition of B'
$\rightarrow (\exists x)\left[(\exists r \in \mathbb{R})(N_r(x) \subseteq A) \wedge (x \text{ is a limit point of } B)\right]$	definition of interior point
$\rightarrow (\exists x)\left[(\exists r \in \mathbb{R})(N_r(x) \subseteq A) \wedge (\forall s \in \mathbb{R})(\exists y \in B)(y \in N_s(x))\right]$	definition of limit point

If something is true for all $s \in \mathbb{R}$, it's true for any particular $\mathbb{R}$. So choose $s = r$.

$\rightarrow (\exists x)\left[(\exists r \in \mathbb{R})(\ (N_r(x) \subseteq A) \wedge (\exists y \in B)(y \in N_r(x))\ )\right]$	substitution of $s = r$
$\rightarrow (\exists x)\left[(\exists r \in \mathbb{R})(\exists y \in B)(y \in N_r(x) \wedge N_r(x) \subseteq A)\right]$	rearrangement of terms for clarity
$\rightarrow (\exists x)\left[(\exists r \in \mathbb{R})(\exists y \in B)(y \in A)\right]$	definition of subset

This last step establishes a contradiction. The sets A and B are disjoint, so there cannot be any possible choice of variables such that $y \in B$ and $y \in A$. Our assumption must have been incorrect: $A \cap \overline{B}$ is, in fact, empty. If we swap the roles of A and B, this same proof also shows us that $\overline{A} \cap B$ is empty. By definition, then, A and B are separated.

Exercise 2.19 c

A **is open in** X**:** The set A is, by definition 2.18(a), a neighborhood of p. By theorem 2.19, then, A is an open subset of X.

B **is open in** X**:** Let x be an arbitrary point in B. Let $r = d(p, x) - \delta$. Let y be an arbitrary element in $N_r(x)$. Proof that $y \in B$:

$\mapsto d(p, y) + d(y, x) \geq d(p, x)$	Property 2.15(c) of metric spaces
$\rightarrow d(p, y) + d(y, x) - d(p, x) \geq 0$	
$\rightarrow d(p, y) + d(y, x) - d(p, x) \geq 0 \wedge r > d(x, y)$	we know $d(x, y) < r$ because $y \in N_r(x)$
$\rightarrow d(p, y) + d(y, x) - d(p, x) \geq 0 \wedge d(p, x) - \delta > d(x, y)$	definition of r
$\rightarrow d(p, y) + d(y, x) - d(p, x) \geq 0 \wedge d(p, x) - \delta - d(x, y) > 0$	
$\rightarrow (d(p, y) + d(y, x) - d(p, x)) + (d(p, x) - \delta - d(x, y)) > 0$	the sum of positive numbers is positive
$\rightarrow d(p, y) - \delta > 0$	cancellation of like terms
$\rightarrow d(p, y) > \delta$	
$\rightarrow y \in B$	definition of membership in B

Our choice for y was arbitrary, so every point in this neighborhood of x is a member of B: by definition, then, x is an interior point of B. But our choice of $x \in B$ was also arbitrary, so this proves that every $x \in B$ is an interior point of B. By definition, this shows that B is open in X.

A **and** B **are disjoint:** If there were some $x \in A \cap B$, then $d(x, p) > \delta$ and $d(x, p) < \delta$, which violates the trichotomy rule for order relations.

A **and** B **are separate:** We've shown that A and B are disjoint open sets, so by exercise 2.19(b) we know that A and B are separate.

Exercise 2.19 d

If we are given any metric space X with a countable or finite number of elements, we can always find some distance δ such that there are no elements $x, y \in X$ with $d(x, y) = \delta$ (proof follows). This allows us to choose

some arbitrary $p \in X$ and then use the results from part (c) to completely partition X into separated sets $A = \{x \in X : d(x,p) > \delta\}$ and $B = \{x \in X : d(x,p) < \delta\}$. This proves that ($X$ is at most countable $\rightarrow X$ is not connected). By contrapositive, this proves that (X is connected $\rightarrow X$ is uncountable), which is what we were asked to prove.

Proof that an at most countable metric space X has some distance δ such that, for all $x, y \in X$, $d(x,y) \neq \delta$: Let X be an at most countable metric space with elements $a_1, a_2, \ldots a_n$. Then from theorem 2.13, we know that the set of all order pairs $\{(a_i, a_j) : a_i, a_j \in X\}$ is at most countable. And there is a clear one-to-one correspondence between this set and the set $\{(d(a_i, a_j) : a_i, a_j \in X\}$: so the set of all distances between all combinations of points in X is at most countable.

Distances in metric spaces are always real numbers (definition 2.15), which are of course uncountable. Because there are an at-most countable number of distances in the set $\{(d(a_i, a_j) : a_i, a_j \in X\}$, we can choose a real number δ that is not in this set (otherwise we would have an at most countable set with $\mathbb{R}$ as a proper subset).

But this isn't quite enough: if δ is so large that there are no elements in the set $\{x \in X : d(x,y) > \delta\}$, then one of our partitions will be empty. We can avoid this problem (as long as X has at least two elements) by picking arbitrary $x, y \in X$ and then choosing delta from the interval $(0, d(x,y))$. This interval is still uncountable, so we are still able to choose a δ that is not in this set. And we know that $d(x,y) > \delta$ and $d(x,x) < \delta$, so our partitions A and B will both be nonempty.

Exercise 2.20 a

Closures of connected sets are always connected. If A and B are connected, then there is either some p in $A \cap \overline{B}$ or some q in $\overline{A} \cap B$. Clearly, then, we have either either $p \in (A' \cup A) \cap \overline{B}$ or $q \in \overline{A} \cap (B' \cup B)$. Therefore $\overline{A} \cap \overline{B}$ is nonempty, so these two closed sets are connected.

Exercise 2.20 b

Consider the segment $(0,1)$ of the real line. Although this segment is open in $\mathbb{R}^1$, it contains no interior points in $\mathbb{R}^2$ since every neighborhood $N_r(x,0)$ contains the point $(x, \frac{r}{2})$. So let $r = \frac{1}{4}$ and let E be the set $N_r(0,0) \cup \{(x,0) : 0 < x < 1\} \cup N_r(1,0)$.
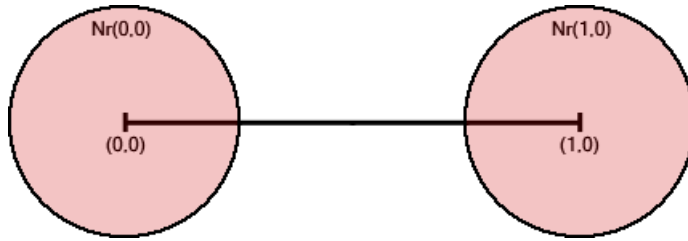
Since the line segment $\{(x,0) : 0 < x < 1\}$ contains no interior points while every point in a neighborhood is an interior point (theorem 2.19), the interior of E is just $N_r(0,0) \cup N_r(1,1)$, and it is trivial to show that this is the union of two nonempty separated sets.

Exercise 2.21 a

The function $p(t)$ can be thought of as the parameterization of a straight line connecting the points a and b. For instance, consider the following sets in $\mathbb{R}^2$:
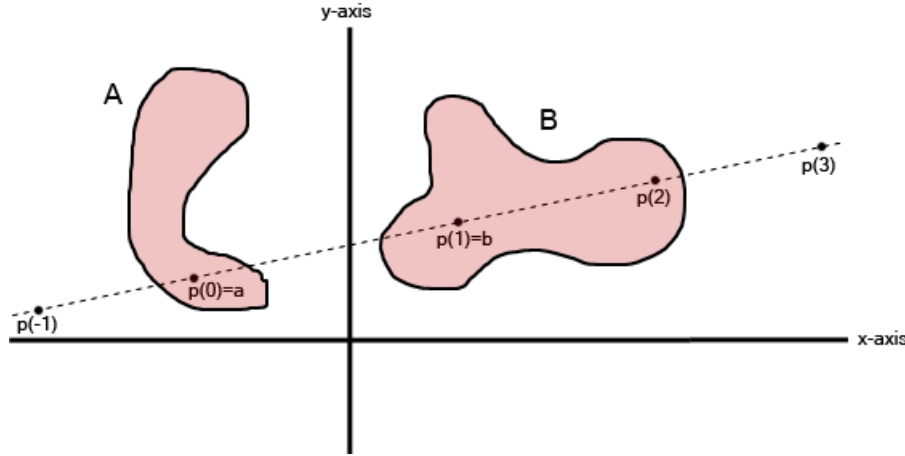
The set A_0 is the set of all t such that $p(t) \in A$, and the set B_0 is the set of all t such that $p(t) \in B$. We're told that A and B are separated and are asked to prove that A_0 and B_0 are separated. Proof by contrapositive:

• **There is some element t_A that is an element of A_0 and a limit point of B_0.**
Assume that A_0 and B_0 are not separated. By definition 2.45, either $\overline{A_0} \cap B_0$ or $A_0 \cap \overline{B_0}$ is nonempty. Assume that $A_0 \cap \overline{B_0}$ is nonempty (the sets are interchangeable, so the proof for $\overline{A_0} \cap B_0 \neq \varnothing$ is almost identical). Let t_A be one element of $A_0 \cap \overline{B_0}$: then either $t_A \in A_0 \cap B_0$ or $t_A \in B_0'$. It can't be the case that $t_A \in A_0 \cap B_0$, because this would imply $p(t_A) \in A \cap B$ which is impossible because A and B are separated sets. So it must be the case that $t_A \in A_0 \cap B_0'$.

• **There is a proportional relationship between $d(t, t+r)$ and $d(p(t), p(t+r))$.**
The distance between $p(t)$ and $p(t+r)$ is the vector norm of $p(t+r) - p(t)$:

$$d(p(t), p(t+r)) = |p(t+r) - p(t)| = |[(1-(t+r))a + (t+r)b] - [(1-t)a + tb]| = |r(b-a)| = |r||(b-a)| = d(t, t+r)d(a, b)$$

• **$p(t_A)$ is an element of A and a limit point of B.**
We've shown that t_A is a limit point of B_0. So for any arbitrarily small r, there is some $t_B \in B_0$ such that $d(t_A, t_B) < r$. And this means that for any arbitrarily small r, there is some $p(t_B) \in B$ such that $d(p(t_A), p(t_B)) < |r(b-a)|$. And since $p(t_A)$ is in A and each $p(t_B)$ is in B, this means that there is an element of A that is a limit point of B: So $A \cap \overline{B}$ is nonempty, which means that A and B are not separated.

This shows that if A_0 and B_0 are not separated, then A and B are not separated. By contrapositive, then, if A and B are separated then A_0 and B_0 are separated. And this is what we were asked to prove.

Exercise 2.21 b

We are asked to find $\alpha \in (0, 1)$ such that $p(\alpha) \notin A \cup B$: from the definition of the function p, this is equivalent to finding $\alpha \in (0, 1)$ such that $\alpha \notin A_0 \cup B_0$. Proof that such a α exists:

Let E be defined as $E = \{d(0, x) : x \in B_0\}$. That is, E is the set of all distances between the point 0 (which is in A_0, since $p(0) = a$) and elements of B_0. The set $\mathbb{R}$ has the greatest lower bound property and E is a subset of $\mathbb{R}$ with a lower bound of 0, so the set E has a greatest lower bound. Let α represent this greatest lower bound.

This intersection contains elements not in the set $A_0 \cup B_0$.
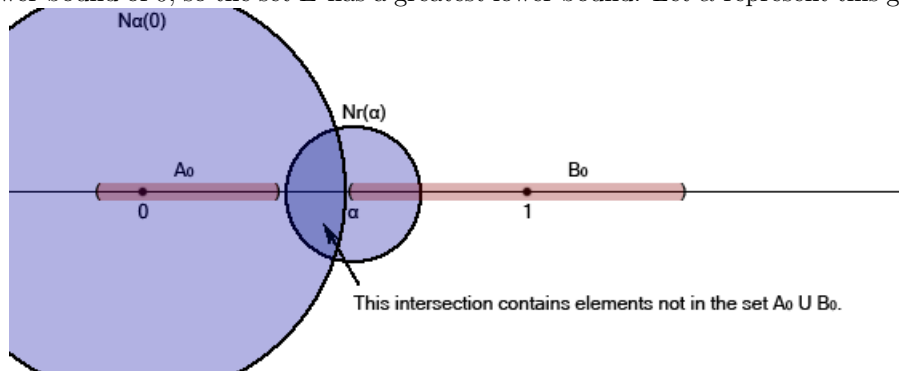
- **$0 < \alpha < 1$:**

We know that $\alpha > 0$ because α is a distance. We know that $\alpha \neq 0$, because this would mean that $d(0,0) \in E$ which means that $0 \in B_0$, which is false. And we know that $\alpha < 1$, because $1 \in B_0$ and B_0 is an open set: so 1 is an interior point of B_0, which means that there is some small r such that $1 - r \in B_0$.

We now need to show that we can use α to construct elements that are not in $A_0 \cup B_0$. We know that either $\alpha \in B_0, \alpha \in A_0$, or $\alpha \notin A_0 \cup B_0$:

- **if $\alpha \in B_0$:**

We know that α is not a limit point of A_0 (otherwise, $x_\alpha \in B_0 \cap \overline{A_0}$ which contradicts the fact from part (a) that B_0 and A_0 are separated). So there is some neighborhood $N_r(\alpha)$ that contains no points in A_0. Now consider the points p in the range $(\alpha - r, \alpha)$. Each p is in the neighborhood $N_r(\alpha)$, so they aren't members of A_0. And for each p we see that $d(0, p) < \alpha$, so they aren't members of B_0 (otherwise α wouldn't have been a lower bound of E). So we see that for every $p \in (\alpha - r, \alpha)$, we have $p \notin A_0 \cup B_0$.

- **if $\alpha \in A_0$:**

This assumption leads to a contradiction. We've shown in part (a) that A_0 and B_0 are separated, so it can't be the case that α is a limit point of B_0 (otherwise $A_0 \cap \overline{B_0}$ would not be empty). So there is some neighborhood $N_r(\alpha)$ that contains no points of B_0. But if the interval $(0, \alpha)$ contains no points of B_0 and the interval $(\alpha - r, \alpha + r)$ contains no points of B_0, then their union $(0, \alpha + r)$ contains no points of B_0. But this contradicts our definition of α as the greatest lower bound of E. So our initial assumption must have been incorrect: it cannot have been the case that $\alpha \in A_0$.

- **if $\alpha \notin A_0$ and $\alpha \notin B_0$:**

Under this assumption, we clearly have $\alpha \notin A_0 \cup B_0$.

Whatever assumption we make about the set containing α, we see that there will always be at least one element α such that $0 < \alpha < 1$ and $\alpha \notin A_0 \cup B_0$. And, from the definition of the function p, this means that $p(\alpha) \notin A \cup B$. And this is what we were asked to demonstrate.

Exercise 2.21 c

Proof by contrapositive. Assume that E is not connected: then E could be described as the union of two separated sets (i.e, $E = A \cup B$). From part (b), we could then choose a, b, and t such that $(1 - t)a + (t)b \notin A \cup B = E$. And this would mean that E is not convex. By contrapositive, if E is convex then E is connected.

Exercise 2.22

The metric space $\mathbb{R}^k$ clearly contains $\mathbb{Q}^k$ as a subset. We know that $\mathbb{Q}^k$ is countable from theorem 2.13. To prove that $\mathbb{Q}^k$ is dense in $\mathbb{R}^k$, we need to show that every point in $\mathbb{R}^k$ is a limit point of $\mathbb{Q}^k$:

Let $\mathbf{a} = (\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_k})$ be an arbitrary point in $\mathbb{R}^k$ and let $N_r(\mathbf{a})$ be an arbitrary neighborhood of $\mathbf{a}$. Let $\mathbf{b} = (\mathbf{b_1}, \mathbf{b_2}, \ldots, \mathbf{b_k})$ where b_i is chosen to be a rational number such that $a_i < b_i < a_i + \frac{r}{\sqrt{k}}$ (possible via theorem 1.20(b)). The point $\mathbf{b}$ is clearly in $\mathbb{Q}^k$, and

$$d(a, b) = \sqrt{(a_1 - b_i)^2 + \ldots + (a_k - b_k)^2} < \sqrt{\frac{r^2}{k} + \ldots + \frac{r^2}{k}} = \sqrt{\frac{kr^2}{k}} = r$$

This shows that every point in $\mathbb{R}^k$ is a limit point of $\mathbb{Q}^k$, which completes the proof that $\mathbb{R}^k$ is separable.

Exercise 2.23

Let X be a separable metric space. From the definition of *separable* in the previous exercise, we know that X contains a countable dense subset E. For each $e_i \in E$, let $N_q(e_i)$ be a neighborhood with rational radius q around point e_i. Let $\{V_\alpha\} = \{N_q(e_i) : q \in \mathbb{Q}, i \in \mathbb{N}\}$ be the collection of all neighborhoods with rational radius centered around members of E. This is a countable collection of countable sets, and therefore $\{V_\alpha\}$ has

countably many elements.

Let x be an arbitrary point in X, and let G be an arbitrary open set in X such that $x \in G$. Because G is open, we know that x is an interior point of G. So there is some neighborhood $N_r(x)$ such that $N_r(x) \subseteq G$. But we can choose a rational q such that $0 < q < \frac{r}{2}$, so that $x \in N_q(x) \subseteq N_r(x) \subseteq G$.
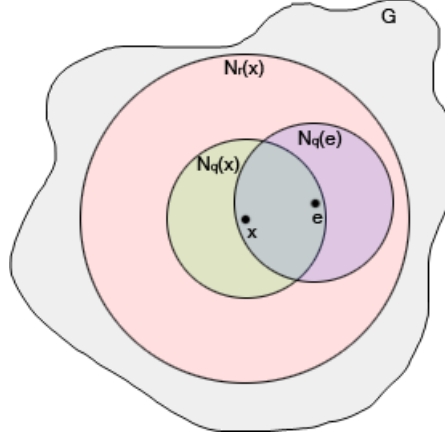
Because E is dense in x, every neighborhood of x contains some $e \in E$. So $e \in N_q(x)$, which means that $d(e, x) < q$. But, on the other hand, this also means that $d(x, e) < q$ so that $x \in N_q(e)$. And $N_q(e) \in \{V_\alpha\}$, by definition.

Having shown that $x \in N_q(e)$, we need to prove that $N_q(e) \subseteq G$. Let y be any point in $N_q(e)$. We know that $d(x, e) < q$ and $d(e, y) < q$. So $d(x, y) < d(x, e) + d(e, y) = 2q$ by definition 2.15(c) of metric spaces. But we defined q so that $0 < q < \frac{r}{2}$, so we know that $d(x, y) < r$. This means that every $y \in N_q(e) \rightarrow y \in N_r(x)$, or $N_q(e) \subseteq N_r(x)$. And we chose r so that $N_r(x) \subseteq G$, so by transitivity we know that $N_q(e) \subseteq G$.

We started by choosing an arbitrary element x in an arbitrary open set $G \subseteq X$, and proved that there was an element $N_q(e) \in \{V_\alpha\}$ such that $x \in N_q(e) \subset G$. We've shown that $\{V_\alpha\}$ has a countable number of elements, and each element is an open neighborhood. This proves that $\{V_\alpha\}$ is a base for X.

Exercise 2.24

We're told that X is a metric space in which every infinite subset has a limit point. Choose some arbitrary δ and construct a set $\{x_i\}$ as described in the exercise.

- **The set $\{x_i\}$ must be finite:** if this set were infinite, then it would have a limit point. And if it has a limit point, then there would be multiple points of $\{x_i\}$ in some neighborhood of radius $\delta/2$, which contradicts our assumption that $d(x_i, x_j) > \delta$ for each pair of points in $\{x_i\}$.

- **The neighborhoods of $\{x_i\}$ form a cover of X:** Every $x \in X$ must be contained in $N_\delta(x_i)$ for some $x_i \in \{x_i\}$. If it weren't, this would imply that $d(x, x_i) > \delta$ for each $x_i \in \{x_i\}$, so that we could have chosen x to be an additional element in $\{x_i\}$.

Let E_i represent the set of $\{x_i\}$ constructed in the above way when $\delta = \frac{1}{i}$, and consider the set $E = \bigcup_{i=1}^{\infty} E_i$. We will show that E is a countable dense subset of X. This union E is a countable union of nonempty finite sets, so E is countable. Each element of E was chosen from X, so clearly $E \subset X$. Proof that E is dense in X:

Choose an arbitrary $x \in X$. Choose an arbitrarily small radius r. From the archimedian principle, we can find an integer n such that $\frac{1}{n} < r$. We've shown that the neighborhoods of E_n cover X, so there is some $e \in E_n$ such that $x \in N_{1/n}(e)$. But if $d(e, x) < \frac{1}{n}$, then $d(x, e) < \frac{1}{n}$ and so $e \in N_{1/n}(x)$. And since $\frac{1}{n} < r$, this implies that $e \in N_r(x)$.

So we've chosen an arbitrary x and an arbitrary radius r, and shown that $N_r(x)$ will always contain some $e \in E_n \subset E$. This proves that every x is a limit point of E, which by definition means that E is dense in X.

This proves that X contains a countable dense subset, so by the definition in exercise 22 we have proven that X is separable.

Exercise 2.25 a

Theorem 2.41 tells us that the set K, by virtue of being a compact space, has the property that every infinite subset of K has a limit point. By exercise 24, this means that K is separable. And exercise 23 proves that every separable metric spaces has a countable base.

I'm not sure that this an appropriate proof, though. The question wants us to prove that K has a countable base and that *therefore* K is separable, whereas this is a proof that K is separable and therefore has a countable base. An alternate proof follows.

Exercise 2.25 b

We're told that K is compact. Consider the set $E_n = \{N_{1/n}(k) : k \in K\}$, the set of neighborhoods of radius $1/n$ around every element in K. This is clearly an open cover of K, since $x \in K \to x \in N_(1/n)(x) \to x \in E_n$. And K is compact, so the open cover E_n must have some finite subcover: i.e., K must be covered by some finite number of neighborhoods from E_n. Let $\{V_n\}$ represent a finite subcover of the open cover E_n.

Now consider the union $V = \bigcup_{n=1}^{\infty} V_n$, a countable collection of finite covers of K. We will prove that V is a base for K.

Choose an arbitrary $x \in K$, and an arbitrary open set G such that $x \in G \subset K$. Because G is open and $x \in G$, x must be an interior point of G. And since x is an interior point, there is some neighborhood $N_r(x)$ such that $N_r(x) \subset G$. Now choose an integer m such that $\frac{1}{m} < \frac{r}{2}$. Because V_m is an open cover of K, there is some $k \in K$ and neighborhood $N_{1/m}(k)$ in the open cover V_m such that $x \in N_{1/m}(k)$. Proof that this neighborhood $N_{1/m}(k)$ is a subset of G:

Assume that y is an element of $N_{1/m}(k)$. From the fact that $x \in N_{1/m}(k)$, we know that $d(x,k) < 1/m$. And from the fact that $y \in N_{1/m}(k)$, we know that $d(k,y) < 1/m$. And from the definition of metric spaces, we know that $d(x,y) \leq d(x,k) + d(k,y)$, which means that $d(x,y) \leq 2/m \leq r$ (since we chose $\frac{1}{m} < \frac{r}{2}$). This proves that every element of $N_{1/m}(k)$ is in the neighborhood $N_r(x)$, so $N_{1/m}(k) \subseteq N_r(x) \subseteq G$. And this shows that $N_{1/m}(k) \subseteq G$.

We've chosen an arbitrary element x and an arbitrary element G, and have shown that there is some $N_{1/m}(k) \in V$ such that $x \in N_{1/m}(k) \subset G$. By the definition in exercise 23, this proves that K has a base. And this base is a countable collection of finite sets, so it's a countable base.

This countable base V is identical to the set E constructed in exercise 24. We saw there that this base was a dense subset, so by the definition of "separable" in exercise 23 we know that K is separable.

Exercise 2.26

Let $\{V_n\}$ be the countable base we constructed in exercises 24 and 25.

$\{V_n\}$ acts as a countable subcover to any open cover of X. Proof: let G be an arbitrary open set from an arbitrary open cover of X. Choose any $x \in G$. We can find an element $N_{1/m}(k) \in V$ such that $x \in N_{1/m}(k) \subseteq G$ by the same method used in the second half of exercise 2.25(b).

We must now prove that a finite subset of $\{V_n\}$ can always be chosen to act as a subcover for any open cover. We'll do this with a proof by contradiction.

Let $\{W_\alpha\}$ be an open cover of X and assume that there is no finite subset of $\{V_n\}$ that acts as a subcover of $\{W_\alpha\}$. Construct F_n and E as described in the exercise. Because E is an infinite subset, it has a limit point: let x be this limit point. Every neighborhood $N_r(x)$ contains infinitely many points of E (theorem 2.20), and we can prove that this arbitrary neighborhood $N_r(x)$ must contain every point of E:

We have constructed the F_n sets in such a way that $F_{n+1} \subseteq F_n$ for each n. So if $N_r(x)$ doesn't contain any points from F_j, then it doesn't contain any points from F_{j+1}, F_{j+2},etc because $y \in F_{j+1} \to y \in F_j$. So if $N_r(x)$ doesn't contain any points from F_j, then it has no more than $j - 1$ points of E: i.e, $N_r(x) \cap E$ would be finite. But x is a limit point of E, so $N_r(x)$ must contain infinitely many points of E. So $N_r(x) \cap E$ contains one point from each F_n.

But this means that $E = \{x\}$, a set with one element. If it contained a second point, we could find a neighborhood $N_r(x)$ that failed to contain this second point, which would contradict our finding that every neighborhood of x contains E. And this means that we chose the same element from each F_n: i.e., x was in each F_n. So $x \in \bigcap F_n$, which contradicts our assumption that $\bigcap F_n$ was empty.

Exercise 2.27

To prove that P is perfect, we must show that every limit point of P is a point of P, and vice-versa.

Proof that every limit point of P is a point of P: assume that x is a limit point of P. Choose some arbitrary r and let $s = \frac{r}{2}$. By the definition of limit point, every neighborhood $N_s(x)$ contains some $y \in P$. And from the fact that $y \in P$, we know that $N_s(y)$ contains uncountably many points of P. But for each $p \in N_s(y)$, we see that $d(x, p) \leq d(x, y) + d(y, p)$ so that $d(x, p) \leq 2s = r$. So the neighborhood $N_r(x)$ contains uncountably many points of P. But $N_r(x)$ was an arbitrary neighborhood of x, so this proves that $x \in P$.

Proof that every point of P is a limit point of P: assume that $x \in P$. By the definition of P, this means that every neighborhood of x contains uncountably many points of P. So clearly every neighborhood of x contains at least one point of P, which means that x is a limit point of P.

This proves that P is perfect. To prove that $P^c \cap E$ is countable, we let $\{V_n\}$ be a countable base for X and let W be the union of all V_n for which $E \cap V_n$ is countable. We will show that $W^c = P$:

Proof that $P \subseteq W^c$: Assume that $x \in P$. By the definition of membership in P, we know that x is a condensation point of E. By the definition of condensation point, then, we know that every neighborhood $N_r(x)$ contains uncountably many points of E. And this means that every open set V_n containing x contains uncountably many points of E. So x is not a member of any countable $V_n \cap E$, which means that $x \in W^c$.

Proof that $W^c \subseteq P$: Assume that $x \in W^c$. Choose any arbitrary neighborhood $N_r(x)$: by the definition of the base, there is some $V_n \subset N_r(x)$ such that $x \in V_n$. By definition of W^c, the fact that $x \in V_n$ means that $E \cap V_n$ is uncountable, which of course means that V_n has uncountably many elements of E. And we've established that $V_n \subseteq N_r(x)$, which means that $N_r(x)$ has uncountably many elements of E. But $N_r(x)$ was an arbitrary neighborhood of x, so this proves that every neighborhood of x has uncountably many elements of E: by definition, then, x is a condensation point and therefore $x \in P$.

So we have proven that $W^c = P$. Taking the complement of each of these, we see that $W = P^c$. And the set W is the union of countably many sets of the form $E \cap V_n$, each of which contains countably many elements: so W is countable. And $W = P^c$, so P^c is countable: this proves that countably many points are P^c. And from this, it's trivial to show that there are countably many points in $E \cap P^c$, which is what we were asked to prove.

Note that this proof assumed only that X had a countable base, so it is valid for any separable metric space.

Exercise 2.28

Let E be a closed set in a separable metric space X. Let P be the set of all condensation points for E.

Proof that $E \cap P$ is perfect: assume that $x \in E \cap P$. Clearly x is a point of P, which means that x is a limit point of E (from the definition of P) and a limit point of P (because P was shown to be perfect in exercise 27). So x is a limit point of $E \cap P$. Now, assume that x is a limit point of $E \cap P$. From this, we know that x is a point of E (because E is closed) and a point of P (because P is perfect). So we have shown that every point of $E \cap P$ is a limit point of $E \cap P$ and vice-versa: this proves that $E \cap P$ is perfect.

Proof that $E \cap P^c$ is at most countable: this was proven in exercise 27.

And $E = (E \cap P) \cup (E \cap P^c)$: that is, E is the union of a perfect set and a countable set. And this is what we were asked to prove.

Exercise 2.29

Let E be an arbitrary open set in $\mathbb{R}^1$ and let $\{G_i\}$ be an arbitrary collection of disjoint segments such that $\bigcup G_i = E$. Each G_i is open and nonempty, so each G_i contains at least one rational point. Therefore there can't be more G_i elements than rational numbers, which means that there are at most a countable number of elements in $\{G_i\}$. But $\{G_i\}$ was an arbitrary set of disjoint segements whose union is E, therefore E cannot be the union of an uncountable number of disjoint segments.

Exercise 2.30

Proof by contradiction: Let $\{F_n\}$ be a collection of sets such that $\bigcup F_n = \mathbb{R}$ and suppose that each F_n has an empty interior.

$$\mapsto (\forall n)(F_n^\circ = \varnothing) \quad \text{hypothesis of contradiction}$$
$$\rightarrow \bigcup F_n^\circ = \varnothing$$
$$\rightarrow (\bigcup F_n^\circ)^c = \mathbb{R} \quad \text{taking the complement of both sides}$$
$$\rightarrow \bigcap (F_n^\circ)^c = \mathbb{R} \quad \text{theorem 2.22}$$
$$\rightarrow \bigcap \overline{F_n^c} = \mathbb{R} \quad \text{exercise 2.9d}$$
$$\rightarrow (\forall n)(\overline{F_n^c} = \mathbb{R})$$

This last step tells us that F_n^c is dense in $\mathbb{R}$ for every n. And F_n was closed, so F_n^c is open. Appealing to the proof of Baire's theorem in exercise 3.22 (which uses only terms and concepts introduced in chapter 2), we see that $\bigcap F_n^c$ is nonempty. Therefore:

$$\rightarrow \bigcap F_n^c \neq \varnothing$$
$$\rightarrow (\bigcup F_n)^c \neq \varnothing \quad \text{theorem 2.22}$$
$$\rightarrow \bigcup F_n \neq \mathbb{R} \quad \text{taking the complement of both sides}$$

And this contradicts our original claim that $\bigcup F_n = \mathbb{R}$ so one of our initial assumptions must be wrong. And our only assumption was that each F_n has an empty interior, so by contradiction we have proven that at least one F_n must have a nonempty interior.

Exercise 3.1

The exercise does not explicitly say that we're operating in the metric space $\mathbb{R}^k$, but there are two reasons to assume that we are. First, we are supposed to understand the meaning of absolute value in this metric space, and $\mathbb{R}^k$ is the only metric space we've encountered so far for which absolute value has been defined. Second, Rudin appears to use s_n to represent series in $\mathbb{R}^k$ and p_n to represent series in arbitrary metric spaces. So we'll assume that we're operating in $\mathbb{R}^k$ for this exercise.

We're told that the sequence $\{s_n\}$ converges to some value s: that is, for any arbitrarily small ϵ there is some integer N such that $n > N$ implies $d(s, s_n) < \epsilon$. But it's always the case that $d(|s|, |s_n|) \leq d(s, s_n)$ (exercise 1.13), so for this same ϵ and N we see that $d(|s|, |s_n|) \leq d(s, s_n) < \epsilon$. By transitivity, this means that for any choice of ϵ there is some integer N such that $n > N$ implies $d(|s|, |s_n|) < \epsilon$. By definition of convergence, this means that the sequence $\{|s_n|\}$ converges to $|s|$.

The converse is not true. If we let $s_k = (-1)^k$, the sequence $\{|s_n|\}$ clearly converges while the sequence $\{s_n\}$ clearly does not.

Exercise 3.2

The exercise asks us to calculate the limit rather than prove it rigorously, so we might be able to manipulate the expression algebraically:

$$\sqrt{n^2+n}-n = (\sqrt{n^2+n}-n)\left(\frac{\sqrt{n^2+n}+n}{\sqrt{n^2+n}+n}\right) = \frac{n}{\sqrt{n^2+n}+n} = \frac{n}{n\sqrt{1+1/n}+n} = \frac{1}{\sqrt{1+1/n}+1}$$

and then use basic calc 2 techniques to show that this limit is $\frac{1}{2}$. If we need to prove it more rigorously, we can use theorem 3.14. We will first need to prove that this sequence has a limit, which theorem 3.14 tells us can be done by proving that the sequence is bounded and monotonically increasing. We can then find the limit by finding the least upper bound of the sequence.

The sequence is bounded

The quantitity $(\sqrt{n^2+n}-n)$ is bounded below: from the fact that $\sqrt{n^2+n} > \sqrt{n^2} = n$, we know that $\sqrt{n^2+n}-n > 0$. And it's bounded above by $\frac{1}{2}$:

$\mapsto \sqrt{n^2+n}-n \geq \frac{1}{2}$ hypothesis to be contradicted

$\rightarrow \sqrt{n^2+n} \geq \frac{1}{2}+n$

$\rightarrow n^2+n \geq \frac{1}{4}+n+n^2$ both sides are positive, so the sign doesn't change

$\rightarrow 0 \geq \frac{1}{4}$ subtract $n+n^2$ from both sides

This last step is clearly false, so our initial assumption must be false: this shows that $(\sqrt{n^2+n}-n)$ is bounded above by $\frac{1}{2}$.

The sequence is monotonically increasing

$\mapsto s_{n+1} > s_n$

$\leftrightarrow \sqrt{(n+1)^2+(n+1)}-(n+1) > \sqrt{n^2+n}-n$ definition of s_n

$\leftrightarrow \sqrt{(n+1)^2+(n+1)}-\sqrt{n^2+n} > (n+1)-n$ rearrange terms

$\leftrightarrow \sqrt{n^2+3n+2}-\sqrt{n^2+n} > 1$ some algebra

$\leftrightarrow (n^2+3n+2)+(n^2+n)-2\sqrt{n^2+3n+2}\sqrt{n^2+n} > 1$ square both sides

$\leftrightarrow (2n^2+4n+2)-2\sqrt{n^4+4n^3+5n^2+2n} > 1$ simplify

$\leftrightarrow -2\sqrt{n^4+4n^3+5n^2+2n} > 1-(2n^2+4n+2)$ rearrange terms

$\leftrightarrow 2\sqrt{n^4+4n^3+5n^2+2n} < 2n^2+4n+1$ multiply both sides by -1 and simplify

$\leftrightarrow 4n^4+16n^3+20n^2+8n < 4n^4+16n^3+20n^2+8n+1$ square both sides

$\leftrightarrow 0 < 1$ subtract $4n^4+16n^3+20n^2+8n$ from each side

The sequence has a least upper bound of $\frac{1}{2}$

We've already shown that $\frac{1}{2}$ is an upper bound. To show that it is the least such upper bound, choose any $x \in [0, \frac{1}{2})$ and assume that it is an upper bound for $\{s_n\}$.

$\mapsto (\forall n \in \mathbb{N})(s_n \leq x)$ hypothesis to be contradicted

$\rightarrow (\forall n \in \mathbb{N})(\sqrt{n^2+n}-n \leq x)$ definition of s_n

$\rightarrow (\forall n \in \mathbb{N})(\sqrt{n^2+n} \leq x+n)$

$\rightarrow (\forall n \in \mathbb{N})(n^2+n \leq x^2+2xn+n^2)$ square both sides

$\rightarrow (\forall n \in \mathbb{N})(0 \leq x^2+(2x-1)n)$ subtract n^2+n from both sides

$\rightarrow (\forall n \in \mathbb{N})(-x^2 \leq (2x-1)n)$ subtract x^2 from both sides

$\rightarrow (\forall n \in \mathbb{N})(\frac{-x^2}{2x-1} \leq n)$ $2x-1 < 0$ when $x < \frac{1}{2}$

This last step must be false: by the Archimedian principle, we can always find some $n \in \mathbb{N}$ greater than any specific quantity. So, by contradiction, we can find $s_n > x$ for every $x \in [0, \frac{1}{2})$: this means that the upper bound cannot be less than $\frac{1}{2}$. But $\frac{1}{2}$ is an upper bound, so we see that $\frac{1}{2}$ is the least such upper bound.

We have shown that $\{s_n\}$ is a monotonically increasing bounded sequence with a least upper bound of $\frac{1}{2}$: by theorem 3.14, this proves that the limit of $\{s_n\}$ is $\frac{1}{2}$.

Exercise 3.3

Note that we are not asked to find the limit of this sequence. We are only asked to show that it converges and that 2 is an upper bound. By theorem 3.14, we can prove convergence by proving that the sequence is bounded and monotonically increasing.

The sequence is bounded above by $\sqrt{2 + \sqrt{2}}$

Although we're asked to show that the sequence has an upper bound of 2, we can prove the stronger result that it has an upper bound of $\sqrt{2 + \sqrt{2}}$. We can prove this by induction. We see that $0 < s_1 = \sqrt{2} < 2$. Now assume that $0 < s_n < 2$:

$$\mapsto 0 < s_n < 2 \qquad \text{hypothesis of induction}$$
$$\to 0 < \sqrt{s_n} < \sqrt{2}$$
$$\to 0 < 2 + \sqrt{s_n} < 2 + \sqrt{2}$$
$$\to \sqrt{2 + \sqrt{s_n}} < \sqrt{2 + \sqrt{2}}$$
$$\to s_{n+1} < \sqrt{2 + \sqrt{2}}$$

By induction, then, $s_n < \sqrt{2 + \sqrt{2}}$ for all n.

The sequence is monotonically increasing

Proof by induction. We can immediately see that $s_1 = \sqrt{2} < \sqrt{2 + \sqrt{2}} = s_2$. Now, assume that we have proven $s_{i-1} < s_i$ for $i = 1 \ldots n$.

$$\mapsto s_n = \sqrt{2 + \sqrt{s_{n-1}}} \qquad \text{definition of } s_n$$
$$\to s_n > \sqrt{2 + \sqrt{s_{n-2}}} \qquad \text{our hypothesis of induction tells us } s_{n-1} > s_{n-2}$$
$$\to s_n > s_{n-1} \qquad \text{definition of } s_{n-1}$$

We've shown that s_n is a monotonically increasing function that is bounded above. By theorem 3.14, this is sufficient to prove that it converges.

Exercise 3.4

From the recursive definition we are given, we can see that $s_2 = 0$ and $s_{2m} = \frac{1}{4} + \frac{1}{2}s_{2m-2}$. So we can use theorem 3.26 to find the limit of s_{2m}.

$$\lim_{m \to \infty} s_{2m} = \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots = \left(\sum_{n=0}^{\infty} \frac{1}{2^n} \right) - 1 - \frac{1}{2} = \frac{1}{1 - \frac{1}{2}} - 1 - \frac{1}{2} = \frac{1}{2}$$

From the same recursive definition, we see that $s_1 = 0$ and $s_{2m+1} = \frac{1}{2} + \frac{1}{2}s_{2m-1}$. So the limit of s_{2m+1} is given by

$$\lim_{m \to \infty} s_{2m+1} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = \left(\sum_{n=0}^{\infty} \frac{1}{2^n} \right) - 1 = \frac{1}{1 - \frac{1}{2}} - 1 = 1$$

This shows that as n increases, the terms of $\{s_n\}$ are either arbitrarily close to $\frac{1}{2}$ or arbitrarily close to 1, so these are our upper and lower bounds for the sequence.

Exercise 3.5

Let $a^* = \lim_{n \to \infty} \sup a_n$, let $b^* = \lim_{n \to \infty} \sup b_n$, and let $s^* = \lim_{n \to \infty} \sup(a_n + b_n)$.

If a^* and b^* are both finite:

From theorem 3.17(a) we know that there are some subsequences $\{a_j\}, \{b_k\}$ such that $\lim_{j \to \infty} a_j + \lim_{k \to \infty} b_k =$

s^* (that is, the supremum of E as defined above is also member of E: E is a closed set). The limit of a_j must be less than or equal to the supremum a^* and the limit of b_j must be less than or equal to the supremum b^*, so

$$s^* = \lim_{j \to \infty} a_j + \lim_{k \to \infty} b_k \leq \lim_{n \to \infty} \sup a_n + \lim_{n \to \infty} \sup b_n = a^* + b^*$$

By transitivity, this shows that $s^* \leq a^* + b^*$: and, by the definitions of these terms, this proves that

$$\lim_{n \to \infty} \sup(a_n + b_n) \leq \lim_{n \to \infty} \sup a_n + \lim_{n \to \infty} \sup b_n$$

If either a^* or b^* is infinite:
We're asked to discount the possibility that $a^* = \infty$ and $b^* = -\infty$. In all other cases when one or both of these values is infinite, the inequality can be easily shown to resolve to $\infty = \infty$ or $-\infty = -\infty$.

Exercise 3.6a

$$a_n = \sqrt{n+1} - \sqrt{n} = \left(\sqrt{n+1} - \sqrt{n}\right)\left(\frac{\sqrt{n+1} + \sqrt{n}}{\sqrt{n+1} + \sqrt{n}}\right) = \frac{1}{\sqrt{n+1} + \sqrt{n}}$$

We can compare this last term to a known series:

$$a_n = \frac{1}{\sqrt{n+1} + \sqrt{n}} > \frac{1}{2\sqrt{n+1}} > \frac{1}{2}\left(\frac{1}{n}\right)^{1/2}$$

We know that the series $\sum \left(\frac{1}{n}\right)^{1/2}$ diverges by theorem 3.28, so the terms of $\{a_n\}$ are larger than the terms of a divergent series. By the comparison theorem 3.25, this tells us that $\sum a_n$ is divergent.

Exercise 3.6b

$$a_n = \frac{\sqrt{n+1} - \sqrt{n}}{n} = \frac{\sqrt{n+1} - \sqrt{n}}{n}\left(\frac{\sqrt{n+1} + \sqrt{n}}{\sqrt{n+1} + \sqrt{n}}\right) = \frac{1}{\sqrt{n^3 + n^2} + \sqrt{n^2}}$$

We can compare this last term to a known series:

$$\frac{1}{\sqrt{n^3 + n^2} + \sqrt{n^3}} < \frac{1}{\sqrt{n^3} + \sqrt{n^3}} = \frac{1}{2}\left(\frac{1}{n}\right)^{3/2}$$

We know that the series $\sum \left(\frac{1}{n}\right)^{3/2}$ converges by theorem 3.28, so the terms of $\{a_n\}$ are smaller than the terms of a convergent series. By the comparison theorem 3.25, this tells us that $\sum a_n$ is convergent.

Exercise 3.6c

For $n \geq 1$, we can show that $0 \leq (\sqrt[n]{n} - 1) < 1$:

$$\mapsto (\forall n \in \mathbb{N})(1 \leq n < 2^n) \qquad \text{the } < 2^n \text{ can easily be proven via induction}$$
$$\to (\forall n \in \mathbb{N})(1 \leq \sqrt[n]{n} < 2) \qquad \text{take the } n^{th} \text{ root of each term}$$
$$\to (\forall n \in \mathbb{N})(0 \leq \sqrt[n]{n} - 1 < 1) \quad \text{subtract 1 from each term}$$

We know that the series $\sum x^n$ converges when $0 \leq x < 1$, so by the comparison theorem 3.25 we know that $\sum a_n$ is convergent.

Exercise 3.6d

This probably is more difficult than it looks at first, mainly because for complex z it's not true that $1 + z^n > z^n$ – in fact, without absolute value signs the inequality $z_1 > z_2$ has no meaning whatsoever (see exercise 1.8). It's also not true that $|1 + z^n| > |z^n|$. The best we can do is appeal to the triangle inequality.

If $|z| < 1$ then by the triangle inequality we have

$$\lim_{n \to \infty} \left| \frac{1}{1+z^n} \right| \geq \lim_{n \to \infty} \frac{1}{|1| + |z^n|} = 1$$

and therefore by theorem 3.23 the series doesn't converge. If $|z| = 1$ then we similarly have

$$\lim_{n \to \infty} \left| \frac{1}{1+z^n} \right| \geq \lim_{n \to \infty} \frac{1}{|1| + |z^n|} = \frac{1}{2}$$

and by theorem 3.23 the series again fails to converge. If $|z| > 1$ then we can use the ratio test:

$$\lim_{n \to \infty} \left| \frac{1+z^n}{1+z^{n+1}} \right| = \lim_{n \to \infty} \left| \frac{z^{-n}}{z^{-n}} \frac{1+z^n}{1+z^{n+1}} \right| = \lim_{n \to \infty} \left| \frac{z^{-n}+1}{z^{-n}+z} \right| = \left| \frac{1}{z} \right| < 1$$

and by theorem 3.34 the series converges. I'm not totally satisfied with this proof because I can't totally justify the claim that $\lim_{n \to \infty} z^{-n} = 0$ without appealing to the polar representation $z^{-n} = r^{-n} e^{-in\theta}$ and the fact that $|e^{ix\theta}| = 1$ for all x.

Exercise 3.7

Define the partial sum t_n to be

$$t_n = \sum_{k=1}^{n} \frac{\sqrt{a_k}}{k}$$

We can show that the series $\sum \sqrt{a_n}/n$ converges by showing that the sequence $\{t_n\}$ converges; we can do this by showing that $\{t_n\}$ is bounded and monotonically increasing (theorem 3.14).

the sequence is monotonically increasing:
From the definition of partial sums, we know that $t_n = \sqrt{a_n}/n + t_{n-1}$ for all n. And we're told $a_n > 0$ for every n, so $t_n > t_{n-1}$ for every n.

the sequence is bounded:
From the Cauchy-Schwartz inequality, we know that $\sum(ab) \leq \sqrt{\sum a^2 \sum b^2}$. Applying this to the given series, we see that

$$\sum \left(\sqrt{a_n} \frac{1}{n} \right) \leq \sqrt{\sum a_n \sum \frac{1}{n^2}}$$

We are told that $\sum a_n$ converges, and from theorem 3.28 we know that that $\sum 1/n^2$ converges. Therefore (by theorem 3.50) we know that their product converges to some α, and so the right-hand side of the above inequality converges to $\sqrt{\alpha}$. This shows us that the series $\sum \sqrt{a_n}/n$ is bounded by $\sqrt{\alpha}$, and therefore the sequence $\{t_n\}$ is bounded by $\sqrt{\alpha}$.

By assuming that $\sum a_n$ is convergent and that $a_n > 0$, we've shown that t_n is bounded and monotonically increasing. By theorem 3.14 this is sufficient to show that t_n converges. And this is what we were asked to prove.

Exercise 3.8

We're told that $\{b_n\}$ is bounded. Let α be the upper bound of $\{|b_n|\}$. We're also told that $\sum a_n$ converges: so for any arbitrarily small ϵ, we can find an integer N such that

$$\left| \sum_{k=m}^{n} a_k \right| \leq \frac{\epsilon}{\alpha} \text{ for all } n, m \text{ such that } n \geq m \geq N$$

which is algebraically equivalent to

$$\left| \sum_{k=m}^{n} a_k \, \alpha \right| \leq \epsilon \text{ for all } n, m \text{ such that } n \geq m \geq N$$

which, since $|b_k| \leq \alpha$ for every k, means that

$$\left| \sum_{k=m}^{n} a_k b_k \right| \leq \left| \sum_{k=m}^{n} a_k \alpha \right| \leq \epsilon \text{ for all } n, m \text{ such that } n \geq m \geq N$$

By theorem 3.22, this is sufficient to prove that $\sum a_n b_n$ converges.

Exercise 3.9a

$$\alpha = \limsup \sqrt[n]{|n^3|} = \limsup |n^{3/n}| = |n^0| = 1$$

So the radius of convergence is $R = 1/\alpha = 1$.

Exercise 3.9b

Example 3.40(b) mentions that "the ratio test is easier to apply than the root test", although we're never explicitly told what the ratio test is, how it might relate to the ratio test for series convergence, or how to apply the ratio test to example (b). According to some course handouts I found online [2], we can't simply apply the ratio test in the same way we use the root test. That is, we can't use the fact that

$$\beta = \limsup \left| \frac{2^{n+1}}{(n+1)!} \frac{n!}{2^n} \right| = \limsup \left| \frac{2}{n+1} \right| = 0$$

to assume that the radius of convergence is $1/\beta = \infty$. Instead, we look at theorem 3.37, which tells us that $\alpha \leq \beta$, so that

$$R = \frac{1}{\alpha} \geq \frac{1}{\beta} = \infty$$

Exercise 3.9c

$$\alpha = \limsup \sqrt[n]{\left| \frac{2^n}{n^2} \right|} = \limsup \frac{2}{\left| \sqrt[n]{n^2} \right|} = \frac{2}{(\limsup \sqrt{n})^2}$$

The last step of this equality is justified by theorem 3.3c. From theorem 3.20c, we know the limit of this last term:

$$\frac{2}{(\limsup \sqrt{n})^2} = \frac{2}{1}$$

So $\alpha = 2$, which gives us a radius of convergence of $1/2$.

Exercise 3.9d

$$\alpha = \limsup \sqrt[n]{|n^3 3^n|} = \limsup \frac{\sqrt[n]{n^3}}{3} = \frac{(\limsup \sqrt[n]{n})^3}{3}$$

The last step of this equality is justified by theorem 3.3c. From theorem 3.20c, we know the limit of this last term:

$$\frac{(\limsup \sqrt[n]{n})^3}{3} = \frac{1}{3}$$

So $\alpha = 1/3$, which gives us a radius of convergence of 3.

Exercise 3.10

We're told that infinitely many of the coefficients of $\sum a_n z^n$ are positive integers. This means that for every N, there is some $n > N$ such that $a_n \geq 1$. Therefore $\limsup_{n \to \infty} |a_n| \geq 1$

$$\to \limsup_{n \to \infty} \sqrt[n]{|a_n|} \geq 1 \quad \text{from the fact that } k > 1 \to k^{1/n} > 1$$
$$\to \frac{1}{R} \geq 1 \quad\quad\quad\quad\quad\quad \text{theorem 3.39}$$
$$\to 1 \geq R$$

This final step indicates that the radius of convergence is at most 1.

---

[2]http://math.berkeley.edu/~gbergman/

Exercise 3.11a

Proof by contrapositive: we will assume that $\sum a_n/(1 + a_n)$ converges and show that this implies that $\sum a_n$ converges.

To simplify the form of $\sum a_n/(1 + a_n)$, we can multiply the numerator and denominator of each term of the series by $1/a_n$ to get

$$\sum \frac{1}{\frac{1}{a_n} + 1}$$

For this to converge, it is necessary that

$$\lim_{n \to \infty} \frac{1}{\frac{1}{a_n} + 1} = 0$$

which can only happen if the limit of the denominator is ∞; and this can only happen if the limit of $1/a_n$ is ∞. That is, it must be the case that

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

This alone is not sufficient to prove that $\sum a_n$ converges. It does, however, allow us to make the terms of a_n as arbitrarily small as we like. For this proof, we it's helpful to recognize that there is some integer N_1 such that $a_n < 1$ for all $n > N_1$.

From the assumption that $\sum a_n/(1 + a_n)$ converges, we know that for every ϵ there is some N_ϵ such that

$$\sum_{k=m}^{n} \frac{a_k}{1 + a_k} < \epsilon \ \text{ for every } \ n \geq m \geq N_\epsilon$$

Let N be the larger of $\{N_1, N_\epsilon\}$. For any ϵ and any $n \geq m \geq N$, we can produce two inequalities:

$$\sum_{k=m}^{n} \frac{a_k}{1 + 1} < \sum_{k=m}^{n} \frac{a_k}{1 + a_k}$$

$$\sum_{k=m}^{n} \frac{a_k}{1 + a_k} < \epsilon \ \text{ for every } \ n \geq m \geq N$$

The first is true because each k is larger than N_1 (so that $a_k < 1$); the second is true because each k is larger than N_ϵ. Together, these inequalities tell us that

$$\sum_{k=m}^{n} \frac{a_k}{2} < \epsilon \ \text{ for every } \ n \geq m \geq N$$

or, equivalently, that

$$\sum_{k=m}^{n} a_k < 2\epsilon \ \text{ for every } \ n \geq m \geq N$$

And our choice of ϵ was arbitrary, so we have shown that for every ϵ there is some N that makes the last statement true: and this is the definition of convergence for the series $\sum a_n$.

We've shown that the convergence of $\sum \frac{a_n}{1 + a_n}$ implies the convergence of $\sum a_n$. By contrapositive, the fact that $\sum a_n$ diverges is proof that $\sum \frac{a_n}{1 + a_n}$ diverges.

Exercise 3.11b

From the definition of s_n, we can see that

$$s_{n+1} = a_1 + \ldots + a_{n+1} = s_n + a_{n+1}$$

and we're told that every $a_n > 0$, so we know that $s_{n+1} > s_n$. By induction, we also know that $s_n \geq s_m$ whenever $n \geq m$. And from this, we know that $\frac{1}{s_m} \geq \frac{1}{s_n}$ whenever $n \geq m$. Therefore:

$$\frac{a_{N+1}}{s_{N+1}} + \ldots + \frac{a_{N+k}}{s_{N+k}} \geq \frac{a_{N+1}}{s_{N+k}} + \ldots + \frac{a_{N+k}}{s_{N+k}} = \frac{a_{N+1} + \ldots + a_{N+k}}{s_{N+k}}$$

A bit of algebraic manipulation shows us that $a_{N+1} + \ldots + a_{N+k} = s_{N+k} - s_N$, so this last inequality is equivalent to

$$\frac{a_{N+1}}{s_{N+1}} + \ldots + \frac{a_{N+k}}{s_{N+k}} \geq \frac{s_{N+k} - s_N}{s_{N+k}} = 1 - \frac{s_N}{s_{N+k}}$$

which is what we were asked to prove.

To show that $\sum \frac{a_n}{s_n}$ is divergent, we once again look at the sequence $\{s_n\}$. We've already determined that $\{s_n\}$ is an increasing sequence, and we know that's it's not convergent (otherwise, by the definition 3.21 of "convergent series", $\sum a_n$ would be convergent). Therefore, we know that $\{s_n\}$ is not bounded (from theorem 3.14, which says that a bounded monotonic series is convergent).

Now let s_N be an arbitrary element of $\{s_n\}$ and let ϵ be an arbitrarily small real. From the fact that $\{s_n\}$ is not bounded, we can make s_{N+k} arbitrarily large by choosing a sufficiently large k. This means that, for any N and any arbitrarily small ϵ, we can make $\frac{s_N}{s_{N+k}} < \epsilon$ by choosing a sufficiently large k. And we've established that

$$\frac{a_{N+1}}{s_{N+1}} + \ldots + \frac{a_{N+k}}{s_{N+k}} \geq 1 - \frac{s_N}{s_{N+k}}$$

which means that, by choosing sufficiently large k, we get the inequality

$$\sum_{k=N+1}^{N+k} \frac{a_k}{s_k} \geq 1 - \frac{s_N}{s_{N+k}} \geq 1 - \epsilon$$

We've now established everything we need to show that $\sum \frac{a_n}{s_n}$ is divergent. Choose any ϵ such that $0 < \epsilon < \frac{1}{2}$. From theorem 3.22, in order for $\sum \frac{a_n}{s_n}$ to be convergent we must be able to find some integer N such that

$$\sum_{k=m}^{n} \frac{a_k}{s_k} < \epsilon \quad \text{for all } n \geq m \geq N$$

But there can be no such N, because for every N we've shown that there is some k such that

$$\sum_{k=N+1}^{N+k} \frac{a_k}{s_k} > 1 - \epsilon > \epsilon$$

(note that $1 - \epsilon > \epsilon$ because $0 < \epsilon < \frac{1}{2}$). And this is sufficient to show that the series does not converge.

Exercise 3.11c

To prove the inequality:

$$
\begin{aligned}
&s_n \geq s_{n-1} && \{s_n\} \text{ is an increasing sequence (see part b)} \\
&\rightarrow \tfrac{1}{s_n} \leq \tfrac{1}{s_{n-1}} \\
&\rightarrow \tfrac{a_n}{s_n^2} \leq \tfrac{a_n}{s_n s_{n-1}} && \text{multiply both sides by } a_n/s_n \\
&\rightarrow \tfrac{a_n}{s_n^2} \leq \tfrac{s_n - s_{n-1}}{s_n s_{n-1}} && s_n - s_{n-1} = a_n \text{ (see part b)} \\
&\rightarrow \tfrac{a_n}{s_n^2} \leq \tfrac{1}{s_{n-1}} - \tfrac{1}{s_n} && \text{algebra}
\end{aligned}
$$

Now consider the summation

$$\sum_{k=2}^{n} \frac{1}{s_{k-1}} - \frac{1}{s_k} = \left(\frac{1}{s_1} - \frac{1}{s_2} \right) + \left(\frac{1}{s_2} - \frac{1}{s_3} \right) + \ldots + \left(\frac{1}{s_{n-1}} - \frac{1}{s_n} \right)$$

Most of the terms in this summation cancel one another out: the summation "telescopes down" and simplifies to

$$\sum_{k=2}^{n} \frac{1}{s_{k-1}} - \frac{1}{s_k} = \frac{1}{s_1} - \frac{1}{s_k}$$

As we saw in part (b), the terms of $\{s_n\}$ increase without bound as $n \to \infty$, so that

$$\sum_{k=2}^{\infty} \frac{1}{s_{k-1}} - \frac{1}{s_k} = \frac{1}{s_1}$$

Now, by the inequality we proved above, we know that

$$\sum_{k=2}^{\infty} \frac{a_n}{s_n^2} \leq \sum_{k=2}^{\infty} \frac{1}{s_{k-1}} - \frac{1}{s_k} = \frac{1}{s_1}$$

We can add one term to each side so that our summation starts at 1 instead of at 2 to get

$$\sum_{k=1}^{\infty} \frac{a_n}{s_n^2} \leq \frac{a_1}{s_1^2} + \frac{1}{s_1}$$

We can now show that the series converges. Define the partial sum of this series to be

$$\{t_n\} = \sum_{k=1}^{n} \frac{1}{s_{k-1}} - \frac{1}{s_k}$$

We've shown that $\{t_n\}$ is bounded above by $a_1/s_1^2 + 1/s_1$, and we know that $\{t_n\}$ is monotonically increasing because each a_n is positive. Therefore, by theorem 3.14 we know that $\{t_n\}$ converges; and so by definition 3.21 we know that its associated series converges. And this is what we were asked to prove.

Exercise 3.11d

The series $\sum a_n/(1 + n^2 a_n)$ always converges. From the fact that $a_n > 0$, we can establish the following chain of inequalities:

$$\frac{a_n}{1 + n^2 a_n} = \frac{1/a_n}{1/a_n} \frac{a_n}{1 + n^2 a_n} = \frac{1}{\frac{1}{a_n} + n^2} < \frac{1}{n^2}$$

From this, we see that

$$\sum_{n=0}^{\infty} \frac{a_n}{1 + n^2 a_n} < \sum_{n=0}^{\infty} \frac{1}{n^2}$$

We know that $\sum 1/n^2$ converges (theorem 3.28), and therefore $\sum a_n/(1 + n^2 a_n)$ converges by the comparison test of theorem 3.25.

The series $\sum a_n/(1 + na_n)$ may or may not converge. If $a_n = 1/n$, for instance, the summation becomes

$$\sum \frac{a_n}{1 + na_n} = \sum \frac{\frac{1}{n}}{2} = \frac{1}{2} \sum \frac{1}{n}$$

which is divergent by theorem 3.28.

To construct a convergent series, let a_n be defined as

$$a_n = \begin{cases} 1 & \text{if } n = 2^m - 1 \ (m \in \mathbb{Z}) \\ 0 & \text{otherwise} \end{cases}$$

The series $\sum a_n$ is divergent, since there are infinitely many integers of the form $2^m - 1$. But the series $\sum a_n/(1 + na_n)$ is convergent:

$$\sum_{n=0}^{\infty} \frac{a_n}{1 + na_n} = \sum_{m=0}^{\infty} \frac{1}{2^m} = \sum \left(\frac{1}{2}\right)^m$$

This series is convergent to 2 by theorem 3.26.

Exercise 3.12a

establishing the inequality

From the definition of r_n, we see that

$$r_k = \sum_{m=k}^{\infty} a_m = a_k + \sum_{m=k+1}^{\infty} a_m = a_k + r_{k+1}$$

so that $a_k = r_k - r_{k+1}$. And because each a_k is positive, we see that $r_k > r_{k+1}$ which (from transitivity) means that $r_m > r_n$ when $n > m$ (that is, $\{r_n\}$ is a decreasing sequence). With these equalities, we can form our proof.

Take $m \leq n$ and consider the sum

$$\sum_{k=m}^{n} \frac{a_k}{r_k} = \frac{a_m}{r_m} + \ldots + \frac{a_n}{r_n}$$

From the fact that $r_m \geq r_n$, we know that $a_k/r_m \leq a_k/r_n$ for all k, so that

$$\frac{a_m + \ldots + a_n}{r_m} \leq \frac{a_m}{r_m} + \ldots + \frac{a_n}{r_n} \leq \frac{a_m + \ldots + a_n}{r_n}$$

Now, from the fact that $a_k = r_k - r_{k+1}$, we see that this inequality is equivalent to

$$\frac{(r_m - r_{m+1}) + (r_{m+1} - r_{m+2}) + \ldots + (r_n - r_{n+1})}{r_m} \leq \frac{a_m}{r_m} + \ldots + \frac{a_n}{r_n} \leq \frac{(r_m - r_{m+1}) + (r_{m+1} - r_{m+2}) + \ldots + (r_n - r_{n+1})}{r_n}$$

Notice that most of the terms of these numerators cancel one another out, leaving us with

$$\frac{r_m - r_{n+1}}{r_m} \leq \frac{a_m}{r_m} + \ldots + \frac{a_n}{r_n} \leq \frac{r_m - r_{n+1}}{r_n}$$

Taking the two leftmost terms of this inequality and performing some simple algebra then gives us

$$1 - \frac{r_{n+1}}{r_m} \leq \frac{a_m}{r_m} + \ldots + \frac{a_n}{r_n}$$

This is close to what we want to prove, but our index is off by one. But each a_n is positive and each r_n is the sum of positive terms, so a_{n+1}/r_{n+1} is strictly positive. By adding this term to the right-hand side of the inequality, we accomplish two things: we correct our index and we make this a strict inequality ($<$) instead of a non-strict inequality ($\leq$).

$$1 - \frac{r_{n+1}}{r_m} < \frac{a_m}{r_m} + \ldots + \frac{a_n}{r_n} + \frac{a_{n+1}}{r_{n+1}}$$

This last statement is true whenever $n \geq m$, which means that it's true whenever $n + 1 > m$. By a simple replacement of variables, then, we know that

$$1 - \frac{r_{n'}}{r_m} < \frac{a_m}{r_m} + \ldots + \frac{a'_n}{r'_n}$$

whenever $n' > m$, which is what we were asked to prove.

proof of divergence

We're told that $\sum a_n$ converges, so $\sum a_n = \alpha$ for some $\alpha \in \mathbb{R}$. From the definition of r_n, we see that there is a relationship between $\{r_n\}$ and α:

$$r_n = \sum_{k=n}^{\infty} a_k = \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{n-1} a_k = \alpha - \sum_{k=1}^{n-1} a_k$$

As $n \to \infty$, this last term approaches $\alpha - \alpha = 0$, and so we see that $\lim_{n \to \infty} r_n = 0$.

Now choose any arbitrarly small ϵ from the interval $(0, \frac{1}{2})$ and choose any arbitrary integer N. From the inequality we verified in part (a1), we see that

$$\frac{a_N}{r_N} + \ldots + \frac{a_{N+n}}{r_{N+n}} > 1 - \frac{r_{N+n}}{r_N}$$

We can choose n to be arbitrarily large; in doing so, r_{N+n} approaches zero while r_N remains fixed. So, by choosing a sufficiently large n, we can guarantee that $r_{N+n}/r_N < \epsilon$. This choice of n gives us the inequality

$$\frac{a_N}{r_N} + \ldots + \frac{a_{N+n}}{r_{N+n}} > 1 - \frac{r_{N+n}}{r_N} > 1 - \epsilon > \epsilon$$

where the last step of this chain of inequalities is justified by our choice of $0 < \epsilon < \frac{1}{2}$.

We've shown that there is some ϵ such that, for every N, we can find

$$\sum_{k=N}^{n} \frac{a_n}{r_n} > \epsilon$$

From theorem 3.22, this is sufficient to show that $\sum a_n/r_n$ diverges (since we've proven the negation of the "for every $\epsilon > 0$ there is some $N \ldots$" statement of the theorem).

Exercise 3.12b

$r_n > r_{n+1} > 0$	established in part (a)
$\rightarrow \sqrt{r_n} > \sqrt{r_{n+1}}$	take square root of each side
$\rightarrow 2\sqrt{r_n} > \sqrt{r_n} + \sqrt{r_{n+1}}$	add $\sqrt{r_n}$ to each side
$\rightarrow 2 > \frac{\sqrt{r_n} + \sqrt{r_{n+1}}}{\sqrt{r_n}}$	divide by $\sqrt{r_n}$
$\rightarrow 2(\sqrt{r_n} - \sqrt{r_{n+1}}) > \frac{(\sqrt{r_n} + \sqrt{r_{n+1}})(\sqrt{r_n} - \sqrt{r_{n+1}})}{\sqrt{r_n}}$	multiply each side by a positive term
$\rightarrow 2(\sqrt{r_n} - \sqrt{r_{n+1}}) > \frac{r_n - r_{n+1}}{\sqrt{r_n}}$	simply the right-hand side
$\rightarrow 2(\sqrt{r_n} - \sqrt{r_{n+1}}) > \frac{a_n}{\sqrt{r_n}}$	we established $a_n = r_n - r_{n+1}$ in part (a)

Having established this inequality, we see that

$$\sum_{k=1}^{n} \frac{a_k}{\sqrt{r_k}} < \sum_{k=1}^{n} 2(\sqrt{r_k} - \sqrt{r_{k+1}})$$

and many of the terms in the right-hand summation cancel one another out, so that the series "telescopes down" and simplifies to

$$\sum_{k=1}^{n} \frac{a_k}{\sqrt{r_k}} < 2\left(\sqrt{r_1} - \sqrt{r_{k+1}}\right)$$

so that

$$\lim_{n \to \infty} \sum_{k=1}^{n} \frac{a_k}{\sqrt{r_k}} < \lim_{n \to \infty} 2\left(\sqrt{r_1} - \sqrt{r_{k+1}}\right) = 2\sqrt{r_1}$$

This shows that this sum of nonnegative terms is bounded, which is sufficient to show that it converges by theorem 3.24.

Exercise 3.13

Let $\sum a_n$ be a series that converges absolutely to α and let $\sum b_n$ be a series that converges absolutely to β. We can prove that the Cauchy product of these two series (definition 3.48) is bounded.

$\mapsto \sum_{k=0}^{n} |c_k|$

$\quad = \sum_{k=0}^{n} \left| \sum_{j=0}^{k} a_j b_{k-j} \right|$ $\hspace{3cm}$ definition 3.48 of c_k

$\quad \leq \sum_{k=0}^{n} \sum_{j=0}^{k} |a_j b_{k-j}|$ $\hspace{3cm}$ triangle inequality

$\quad = \sum_{k=0}^{n} \sum_{j=0}^{k} |a_j||b_{k-j}|$ $\hspace{3cm}$ 1.33(c)

We can expand this summation out to get:

$$= |a_0||b_0| + (|a_0||b_1| + |a_1||b_0|) + \ldots + (|a_0||b_n| + |a_1||b_{n-1}| + \ldots + |a_n||b_0|)$$

Define B_n to be $\sum_{k=0}^{n} |b_k|$ and A_n to be $\sum_{k=0}^{n} |a_k|$. We can rearrange these terms to get:

$$= |a_0|B_n + |a_1|B_{n-1} + |a_2|B_{n-2} + \ldots + |a_n|B_0$$

Now we can add several nonnegative terms to get

$$\leq |a_0|B_n + |a_1|(B_{n-1} + |b_n|) + |a_2|(B_{n-2} + |b_{n-1}| + |b_n|) + \ldots + |a_n|(B_0 + |b_1| + \ldots + |b_n|)$$

$$= |a_0|B_n + |a_1|B_n + |a_2|B_n + \ldots + |a_n|B_n$$
$$= A_n B_n$$
$$= \left(\sum_{k=0}^{n} |a_k|\right)\left(\sum_{k=0}^{n} |b_k|\right)$$
$$= \alpha\beta$$

This shows that each partial sum of $\sum |c_k|$ is bounded above by $\alpha\beta$ and below by 0 (since it's a series of nonnegative terms). By theorem 3.24, this is sufficient to prove that $\sum |c_k|$ converges; and since $\sum c_k$ is the Cauchy product, we have proved that the Cauchy product $\sum c_k$ converges absolutely.

Exercise 3.14a

Choose any arbitrarily small $\epsilon > 0$. We're told that $\lim\{s_n\} = s$, which means that there is some N such that $d(s, s_n) < \epsilon$ whenever $n > N$. So we'll rearrange the terms of σ_n a bit:

$$\sigma_n = \frac{\sum_{k=0}^{n} s_k}{n+1} = \frac{\sum_{k=0}^{N} s_k + \sum_{k=N+1}^{n} s_k}{n+1}$$

Whenever $n > N$ we know that $d(s, s_n) < \epsilon$, which means that $-\epsilon < s_n - s < \epsilon$, or that $s - \epsilon < s_n < s + \epsilon$. This gives us the inequality

$$\frac{\sum_{k=0}^{N} s_k + \sum_{k=N+1}^{n}(s-\epsilon)}{n+1} < \sigma_n < \frac{\sum_{k=0}^{N} s_k + \sum_{k=N+1}^{n}(s+\epsilon)}{n+1}$$

The terms in some of these summations don't depend on k, so we can further rewrite this as

$$\frac{\sum_{k=0}^{N} s_k + (n-(N+1))(s-\epsilon)}{n+1} < \sigma_n < \frac{\sum_{k=0}^{N} s_k + (n-(N+1))(s+\epsilon)}{n+1}$$

$$\frac{\sum_{k=0}^{N} s_k}{n+1} + \frac{n(s-\epsilon)}{n+1} - \frac{(N+1)(s-\epsilon)}{n+1} < \sigma_n < \frac{\sum_{k=0}^{N} s_k}{n+1} + \frac{n(s+\epsilon)}{n+1} - \frac{(N+1)(s+\epsilon)}{n+1}$$

For many of these terms, the numerators are constant with respect to n. Therefore many of these terms will approach zero as $n \to \infty$.

$$\lim_{n\to\infty} \frac{n(s-\epsilon)}{n+1} < \lim_{n\to\infty} \sigma_n < \lim_{n\to\infty} \frac{n(s+\epsilon)}{n+1}$$

$$s - \epsilon < \lim_{n\to\infty} \sigma_n < s + \epsilon$$

And this is just another way of saying $d(s, \lim \sigma_n) < \epsilon$. And ϵ was arbitrarily small, so we have shown that $\lim_{n\to\infty} \sigma_n = s$.

Exercise 3.14b

Let $s_n = (-1)^n$. Then $\sum s_n$ is either equal to 0 or 1, depending on n, and

$$\frac{0}{n+1} \leq \sigma_n \leq \frac{1}{n+1}$$

Taking the limit as $n \to \infty$ gives us

$$0 \leq \lim_{n \to \infty} \sigma_n \leq 0$$

which can only be true if

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

Exercise 3.14c

Define s_n to be

$$s_n = \begin{cases} \left(\frac{1}{2}\right)^n + k & \text{if } n = k^3 \ (k \in \mathbb{Z}) \\ \left(\frac{1}{2}\right)^n & \text{otherwise} \end{cases}$$

There are no more than $\sqrt[3]{n}$ perfect cubes within the first n integers, so $\{s_n\}$ will contain no more than $\lfloor \sqrt[3]{n} \rfloor$ terms of the form $(1/2)^n + k$. This gives us the inequality

$$\sum_{k=0}^{n} s_n = \sum_{k=0}^{n} \left(\frac{1}{2}\right)^n + \sum_{k=0}^{\lfloor \sqrt[3]{n} \rfloor} k \leq 2 + \frac{\sqrt[3]{n}(\sqrt[3]{n} + 1)}{2}$$

where the last step in this chain of inequalities is justified by the common summation $\sum_{k=1}^{n} k = k(k+1)/2$. Continuing, we see that

$$\sum_{k=0}^{n} s_n \leq 2 + \frac{\sqrt[3]{n}(\sqrt[3]{n} + 1)}{2} \leq 2 + \frac{\sqrt[3]{n}(\sqrt[3]{n} + \sqrt[3]{n})}{2} = 2 + \sqrt[3]{n^2}$$

We can now analyze the value of σ_n.

$$\sigma_n = \frac{\sum_{k=0}^{n} s_n}{n+1} \leq \frac{2 + \sqrt[3]{n^2}}{n+1} = \frac{\frac{2}{\sqrt[3]{n^2}} + 1}{\sqrt[3]{n} + \frac{1}{\sqrt[3]{n^2}}}$$

Taking the limit of each side as $n \to \infty$ gives us

$$\lim_{n \to \infty} \sigma_n \leq \lim_{n \to \infty} \frac{\frac{2}{\sqrt[3]{n^2}} + 1}{\sqrt[3]{n} + \frac{1}{\sqrt[3]{n^2}}} = \lim_{n \to \infty} \frac{1}{\sqrt[3]{n}} = 0$$

Every term of $\{s_n\}$ was greater than zero, so we know that the arithmetic average σ_n is greater than zero. Therefore $0 \leq \lim \sigma_n \leq 0$, and therefore $\lim \sigma_n = 0$.

Exercise 3.14

proving the equality

$$\sum_{k=1}^{n} k a_k = (s_1 - s_0) + 2(s_2 - s_1) + \ldots + n(s_n - s_{n-1})$$

$$= -s_0 + (1-2)s_1 + (2-3)s_2 + \ldots + ((n-1)-n)s_{n-1} + s_n(n)$$

$$= -s_0 - s_1 - \ldots - s_n + (n+1)s_n$$

Therefore, if we divide this by $n+1$, we get

$$\frac{1}{n+1} \sum_{k=1}^{n} k a_k = \frac{-s_0 - s_1 - \ldots - s_n}{n+1} + s_n = -\sigma_n + s_n$$

establishing convergence

We're told that $\lim_{n\to\infty} na_n = 0$, so by part a we know that

$$\lim_{n\to\infty} \frac{\sum_{k=1}^n ka_k}{n+1} = 0$$

We're also told that $\{\sigma_n\}$ converges to some value σ, so by theorem 3.3(a) we know that

$$\lim_{n\to\infty} \left(\frac{\sum_{k=1}^n ka_k}{n+1} + \sigma_n \right) = (0 + \sigma) = \sigma$$

And since $s_n = \frac{\sum_{k=1}^n ka_k}{n+1} + \sigma_n$, this is sufficient to prove that $\lim_{n\to\infty} s_n = \sigma$.

Exercise 3.15

If you think I'm going to go through this tedious exercise, you can lick me where I shit.

Exercise 3.16a

$\{x_n\}$ is monotonically decreasing

From the fact that $0 < \sqrt{\alpha} < x_n$ we know that $\alpha < x_n^2$, and therefore

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{\alpha}{x_n} \right) < \frac{1}{2}\left(x_n + \frac{x_n^2}{x_n} \right) = x_n$$

This shows that $x_n > x_{n+1}$ for all n, which proves that $\{x_n\}$ is monotonically decreasing.

The limit of $\{x_n\}$ exists and is not less than $\sqrt{a}$

First, we show that $\{x_n\}$ is bounded below by $\sqrt{a}$. We know that $x_0 > \sqrt{a}$ because we chose it to be. And if $x_n > \sqrt{a}$ for any n, we have

$$
\begin{aligned}
&\mapsto x_n \neq \sqrt{a} && \text{assumed} \\
&\to x_n - \sqrt{a} \neq 0 \\
&\to (x_n - \sqrt{a})^2 \geq 0 && \text{1.18(d)} \\
&\to x_n^2 - 2x_n\sqrt{a} + a \geq 0 \\
&\to x_n^2 + a \geq 2x_n\sqrt{a} \\
&\to \frac{x_n^2 + a}{2x_n} \geq \sqrt{a} \\
&\to \frac{1}{2}\left(x_n + \frac{a}{x_n} \right) \geq \sqrt{a} \\
&\to x_{n+1} \geq \sqrt{a} && \text{definition of } x_{n+1}
\end{aligned}
$$

So by induction, we know that $x_n \geq \sqrt{a}$ for all n. We've now demonstrated that $\{x_n\}$ is monotonically decreasing and is bounded below by $\sqrt{a}$, so we're guaranteed that the limit of $\{x_n\}$ exists and that $\lim\{x_n\} \geq \sqrt{a}$.

The limit of $\{x_n\}$ is not greater than $\sqrt{a}$

Now we show that $\lim\{x_n\} \leq \sqrt{a}$. Assume that $\lim\{x_n\} = \sqrt{b}$ for some $\sqrt{b} > \sqrt{a}$. By the definition of "limit", we can find N such that $n > N$ implies $d(x_n, \sqrt{b}) < \epsilon$ for any arbitrarily small ϵ. We'll choose $\epsilon = \sqrt{b} - a$ (which is positive since $\sqrt{b} > \sqrt{a} \geq 1$ implies $b > a$).

$$
\begin{aligned}
&\mapsto d(x_n, \sqrt{b}) < \epsilon \\
&\to d(x_n, \sqrt{b}) < \sqrt{b} - a && \text{chosen value of } \epsilon \\
&\to x_n - \sqrt{b} < \sqrt{b} - a && \text{metric on } \mathbb{R}^1 \\
&\to x_n < \sqrt{b} - a + \sqrt{b}
\end{aligned}
$$

We can then calculate x_{n+1} in terms of x_n to get a chain of inequalities:

$$x_{n+1} = \frac{x_n^2 + a}{2x_n} < \frac{(\sqrt{b-a} + \sqrt{b})^2 + a}{2(\sqrt{b-a} + \sqrt{b})} = \frac{(b-a) + 2\sqrt{b}\sqrt{b-a} + b + a}{2(\sqrt{b-a} + \sqrt{b})}$$

$$= \frac{2b + 2\sqrt{b}\sqrt{b-a}}{2(\sqrt{b-a} + \sqrt{b})} = \frac{\sqrt{b}(\sqrt{b} + \sqrt{b-a})}{\sqrt{b} + \sqrt{b-a}} = \sqrt{b}$$

This shows us that if the distance between x_n and $\sqrt{b}$ becomes less than $\sqrt{b-a}$ (which it is guaranteed to do eventually because $\lim\{x_n\} = \sqrt{b}$), then the next term in the sequence $\{x_n\}$ will be *less* than $\sqrt{b}$. And since we've already shown that the sequence is monotonically decreasing, this means that every subsequent term will be even farther from $\sqrt{b}$, which contradicts our assumption that $\lim\{x_n\} = \sqrt{b}$. Therefore the limit is not $\sqrt{b}$: but $\sqrt{b}$ was an arbitrary value greater than $\sqrt{a}$, so we've shown that the limit cannot be any value greater than $\sqrt{a}$.

Having shown that $\lim\{x_n\}$ exists, that the limit is not less than $\sqrt{a}$, and that the limit is not greater than $\sqrt{a}$, we have therefore proven that $\lim\{x_n\} = \sqrt{a}$.

Exercise 3.16b

From the definition of x_{n+1} and e_n, we have

$$e_{n+1} = x_{n+1} - \sqrt{a} = \frac{x_n^2 + \sqrt{a}}{2x_n} - \sqrt{a} = \frac{x_n^2 - 2x_n\sqrt{a} + a}{2x_n} = \frac{(x_n - \sqrt{a})^2}{2x_n}$$

$$= \frac{e_n^2}{2x_n} < \frac{e_n^2}{2\sqrt{a}}$$

where the last step is justified by the fact that $x_n > \sqrt{a}$ (see part (a)).

The next part can be proven by induction. Setting $n = 1$ in the above inequality, we have

$$e_2 < \frac{e_1^2}{2\sqrt{a}} = \frac{e_1^2}{\beta} = \beta \left(\frac{e_1}{\beta}\right)^{2^1}$$

And, if the statement is true for n, then

$$e_{n+1} < \frac{e_n^2}{2\sqrt{a}} = \frac{1}{\beta}(e_n^2) < \frac{1}{\beta}\left[\beta\left(\frac{e_1}{\beta}\right)^{2^{n-1}}\right]^2 = \frac{1}{\beta}\left[\beta^2\left(\frac{e_1}{\beta}\right)^{2^n}\right] = \beta\left(\frac{e_1}{\beta}\right)^{2^n}$$

which shows that it is therefore true for $n + 1$. By induction, this is sufficient to show that it is true for all $n \in \mathbb{N}$.

Exercise 3.16c

When $a =$ and $x_1 = 2$, we have

$$\frac{e_1}{\beta} = \frac{x_1 - \sqrt{a}}{2\sqrt{a}} = \frac{2 - \sqrt{3}}{2\sqrt{3}}$$

which can be shown to be less than $1/10$ through simple algebra. Then, by part (b), we have

$$e_n < \beta\left(\frac{e_1}{\beta}\right)^{2^n} < 2\sqrt{3}\left(\frac{1}{10}\right)^{2^n} < 4\left(\frac{1}{10}\right)^{2^n}$$

Exercise 3.17

Lemma 1: $x_n > \sqrt{a} \to x_{n+1} < \sqrt{a}$, **and** $x_n < \sqrt{a} \to x_{n+1} > \sqrt{a}$

$\mapsto x_n > \sqrt{a}$

$\leftrightarrow x_n(\sqrt{a} - 1) > \sqrt{a}(\sqrt{a} - 1)$

$\leftrightarrow x_n\sqrt{a} - x_n > a - \sqrt{a}$

$\leftrightarrow x_n\sqrt{a} + \sqrt{a} > a + x_n$

$\leftrightarrow \sqrt{a}(x_n + 1) > a + x_n$

$\leftrightarrow \sqrt{a} > \frac{a + x_n}{x_n + 1}$

$\leftrightarrow \sqrt{a} > x_{n+1}$ definition of x_{n+1}

This shows that $x_n > \sqrt{a} \to x_{n+1} < \sqrt{a}$. If ">" is replaced with "<" in each of the above steps, we will have also constructed a proof that $x_n < \sqrt{a} \to x_{n+1} > \sqrt{a}$.

Lemma 2: $x_n > \sqrt{a} \to x_{n+2} < x_n$ **and** $x_n < \sqrt{a} \to x_{n+2} > x_n$

$\mapsto x_n > \sqrt{a}$

$\leftrightarrow x_n - \sqrt{a} > 0$

$\leftrightarrow 2(x_n - \sqrt{a})(x_n + \sqrt{a}) > 0$ multiplied by positive terms, so it's still > 0

$\leftrightarrow 2(x_n^2 - a) > 0$

At this point, the algebraic steps become bizarre and seemingly nonsensical. But bear with me.

$\leftrightarrow 2x_n^2 + 0x_n - 2a) > 0$

$\leftrightarrow 2x_n^2 + (1 + a - 1 - a)x_n - 2a) > 0$

$\leftrightarrow x_n + x_n^2 + ax_n + x_n^2 > a + ax_n + a + x_n$

$\leftrightarrow x_n(1 + x_n) + x_n(a + x_n) > a(1 + x_n) + a + x_n$

$\leftrightarrow \frac{x_n(1 + x_n) + x_n(a + x_n)}{1 + x_n} > \frac{a(1 + x_n) + a + x_n}{1 + x_n}$ division by a positive term

$\leftrightarrow x_n\left(1 + \frac{a + x_n}{1 + x_n}\right) > a + \frac{a + x_n}{1 + x_n}$

$\leftrightarrow x_n(1 + x_{n+1}) > a + x_{n+1}$

$\leftrightarrow x_n > \frac{a + x_{n+1}}{1 + x_{n+1}}$

$\leftrightarrow x_n > x_{n+2}$ definition of x_{n+2}

This shows that $x_n > \sqrt{a} \to x_{n+2} < x_n$. If ">" is replaced with "<" in each of the above steps, we will have also constructed a proof that $x_n < \sqrt{a} \to x_{n+2} > x_n$.

Exercise 3.17a

We're forced to choose x_1 such that $x_1 > \sqrt{a}$, so we can use induction with lemma 2 to show than $\{x_{2k+1}\}$ (the subsequence of $\{x_n\}$ consisting of elements with odd indices) is a monotonically decreasing sequence.

Exercise 3.17b

Because we chose x_1 such that $x_1 > \sqrt{a}$, lemma 1 tells us that $x_2 < \sqrt{a}$. We can then use induction with lemma 2 to show that $\{x_{2k}\}$ (the subsequence of $\{x_n\}$ consisting of elements with even indices) is a monotonically increasing sequence.

Exercise 3.17c

We can show than $\{x_n\}$ converges to $\sqrt{a}$ by showing that we can make $d(x_n, \sqrt{a})$ arbitrarily small. We do this by demonstrating the relationship between $d(x_n, \sqrt{a})$ and $d(x_1, \sqrt{a})$.

$$e_{n+2} = x_{n+2} - \sqrt{a} = \frac{a + x_{n+1}}{1 + x_{n+1}} - \sqrt{a} = \frac{a + x_{n+1} - \sqrt{a} - x_{n+1}\sqrt{a}}{1 + x_{n+1}}$$

$\mapsto e_{n+2} = x_{n+2} - \sqrt{a}$

$\rightarrow e_{n+2} = \frac{a + x_{n+1}}{1 + x_{n+1}} - \sqrt{a}$

$\rightarrow e_{n+2} = \frac{a + x_{n+1} - \sqrt{a} - x_{n+1}\sqrt{a}}{1 + x_{n+1}}$

$\rightarrow e_{n+2}(1 + x_{n+1}) = a + x_{n+1} - \sqrt{a} - x_{n+1}\sqrt{a}$

$\rightarrow e_{n+2}(1 + x_{n+1}) = x_{n+1}(1 - \sqrt{a}) - \sqrt{a}(1 - \sqrt{a})$

$\rightarrow e_{n+2}(1 + x_{n+1}) = (x_{n+1} - \sqrt{a})(1 - \sqrt{a})$

$\rightarrow e_{n+2}(1 + x_{n+1}) = e_{n+1}(1 - \sqrt{a})$

$\rightarrow e_{n+2} = e_{n+1}\left(\frac{1 - \sqrt{a}}{1 + x_{n+1}}\right)$

This tells us how to express e_{n+2} in terms of e_{n+1}. We can use this same equality to express e_{n+1} in terms of e_n, giving us

$$e_{n+2} = e_{n+1}\left(\frac{1 - \sqrt{a}}{1 + x_{n+1}}\right) = e_n\left(\frac{1 - \sqrt{a}}{1 + x_n}\right)\left(\frac{1 - \sqrt{a}}{1 + x_{n+1}}\right)$$

$$= e_n\left(\frac{1 - \sqrt{a}}{1 + x_n}\right)\left(\frac{1 - \sqrt{a}}{1 + \frac{a + x_n}{1 + x_n}}\right) = e_n\left(\frac{1 - \sqrt{a}}{1 + x_n}\right)\left(\frac{1 - \sqrt{a}}{\frac{1 + x_n + a + x_n}{1 + x_n}}\right)$$

$$= e_n\left(\frac{1 - \sqrt{a}}{1 + x_n}\right)\left(\frac{(1 + x_n)(1 - \sqrt{a})}{1 + 2x_n + a}\right) = e_n\left(\frac{(1 - \sqrt{a})^2}{1 + 2x_n + a}\right)$$

$$< e_n\left(\frac{(\sqrt{a} - 1)^2}{a - 1}\right)$$

The last step in this chain of inequalities is justified by the fact that $a + 2x_n + 1 > a - 1 > 1$. Continuing, we have

$$e_{n+2} < e_n\left(\frac{(\sqrt{a} - 1)^2}{a - 1}\right) = e_n\left(\frac{\sqrt{a} - 1}{\sqrt{a} + 1}\right) = ce_n, \quad 0 < c < 1$$

This tells us how to express e_{n+2} in terms of e_n. We can use this same inequality to express e_{2n+2} in terms e_2 and e_{2n+1} in terms of e_1:

$$e_{2n+2} < ce_{2n} < c(ce_{2n-2}) < \ldots < c^n e_2$$

$$e_{2n+1} < ce_{2n-1} < c(ce_{2n-3}) < \ldots < c^n e_1$$

And since $0 < c < 1$, this means that we can make c^n arbitrarily small by taking sufficiently large n; and $e_n < c^2[maxe_1, e_2]$, so we can make e_n arbitrarily small by taking sufficiently large n; and this shows that $\lim_{n\to\infty} e_n = 0$. Finally, remember that we defined e_n to be $d(x_n, \sqrt{a})$. So $\lim_{n\to\infty} d(x_n, \sqrt{a}) = 0$ which is sufficient to prove that $\lim_{n\to\infty}\{x_n\} = \sqrt{a}$.

Exercise 3.17d

As shown above, every two iterations reduces the error term by a factor of $(\sqrt{a}-1)/(\sqrt{a}+1)$ (linear convergence). For the algorithm in exercise 16, the nth iteration reduced the error term by a factor of 10^{-2^n} (quadratic convergence).

Exercise 3.18

Lemma 1 : $\{x_n\}$ is decreasing.

We're asked to choose that $x_1 > \sqrt[p]{a}$. And if $x_n > \sqrt[p]{a}$, we have

$\mapsto x_n > \sqrt[p]{a}$

$\to x_n^p > a$

$\to 0 > a - x_n^p$

$\to px_n^p > px_n^p + a - x_n^p$

$\to px_n^p > (p-1)x_n^p + a$

$\to \frac{px_n^p}{px_n^{p-1}} > \frac{(p-1)x_n^p + a}{px_n^{p-1}}$

$\to x_n > \frac{(p-1)x_n}{p} + \frac{a}{p}x_n^{p-1}$

$\to x_n > x_{n+1}$

Lemma 2 : $0 < k < 1 \to p(1-k) > 1 - k^p$

Let k be a positive number less than 1 and let p be a positive integer.

$\mapsto p = 1^0 + 1^2 + 1^3 + \ldots + 1^{p-1}$

$\to p > k + k^2 + k^3 + \ldots + k^{p-1}$

$\to p > \frac{(k-1)(k+k^2+k^3+\ldots+k^{p-1})}{k-1}$

$\to p > \frac{k^p - 1}{k-1}$

$\to p(k-1) > k^p - 1$

Lemma 3 : $x_n > \sqrt[p]{a}$

We know that $x_1 > \sqrt[p]{a}$ because we chose it. And if $x_n > \sqrt[p]{a}$, we have:

$\mapsto x_n > \sqrt[p]{a} > 0$

$\to 1 > \frac{\sqrt[p]{a}}{x_n} > 0$

$\to p\left(1 - \frac{\sqrt[p]{a}}{x_n}\right) > 1 - \left(\frac{\sqrt[p]{a}}{x_n}\right)^p$ from lemma 2

$\to p - 1 > p\frac{\sqrt[p]{a}}{x_n} - \frac{a}{x_n^p}$ expand and rearrange the terms

$\to x_n^p(p-1) > x_n^p\left(p\frac{\sqrt[p]{a}}{x_n} - \frac{a}{x_n^p}\right)$ multiply both sides by x_n^p

$\to px_n^p - x_n^p > p\sqrt[p]{a}x_n^{p-1} - a$

$\to (p-1)x_n^p + a > p\sqrt[p]{a}x_n^{p-1}$ rearrange the terms and simplify

$\to \frac{(p-1)x_n^p + a}{px_n^{p-1}} > \frac{p\sqrt[p]{a}x_n^{p-1}}{px_n^{p-1}}$ divide both sides by px_n^{p-1}

$\to \frac{(p-1)x_n}{p} + \frac{a}{px_n^{p-1}} > \sqrt[p]{a}$ simplify

$\to x_{n+1} > \sqrt[p]{a}$ definition of x_{n+1}

Finally : $\lim_{n\to\infty}\{x_n\} = \sqrt[p]{a}$

We've shown that $\{x_n\}$ is decreasing (lemma 1) and that it's bounded below (lemma 3), which is sufficient to show that it has some limit x^* with $x* \geq \sqrt[p]{a}$. From the definition of limit, we can find N such that $n > N \to d(x_n, x^*) < \frac{\epsilon}{2p}$. From this, we see that we can find x_n and x_{n+1} such that:

$$\mapsto d(x_n, x^*) < \epsilon \wedge d(x_{n+1}, x^*) < \tfrac{\epsilon}{2p}$$
$$\rightarrow d(x_n, x_{n+1}) < d(x_n, x^*) + d(x^*, x_{n+1}) = \tfrac{\epsilon}{p} \qquad \text{triangle inequality}$$
$$\rightarrow x_n - x_{n+1} < \tfrac{\epsilon}{p} \qquad\qquad\qquad\qquad \text{lemma 1: } x_n > x_{n+1}$$
$$\rightarrow x_n - \left[x_n - \tfrac{x_n}{p} + \tfrac{a}{px_n^{p-1}} \right] < \tfrac{\epsilon}{p} \qquad \text{definition of } x_{n+1}$$
$$\rightarrow \tfrac{x_n}{p} - \tfrac{a}{px_n^{p-1}} < \tfrac{\epsilon}{p}$$
$$\rightarrow \tfrac{x_n^p - a}{px_n^{p-1}} < \tfrac{\epsilon}{p}$$
$$\rightarrow x_n - \tfrac{a}{x_n^{p-1}} < \epsilon$$

This last statement tells us that $\lim_{n\to\infty} x_n - \tfrac{a}{x_n^{p-1}} = 0$, or that

$$\mapsto \lim_{n\to\infty} x_n - \tfrac{a}{x_n^{p-1}} = 0 \qquad\qquad \text{assumed}$$
$$\rightarrow \lim_{n\to\infty} x_n(1 - a/x_n^p) = 0 \qquad \text{theorem 3.3c}$$
$$\rightarrow x^* \left(1 - \left(\tfrac{\sqrt[p]{a}}{x^*} \right)^p \right) = 0 \qquad \text{theorem 3.3c}$$
$$\rightarrow \sqrt[p]{a}/x^* = 1$$
$$\rightarrow \sqrt[p]{a} = x^*$$

From the definition of x^* as the limit of $\{x_n\}$, this tells us that $\lim_{n\to\infty}\{x_n\} = \sqrt[p]{a}$ which is what we wanted to prove.

Exercise 3.19

The idea behind this proof is to consider the elements of the line segment $[0, 1]$ in the form of their base-3 expansion. When we do this, we notice that the mth iteration of the Cantor set eliminates every number with a 1 as the mth digit of its ternary decimal expansion.

From equation 3.24 in the book, we know that a real number r is not in the Cantor set iff it is in any interval of the form

$$\left(\frac{3k+1}{3^m}, \frac{3k+2}{3^m} \right) \qquad m, k \in \mathbb{N}$$

Let $\{\alpha_n\}$ represent an arbitrary ternary sequence (that is, each digit is either 0,1, or 2). We can determine the necessary and sufficient conditions for $x(\alpha)$ to fall into such an interval:

$$x(\alpha) \notin \text{Cantor} \ \ \text{iff} \ \ x(\alpha) \in \left(\frac{3k+1}{3^m}, \frac{3k+2}{3^m} \right)$$

$$\text{iff } (\exists m) \ \ \sum_{n=1}^{\infty} \frac{\alpha_n}{3^n} \in \left(\frac{3k+1}{3^m}, \frac{3k+2}{3^m} \right)$$

$$\text{iff } (\exists m) \ \ \sum_{n=1}^{\infty} \frac{\alpha_n}{3^{n-m}} \in (3k+1, 3k+2)$$

We then split up the summation into three distinct parts.

$$\text{iff } (\exists m) \ \ \sum_{n=1}^{m-1} \frac{\alpha_n}{3^{n-m}} + \sum_{n=m}^{m} \frac{\alpha_n}{3^{n-m}} + \sum_{n=m+1}^{\infty} \frac{\alpha_n}{3^{n-m}} \in (3k+1, 3k+2)$$

$$\text{iff } (\exists m) \ \ \sum_{n=1}^{m-1} \alpha_n 3^{m-n} + \alpha_m + \sum_{n=m+1}^{\infty} \frac{\alpha_n}{3^{n-m}} \in (3k+1, 3k+2)$$

Every term in the leftmost sum is divisible by three, so the leftmost sum is itself divisible by three. This gives us, for some $j \in \mathbb{N}$,

$$\text{iff } (\exists m) \ \ 3j + \alpha_m + \sum_{n=m+1}^{\infty} \frac{\alpha_n}{3^{n-m}} \in (3k+1, 3k+2), \ \ j \in \mathbb{N}$$

Without specifying a certain sequence $\{\alpha_n\}$, we can't evaluate the rightmost summation. But we can establish bounds for it. Each α_n is either 0,1, or 2. So the summation is largest whenever $\alpha_n = 2$ for all n, and it's smallest when $\alpha_n = 0$ for all n. This gives us the bound

$$0 = \sum_{n=m+1}^{\infty} \frac{0}{3^{n-m}} < \sum_{n=m+1}^{\infty} \frac{\alpha_n}{3^{n-m}} < \sum_{n=m+1}^{\infty} \frac{2}{3^{n-m}} = 2\sum_{n=1}^{\infty} \frac{1}{3^n} = 1$$

Note that the upper bound is 1, not 3, since the index is from 1 instead of 0. Continuing with our chain of "iff" statements, we conclude that $x(a)$ is not member of the Cantor set iff:

$$x(a) \notin \text{Cantor} \ \text{ iff } (\exists m) \quad 3j + \alpha_m + \delta \in (3k+1, 3k+2), \quad j \in \mathbb{N}, 0 < \delta < 1$$

From the bounds on α_m and δ, we know that $0 < \alpha_m + \delta < 3$: their sum is never a multiple of 3. So the only way that the set membership in the previous statement can be true is

$$\text{iff } (\exists m) \quad \alpha_m + \delta \in (1,2), \quad 0 < \delta < 1$$

$$\text{iff } (\exists m) \quad \alpha_m \in (0,2)$$

We know that α_m is an integer, and there's only one integer in the open interval $(0,2)$:

$$\text{iff } (\exists m) \quad \alpha_m = 1$$

$$\text{iff } \ x(\alpha) \notin x(a)$$

where $x(a)$ is as defined in the exercise. This shows that any real number $x(\alpha)$ is a member of the Cantor set if and only if it is a member of $x(a)$: this is sufficient to prove that the Cantor set is equal to $x(a)$.

Exercise 3.20

Choose some arbitrarily small ϵ. From the definition of Cauchy convergence, there is some N such that $i, n > N$ implies $d(p_i, p_n) < \epsilon/2$. From the convergence of the subsequence to p^*, there is some M such that $i > M$ implies $d(p^*, pi) < \epsilon/2$.

So, for $i, n > \max(M, N)$ we have

$$d(p^*, p_n) \leq d(p^*, p_i) + d(p_i, p_n) = \epsilon$$

which, because ϵ is arbitrarily small, is sufficient to prove that $\{p_n\}$ converges to p^*.

Exercise 3.21

We know that each $E_i \in \bigcap E_n$ is nonempty (see note 1), so we can construct at least one sequence whose ith element is an arbitrary element of E_i. Let $\{s_n\}$ be an arbitrary sequence constructed in this way. We can immediately say three things about this sequence.

1) The sequence is Cauchy. This comes from the fact that $\lim \text{diam } E_n = 0$: see the text below definition 3.9.

2) The sequence is convergent. This comes from the fact that it's a sequence in a complete metric space X: see definition 3.12.

3) The sequence converges to some $s* \in \bigcap E_n$. The set $\bigcap E_n$ is a intersection of closed sets, and is therefore closed (thm 2.24b), so any limit point of $\bigcap E_n$ is a point of $\bigcap E_n$, and therefore s^* (being a limit point of $\bigcap E_n$) is an element of $\bigcap E_n$.

This shows that $\bigcap E_n$ is nonempty: it contains at least one element s^*. We can then follow the proof of theorem 3.10b to conclude that $\bigcap E_n$ contains only one point.

note 1

We're told that $E_i \supset E_{i+1}$ for all i. If E_i were empty, then E_k would be empty for all $k > i$. This would mean that

$$\lim \operatorname{diam} E_n = \operatorname{diam}\varnothing = \operatorname{diam} \sup \{d(p,q) : p, q \in \varnothing\} = \sup \varnothing$$

To find the supremum of the empty set in $\mathbb{R}$, we need to rely on the definition of supremum:

$$\sup \varnothing = \text{least upper bound of } \varnothing \text{ in } \mathbb{R} = \min \{x \in \mathbb{R} : a \in \varnothing \to a \leq x\}$$

The set for which we're seeking a minimum contains every $x \in \mathbb{R}$ (because of false antecedent $a \in \varnothing$), so the supremum of the empty set in $\mathbb{R}$ is the minimum of $\mathbb{R}$ itself. Regardless of how we define this minimum (or if it's defined at all), it certainly isn't equal to zero. But we're told that

$$\lim \operatorname{diam} E_n = 0$$

so our initial assumption must have been wrong: there is no empty $E_i \in \bigcap E_n$.

Exercise 3.22

The set G_1 is an open set. Choose $p_1 \in G_1$ and find some open neighborhood $N_{r_1}(p_1) \subseteq G_1$. Choose a smaller neighborhood $N_{s_1}(p_1)$. Not only do we have $N_{s_1} \subset N_{r_1} \subseteq G_1$, we can take the closure of N_{s_1} and still have $\overline{N_{s_1}} \subset N_{r_1} \subseteq G_1$ (since $\delta \leq s_1 < r_1$). Define E_1 to be the neighborhood N_{s_1} (*not* its closure). At this point, we have

$$E_1 \subset \overline{E_1} \subseteq G_1$$

The set G_2 is dense in X, so p_1 is either an element of G_2 or is a limit point for G_2. In either case, every neighborhood of p_1 contains some point $p_2 \in G_2$. More specifically, the neighborhood E_1 contains some point $p_2 \in G_2$. Since p_2 is in both E_1 and G_2, both of which are open sets, there is some neighborhood $N_{r_2}(p_2)$ that's a subset of both E_1 and G_2. We now choose an even smaller radius $N_{s_2}(p_2)$. Not only do we have $N_{s_2} \subset N_{r_2} \subseteq G_2$ and $N_{s_2} \subset N_{r_2} \subseteq E_1$, we could take the closure $\overline{N_{s_2}}$. Define E_2 to be the neighborhood N_{s_2} (*not* its closure). At this point, we have

$$E_2 \subset \overline{E_2} \subset G_2, \quad E_2 \subset \overline{E_2} \subset E_1 \subset \overline{E_1}$$

We can continue in this same way, constructing a series of nested sets

$$\overline{E_1} \supset \overline{E_2} \supset \cdots \supset \overline{E_n}$$

This is a series of closed, bounded, nonempty, nested sets. So (from exercise 21) we know that $\bigcap \overline{E_n}$ contains a single point x. This single point x must be in every $\overline{E_i} \in \bigcap \overline{E_n}$, and $\overline{E_i} \subset G_i$, and x must be in every $G_i \in \bigcap G_n$. Therefore $\bigcap G_n$ is nonempty.

Exercise 3.23

We're aren't given enough information about the metric space X to assume that the Cauchy sequences converge to elements in X. All we can say is that, for any arbitrarily small ϵ, there exists some $M, N \in \mathbb{R}$ such that

$$n, m > M \to d(p_n, p_m) \leq \epsilon/2$$

$$n, m > N \to d(q_n, q_m) \leq \epsilon/2$$

From multiple applications of the triangle inequality, we have

$$d(p_n, q_n) \leq d(p_n, p_m) + d(p_m, q_n) \leq d(p_n, p_m) + d(p_m, q_m) + d(q_m, q_n)$$

which, for $n, m > \max\{N, M\}$, becomes

$$d(p_n, q_n) \leq \epsilon + d(p_m, q_m)$$

or

$$d(p_n, q_n) - d(p_m, q_m) \leq \epsilon$$

Exercise 3.24a

reflexivity

For all sequences $\{p_n\}$, we have

$$\lim_{n \to \infty} d(p_n, p_n) = \lim_{n \to \infty} |p_n - p_n| = \lim_{n \to \infty} 0 = 0$$

so $\{p_n\} = \{p_n\}$.

symmetry

If $\{p_n\} = \{q_n\}$, we have

$$\lim_{n \to \infty} d(p_n, q_n) = \lim_{n \to \infty} |p_n - q_n| = \lim_{n \to \infty} |q_n - p_n| = \lim_{n \to \infty} d(q_n, p_n)$$

so $\{q_n\} = \{p_n\}$.

transitivity

If $\{p_n\} = \{q_n\}$ and $\{q_n\} = \{r_n\}$, the triangle inequality gives us

$$\lim_{n \to \infty} d(p_n, r_n) \leq \lim_{n \to \infty} d(p_n, q_n) + d(q_n, r_n) = 0 + 0$$

This tells us that $\lim_{n \to \infty} d(p_n, r_n) = 0$, so $\{p_n\} = \{r_n\}$.

Exercise 3.24b

Let $\{a_n\} = \{p_n\}$ and let $\{b_n\} = \{q_n\}$. From the triangle inequality, we have

$$\Delta(P, Q) = \lim_{n \to \infty} d(p_n, q_n) \leq \lim_{n \to \infty} d(p_n, a_n) + d(a_n, b_n) + d(b_n, q_n)$$

Which, from the definition of equality established in part (a), gives us

$$\Delta(P, Q) = \lim_{n \to \infty} d(p_n, q_n) \leq \lim_{n \to \infty} d(a_n, b_n) \tag{5}$$

A similar application of the triangle inequality gives us

$$\lim_{n \to \infty} d(a_n, b_n) \leq \lim_{n \to \infty} d(a_n, p_n) + d(p_n, q_n) + d(q_n, b_n)$$

Which, from the definition of equality established in part (a), gives us

$$\lim_{n \to \infty} d(a_n, b_n) \leq \lim_{n \to \infty} d(p_n, q_n) \tag{6}$$

Combining equations (5) and (6), we have

$$\Delta(P, Q) = \lim_{n \to \infty} d(p_n, q_n) = \lim_{n \to \infty} d(a_n, b_n)$$

Exercise 3.24c

Define X^* to be the set of equivalence classes from part (b). Let $\{P_n\}$ be a Cauchy sequence in (X^*, Δ). This is an unusual metric space, so to be clear: The sequence $\{P_n\}$ is a sequence $\{P_0, P_1, P_2 \ldots\}$ of equivalence classes in X^* that get "closer" to one another with respect to distance function Δ. Each $P_i \in \{P_n\}$ is an equivalence class of Cauchy sequences of the form $\{p_{i0}, p_{i1}, p_{i2}, \ldots\}$ containing elements of X that get "closer together" with respect to distance function d.

Each $P_i \in \{P_n\}$ is a set of equivalent Cauchy sequences in X. From each of these, choose some sequence $\{p_{in}\} \in P_i$. For each i we have

$$(\exists N_i \in \mathbb{R}) \left(m, n > N_i \to d(p_{im}, p_{in}) < \frac{\epsilon}{i} \right)$$

We'll define a new sequence $\{q_n\}$ by letting $q_i = p_{im}$ for some $m > N_i$. Let Q be the equivalence class containing $\{q_n\}$. We can show that $\lim_{n\to\infty} P_n = Q$.

$$0 \le \lim_{n\to\infty} \Delta(P_n, Q) = \lim_{n\to\infty} \left[\lim_{k\to\infty} d(p_{nk}, q_k) \right] \le \lim_{n\to\infty} \frac{\epsilon}{n}$$

So, by the squeeze theorem, we have

$$\lim_{n\to\infty} \Delta(P_n, Q) = 0 \tag{7}$$

which, by the definition of equality in X^*, means that $\lim_{n\to\infty} P_n = Q$.

Proof that $Q \in X^*$:

Choose any arbitrarily small values $\epsilon_1, \epsilon_2, \epsilon_3$. From (7) we have

$$\exists X : n, k > X \to d(p_{nk}, q_n) < \epsilon_1$$

And $\{P_n\}$ is Cauchy, so we have

$$\exists Y : n > Y \to \Delta(P_n, P_{n+m}) < \epsilon_2$$

which, after choosing appropriately large n, gives us

$$\exists Z : k > Z \to d(p_{nk}, p_{(n+m)k}) < \epsilon_3$$

Therefore, by the triangle inequality:

$$d(q_n, q_{n+m}) \le d(q_n, p_{nk}) + d(p_{nk}, p_{(n+m)k}) + d(p_{(n+m)k}, q_{n+m})$$

Taking the limits of each side gives us

$$\lim_{n\to\infty} d(q_n, q_{n+m}) \le \epsilon_1 + \epsilon_3 + \epsilon_1$$

But these epsilon values were arbitrarily small and m was an arbitrary integer. This shows that for every ϵ there exists some integer N such that $n, n+m > N$ implies $d(q_n, q_{n+m}) < \epsilon$. And this is the definition of $\{q_n\}$ being a Cauchy sequence. Therefore $Q \in X^*$.

Exercise 3.24d

Let $\{p_n\}$ represent the sequence whose terms are all p, and let $\{q_n\}$ represent the sequence whose terms are all q.

$$\Delta(P_p, P_q) = \lim_{n\to\infty} d(p_n, q_n) = \lim_{n\to\infty} d(p, q) = d(p, q)$$

Exercise 3.24e : proof of density

Let $\{a_n\}$ be an arbitrary Cauchy sequence in X and let $A \in X^*$ be the equivalence class containing $\{a_n\}$. If the sequence $\{a_n\}$ converges to some element $x \in X$, then $A = P_x = \varphi(x)$ and therefore $A \in \varphi(X)$ (see exercise 3.24 for the definition of P_x).

If the sequence $\{a_n\}$ does not converge to some element $x \in X$, then choose an arbitrarily small ϵ. The sequence $\{a_n\}$ is Cauchy, so we're guaranteed the existence of K such that

$$j, k > K \to d(a_j, a_k) < \epsilon$$

From this, we can consider the sequence whose terms are all a_k. This sequence is a member of the equivalence class $P_{a_k} = \varphi(a_k)$ and

$$\Delta(A, P_{a_k}) = \lim_{n\to\infty} d(a_n, a_k) < \epsilon$$

This shows that we can find some element of $\varphi(X)$ to be arbitrarily close to A, which means that A is a limit point of $\varphi(X)$.

We have shown that an arbitrary Cauchy sequence $A \in X^*$ is either an element of $\varphi(X)$ or a limit point of $\varphi(X)$. By definition, this means that $\varphi(X)$ is dense in X^*.

Exercise 3.24e : if X is complete

If X is complete then every arbitrary Cauchy sequence $A \in X^*$ converges to some point $a \in X$, so that $A = P_a = \varphi(a) \in \varphi(X)$. This shows that $X^* \subseteq \varphi(X)$. And for every $\varphi(b) \in \varphi(X)$ there is some Cauchy sequence in X^* whose every element is b, so that $\varphi(b) = P_b \in X^*$. This shows that $\varphi(X) \subseteq X^*$.

This shows that $X^* \subseteq \varphi(X) \subseteq X^*$, or that $\varphi(X) = X^*$.

Exercise 3.25

The completion of the set of rational numbers is a set that's isomorphic to $\mathbb{R}$.

Exercise 4.1

Consider the function

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

This function satisfies the condition that

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

for all x, but the function is not continuous at $x = 0$: We can choose $\epsilon < 1$, and every neighborhood $N_\delta(0)$ will contain a point p for which

$$d(f(p), f(0)) = 1$$

Therefore we can't pick δ such that $d(p, 0) < \delta \to d(f(p), f(0)) < 1$, which means that f is not continuous by definition 4.5.

Exercise 4.2

Let X be a metric space, let E be an arbitrary subset of X, and let $\overline{E}$ represent the closure of E. We want to prove that $f(\overline{E}) \subseteq \overline{f(E)}$. To do this, assume $y \in f(\overline{E})$. This means that $y = f(e)$ for some $e \in (E \cup E')$.

case 1: $e \in E$

If $e \in E$, then $y \in f(E)$ and therefore $y \in \overline{f(E)}$.

case 2: $e \in E'$

If $e \in E'$, then every neighborhood of e contains infinitely many points of E. Choose an arbitrarily small neighborhood $N_\epsilon(y)$. We're told that f is continuous so, by definition 4.1, we're guaranteed the existence of δ such that $f(x) \in N_\epsilon(y)$ whenever $x \in N_\delta(e)$. But there are infinitely many elements of E in the neighborhood $N_\delta(e)$, so there are infinitely many elements of $f(E)$ in $N_\epsilon(y)$. This means that y is a limit point of $f(E)$.

We've shown that every arbitrary element $y \in f(\overline{E})$ is either a member of $f(E)$ or a limit point of $f(E)$, which means that $y \in (f(E) \cup f(E)') = \overline{f(E)}$. This proves that $f(\overline{E}) \subseteq \overline{f(E)}$.

A function f for which $f(\overline{E})$ is a proper subset $\overline{f(E)}$

Let X be the metric space consisting of the interval $(0, 1)$ with the standard distance metric. Let Y be the metric space $\mathbb{R}^1$. Define the function $f : X \to Y$ as $f(x) = x$. The interval $(0, 1)$ is closed in X but open in Y, so we have

$$f(\overline{X}) = f(X) = (0, 1) \neq \overline{(0, 1)}$$

Exercise 4.3

If we consider the image of $Z(f)$ under f, we have $f(Z(f)) = \{0\}$. This range is a finite set, and is therefore a closed set. By the corollary of theorem 4.8, we know that $f^{-1}(\{0\}) = Z(f)$ must also be a closed set.

Exercise 4.4: $f(E)$ is dense in $f(X)$

To show that $f(E)$ is dense in $f(X)$ we must show that every element of $f(X)$ is either an element of $f(E)$ or a limit point of $f(E)$.

Assume $y \in f(X)$. Then $p = f^{-1}(y) \in X$. We're told that E is dense in X, so either $p \in E$ or $p \in E'$.

case 1: $p \in E$

If $p \in E$, then $y = f(p) \in f(E)$.

case 2: $f^{-1}(y) \in E'$

If p is a limit point of E, then there is a sequence $\{e_n\}$ of elements of E such that $e_n \neq p$ and $\lim_{n \to \infty} e_n = p$. We're told that f is continuous, so by theorem 4.2 we know that $\lim_{n \to \infty} f(e_n) = f(p) = y$. Using definition 4.2 again, we know that there is a sequence $\{f(e_n)\}$ of elements of $f(E)$ From theorem 4.2, this tells us that $\lim_{x \to p} f(x) = f(p) = y$. Therefore y is a limit point of $f(E)$.

We've shown that every element $y \in f(X)$ is either an element of $f(E)$ or a limit point of $f(E)$. By definition, this means that $f(X)$ is dense in $f(E)$.

Exercise 4.4b

Choose an arbitrary $p \in X$. We're told E is dense in X, so p is either an element of E or a limit point of E.

Case 1: $p \in E$

If $p \in E$, then we're told that $f(p) = g(p)$.

Case 2: $p \in E'$

If p is a limit point of E, then there is a sequence $\{e_n\}$ of elements of E such that $e_n \neq p$ and $\lim_{n \to \infty} e_n = p$. We're told that f and g are continuous, so by theorem 4.2 we know that $\lim_{n \to \infty} f(e_n) = f(p)$ and $\lim_{n \to \infty} g(e_n) = g(p)$. But each e_n is an element of E, so we have $f(e_n) = g(e_n)$ for all n. This tells us that

$$g(p) = \lim_{n \to \infty} g(e_n) = \lim_{n \to \infty} f(e_n) = f(p)$$

We see that $f(p) = g(p)$ in either case. This proves that $f(p) = g(p)$ for all $p \in X$.

Exercise 4.5

The set E^C can be formed from an at most countable number of disjoint open intervals

We're told that E is closed, so E^C is open. Exercise 2.29 tells us that E^C contains an at-most countable number of disjoint segments. Each of these segments must be open (if any of them contained a non-interior point, it would be a non-interior point of E^C, but open sets have no non-interior points). Let $\{(a_n, b_n)\}$ be the at-most countable collection of disjoint open segments.

constructing the function

We must separately consider the cases for $x \in E$ and $x \notin E$; and if $x \notin E$ we must consider the possibility that E contains an interval of the form $(-\infty, b)$ or (a, ∞). We'll define the function to be

$$g(x) = \begin{cases} f(x), & x \in E \\ f(b_i), & x \in (a_i, b_i) \wedge a_i = -\infty \\ f(a_i), & x \in (a_i, b_i) \wedge b_i = \infty \\ f(a_i) + \frac{f(b_i) - f(a_i)}{b_i - a_i}(x - a_i), & x \in (a_i, b_i) \wedge -\infty < a_i < b_i < \infty \end{cases} \tag{8}$$

This function is the one mentioned in the hint: the graph of g is a straight line on each closed interval $[b_i, a_{i+1}] \in E^C$. This function can easily (albeit tediously) be shown to be continuous on $\mathbb{R}^1$.

failure if "closed" is omitted

Consider the following function defined on the open set $(-\infty, 0) \cap (0, \infty)$:

$$f(x) = \begin{cases} 1, & x > 0 \\ -1 & x < 0 \end{cases}$$

This function will be discontinuous at $x = 0$ no matter how we define the function at this point.

Extending this to vector-valued functions

Let E be a closed subset of $\mathbb{R}$ and let $f : E \to \mathbb{R}^k$ be a vector-valued function defined by

$$f(x) = (f_1(x), f_2(x), \ldots, f_k(x))$$

For each f_i we can define a function g_i as in equation (8) to create a new vector-valued function $g : \mathbb{R} \to \mathbb{R}^\daleth$ defined by

$$g(x) = (g_1(x), g_2(x), \ldots, g_k(x))$$

We've shown that each of these g_i functions are continuous, therefore by theorem 4.10 the function $g(x)$ is continuous.

Exercise 4.6

Let G represent the graph of f. Let $g : E \to G$ be defined as as $g(x) = (x, f(x))$. We can make G into a metric space by defining a distance function: the set G is a subset of $\mathbb{R} \times \mathbb{R}$, so the natural choice is to use the metric $d(g(x), g(y)) = d((x, f(x)), (y, f(y)) = \sqrt{(x-y)^2 + (f(x) - f(y))^2}$. This choice allows us to treat G as a subset of $\mathbb{R}^2$.

If f is continuous

Both $x \mapsto x$ and $x \mapsto f(x)$ are continuous mappings, so $g(x, f(x))$ is a continuous mapping by theorem 4.10. The domain of g is a compact metric space so by theorem 4.14 (or theorem 4.15) G is compact.

If the graph is compact

Define g as above. The inverse of g is $g^{-1}(x, f(x)) = x$. It's clear that g^{-1} is a one-to-one and onto function from G to E. Therefore by theorem 4.17 the inverse of g^{-1} – that is, g itself – is continuous. We can then appeal to theorem 4.10 once again to conclude that f is continuous.

Exercise 4.7

f is bounded

If $x = 0$, then $f(x, y) = 0$ for any value of y. If $x > 0$:

$$
\begin{aligned}
& (x - y^2)^2 \geq 0 && \text{squares are positive} \\
\to\ & x^2 - 2xy^2 + y^4 \geq 0 && \text{expanding the squared term} \\
\to\ & x^2 - xy^2 + y^4 \geq 0 && \text{LHS remains positive after adding the nonnegative term } (xy^2) \\
\to\ & x^2 + y^4 \geq xy^2 && \text{add } xy^2 \text{ to both sides} \\
\to\ & 1 \geq \frac{xy^2}{x^2+y^4} && \text{divide both sides by positive term } x^2 + y^4
\end{aligned}
$$

If $x < 0$:

$$
\begin{aligned}
& (x + y^2)^2 \geq 0 && \text{squares are positive} \\
\to\ & x^2 + 2xy^2 + y^4 \geq 0 && \text{expanding the squared term} \\
\to\ & x^2 - xy^2 + y^4 \geq 0 && \text{LHS remains positive after adding the nonnegative term } (-3xy^2) \\
\to\ & x^2 + y^4 \geq xy^2 && \text{add } xy^2 \text{ to both sides} \\
\to\ & 1 \geq \frac{xy^2}{x^2+y^4} && \text{divide both sides by positive term } x^2 + y^4
\end{aligned}
$$

g is unbounded near (0,0)

If we let $x = n^\alpha$ and let $y = n^\beta$, we have

$$g(n^\alpha, n^\beta) = \frac{n^{\alpha + 2\beta}}{n^{2\alpha} + n^{6\beta}}$$

We can divide the numerator and denominator by $n^{\alpha + 2\beta}$ to get

$$g(n^\alpha, n^\beta) = \frac{1}{n^{\alpha - 2\beta} + n^{4\beta - \alpha}}$$

If we let $\alpha = -3, \beta = -1$ this becomes

$$g\left(\frac{1}{n^3}, \frac{1}{n}\right) = \frac{1}{n^{-1} + n^{-1}} = \frac{n}{2}$$

Taking limits, we have

$$g(0,0) = \lim_{n \to \infty} g\left(\frac{1}{n^3}, \frac{1}{n}\right) = \lim_{n \to \infty} \frac{n}{2}$$

The rightmost limit is $+\infty$.

f is not continuous at (0,0)

Choose $0 < \delta < 1/2$. For f to be continuous we must be able to choose some $\epsilon > 0$ such that

$$d\left((0,0), (x,y)\right) < \epsilon \to d\left(f(0,0), f(x,y)\right) < \delta$$

But there can be no such epsilon. To see this, choose an arbitrary $\epsilon > 0$, choose $x > 0$ such that $x^2 + x < \epsilon^2$, and then choose $y = \sqrt{x}$. This gives us

$$d\left((0,0), (x,y)\right) = \sqrt{x^2 + y^2} = \sqrt{x^2 + x} < \epsilon$$

but

$$d\left(f(0,0), f(x,y)\right) = \frac{xy^2}{x^2 + y^4} = \frac{x^2}{2x^2} = \frac{1}{2}$$

The restriction of f to a straight line is continuous

Any straight line that doesn't pass through $(0,0)$ doesn't encounter any of the irregularities that occur at the origin; it's trivial but tedious to show that the restriction of f to such a line is continuous. So we need only consider lines that pass through the origin: that is, lines of the form $y = cx$ or $x = 0$ for some constant c.

For the line $y = cx$: let $\epsilon > 0$ be given and let $\delta = \epsilon/c^2$. Choose x such that $0 < x < \delta$. Then:

$$d(f(0,0), f(x,y)) = \left|\frac{xy^2}{x^2 + y^4}\right| = \left|\frac{c^2 x^3}{x^2 + c^4 x^4}\right| = \left|\frac{c^2 x}{1 + c^4 x^2}\right| < \left|\frac{c^2 x}{1}\right| < \left|c^2 \delta\right| = \epsilon$$

And ϵ was arbitrary, so this proves that f is continuous on this line.

For the line $x = 0$: let $\epsilon > 0$ be given choose any $y \neq 0$. Regardless of our choice of δ, we have

$$d(f(0,0), f(x,y)) = \left|\frac{xy^2}{x^2 + y^4}\right| = \left|\frac{0}{y^4}\right| = 0 < \epsilon$$

And ϵ was arbitrary, so this proves that f is continuous on this line.

The restriction of g to a straight line is continuous

The proof for the continuity of the restriction of g is almost identical to that of the continuity of the restriction of f.

Exercise 4.8

To show that f is bounded, we need to show that there exists some M such that $|f(p)| < M$ for every $p \in \mathbb{R}^1$. Let p be an arbitrary element of $\mathbb{R}^1$. We're told that E is a bounded subset of $\mathbb{R}^1$, so we know that E has a lower bound α. Choose any $\epsilon > 0$. From the definition of uniform continuity there exists some $\delta > 0$ such that, for all $p, q \in \mathbb{R}^1$,

$$d(p, q) < \delta \rightarrow d(f(p), f(q)) < \epsilon$$

Let n be the smallest integer such that $n\delta > p - \alpha$ (which exists from the Archimedean property of the reals) and define $\gamma = (p - \alpha)/n$. This allows us to divide the interval (α, p) into n intervals of length γ, each of which is smaller than δ. We can then apply the triangle inequality multiple times:

$$
\begin{aligned}
d(f(\alpha), f(p)) \quad &\leq \quad d(f(\alpha), f(\alpha + \gamma)) + d(f(\alpha + \gamma), f(\alpha + 2\gamma)) + \ldots + d(f(\alpha + (n-1)\gamma), f(\alpha + n\gamma)) \\
&< \quad \epsilon + \epsilon + \ldots + \epsilon \quad (n \text{ terms}) \\
&= \quad n\epsilon
\end{aligned}
$$

This shows us that $|f(p)|$ is bounded by $|f(\alpha) \pm n\epsilon|$ for all $p \in \mathbb{R}^1$, so by definition 4.13 we know that f is bounded.

if E is not bounded

If E is not bounded below, then we can revise the previous proof using the upper bound β in place of the lower bound α. If E is not bounded above or below, then we will not be able to use the Archimedean property to find n such that $n\delta > p - (-\infty)$ and the proof fails. For an example of a real uniformly continuous function that is not bounded, we can simply look to $f(x) = x$.

Exercise 4.9

Definition 4.18 says that f is *uniformly continuous* on X if for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$d_X(p, q) < \delta \rightarrow d_Y(f(p), f(q)) < \epsilon \tag{9}$$

We want to show that this conditional statement is true iff for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$\text{diam } E < \delta \rightarrow \text{diam } f(E) < \epsilon \tag{10}$$

(10) **implies** (9)

Let equation (10) hold and suppose $d(p, q) < \delta$.

$d(p, q) < \delta$	assumed
$\rightarrow \text{diam } \{p, q\} < \delta$	we're letting E be the set containing only p and q.
$\rightarrow \text{diam } \{f(p), f(q)\} < \epsilon$	from (10)
$\rightarrow d(f(p), f(q)) < \epsilon$	definition of diameter

Therefore (9) holds.

(9) **implies** (10)

Let equation (9) hold, let E be a nonempty subset of X, and suppose diam $E < \delta$.

$\text{diam } E < \delta$	assumed
$\rightarrow (\forall p, q \in E)(d(p, q) < \delta)$	definition of diameter
$\rightarrow (\forall p, q \in E)(d(f(p), f(q)) < \epsilon)$	definition of diameter
$\rightarrow (\forall f(p), f(q) \in f(E))(d(f(p), f(q)) < \epsilon)$	$p \in E \rightarrow f(p) \in f(E)$
$\rightarrow \text{diam } f(E) < \epsilon$	definition of diameter

Therefore (10) holds.

Exercise 4.10

We want to prove the converse of theorem 4.19: that is, we want to prove that if f is not uniformly continuous on X then either X is not a compact metric space or f is not a continuous function.

If f is not uniformly continuous, then the converse of definition 4.18 tells us that there exists some ϵ such that, for every $\delta > 0$, we can find $p, q \in X$ such that $d(p,q) < \delta$ but $d(f(p), f(q)) \geq \epsilon$. We'll be contradicting this claim, so we'll restate it formally in a numbered equation:

$$(\exists \epsilon > 0)(\forall \delta > 0) \quad d(p,q) < \delta \wedge d(f(p), f(q)) \geq \epsilon \tag{11}$$

From the fact that we can find such p, q for every value of δ, we can set $\delta_n = \frac{1}{n}$ and construct two sequences $\{p_n\}$ and $\{q_n\}$ where $d(p_n, q_n) < \frac{1}{n}$ but $d(f(p_n), f(q_n)) \geq \epsilon$.

If the set X is compact (**supposition 1**) the sequence $\{p_n\}$ must have at least one subsequential limit p (theorem 3.6a or 2.37). And, since $d(p_n, q_n) < \frac{1}{n}$, the sequence $\{q_n\}$ must also have p as a subsequential limit since

$$d(q_n, p) \leq d(q_n, p_n) + d(p_n, p) < \frac{1}{n} + \epsilon$$

But then, from theorem 4.2, it must be the case that both $\{f(p_n)\}$ and $\{f(q_n)\}$ have subsequential limits of $f(p)$. If the function f is continuous without being uniformly continuous (**supposition 2**) this would imply that, for some sufficiently large n,

$$d(f(p_n), f(q_n)) \leq d(f(p_n), f(p)) + d(f(p), f(q_n)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

which contradicts (11).

We've established a contradiction from our initial assumption that f is uniformly continuous, so one of our suppositions must be incorrect: either X is not compact or f is not continuous. By converse, if X is compact and f is continuous then f is uniformly continuous. And this is what we were asked to prove.

Exercise 4.11

Choose an arbitrary $\epsilon > 0$. We're told that f is uniformly continuous, so

$$(\exists \delta > 0): \quad d(x_m, x_n) < \delta \rightarrow d(f(x_m), f(x_n)) < \epsilon$$

And we're told that $\{x_n\}$ is Cauchy, so

$$(\exists N): \quad n, m > N \rightarrow d(x_m, x_n) < \delta$$

Together, these two conditional statements tell us that, for any ϵ,

$$(\exists N): \quad n, m > N \rightarrow d(f(x_m), f(x_n)) < \epsilon$$

which, by definition, shows that $\{f(x_n)\}$ is Cauchy. See exercise 4.13 for the proof that uses this result.

Exercise 4.12

We're asked to prove that the composition of uniformly continuous functions is uniformly continuous. Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be uniformly continuous functions. Choose an arbitrary $\epsilon > 0$. Because g is uniformly continuous, we have

$$(\exists \alpha > 0): \quad d(y_1, y_2) < \alpha \rightarrow d(g(y_1), g(y_2)) < \epsilon$$

Because f is uniformly continuous, we have

$$(\exists \delta > 0): \quad d(x_1, x_2) < \delta \rightarrow d(f(x_1), f(x_2)) < \alpha$$

Since the elements $f(x_1), f(x_2)$ are elements of Y these two conditional statements together tell us that, for arbitrary $\epsilon > 0$,

$$(\exists \delta > 0): \quad d(x_1, x_2) < \delta \rightarrow d(f(x_1), f(x_2)) < \alpha \rightarrow d(g(f(x_1)), g(f(x_2))) < \epsilon$$

which, by definition means that the composite function $(g \circ f) : X \rightarrow Z$ is uniformly continuous.

Exercise 4.13 (Proof using exercise 4.11)

For every $p \in X$ it is either the case that $p \in E$ or $p \notin E$. If $p \notin E$, then p is a limit point of E (because of density) and therefore we can construct a sequence $\{e_n\}$ that converges to p. We now define the function $g : X \to Y$ as

$$g(p) = \begin{cases} \lim_{n \to \infty} f(p), & \text{if } p \in E \\ \lim_{n \to \infty} f(e_n), & \text{if } p \notin E \end{cases}$$

We need to prove that this is actually a well-defined function and that it's continuous.

The function g is well-defined

It's clear that every element of X is mapped to at least one element of Y, but it's not immediately clear that each element of X is mapped to *only* one element of Y. We need to show that if $\{p_n\}$ and $\{q_n\}$ are sequences in E that converge to p, then $\{f(p_n)\}$ and $\{f(q_n)\}$ both converge to the same element.

Let $\{p_n\}$ and $\{q_n\}$ be arbitrary sequences in E that converge to p. From the definition of convergence (or from exercise 4.11), we know that

$$(\forall \delta > 0)(\exists N): \quad n > N \to d(p_n, p) < \frac{\delta}{2}$$

$$(\forall \delta > 0)(\exists M): \quad m > M \to d(q_m, p) < \frac{\delta}{2}$$

We can then use the triangle inequality to show that

$$(\forall \delta > 0)(\exists M, N): \quad n > \max\{M, N\} \to d(p_n, q_n) \leq d(p_n, p) + d(p, q_n) < \delta \tag{12}$$

We know that the function f is continuous on E, so for any $\epsilon > 0$ there exists some $\delta > 0$ such that

$$d(p_n, q_n) < \delta \to d(f(p_n), f(q_n)) < \epsilon \tag{13}$$

Combining equations (12) and (13) we see that for any $\epsilon > 0$ we can find some integer N such that

$$n > N \to (d(p_n, q_n) < \delta) \to (d(f(p_n), f(q_n)) < \epsilon)$$

This tells us that $\{f(p_n)\}$ and $\{f(q_n)\}$ don't converge to different points, but without being able to conclude that Y is compact it's entirely possible that $\{f(p_n)\}$ and $\{f(q_n)\}$ are merely Cauchy without actually converging to any point in Y at all. The exercise asks us to assume that Y is $\mathbb{R}^1$, but we'll try for a more general solution and assume only that Y is a complete metric space (see definition 3.12). This allows us to conclude that $\{f(p_n)\}$ and $\{f(q_n)\}$ both converge to the same point of Y, and we call this point $g(p)$. Therefore the function $g(p)$ has a unique value for every $p \in X$.

The function g is continuous

$$d(p,q) < \delta/2 \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{assumed}$$

Let $\{p_n\}$ be a sequence that converges to p and let $\{q_n\}$ be a sequence that converges to q.

$\rightarrow d(p_n, q_n) \leq d(p_n, p) + d(p, q) + d(q, q_n)$

$\rightarrow d(p_n, q_n) < d(p_n, p) + \delta/2 + d(q, q_n) \qquad\qquad\qquad$ from our initial assumption

This holds true for any n, even when n is sufficiently large to make $d(q, q_n) < \delta/4, d(p, p_n) < \delta/4$ in which case we have

$\rightarrow d(p_n, q_n) < \delta$

We're told that f is continuous on E, so we can make $d(f(p_n), f(q_n))$ arbitrarily small if we can make $d(p_n, q_n)$ arbitrarily small. And we *can* make δ arbitrarily small, so for any ϵ we can choose p_n, q_n such that

$\rightarrow d(f(p_n), f(q_n)) < \epsilon/2$

By definition, we have $g(p) = f(p)$ for $p \in E$, so

$\rightarrow d(g(p_n), g(q_n)) < \epsilon/2$

$\rightarrow d(g(p), g(q)) \leq d(g(p), g(p_n)) + d(g(p_n), g(q_n)) + d(g(q_n), g(q)) \quad$ triangle inequality

$\rightarrow d(g(p), g(q)) < d(g(p), g(p_n)) + \epsilon/2 + d(g(q_n), g(q)) \qquad$ established bound for $d(g(p_n), g(q_n))$

This holds true for any n, even when n is sufficiently large to make $d(q, q_n) < \epsilon/4, d(p, p_n) < \epsilon/4$ (we know that $\{g(p_n)\}$ converges to $g(p)$ and that $\{g(q_n)\}$ converges to $g(q)$ because we defined $g(p), g(q)$ specifically so that this would be true). For such values of n we have

$$d(g(p), g(q)) < \epsilon$$

This shows that we can constrain the range of g by constraining its domain, and therefore g is continuous.

Could we replace the range with $\mathbb{R}^k$ or any compact metric space? By any complete metric space?

Yes. Definition 3.12 tells us that all compact metric spaces and all Euclidean metric spaces are also complete metric spaces. Our proof didn't depend on the range of g being $\mathbb{R}^1$: we assumed only that the range was a complete metric space.

Could we replace the range with any metric space?

No. Consider the function $f(x) = \frac{1}{x}$ with a domain of $E = \left\{ \frac{1}{n} : n \in \mathbb{N} \right\}$. The set E is dense in $X = E \cup \{0\}$. This function is continuous at every $x \in E$ (we can find a neighborhood $N_\delta(x)$ that contains only x, which guarantees that $d(f(y), f(x)) = 0$ for any y in $N_\delta(x)$). However, there is no way to define $f(0)$ to make this a continuous function. This can be seen intuitively by recognizing that the range of $f(x) = 1/x$ is unbounded in every neighborhood of $x = 0$. If you're satisfied with an "intuitive proof", we're done. Otherwise, buckle up and get ready for a few more paragraphs of inequalities and Greek letters.

To prove that f is discontinuous at $x = 0$, choose any arbitrary neighborhood around 0 with radius δ. We can find an integer n such that $n > \frac{1}{\delta}$ (Archimedian property of the reals) so that $d(\frac{1}{n}, 0) < \delta$. For any positive integer k we also have $d(\frac{1}{n+k}, 0) < \delta$. Now, by the triangle inequality, we have

$$d\left[f\left(\frac{1}{n} \right), f\left(\frac{1}{n+k} \right) \right] \leq d\left[f\left(\frac{1}{n} \right), f(0) \right] + d\left[f(0), f\left(\frac{1}{n+k} \right) \right]$$

We haven't specificied a value for $f(0)$ yet, but we do know that $|f(1/n) - f(1/n+k)| = |n - (n+k)| = k$. This last inequality is therefore equivalent to

$$k \leq d\left[f\left(\frac{1}{n} \right), f(0) \right] + d\left[f(0), f\left(\frac{1}{n+k} \right) \right]$$

Every term of this inequality is nonnegative, so one of the two terms on the right-hand side of this inequality must be $\geq k/2$. But this is true for arbitrarily large k and arbitrarily small δ.

We can now show that f is not continuous at $x = 0$. Choose any $\epsilon > 0$. For all possible choice of $\delta > 0$ we can find integers $n > \frac{1}{\delta}$ and $k > 2\epsilon$, so that by the previous method we can be sure that either $d(f(\frac{1}{n}), f(0)) \geq k/2 = \epsilon$ or $d(f(\frac{1}{n+k}), f(0)) \geq k/2 > \epsilon$. By definition 4.5 of "continuous", this proves that f is not continuous at 0. And we never specified the value of $f(0)$, so we've proven that f will be discontinuous however we define $f(0)$. This means that there is no continuous extension from E to X.

Exercise 4.14 proof 1

If $f(0) = 0$ or $f(1) = 1$, we're done. Otherwise, we know that $f(0) > 0$ and $f(1) < 1$.

We'll now inductively define a sequence of intervals $\{[a_i, b_i]\}$ as follows[3]: Let $[a_0, b_0] = [0, 1]$. Now assume that we have defined $\{[a_i, b_i]\}$ for $i = 0, 1, \ldots, n$. Let $m = (a_n + b_n)/2$. We define $[a_{n+1}, b_{n+1}]$ as

$$[a_n, m], \quad \text{if } f(m) \leq m$$
$$[m, b_n], \quad \text{if } f(m) \geq m$$

(This procedure is underdefined for the case of $f(m) = m$, but when $f(m) = m$ we've found the point we're looking for anyway). The interval $[a_{n+1}, b_{n+1}]$ has diameter $2^{-(n+1)}$ with $f(a_{n+1}) \geq a_{n+1}, f(b_{n+1}) \leq b_{n+1}$. We therefore have a nested sequence of compact sets $\{[a_n, b_n]\}$ whose diameter converges to zero, so by exercise 3.21 we know that $\bigcap\{[a_n, b_n]\}$ consists of just one point $i \in I$. Our method of constructing $\{[a_n, b_n]\}$ guarantees that $f(i) \leq i$ and $f(i) \geq i$, therefore $f(i) = i$ and this is the point we were asked to find.

Exercise 4.14 proof 2

This proof, which is much clearer and more concise than mine, was taken from the homework of Men-Gen Tsai (b89902089@ntu.edu.tw).

Define a new function $g(x) = f(x) - x$. Both $f(x)$ and x are continuous, so $g(x)$ is continuous by theorem 4.9. If $g(0) = 0$ or $g(1) = 0$, then we're done. Otherwise, we have $g(0) > 0$ and $g(1) < 0$. Therefore by theorem 4.23 there must be some intermediate point in the interval $[0, 1]$ for which $g(x) = 0$: at this point, $f(x) = x$.

Exercise 4.15

Although the details of this proof might be ugly, the general idea is simple. If the function f is continuous but not monotonic then we can find some open interval (x_1, x_2) on which $f(x)$ obtains a local maximum or minimum. This point $f(x)$ is not an interior point of $f((x_1, x_2))$, so $f((x_1, x_2))$ is not an open set.

Let $f : \mathbb{R}^1 \to \mathbb{R}^1$ be a continuous mapping and assume that f is not monotonically increasing or decreasing. Then we can find $x_1 < x_2 < x_3$ such that $f(x_2) > f(x_1)$ and $f(x_2) > f(x_3)$, or such that $f(x_2) < f(x_1)$ and $f(x_2) < f(x_3)$. The proofs for either case are analogous, so we'll focus only on the case that $f(x_2) > f(x_1)$ and $f(x_2) > f(x_3)$.

We can't assume that the supremum of $f((x_1, x_3))$ occurs at $f(x_2)$: there could be some $x_{2.5}$ for which $f(x_{2.5}) > f(x_2)$. We want to construct a closed subinterval $[x_a, x_b] \subset (x_1, x_3)$ containing the x for which $f(x)$ obtains its local maximum. To do this, we define ϵ to be

$$\epsilon = \min\{d(f(x_2), f(x_1)), d(f(x_2), f(x_3))\}$$

Because f is continuous, we know that there exists some δ_a and δ_b such that

$$d(x_1, x) < \delta_a \to d(f(x_1), f(x)) < \epsilon$$

$$d(x_3, x) < \delta_b \to d(f(x_3), f(x)) < \epsilon$$

_____

[3]Our method of defining this sequence is similar to Newton's method or the Bisection method of root finding.

We can now let $x_a = x_1 + \frac{1}{2}\delta_a$ and $x_b = x_3 - \frac{1}{2}\delta_b$. We know that f doesn't obtain its local maximum for any x in the intervals (x_1, x_a) or (x_b, x_3) because all of these xs are sufficiently close (within δ) to x_1 or x_3 so that $f(x)$ is less (by at least $\frac{1}{2}\delta$) than $f(x_2)$. And $[x_a, x_b]$ is a closed subset of $\mathbb{R}^1$ and is therefore a compact set, so we know that $f([x_a, x_b])$ is a closed and bounded subset of $\mathbb{R}^1$ (theorem 4.15), and therefore $f([x_a, x_b])$ contains its supremum (which may or may not occur at x_2, but we only care that the supremum exists for some $x \in [x_a, x_b]$). This is all getting a bit muddled, so to recap our results so far:

1) $[x_a, x_b] \subset (x_1, x_3)$

2) $f([x_a, x_b]) \subset f((x_1, x_3))$

3) $(x \in (x_1, x_a) \cup (x_b, x_3)) \to f(x) < f(x_2)$

4) $f(x)$ has a local maximum for some $x \in [x_a, x_b]$

Together these four facts tell us that $f(\ (x_1, x_3)\ )$ contains its own supremum for some $x \in [x_a, x_b]$. But the supremum of a set can't be an interior point of the set, so $f(\ (x_1, x_3)\ )$ is not open even though (x_1, x_3) is open, so f is not an open mapping (theorem 4.8) We've shown that "f is not monotonic" implies "f is not a continuous open mapping from $\mathbb{R}^1$ to $\mathbb{R}^1$". By converse, then, we have shown that if f is a continuous open mapping from $\mathbb{R}^1$ to $\mathbb{R}^1$ then f is monotonic. And this is what we were asked to prove.

Exercise 4.16

The functions $[x]$ and (x) are both discontinuous for each $x \in \mathbb{N}$. If $x \in \mathbb{N}$ then, for every $0 < \delta < \frac{1}{2}$, we have

$$d([x - \delta], [x]) = |(x - 1) - x| = 1$$

$$d((x - \delta), (x)) = |(1 - \delta) - 0| = 1 - \delta > \frac{1}{2}$$

This shows that we can't constrain the range of these functions by restraining the domain around $x \in \mathbb{N}$, which by definition means that the functions are discontinuous for integer values of x.

Note that the the sum of these discontinuous functions, $[x] + (x)$, is the continuous function $f(x) = x$.

Exercise 4.17

Define E as in the hint. To the three criteria listed in the exercise, we add a fourth:

(d) $a < q < x < r < b$

Lemma 1: For each $x \in E$ we can find p, q, r such that the four criteria are met

We're assuming that $f(x-) < f(x+)$, so by the density of $\mathbb{Q}$ in $\mathbb{R}$ (theorem 1.20b) we can find some rational p such that $f(x-) < p < f(x+)$. From our definition of the set E we know that $f(x-)$ exists for all $x \in E$, so we can find some neighborhood $N_\delta(x)$ containing a rational q that fulfills criteria (b). More formally,

$$(\exists q \in N_\delta(x) \cap (a, x))(a < q < x \land [q < t < x \to f(t) < p]$$

Similarly, the existence of $f(x+)$ tells us that there is some neighborhood $N_{\delta 2}(x)$ containing a rational r such that

$$(\exists r \in N_{\delta 2}(x) \cap (x, b))(x < r < b \land [x < t < r \to f(t) > p]$$

Therefore each $x \in E$ can be associated with at least one (p, q, r) rational triple with $q < x < r$. $\square$

Lemma 2: If a given (p, q, r) triple fulfills the criteria for x and y, then $x = y$

Suppose x, y are two elements of E that both meet the four criteria. Assume, without loss of generality, that $x < y$. The density of the reals guarantees that there exists some rational w such that $x < w < y$. But this means that $f(w) < p$ (by criteria (b) since $q < w < y$) and that $f(w) > p$ (by criteria (c) since $x < w < r$). Reversing the roles of x and y establishes the same results for the case of $x > y$. This is clearly a contradiction, so one of our assumptions must be wrong: for any given triple (p, q, r) it's not possible to find two distinct elements of E that fulfill all four criteria. Note that we haven't proven that each triple can be associated with a unique $x \in E$: we've only shown that each triple can be associated with *at most* one $x \in E$. $\square$

Lemma 3: E is at most countable

The set of rational triples is countable (theorem 2.13). Lemma 2 tells us that we can create a function from the set of triples onto E, so the cardinality of E is not more than the cardinality of the set of rational triples. Therefore the cardinality of E is at most countable. $\qquad\square$

Lemma 4: F is at most countable

If we define F to be the set of x for which $f(x+) < f(x-)$ we can prove that F is at most countable with only trivial modifications to the previous lemmas. $\qquad\square$

Lemma 5: G is at most countable

Define G to be the set of x for which $f(x+) = f(x-) = \alpha$ but $f(x) \neq \alpha$. Let x be an arbitrary element of G. If every neighborhood of x contained another element of G, we could construct a sequence $\{t_n\}$ for which $\{t_n\} \to x$ but $\{f(t_n)\} \nrightarrow \alpha$. But this would mean that either $f(x-)$ or $f(x+)$ wouldn't exist, which contradicts our definition of G. Therefore each x is an isolated point of G: we can find some radius δ such that $N_\delta(x) \cap G = x$ and $N_\delta(x)$ contains rational numbers q, r such that $q < x < r$.

Now consider all of the (q, r) pairs of rational numbers. Each of these pairs will either have one, more than one, or zero elements of G in the associated open interval (q, r). Let $\mathcal{G}$ be the set of rational pairs (q, r) for which there is exactly one element of g in the open interval (q, r). The set $\mathcal{G}$ is a subset of $\mathbb{Q} \times \mathbb{Q}$, so it's at most countable (theorem 2.13). And each $x \in G$ can be associated with at least one $(p, q) \in \mathcal{G}$, so we can create a function from $\mathcal{G}$ onto G. This proves that G is at most countable. $\qquad\square$

The sets E,F,and G exhaust all the different types of simple discontinuities and all of these sets are countable. Therefore their union is countable (theorem 2.12). And this is what we were asked to prove.

Exercise 4.18

f is continuous at every irrational point

Let x be irrational. Choose any $\epsilon > 0$ and let n be the smallest integer such that $n > \frac{1}{\epsilon}$. If we want to constrain the range of f to $f(x) \pm \epsilon$, we will need to find a neighborhood of x that contains no rational numbers in the set

$$\left\{ \frac{p}{q} : p, q \in \mathbb{Q}, q \leq n \right\}$$

For each $i \leq n$ the product xi is irrational, so there exists some integer m_i such that

$$m < xi < m + 1$$

which means that

$$\frac{m}{i} < x < \frac{m+1}{i}$$

So we can define a neighborhood around x with radius δ_i where

$$\delta_i = \min\left\{ d\left(x, \frac{m}{i}\right), d\left(x, \frac{m+1}{i}\right) \right\}$$

The neighborhood $(x - \delta_i, x + \delta_i)$ contains no rationals of the form $\frac{k}{i}$. So if we have constructed neighborhoods $\delta_1, \delta_2, \ldots, \delta_n$ we can let $\delta = \min\{\delta_i\}$ (this minimum exists because n is finite), and we will have constructed a neighborhood $(x - \delta, x + \delta)$ that contains no rational numbers with denominators $\leq n$. Therefore, for all rational numbers y, we have

$$d(x, y) < \delta \to d(f(x), f(y)) < \frac{1}{n} < \epsilon$$

and for irrational numbers y we have

$$d(x, y) < \delta \to d(f(x), f(y)) = 0 < \epsilon$$

Which, by definition, means that f is continuous at x.

f has a simple discontinuity at every rational point

Let x be a rational point. If we follow the previous method of constructing δ we immediately see that $f(x+) = f(x-) = 0$. When x is rational, though, we no longer have $f(x) = 0$ and therefore f has a simple discontinuity at x.

Exercise 4.19

Suppose that f is discontinous at some point p. By definition 4.5 of "continuity", we know that there exists some $\epsilon > 0$ such that, for every $\delta > 0$, we can find some x such that $d(x, p) < \delta$ but $d(f(x), f(p)) \geq \epsilon$. Therefore we are able construct a sequence $\{x_n\}$ where each term of $\{x_n\}$ gets closer to p even though every term of $\{f(x_n)\}$ differs from $f(p)$ by at least ϵ. More formally, we choose each element of $\{x_n\}$ so that $d(p, x_n) < \frac{1}{n}$ but $d(f(p), f(x_n)) \geq \epsilon$. By constructing the sequence in this way we see that $\lim x_n \to p$ but $d(f(x_n), f(p)) \geq \epsilon$ for all n.

Assume, without loss of generality, that $f(x_n) < f(p)$ for every term of $\{x_n\}$ (see note 1 for justification of WLOG). From the density property of the reals we can find a rational number r such that $f(x_n) < r < f(p)$. We're told that f has the intermediate value property, so for each x_n there is some y_n between x_n and p such that $f(y_n) = r$. This sequence $\{y_n\}$ converges to p (since the terms of $\{y_n\}$ are squeezed between the terms of $\{x_n\}$ and p) and $\{f(y_n)\}$ clearly converges to r (since $f(y_n) = r$ for all n). But the fact that $\{y_n\}$ converges to p means that p is a limit point of $\{y_n\}$, which means that p is a limit point of the closed set described in the exercise (the set of all x with $f(x) = r$). Therefore p is an element of this set, which means that $f(p) = r$. And this is a contradiction since we specifically chose r so that $f(p) \neq r$.

We've established a contradiction, so our initial supposition must be false: f is not discontinous at any point p. Therefore f is continuous.

Note 1: justification for WLOG claim

For each term of the sequence $\{f(x_n)\}$ either $f(p) < f(x_n)$ or $f(x_n) < f(p)$. Therefore $\{f(x_n)\}$ either contains an infinite subsequence where $f(x_n) < f(p)$ for all n or $f(x_n) > f(p)$ for all n (or both). If it contains an infinite subsequence where $f(x_n) < f(p)$ for all n then we use this subsequence in place of $\{x_n\}$ in the proof. If it does not contain such a subsequence, then it contains an infinite subsequence where $f(x_n) > f(p)$ for all n: the proof for this case requires only the trivial twiddling of a few inequality signs. Therefore we can assume, without loss of generality, that $f(x_n) < f(p)$ for each term of $\{x_n\}$.

Exercise 4.20a

If $x \notin \overline{E}$ then x is an element of the open set $\overline{E}^C$, therefore there is some radius ϵ such that $d(x, y) < \epsilon \to y \notin \overline{E}$, and therefore $\inf_{z \in E} d(x, z) \geq \epsilon > 0$.

If $x \in \overline{E}$ then for every $\epsilon > 0$ we can find some $z \in E$ such that $d(x, z) < \epsilon$ and therefore $\inf_{z \in E} d(x, z) = 0$.

Exercise 4.20b

Case 1: $x, y \in E$

If x and y are both elements of E then from part (a) we have $\rho_E(x) = \rho_E(y) = 0$ and therefore

$$|\rho_E(x) - \rho_E(y)| = 0 \leq d(x, y)$$

Case 2: $x \in E, y \notin E$

If $y \in E$ but $x \notin E$ then $\rho_E(y) = 0$ and $d(x, y) \in \{d(x, z) : z \in E)\}$. Therefore

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

Case 3: $x, y \notin E$

If neither x nor y are elements of E then, for arbitrary $z \in E$ we have

$$\rho_E(x) \le d(x, z) \le d(x, y) + d(y, z)$$

This must hold for any choice of z. By choosing z to make $d(y, z)$ arbitrarily close to $\inf_{z \in E} d(x, z)$ we have

$$\rho_E(x) \le d(x, y) + \rho_E(y) + \epsilon$$

Our choice of ϵ can be made arbitrarily small (possibly even zero, if y is a limit point of E), so this is equivalent to

$$\rho_E(x) \le d(x, y) + \rho_E(y) \qquad \rightarrow \qquad \rho_E(x) - \rho_E(y) \le d(x, y)$$

By changing the roles of x and y we can similarly show that

$$\rho_E(y) - \rho_E(x) \le d(x, y)$$

Together, these last two inequalities show us that

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

These three cases exhaust the possibilities for x, y. This shows us that for every $\epsilon > 0$ we have $d(\rho_E(x), \rho_E(y)) < \epsilon$ whenever $d(x, y) < \epsilon$. And this is just definition 4.18 for uniform continuity with $\delta = \epsilon$.

Exercise 4.21

K is compact and ρ_F is continuous on K, so $\rho_F(K)$ is compact (and is therefore both closed and bounded from theorem 2.41). The results of exercise 20 tell us that 0 is not an element of $\rho_F(K)$, and therefore 0 is not a limit point of $\rho_F(K)$ (since $\rho_F(K)$ is closed). Therefore we can find some neighborhood around 0 with radius δ such that none of the elements in the interval $(0 - \delta, 0 + \delta)$ are in $\rho_F(K)$.

Choose $p \in K, q \in F$. The distance $d(p, q)$ is an element of $\rho_F(K)$ and therefore $d(p, q) \notin (0 - \delta, 0 + \delta)$. Therefore $d(p, q) > \delta$.

The conclusion fails if neither set is compact

Consider the sets[4]

$$K = \left\{ n + \frac{1}{n} : n \in \mathbb{N} \right\}$$

$$F = \{ n : n \in \mathbb{N} \} - \{2\}$$

The set F is closed, neither K nor F is compact, and we can choose $p \in K, q \in F$ such that $d(p, q) < \frac{1}{n}$ for any n.

Exercise 4.22

Exercise 20a shows us that $\rho_A(p) = 0$ iff $p \in A$ and $\rho_B(p) = 0$ iff $p \in B$, so it's clear that $f(p) = 0$ iff $p \in A$ and $f(p) = 1$ iff $p \in B$. This means that the $\rho_A(p) + \rho_B(p)$ is never zero, so f is continuous from theorem 4.9.

The intervals $[0, \frac{1}{2})$ and $(\frac{1}{2}, 1]$ are open in $[0, 1]$; therefore V and W are open by theorem 4.8. The sets V and W are clearly disjoint since $f(p)$ can have only a single value for any given p. From the definition of A and B we have

$$A = f^{-1}(0) \subseteq f^{-1}\left(\left[0, \frac{1}{2} \right) \right) = V$$

$$B = f^{-1}(1) \subseteq f^{-1}\left(\left(\frac{1}{2}, 1 \right] \right) = W$$

---

[4]This counterexample was taken from the homework of Men-Gen Tsai, b89902089@ntu.edu.tw

Exercise 4.23a: proof of continuity

Choose $x \in (a, b)$. Assume without loss of generality that $f(x) \neq f(a)$ and $f(x) \neq f(b)$ (see below for justification of WLOG). Choose some $\epsilon > 0$, and choose ϵ' such that

$$0 < \epsilon' < \min\{\epsilon, |f(b) - f(x)|, |f(a) - f(x)|\}$$

Choose δ such that

$$0 < \delta < \min\left\{\frac{\epsilon'(b - x)}{|f(b) - f(x)|}, \frac{\epsilon'(x - a)}{|f(a) - f(x)|}\right\}$$

Define λ_a, λ_b as

$$\lambda_a = 1 - \frac{\delta}{x - a}$$

$$\lambda_b = 1 - \frac{\delta}{b - x}$$

The method we used to select ϵ' and δ ensure that both of these λ values are in the interval $(0, 1)$. Notice that we can express δ in terms of λ_a or λ_b as

$$\delta = (1 - \lambda_a)(x - a) = (1 - \lambda_b)(b - x)$$

If we restrict the domain of f to $(x - \delta, x + \delta)$ we have

$$
\begin{aligned}
&|f(x + \delta) - f(x)| \\
&= |f(x + (1 - \lambda_b)(b - x)) - f(x)| && \text{from our definition of } \delta \\
&= |f(\lambda_b x + (1 - \lambda_b)b) - f(x)| && \text{algebraic rearrangement} \\
&\leq |\lambda_b f(x) + (1 - \lambda_b)f(b) - f(x)| && \text{from the convexity of } f \\
&= |(1 - \lambda_b)(f(b) - f(x))| && \text{algebraic rearrangement} \\
&= |\tfrac{\delta}{b-x}(f(b) - f(x))| && \text{from our definition of } \lambda \\
&< \tfrac{\epsilon'(b-x)}{(b-x)|f(b)-f(x)|}|(f(b) - f(x))| && \text{from our definition of } \delta \\
&= \epsilon' && \text{algebra} \\
&< \epsilon && \text{from our definition of } \epsilon
\end{aligned}
$$

and also

$$
\begin{aligned}
&|f(x - \delta) - f(x)| \\
&= |f(x + (1 - \lambda_a)(x - a)) - f(x)| && \text{from our definition of } \delta \\
&= |f(\lambda_a x + (1 - \lambda_a)a) - f(x)| && \text{algebraic rearrangement} \\
&\leq |\lambda_a f(x) + (1 - \lambda_a)f(a) - f(x)| && \text{from the convexity of } f \\
&= |(1 - \lambda_a)(f(a) - f(x))| && \text{algebraic rearrangement} \\
&= |\tfrac{\delta}{x-a}(f(a) - f(x))| && \text{from our definition of } \lambda \\
&< \tfrac{\epsilon'(x-a)}{(x-a)|f(a)-f(x)|}|(f(a) - f(x))| && \text{from our definition of } \delta \\
&= \epsilon' && \text{algebra} \\
&< \epsilon && \text{from our definition of } \epsilon
\end{aligned}
$$

We chose an arbitrarily small ϵ and showed how to find a nonzero upper bound for δ such that $d(x, y) < \delta \rightarrow d(f(x), f(y)) < \epsilon$. By definition, this means that f is continuous.

Justifying "without loss of generality": special case of $f(a) = f(x)$ or $f(b) = f(x)$

Let x be an arbitrary point in the interval (a, b). Choose $\alpha \in (a, x)$ such that $f(\alpha) \neq f(x)$: if this is not possible, then choose $\alpha = x$. Choose $\beta \in (x, b)$ such that $f(\beta) \neq f(x)$: if this is not possible, then choose $\beta = x$.

If $\alpha < x < \beta$, we can use the previous method to prove that f is continuous on an interval (α, β) containing x such that $(\alpha, \beta) \subseteq (a, b)$. Since x was an arbitrary element of (a, b) this is sufficient to prove that f is continuous on (a, b).

If $\alpha = x < \beta$, we can use the previous method to prove that f is continuous on an interval (x, β) containing x such that $(x, \beta) \subseteq (a, b)$. We only have $\alpha = x$ if there was no $\alpha \in (a, x)$ such that $f(\alpha) \neq f(x)$: this can only happen if f is constant (and therefore continuous) on the interval $(a, x]$. Therefore f is continuous on the interval (a, β) containing our arbitrary x; this is sufficient to prove that f is continuous on (a, b).

If $\alpha < x = \beta$, we can use the previous method to prove that f is continuous on an interval (α, x) containing x such that $(\alpha, x) \subseteq (a, b)$. We only have $\beta = x$ if there was no $\beta \in (x, b)$ such that $f(\beta) \neq f(x)$: this can only happen if f is constant (and therefore continuous) on the interval $[x, b)$. Therefore f is continuous on the interval (α, b) containing our arbitrary x; this is sufficient to prove that f is continuous on (a, b).

If $\alpha = x = \beta$ then f is a constant function on the entire interval (a, b) and is therefore trivially continuous.

Exercise 4.23b: increasing convex functions of convex functions are convex

Let g be an increasing convex function, let f be a convex function, and define their composite to be $h = g \circ f$.

$$
\begin{aligned}
& f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) && \text{from convexity of } f \\
\rightarrow\ & g(f(\lambda x + (1 - \lambda)y)) \leq g(\lambda f(x) + (1 - \lambda)f(y)) && g \text{ is an increasing function} \\
\rightarrow\ & g(f(\lambda x + (1 - \lambda)y)) \leq g(\lambda f(x) + (1 - \lambda)f(y)) \leq \lambda g(f(x)) + (1 - \lambda)g(f(y)) && \text{from convexity of } g \\
\rightarrow\ & g(f(\lambda x + (1 - \lambda)y)) \leq \lambda g(f(x)) + (1 - \lambda)g(f(y)) && \text{transitivity} \\
\rightarrow\ & h(\lambda x + (1 - \lambda)y) \leq \lambda h(x) + (1 - \lambda)h(y) && \text{definition of } h
\end{aligned}
$$

x and y were arbitrary, so this shows that h is a convex function.

Exercise 4.23c: the ugly-looking inequality

We're given that $s < t < u$, so we know that $0 < t - s < u - s$. Define λ to be

$$
\lambda = \frac{t - s}{u - s}
$$

From this definition we immediately see that $0 < \lambda < 1$. We can also see that

$$
(1 - \lambda) = 1 - \frac{t - s}{u - s} = \frac{u - t}{u - s}
$$

From this, we can establish an inequality:

$$
\begin{aligned}
& \lambda f(u) + (1 - \lambda)f(s) \geq f(\lambda u + (1 - \lambda)s) && \text{from convexity of } f \\
\rightarrow\ & \lambda f(u) + (1 - \lambda)f(s) \geq f\left(\frac{t - s}{u - s}u + \frac{u - t}{u - s}s\right) && \text{definition of } \lambda, (1 - \lambda) \\
\rightarrow\ & \lambda f(u) + (1 - \lambda)f(s) \geq f\left(\frac{ut - us + su - st}{u - s}\right) && \text{algebra} \\
\rightarrow\ & \lambda f(u) + (1 - \lambda)f(s) \geq f\left(\frac{t(u - s)}{u - s}\right) && \text{algebra} \\
\rightarrow\ & \lambda f(u) + (1 - \lambda)f(s) \geq f(t) && \text{algebra}
\end{aligned}
$$

From this last inequality, we can derive two results.

$$
\begin{aligned}
& \lambda f(u) + (1 - \lambda)f(s) \geq f(t) && \text{previously derived} \\
\rightarrow\ & \lambda f(u) - \lambda f(s) \geq f(t) - f(s) && \text{subtract } f(s) \text{ from each side} \\
\rightarrow\ & \frac{t - s}{u - s}(f(u) - \lambda f(s)) \geq f(t) - f(s) && \text{definition of } \lambda
\end{aligned}
$$

Dividing both sides of this last inequality by $t - s$ gives us

$$\frac{f(u) - f(s)}{u - s} \geq \frac{f(t) - f(s)}{t - s} \tag{14}$$

Similarly, we have:

$$
\begin{array}{lll}
& \lambda f(u) + (1 - \lambda)f(s) \geq f(t) & \text{previously derived} \\
\rightarrow & -\lambda f(u) - (1 - \lambda)f(s) \leq -f(t) & \text{multiply both sides by } -1 \\
\rightarrow & (1 - \lambda)f(u) - (1 - \lambda)f(s) \leq f(u) - f(t) & \text{add } f(u) \text{ to both sides} \\
\rightarrow & (1 - \lambda)(f(u) - f(s)) \leq f(u) - f(t) & \text{add } f(u) \text{ to both sides} \\
\rightarrow & \frac{u-t}{u-s}(f(u) - f(s)) \leq f(u) - f(t) & \text{definition of } \lambda
\end{array}
$$

Dividing both sides of this last equation by $u - t$ gives us

$$\frac{f(u) - f(s)}{u - s} \leq \frac{f(u) - f(t)}{u - t} \tag{15}$$

Combining (14) and (15) gives us

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(s)}{u - s} \leq \frac{f(u) - f(t)}{u - t}$$

which is what we were asked to prove.

Exercise 4.24

To prove that f is convex on (a, b) we must show that

$$(\forall x, y \in (a, b), \lambda \in [0, 1]) : f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \tag{16}$$

To do this, choose an arbitrary $x, y \in (a, b)$ (this choice of x and y will remain fixed for the majority of this proof). Let Λ represent the set of values for λ for which (16) holds. It's immediately clear that $0 \in \Lambda$ and $1 \in \Lambda$. We're also told that f has the property

$$f\left(\frac{x + y}{2}\right) \leq \frac{f(x) + f(y)}{2} \text{ for all } x, y \in (a, b) \tag{17}$$

which is just (16) with $\lambda = \frac{1}{2}$; therefore $\frac{1}{2} \in \Lambda$. To prove that f is convex we must show that $[0, 1] \subset \Lambda$.

Lemma 1: If $j, k \in \Lambda$ then $(j + k)/2 \in \Lambda$

Assume that $j, k \in \Lambda$ and let $m = \frac{j+k}{2}$. Proof that $m \in \Lambda$:

$$
\begin{array}{ll}
f(mx + (1 - m)y) = f\left(\frac{j+k}{2}x + \frac{2-j-k}{2}y\right) & \text{definition of } m \\
\quad = f\left(\frac{jx+kx+2y-jy-ky}{2}\right) & \text{algebra} \\
\quad = f\left(\frac{jx-jy+y}{2} + \frac{kx-ky+y}{2}\right) & \text{algebra} \\
\quad = f\left(\frac{jx+(1-j)y}{2} + \frac{kx+(1-k)y}{2}\right) & \text{algebra}
\end{array}
$$

j is in Λ, so $j \in [0, 1]$, so $(jx + (1 - j)y) \in (a, b)$. The same is true for k. So we can apply (17).

$$
\begin{array}{ll}
\quad \leq \frac{f(jx+(1-j)y)+f(kx+(1-k)y)}{2} & \\
\quad \leq \frac{jf(x)+(1-j)f(y)+kf(x)+(1-k)f(y)}{2} & \text{We can apply (16) because } j, k \in \Lambda \\
\quad = \frac{j+k}{2}f(x) + \frac{2-j-k}{2}f(y) & \text{algebra} \\
\quad = mf(x) + (1 - m)f(y) & \text{definition of } m
\end{array}
$$

This shows that $f(mx + (1 - m)y) \leq mf(x) + (1 - m)f(y)$, therefore $m \in \Lambda$. $\qquad \square$

Lemma 2: all rationals of the form $m/2^n$ with $0 \le m \le 2^n$ are members of Λ

This can be proven by induction. Let E be the set of all n for which the lemma is true. We know that $\{0, \frac{1}{2}, 1\} \subset \Lambda$ so we have $0 \in E, 1 \in E$. Now assume that E contains $1, 2, \ldots, k$. Choose an arbitrary m such that $0 \le m \le 2^{k+1}$.

case 1: m is even
If m is even then $m = 2\alpha$ for some $\alpha \in \mathbb{Z}, 0 \le \alpha \le 2^k$ and therefore $m/2^{k+1} = \alpha/2k$. By the hypothesis of induction $\alpha/2k \in \Lambda$ therefore $m/2^{k+1} \in \Lambda$.

case 2: m is odd
If m is odd then $(m-1)/2 = \alpha$ and $(m+1)/2 = \beta$ for some $\alpha, \beta \in \mathbb{Z}$ with $\alpha, \beta \in [0, 2^k]$. From this we have

$$\frac{m}{2^{k+1}} = \frac{m+1}{2^{k+2}} + \frac{m-1}{2^{k+2}} = \frac{1}{2}\left(\frac{\alpha}{2^k} + \frac{\beta}{2^k}\right)$$

By the hypothesis of induction $\alpha/2^k \in \Lambda$ and $\beta/2^k \in \Lambda$. Therefore, by lemma 1, $m/2^{k+1} \in \Lambda$.

This covers all possible cases for m, so $k + 1 \in E$. This completes the inductive step, therefore $E = \mathbb{N}$. $\quad\square$

Lemma 3: every element of $[0, 1]$ is a member of Λ

Choose any $p \in [0, 1]$. From lemma 2, we see that Λ is dense in $[0, 1]$ so we can construct a sequence $\{\lambda_n\}$ in Λ such that $\lim_{n \to \infty} \lambda_n = p$. Each λ_n is an element of Λ, so for each n we have

$$f(\lambda_n x + (1 - \lambda_n)y) \le \lambda_n f(x) + (1 - \lambda_n)f(y)$$

The function f is continuous, so by theorem 4.2 we can take the limit of both sides as $n \to \infty$ to get

$$f(px + (1 - p)y) \le pf(x) + (1 - p)f(y)$$

which means that $p \in \Lambda$. $\quad\square$

By lemma 3 we know that

$$(\forall \lambda \in [0, 1])f(\lambda x + (1 - \lambda)y) \le \lambda f(x) + (1 - \lambda)f(y)$$

But we chose x and y arbitrarily from the interval (a, b), so we have proven that

$$(\forall x, y \in (a, b), \lambda \in [0, 1])f(\lambda x + (1 - \lambda)y) \le \lambda f(x) + (1 - \lambda)f(y)$$

which is (16), the definition of convexity. This shows that f is convex, which is what we were asked to prove.

Exercise 4.25a

The "hint" given for this problem is actually a proof. The set $F = \mathbf{z} - C$ is clearly closed because $\mathbf{z}$ is a fixed element and C is closed. The sets K and F are disjoint:

$$
\begin{aligned}
&k \in K \cap F && \text{assumed} \\
\to\ &k \in K \wedge k \in F && \text{def. of intersection} \\
\to\ &k \in K \wedge (\exists c \in C)(k = z - c) && \text{def. of } F \\
\to\ &k \in K \wedge (\exists c \in C)(k + c = z) && \\
\to\ &(\exists k \in K, c \in C)k + c = z && \text{rearrangement of quantifiers} \\
\to\ &z \in K + C && \text{def. of } K + C
\end{aligned}
$$

From the definition of the set F, we see that $(\forall c \in C)(z - c \in F)$. And we've established that F is closed and that and K is compact, so by exercise 21 we know that there exists some $\delta > 0$ such that

$$
\begin{aligned}
& (\forall c \in C, k \in K)(d(z-c,k) > \delta) && \text{by exercise 21} \\
\to\ & (\forall c \in C, k \in K)(|z-c-k| > \delta) && \text{def. of distance function in } \mathbb{R}^k \\
\to\ & (\forall c \in C, k \in K)(|z-(c+k)| > \delta) && \text{algebra} \\
\to\ & (\forall c \in C, k \in K)(d(z,c+k) > \delta) && \text{def. of distance function in } \mathbb{R}^k
\end{aligned}
$$

The set of $c+k$ for all $c \in C, k \in K$ is simply the set $C+K$: so we have

$$
\begin{aligned}
\to\ & (\forall y \in C+K)(d(z,y) > \delta) && \text{def. of distance function in } \mathbb{R}^k
\end{aligned}
$$

This shows us that z is an interior point of $\overline{C+K}$. But z was an arbitrary point of $\overline{C+K}$, so we have shown that $\overline{C+K}$ is open. By theorem 2.23 the complement of an open set is closed, so $C+K$ is closed. And this is what we were asked to prove.

Exercise 4.25b

Let α be an irrational number and let $C_1 + C_2$ be defined as the set

$$
C_1 + C_2 = \{m + n\alpha : m, n \in \mathbb{N}\} \tag{18}
$$

We'll first prove that $C_1 + C_2$ is dense in $[0, 1)$ and then extend this proof out to prove that $C_1 + C_2$ is dense in $\mathbb{R}^1$. Define the set Δ to be the set of radii δ such that, for every $x \in [0, 1)$, every neighborhood of radius $> \delta$ contains a point of $C_1 + C_2$. More formally, we define it to be

$$
\Delta = \{\delta : (\forall x \in [0,1))(N_\delta(x) \cap C_1 + C_2 \neq \varnothing)\}
$$

We want to prove that $C_1 + C_2$ is dense in $[0, 1)$: we can do this by proving that Δ has a greatest lower bound of 0.

Lemma 1: Each element of $C_1 + C_2$ has a unique representation of the form $m + n\alpha$

Assume that $m + n\alpha$ and $p + q\alpha$ are two ways of describing the same element of $C_1 + C_2$.

$$
\begin{aligned}
& m + n\alpha = p + q\alpha && \text{assumption of equality} \\
\to\ & (m-p) + (n-q)\alpha = 0 && \text{algebra} \\
\to\ & (n-q)\alpha = (p-m) && \text{algebra}
\end{aligned}
$$

If both sides of this equation are zero, then $n = q$ and $p = m$ so that our two representations are not unique. If both sides are nonzero, we can divide by $p - m$:

$$
\begin{aligned}
\to\ & \alpha = \frac{p-m}{n-q} && \text{algebra}
\end{aligned}
$$

But m, n, p, q are integers so this last statement implies that α is a rational number. By contradiction, each element of $C_1 + C_2$ has a unique representation of the form $m + n\alpha$. $\qquad\square$

Lemma 2: For each $n \in \mathbb{N}$, there is exactly one $m \in \mathbb{N}$ such that $m + n\alpha \in [0, 1)$.

Assume that $p + n\alpha$ and $q + n\alpha$ are both in the interval $[0, 1)$ Then $|(p + n\alpha) - (q + n\alpha| = |p - q| \in [0, 1)$. And p and q are both integers, so it must be the case that $p = q$.

Lemma 3: If $d(x, y) = \delta$ for any $x, y \in [0, 1) \cap C_1 + C_2$ then $\delta \in \Delta$

Assume $d(x, y) = \delta$ for some $0 \leq x < y < 1$ with $x, y \in C_1 + C_2$. Then we have

$$
d(x,y) = y - x = p + m\alpha - q + n\alpha = (p-q) + (m-n)\alpha = \delta
$$

So that δ is itself an element of $C_1 + C_2$. It's also clear that any integer multiple of δ is an element of $C_1 + C_2$. Now, choose an arbitrary $p \in [0, 1)$ that is not a multiple of δ. Every real number lies between two integers, so there exists some a such that

$$
a < \frac{p}{\delta} < a + 1
$$

71

which implies

$$a\delta < p < (a+1)\delta$$

which shows that p is in a neighborhood of radius δ of some element of $C_1 + C_2$. But p was an arbitrary element of $[0, 1)$, so every element of $[0, 1)$ lies in such a neighborhood of radius δ, and therefore $\delta \in \Delta$. $\qquad\square$

Lemma 4: if $\delta \in \Delta$, then $\frac{1}{2}\delta \in \Delta$

Proof by contradiction. Assume that $\delta \in \Delta$ but $\frac{1}{2}\delta \notin \Delta$. By lemma 3 we know that $d(x, y) > \delta$ for all $x, y \in [0, 1) \cap C_1 + C_2$. This gives us a maximum size for the set $[0, 1) \cap C_1 + C_2$:

$$|[0, 1) \cap C_1 + C_2| \leq \frac{1}{\delta}$$

But this contradicts lemmas 1 and 2, which tell us that there is one unique element of $[0, 1) \cap C_1 + C_2$ for each $n \in \mathbb{N}$. By contradiction, then, we must have $\frac{1}{2}\delta \in \Delta$ whenever $\delta \in \Delta$. $\qquad\square$

The set $C_1 + C_2$ is dense in $[0, 1)$

We can now use induction to show that $\inf \Delta = 0$. We know that $1 \in \Delta$ because a neighborhood of radius $\delta = 1$ around any $x \in [0, 1)$ will contain $0 \in C_1 + C_2$. Therefore $1 \in \Delta$. Using induction with lemma 4 tells us that $(\forall n \in \mathbb{N})(2^{-n} \in \Delta)$. This is sufficient to prove that $\inf \Delta \leq 0$. Each element of Δ is a distance, so $\inf \Delta \geq 0$. Therefore $\inf \Delta = 0$.

By the definition of Δ, this means that every neighborhood of every element of $[0, 1)$ has an element of $C_1 + C_2$. This allows us to conclude that every element of $[0, 1)$ is a limit point of $C_1 + C_2$, which means that $C_1 + C_2$ is dense in $[0, 1)$. $\qquad\square$

The set $C_1 + C_2$ is dense in $\mathbb{R}^1$

Choose an arbitrary $p \in \mathbb{R}^1$. Let m be the integer such that $m \leq p < m+1$. This means that $0 \leq p - m < 1$, and we know that $C_1 + C_2$ is dense in $[0, 1)$. Therefore we can construct a sequence $\{c_n\}$ of elements of $[0, 1) \cap C_1 + C_2$ that converges to $p - m$. The definition of $C_1 + C_2$ guarantees that $\{c_n + m\}$ is also a sequence in $C_1 + C_2$, and theorem 3.3 tells us that $\{c_n + m\}$ converges to p. Therefore p is a limit point of $C_1 + C_2$. But p was an arbitrary element of $\mathbb{R}^1$, so every element of $\mathbb{R}^1$ is a limit point of $C_1 + C_2$. And this proves that $C_1 + C_2$ is dense in $\mathbb{R}^1$. $\qquad\square$

The sets C_1 and C_2 are closed, but $C_1 + C_2$ is not closed.

The sets C_1 and C_2 are closed because they have no limit points, so it's trivially true that they contain all of their limit points. The set $C_1 + C_2$ doesn't contain any non-integer rational numbers, but every real number is a limit point of $C_1 + C_2$. Therefore $C_1 + C_2$ doesn't contain all of its limit points which means it is not closed.

Exercise 4.26

We're told that Y is compact and that g is continuous and one-to-one. We conclude that $g(Y)$ is a compact subset of Z (theorem 4.14) and therefore g is uniformly continuous on Y (theorem 4.19). The fact that g is one-to-one tells us that g is one-to-one and onto $g(Y)$, so by theorem 4.17 we conclude that $g^{-1}(g(Y))$ is a continuous mapping from compact space $g(Y)$ to compact space X. So by theorem 4.19 we see that g^{-1} is uniformly continuous on $g(Y)$.

In exercise 12 we proved that the composition of uniformly continuous functions is uniformly continuous, therefore $f(x) = g^{-1}(h(x))$ is uniformly continuous if h is uniformly continuous. Theorem 4.7 tells us that the composition of continuous functions is continuous, therefore $f(x) = g^{-1}(h(x))$ is continuous if h is continuous.

To construct the counter example, define $X = Z = [0, 1]$ and $Y = [0, \frac{1}{2}) \cup [1, 2]$. Define the functions $f : X \to Y$ and $g : Y \to Z$ as

$$f(x) = \begin{cases} x & \text{if } x < \frac{1}{2} \\ 2x & \text{if } x \geq \frac{1}{2} \end{cases}$$

$$g(y) = \begin{cases} y & \text{if } y < \frac{1}{2} \\ \frac{y}{2} & \text{if } y \geq \frac{1}{2} \end{cases}$$

We can easily demonstrate that f fails to be continuous at $x = \frac{1}{2}$ and that g is continous at every point. The composite function $h : X \to Z$ is just $h(x) = x$, which is clearly continuous.

Exercise 5.1

Choose arbitrary elements $x, y \in \mathbb{R}^1$. We're told that f is continuous and that

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

Dividing the leftmost and rightmost terms by $|x - y|$ we have

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

Taking the limit of each side as $y \to x$ gives us

$$|f'(x)| \leq 0$$

But our choice of x was arbitrary, so the derivative of f is zero at every point. By theorem 5.11b, this proves that f is a constant function.

Exercise 5.2

Lemma 1: $g(f(x)) = x$

We're told that $f'(x) > 0$ for all x, so f is strictly increasing in (a, b) (this result is a trivial extension of the proofs for theorem 5.11). By definition, this means that $x < y \leftrightarrow f(x) < f(y)$ and $x > y \leftrightarrow f(x) > f(y)$: from this we conclude

$$x \neq y \leftrightarrow f(x) \neq f(y)$$

By contrapositive this is equivalent to

$$x = y \text{ iff } f(x) = f(y)$$

which, by definition, means that f is one-to-one. Therefore $g(f(x)) = x$ (theorem 4.17). $\qquad \square$

$g'(x)$ exists for all $x \in f(\ (a, b)\ )$

Using lemma 1 we see that $f'(x)$ is inversely proportional to $g'(x)$:

$$f'(x) = \lim_{t \to x} \frac{f(x) - f(t)}{x - t} = \frac{f(x) - f(t)}{g(f(x)) - g(f(t))} = \lim t \to x \frac{1}{\frac{g(f(x)) - g(f(t))}{f(x) - f(t)}} = \frac{1}{g'(x)}$$

We're told that the derivative $f'(x)$ is defined for all $x \in (a, b)$, so $1/g'(x)$ must also be defined.

bonus proofs: a big wad of properties for f and g

We can show that both f and g are uniformly continous, injective, differentiable, strictly increasing functions whose domains and ranges are compact.

From lemma 1, we know that f is a one-to-one function that is strictly increasing on (a, b). We're told that f is differentiable on (a, b), therefore f is continuous on $[a, b]$ (theorem 5.2). Because the domain of f is a compact metric space we can conclude that f is uniformly continuous (theorem 4.19). From the continuity and injectiveness of f we can conclude that g is continuous on $f(\ [a, b]\ )$ (theorem 4.17). The domain of g is the range of f, so the domain of g is a compact space (thoerem 4.14) and therefore g is uniformly continuous (theorem 4.19). From lemma 1 we know that $g(f(x)) = x$ for all x, therefore the inequalities in lemma 1 are equivalent to

$$g(f(x)) > g(f(y)) \leftrightarrow f(x) > f(y)$$
$$g(f(x)) < g(f(y)) \leftrightarrow f(x) < f(y)$$
$$g(f(x)) = g(f(y)) \leftrightarrow f(x) = f(y)$$

Which means that g is a one-to-one function and is strictly increasing on $f(\ (a, b)\ )$.

Exercise 5.3

Choose ϵ such that $|\epsilon| < \frac{1}{M}$. The function f is the sum of two differentiable functions, so by theorem 5.3 f is itself differentiable:

$$f'(x) = \lim_{t \to x} \frac{x - t}{x - t} + \lim_{t \to x} \frac{\epsilon g(x) - \epsilon g(t)}{x - t} = 1 + \epsilon g'(x)$$

From our bounds on ϵ and $g'(x)$ this gives us

$$1 - \left(\frac{1}{M} M \right) < f'(x) < 1 + \left(\frac{1}{M} M \right)$$

These are strict inequalities, so we conclude that $f'(x)$ is always positive for this choice of ϵ. Therefore f is an increasing function and is one-to-one (see lemma 1 of the previous exercise).

Exercise 5.4

Consider the function

$$f(x) = C_0 x + \frac{C_1 x^2}{2} + \cdots + \frac{C_n x^{n+1}}{n + 1}$$

When $x = 1$ this evaluates to the function given to us in the exercise, so $f(1) = 0$. When $x = 0$ every term evaluates to zero, so $f(0) = 0$. From example 5.4 we know that $f(x)$ is differentiable and its derivative is given by

$$f'(x) = C_0 + C_1 x + C_2 x^2 + \cdots + C_n x^n$$

By theorem 5.10, the fact that $f(0) = f(1) = 0$ means that $f'(x) = 0$ for some $x \in (0, 1)$. Therefore $\sum C_n x^n$ has a real root in $(0, 1)$, and this is what we were asked to prove.

Exercise 5.5

Choose an arbitrary $\epsilon > 0$. We're told that $f'(t) \to 0$ as $t \to \infty$, so there exists some N such that

$$t > N \;\; \to \;\; |f'(t)| < \epsilon$$

With some algebraic manipulation and the use of the the mean value theorem (theorem 5.10) we can express g as

$$g(x) = f(x + 1) - f(x) = \frac{f(x + 1) - f(x)}{(x + 1) - (x)} = f'(t) \;\; \text{for some } t \in (x, x + 1)$$

This must be true for all possible values of x, so choose $x > N$. We now have $t > x > N$, so the $f'(t)$ term in the previous equation is now less than ϵ.

$$|g(x)| = |f'(t)| \leq \epsilon$$

which means that $|g(x)| < \epsilon$ for all $x > N$. And ϵ was an arbitrary positive real, which by definition means that $g(x) \to 0$ as $x \to \infty$. And this is what we were asked to prove.

Exercise 5.6

The function g is differentiable (theorem 5.3c) and its derivative is

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

We want to prove that g is monotonically increasing. This is true iff $g'(x) > 0$ for all x, which is true iff

$$\frac{x f'(x) - f(x)}{x^2} > 0$$

which is true iff

$$x f'(x) > f(x)$$

which is true iff

$$f'(x) > \frac{f(x)}{x} \tag{19}$$

To show that (19) holds for all x, choose an arbitrary $x \in \mathbb{R}$. From the fact that $f(0) = 0$ we know from theorem 5.10 that

$$\frac{f(x)}{x} = \frac{f(x) - f(0)}{x - 0} = f'(c) \quad \text{for some } c \in (0, x)$$

We're told that f' is monotonically increasing, so $f'(x) > f'(c)$. Therefore:

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

Therefore (19) holds, and we've shown that this occurs iff $g'(x) > 0$, which means that g is monotonically increasing. And this is what we were asked to prove.

Exercise 5.7

From the fact that $f(x) = g(x) = 0$ we see that

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

Exercise 5.8

We're told that f' is a continuous function on the compact space $[a, b]$ therefore f' is uniformly continuous (theorem 4.19). Choose any $\epsilon > 0$: by the definition of uniform continuity there exists some δ such that $0 < |t - x| < \delta \rightarrow |f'(t) - f'(x)| < \epsilon$. Choose $t, x \in [a, b]$: by the mean value theorem there exists $c \in (t, x)$ such that

$$f'(c) = \frac{f(t) - f(x)}{t - x}$$

From the fact that $c \in (t, x)$ we know that $|c - x| < |t - x|\delta$. Therefore $|f'(c) - f'(x)| < \epsilon$. Therefore, from our definition of c, we have

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

Our initial choice of t, x, and ϵ was arbitrary, some $\delta > 0$ must exist so that this previous inequality is true for all t, x, and ϵ. And this is what we were asked to prove.

Does this hold for vector-valued functions?

Yes. Choose an arbitrary $\epsilon > 0$ and define the vector-valued function f to be

$$f(x) = (f_1(x), f_2(x), \dots, f_n(x))$$

Assume that f is differentiable on $[a, b]$. Then each f_i is differentiable on $[a, b]$ (remark 5.16) and $[a, b]$ is compact so by the preceeding proof we know that for each f_i there exists some $\delta_i > 0$ such that

$$|t - x| < \delta_i \quad \rightarrow \quad \left| \frac{f_i(t) - f_i(x)}{t - x} - f_i'(x) \right| < \frac{\epsilon}{n}$$

Define $\delta = \min\{\delta_1, \delta_2, \dots, \delta_n\}$. For $|t - x| < \delta$ we now have

$$
\begin{aligned}
\left| \frac{f(t) - f(x)}{t - x} - f'(x) \right| &= \left| \frac{(f_1(t) + f_2(t) + \cdots + f_n(t)) - (f_1(x) + f_2(x) + \cdots + f_n(x))}{t - x} - (f_1'(x) + f_2'(x) + \cdots + f_n'(x)) \right| \\
&= \left| \left(\frac{f_1(t) - f_1(x)}{t - x} - f_1'(x) \right) + \left(\frac{f_2(t) - f_2(x)}{t - x} - f_2'(x) \right) + \cdots + \left(\frac{f_n(t) - f_n(x)}{t - x} - f_n'(x) \right) \right| \\
&\leq \left| \left(\frac{f_1(t) - f_1(x)}{t - x} - f_1'(x) \right) \right| + \left| \left(\frac{f_2(t) - f_2(x)}{t - x} - f_2'(x) \right) \right| + \cdots + \left| \left(\frac{f_n(t) - f_n(x)}{t - x} - f_n'(x) \right) \right| \\
&< \frac{\epsilon}{n} + \frac{\epsilon}{n} + \cdots + \frac{\epsilon}{n} = \epsilon
\end{aligned}
$$

Exercise 5.9

We're asked to show that $f'(0)$ exists. From the definition of the derivative, we know that

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

The function f is continuous, so $\lim_{x \to 0} f(x) - f(0) = 0$ and $\lim_{x \to 0} x = 0$. Therefore we can use L'Hopital's rule.

$$f'(0) = \lim_{x \to 0} \frac{f(x) - f(0)}{x} = \lim_{x \to 0} f'(x) - 01 - 0 = \lim_{x \to 0} f'(x)$$

We're told that the right-hand limit exists and is equal to 3, therefore the leftmost term $(f'(0))$ exists and is equal to 3. And this is what we were asked to prove.

Exercise 5.9 : Alternate proof

If $\lim_{x \to 0} f'(0) = 3$ and $f'(0) \neq 3$, then f' would have a simple discontinuity at $x = 0$. Therefore $f'(0) = 3$ as an immediate consequence of the corollary to theorem 5.12.

Exercise 5.10

Let f_1, g_1 represent the real parts of the functions f, g and let f_2, g_2 represent their imaginary parts: that is, $f(x) = f_1(x) + if_2(x)$ and $g(x) = g_1(x) + ig_2(x)$. We're told that f and g are differentiable, therefore each of these dependent functions is differentiable (see Rudin's remark 5.16). Applying the hint given in the exercise, we have

$$\lim_{x \to 0} \frac{f(x)}{g(x)} = \lim_{x \to 0} \left\{ \frac{f_1(x) + if_2(x)}{x} - A \right\} \cdot \frac{x}{g_1(x) + ig_2(x)} + A \cdot \frac{x}{g_1(x) + ig_2(x)}$$

Each of these functions is differentiable and the denominators all tend to 0, so we can apply L'Hopital's rule.

$$
\begin{aligned}
\lim_{x \to 0} \frac{f(x)}{g(x)} &= \lim_{x \to 0} \left\{ \frac{f_1'(x) + if_2'(x)}{1} - A \right\} \cdot \frac{1}{g_1'(x) + ig_2'(x)} + A \cdot \frac{1}{g_1'(x) + ig_2'(x)} \\
&= \lim_{x \to 0} \left\{ \frac{f'(x)}{1} - A \right\} \cdot \frac{1}{g'(x)} + A \cdot \frac{1}{g'(x)} \\
&= \{A - A\} \cdot \frac{1}{B} + A \cdot \frac{1}{B} = \frac{A}{B}
\end{aligned}
$$

Exercise 5.11

The denominator of the given ratio tends to 0 as $h \to 0$, so we can use L'Hopital's rule (differentiating with respect to h):

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

$$= \lim_{h \to 0} \frac{1}{2} \left(\frac{f'(x + h) - f'(x)}{h} + \frac{f'(x) - f'(x - h)}{h} \right)$$

We're told that $f''(x)$ exists, so this limit exists and is equal to

$$= \lim_{h \to 0} \frac{1}{2} \left(f''(x) + f''(x) \right) = f''(x)$$

A function for which this limit exists although $f''(x)$ does not

For a counterexample, we need only find a differentiable function for which $f''(x) = 1$ when $x > 0$ and $f''(x) = -1$ when $x < 0$. These criteria are met by

$$f(x) = \begin{cases} x^2, & \text{if } x > 0 \\ -(x^2), & \text{if } x < 0 \\ 0, & \text{if } x = 0 \end{cases}$$

This function is continuous and differentiable, but $f''(x)$ does not exist at $x = 0$.

Exercise 5.12

From the definition of $f'(x)$, we have

$$f'(x) = \lim_{h \to 0} \frac{|x + h|^3 - |x|^3}{h} \qquad (20)$$

If $x > 0$ then the terms in the numerator are positive and (20) resolves to

$$f'(x) = \lim_{h \to 0} \frac{(x + h)^3 - (x)^3}{h} = \lim_{h \to 0} 3x^2 + 3xh + h^2 = 3x^2$$

If $x < 0$ then the terms in the numerator are negative and (20) resolves to

$$f'(x) = \lim_{h \to 0} \frac{-(|x| + h)^3 + (|x|)^3}{h} = \lim_{h \to 0} -(3|x|^2 + 3|x|h + h^2) = -(3|x^2|)$$

It's clear from the above results that $f'(x) \to 0$ as $x \to 0$, and this agrees with $f'(0)$:

$$f'(0) = \lim_{h \to 0} \frac{|h^3|}{h} = 0$$

So $f'(x) = 3x|x|$ for all x.

From the definition of $f''(x)$, we have

$$f''(x) = \lim_{h \to 0} \frac{|3(x + h)^2| - |3x^2|}{h} \qquad (21)$$

If $x > 0$ then the terms in the numerator are positive and (21) resolves to

$$f''(x) = \lim_{h \to 0} \frac{3(x + h)^2 - 3x^2}{h} = \lim_{h \to 0} 6x + 3h = 6x$$

If $x < 0$ then the terms in the numerator are negative and (21) resolves to

$$f''(x) = \lim_{h \to 0} \frac{-3(|x| + h)^2 - 3|x^2|}{h} = \lim_{h \to 0} 6|x| + 3h = 6|x|$$

It's clear from the above results that $f''(x) \to 0$ as $x \to 0$, and this agrees with $f''(0)$:

$$f''(0) = \lim_{h \to 0} \frac{|3h^2|}{h} = 0$$

So $f''(x) = |6x|$ for all x.

From the definition of $f'''(x)$, we have

$$f'''(x) = \lim_{h \to 0} \frac{|6(x + h)| - |6x|}{h} \qquad (22)$$

If $x = 0$, then when $h > 0$ we have

$$\lim_{h \to 0} \frac{|6h|}{h} = 6$$

and when $h < 0$ we have

$$\lim_{h \to 0} \frac{|6h|}{h} = -6$$

So the limit in (22) (which is $f^{(3)}(0)$) doesn't exist.

Exercise 5.13a

f is continuous when $x \neq 0$

The proof that x^a is continuous when $x \neq 0$ is trivial. The sin function hasn't been well-defined yet, but we can assume that it's a continuous function[5]. Therefore their product $x^a \sin(x^{-c})$ is continuous wherever it's defined (theorem 4.9), which is everywhere but $x = 0$.

---

[5]Example 5.6 says that we should assume without proof that sin is differentiable, so we can also assume that it's continuous (theorem 5.2).

f is continous at $x = 0$ if $a > 0$

We have $f(0) = 0$ by definition. For f to be continuous at $x = 0$ it must be the case that $\lim_{x\to 0} f(x) = 0$. The range of the sin function is $[-1, 1]$, so if $a > 0$ we have

$$x^a(-1) \le x^a \sin(x^{-c}) \le x^a(1)$$

Taking the limit of each of these terms as $x \to 0$ gives us

$$0 \le \lim_{x\to 0} x^a \sin(x^{-c}) \le 0$$

which shows that $\lim_{x\to 0} f(x) = 0$, and therefore f is continuous at $x = 0$.

f is discontinous at $x = 0$ if $a = 0$

To show that f is not continuous at $x = 0$, it's sufficient to construct a sequence $\{x_n\}$ such that $\lim x_n = 0$ but $\lim f(x_n) \ne f(0)$ (theorem 4.2). Define the terms of $\{x_n\}$ to be

$$x_n = \left(\frac{1}{2n\pi + \pi/2} \right)^{\frac{1}{c}}$$

This sequence clearly has a limit of 0, but

$$f(x_n) = x_n^0 \sin(x_n^{-c}) = \sin(2n\pi + \pi/2) = 1$$

so that $\lim\{f(x_n)\} = 1$. Note that we're making lots of unjustified assumptions about the sin function and the properties of the as-of-yet undefined symbol π.

f is discontinous at $x = 0$ if $a < 0$

Define the terms of $\{x_n\}$ to be

$$x_n = \left(\frac{1}{2n\pi + \pi/2} \right)^{\frac{1}{c}}$$

This sequence clearly has a limit of 0, but

$$f(x_n) = x_n^a \sin(x_n^{-c}) = \left(\frac{1}{2n\pi + \pi/2} \right)^{\frac{a}{c}} \sin(2n\pi + \pi/2) = \left(\frac{1}{2n\pi + \pi/2} \right)^{\frac{a}{c}}$$

We're told that $c > 0$, therefore $a/c < 0$ and we have

$$f(x_n) = \left(\frac{1}{2n\pi + \pi/2} \right)^{\frac{a}{c}} = (2n\pi + \pi/2)^{\frac{-a}{c}}$$

where $-a/c > 0$. By theorem 3.20a we see that $\lim f(x_n) = \infty$, which means that $\lim x_n = 0$ but $\lim f(x_n) \ne 0$ and therefore f is not continuous at $x = 0$.

These cases show that f is continous iff $a > 0$.

Exercise 5.13b

From the definition of limit we have

$$f'(0) = \lim_{h\to 0} \frac{f(0 + h) - f(0)}{h} = \lim_{h\to 0} \frac{h^a \sin(h^{-c}) - 0}{h} = \lim_{h\to 0} h^{a-1} \sin(h^{-c}) \tag{23}$$

We can evaluate the rightmost term by noting that sin is bounded by $[-1, 1]$ so that

$$|h^{a-1}|(-1) \le h^{a-1} \sin(h^{-c}) \le |h^{a-1}|(1) \tag{24}$$

Lemma 1: f is differentiable when $x \neq 0$

The proof that x^a and x^{-c} are differentiable when $x \neq 0$ is trivial. The sin function hasn't been well-defined yet, but Rudin asks us in example 5.6 to assume that it's differentiable. Therefore $\sin(x^{-c})$ is differentiable when $x \neq 0$ (theorem 5.5, the chain rule) and therefore $x^a \sin(x^{-c})$ is differentiable when $x = 0$ (theorem 5.3b, the product rule).

case 1: $f'(0)$ exists when $a > 1$

When $a > 1$ we have $a - 1 > 0$ and therefore taking the limits of (24) as $h \to 0$ gives us

$$0 \leq \lim_{h \to 0} |h^{a-1} \sin(h^{-c})| \leq 0$$

which means that (23) becomes

$$f'(0) = \lim_{h \to 0} h^{a-1} \sin(h^{-c}) = 0$$

which shows that $f'(0)$ is defined.

case 2: $f'(0)$ does not exist when $a = 1$

Define the sequences $\{h_n\}$ and $\{j_n\}$ such that

$$h_n = \left(\frac{1}{2n\pi + \frac{\pi}{2}} \right)^{1/c}$$

$$j_n = \left(\frac{1}{(2n+1)\pi + \frac{\pi}{2}} \right)^{1/c}$$

When $a = 1$ we have $a - 1 = 0$ and therefore equation (23) gives us

$$\lim_{h \to 0} h^{a-1} \sin(h^{-c}) = \lim_{n \to \infty} \sin(h_n^{-c}) = 1$$

$$\lim_{j \to 0} j^{a-1} \sin(j^{-c}) = \lim_{n \to \infty} \sin(j_n^{-c}) = -1$$

We know the sequences $\{f'(h_n)\}$ and $\{j'(h_n)\}$ are well-defined because of lemma 1, therefore we have conflicting definitions of $f'(0)$. This means that the limit in (23) (and therefore $f'(0)$ itself) does not exist.

case 3: $f'(0)$ does not exist when $a < 1$

Define the sequences $\{h_n\}$ and $\{j_n\}$ such that

$$h_n = \left(\frac{1}{2n\pi + \frac{\pi}{2}} \right)^{1/c}$$

$$j_n = \left(\frac{1}{(2n+1)\pi + \frac{\pi}{2}} \right)^{1/c}$$

When $a < 1$ we have $a - 1 < 0$ and therefore equation (23) gives us

$$\lim_{h \to 0} h^{a-1} \sin(h^{-c}) = \lim_{n \to \infty} h_n^{a-1} = \infty$$

$$\lim_{j \to 0} j^{a-1} \sin(j^{-c}) = \lim_{n \to \infty} -j_n^{a-1} = -\infty$$

We know the sequences $\{f'(h_n)\}$ and $\{j'(h_n)\}$ are well-defined because of lemma 1, therefore we have conflicting definitions of $f'(0)$. This means that the limit in (23) (and therefore $f'(0)$ itself) does not exist.

These cases show that $f'(0)$ exists iff $a > 1$.

Exercise 5.13c

Note that we've only defined f on the domain $[-1, 1]$, so we only need to show that f is bounded on this domain.

case 1: f' is unbounded when $a < 1$

We saw in case (3) of part (b) that f' is unbounded near 0 when $a < 1$.

case 2: f' is unbounded when $1 \leq a < c + 1$

By the lemma of part (b) we know that $f'(x)$ is defined for all $x \in [1, 1]$ except for possibly $x = 0$. By the chain rule and product rule we know that the derivative of f when $x \neq 0$ is

$$f'(x) = ax^{a-1} \sin(x^{-c}) - cx^{a-(c+1)} \cos(x^{-c}) \tag{25}$$

Define the sequence $\{h_n\}$ such that

$$h_n = (2n\pi)^{-1/c}$$

Evaluating the derivative in (25) at $x = h_n$ gives us

$$f'(h_n) = (2n\pi)^{\frac{1-a}{c}} \sin(2n\pi) - c(2n\pi)^{\frac{(1+c)-a}{c}} \cos(2n\pi) = -c(2n\pi)^{\frac{(1+c)-a}{c}}$$

We're assuming in this case that $a < c + 1$, so taking the limits of this last equation as $n \to \infty$ gives us

$$\lim_{n \to \infty} f'(h_n) = \lim_{n \to \infty} -c(2n\pi)^{\frac{(1+c)-a}{c}} = -\infty$$

This doesn't prove anything about $f'(0)$ itself, but it does show that $f'(x)$ is unbounded near 0.

case 3: f' is bounded when $a \geq c + 1$

If $a \geq c + 1$ then clearly $a > 1$, so by the lemma of part (b) we know that $f'(x)$ is defined for all $x \in [1, 1]$ including $x = 0$. By the chain rule and product rule we know that the derivative of f is

$$f'(x) = ax^{a-1} \sin(x^{-c}) - cx^{a-(c+1)} \cos(x^{-c}) \tag{26}$$

Since $f'(x)$ is defined for $x = 0$, we can take the limit of (26) as $x \to 0$:

$$f'(0) = \lim_{x \to 0} f'(x) = \lim_{x \to 0} ax^{a-1} \sin(x^{-c}) - cx^{a-(c+1)} \cos(x^{-c})$$

Because x is bounded by $[-1, 1]$ we know that $x^{a-1}, x^{a-(c+1)}, \sin$, and $\cos$ are all bounded by $[-1, 1]$. Therefore the rightmost limit of the previous equation is bounded by

$$-(a + c) \leq \lim_{x \to 0} ax^{a-1} \sin(x^{-c}) - cx^{a-(c+1)} \cos(x^{-c}) \leq a + c$$

Which, of course, means that $f'(x)$ is also bounded by $[-(a+c), a+c]$ for $x \in [-1, 1]$. We could find stricter bounds for $f'(x)$, but it's not necessary.

These three cases show that f' is bounded iff $a \geq c + 1$.

Exercise 5.13d

lemma: f' is continuous when $x \neq 0$

From lemma 1 of part (b) we know that f' exists for all $x \neq 0$ and its derivative is given by

$$f'(x) = ax^{a-1} \sin(x^{-c}) - cx^{a-(c+1)} \cos(x^{-c}) \tag{27}$$

Rudin asks us to assume that sin and cos are continuous functions and it's trivial to show that $x^{\pm \alpha}$ is continuous when $x \neq 0$ for any α, so we can use the chain rule (theorem 5.5) and product rule (theorem 5.3b) to show that f' is continuous when $x \neq 0$.

case 1: f' is continuous at $x = 0$ when $a > 1 + c$

We've shown that $f'(0) = 0$ when $a > 1$ (case 1 of part (b)). For f' to be continuous at $x = 0$ it must be the case that $\lim_{x \to 0} f'(x) = 0$. We can algebraically rearrange (27) to obtain

$$f'(x) = x^{a-(c+1)} ax^c \sin(x^{-c}) - c \cos(x^{-c})$$

The range of the cosine and sine functions are $[-1, 1]$, so we can establish a bound on this function.

$$|f'(x)| = |x^{a-(c+1)} ax^c \sin(x^{-c}) - c \cos(x^{-c})| \leq |x^{a-(c+1)}| \cdot |ax^c + c| \leq |x^{a-(c+1)}| \cdot (|ax^c| + |c|)$$

Because $a > c + 1$, we have $|x^{a-(c+1)}| \to 0$ and $|ax^c| \to 0$ as $x \to 0$. Taking the limits of this last inequality as $x \to 0$ therefore gives us

$$\lim_{x \to 0} |f'(x)| \leq 0 \cdot (0 + c) = 0$$

This shows that $\lim_{x \to 0} f'(x) = f'(0)$, therefore f' is continuous at $x = 0$.

case 2: f' is not continuous at $x = 0$ when $a = 1 + c$

To show that f' is not continuous at $x = 0$ it's sufficient to construct a sequence $\{x_n\}$ such that $\lim x_n = 0$ but $\lim f'(x_n) \neq f'(0) = 0$ (theorem 4.2). Define the terms of x_n to be

$$x_n = \left(\frac{1}{2n\pi} \right)^{\frac{1}{c}}$$

This sequence clearly has a limit of 0, but $\sin(x_n) = 0$ and $\cos(x_n) = 1$ so that the terms of $\{f'(x_n)\}$ are

$$f'(x_n) = ax_n(0) - cx^0(1) = -c$$

so that $\lim\{f'(x_n)\} = -c \neq f'(0)$.

case 3: f' is not continuous at $x = 0$ when $a < 1 + c$

If $a < 1+c$ we know that f' is not bounded on $[-1, 1]$ (part (c)) therefore f' is not continuous on $[-1, 1]$ (theorem 4.15). From lemma 1, we know that the discontinuity must occur at the point $x = 0$.

For an alternative proof we could use the sequence $\{x_n\}$ established in case 2 and show that $f'(x_n) \to \infty \neq f'(0)$ as $x_n \to 0$ when $a < 1 + c$.

Exercise 5.13e

Lemma 1: f' is differentiable when $x \neq 0$

We established in part (b) that f' exists for $x \neq 0$ and is given by

$$f'(x) = ax^{a-1} \sin(x^{-c}) - cx^{a-(c+1)} \cos(x^{-c}) \tag{28}$$

We know that all of the exponential powers of x are differentiable when $x \neq 0$ and Rudin asks us to assume that sin and cos are differentiable, so we can use theorem 5.5 (the chain rule) and theorem 5.3(the product rule) to show that f' is differentiable when $x \neq 0$.

case 1: $f''(0)$ exists when $a > 2 + c$

From the definition of limit we know that

$$f''(0) = \lim_{h \to 0} \frac{f'(0 + h) - f'(0)}{h} = \lim_{h \to 0} \frac{ah^{a-1} \sin(h^{-c}) - ch^{a-(c+1)} \cos(h^{-c})}{h}$$

$$= \lim_{h \to 0} \left[(ah^{a-2}) \sin(h^{-c}) - (ch^{a-(c+2)}) \cos(h^{-c}) \right] \tag{29}$$

81

The range of the sin and cos functions is $[-1, 1]$ so we can establish bounds for the limited term.

$$-(|ah^{a-2}| + |ch^{a-(c+2)}|) \le ah^{a-2}\sin(h^{-c}) - ch^{a-(c+2)}\cos(h^{-c}) \le |ah^{a-2}| + |ch^{a-(c+2)}|$$

When $a > (2+c) > 2$, the powers of h tend tend to zero as $h \to 0$. Taking the limits of the previous inequality as $h \to 0$, we have

$$0 \le \lim_{h \to 0} ah^{a-2}\sin(h) - ch^{a-(c+2)}\cos(h^{-c}) \le 0$$

This shows that the limit in (29) (and therefore $f''(0)$) exists.

case 2: $f''(0)$ does not exist when $a = 2 + c$

Define the sequences $\{h_n\}$ and $\{j_n\}$ such that

$$h_n = \left(\frac{1}{2n\pi}\right)^{1/c}$$

$$j_n = \left(\frac{1}{(2n+1)\pi}\right)^{1/c}$$

When $a = 2 + c$ we have $a - (c+2) = 0$ and therefore equation (28) gives us

$$\lim_{h \to 0} f''(h) = \lim_{n \to \infty} f''(h_n) = \left[0 - (ch_n^0)(1)\right] = -c$$

$$\lim_{j \to 0} f''(j) = \lim_{n \to \infty} f''(j_n) = \left[0 - (cj_n^0)(-1)\right] = c$$

We know the sequences $\{f''(h_n)\}$ and $\{j''(h_n)\}$ are well-defined because of lemma 1, therefore we have conflicting definitions of $f''(0)$. This means that the limit in (28) (and therefore $f''(0)$ itself) does not exist.

case 2: $f''(0)$ does not exist when $a < 2 + c$

Define the sequences $\{h_n\}$ and $\{j_n\}$ such that

$$h_n = \left(\frac{1}{2n\pi}\right)^{1/c}$$

$$j_n = \left(\frac{1}{(2n+1)\pi}\right)^{1/c}$$

When $a < 2 + c$ we have $a - (c+2) < 0$ and therefore $h_n^{a-(c+2)} \to \infty$. This means that equation (28) gives us

$$\lim_{h \to 0} f''(h) = \lim_{n \to \infty} f''(h_n) = \left[0 - (ch_n^{a-(c+2)})(1)\right] = -\infty$$

$$\lim_{j \to 0} f''(j) = \lim_{n \to \infty} f''(j_n) = \left[0 - (cj_n^{a-(c+2)})(-1)\right] = \infty$$

We know the sequences $\{f''(h_n)\}$ and $\{j''(h_n)\}$ are well-defined because of lemma 1, therefore we have conflicting definitions of $f''(0)$. This means that the limit in (28) (and therefore $f''(0)$ itself) does not exist.

These three cases show that $f''(0)$ exists iff $a > 2 + c$.

Exercise 5.13f

Lemma 1: to hell with this

By the lemma of part (e) we know that $f''(x)$ is defined for all $x \in [1, 1]$ except for possibly $x = 0$. By the chain rule and product rule we know that the derivative of f when $x \ne 0$ is

$$f''(x) = \left[(a^2 - a)x^{a-2} - c^2 x^{a-(2+2c)}\right]\sin(x^{-c}) + \left[(c^2 + c - ca)x^{a-(2+c)} - cax^{a-(1+c)}\right]\cos(x^{-c}) \quad (30)$$

And I'm not going to screw around with limits and absolute values of something so annoying to type out, so I'll conclude with "the proof is similar to that of part(c)".

Exercise 5.13g

The proof is similar to that of part (d), and the sentiment is similar to that of lemma (1) of part (f).

Exercise 5.14

Lemma 1: If f is not convex then the convexity condition fails for all λ for some $s, t \in (a, b)$

This could also be stated as "if f is not convex on (a, b) then it is strictly concave on some interval $(s, t) \subset (a, b)$". Assume that f is continuous on the interval (a, b) and that f is not convex. By definition, this means that we can choose $c, d \in (a, b)$ and $\lambda \in (0, 1)$ such that

$$f(\lambda c + (1 - \lambda)d) > \lambda f(c) + (1 - \lambda)f(d) \tag{31}$$

Having fixed our choice of c, d, and λ so that the previous equation is true, we define the function $g(x)$ to be

$$g(x) = f((x)c + (1 - x)d) - (x)f(c) + (1 - x)f(d)$$

We know that g is continuous on $[0, 1]$ because f is continuous on $[c, d] \subset (a, b)$ (theorem 4.9). And we know that there is at least one $p \in [0, 1]$ such that $g(p) > 0$ because we can choose $p = \lambda$ which causes the previous equation to simplify to

$$g(p) = f(\lambda c + (1 - \lambda)d) - [\lambda f(c) + (1 - \lambda)f(d)]$$

which is > 0 from (31). Let Z_1 be the set of all $p \in [0, \lambda)$ for which $g(p) = 0$ and let Z_2 be the set of all $p \in (\lambda, 1]$ for which $g(p) = 0$ It's immediately clear that these sets are nonempty since $g(0) = g(1) = 0$. These sets are closed (exercise 4.3) and therefore contain their supremums and infimums. Let $\alpha = \sup\{Z_1\}$ and let $\beta = \inf\{Z_2\}$. We can now claim that equation (31) holds for all $\lambda \in (\alpha, \beta)$. And this is the same as saying that

$$f(\lambda s + (1 - \lambda)t) > \lambda f(s) + (1 - \lambda)f(t)$$

for all $\lambda \in (0, 1)$ for

$$s = \alpha c + (1 - \alpha)d, \qquad t = \beta c + (1 - \beta)d$$

which, by definition, means that the convexity conditions fails on the inteval $(s, t) \in (a, b)$ for all $\lambda \in (0, 1)$.

The "if" case: f' is monotonically increasing if f is convex

Let f be a function that is convex on the interval (a, b) (see exercise 4.23) and choose $x, y \in (a, b)$ with $y \geq x$. By definition this means that for all $x, y \in (a, b)$ and for all $\lambda \in (0, 1)$ it must be the case that

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \tag{32}$$

We want to express the left-hand side of this inequality as $f(x + h)$: we can do this by defining h such that

$$h = (\lambda - 1)(x - y)$$

Rearranging this algebraically allows us to express λ and $1 - \lambda$ as

$$\lambda = 1 - \frac{h}{y - x}, \qquad (1 - \lambda) = \frac{h}{x - y}$$

Substituting these values of λ and $\lambda - 1$ into (32) results in

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

which is algebraically equivalent to

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

Equation (32) had to be true for any value of $\lambda \in (0, 1)$. As $\lambda \to 1$ we see that $h \to 0$. Taking the limit of both sides of the previous equation as $h \to 0$ gives us

$$f'(x) \le \frac{f(y) - f(x)}{y - x} \tag{33}$$

Having established (33), we now want to express the left-hand side of (32) as $f(y - h)$: we can do this by redefining h such that

$$h = -(\lambda)(x - y)$$

Rearranging this algebraically allows us to express λ and $1 - \lambda$ as

$$\lambda = \frac{h}{y - x}, \quad (1 - \lambda) = 1 - \frac{h}{x - y}$$

Substituting these values of λ and $1 - \lambda$ into (32) results in

$$f(y - h) \le \frac{h \, f(x)}{y - x} + f(y) - \frac{h \, f(y)}{y - x}$$

which is algebraically equivalent to

$$\frac{f(y) - f(y - h)}{h} \ge \frac{f(y) - f(x)}{y - x}$$

Equation (32) had to be true for any value of $\lambda \in (0, 1)$. As $\lambda \to 0$ we see that $h \to 0$. Taking the limit of both sides of the previous equation as $h \to 0$ gives us

$$f'(y) \ge \frac{f(y) - f(x)}{y - x} \tag{34}$$

Combining equations (33) and (34), we have have shown that

$$f'(y) \ge f'(x)$$

We assumed only that f was convex and that $y > x$; we concluded that $f'(y) \ge f'(x)$. By definition, this means that f' is monotonically increasing if f is convex.

The "only if" case: f' is monotonically increasing only if f is convex

Assume that f is not convex. By lemma 1, we can find some subinterval $(s, t) \in (a, b)$ such that

$$f(\lambda s + (1 - \lambda)t) > \lambda f(s) + (1 - \lambda)f(t) \tag{35}$$

is true for all $\lambda \in (0, 1)$. We can now follow the logic of the "if" case and define

$$h = (\lambda - 1)(s - t)$$

Rearranging this algebraically allows us to express λ and $1 - \lambda$ as

$$\lambda = 1 - \frac{h}{t - s}, \quad (1 - \lambda) = \frac{h}{s - t}$$

Substituting these values of λ and $\lambda - 1$ into (35) results in

$$f(s + h) > f(s) - \frac{h \, f(s)}{t - s} + \frac{h \, f(t)}{t - s}$$

which is algebraically equivalent to

$$\frac{f(s + h) - f(s)}{h} > \frac{f(t) - f(t)}{t - s}$$

Equation (35) had to be true for any value of $\lambda \in (0, 1)$. As $\lambda \to 1$ we see that $h \to 0$. Taking the limit of both sides of the previous equation as $h \to 0$ gives us

$$f'(s) > \frac{f(t) - f(s)}{t - s} \tag{36}$$

Having established (36), we now redefine h such that

$$h = -(\lambda)(s - t)$$

Rearranging this algebraically allows us to express λ and $1 - \lambda$ as

$$\lambda = \frac{h}{t - s}, \quad (1 - \lambda) = 1 - \frac{h}{s - t}$$

Substituting these values of λ and $1 - \lambda$ into (35) results in

$$f(t - h) > \frac{h\, f(s)}{t - s} + f(t) - \frac{h\, f(t)}{t - s}$$

which is algebraically equivalent to

$$\frac{f(t) - f(t - h)}{h} < \frac{f(t) - f(s)}{t - s}$$

Equation (35) had to be true for any value of $\lambda \in (0, 1)$. As $\lambda \to 0$ we see that $h \to 0$. Taking the limit of both sides of the previous equation as $h \to 0$ gives us

$$f'(t) < \frac{f(t) - f(s)}{t - s} \tag{37}$$

Combining equations (36) and (37), we have have shown that

$$f'(t) < f'(s)$$

for some $t > s$. We assumed only that f was not convex; we concluded that f' was not monotonically increasing. By contrapositive, this means that f' is monotonically increasing only if f is convex.

f is convex iff $f''(x) \geq 0$ for all $x \in (a, b)$

We've shown that f is convex iff f' is monotonically inreasing, and theorem 5.11 tells us that f' is monotonically increasing iff $f''(x) \geq 0$ for all $x \in (a, b)$.

Exercise 5.15

Note on the bounds of f and its derivatives

When Rudin asks us to assume that $|f|$ and its derivatives have upper bounds of M_0, M_1, and M_2 it appears that we must assume that these bounds are finite. Otherwise the function $f(x) = x$ would be a counterexample to the claim that $M_1^2 \leq 4M_0 M_2$.

Proof that $M_1^2 \leq 4M_0 M_2$ for real-valued functions

Choose any $h > 0$. Using Taylor's theorem (theorem 5.15) we can express $f(x + 2h)$ as

$$f(x + 2h) = f(x) + 2hf'(x) + \frac{4h^2 f''(\xi)}{2}, \quad \xi > a \tag{38}$$

which can be algebraically arranged to give us

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

$$|f'(x)| = |\frac{f(x + 2h)}{2h} - \frac{f(x)}{2h} - hf''(\xi)| \leq |\frac{f(x + 2h)}{2h}| + |\frac{f(x)}{2h}| + |hf''(\xi)|$$

We're given upper bounds for $|f(x)|$ and $|f''(x)|$; these bounds give us the inequality

$$|f'(x)| \leq \frac{2M_0}{h} + hM_2$$

85

This inequality must hold for all x, even when $|f(x)|$ approaches its upper bound, so we have

$$M_1 \leq \frac{2M_0}{h} + hM_2$$

We can multiply both sides by h and then algebraically rearrange this into a quadratic equation in h.

$$h^2 M_2 - hM_1 + M_0 \geq 0 \qquad (39)$$

The quadratic solution to this equation is

$$h = \frac{M_1 \pm \sqrt{M_1^2 - 4M_0 M_2}}{2M_2} \qquad (40)$$

We want to make sure that there are *not* two solutions to (40): we want (39) to hold for all values of h, and if (40) had two solutions then we would have $h^2 M_2 - hM_1 + M_0 < 0$ on some interval of h. To make sure that there is at most a single real solution we need to make sure that the discriminant of (40) is either zero (one solution) or negative (zero solutions). This occurs exactly when

$$M_1^2 \leq 4M_0 M_2$$

Does this apply to vector-valued functions?

Yes. Let f be a vector-valued function that is continuous on (a, ∞) and is defined by

$$f(x) = (\ f_1(x), f_2(x), \ldots, f_n(x)\ )$$

Assume that $|f|, |f'|$, and $|f''|$ have finite upper bounds of (respectively) M_0, M_1, and M_2. If we evaluate f at $x + 2h$ we have

$$f(x + 2h) = (\ f_1(x + 2h), f_2(x + 2h), \ldots, f_n(x + 2h)\ )$$

Taking the Taylor expansion of each of these terms, we have

$$
\begin{aligned}
f(x + 2h) &= (\ [f_1(x) + 2hf_1'(x) + 2h^2 f_1''(\xi_1)], [f_2(x) + 2hf_2'(x) + 2h^2 f_2''(\xi_1)], \ldots, [f_n(x) + 2hf_n'(x) + 2h^2 f_n''(\xi_1)]\ ) \\
&= (\ f_1(x), f_2(x), \ldots, f_n(x)\ ) + (\ 2hf_1'(x), 2hf_2'(x), \ldots, 2hf_n'(x)\ ) + (\ 2h^2 f_1''(\xi), 2h^2 f_2''(\xi), \ldots, 2h^2 f_n''(\xi)\ ) \\
&= f(x) + 2hf'(x) + 2h^2 f''(\xi)
\end{aligned}
$$

This tells us that equation (38) holds for vector-valued functions. But nothing in the proof following equation (38) requires us to assume that f is real-valued instead of vector-valued. So the proof following (38) still suffices to prove that

$$M_1^2 \leq 4M_0 M_2$$

Exercise 5.16 proof 1

In exercise 5.15 we established the inequality

$$|f'(x)| \leq |\frac{f(x + 2h)}{2h}| + |\frac{f(x)}{2h}| + |hf''(\xi)|$$

We are given an upper bound of M_2 for $|f''(\xi)|$, so this inequality can be expressed as

$$|f'(x)| \leq |\frac{f(x + 2h)}{2h}| + |\frac{f(x)}{2h}| + hM_2$$

We're told that f is twice-differentiable, so we know that both f and f' are continuous (theorem 5.2), so we can take the limit of both sides of this equation as $x \to 0$ (theorem 4.2).

$$\lim_{x \to \infty} |f'(x)| \leq \lim_{x \to \infty} \left(|\frac{f(x + 2h)}{2h}| + |\frac{f(x)}{2h}| + hM_2 \right)$$

We're told that $f(x) \to 0$ as $x \to 0$ so this becomes

$$\lim_{x \to \infty} |f'(x)| \leq \lim_{x \to \infty} hM_2$$

This must be true for all h, so we can take the limit of both sides of this as $h \to 0$ (we can do this because both $|f'(x)|$ and hM_2 are continuous functions with respect to the variable h).

$$\lim_{h \to 0} \lim_{x \to \infty} |f'(x)| \leq \lim_{h \to 0} \lim_{x \to \infty} hM_2$$

The left-hand side of the previous inequality is independent of h and therefore doesn't change; the right-hand side becomes 0.

$$\lim_{x \to \infty} |f'(x)| = \lim_{h \to 0} \lim_{x \to \infty} |f'(x)| \leq \lim_{h \to 0} \lim_{x \to \infty} hM_2 = 0$$

$$\lim_{x \to \infty} |f'(x)| \leq 0$$

This show that $f'(x) \to 0$ as $x \to \infty$, which is what we were asked to prove.

Exercise 5.16 proof 2

In exercise 5.15 we established the inequality

$$M_1^2 \leq 4M_0 M_2$$

where each of the M terms represented a supremum on the interval (a, ∞). This inequality was proven to hold for all a such that f was twice-differentiable on the interval (a, ∞). To show more explicitly that the value of these M terms depends on a, we might express the previous inequality as

$$\sup_{x > a} |f'(x)|^2 \leq \sup_{x > a} 4|f(x)||f''(x)|$$

Each of these terms is continuous with respect to a (the proof of this claim is trivial but tedious) so we can take the limit of both sides as $a \to \infty$.

$$\lim_{a \to \infty} \sup_{x > a} |f'(x)|^2 \leq \lim_{a \to \infty} \sup_{x > a} 4|f(x)||f''(x)|$$

We're told that $|f''(x)|$ has an upper bound of M_2. We're also told that $f(x) \to 0$ as $x \to \infty$. The last inequality therefore allows us to conclude that

$$\lim_{a \to \infty} \sup_{x > a} |f'(x)|^2 \leq \lim_{a \to \infty} \sup_{x > a} 4|0||M_2| = 0$$

It's clear that $|f'(x)|$ must be less than or equal to the supremum of a set containing $|f'(x)|$, so we have

$$\lim_{x \to \infty} |f'(x)| \leq \lim_{a \to \infty} \sup_{x > a} |f'(x)|^2 \leq 0$$

$$\lim_{x \to \infty} |f'(x)| \leq 0$$

This shows that $f'(x) \to 0$ as $x \to \infty$, which is what we were asked to prove.

Exercise 5.17

We're told that f is three-times differentiable on the interval $(-1, 1)$ so we can take the Taylor expansions of $f(-1)$ and $f(1)$ around $x = 0$:

$$f(-1) = f(0) + f'(0)(-1) + \frac{f''(0)(-1)^2}{2} + \frac{f'''(\xi_1)(-1)^3}{6}, \quad \xi_1 \in (-1, 0)$$

$$f(1) = f(0) + f'(0)(1) + \frac{f''(0)(1)^2}{2} + \frac{f'''(\xi_2)(1)^3}{6}, \quad \xi_2 \in (0, 1)$$

When we evaluate $f(1) - f(-1)$, many of these terms cancel out and we're left with

$$f(1) - f(-1) = 2f'(0) + \frac{f'''(\xi_1) + f'''(\xi_2)}{6}, \quad \xi_1 \in (-1, 0) \xi_2 \in (0, 1)$$

We're given the values of $f(1), f(-1)$, and $f'(0)$ so this last equation becomes

$$1 = \frac{f'''(\xi_1) + f'''(\xi_2)}{6}, \quad \xi_1 \in (-1, 0) \xi_2 \in (0, 1)$$

which is algebraically equivalent to

$$f'''(\xi_2) = 6 - f'''(\xi_1), \quad \xi_1 \in (-1, 0) \xi_2 \in (0, 1)$$

If $f'''(\xi_1) \leq 3$ then $f'''(\xi_2) \geq 3$, and vice-versa: so $f'''(x) \geq 3$ for either ξ_1 or $\xi_2 \in (-1, 1)$. And this is what we were asked to prove.

Exercise 5.18

The nth derivative of $f(t)$

The exercise tells gives us the following formula for $f(t)$:

$$f(t) = f(\beta) - (\beta - t)Q(t)$$

Since β is a fixed constant, the first two derivatives of f with respect to t are

$$f'(t) = Q(t) - (\beta - t)Q''(t)$$

$$f''(t) = 2Q'(t) - (\beta - t)Q'''(t)$$

And, in general, we have

$$f^{(n)}(t) = nQ^{(n-1)}(t) - (\beta - t)Q^{(n)}(t)$$

which, after multiplying by $(\beta - \alpha)^n/n!$ and setting $t = \alpha$, becomes

$$\frac{f^{(n)}(t)}{n!}(\beta - \alpha)^n = \frac{Q^{(n-1)}(t)}{(n-1)!}(\beta - \alpha)^n - \frac{Q^{(n)}(t)}{n!}(\beta - \alpha)^{n+1} \tag{41}$$

Modifying the Taylor formula

The formula for the Taylor expansion of $f(\beta)$ around the point $f(\alpha)$ (theorem 5.15) includes a $P(\beta)$ term defined as

$$P(\beta) = \sum_{k=0}^{n-1} \left[\frac{f^{(k)}(\alpha)}{k!}(\beta - \alpha)^k \right]$$

If we isolate the $k = 0$ case this becomes

$$P(\beta) = f(\alpha) + \sum_{k=1}^{n-1} \left[\frac{f^{(k)}(\alpha)}{k!}(\beta - \alpha)^k \right]$$

We can now use equation (41) to express the terms of the summation as functions of Q.

$$P(\beta) = f(\alpha) + \sum_{k=1}^{n-1} \left[\frac{Q^{(k-1)}(\alpha)}{(k-1)!}(\beta - \alpha)^k \right] - \sum_{k=1}^{n-1} \left[\frac{Q^{(k)}(\alpha)}{k!}(\beta - \alpha)^{k+1} \right]$$

If we extract the $k = 1$ term from the leftmost summation and the $k = n-1$ term from the rightmost summation, we have

$$P(\beta) = f(\alpha) + Q(\alpha)(\beta - \alpha) + \sum_{k=2}^{n-1} \left[\frac{Q^{(k-1)}(\alpha)}{(k-1)!}(\beta - \alpha)^k \right] - \sum_{k=1}^{n-2} \left[\frac{Q^{(k)}(\alpha)}{k!}(\beta - \alpha)^{k+1} \right] - \frac{Q^{(n-1)}(\alpha)}{(n-1)!}(\beta - \alpha)^n$$

88

We can then re-index the leftmost summation to obtain

$$P(\beta) = f(\alpha) + Q(\alpha)(\beta - \alpha) + \sum_{k=1}^{n-2} \left[\frac{Q^{(k)}(\alpha)}{(k)!} (\beta - \alpha)^{k+1} \right] - \sum_{k=1}^{n-2} \left[\frac{Q^{(k)}(\alpha)}{k!} (\beta - \alpha)^{k+1} \right] - \frac{Q^{(n-1)}(\alpha)}{(n-1)!} (\beta - \alpha)^n$$

The two summations cancel one another, leaving us with

$$P(\beta) = f(\alpha) + Q(\alpha)(\beta - \alpha) - \frac{Q^{(n-1)}(\alpha)}{(n-1)!} (\beta - \alpha)^n$$

If we replace $Q(\alpha)$ with the definition of Q given in the exercise, we see that this previous equation evaluates to

$$P(\beta) = f(\alpha) + \frac{f(\alpha) - f(\beta)}{\alpha - \beta} (\beta - \alpha) - \frac{Q^{(n-1)}(\alpha)}{(n-1)!} (\beta - \alpha)^n$$

This simplifies to

$$P(\beta) = f(\alpha) + (f(\beta) - f(\alpha)) - \frac{Q^{(n-1)}(\alpha)}{(n-1)!} (\beta - \alpha)^n$$

A simple algebraic rearrangement of these terms gives us

$$f(\beta) = P(\beta) - \frac{Q^{(n-1)}(\alpha)}{(n-1)!} (\beta - \alpha)^n$$

which is the equation we were asked to derive.

Exercise 5.19a

The given expression for D_n is algebraically equivalent to

$$D_n = \left[\frac{f(\beta_n) - f(0)}{\beta_n - 0} \frac{\beta_n}{\beta_n - \alpha_n} \right] + \left[\frac{f(0) - f(\alpha_n)}{0 - \alpha_n} \frac{-\alpha_n}{\beta_n - \alpha_n} \right]$$

We really want to be able to evaluate this by taking the limit of each side of this previous equation in the following manner:

$$\lim_{n \to \infty} D_n = \lim_{n \to \infty} \left[\frac{f(\beta_n) - f(0)}{\beta_n - 0} \right] \cdot \lim_{n \to \infty} \left[\frac{\beta_n}{\beta_n - \alpha_n} \right] + \lim_{n \to \infty} \left[\frac{f(0) - f(\alpha_n)}{0 - \alpha_n} \right] \cdot \lim_{n \to \infty} \left[\frac{-\alpha_n}{\beta_n - \alpha_n} \right] \quad (42)$$

There two conditions that must be met in order for this last step to be justified:

condition 1: theorem 3.3 tells us that each of the limits must actually exist (and must not be $\pm\infty$)

condition 2: and theorem 4.2 tells us that we must have $\lim f(\alpha_n) = \lim f(\beta_n) = f(0)$.

The fact that $\alpha_n < 0 < \beta_n$ guarantees that $0 < \beta_n/(\beta_n - \alpha_n) < 1$ and $0 < -\alpha_n/(\beta_n - \alpha_n) < 1$, which tells us that at 2 of the 4 limits in (42) exist. The other two limits exist because they're equal to $f'(0)$, and we're told that $f'(0)$ exists. Therefore condition 1 is met. The fact that $f'(0)$ exists tells us that f is continuous at $x = 0$ (theorem 5.2) and therefore condition 2 is met (theorem 4.2). Therefore we're justified in taking the limits in (42), giving us

$$\lim_{n \to \infty} D_n = f'(0) \cdot \lim_{n \to \infty} \left[\frac{\beta_n}{\beta_n - \alpha_n} \right] + f'(0) \cdot \lim_{n \to \infty} \left[\frac{-\alpha_n}{\beta_n - \alpha_n} \right]$$

$$= \lim_{n \to \infty} f'(0) \frac{\beta_n - \alpha_n}{\beta_n - \alpha_n} = f'(0)$$

Exercise 5.19b

The given expression for D_n is algebraically equivalent to

$$D_n = \left[\frac{f(\beta_n) - f(0)}{\beta_n - 0} \frac{\beta_n}{\beta_n - \alpha_n} \right] + \left[\frac{f(\alpha_n) - f(0)}{\alpha_n - 0} \left(1 - \frac{\beta_n}{\beta_n - \alpha_n} \right) \right]$$

We want to evaluate this by taking the limits of the individual terms as we did in part (a):

$$\lim_{n \to \infty} D_n = \lim_{n \to \infty} \left[\frac{f(\beta_n) - f(0)}{\beta_n - 0} \right] \cdot \lim_{n \to \infty} \left[\frac{\beta_n}{\beta_n - \alpha_n} \right] + \lim_{n \to \infty} \left[\frac{f(\alpha_n) - f(0)}{\alpha_n - 0} \right] \lim_{n \to \infty} \left[\left(1 - \frac{\beta_n}{\beta_n - \alpha_n} \right) \right] \qquad (43)$$

In order for this to be justified we must once again meet the two conditions mentioned in part (a). We're told that $0 < \beta_n/(\beta_n - \alpha_n) < M$ for some M, which tells us that at 2 of the 4 limits in (43) exist. The other two limits exist because they're equal to $f'(0)$, and we're told that $f'(0)$ exists. Therefore condition 1 is met. The fact that $f'(0)$ exists tells us that f is continuous at $x = 0$ (theorem 5.2) and therefore condition 2 is met (theorem 4.2). Therefore we're justified in taking the limits in (43), giving us

$$
\begin{aligned}
\lim_{n \to \infty} D_n &= f'(0) \cdot \lim_{n \to \infty} \left[\frac{\beta_n}{\beta_n - \alpha_n} \right] + f'(0) \cdot \lim_{n \to \infty} \left[1 - \frac{\beta_n}{\beta_n - \alpha_n} \right] \\
&= f'(0) \cdot \lim_{n \to \infty} \left[\frac{\beta_n}{\beta_n - \alpha_n} + 1 - \frac{\beta_n}{\beta_n - \alpha_n} \right] \\
&= f'(0)
\end{aligned}
$$

Exercise 5.19c proof 1

Define the sequence $\{h_n\}$ where $h_n = \beta_n - \alpha_n$. We can now express D_n as

$$D_n = \frac{f(\alpha_n + h_n) - f(\alpha_n)}{h_n}$$

We know that $\alpha_n \to 0$ and $\beta_n \to 0$ as $n \to \infty$, so clearly $h_n \to 0$ as $n \to \infty$. We're also told that f' is continuous on $(-1, 1)$, so we know that f' is defined on this interval. Therefore we have

$$
\begin{aligned}
\lim_{n \to \infty} D_n &= \lim_{n \to \infty} \lim_{h \to 0} \frac{f(\alpha_n + h_n) - f(\alpha_n)}{h_n} \\
&= \lim_{n \to \infty} f'(\alpha_n)
\end{aligned}
$$

We're told that f' is continuous on the interval $(-1, 1)$ and that $\lim \alpha_n = 0$ so by theorem 4.2 we have

$$\lim_{n \to \infty} D_n = \lim_{n \to \infty} f'(\alpha_n) = f'(0)$$

Exercise 5.19c proof 2

The mean value theorem (theorem 5.10) allows us to construct a sequence $\{\gamma_n\}$ as follows: for each n, choose some $\gamma_n \in (\alpha_n, \beta_n)$ such that

$$D_n = \frac{f(\beta_n) - f(\alpha_n)}{\beta_n - \alpha_n} = f'(\gamma_n) \qquad (44)$$

Each γ_n is in the interval (α_n, β_n). We're told that $\lim \alpha_n = \lim \beta_n = 0$. Therefore we can use the squeeze theorem to determine $\lim \gamma_n$.

$$0 = \lim_{n \to \infty} \alpha_n \leq \lim_{n \to \infty} \gamma_n \leq \lim_{n \to \infty} \beta_n = 0$$

which means that $\lim \gamma_n \to 0$. We're told that f' is continuous so by theorem 5.2 we can take the limit of equation (44).

$$\lim_{n \to \infty} D_n = \lim_{n \to \infty} f'(\gamma_n) = f'(0)$$

Exercise 5.20

Exercise 5.21

We saw in exercise 2.29 that any open set in $\mathbb{R}^1$ can be represented as a countable number of disjoint open segments. Let $\{(a_n, b_n)\}$ represent such a countable collection of disjoint sets such that $\bigcup(a_i, b_i) = E^C$. Define $f : \mathbb{R}^1 \to \mathbb{R}^1$ as

$$f(x) = \begin{cases} 0, & x \in E \\ (x - b_i)^2, & x \in (a_i, b_i) \wedge a_i = -\infty \\ (x - a_i)^2, & x \in (a_i, b_i) \wedge b_i = \infty \\ (x - a_i)^2(x - b_i)^2, & x \in (a_i, b_i) \wedge -\infty < a_i < b_i < \infty \end{cases}$$

It's easy but tedious to verify that this a continuous function which is differentiable and that $f(x) = 0$ iff $x \in E$. Note that the derivative also has the property that $f'(x) = 0$ iff $x \in E$. For the second part of the exercise we can define a function f to be

$$f(x) = \begin{cases} 0, & x \in E \\ (x - b_i)^{n+1}, & x \in (a_i, b_i) \wedge a_i = -\infty \\ (x - a_i)^{n+1}, & x \in (a_i, b_i) \wedge b_i = \infty \\ (x - a_i)^{n+1}(x - b_i)^{n+1}, & x \in (a_i, b_i) \wedge -\infty < a_i < b_i < \infty \end{cases}$$

It's similarly easy but tedious to verify that this is a continuous function that is n times differentiable and that $f(x) = 0$ iff $x \in E$. It can also be seen that when $k \le n$ we have $f^{(k)}(x) = 0$ iff $x \in E$. Finally[6], we can pretend that we've defined the exponential function and define f to be

$$f(x) = \begin{cases} 0, & x \in E \\ \exp\left(\frac{-1}{(x-b_i)^2}\right), & x \in (a_i, b_i) \wedge a_i = -\infty \\ \exp\left(\frac{-1}{(x-a_i)^2}\right), & x \in (a_i, b_i) \wedge b_i = \infty \\ \exp\left(\frac{-1}{(x-a_i)^2(x-b_i)^2}\right), & x \in (a_i, b_i) \wedge -\infty < a_i < b_i < \infty \end{cases}$$

It's fairly easy to confirm that this function is continuous and that $f(x) = 0$ iff $x \in E$. Looking at the derivative with respect to x we have

$$f'(x) = \begin{cases} 0, & x \in E \\ \frac{-2}{(x-b_i)^3} \exp\left(\frac{-1}{(x-b_i)^2}\right), & x \in (a_i, b_i) \wedge a_i = -\infty \\ \frac{-2}{(x-a_i)^3} \exp\left(\frac{-1}{(x-a_i)^2}\right), & x \in (a_i, b_i) \wedge b_i = \infty \\ \frac{-2(b_i-a_i)}{(x-a_i)^2(x-b_i)^2} \exp\left(\frac{-1}{(x-a_i)^2(x-b_i)^2}\right), & x \in (a_i, b_i) \wedge -\infty < a_i < b_i < \infty \end{cases}$$

To calculate the limit of $f'(x)$ in the first case, we use L'Hopital's rule.

$$\lim_{x \to a} f'(x) = \lim_{x \to a} \frac{\exp\left(\frac{-1}{(x-b_i)^2}\right)}{(x - b_i)^3}$$

The numerator and denominator of this last term both tend to $\pm\infty$, so L'Hopital's rule is applicable. Repeated applications of L'Hopital's rule will eventually give us a constant term in the numerator and a term that tends to $\pm\infty$ in the denominator, so we see that $\lim f'(x) = 0$. Similar results hold for the limits in the other two cases. The general idea is that, for any polynomial term $p(n)$, the exponential limit $\lim_{n \to \infty} exp(-1/n)$ will tend towards zero faster than $\lim_{n \to \infty} p(n)$ tends towards ∞. Therefore $p(n)exp(-1/n)$ will tend to zero as $n \to \infty$. Every term of every derivative of $f(x)$ will consist only of polynomial multiples of the exponential term, so it will hold that for all k:

$$\lim_{x \to a_i} f^{(k)}(x) = \lim_{x \to b_i} f^{(k)}(x) = 0$$

It seems clear in a vague, poorly-defined Calc II-ish way that f is infinitely differentiable and that, for all n, $f^{(n)}(x) = 0$ iff $x \in E$. I have no idea how to prove this.

---

[6]this example was provided by Boris Shekhtman, University of South Florida

Exercise 5.22a

Assume that $f(x_1) = x_1$ and $f(x_2) = x_2$ with $x_1 \neq x_2$. Then

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{x_2 - x_1}{x_2 - x_1} = 1$$

By theorem 5.10 this means that $f'(x) = 1$ for some x between x_1 and x_2. This contradicts the presumption that $f'(x) \neq 1$, so our initial assumption must be wrong: there cannot be two fixed points.

Exercise 5.22b

For $f(t)$ to have a fixed point it would be necessary that $f(t) = t$, in which case we would have

$$t = t + (1 - e^t)^{-1} \quad \longrightarrow \quad \frac{1}{1 - e^t} = 0$$

This statement is not true for any t.

Exercise 5.22c

Choose an arbitrary value for x_0 and let $\{x_n\}$ be the sequence recursively defined by $x_{n+1} = f(x_n)$. We're told that $|f'(t)| \leq A$ for all real t. So for any n we have

$$\left| \frac{f(x_{n-1}) - f(x_{n-2})}{x_{n-1} - x_{n-2}} \right| \leq A \quad \longrightarrow \quad |f(x_{n-1}) - f(x_{n-2})| \leq A\,|x_{n-1} - x_{n-2}|$$

which, since $f(x_n) = x_{n+1}$, gives us

$$|x_n - x_{n-1}| \leq A|x_{n-1} - x_{n-2}|$$

But this holds for *all* n. So:

$$|x_n - x_{n-1}| \leq A|x_{n-1} - x_{n-2}| \leq A^2|x_{n-2} - x_{n-3}| \leq A^3|x_{n-3} - x_{n-4}| \leq \cdots \leq A^{n-2}|x_2 - x_1|$$

We can use this general formula to determine the difference between x_{n+k} and x_n.

$$
\begin{aligned}
|x_{n+k} - x_n| &= |(x_{n+k} - x_{n+k-1}) + (x_{n+k-1} - x_{n+k-2}) + \cdots + (x_{n+2} - x_{n+1}) + (x_{n+1} - x_n)| \\
&\leq |x_{n+k} - x_{n+k-1}| + |x_{n+k-1} - x_{n+k-2}| + \cdots + |x_{n+2} - x_{n+1}| + |x_{n+1} - x_n)| \\
&\leq A^{n+k-2}|x_2 - x_1| + A^{n+k-3}|x_2 - x_1| + \cdots + A^n|x_2 - x_1| + A^{n-1}|x_2 - x_1)| \\
&\leq A^n|x_2 - x_1|
\end{aligned}
$$

This shows us that

$$|x_{n+k} - x_n| \leq A^n|x_2 - x_1|$$

We're told that $A < 1$, so taking the limit of eachs ide of this inequality as $n \to \infty$ gives us

$$\lim_{n \to \infty} |x_{n+k} - x_n| \leq \lim_{n \to \infty} A^n|x_2 - x_1| = 0$$

And this is just the Cauchy criterion for convergence for the sequence $\{x_n\}$. And we're converging in $\mathbb{R}^1$ so by theorem 3.11 we know that the sequence converges to some element $x \in \mathbb{R}^1$.

$$\exists x \in \mathbb{R} : \lim_{n \to \infty} x_n = x$$

But the elements of $\{x_n\}$ are just the elements of $\{f(x_n)\}$, so we can also conclude that

$$\exists x \in \mathbb{R} : \lim_{n \to \infty} f(x_n) = x$$

And we're told that f is continuous, so by theorem 4.6 we see that

$$\exists x \in \mathbb{R} : f(x) = x$$

Exercise 5.23a

If $x < \alpha$, then we can express x as $x = \alpha - \delta$ for some $\delta > 0$. This gives us

$$
\begin{aligned}
f(x) &= f(\alpha - \delta) && \text{definition of } x \\
&= \tfrac{(\alpha-\delta)^3 - 1}{3} && \text{definition of } f \\
&= \tfrac{\alpha^3 - 3\alpha^2\delta + 3\alpha\delta^2 - \delta^3 + 1}{3} && \text{algebra} \\
&= \tfrac{\alpha^3 + 1}{3} + \tfrac{-3\alpha^2\delta + 3\alpha\delta^2}{3} - \tfrac{\delta^3}{3} && \text{algebra} \\
&= f(\alpha) - \alpha\delta(\alpha - \delta) - \tfrac{1}{3}\delta^3 && \text{algebra} \\
&= f(\alpha) - \alpha\delta x - \tfrac{1}{3}\delta^3 && \text{definition of } x \\
&= \alpha - \alpha\delta x - \tfrac{1}{3}\delta^3 && \alpha \text{ is a fixed point}
\end{aligned}
$$

We know that $x < \alpha < -1$, so $\alpha x > 1$ and therefore $\alpha\delta x > \delta$. From this we have an inequality:

$$
\begin{aligned}
&< \alpha - \delta - \tfrac{1}{3}\delta^3 \\
&< \alpha - \delta \\
&= x && \text{definition of } x
\end{aligned}
$$

This establishes that $x < \alpha \to f(x) < x$. Therefore if $x_1 < \alpha$ then the sequence $\{x_n\}$ will be a nonincreasing sequence. We know that this sequence doesn't converge because we're told that f has no fixed points less than α.

Exercise 5.23b

The first derivative of f is $f'(x) = x^2$. This is nonnegative, so we know that f is monotonically increasing (theorem 5.11). The second derivative is $f'(x) = 2x$. This is negative for $x < 0$ and positive for $x > 0$: therefore f is strictly convex on $(0, \gamma)$ and is strictly concave on $(\alpha, 0)$. Now let x_k be chosen from the interval (α, γ).

Case 1: $x_k \in (\alpha, 0)$

Let $x_k \in (\alpha, 0)$ be chosen. We can express this x as $x_k = \lambda\alpha + (1-\lambda)0$ for some $\lambda \in (0, 1)$. The second derivative of $f''(x) = 2x$ is negative on the interval $(\alpha, 0)$ and so the function is concave on this interval (corollary of exercise 5.14). Therefore we have

$$
\begin{aligned}
x_{k+1} &= f(x_k) && \text{definition of } x_{k+1} \\
&= f(\lambda\alpha + (1-\lambda)0) && \text{definition of } x_k \\
&> \lambda f(\alpha) + (1-\lambda)f(0) && f \text{ is concave on } (\alpha, 0) \\
&= \lambda\alpha + (1-\lambda)(1/3) && \alpha \text{ is a fixed point, } f(0) = 1/3 \\
&= x_k + (1-\lambda)(1/3) && \text{definition of } x_k \\
&> x_k
\end{aligned}
$$

We see that $x_k < x_{k+1}$; from our choice of x_k we have $\alpha < x_k$; from the monotonic nature of f we have $x_k < \beta \to x_{k+1} = f(x_k) < f(\beta) = \beta$. Combining these inequalities yields

$$
\alpha < x_k < x_{k+1} < \beta
$$

Therefore our initial choice of x_k gives us an increasing sequence $\{x_n\}$ with an upper bound of β. Therefore it converges to a some fixed point in the interval $(x_1, beta]$, and this point must be β.

Case 2: $x \in (\beta, \gamma)$

Let $x_k \in (\beta, \gamma)$ be chosen. We can express this as $x_k = \lambda\beta + (1-\lambda)\gamma$ for some $\lambda \in (0, 1)$. The second derivative of $f''(x) = 2x$ is positive on the interval (β, γ) and so the function is convex on this interval (exercise 5.14). There-

fore we have

$$
\begin{aligned}
x_{k+1} &= f(x_k) & \text{definition of } x_{k+1} \\
&= f(\lambda\beta + (1-\lambda)\gamma) & \text{definition of } x_k \\
&< \lambda f(\beta) + (1-\lambda)f(\gamma) & f \text{ is convex on } (\beta,\gamma) \\
&= \lambda\beta + (1-\lambda)\gamma & \beta \text{ and } \gamma \text{ are fixed points} \\
&= x_k & \text{definition of } x_k
\end{aligned}
$$

We see that $x_{k+1} < x_k$; from our choice of x_k we have $x_k < \gamma$; from the monotonic nature of f we have $\beta < x_k \to \beta = f(\beta) < f(x_k) = x_{k+1}$. Combining these inequalities yields

$$\beta < x_{k+1} < x_k < \gamma$$

Therefore our initial choice of x_k gives us an decreasing sequence $\{x_n\}$ with an lower bound of β. Therefore it converges to a some fixed point in the interval $[\beta, x_1)$, and this point must be β.

Case 3: $x \in (0, \beta)$

Let $x_k \in (0, \beta)$ be chosen. We can express this as $x_k = \lambda 0 + (1-\lambda)\beta$ for some $\lambda \in (0,1)$. The second derivative of $f''(x) = 2x$ is positive on the interval $(0, \beta)$ and so the function is convex on this interval (exercise 5.14). Therefore we have

$$
\begin{aligned}
x_{k+1} &= f(x_k) & \text{definition of } x_{k+1} \\
&= f(\lambda(0) + (1-\lambda)\beta) & \text{definition of } x_k \\
&< \lambda f(0) + (1-\lambda)f(\beta) & \text{definition of convexity} \\
&= \lambda(1/3) + (1-\lambda)\beta & \beta \text{ is a fixed point}, f(0) = 1/3 \\
&= \lambda(1/3) + x & \text{definition of } x_k \\
&> x_k
\end{aligned}
$$

We see that $x_k < x_{k+1}$; from our choice of x_k we have $\alpha < x_k$; from the monotonic nature of f we have $x_k < \beta \to x_{k+1} = f(x_k) < f(\beta) = \beta$. Combining these inequalities yields

$$\alpha < x_k < x_{k+1} < \beta$$

Therefore our initial choice of x_k gives us an increasing sequence $\{x_n\}$ with an upper bound of β. Therefore it converges to a some fixed point in the interval $(x_1, beta]$, and this point must be β.

Case 4: $x_k = \beta$ or $x_k = 0$

If $x_k = \beta$ then every term of $\{x_n\}$ is β, so this sequence clearly converges to β. If $x_k = 0$ then $x_{k+1} = f(0) = (1/3)$ and the remainder of the sequence $\{x_n\}$ converges to β by one of the previous cases.

Exercise 5.23c

If $x > \gamma$ then we can express x as $x = \gamma + \delta$ for some $\delta > 0$. This gives us

$$
\begin{aligned}
f(x) &= f(\gamma + \delta) & \text{definition of } x \\
&= \frac{(\gamma+\delta)^3 - 1}{3} & \text{definition of } f \\
&= \frac{\gamma^3 + 3\gamma^2\delta + 3\gamma\delta^2 + \delta^3 + 1}{3} & \text{algebra} \\
&= \frac{\gamma^3 + 1}{3} + \frac{3\gamma^2\delta + 3\gamma\delta^2}{3} + \frac{\delta^3}{3} & \text{algebra} \\
&= f(\gamma) + \gamma\delta(\gamma + \delta) + \tfrac{1}{3}\delta^3 & \text{algebra} \\
&= f(\gamma) - \gamma\delta x + \tfrac{1}{3}\delta^3 & \text{definition of } x \\
&= \gamma + \gamma\delta x + \tfrac{1}{3}\delta^3 & \gamma \text{ is a fixed point}
\end{aligned}
$$

We know that $\gamma > 1$ and $x > 1$, so $\gamma x \delta > \delta$. From this we have an inequality :

$$
\begin{aligned}
&> \gamma + \delta + \tfrac{1}{3}\delta^3 \\
&> \gamma + \delta \\
&= x & \text{definition of } x
\end{aligned}
$$

This establishes that $x > \gamma \rightarrow f(x) > x$. Therefore if $x_1 > \gamma$ then the sequence $\{x_n\}$ will be a nonincreasing sequence. We know that this sequence doesn't converge because we're told that f has no fixed points greater than γ.

Exercise 5.24

The function $f(x)$ has a derivative of zero at its fixed point, so when x_k and x_{k+1} are both close to $\sqrt{a}$ the mean value theorem guarantees us that $f(x_k)$ and $f(x_{k+1})$ will be very near one another:

$$|f(x_k) - f(x_{k+1})| = |(x_k - x_{k+1})f'(\xi)| \approx 0$$

The lefthand term of this equation converges very rapidly to 0 because both $|x_k - x_{k+1}|$ and $f'(x)$ are converging toward zero. The function $g(x)$ does not have a derivative of zero at its fixed point and therefore does not have this property (although, as we saw in exercise 3.17, it still converges albeit more slowly).

Exercise 5.25a

Each x_{n+1} is chosen to be the point where the line tangent tangent to $f(x_n)$ crosses the x-axis.

Exercise 5.25b

Lemma 1: $x_{n+1} < x_n$ if $x_n > \xi$

We're told that $f(b) > 0$ and that $f(x) = 0$ only at $x = \xi$. We know that f is continuous since it is differentiable (theorem 5.2), so by the intermediate value theorem (theorem 4.23) we know that $f(x) > 0$ when $x > \xi$ (otherwise we'd have $f(x) = 0$ for some second $x \in (\xi, b)$). We're also told that $f'(x) > \delta > 0$ for all $x \in (a, b)$ (without the δ it might be the case that $f'(x) \neq 0$ but $\lim f'(x) = 0$). Therefore the ratio $f(x_n)/f'(x_n)$ is defined at each x and is positive when $x > \xi$. Therefore we have

$$x_n - \frac{f(x_n)}{f'(x_n)} < x_n, \quad x_n > \xi$$

This of course means that $x_{n+1} < x_n$ when $x_n > \xi$.

Lemma 2: $x_{n+1} > \xi$ if $x_n > \xi$

We're told that $f''(x) \geq 0$, which means that $f'(x)$ is monotonically increasing, which means that $c < x_n$ implies $f'(c) \leq f'(x_n)$, which means that

$$\frac{f'(x_n) - f(\xi)}{x_n - \xi} \leq f'(x_n)$$

because the LHS of this inequality is equal to $f'(c)$ for some $c < x_n$. Using the fact that $f(\xi) = 0$ and a bit of algebraic manipulation, this equation is equivalent to

$$x_n - \frac{f(x_n)}{f'(x_n)} \geq \xi$$

which of course means that $x_{n+1} \geq \xi$.

Lemma 3: if $\{x_n\} \rightarrow \kappa$ then $f(\kappa) = 0$

Suppose that $\lim_{n\to\infty} x_{n+1} = \kappa$. Then by the Cauchy criterion we have $\lim_{n\to\infty} x_n - x_{n+1} = 0$ which is equivalent to

$$\lim_{n\to\infty} x_n - \left(x_n - \frac{f(x_n)}{f'(x_n)} \right) = 0$$

or

$$\lim_{n\to\infty} \frac{f(x_n)}{f'(x_n)} = 0$$

For this to hold it must either be the case that $f(x_n) \rightarrow 0$ or $f'(x_n) \rightarrow \pm\infty$: but we know that $f'(x_n)$ is bounded from the fact that $f''(x_n)$ is bounded (mean value theorem : if $f'(x_n)$ were unbounded, then

$(f'(x_n) - f'(x_1))/x_n - x_1 = f''(c)$ would be unbounded). Therefore it must be the case that $f(x_n) \to 0$. So we have $\lim_{n \to \infty} x_n = \kappa$ and $\lim_{n \to \infty} f(x_n) = 0$. Therefore by theorem 4.2 we have

$$f(\kappa) = \lim_{x \to \kappa} f(x) = 0$$

The actual proof

Choose $x_1 > \xi$. By induction, using lemma 2, we know that every element of $\{x_n\}$ will be $> \xi$. Therefore, by lemma 1, $x_{n+1} < x_n$ for all n. This means that $\{x_n\}$ is a decreasing sequence with a lower bound of ξ. Therefore, by theorem 3.14, the sequence converges to some point κ. By lemma 3 we have $f(\kappa) = 0$. But we're told that f has only one zero, so it must be the case that $\kappa = \xi$. This means that $\{x_n\}$ converges to ξ.

Exercise 5.25c

Using Taylor's theorem to expand $f(\xi)$ around $f(x_n)$, we have

$$f(\xi) = f(x_n) + f'(x_n)(\xi - x_n) + \frac{f''(t_n)(\xi - x_n)^2}{2}$$

Subtracting $f(x_n)$ from both sides then dividing by $f'(x_n)$ (we're told $f'(x) \geq \delta > 0$) gives us

$$\frac{f(\xi) - f(x_n)}{f'(x_n)} = (\xi - x_n) + \frac{f''(t_n)(\xi - x_n)^2}{2f'(x_n)}$$

Rearranging some terms and recognizing that $f(\xi) = 0$, we have

$$x_n - \frac{f(x_n)}{f'(x_n)} - \xi = \frac{f''(t_n)(x_n - \xi)^2}{2f'(x_n)}$$

Which, by the definition of x_{n+1}, is equivalent to

$$x_{n+1} - \xi = \frac{f''(t_n)(x_n - \xi)^2}{2f'(x_n)}$$

Exercise 5.25d

We're told that $f'(x) \geq \delta$ and $f''(x) \leq M$, so the inequality in part (c) guarantees us that

$$x_{n+1} - \xi \leq \frac{M(x_n - \xi)^2}{2\delta} = A(x_n - \xi)^2$$

This allows us to recursively construct a chain of inequalities.

$$x_{n+1} - \xi \leq A^1(x_n - \xi)^2 \leq A^1 A^2(x_{n-1} - \xi)^4 \leq A^1 A^2 A^4(x_{n-2} - \xi)^8 \leq \cdots \leq A^1 A^2 \ldots A^{2^{n-1}}(x_1 - \xi)^{2^n}$$

Collapsing the exponents of the rightmost term, we have

$$x_{n+1} - \xi \leq A^{2^n - 1}(x_1 - \xi)^{2^n} = \frac{1}{A}[A(x_1 - \xi)]^{2^n}$$

Exercise 5.25e

How does g behave near ξ?

We're told f and f' are differentiable, and clearly x is differentiable, therefore g is differentiable (theorem 5.3). Using the quotient rule (theorem 5.3 again), the derivative of g is given by

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

96

Taking the absolute value of each side, we have

$$|g'(x)| = \left| \frac{f(x)f''(x)}{f'(x)} \right|$$

Using the inequality we established in part (d) we have

$$|g'(x)| \leq A|f(x)|$$

When we take the limit of each side of this equation as $x \to \xi$, the RHS reduces to $f(\xi) = 0$ because f is continuous.

$$\lim_{x \to \xi} |g'(x)| \leq \lim_{x \to \xi} A|f(x)| = 0$$

Which immediately implies

$$\lim_{x \to \xi} g'(x) = 0$$

and this describes the behavior of g' as $x \to \xi$, which is what we were asked to describe.

Show that Newton's method involves finding a fixed point of g

This is a slightly modified version of lemma 3 for part (b). Suppose that we have found a fixed point κ such that $g(\kappa) = \kappa$. From the definition of g this would mean that

$$g(\kappa) - \kappa = -\frac{f(\kappa)}{f'(\kappa)} = 0$$

For this to hold it must either be the case that $f(\kappa) = 0$ or $f'(\kappa) = \pm\infty$: but we know that f' is bounded from the fact that f'' is bounded (mean value theorem : if f' were unbounded, then $\frac{f'(x)-f'(y)}{x-y} = f''(c)$ would be unbounded for some x, y). Therefore it must be the case that $f(\kappa) = 0$ and therefore κ must be the unique point for which $f(\kappa) = 0$.

Exercise 5.25f

We'll consider the more general case in which $f(x) = x^m$. In this case the single real zero occurs at $x = 0$. Newton's formula gives us the step function

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^m}{mx_n^{m-1}} = x_n - \frac{x_n}{m} = \left(\frac{m-1}{m} \right) x_n$$

This gives us a recursive definition of x_n.

$$x_{n+1} = \left(\frac{m-1}{m} \right) x_n = \left(\frac{m-1}{m} \right)^2 x_{n-1} = \left(\frac{m-1}{m} \right)^3 x_{n-2} = \cdots = \left(\frac{m-1}{m} \right)^n x_1$$

Taking the limit of each side as $n \to \infty$, we have

$$\lim_{n \to \infty} x_{n+1} = \lim_{n \to \infty} \left(\frac{m-1}{m} \right)^n x_1 \tag{45}$$

In order to have $\{x_n\}$ converge to 0 we must either have $x_1 = 0$ or $|m - 1/m| < 1$. The latter case occurs when

$$|m - 1| < |m|$$

This inequality holds iff $m > \frac{1}{2}$. If m is larger than $\frac{1}{2}$ the limit in equation (45) exists and is equal to zero and so $\{x_n\} \to 0$; if m is smaller than $\frac{1}{2}$ then the limit in equation (45) is unbounded and so $\{x_n\} \to \pm\infty$; if $m = \frac{1}{2}$ then equation (45) becomes

$$\lim_{n \to \infty} x_{n+1} = \lim_{n \to \infty} (-1)^n x_1$$

This limit clearly doesn't exist when $x_1 \neq 0$.

In the specific case given in the exercise we have $m = \frac{1}{3}$ and therefore the sequence $\{x_n\}$ fails to converge.

Exercise 5.26

Following the hint given in the exercise, let $M_0 = \sup|f(x)|$ and let $M_1 = \sup|f'(x)|$ for $x \in [a, x_0]$. We know that

$$|f(x)| \le M_1(x_0 - a) \tag{46}$$

because otherwise we'd have

$$\left|\frac{f(x)}{x_0 - a}\right| = \left|\frac{f(x) - f(a)}{x_0 - a}\right| = f'(c) \text{ for some } c \in (a, b) > M_1 = \sup|f'(x)|$$

which is clearly contradictory. Additionally, we know that

$$M_1 = \sup|f'(x)| \le \sup|Af(x)| = A\sup|f(x)| = AM_0 \tag{47}$$

Combining (46) and (47) gives us

$$|f(x)| \le M_1(x_0 - a) \le AM_0(x_0 - a) \quad, x \in [a, x_0] \tag{48}$$

From the fact that $f(a) = 0$ we know that $M_1(x_0 - a)$ is the maximum possible value that f could possibly obtain at $f(x_0)$ (otherwise, if $f(x_0) > M_1(x_0 - a)$, then the mean value theorem would give us some $f'(c) > M_1$). In comparison, M_0 is the maximum value that f *actually does* obtain at some $x \in [a, b]$. Therefore we have

$$M_0 \le M_1(x_0 - a) \tag{49}$$

Suppose $0 < A(x_0 - a) < \delta < 1$. Then (48) would give us

$$|f(x)| \le M_1(x_0 - a) \le AM_0(x_0 - a) \le \delta M_0$$

which contradicts (49) unless $M_1 = M_0 = 0$. And since we can force $0 < A(x_0 - a) < 1$ by choosing an appropriate x_0, it must be the case that $M_1 = M_0 = 0$. This turns (48) into

$$|f(x)| \le 0 \quad, x \in [a, x_0]$$

which shows that $f(x) = 0$ on the interval $[a, x_0]$. We can now repeat these steps using the interval $[x_0, b]$. We need only perform this a finite number of times (specifically, a maximum of of $[(b-a)/(x_0-a)]+1$ times) before we have covered the entire interval. Therefore $f(x) = 0$ on the entire interval $[a, b]$.

Exercise 5.27

The given hint is pretty much the entire solution. Let y_1 and y_2 be two solutions to the initial-value problem and define the function $f(x) = y_2(x) - y_1(x)$. The function f meets all of the prerequisites spelled out in exercise 5.26: it's differentiable, $f(a) = 0$, and we're told that there is a real number A such that

$$|f'(x)| = |y_2' - y_1'| = |\phi(x, y_2) - \phi(x, y_1)| \le A|y_2 - y_1| = A|f(x)|$$

Therefore, by exericse 5.26, we have $y_2(x) - y_1(x) = f(x) = 0$ for all x; therefore $y_1 = y_2$. But y_1 and y_2 were arbitrary solutions for the intial-value problem, so we have proven that there is only one unique solution.

Exercise 5.28

Let $\mathbf{y_a}$ and $\mathbf{y_b}$ be two solution vectors to the initial-value problem and define the vector-valued function $f(x) = \mathbf{y_b}(x) - \mathbf{y_a}(x)$. As in the previous problem, we see that $f(x)$ is differentiable and that $f(a) = a$. Therefore, if there is a real number A such that

$$|f'(x)| = |\mathbf{y_b'} - \mathbf{y_a'}| = |\phi(x, \mathbf{y_b}) - \phi(x, \mathbf{y_a})| \le A|y_2 - y_1| = A|f(x)|$$

then, by exercises 5.26 and 5.27 the initial-value problem has a unique solution.

Exercise 5.29

Note: there's probably a more subtle answer here, probably involving Taylor's theorem.

Let $\mathbf{v}$ and $\mathbf{v}$ be two solution vectors to the initial value problem and define the vector valued function $f(x) = \mathbf{w}(x) - \mathbf{v}(x)$:

$$f(x) = [w_1(x) - v_1(x), w_2(x) - v_2(x), \ldots, w_{k-1}(x) - v_{k-1}(x), w_k(x) - v_k(x)] \tag{50}$$

The derivative of this is given by

$$f'(x) = \left[w_1'(x) - v_1'(x), w_2'(x) - v_2'(x), \ldots, w_{k-1}'(x) - v_{k-1}'(x), w_k'(x) - v_k'(x)\right]$$

which, by the equivalences given in the exercise, becomes

$$f'(x) = \left[w_2(x) - v_2(x), w_3(x) - v_3(x), \ldots, w_k(x) - v_k(x), \sum_{j=1}^{k} g_j(w_j - v_j)\right] \tag{51}$$

From exercises 5.26-5.28 we know that we want an inequality of the form $|f'(x)| \leq A|f'(x)|$. Using (50) and (51) we see that this inequality will hold if there exists some A such that

$$|w_1(x) - v_1(x), w_2(x) - v_2(x), \ldots, w_{k-1}(x) - v_{k-1}(x), w_k(x) - v_k(x)|$$

$$\leq A\left|w_2(x) - v_2(x), w_3(x) - v_3(x), \ldots, w_k(x) - v_k(x), \sum_{j=1}^{k} g_j(w_j - v_j)\right|$$

The order of the components don't matter when we're taking the norm, and the rightmost term is a dot product, so this is equivalent to

$$|w_2(x) - v_2(x), w_3(x) - v_3(x), \ldots, w_k(x) - v_k(x), w_1(x) - v_1(x)| \leq A|w_2(x) - v_2(x), w_3(x) - v_3(x), \ldots, w_k(x) - v_k(x), g(x) \cdot (w($$

Exercise 6.1 (Proof 1)

Outline of the proof

We'll see that $L(P, f) = 0$ for every partition; therefore $\sup L(P, f) = 0$. By choosing an appropriate partition we can make $U(P, f)$ arbitrarily small; therefore $\inf U(P, f) = 0$. We conclude that $\int f = 0$.

sup L(P,f) = 0

Although it's easy to construct a partition such that $L(P, f) = 0$, we have to show that 0 is the supremum of the set of all $L(P, f)$. To do this, let P be an arbitrary partition of $[a, b]$ and let m_i be an arbitrary interval of this arbitrary partition. If m_i contains any point other than x_0 then $\inf f(x) = 0$ on this interval, so $m_i \Delta \alpha_i = 0$. If m_i contains only the point x_0 (that is, if the interval is $[x_0, x_0]$) then $\Delta \alpha_i = [\alpha(x_0) - \alpha(x_0)] = 0$ and therefore $m_i \Delta \alpha_i = 0$. Therefore $m_i \Delta \alpha_i = 0$ for all i, which means $L(P, f) = 0$. But P was an arbitrary partition, so $L(P, f) = 0$ for all P. Therefore

$$0 = \sup\{L(P, f) : P \text{ is a partition of } [a, b]\} = \int_{\underline{a}}^{b} f \, dx$$

inf U(P,f)=0

We need to show that for any $\epsilon > 0$ we can find some partition P such that $0 \leq U(P, f) < \epsilon$. So let $\epsilon > 0$ be given.

We're told that α is continuous at x_0, so by the definition of continuity we can find some $\delta > 0$ such that

$$d(x, x_0) < \delta \rightarrow d(\alpha(x), \alpha(x_0)) < \frac{\epsilon}{2} \tag{52}$$

Let $0 < \mu < \frac{\delta}{2}$ and let P be the partition $\{a, x_0 - \mu, x_0 + \mu, b\}$. We now calculate each M_i:

$$
\begin{array}{rclcl}
M_1 \Delta \alpha_1 & = & M_1[\alpha(x_0 - \mu) - \alpha(a)] & = & 0 \cdot [\alpha(x_0 - \mu) - \alpha(a)] = 0 \\
M_2 \Delta \alpha_2 & = & M_2[\alpha(x_0 + \mu) - \alpha(x_0 - \mu)] & = & 1 \cdot [\alpha(x_0 + \mu) - \alpha(x_0 - \mu)] \\
M_3 \Delta \alpha_3 & = & M_3[\alpha(b) - \alpha(x_0 + \mu)] & = & 0 \cdot [\alpha(b) - \alpha(x_0 - \mu)] = 0
\end{array}
$$

To determine a bound for $M_2 \Delta \alpha_2$ we apply (52) which allows us to conclude that $\alpha(x_0 + \mu) - \alpha(x_0 - \mu) < \epsilon$ from the fact that $d(x_0 - \mu, x_0 + \mu) = 2\mu < \delta$.

$$
M_2 \Delta \alpha_2 = \alpha(x_0 + \mu) - \alpha(x_0 - \mu) \le \epsilon
$$

From this we have

$$
U(P, f) \sum_{i=1^3} M_i \Delta \alpha_i = 0 + \epsilon + 0 = \epsilon
$$

But this epsilon is arbitrarily small, so

$$
0 = \inf\{U(P, f) : P \text{ is a partition of } [a, b]\} = \overline{\int_a^b} f \, dx
$$

Exercise 6.1 (Proof 2)

Thanks to Helen Barclay (hbarcla2@mail.usf.edu) for this proof. The function f is discontinuous at only one point and α is continous at that point, so $f \in \mathscr{R}$ by theorem 6.10. Every partition of $[a, b]$ will contain a point other than x_0 so $m_i = 0$ for every interval; therefore the infimum of $L(P, f)$ is zero; therefore

$$
\overline{\int_a^b} f \, d\alpha = \underline{\int_a^b} f \, d\alpha = 0
$$

Exercise 6.2

Outline of the proof

If we assume that $f(x) \ne 0$ for some $x \in [a, b]$, then we can construct a partition P for which $L(P, f) > 0$. From this we conclude that $\sup L(P, f) > 0$ and therefore $\int f \ne 0$.

The proof

Suppose, for purposes of contradiction, that $f(x_0) = \kappa$ for some $x_0 \in [a, b]$ where κ is some arbitrary nonzero number. We're told that f is continuous on a closed set, therefore f is uniformly continuous (theorem 4.19). Therefore there exists some $\delta > 0$ such that $|x_0 - x| < \delta \to |f(x_0) - f(x)| < \frac{\kappa}{2}$. Now let $0 < \mu < \delta$ and let P be the four-element partition $\{a, x_0 - \mu, x_0 + \mu, b\}$. For clarity, let $X = (x_0 - \mu, x_0 + \mu)$.

Since $|x_0 - x| \le \mu < \delta$ for all $x \in X$, we have $|f(x_0) - f(x)| = |\kappa - f(x)| < \frac{\kappa}{2}$ for all $x \in X$. For this inequality to hold for all $f(x)$ it must be the case that $\min f(X) > \frac{\kappa}{2}$. This means that

$$
L(P, f) = \sum m_i \Delta x_i \ge m_2 \Delta x_2 = \min f(X)[(x_0 + \mu) - (x_0 - \mu)] = \min f(X)(2\mu) > \frac{\kappa}{2}(2\mu) > 0
$$

Since $L(P, f) > 0$ for this particular P, it must be the case that $\sup L(P, f) > 0$ and therefore $\int f \, dx \ne 0$.

We assumed that $f(x) \ne 0$ for some $x \in [a, b]$ and concluded that $\int f \, dx \ne 0$: by contrapositive, if $\int f \, dx = 0$ then $f(x) = 0$ for all $x \in [a, b]$.

Exercise 6.3

Outline of the proof

We'll see that $U(P, f) - L(P, f)$ is equivalent to $M_i - m_i$ on a single arbitrarily small interval around 0. To make this difference arbitrarily small we will need to make $\sup f(x) - \inf f(x)$ arbitrarily small by restricting x to a sufficiently small neighborhood of 0: this is possible iff f is continuous at 0.

Lemma 1: For each of the β functions, the upper and lower Riemann sums depend entirely on the intervals containing 0

Let P be an arbitrary partition of $[-1, 1]$. By theorem 6.4 we can assume, without loss of generality, that 0 is an element of this partition. We define α to be the partition element immediately preceding 0 and define ω to be the partition element immediately following 0 (so our partition has the form $P = \{x_0 < x_1 < \ldots < \alpha < 0 < \omega < \ldots < x_n < x_{n+1}\}$). Note that α and ω must both exist since P is a finite set, although it's possible that $\alpha = -1$ or $\omega = 1$. For every $x_{i-1} < x_i < 0$ we have $\beta(x_{i-1}) = \beta(x_i) = 0$, therefore $\Delta\beta = 0$, therefore $M_i\Delta\beta = m_i\Delta\beta = 0$. Similarly, for every $0 < x_{i-1} < x_i$ we have $\beta(x_{i-1}) = \beta(x_i) = 1$, therefore $\Delta\beta = 0$, therefore $M_i\Delta\beta = m_i\Delta\beta = 0$.

This shows that $M_i\Delta\beta = m_i\Delta\beta = 0$ for every interval that doesn't contain 0. Therefore the values of $U(P, f)$ and $L(P, f)$ depend entirely on the intervals $[\alpha, 0]$ and $[0, \omega]$. This holds for arbitrary P.

The general form of $M_i\Delta\beta$ and $m_i\Delta\beta$ on $[\alpha, 0]$ and $[0, \omega]$

Let P be an arbitrary partition. Let M_α denote the supremum of $f(x)$ on the interval $[\alpha, 0]$; similarly define $m_\alpha, M_\omega, m_\omega$, etc. For all three of the β functions we have

$$M_\alpha\Delta\beta_\alpha = M_\alpha[\beta(0) - \beta(\alpha)] = M_\alpha\beta(0) \tag{53}$$
$$M_\omega\Delta\beta_\omega = M_\omega[\beta(\omega) - \beta(0)] = M_\omega[1 - \beta(0)] \tag{54}$$
$$m_\alpha\Delta\beta_\alpha = m_\alpha[\beta(0) - \beta(\alpha)] = m_\alpha\beta(0) \tag{55}$$
$$m_\omega\Delta\beta_\omega = m_\omega[\beta(\omega) - \beta(0)] = m_\omega[1 - \beta(0)] \tag{56}$$
$$\tag{57}$$

Case 1: $\beta(0) = 0$

When $\beta(0) = 0$ we have

$$\sum M_i\Delta\beta = M_\alpha\beta(0) + M_\omega[1 - \beta(0)] = M_\omega$$

$$\sum m_i\Delta\beta = m_\alpha\beta(0) + m_\omega[1 - \beta(0)] = m_\omega$$

Proof that continuity of f implies integrability: let $\epsilon > 0$ be given. We're told that $f(0+) = f(0)$, so by definition of continuity at this point there exists some $\delta > 0$ such that

$$d(0, x) < \delta \rightarrow d(f(0), f(x)) < \frac{\epsilon}{2}$$

Construct a partition $P = \{-1, 0, \delta, 1\}$. The interval $[0, \delta]$ is compact, therefore by theorem 4.16 the points $f^{-1}(M_\omega)$ and $f^{-1}(m_\omega)$ both exist in $[0, \delta]$ and therefore:

$$d(f^{-1}(M_\omega), 0) < \delta \rightarrow d(M_\omega, f(0)) < \frac{\epsilon}{2}$$

$$d(f^{-1}(m_\omega), 0) < \delta \rightarrow d(m_\omega, f(0)) < \frac{\epsilon}{2}$$

and therefore, by the triangle inequality,

$$d(M_\omega, m_\omega) \leq d(M_\omega, f(0)) + d(f(0), m_\omega) < \epsilon$$

and therefore

$$U(P, f) - L(P, f) = M_\omega - m_\omega < \epsilon$$

And since ϵ was arbitrary, this is sufficient to show that $\int f\, dx$ exists. $\qquad\square$

Proof by contrapositive that integrability of f implies right-hand continuity: Assume that $f(0+) \neq f(0)$. From the negation of the definition of continuity we therefore know that there exists some $\epsilon > 0$ such that for all δ we can find some $0 < x < \delta$ such that $|f(x) - f(0)| \geq \epsilon$. So for any partition P let a be a point at which $f(a) = M_\omega$ and let b be a point for which $|f(b) - f(0)| \geq \epsilon$. This gives us $f(b) \leq f(a) = M_\omega$ and $m_\omega \leq f(0)$, therefore

$$U(P, f) - L(P, f) = M_\omega - m_\omega = f(a) - m_\omega \geq f(a) - f(0) \geq f(b) - f(0) \geq \epsilon$$

The partition P was arbitrary so this inequality must be true for all possible partitions, so we can never find P such that

$$\inf U(P, f) - \sup L(P, f) \le \epsilon$$

and therefore $\int f \, dx$ does not exist. □

Proof that this integral, if it exists, is equal to $f(0)$: from the fact that 0 is an element of $[0, \delta]$ we know that $M_\omega \ge f(0)$. We also know that $L(P, f) \le \int f$ (by theorem 6.4 and/or the definition of $\int f$). This gives us the inequalities

$$M_\omega = U(P, f) \ge f(0)$$

$$L(P, f) \le \int f \, d\beta$$

from which we have

$$U(P, f) - L(P, f) \ge f(0) - \int f \, d\beta \tag{58}$$

But we also know that $m_\omega \le f(0)$ and that $U(P, f) \ge \int f$. This gives us

$$U(P, f) \ge \int f \, d\beta$$

$$m_\omega = L(P, f) \le f(0)$$

from which we have

$$U(P, f) - L(P, f) \ge \int f \, d\beta - f(0) \tag{59}$$

Combining (58) and (59) gives us

$$\left| \int f \, d\beta - f(0) \right| \le U(P, f) - L(P, f) = \epsilon$$

If this is to be true for all $\epsilon > 0$ we must have $\int f \, d\beta = f(0)$. □

Case 2: $\beta(0) = 1$

When $\beta(0) = 1$ we have

$$\sum M_i \Delta \beta = M_\alpha \beta(0) + M_\omega [1 - \beta(0)] = M_\alpha$$

$$\sum m_i \Delta \beta = m_\alpha \beta(0) + m_\omega [1 - \beta(0)] = m_\alpha$$

From here the proof that $\int f \, d\beta = f(0)$ iff $f(0) = f(0-)$ is almost identical to the previous proof with α in the place of ω.

Case 3: $\beta(0) = \frac{1}{2}$

When $\beta(0) = 1/2$ we have

$$\sum M_i \Delta \beta = M_\alpha \beta(0) + M_\omega [1 - \beta(0)] = \frac{1}{2}[M_\alpha + M_\omega]$$

$$\sum m_i \Delta \beta = m_\alpha \beta(0) + m_\omega [1 - \beta(0)] = \frac{1}{2}[m_\alpha + m_\omega]$$

Therefore

$$U(P, f) - L(P, f) = \frac{1}{2}[M_\alpha - m_\alpha] + \frac{1}{2}[M_\omega - m_\omega]$$

The function f is integrable iff this term can be made arbitrarily small; we saw in case (1) that $[M_\omega - m_\omega]$ can be made arbitrarily small iff $f(0) = f(0+)$, and we saw in case (2) that $[M_\alpha - m_\alpha]$ can be made arbitrarily small iff $f(0) = f(0-)$. Therefore we can minimize $U(P, f) - L(P, f)$ in this case iff $f(0+) = f(0-) = f(0)$; that is, iff f is continuous at 0. □

Proof that this integral, if it exists, is equal to $f(0)$: from the fact that 0 is an element of $[0, \delta]$ we know that $M_\omega \geq f(0)$ and $M_\alpha \geq f(0)$. We also know that $L(P, f) \leq \int f$ (by theorem 6.4 and/or the definition of $\int f$). This gives us the inequalities

$$\frac{1}{2}[M_\alpha + M_\omega] = U(P, f) \geq \frac{1}{2}[f(0) + f(0)] = f(0)$$

$$L(P, f) \leq \int f \, d\beta$$

from which we have

$$U(P, f) - L(P, f) \geq f(0) - \int f \, d\beta \tag{60}$$

But we also know that $m_\omega \leq f(0)$ and that $m_\alpha \leq f(0)$ and that $U(P, f) \geq \int f$. This gives us

$$U(P, f) \geq \int f \, d\beta$$

$$\frac{1}{2}[m_\alpha + m_\omega] = L(P, f) \leq \frac{1}{2}[f(0) + f(0)] = f(0)$$

from which we have

$$U(P, f) - L(P, f) \geq \int f \, d\beta - f(0) \tag{61}$$

Combining (58) and (59) gives us

$$\left| \int f \, d\beta - f(0) \right| \leq U(P, f) - L(P, f) = \epsilon$$

If this is to be true for all $\epsilon > 0$ we must have $\int f \, d\beta = f(0)$. $\qquad \square$

Part (d) of the exercise

If f is continuous at 0 then we have $f(0) = f(0+) = f(0-)$, so by case (1) we have $\int f \, d\beta_1 = f(0)$ and by case (2) we have $\int f \, d\beta_2 = f(0)$ and by case (3) we have $f \, d\beta_3 = f(0)$.

Exercise 6.4

Let P be an arbitrary partition of $[a, b]$. Let $n = |P| - 1$, so that n represents the number of intervals in the partition P. Every interval will contain at least one rational number (theorem 1.20b) and at least one irrational number (by pigeonhole principle, since $|(x_i, x_{i+1})| = |\mathbb{R}| > |\mathbb{Q}|$). Therefore we have $M_i = 1, m_i = 0$ for all i. This gives us

$$\sum M_i \Delta x_i = \sum_1^n 1 \cdot (x_{i+1} - x_i) = x_{n+1} - x_0 = b - a$$

$$\sum m_i \Delta x_i = \sum_1^n 0 \cdot (x_{i+1} - x_i) = 0$$

But P was an arbitrary partition, so this holds for all partitions, and therefore

$$\sup L(P, f) = 0, \quad \inf U(P, f) = 1$$

and therefore $f \notin \mathscr{R}$.

Exercise 6.5

Is $f \in \mathscr{R}$ if $f^2 \in \mathscr{R}$?

No. Consider the function

$$f(x) = \begin{cases} -1, & x \in \mathbb{Q} \\ 1, & x \notin \mathbb{Q} \end{cases}$$

We saw that $f \notin \mathscr{R}$ in exercise 6.4, but clearly $f^2(x) = 1 \in \mathscr{R}$.

Is $f \in \mathscr{R}$ if $f^2 \in \mathscr{R}$?

Yes. The function $\phi(x) = \sqrt[3]{x}$ is a continuous function, so we let $h(x) = \phi(f^3(x)) = f(x)$ and appeal to theorem 6.11 to claim that

$$f^3 \in \mathscr{R} \to h \in \mathscr{R} \to f \in \mathscr{R}$$

Note that the claim that $h(x) = \phi(f^3(x)) = f(x)$ relied on the fact that $x \to x^3$ is a one-to-one mapping, so that $\sqrt[3]{x^3} = x$. In contrast, the mapping $x \to x^2$ is *not* one-to-one as can be seen by the fact that $\sqrt{(-1)^2} \neq -1$.

Exercise 6.6

Outline of the proof

We can define a partition for which f is discontinuous on 2^n intervals each of which has length 3^{-n}, so the Riemann sum across these intervals is proportional to $(2/3)^n$. This sum approaches 0 as $n \to \infty$.

The proof

Let E_i be defined as in sec. 2.44. Note that E_0 consists of one interval of length 1; E_1 consists of two intervals of length $\frac{1}{3}$; E_2 consists of 4 intervals of length $\frac{1}{9}$; and, in general, E_n will consist of 2^n intervals of length 3^{-n}.

With each E_n we can associate a set F_n that contains the endpoints of the intervals contained in E_n. That is, we define

$$F_n = \{a_1 < b_1 < a_2 < b_2 < \ldots < a_{2^n} < b_{2^n}\}$$

where each $[a_i, b_i]$ is an interval in E_n. Since every point of the Cantor set – and therefore every point at which f might be discontinuous – is in an interval of the form $[a_i, b_i]$ we'll choose a partition that lets us isolate these intervals.

Let m and M represent the lower and upper bounds of f on $[0, 1]$. Choose an arbitrary $\epsilon > 0$. Choose n large enough that

$$3^{n-1} > \frac{2^{n+1}(M - m)}{\epsilon}$$

and choose δ such that

$$0 < \delta < \frac{1}{3^{n+1}}$$

These choice will be justified later. Define P_n to be

$$P_n = \{0 = a_1 < (b_1 + \delta) < (a_2 - \delta) < (b_2 + \delta) < (a_3 - \delta) < (b_3 + \delta) < \ldots < (a_{2^n} - \delta) < b_{2^n} = 1\}$$

This partition contains 2^{n+1} points and therefore contains $2^{n+1} - 1$ segments. We must now show that we can make $U(P_n, f) - L(P_n, f)$ arbitrarily small.

$$U(P_n, f) - L(P_n, f) = \sum_{i=1}^{2^{n+1}-1} (M_i - m_i)\Delta x_i$$

We'll separate this sum into the intervals on which f is continuous ($i = 2, 4, 6, \ldots$) and the intervals on which f might contain discontinuities ($i = 1, 3, 5, \ldots$).

$$= \sum_{i=1}^{2^n}(M_{2i} - m_{2i})\Delta x_{2i} + \sum_{i=0}^{2^n}(M_{2i+1} - m_{2i+1})\Delta x_{2i+1}$$

The function f is continuous on every interval of the form $[b_i + \delta, a_{i+1} - \delta]$, and these intervals are represented by the lefthand summation. We can therefore refine P_n such that

$$\leq \frac{\epsilon}{2} + \sum_{i=1}^{2^n}(M_{2i+1} - m_{2i+1})\Delta x_{2i+1}$$

We know the exact value of the Δx terms. For P_n, we have $d(a_i, b_i) = 3^{-n}$ and therefore $\Delta x_{2i+1} = 3^{-n} + 2\delta$. Similarly, we know that $M_i \le M$ and $m_i \ge m$, so we have

$$\le \frac{\epsilon}{2} + \sum_{i=1}^{2^n} (M - m)\left(\frac{1}{3^n} + 2\delta\right)$$

From our choice of $\delta < \frac{1}{3^{n+1}} < \frac{1}{3n}$ this becomes

$$\le \frac{\epsilon}{2} + \sum_{i=1}^{2^n} (M - m)\left(\frac{1}{3^{n-1}}\right)$$

This summation is constant with respect to i so it becomes simply

$$= \frac{\epsilon}{2} + 2^n(M - m)\left(\frac{1}{3^{n-1}}\right)$$

From our choice of n we have $3^{n-1} > \frac{2^{n+1}(M-m)}{\epsilon}$ and therefore $3^{-(n-1)} < \epsilon/2^{n+1}(M - m)$.

$$\le \frac{\epsilon}{2} + \frac{2^n(M - m)\epsilon}{2^{n+1}(M - m)}$$

$$=\le \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Exercise 6.7a

If $f \in \mathscr{R}$, then from theorem 6.12 we have

$$\int_0^1 f\, dx = \int_0^c f\, dx + \int_c^1 f\, dx \tag{62}$$

Now consider the partition P of $[0, c]$ defined by $P = \{0, c\}$. This partition has only a single interval, so we have

$$\inf_{0<x<c} f(x) \cdot c = L(P, f) \le \int_0^c f\, dx \le U(P, f) = \sup_{0<x<c} f(x) \cdot c$$

From (62), adding $\int_c^1 f\, dx$ to each term of this inequality gives us

$$\int_c^1 f\, dx + \inf_{0<x<c} f(x) \cdot c \le \int_0^1 f\, dx \le \int_c^1 f\, dx + \sup_{0<x<c} f(x) \cdot c$$

Taking the limit of each term of this inequality as $c \to 0$, we see that $f(x) \cdot c \to 0$ while the center term remains constant. The resulting inequality is

$$\lim_{c \to 0} \int_c^1 f\, dx \le \int_0^1 f\, dx \le \lim_{c \to 0} \int_c^1 f\, dx$$

which of course implies that

$$\lim_{c \to 0} \int_c^1 f\, dx = \int_0^1 f\, dx$$

which is what we wanted to prove.

Exercise 6.7b

Outline of the proof

We construct a function f for which $\int f = \sum_{i=1}^n \frac{(-1)^i}{i}$: this is the alternating harmonic series, which converges (see Rudin's remark 3.46). We then see that $\int |f| = \sum_{i=1}^n \frac{1}{i}$: this is the harmonic series, which diverges (theorem 3.28).

The proof

Choose any $c \in (0, 1)$ and consider the following function defined on $[c, 1]$.

$$f(x) = \begin{cases} (-1)^n(n+1), & \frac{1}{n+1} < x < \frac{1}{n} \text{ for some } n \in \mathbb{N} \\ 0, & \text{otherwise} \end{cases}$$

Claim 1 : f is integrable on any interval $[c, 1]$

Choose an arbitrarily small $\epsilon > 0$. Let $\frac{1}{N}$ be the smallest harmonic number greater than c. We want to choose δ such that each harmonic number in $[c, 1]$ is contained in a distinct interval of radius δ: so choose δ such that $0 < \delta < \frac{1}{2}[\frac{1}{N} - c]$. We must also make sure that $\delta < \frac{\epsilon}{2(N+1)(N+2)}$ for reasons that will become clear later. Let H_c represent the partition containing the elements $\{c\} \cup \{\frac{1}{n} \pm \delta : n \in \mathbb{N}, n < \frac{1}{c}\} \cap [0, 1]$.

- The partition H_c contains one interval of the form $[c, \frac{1}{N} - \delta]$.

$$M_1 \Delta x_1 = \sup f(x)\Delta x_1 = (-1)^N(N+1)\Delta x_1 = (-1)^N(N+1)\left(\frac{1}{N} - \delta - \frac{1}{N+1}\right)$$

$$= (-1)^N(N+1)\left(\frac{1}{N(N+1)} - \delta\right) = (-1)^N\left(\frac{1}{N} - (N+1)\delta\right)$$

The function is constant over this interval, so we have

$$m_1 \Delta x_1 = \inf f(x)\Delta x_1 = \sup f(x)\Delta x_1 = M_1$$

therefore

$$|M_1 - m_1|\Delta x_1 = 0 \tag{63}$$

- The partition H_c contains one interval of the form $[1 - \delta, 1]$ for which

$$M_n \Delta x_n = \sup f(x)\Delta x_n = 0\Delta x = 0$$

$$m_n \Delta x_n = \inf f(x)\Delta x_n = -2\Delta x = -2\delta$$

therefore

$$|M_n - m_n|\Delta x_n = 2\delta \tag{64}$$

- The partition contains $N - 1$ intervals of the form $\left[\frac{1}{i} - \delta, \frac{1}{i} + \delta\right]$ for which

$$M_i \Delta x_i = \sup f(x)\Delta x_i \leq \sup |f(x)|2\delta \leq |i+1|2\delta$$

$$m_i \Delta x_i = \inf f(x)\Delta x_i \geq \inf -|f(x)|2\delta \geq -|i+1|2\delta$$

therefore

$$|M_i - m_i|\Delta x_i \leq |M_i| + |m_i| \leq 4\delta|i+1| \tag{65}$$

- The partition contains $N - 2$ intervals of the form $\left[\frac{1}{j+1} + \delta, \frac{1}{j} - \delta\right]$ for which

$$M_j \Delta x_j = \sup f(x)\Delta x_j = (-1)^j(j+1)\left(\frac{1}{j} - \delta - \frac{1}{j+1} - \delta\right) = (-1)^j(j+1)\left(\frac{1}{j(j+1)} - 2\delta\right) = (-1)^j\left(\frac{1}{j} - 2\delta(j+1)\right)$$

The function is constant over this interval, so we have

$$m_j \Delta x_j = \inf f(x)\Delta x_j = \sup f(x)\Delta x = M = (-1)^j\left(\frac{1}{j} - 2\delta(j+1)\right)$$

therefore

$$(M_j - m_j)\Delta x_j = 0 \tag{66}$$

106

Summing equations (63)-(66), we have

$$
\begin{aligned}
|U(P,f) - L(P,f)| &= |(M_1 - m_1)\Delta x_1 + (M_n - m_n)\Delta x_n + \sum_{i=2}^{N}(M_i - m_i)\Delta x_i + \sum_{j=2}^{N-1}(M_j - m_j)\Delta x_j| \\
&\leq |(M_1 - m_1)|\Delta x_1 + |(M_n - m_n)|\Delta x_n + \sum_{i=2}^{N}|M_i - m_i|\Delta x_i + \sum_{j=2}^{N-1}|M_j - M_j|\Delta x_j \\
&\leq 0 + 2\delta + \sum_{i=2}^{N}4\delta|i+1| + \sum_{j=2}^{N-1}0 \\
&= 2\delta + 4\delta\,\frac{(N+1)(N+2) - 1}{2} \\
&= 2\delta(N+1)(N+2)
\end{aligned}
$$

Earlier in the proof we chose δ such that $\delta < \frac{\epsilon}{2(N+1)(N+2)}$, so we obtain the inequality

$$|U(P,f) - L(P,f)| \leq \epsilon$$

And ϵ was arbitrary, so this proves that $f \in \mathscr{R}$.

Claim 2 : The value of $\lim_{c \to 0}\int_c^1 f$ is defined

Adding the values for the M terms, we have

$$
\begin{aligned}
U(P,f) &= M_1 + M_n + \sum_{i=2}^{N}M_i + \sum_{i=2}^{N-1}M_j \\
&\leq (-1)^N\left(\frac{1}{N} - (N+1)\delta\right) + 0 + \sum_{i=2}^{N}|i+1|2\delta + \sum_{j=2}^{N-1}(-1)^j\left(\frac{1}{j} - 2\delta(j+1)\right) \\
&= (-1)^N\left(\frac{1}{N} - (N+1)\delta\right)\delta[(N+1)(N+2) - 6] + \sum_{j=2}^{N-1}(-1)^j\left(\frac{1}{j} - 2\delta(j+1)\right)
\end{aligned}
$$

Remember that we first chose a value for c, which forced us to use a particular value for N; our choice of δ came afterward, so we're still free to select δ small enough to make this last inequality become:

$$U(P_c, f) \leq \frac{(-1)^N}{N} + \sum_{j=2}^{N-1}(-1)^j\frac{1}{j} + \epsilon$$

We defined $\frac{1}{N}$ to be the smallest harmonic number greater than c, so $\frac{1}{N} \to 0$ and $N \to \infty$ as $c \to 0$. Therefore taking the limit of both sides as $c \to 0$ gives us

$$\lim_{c \to 0}U(P_c, f) = \sum_{j=2}^{\infty}(-1)^j\frac{1}{j} + \epsilon = 1 - \ln(2) + \epsilon$$

which proves that

$$\lim_{c \to 0}\int_c^{\overline{1}} f\,dx \geq 1 - \ln(2)$$

The values for the m terms are almost identical to the M terms, and adding them gives us the inequality

$$\lim_{c \to 0}\int_{\underline{c}}^1 f\,dx \leq 1 - \ln(2)$$

Together, of course, these two inequalities prove that

$$\lim_{c \to 0}\int_c^1 f\,dx = 1 - \ln(2)$$

107

Claim 3 : The value of $\lim_{c \to 0} \int_c^1 |f|$ is not defined

If we follow the logic from claims 1 and 2, we find that adding the values for the M terms gives us

$$
\begin{aligned}
U(P, f) &= M_1 + M_n + \sum_{i=2}^{N} M_i + \sum_{i=2}^{N-1} M_j \\
&\leq \left(\frac{1}{N} - (N+1)\delta \right) + \frac{1}{2} + \sum_{i=2}^{N} |i+1| 2\delta + \sum_{j=2}^{N-1} \left(\frac{1}{j} - 2\delta(j+1) \right) \\
&= \left(\frac{1}{N} - (N+1)\delta \right) + \frac{1}{2}\delta[(N+1)(N+2) - 6] + \sum_{j=2}^{N-1} \left(\frac{1}{j} - 2\delta(j+1) \right)
\end{aligned}
$$

Again following the logic of part 2, we select δ small enough to gives us

$$
\lim_{c \to 0} \int_c^1 f \, dx = \sum_{j=2}^{\infty} \frac{1}{j} + \epsilon
$$

But we know from chapter 3 that this series doesn't converge, so the lefthand limit doesn't exist and therefore $\lim_{c \to 0} \int_c^1 f \, dx$ does not exist.

Exercise 6.8

Choose an arbitrary real number $c > 1$. Let n be the greatest integer such that $n < c$. Define the partition P_n of $[0, n]$ to be $P_n = \{x_0 = 1, 2, 3, 4, \ldots, n = x_{n-1}\}$.

First, assume that $\sum f(n)$ diverges. Because $f(x) > 0$ this means that $\sum f(n)$ is unbounded (theorem 3.24). The fact that f is monotonically decreasing allows us to derive the following chain of inequalities.

$$
\begin{aligned}
\int_1^c f(x) \, dx &\geq \int_1^n f(x) \, dx \\
&\geq L(P_n, f) \\
&= \sum_{i=1}^{n-2} m_i \Delta x_i \\
&= \sum_{i=1}^{n-2} f(i+1)
\end{aligned}
$$

(Note: the index is going to $n-2$ instead of $n-1$ because of the awkward numbering of the partition: note that the partition only goes to x_{n-1}). Taking the limit as $c \to \infty$, we have

$$
\int_1^\infty f(x) \, dx = \lim_{c \to \infty} \int_1^c f(x) \, dx \geq \lim_{n \to \infty} \sum_{i=2}^{n} f(i+1)
$$

The integral on the left-hand side of this chain of inequalities is greater than the unbounded sum on the right-hand side, so we conclude that the integral does not converge.

Next assume that $\sum f(n)$ converges to some $\kappa \in \mathbb{R}$. Choose c, n, P_{n+1} as defined above. The fact that f is monotonically decreasing allows us to derive the following chain of inequalities.

$$
\begin{aligned}
\int_1^c f(x) \, dx &\leq \int_1^{n+1} f(x) \, dx \\
&\leq U(P_{n+1}, f) \\
&= \sum_{i=1}^{n-1} M_i \Delta x_i \\
&= \sum_{i=1}^{n-1} f(i)
\end{aligned}
$$

Taking the limit as $c \to \infty$, we have

$$\int_1^\infty f(x)\, dx = \lim_{c \to \infty} \int_1^c f(x)\, dx \leq \lim_n \to \infty \sum_{i=1}^{n-1} f(i)$$

We're told that $f(x) > 0$, so we know that $\int_0^c f(x)$ is a monotonically increasing function of c. The previous inequality tells us that $\int_1^c$ is bounded above. Therefore by theorem 3.14 we know that $\lim_{c \to \infty} \int_1^c f\, dx$ is defined.

Exercise 6.9

By the definition given in exercise 6.8:

$$\int_0^\infty f(x)g'(x)\, dx = \lim_{c \to \infty} \int_0^c f(x)g'(x)\, dx$$

The integral from 0 to c is finite, so we can use integration by parts:

$$\int_0^\infty f(x)g'(x)\, dx = \lim_{c \to \infty} f(c)g(c) - f(0)g(0) - \int_0^c f'(x)g(x)\, dx$$

Applying this to the given function with $f(x) = (1+x)^{-1}$ and $g(x) = \sin x$, we have $f'(x) = -(1+x)^{-2}$ and $g'(x) = \cos x$. Integration by parts therefore yields

$$\int_0^\infty \frac{\cos x}{1+x}\, dx = \lim_{c \to \infty} \frac{\sin c}{1+c} - \frac{\sin 0}{1} + \int_0^c \frac{\sin x}{(1+x)^2}\, dx$$

Since $-1 \leq \sin c \leq 1$, the non-integral terms both tend to zero as $c \to \infty$. This leaves us with

$$\int_0^\infty \frac{\cos x}{1+x}\, dx = \lim_{c \to \infty} \int_0^c \frac{\sin x}{(1+x)^2}\, dx$$

By the definition in exercise 6.8 this is equivalent to

$$\int_0^\infty \frac{\cos x}{1+x}\, dx = \int_0^\infty \frac{\sin x}{(1+x)^2}\, dx$$

Exercise 6.10 a: method 1

This proof was due to Helen Barclay (hbarcla2@mail.usf.edu). We rewrite uv as the exponential of a natural log:

$$uv = (u^p)^{1/p}(v^q)^{1/q} = \exp(\ln((u^p)^{1/p}(v^q)^{1/q})) = \exp\left(\frac{1}{p}\ln u^p + \frac{1}{q}\ln v^q\right)$$

The exponential function $f(x) = e^x$ has a strictly positive second derivative; therefore its first derivative is monotonically increasing (theorem 5.11); therefore $f(x)$ is a convex function (proof in exercise 5.14, definition of convex in exercise 4.23). By definition of convexity with $\lambda = 1/p, 1 - \lambda = 1/q$ we have

$$\exp\left(\frac{1}{p}\ln u^p + \frac{1}{q}\ln v^q\right) \leq \frac{1}{p}\exp\left(\ln u^p\right) + \frac{1}{q}\exp\left(v^q\right) = \frac{u^p}{p} + \frac{v^q}{q}$$

Combining these last two inequalities, we have

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

which is what we were asked to prove.

Exercise 6.10 a: method 2

Using the variable substitutions $s = u^p, t = v^q, a = \frac{1}{p}, (1 - a) = \frac{1}{q}$ we can rewrite

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q} \tag{67}$$

as

$$s^a t^{1-a} \leq as + (1-a)t$$

Multiplying both sides by $s^{-a}t^{a-1}$ gives us

$$1 \leq as^{1-a}t^{a-1} + (1-a)s^{-a}t^a$$

We can rewrite this previous inequality in two equivalent ways:

$$1 \leq a\left(\frac{t}{s}\right)^{a-1} + (1-a)\left(\frac{t}{s}\right)^a \tag{68}$$

$$1 \leq a\left(\frac{s}{t}\right)^{1-a} + (1-a)\left(\frac{s}{t}\right)^{-a} \tag{69}$$

If $s \leq t$ then $t/s \geq 1$ and therefore

$$a\left(\frac{t}{s}\right)^{a-1} + (1-a)\left(\frac{t}{s}\right)^a \geq a(1)^{a-1} + (1-a)(1)^a = 1$$

so the inequality in (68) holds. If $s \geq t$ then $s/t \geq 1$, and therefore

$$a\left(\frac{s}{t}\right)^{1-a} + (1-a)\left(\frac{s}{t}\right)^{-a} \geq a(1)^{1-a} + (1-a)(1)^{-a} = 1$$

so the inequality in (69) holds. But both of these inequalities are just (67) with a variable change, so we see that (67) holds when $s \geq t$ or when $t \geq s$ – that is, it *always* holds. And this is what we were asked to prove.

Exercise 6.10 a: method 3

This proof was due to Boris Shektman (don't email him). Define the function $f(u)$ to be

$$f(u) = \frac{u^p}{p} + \frac{v^q}{q} - uv$$

This function has the derivative

$$f'(u) = u^{p-1} - v$$

Note that $f'(u) = 0$ has only one positive real solution, which occurs at $u = v^{1/(p-1)}$. We can show algebraically that $\frac{1}{p-1} = \frac{q}{p}$. For $u < v^{q/p}$ we have $f'(u) < 0$ and for $u > v^{q/p}$ we have $f'(u) > 0$, so the the point $u = v^{q/p}$ is the unique global minimum of $f(u)$. Evaluating the function at this value of u gives us

$$f(v^{q/p}) = \frac{v^q}{p} + \frac{v^q}{q} - v^{q/p}v$$

$$= \left(\frac{1}{p} + \frac{1}{q}\right)v^q - (v^{(p+q)/p})$$

We're given that $1/p + 1/q = 1$ and, from this, we also know that $p + q = pq$. Making these substitutions gives us

$$= v^q - v^{pq/p} = v^q - v^q 0$$

Therefore the unique minimum of f is $f(v^{q/p}) = 0$, and $f(u) \geq 0$ for all other u.

Exercise 6.10b

From part (a), we have

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

Exercise 6.10c

Define κ and λ to be

$$\kappa = \int_a^b |f|^p \, d\alpha, \quad \lambda = \int_a^b |g|^q \, d\alpha$$

Define $\hat{f}$ and $\hat{g}$ to be

$$\hat{f}(x) = \frac{f(x)}{\kappa^{1/p}}, \quad \hat{g}(x) = \frac{g(x)}{\lambda^{1/q}},$$

These two functions are "normalized" in the sense that

$$\int_a^b |\hat{f}|^p \, d\alpha = \int_a^b \frac{|f|^p}{\kappa} \, d\alpha = \frac{1}{\kappa} \int_a^b |f|^p \, d\alpha = \frac{\kappa}{\kappa} = 1$$

$$\int_a^b |\hat{g}|^q \, d\alpha = \int_a^b \frac{|g|^q}{\lambda} \, d\alpha = \frac{1}{\lambda} \int_a^b |g|^q \, d\alpha = \frac{\lambda}{\lambda} = 1$$

We know that $\hat{f}$ and $\hat{g}$ are Riemann integrable by theorem 6.11, and therefore by theorem 6.13 and part (b) of this exercise we know that $|\hat{f}|$ and $|\hat{g}|$ are Riemann integrable and that

$$\left| \int_a^b \hat{f}\hat{g} \, d\alpha \right| \le \int_a^b |\hat{f}||\hat{g}| \, d\alpha \le 1$$

By definition of $\hat{f}$ and $\hat{g}$ this becomes:

$$\left| \int_a^b \frac{f}{\kappa^{1/p}} \frac{g}{\lambda^{1/q}} \, d\alpha \right| \le \int_a^b \frac{|f|}{\kappa^{1/p}} \frac{|g|}{\lambda^{1/q}} \, d\alpha \le 1$$

Multiplying both sides by the constant $\kappa^{1/p}\lambda^{1/q}$ gives us

$$\left| \int_a^b fg \, d\alpha \right| \le \int_a^b |fg| \, d\alpha \le \kappa^{1/p}\lambda^{1/q}$$

which, by the definition of κ and λ, is equivalent to

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

which is what we wanted to prove.

Exercise 6.10d

Proof by contrapositive. If the inequality were false for the improper integrals, we would be able to find some function f such that

$$\lim_{c \to \infty} \left| \int_a^c fg \, d\alpha \right| > \lim_{c \to \infty} \left\{ \int_a^c |f|^p \, d\alpha \right\}^{1/p} \left\{ \int_a^c |g|^q \, d\alpha \right\}^{1/q}$$

or

$$\lim_{c \to x_0} \left| \int_{x_0}^b fg \, d\alpha \right| > \lim_{c \to x_0} \left\{ \int_{x_0}^b |f|^p \, d\alpha \right\}^{1/p} \left\{ \int_{x_0}^b |g|^q \, d\alpha \right\}^{1/q}$$

But this is a strict inequality, so in either of these cases we would have to be able to find a neighborhood of ∞ or x_0 for which this inequality still held. And this would give us a *proper* integral for which Holder's inequality doesn't hold. By contrapositive, the fact that Holder's inequality is valid for proper integrals shows that it's true for improper integrals.

Exercise 6.11

If we can prove that $||f+g|| \leq ||f||+||g||$, then we can prove the given inequality by letting $f = f-g$ and $g = g-h$.

$$||f+g||^2 = \int_a^b |f+g|^2 \, d\alpha \qquad\qquad\qquad \text{definition of } ||\cdot||$$

$$\leq \int_a^b (|f|+|g|)^2 \, d\alpha \qquad\qquad\qquad \text{The triangle inequality is established for } |\cdot|$$

$$= \int_a^b |f|^2 + 2|fg| + |g|^2 \, d\alpha$$

$$= \int_a^b |f|^2 \, d\alpha + 2\int_a^b |fg| \, d\alpha + \int_a^b |g|^2 \, d\alpha \qquad \text{properties of integrals: theorem 6.12}$$

$$\leq \int_a^b |f|^2 \, d\alpha + 2\left\{\int_a^b |f|^2 \, d\alpha\right\}^{1/2} \left\{\int_a^b |g|^2 \, d\alpha\right\}^{1/2} + \int_a^b |g|^2 \, d\alpha \quad \text{Holder's inequality (exercise 6.10)}$$

$$= ||f||^2 + 2||f|| \, ||g|| + ||g||^2 \qquad\qquad \text{definition of } ||\cdot||$$

$$= (||f|| + ||g||)^2$$

Taking the square root of the first and last terms of this chain of inequalities, we have

$$||f+g|| \leq ||f|| + ||g||$$

This inequality must hold for any $f, g \in \mathscr{R}$. Letting $f = f - g$ and $g = g - h$ this becomes

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

which is what we were asked to prove.

Exercise 6.12

We're told that f is continuous, so it has upper and lower bounds $m \leq f(x) \leq M$. Choose any $\epsilon > 0$. Since $f \in \mathscr{R}(\alpha)$, we can find some partition P such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \frac{\epsilon^2}{M-m}$$

Having now determined a particular partition P we can define g as suggested in the hint. This function is simply a series of straight lines connecting $f(x_{i-1})$ to $f(x_i)$. This becomes clearer if we rewrite g in an algebraically equivalent form:

$$g(t) = f(x_{i-1}) + \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}(t - x_{i-1})$$

We see that on any interval $[x_{i-1}, x_i]$ the function $g(t)$ is bounded between $f(x_{i-1})$ and $f(x_i)$. That is,

$$m_i \min\{f(x_{i-1}), f(x_i)\} \leq g(t), t \in [x_{i-1}, x_i] \leq \max\{f(x_{i-1}), f(x_i)\} \leq M_i$$

And of course we have similar bounds on f:

$$m_i \leq f(t) \leq M_i$$

and therefore, since $f(t), g(t) \leq M_i$ and $-f(t), -g(t) \leq -m_i$, we have

$$|f(t) - g(t)| \leq |M_i - m_i| = M_i - m_i \tag{70}$$

Similarly, since $M_i \leq M$ and $-m_i \leq -m$ we have

$$M_i - m_i \leq M - m \tag{71}$$

We've now established all of the inequalities we need to complete our proof.

$$\begin{aligned}
||f - g||^2 &= \int_a^b |f(x) - g(x)|^2 \, d\alpha & \text{definition of } ||\cdot|| \\
&= \sum_{i=0}^n \int_{x_i}^{x_{i+1}} |f(x) - g(x)|^2 \, d\alpha & \text{integral property 6.12c} \\
&\leq \sum_{i=0}^n \int_{x_i}^{x_{i+1}} |M_i - m_i|^2 \, d\alpha & \text{from (70)} \\
&= \sum_{i=0}^n |M_i - m_i|^2 \int_{x_i}^{x_{i+1}} d\alpha & \text{integral property 6.12a} \\
&= \sum_{i=0}^n |M_i - m_i|^2 \, \Delta\alpha(x_i) \\
&\leq \sum_{i=0}^n (M - m)(M_i - m_i) \, \Delta\alpha(x_i) & \text{from (71)} \\
&= (M - m) \sum_{i=0}^n (M_i - m_i) \, \Delta\alpha(x_i) & \text{integral property 6.12a} \\
&= (M - m)[U(P, f, \alpha) - L(P, f, \alpha)] & \text{Definition of } U \text{ and } L \\
&\leq (M - m) \frac{\epsilon^2}{M - m} & \text{we chose } P \text{ so that this would hold} \\
&= \epsilon^2 & \text{we chose } \epsilon \text{ so that this would hold}
\end{aligned}$$

Taking the square root of the first and last terms gives us

$$||f - g|| \leq \epsilon$$

which is what we were asked to prove.

Exercise 6.13a

Letting $u = t^2$ we have $du/dt = 2t$ and therefore

$$dt = \frac{du}{2t} = \frac{du}{2\sqrt{u}}$$

Using the change of variables theorem (6.19) we see that the given integral is equivalent to

$$f(x) = \int_{x^2}^{(x+1)^2} \frac{\sin u}{2\sqrt{u}} \, du \tag{72}$$

Integration by parts (theorem 6.22) with $F = 1/2\sqrt{u}$ and $g = \sin u \, du$ gives us

$$f(x) = \frac{-\cos u}{2\sqrt{u}} \Bigg|_{x^2}^{(x+1)^2} - \int_{x^2}^{(x+1)^2} \frac{\cos u}{4u^{3/2}} \, du$$

which expands to

$$f(x) = \frac{-\cos(x+1)^2}{2(x+1)} + \frac{\cos x^2}{2x} - \int_{x^2}^{(x+1)^2} \frac{\cos u}{4u^{3/2}} \, du$$

By the triangle inequality this gives us

$$\begin{aligned}
|f(x)| &= \left| \frac{-\cos(x+1)^2}{2(x+1)} + \frac{\cos x^2}{2x} - \int_{x^2}^{(x+1)^2} \frac{\cos u}{4u^{3/2}} \, du \right| \\
&\leq \left| \frac{-\cos(x+1)^2}{2(x+1)} \right| + \left| \frac{\cos x^2}{2x} \right| + \left| \int_{x^2}^{(x+1)^2} \frac{\cos u}{4u^{3/2}} \, du \right| \\
&< \left| \frac{1}{2(x+1)} \right| + \left| \frac{1}{2x} \right| + \left| \int_{x^2}^{(x+1)^2} \frac{1}{4u^{3/2}} \, du \right| \\
&= \left| \frac{1}{2(x+1)} \right| + \left| \frac{1}{2x} \right| + \left| \frac{1}{2} \left[\frac{1}{x+1} - \frac{1}{x} \right] \right| \\
&= \frac{1}{2(x+1)} + \frac{1}{2x} + \frac{1}{2x(x+1)} \\
&= \frac{x + (x+1) + 1}{2x(x+1)} \\
&= \frac{2(x+1)}{2x(x+1)} = \frac{1}{x}
\end{aligned}$$

Note that the strict inequality is justified by the fact that $\cos u \leq 1$ for all $u \in [x^2, (x+1)^2]$ but $\cos$ is not a constant function so $\cos u < 1$ for some $u \in [x^2, (x+1)^2]$.

Exercise 6.13b

In the previous problem we used integration by parts to determine that

$$f(x) = \frac{-\cos(x+1)^2}{2(x+1)} + \frac{\cos x^2}{2x} - \int_{x^2}^{(x+1)^2} \frac{\cos u}{4u^{3/2}}$$

Multiplying by $2x$ gives us

$$2xf(x) = \frac{-x\cos(x+1)^2}{x+1} + \cos x^2 - 2x \int_{x^2}^{(x+1)^2} \frac{\cos u}{4u^{3/2}} \, du$$

which is algebraically equivalent to

$$2xf(x) = \cos x^2 - \cos[(x+1)^2] + \frac{\cos(x+1)^2}{x+1} - 2x \int_{x^2}^{(x+1)^2} \frac{\cos u}{4u^{3/2}} \, du$$

Letting $r(x)$ be defined as

$$r(x) = \frac{\cos(x+1)^2}{x+1} - 2x \int_{x^2}^{(x+1)^2} \frac{\cos u}{4u^{3/2}} \, du$$

we have, by the triangle inequality,

$$
\begin{aligned}
|r(x)| &\leq \left| \frac{\cos(x+1)^2}{x+1} \right| + \left| 2x \int_{x^2}^{(x+1)^2} \frac{\cos u}{4u^{3/2}} \, du \right| \\
&\leq \left| \frac{1}{x+1} \right| + \left| 2x \int_{x^2}^{(x+1)^2} \frac{1}{4u^{3/2}} \, du \right| \\
&= \left| \frac{1}{x+1} \right| + \left| 2x\frac{1}{2} \left[\frac{1}{x+1} - \frac{1}{x} \right] \right| \\
&= \frac{1}{x+1} + x\frac{1}{x(x+1)} \\
&= \frac{1}{x+1} + \frac{1}{x+1} \\
&= \frac{2}{x+1} \\
&< \frac{2}{x}
\end{aligned}
$$

So we see that

$$2xf(x) = \cos x^2 - \cos[(x+1)^2] + r(x)$$

where $|r(x)| < 2/x$.

Exercise 6.13c

We're asked to find the $\limsup$ and $\liminf$ of the function

$$xf(x) = \frac{1}{2} \left[\cos(x^2) - \cos([x+1]^2) + r(x) \right]$$

We established in part (b) that $r(x) \to 0$ as $x \to \infty$. It's also clear that this function is bounded above by 1 and bounded below by -1, but it's not immediately clear that these bounds are the $\limsup$ and the $\liminf$.

Remark: The supremum and infimum of $\cos(x^2) - \cos([x+1]^2)$ are never obtained

The supremum would be obtained if we could find x such that

$$\cos(x^2) - \cos([x+1]^2) = 2 \tag{73}$$

which would occur precisely when

$$x^2 = 2k\pi, \quad [x+1]^2 = (2j+1)\pi, \quad j, k \in \mathbb{N} \tag{74}$$

After some tedious algebra, we can see that these two equalities hold when

$$\left(\frac{\pi[2(j-k)+1] - 1}{2\sqrt{2\pi}} \right)^2 = k \tag{75}$$

But this equation can't hold. If it did, we could rearrange it algebraically to give us

$$\pi^2[2(j-k)+1]^2 - 2\pi[2j+3k+1] + 1 = 0$$

This would allow us to use the quadratic formula to give an algebraic expression for π; but this is impossible as π is a transcendental number. The infimum of $\cos(x^2) - \cos([x+1]^2)$ is not obtained for the same reason.

Proof: The $\limsup$ of $xf(x)$ is 1

Let $1 > \epsilon > 0$ and $N \in \mathbb{N}$ be given. Let δ be chosen so that

$$0 < \delta < \frac{\cos^{-1}(1-\epsilon)}{2\pi}$$

Define the function $p(m)$ to be

$$p(m) = \left(\frac{\pi[2m+1] - 1}{2\sqrt{2\pi}} \right)^2$$

Its derivative with respect to m is

$$p'(m) = \sqrt{2\pi}(\pi[2m+1] - 1)$$

which is strictly increasing. Therefore we can make the derivative as large as we want by choosing a sufficiently large m. Specifically, we are able to choose $M \in \mathbb{N}$ such that $p'(M)\delta > 1$ and $M > N$. By the mean value theorem we have

$$p(M+\delta) - p(M) = p'(\xi)\delta, \quad \xi \in (M, M+\delta)$$

From the strictly increasing nature of p' and our choice of M we have

$$p(M+\delta) - p(M) = p'(\xi)\delta > p'(M)\delta > 1$$

Therefore p(x) must take an integer value for at least one $x \in [M, M+\delta]$. Let κ represent one such x. The two important properties of κ are that $p(\kappa)$ is an integer and that $\kappa - M < \delta$ (so that κ itself is "almost" an integer). We now have

$$p(\kappa) = p(M+\delta) = \left(\frac{\pi[2(M+\delta)+1] - 1}{2\sqrt{2\pi}} \right)^2 = k \in \mathbb{N}$$

If we define j to be $j = k + M$ this becomes

$$\left(\frac{\pi[2(j-k+\delta)+1] - 1}{2\sqrt{2\pi}} \right)^2 = k \in \mathbb{N}$$

Reversing the algebraic steps that led from (74) to (75) tells us that there exists some $x \in \mathbb{R}$ such that

$$x^2 = 2k\pi, \quad [x+1]^2 = (2(j+\delta)+1)\pi, \quad j, k \in \mathbb{N}$$

Using this value of x in our original function, we have

$$xf(x) = \frac{1}{2}\left[\cos(x^2) - \cos([x+1]^2) + r(x) \right] = \frac{1}{2}\left[\cos(2k\pi) - \cos(2(j+1)\pi + 2\delta\pi) + r(x) \right]$$

Using trig identities, this becomes

$$xf(x) = \frac{1}{2}\left[1 - \{\cos(2(j+1)\pi)\cos(2\epsilon\pi) - \sin(2(j+1)\pi)\sin(2\epsilon\pi)\} + r(x)\right]$$

$$= \frac{1}{2}\left[1 - (-1)\cos(2\epsilon\pi) - 0 + r(x)\right]$$

Finally, from our original choice of δ as an inverse cosine, this becomes

$$xf(x) = \frac{1}{2}\left[1 + (1 - \epsilon) + r(x)\right]$$

$$= 1 + \frac{1}{2}\left[r(x) - \epsilon\right]$$

From part (b), we know that $r(x) \to 0$ as $x \to \infty$:

$$\lim_{x\to\infty} xf(x) = 1 - \frac{\epsilon}{2}$$

And ϵ was arbitrary, so the supremum is

$$\lim_{x\to\infty} \sup xf(x) = 1$$

The proof that $\liminf xf(x) = -1$ is similar.

Exercise 6.13d

$$\int_0^\infty \sin(t^2)\,dt \quad = \quad \sum_0^\infty \int_{\sqrt{n\pi}}^{\sqrt{(n+1)\pi}} \sin(t^2)\,dt \tag{76}$$

$$= \quad \sum_0^\infty \int_{\sqrt{n\pi}}^{\sqrt{(n+1)\pi}} (-1)^n |\sin(t^2)|\,dt$$

$$= \quad \sum_0^\infty (-1)^n \int_{\sqrt{n\pi}}^{\sqrt{(n+1)\pi}} |\sin(t^2)|\,dt$$

$$\leq \quad \sum_0^\infty (-1)^n \int_{\sqrt{n\pi}}^{\sqrt{(n+1)\pi}} 1\,dt \quad \text{(and is similarly bounded below)}$$

$$= \quad \sqrt{\pi}\sum_0^\infty (-1)^n \left(\sqrt{n+1} - \sqrt{n}\right) \tag{77}$$

$$\tag{78}$$

We saw in exercise 3.6.a that the sequence $\{\sqrt{n+1} - \sqrt{n}\}$ is decreasing. Therefore, by the alternating series theorem (3.43) the series in (77) converges. Therefore, by the comparison test (3.25) the series in (76) converges and therefore the integral in (76) converges.

Exercise 6.14a

Letting $u = e^t$ we have $du/dt = e^t = u$ and therefore

$$dt = \frac{du}{e^t} = \frac{du}{u}$$

Using the change of variables theorem (6.19) we see that the given integral is equivalent to

$$f(x) = \int_{e^x}^{e^{x+1}} \frac{\sin(u)}{u}\,du \tag{79}$$

116

We use integration by parts as we did in 6.13(a).

$$f(x) = \frac{-\cos u}{u}\bigg|_{e^x}^{e^{x+1}} - \int_{e^x}^{e^{x+1}} \frac{\cos u}{u^2}\, du$$

which expands to

$$f(x) = \frac{\cos e^x}{e^x} - \frac{\cos e^{x+1}}{e^{x+1}} - \int_{e^x}^{e^{x+1}} \frac{\cos u}{u^2}\, du$$

By the triangle inequality this gives us

$$
\begin{aligned}
|f(x)| \;&=\; \left| \frac{\cos e^x}{e^x} - \frac{\cos e^{x+1}}{e^{x+1}} - \int_{e^x}^{e^{x+1}} \frac{\cos u}{u^2}\, du \right| \\[2mm]
&\le\; \left| \frac{\cos e^x}{e^x} \right| + \left| \frac{\cos e^{x+1}}{e^{x+1}} \right| + \left| \int_{e^x}^{e^{x+1}} \frac{\cos u}{u^2}\, du \right| \\[2mm]
&<\; \left| \frac{1}{e^x} \right| + \left| \frac{1}{e^{x+1}} \right| + \left| \int_{e^x}^{e^{x+1}} \frac{1}{u^2}\, du \right| \\[2mm]
&<\; \left| \frac{1}{e^x} \right| + \left| \frac{1}{e^{x+1}} \right| + \left| \frac{1}{e^x} - \frac{1}{e^{x+1}} \right| \\[2mm]
&=\; \frac{1}{e^x} + \frac{1}{e^{x+1}} + \frac{e-1}{e^{x+1}} \\[2mm]
&=\; \frac{2e}{e^{x+1}} \\[2mm]
&=\; \frac{2}{e^x}
\end{aligned}
$$

Therefore $e^x |f(x)| < 2$. Note that the strictness of the inequality is justified by the fact that $\cos u \le 1$ for all $u \in [e^x, e^{x+1}]$ but $\cos$ is not a constant function so $\cos u < 1$ for some $u \in [e^x, e^{x+1}]$.

Exercise 6.14b

In the previous exercise we used integration by parts to determine that

$$f(x) = \frac{\cos e^x}{e^x} - \frac{\cos e^{x+1}}{e^{x+1}} - \int_{e^x}^{e^{x+1}} \frac{\cos u}{u^2}\, du$$

Multiplying by e^x gives us

$$e^x f(x) = \cos e^x - e^{-1} \cos e^{x+1} - e^x \int_{e^x}^{e^{x+1}} \frac{\cos u}{u^2}\, du$$

Letting $r(x)$ be defined as

$$r(x) = -e^x \int_{e^x}^{e^{x+1}} \frac{\cos u}{u^2}\, du$$

we have, by the triangle inequality,

$$|r(x)| \leq \left| -e^x \int_{e^x}^{e^{x+1}} \frac{\cos u}{u^2} \, du \right|$$

$$= \left| -e^x \int_{e^x}^{e^{x+1}} \frac{1}{u^2} \, du \right|$$

$$= |e^x| \left| \int_{e^x}^{e^{x+1}} \frac{1}{u^2} \, du \right|$$

$$= |e^x| \left| \frac{1}{e^x} - \frac{1}{e^{x+1}} \right|$$

$$= \frac{e-1}{e}$$

$$< \frac{2}{e}$$

So we see that

$$e^x f(x) = \cos e^x - e^{-1} \cos e^{x+1} + r(x)$$

where $|r(x)| < 2e^{-1}$.

Exercise 6.15a

Using integration by parts with $F = f^2(x)$ and $g = dx$ gives us

$$1 = \int_a^b f^2(x) \, dx = x f^2(x) \big|_a^b - \int_a^b 2f(x) f'(x) x \, dx$$

which evaluates to

$$1 = b f^2(b) - a f^2(a) - \int_a^b 2f(x) f'(x) x \, dx$$

which, since $f(a) = f(b) = 0$, becomes

$$1 = - \int_a^b 2f(x) f'(x) x \, dx$$

Dividing both sides by -2 gives us

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

which is the desired equality.

Exercise 6.15b

Applying Holder's inequality to the last equation in part (a) gives us

$$\left| \frac{-1}{2} \right| = \left| \int_a^b f(x) f'(x) x \, dx \right| \leq \left\{ \int_a^b |x^2 f^2(x)| \, dx \right\}^{1/2} \cdot \left\{ \int_a^b |[f'(x)]^2| \, dx \right\}^{1/2}$$

Since $f(x) f'(x)$ must be negative at some point (by mean value theorem, since $f(a) = f(b) = 0$) while $x^2 f^2$ and $[f'(x)]^2$ are strictly positive, this inequality must be strict:

$$\left| \frac{-1}{2} \right| = \left| \int_a^b f(x) f'(x) x \, dx \right| < \left\{ \int_a^b |x^2 f^2(x)| \, dx \right\}^{1/2} \cdot \left\{ \int_a^b |[f'(x)]^2| \, dx \right\}^{1/2}$$

Squaring the first and last term in this chain of inequalities, we have

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

All of the normed values are squares, so the norms are redundant:

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

Exercise 6.16a

We can integrate the given function separately over infinitely many intervals of length 1:

$$s \int_1^\infty \frac{[x]}{x^{s+1}}\, dx = \sum_{n=1}^\infty s \int_n^{n+1} \frac{[x]}{x^{s+1}}\, dx$$

We have $[x] = n$ on the interval $[n, n+1)$ so this becomes

$$= \sum_{n=1}^\infty s \int_n^{n+1} \frac{n}{x^{s+1}}\, dx$$

$$= \sum_{n=1}^\infty sn \left[\frac{-1}{sx^s} \right]_{x=n}^{n+1}$$

$$= \sum_{n=1}^\infty n \left[\frac{1}{n^s} - \frac{1}{(n+1)^s} \right]$$

We then split up the summation into three parts.

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

Evaluating the first summation and changing the index of the third gives us

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

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

$$= 1 + \sum_{n=2}^\infty \frac{1}{n^s}$$

And since 1 is clearly equal to $1/n^s$ when $n = 1$, this is equivalent to

$$= \sum_{n=1}^\infty \frac{1}{n^s}$$

Exercise 6.16b

We're asked to evaluate the integral

$$\frac{s}{s-1} - s \int_1^\infty \frac{x - [x]}{x^{s+1}}\, dx$$

Having determined that $[x]$ was integrable in part (a) we can split up the integral as follows:

$$= \frac{s}{s-1} - s \int_1^\infty \frac{1}{x^s}\, dx + s \int_1^\infty \frac{[x]}{x^{s+1}}\, dx$$

Elementary calculus allows us to calculate the left integral:

$$= \frac{s}{s-1} - \frac{s}{s-1} + s \int_1^\infty \frac{[x]}{x^{s+1}}\, dx$$

$$= s \int_1^\infty \frac{[x]}{x^{s+1}}\, dx$$

This, as we saw in part (a), is equivalent to

$$= \zeta(s)$$

Exercise 6.17

I'm going to change the notation of this problem a bit to make it clearer (to me, at least) by letting $f = G$ and $f' = g$. We're told that α is a monotonically increasing function on $[a, b]$ and that f is continuous. We're asked to prove that

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

Most of the work is done for us by the theorem 6.22 (integration by parts) which tells us that

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

So our proof is reduced to simply proving that

$$\int_a^b f(x) \, d\alpha = \int_a^b f(x)\alpha'(x) \, dx \tag{80}$$

It's tempting to appeal to elementary calculus to show that $d\alpha = \alpha'(x) \, dx$, but we can prove this more formally.

Proving this more formally

Let $\epsilon > 0$ be given. We're told that f is continuous and that α is monotonically inreasing; therefore $f \in \mathscr{R}(\alpha)$. Let P be an arbitrary partition of $[a, b]$ such that $U(P, f, \alpha) - L(P, f, \alpha) < \frac{\epsilon}{2}$.

On any interval $[x_{i-1}, x_i]$ the mean value theorem tells us that there is some $t_i \in [x_{i-1}, x_i]$ such that

$$\alpha(x_{i-1}) - \alpha(x_i) = \alpha'(t_i)[x_{i-1} - x_i]$$

or, to express the same equation with different notation,

$$\Delta\alpha_i = \alpha'(t_i)\Delta x_i \tag{81}$$

By theorem 6.7 (c) we have

$$\left| \sum_{i=1}^n f(t_i)\Delta\alpha_i - \int_a^b f(x) \, d\alpha \right| < \frac{\epsilon}{2} \tag{82}$$

and also

$$\left| \sum_{i=1}^n f(t_i)\alpha'(t_i)\Delta x_i - \int_a^b f(x)\alpha'(x) \, dx \right| < \frac{\epsilon}{2} \tag{83}$$

By the inequality (81) we have

$$\sum_{i=1}^n f(t_i)\Delta\alpha_i = \sum_{i=1}^n f(t_i)\alpha'(t_i)\Delta x_i$$

Therefore, by (82) and (83) and the triangle inequality, we have

$$\left| \int_a^b f(x) \, d\alpha - \int_a^b f(x)\alpha'(x) \, dx \right| < \epsilon$$

which, in order to be true for any $\epsilon > 0$, requires that

$$\int_a^b f(x) \, d\alpha = \int_a^b f(x)\alpha'(x) \, dx$$

This proves (80), and this completes the proof.

Exercise 6.18

The derivative $\gamma_1'(t) = ie^{it}$ is continuous, therefore $\gamma_1(t)$ is rectifiable (theorem 6.27) and its length is given by

$$\int_0^{2\pi} |ie^{it}| = \int_0^{2\pi} 1 = 2\pi$$

The derivative $\gamma_2'(t) = 2ie^{it}$ is continuous, therefore $\gamma_2(t)$ is rectifiable and its length is given by

$$\int_0^{2\pi} |2ie^{it}| = \int_0^{2\pi} 2 = 4\pi$$

The arc γ_3 is not rectifiable

Proof by contradiction. If γ_3 were rectifiable then its length would be given by the integral

$$\int_0^{2\pi} |\gamma'(t)|\, dt = \int_0^{2\pi} \left| \left[2\pi i \sin\left(\frac{1}{t}\right) + \frac{-1}{t^2} 2\pi it \cos\left(\frac{1}{t}\right) \right] e^{2\pi it \sin(1/t)} \right|\, dt$$

$$= \int_0^{2\pi} 2\pi \left| \sin\left(\frac{1}{t}\right) - \frac{1}{t} \cos\left(\frac{1}{t}\right) \right|\, dt \tag{84}$$

But this integral isn't defined, since Riemann integrals are defined only for bounded functions and this function is unbounded:

$$\lim_{t \to 0} |\gamma'(t)\, dt| = \lim_{k \to \infty} \left| \gamma'\left(\frac{1}{2k\pi}\right) \right| = \lim_{k \to \infty} 2\pi \left| \sin(2k\pi) - 2k\pi \cos(2k\pi) \right|$$

$$= \lim_{k \to \infty} 2\pi \left| \pm 2k\pi \right|$$

$$= \lim_{k \to \infty} 4\pi^2 k = \infty$$

Therefore the integral in (84) doesn't exist, which means the arc length of γ_3 is not defined, which means that γ_3 is not rectifiable. $\square$

Exercise 6.19

Lemma 1: If $f : A \to B$ and $G : B \to C$ are both one-to-one, then $g \circ f : A \to C$ is one-to-one

By definition of "one-to-one", we have

$$f(x) = f(y) \to x = y, \qquad g(s) = g(t) \to s = t$$

Therefore

$$g(f(x)) = g(f(y)) \to f(x) = f(y) \to x = y$$

and so $g \circ f$ is one-to-one. $\square$

Lemma 2: If $g \circ f : A \to C$ is one-to-one and $f : A \to B$ is one-to-one and onto, then $g : B \to C$ is one-to-one.

Proof by contrapositive. Suppose that g is not one-to-one. Then we could find $x, y \in B$ such that $g(x) = g(y)$ but $x \neq y$. But f is one-to-one and onto, so there exist unique $s, t \in A$ such that $f(s) = x, f(t) = y$. So we have $g(f(s)) = g(f(t))$ but $f(s) \neq f(t)$ and therefore $s \neq t$. So $g \circ f$ is not one-to-one. By contrapositive, the lemma is proven. $\square$

Lemma 3: If $f : [a, b] \to [c, d]$ is a continuous, one-to-one, real function then f is either strictly decreasing or strictly increasing

Proof by contrapositive. Let f be a continuous real function. If f were not strictly increasing and not strictly decreasing then we could find some $x < y < z$ such that either $f(y) \leq f(x)$ and $f(y) \leq f(z)$ or such that $f(y) \geq f(x)$ and $f(y) \geq f(z)$. From the intermediate value property of continuous functions we know that we must then be able to find x', y' such that

$$x \leq x' < y < z' < z, \qquad f(x') = f(z')$$

so that f is not one-to-one. By contrapositive, the lemma is proven. $\square$

Lemma 4: If $P = \{x_i\}$ is a partition of γ_2, then $P' = \{\phi(x_i)\}$ is a partition of γ_1

From lemma 3 we know that ϕ is strictly increasing or decreasing, and since $\phi(c) = a$ it must be strictly increasing. And ϕ is onto, so $\phi(d) = b$. Let P be an arbitrary partition of $[c, d]$:

$$P = \{x_i\} = \{c = x_0, x_1, \ldots, x_n, x_{n+1} = d\}$$

From the properties of ϕ we have

$$a = \phi(c) = \phi(x_0) < \phi(x_1) < \phi(x_2) < \ldots < \phi(x_n) < \phi(x_{n+1}) = \phi(d) = b$$

and therefore

$$\{\phi(x_i)\} = \{a = \phi(x_0), \phi(x_1), \phi(x_2), \ldots, \phi(x_n), \phi(x_{n+1}) = b\}$$

which is a partition of $[a, b]$. $\qquad\square$

Lemma 5: If $P = \{x_i\}$ is a partition of γ_1, then $P' = \phi^{-1}(x_i)$ is a partition of γ_2

We're told that ϕ is a continuous one-to-one function from $[a, b]$ to $[c, d]$, so its inverse exists and is a continuous mapping from $[a, b]$ onto $[c, d]$ (theorem 4.17). By lemma 3, this means that ϕ^{-1} is either strictly increasing or strictly decreasing, and since $\phi^{-1}(a) = c$ it must be strictly increasing. And ϕ^{-1} is onto, so $\phi^{-1}(b) = d$. Let P be an arbitrary partion of $[a, b]$:

$$P = \{x_i\} = \{a = x_0, x_1, \ldots, x_n, x_{n+1} = b\}$$

From the properties of ϕ^{-1} we have

$$c = \phi^{-1}(a) = \phi^{-1}(x_0) < \phi^{-1}(x_1) < \phi^{-1}(x_2) < \ldots < \phi^{-1}(x_n) < \phi^{-1}(x_{n+1}) = \phi^{-1}(b) = d$$

and therefore

$$\{\phi^{-1}(x_i)\} = \{c = \phi^{-1}(x_0), \phi^{-1}(x_1), \phi^{-1}(x_2), \ldots, \phi^{-1}(x_n), \phi^{-1}(x_{n+1}) = d\}$$

which is a partition of $[c, d]$. $\qquad\square$

Proof 1: γ_1 is an arc iff γ_2 is an arc

If γ_1 is an arc then γ_1 is a one-to-one function. We're told that ϕ is one-to-one, therefore by lemma 1 $\gamma_1 \circ \phi$ is one-to-one. And $\gamma_2 = \gamma_1 \circ \phi$ so γ_2 is one-to-one. There γ_2 is an arc.

If γ_2 is an arc then $\gamma_2 = \gamma_1 \circ \phi$ is one-to-one. We're told that ϕ is a one-to-one and onto function, therefore by lemma 2 we know that γ_1 is one-to-one. Therefore γ_1 is an arc. $\qquad\square$

Proof 2: γ_1 is a closed curve iff γ_2 is a closed curve

Assume that γ_1 is a closed curve so that $\gamma_1(a) = \gamma_1(b)$. We saw in lemma 4 that $\phi(c) = a$ and $\phi(d) = b$, so we have

$$\gamma_2(c) \equiv \gamma_1(\phi(c)) = \gamma_1(a) = \gamma_1(b) = \gamma_1(\phi(d)) \equiv \gamma_2(d)$$

Therefore $\gamma_2(c) = \gamma_2(d)$, which means that γ_2 is a closed curve.

Now assume that γ_2 is a closed curve so that $\gamma_2(c) = \gamma_2(d)$. We saw in lemma 5 that $\phi^{-1}(a) = c$ and $\phi^{-1}(b) = d$ so we have

$$\gamma_1(a) = \gamma_1(\phi(\phi^{-1}(a))) \equiv \gamma_2(\phi^{-1}(a)) = \gamma_2(c) = \gamma_2(d) = \gamma_2(\phi^{-1}(b)) \equiv \gamma_1(\phi(\phi^{-1}(b))) = \gamma_1(b)$$

Therefore $\gamma_1(a) = \gamma_1(b)$, which means that γ_1 is a closed curve. $\qquad\square$

Proof 3: γ_1 and γ_2 have the same length

Let P be an arbitrary partitioning of $[a, b]$. Using the notation from definition 6.26, we have

$$\Lambda(P, \gamma_2) = \sum_{i=1}^{n} |(\gamma_2(x_i) - \gamma_2(x_{i-1})| < M$$

From the definition of γ_2, this becomes

$$\Lambda(P, \gamma_2) = \sum_{i=1}^{n} |(\gamma_1(\phi(x_i)) - \gamma_1(\phi(x_{i-1}))| < M$$

From lemma 4 we see that $\{\phi(x_i)\}$ describes a partitioning of $[c, d]$:

$$\Lambda(P, \gamma_2) = \Lambda(\phi(P), \gamma_1)$$

Similarly, if we let P' be an arbitrary partition of $[a, b]$ we can use lemma 5 to show that

$$\Lambda(P', \gamma_1) = \Lambda(\phi^{-1}(P), \gamma_2)$$

This means that the set $\{\Lambda(P, \gamma_1)\}$ is identical to the set of $\{\Lambda(P, \gamma_2)\}$ so clearly these sets have the same supremums, which means that $\Lambda(\gamma_1) = \Lambda(\gamma_2)$. $\qquad\square$

Proof 4: γ_1 is rectifiable iff γ_2 is rectifiable

Having shown previously that $\Lambda(\gamma_1) = \Lambda(\gamma_2)$, it's clear that $\Lambda(\gamma_1)$ is finite iff $\Lambda(\gamma_2)$ is finite and so γ_1 is rectifiable iff γ_2 is rectifiable. $\qquad\square$

Exercise 7.1

Let $\{f_n\}$ be a sequence of functions that converges uniformly to f. Let $M_n = \sup |f_n|$ and let $M = \sup |f|$. Let $\epsilon > 0$ be given. Because of the uniform convergence of $\{f_n\} \to f$ we can find some N such that

$$n > N \to |f_n(x)| = |f_n(x) - f(x) + f(x)| \leq |f_n(x) - f(x)| + |f(x)| \leq \epsilon + M$$

This tells us that for all $n > N$ the function f_n is bounded above by $M + \epsilon$. There are only finitely many $n \leq N$, each of which is bounded by some M_n. So by choosing

$$\max\{M + \epsilon, M_1, M_2, \ldots, M_N\}$$

we have found a bound for all f_n.

Exercise 7.2a

Let $\epsilon > 0$ be given. We're told that $\{f_n\}$ and $\{g_n\}$ converge uniformly on E so we can find N, M such that

$$n, m > N \to |f_n(x) - f_m(x)| < \frac{\epsilon}{2}, \quad n, m > M \to |g_n(x) - g_m(x)| < \frac{\epsilon}{2}$$

By choosing $N^* > \max\{N, M\}$ we have

$$n, m > N^* \to |(f_n(x) + g_n(x)) - (f_m(x) - g_m(x))| \leq |f_n(x) - f_m(x)| + |g_n(x) - g_m(x)| < \epsilon$$

which shows that $f_n + g_n$ converges uniformly on E.

Exercise 7.2b

If $\{f_n\}$ and $\{g_n\}$ are bounded functions then by exercise 7.1 they are uniformly bounded, say by F and G. Let $\epsilon > 0$ be given and choose δ such that

$$\delta < \min\left\{\epsilon, \sqrt{\frac{\epsilon}{3}}, \frac{\epsilon}{3F}, \frac{\epsilon}{3G}\right\}$$

We're told that $\{f_n\}$ and $\{g_n\}$ converge uniformly on E so we can find N, M such that

$$n, m > N \rightarrow |f_n(x) - f_m(x)| < \delta, \quad n, m > M \rightarrow |g_n(x) - g_m(x)| < \delta$$

$$|f_n g_n - f_m g_m| = |(f_n - f_m)(g_n - g_m) + f_m(g_n - g_m) + g_m(f_n - f_m)|$$

$$\leq |(f_n - f_m)||(g_n - g_m)| + |f_m||(g_n - g_m)| + |g_m||(f_n - f_m)| \leq \delta^2 + |f_m|\delta + |g_m|\delta \leq \delta^2 F\delta + G\delta$$

$$\leq \left(\sqrt{\frac{\epsilon}{3}}\right)^2 + F\left(\frac{\epsilon}{3F}\right) + G\left(\frac{\epsilon}{3G}\right) = \epsilon$$

Exercise 7.3

Let $\{f_n\}$ be a sequence such that $f_n(x) = x$. This function obviously converges uniformly to the function $f(x) = x$ on the set $\mathbb{R}$. Let $\{g_n\}$ be a sequence of constant functions such that $g_n(x) = \frac{1}{n}$. This sequence obviously converges uniformly to the function $g(x) = 0$. Their product is the sequence $\{f_n g_n\}$ where $f_n(x)g_n(x) = x/n$. It's clear that this sequence converges pointwise to $f_n(x)g_n(x) = 0$.

To show that $f_n g_n$ is not uniformly convergent, let $\epsilon > 0$ be given. Choose an arbitrarily large $n \in \mathbb{Z}$ and choose $t \in \mathbb{R}$ such that $t > \epsilon(n(n+1))$. We now have

$$|f_n(t)g_n(t) - f_{n+1}(t)g_{n+1}(t)| = \left|\frac{t}{n} - \frac{t}{n+1}\right| = \left|\frac{t}{n(n+1)}\right| > \left|\frac{\epsilon(n(n+1))}{n(n+1)}\right| = \epsilon$$

This shows that the necessary requirements for uniform convergence given in theorem 7.8 do not hold.

Exercise 7.4

Holy shit this exercise is a mess.

For what values of x does the series converge absolutely? (Incomplete)

For values of $x > 0$ we have

$$\sum\left|\frac{1}{1 + n^2 x}\right| = \left|\frac{1}{x}\right|\sum\left|\frac{1}{1/x + n^2}\right| \leq \left|\frac{1}{x}\right|\sum\left|\frac{1}{n^2}\right|$$

By the comparison test this shows that $f(x)$ converges absolutely. If $x = 0$ then we have

$$\sum\left|\frac{1}{1 + n^2 x}\right| = \sum|1| = \infty$$

This series clearly doesn't converge, absolutely or otherwise. If $x < 0$ things get more complicated: If $x = -1/n^2$ for any $n \in \mathbb{N}$ then the nth term of the series is undefined and therefore $f(x)$ is undefined. If $x \neq -1/n^2$ for any $n \in \mathbb{N}$ then we can use the fact that

For what intervals does the function converge uniformly?

If E is any interval of the form $[a, b]$ with $a > 0$ then we have

$$\sup|f_n(x)| = f_n(a) = \frac{1}{1 + n^2 a}$$

And therefore we have

$$\sum \sup |f_n(x)| = \sum \frac{1}{1 + n^2 a} \le \frac{1}{a} \sum \frac{1}{n^2}$$

This shows that $\sum \sup |f_n(x)|$ converges by the comparison test, and so by theorem 7.10 we see that the series $\sum f_n(x)$ converges uniformly on E.

If E is any interval of the form $[a, b]$ with $b < 0$ that does not contain any elements of the form $-1/n^2$, $n \in \mathbb{N}$

For what intervals does the function fail to converge uniformly?

Define the set $X = \{x_n\}$ with $x_n = -1/n^2$. The function f will fail to converge uniformly on any interval that contains an element of $X \cup 0$ or has an element of $X \cup 0$ as a limit point of E.

Proof: Let E be an arbitrary interval. If E contains any $x_n \in X$ then $f(x_n)$ undefined and so f fails to converge uniformly on E. If E contains 0 then $f(0) = \sum 1 = \infty$, but we will never find some finite N such that $\left| \sum_1^N 1 - f(0) \right| < \epsilon$ so it's clear that f fails to converge uniformly on E.

Now suppose that some $x_n \in X$ is a limit point of E. The nth term of f is unbounded near x_n, so f is unbounded near x_n, and therefore $\lim_{t \to x_n} f(t) = \infty$. From this we have

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

On the other hand, if we first fix a value of t and take the limit of f as $n \to \infty$ we have

$$\lim_{n \to \infty} f(t) = \lim_{n \to \infty} \frac{1}{1 + n^2 t} = 0$$

and therefore

$$\lim_{t \to x_n} \lim_{n \to \infty} f(t) = \lim_{t \to x_n} 0 = 0$$

Exercise 7.5

If $x \le 0$ or $x > 1$ then $x = 0$ for all n and therefore in these cases $\lim_{n \to \infty} f_n(x) = 0$. For any other x, choose an integer N large enough so that $N > 1/x$. For all $n > N$ we now have $x > 1/n$ and therefore $f_n(x) = 0$, so for this case we have $\lim_{n \to \infty} f_n(x) = 0$. This exhausts all possible values of x, so $\{f_n\}$ converges pointwise to the continuous function $f(x) = 0$.

To show that this function doesn't converge uniformly to $f(x) = 0$ let $1 > \epsilon > 0$ be given and let n be an arbitrarily large integer. We can easily verify that

$$f_n \left(\frac{1}{n + 1/2} \right) = \sin^2(n\pi + \pi/2) = 1$$

and therefore the definition of uniform convergence is not satisfied.

The last part of the question asks us to use the series $\sum f_n$ to show that absolute convergence for all x does not imply uniform convergence. The proof of this is simple: we've already shown that $\{f_n\}$ is not uniformly convergent but $\sum |f_n(x)|$ converges to $f(x) = |\sin^2(\pi/x)|$ so $\sum |f_n|$ is absolutely convergent.

Exercise 7.6

To show that the series doesn't converge absolutely for any x:

$$\sum \left| (-1)^n \frac{x^2 + n}{n^2} \right| = \sum \left| \frac{x^2}{n^2} + \frac{1}{n} \right| \ge \sum \left| \frac{1}{n} \right|$$

The rightmost sum is the harmonic series, which is known to diverge. By comparison test the leftmost series must also diverge.

To show that the series converges uniformly in every bounded interval $[a, b]$, let $\epsilon > 0$ be given. Define X to be $\sup\{|a|, |b|\}$. Define the partial sum f_m to be

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

We can rearrange this algebraically to form

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

We know that the alternating harmonic series converges (theorem 3.44, or example 3.40(d)) and that $\sum 1/n^2$ converges (theorem 3.28). Therefore by the Cauchy criterion for convergence we can find N such that

$$p > q > N \rightarrow |\sum_{n=q}^{p} (-1)^n \frac{1}{n}| < \frac{\epsilon}{2}$$

and we can also find M such that

$$p > q > M \rightarrow |\sum_{n=q}^{p} (-1)^n \frac{1}{n^2}| < \frac{\epsilon}{2|X|^2}$$

So, by choosing $p > q > \max\{N, M\}$ we have (for all $x \in [a, b]$):

$$
\begin{aligned}
|f_p(x) - f_q(x)| &= \left| \sum_{n=q}^{p} (-1)^n \frac{1}{n} + x^2 \sum_{n=q}^{p} (-1)^n \frac{1}{n^2} \right| \\
&\leq \left| \sum_{n=q}^{p} (-1)^n \frac{1}{n} \right| + \left| x^2 \sum_{n=q}^{p} (-1)^n \frac{1}{n^2} \right| \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon x^2}{2X^2} \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}
$$

Exercise 7.7

We can establish a global maximum for $|f_n(x)|$

The derivative of $f_n(x)$ is

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

This derivative is zero only when $nx^2 = 1$, which occurs only when $x = \pm 1/\sqrt{n}$, at which point we have

$$f_n\left(\frac{\pm 1}{\sqrt{n}}\right) = \frac{\pm 1}{2\sqrt{n}}$$

These extrema must be the global extrema for the function, since f_n has no asymptotes and $f_n(x) \to 0$ as $x \to \pm\infty$. Therefore $|f_n(x)| \leq 1/(2\sqrt{n})$.

f_n converges uniformly to $f(x) = 0$

Clearly f_n converges pointwise to $f(x) = 0$, since for any x we have

$$\lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \frac{x}{1 + nx^2} = 0$$

Now let $\epsilon > 0$ be given and choose n sufficiently large so that $1/(2\sqrt{n}) < \epsilon$: this can be done by choosing $n > \frac{1}{4\epsilon^2}$. From the previously established bounds, we now have

$$|f_n(x) - f(x)| = |f_n(x) - 0| = |f_n(x)| \leq \frac{1}{2\sqrt{n}} < \epsilon$$

By theorem 7.9 this is sufficient to prove that f_n converges uniformly to $f(x) = 0$.

When does $f'(x) = \lim f_n'(x)$?

We've established that $f(x) = 0$, so clearly $f'(x) = 0$ for all x. The limit of f_n' is given by:

$$\lim_{n\to\infty} f_n'(x) = \lim_{n\to\infty} \frac{1 - nx^2}{n^2 x^4 + 2nx^2 + 1} = \begin{cases} 0 & x \neq 0 \\ 1 & x = 0 \end{cases}$$

This shows that it's not necessarily true that $[\lim f_n]' = \lim[f_n']$ even if f_n converges uniformly.

Exercise 7.8

Proof of uniform convergence

Let $\epsilon > 0$ be given. Let f_m represent the partial sum

$$f_m = \sum_{n=1}^{m} c_n I(x - x_i)$$

We're told that $\sum |c_n|$ converges, so $\sum |c_n|$ satisfies the Cauchy convergence criterion, so we can find N such that $p > q > N$ implies

$$\sum_{n=q}^{p} |c_n| < \epsilon$$

For the same values of $p > q > N$ we also have, by the triangle inequality and the fact that $I(x - x_n) \leq 1$,

$$|f_p - f_q| = \left| \sum_{n=q}^{p} c_n I(x - x_n) \right| \leq \sum_{n=q}^{p} |c_n I(x - x_n)| \leq \sum_{n=q}^{p} |c_n| < \epsilon$$

so $\{f_n\}$ converges uniformly by theorem 7.8

Proof of continuity

Let t be an arbitrary point (not necessarily in the interval (a, b)). This t either is or isn't a limit of some subsequence of $\{x_n\}$.

If t is not a limit point of some subsequence then we can find some neighborhood around t that contains no points of $\{x_n\}$; the function $I(t - x_n)$ is constant on this interval for all n and therefore $f(x) = \sum c_n I(x - x_n)$ is constant on this interval for all n. (see below for a a more thorough justification of this claim). And if f is constant on an interval around t then it is clearly continuous at t.

Now suppose that t a limit point of $\{x_n\}$. By the Cauchy convergence of $\sum |c_n|$ we can find N such that $\sum_{n=N}^{\infty} |c_n| < \epsilon$. Choose a neighborhood around t small enough that it does not contain the first N terms of $\{x_n\}$; let δ represent the radius of this neighborhood. If we choose s such that $|s - t| < \delta$, then $I(s - x_n) = I(t - x_n)$ for $n = 1, 2, \ldots, N$. From this we have

$$|s - t| < \delta \to f(s) - f(t) \leq \left| \sum_{n=N+1}^{\infty} (c_n I(s - x_n) - c_n I(t - x_n)) \right| \leq \sum_{n=N+1}^{\infty} |c_n [I(s - x_n) - I(t - x_n)]|$$

Each $I(x - x_n)$ is either 0 or 1 so $|I(s - x_n) - I(t - x_n)| \leq 1$, so we have

$$|s - t| < \delta \to f(s) - f(t) \leq \sum_{n=N+1}^{\infty} |c_n [I(s - x_n) - I(t - x_n)]| \leq \sum_{n=N+1}^{\infty} |c_n| < \epsilon$$

That is, we've found δ such that

$$|s - t| < \delta \to f(s) - f(t) < \epsilon$$

which, by definition, means that f is continuous at t.

We have therefore shown that f is continuous at t under all cases. But t was an arbitrary point (not necessarily confined to (a, b)) and therefore f is continuous at all points.

Justification of $I(x - x_n)$ being constant on some interval

As mentioned above, if t is not a limit point of some subsequence then we can find some neighborhood around t that contains no points of $\{x_n\}$. Let A be the set of elements of $\{x_n\}$ that are greater than t, and let B be the set of elements of $\{x_n\}$ that are smaller than t. If $|s - t| < \delta$ then there are no elements of $\{x_n\}$ between s and t and so every element of A is greater than both t and s and every element of B is smaller than both t and s. From this we see that $I(s - x_n) = I(t - x_n) = 0$ if $x_n \in A$ and $I(s - x_n) = I(t - x_n) = 1$ if $x_n \in B$. This means that $I(x - x_n)$ has the same value for every $x \in N_\delta(t)$; so $\sum c_n I(x - x_n)$ has the same value for every $x \in N_\delta(t)$; and therefore $f(x)$ has the same value for every $x \in N_\delta(t)$.

Exercise 7.9

We're told that $\{f_n\} \to f$ uniformly on E, so we can find N such that

$$n > N \to |f_n(t) - f(t)| < \frac{\epsilon}{2} \quad \text{for all } t \in E$$

This holds for all $t \in E$, so it must also hold for the elements of $\{x_n\}$, which means that

$$n > N \to |f_n(x_n) - f(x_n)| < \frac{\epsilon}{2} \tag{85}$$

We're also told that $\{f_n\}$ is a uniformly convergent sequence of continuous functions, so f is continuous; and since $\{x_n\} \to x$ we can find M such that

$$n > M \to |f(x_n) - f(x)| < \frac{\epsilon}{2} \tag{86}$$

Using the triangle inequality and equations (85) and (86), we have

$$|f_n(x_n) - f(x)| \leq |f_n(x_n) - f(x_n)| + |f(x_n) - f(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Converse statement 1

Suppose the converse is "Let $\{f_n\}$ be a sequence of continuous functions that converges uniformly to f. Is it true that if $\lim f_n(x_n) = f(x)$ then $\{x_n\} \to x$?". The answer to this question is "no". Consider the function $f_n(x) = x/n$ and the sequence $\{x_n\} = \{1\}$ (an infinite sequence of 1s). The sequence $\{f_n\}$ converges uniformly to $f(x) = 0$ on the set $[0, 1]$, so we have

$$\lim_{n \to \infty} f_n(x_n) = \lim_{n \to \infty} f_n(1) = 0 = f(0)$$

It's clear that $\{x_n\} \nrightarrow 0$, so the converse does not hold.

Converse statement 2

Suppose the converse is "Let $\{f_n\}$ be a sequence such that $\lim f_n(x_n) = f(x)$ for some function f and for all sequences of points $x_n \in E$ such that $\{x_n\} \to x \in E$. Is it true that $\{f_n\}$ converges uniformly to f on E?

The answer to this question is "no". Let $f_n(x) = 1/(nx)$. Let E be the harmonic set $\{1/n\}$. Then f_n converges pointwise on E to $f(x) = 0$. Although 0 is a limit point of E, 0 is not an element of E and this converse statement is only concerned with sequences $\{x_n\}$ that converge to some $x \in E$. So the only sequences of points $x_n \in E$ such that $\{x_n\} \to x \in E$ are sequences where every term of the sequence is eventually just x itself (that is, sequences for which there is some N such that $n > N \to x_n = x$). So clearly for every such sequence we must have $\lim f_n(x_n) = \lim f_n(x) = f(x)$. But, despite the fact that $\lim f_n(x_n) = f(x)$ whenever $\{x_n\} \to x \in E$, it's clear that $f_n(x)$ does not converge uniformly to $f(x) = 0$ on E (to prove this, choose any ϵ and any n and then choose $x < (n\epsilon)^{-1}$).

Exercise 7.10a

f is continuous at every irrational number

We know that $\sum 1/n^2$ converges, so by the Cauchy criterion we can choose N such that

$$b > a > N \to \sum_a^b \frac{1}{n^2} < \frac{\epsilon}{4}$$

The fractional part of (nx) will never be zero because x is irrational, so we can choose δ such that

$$0 < \delta < \frac{1}{2}\min\{(nx), 1 - (nx) : n < N\}$$

This guarantees that $0 < (n[x - \delta]) < (nx)$ for all $n < N$ and $(nx) < (n[x + \delta]) < 1$ for all n. More importantly, this choice of δ guarantees that $(n[x - \delta]) = (nx) - (n\delta)$ for $n < N$. We now derive the following chain of inequalities:

$$|f(x) - f(x \pm \delta)| = \left| \sum_1^N \left[\frac{(nx)}{n^2} - \frac{(nx \pm n\delta)}{n^2} \right] + \left[\sum_{N+1}^\infty \frac{(nx)}{n^2} - \frac{(nx \pm n\delta)}{n^2} \right] \right|$$

$$\leq \left| \sum_1^N \left[\frac{(nx)}{n^2} - \frac{(nx \pm n\delta)}{n^2} \right] \right| + \left| \sum_{N+1}^\infty \frac{(nx)}{n^2} \right| + \left| \sum_{N+1}^\infty \frac{(nx \pm n\delta)}{n^2} \right|$$

$$< \left| \sum_1^N \left[\frac{(nx)}{n^2} - \frac{(nx \pm n\delta)}{n^2} \right] \right| + \frac{\epsilon}{4} + \frac{\epsilon}{4}$$

$$< \left| \sum_1^N \left[\frac{(nx)}{n^2} - \frac{(nx)}{n^2} \pm \frac{(n\delta)}{n^2} \right] \right| + \frac{\epsilon}{2}$$

$$= \left| \sum_1^N \left[\frac{\pm(n\delta)}{n^2} \right] \right| + \frac{\epsilon}{2}$$

We chose our value of δ after we fixed a particular value of N, so we could have chosen δ small enough to make this sum less than $\epsilon/2$. In doing so, we have

$$|f(x) - f(x \pm \delta)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

And this would hold not just for $x \pm \delta$, but for all y such that $|x - y| < \delta$. That is:

$$|x - y| < \delta \to |f(x) - f(y)| < \epsilon$$

which means that f is continuous at x.

Lemma 1: $\sum (nx)/n^2$ **converges uniformly to** f

This is an immediate consequence of the fact that $(nx) < 1$ for all n, x and that $\sum 1/n^2$ converges. Let $\epsilon > 0$ be given. By the Cauchy criterion we can choose N such that

$$b > a > N \rightarrow \sum_a^b \frac{1}{n^2} < \epsilon$$

and therefore we have

$$\left| f(x) - \sum_{n=1}^{N} \frac{(nx)}{n^2} \right| = \left| \sum_{n=1}^{\infty} \frac{(nx)}{n^2} - \sum_{n=1}^{N} \frac{(nx)}{n^2} \right|$$

$$= \left| \sum_{n=N+1}^{\infty} \frac{(nx)}{n^2} \right| \leq \sum_{n=N+1}^{\infty} \left| \frac{(nx)}{n^2} \right| < \sum_{n=N+1}^{\infty} \left| \frac{1}{n^2} \right| < \epsilon$$

f **is discontinuous at every rational number**

When nx is not an integer the following limits are identical:

$$\lim_{t \to x+} \frac{(nt)}{n^2} = \frac{(nx)}{n^2} = \lim_{t \to x-} \frac{(nt)}{n^2}$$

When nx is an integer then this equality no longer holds. Instead, we have

$$\lim_{t \to x+} \frac{(nt)}{n^2} = \frac{0}{n^2}, \quad \lim_{t \to x-} \frac{(nt)}{n^2} = \frac{1}{n^2}$$

We know by lemma 1 that $\sum (nx)/n^2$ converges uniformly to f, so we also have (by theorem 7.11):

$$\sum_{n=1}^{\infty} f_n(x+) = f(x+), \quad \sum_{n=1}^{\infty} f_n(x-) = f(x-)$$

We know that $\sum 1/n^2$ converges, so by the Cauchy criterion we can choose N such that

$$b > a > N \rightarrow \sum_a^b \frac{1}{n^2} < \frac{\epsilon}{4}$$

If f were continuous at our rational point x, then we would have $|f(x+) - f(x-)| = 0$. But when we actually calculate this difference we find:

$$|f(x+) - f(x-)| = \left| \sum_{n=1}^{\infty} f_n(x+) - \sum_{n=1}^{\infty} f_n(x-) \right|$$

We determined that $f_n(x+) = f_n(x-)$ unless nx is an integer; this occurs when n is a multiple of q, at which point $nx = [mq]x = mq[p/q] = mp$. Most of the terms of the summations cancel out, leaving us with

$$|f(x+)-f(x-)| = \left| \sum_{m=1}^{\infty} f_{mq}(x+) - \sum_{m=1}^{\infty} f_{mq}(x-) \right| = \left| \sum_{m=1}^{\infty} \lim_{t \to x+} \frac{(mq)}{[mq]^2} - \lim_{t \to x-} \frac{(mq)}{[mq]^2} \right| = \left| \sum_{m=1}^{\infty} \frac{(0)}{n^2} - \frac{1}{[mq]^2} \right| = \left| \sum_{m=1}^{\infty} \frac{1}{[mq]^2} \right|$$

This is clearly not equal to zero and it can't be made arbitrarily small because we have no freedom to select particular values for m or q. Therefore $f(x+) \neq f(x-)$ and therefore f is not rational at x. But x was an arbitrary rational number, and therefore f is not continuous at any rational number.

Exercise 7.10b

We've shown that the discontinuities of f are the rational numbers, and these are clearly countable and clearly dense in $\mathbb{R}$.

Exercise 7.10c

To prove that $f \in \mathscr{R}$ we need only show that $(nx)/n^2 \in \mathscr{R}$ for any fixed value of n. We can then use the fact that $\sum (nx)/n^2$ converges uniformly to f (lemma 1) and theorem 7.16 to show that $\sum \int (nx)/n^2 = \int \sum (nx)/n^2 = \int f$.

Let $n \in \mathbb{N}$ be given. Let $[a, b]$ be an arbitrary interval. Assume without loss of generality that a and b are integers.

$$\int_a^b \frac{(nx)^2}{n}\, dx = \sum_{k=a}^{b-1} \int_k^{k+1} \frac{(nx)}{n^2}\, dx$$

For $0 \leq \delta < 1$ we have $(n(k+\delta)) = (nk + n\delta) = (n\delta)$, so we can integrate over $(0,1)$ instead of $(k, k+1)$ without changing the value of the integral.

$$\int_a^b \frac{(nx)^2}{n}\, dx = \sum_{k=a}^{b-1} \int_0^1 \frac{(nx)}{n^2}\, dx$$

As x ranges from 0 to 1, nx ranges from 0 to n. So we split up the interval $(0,1)$ into n intervals of length $1/n$:

$$\int_a^b \frac{(nx)^2}{n}\, dx = \sum_{k=a}^{b-1} \sum_{j=0}^{n-1} \int_{j/n}^{[j+1]/n} \frac{(nx)}{n^2}\, dx$$

To make this integral a bit more manageable we make the variable substitution $u = nx$. When $x = j/n$ we have $u = j$; when $x = [j+1]/n$ we have $u = j+1$; and we have $dx = \frac{1}{n}\, du$. Using this variable substitution in the previous integral:

$$\int_a^b \frac{(nx)^2}{n}\, dx = \sum_{k=a}^{b-1} \sum_{j=0}^{n-1} \int_j^{j+1} \frac{(u)}{n^3}\, du$$

Again, there is no difference between the value of (u) on the intervals $(j, j+1)$ and the value of (u) on the interval $(0,1)$:

$$\int_a^b \frac{(nx)^2}{n}\, dx = \sum_{k=a}^{b-1} \sum_{j=0}^{n-1} \int_0^1 \frac{(u)}{n^3}\, du$$

On the interval $(0,1)$ we have $(u) = u$, so we can finally start integrating this thing. After a bunch of trivial calculus, we have

$$\int_a^b \frac{(nx)^2}{n}\, dx = [b - a - 1][n]\frac{u^2}{2n^3} = \frac{b - a - 1}{n^2}$$

And therefore

$$\int_a^b f(x) = \int_a^b \sum_{n=1}^{\infty} \frac{(nx)}{n^2} = \sum_{n=1}^{\infty} \int_a^b \frac{(nx)}{n^2} = \sum_{n=1}^{\infty} \frac{b - a - 1}{n^2}$$

We know this rightmost sum converges (specifically, it converges to $\pi^2 [b - a - 1]/6$), therefore $\int_a^b f(x)$ exists and therefore $f \in \mathscr{R}$.

Exercise 7.11

Let $\epsilon > 0$ be given. We're told that $\sum f_n$ has uniformly bounded partial sums, so let M represent the upper bound of the partial sums of $|\sum f_n|$. We're told that $g_n \to 0$ uniformly, so we can find N such that

$$n > N \to g_n(x) < \frac{\epsilon}{M} \quad \text{for all } x$$

Let A_n represent the partial sum $\sum_{k=1}^n f_k g_k$. Our result is proven if we can prove that $\{A_n\}$ satisfies the Cauchy converge criterion. To do this, we choose $q > p > N$. Following the logic of theorem 3.42, we have

$$|A_q - A_{p-1}| = \left| \sum_{k=p}^q f_k g_k \right| = \left| \left[\sum_{k=p}^{q-1} A_n(g_n - g_{n+1}) \right] + A_q g_q - A_{p-1} g_p \right| \leq M \left| \left[\sum_{k=p}^{q-1} (g_n - g_{n+1}) \right] + g_q - g_p \right|$$

The $(g_n - g_{n+1})$ terms telescope down, leaving us with

$$|A_q - A_{p-1}| \leq M \left| \left[g_p - \sum_{k=p}^{q-1}(g_n - g_{n+1}) \right] + g_q + g_p \right| = 2M|g_p| \leq 2M|g_N| = \epsilon$$

Exercise 7.12

Let $\epsilon > 0$ be given. We're told that $0 \leq f_n, f \leq g$ and that $\int g < \infty$. Let $M = \int_0^\infty g$. Define $G(t)$ to be

$$G(t) = \int_0^c g(x)\, dx$$

Since $G(t)$ is continuous (theorem 6.30) and is strictly increasing (because $g > 0$) and has an upper bound of M, we know that

$$\lim_{t \to \infty} G(t) = M$$

we can find some N such that

$$n > N \to M - G(n) < \frac{\epsilon}{2}$$

which is equivalent to saying that

$$n > N \to \int_0^\infty g - \int_0^c g = \int_c^\infty g < \frac{\epsilon}{2} \tag{87}$$

We're given that $f_n \to f$ uniformly, so for all x and any specified value for c we can find some M such that

$$m > M \to |f_m(x) - f(x)| < \frac{\epsilon}{2c} \tag{88}$$

from which we can conclude that

$$m > M \to \left| \int_0^c f_m - f \right| \leq \int_0^c |f_m - f| \leq c\frac{\epsilon}{2c} = \frac{\epsilon}{2} \tag{89}$$

So, for the given value of $\epsilon > 0$ we choose c large enough that (88) holds and then, based on this choice of c, we choose n to be large enough that (89) holds. By the triangle inequality we then have

$$\left| \int_0^\infty f_n - f \right| \leq \left| \int_0^c f_n - f \right| + \left| \int_c^\infty f_n - f \right|$$

$$\leq \int_0^c |f_n - f| + \int_c^\infty |f_n - f|$$

$$\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

We can make this hold for arbitrary ϵ by taking n sufficiently large, which is simply saying that

$$\lim_{n \to \infty} \left| \int_0^\infty f_n - f \right| = 0$$

from which we conclude

$$\lim_{n \to \infty} \int_0^\infty f_n = \int_0^\infty f$$

Exercise 7.13a

For any choice of $x_1 \in [0, 1]$, the sequence $\{f_n^{(1)}(x_1)\}$ is a bounded sequence of real numbers in the compact domain $[0, 1]$. Therefore there exists some subsequence of functions $\{f_n^{(1)}\}$ for which the subsequence of real numbers $\{f_n^{(1)}(x_1)\}$ converges to some point in $[0, 1]$ (theorem 3.6a). Call this subsequence of functions $\{f_n^{(2)}\}$ and choose any $x_2 \in [0, 1]$. The sequence $\{f^{(2)}(x_2)$ is still a bounded sequence of real numbers in the compact domain $[0, 1]$, so we can find some subsequence of functions $\{f_n^{(3)}\}$ for which $\{f_n(x_1)\}$ and $\{f_n(x_2)\}$ both converge. This can be repeated a countable number of times to construct a sequence of functions $\{F_n\}$ for which $\{F_n(x)\}$ converges at all rational numbers and for all points of discontinuity for each f_n (this set is countable by theorems 4.30, 2.12, and 2.13).

Now let t be a rational number or a point of discontinuity for some f_n. We have constructed $\{F_n\}$ so that $\{F_n(t)\}$ converges to some point, so we simply define $f(t)$ for rational t to be

$$f(t) = \lim_{n \to \infty} F_n(t), \quad \text{if } t \text{ is rational or a point of discontinuity}$$

Now t be a rational number and a point at which f_n is continuous for all n. The rational numbers are dense in $\mathbb{R}$ so we can construct a sequence $\{r_n\}$ of rational numbers such that $\{r_n\} \to t$. Each F_n is continuous at t, so $\lim_{k \to \infty} F_n(r_k) = F_n(t)$ for every n. Therefore we simply *define* the value of $f(t)$ to be

$$f(t) = \lim_{n \to \infty} F_n(t), \quad \text{if } t \text{ is irrational and each } F_n \text{ is continuous at } t$$

Exercise 7.13: example

An example where a sequence $\{f_n\}$ of monotonically increasing functions converges pointwise to f and f is not continuous:

$$f_n(x) = \begin{cases} 0 & x \leq 0 \\ 1 - \frac{1}{1+n} & x \geq 1 \\ 1 - \frac{1}{1+nx} & 0 < x < 1 \end{cases}$$

$$\lim_{n \to \infty} f_n(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

Exercise 7.13b

Let $\alpha = \inf f(x)$ and let $\beta = \sup f(x)$ (these values may be $\pm\infty$). Let $\epsilon > 0$ be given. The function f must come arbitrarily close to its supremum and infimum on $\mathscr{R}$. Let a be a point at which $|f(a) - \alpha| < \epsilon$ and let b be a point at which $|f(b) - \beta| < \epsilon$. From the monotonicity of f we have

$$x < a \to |f(x) - \alpha| < \epsilon, \quad x > b \to |f(x) - \beta| < \epsilon \tag{90}$$

Proving uniform convergence is now a matter of proving uniform convergence on $[a, b]$.

We're given that f is continuous on $[a, b]$; therefore it's uniformly continuous on $[a, b]$ (theorem 4.19); therefore we can find some δ such that

$$|x - y| < \delta \to |f(x) - f(y) < \epsilon \tag{91}$$

We can cover $[a, b]$ with finitely many intervals of length $\delta/2$ ($[a, b]$ is finite, so we don't even need to rely on compactness for this). From each interval we choose a rational number q_i. We know that f converges pointwise at each rational number, and the set of q_is is a finite set, so we can find some integer N such that

$$n > N \to |F_n(q_i) - f(q_i)| < \epsilon \quad \text{for each } q_i \tag{92}$$

Now choose any $x \in [a, b]$. If $x = q_i$ for some i then by (92) we have $n > N \to |F_n(q_i) - f(q_i)| < \epsilon$. If $x \neq q_i$ then we can find q_i, q_{i+1} such that $q_i < x < q_{i+1}$ and $|q_{i+1} - q_i| < \delta$. By the triangle inequality:

$$|F_n(x) - f(x)| \leq |F_n(x) - F_n(q_{i+1})| + |F_n(q_{i+1}) - f(q_{i+1})| + |f(q_{i+1}) - f(x)|$$

Each F_n is monotonic, so we have $|F_n(x) - F_n(q_{i+1})| \leq |F_n(q_i) - F_n(q_{i+1})|$ so this last inequality becomes

$$|F_n(x) - f(x)| \leq |F_n(q_i) - F_n(q_{i+1})| + |F_n(q_{i+1}) - f(q_{i+1})| + |f(q_{i+1}) - f(x)|$$

The term $|F_n(q_{i+1}) - f(q_{i+1})|$ is $< \epsilon$ by (92). The term $|f(q_{i+1}) - f(x)|$ is $< \epsilon$ by (91). This leaves us with

$$|F_n(x) - f(x)| \leq |F_n(q_i) - F_n(q_{i+1})| + 2\epsilon$$

Additional applications of the triangle inequality gives us

$$|F_n(x) - f(x)| \leq |F_n(q_i) - f(q_i)| + |f(q_i) - f(q_{i+1})||f(q_i) - F_n(q_{i+1})| + 2\epsilon$$

The terms $|F_n(q_i) - f(q_i)|$ and $|f(q_i) - F_n(q_{i+1})|$ are both $< \epsilon$ by (92). The term $|f(q_i) - f(q_{i+1})|$ is $< \epsilon$ by (91). This leave us with

$$|F_n(x) - f(x)| \leq 5\epsilon$$

Exercise 7.14

Exercise 7.15

Let $\epsilon > 0$ be given. Each f_n is equicontinous on $[0, 1]$, which means that there exists some δ such that, for all n,

$$|0 - y| < \delta \rightarrow |f_n(0) - f_n(y)| < \epsilon$$

By the definition of f this means that, for all $y \in (0, 1)$ and all $n \in \mathbb{N}$,

$$|y| < \delta \rightarrow |f(0) - f(ny)| < \epsilon$$

By taking n sufficiently large and choosing y appropriately we can cause ny to take on any value in $(0, \infty)$. So we conclude that f is a constant function on $[0, \infty)$.

Exercise 7.16

See part (b) of exercise 7.13

Exercise 7.17

Exercise 7.18

Lemma 1: $\{F_n\}$ is uniformly bounded

We're told that $\{f_n\}$ is a uniformly bounded sequence: let M be the upper bound of $\{|f_n|\}$. For each n, we have

$$|F_n(x)| = \left| \int_a^x f_n(t)\, dt \right| \leq \int_a^x |f_n(t)|\, dt \leq \int_a^b |f_n(t)|\, dt \leq (b-a)M$$

and therefore $\{|F_n|\}$ is uniformly bounded. $\square$

Lemma 2: $\{F_n\}$ is equicontinuous

Let $\epsilon > 0$ be given. Let $\delta = \epsilon/M$.

$$|y - x| < \delta \rightarrow |F_n(y) - F_n(x)| = \left| \int_x^y f_n(t)\, dt \right| \leq \int_x^y |f_n(t)\, dt| \leq (y-x)M < \delta M < \epsilon$$

$\square$

Therefore, by theorem 7.25, we know that $\{F_n\}$ contains a uniformly convergent subsequence.

Exercise 7.19

This was solved during class

Exercise 7.21

We know that $\mathscr{A}$ separates points because $f(e^{i\theta}) = e^{i\theta} \in \mathscr{A}$. For this same function, $|f(e^{i\theta})| = |e^{i\theta}| = 1$ and so $\mathscr{A}$ vanishes at no point of K. But the function $f(e^{i\theta}) = -e^{i\theta}$ is a continuous function that is not in the closure of $\mathscr{A}$.