# Real Analysis I 

## Same Questions and Answers Preparation Document

Sept - Dec 2023

Prof. Dr. Marc Paolella

- The course covers the standard topics in a first (undergraduate for mathematics or applied mathematics majors) course in real analysis, the ordering of which is nearly standard in most books. In this first semester course, we reach and cover univariate Riemann integration. In this document, there are four sections, partitioned based on the four in-class exams given in Fall 2023, and using pieces of the two books, Fitzpatrick, and Stoll, about equally. The first section is the shortest, covering the basics of sets and functions (e.g., injection, surjection, inverse), and the concepts of completeness, supremum and infimum.

Sections 2 and 3 are more substantial, covering (sub-)sequences, Cauchy sequences and series, sequential and topological compactness, Bolzano-Weierstrass, Heine-Borel, and much of this done so in the more general context of metric spaces. Section 4 embodies a solid amount of material on (uniform) continuity of functions and their relation with compactness; and several questions on differentiation, with highlights being a deeper study of Darboux' theorem in exercise 11; our first and only taste of numerical analysis, in exercise 23; and an introduction to convexity and concavity, in exercise 24.

- References are made to other books in abbreviated form, as opposed to a correct citation and having a bibliography. For example, "Heil, Metrics" refers to Christopher Heil's book entitled Metrics, Norms, Inner Products, and Operator Theory, and Field refers to Michael Field's Essential Real Analysis. Often, just the author's or authors' names are given, e.g., Al-Gwaiz \& Elsanousi, Ash, Bartle \& Sherbert, Conway, Field, Fitzpatrick, Garling, Ghorpade \& Limaye, Giv, Jacob \& Evans, Junghenn, Kuttler, Laczkovich \& Sós, Lebl, Loya, Mattuck, Nair, Petrovic, Pons, Sasane, Stoll, Tao, Terrell, Thomson, Bruckner \& Bruckner, Trench, Zorich, etc.. Many of these authors have two or more analysis books that form a set, and unless otherwise specified, I mean the first book in their set, e.g., Laczkovich and Sós refers to their first volume, Real Analysis: Foundations and Functions of One Variable.
- I sincerely thank Ralf Blöchlinger and Chen Lan (students no less, in my 2023 course) for having gone carefully through this entire document, finding and fixing numerous issues, too many to list, as well as adding proofs. Professor Christopher Heil also contributed proofs and finding errors (such as my gaffe in exercise 8 in section 3). I also wish to advertise his - outstanding and highly accessible - books on metric spaces, and on measure theory. Professor Manfred Stoll also kindly and patiently answers all my questions, further making his book a pleasure to use for our class.


## Contents

1 Stoll Chapter 1; Fitzpatrick Chapter 1 ..... 2
2 Fitzpatrick, Chapters 2 and 3 ..... 8
3 Stoll, Chapter 3 ..... 24
4 Stoll Chapters 4, 5, 6 (6.1 to 6.4), and Fitzpatrick Chapter 4 ..... 54
4.1 Stoll Chapter 4 ..... 54
4.2 Stoll Chapter 5; Fitzpatrick Chapter 4 (Differentiation) ..... 71
4.3 Stoll Chapter 6 (Integration, FTC) ..... 91

## 1 Stoll Chapter 1; Fitzpatrick Chapter 1

Every day you may make progress. Every step may be fruitful. Yet there will stretch out before you an ever-lengthening, ever-ascending, ever-improving path. You know you will never get to the end of the journey. But this, so far from discouraging, only adds to the joy and glory of the climb.

Sir Winston Churchill
Tell me, tell me what you're after. I just want to get there faster. Smashing Pumpkins, Siva

1. State the definition of a function (in terms of sets). Define the domain, codomain, range, image, preimage, and what it means for a function to be one-to-one, onto, and to have an inverse. State the Archimedean property. What are the other words for one-to-one, onto, and having an inverse?
ANS: See e.g., Fitzpatrick, Ch. 1; and Stoll, Sec. 1.2.
2. Define supremum (also called least upper bound), infimum (greatest lower bound), and state the completeness axiom. Give an example of a set that does not contain its supremum, but does contain its infimum.

ANS: See e.g., Fitzpatrick, Sec. 1.1. Example is $[0,1)$.
3. Give two examples dense sets in $\mathbb{R}$, and an infinite set that is not dense in $\mathbb{R}$.

ANS: The rationals $\mathbb{Q}$, the irrationals $\mathbb{I}=\mathbb{R} \backslash \mathbb{Q}$; and the natural numbers $\mathbb{N}$.
4. State and prove the "reverse triangle inequality".

ANS: For $a, b \in \mathbb{R},||a|-|b|| \leq|a+b|$. To prove this: $|a|=|(a+b)+(-b)|<=|a+b|+|b|$, so $|a|-|b|<=|a+b|$. Switching $a$ and $b$ gives $|b|-|a|<=|a+b|$. Combining gives $||a|-|b|| \leq|a+b|$. Further note that replacing $b$ with $-b$ gives $\| a|-|b|| \leq|b-a|$.
5. Let $f: X \rightarrow Y$, and let $\left\{A_{i}\right\}$ be a family of (possibly uncountably many) subsets of $Y$. Prove:

$$
\begin{equation*}
f^{-1}\left(\bigcup_{i} A_{i}\right)=\bigcup_{i} f^{-1}\left(A_{i}\right), \quad f^{-1}\left(\bigcap_{i} A_{i}\right)=\bigcap_{i} f^{-1}\left(A_{i}\right) \tag{1}
\end{equation*}
$$

and

$$
\begin{equation*}
f^{-1}\left(A_{i}^{c}\right)=\left[f^{-1}\left(A_{i}\right)\right]^{c} \tag{2}
\end{equation*}
$$

ANS: (Ash, Thm 4.1.5) ${ }^{1}$

$$
\text { - } \begin{aligned}
x \in f^{-1}\left(\bigcup_{i} A_{i}\right) & \Longleftrightarrow f(x) \in \bigcup_{i} A_{i} \Longleftrightarrow \exists i: f(x) \in A_{i} \\
& \Longleftrightarrow \exists i: x \in f^{-1}\left(A_{i}\right) \Longleftrightarrow x \in \bigcup_{i} f^{-1}\left(A_{i}\right) \\
\bullet x \in f^{-1}\left(\bigcap_{i} A_{i}\right) & \Longleftrightarrow f(x) \in \bigcap_{i} A_{i} \Longleftrightarrow \forall i: f(x) \in A_{i} \\
& \Longleftrightarrow \forall i: x \in f^{-1}\left(A_{i}\right) \Longleftrightarrow x \in \bigcap_{i} f^{-1}\left(A_{i}\right) \\
\bullet x \in f^{-1}\left(A_{i}^{c}\right) & \Longleftrightarrow f(x) \in A_{i}^{c} \Longleftrightarrow f(x) \notin A_{i} \\
& \Longleftrightarrow x \notin f^{-1}\left(A_{i}\right) \Longleftrightarrow x \in\left[f^{-1} A_{i}\right]^{c}
\end{aligned}
$$

[^0]6. Let $f: X \rightarrow Y$. Prove:
(a) If $A \subseteq X$, then $A \subset f^{-1}[f(A)]$. (Ash, Thm 4.1.8)

ANS: $x \in A \Rightarrow f(x) \in f(A) \Rightarrow x \in f^{-1}[f(A)]$.
(b) If $B \subseteq Y$, then $f\left[f^{-1}(B)\right] \subset B$. Give a strict inequality example. (Ash, Thm 4.1.8)

ANS: Recall $A \wedge B$ means conditions $A$ and $B$ both hold. ( $\vee$ is or).

$$
\begin{aligned}
y \in f\left[f^{-1}(B)\right] & \Rightarrow \exists x_{y}:\left[x_{y} \in f^{-1}(B)\right] \wedge\left[y=f\left(x_{y}\right)\right] \\
& \Rightarrow f\left(x_{y}\right) \in B \Rightarrow y \in B
\end{aligned}
$$

Example: We need a non-onto function, e.g., $X=Y=\mathbb{R}, f(x)=x^{2}$, so with $B=(-1,1)$, $f^{-1}(B)=[0,1)$, and $f\left(f^{-1}(B)\right)=[0,1) \subset B$.
(c) $[f$ onto $Y] \Longleftrightarrow\left[\forall B \subset Y: f\left(f^{-1}(B)\right)=B\right]$. (Terrell, Thm 1.2.3; Exercise 1.2.1)

ANS: Let $f: X \rightarrow Y$ be onto $Y$; and let $B \subset Y$. We need to prove $f\left[f^{-1}(B)\right] \supset B$.

$$
\begin{align*}
{[y \in B \subset Y] \wedge[f \text { onto }] } & \Rightarrow \exists x_{y} \in X: y=f\left(x_{y}\right) \Rightarrow x_{y} \in\left\{f^{-1}(\{y\})\right\}  \tag{3}\\
& \Rightarrow f\left(x_{y}\right) \subset f\left(f^{-1}(\{y\})\right) \subset f\left(f^{-1}(B)\right),
\end{align*}
$$

i.e., $B \subset f\left[f^{-1}(B)\right]$. Observe in (3) we need to write $x_{y} \in\left\{f^{-1}(\{y\})\right\}$ instead of $x_{y}=f^{-1}(y)$ because $f$ may not be one-to-one. If $f$ is additionally one-to-one, then it is a bijection, and $A=f^{-1}[f(A)]$ and $B=f\left[f^{-1}(B)\right]$.
7. Let $f: X \rightarrow Y$, and $A_{1}, A_{2} \subset X$. Give an example to show that $f\left(A_{1} \cap A_{2}\right) \neq f\left(A_{1}\right) \cap f\left(A_{2}\right)$.

ANS: Let $f: \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ be given by $f(x)=x^{2}$, noting this is onto, but not one-to-one. Let $A_{1}=(-1,0]$ and $A_{2}=[0,1)$, so that $A_{1} \cap A_{2}=\{0\}$ and $f\left(A_{1} \cap A_{2}\right)=0$; and $f\left(A_{1}\right)=f\left(A_{2}\right)=$ $[0,1)=f\left(A_{1}\right) \cap f\left(A_{2}\right)$.
8. Prove (a particular version of) Bernoulli's inequality, namely

$$
\begin{equation*}
\forall h \geq 0, \forall n \in \mathbb{N},(1+h)^{n} \geq 1+n h, \tag{4}
\end{equation*}
$$

in two ways:
(a) By induction.

ANS: As in Stoll, p. 18 (and noting (4) actually holds for $b>-1$ ): For $n=1,(1+h)^{1}=1+h$. As equality holds, the inequality is certainly valid. Assume that the inequality is true when $n=k, k \geq 1$. Then for $n=k+1,(1+h)^{k+1}=(1+h)^{k}(1+h)$. By the induction hypothesis and the fact that $(1+h)>0$, we have

$$
\begin{aligned}
(1+h)^{k+1} & =(1+h)^{k}(1+h) \\
& \geq(1+k h)(1+h)=1+(k+1) h+k h^{2} \\
& \geq 1+(k+1) h .
\end{aligned}
$$

Thus, by the principle of mathematical induction, (4) holds for all $n \in \mathbb{N}$.
(b) Use of the binomial formula. (Fitzpatrick, p. 20, \#11)

ANS:

$$
(1+h)^{n}=\sum_{i=0}^{n}\binom{n}{i} h^{i}=1+n h+\binom{n}{2} h^{2}+\cdots+n h^{n-1}+h^{n} \geq 1+n h .
$$



Figure 1: From Lebl, Basic Analysis I, p. 18
9. Show a tabular array to indicate why the set $\mathbb{N} \times \mathbb{N}$ is countable.

ANS: See Figure 1.
10. Make a simple argument that the irrational numbers are uncountable.

ANS: $\mathbb{R}$ is the union of the rationals, $\mathbb{Q}$, and the irrationals, $\mathbb{I}$. If $\mathbb{I}$ were countable, then $\mathbb{R}$ would be the union of two countable sets, $\mathbb{Q}$ and $\mathbb{I}$, which is countable. Thus, $\mathbb{I}$ is uncountable.
11. (Stoll, p. 29, \#14) Let $A$ and $B$ be subsets of $\mathbb{R}$. Define

$$
A+B=\{a+b: a \in A, b \in B\} \quad \text { and } \quad A \cdot B=\{a b: a \in A, b \in B\} .
$$

(a) If $A$ and $B$ are nonempty and bounded above, prove $\sup (A+B)=\sup A+\sup B$.
(b) Give an example of two nonempty bounded sets $A$ and $B$ for which $\sup (A \cdot B) \neq(\sup A)(\sup B)$.

ANS: For (a), as in Stoll:

- $\sup (A+B) \leq \sup A+\sup B$ :

Both $A$ and $B$ are non-empty and bounded above, so $\alpha=\sup A$ and $\beta=\sup B$ exist in $\mathbb{R}$. Therefore, $a+b \leq \alpha+\beta$ for all $a \in A$ and $b \in B$. This means $\alpha+\beta$ is an upper bound for $A+B$, and, by definition of $\sup , \gamma=\sup (A+B) \leq \alpha+\beta$.

- $\sup (A+B) \geq \sup A+\sup B$ :

Let $\gamma=\sup (A+B)$, which is an upper bound for $A+B$. Then $a+b \leq \gamma$ for all $a \in A$ and $b \in B$. Let $b \in B$ be arbitrary but fixed. Then $a \leq \gamma-b$ for all $a \in A$. Thus $\gamma-b$ is an upper bound for $A$ and, hence, $\sup A=\alpha \leq \gamma-b$. As this holds for all $b \in B$, we also have $b \leq \gamma-\alpha$ for all $b \in B$. Thus $\sup B=\beta \leq \gamma-\alpha$; i.e., $\alpha+\beta \leq \gamma$.
For (b): $A=B=[-1,0]$, non-empty and bounded. $(\sup A)(\sup B)=0$, but $\sup (A \cdot B)=1$.
12. (Stoll, p. 29, \#15) Let $f$ and $g$ be real-valued functions defined on a nonempty set $X \subset \mathbb{R}$ with bounded ranges. Prove:
(a) $\sup \{f(x)+g(x): x \in X\} \leq \sup \{f(x): x \in X\}+\sup \{g(x): x \in X\}$.
(b) Provide an example for which equality does not hold in (a).
(c) If $f(x) \leq g(x)$ for all $x \in X$, then

$$
\begin{equation*}
\sup \{f(x): x \in X\} \leq \sup \{g(x): x \in X\} \tag{5}
\end{equation*}
$$

This result is easy and important; we can call it monotonicity of the supremum.
(d) $\sup \{f(x)+g(y): x \in X, y \in X\}=\sup \{f(x): x \in X\}+\sup \{g(x): x \in X\}$.

ANS: For (a), as in Stoll: Let $\alpha=\sup \{f(x): x \in X\}$ and $\beta=\sup \{g(x): x \in X\}$. The ranges of $f$ and $g$ are bounded, so $\alpha$ and $\beta$ are finite, with $f(x)+g(x) \leq \alpha+\beta$ for every $x \in X$. Thus, $\alpha+\beta$ is an upper bound for $\{f(x)+g(x): x \in X\}$, and $\sup \{f(x)+g(x): x \in X\} \leq \alpha+\beta$.
An answer to (b) can be: $X=[0,1], f(x)=x$, and $g(x)=-x$. The lhs is 0 ; the rhs is $1+0$.
For (c), as in Stoll: Let $\alpha=\sup \{g(x): x \in X\}$, so that $g(x) \leq \alpha$ for all $x \in X$. Thus, by hypothesis, $f(x) \leq \alpha$ for all $x \in X$. Therefore, $\alpha$ is an upper bound for $\{f(x): x \in X\}$, and, as a consequence, $\sup \{f(x): x \in X\} \leq \alpha=\sup \{g(x): x \in X\}$.
For part (d): Apply 11a, taking $A$ to be the range of $f$; and $B$ to be the range of $g$.
13. Let $a>b \geq 0$ and $x \in(0,1)$. Show $\left(1+x^{b}\right)^{a}>\left(1+x^{a}\right)^{b}$.

From: Problems and Solutions, The College Mathematics Journal (1998) 29:1, 66-72.
DOI: 10.1080/07468342.1998.11973918.
ANS: $x^{a}<x^{b}$, so $\left(1+x^{a}\right)^{b}<\left(1+x^{b}\right)^{b}<\left(1+x^{b}\right)^{a}$.
14. Prove: If $f$ and $g$ are injective, then so is $g \circ f$.

ANS: Both $f$ and $g$ are injective, so $f\left(x_{1}\right)=f\left(x_{2}\right) \Longrightarrow x_{1}=x_{2}$ and $g\left(y_{1}\right)=g\left(y_{2}\right) \Longrightarrow y_{1}=y_{2}$. Therefore, $(g \circ f)\left(x_{1}\right)=(g \circ f)\left(x_{2}\right) \Longrightarrow g\left(f\left(x_{1}\right)\right)=g\left(f\left(x_{2}\right)\right) \Longrightarrow f\left(x_{1}\right)=f\left(x_{2}\right) \Longrightarrow x_{1}=x_{2}$. Hence, $g \circ f$ is also injective.
15. Prove: If $f$ and $g$ are surjective, then so is $g \circ f$.

ANS: Both $f$ and $g$ are surjective, so $\forall y \in Y, \exists x \in X$ such that $f(x)=y$ and $\forall z \in Z, \exists y \in Y$ such that $g(y)=z$. Thus, $\forall z \in Z, \exists x \in X$ such that $g(f(x))=z$.
16. If $g \circ f$ is injective, then $f$ is injective.

ANS: $(g \circ f)\left(x_{1}\right)=(g \circ f)\left(x_{2}\right) \Longrightarrow x_{1}=x_{2}$. Thus $f\left(x_{1}\right)=f\left(x_{2}\right) \Longrightarrow g\left(f\left(x_{1}\right)\right)=g\left(f\left(x_{2}\right)\right) \Longrightarrow$ $(g \circ f)\left(x_{1}\right)=(g \circ f)\left(x_{2}\right) \Longrightarrow x_{1}=x_{2}$.
17. If $g \circ f$ is surjective, then $g$ is surjective.

ANS: With $y=f(x)$,

$$
\forall z \in Z: \exists x \in X \text { such that }(g \circ f)(x)=z \quad \Longrightarrow \quad \forall z \in Z, \exists y \in Y \text { such that } g(y)=z .
$$

18. Here we state some results that are occasionally of great use. They appear in many books; I took them from Junghenn, p. $7, \# 4$ (first three); and p. 18, \#15 (last one). Let $n \in \mathbb{N}$. Prove the following identities without using mathematical induction:
(a) $x^{n}-y^{n}=(x-y) \sum_{j=1}^{n} x^{n-j} y^{j-1}$.
(b) $x^{n}+y^{n}=(x+y) \sum_{j=1}^{n}(-1)^{j-1} x^{n-j} y^{j-1}$ if $n$ is odd.
(c) $x^{-n}-y^{-n}=(y-x) \sum^{n} x^{j-n-1} y^{-j}$ if $x \neq 0$ and $y \neq 0$.
(d) For $a, b>0$ and $n \in \mathbb{N}, a^{1 / n}-b^{1 / n}=(a-b)\left(\sum_{j=1}^{n} a^{1-j / n} b^{(j-1) / n}\right)^{-1}$.

ANS: For (a),

$$
(x-y) \sum_{j=1}^{n} x^{n-j} y^{j-1}=\sum_{j=1}^{n} x^{n-j+1} y^{j-1}-\sum_{j=1}^{n} x^{n-j} y^{j}=\sum_{j=0}^{n-1} x^{n-j} y^{j}-\sum_{j=1}^{n} x^{n-j} y^{j}=x^{n}-y^{n} .
$$

For (b), replace $y$ in part (a) by $-y$. For (c), replace $x$ and $y$ in part (a) by $x^{-1}$ and $y^{-1}$, respectively. For (d), use part (a) with $x=a^{1 / n}, y=b^{1 / n}$.
19. (AM-GM Inequality). Let $n \in \mathbb{N}$ and let $a_{1}, \ldots, a_{n}$ be nonnegative real numbers. Then the arithmetic mean of $a_{1}, \ldots, a_{n}$ is greater than or equal to their geometric mean, that is,

$$
\begin{equation*}
\frac{a_{1}+\cdots+a_{n}}{n} \geq \sqrt[n]{a_{1} \cdots a_{n}} \tag{6}
\end{equation*}
$$

Moreover, equality holds if and only if $a_{1}=\cdots=a_{n}$. Prove (6).
NOTE: This is generalized to non-equal weights in (76), where we will see the utility of the concept of convexity and Jensen's inequality.
ANS: (Ghorpade and Limaye, Prop 1.11) If some $a_{i}=0$, then the result is obvious. Assume $a_{i}>0$. Let $g=\left(a_{1} \cdots a_{n}\right)^{1 / n}$ and $b_{i}=a_{i} / g$ for $i=1, \ldots, n$. Then $b_{1}, \ldots, b_{n}$ are positive and $b_{1} \cdots b_{n}=1$. We shall now show, using induction on $n$, that $b_{1}+\cdots+b_{n} \geq n$. This is clear if $n=1$ or if each of $b_{1}, \ldots, b_{n}$ equals 1. Suppose $n>1$ and not every $b_{i}$ equals 1 . Then $b_{1} \cdots b_{n}=1$ implies that among $b_{1}, \ldots, b_{n}$ there is a number $<1$ as well as a number $>1$. Relabeling $b_{1}, \ldots, b_{n}$ if necessary, we may assume that $b_{1}<1$ and $b_{n}>1$. Let $c_{1}=b_{1} b_{n}$. Then $c_{1} b_{2} \cdots b_{n-1}=1$, and hence by the induction hypothesis $c_{1}+b_{2}+\cdots+b_{n-1} \geq n-1$. Now observe that

$$
\begin{aligned}
b_{1}+\cdots+b_{n} & =\left(c_{1}+b_{2}+\cdots+b_{n-1}\right)+b_{1}+b_{n}-c_{1} \\
& \geq(n-1)+b_{1}+b_{n}-b_{1} b_{n}=n+\left(1-b_{1}\right)\left(b_{n}-1\right)>n
\end{aligned}
$$

because $b_{1}<1$ and $b_{n}>1$. This proves that $b_{1}+\cdots+b_{n} \geq n$, and moreover, the inequality is strict unless $b_{1}=\cdots=b_{n}=1$. Substituting $b_{i}=a_{i} / g$, we obtain the desired result.
20. (Cauchy-Schwarz Inequality). Let $n \in \mathbb{N}$ and let $a_{1}, \ldots, a_{n}$ and $b_{1}, \ldots, b_{n}$ be any real numbers. Then

$$
\sum_{i=1}^{n} a_{i} b_{i} \leq\left(\sum_{i=1}^{n} a_{i}^{2}\right)^{1 / 2}\left(\sum_{i=1}^{n} b_{i}^{2}\right)^{1 / 2}
$$

Moreover, equality holds if and only if $a_{1}, \ldots, a_{n}$ and $b_{1}, \ldots, b_{n}$ are proportional to each other, that is, if $a_{i} b_{j}=a_{j} b_{i}$ for all $i, j=1, \ldots, n$.
This result is fundamental, and we will use it when we arrive at metric spaces and the Euclidean distance for $\mathbb{R}^{n}$. It can be proved in several ways. One fast way is to use the non-obvious and anyway interesting Lagrange Identity

$$
\begin{equation*}
\left(\sum_{i=1}^{n} a_{i}^{2}\right)\left(\sum_{j=1}^{n} b_{j}^{2}\right)-\left(\sum_{i=1}^{n} a_{i} b_{i}\right)^{2}=\sum_{1 \leq i<j \leq n}\left(a_{i} b_{j}-a_{j} b_{i}\right)^{2} \tag{7}
\end{equation*}
$$

but there are easier methods of proof for Cauchy-Schwarz that obviously then imply (7).
ANS: (Ghorpade and Limaye, Prop 1.12) Let $\alpha:=\left(\sum_{i=1}^{n} a_{i}^{2}\right)^{1 / 2}$ and $\beta:=\left(\sum_{i=1}^{n} b_{i}^{2}\right)^{1 / 2}$. If $\alpha=0$ or $\beta=0$, then $a_{1}=\cdots=a_{n}=0$ or $b_{1}=\cdots=b_{n}=0$, and the desired inequality as well as the assertion about equality is clear. So assume that $\alpha \neq 0$ and $\beta \neq 0$. Now for all $a, b \in \mathbb{R}$. by
considering $(a-b)^{2}$, we see that $a b \leq \frac{1}{2}\left(a^{2}+b^{2}\right)$ and that equality holds if and only if $a=b$. Thus, for each $i=1, \ldots, n, \frac{a_{i}}{\alpha} \cdot \frac{b_{i}}{\beta} \leq \frac{1}{2}\left(\frac{a_{i}^{2}}{\alpha^{2}}+\frac{b_{i}^{2}}{\beta^{2}}\right)$, and equality holds if and only if $\frac{a_{i}}{\alpha}=\frac{b_{i}}{\beta}$.
By summing from $i=1$ to $n$, we obtain

$$
\sum_{i=1}^{n} a_{i} b_{i} \leq \frac{\alpha \beta}{2}\left(\frac{\alpha^{2}}{\alpha^{2}}+\frac{\beta^{2}}{\beta^{2}}\right)=\alpha \beta
$$

Equality holds iff $a_{i} / \alpha=b_{i} / \beta$ for all $i=1, \ldots, n$. This yields the desired result.

## 2 Fitzpatrick, Chapters 2 and 3

Ignorance may be quickly eliminated by the spread of knowledge, while a prejudice is relatively independent of knowledge.

Gary S. Becker (1957) ${ }^{2}$
Only a warrior chooses pacifism. Others are condemned to it.
Eduardo Garcia

1. State the definition of the convergence of a sequence.

ANS: Fitzpatrick, page 26, definition.
2. State the definition of a closed set in terms of sequences.

ANS: Fitzpatrick, page 37, definition.
3. (Dense sets)

Note: For the exam, either this question is asked, or the next question, number 4. If you ask this question number 3 , then you need to tell the students to know the definition of dense from Fitzpatrick chapter 1, and memorize Fitzpatrick, page 36, Prop 2.19, and its proof.
(a) State the definition of dense (Fitzpatrick chap 1; not using sequences).

ANS: (Fitzpatrick, p. 15, Def) A set $S$ of real numbers is said to be dense in $\mathbb{R}$ provided that every interval $I=(a, b)$, where $a<b$, contains a member of $S$.
(b) Prove: A set $S$ is dense in $\mathbb{R}$ if and only if every number $x \in \mathbb{R}$ is the limit of a sequence in $S$. ANS: (Fitzpatrick, p. 36, Prop 2.19) First, assume that the set $S$ is dense in $\mathbb{R}$. Fix a number $x$. Let $n$ be an index. By the denseness of $S$ in $\mathbb{R}$, there is a member of $S$ in the interval $(x, x+1 / n)$. Choose a member of $S$ that belongs to this interval and label it $s_{n}$. This defines a sequence $\left\{s_{n}\right\}$ that has the property that $\left|s_{n}-x\right|<1 / n$, for every index $n$. Since the sequence $\{1 / n\}$ converges to 0 , it follows from the Comparison Lemma (exercise 18m; Fitzpatrick, p. 28, Lemma 2.9) that $\left\{s_{n}\right\}$ converges to $x$, and, by the above choice, $\left\{s_{n}\right\}$ is a sequence in $S$.
It remains to prove the converse. Suppose that the set $S$ has the property that every number is the limit of a sequence in $S$. We will show that $S$ is dense in $\mathbb{R}$. Indeed, consider an interval $(a, b)$. We must show that this interval contains a point of $S$. Consider the midpoint $s=(a+b) / 2$ of the interval. By assumption, there is a sequence $\left\{s_{n}\right\}$ of points in $S$ that converges to $s$. Define $\epsilon \equiv(b-a) / 2$. Then $\epsilon>0$. By the definition of a convergent sequence, there is an index $N$ such that $s_{n}$ belongs to $(s-\epsilon, s+\epsilon)$ for each index $n \geq N$. However,

$$
(s-\epsilon, s+\epsilon)=(a, b)
$$

The point $s_{N}$ belongs to $S$ and also belongs to $(a, b)$. Thus, $S$ is dense in $\mathbb{R}$.
4. First prove: For any open interval $J=(a, b)$ with $a, b \in \mathbb{R}, J$ contains a rational number. Second (and leave some space, and label your solutions clearly): Show that every $x \in \mathbb{R}$ is the limit of a sequence of rational numbers. Hint: Try to recall a useful theorem-you do not need to prove it, but rather just state it exactly, and apply it.
If this question is asked, instead of the previous one, then the students need to be told to memorize the definition of dense from Fitzpatrick. Further, they need to know and memorize the proof of Fitzpatrick Theorem 1.9.

[^1]ANS: The first is Fitzpatrick Theorem 1.9.
For the second question: This is trivially true if $x$ is rational. Recall the definition of dense (as given in Fitzpatrick): A subset $S$ of $\mathbb{R}$ is dense in $\mathbb{R}$ if every open interval $(a, b)$ contains a point in $S$. Then Prop 2.19 says: A set $S$ is dense in $\mathbb{R}$ if and only if every number $x$ in $\mathbb{R}$ is the limit of a sequence in $S$. The student of course does not need to know the number of the theorem in Fitzpatrick, but instead just state the theorem. Note it is also not required to prove Prop 2.19.
So, the question is asking to show that the rationals are dense in $\mathbb{R}$, and this means we need to show that, for any open interval $J=(a, b)$ with $a, b \in \mathbb{R}, J$ contains a rational number. This (Theorem 1.9 Fitzpatrick) was the first question.
5. Derive the limit as $n \rightarrow \infty$ of $n^{1 / n}$, for $n \in \mathbb{N}$. (Fitzpatrick, page 33, \#10)

ANS: From Stoll, page 91: Let $x_{n}=n^{1 / n}-1$. Since $x_{n}$ is positive, by the binomial theorem $n=\left(1+x_{n}\right)^{n} \geq \frac{n(n-1)}{2} x_{n}^{2}$ for all $n \geq 2$. Therefore, $x_{n}^{2} \leq \frac{2}{n-1}$ for all $n \geq 2$, and as a consequence $0 \leq x_{n} \leq \sqrt{\frac{2}{n-1}}$. The Squeeze Theorem implies $\lim _{n \rightarrow \infty} n^{1 / n}=1$.
6. Let $E$ be a nonempty subset of $\mathbb{R}$. Assume $E$ is bounded above and $a:=\sup E$. Prove there is a sequence $\left\{a_{n}\right\}$ such that $a_{n} \in E$ for all $n \in \mathbb{N}$ and $a_{n} \rightarrow a$. (Ghorpade and Limaye, Coro 2.6)
ANS: Suppose E is bounded above. Let $a:=\sup E$. Then for every $n \in \mathbb{N}$, there is $a_{n} \in E$ such that $a_{n}>a-\left(\frac{1}{n}\right)$. Also since $a_{n} \leq a$ for all $n \in \mathbb{N}$, by the Squeeze Theorem, we see that $a_{n} \rightarrow a$.
7. Let $a \in \mathbb{R}$ with $|a|<1$. Prove: $\lim _{n \rightarrow \infty} a^{n}=0$ and $\lim _{n \rightarrow \infty} n a^{n}=0$.
(Ghorpade and Limaye, Examples 2.7)
ANS: If $a=0$, this is obvious. Suppose $a \neq 0$. Write $\frac{1}{|a|}=1+h$. Then $h>0$, and by the Binomial Theorem, $\frac{1}{|a|^{n}}=(1+h)^{n}=1+n h+\frac{n(n-1)}{2} n h^{2}+\cdots+h^{n}$ for all $n \in \mathbb{N}$. Consequently, $\frac{1}{|a|^{n}}>n h$ for all $n \in \mathbb{N}$ and $\frac{1}{|a|^{n}}>\frac{n(n-1)}{2} h^{2}$ for all $n \in \mathbb{N}$ with $n>1$. Hence, $0<|a|^{n}<\frac{1}{n h}$ and $0<n|a|^{n}<\frac{2}{(n-1) h^{2}}$ for all $n \in \mathbb{N}$ with $n>1$. So, $1 / n h \rightarrow 0$ and $\frac{2}{(n-1) h^{2}} \rightarrow 0$. Therefore, by the Squeeze Theorem, $|a|^{n} \rightarrow 0$ and $n|a|^{n} \rightarrow 0$, and thus $a^{n} \rightarrow 0$ and $n a^{n} \rightarrow 0$.
8. We know: Every bounded sequence in $\mathbb{R}$ has a convergent subsequence. Prove this in two ways, namely via "bisection" and the way in Fitzpatrick. For both proofs, explicitly state (but not prove) all invoked required theorems.
ANS: For the bisection method, the key relevant theorem is the nested interval theorem:
For each $n \in \mathbb{N}$, let $a_{n}$ and $b_{n}$ be such that $a_{n}<b_{n}$, and define the sequence of closed intervals $I_{n}=\left[a_{n}, b_{n}\right]$, with $I_{n+1} \subseteq I_{n}$. Under the assumption that $\lim _{n \rightarrow \infty}\left[b_{n}-a_{n}\right]=0$, there is exactly one point $x$ that belongs to the interval $I_{n}$ for all $n$, and both sequences $\left\{a_{n}\right\},\left\{b_{n}\right\}$ converge to this point.

The most detailed presentation I have found is in Pons, Thm 2.3.7, ${ }^{3}$ and replicated here, including his graphic, copied here as Figure 2.
Let $\left(a_{n}\right)$ be a bounded sequence of real numbers. Using this hypothesis, we can find $M>0$ such that $a_{n} \in[-M, M]$ for all $n \in \mathbb{N}$. To construct a sequence of closed, bounded, nested intervals with lengths decreasing to zero, we will successively cut the interval $[-M, M]$ into halves. For $I_{1}$, we bisect $[-M, M]$ into the closed subintervals $[-M, 0]$ and $[0, M]$. The fact that sequences must have countably many terms guarantees us that at least one of these two subintervals must contain

[^2]

Figure 2: From Pons, p. 78
countably many terms of the sequence $\left(a_{n}\right)$. Choose $I_{1}$ to be one such subinterval (if both $[-M, 0]$ and $[0, M]$ contain infinitely many terms of the sequence, then the choice is completely arbitrary). Note that the length of $I_{1}$ is $M$.
To select $I_{2}$, we bisect $I_{1}$ into closed subintervals of length $M / 2=M / 2^{1}$. Again, at least one of these subintervals must contain countably many terms of the sequence $\left(a_{n}\right)$ and we choose $I_{2}$ to be one such closed interval. Notice also that $I_{2} \subseteq I_{1}$. Continuing inductively, we can construct a collection of closed, bounded, nested intervals $\left\{I_{j}\right\}$ (Fig. 2.3) with the property that each interval contains countably many terms of the sequence ( $a_{n}$ ) and the length of $I_{j}$ is $M / 2^{j-1}$. Sequence $M / 2^{j-1}$ converges to zero as $j \rightarrow \infty$. From the nested intervals theorem, it follows that the intersection $\cap_{j=1}^{\infty} I_{j}$ contains a unique element which we shall denote by $a$. The convergence of the sequence $M / 2^{j-1}$ will also be key in the final component of the proof.
At this point we are ready to define a subsequence; this is also done in an inductive manner. Choose $n_{1} \in \mathbb{N}$ so that $a_{n_{1}} \in I_{1}$. Next, choose $n_{2} \in \mathbb{N}$ so that $n_{2}>n_{1}$ and $a_{n_{2}} \in I_{2}$, which is permissible since $I_{2}$ contains countably many terms of the sequence $\left(a_{n}\right)$. Continuing, for each $j \in \mathbb{N}$, we choose $n_{j} \in \mathbb{N}$ so that $n_{j}>n_{j-1}$ and $a_{n_{j}} \in I_{j}$.
We claim now that the subsequence $\left(a_{n_{j}}\right)$ converges to $a$. Let $\epsilon>0$ and choose $J \in \mathbb{N}$ such that $M / 2^{J}<\epsilon$. Then if $j \geq J+1$, we have that $j-1 \geq J$ and $M / 2^{j-1} \leq M / 2^{J}<\epsilon$. In other words, for $j \geq J+1$, the length of the interval $I_{j}$ is less than $\varepsilon$. Finally, if we consider a term $a_{n_{j}}$ with $j \geq J+1$, then both $a_{n_{j}}$ and $a$ are in $I_{j}$ and, using the restriction on the length of this interval, it is clear that

$$
\left|a_{n_{j}}-a\right| \leq \frac{M}{2^{j-1}} \leq \frac{M}{2^{J}}<\epsilon
$$

Thus we conclude that $\left(a_{n_{j}}\right) \rightarrow a$ as desired.
For the second way, see Fitzpatrick, page 45, Theorem 2.33. It requires:
(a) every sequence has a monotone subsequence;
(b) the monotone convergence theorem: a monotone sequence converges if and only if it is bounded.

Both of these results are in Fitzpatrick. So, let $\left\{x_{n}\right\}$ be a bounded sequence in $\mathbb{R}$, and let $\left\{x_{n_{k}}\right\}$ be a monotone subsequence. This subsequence is then bounded and monotone, which means it is convergent.
9. Sequential compactness
(i) State the definition of sequential compactness. (Remember, it must be exact to get any credit.)
(ii) State and provide the full proof, as given in Fitzpatrick, of the sequential compactness theorem.
(iii) State the other very common name for this theorem.

ANS: Fitzpatrick, page 46, definition; Theorem 2.36; and the comment provided thereafter that says this is called "Bolzano-Weierstrass".
10. State the definition exactly as in Fitzpatrick of compact, where here is meant what I refer to as topologically compact.
ANS: Fitzpatrick, page 49, definition.
11. What is the flaw in the following argument? Let $S=(0,1]$, and let $S_{n}=\left(0,1+\frac{1}{n}\right)$. Then, any $S_{i}$ is a finite cover for $S$, and thus $S$ is (topologically) compact.
ANS: For compactness of $S$, we need that *any* cover of $S$ by a collection $\left\{I_{n}\right\}_{n=1}^{\infty}$ of open intervals has a finite subcover. Take $S_{n}=\left(\frac{1}{n}, 2\right)$. Clearly, $S \subseteq \cup_{n=1}^{\infty} S_{n}$. Thus, the infinite union of $S_{n}$ is an open cover of $S$, however, it will not have a finite subcover, because, for any index $N \in \mathbb{N}$, the numbers between 0 and $1 / N$ will not be covered.
12. State (not prove) the Heine Borel theorem. (Fitzpatrick, page 52).

ANS: A closed and bounded subset of $\mathbb{R}$ is (topologically) compact.
13. Fitzpatrick shows the proof that sequential compactness implies topological compactness. Your task here is to prove the converse, but the proof is short and easy, because you are allowed to use theorems from his book. Of course, you do not need to know the numbers of the theorems in his book, but you need to know, and give, the statements of the theorems, and specify them precisely.
ANS: Fitzpatrick, Prop 2.40: If $S$ is a topologically compact subset of $\mathbb{R}$, then $S$ is closed and bounded. Fitzpatrick, Prop 2.37: If $S$ is a closed and bounded subset of $\mathbb{R}$, then $S$ is sequentially compact.
14. For compact subsets $A$ and $B$ of $\mathbb{R}$, either prove, or refute (by giving a counter-example) that $A \backslash B$ is compact.
ANS: Take, for example, $A=[0,1]$ and $B=[1 / 2,1]$. Then $A \backslash B=[0,1 / 2)$, which is not closed, and therefore by Heine-Borel, is not compact.
15. For compact $A, B \subset \mathbb{R}$, either prove, or refute by giving a counter-example that:
(a) $A \cup B$ is compact.
(b) $A \cap B$ is compact.
(Fitzpatrick, page 52, \#7; among other sources)
NOTE: If you believe a result is true, and wish to prove it, then you are required to give proofs that invoke only material (definitions and theorems) so far seen in Fitzpatrick, chapters 1 and 2, in order to obtain full credit. You *may* provide a proof invoking technology we have not yet seen (but will see soon), for partial credit. You can also provide both methods of proof for extra credit.

ANS:

NOTE: Proofs for both parts (a) and (b) are trivial, because both $A \cup B$ and $A \cap B$ are closed and bounded, and thus Heine-Borel applies. Indeed, Fitzpatrick, Sec. 2.5 proves the equivalence of:
(i) a set in $\mathbb{R}$ is closed and bounded,
(ii) sequential compactness,
(iii) topological compactness (an open interval cover admits a finite subcover).

Now, what makes it interesting is that, up to this point in the book, Fitzpatrick has not defined or proven that a closed set is the complement of an open set, nor anything about finite and countable unions and intersections of open and closed sets. That limits the tools we can use! All we know is that $A$ and $B$ are compact (and again, restrict ourselves just to $\mathbb{R}$ ), and the equivalence of sequential and topological compactness.
Thus, we can envision two tasks: Prove the two results using only sequential compactness; and prove the two results using only topological compactness; and not using the aforementioned results that both $A \cup B$ and $A \cap B$ are closed and bounded, and applying Heine-Borel.
Proof of compactness of $A \cup B$ via Sequential Compactness:
Let $\left\{u_{n}\right\}$ be any sequence in $A \cup B$. There are only three, distinct, possibilities:

- Sequence $\left\{u_{n}\right\}$ eventually only contains elements of set $A$, i.e., there is an $N \in \mathbb{N}$ such that, for $n \geq N, u_{n} \in A$.
- Sequence $\left\{u_{n}\right\}$ eventually only contains elements of set $B$.
- Neither of the above cases.

In the first case, there exists a subsequence of $\left\{u_{n}\right\}$ such that all its elements are in $A$. As $A$ is compact, there is a subsequence of this subsequence that converges to a point in $A$. As $A \subset(A \cup B)$, this subsubsequence converges in $A \cup B$ as well, and thus, for this case, $A \cup B$ is sequentially compact.
The second case is similar. For the third case, there is a subsequence of $\left\{u_{n}\right\}$ such that all its elements are in $A$ (and likewise in $B$, but we only need one of these two cases). As previous, as $A$ is compact, there is a subsequence of this subsequence that converges to a point in $A$; and as $A \subset(A \cup B), A \cup B$ is compact.

Proof of compactness of $A \cap B$ via Sequential Compactness:
Let $\left\{u_{n}\right\}$ be any sequence in $A \cap B$. As $(A \cap B) \subset A,\left\{u_{n}\right\}$ is obviously a sequence in $A$. As $A$ is compact, $\left\{u_{n}\right\}$ has a subsequence $\left\{u_{n_{k}}\right\}$ that converges to a point, say $p_{A}$, in $A$. As all $u_{n_{k}} \in$ $(A \cap B) \subset B$, and $B$ is compact, there is a subsequence of this subsequence that converges to a point in $B$, say $p_{B}$. But $\left\{u_{n_{k}}\right\}$ is a convergent (sub)sequence, and thus its subsequences converge to the same limit, i.e., $p:=p_{A}=p_{B}$. As $p_{A} \in A$ and $p_{B} \in B, p \in(A \cap B)$. Thus, we have found a (sub)subsequence $A \cap B$ that converges to a point in $A \cap B$, showing that $A \cap B$ is (sequentially) compact.

Proof of compactness of $A \cup B$ via Topological Compactness: We give two proofs that differ only slightly. The first avoids the use of double subscripts, while the second results in a minimal subcover.

- Proof 1: Let $\cup_{n=1}^{\infty} I_{n}$ be any arbitrary cover of $A \cup B$ where for each $n, I_{n}$ is an open interval. Note that $A \subset A \cup B$ and $B \subset A \cup B$, therefore $\cup_{n=1}^{\infty} I_{n}$ is an open cover of $A$ and of $B$. By compactness of $A$ and $B$ there exist $N_{1}, N_{2} \in \mathbb{N}$ such that

$$
A \subset \cup_{n=1}^{N_{1}} I_{n} \text { and } B \subset \cup_{n=1}^{N_{2}} I_{n}
$$

Take the union of the finite subcovers. This union is again finite and a cover of $A \cup B$ :

$$
A \cup B \subset\left[\cup_{n=1}^{N_{1}} I_{n}\right] \cup\left[\cup_{n=1}^{N_{2}} I_{n}\right] .
$$

We have found a finite subcover for $A \cup B$ for an arbitrary cover of open intervals. Hence, $A \cup B$ is compact.

- Proof 2: Let $\cup_{n=1}^{\infty} I_{n}$ be any arbitrary cover of $A \cup B$ where for each $n, I_{n}$ is an open interval. Note that $A \subset A \cup B$ and $B \subset A \cup B$, therefore $\cup_{n=1}^{\infty} I_{n}$ is an open cover of $A$ and of $B$. By compactness of $A$, there exists a $k_{A} \in \mathbb{N}$ and a set of natural numbers $S_{A}:=\left\{n_{1}, \ldots, n_{k_{A}}\right\}$ such that $I_{n_{k}}, k=1, \ldots, k_{A}$, is a finite subcover of $A$. Likewise, as $B$ is compact, similarly define $k_{B}$ and $S_{B}:=\left\{m_{1}, \ldots, m_{k_{B}}\right\}$. Then $S:=S_{A} \cup S_{B}$ forms a finite subcover of $A \cup B$.

Proof of compactness of $A \cap B$ via Topological Compactness: We give three correct proofs, and one that is wrong, but might entice someone because it initially seem appealing.

- Proof 0: (Intentionally Erroneous)
"A finite subcover of $A \cup B$ also covers $A \cap B$, and we just proved that $A \cup B$ is compact. QED" This is wrong because we need to start with *any* countable cover of $A \cap B$. Say, for example, $A=[-1,2]$ and $B=[1,3]$, and $A \cap B=[1,2]$. The cover $\left\{I_{n}\right\}$, with $I_{n}=(0, n)$ clearly covers $A \cap B$, but it is not a cover for $A \cup B$. So, this "proof" fails.
- Proof 1: (Contributed by Christopher Heil) ${ }^{4}$

This requires knowing the following fact, which can be found in our course material, e.g., Stoll, Theorem 2.3.5(b): Every closed subset of a compact set is compact. The proof of this result is short, and so I provide it here:

Proof: (Stoll, p. 72) Let $F$ be a closed subset of the compact set $K$ and let $\left\{O_{\alpha}\right\}_{\alpha \in A}$ be an open cover of $F$. Then $\left\{O_{\alpha}\right\}_{\alpha \in A} \cup\left\{F^{c}\right\}$ is an open cover of $K$.
(This uses the facts that (i) the complement of a closed set is open; and (ii) the union of open sets is open.) Since $K$ is compact, a finite number of these will cover $K$, and hence also $F$.

Proof of main result: $A \cap B$ is a closed subset of $A$, which is compact, so $A \cap B$ is compact. In addition to the previous stated result, this also requires the fact that the intersection of closed sets is closed; see, e.g., Stoll, Theorem 2.2.10.

- Proof 2: (Contributed by Christopher Heil)

Let $\left\{I_{n}\right\}$ be any cover of $A \cap B$ by open sets. Add the set $B^{c}$ to the collection. Then $\left\{I_{n}, B^{c}\right\}$ is an open cover of $A$, so it has a finite subcover. That subcover might include $B^{c}$, but if you remove that one set, then the remaining finitely many must cover $A \cap B$.

- Proof 3: (Contributed by Ralf Blöchlinger)

This proof is inspired from https://math.stackexchange.com/questions/3481414. It requires knowing that an open set is the complement of a closed set; and that each open subset of $\mathbb{R}$ can be expressed as a countable union of open intervals. It further requires the above result that $A \cup B$ is compact.
Let $\cup_{n=1}^{\infty} I_{n}$ be any arbitrary cover of $A \cap B$ where, for each $n, I_{n}$ is an open interval. Note that $(A \cap B)^{c}=\mathbb{R} \backslash(A \cap B)$. Further note that $\cup_{n=1}^{\infty} I_{n} \cup[\mathbb{R} \backslash(A \cap B)]=\mathbb{R}$. Therefore, $\cup_{n=1}^{\infty} I_{n} \cup[\mathbb{R} \backslash(A \cap B)]$ is a union of open intervals that covers $\mathbb{R}$ and thus a valid cover of

[^3]$A \cup B$. Rewrite $[\mathbb{R} \backslash(A \cap B)]=\cup_{n=1}^{\infty} J_{n}$ where each $J_{n}$ is an open interval. As $A \cup B$ is compact, $\exists N \in \mathbb{N}$ such that
$$
A \cup B \subset\left[\cup_{n=1}^{N} I_{n}\right] \cup\left[\cup_{n=1}^{N} J_{n}\right] .
$$

Note $(A \cap B) \subset(A \cup B)$ and $[A \cap B] \cap[\mathbb{R} \backslash(A \cap B)]=\emptyset$. The finite subcover of $A \cup B$ we produced is also a finite subcover of $A \cap B$, but all intervals in $\cup_{n=1}^{N} J_{n}$ are disjoint from $A \cap B$. We can thus remove these intervals and still have a finite subcover of $A \cap B$. That is, $A \cap B \subset \cup_{n=1}^{N} I_{n}$. We have found a finite subcover for any arbitrary open cover of $A \cap B$ and thus conclude that $A \cap B$ is compact.
16. Prove: If $K$ is sequentially compact and if $H$ is a closed subset of $K$, then $H$ is sequentially compact. Use only the concept of sequential compactness and not topological compactness.
ANS: Recall:

- Fitzpatrick, p. 37, definition: A subset $S$ of $\mathbb{R}$ is said to be closed provided that, if $\left\{a_{n}\right\}$ is a sequence in $S$ that converges to a number $a$, then $a \in S$.
- Recall Fitzpatrick, p. 46, definition: A set of real numbers $S$ is said to be sequentially compact provided that every sequence $\left\{a_{n}\right\}$ in $S$ has a subsequence that converges to a point that belongs to $S$.

Let $\left\{x_{n}\right\} \subseteq H$. Then since $K$ is sequentially compact, from the above definition, there is a subsequence, $\left\{x_{n_{k}}\right\}$ that converges to a point, $x \in K$. But these $x_{n_{k}}$ are in the closed set $H$ and so $x \in H$ from the above definition of closed.
17. Let $x_{n}$ and $y_{n}$ be sequences such that $\lim _{n \rightarrow \infty} x_{n}=x$ and $\lim _{n \rightarrow \infty} y_{n}=y$. Prove: If $x_{n} \leq y_{n}$ for all $n$ sufficiently large, then $x \leq y$. (Stoll, p. 87, \#8; Petrovic, Coro 2.3.12)
ANS: $\left|x_{n}-x\right|<\epsilon$, so $-\epsilon<x_{n}-x<\epsilon$, implying $-x>-x_{n}-\epsilon$. Similarly, $\left|y_{n}-y\right|<\epsilon$, or $-\epsilon<y_{n}-y<\epsilon$, implying $y>y_{n}-\epsilon$. Adding the two inequalities gives

$$
y-x \geq\left(y_{n}-\epsilon\right)-\left(x_{n}+\epsilon\right)=\left(y_{n}-x_{n}\right)-2 \epsilon \geq-2 \epsilon .
$$

As $\epsilon$ is arbitrary, it follows that $y-x \geq 0$.
18. Let $f$ be a continuous function on (nonempty) interval $(a, b) \in \mathbb{R}$.
(Petrovic, Thm 3.6.12 and Remark 3.6.13)
(a) Let $c \in(a, b)$ such that $f(c)>0$. Prove: $\exists \delta>0$ such that $f(x)>0$ for $x \in(c-\delta, c+\delta)$.

ANS: Suppose to the contrary that no such $\delta$ exists. Then, for every $\delta>0$, there exists $x \in(c-\delta, c+\delta)$ such that $f(x) \leq 0$. If we take $\delta=1 / n$, we obtain a sequence $x_{n}$ in $(c-1 / n, c+1 / n)$ with $f\left(x_{n}\right) \leq 0$. The inequality $\left|x_{n}-c\right|<1 / n$ shows that the sequence $x_{n}$ converges to $c$, and the continuity of $f$ implies that the sequence $f\left(x_{n}\right)$ converges to $f(c)$. From exercise 17 and that $f\left(x_{n}\right) \leq 0, f(c) \leq 0$. This contradicts the assumption that $f(c)>0$.
(b) Devise an example to show that the converse of the previous question is not true.

Hint: This means you need to demonstrate a continuous function $f$ such that

$$
\forall x \in\{x: 0<|x-c|<\delta\}, \quad f(x)>0, \quad f(c) \leq 0
$$

ANS: Take $f(x)=x^{2}, c=0$.
19. Absolute value and continuous functions
(a) Let $\left\{a_{n}\right\}$ be a convergent sequence with $a_{n} \rightarrow a \in \mathbb{R}$. Prove: $\left|a_{n}\right| \rightarrow|a|$.

ANS: The reverse triangle inequality implies $0 \leq\left|\left|a_{n}\right|-|a|\right| \leq\left|a_{n}-a\right|$. This holds $\forall n \in \mathbb{N}$, so $\left|a_{n}\right| \rightarrow|a|$.
(b) Let function $f: D \rightarrow \mathbb{R}$ be continuous. Define $|f|: D \rightarrow \mathbb{R}$ by $|f|(x)=|f(x)|$. Prove that $|f|$ is also continuous. (Fitzpatrick, p. $58 \# 12$ )
ANS: Let $\left\{x_{n}\right\}$ be a sequence in $D$ such that $x_{n} \rightarrow c$. From continuity, $f\left(x_{n}\right) \rightarrow f(c)$. The previous result implies $\left|f\left(x_{n}\right)\right| \rightarrow|f(c)|$.
20. State (not prove) the Extreme Value Theorem, and the Intermediate Value Theorem.

ANS: As in Fitzpatrick, p. 60 (EVT) and p. 62 (IVT):
(a) EVT: A continuous function defined on a closed bounded interval, $f:[a, b] \rightarrow \mathbb{R}$, attains both a minimum and a maximum value.
(b) IVT: Suppose that the function $f:[a, b] \rightarrow \mathbb{R}$ is continuous. Let $c$ be such that $f(a)<c<f(b)$ or $f(b)<c<f(a)$. Then $\exists x_{0} \in(a, b)$ for which $f\left(x_{0}\right)=c$.
21. State the definitions, in terms of sequences, of continuity and uniform continuity. State the equivalent formulations of these two concepts in terms of $\epsilon-\delta$ formulations.
ANS: Fitzpatrick, p. 53 and p. 70 for continuity; p. 66 and p. 72 for uniform continuity.
For (regular) continuity:
(a) A function $f: D \rightarrow \mathbb{R}$ is said to be continuous at $x_{0} \in D$ provided that, for sequence $\left\{x_{n}\right\} \in D$ with $x_{n} \rightarrow x_{0} \in D, f\left(x_{n}\right) \rightarrow f\left(x_{0}\right)$. The function $f: D \rightarrow \mathbb{R}$ is said to be continuous provided that it is continuous at every point in $D$.
(b) A function $f: D \rightarrow \mathbb{R}$ is said to satisfy the $\epsilon-\delta$ criterion at $x_{0} \in D$ provided that, for each $\epsilon>0$, $\exists \delta>0$ such that, for $x \in D,\left|f(x)-f\left(x_{0}\right)\right|<\epsilon$ if $\left|x-x_{0}\right|<\delta$.

For uniform continuity:
(a) A function $f: D \rightarrow \mathbb{R}$ is said to be uniformly continuous provided that, for $\left\{u_{n}\right\} \in D$ and $\left\{v_{n}\right\} \in D$,

$$
\lim _{n \rightarrow \infty}\left[u_{n}-v_{n}\right]=0 \Rightarrow \lim _{n \rightarrow \infty}\left[f\left(u_{n}\right)-f\left(v_{n}\right)\right]=0
$$

(b) A function $f: D$ to $\mathbb{R}$ is said to satisfy the $\epsilon-\delta$ criterion on the domain $D$ provided that, for each positive number $\epsilon, \exists \delta>0$ such that

$$
u, v \in D,|u-v|<\delta \Rightarrow|f(u)-f(v)|<\epsilon
$$

22. Provide an example of a continuous function that is not uniformly continuous, and demonstrate the latter.

ANS: See, e.g., Fitzpatrick, p. 67, Examples 3.15 and 3.16.
23. For a function $f: D \rightarrow \mathbb{R}$, a solution of the equation: $f(x)=x, x \in D$, is called a fixed point of $f$. A fixed point corresponds to a point at which the graph of the function $f$ intersects the line $y=x$. If $f:[-1,1] \rightarrow \mathbb{R}$ is continuous, $f(-1)>-1$ and $f(1)<1$, show that $f:[-1,1] \rightarrow \mathbb{R}$ has a fixed point. (Fitzpatrick, p. 65, \#4)

ANS: As in Garling: Suppose that $[a, b]$ is a closed bounded interval and that $f:[a, b]$ to $[a, b]$ is continuous. Then there exists $c \in[a, b]$ with $f(c)=c$. Proof: If $f(a)=a$ or if $f(b)=b$, there is nothing to prove. Otherwise, let $g(x)=x-f(x)$. Then $g(a)=a-f(a)<0$ and $g(b)=b-f(b)>0$. By the intermediate value theorem, there exists $c \in[a, b]$ with $g(c)=c-f(c)=0$.
24. Show that uniformly continuous functions defined on the same domain form a vector space. That means, if $f, g: D \rightarrow \mathbb{R}$ are uniformly continuous functions, then $c f+d g$ is uniformly continuous, where $c, d \in \mathbb{R}$. (Fitzpatrick, p. 69, \#2 and \#3)
ANS: Show additivity and scalar multiplication (homogeneity) separately:
(a) Additivity: Given $\epsilon>0$ there are $\delta_{1}>0$ and $\delta_{2}>0$ such that for any $x, y \in D$, if $|x-y|<\delta_{1}$, then $|f(x)-f(y)|<\epsilon / 2$ and if $|x-y|<\delta_{2}$, then $|g(x)-g(y)|<\epsilon / 2$. Therefore, if $|x-y|<$ $\min \left\{\delta_{1}, \delta_{2}\right\}$, then, from the triangle inequality,

$$
|(f+g)(x)-(f+g)(y)| \leq|f(x)-f(y)|+|g(x)-g(y)|<\epsilon / 2+\epsilon / 2=\epsilon
$$

(b) Homogeneity: By uniform continuity of $f$, we have that, given $\epsilon>0$, there is a $\delta>0$ such that, for any $x, y \in D$, if $|x-y|<\delta$, then $|f(x)-f(y)|<\epsilon /|c|$, for $c \in \mathbb{R} \backslash\{0\}$. Therefore, if $|x-y|<\delta$,

$$
|c f(x)-c f(y)| \leq|c||f(x)-f(y)|<|c| \epsilon /|c|=\epsilon .
$$

Combining the two statements gives the result.
25. Show that it is not necessarily the case that, if $f$ and $g$ are uniformly continuous functions, then so is the product $h:=f g$. (Fitzpatrick, p. 69, \#6)

ANS: Take $f, g: \mathbb{R} \rightarrow \mathbb{R}$ with $f(x)=g(x)=x$ as a counterexample. Function $h(x)=x^{2}$ is continuous, but not uniformly continuous on this domain, as shown in Fitzpatrick, p. 67, Example 3.15 .
26. If $f: D \rightarrow D$ is uniformly continuous, show that $f(f(x))$ is also uniformly continuous.

ANS: Fix $\epsilon>0$. We need to show:

$$
\exists \delta>0:|x-y|<\delta \Rightarrow|f(f(x))-f(f(y))|<\epsilon
$$

By uniform continuity of $f$ applied to $\epsilon$,

$$
\exists \epsilon_{1}:|w-z|<\epsilon_{1} \Rightarrow|f(w)-f(z)|<\epsilon
$$

By the uniform continuity of $f$ again, applied to $\epsilon_{1}$,

$$
\exists \delta:|x-y|<\delta \Rightarrow|f(x)-f(y)|<\epsilon_{1}
$$

Combining gives $|f(f(x))-f(f(y))|<\epsilon$.
27. Suppose that the functions $f: D \rightarrow \mathbb{R}$ and $g: D \rightarrow \mathbb{R}$ are uniformly continuous and bounded. Prove that the product $f g: D \rightarrow \mathbb{R}$ also is uniformly continuous. (Fitzpatrick, p. 69, \#7)
ANS: Because $f$ and $g$ are bounded, $\exists M_{f}>0$ such that, $\forall u \in D,|f(u)|<M_{f}$; and $\exists M_{g}>0$ such that $\forall u \in D,|g(u)|<M_{g}$. Set $M:=\max \left\{M_{f}, M_{g}\right\}$. Because $f$ and $g$ are uniformly continuous on $D$, given $\epsilon>0, \exists \delta>0$ such that, if $u, v \in D$ and $|u-v|<\delta$, then

$$
\begin{equation*}
|f(u)-f(v)|<\frac{\epsilon}{2 M} \quad \text { and } \quad|g(u)-g(v)|<\frac{\epsilon}{2 M} \tag{8}
\end{equation*}
$$

As in the hint of Fitzpatrick, write

$$
f(u) g(u)-f(v) g(v)=f(u)[g(u)-g(v)]+g(v)[f(u)-f(v)] .
$$

Taking the absolute value and applying the triangle inequality, we have, for $|u-v|<\delta$,

$$
\begin{aligned}
|f(u)[g(u)-g(v)]+g(v)[f(u)-f(v)]| & \leq|f(u)[g(u)-g(v)]|+|g(v)[f(u)-f(v)]| \\
& \leq|f(u)||g(u)-g(v)|+|g(v)||f(u)-f(v)| \\
& \leq M|g(u)-g(v)|+M|f(u)-f(v)| \\
& <M \frac{\epsilon}{2 M}+M \frac{\epsilon}{2 M}=\epsilon,
\end{aligned}
$$

where we used the fact that $f$ and $g$ are bounded by $M$; and applying (8).
28. If $f: \mathbb{R} \rightarrow \mathbb{R}$ is a function such that $[f(x)]^{2}$ is continuous, is $f$ continuous?

ANS: No: Take $f(x)=1$ if $x$ is rational, and -1 otherwise.
29. Assume $f$ is a continuous function.
(a) Is $[f(x)]^{2}$ continuous?

ANS: The product of two continuous functions is continuous by Fitzpatrick Thm 3.4.
(b) Now assume further that $f$ is a uniformly continuous function. Is $[f(x)]^{2}$ uniformly continuous? ANS: Not necessarily: Take $f, g: \mathbb{R} \rightarrow \mathbb{R}$ with $f(x)=g(x)=x$ as a counterexample. Function $h(x)=x^{2}$ is clearly continuous but not uniformly continuous on this domain, as shown in Fitzpatrick, Example 3.15.
30. Let $K \subset \mathbb{R}$ be compact. Let $f: K \rightarrow \mathbb{R}$ be a continuous function on $K$. Prove that $f(K)$, the image of $f$, is compact. This is to be done in the following two ways:
(a) Show set $f(K)$ obeys sequential compactness.
(b) Show set $f(K)$ obeys topological compactness. (Fitzpatrick, Thm 11.20)

Note: For part (b), you will need the following fact (e.g., Stoll, Theorem 4.2.6):
Let $E$ be a subset of a metric space $X$ and let $f$ be a real-valued function on $E$. Then: function $f$ is continuous $\Longleftrightarrow f^{-1}(V)$ is open in $E$ for every open $V \subset \mathbb{R}$.

The two ways $(a)$ and $(b)$ are of course equivalent: Recall that, in this context of $\mathbb{R}$ (and $\mathbb{R}^{n}$ ) that (Fitzpatrick, Thm 2.42) Sequentially compact $\Longleftrightarrow$ topologically compact.
ANS:
(a) First recall for convenience the following definitions from Fitzpatrick:

- Fitzpatrick, p. 46: A set of real numbers $S$ is said to be sequentially compact provided that every sequence $\left\{a_{n}\right\}$ in $S$ has a subsequence that converges to a point that belongs to $S$.
- Fitzpatrick, p. 53: A function $f: D \rightarrow \mathbb{R}$ is said to be continuous at the point $x_{0}$ in $D$ provided that whenever $\left\{x_{n}\right\}$ is a sequence in $D$ that converges to $x_{0}$, the image sequence $\left\{f\left(x_{n}\right)\right\}$ converges to $f\left(x_{0}\right)$.
We need to show that a sequence $\left\{y_{n}\right\}$ in $f(K)$ has a convergent subsequence whose limit is in $f(K)$. With $y_{n} \in f(K)$ there is (at least one) $x_{n} \in K$ such that $y_{n}=f\left(x_{n}\right)$. This determines sequence $\left\{x_{n}\right\}$. $K$ compact implies $x_{n}$ has a convergent subsequence $\left\{x_{n_{k}}\right\}$ whose limit $x \in K$. As $\left\{x_{n_{k}}\right\} \rightarrow x$ and $f$ is continuous, $y_{n_{k}}=f\left(x_{n_{k}}\right) \rightarrow f(x)$. As $x \in K, f(x) \in f(K)$, so that $\left\{y_{n}\right\}$ has a convergent subsequence $\left\{y_{n_{k}}\right\}$ whose limit $f(x)$ is in $K$.
(b) Suppose that $K \subseteq \mathbb{R}$ is compact, and let $\left\{V_{i}\right\}, i \in J$, be any open cover of $f(K)$. Since $f$ is continuous, each inverse image $U_{i}=f^{-1}\left(V_{i}\right)$ is open. Consequently, $\left\{U_{i}\right\}, i \in J$, is an open cover of $K$. As $K$ is compact, there must exist some finite subcover of $K$, say $\left\{U_{i_{1}}, \ldots, U_{i_{N}}\right\}$. But then $\left\{V_{i_{1}}, \ldots, V_{i_{N}}\right\}$ is a finite cover of $f(K)$, so $f(K)$ is compact.

31. Prove that a closed subset of a compact set is compact.

ANS: See, e.g., Stoll, Theorem 2.3.5.
32. Let $f: D \rightarrow \mathbb{R}, f(x)=\sqrt{x}$, with $D=[0,1]$. A function $f: D \rightarrow \mathbb{R}$ is said to be Lipschitz, provided $\exists C \in \mathbb{R}_{>0}$ such that $\forall u, v \in D,|f(u)-f(v)| \leq C|u-v|$. (Fitzpatrick, p. 74, \#7; Stoll, p. 161, \#5)
(a) Prove that $f$ is continuous.
(b) Prove that $f$ is uniformly continuous.
(c) Prove that $f$ is not Lipschitz.
(d) Determine whether or not $f:[1,+\infty) \rightarrow \mathbb{R}, f(x)=\sqrt{x}$, is uniformly continuous. Hint:

$$
|\sqrt{x}-\sqrt{y}|=|\sqrt{x}-\sqrt{y}| \frac{\sqrt{x}+\sqrt{y}}{\sqrt{x}+\sqrt{y}}
$$

ANS: In part (a), note that $|x-p|<\delta \Leftrightarrow p-\delta<x<p+\delta$, and $x$ is restricted to $D=[0,1]$. This is why $\delta$ is taken to be $\delta=\min \{p, \sqrt{p} \epsilon\}$.
(a) (Stoll, p. 156, \#4) If $p>0,|f(x)-f(p)|=|\sqrt{x}-\sqrt{p}|=|x-p| /(\sqrt{x}+\sqrt{p})<\frac{1}{\sqrt{p}}|x-p|$. Let $\epsilon>0$ be given. Set $\delta=\min \{p, \sqrt{p} \epsilon\}$. Then $|x-p|<\delta$ implies that $|f(x)-f(p)|<\epsilon$. Therefore $f$ is continuous at $p$.
If $p=0,|f(x)-f(p)|=|\sqrt{x}-\sqrt{p}|=|x-p| /(\sqrt{x}+\sqrt{p})$ set $\delta=\epsilon^{2}$. Alternately use Stoll, Thm 4.1.3 and Stoll, Exercise 6 of Section 3.1.
(b) Fitzpatrick, Thm 3.17: A continuous function with domain a closed, bound interval is uniformly continuous.
(c) One argument using the derivative is: Function $f$ is Lipschitz if $|f(x)-f(y)| \leq C|x-y|$, which implies $\left|\frac{f(x)-f(y)}{x-y}\right| \leq C$, . The left-hand side term is a difference quotient (or a growth of rate) of a function, or graphically the slope of the line joining $(x, f(x))$ and $(y, f(y))$. Thus $f$ is Lipschitz if all the secant lines are of bounded slope.
A second argument not invoking derivatives is to consider sequence $x_{n}=1 / n$ for $n \in \mathbb{N}$. Observe

$$
\frac{\sqrt{1 / n}-\sqrt{0}}{1 / n-0}=\frac{1 / \sqrt{n}}{1 / n}=\sqrt{n}
$$

This ratio can be made arbitrarily large as $n \rightarrow \infty$. Therefore, the square-root function fails to be Lipschitz.
(d) The function is uniformly continuous on $[1,+\infty)$. Let $\epsilon>0$, and take $\delta=\epsilon$. If $|x-y|<\delta$, then

$$
|\sqrt{x}-\sqrt{y}|=|\sqrt{x}-\sqrt{y}| \frac{\sqrt{x}+\sqrt{y}}{\sqrt{x}+\sqrt{y}}=\frac{|x-y|}{\sqrt{x}+\sqrt{y}} \leq \frac{|x-y|}{2}<|x-y|<\delta=\epsilon
$$

33. In Fitzpatrick, Sec. 3.6, there is a theorem that does not assume continuity of function $f$, but rather concludes it.
(a) State this theorem (not the proof).
(b) Provide an example analytically and also a sketch of your function such that when you draw the function, it cannot be drawn "without removing your pen from the page".
(c) Investigate at least one alternative proof of this result from other analysis books.

ANS:
(a) Fitzpatrick, Thm 3.23: Suppose that the function $f: D \rightarrow \mathbb{R}$ is monotone, and its image $f(D)$ is an interval. Then $f$ is continuous.
(b) Example: let $D=[0,1] \cup(2,3], f(x)=x$ for $0 \leq x \leq 1$, and $f(x)=x-1$, for $2<x \leq 3$. Then $f$ is (strictly) monotone increasing, and $f(D)=[0,2]$.
(c) (Loya, Theorem 4.27) A monotone function on $\mathbb{R}$ is continuous on $\mathbb{R}$ if and only if its range is an interval.
NOTE: $f(c-)$ and $f(c+)$ refer to left and right function limits; see exercise 22.
Proof: By the intermediate value theorem, we already know that the range of every (in particular, a monotone) continuous function on $\mathbb{R}$ is an interval. Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be monotone and suppose, for concreteness, that $f$ is nondecreasing, since the case for a nonincreasing function is similar. It remains to prove that if the range of $f$ is an interval, then $f$ is continuous. We shall prove the contrapositive, so assume that $f$ is not continuous at some point $c$. Then one of the equalities in

$$
f(c-)=f(c)=f(c+)
$$

must fail. Since $f$ is nondecreasing, we have $f(c-) \leq f(c) \leq f(c+)$, and therefore, one of the intervals $(f(c-), f(c)),(f(c), f(c+))$ is nonempty. Whichever interval is nonempty is not contained in the range of $f$. By [Lemma 4.22; see below], the range of $f$ cannot be an interval.

Loya, Lemma 4.22: A set $A \subset \mathbb{R}$ is an interval $\Leftrightarrow$ given points $a<b$ in $A$, we have $[a, b] \subseteq A$. That is, $A$ is an interval $\Leftrightarrow$ given points $a, b$ in $A$ with $a<b$, all points between $a$ and $b$ also lie in $A$.
34. In Fitzpatrick Sec. 3.6, there is a theorem that concludes that the inverse function $f^{-1}$ is continuous.
(a) State and prove it, and locate the same theorem in Stoll.
(b) Provide an example analytically and also a sketch of your function such that $f$ has a jump discontinuity. (A jump discontinuity is shown in Stoll, Fig. 4.9.) Plot also $f^{-1}(y)$.
(c) (Stoll, p. 176, \#17) Let $E=[0,1] \cup[2,3)$, and for $x \in E$ set

$$
f(x)=\left\{\begin{array}{cl}
x^{2}, & 0 \leq x \leq 1 \\
4-x, & 2 \leq x<3
\end{array}\right.
$$

Notice that, as subsequently asked, the image of $f$ is an interval, and $f$ is continuous. However, it is not monotone, and indeed, $f^{-1}$ is not continuous on all of the image $f(E)$.
i. Sketch the graph of $f$.
ii. Show that $f$ is one-to-one and continuous on $E$.
iii. Show that $f(E)=[0,2]$.
iv. Find $f^{-1}(y)$ for $y \in[0,2]$, and show that $f^{-1}$ is not continuous at $y_{o}=1$.

ANS:
(a) - Fitzpatrick, Thm 3.29: Let $I$ be an interval and suppose that the function $f: I \rightarrow \mathbb{R}$ is strictly monotone. Then the inverse function $f^{-1}: f(I) \rightarrow \mathbb{R}$ is continuous.

- Stoll, Thm 4.4.12: Let $I \subset \mathbb{R}$ be an interval and let $f: I \rightarrow \mathbb{R}$ be strictly monotone and continuous on $I$. Then $f^{-1}$ is strictly monotone and continuous on $D=f(I)$.
- Proof: Function $f^{-1}: D \rightarrow \mathbb{R}$ is monotone with interval $I$ as its image. From Theorem [Fitz 3.23; see exercise 33] $f^{-1}: D \rightarrow \mathbb{R}$ is continuous.
(b) Example is $f:[0,2] \rightarrow \mathbb{R}$, with $f(x)=x^{2}$, for $0 \leq x \leq 1$; and $f(x)=1+x^{2}$, for $1<x \leq 2$. See Figure 3.


Figure 3: $I=[0,2] ; f: I \rightarrow \mathbb{R}$ as given in the text, with $D=f(I)=[0,1] \cup(2,5]$
(c) The question from Stoll, p. 176, \#17, is straightforward and illustrates the necessity of the conditions in the theorem.
35. Suppose that the function $f:[0,1] \rightarrow \mathbb{R}$ is continuous, $f(0)>0$, and $f(1)=0$. Prove that $\exists x_{0} \in(0,1]$ such that $f\left(x_{0}\right)=0$ and $f(x)>0$ for $0 \leq x<x_{0}$; that is, there is a smallest point in the interval $[0,1]$ at which the function $f$ attains zero or negative values.
(Fitzpatrick, Sec. 3.2, \#7)
Note: This is easy to graphically motivate using our intuition of continuity. Note that the IVT, in subsequent Sec. 3.3 of Fitzpatrick, tells us that, if, under the conditions stated here, namely $f$ continuous, $f(0)>0, f(1)=0$, and assuming $f$ takes on negative values, then there must be an $x_{0} \in(0,1)$ with $f\left(x_{0}\right)=0$, but this does not yield what we want, namely, that $\exists x_{0} \in(0,1]$ such that $f\left(x_{0}\right)=0$ and $f(x)>0$ for $0 \leq x<x_{0}$; that is, there is a smallest point in the interval $[0,1]$ at which the function $f$ attains zero.
ANS: Several proofs are given.
(a) Marc Proof 1: For this short proof, we need to invoke two well-known results not presented in Fitzpatrick up to his Sec. 3.
i. Let $f: E \subset \mathbb{R}^{n} \rightarrow \mathbb{R}$. The function $f$ is continuous on $E$ if and only if the inverse image $f^{-1}(W)$ of every closed set $W$ is closed relative to $E$. If the domain $E$ is a closed set in
$\mathbb{R}^{n}$, then $f$ is continuous on $E$ if and only if the inverse image $f^{-1}(W)$ of every closed set $W$ is closed. See, e.g., Terrell, A Passage to Modern Analysis, Theorem 8.10.10 page 261; and Tao, Analysis II, 4 th ed., Theorem 2.1.5(d).
ii. Let $K$ be a compact subset of $\mathbb{R}$, and let $f: K \rightarrow \mathbb{R}$ be a continuous function on $K$. The image $f(K)$ is compact. (See below Addendum for proof.)
We also require this result:
iii. A compact set $K \subset \mathbb{R}$ contains its minimum (and maximum).

Proof (Marc): Let $g: K \rightarrow K, g(x)=x$. Then, the Extreme Value Theorem (Fitzpatrick, Sec. 3.2, Theorem 3.9, and thus known to the students at this point) implies $g(x)=x$ obtains its minimum (and maximum).
Note: Other proofs, besides the one I give above and believe correct, can be found on
https://math.stackexchange.com/questions/
553905/compact-set-always-contains-its-supremum-and-infimum
Proof of the main result: Let $D=[0,1]$, which is compact (Heine-Borel), and let $f: D \rightarrow \mathbb{R}$ be continuous, with $f(0)>0$, and $f(1)=0$. Then, from (ii) above, image $f(D)$ is compact. The set $Z=\{0\} \subset f(D)$ is closed, so, from (i) above, $f^{-1}(Z)$ is closed. $f^{-1}(Z)$ is obviously bounded, as $f^{-1}(Z) \subset D$. So, $f^{-1}(Z)$ is compact, and, from (iii) above, contains its minimum. This is the desired $x_{0}$.
Addendum: I provides here a proof of the second result stated above.
Let $K$ be a compact subset of $\mathbb{R}$, and let $f: K \rightarrow \mathbb{R}$ be a continuous function on $K$. Prove that $f(K)$, the image of $f$, is compact. This is to be done in the following two ways:
i. Show set $f(K)$ obeys sequential compactness.
ii. Show set $f(K)$ obeys topological compactness.

For part (ii), you will need the following fact (that will be proven later, e.g., Stoll, Theorem 4.2.6):

Let $E$ be a subset of a metric space $X$ and let $f$ be a real-valued function on $E$. Then: function $f$ is continuous $\Longleftrightarrow f^{-1}(V)$ is open in $E$ for every open subset $V$ of $\mathbb{R}$.
The two ways $(a)$ and $(b)$ are of course equivalent. See, e.g., Fitzpatrick, p. 51, Thm 2.42: sequentially compact $\Longleftrightarrow$ topologically compact; and Fitzpatrick, Thm 11.20.
i. First recall for convenience the following.

Fitzpatrick, p. 46, definition: A set of real numbers $S$ is said to be sequentially compact provided that every sequence $\left\{a_{n}\right\}$ in $S$ has a subsequence that converges to a point that belongs to $S$.
Fitzpatrick, p. 53, definition: A function $f: D \rightarrow \mathbb{R}$ is said to be continuous at the point $x_{0}$ in $D$ provided that whenever $\left\{x_{n}\right\}$ is a sequence in $D$ that converges to $x_{0}$, the image sequence $\left\{f\left(x_{n}\right)\right\}$ converges to $f\left(x_{0}\right)$.
For sequential compactness, we need to show that a sequence $\left\{y_{n}\right\}$ in $f(K)$ has a convergent subsequence whose limit is in $f(K)$. With $y_{n} \in f(K)$ there is (at least one) $x_{n} \in K$ such that $y_{n}=f\left(x_{n}\right)$. This determines sequence $\left\{x_{n}\right\}$. $K$ compact implies $x_{n}$ has a convergent subsequence $\left\{x_{n_{k}}\right\}$ whose limit $x \in K$. Since $\left\{x_{n_{k}}\right\} \rightarrow x$ and $f$ is continuous, $y_{n_{k}}=f\left(x_{n_{k}}\right) \rightarrow f(x)$. As $x \in K, f(x) \in f(K)$, so that $\left\{y_{n}\right\}$ has a convergent subsequence $\left\{y_{n_{k}}\right\}$ whose limit $f(x)$ is in $K$.
ii. Suppose that $K \subseteq \mathbb{R}$ is compact, and let $\left\{V_{i}\right\}, i \in J$, be any open cover of $f(K)$. Since $f$ is continuous, each inverse image $U_{i}=f^{-1}\left(V_{i}\right)$ is open. Consequently, $\left\{U_{i}\right\}, i \in J$, is an open cover of $K$. As $K$ is compact, there must exist some finite subcover of $K$, say $\left\{U_{i_{1}}, \ldots, U_{i_{N}}\right\}$. But then $\left\{V_{i_{1}}, \ldots, V_{i_{N}}\right\}$ is a finite cover of $f(K)$, so $f(K)$ is compact.
(b) Ralf Proof: Inspired by: https://math.stackexchange.com/q/1506975

Since $f(1)=0$, there is at least one point $x_{0} \in(0,1]$ such that $f(x)=0$. Potentially there are multiple roots of $f(x)$ in the interval. We want to show that there is a smallest root $x_{0}^{*}$.
Let $A=\{x \in(0,1]: f(x)=0\}$. Note that $A$ is a nonempty subset of $\mathbb{R}$ that is bounded below by 0 . Thus $\inf A$ exists by completeness.
To show that there is a smallest $x$ such that $f(x)=0$, we need to show that $A$ is closed, because then, $\inf A \in A$. If $A$ is finite, it automatically has a smallest element. ${ }^{5}$
Therefore, let $A$ be infinite and let $\left\{a_{n}\right\}$ be a sequence in $A$ that converges to a limit $a .{ }^{6}$ As $f$ is continuous, $\lim _{n \rightarrow \infty} f\left(a_{n}\right)=f(a)$ and $f(a)=0$ because $f\left(a_{n}\right)=0, \forall n$. Observe $f(a)=0$, by construction of the set $A, a \in A$. Thus $A$ is closed and $\inf A=\min A=x_{0}^{*}$.
(c) Proof by author Patrick Fitzpatrick:

Let $I$ be the set of points in $[a, b]$ at which $f$ is less than or equal to 0 . By assumption, 1 belongs to $I$. By continuity, $I$ is closed. By completeness, $I$ has an infimum; denote it by $x_{0}$. Since $f(0)>0$, by continuity, there is a neighborhood of 0 on which $f$ is positive. Therefore, $x_{0}>0$. Since $x_{0}$ is a lower bound for $I, f$ is positive on $\left[0, x_{0}\right)$. Since $x_{0}$ is the greatest lower bound, by continuity, $f\left(x_{0}\right)=0$.
One can fill in a few gaps (say, $f\left(x_{0}\right)=0$ ) by inserting a sequential argument. Also, an epsilon-delta argument can be inserted to show $f$ is positive on $s$ neighborhood of 0 .

## (d) Marc Proof 2

Note: This "proof" is neither short-and-elegant, nor guaranteed to be valid. It is included for instructional purposes. Books only contain perfect, optimized proofs, and correctly so. It is however sometimes useful to see "thinking in progress", even if faulty. I would be grateful if someone can build on this and ensure correctness.
Thoughts from Christopher Heil: I took a quick look at your second solution. I'm hesitant about these types of arguments. You argue that you can take a step leftward, but do you have control on the size of the step? If the stepsize is shrinking, then you could take infinitely many steps and only move epsilon in distance. I don't think I'm ready to sign off on this argument :) Proof: We will require this well-known result: Let $f: D \rightarrow \mathbb{R}$, with $D \subset \mathbb{R}$, be a function that is continuous at some $a \in D$, and let $f(a)>0$. If $a$ is a limit point of $D$, then there exists $\delta>0$ such that, for every $x \in N_{\delta}(a) \cap D, f(x)>0$. So, we know from this that, if $x_{0}$ exists, then $x_{0}>0$, and in fact $x_{0} \geq \delta$, and so is bounded away from zero.
Let $D=[0,1]$. We are told $f(1)=0$, so $\{x: f(x)=0\} \neq \emptyset$. If $f>0$ on $D \backslash\{1\}=[0,1)$, then $x_{0}=1$. Now assume image $f(D \backslash\{1\})$ contains at least one non-positive number. Choose a value, defined as both $x_{N, 1}$ and $x_{N}$ (the $N$ for "negative or zero"), in $D \backslash\{1\}$ s.t. $f\left(x_{N}\right) \leq 0$. If $f\left(x_{N}\right)=0$, then we are in the same situation as initially, and can work with $\left.f\right|_{D_{1}}$, i.e., the restriction of $f$ to domain $D_{1}:=\left[0, x_{N}\right]$, with, as before, $f(0)>0$ and $f\left(x_{N}\right)=0$. In this case, if $f\left(D_{1} \backslash\left\{x_{N}\right\}\right)>0$, we are finished, with $x_{0}=x_{N}$. Otherwise, $f\left(D_{1} \backslash\left\{x_{N}\right\}\right)$ contains at least one non-positive number. Choose $x_{N, 2} \in\left[\delta, x_{N}\right)$ s.t. $f\left(x_{N, 2}\right) \leq 0$. We claim, as presented below, that we can repeat this, moving leftward, a finite number, say $n$, times, constructing sets $D_{1}, \ldots, D_{n}$ and points $x_{N, 1}, \ldots, x_{N, n}$, such that $f\left(D_{n} \backslash\left\{x_{N, n}\right\}\right)>0$. This implies the desired $x_{0}=x_{N, n}$.
Assuming the claim, it still remains to show that $f\left(x_{N, n}\right)=0$. Since in this scheme, continuity was not so far invoked, it seems it would be needed to draw this conclusion. That the iteration stopped implies $f\left(D_{n} \backslash\left\{x_{N, n}\right\}\right)>0$, so $f\left(x_{N, n}\right)$ cannot be negative, because of sequential

[^4]continuity from the left of $x_{0}$. It cannot be positive either, because if it were, then there exists a $\delta_{1}>0$ such that $f(x)>0$ for every $x \in N_{\delta_{1}}\left(x_{N, n}\right)$, and by construction, $f\left(x_{N, n}\right) \leq 0$. So, $f\left(x_{N, n}\right)=0$.

## Proof of claim that $n$ is finite and $x_{0}=x_{N, n}$

We know $f(0)>0$, and if $x_{0}$ exists, then $x_{0} \geq \delta>0$. So, we can rule out the case that $f$ has infinite variation at zero. Further assume $f$ is a function of bounded variation, which (I believe) implies it can cross the x-axis only a finite number of times. First consider a specification of $f$ such that the above proof is easily seen to be valid: Assume $f: D \rightarrow \mathbb{R}$ is continuous, $f(0)>0$, $f(1)=0, f$ stays positive for $x \in\left[0, x_{0}^{*}\right)$ for some $x_{0}^{*} \in[\delta, 1)$, and $f(x)=0$ for $x \in\left[x_{0}^{*}, 1\right]$. Then the desired $x_{0}=x_{0}^{*}$. In this case, the part of the above proof that says: " $f\left(D_{1} \backslash\left\{x_{N}\right\}\right)$ contains at least one non-positive number" applies, assuming we picked a value $1>x_{N}>x_{0}$, which we now show, can be guaranteed. Recall in this case, $f(x)=0$ for $x \in\left[x_{0}, 1\right]$. We can ascertain if $x_{0}<x_{N}$ because, if $x_{0}=x_{N}$, then $f\left(x_{N}-\epsilon\right)>0$ for all $0<\epsilon \leq x_{N}$. Now, if indeed $x_{0}<x_{N}$, then there exists $\delta_{1}>0$ such that, for $x \in B_{\delta_{1}}\left(x_{N, 1}\right) \cap D, f(x)=0$. This implies that we can choose $x_{N, 2} \in\left[\delta, x_{N, 1}\right)$ as $x_{N, 2}=x_{N, 1}-\delta_{1}$.
Notice $x_{N, 2}$ is not in the (open) ball $B_{\delta_{1}}\left(x_{N, 1}\right) \cap D$, but rather in its closure, so that $x_{N, 2}$ could equal $x_{0}$. This is repeated at most a finite number of iterations, namely $n$, until $f\left(D_{n} \backslash\left\{x_{N, n}\right\}\right)>$ 0 ; and $x_{N, n}=x_{0}$.
Now imagine something like $f$ being an appropriately scaled, damped cosine function, that crosses the x -axis a finite number of times, and such that, as given in the exercise, $f:[0,1] \rightarrow \mathbb{R}$ is continuous, $f(0)>0$, and $f(1)=0$. The proof starts as before determining $x_{N, 1}$, and assume " $f\left(D_{1} \backslash\left\{x_{N, 1}\right\}\right)$ contains at least one non-positive number". From continuity of $f$, only one of three things can occur, as follows. Let $x_{N}=x_{N, 1}$ for convenience.
i. From continuity of $f$, there exists $\delta_{1}>0$ such that, for $x \in B_{\delta_{1}}\left(x_{N}\right) \cap D, f(x)=0$. This case was addressed above, and we can "move to the left" by a positive increment $\delta_{1}$ to choose $x_{N, 2} \in\left[\delta, x_{N}\right)$ such that $f\left(x_{N, 2}\right) \leq 0$. (Or, there is no such value, and the algorithm terminates, having found $x_{0}$, but we ruled this out because we have assumed " $f\left(D_{1} \backslash\left\{x_{N}\right\}\right)$ contains at least one non-positive number".)
ii. From continuity of $f$, there exists $\delta_{1}>0$ such that, for $x \in B_{\delta_{1}}\left(x_{N}\right) \cap\left(\delta, x_{N}\right), f(x)<0$. In this case, we can again "move to the left" by a positive increment $\delta_{1}$.
iii. Same as previous, but $f(x)>0$, in which case, we "move to the left" by an amount larger than $\delta_{1}$, because (either the algorithm terminates here, or) we choose $x_{N, 2} \in\left[\delta, x_{N}\right)$ s.t. $f\left(x_{N, 2}\right) \leq 0$.

In all cases, we "move to the left" by a positive amount, ensuring a finite number, $n$, of steps.

## 3 Stoll, Chapter 3

Today I will do what others won't, so tomorrow I can do what others can't.
Spirits fly on dangerous missions, imaginations on fire.
Focused high on soaring ambitions, consumed in a single desire.
In the grip of a nameless possession - a slave to the drive of obsession.
A spirit with a vision, is a dream with a mission
Rush, Mission

1. Definitions and basics from Stoll:
(a) Define a metric, and a metric space.

ANS: Let $X$ be a nonempty set. A real valued function $d$ defined on $X \times X$ satisfying
i. $d(x, y) \geq 0$ for all $x, y \in X$,
ii. $d(x, y)=0 \Longleftrightarrow x=y$,
iii. $d(x, y)=d(y, x)$,
iv. $d(x, y) \leq d(x, z)+d(z, y)$ for all $x, y, z \in X$,
is called a metric on $X$. The set $X$ with metric $d$ is called a metric space, and is denoted by $(X, d)$.
(b) Let $(X, d)$ be a metric space. Define an $\epsilon$-neighborhood and an open ball.

ANS: Stoll defines: Let $(X, d)$ be a metric space and let $p \in X$. For $\epsilon>0$, an $\epsilon$-neighborhood of $p$ is the set $N_{\epsilon}(p)=\{x \in X: d(p, x)<\epsilon\}$. Other books use the open ball notation, e.g., Heil, Metrics, p. 53, who defines: Let $X$ be a metric space. Given a point $x \in X$ and given a positive number $r>0$, the open ball in $X$ with radius $r$ centered at $x$ is $B_{r}(x)=\{y \in X: d(x, y)<r\}$. (He and other authors define an open neighborhood to be an open set that contains point $x$. Notice the difference.)
(c) Define interior point and interior of a set.

ANS: Let $E$ be a subset of $X$. A point $p \in E$ is called an interior point of $E$ if there exists an $\epsilon>0$ such that $N_{\epsilon}(p) \subset E$. The set of interior points of $E$ is denoted by $\operatorname{Int}(E)$, or $E^{\circ}$, and is called the interior of $E$.
(d) Define open and closed sets, using the definition in Stoll.

ANS: A subset $O$ of $\mathbb{R}$ is open if every point of $O$ is an interior point of $O$. A subset $F$ of $\mathbb{R}$ is closed if $F^{c}=\mathbb{R} \backslash F$ is open.
(e) Let $E$ be a subset of a metric space $X$. Define a limit point, and isolated point, of $E$. What two other names are also used, by other authors, for limit points?
ANS: From Stoll, Def 2.2.12: Let $E$ be a subset of a metric space $X$. A point $p \in X$ is a limit point of $E$ if every $\epsilon$-neighborhood $N_{\epsilon}(p)$ of $p$ contains a point $q \in E$ with $q \neq p$. A point $p \in E$ that is not a limit point of $E$ is called an isolated point of $E$.
Limit points are also called accumulation points (e.g., in the books by Heil, Junghenn, Jacob and Evans, etc.), or cluster points (e.g., Terrell). For example, from Jacob and Evans (A Course in Analysis, Vol. II, p. 23), point $y \in E$ is an accumulation point of the sequence $\left\{x_{k}\right\}_{k \in \mathbb{N}}, x_{k} \in E$, if a subsequence $\left\{x_{k_{m}}\right\}_{m \in \mathbb{N}}$ of $\left\{x_{k}\right\}_{k \in \mathbb{N}}$ converges to $y$.
(f) Define the derived set of a metric space.

ANS: If $E$ is a subset of a metric space $X$, the derived set, denoted $E^{\prime}$, is the set of all limit points of $E$.
(g) State the definition of dense for metric spaces, given in Stoll, and state a definition of dense for metric spaces in terms of sequences.
ANS: Stoll, p. 65, Def 2.2.19: For metric space $X, D \subset X$ is dense in $X$ if the closure of $D$, $\bar{D}=X$. Sequential definition: Let $(X, d)$ be a metric space, with $A \subset X$. Set $A$ is dense in $X$ if, for each $x \in X$ there is a sequence in $A$ converging to $x$.
(h) Prove: If $O \subset \mathbb{R}$ is open, then $O^{c}$ is closed.

ANS: Stoll, Def 2.2.5: For $F \subset \mathbb{R}$ and $F^{c}=\mathbb{R} \backslash F,\{F$ closed $\} \Leftrightarrow\left\{F^{c}\right.$ open $\}$. So, trivially, with $O=F^{c}$ open, $O^{c}=F$ is closed.
(i) True or False: If set $A$ is not open, its complement is not closed. Prove your answer. ANS: True. $\{A$ open $\} \Leftrightarrow\left\{A^{c}\right.$ closed $\} \Longleftrightarrow \sim\{A$ open $\} \Leftrightarrow \sim\left\{A^{c}\right.$ closed $\}$.
2. (Definitions of limit point) Show the equivalence of:

- Fitzpatrick, p 81, Definition: For a set $E \subset \mathbb{R}$, the number $p$ is called a limit point of $E$ provided that there is a sequence of points in $E \backslash\{p\}$ that converges to $p$.
- Stoll, Def 2.2.12: Let $E$ be a subset of a metric space $X$. A point $p \in X$ is a limit point of $E$ if every $\epsilon$-neighborhood $N_{\epsilon}(p)$ of $p$ contains a point $q \in E$ with $q \neq p$.

ANS: Equivalence means proving both directions.
$(\Rightarrow)$ Let $p$ be a limit point of $E \subset \mathbb{R}$. Then $\exists\left\{q_{n}\right\}$ such that, $\forall n \in \mathbb{N}, q_{n} \in E \backslash\{p\}$ such that $d\left(q_{n}, p\right)<1 / n$. Fix $\epsilon>0$. Then, for $n>1 / \epsilon, q_{n} \in N_{\epsilon}(p)$.
$(\Leftarrow)$ Let $p \in X$ be a limit point of $E$. For $n \in \mathbb{N}, \exists q_{n} \in\left(N_{1 / n}(p) \cap E\right)$ with $q_{n} \neq p$. Then $\left\{q_{n}\right\}$ is a sequence in $E \backslash\{p\}$ such that, $\forall n \in \mathbb{N}, d\left(q_{n}, p\right)<1 / n$. Thus, $q_{n} \rightarrow p$.
3. Metrics $d_{1}, d_{\infty}$, and $d_{2}$

Let $\mathbf{x}=\left(x_{1}, \ldots, x_{n}\right) \in \mathbb{R}^{n}$ and $\mathbf{y}=\left(y_{1}, \ldots, y_{n}\right) \in \mathbb{R}^{n}$.
(a) State the definition of metric $d_{1}$ for vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n}$. Prove it is indeed a metric.

ANS: The "taxicab" or "Manhattan" metric $d_{1}$ is defined right below. To see $d_{1}$ is a metric, we need to verify the list in exercise 1a, the first three of which are all obvious in this case, while for the fourth, namely the triangle inequality, with $\mathbf{r}=\left(r_{1}, \ldots, r_{n}\right) \in \mathbb{R}^{n}$,

$$
\begin{aligned}
d_{1}(\mathbf{x}, \mathbf{y}) & =\sum_{i=1}^{n}\left|x_{i}-y_{i}\right|=\sum_{i=1}^{n}\left|x_{i}-r_{i}+r_{i}-y_{i}\right| \\
& \leq \sum_{i=1}^{n}\left|x_{i}-r_{i}\right|+\sum_{i=1}^{n}\left|r_{i}-y_{i}\right|=d_{1}(\mathbf{x}, \mathbf{r})+d_{1}(\mathbf{r}, \mathbf{y}) .
\end{aligned}
$$

(b) Same, for $d_{\infty}$.

ANS: The first three parts of the definition of metric are easy to verify for the sup-metric $d_{\infty}$. For the triangle inequality with $d_{\infty}$, note that, for positive $a_{i}$ and $b_{i}$,

$$
\max \left\{\left(a_{1}+b_{1}\right),\left(a_{2}+b_{2}\right)\right\} \leq \max \left\{a_{1}, a_{2}\right\}+\max \left\{b_{1}, b_{2}\right\}
$$

Thus,

$$
\begin{aligned}
d_{\infty}(\mathbf{x}, \mathbf{y}) & =\max \left\{\left|x_{i}-y_{i}\right|\right\}=\max \left\{\left|x_{i}-r_{i}+r_{i}-y_{i}\right|\right\} \\
& \leq \max \left\{\left|x_{i}-r_{i}\right|+\left|r_{i}-y_{i}\right|\right\} \\
& \leq \max \left\{\left|x_{i}-r_{i}\right|\right\}+\max \left\{\left|r_{i}-y_{i}\right|\right\}=d_{\infty}(\mathbf{x}, \mathbf{r})+d_{\infty}(\mathbf{r}, \mathbf{y})
\end{aligned}
$$

(c) Same, for $d_{2}$; except that you are not expected to prove the triangle inequality, but rather just state the names of the famous inequalities that are required for their proof, and ideally, state these inequalities.
ANS: Euclidean metric $d_{2}$ is given by $d(\mathbf{x}, \mathbf{y})=\left(\sum_{i=1}^{n}\left|x_{i}-y_{i}\right|^{2}\right)^{1 / 2}$. The first three metric space definitions are easy to verify for $d_{2}$. It satisfies the triangle inequality for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n}$, which is a special case of Minkowski's inequality. The proof of triangle requires what is referred to as the Cauchy-Schwarz, or Cauchy-Bunyakovski-Schwarz (CBS) inequality, which in turn is a special case of Hölder's inequality. CBS is given by

$$
\sum_{i=1}^{n}\left|x_{i}\right|\left|y_{i}\right| \leq\left(\sum_{i=1}^{n}\left|x_{i}\right|^{2}\right)^{1 / 2}\left(\sum_{i=1}^{n}\left|y_{i}\right|^{2}\right)^{1 / 2}
$$

while triangle is given by

$$
\left(\sum_{i=1}^{n}\left|x_{i}+y_{i}\right|^{2}\right)^{1 / 2} \leq\left(\sum_{i=1}^{n}\left|x_{i}\right|^{2}\right)^{1 / 2}+\left(\sum_{i=1}^{n}\left|y_{i}\right|^{2}\right)^{1 / 2}
$$

Both Hölder and Minkowski have generalizations to infinite sums, and also to Riemann and Lebesgue integrals.
4. Assume $d_{1}$ and $d_{2}$ are metrics on $X$. Let $\mathbf{x}=\left(x_{1}, x_{2}\right), \mathbf{y}=\left(y_{1}, y_{2}\right)$, and define

$$
d(\mathbf{x}, \mathbf{y})=\sqrt{d_{1}\left(x_{1}, y_{1}\right)^{2}+d_{2}\left(x_{2}, y_{2}\right)^{2}} .
$$

Prove that $d$ is a metric on $X^{2}$.
(Question and answer found here: https://math.stackexchange.com/questions/1211661/)
The result is most elegantly obtained by invoking the triangle inequality (special case of Minkowski) of a norm. From, e.g., Stoll, Def 7.4.8: Let $X$ be a vector space of $\mathbb{R}$. A function $\|\cdot\|: X \rightarrow \mathbb{R}$ on space $X$ satisfying
(a) $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x} \in X$,
(b) $\|\mathbf{x}\|=0 \Leftrightarrow \mathbf{x}=\mathbf{0}$,
(c) $\|c \mathbf{x}\|=|c|\|\mathbf{x}\|$ for all $c \in \mathbb{R}$ and $\mathbf{x} \in X$,
(d) $\|\mathbf{x}+\mathbf{y}\| \leq\|\mathbf{x}\|+\|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in X$.
is called a norm on $X$. We only require a simple version of Minkowski, namely Stoll (p. 331, Eq. 6). ANS: Consider the function $f:[0, \infty) \times[0, \infty) ; f(a, b)=\sqrt{a^{2}+b^{2}}$. It admits the following two results.
(i) $f(a, b) \leq f(c, d)$, when $a \leq c, b \leq d$.

Proof: For $a \leq c, b \leq d, f(a, b)=\sqrt{a^{2}+b^{2}} \leq \sqrt{c^{2}+d^{2}}=f(c, d)$.
(ii) $f(a+c, b+d) \leq f(a, b)+f(c, d)$.

Proof: For $(a, b),(c, d) \in \mathbb{R}^{2}$, and using the euclidean norm in $\mathbb{R}^{2}$, we have $\|(a, b)\|=$ $\sqrt{a^{2}+b^{2}}=f(a, b)$ and, similarly, $\|(c, d)\|=\sqrt{c^{2}+d^{2}}=f(c, d)$. Then

$$
\|(a, b)+(c, d)\|=\|(a+c, b+d)\|=\sqrt{(a+c)^{2}+(b+d)^{2}}=f(a+c, b+d)
$$

From (d), $\|(a, b)+(c, d)\| \leq\|(a, b)\|+\|(c, d)\|$, so that $f(a+c, b+d) \leq f(a, b)+f(c, d)$.

We can now prove the main result. Given $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X^{2}$, where $\mathbf{x}=\left(x_{1}, x_{2}\right), \mathbf{y}=\left(y_{1}, y_{2}\right), \mathbf{z}=\left(z_{1}, z_{2}\right)$,

$$
\begin{aligned}
d(x, z) & =f\left(d_{1}\left(x_{1}, z_{1}\right), d_{2}\left(x_{2}, z_{2}\right)\right) \\
& \leq f\left(d_{1}\left(x_{1}, y_{1}\right)+d_{1}\left(y_{1}, z_{1}\right), d_{2}\left(x_{2}, y_{2}\right)+d_{2}\left(y_{2}, z_{2}\right)\right) \\
& \leq f\left(d_{1}\left(x_{1}, y_{1}\right), d_{2}\left(x_{2}, y_{2}\right)\right)+f\left(d_{1}\left(y_{1}, z_{1}\right), d_{2}\left(y_{2}, z_{2}\right)\right. \\
& =d(x, y)+d(y, z),
\end{aligned}
$$

where the first inequality follows from (i) and the triangle inequalities for $d_{1}$ and $d_{2}$; and the second inequality follows from (ii). Thus, $d$ satisfies the triangle inequality.
5. If $d$ is a metric on $X$, prove that $|d(x, z)-d(z, y)| \leq d(x, y)$.

ANS: Let $x, y, z \in X$. By triangle, $d(x, z) \leq d(x, y)+d(y, z)$, or $d(x, z)-d(y, z) \leq d(x, y)$. Again, by triangle, $d(y, z) \leq d(y, x)+d(x, z)$, or $-(d(x, z)-d(y, z)) \leq d(y, x)=d(x, y)$. Combining yields $|d(x, z)-d(y, z)| \leq d(x, y)$.
6. (Closure of a set) Let $E$ be a subset of a metric space $X$.
(a) Define $\bar{E}$, the closure of $E$. Does $\bar{E}$ contain the isolated points of $E$ ?

ANS: $\bar{E}=E \cup E^{\prime}$, where $E^{\prime}$ is the derived set: the set of all limit points of $E$. By definition (see exercise 1e), $E^{\prime}$ does not contain the isolated points of $E$, but $\bar{E}=E \cup E^{\prime}$ does.
(b) Is $\bar{E}$ necessarily closed? (No proof required.)

ANS: It is closed, e.g., Stoll, Thm 2.2.18(a).
(c) Prove: $\bar{E}=E$ if and only if $E$ is closed.

ANS:

- If $E=\bar{E}$, then $E$ is closed (Stoll, Thm 2.2.18(a)). Conversely, if $E$ is closed, then $E^{\prime} \subset E$ (because Stoll, Thm 2.2.14: A subset $F$ of a metric space $X$ is closed if and only if $F$ contains all its limit points) and thus $\bar{E}:=E \cup E^{\prime}=E$.
- (Chen Lan) Let $(X, d)$ be a metric space. Let $F$ be a subset of $X$. Recall the definitions of boundary points, exterior points, and limit points. Note by contrapositives that(i) a point in $F^{c}$ is an exterior point of $F$ if and only if it is not a boundary point of $F$; (ii) a point in $F^{c}$ is an exterior point of $F$ if and only if it is not a limit point of $F$. Then, via the definitions, we have

$$
\begin{aligned}
F \text { is closed } & \Longleftrightarrow F^{c} \text { is open (i.e., every point in } F^{c} \text { is an interior point of } F^{c} \text { ) } \\
& \Longleftrightarrow \text { every point in } F^{c} \text { is an exterior point of } F \\
& \Longleftrightarrow F \text { contains all its boundary points } \\
\text { by (i) } & \Longleftrightarrow F \text { contains all its limit points } \\
\bar{F}=F \cup F^{\prime} & \Longleftrightarrow F=\bar{F} .
\end{aligned}
$$

Note that the last three statements are equivalent characterization of closed sets.
(d) Let $X=\mathbb{R}$, standard metric. Exhibit a proper subset $E \subset X$ such that $E \subset \bar{E}$.

ANS: $(a, b)=: E \subset \bar{E}=[a, b]$.
(e) Let $X=\mathbb{R}$, standard metric. Exhibit a proper subset $E \subset X$ such that $E^{\prime} \subset \bar{E}$, where $E^{\prime}$ is the derived set.
ANS: Let $E=[a, b] \cup\{c\}$, for $c>b$. Then $E^{\prime}=[a, b] \subset E=\bar{E}$. Note $E$ is closed (a finite union of closed set is closed) and $\bar{E}=E$, as stated above.
(f) Let $E, F$ be subsets of a metric space $X$ such that $E \subset F$ and $F$ is closed. Prove that $\bar{E} \subset F$. (Stoll, Thm 2.2.18(c))
ANS: Stoll states: If $E \subset F$ and $F$ is closed, then $E^{\prime} \subset F$. Thus $\bar{E} \subset F$.
We give here a bit more detail.
i. (Marc) We need to show $E^{\prime} \subset F$, where the derived set $E^{\prime}$ is the set of all limit points of $E$. If so, then $\bar{E}=\left(E \cup E^{\prime}\right) \subset F$. This in fact is easy, if we appeal to a definition common in some books, namely:

The closure $\bar{E}$ is defined as the smallest closed set containing $E$.
This implies, with $E \subset F$ and $F$ closed, $E^{\prime} \subset\left(E \cup E^{\prime}\right)=\bar{E} \subset F$.
This also begs the question to prove the equivalence of the two definitions of $\bar{E}$.
ii. (Ralf) We are given that $E \subset F$, and need to show $E^{\prime} \subset F$, because $\bar{E}=\left(E \cup E^{\prime}\right)$. We argue by contradiction. Assume $p \in F^{c}$ is a limit point of $E$. Then for every $\epsilon>0, \exists q \in E$ with $q \neq p$ such that $q \in N_{\epsilon}(p)$. As $E \subset F, q \in F$, so that $p$ is also a limit point of $F$. But as $F$ is closed, $p \in F$. (Recall: A closed set contains all its limit points; e.g., Stoll Thm 2.2.14). This contradiction implies $p \in F^{c}$ cannot be a limit point of $E$, and thus, as $p$ was arbitrary, $E^{\prime} \subset F$ and, thus, $\bar{E} \subset F$.
iii. (Chen Lan) Take any $p \in E^{\prime}$. This implies the existence of a sequence in $E \backslash\{p\} \subset F$ that converges to $p$ (Stoll, Thm 3.1.4 (c)). As $F$ is closed, $p \in F$ (Fitzpatrick's definition of closed sets).
7. Alternative definitions of open and closed; interior, exterior, boundary
(a) Define a boundary point for set $E \subset \mathbb{R}$; and $\partial E$, the boundary of $E$. Does $\partial E$ contain the isolated points of $E$ ? Does $\bar{E}$ contain the isolated points of $E$ ?
ANS: From Stoll, p. $68 \# 9$, a point $p \in \mathbb{R}$ is a boundary point of $E$ if, for every $\epsilon>0, N_{\epsilon}(p)$ contains both points of $E$ and points of $E^{c}$. The boundary of $E$ is the union of all boundary points of $E$.
Regarding whether $\partial E$ contain the isolated points of $E$, first imagine the metrics $d_{1}, d_{\infty}$, and $d_{2}$, for, as stated above, $X=\mathbb{R}$, and more generally, for $X=\mathbb{R}^{n}$. It is a standard result that these norms are equivalent. ${ }^{7}$ Based on the definition, $\partial E$ contains the isolated points of $E$.
However, for metric space $(X, d)$ with $d$ the trivial, or discrete, metric, this is not the case. From exercise 19 part (a): Every element of $E \subset X$ is an interior point of $E$; every element of $E$ is an isolated point; every point not contained in $E$ is an exterior point of $E$; and there are no boundary points, i.e., $\partial E=\emptyset$. As such, the set of isolated points is not contained in the boundary of $E$. Thus, whether or not $\partial E$ contain the isolated points of $E$ depends on $(X, d)$.
Next, recall (e.g., exercise 6a) that $\bar{E}=E \cup E^{\prime}$, where $E^{\prime}$ is the derived set; and that (exercise $1 \mathrm{e}), E^{\prime}$ does not contain the isolated points of $E$, but $\bar{E}=E \cup E^{\prime}$ clearly does.
(b) Recall the interior of subset $E$ is denoted $\operatorname{Int}(E)$ or $E^{\circ}$. Prove that

$$
\begin{equation*}
\partial E=\bar{E} \backslash E^{\circ} \quad \text { (equivalently) } \quad \bar{E}=\partial E \cup E^{\circ} . \tag{9}
\end{equation*}
$$

Conclude that a closed set is the union of its interior and its boundary.
ANS: The two statements in (9), if true, are clearly equivalent. We show the latter, i.e., $\left(E \cup E^{\prime}\right)=\left(\partial E \cup E^{\circ}\right)$. Point $x \in E$ is either a limit point, or an isolated point, of $E$, and so we consider these two disjoint cases separately.

[^5]- Let $x$ be a limit point of $E$. Then $x$ is contained in one of either $E$ or $E^{\prime} \backslash E$. Next observe that $x \in E^{\prime} \backslash E \Rightarrow x \in \partial E$, because $x \notin E$ but $x$ is a limit point, i.e., for every $\epsilon>0, \exists q \in E$ such that $q \in N_{\epsilon}(x)$. Clearly, $E^{\circ} \subset E$. Thus,

$$
\begin{equation*}
x \in\left(E \cup E^{\prime}\right) \Longleftrightarrow x \in\left(\partial E \cup E^{\circ}\right) . \tag{10}
\end{equation*}
$$

- Let $x \in E$ be an isolated point.

To show $\Longleftarrow$, as $x$ is an isolated point, it must be that $x \in \partial E$ (and $x \notin E^{\circ}$ ). But by assumption, $x \in E$.
To show $\Longrightarrow$, first consider the more common metrics $d_{1}, d_{\infty}$, and $d_{2}$ for $\mathbb{R}^{n}$. Then $x \in \partial E$, confirming $\Longrightarrow$. Now consider the discrete metric, for which (see exercise 19) every element of $E$ is an interior point of $E$; every point not contained in $E$ is an exterior point of $E$; and there are no boundary points. Thus, $x \in E \Rightarrow x \in E^{\circ}$.
This shows (10) holds.

- From exercise 6 c, $\bar{E}=E$ if and only if $E$ is closed. So, $E$ closed implies $\bar{E}=E$, and using (9), $E=\bar{E}=\partial E \cup \operatorname{Int}(E)$.
(c) For a metric space $(X, d)$ with $E \subset X$, conjecture the definition of an exterior point of $E$. The collection of all exterior points of $E$ is then defined as the exterior of $E$, denoted Ext $(E)$. Hint:

$$
\begin{equation*}
\text { For } E \subset X, X \text { can be partitioned into disjoint sets as } X=\operatorname{Int}(E) \cup \operatorname{Ext}(E) \cup \partial E \text {. } \tag{11}
\end{equation*}
$$

ANS:

- (Tao, Analysis II, Def 1.2.5) Let $(X, d)$ be metric space and $E \subset X$. A point $x$ is an exterior point of $E$ if $\exists r>0$ such that $B_{r}(x) \cap E=\emptyset$.
- (Heil, MTSE, forthcoming, p. 53): The exterior of $E$ is $\operatorname{Int}\left(E^{c}\right)$.
(d) For a metric space $(X, d)$ with $E \subset X$, Prove that $\partial E$ is closed.

ANS: From partition (11), $\partial E=(\operatorname{Int}(E) \cup \operatorname{Ext}(E))^{c}$. As both $\operatorname{Int}(E)$ and $\operatorname{Ext}(E)$ are open, and the union of open sets is open, it follows that the complement of their union, $\partial E$, is closed. Alternatively, from (9), $\partial E=\bar{E} \cap(\operatorname{Int}(E))^{c}$, which, as $\operatorname{Int}(E)$ is open, is the intersection of two closed sets, and thus closed.
(e) Prove that a set $E \subset \mathbb{R}$ is open if and only if $E$ does not contain any of its boundary points. (Stoll, p. 68, \#10a)
ANS:
$(\Rightarrow)$ Assume $E$ is open. By definition (Stoll, Def 2.2.5(a)), every point of $E$ is an interior point of $E$, and thus cannot contain any boundary points.
$(\Leftarrow)$ Assume $E$ does not contain any of its boundary points. Clearly, $E$ does not contain any of its exterior points. So, every point $a \in E$ must be an interior point, i.e., for every $a \in E, \exists \epsilon_{a}$ such that $B_{\epsilon_{a}}(a) \subset A$. Thus, $E$ is open.
(f) Prove that a set $E \subset \mathbb{R}$ is closed if and only if $E$ contains all its boundary points.

ANS: Recall Stoll, Def 2.2.5(b): A subset $F$ of $\mathbb{R}$ is closed if $F^{c}=\mathbb{R} \backslash F$ is open. Also, Stoll, Thm 2.2.14: A subset $F$ of a metric space $X$ is closed if and only if $F$ contains all its limit points.
$(\Rightarrow)$ Assume $E$ is closed. Then $E^{c}$ is open, and, as just shown, does not contain any of its boundary points. Thus, the set of boundary points must be contained in $X \backslash E^{c}=E$.
$(\Leftarrow)$ Assume $E$ contains all its boundary points. Thus, $E^{c}$ does not contain any boundary points of $E$, but the boundary of $E$ is the same as the boundary of $E^{c}$, so, as shown above, $E^{c}$ is open, and thus $E$ is closed.
8. Consider the set $\mathbb{Q}$, the set of rational numbers in $\mathbb{R}$, and the set of irrationals, $\mathbb{I}:=\mathbb{R} \backslash \mathbb{Q}$.
(a) Determine $\operatorname{Int}(\mathbb{Q})$ and $\partial \mathbb{Q}$.

ANS: Recall the definition of interior point from exercise 1c. We know every $\epsilon$-neighborhood of $q \in \mathbb{Q}$ contain irrational numbers, because the irrationals are dense in $\mathbb{R}$ (see, e.g., Fitzpatrick, p. 15) and thus no point of $\mathbb{Q}$ is an interior point, i.e., $\operatorname{Int}(\mathbb{Q})=\emptyset$.

From $(9), \partial \mathbb{Q}=\overline{\mathbb{Q}} \backslash \operatorname{Int}(\mathbb{Q})=\mathbb{R} \backslash \emptyset=\mathbb{R}$.
(b) Demonstrate with full clear arguments whether or not $\mathbb{Q}$ is open, and whether or not $\mathbb{Q}$ is closed. Same for $\mathbb{I}:=\mathbb{R} \backslash \mathbb{Q}$.
ANS: For learning purposes, I begin with my initial answer regarding if $\mathbb{Q}$ is open, which is wrong, as indicated, and then show the correct way, as well as a second proof. Finally, Chen gives the shortest, simplest proof, leveraging previous results.

- (Marc) (Intentionally Erroneous)

No point of $\mathbb{Q}$ is an interior point of $\mathbb{Q}$, because $\mathbb{Q}$ is dense in $\mathbb{R}$, so $\operatorname{Int}(\mathbb{Q})=\emptyset$. By definition, a subset $O$ of $\mathbb{R}$ is open if every point of $O$ is an interior point of $O$. So, $\mathbb{Q}$ is not open.

- (Christopher Heil, commenting on my above "proof")

Not quite true. $\mathbb{R}$ is dense in $\mathbb{R}$ and every point of $\mathbb{R}$ is an interior point. It is true that no point of $\mathbb{Q}$ is an interior point, but that is not because it is dense.

- (Marc, upgrade; correct proof)
$O \subset \mathbb{R}$ is open if every point of $O$ is an interior point of $O$. From part (a), $\operatorname{Int}(\mathbb{Q})=\emptyset$, i.e., no point of $\mathbb{Q}$ is an interior point of $\mathbb{Q}$, so $\mathbb{Q}$ is not open.
- (Marc) The set of irrationals, $\mathbb{I}=\mathbb{R} \backslash \mathbb{Q}$ is not closed, because, either:
- Recall: A subset $F$ of a metric space $X$ is closed if and only if $F$ contains all its limit points. (See exercise 11). $\mathbb{I}$ does not contain all its limit points, e.g., $a_{n}=\sqrt{2} / n$.
- Its closure $\overline{\mathbb{I}}=\mathbb{R} \neq \mathbb{I}$, and thus $\mathbb{I}$ is not closed. (See exercise 6 c ).

Thus, $\{\mathbb{I}$ not closed $\} \Leftrightarrow\left\{\mathbb{I}^{c}=\mathbb{Q}\right.$ not open $\}$. (See exercise 1h.)
Set $\mathbb{Q}$ is also not closed. Here are three ways of seeing this.
i. The complement of $\mathbb{Q} \in \mathbb{R}$ is $\mathbb{I}$, the irrationals, which, for the same reason as the rationals, is not open. Thus, $\{\mathbb{I}$ not open $\} \Leftrightarrow\left\{\mathbb{I}^{c}=\mathbb{Q}\right.$ not closed $\}$.
ii. Recall that a subset $S$ of $\mathbb{R}$ is closed if, for every sequence $\left\{a_{n}\right\} \in S$ that converges to a number $a, a \in S$. This is clearly not fulfilled for $\mathbb{Q}$.
iii. Recalling exercise 1 g , a set $A$ is dense in $\mathbb{R}$ if $\bar{A}=\mathbb{R}$ and it is closed if $A=\bar{A}$. As $\mathbb{Q} \neq \overline{\mathbb{Q}}=\mathbb{R}, \mathbb{Q}$ fails to be closed.

- (Chen Lan)

Using exercises 7 e and 7 f regarding the boundary, it follows that both $\mathbb{Q}$ and $\mathbb{I}$ are neither open, nor closed.
9. $\cup O_{i}$ and $\cap F_{i}$
(a) Prove: The union of any collection (finite, countable, or uncountable) of open sets in $\mathbb{R}$ is an open set in $\mathbb{R}$.
ANS: Let $A$ be any index set and let $O:=\cup_{\alpha \in A} O_{\alpha}$ where, for each $\alpha \in A, O_{\alpha}$ is an open set of real numbers. If $x \in O$, then $x \in O_{\alpha_{0}}$ for some $\alpha_{0} \in A$. As $O_{\alpha_{0}}$ is open, there is an $\epsilon>0$ such that $(x-\epsilon, x+\epsilon) \subset O_{\alpha_{0}} \subset O$. Therefore, every point of the union is an interior point and thus, by definition, $O$ is open.
(b) Give an example of a countable intersection of open sets that is not open.

ANS: Take, for example $\{1\}=\cap_{j=1}^{\infty}(1-1 / j, 1+1 / j)$.
(c) Prove that intersections of any collection (finite, countable, or uncountable) of closed sets is closed.
ANS: Let $F_{\alpha}$ be a closed set of real numbers for each $\alpha$ in some index set $A$, and let $F=\cap_{\alpha \in A} F_{\alpha}$. Set $F$ is closed if $F^{c}$ is open. To see the latter, DeMorgan's rule implies $F^{c}=\cup_{\alpha \in A} F_{\alpha}^{c}$. This is open, as a union of open sets, so $F^{c}$ is open.
(d) Give an example of a countable union of closed sets that is not closed.

ANS: $\cup_{j=2}^{\infty}[1 / j, 1-1 / j]=(0,1)$.
10. Is $\mathbb{Z}$ (the set of integers) closed (in ambient set $\mathbb{R}$ )?

ANS: Yes.
Proof I: complement $C:=\mathbb{R} \backslash \mathbb{Z}$ is a countable union of open sets, so is open (from, e.g., Stoll, Thm 2.2.9(a)). Thus, $\mathbb{Z}=C^{c}$ is closed, by definition (Stoll, Def 2.2.5(b)).

Proof II: Recall exercises 6 a and $6 \mathrm{c} . \mathbb{Z}$ consists strictly of isolated points (with respect to ambient space $\mathbb{R}$ ) and has no limit points. So, $\bar{E}=E$, implying $E$ is closed.
Proof III: (Chen Lan) From exercise $1 \mathrm{~g}, \mathbb{Z}$ is not dense in $\mathbb{R}$. Thus no point of $\mathbb{R} \backslash \mathbb{Z}$ is a boundary point of $\mathbb{Z}$, and so $\mathbb{Z}$ contains all its boundary points. From exercise $7 \mathrm{f}, \mathbb{Z}$ is closed.
11. State (possibly again) the definition of a closed set as given in Stoll.

Prove: A subset $F$ of a metric space $X$ is closed if and only if $F$ contains all its limit points.
ANS: This is Stoll, Thm 2.2.14; and is the definition as given in Fitzpatrick p. 37.
12. For metric space $(X, d)$ and $A \subset X$, the distance from $x \in X$ to $A$ is defined by $d(x, A):=$ $\inf _{a \in A} d(x, a)$. Further, as a definition, $d(x, \emptyset):=+\infty$. Observe, $\forall a \in A, 0 \leq d(x, A) \leq d(x, a)$.
(a) Let $A \subset B \subset X$. Determine a space $X$, a metric $d$, and an $A$ and $B$ where $A$ is a strict subset of $B$, such that, $\forall x \in X, d(x, A)=d(x, B)$.
ANS: Take $X=\mathbb{R}$, usual metric, $a, b \in \mathbb{R}, a<b, A=(a, b), B=[a, b]$.
(b) (Re-)state the definition of open set for metric spaces, and conjecture, without a detailed proof, an equivalent definition based on the above-defined distance from $x \in X$ to $A \subset X$.
ANS: A subset $A$ of $\mathbb{R}$ is open if every point of $A$ is an interior point of $A$. Consider the following equivalent definition: A subset $A$ of metric space $(X, d)$ is open if $d(a, X \backslash A)>0$. A formal proof of equivalence can be found in Field, Essential Real Analysis, Prop 7.4.7.
(c) The distance between nonempty subsets $A, B \in X$ is defined as

$$
\begin{equation*}
d(A, B):=\inf \{d(x, y): x \in A, y \in B\} \tag{12}
\end{equation*}
$$

Suppose $A, B \subset \mathbb{R}$ with $A$ compact, $B$ closed, and $A \cap B=\emptyset$. Show $d(A, B)>0$.
ANS: Answer from https://math.stackexchange.com/questions/2243887
Assume $d(A, B)=0$. Then $\exists\left\{a_{n}\right\} \subset A$ and $\exists\left\{b_{n}\right\} \subset B$ such that $\left|a_{n}-b_{n}\right| \rightarrow 0$. The latter implies that, for any subsequences $\left\{a_{n_{k}}\right\}$ and $\left\{b_{n_{k}}\right\},\left|a_{n_{k}}-b_{n_{k}}\right| \rightarrow 0$. As $A$ is compact (recall Fitzpatrick, p. 46, definition of sequential compactness), $\exists\left\{a_{n_{k}}\right\}$ such that $a_{n_{k}} \rightarrow a \in A$, so that

$$
\left|a-b_{n_{k}}\right| \leq\left|a-a_{n_{k}}\right|+\left|a_{n_{k}}-b_{n_{k}}\right| \xrightarrow{k \rightarrow \infty} 0 .
$$

Thus, $a$ is a limit point of $\left\{b_{n}\right\}$, but as $B$ is closed, $a \in B$. But $a \in A$, contradicting that $A \cap B=\emptyset$. So, $d(A, B) \neq 0$, which means $d(A, B)>0$.
(d) Give an example of two non-compact sets $A, B \in X$ such that $d(A, B)>0$. More challenging is to give an example of $A, B \in X$ with $A \cap B=\emptyset, B$ is closed, and $A$ is closed but not necessarily compact, and such that $d(A, B)=0$. Thus, show that we cannot relax the condition that $A$ is compact.
ANS: For the first part, trivially, take $A=(-\infty, 0)$ and $B=(1, \infty)$. For the second part, consider the following.

- From Exam 2 solutions: https://www.math.colostate.edu/~clayton/teaching/m317f10/

Let $A=\{n: n=2,3,4, \ldots\}$ and $B=\{n+1 / n: n=2,3,4, \ldots\}$. Both $A$ and $B$ consist entirely of isolated points (in each case, the distance between distinct elements is at least $1 / 2$ ), so both $A$ and $B$ are closed. Fix $\epsilon>0$. From the Archimedean Property, $\exists N \in \mathbb{N}$ such that $N>\max (1 / \epsilon, 1)$. Now, $N \in A$ and $N+1 / N \in B$, and $|N-(N+1 / N)|=1 / N<\epsilon$. Such an $N$ exists for every $\epsilon>0$, so $\nexists \epsilon>0$ such that $\forall(a \in A, b \in B),|a-b|>\epsilon$.

- (Chen Lan) Consider any $x=n_{1} \in A$ and $y=\left(n_{2}+1 / n_{2}\right) \in B, n_{1}, n_{2}=2,3, \ldots$. If $n_{1} \neq n_{2}$, then $|x-y|>1$. If $n_{1}=n_{2}=n, n=2,3, \ldots$, then $|x-y|=n^{-1}>0$ and $|x-y| \rightarrow 0$ as $n \rightarrow \infty$. Thus, $d(A, B)=0$.

13. If $A \subset \mathbb{R}$, do $A$ and $\bar{A}$ always have the same interior? Do $A$ and $\operatorname{Int}(A)$ always have the same closure?

ANS: Exam 2 solutions: https://www.math.colostate.edu/~clayton/teaching/m317f10/

- No. Let $A=(-\infty, 0) \cup(0, \infty)$. Then $A=\operatorname{Int}(A)$, but $\bar{A}=\mathbb{R}$, so $\operatorname{Int}(\bar{A})=\mathbb{R}$.
- No. Let $A=\{1\}$. There is no $\epsilon>0$ such that $B_{\epsilon}(1) \subset A$, so $A$ has no interior points, i.e., $\operatorname{Int}(A)=\emptyset$, so $\overline{\operatorname{Int}(A)}=\emptyset$. As $A$ is closed, $\bar{A}=A=\{1\}$.

14. Let $A, B$ be subsets of $\mathbb{R}$. (Stoll, p. 69, \#18)
(a) If $A \subset B$, show $\operatorname{Int}(A) \subset \operatorname{Int}(B)$.

ANS: If $a \in \operatorname{Int}(A)$, then $\exists \epsilon>0$ such that $N_{\epsilon}(a) \subset A \subset B$, so $a \in \operatorname{Int}(B)$.
(b) Show $\operatorname{Int}(A \cap B)=\operatorname{Int}(A) \cap \operatorname{Int}(B)$.

ANS: Need to prove inclusion in both directions.

- $\operatorname{Int}(A \cap B) \subset \operatorname{Int}(A) \cap \operatorname{Int}(B)$ :
$p \in \operatorname{Int}(A \cap B)$ implies $\exists \epsilon>0$ such that $N_{\epsilon}(p) \subset(A \cap B) \subset A$, so $p \in \operatorname{Int}(A)$; likewise $p \in \operatorname{Int}(B)$; and thus $p \in \operatorname{Int}(A) \cap \operatorname{Int}(B)$.
- $\operatorname{Int}(A) \cap \operatorname{Int}(B) \subset \operatorname{Int}(A \cap B)$ :
$p \in \operatorname{Int}(A) \cap \operatorname{Int}(B)$ implies $p \in \operatorname{Int}(A)$ and $p \in \operatorname{Int}(B)$, i.e., $\exists \epsilon_{A}>0$ such that $N_{\epsilon_{A}}(p) \subset A$ and $\exists \epsilon_{B}>0$ such that $N_{\epsilon_{B}}(p) \subset B$. With $\epsilon:=\min \left(\epsilon_{A}, \epsilon_{B}\right), N_{\epsilon}(p) \subset A$ and $N_{\epsilon}(p) \subset B$, i.e., $N_{\epsilon}(p) \subset(A \cap B)$, showing that $p \in \operatorname{Int}(A \cap B)$.
(c) Is $\operatorname{Int}(A \cup B)=\operatorname{Int}(A) \cup \operatorname{Int}(B)$ ?

ANS: No. Let $A=\mathbb{Q}, B=\mathbb{I}=\mathbb{R} \backslash \mathbb{Q} \cdot \operatorname{Int}(A \cup B)=\mathbb{R}, \operatorname{Int}(\mathbb{Q})=\operatorname{Int}(\mathbb{I})=\emptyset$.
(d) Are the results of (a) and (b) still true if $A$ and $B$ are subsets of a metric space $X$ ?

ANS: Yes. (Answer confirmed by Manfred Stoll.)
15. Let $A, B$ be subsets of a metric space $X$. (Stoll, p. 69, \#19)
(a) Show that $\overline{A \cup B}=\bar{A} \cup \bar{B}$.

ANS: We need to show $\overline{A \cup B} \subset \bar{A} \cup \bar{B}$ and $\overline{A \cup B} \supset \bar{A} \cup \bar{B}$.

- To show $\overline{A \cup B} \subset \bar{A} \cup \bar{B}$ : By definition, $\bar{A}=A \cup A^{\prime}$, so $A \subset \bar{A}$. Likewise, $B \subset \bar{B}$, so that $A \cup B \subset \bar{A} \cup \bar{B}$. Note $\bar{A} \cup \bar{B}$, as a finite union of closed sets, is closed. Recall (Stoll, Thm 2.2.18(c)):

For $E, F$ subsets of a metric space $X$ such that $E \subset F$ and $F$ is closed, $\bar{E} \subset F$.
This implies $\overline{A \cup B} \subset \bar{A} \cup \bar{B}$.

- Proof I: To show $\overline{A \cup B} \supset \bar{A} \cup \bar{B}$, we first show that $A \subset(A \cup B) \Rightarrow \bar{A} \subset \overline{A \cup B}$. This follows by inspecting the two mutually exclusive and exhaustive cases.
Case 1: Let $a$ be a limit point of $A$. Then, by definition, $a \in \bar{A}$, and $a$ is also a limit point of the larger set $A \cup B$, i.e., $a \in \overline{A \cup B}$.
Case 2: Let $a$ be an isolated point of $A$. Then $a \in A \subset\left(A \cup A^{\prime}\right)=\bar{A}$.
But also $a \in(A \cup B) \subset \overline{A \cup B}$.
So, $A \subset(A \cup B) \Rightarrow \bar{A} \subset \overline{A \cup B}$, and, likewise, $\bar{B} \subset \overline{A \cup B}$, so $\bar{A} \cup \bar{B} \subset \overline{A \cup B}$.
- Proof II (Ralf and Chen): $A \subset A \cup B \subset \overline{A \cup B}$ and $\overline{A \cup B}$ is closed $\Rightarrow \bar{A} \subset \overline{A \cup B}$ (Stoll, Thm 2.2.18 (c)). Also, $\bar{B} \subset \overline{A \cup B}$. Thus, $\bar{A} \cup \bar{B} \subset \overline{A \cup B}$.
(b) Show that $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$, and give an example for which the containment is proper.

ANS: Similar to the above result, with proof, that $A \subset(A \cup B) \Rightarrow \bar{A} \subset \overline{A \cup B}$, we assume without proof: $(A \cap B) \subset A \Rightarrow \overline{A \cap B} \subset \bar{A}$ and, likewise, $\overline{A \cap B} \subset \bar{B}$. These two together imply (make a Venn diagram if needed) $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$.
As an example, let $X=\mathbb{R}, A=(0,1), B=(1,2)$, with $A \cap B=\emptyset$, and closure $\overline{A \cap B}=\emptyset$. But $\bar{A}=[0,1], \bar{B}=[1,2]$, and $\bar{A} \cap \bar{B}=\{1\}$.
16. Let $K \subset \mathbb{R}$ be nonempty and compact.

Prove: $\sup K$ and $\inf K$ exist and are in $K$. (Stoll, p. 73, \#4)
Do so by using only technology up to and including Stoll, Sec. 2.3, and then state how results from his Sec. 2.4 could also be used.
ANS: First show $K$ is bounded.

- Using, e.g., Stoll, Them 2.3.5, $K$ compact $\Rightarrow K$ closed. To see that $K$ is bounded using only the material up to and including Stoll, Sec. 2.3, we can argue as follows. If $K$ were unbounded, then we can find a countable cover of open sets such that it does not admit a finite subcover, and thus $K$ could not be compact. (See, e.g., Stoll, Example 2.3.4 (c).) Thus, compact implies bounded.
- (Chen Lan) Suppose that $K$ is compact, i.e, every open cover of $K$ has a finite subcover. Let $I_{n}=(-n, n), \forall n \in \mathbb{N}$. Then $\left\{I_{n}\right\}_{n=1}^{\infty}$ is an open cover of $\mathbb{R}$, and thus also of $K . K$ is compact $\Rightarrow \exists N \in \mathbb{N}$ such that $K \subset \cup_{n=1}^{N} I_{n}=(-N, N)$. Then, $\forall x \in K,|x|<N$; that is, $K$ is bounded.
- We could have also invoked the Heine-Borel theorem from (Stoll, Sec. 2.4, Thm 2.4.2), which implies $K$ is (closed and) bounded.

Existence of $\alpha:=\sup K \in \mathbb{R}$ follows because $K$ is nonempty and bounded. To show $\alpha \in K$, assume the contrary, $\alpha \in K^{c}$. As $K^{c}$ is open, $\exists \epsilon>0$ such that $B_{\epsilon}(\alpha) \in K^{c}$. By definition of sup, $\exists a_{n} \in K$ such that, $\forall n \in \mathbb{N}, \alpha-1 / n<a_{n} \leq \alpha$. But for $n>1 / \epsilon, a_{n} \in B_{\epsilon}(\alpha) \in K^{c}$, which is a contradiction.

- We can conclude this from Bolzano-Weierstrass, Stoll, Sec. 2.4.

The proof for the infimum is similar.
17. Let $A=\{1 / n: n \in \mathbb{N}\}$. (Stoll, p. 73, \#1)
(a) Show that $A$ is not compact.

ANS: As in Stoll, Example 2.2.13(b), for each $n \in \mathbb{N}$, choose $0<\epsilon_{n}<n^{-1}-(n+1)^{-1}$, in which case $N_{\epsilon_{n}}(1 / n) \cap A=\{1 / n\}$. This forms a countably infinite cover of $A$ such that no finite subcover can cover $A$, and thus $A$ is not compact.

Notice 0 is the only limit point of $A$ (and all points of $A$ are isolated points), and, as $0 \notin A, A$ cannot be closed, and thus not compact.
(b) Prove directly (using the definition) that $K=A \cup\{0\}$ is compact.

ANS: For any $\epsilon>0, B_{\epsilon}(0)$ covers $\{0\}$, and an infinite number of the $1 / n$ sequence in $A$ (because 0 is a limit point of $A$ ). The finite rest of the elements of $A$ can be covered by a finite number of open sets as in part (a).
NOTE: This proof was based, as requested, on the definition of (topological) compactness. We could also use Heine-Borel, or Bolzano-Weierstrass, to see that, as $K$ is closed and bounded, it must be compact.
18. Definitions and results for sequences
(a) State the definition of sequence convergence for metric space $(X, d)$ in terms of neighborhoods (as in Stoll), and also the distance measure.
ANS: (Stoll, Def 3.1.1) A sequence $\left\{p_{n}\right\}_{n=1}^{\infty}$ is said to converge if there exists a point $p \in X$ such that, for every $\epsilon>0$, there exists a positive integer $n_{o}=n_{o}(\epsilon)$ such that $p_{n} \in N_{\epsilon}(p)$ for all $n \geq n_{o}$. If this is the case, we say $\left\{p_{n}\right\}$ converges to $p$ or that $p$ is the limit of the sequence $\left\{p_{n}\right\}$, and we write $p_{n} \rightarrow p$.
The neighborhood condition $p_{n} \in N_{\epsilon}(p)$ for all $n \geq n_{o}$ is equivalent to distance expression $d\left(p_{n}, p\right)<\epsilon$ for all $n \geq n_{o}$.
(b) Let $\left\{a_{n}\right\} \subset \mathbb{R}$ be a convergent sequence with $\lim a_{n}=a$, and suppose $a_{n} \geq 0$ for all $n \in \mathbb{N}$. Prove that $a \geq 0$. (Petrovic, Prop 2.3.11)
ANS: If $a<0$, then, with $\epsilon=-a / 2, N_{\epsilon}(a)=(a-\epsilon, a+\epsilon)$ does not contain any $a_{n}$, and thus $\left\{a_{n}\right\}$ cannot converge to $a$, a contradiction.
(c) Give an example showing that, for sequence $\left\{a_{n}\right\} \in \mathbb{R}$ with $a_{n}>0$ for all $n \in \mathbb{N}$, it need not be true that $\lim a_{n}>0$.
ANS: Take $a_{n}=1 / n$.
(d) Let $\left\{a_{n}\right\},\left\{b_{n}\right\} \in \mathbb{R}$ be two convergent sequences, with $\lim a_{n}=a, \lim b_{n}=b$, and suppose $a_{n} \leq b_{n}$ for all $n \in \mathbb{N}$. Prove:

$$
\begin{equation*}
a \leq b \tag{13}
\end{equation*}
$$

(Petrovic, Prop 2.3.12; Laczkovich and Sós, Thm 5.10; Boules, Example 5, p. 12) Do so in two ways. First, using part (b); and then not using part (b).
ANS:
i. Let $c_{n}=b_{n}-a_{n}$. Then $c_{n} \geq 0$ and $\lim c_{n}=\lim b_{n}-\lim a_{n}=b-a$. By part (b), $\lim c_{n} \geq 0$, so $b \geq a$.
ii. Suppose $b<a$. Let $\epsilon=(a-b) / 3$. Then $b-\epsilon<b+\epsilon<a-\epsilon<a+\epsilon$, and $\exists N_{1}, N_{2} \in \mathbb{N}$ such that, for $n>N_{1}, b_{n} \in(b-\epsilon, b+\epsilon)$, and, for $n>N_{2}, a_{n} \in(a-\epsilon, a+\epsilon)$. Let $N=\max \left(N_{1}, N_{2}\right)$. Thus, $\forall n>N, b_{n}<a_{n}$.
(e) Let $\left\{a_{n}\right\} \in \mathbb{R}$ be a convergent sequence with $\lim a_{n}=a$. Prove: $\lim a_{n}^{2}=a^{2}$. (Stoll, p. 87, \#5) Note: The result follows from continuity of $f: \mathbb{R} \rightarrow[0, \infty) ; f(x)=x^{2}$, but this result is forthcoming in our course, so prove the result without its use.
ANS: Need show that, for any $\epsilon>0, \exists N \in \mathbb{N}$ such that, for $n>N,\left|a_{n}^{2}-a^{2}\right|<\epsilon$.
Note $a_{n}^{2}-a^{2}=\left(a_{n}+a\right)\left(a_{n}-a\right)$. As $\left\{a_{n}\right\}$ is convergent, it is bounded, so $\exists M \in \mathbb{R}_{+}$such that, $\forall n \in \mathbb{N},\left|a_{n}\right| \leq M$. Further, for any $\epsilon_{1}>0, \exists N \in \mathbb{N}$ such that, for $n>N,\left|a_{n}-a\right|<\epsilon_{1}$. The (regular) triangle inequality then implies that, for $n>N$,

$$
\left|a_{n}^{2}-a^{2}\right| \leq\left|a_{n}+a\right| \times\left|a_{n}-a\right| \leq\left(\left|a_{n}\right|+|a|\right) \times\left|a_{n}-a\right|<(M+|a|) \epsilon_{1} .
$$

Taking $\epsilon_{1}=\epsilon /(M+|a|)$ shows $\left|a_{n}^{2}-a^{2}\right|<\epsilon$.
(f) Let $\left\{a_{n}\right\} \in \mathbb{R}_{\geq 0}$ be a convergent sequence with $\lim a_{n}=a$. Prove: $\lim \sqrt{a_{n}}=\sqrt{a}$.
(Stoll, p. 87, \#6; Junghenn, Thm 2.1.11(e))
Note: The result follows from continuity of $f:[0, \infty) \rightarrow[0, \infty) ; f(x)=x^{1 / 2}$, but this result is forthcoming in our course, so prove the result without its use.
ANS: From part (b), $a \geq 0$. Choose any $\epsilon>0$. If $a=0$, choose $N$ such that $a_{n}<\epsilon^{2}$ for all $n \geq N$. If $a>0$, choose $N$ such that $\left|a_{n}-a\right|<\epsilon \sqrt{a}$ for all $n \geq N$. For such $n$,

$$
\left|\sqrt{a_{n}}-\sqrt{a}\right|=\frac{\left|a_{n}-a\right|}{\sqrt{a_{n}}+\sqrt{a}} \leq \frac{\left|a_{n}-a\right|}{\sqrt{a}}<\epsilon
$$

(g) Sequences involving $n$th roots
i. Let $n \in \mathbb{N}$ and $a, b \in \mathbb{R}_{\geq 0}$. Prove: $\left|a^{1 / n}-b^{1 / n}\right| \leq|a-b|^{1 / n}$.
(Ghorpade and Limaye, Prop 1.9(ii))
ANS: Assume without loss of generality (w.l.o.g.) that $a \geq b$. Let $c=a^{1 / n}$ and $d=b^{1 / n}$. Then $c-d \geq 0$ and, by the Binomial Theorem,

$$
c^{n}=[(c-d)+d]^{n}=(c-d)^{n}+\cdots+d^{n} \geq(c-d)^{n}+d^{n} .
$$

Therefore, $a-b=c^{n}-d^{n} \geq(c-d)^{n}=\left[a^{1 / n}-b^{1 / n}\right]^{n}$.
ii. Let $\left\{a_{n}\right\} \in \mathbb{R}_{\geq 0}$ be a sequence with $\lim a_{n}=a$. Prove that $a \geq 0$ and, for every $k \in \mathbb{N}$, $a_{n}^{1 / k} \rightarrow a^{1 / k}$. As with the questions above involving $a_{n}^{2}$ and $\sqrt{a_{n}}$, continuity could be invoked, but do not use that. Hint: Use part (i).
ANS: That $a \geq 0$ follows from (13). (Ghorpade and Limaye, Prop 2.4(iv)) Let $k \in \mathbb{N}$ and $\epsilon>0$ be given. As $\epsilon^{k}>0, \exists n_{2} \in \mathbb{N}$ such that, $\forall n \geq n_{2},\left|a_{n}-a\right|<\epsilon^{k}$. The result in part (i) then implies

$$
\left|a_{n}^{1 / k}-a^{1 / k}\right| \leq\left|a_{n}-a\right|^{1 / k}<\epsilon, \quad \forall n \geq n_{2} .
$$

Thus, $a_{n}^{1 / k} \rightarrow a^{1 / k}$.
(h) Let $\left\{a_{n}\right\},\left\{b_{n}\right\} \in \mathbb{R}$ be sequences such that $\left\{b_{n}\right\}$ is bounded and $\lim a_{n}=0$. Prove $\lim a_{n} b_{n}=0$. (Stoll, Thm 3.2.3)
ANS: As $\left\{b_{n}\right\}$ is bounded, $\exists M \in \mathbb{R}_{+}$such that, $\forall n \in \mathbb{N},\left|b_{n}\right| \leq M$. As $\left\{a_{n}\right\}$ is convergent, $\exists N \in \mathbb{N}$ such that, for any given $\epsilon>0,\left|a_{n}\right|<\epsilon / M$. Then $0 \leq\left|a_{n} b_{n}\right|=\left|a_{n}\right| \times\left|b_{n}\right|<\epsilon$.
(i) (Squeeze Theorem) Suppose $\left\{a_{n}\right\},\left\{b_{n}\right\},\left\{c_{n}\right\} \in \mathbb{R}$ are sequences for which there exists $n_{o} \in \mathbb{N}$ such that $a_{n} \leq b_{n} \leq c_{n}$ for all $n \in \mathbb{N}, n \geq n_{o}$, and that $\lim _{n \rightarrow \infty} a_{n}=\lim _{n \rightarrow \infty} c_{n}=L$. Prove: the sequence $\left\{b_{n}\right\}$ converges and $\lim _{n \rightarrow \infty} b_{n}=L$.
ANS: For any $\epsilon>0, \exists n_{a} \in \mathbb{N}$ such that, $\forall n \geq n_{a}, a_{n} \in N_{\epsilon}(L)$; and $\exists n_{c} \in \mathbb{N}$ such that, $\forall n \geq$ $n_{c}, c_{n} \in N_{\epsilon}(L)$. Set $N=\max \left(n_{o}, n_{a}, n_{c}\right)$, so that $\forall n \geq N,\left\{a_{n}, c_{n}\right\} \in N_{\epsilon}(L)$ and $a_{n} \leq b_{n} \leq c_{n}$. Then $b_{n} \in N_{\epsilon}(L)$ for $n \geq N$, which is the definition of convergence of sequence $\left\{b_{n}\right\}$.
(j) For $a \in \mathbb{R}_{+}$, i.e., $a>0$, prove that $\lim _{n \rightarrow \infty} \sqrt[n]{a}=1$.

ANS: (Petrovic, Example 2.9.1) First let $a \geq 1$. Recall Bernoulli's inequality (which is just the first term in the binomial theorem expansion, or can be proven by induction): For $x>-1$ and $n \in \mathbb{N},(1+x)^{n} \geq 1+n x$. With $x:=\sqrt[n]{a}-1 \geq 0$,

$$
a=(1+x)^{n} \geq 1+n x=1+n(\sqrt[n]{a}-1), \quad \text { or } \quad 0 \leq \sqrt[n]{a}-1 \leq \frac{a-1}{n}
$$

The squeeze Theorem implies $\lim a_{n}=1$.

Now consider the case for which $0<a<1$. Let $b=1 / a>1$. From the previous result, $1=\lim \sqrt[n]{b}=\lim b^{1 / n}$. Recalling the result for limits of ratios (e.g., Stoll, Thm 3.2.1(c)),

$$
\lim \sqrt[n]{a}=\lim \frac{1}{\sqrt[n]{b}}=\frac{\lim 1}{\lim \sqrt[n]{b}}=\frac{1}{1}=1
$$

(k) For $n \in \mathbb{N}$, set $p_{n}=n^{1 / n}$. (Stoll, p. $106 \# 6$ )
i. Show that, for $n \geq 3,1<p_{n+1}<p_{n}$.

ANS: To show $1<p_{n}$, observe: $1<p_{n} \Leftrightarrow 1=1^{n}<n$ for $n \geq 2$. To show $p_{n+1}<p_{n}$, I found a very nice solution on https://math.stackexchange.com/questions/2503266: For $n \geq 3$,

$$
p_{n}>p_{n+1} \Leftrightarrow n^{1 / n}>(n+1)^{1 /(n+1)} \Leftrightarrow n^{n+1}>(n+1)^{n} \Leftrightarrow\left(\frac{n+1}{n}\right)^{n}<n .
$$

Thus, we need to show, for $n \geq 3$,

$$
\sum_{i=0}^{n}\binom{n}{i}\left(\frac{1}{n}\right)^{i}=\left(1+\frac{1}{n}\right)^{n}=\left(\frac{n+1}{n}\right)^{n}<n
$$

but the lhs is

$$
\sum_{i=0}^{n} \frac{n(n-1) \ldots(n-i+1)}{i!n^{i}} \leq \sum_{i=0}^{n} \frac{1}{i!} \leq 1+\sum_{i=1}^{n} \frac{1}{2^{i-1}} \leq 1+2 \leq n
$$

ii. Let $p=\lim p_{n}$. Use the fact that the subsequence $\left\{p_{2 n}\right\}$ also converges to $p$ to conclude that $p=1$.
ANS: From exercise 18 f above, and the results for products of convergent sequences, and using Stoll Example 3.4.4(a) as a template,

$$
p=\lim _{n \rightarrow \infty} p_{2 n}=\lim _{n \rightarrow \infty} 2^{1 / 2 n} \lim _{n \rightarrow \infty}\left[n^{1 / n}\right]^{1 / 2}=p^{1 / 2}
$$

with (real) solutions $p=0$ and $p=1$. From part (i), only the latter is valid.
(l) Suppose $\left\{a_{n}\right\},\left\{b_{n}\right\} \in \mathbb{R}$ are convergent sequences with limits $a$ and $b$, respectively. Define sequence $M_{n}=\max \left(a_{n}, b_{n}\right)$. Prove that $\lim M_{n}=\max (a, b)$.
ANS: I consider the three cases $a=b=L, a>b$, and $a<b$.

- If $a=b=L$, then $a_{n} \rightarrow L$ and $b_{n} \rightarrow L$, so that, for any $\epsilon>0, \exists n_{a} \in \mathbb{N}$ such that, $\forall n \geq n_{a}, a_{n} \in N_{\epsilon}(L)$; and $\exists n_{b} \in \mathbb{N}$ such that, $\forall n \geq n_{b}, b_{n} \in N_{\epsilon}(L)$. Set $N=\max \left(n_{a}, n_{b}\right)$. Then, $\forall n \geq N, \max \left(a_{n}, b_{n}\right) \in N_{\epsilon}(L)$.
- (Petrovic, \#2.3.9) Suppose $a>b$, so $\max (a, b)=a$. Let $\epsilon=(a-b) / 2>0$. As $\left\{a_{n}\right\}$ and $\left\{b_{n}\right\}$ are convergent, $\exists N_{1} \in \mathbb{N}$ such that, $\forall n \geq N_{1},\left|a_{n}-a\right|<\epsilon$; and $\exists N_{2} \in \mathbb{N}$ such that, $\forall n \geq N_{2}$, $\left|b_{n}-b\right|<\epsilon$. Let $N=\max \left(N_{1}, N_{2}\right)$. For $n \geq N$,

$$
\begin{aligned}
a-a_{n} & \leq\left|a_{n}-a\right|<\epsilon=(a-b) / 2 \Rightarrow a_{n}>a-(a-b) / 2=(a+b) / 2 \\
b_{n}-b & \leq\left|b_{n}-b\right|<\epsilon=(a-b) / 2 \Rightarrow b_{n}<b+(a-b) / 2=(a+b) / 2 .
\end{aligned}
$$

Thus, $a_{n}>b_{n}$, and $M_{n}=\max \left(a_{n}, b_{n}\right)=a_{n} \rightarrow a=\max (a, b)$. The case of $a<b$ is similar.

- (Chen Lan) Note that $\max (x, y)=\frac{1}{2}(x+y+|x-y|)$. Let $M:=\max (a, b)=\frac{1}{2}(a+b+|a-b|)$. $a_{n} \rightarrow a, b_{n} \rightarrow b \Rightarrow \forall \epsilon>0, \exists N_{1}, N_{2} \in \mathbb{N}$ such that $\forall n \geq N_{1},\left|a_{n}-a\right|<\epsilon ;$ and $\forall n \geq N_{2}$, $\left|b_{n}-b\right|<\epsilon$. Let $N=\max \left(N_{1}, N_{2}\right)$. Using the triangle and reverse triangle inequalities, for $n \geq N$,

$$
\begin{aligned}
2\left|M_{n}-M\right| & =\left|\left(a_{n}+b_{n}+\left|a_{n}-b_{n}\right|\right)-(a+b+|a-b|)\right| \\
& =\left|\left(a_{n}-a\right)+\left(b_{n}-b\right)+\left(\left|a_{n}-b_{n}\right|-|a-b|\right)\right| \\
& \leq\left|a_{n}-a\right|+\left|b_{n}-b\right|+\left|\left|a_{n}-b_{n}\right|-|a-b|\right| \\
& \leq\left|a_{n}-a\right|+\left|b_{n}-b\right|+\left|\left(a_{n}-b_{n}\right)-(a-b)\right| \\
& =\left|a_{n}-a\right|+\left|b_{n}-b\right|+\left|\left(a_{n}-a\right)-\left(b_{n}-b\right)\right| \\
& \leq\left|a_{n}-a\right|+\left|b_{n}-b\right|+\left|a_{n}-a\right|+\left|b_{n}-b\right| \\
& <4 \epsilon .
\end{aligned}
$$

Thus, $M_{n} \rightarrow M$.
(m) First state and prove the "Comparison Lemma" (Fitzpatrick p. 28). Next prove the equivalence (that means, $\Leftrightarrow$ ) of the Comparison Lemma and the Squeeze Theorem.
ANS:
Comparison Lemma: Let sequence $\left\{a_{n}\right\}$ converge to $a$. Sequence $\left\{b_{n}\right\}$ converges to $b$ if there is a $C \geq 0$ and $N_{1} \in \mathbb{N}$ such that $\left|b_{n}-b\right| \leq C\left|a_{n}-a\right|, \forall n \geq N_{1}$.
Proof: Let $\epsilon>0$, and assume $C>0$. Choose $N_{2}$ such that $\forall n \geq N_{2},\left|a_{n}-a\right|<\epsilon / C$. Let $N=\max \left(N_{1}, N_{2}\right)$. Then $\forall n>N,\left|b_{n}-b\right| \leq C\left|a_{n}-a\right|<C \cdot \epsilon / C=\epsilon$.
Comparison $\Rightarrow$ Squeeze:
Suppose $\left\{a_{n}\right\},\left\{b_{n}\right\}$, and $\left\{c_{n}\right\}$ are sequences of real numbers for which there exists $n_{o} \in \mathbb{N}$ such that $a_{n} \leq b_{n} \leq c_{n}$ for all $n \in \mathbb{N}, n \geq n_{o}$, and that $\lim _{n \rightarrow \infty} a_{n}=\lim _{n \rightarrow \infty} c_{n}=L$. Then,

$$
\forall n \geq n_{o}, a_{n}-L \leq b_{n}-L \leq c_{n}-L \leq C\left|c_{n}-L\right|, \text { where } C=1
$$

Similarly, multiplying by -1 , we obtain

$$
\forall n \geq n_{o}, L-c_{n} \leq L-b_{n} \leq L-a_{n} \leq C\left|a_{n}-L\right|, \text { where } C=1
$$

Thus, $\left|b_{n}-L\right| \leq C \max \left(\left|a_{n}-L\right|,\left|c_{n}-L\right|\right)$. From the previous question, the rhs is a convergent sequence. Thus, the Comparison Lemma implies that $b_{n} \rightarrow b$.
Squeeze $\Rightarrow$ Comparison:
Let sequence $\left\{a_{n}\right\}$ converge to $a$. Assume $\exists C \geq 0$ and $\exists N_{1} \in \mathbb{N}$ such that, $\forall n \geq N_{1},\left|b_{n}-b\right| \leq$ $C\left|a_{n}-a\right|$. Then

$$
\forall n \geq N_{1}, \quad-C\left|a_{n}-a\right| \leq b_{n}-b \leq C\left|a_{n}-a\right|
$$

As $a_{n} \rightarrow a,\left|a_{n}-a\right| \rightarrow 0$, and the Squeeze Theorem implies $b_{n}-b \rightarrow 0$, or $b_{n} \rightarrow b$.
(n) A null sequence is any real-valued sequence $\left\{h_{n}\right\}_{n=0}^{\infty}$ that converges to 0 as $n \rightarrow \infty$. That is, for any $\epsilon>0, \exists n_{0}=n_{0}(\epsilon)$ such that $\left|h_{n}\right|<\epsilon$ for $n \geq n_{0}$. Examples of positive such sequences include $h_{k}=1 / k$ and $h_{k}=1 / 2^{k}$.
Prove: Let $\left\{\epsilon_{n}\right\}_{n=0}^{\infty}$ be a null sequence of positive numbers. Sequence $\left\{a_{n}\right\}_{n=0}^{\infty}$ converges to $\ell$ iff $\forall k, \exists n_{k}$ such that $\left|a_{n}-\ell\right|<\epsilon_{k}$ for $n \geq n_{k}$. (Garling, Prop 3.2.8)
ANS:
Necessary $(\Leftarrow)$ : As $\left\{\epsilon_{n}\right\}_{n=0}^{\infty}$ is a null sequence of positive numbers, for any given $\epsilon, \exists k \in \mathbb{N}$ such $\overline{\text { that } 0<\epsilon_{k}<\epsilon}$. The condition then implies that, for $n>n_{k},\left|a_{n}-\ell\right|<\epsilon_{k}<\epsilon$, this being the definition of sequence convergence.

Sufficient $(\Rightarrow)$ : Assume sequence $\left\{a_{n}\right\}_{n=0}^{\infty}$ converge to $\ell$. Let $\epsilon_{1}>0$ be given. Then, as $\ell$ is the limit point of $\left\{a_{n}\right\}, \exists n_{1} \in \mathbb{N}$ such that $\left|a_{n}-\ell\right|<\epsilon_{1}, \forall n \geq n_{1}$. Similarly, given $\epsilon_{2}$ such that $0<\epsilon_{2}<\epsilon_{1}, \exists n_{2} \in \mathbb{N}, n_{2}>n_{1}$, such that $\left|a_{n}-\ell\right|<\epsilon_{2}, \forall n \geq n_{2}$. Continuing, we obtain, as required, a strictly increasing sequence $\left\{n_{k}\right\} \in \mathbb{N}$ and strictly decreasing sequence $\left\{\epsilon_{k}\right\} \in \mathbb{R}_{>0}$ such that, for each $k \in \mathbb{N},\left|a_{n}-\ell\right|<\epsilon_{k}, \forall n \geq n_{k}$.
19. Recall the "trivial", or "discrete" metric, namely: Let $X$ be any nonempty set. For $p, q \in X$, $d(p, q)=1$ for $p \neq q$, and $d(p, q)=0$ if $p=q$. Let $E \subset X$ be arbitrary.
(a) Show that every element of $E$ is an interior point of $E$; every point not contained in $E$ is an exterior point of $E$; and there are no boundary points, i.e., $\partial E=\emptyset$.
(Tao II, Exercise 1.2.1 for his Example 1.2.8)
ANS: Recall (e.g., Tao II, Def 1.2.5): Let $(X, d)$ be a metric space, let $E$ be a subset of $X$, and let $x_{0}$ be a point in $X$. We say that $x_{0}$ is an interior point of $E$ if there exists a radius $r>0$ such that $B_{r}\left(x_{0}\right) \subset E$.
For the discrete metric on $X$, and $E$ any subset of $X$ : for $x_{0} \in E$ and $0<r \leq 1, B_{r}\left(x_{0}\right)=$ $\left\{x_{0}\right\} \subset E$. So, by definition, every element of $E$ is an interior point of $E$.
To show $\partial E=\emptyset$, assume $b \in \partial E$ and $0<r \leq 1$. Then $B_{r}(b)=\{b\}$, and either $b \in E$ or $b \in E^{c}$, and thus $b$ cannot be a boundary point. As every element of $E$ is an interior point of $E$, and $\partial E=\emptyset, x \in E^{c}$ must be an exterior point of $E$. See also (11).
(b) Show that all subsets of $X$ are both open and closed.

ANS: Recall: (e.g., Tao II, Def 1.2.12): Subset $E \subset X$ is open if it contains none of its boundary points. Equivalently (e.g., Stoll, Def $2.2 .5(\mathrm{a})) E$ is open if every point of $E$ is an interior point of $E$.
Both definitions, along with the previous results, show that $E$ is open. As $E$ was any subset of $X$, all subsets of $X$ are open. We can use a third definition of open, namely that from question 12: First, let $E$ be a proper subset of $X$. Then, $\forall u \in E, d(u, X \backslash E)=1>0$. For $E=X$, $X \backslash X=\emptyset$, and so, by definition, $d(x, \emptyset)=+\infty>0$.
Now we turn to closed. As $E^{c} \in X, E^{c}$ is open, which by definition implies $E$ is closed. Thus, all subsets of $X$ are both open and closed.
(c) For $E \subset X$, characterize the limit and isolated points of $E$.

ANS: Recalling the definitions of limit and isolated points, and by the above result that, for $x_{0} \in E$ and $0<r \leq 1, B_{r}\left(x_{0}\right)=\left\{x_{0}\right\}$, we see that $E$ has no limit points, and all its points are isolated.
Notice that, from the result (e.g., Stoll, Thm 2.2.14): A subset $F$ of a metric space $X$ is closed if and only if $F$ contains all its limit points, we see that this is vacuously fulfilled, and thus $E$ is closed. As $E^{c}$ is also a subset of $X$ and is thus closed, $E$ is open, so that, as above, all subsets of $X$ are both open and closed.
(d) For $X=\mathbb{R}, a, b \in \mathbb{R}$ with $a<b$, let $E=[a, b]$. Determine if $E$ is compact.

ANS: Let $\mathcal{O}=\{\{x\}: x \in[a, b]\}$ be an (uncountable) cover of $E$. Clearly, $\mathcal{O}$ cannot admit a finite subcover of $E$, so $E$ is not compact. ${ }^{8}$

[^6]20. (Cesàro summation) For sequence $\left\{a_{k}\right\}_{k=1}^{\infty} \in \mathbb{R}$, let $s_{n}=n^{-1} \sum_{k=1}^{n} a_{k}, n \in \mathbb{N}$. (Parts (a) and (b) are from Stoll, p. 95, \#14)
(a) Prove that, if $\lim _{k \rightarrow \infty} a_{k}=a$, then $\lim _{n \rightarrow \infty} s_{n}=a$.

ANS: This is detailed in, e.g., Goldberg (1976, Sec. 2.11, 3.9), with proof in his Thm 2.11B. From https://math.stackexchange.com/questions/155839:
Given $\epsilon>0$ there exists $n_{0}$ such that, if $n \geq n_{0}$, then $\left\|a_{n}-a\right\|<\epsilon$. Then

$$
\begin{aligned}
0 & \leq\left\|s_{n}-a\right\| \leq\left\|\frac{a_{1}+\cdots+a_{n}-n a}{n}\right\| \\
& \leq \frac{\left\|a_{1}-a\right\|}{n}+\cdots+\frac{\left\|a_{n_{0}-1}-a\right\|}{n}+\frac{\left\|a_{n_{0}}-a\right\|}{n}+\cdots+\frac{\left\|a_{n}-a\right\|}{n} \\
& \leq \frac{1}{n} \sum_{i=1}^{n_{0}-1}\left\|a_{i}-a\right\|+\frac{n-n_{0}+1}{n} \epsilon
\end{aligned}
$$

As $\left\{a_{k}\right\}_{k=1}^{\infty}$ is convergent, all its terms are bounded. Thus, the first $n_{0}-1$ terms $\left\|a_{n_{0}}-a\right\|$ are bounded, by say $M$. Let $N_{0}=\left(n_{0}-1\right) M / \epsilon$. For $n \geq N_{0}$, we have

$$
\frac{1}{n} \sum_{i=1}^{n_{0}-1}\left\|a_{n}-a\right\| \leq \frac{1}{n}\left(n_{0}-1\right) M \leq \epsilon
$$

Thus, for $n \geq N_{0},\left\|s_{n}-a\right\|<2 \epsilon$.
(b) Give an example of a sequence $\left\{a_{k}\right\}$ which diverges, but for which $\left\{s_{n}\right\}$ converges. Hint: Recall diverge means, does not converge, and does not necessarily mean the sequence tends to infinity. ANS: Let $a_{k}=(-1)^{k}$.
(c) Consider taking $\left\{a_{k}\right\}$ to be the harmonic series, i.e., $a_{k}=\sum_{j=1}^{k} 1 / j$, which diverges. Is $\left\{s_{n}\right\}$ convergent?
ANS: Here,

$$
s_{n}=n^{-1} \sum_{k=1}^{n} a_{k}=\sum_{k=1}^{n} \sum_{j=1}^{k} \frac{1}{n j},
$$

i.e., the sum of terms in the following tabular representation, with vertical entries corresponding to $k=1,2, \ldots, n$, and horizonal corresponding to $j=1,2, \ldots, k$ :

$$
\begin{array}{cccccc}
1 / n & & & & & \\
1 / n & 1 / 2 n & & & & \\
1 / n & 1 / 2 n & 1 / 3 n & & & \\
1 / n & 1 / 2 n & 1 / 3 n & 1 / 4 n & & \\
\vdots & \vdots & \vdots & \vdots & & \\
1 / n & 1 / 2 n & 1 / 3 n & 1 / 4 n & \ldots & 1 / n^{2}
\end{array}
$$

The column sums are $1,(n-1) / 2 n,(n-2) / 3 n, \ldots, 1 / n^{2}$, which sum to

$$
s_{n}=\sum_{h=1}^{n} \frac{n-h+1}{h n}=-1+\left(\frac{n+1}{n}\right) \sum_{h=1}^{n} \frac{1}{h},
$$

which clearly diverges.
21. Let $x_{n}=1 /(n+1)+1 /(n+2)+\cdots+1 /(2 n)$, a sum of $n$ terms, i.e., $x_{1}=1 / 2, x_{2}=1 / 3+1 / 4=$ $7 / 12>x_{1}$, etc.. Show $\left\{x_{n}\right\}$ is (strictly) monotone increasing and bounded, and thus has a limit. (Q\&A from https://math.stackexchange.com/questions/525749)

ANS: Bounded follows because $x_{n} \leq n /(n+1) \leq 1$. Monotone follows because

$$
x_{n+1}-x_{n}=\frac{1}{2 n+2}+\frac{1}{2 n+1}-\frac{1}{n+1}=\frac{1}{2(2 n+1)(n+1)}>0 .
$$

Passing to the limit of a Riemann integral seems to be the fastest and easiest way to determine the limiting sum. From the aforementioned webpage,

$$
x_{n}=\sum_{k=1}^{n} \frac{1}{n+k}=\sum_{k=1}^{n} \frac{1 / n}{1+k / n} \rightarrow \int_{0}^{1} \frac{d x}{1+x}=\log 2 .
$$

22. Recursive expressions for computing roots of positive numbers.

NOTES: Stoll, p. 101, \#6, poses this as follows: Let $\alpha>0$. Choose $x_{1}>\sqrt{\alpha}$. For $n \in \mathbb{N}$, define $x_{n+1}=\left(x_{n}+\alpha / x_{n}\right) / 2$. He then asks: (a) Show that $x_{n}$ is monotone and bounded. (b) Prove that $\lim _{n \rightarrow \infty} x_{n}=\sqrt{\alpha}$. (c) Prove that $0 \leq x_{n}-\sqrt{\alpha} \leq\left(x_{n}^{2}-\alpha\right) / x_{n}$.
It appears from " $x_{1}>\sqrt{\alpha}$ " that we require knowledge of the answer, but note that, for $\alpha>1$, $\sqrt{\alpha}<\alpha$, so one could choose $x_{1}=\alpha$. For $\alpha<1$, we can choose $x_{1}=1$.
This recursion is in fact Newton-Raphson for solving the equation $x^{2}=\alpha$. It can be proven to have quadratic convergence, which means, with each step, the number of correct decimals is about doubled. This result is proven in many textbooks, including Stoll, Example 5.4.1.
See also https://math.stackexchange.com/questions/786453/.
The first question just involves developing a computer program to implement this square-root recursion. The next part considers the more general $k$ th root, $k \in \mathbb{N}, k \geq 2$.
(a) First write a simple computer program to compute the recursion, just for the $k=2$ case, stopping after obtaining six significant digits after the decimal point, and show the resulting $x_{i}$. Choose $\alpha=400$ and start with $x_{1}=400$.
ANS: In Matlab, this is accomplished with

```
function theroot = mysqrt(a)
    a=abs(a); maxiter=20; tol=1e-6; X=zeros(maxiter,1); X(1)=a;
    notyet=1; loop=2;
    while notyet && (loop<maxiter)
        x = X(loop-1); X(loop)=(x+a/x)/2;
        notyet = abs(X(loop)-x) > tol; loop=loop+1;
    end
    disp(X(1:(loop-1))), theroot=X(loop-1);
```

The output is
400.0000, 200.5000, 101.2475, 52.5991, 30.1019, 21.6950, 20.0662, 20.0001, 20.0000, 20.0000
(b) Below, we will crucially require one of the variants of Bernoulli's inequality. One is given in (4). We require the following version: Suppose $0<h<1$ and $n \in \mathbb{N}$. Then

$$
\begin{equation*}
(1-n h) \leq(1-h)^{n} . \tag{14}
\end{equation*}
$$

Prove this result using induction.
ANS: (Garling p. 74) The result is clearly true for $n=1$. Suppose it is true for $n$. Then

$$
\begin{aligned}
(1-h)^{n+1} & =(1-h)(1-h)^{n} \\
& \geq(1-h)(1-n h)=1-(n+1) h+n h^{2}>1-(n+1) h
\end{aligned}
$$

(c) Recursions for computing the $k$ th root are given in, e.g., Jacob and Evans, Vol. I, p. 241, \#8; and Garling, Vol. I, p. 88. The expression in Garling reduces to the above when $k=2$.
To proceed, we need to know that, for $y$ a positive real number, and $k \in \mathbb{N}, k \geq 2$, there exists a unique positive real number $s$ such that $s^{k}=y$, written as $s=y^{1 / k}$. Ghorpade and Limaye have a proof, as does Garling, Thm 6.4.6.
From Garling, for $y>0$, the recursion to obtain $s=y^{1 / k}$ is given as follows. Let $a_{0}=\max (1, y)$, so $a_{0}^{k} \geq y$. Define sequence $\left\{a_{n}\right\}$ recursively as

$$
a_{n+1}=\frac{1}{k}\left((k-1) a_{n}-\frac{y}{a_{n}^{k-1}}\right)=a_{n}\left(1-\frac{a_{n}^{k}-y}{k a_{n}^{k}}\right), \quad n \in \mathbb{N} .
$$

Observe this reduces to the above equation when $k=2$. We wish to show $\left\{a_{n}\right\}$ is convergent, with limit $s=y^{1 / k}$. To do this, prove:
i. $\left\{a_{n}\right\}$ is monotone decreasing, i.e., show $0<a_{a+1} \leq a_{n}$,
ii. $\left\{a_{n}\right\}$ is bounded below by $s=y^{1 / k}$, i.e., show $y \leq a_{n}^{k}$, and
iii. the limit of the sequence is unique.

NOTE 1: In the below proof, Garling says " $a_{n+1}$ is properly defined because $a_{n}>0$ ". Under the below stated assumptions that, for $n \geq 1,0<a_{n} \leq a_{n-1}$ and $y \leq a_{n}^{k}$, then from the lhs of the above recursion, and the assumption of $k \geq 2$, we need to confirm that

$$
(k-1) a_{n}>\frac{y}{a_{n}^{k-1}} \text { or } k-1>\frac{y}{a_{n}^{k}},
$$

which is, indeed with $k \geq 2$, true.
NOTE 2: In the below proof, we need to confirm the above rhs fraction is less than one, i.e., that $a_{n}^{k}-y<k a_{n}^{k}$. By assumption, $y \leq a_{n}^{k}$, or $a_{n}^{k}-y>0$, so, indeed for $k \geq 2$,

$$
a_{n}^{k}-y \leq(k-1) a_{n}^{k}-y=k a_{n}^{k}-\left(a_{n}^{k}+y\right)<k a_{n}^{k} .
$$

NOTE 3: We require this below: From the product rule for convergent sequences, if $a_{n+1}$ converges, then $a_{n+1}^{k}$ also converges, and converges to $\ell^{k}$. As $a_{n+1}^{k} \geq y$, it follows from (13) that $\ell^{k} \geq y$.

ANS: (Garling p. 88) Suppose we have defined $a_{n}$, and shown that, for $n \geq 1,0<a_{n} \leq a_{n-1}$ and $y \leq a_{n}^{k}$. Then $a_{n+1}$ is properly defined because $a_{n}>0$ (see my NOTE 1 above). Next, as $k a_{n}^{k}>a_{n}^{k}-y \geq 0$ (see my NOTE 2 above), we have $0<a_{n+1} \leq a_{n}$. In order to show that $a_{n+1}^{k} \geq y$, we use Bernoulli's inequality in (14) above: For $0<t<1,(1-t)^{k} \geq 1-k t$. Thus

$$
a_{n+1}^{k}=a_{n}^{k}\left(1-\frac{a_{n}^{k}-y}{k a_{n}^{k}}\right)^{k} \geq a_{n}^{k}\left(1-\frac{a_{n}^{k}-y}{a_{n}^{k}}\right)=y .
$$

As $\left\{a_{n}\right\}_{n=0}^{\infty}$ is monotone decreasing and bounded below, it converges to a limit $\ell$ as $n \rightarrow \infty$. Thus, (see my NOTE 3 above) $\ell^{k} \geq y$. As $y>0, \ell>0$, and thus, from the basic produce and quotient properties of convergent sequences, as $n \rightarrow \infty$,

$$
a_{n+1}=a_{n}\left(1-\frac{a_{n}^{k}-y}{k a_{n}^{k}}\right) \rightarrow \ell\left(1-\frac{\ell^{k}-y}{k \ell^{k}}\right) .
$$

Clearly, as $\left\{a_{n}\right\}$ is convergent with $\lim a_{n}=\ell, \lim a_{n+1}=\lim a_{n}=\ell$, so

$$
\ell=\ell\left(1-\frac{\ell^{k}-y}{k \ell^{k}}\right),
$$

so that $\ell^{k}=y$.
Finally, if $\ell^{k}=m^{k}$, then, from the difference of powers formula (Fitzpatrick, p. 18), $0=$ $\ell^{k}-m^{k}=(\ell-m)\left(\ell^{k-1}+\ell^{k-2} m+\cdots+m^{k-1}\right)$, so that $\ell=m$.
(d) A recursion for the square root based on continued fractions is nicely developed in Blyth and Robertson, Basic Linear Algebra (2nd edition, 2002, second printing, 2005). Another place to find this is in Elaydi, An Introduction to Difference Equations, 2005, p. 408.
For $\sqrt{g}, g>0$, the continued fraction is, when expressed as a sequence $\left\{t_{n}\right\}, t_{1}=2, t_{n+1}=$ $2-a / t_{n}$, where $a=1-g$; this converging to $1+\sqrt{g}$. Program this such that it iterates until the solution is within some fixed tolerance, and plot the required number of iterations, say $r=r(g)$, as a function of $g$. Show that $r$ increases monotonically with $g$.
When I did it, for $g=2$ and a tolerance of $10^{-15}$, i.e., approximate (usual) machine precision, we require $r=21$. With $g=200, r=262$.
As the tolerance gets smaller, the algorithm fails to converge (presumably due to round off error). For example, with the tolerance of $10^{-15}$, the algorithm fails for $g \geq 227$. With a tolerance of (the still quite respectable) $10^{-12}$, the algorithm works for $g=30,000$, but fails if the tolerance is changed to $10^{-13}$.
To enable usage for all positive numbers (that can be represented in a computer), and also keep the required number of iterations as low as possible, make an algorithm to determine the smallest required $g^{*}$ based on integer square roots, e.g., if $100<g<121$, then, letting $b=100=10^{2}, g^{*}=g / 100$, and we compute $\sqrt{b \times g / b}=10 \sqrt{g^{*}}$. Trivially, $(n+1)^{2} / n^{2} \rightarrow 1$ as $n \rightarrow \infty$, so that $g^{*}$ is bounded and approaches 1 , and thus this method is valid for all positive reals, and ensures that calculation of $\sqrt{g^{*}}$ to machine precision is reliable and fast. However, the determination of base $b$ would presumably be the bottleneck in the algorithm.
Program this method, illustrate the output, and comment on the performance compared to the above method of computing square roots.
(e) Blyth and Robertson, Example 9.12, analyze yet another iterative example for computing $\sqrt{2}$ : Let $q_{1}$ a positive rational number, and set $q_{2}=\left(2+q_{1}\right) /\left(1+q_{1}\right)$. Continue iterating on this recursion. Numerically confirm it works.
23. Compactness
(a) Show that a finite union of compact sets in $\mathbb{R}^{n}$ is compact.

ANS: It suffices to show that, for compact $K_{1}, K_{2} \in \mathbb{R}^{n}, K:=K_{1} \cup K_{2}$ is compact. From Heine-Borel, each $K_{i}$ is closed and bounded. It is easy to see that $K$ is bounded. A finite union of closed sets is closed, so $K$ is closed and bounded and thus compact. The general result follows by induction.
(b) Show that a finite union of compact sets in general metric space $X$ is compact.

ANS: As in the previous question, it suffices to show for two sets. Let $K_{1}$ and $K_{2}$ be compact subsets of $X$, and let $K:=K_{1} \cup K_{2}$. Let $O=\left\{O_{i}\right\}$ be any open cover of $K$. As $K_{i} \subset K$, $i=1,2, O$ is an open cover of both $K_{1}$ and $K_{2}$, so that there exists finite collections of sets in $O$, say $O_{1,1}, \ldots O_{1, k_{1}}$ and $O_{2,1}, \ldots O_{2, k_{2}}$ that are finite subcovers of $K_{1}$ and $K_{2}$, respectively. Thus, $\left(\cup_{i} O_{1, k_{i}}\right) \cup\left(\cup_{j} O_{2, k_{j}}\right)$ is a finite subcover of $K=K_{1} \cup K_{2}$, showing $K$ is compact.
24. (lim inf and lim sup)
(a) State the definitions, from Stoll, of lim inf and lim sup, for sequences of real numbers.

ANS: For sequence $\left\{s_{n}\right\} \in \mathbb{R}$, let $a_{k}=\inf \left\{s_{n}: n \geq k\right\}$ and $b_{k}=\sup \left\{s_{n}: n \geq k\right\}$, for each $k \in \mathbb{N}$. Note, from the definitions of inf and sup, $a_{k} \leq b_{k}, a_{k} \leq a_{k+1}$, and $b_{k} \geq b_{k+1}$. From the latter two inequalities, $\left\{a_{k}\right\}$ and $\left\{b_{k}\right\}$ have limits in the extended real line, and limits in $\mathbb{R}$ if $\left\{s_{n}\right\}$ is bounded.
The limit inferior and limit superior of $\left\{s_{n}\right\}$ are, respectively,

$$
\begin{aligned}
& \liminf s_{n}:=\lim _{k \rightarrow \infty} a_{k} \\
&=\sup _{k \in \mathbb{N}} \inf \left\{s_{n}: n \geq k\right\} \\
& \limsup s_{n}:=\lim _{k \rightarrow \infty} b_{k}
\end{aligned}=\inf _{k \in \mathbb{N}} \sup \left\{s_{n}: n \geq k\right\} . ~ \$
$$

(b) For sequence $\left\{s_{n}\right\}$, let $\beta=\limsup s_{n} \in \mathbb{R}$. State the two results that pertain to $s_{n}<\beta+\epsilon$ and $s_{k}>\beta-\epsilon$.
ANS: This is, e.g., Stoll, Thm 3.5.3(a).
(i) $\exists n_{o} \in \mathbb{N}$ such that $\forall n \geq n_{o}, s_{n}<\beta+\epsilon$. In words, eventually, all $s_{n}<\beta+\epsilon$.
(ii) Given $n \in \mathbb{N}, \exists k \in \mathbb{N}$ with $k \geq n$ such that $s_{k}>\beta-\epsilon$. In words, $s_{n}>\beta-\epsilon$ for infinitely many $n$.
(c) A common alternative definition (e.g., Petrovic, Def 2.8.9) of lim inf (lim sup) is:

If $\left\{a_{n}\right\}$ is a bounded sequence, then $\lim \inf \left\{a_{n}\right\}\left(\lim \sup \left\{a_{n}\right\}\right)$ is the smallest (largest) accumulation point of $\left\{a_{n}\right\}$.

In turn, we require:
Point $c$ is an accumulation point of sequence $\left\{a_{n}\right\}$ if, for convergent subsequence $\left\{a_{n_{k}}\right\}$, $\lim a_{n_{k}}=c$.
Their equivalence is shown in Stoll, Thm 3.5.7, but can also be found (and I prefer) in Jacob and Evans, Thm 19.22A; and Terrell, Thms 3.10.3 and 3.10.4.
Assume the equivalence of these two formulations. Prove: A sequence $\left\{a_{n}\right\}$ converges to $a$ iff $\liminf \left\{a_{n}\right\}=\limsup \left\{a_{n}\right\}=a$.
ANS: Almost trivially, if $\left\{a_{n}\right\}$ converges to $a$, then $\left\{a_{n}\right\}$ has only $a$ as its accumulation point.
And, going the other way, if $\lim \inf \left\{a_{n}\right\}=\lim \sup \left\{a_{n}\right\}=a$, then, these being the smallest and largest accumulation points, $a$ is the only accumulation point of $\left\{a_{n}\right\}$, and thus the limit of $\left\{a_{n}\right\}$.
(d) Let $\left\{a_{n}\right\},\left\{b_{n}\right\} \in \mathbb{R}_{\geq 0}$. We consider the relation

$$
\begin{equation*}
\limsup \left\{a_{n} b_{n}\right\} \leq \lim \sup \left\{a_{n}\right\} \lim \sup \left\{b_{n}\right\} \tag{15}
\end{equation*}
$$

(Found in nearly all books, e.g., Stoll, p. 112, \#7(a); Petrovic, p. 50, \#2.8.9(a); Junghenn, Prop 2.4.1; etc.)
i. Give an example for which (15) is strict.

ANS: For an example in which (15) is strict: Take $a_{n}=\left((-1)^{n+1}+1\right) / 2$, i.e., $\left\{a_{n}\right\}=$ $\{1,0,1,0, \ldots\}$; and $b_{n}=\left((-1)^{n}+1\right) / 2$, i.e., $\left\{b_{n}\right\}=\{0,1,0,1, \ldots\}$. This results in a strict inequality.
ii. Prove (15).

ANS: We provide two different forms of proof.

- Proof I: We first show the anyway useful

$$
\begin{equation*}
\sup _{k \geq n} a_{k} b_{k} \leq \sup _{k \geq n} a_{k} \sup _{k \geq n} b_{k} . \tag{16}
\end{equation*}
$$

For each $n \in \mathbb{N}, a_{n} \leq \sup _{k \geq n} a_{k}$ and $b_{n} \leq \sup _{k \geq n} b_{k}$, so that, as $\left\{a_{n}\right\},\left\{b_{n}\right\} \in \mathbb{R}_{\geq 0}, a_{n} b_{n} \leq$ $\left(\sup _{k \geq n} a_{k}\right)\left(\sup _{k \geq n} b_{k}\right)$. As the rhs is an upper bound, (16) follows.
For the main result: Let $c_{k}=a_{k} b_{k}$ and $\bar{c}_{n}=\sup _{k \geq n} c_{k}, \bar{a}_{n}=\sup _{k \geq n} a_{k}$, and $\bar{b}_{n}=\sup _{k \geq n} b_{k}$. Then (16) reads, for sequences $\left\{\bar{c}_{n}\right\},\left\{\bar{a}_{n}\right\}$, and $\left\{\bar{b}_{n}\right\}$,

$$
\begin{equation*}
\forall n \in \mathbb{N}, \quad \bar{c}_{n} \leq \bar{a}_{n} \bar{b}_{n} \tag{17}
\end{equation*}
$$

Then, from (17), the result in exercise 17, and limit of product is product of limits,

$$
\begin{aligned}
\limsup \left\{a_{n} b_{n}\right\}=\lim _{n \rightarrow \infty} \sup _{k \geq n} a_{k} b_{k}=\lim _{n \rightarrow \infty} \bar{c}_{n} & \leq \lim _{n \rightarrow \infty} \bar{a}_{n} \bar{b}_{n} \\
& =\lim _{n \rightarrow \infty} \sup _{k \geq n} a_{k} \sup _{k \geq n} b_{k}=\lim \sup \left\{a_{n}\right\} \lim \sup \left\{b_{n}\right\}
\end{aligned}
$$

which is (15).

## - Proof II:

Let $L_{1}=\lim \sup a_{n}, L_{2}=\limsup b_{n}$. We will make use of the aforementioned standard result for lim sup, namely Stoll, Theorem 3.5.3(a), which we reproduce here again:

Let $\left\{s_{n}\right\}_{n=1}^{\infty}$ be a sequence in $\mathbb{R}$. Suppose $\lim \sup s_{n} \in \mathbb{R}$. Then

$$
\beta=\limsup s_{n} \Leftrightarrow \forall \epsilon>0, \text { (i) and (ii) hold, }
$$

where
(3.5.3i) $\exists n_{o} \in \mathbb{N}$ such that, $\forall n \geq n_{o}, s_{n}<\beta+\epsilon$.
(3.5.3ii) Given $n \in \mathbb{N}, \exists k \in \mathbb{N}$ with $k \geq n$ such that $s_{k}>\beta-\epsilon$.

We distinguish the following four cases:

1. $L_{1}>0$ and $L_{2}>0$
2. $L_{1}=L_{2}=0$
3. $L_{1}=0$ and $L_{2}>0$ (or vice versa)
4. one or both of $L_{1}, L_{2}$, are infinite

Case 1: $L_{1}>0$ and $L_{2}>0$ (As in Petrovic, solution to his exercise 2.8.9)
Fix $\epsilon>0$. From (3.5.3i), $\exists N_{1} \in \mathbb{N}$ such that, $\forall n \geq N_{1}, a_{n}<L_{1}(1+\epsilon)$; and $\exists N_{2} \in \mathbb{N}$ such that, $\forall n \geq N_{2}, b_{n}<L_{2}(1+\epsilon)$. Set $N:=\max \left(N_{1}, N_{2}\right)$, so that $a_{n} b_{n}<L_{1} L_{2}(1+\epsilon)^{2}$. Thus, $\lim \sup \left\{a_{n} b_{n}\right\} \leq L_{1} L_{2}(1+\epsilon)^{2}$ and, this being true for any $\epsilon>0, \lim \sup \left\{a_{n} b_{n}\right\} \leq L_{1} L_{2}$.
Case 2: $L_{1}=0$ and $L_{2}=0$
Clearly, the rhs of (15) is 0 . To show the lhs is 0 , we need to show (3.5.3i) and (3.5.3ii) for sequence $\left\{a_{n} b_{n}\right\}$. The latter is trivially fulfilled, as both $\left\{a_{n}\right\}$ and $\left\{b_{n}\right\}$ are nonnegative. To see (3.5.3i), apply (3.5.3i) to both $\left\{a_{n}\right\}$ and $\left\{b_{n}\right\}: L_{1}=L_{2}=0$ implies that, for a fixed $\epsilon>0, \exists n_{1}, n_{2} \in \mathbb{N}$ such that, with $n_{0}=\max \left\{n_{1}, n_{2}\right\}, \forall n \geq n_{0}, a_{n}<\sqrt{\epsilon}$ and $b_{n}<\sqrt{\epsilon}$, i.e., $\forall n \geq n_{0}, a_{n} b_{n}<\epsilon$. Thus, $\lim \sup \left(a_{n} b_{n}\right)=0$ and the inequality holds.
Case 3: $L_{1}=0$ and $L_{2}>0$ (or vice versa)
Assume that $\left\{b_{n}\right\}$ is bounded, ${ }^{9}$ i.e., $\exists M>0$ such that $\forall n \in \mathbb{N}, b_{n}<M$.

[^7]In that case, as $L_{2}$ is the largest subsequential limit of $\left\{b_{n}\right\}, L_{2} \leq M$, and the rhs of (15) equals 0 . To see that its lhs is also zero, we require (as in case 2 above), that, for all $\epsilon>0$, there exists $n_{o} \in \mathbb{N}$ such that $a_{n} b_{n}<\epsilon$ for all $n \geq n_{o}$. From (3.5.3i) applied to $\left\{a_{n}\right\}$, for a fixed $\epsilon>0, \exists n_{1} \in \mathbb{N}$ such that, $\forall n \geq n_{1}, a_{n}<\epsilon / M$. Thus, $\forall n \geq n_{1}, a_{n} b_{n}<\epsilon$, $\lim \sup \left(a_{n} b_{n}\right)=0$, and the inequality holds.
Case 4: One or both of $L_{1}, L_{2}$, are infinite
Assume both $L_{1}=\infty$ and $L_{2}=\infty$; or, choosing one of the two possible combinations, say $0<L_{1}<\infty$ and $L_{2}=\infty$. Then (15) clearly holds. The case of concern is $L_{1}=0$ and $L_{2}=\infty$, this case having been ruled out in Stoll's posing of the question. Thus, it is of interest to find cases that give discrepant results, showing that, in this case, (15) is not true in general. Take $a_{n}=n^{-1}$ and $b_{n}=n$, so that $\lim \sup \left\{a_{n} b_{n}\right\}=0$, and the inequality holds, irrespective of how one defines $0 \cdot \infty$. Now take $a_{n}=n^{-1 / 2}$ and $b_{n}=n$, so that $\lim \sup \left\{a_{n} b_{n}\right\}=\infty$, and the inequality holds only if we define $0 \cdot \infty=\infty$.
iii. Stoll additionally imposes that the rhs cannot be of the form $0 \cdot \infty$. Give a case in which the necessity of this becomes clear.
ANS: (Chen) The rhs cannot be $0 \cdot \infty$ for the inequality to hold. Here is a counterexample in which we arrive at $4 \leq 0 \cdot \infty$. First consider $S_{n}=n\left[1+(-1)^{n}\right]$. Then

$$
\left\{S_{n}: n \geq k\right\}=\left\{\begin{array}{l}
\{0,2(k+1), 0,2(k+3), \ldots\}, \text { if } k \text { is odd } \\
\{2 k, 0,2(k+2), 0,2(k+4), \ldots\}, \text { if } k \text { is even. }
\end{array}\right.
$$

## Observe

$$
\left\{\begin{array}{l}
a_{k}=\inf \left\{S_{n}: n \geq k\right\}=0, \forall k . \quad \liminf S_{n}=\lim _{k \rightarrow \infty} a_{k}=0 \\
b_{k}=\sup \left\{S_{n}: n \geq k\right\}=\infty, \forall k . \quad \underline{\lim \sup S_{n}=\lim _{k \rightarrow \infty} b_{k}=\infty}
\end{array}\right.
$$

Next consider $T_{n}=n^{-1}\left[1+(-1)^{n}\right]$. Then

$$
\left\{T_{n}: n \geq k\right\}=\left\{\begin{array}{l}
\left\{0, \frac{2}{k+1}, 0, \frac{2}{k+3}, \cdots\right\}, \text { if } k \text { is odd } \\
\left\{\frac{2}{k}, 0, \frac{2}{k+2}, 0, \frac{2}{k+4}, \cdots\right\}, \text { if } k \text { is even } .
\end{array}\right.
$$

Observe $a_{k}=\inf \left\{T_{n}: n \geq k\right\}=0$, so that $\liminf T_{n}=\lim _{k \rightarrow \infty} a_{k}=0$, and

$$
b_{k}=\sup \left\{T_{n}: n \geq k\right\}= \begin{cases}2 /(k+1), & \text { if } k \text { odd } \\ 2 / k, & \text { if } k \text { even }\end{cases}
$$

so that $\limsup T_{n}=\lim _{k \rightarrow \infty} b_{k}=0$. This can also be seen from $\lim _{k \rightarrow \infty} T_{2 k}=\lim _{k \rightarrow \infty} k^{-1}=$ $0=\lim _{k \rightarrow \infty} T_{2 k-1}$. So, $\lim _{n \rightarrow \infty} T_{n}=0=\limsup T_{n}=\liminf T_{n}$.
Observe $S_{n} T_{n}=\left[1+(-1)^{n}\right]^{2}=2\left[1+(-1)^{n}\right]$, with $\lim \inf S_{n} T_{n}=0$ and $\limsup S_{n} T_{n}=4$. Thus, $\lim \sup S_{n} T_{n}=4$ but $\limsup S_{n} \cdot \lim \sup T_{n}=0 \cdot \infty$, this being an indeterminate form (https://en.wikipedia.org/wiki/Indeterminate_form) and not necessarily equal to zero (as $\infty \notin \mathbb{R}$ ). Books on measure theory typically define it to be zero, as that works well when describing the area of a box and for enabling other results. ${ }^{10}$ It appears safe to say, we cannot conclude $4 \leq 0 \cdot \infty$, as (15) suggests.

[^8](e) There are also lim inf and lim sup definitions for sets. These are not difficult, and will be essential when learning measure theory, the Lebesgue integral, and probability theory. In the following, we do not need to define measurable spaces or measure spaces, so just ignore this terminology for now.
Let $\left\{E_{k}\right\}_{k=1}^{\infty}$ be a countable family of measurable subsets of measurable space $(X, \mathcal{A})$. Define
\[

$$
\begin{equation*}
\liminf _{k \rightarrow \infty} E_{k}:=\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_{k}, \quad \limsup _{k \rightarrow \infty} E_{k}:=\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_{k} \tag{18}
\end{equation*}
$$

\]

Equivalent definitions are

$$
\begin{align*}
\liminf _{k \rightarrow \infty} E_{k} & :=\left\{x \in X: x \in E_{k} \text { for all but finitely many } k\right\} \\
\limsup _{k \rightarrow \infty} E_{k} & :=\left\{x \in X: x \in E_{k} \text { for infinitely many } k\right\} \tag{19}
\end{align*}
$$

For $\lim \inf E_{k}$, "for all but finitely many $k$ " means, $\exists k_{0} \in \mathbb{N}$ such that, $\forall k \geq k_{0}, x \in E_{k}$. From (19), it is apparent that $\lim \inf E_{k} \subset \lim \sup E_{k}$. Prove the equivalence of (18) and (19).

ANS: We begin with $(\Rightarrow)$ and $(\Leftarrow)$ for lim sup.
$(\Rightarrow)$ Let $x \in \lim \sup$ in (18), so that

$$
\begin{equation*}
x \in\left(\bigcup_{k=1}^{\infty} E_{k}\right) \cap\left(\bigcup_{k=2}^{\infty} E_{k}\right) \cap\left(\bigcup_{k=3}^{\infty} E_{k}\right) \cap \cdots . \tag{20}
\end{equation*}
$$

Suppose $x \notin\left\{x \in X: x \in E_{k}\right.$ for infinitely many $\left.k\right\}$. Then $\exists k_{0} \in \mathbb{N}$ such that, $\forall k \geq k_{0}, x \notin$ $E_{k}$, which implies $x \notin \bigcup_{k=k_{0}}^{\infty} E_{k}$, which contradicts (20).
$(\Leftarrow)$ Let $x \in \lim$ sup in (19). Then $x$ never stops reappearing in $\left\{E_{k}\right\} \Leftrightarrow x \in \bigcup_{k=1}^{\infty} E_{k}, x \in$ $\bigcup_{k=2}^{\infty} E_{k}$, etc. $\Leftrightarrow$

$$
x \in\left(\bigcup_{k=1}^{\infty} E_{k}\right) \cap\left(\bigcup_{k=2}^{\infty} E_{k}\right) \cap \cdots \Leftrightarrow x \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_{k}
$$

$(\Rightarrow)$ Let $x \in \lim \inf$ in (18), so that

$$
\begin{equation*}
x \in\left(\bigcap_{k=1}^{\infty} E_{k}\right) \cup\left(\bigcap_{k=2}^{\infty} E_{k}\right) \cup\left(\bigcap_{k=3}^{\infty} E_{k}\right) \cup \cdots \tag{21}
\end{equation*}
$$

Suppose $x \notin\left\{x \in X: x \in E_{k}\right.$ for all but finitely many $\left.k\right\}$. That means, $\forall n \in \mathbb{N}, \exists j>n$ such that $x \notin \bigcap_{k=j}^{\infty} E_{k}$, which contradicts (21).
$(\Leftarrow)$ Let $x \in \lim \inf$ in (19). That means $\exists j_{x}$ such that

$$
x \in \bigcap_{k=j_{x}}^{\infty} E_{k} \Longleftrightarrow x \in\left(\bigcap_{k=1}^{\infty} E_{k}\right) \cup\left(\bigcap_{k=2}^{\infty} E_{k}\right) \cup\left(\bigcap_{k=3}^{\infty} E_{k}\right) \cup \cdots=\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_{k}
$$

25. Use the nested intervals property to prove that $[0,1]$ is uncountable. (Stoll, p. 102, \#22)

ANS: (The solution follows
mathonline.wikidot.com/the-set-of-real-numbers-is-uncountable-nested-intervals-proo)

Suppose $[0,1]$ is countable. Then we can enumerate it as $[0,1]=\left\{x_{1}, x_{2}, \ldots\right\}$. Let $I_{1} \subset[0,1]$ be a closed interval such that $x_{1} \notin I_{1}$; let $I_{2} \subset I_{1}$ be a closed interval such that $x_{2} \notin I_{2}$; etc., i.e., for each $n \in \mathbb{N}$, let $I_{n} \subset I_{n-1}$ be a closed interval such that $x_{n} \notin I_{n}$. Observe that, $\forall n \in \mathbb{N}, I_{n} \subset[0,1]$ is bounded. Thus, we obtain the sequence of nested closed and bound intervals $I_{1} \supset I_{2} \supset \cdots$, but by the nested intervals theorem, $\cap_{n=1}^{\infty} \neq \emptyset$. Thus, $\exists m \in \mathbb{N}$ such that $x_{m} \in I_{n}$ for every $n \in \mathbb{N}$. But, by construction, $x_{m} \notin I_{n}$ for all $n \geq m$. Therefore, by contradiction, $[0,1]$ is not countable.
26. Let $A \subset \mathbb{R}$ be uncountable. Prove: some point of $A$ is a limit point of $A$. (Stoll, p. 80, \#3)

ANS: We will need Stoll, Thm 2.2.14, as stated in problem 11 above; and also Stoll, Thm 2.4.2 (Heine-Borel-Bolzano-Weierstrass), parts (b) and (c):

$$
K \text { compact } \Leftrightarrow \text { Every infinite subset of } K \text { has a limit point in } K \text {. }
$$

The latter is just the statement of sequential compactness: (Fitzpatrick, p. 46) A set of real numbers $S$ is said to be sequentially compact provided that every sequence $\left\{a_{n}\right\} \in S$ has a subsequence that converges to a point that belongs to $S$.

Proof I: Exam 2 solutions: https://www.math.colostate.edu/~clayton/teaching/m317f10/
We begin with a Lemma: Let $B \subset \mathbb{R}$ be bounded. If $B$ is infinite, then $B$ has a limit point.
NOTE: This is in fact Stoll, Thm 2.4.3, and is one of the two formulations of Bolzano-Weierstrass. (See, e.g., Duren, who explicitly states and juxtaposes the two of them.) We still include the proof given in the above-mentioned source.

Proof of Lemma: Supposing $B$ is infinite, then $\exists\left\{b_{n}\right\} \in B$ such that $b_{i} \neq b_{j}$ when $i \neq j$. As $\left\{b_{n}\right\}$ is bounded, Bolzano-Weierstrass implies that it contains a convergent subsequence $\left\{b_{n_{k}}\right\}$, with limit, say, $L$. As the $b_{n_{k}}$ are all distinct, at most one term can equal $L$. Form the sub-subsequence $\left\{b_{n_{k_{\ell}}}\right\}$ from $\left\{b_{n_{k}}\right\}$ by deleting the term equal to $L$, if it exists. Thus, $\left\{b_{n_{k_{\ell}}}\right\}$ is a sequence contained in $B$ and converging to $L$, with $b_{n_{k_{\ell}}} \neq L$ for all $\ell$. Therefore, $L$ is a limit point of $B$, so $B$ does indeed have a limit point.

Proof of main result: Suppose $A \subset \mathbb{R}$ has no limit points. Let bounded set $A_{n}:=[-n, n] \cap A, n \in \mathbb{N}$, which, as $A_{n} \subset A$, also has no limit points. The (contrapositive of the) Lemma shows that $A_{n}$ is finite. As $\cup_{n=1}^{\infty} A_{n}=A$, a union of finite sets, $A$ is at most countable, and thus not uncountable.

Proof II: This is a more general proof, applicable to $\mathbb{R}^{n}$, as outlined and with further results here: https://planetmath.org/limitpointsofuncountablesubsetsofrn
The proof is definitely correct for the usual Euclidean metric, without needing to invoke a normed vector space, as the author above does, this being material we have not seen yet in our class. (In case you are curious: Given a norm $\|\cdot\|$, you can always define a metric $d(x, y)=\|x-y\|$. Given a metric $d(x, y)$ on a vector space, you can define a function $d(x, 0)$, and sometimes it is a norm, such as for Euclidean distance and $\mathbb{R}^{n}$, and sometimes it is not.)
Proof: Let closed ball $\bar{B}_{k}(\mathbf{0})=\left\{\mathbf{v} \in \mathbb{R}^{n}: d(\mathbf{v}, \mathbf{0}) \leq k\right\}, \forall k \in \mathbb{N}$. Define $V_{k}=\bar{B}_{k}(\mathbf{0}) \cap A$; and note $\cup_{k=1}^{\infty} V_{k}=A$. If each $V_{k}$ is finite (imagine $A=\mathbb{N}^{2} \subset \mathbb{R}$ ), then $A$ is at most countable, which contradicts the assumption on $A$. Thus, $\exists k_{0} \in \mathbb{N}$ such that $V_{k_{0}}$ is infinite. As (i) $V_{k_{0}} \subset \bar{B}_{k_{0}}$; (ii) $\bar{B}_{k_{0}}$ is compact; and (iii) $V_{k_{0}}$ is infinite, there exists a limit point of $V_{k_{0}} \in \mathbb{R}^{n}$, using Heine-Borel-BolzanoWeierstrass given above.
27. Recall the Bolzano-Weierstrass theorem (e.g., Stoll, Coro 3.4.6; Fitzpatrick, Thm 2.33): Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.

State the Nested Intervals Property (e.g., Stoll, Coro 3.3.3; Fitzpatrick, Thm 2.29). Prove: The Bolzano-Weierstrass theorem implies the nested intervals property. (Stoll p. 106 \#11)
ANS: The Nested Intervals Property: If $\left\{I_{n}\right\}_{n=1}^{\infty}$ is a sequence of closed and bounded intervals in $\mathbb{R}$ with $I_{n} \supset I_{n+1}$ for all $n \in \mathbb{N}$, then $\cap_{n=1}^{\infty} I_{n} \neq \emptyset$.
(From the exam 1 solutions of https://www.math.colostate.edu/~clayton/teaching/m317f10/)
Let $I_{n}=\left[a_{n}, b_{n}\right]$ for each $n \in \mathbb{N}$, such that $I_{n} \supset I_{n+1}$. Then, for each $n \in \mathbb{N}, a_{1}<b_{n} \leq b_{1}$, so $\left\{b_{n}\right\}$ is bounded. The Bolzano-Weierstrass theorem implies it contains a convergent subsequence $b_{n_{k}} \rightarrow b$.
First show that $b \leq b_{n}$ for each $n \in \mathbb{N}$ : Both $\left\{b_{n}\right\}$ and $\left\{b_{n_{k}}\right\}$ are decreasing, because $I_{n} \supset I_{n+1}$. If $b>b_{m}$ for some $m \in \mathbb{N}$, then, $\forall k \geq m, n_{k} \geq n_{m} \geq m$, which implies $b_{n_{k}} \leq b_{n_{m}} \leq b_{m}<b$ and

$$
\left|b_{n_{k}}-b\right|=b-b_{n_{k}} \geq b-b_{m}
$$

But $b-b_{m}$ is a fixed positive number and $b_{n_{k}} \rightarrow b$, so by contradiction, $b \leq b_{m}$.
Next show $b \geq a_{n}$ for each $n \in \mathbb{N}$ : For each $n \in \mathbb{N}, b_{n_{k}} \geq a_{n}$. Taking the limit in $k$ and using (13) gives $b \geq a_{n}$.
Thus, $\forall n \in \mathbb{N}$, $a_{n} \leq b \leq b_{n}$, so, $\forall n \in \mathbb{N}, b \in I_{n}$, or $b \in \cap_{i=1}^{\infty} I_{n}$, implying the intersection is non-empty. As the choice of nested $\left\{I_{n}\right\}$ was arbitrary, the Nested Interval Property is proven.
28. (Convergent sequences and subsequences in a metric space)
(a) Let $(X, d)$ be a metric space and let $\left\{p_{n}\right\}$ be a sequence in $X$. Prove: If $\left\{p_{n}\right\}$ converges to $p$, then every subsequence of $\left\{p_{n}\right\}$ also converges to $p$.
(This is in all books, e.g., Stoll, Thm 3.4.3.)
ANS: Let $\left\{p_{n_{k}}\right\}$ be a subsequence of $\left\{p_{n}\right\}$, and fix $\epsilon>0$. As $p_{n} \rightarrow p, \exists n_{o} \in \mathbb{N}$ such that $\forall n \geq n_{o}, d\left(p_{n}, p\right)<\epsilon$. For $k \geq n_{o}, n_{k} \geq n_{o}$, and $d\left(p_{n_{k}}, p\right)<\epsilon$, i.e., $p_{n_{k}} \rightarrow p$.
(b) Prove the converse of part (a) when $X$ compact: That is, for a compact metric space $(X, d)$, if every convergent subsequence of a sequence $\left\{x_{n}\right\}$ converges to the same limit, say $L$, then the original sequence also converges to $L$.
Hint: Use contradiction, and the (sequential) compactness of $X$.
ANS: (https://math.stackexchange.com/questions/461610/)
Suppose $\left\{x_{n}\right\} \nrightarrow L$. Then $\exists \epsilon>0$ such that, $\forall n_{0} \in \mathbb{N}, \exists n>n_{0}$ such that $d\left(x_{n}, L\right) \geq \epsilon$. Thus, we get a subsequence $\left\{x_{n_{k}}\right\}$ such that $d\left(x_{n_{k}}, L\right) \geq \epsilon$ for every $k \in \mathbb{N}$. As $X$ is compact, $\left\{x_{n_{k}}\right\}$ has a convergent (to a value in $X$ ) subsequence $\left\{x_{n_{k_{m}}}\right\}$, and note $\left\{x_{n_{k_{m}}}\right\}$ cannot converge to $L$. But $\left\{x_{n_{k_{m}}}\right\}$ is also a subsequence of $\left\{x_{n}\right\}$, which is convergent, and so, from part (a), converges to $L$, which is a contradiction.
(c) Consider the following proof that part (b) is true. Observe it does not invoke compactness. Is the proof correct, and if not, where is it wrong? (Contributed by Chen and Ralf)
"Proof": Suppose that every convergent subsequence of $\left\{x_{n}\right\}$ converges to $L \in X$. In particular, $x_{2 n-1} \rightarrow L$ and $x_{2 n} \rightarrow L$. Then, $\forall \epsilon>0, \exists N_{1}, N_{2} \in \mathbb{N}$ such that

$$
x_{2 n-1} \in N_{\epsilon}(L), \forall n \geq N_{1}, \text { and } x_{2 n} \in N_{\epsilon}(L), \forall n \geq N_{2}
$$

Let $N=\max \left(N_{1}, N_{2}\right)$. Then, $x_{n} \in N_{\epsilon}(L), \forall n \geq 2 N$. That is, $x_{n} \rightarrow L$.
ANS: The proof is not valid: We do not know that $\left\{x_{2 n}\right\}$ and $\left\{x_{2 n+1}\right\}$ are convergent subsequences of $\left\{x_{n}\right\}$. Note that all convergent subsequences converge to the same limit, but not all subsequences in general. For instance, take a sequence $n$ but replace every second element with
a zero. Then all convergent subsequences converge to 0 , but the sequence as such diverges to infinity.
NOTE: An open question is, can we still obtain the result, with relaxing (not imposing) compactness, but rather only boundedness of every subsequence of a sequence $\left\{x_{n}\right\}$ ?
29. (Cauchy sequences and series)
(a) State the definition of a Cauchy sequence in a general metric space ( $X, d$ ), and simplify it for $X=\mathbb{R}$ endowed with the usual metric.
ANS: (Stoll, Def 3.6.1) Let $(X, d)$ be a metric space. A sequence $\left\{p_{n}\right\}$ in $X$ is a Cauchy sequence if, for every $\epsilon>0, \exists N \in \mathbb{N}$ such that $d\left(p_{n}, p_{m}\right)<\epsilon$ for all integers $n \geq N, m \geq N$.
For $(X, d)=(\mathbb{R},|\cdot|)$, this simplifies to (Fitzpatrick, p. 228): A sequence $\left\{p_{n}\right\}$ in $\mathbb{R}$ is a Cauchy sequence if, for every $\epsilon>0, \exists N \in \mathbb{N}$ such that $\left|p_{n}-p_{m}\right|<\epsilon$ for all integers $n \geq N, m \geq N$.
(b) State the definition of a complete metric space.

ANS: (Stoll, Def 3.6.6) A metric space $(X, d)$ is said to be complete if every Cauchy sequence in $X$ converges to a point in $X$.
(c) Let $\left\{x_{n}\right\} \in \mathbb{R}$ such that $\forall n \in \mathbb{N},\left|x_{n}-x_{n+1}\right| \leq 2^{-n}$. Show $\left\{x_{n}\right\}$ converges.
(From exam 1 solutions of https://www.math.colostate.edu/~clayton/teaching/m317f10/)
ANS: Let $\epsilon>0$. Choose $N \in \mathbb{N}$ such that $N>1-\log _{2} \epsilon$. Then for any $n>m \geq N$,

$$
\begin{aligned}
\left|x_{m}-x_{n}\right| & =\left|x_{m}-x_{m+1}+x_{m+1}-x_{m+2}+\cdots+x_{n-1}-x_{n}\right| \\
& \leq\left|x_{m}-x_{m+1}\right|+\left|x_{m+1}-x_{m+2}\right|+\cdots+\left|x_{n-1}-x_{n}\right| \\
& \leq \frac{1}{2^{m}}+\frac{1}{2^{m+1}}+\cdots+\frac{1}{2^{n-1}} \\
& =\frac{1}{2^{m}}\left(1+\frac{1}{2}+\cdots+\frac{1}{2^{n-m-1}}\right) \\
& \leq \frac{1}{2^{m-1}} \leq \frac{1}{2^{N-1}}<\epsilon
\end{aligned}
$$

Thus, $\left\{x_{n}\right\}$ is Cauchy, and thus convergent.
(d) (Mattuck, Example 6.4) ${ }^{11}$ The sequence of Fibonacci fractions is defined recursively by

$$
\begin{equation*}
a_{1}=1, \quad a_{n+1}=\frac{1}{a_{n}+1}, \quad n=2,3, \ldots \tag{22}
\end{equation*}
$$

It starts: $\frac{1}{1}, \frac{1}{2}, \frac{2}{3}, \frac{3}{5}, \frac{5}{8}, \frac{8}{13}, \ldots$ Prove it converges, and determine its limit.
ANS: Solution. To prove convergence, we show the sequence is a Cauchy sequence. To see that $\left|a_{n}-a_{m}\right|$ is small, we first use (22) to get an estimate of how rapidly the difference $\left|a_{n-1}-a_{n}\right|$ decreases as you move from one pair of successive terms to the next. We have by (22):

$$
\begin{align*}
\left|a_{n}-a_{n+1}\right| & =\left|\frac{1}{a_{n-1}+1}-\frac{1}{a_{n}+1}\right| \\
& =\frac{\left|a_{n-1}-a_{n}\right|}{\left(a_{n}+1\right)\left(a_{n-1}+1\right)} \tag{23}
\end{align*}
$$

To estimate this, it looks from (22) as if $a_{n} \geq \frac{1}{2}$ for all $n$; assuming this for the moment, we get for the denominator

$$
\left(a_{n}+1\right)\left(a_{n-1}+1\right) \geq \frac{3}{2} \cdot \frac{3}{2}>2
$$

[^9]so that from (23) we get
\[

$$
\begin{equation*}
\left|a_{n}-a_{n+1}\right| \leq \frac{1}{2}\left|a_{n-1}-a_{n}\right|, \text { for all } n \geq 2 \tag{24}
\end{equation*}
$$

\]

We now use recursion, applying (24) in turn to $\left|a_{n-1}-a_{n}\right|$; this gives

$$
\left|a_{n}-a_{n+1}\right| \leq \frac{1}{4}\left|a_{n-2}-a_{n-1}\right|
$$

and continuing in this way, we get finally

$$
\begin{equation*}
\left|a_{n}-a_{n+1}\right| \leq \frac{1}{2^{n-1}}\left|a_{1}-a_{2}\right| \leq \frac{1}{2^{n}} \tag{25}
\end{equation*}
$$

Now we use (24) to estimate $\left|a_{n}-a_{m}\right|$, for $m>n$. From the (extended) triangle inequality,

$$
\begin{align*}
\left|a_{n}-a_{m}\right| & \leq\left|a_{n}-a_{n+1}\right|+\left|a_{n+1}-a_{n+2}\right|+\ldots+\left|a_{m-1}-a_{m}\right|  \tag{26}\\
& \leq \frac{1}{2^{n}}+\frac{1}{2^{n+1}}+\ldots+\frac{1}{2^{m-1}}, \quad \text { using }(25)  \tag{27}\\
& \leq \frac{1}{2^{n}}\left(1+\frac{1}{2}+\frac{1}{4}+\ldots\right)  \tag{28}\\
& \leq \frac{1}{2^{n-1}}<\epsilon, \quad \text { for } m>n \gg 1 . \tag{29}
\end{align*}
$$

This proves $\left\{a_{n}\right\}$ is a Cauchy sequence, which converges therefore to some limit $L$. As subsequences $a_{n} \rightarrow L$ and $a_{n+1} \rightarrow L$, it follows that $L=1 /(L+1)$; cross-multiplying turns this into a quadratic equation whose unique positive root is $L=(\sqrt{5}-1) / 2$. The proof used $a_{n} \geq 1 / 2$ for all $n$; to show this, use recursion: $a_{1}=1$, and

$$
\begin{aligned}
\frac{1}{2} \leq a_{n} \leq 1 & \Rightarrow \frac{3}{2} \leq a_{n}+1 \leq 2 \\
& \Rightarrow \frac{1}{2} \leq \frac{1}{a_{n}+1} \leq \frac{2}{3}, \quad \text { by the inequality laws; } \\
& \Rightarrow \frac{1}{2} \leq a_{n+1} \leq 1, \quad \text { by }(22)
\end{aligned}
$$

(e) State the Cauchy criterion for series convergence, and prove it.

ANS: (In every book, e.g., Stoll, Thm 3.7.3; Fitzpatrick, Thm 9.17)
The series $\sum_{k=1}^{\infty} a_{k}$ converges iff given $\epsilon>0, \exists n_{o} \in \mathbb{N}$ such that

$$
\left|\sum_{k=n+1}^{m} a_{k}\right|<\epsilon, \quad \forall m>n \geq n_{o} .
$$

Proof: $\left|\sum_{k=n+1}^{m} a_{k}\right|=\left|s_{m}-s_{n}\right|$, where $s_{n}=\sum_{k=1}^{n} a_{k}$ is the $n$th partial sum of the series.
Next, recall
i. Stoll, Thm 3.6.2(a): Let $(X, d)$ be a metric space. Every convergent sequence in $X$ is a Cauchy sequence.
ii. Stoll, Thm 3.6.5: Every Cauchy sequence of real numbers converges.

Thus, $(\Rightarrow)$ follows from (a); and $(\Leftarrow)$ follows from (b).
(f) Show that convergence of $\sum_{n=1}^{\infty} n a_{n}$ implies that of $\sum_{n=1}^{\infty} a_{n}$.
(Hata, Problems and Solutions in Real Analysis, \#2.5)
ANS: Let $b_{n}:=a_{1}+2 a_{2}+\cdots+n a_{n}, M:=\sup \left|b_{n}\right|$, and fix $\epsilon>0$. Choose any $p, q \in \mathbb{N}$ such that $p>q>2 M / \epsilon$. Begin for illustration with $p=q+2$. Then

$$
\begin{aligned}
\left|\sum_{n=q}^{q+2} a_{n}\right| & =\left|\frac{b_{q}-b_{q-1}}{q}+\frac{b_{q+1}-b_{q}}{q+1}+\frac{b_{q+2}-b_{q+1}}{q+2}\right| \\
& =\left|-\frac{b_{q-1}}{q}+\left(\frac{1}{q}-\frac{1}{q+1}\right) b_{q}+\left(\frac{1}{q+1}-\frac{1}{q+2}\right) b_{q+1}+\frac{b_{q+2}}{q+2}\right| \\
& \leq(p-q+2) M\left[\left|-\frac{1}{q}\right|+\left|\left(\frac{1}{q}-\frac{1}{q+1}\right)\right|+\left|\left(\frac{1}{q+1}-\frac{1}{q+2}\right)\right|+\frac{1}{q+2}\right] \\
& \leq \frac{(p-q+2) M}{(p-q+2) q}=\frac{M}{q} \leq M \frac{\epsilon}{2 M}=\frac{\epsilon}{2}<\epsilon .
\end{aligned}
$$

In the general case,

$$
\begin{aligned}
\left|\sum_{n=q}^{p} a_{n}\right| & =\left|\frac{b_{q}-b_{q-1}}{q}+\frac{b_{q+1}-b_{q}}{q+1}+\cdots+\frac{b_{p}-b_{p-1}}{p}\right| \\
& =\left|-\frac{b_{q-1}}{q}+\left(\frac{1}{q}-\frac{1}{q+1}\right) b_{q}+\cdots+\left(\frac{1}{p-1}-\frac{1}{p}\right) b_{p-1}+\frac{b_{p}}{p}\right|<\epsilon
\end{aligned}
$$

This is the Cauchy criterion for convergence of series $\sum_{n=1}^{\infty} a_{n}$.
(g) Show that $n^{-1} \sum_{k=1}^{\infty} a_{k}$ converges to zero when $\sum_{n=1}^{\infty} a_{n} / n$ converges.
(Hata, Problems and Solutions in Real Analysis, \#2.6)
ANS: For this, we will need the definition of null sequence (exercise 18n) and the convergence result of Cesàro summation (exercise 20a).
Let $\left\{\epsilon_{n}\right\}$ be a null sequence, i.e., it converges to zero. Define $\alpha, \epsilon_{n}$, and $\sigma_{n}$ as follows:

$$
\frac{a_{1}}{1}+\frac{a_{2}}{2}+\cdots+\frac{a_{n}}{n}=\alpha+\epsilon_{n}, \quad \text { and } \quad \sigma_{n}=\frac{a_{1}+a_{2}+\cdots+a_{n}}{n} .
$$

If we can show

$$
\begin{equation*}
\sigma_{n}=\frac{\alpha}{n}-\frac{\epsilon_{1}+\epsilon_{2}+\cdots+\epsilon_{n-1}}{n}+\epsilon_{n} \tag{30}
\end{equation*}
$$

then the result follows, i.e., $\sigma_{n} \rightarrow 0$ as $n \rightarrow \infty$. To prove (30), use induction: The case $n=1$ is true because $\sigma_{1}=a_{1}=\alpha+\epsilon_{1}$. Suppose next that (30) holds for $n=m$. Then

$$
\begin{aligned}
\sigma_{m+1} & =\frac{m}{m+1} \sigma_{m}+\frac{a_{m+1}}{m+1} \\
& =\frac{m}{m+1}\left(\frac{\alpha}{m}-\frac{\epsilon_{1}+\cdots+\epsilon_{m-1}}{m}+\epsilon_{m}\right)+\epsilon_{m+1}-\epsilon_{m} \\
& =\frac{\alpha}{m+1}-\frac{\epsilon_{1}+\cdots+\epsilon_{m}}{m+1}+\epsilon_{m+1},
\end{aligned}
$$

as required.
30. (Box covers) In this exercise, unless otherwise allowed, $Q_{j} \subset \mathbb{R}$ will refer to a closed, finite (bounded), nonempty interval on the real line, e.g., $[a, b]$, for $a, b \in \mathbb{R}$, with $a<b$. (A box generalizes this to closed hyper-rectangles in $\mathbb{R}^{d}$, but for $d=1$, a box reduces to said interval.)

Let $J$ be an at most countable set. Common examples include $J=\{1,2, \ldots, N\}$ for $N \in \mathbb{N} ; J=\mathbb{N}$; and $J=\mathbb{Z}$. The collection $\left\{Q_{k}\right\}_{k \in J}$ is said to be nonoverlapping if their interiors are disjoint, i.e., $\forall j, k \in J, j \neq k \Rightarrow Q_{j}^{\circ} \cap Q_{k}^{\circ}=\emptyset$. As an example, with $Q_{k}=[k, k+1]$, the elements of $\left\{Q_{k}\right\}_{k \in \mathbb{Z}}$ are nonoverlapping, though not disjoint. Note that $\left\{Q_{k}\right\}_{k \in \mathbb{Z}}=\mathbb{R}$.
Prove the following statements. (Heil, Measure Theory for Scientists and Engineers, forthcoming, Exercise 2.1.7)
(a) $(0,1]=\bigcup_{k=0}^{\infty}\left[2^{-k-1}, 2^{-k}\right]$.

ANS: We need to show both directions of set inclusion.
To show $(0,1] \supset \cup_{k=0}^{\infty}\left[2^{-k-1}, 2^{-k}\right]$, note that $\forall k \in(\{0\} \cup \mathbb{N}),\left[2^{-k-1}, 2^{-k}\right] \subset(0,1]$.
To show $(0,1] \subset \cup_{k=0}^{\infty}\left[2^{-k-1}, 2^{-k}\right]$, observe that:

- $\lim _{k \rightarrow \infty} 2^{-k-1}=\lim _{k \rightarrow \infty} 1 / 2^{k}=0$, but $\forall k \in \mathbb{N}$, limit point $0 \notin\left[2^{-k-1}, 2^{-k}\right]$.
- For $k=0,1 \in\left[2^{-k-1}, 2^{-k}\right]=[1 / 2,1]$. Thus, $1 \in \cup_{k=0}^{\infty}\left[2^{-k-1}, 2^{-k}\right]$.
- Recall the Archimedean property: $\forall x>0, \exists n \in \mathbb{N}$ such that $1 / n<x$. Both $\{1 / n\}$ and $\left\{1 / 2^{k+1}\right\}$ are positive, strictly monotone decreasing null sequences (see exercise 18 n ), and so the Archimedean property also applies to the latter, namely: $\forall x>0, \exists k \in \mathbb{N}$ such that $1 / 2^{k+1}<x$. - Let $x \in(0,1)$. The set $A_{x}:=\left\{k \in \mathbb{N}: 1 / 2^{k+1} \leq x\right\}$ is thus nonempty, and it is bounded below. From the Well-Ordering principle, it has a smallest element, $k_{x}$. By construction, $k_{x}-1 \notin A_{x}$, and thus $1 / 2^{k_{x}+1} \leq x<1 / 2^{k_{x}}$, i.e.,

$$
x \in\left[1 / 2^{k_{x}+1} \leq x<1 / 2^{k_{x}}\right) \subset\left[1 / 2^{k_{x}+1} \leq x<1 / 2^{k_{x}}\right] \subset \cup_{k=0}^{\infty}\left[2^{-k-1}, 2^{-k}\right]
$$

showing the desired inclusion.
(b) $[1, \infty)=\bigcup_{k=0}^{\infty}\left[2^{k}, 2^{k+1}\right]$.

ANS: We need to show both directions of set inclusion.
To show $[1, \infty) \supset \cup_{k=0}^{\infty}\left[2^{k}, 2^{k+1}\right]$, note that $\forall k \in(\{0\} \cup \mathbb{N}),\left[2^{k}, 2^{k+1}\right] \subset[1, \infty)$.
To show $[1, \infty) \subset \cup_{k=0}^{\infty}\left[2^{k}, 2^{k+1}\right]$, observe that:

- For $k=0,1 \in\left[2^{k}, 2^{k+1}\right]=[1,2]$. Thus $1 \in \cup_{k=0}^{\infty}\left[2^{k}, 2^{k+1}\right]$.
- Let $x>1$. To see that $x \in \cup_{k=0}^{\infty}\left[2^{k}, 2^{k+1}\right]$, note that the $\left[2^{k}, 2^{k+1}\right]$ are nonoverlapping, but adjacent and not disjoint. Their union (starting from $k=0$ ) clearly covers $(1, \infty)$, because $\lim _{k \rightarrow \infty} 2^{k+1}=\lim _{k \rightarrow \infty} 2^{k}=\infty$.
(c) Every finite open interval $(a, b)$ is a union of countably many nonoverlapping intervals $\left\{Q_{k}\right\}_{k \in \mathbb{N}}$. ANS:
- (Marc) First consider the situation when $b-a>2$. Observe that $(a, b)$ can be written as the union of overlapping closed intervals $(a, b)=\cup_{n=1}^{\infty} H_{k}$, where $H_{k}=[a+1 / k, b-1 / k]$. This motivates defining

$$
Q_{0}=[a+1, b-1], \quad Q_{a, k}=[a+1 /(k+1), a+1 / k], \quad Q_{b, k}=[b-1 / k, b-1 /(k+1)],
$$

so that $(a, b)=Q_{0} \cup\left(\cup_{k=1}^{\infty} Q_{a, k}\right) \cup\left(\cup_{k=1}^{\infty} Q_{b, k}\right)$, which, being a finite union of countable unions, is itself a countable union, of nonoverlapping, closed, bounded, nonempty intervals.
Now consider the case when $b-a \leq 2$. Scale $(a, b)$ to have length greater than two, e.g., let $a^{*}=\kappa a /(b-a)$ and $b^{*}=\kappa b /(b-a)$, where $\kappa>2$. Then the previous result is applicable to $\left(a^{*}, b^{*}\right)$, producing the desired nonoverlapping set of intervals $Q_{0}^{*},\left\{Q_{a, k}^{*}\right\}_{k \in \mathbb{N}}$, and $\left\{Q_{b, k}^{*}\right\}_{k \in \mathbb{N}}$. Scale these by multiplying each by $(b-a) / \kappa$ to get the result.
NOTE: For fun, observe $[a, b]=\cap_{n=1}^{\infty}(a-1 / n, b+1 / n)$.

- (Christopher Heil) Part (a) shows how to cover ( 0,1 ] by nonoverlapping boxes, and a symmetric construction tells us how to cover $[1,2)$. The union of these two covers gives us a cover of $(0,2)$. Rescaling and translating gives us a cover of any finite interval $(a, b)$.
- (Ralf) Let

$$
Q_{k}=\left[a+\frac{b-a}{2} 2^{-k-1}, a+\frac{b-a}{2} 2^{-k}\right], \quad Q_{j}=\left[b-\frac{b-a}{2} 2^{-k}, b-\frac{b-a}{2} 2^{-k-1}\right],
$$

$\forall k, j \in \mathbb{N}$. Then $\cup Q_{k}=\left(a, \frac{b+a}{2}\right]$ and $\cup Q_{j}=\left[\frac{b+a}{2}, b\right)$. The exercise requires finding a countable collection of nonoverlapping closed intervals. A finite union of countable sets is countable.
(d) Every infinite open interval is a union of countably many nonoverlapping intervals $\left\{Q_{k}\right\}_{k \in \mathbb{N}}$. Hint: The question does not say that $Q_{k}$ needs to be bounded.
ANS: Assume the interval is of the form $(b, \infty)$, with the case of $(-\infty, a)$ being similar. Then we can write $(b, \infty)=[b+1 / 2, \infty) \cup\left(\cup_{k=1}^{\infty}\left[b+1 / 2^{k+1}, b+1 / 2^{k}\right]\right)$.
(e) If $U \subset \mathbb{R}$ is open, then there exists a countable collection of nonoverlapping intervals $\left\{Q_{k}\right\}_{k \in \mathbb{N}}$ such that $U=\cup Q_{k}$.
ANS: We require the well-known "Characterization of the Open Subsets of $\mathbb{R}$ " (e.g., Stoll, Thm 2.2.20; Terrell, Thm 4.1.6):

If $U$ is an open subset of $\mathbb{R}$, then there exists a finite or countable collection $\left\{I_{n}\right\}$ of pairwise disjoint open intervals such that $U=\cup_{n} I_{n}$.
It might be required that one or two of the $I_{n}$ are infinite. The result follows from (i) this characterization; (ii) parts (c) and (d), whereby each $I_{n}$ can be represented as union of countably many nonoverlapping intervals; (iii) the $\left\{I_{n}\right\}$ are disjoint; and (iv) that a countable union of countable unions is itself a countable union.
31. (Optional) Study Terrell, Thm 3.5.2: The Euler number $e$ is irrational.

## 4 Stoll Chapters 4, 5, 6 (6.1 to 6.4), and Fitzpatrick Chapter 4

Every day you face battles - that is the reality for all creatures in their struggle to survive. But the greatest battle of all is with yourself-your weaknesses, your emotions, your lack of resolution in seeing things through to the end. You must declare unceasing war on yourself.

Robert Greene
Si vis pacem, para bellum. If you want peace, prepare for war.

### 4.1 Stoll Chapter 4

1. State and prove the Squeeze Theorem for Functions. (Stoll, Thm 4.1.9; Giv, Thm 3.34) ${ }^{12}$

ANS: The Squeeze Theorem for Functions is, here stated as in Giv: Let $f, g$, and $h$ be functions defined on $E \subset \mathbb{R}$ such that $\lim _{x \rightarrow a} g(x)=\lim _{x \rightarrow a} h(x)=L$, where $a$ is a limit point of $E$. If for every $x \in E \backslash\{a\}$ which is sufficiently close to $a$,

$$
\begin{equation*}
g(x) \leq f(x) \leq h(x), \tag{31}
\end{equation*}
$$

then

$$
\begin{equation*}
\lim _{x \rightarrow a} f(x)=L \tag{32}
\end{equation*}
$$

We begin by noting that, when Giv writes (correctly, and as other authors also do, such at Mattuck, but not in the most precise way) that "If for every $x \in E \backslash\{a\}$ which is sufficiently close to $a$," this means:

$$
\begin{equation*}
\exists \delta_{0}>0 \text { such that, } \forall x \in\left(B_{\delta_{0}}(a) \backslash\{a\}\right) \cap E, \quad g(x) \leq f(x) \leq h(x) . \tag{33}
\end{equation*}
$$

- Proof I: By the equivalence of the $\epsilon-\delta$ and sequential limit criterion of function limits (e.g., Stoll, Thm 4.1.3), to prove (32), it is sufficient to show that, for $\left\{a_{n}\right\}$ any sequence in $E \backslash\{a\}$ that converges to $a$,

$$
\begin{equation*}
\lim _{n \rightarrow \infty} f\left(a_{n}\right)=L \tag{34}
\end{equation*}
$$

From (33) and that $a_{n} \rightarrow a, \exists N \in \mathbb{N}$ such that, $\forall n \geq N, a_{n} \in B_{\delta_{0}}(a)$. Thus, $\forall n \geq N, g\left(a_{n}\right) \leq$ $f\left(a_{n}\right) \leq h\left(a_{n}\right)$, and (34) follows from the Squeeze Theorem for sequences (exercise 18i).

- Proof II (Ralf): Let $\epsilon>0$. Then (as $\left.\lim _{x \rightarrow a} g(x)=\lim _{x \rightarrow a} h(x)=L\right) \exists \delta_{g}, \delta_{h}>0$ such that

$$
\begin{align*}
& \forall x \in E \text { such that } 0<d(x, a)<\delta_{g}, \quad|g(x)-L|<\epsilon ; \text { and }  \tag{35}\\
& \forall x \in E \text { such that } 0<d(x, a)<\delta_{h}, \quad|h(x)-L|<\epsilon . \tag{36}
\end{align*}
$$

Let $\delta=\min \left\{\delta_{0}, \delta_{h}, \delta_{g}\right\}$. Then

$$
\forall x \in E \text { such that } 0<d(x, a)<\delta, \quad-\epsilon<g(x)-L \leq f(x)-L \leq h(x)-L<\epsilon,
$$

which implies $\lim _{x \rightarrow p} f(x)=L$.
2. (Trigonometry)
(a) Prove:

$$
\begin{equation*}
\forall x \in(-\pi / 2, \pi / 2) \backslash\{0\}, \quad \sin x<x \text { and } \cos x<x^{-1} \sin x<1 . \tag{37}
\end{equation*}
$$

ANS: The result for $x \in(0, \pi / 2)$ is detailed in, e.g., Conway, A First Course in Analysis, Lemma 2.2.7; Duren, Invitation to Classical Analysis, p. 111; and Figure 4. For $x \in(-\pi / 2,0)$, let $y=-x$ so that $\cos x<x^{-1} \sin x<1 \Rightarrow \cos y<-y^{-1}(-1) \sin y<1$.

[^10]

Figure 4: From https://en.m.wikipedia.org/wiki/Squeeze_theorem
(b) Prove the well-known angle sum and difference identities:

$$
\begin{array}{ll}
\cos (u-v)=\cos u \cos v+\sin u \sin v, & \cos (u+v)=\cos u \cos v-\sin u \sin v  \tag{38}\\
\sin (x+y)=\sin x \cos y+\sin y \cos x, & \sin (x-y)=\sin x \cos y-\cos x \sin y
\end{array}
$$

ANS: (Kenneth Kuttler, Calculus of One and Many Variables, p. 59) Figures 5-6 show excerpts from Kuttler, providing one of the best proofs I have seen for these results.
(c) Prove:

$$
\begin{equation*}
\cos (2 x)=\cos ^{2}(x)-\sin ^{2}(x) \text { and } 1-\cos x=2 \sin ^{2}(x / 2) . \tag{39}
\end{equation*}
$$

Note that the latter is the same as the more commonly seen $\cos (2 x)=1-2 \sin ^{2}(x)$, which is one of the double angle formulae, as stated (without proof) in, e.g., Loya, Amazing and Aesthetic Aspects of Analysis, p. 324, namely $\cos (2 z)=\cos ^{2} z-\sin ^{2} z=2 \cos ^{2} z-1=1-2 \sin ^{2} z$; and $\sin (2 z)=2 \cos z \sin z$.
ANS: From (38), $\cos (2 x)=\cos (x+x)=\cos (x) \cos (x)-\sin (x) \sin (x)=\cos ^{2}(x)-\sin ^{2}(x)$. Then, using this, we obtain $\cos 2 x=\cos ^{2} x-\sin ^{2} x=\left(1-\sin ^{2} x\right)-\sin ^{2} x=1-2 \sin ^{2} x$.
(d) Verify the following equalities. (Stoll, p. 143, \#5; Giv, Example 3.35)
i. $\lim _{x \rightarrow 0} \cos x=1$.
ii. $\lim _{x \rightarrow 0}(\sin x) / x=1$.
iii. $\lim _{x \rightarrow 0}(1-\cos x) / x=0$.

ANS:

- For (i), we could appeal to continuity of sin, cos, square and square root, and write $\cos (x)=$ $\sqrt{1-\sin ^{2}(x)}$; and then use the exchangeability of limits and continuous functions, and that $\lim _{h \rightarrow 0} \sin (h)=0$ (which follows from (37) and/or geometrically from Figure 4). However, we have not yet seen the result on exchange of limit and continuous functions.
Instead, we can use Stoll, Thm 4.1.6(b), namely:
Suppose $E$ is a subset of a metric space $X ; f, g: E \rightarrow \mathbb{R}$; and $p$ is a limit point of $E$.
If $\lim _{x \rightarrow p} f(x)=A$ and $\lim _{x \rightarrow p} g(x)=B$, then $\lim _{x \rightarrow p} f(x) g(x)=A B$.
So, using basic properties of limits, negate both sides of $\lim _{x \rightarrow 0} \cos x=1$ and add 1 , and use (39), so that the result is equivalent to showing $\lim _{x \rightarrow 0} \sin ^{2}(x / 2)=0$, or $\lim _{x \rightarrow 0} \sin (x / 2)=0$, or $\lim _{h \rightarrow 0} \sin (h)=0$. This latter result follows from (37) and/or geometrically from Figure 4.
- For (ii), let $x \rightarrow 0$ in (37), with $\lim _{x \rightarrow 0} \cos x=1$ from part (i); and use (32).


## Theorem 2.3.5 Let $x, y \in \mathbb{R}$. Then

$$
\begin{equation*}
\cos (x+y) \cos (x)+\sin (x+y) \sin (x)=\cos (y) . \tag{2.4}
\end{equation*}
$$

Proof: Recall that for a real number $z$, there is a unique point $p(z)$ on the unit circle and the coordinates of this point are $\cos z$ and $\sin z$. Now it seems geometrically clear that the length of the arc between $p(x+y)$ and $p(x)$ has the same length as the arc between $p(y)$ and $p(0)$. As in the following picture.


Also from geometric reasoning, rigorously examined later, the distance between the points $p(x+y)$ and $p(x)$ must be the same as the distance from $p(y)$ to $p(0)$. In fact, the two triangles have the same angles and the same sides. Writing this in terms of the definition of the trig functions and the distance formula,

$$
(\cos (x+y)-\cos x)^{2}+(\sin (x+y)-\sin x)^{2}=(\cos y-1)^{2}+\sin ^{2} x .
$$

Expanding the above,
$\cos ^{2}(x+y)+\cos ^{2} x-2 \cos (x+y) \cos x+\sin ^{2}(x+y)+\sin ^{2} x-2 \sin (x+y) \sin x$
$=\cos ^{2} y-2 \cos y+1+\sin ^{2} y$
Now using that $\cos ^{2}+\sin ^{2}=1$,

$$
2-2 \cos (x+y) \cos (x)-2 \sin (x+y) \sin (x)=2-2 \cos (y) .
$$

Therefore, $\cos (x+y) \cos (x)+\sin (x+y) \sin (x)=\cos (y)$

Figure 5: Kuttler, p. 59. Continued in Figure 6.

Recall that the length of the unit circle is defined as $2 \pi$. This started with Euler who decided that $\pi$ should be such that $2 \pi$ is the length of the unit circle. Thus it becomes obvious what the sine and cosine are for certain special angles. For example, $\sin \left(\frac{\pi}{2}\right)=1, \cos \left(\frac{\pi}{2}\right)=0$. Letting $x=\pi / 2,(2.4)$ shows that

$$
\begin{equation*}
\sin (y+\pi / 2)=\cos y \tag{2.5}
\end{equation*}
$$

Now let $u=x+y$ and $v=x$. Then (2.4) implies

$$
\begin{equation*}
\cos u \cos v+\sin u \sin v=\cos (u-v) \tag{2.6}
\end{equation*}
$$

Also, from this and (2.3),

$$
\begin{align*}
\cos (u+v) & =\cos (u-(-v))=\cos u \cos (-v)+\sin u \sin (-v) \\
& =\cos u \cos v-\sin u \sin v \tag{2.7}
\end{align*}
$$

Thus, letting $v=\pi / 2$,

$$
\begin{equation*}
\cos \left(u+\frac{\pi}{2}\right)=-\sin u \tag{2.8}
\end{equation*}
$$

It follows

$$
\begin{align*}
\sin (x+y) & =-\cos \left(x+\frac{\pi}{2}+y\right) \\
& =-\left[\cos \left(x+\frac{\pi}{2}\right) \cos y-\sin \left(x+\frac{\pi}{2}\right) \sin y\right] \\
& =\sin x \cos y+\sin y \cos x \tag{2.9}
\end{align*}
$$

Then using (2.3), that $\sin (-y)=-\sin (y)$ and $\cos (-x)=\cos (x)$, this implies

$$
\begin{equation*}
\sin (x-y)=\sin x \cos y-\cos x \sin y . \tag{2.10}
\end{equation*}
$$

Figure 6: Kuttler, p. 60

- For (iii), use $1-\cos x=2 \sin ^{2}(x / 2)$ from exercise 2c and $|\sin x| \leq|x|$ from (37), so that

$$
\forall x>0, \quad 0 \leq \frac{1-\cos x}{x}=\frac{2 \sin ^{2}(x / 2)}{x} \leq \frac{2(x / 2)^{2}}{x}=\frac{x}{2}
$$

Now apply (32). For $y=-x,(1-\cos y) / y=-(1-\cos x) / x$.
3. Let $(X, d)$ be a metric space, $E$ be a subset of $X$ and $f$ a real-valued function with domain $E$. State the definition of limit of $f$ as in Stoll, and also using "punctured neighborhood" notation.
ANS: (Stoll, Def 4.1.1) Suppose that $p$ is a limit point of $E$. The function $f$ has a limit at $p$ if there exists a number $L \in \mathbb{R}$ such that, given any $\epsilon>0, \exists \delta>0$ for which $|f(x)-L|<\epsilon$ for all points $x \in E$ satisfying $0<d(x, p)<\delta$. The constraint on $x$ can be written in terms of a "punctured neighborhood" of $p$ as (with $\ni$ being a shortcut for "such that")

$$
\exists L \in \mathbb{R} \ni \forall \epsilon>0, \exists \delta>0 \ni \forall x \in B_{\delta}(p) \cap(E \backslash\{p\}),|f(x)-L|<\epsilon
$$

4. (Stoll, p. 143, \#14) Let $f, g$ be defined on $E \subset \mathbb{R}$ and let $p$ be a limit point of $E$.
(a) If $\lim _{x \rightarrow p} f(x)$ and $\lim _{x \rightarrow p}[f(x)+g(x)]$ exist, prove that $\lim _{x \rightarrow p} g(x)$ exists.

ANS: As the limit of a sum of functions equals the sum of the respective limits (Stoll, Thm 4.1.6(a)),

$$
\lim _{x \rightarrow p} g(x)=\lim _{x \rightarrow p}[f(x)+g(x)]-\lim _{x \rightarrow p} f(x)
$$

and both terms on the rhs are assumed to exist.
(b) If $\lim _{x \rightarrow p} f(x)$ and $\lim _{x \rightarrow p}[f(x) g(x)]$ exist, does it follow that $\lim _{x \rightarrow p} g(x)$ exists?

ANS: Similar to (a), and appealing to Stoll, Thm 4.1.6(c),

$$
\lim _{x \rightarrow p} g(x)=\lim _{x \rightarrow p} \frac{f(x) g(x)}{f(x)}=\frac{\lim _{x \rightarrow p} f(x) g(x)}{\lim _{x \rightarrow p} f(x)}
$$

and both terms on the rhs are assumed to exist, but in order for $\lim _{x \rightarrow p} g(x)$ to exist in $\mathbb{R}$, the denominator cannot be zero. So, $\lim _{x \rightarrow p} g(x)$ exists if $\lim _{x \rightarrow p} f(x) \neq 0$. As an example, let $E=\mathbb{R}, f(x)=\sin (x), g(x)=1 / x$, and $p=0$. Then, from exercise $2 \mathrm{~d}, \lim _{x \rightarrow 0} f(x) g(x)=1$ exists, but $\lim _{x \rightarrow 0} g(x)$ does not exist in $\mathbb{R}$.
If we allow $p$ to tend to infinity (not a limit point of $\mathbb{R}$ ), we can take $E=\mathbb{R}, f(x)=\exp (-x)$, $g(x)=\exp (x)$, so that $(f g)(x) \equiv 1$, and $\lim _{x \rightarrow \infty} g(x)=\infty \notin \mathbb{R}$, so the result does not hold.
5. (Stoll, p. 143, \#9) Suppose $f:(a, b) \rightarrow \mathbb{R}, p \in[a, b]$, and $\lim _{x \rightarrow p} f(x)>0$. Prove $\exists \delta>0$ such that, $\forall x \in(a, b)$ with $0<|x-p|<\delta, f(x)>0$. (Notice continuity is not assumed.)

ANS: (Stoll) Let $L=\lim _{x \rightarrow p} f(x)>0$. Take $\epsilon=L / 2$. From the definition of limit (exercise 3), $\exists \delta>0$ such that, $\forall x \in B_{\delta}(p) \cap((a, b) \backslash\{p\})$,

$$
L / 2<f(x)<3 L / 2 \Leftrightarrow-L / 2<f(x)-L<L / 2 \Leftrightarrow 0<|f(x)-L|<L / 2=\epsilon .
$$

6. (Monotonicity of limits of functions) Let $A \subset \mathbb{R}$ and $f, g: A \rightarrow \mathbb{R}$ such that

$$
\lim _{\substack{x \rightarrow \alpha \\ x \in A}} f(x)=b \quad \text { and } \quad \lim _{\substack{x \rightarrow \alpha \\ x \in A}} g(x)=c
$$

(a) (Laczkovich and Sós, Thm 10.30) Prove: If $b<c$, then there exists a punctured neighborhood $\dot{U}$ of $\alpha$, denoted $\dot{U}(\alpha)$, such that, $\forall x \in \dot{U} \cap A, f(x)<g(x)$. Is the converse true?
ANS: Let $\epsilon:=(c-b) / 2$. Then $\exists \dot{U}_{1}(\alpha)$ such that, for $x \in A \cap \dot{U}_{1}(\alpha),|f(x)-b|<\epsilon$; and $\exists \dot{U}_{2}(\alpha)$ such that, for $x \in A \cap \dot{U}_{2}(\alpha),|g(x)-c|<\epsilon$. Let $\dot{U}(\alpha)=\dot{U}_{1}(\alpha) \cap \dot{U}_{2}(\alpha)$. Then

$$
x \in A \cap \dot{U}(\alpha) \Longrightarrow f(x)<b+\epsilon=b+\frac{c-b}{2}=\frac{b+c}{2}=c-\frac{c-b}{2}=c-\epsilon<g(x) .
$$

The converse is not true: If $f(x)<g(x)$ holds on a punctured neighborhood of $\alpha$, then we cannot conclude that $\lim _{x \rightarrow \alpha} f(x)<\lim _{x \rightarrow \alpha} g(x)$. If, for example, $f(x)=0$ and $g(x)=|x|$, then $f(x)<g(x)$ for all $x \neq 0$, but $\lim _{x \rightarrow 0} f(x)=\lim _{x \rightarrow 0} g(x)=0$.
(b) (Laczkovich and Sós, Thm 10.31) Now assume $f(x) \leq g(x)$ holds for all $x \in A \cap \dot{U}(\alpha)$. Prove: $b \leq c$. Is the converse true?
ANS: Let $\dot{U}(\alpha)$ be such that, $\forall x \in A \cap \dot{U}(\alpha), f(x) \leq g(x)$. Suppose that $b>c$. Then by part (a), $\exists \dot{V}(\alpha)$ such that, $\forall x \in A \cap \dot{V}(\alpha), f(x)>g(x)$. This, however, is impossible, because the set $A \cap \dot{U}(\alpha) \cap \dot{V}(\alpha)$ is nonempty, and $\forall x \in(A \cap \dot{U}(\alpha) \cap \dot{V}(\alpha)), f(x) \leq g(x)$.
The converse is not true: If $\lim _{x \rightarrow \alpha} f(x) \leq \lim _{x \rightarrow \alpha} g(x)$, then we cannot conclude that $f(x) \leq$ $g(x)$ holds in a punctured neighborhood of $\alpha$. If, for example, $f(x)=|x|$ and $g(x)=0$, then $\lim _{x \rightarrow 0} f(x) \leq \lim _{x \rightarrow 0} g(x)=0$, but $f(x)>g(x)$ for all $x \neq 0$.
7. (Stoll, Thm 4.1.8, and p. 143, \#12) Suppose $E$ is a subset of a metric space $X, p$ is a limit point of $E$, and $f, g$ are real-valued functions on $E$. Let $g$ be bounded on $E$, and $\lim _{x \rightarrow p} f(x)=0$. Prove: $\lim _{x \rightarrow p} f(x) g(x)=0$.
ANS: (Chen) Note that this is related to exercise 18 h . We can either use the sequential criterion or the definition of the limit of a function. Function $g$ is bounded on $E \Longrightarrow \exists M>0$ such that $\forall x \in E,|g(x)| \leq M$; while $\lim _{x \rightarrow p} f(x)=0 \Longrightarrow \forall \varepsilon>0, \exists \delta>0$ such that, $\forall x \in E, 0<d(x, p)<$ $\delta \Longrightarrow|f(x)|<\epsilon$. Then, $\forall x \in E, 0<d(x, p)<\delta \Longrightarrow|f(x) g(x)|=|f(x) \| g(x)|<M \epsilon$. Thus, $\lim _{x \rightarrow p} f(x) g(x)=0$.
8. (Stoll, p. 144, \#15) Let $E$ be a subset of a metric space, and let $p$ a limit point of $E$. Suppose $f$ is a bounded real-valued function on $E$ having the property that $\lim _{x \rightarrow p} f(x)$ does not exist. Prove that there exist distinct sequences $\left\{p_{n}\right\}$ and $\left\{q_{n}\right\}$ in $E$ with $p_{n} \rightarrow p$ and $q_{n} \rightarrow p$ such that $\lim _{n \rightarrow \infty} f\left(p_{n}\right)$ and $\lim _{n \rightarrow \infty} f\left(q_{n}\right)$ exist, but are not equal.
ANS: We begin with two illustrations, and then provide a proof.

- Here are two illustrations.
(a) Take a bounded function on an interval with a jump discontinuity for some $p \in(a, b)$, but is otherwise continuous. (A jump discontinuity is shown in Stoll, Fig. 4.9.) Let sequence $\left\{p_{n}\right\}$ approach $p$ from the left, and let sequence $\left\{q_{n}\right\}$ approach $p$ from the right.
(b) Let $E$ be any nonempty interval of $\mathbb{R}$, and let $f: E \rightarrow \mathbb{R}$ be the Dirichlet function, namely $f(x)=\chi_{\mathbb{Q}}(x)$. Let $\left\{p_{n}\right\} \in \mathbb{Q}$ with $p_{n} \rightarrow p$; and $\left\{q_{n}\right\} \in \mathbb{R} \backslash \mathbb{Q}$ with $q_{n} \rightarrow p$. Then $\lim _{n \rightarrow \infty} f\left(p_{n}\right)=$ $\lim _{n \rightarrow \infty} 1=1$, while $\lim _{n \rightarrow \infty} f\left(q_{n}\right)=\lim _{n \rightarrow \infty} 0=0$.
- Proof (Chen and Ralf): Recall the equivalence between the sequential criterion of the limit of a function and the definition given by Stoll, namely:
- (Stoll, Def 4.1.1) Let $(X, d)$ be a metric space, $E$ be a subset of $X$ and $f$ a real-valued function with domain $E$. Suppose that $p$ is a limit point of $E$. The function $f$ has a limit at $p$ if there exists a number $L \in \mathbb{R}$ such that given any $\epsilon>0$, there exists a $\delta>0$ for which $|f(x)-L|<\epsilon$ for all points $x \in E$ satisfying $0<d(x, p)<\delta$. If this is the case, we write

$$
\lim _{x \rightarrow p} f(x)=L \quad \text { or } \quad f(x) \rightarrow L \quad \text { as } \quad x \rightarrow p
$$

- (Stoll, Thm 4.1.3) Let $E$ be a subset of a metric space $X, p$ a limit point of $E$, and $f$ a real-valued function defined on $E$. Then

$$
\lim _{x \rightarrow p} f(x)=L \quad \text { if and only if } \quad \lim _{n \rightarrow \infty} f\left(p_{n}\right)=L
$$

for every sequence $\left\{p_{n}\right\}$ in $E$, with $p_{n} \neq p$ for all $n$, and $\lim _{n \rightarrow \infty} p_{n}=p$.
Since $f: E \rightarrow \mathbb{R}$ does not have a limit at the limit point $p \in E$, by the negation of the above theorem, either:
(a) for any sequence $\left\{p_{n}\right\}$ in $E$ converging to $p$ with $p_{n} \neq p \forall n \in \mathbb{N}$, $\lim _{n \rightarrow \infty} f\left(p_{n}\right)$ exists but there exist different sequences $\left\{q_{n}\right\},\left\{r_{n}\right\}$ in $E$ converging to $p$ (again with $q_{n} \neq p$ and $r_{n} \neq p \forall n \in \mathbb{N}$ ), for which $\lim _{n \rightarrow \infty} f\left(q_{n}\right) \neq \lim _{n \rightarrow \infty} f\left(r_{n}\right)$;
(b) there exists at least one sequence $\left\{p_{n}\right\}$ in $E$ converging to $p$ with $p_{n} \neq p \forall n \in \mathbb{N}$, for which $\lim _{n \rightarrow \infty} f\left(p_{n}\right)$ fails to exist.

In the first case, we can take the two sequences $\left\{q_{n}\right\},\left\{r_{n}\right\}$ and arrive at the conclusion of the theorem. In the second case, we use the boundedness of $f$ and the Bolzano-Weierstrass Theorem to produce two such sequences. Recall:
(Stoll, Cor 3.4.6) Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.
By the above corollary, as $f$ and therefore $\left\{f\left(p_{n}\right)\right\}$ is bounded, there exists a subsequence $\left\{p_{n_{k}}\right\}$ of $\left\{p_{n}\right\}$ such that $\left\{f\left(p_{n_{k}}\right)\right\}$ converges to some $a \in \mathbb{R}$. Since $\left\{f\left(p_{n}\right)\right\}$ does not converge to $a, \exists \epsilon>0$ such that for any given $m \in \mathbb{N}, \exists l \geq m$ with $\left|f\left(p_{l}\right)-a\right| \geq \epsilon$. For each $m \in \mathbb{N}$, choose such an index $l$. This way, we construct a subsequence $\left\{p_{l}\right\}$ such that $\left\{f\left(p_{l}\right)\right\}$ never enters $N_{\epsilon}(a)$. By the above corollary, as $\left\{f\left(p_{l}\right)\right\}$ is still bounded, there exists a subsequence $\left\{p_{l_{k}}\right\}$ of $\left\{p_{l}\right\}$ such that $\left\{f\left(p_{l_{k}}\right)\right\}$ converges to some point $b \in \mathbb{R}$. Clearly $a \neq b$ so that $\left\{p_{l_{k}}\right\},\left\{p_{n_{k}}\right\}$ are two sequences in $E$ converging to $p$ with $\lim _{k \rightarrow \infty} f\left(p_{l_{k}}\right) \neq \lim _{k \rightarrow \infty} f\left(p_{n_{k}}\right)$.
9. (Stoll, p. 144, \#19) Let $f: \mathbb{R} \rightarrow \mathbb{R}$ satisfy $f(x+y)=f(x)+f(y)$ for all $x, y \in \mathbb{R}$. If $\lim _{x \rightarrow 0} f(x)$ exists, prove that
(a) $\lim _{x \rightarrow 0} f(x)=0$, and
(b) $\lim _{x \rightarrow p} f(x)$ exists for every $p \in \mathbb{R}$.

ANS: For (a),

$$
\lim _{x \rightarrow 0} f(x)=\lim _{x \rightarrow 0} f\left(\frac{x}{2}+\frac{x}{2}\right)=\lim _{x \rightarrow 0} f\left(\frac{x}{2}\right)+\lim _{x \rightarrow 0} f\left(\frac{x}{2}\right)=2 \lim _{x \rightarrow 0} f(x),
$$

which implies $\lim _{x \rightarrow 0} f(x)=0$. For (b), let $p \in \mathbb{R}$, and note that $f(p+x)=f(p)+f(x)$. Take the limit as $x \rightarrow 0$, which gives (via part (a))

$$
\lim _{x \rightarrow p} f(x)=\lim _{x \rightarrow 0} f(p+x)=\lim _{x \rightarrow 0} f(p)+\lim _{x \rightarrow 0} f(x)=f(p)+\lim _{x \rightarrow 0} f(x)
$$

The rhs exists, so the lhs exists. As $p \in \mathbb{R}$ was arbitrary, the lhs exists for every $p \in \mathbb{R}$.
10. (Continuous function limit results)
(a) State the definition of continuity of a function $f: E \rightarrow \mathbb{R}$ in $\epsilon-\delta$ terms (e.g., Stoll, Def 4.2.1) and in terms of sequences. Justify the expression: $\lim _{x \rightarrow p} f(x)=f\left(\lim _{x \rightarrow p} x\right)$.
ANS:

- $(\epsilon-\delta)$ A function $f: E \rightarrow \mathbb{R}$ is continuous at $p \in E$ if given $\epsilon>0, \exists \delta>0$ such that

$$
\forall x \in N_{\delta}(p) \cap E, \quad f(x) \in N_{\epsilon}(f(p)) .
$$

- ( $\epsilon-\delta)$ One can also write this in terms of distance measures as follows. Let $\left(X, d_{X}\right)$ and $\left(Y, d_{Y}\right)$ be metric spaces; and let $f: X \rightarrow Y$. Let $X=E$ and $Y=\mathbb{R}$. Function $f: E \rightarrow \mathbb{R}$ is continuous at $p \in E$ if

$$
\begin{equation*}
(\forall \epsilon>0)(\exists \delta>0)(\forall y \in E)\left[d_{X}(p, y)<\delta \Longrightarrow d_{Y}(f(p), f(y))<\epsilon\right] . \tag{40}
\end{equation*}
$$

- (sequential) A function $f: E \rightarrow \mathbb{R}$ is continuous at $p \in E$ if

$$
\left\{x_{n}\right\} \text { any sequence in } E \text { such that } x_{n} \rightarrow p \Longrightarrow f\left(x_{n}\right) \rightarrow f(p)
$$

- The function $f$ is continuous on $E$ if $f$ is continuous at every point $p \in E$.
- If $p \in E$ is a limit point of $E$, then $(f$ is continuous at $p) \Leftrightarrow \lim _{x \rightarrow p} f(x)=f\left(\lim _{x \rightarrow p} x\right)$. This follows from the definition of limit in, e.g., exercise 3, with $L=f(p)$.
(b) (Petrovic, Thm 3.6.12) Let $f$ be a continuous function on an interval $(a, b)$ and let $c \in(a, b)$. Prove: If $f(c)>0$, then $\exists \delta>0$ such that $f(x)>0$ for $x \in(c-\delta, c+\delta)$. What can you say about the converse? (See also, e.g., Stoll, p. 157, \#21; Laczkovich and Sós, Coro 10.32; Ghorpade and Limaye, Prop 3.2 (iv); and Sasane, \#3.5.)
NOTE: (Ralf) As it is worded, the converse is actually true. The converse of the given statement is: $\forall x \in(c-\delta, c+\delta), f(x)>0$. Trivially, $c$ is a part of that interval and therefore $f(c)>0$. Instead, rephrase the question: Suppose $\exists \delta>0$ such that $\forall x \in(c-\delta, c+\delta) \backslash\{c\}, f(x)>0$. Can we conclude that $f(c)>0$ ?
ANS:
Proof I, using sequential argument: Suppose to the contrary that no such $\delta$ exists. Then, for every $\delta>0$, there exists $x \in(c-\delta, c+\delta)$ such that $f(x) \leq 0$. If we take $\delta=1 / n$, we obtain a sequence $\left\{x_{n}\right\}$ in $(c-1 / n, c+1 / n)$ with $f\left(x_{n}\right) \leq 0$. The inequality $\left|x_{n}-c\right|<1 / n$ shows that the sequence $\left\{x_{n}\right\}$ converges to $c$, while the continuity of $f$ implies that the sequence $\left\{f\left(x_{n}\right)\right\}$ converges to $f(c)$. As $f\left(x_{n}\right) \leq 0$, the result of exercise 17 implies that $f(c) \leq 0$, which contradicts the assumption that $f(c)>0$.
Proof II (Ralf), using $\epsilon-\delta$ argument: Let $\epsilon:=f(c)>0$. By continuity, $\exists \delta>0$ such that

$$
\begin{equation*}
|f(x)-f(c)|<\underbrace{f(c)}_{=\epsilon}, \quad \forall x \in(a, b) \text { with } d(x, c)<\delta . \tag{41}
\end{equation*}
$$

That is, $\forall x \in(a, b)$ with $d(x, c)<\delta$, we have $0=f(c)-f(c)<f(x)$.
Proof III (Chen), using existing results: Function $f$ is continuous on $(a, b)$ and thus at $c \in(a, b)$. Since $c$ is a limit point of $(a, b), \lim _{x \rightarrow c} f(x)=f(c)>0$. The result now follows from exercise 6a. More specifically, let $g(x) \equiv 0$, and $\lim _{x \rightarrow c} f(x)=f(c)>0=\lim _{x \rightarrow c} g(x)$.
Further, the next question in the exercise, part (c), follows from exercise 6b.
The (correct) converse is almost true. When $f(x)>0$ for all $x, 0<|x-c|<\delta$, it does not follow that $f(c)>0$. Example: $f(x)=x^{2}$, and $c=0$. However, it does follow that $f(c) \geq 0$.
(c) (Petrovic, Thm 3.6.14) Let $f$ be a continuous function at a point $c$ in $(a, b)$. If $f(x) \geq 0$ on $(a, c) \cup(c, b)$ then $f(c) \geq 0$.
ANS: Let $\left\{a_{n}\right\}$ be a sequence in $(a, b)$ converging to $c$, and $a_{n} \neq c$. Then $f\left(a_{n}\right) \geq 0$, and $\lim f\left(a_{n}\right) \geq 0$ from exercise 17. Continuity of $f$ at $c$ implies $f(c) \geq 0$.
11. (Function limit analogue to Cauchy's criterion) (Laczkovich and Sós, Thm 10.34) Let $f$ be defined on a punctured neighborhood of $\alpha$. Prove: $\lim _{x \rightarrow \alpha} f(x)$ exists, and is finite, iff $\forall \epsilon>0, \exists \dot{U}$ such that

$$
\begin{equation*}
\forall x_{1}, x_{2} \in \dot{U}_{\alpha}, \quad\left|f\left(x_{1}\right)-f\left(x_{2}\right)\right|<\epsilon, \tag{42}
\end{equation*}
$$

where $\dot{U}_{\alpha}$ is a punctured neighborhood of $\alpha$.
ANS: $(\Longrightarrow)$ Suppose that $\lim _{x \rightarrow \alpha} f(x)=b \in \mathbb{R}$, and let $\epsilon>0$ be fixed. Then $\exists \dot{U}_{\alpha}$ such that $\forall x \in \dot{U}_{\alpha},|f(x)-b|<\epsilon / 2$. The triangle inequality shows that (42) holds for all $x_{1}, x_{2} \in \dot{U}_{\alpha}$.
$(\Longleftarrow)$ Now suppose (42) holds. If $x_{n} \rightarrow \alpha$ and $x_{n} \neq \alpha$ for all $n \in \mathbb{N}$, then the sequence $f\left(x_{n}\right)$ satisfies the Cauchy criterion. Indeed, for a given $\epsilon$, choose a punctured neighborhood $\dot{U}_{\alpha}$ such that (42) holds for all $x_{1}, x_{2} \in \dot{U}_{\alpha}$. Since $x_{n} \rightarrow \alpha$ and $x_{n} \neq \alpha$ for all $n, \exists N \in \mathbb{N}$ such that, $\forall n \geq N$, $x_{n} \in \dot{U}_{\alpha}$. If $n, m \geq N$, then by (42), we have $\left|f\left(x_{n}\right)-f\left(x_{m}\right)\right|<\epsilon$. From Cauchy's criterion for sequences (e.g., Stoll, Thm 3.6.2(a) and Thm 3.6.5), the sequence $\left\{f\left(x_{n}\right)\right\}$ is convergent.

Fix a sequence $x_{n} \rightarrow \alpha$ that satisfies $x_{n} \neq \alpha$ for all $n$, and let $\lim _{n \rightarrow \infty} f\left(x_{n}\right)=b$. If $y_{n} \rightarrow \alpha$ is another sequence satisfying $y_{n} \neq \alpha$ for all $n$, then the combined sequence $\left(x_{1}, y_{1}, x_{2}, y_{2}, \ldots\right)$ also satisfies this assumption, and so the sequence of function values $s=\left(f\left(x_{1}\right), f\left(y_{1}\right), f\left(x_{2}\right), f\left(y_{2}\right), \ldots\right)$ is also convergent. Since $\left\{f\left(x_{n}\right)\right\}$ is a subsequence of this, the limit of $s$ can only be $b$. On the other hand, $\left\{f\left(y_{n}\right)\right\}$ is also a subsequence of $s$, so $f\left(y_{n}\right) \rightarrow b$. This holds for all sequences $y_{n} \rightarrow \alpha$ for which $y_{n} \neq \alpha$ for all $n$, so $\lim _{x \rightarrow a} f(x)=b$. (This latter result is, e.g., Stoll, Thm 4.1.3 and Coro 4.1.4).
12. (Topological characterization of continuity)
(a) State the topological characterization of continuity in terms of open sets (e.g., Stoll, Thm 4.2.6). Note: This is very important, often taken to be the definition in more advanced presentations. ANS: Let $X, Y$ be metric spaces; $E \subset X$. Then a function $f$ is continuous on $E$ if and only if $f^{-1}(V)$ is open in $E$ for every open subset $V \subset Y$.
(b) Prove the result in part (a) taking, for simplicity, $E=X$.

ANS: (Ash, Thm 4.1.6)
Let $\left(X, d_{X}\right)$ and $\left(Y, d_{Y}\right)$ be metric spaces; and let $f: X \rightarrow Y$. The function $f$ is continuous on $X$ if and only if for each open set $V \subset Y$ the pre-image $f^{-1}(V)$ is an open subset of $X$.
Proof. Assume $f$ continuous. Let $x$ belong to $f^{-1}(V)$, where $V$ is open in $Y$. Then $f(x) \in V$, so for some $\epsilon>0, B_{\epsilon}(f(x)) \subset V$. If $\delta>0$ is as given in (40), then

$$
y \in B_{\delta}(x) \Longrightarrow f(y) \in B_{\epsilon}(f(x)) ; \text { hence, } f(y) \in V
$$

Thus, $y \in f^{-1}(V)$, proving that $B_{\delta}(x) \subseteq f^{-1}(V)$. Therefore $f^{-1}(V)$ is open.
Conversely, assume $V$ open implies $f^{-1}(V)$ open. If $x \in X$, we show that $f$ is continuous at $x$. Given $\epsilon>0, f(x) \in B_{\epsilon}(f(x))$, which is an open set $V$. Thus, $x \in f^{-1}(V)$, which is open by hypothesis, so $B_{\delta}(x) \subset f^{-1}(V)$ for some $\delta>0$. Consequently,

$$
y \in B_{\delta}(x) \Longrightarrow y \in f^{-1}(V) \Longrightarrow f(y) \in V=B_{\epsilon}(f(x))
$$

in other words, $d_{X}(x, y)<\delta \Longrightarrow d_{Y}(f(x), f(y))<\epsilon$. From (40), $f$ is continuous.
(c) Give an example showing that the direct image of an open set under a continuous function need not be open.
ANS: Example: Let $f: \mathbb{R} \rightarrow \mathbb{R}, f(x)=\sin x$, and $A=(0,2 \pi)$ open. Then $f(A)=[-1,1]$.
(d) Prove that the result holds also in terms of closed sets. (Also Stoll, p. 157, \#26)

ANS: We provide two solutions.

- (Field, Thm 7.11.3): Let $U \subset Y$ be open, and $F \subset Y$ be closed. Observe $f^{-1}(Y \backslash U)=$ $f^{-1}(Y) \backslash f^{-1}(U)=X \backslash f^{-1}(U)$, and $f^{-1}(Y \backslash F)=X \backslash f^{-1}(F)$.
NOTE: Augmentation from Chen for Field's rather cryptic answer: Let $\left(X, d_{X}\right),\left(Y, d_{Y}\right)$ be metric spaces, $E \subset X, f: E \rightarrow Y$. We already know that $f$ is continuous on $E \Longleftrightarrow \forall$ open subsets $V$ of $Y, f^{-1}(V)$ is open in $E$.
First show that, $\forall F \subset Y, f^{-1}(Y \backslash F)=E \backslash f^{-1}(F)$. Take any $x \in E$. Then,

$$
x \in f^{-1}(Y \backslash F) \Longleftrightarrow f(x) \in Y \backslash F \Longleftrightarrow f(x) \notin F \Longleftrightarrow x \notin f^{-1}(F) \Longleftrightarrow x \in E \backslash f^{-1}(F)
$$

Next, note $F \subset Y$ closed in $Y \Longleftrightarrow Y \backslash F$ open in $Y$. As $f$ continuous on $E \Longleftrightarrow f^{-1}(Y \backslash F)$ open in $E$, we finally have the desired relation: The topological characterization of continuity in terms of open sets is equivalent to that in terms of closed sets:

$$
F \subset Y \text { closed in } Y \Longleftrightarrow E \backslash f^{-1}(F) \text { open in } E \Longleftrightarrow f^{-1}(F) \text { closed in } E .
$$

- (Ash, Thm 4.1.6) We must show that the preimage of each closed set is closed if and only if the preimage of each open set is open. Suppose that for each closed $C \subseteq \Omega^{\prime}, f^{-1}(C)$ is closed, and assume $V$ is an open subset of $\Omega^{\prime}$. Then $V^{c}$ is closed, so $f^{-1}\left(V^{c}\right)$ is closed. But by (2),

$$
f^{-1}\left(V^{c}\right)=\left[f^{-1}(V)\right]^{c}
$$

Thus, $\left[f^{-1}(V)\right]^{c}$ is closed, so $f^{-1}(V)$ is open. Conversely, if the preimage of each open set is open and $C$ is a closed subset of $\Omega^{\prime}$, then $C^{c}$ is open, and hence $f^{-1}\left(C^{c}\right)=\left[f^{-1}(C)\right]^{c}$ is open. Therefore $f^{-1}(C)$ is closed.
(e) (Stoll, Example 4.2.7) Let $D=[0, \infty)$, and $f: D \rightarrow \mathbb{R}$ given by $f(x)=\sqrt{x}$. Show $f$ is continuous by demonstrating it satisfies the topological characterization of continuity in terms of open sets. Crucially, explicitly state every theorem you invoke. I count four of them.
ANS: The four required theorems are:
i. Stoll, Theorem 1.7.14(b). Let $f$ be a function from $X$ into $Y$, and let $A$ be a nonempty set. If $\left\{B_{\alpha}\right\}_{\alpha \in A}$ is a family of subsets of $Y$, then

$$
f^{-1}\left(\bigcup_{\alpha \in A} B_{\alpha}\right)=\bigcup_{\alpha \in A} f^{-1}\left(B_{\alpha}\right), \quad f^{-1}\left(\bigcap_{\alpha \in A} B_{\alpha}\right)=\bigcap_{\alpha \in A} f^{-1}\left(B_{\alpha}\right)
$$

this being given in (1).
ii. Stoll, Theorem 2.2.9(a). Let $(X, d)$ be a metric space. Then for any collection $\left\{O_{\alpha}\right\}_{\alpha \in A}$ of open subsets of $X, \bigcup_{\alpha \in A} O_{\alpha}$ is open.
iii. Stoll, Theorem 2.2.20. If $U$ is an open subset of $\mathbb{R}$, then there exists a finite or countable collection $\left\{I_{n}\right\}$ of pairwise disjoint open intervals such that $U=\bigcup_{n} I_{n}$.
iv. Stoll, Theorem 2.2.23(a). Let $Y$ be a subset of a metric space $X$. A subset $U$ of $Y$ is open in $Y$ if and only if $U=Y \cap O$ for some open subset $O$ of $X$. (Note $U \subset Y \subset X$ and $O \subset X$.)

Let $V=(a, b)$ with $a<b$. Then

$$
f^{-1}(V)= \begin{cases}\emptyset, & b \leq 0 \\ {\left[0, b^{2}\right),} & a \leq 0<b \\ \left(a^{2}, b^{2}\right), & 0<a\end{cases}
$$

Clearly $\emptyset$ and $\left(a^{2}, b^{2}\right)$ are open subsets of $\mathbb{R}$ and hence also of $[0, \infty)$.
Although $\left[0, b^{2}\right)$ is not open in $\mathbb{R}$, note that $\left[0, b^{2}\right)=\left(-b^{2}, b^{2}\right) \cap[0, \infty)$, and by Theorem 2.2.23(a), $\left[0, b^{2}\right)$ is open in $[0, \infty)$.
If $V$ is an arbitrary open subset of $\mathbb{R}$, then, by Theorem $2.2 .20, V=\bigcup_{n} I_{n}$, where $\left\{I_{n}\right\}$ is a finite or countable collection of open intervals. As $f^{-1}(V)=\bigcup_{n} f^{-1}\left(I_{n}\right)$ (Theorem 1.7.14(b)) and each $f^{-1}\left(I_{n}\right)$ is open in $[0, \infty)$, from Theorem 2.2.9(a), $f^{-1}(V)$ is open in $[0, \infty)$. Therefore $f(x)=\sqrt{x}$ is continuous on $[0, \infty)$.
(f) (Stoll, p. 176, \#14(a)) We know from part 12c that the direct image of an open set under a continuous function need not be open. Let $I \subset \mathbb{R}$ be an open interval and let $f: I \rightarrow \mathbb{R}$ be strictly increasing and continuous on $I$. If $U \subset I$ is open, prove that $f(U)$ is open.
ANS: We require:

- (Stoll, Thm 1.7.14(a)) Let $f: X \rightarrow Y$, and let $A$ be a nonempty set. If $\left\{E_{\alpha}\right\}_{\alpha \in A}$ is a family of subsets of $X$, then $f\left(\cup_{\alpha \in A} E_{\alpha}\right)=\cup_{\alpha \in A} f\left(E_{\alpha}\right)$.
- (Stoll, Thm 2.2.9) For collection $\left\{O_{\alpha}\right\}_{\alpha \in A}$ of open subsets of $(X, d), \cup_{\alpha \in A} O_{\alpha}$ is open.
- (Stoll, Thm 2.2.20) If $U$ is an open subset of $\mathbb{R}$, then there exists a finite or countable collection $\left\{I_{n}\right\}$ of pairwise disjoint open intervals such that $U=\cup_{n} I_{n}$.
- (Stoll, Thm 4.4.7) Let $I \subset \mathbb{R}$ be an open interval and let $f: I \rightarrow \mathbb{R}$ be monotone increasing on $I$. Then $f(p+)$ and $f(p-)$ exists for every $p \in I$ and

$$
\sup _{x<p} f(x)=f(p-) \leq f(p) \leq f(p+)=\inf _{p<x} f(x)
$$

Furthermore, if $p<q, p, q \in I$, then $f(p+) \leq f(q-)$.

- (Stoll, Coro 4.2.12; Fitzpatrick, Thm 3.14) Let $I$ be an interval and suppose that the function $f: I \rightarrow \mathbb{R}$ is continuous. Then its image $f(I)$ also is an interval.

Proof (Stoll, augmented by Chen): First, let $U=(a, b) \subset I$, with $-\infty \leq a<b \leq \infty$. Then $f((a, b))=(f(a+), f(b-))$ is open in $\mathbb{R}$. Next, for any open set $U \subset I$, write $U=\cup_{n} I_{n}$, where $\left\{I_{n}\right\}$ is an at most countable collection of open intervals. Then $f(U)=f\left(\cup_{n} I_{n}\right)=\cup_{n} f\left(I_{n}\right)$ is open.
13. (Uniform Continuity)
(a) State the definition of uniform continuity in $\epsilon-\delta$ terms; and also in terms of sequences.

ANS: Let $X$ and $Y$ be metric spaces with respective metrics $d_{X}$ and $d_{Y}$. Let $D \subset X$ and $f: D \rightarrow Y$. Function $f$ is uniformly continuous on $D$ if:

$$
\begin{align*}
\forall \epsilon>0, \exists \delta>0 & \text { such that } \forall x, y \in D, d_{X}(x, y)<\delta
\end{aligned} \begin{aligned}
& \Longrightarrow d_{Y}(f(x), f(y))<\epsilon  \tag{43}\\
& \text { and } \forall\left\{x_{n}\right\},\left\{y_{n}\right\} \subset D,\left(x_{n}-y_{n}\right)
\end{align*}>0 \Longrightarrow\left[f\left(x_{n}\right)-f\left(y_{n}\right)\right] \rightarrow 0 .
$$

(b) Prove their equivalence.

ANS: (Ghorpade and Limaye, Prop 3.22):
$\Leftarrow$ Suppose $\exists \epsilon>0$ such that, $\forall \delta>0, \exists x, y \in D$ such that $|x-y|<\delta$, but $|f(x)-f(y)| \geq \epsilon$. Considering $\delta:=1 / n$ for $n \in \mathbb{N}$, we obtain sequences $\left\{x_{n}\right\}$ and $\left\{y_{n}\right\}$ in $D$ such that $\left|x_{n}-y_{n}\right|<$ $1 / n$ but $\left|f\left(x_{n}\right)-f\left(y_{n}\right)\right| \geq \epsilon, \forall n \in \mathbb{N}$. Then $x_{n}-y_{n} \rightarrow 0$, but $f\left(x_{n}\right)-f\left(y_{n}\right) \nrightarrow 0$. This is a contradiction.
$\Rightarrow$ Let $\left\{x_{n}\right\}$ and $\left\{y_{n}\right\}$ be any sequences in $D$ such that $\left(x_{n}-y_{n}\right) \rightarrow 0$. Let $\epsilon>0$. Then $\exists \delta>0$ such that $|f(x)-f(y)|<\epsilon$, whenever $x, y \in D$ and $|x-y|<\delta$. As $\left(x_{n}-y_{n}\right) \rightarrow 0$, $\exists n_{0} \in \mathbb{N}$ such that $\forall n \geq n_{0},\left|x_{n}-y_{n}\right|<\delta$. But then $\forall n \geq n_{0},\left|f\left(x_{n}\right)-f\left(y_{n}\right)\right|<\epsilon$. Thus $\left[f\left(x_{n}\right)-f\left(y_{n}\right)\right] \rightarrow 0$.
(c) State a sufficient condition in terms of sequences to confirm that a function is not uniformly continuous.
ANS: As in Pons, Thm 4.4.8: Let $f: A \rightarrow \mathbb{R}$. Then $f$ is not uniformly continuous on $A$ if $\exists\left\{x_{n}\right\},\left\{y_{n}\right\} \in A$ and $\exists M>0$ such that $\left(\left|x_{n}-y_{n}\right|\right) \rightarrow 0$ as $n \rightarrow \infty$ while, $\forall n \in \mathbb{N}$, $\left|f\left(x_{n}\right)-f\left(y_{n}\right)\right| \geq M$.
14. (Continuity and compactness) Let $X$ and $Y$ be metric spaces with respective metrics $d_{X}$ and $d_{Y}$, and assume that $f: X \rightarrow Y$ is continuous.
(a) Prove: If $K$ is a compact subset of $X$, then $f(K)$ is a compact subset of $Y$. Hint: Use topological compactness.
ANS: Let $\left\{V_{i}\right\}_{i \in J}$ be any open cover of $f(K)$. Each set $U_{i}=f^{-1}\left(V_{i}\right)$ is open, and $\left\{U_{i}\right\}_{i \in J}$ is an open cover of $K$. As $K$ is compact, this cover admits a finite subcover $\left\{U_{i_{1}}, \ldots, U_{i_{N}}\right\}$. But then $\left\{V_{i_{1}}, \ldots, V_{i_{N}}\right\}$ is a finite subcover of $f(K)$, so $f(K)$ is compact.
(b) Give an example of a function $f: \mathbb{R} \rightarrow \mathbb{R}$ that is bounded and continuous but not uniformly continuous. Hint: Consider $f(x)=\sin \left(x^{2}\right)$, and recall the Mean Value Theorem (MVT), e.g., Stoll, Thm 5.2.6:

If $f:[a, b] \rightarrow \mathbb{R}$ is continuous on $[a, b]$ and differentiable on $(a, b)$, then $\exists c \in(a, b)$ such that $f(b)-f(a)=f^{\prime}(c)(b-a)$.
ANS: From https://math.stackexchange.com/questions/23443. Let $x, y \in \mathbb{R}$. The MVT implies

$$
\begin{equation*}
\exists k \in(x, y) \cup(y, x) \text { such that }\left|\sin \left(x^{2}\right)-\sin \left(y^{2}\right)\right|=2|k|\left|\cos \left(k^{2}\right)\right||x-y| \tag{44}
\end{equation*}
$$

Let $\epsilon>0$. From (43), we need to show, $\forall x, y \in \mathbb{R}$,

$$
\exists \delta>0 \text { such that }|x-y|=d_{X}(x, y)<\delta \Longrightarrow d_{Y}(f(x), f(y))=\left|\sin \left(x^{2}\right)-\sin \left(y^{2}\right)\right|<\epsilon
$$

Observe from (44) that, as $x, y \rightarrow \infty, k$ is not bounded, i.e., not less than $|x-y|$. Thus $\nexists \delta>0$ that satisfies (43).
15. (IVT)
(a) State (not prove) the Intermediate Value Theorem (IVT), e.g., Stoll, Thm 4.2.11.

ANS: Let $f:[a, b] \rightarrow \mathbb{R}$ be continuous. Suppose $f(a)<f(b)$. If $\gamma$ is a number satisfying $f(a)<\gamma<f(b)$, then $\exists c \in(a, b)$ such that $f(c)=\gamma$.
(b) State why we need continuity at the endpoints of the closed interval.

ANS: We need the left and right limits to agree with the function values at those points: Without this, we could move the points $f(a)$ and $f(b)$, and the result would not hold.
(c) What can we use the contrapositive of the IVT for?

ANS: The contrapositive implies that, if we know there is no $c$ such that $f(c)=0$ ( 0 being a useful special case), then either both $f(a)$ and $f(b)$ are positive, or negative.
(d) Prove: For every real number $\gamma>0$ and every positive integer $n$, there exists a unique positive real number $y$ so that $y^{n}=\gamma$. (Stoll, Coro 4.2.13; Al-Gwaiz and Elsanousi, Thm 2.6)
ANS: Proof. Stoll writes "That $y$ is unique is clear." Before progressing, let's process this.

- Uniqueness 1: Let $\epsilon>0$, and take $y, z \in \mathbb{R}$ such that $y=z+\epsilon>z>0$, and assume $y^{n}=z^{n}=\gamma$. From the Binomial Theorem, we have the contradiction $(z+\epsilon)^{n}>z^{n}$.
- Uniqueness 2: Assume $y_{1}$ and $y_{2}$ satisfy $y_{1}^{n}=y_{2}^{n}=\gamma$. Then $g:=\log (\gamma)$ satisfies $g / n=$ $\log \left(y_{1}\right)=\log \left(y_{2}\right)$, or, $\log \left(y_{1} / y_{2}\right)=0$, i.e., $y_{1}=y_{2}$.
- Uniqueness 3: From Junghenn, p. 18, \#15, and p. 7, \#4a: Page 15 says "The existence of $b^{1 / n}$ is an easy consequence of the IVT. Uniqueness follows from Exercise 1.2.4(a)."
Continuing, let $f(x)=x^{n}$, which is (via, e.g.., Stoll, Thm 4.2.3(b)) continuous on $\mathbb{R}$. For $\gamma>0$, let $a=0$ and $b=\gamma+1$, that $f(a)=0<\gamma<(\gamma+1)^{n}=f(b)$. Then $f$ satisfies the IVT. Thus $\exists y$ with $0<y<\gamma+1$, such that $f(y)=y^{n}=\gamma$.

16. (Maximum and minimum of two functions)
(a) Suppose $f: E \rightarrow \mathbb{R}$; $p$ is a limit point of $E$; and $\lim _{x \rightarrow p} f(x)=L$. Prove: $\lim _{x \rightarrow p}|f(x)|=|L|$. ANS: Let $\epsilon>0$. As $\lim f(x)=L$ exists by hypothesis, $\exists \delta>0$ such that, $\forall x \in E$ with $0<d(x, p)<\delta,|f(x)-L|<\epsilon$. By the reverse triangle inequality,

$$
\begin{equation*}
\forall x \in E \text { with } 0<d(x, p)<\delta, \quad \| f(x)|-|L|| \leq|f(x)-L|<\epsilon . \tag{45}
\end{equation*}
$$

(b) Let $E \subset \mathbb{R}$, and suppose $f: E \rightarrow \mathbb{R}$ is continuous at $p \in E$. Prove that $|f|$ is continuous at $p$. Is the converse true? (Stoll, p. 156, \#6)
ANS: Take $L=f(p)$ in part (a). The converse is not true. Take, e.g., $f(x)$ to be -1 if $x$ is irrational, and 1 if $x$ is rational.
(c) Prove: Let $\left\{a_{n}\right\},\left\{b_{n}\right\} \in \mathbb{R}$ and let $a, b \in \mathbb{R}$ such that $a_{n} \rightarrow a$ and $b_{n} \rightarrow b$. Then $\max \left\{a_{n}, b_{n}\right\} \rightarrow$ $\max \{a, b\}$ and $\min \left\{a_{n}, b_{n}\right\} \rightarrow \min \{a, b\}$. (Ghorpade and Limaye, Prop 2.4(ii))
ANS: Let $\epsilon>0$ be given. Then $\exists n_{1}, n_{2} \in \mathbb{N}$ such that

$$
\forall n \geq n_{1}, \quad a-\epsilon<a_{n}<a+\epsilon, \quad \text { and } \quad \forall n \geq n_{2}, \quad b-\epsilon<b_{n}<b+\epsilon
$$

Let $n_{0}=\max \left\{n_{1}, n_{2}\right\}$. Then $\forall n \geq n_{0}, \max \{a-\epsilon, b-\epsilon\}<\max \left\{a_{n}, b_{n}\right\}<\max \{a+\epsilon, b+\epsilon\}$. As $\max \{a-\epsilon, b-\epsilon\}=\max \{a, b\}-\epsilon$ and $\max \{a+\epsilon, b+\epsilon\}=\max \{a, b\}+\epsilon$, it follows that $\max \left\{a_{n}, b_{n}\right\} \rightarrow \max \{a, b\}$. The proof for min is similar.
(d) Let $D \subset \mathbb{R}$, and let $f, g: D \rightarrow \mathbb{R}$ be functions continuous at $c \in D$. Prove:
(Ghorpade and Limaye, Prop 3.3)
(i) $|f|$ is continuous at $c$.
(ii) $\max \{f, g\}$ and $\min \{f, g\}$ are continuous at $c$.

ANS: Let $\left\{x_{n}\right\}$ be a sequence in $D$ such that $x_{n} \rightarrow c$.
(i) Then $f\left(x_{n}\right) \rightarrow f(c)$, and by part (a), $\left|f\left(x_{n}\right)\right| \rightarrow|f(c)|$.
(ii) Also, $g\left(x_{n}\right) \rightarrow g(c)$ as well, so by part (c),
$\max \left\{f\left(x_{n}\right), g\left(x_{n}\right)\right\} \rightarrow \max \{f(c), g(c)\}$ and $\min \left\{f\left(x_{n}\right), g\left(x_{n}\right)\right\} \rightarrow \min \{f(c), g(c)\}$.
(e) Prove: Suppose $E \subset \mathbb{R}$ and $f, g: E \rightarrow \mathbb{R}$ are continuous at $p \in E$. Prove: $\forall x \in E$, $\max \{f, g\}(x):=\max \{f(x), g(x)\}$ is continuous at $p$. (Stoll, p. 156, \#14(a))
ANS: Use the fact that $\max \{f(x), g(x)\}=\frac{1}{2}(f(x)+g(x)+|f(x)-g(x)|)$ and part (b).
17. (Stoll, p. 157, \#19) Let $E \subset \mathbb{R}$ and let $f: E \rightarrow \mathbb{R}$ be continuous. Let $F=\{x \in E: f(x)=0\}$. First prove that $F$ is closed in $E$. Then consider if $F$ is necessarily closed in $\mathbb{R}$.

Before proceeding, we list some useful results.

- Recall: (Stoll, Thm 2.2.14; Fitzpatrick, definition) $A$ subset $F$ of a metric space $X$ is closed if and only if $F$ contains all its limit points.
- Note we have $F \subset E \subset \mathbb{R}$, and want to know if $F$ is closed in $E$. Example: Consider then set $[13,15) \subseteq[0,15) \subseteq \mathbb{R}$. Clearly $[13,15)$ is not closed in $\mathbb{R}$ because 15 is an open end point. 15 is a limit point of $[13,15)$, but $15 \notin[13,15)$. A set is closed, by definition, if it contains all its limit points: $15 \notin[13,15)$. But $[13,15)$ is closed in $[0,15)$ because all limit points of $[13,15)$ are contained in $[0,15)$, while $15 \notin[0,15)$.
- Recall Stoll, Thm 3.1.4 (notably (c)): Let $(X, d)$ be a metric space.
(a) If a sequence $\left\{p_{n}\right\}$ in $X$ converges, then its limit is unique.
(b) Every convergent sequence in $X$ is bounded.
(c) If $E \subset X$ and $p$ is a limit point of $E$, then $\exists\left\{p_{n}\right\} \in E$ with $p_{n} \neq p$ for all $n$ such that $\lim _{n \rightarrow \infty} p_{n}=p$.

ANS: (Ralf \& Marc) (i) Let $p \in E$ be a limit point of $F$. From Theorem 3.1.4(c), $\exists\left\{p_{n}\right\} \in E$ with $p_{n} \neq p$ for all $n \in \mathbb{N}$ such that $p=\lim _{n \rightarrow \infty} p_{n}$. Note that by continuity, $\lim _{n \rightarrow \infty} f\left(p_{n}\right)=f(p)$. But as $p_{n} \in F$, we have, $\forall n \in \mathbb{N}, f\left(p_{n}\right)=0$. Thus $f(p)=0$; and by construction, $p \in F$. So, $F$ contains all its limit points and is thus closed.
(ii) $F$ is not necessarily closed in $\mathbb{R}$. To see this, assume that $E=(a, b)$ and $f(x)=0, \forall x \in E$. Note: A constant function is continuous. Clearly then $F=(a, b)$, which is open in $\mathbb{R}$.
18. (Compactness and continuity) Recall: If $K$ is a compact subset of a metric space $X$ and if $f: K \rightarrow \mathbb{R}$ is continuous on $K$, then $f(K)$ is compact. This is, e.g., Stoll, Thm 4.2.8, proven by topological compactness, i.e., showing that $f(K)$ admits a finite subcover.
Let $K$ be a compact subset of a metric space $X$ and let $f: K \rightarrow \mathbb{R}$ be continuous on $K$. Prove that $f(K)$ is compact by showing that $f(K)$ is closed and bounded.
(Stoll, p. 157, \# 25)

## Hints:

(1) Recall the Heine-Borel-Bolzano-Weierstrass Theorem, e.g., Stoll, Thm 2.4.2: Let $K$ be a subset of $\mathbb{R}$. Then the following are equivalent: (a) $K$ is closed and bounded. (b) $K$ is compact. (c) Every infinite subset of $K$ has a limit point in $K$.
(2) Also recall the definition of sequentially compact (e.g., Fitzpatrick, p. 46): A set of real numbers $S$ is said to be sequentially compact provided that every sequence $\left\{a_{n}\right\}$ in $S$ has a subsequence that converges to a point that belongs to $S$.

ANS: (Stoll, outline of answer provided in book) First show that $f(K)$ is closed as follows. Let $q$ be a limit point of $f(K)$. Then $\exists\left\{p_{n}\right\} \in K$ such that $f\left(p_{n}\right) \rightarrow q$. From the equivalence of topological and sequential compactness (e.g., Stoll, Thm 3.4.5), and the continuity of $f, q=\lim f\left(p_{n}\right)=f\left(\lim p_{n}\right)$, and as $K$ is (sequentially) compact, $p:=\lim p_{n} \in K$, so $q=f(p) \in f(K)$, showing that $f(K)$ is closed.

Now assume $f(K)$ is not bounded. Then $\exists\left\{p_{n}\right\} \in K$ such that $\left|f\left(p_{n}\right)\right| \rightarrow \infty$. Closure of $f(K)$ and continuity of $f$ (and thus continuity of $|f|$ from exercise 19 b ) imply, with $p:=\lim p_{n}$, $\infty=\lim \left|f\left(p_{n}\right)\right|=|f|\left(\lim \left(p_{n}\right)\right)=|f|(p)$, which violates that $f$ is a real-valued function. This contradiction shows that $f(K)$ is bounded.
19. (Compactness and continuity) Prove (albeit just using known results, so is easy): A continuous real-valued function on a closed and bounded interval $[a, b]$ is uniformly continuous. (Stoll, Coro 4.3.5)

ANS: This is in fact Fitzpatrick, Thm 3.17, proven via sequential compactness arguments. Stoll, Thm 4.3.4 uses topological compactness, and this exercise just makes the connection yet more precise.

The corollary follows from (i) the well-known result that: If $K$ is compact and $f: K \rightarrow \mathbb{R}$ is continuous on $K$, then $f$ is uniformly continuous on $K$ (e.g., Stoll, Thm 4.3.4); and (ii) Heine-Borel: Every closed and bounded interval $[a, b]$ is compact (e.g., Stoll, Thm 2.4.1).
20. (Stoll, p. 157, \#29) Let $K$ be a compact subset of a metric space $X$, and let $f$ be a real-valued function on $K$. Suppose that, $\forall x \in K, \exists \epsilon_{x}>0$ such that $f$ is bounded on $N_{\epsilon_{x}}(x) \cap K$. Prove that $f$ is bounded on $K$.
ANS: (Stoll, solutions) By hypothesis, $\forall x \in K, \exists \epsilon_{x}>0$ and $\exists M_{x}>0$ such that, $\forall y \in N_{\epsilon_{x}}(x) \cap K$, $|f(y)| \leq M_{x}$. The collection $\left\{N_{\epsilon_{x}}(x)\right\}_{x \in K}$ is an open cover of $K$. As $K$ is compact, $\exists N \in \mathbb{N}$ and a set $S=\left\{x_{1}, \ldots, x_{N}\right\}$ such that $\left\{N_{\epsilon_{x}}(x)\right\}_{x \in S}$ covers $K$. Let $M_{S}=\max \left\{M_{x_{1}}, \ldots, M_{x_{N}}\right\}$. Thus, $\forall y \in K$, $|f(y)| \leq M_{S}$.
21. (Junghenn, p. 262, $\# 15$; see also Stoll, p. 158, $\# 30$ ) Let $A$ be a nonempty subset of $X$ and define $d(A, \cdot): X \rightarrow \mathbb{R}$ by $d(A, x)=d(A,\{x\})$ as in (12). Prove the following:
(a) $|d(A, x)-d(A, y)| \leq d(x, y)$, hence $d(A, \cdot)$ is uniformly continuous.
(b) $d(A, x)=0$ iff $x \in \operatorname{cl}(A)$.
(c) If $A$ and $B$ are disjoint closed sets, then the function

$$
F_{A B}(x)=\frac{d(A, x)}{d(A, x)+d(B, x)}, \quad x \in X
$$

is well-defined and continuous, $0 \leq F_{A B} \leq 1$ on $X$, and

$$
A=\left\{x: F_{A B}(x)=0\right\}, \quad B=\left\{x: F_{A B}(x)=1\right\} .
$$

(d) If $A$ and $B$ are disjoint closed sets of $X$, then there exist disjoint open sets $U$ and $V$ such that $A \subseteq U$ and $B \subseteq V$. ( $U$ and $V$ are then said to separate $A$ and $B$.)

ANS: (Junghenn, solutions manual) Note he uses the notation $\operatorname{cl}(A)$ for $\bar{A}$.
(a) For $a \in A, d(A, x) \leq d(a, x) \leq d(a, y)+d(y, x)$, so $d(A, x)-d(y, x) \leq d(a, y)$. Taking the infimum over $a$ yields $d(A, x)-d(y, x) \leq d(A, y)$ or $d(A, x)-d(A, y) \leq d(y, x)$. Interchanging $x$ and $y$ yields (a).
(b) If $x \notin \operatorname{cl}(A), \exists r>0$ such that $B_{r}(x) \cap \operatorname{cl}(A)=\emptyset$. Then $d(a, x) \geq r$ for all $a \in A$ hence $d(A, x)>0$. Conversely, assume $x \in \operatorname{cl}(A)$ and let $a_{n} \in A$ with $a_{n} \rightarrow x$. Note that $0 \leq$ $d(A, x) \leq d\left(a_{n}, x\right) \rightarrow 0$, so $d(A, x)=0$.
(c) By (b), the denominator of $F_{A B}(x)$ is positive, hence $F_{A B}$ is well-defined. Continuity follows from (a), and clearly $0 \leq F_{A B} \leq 1$. The last assertions follow from (b).
NOTE: (Ralf) Regarding continuity of $F_{A B}$, we additionally need to recall that ratios of continuous functions are continuous, provided the denominator is not zero. See, e.g., Stoll, Thm 4.2.3(c).
(d) $U=\left\{x \in X: F_{A B}(x)<1 / 2\right\}, V=\left\{x \in X: F_{A B}(x)>1 / 2\right\}$.
22. (Right and left continuity; discontinuities) What is meant by $f(p+)$ and $f(p-)$, as Stoll defines them? Define left and right continuity of a function $f$ at a point $p$ in the domain of $f$. State the conditions in terms of $f(p+)$ and $f(p-)$ such that $f$ is continuous at $p$. Finally, define discontinuities of the second kind.
ANS: From Stoll, Def 4.4.1: Let $E \subset \mathbb{R}$ and let $f$ be a real-valued function defined on $E$. Suppose $p$ is a limit point of $E \cap(p, \infty)$. The function $f$ has a right limit at $p$ if $\exists L \in \mathbb{R}$ such that $\forall \epsilon>0, \exists \delta>0$ for which

$$
|f(x)-L|<\epsilon \text { for all } x \in E \text { satisfying } p<x<p+\delta
$$

The right limit of $f$, if it exists, is denoted by $f(p+)$, and we write

$$
f(p+)=\lim _{x \rightarrow p^{+}} f(x)=\lim _{\substack{x \rightarrow p \\ x>p}} f(x)
$$

Similarly, if $p$ is a limit point of $E \cap(-\infty, p)$, the left limit of $f$ at $p$, if it exists, is denoted by $f(p-)$, and we write

$$
f(p-)=\lim _{x \rightarrow p^{-}} f(x)=\lim _{\substack{x \rightarrow p \\ x<p}} f(x)
$$

The hypothesis that $p$ is a limit point of $E \cap(p, \infty)$ guarantees that, $\forall \delta>0, E \cap(p, p+\delta) \neq \emptyset$.
From Stoll, Def 4.4.2: Let $E \subset \mathbb{R}$ and let $f$ be a real-valued function on $E$. The function $f$ is right continuous (left continuous) at $p \in E$ if $\forall \epsilon>0, \exists \delta>0$ such that

$$
|f(x)-f(p)|<\epsilon \text { for all } x \in E \text { with } p \leq x<p+\delta \quad(p-\delta<x \leq p)
$$

From Stoll, p. 164, a function $f$ is continuous at $p \in(a, b)$ if and only if (a) $f(p+)$ and $f(p-)$ both exist; and (b) $f(p+)=f(p-)=f(p)$.
From Stoll, p. 165, all discontinuities for which $f(p+)$, and/or $f(p-)$ do not exist are discontinuities of the second kind. Researching this a bit further, from https://www. statisticshowto.com/ calculus-definitions/types-of-discontinuity/, we read:

An essential discontinuity (also called second type or irremovable discontinuity) is a discontinuity that jumps wildly as it gets closer to the limit. Essential discontinuities (i.e., non-removable discontinuities) can be further broken down into two types of discontinuity, based on whether the one-sided limits are bounded or unbounded (Bauldry, 2011):

- Bounded: oscillatory discontinuity. The pattern near the limit bounces up and down, never forming a pattern you can pin down.
- Unbounded: infinite discontinuity. The limits exist, but they are infinite, getting larger as you move closer to the limit.
Simple (removable) discontinuities can also be broken down into two subtypes:
- A removable discontinuity has a gap that can easily be filled in, because the limit is the same on both sides.
- A jump discontinuity at a point has limits that exist, but it's different on both sides of the gap.

See also https://en.wikipedia.org/wiki/Classification_of_discontinuities.
23. (Monotonic functions with discontinuities) First recall Stoll, Thm 4.4.7:

Let $I \subset \mathbb{R}$ be an open interval and let $f: I \rightarrow \mathbb{R}$ be monotone increasing on $I$. Then $f(p+)$ and $f(p-)$ exists for every $p \in I$ and

$$
\sup _{x<p} f(x)=f(p-) \leq f(p) \leq f(p+)=\inf _{p<x} f(x)
$$

Furthermore, if $p<q$, for $p, q \in I$, then $f(p+) \leq f(q-)$.
Next, recall Stoll, Coro 4.4.8:
If $f$ is monotone on an open interval $I$, then the set of discontinuities of $f$ is at most countable.

Prove this important result, Coro 4.4.8, adding a bit of detail (shown in my answer in italics).
ANS: Exactly as in Stoll, with my added comments:
(Step 1): Let $E=\{p \in I: f$ is discontinuous at $p\}$. Suppose $f$ is monotone increasing on $I$. Then

$$
p \in E \quad \text { if and only if } \quad f(p-)<f(p+)
$$

- Recall from Stoll, p. 164: A function $f$ is continuous at $p \in(a, b)$ if and only if (a) $f(p+)$ and $f(p-)$ both exist; and (b) $f(p+)=f(p-)=f(p)$. Here we use the contrapositive.
(Step 2): For each $p \in E$, choose $r_{p} \in \mathbb{Q}$ such that

$$
f(p-)<r_{p}<f(p+)
$$

- This choice can be made because $\mathbb{Q}$ is dense in $\mathbb{R}$ : Recall, e.g., Fitzpatrick, p. 15: Definition $A$ set $S$ of real numbers is said to be dense in $\mathbb{R}$ provided that every interval $I=(a, b)$, where $a<b$, contains a member of $S$; and his Theorem 1.9: $\mathbb{Q}$ is dense in $\mathbb{R}$.
(Step 3): If $p<q$, then $f(p+) \leq f(q-)$.
- This is stated in Stoll, Thm 4.4.7, and repeated directly above.
(Step 4): Therefore, if $p, q \in E$, we have $r_{p} \neq r_{q}$; and thus the function $p \rightarrow r_{p}$ is a one-to-one map of $E$ into $\mathbb{Q}$.
- Recall: $f: X \rightarrow Y$ is one-to-one if: $\forall x_{1}, x_{2} \in X, x_{1} \neq x_{2} \Rightarrow f\left(x_{1}\right) \neq f\left(x_{2}\right)$. (Step 5): Therefore, $E$ is equivalent to a subset of $\mathbb{Q}$ and thus is at most countable.
- Recall: Stoll, Def 1.7.1: Two sets $A$ and $B$ are said to be equivalent (or to have the same cardinality), denoted $A \sim B$, if there exists a one-to-one function of $A$ onto $B$.
- Here, function $p \rightarrow r_{p}$ is a one-to-one map of $E$ into $\mathbb{Q}$. As Stoll writes, " $E$ is equivalent to a subset of $\mathbb{Q}$ and thus is at most countable."

24. (Monotonic function with countably infinite discontinuities) Recall the following theorem, which is not found in most beginning analysis books:

Stoll, Thm 4.4.10: Let $a, b \in \mathbb{R}$ with $a<b$, and let $\left\{x_{n}\right\}_{n \in \mathbb{N}}$ be a countable subset of $(a, b)$. Let $\left\{c_{n}\right\}_{n=1}^{\infty}$ be any sequence of positive real numbers such that $\sum_{n=1}^{\infty} c_{n}$ converges. Then there exists a monotone increasing function $f$ on $[a, b]$ such that
(a) $f(a)=0$ and $f(b)=\sum_{n=1}^{\infty} c_{n}$,
(b) $f$ is continuous on $[a, b] \backslash\left\{x_{n}: n=1,2, \ldots\right\}$,
(c) $f\left(x_{n}+\right)=f\left(x_{n}\right)$ for all $n$; i.e., $f$ is right continuous at all $x_{n}$, and
(d) $f$ is discontinuous at each $x_{n}$ with $f\left(x_{n}\right)-f\left(x_{n}-\right)=c_{n}$.

In the proof, Stoll defines $f(x)=\sum_{n=1}^{\infty} c_{n} I\left(x-x_{n}\right), \forall x \in[a, b]$, this being a sum of monotone increasing functions. On page 169, Stoll asks the reader to verify that $f$ is also monotone increasing. This is easily seen: Copying from https://math.stackexchange.com/questions/1501539:

If $f_{1}, \ldots, f_{N}$ are all increasing functions, then their sum is increasing. In fact, for all $x<y$, you have $f_{i}(x) \leq f_{i}(y)$, so summing up you get $f_{1}(x)+\cdots+f_{n}(x) \leq f_{1}(y)+\cdots+f_{n}(y)$, i.e., $\sum_{i} f_{i}$ is increasing. Moreover if one of them is strictly increasing, then the sum is strictly increasing (you have $<$ instead of $\leq$ ).

Further in the proof, on the top of page 172, we read: Therefore

$$
\begin{equation*}
f\left(x_{n}\right)-f(y)=c_{n} \quad \text { for all } \quad y, x_{n}-\delta<y<x_{n} \tag{46}
\end{equation*}
$$

Confirm this.
ANS: $f(x)=\sum_{k=1}^{\infty} C_{k} I\left(x-x_{k}\right)$ and $f\left(x_{n}\right)-f(y)=\sum_{k=1}^{\infty} C_{k}\left[I\left(x_{n}-x_{k}\right)-I\left(y-x_{k}\right)\right]$.
Consider the two possible cases: Do not read yet. I need to process this further
(a) for $k=n$, and $x_{n}-\delta<y<x_{n}, c_{n}\left[I(0)-I\left(y-x_{n}\right)\right]=c_{n}[1-0]$, because $y<x_{n}$.
(b) for $k \neq n$ and $x_{n}-\delta<y<x_{n}$, the term [•] is zero, from Stoll, page 171, line -3 , namely, for $y \in\left(x_{n}-\delta, x_{n}\right), I\left(y-x_{k}\right)=I\left(x_{n}-x_{k}\right)$.

So, $f\left(x_{n}\right)-f(y)=C_{n}$, for $x_{n}-\delta<y<x_{n}$.

### 4.2 Stoll Chapter 5; Fitzpatrick Chapter 4 (Differentiation)

1. (Derivatives at endpoints of an interval)
(a) Let $I \subset \mathbb{R}$ be an interval. State the definitions of right (left) derivative.
(b) State the conditions under which $f^{\prime}(p)$ exists for (i) $p \in \operatorname{Int}(I)$; and (ii) $p$ a left or right endpoint $I$.
(c) Distinguish between $f_{+}^{\prime}(p)$ and $f^{\prime}(p+)$, and state when they exist. Similar for $f_{-}^{\prime}(p)$ and $f^{\prime}(p-)$.

ANS:
(a) Stoll, Def 5.1.2: Let $I \subset \mathbb{R}$ be an interval and let $f$ be a real-valued function with domain $I$. If $p \in I$ is such that $I \cap(p, \infty) \neq \emptyset$, then the right derivative of $f$ at $p$, denoted $f_{+}^{\prime}(p)$, is defined as

$$
\begin{equation*}
f_{+}^{\prime}(p)=\lim _{h \rightarrow 0^{+}} \frac{f(p+h)-f(p)}{h} \tag{47}
\end{equation*}
$$

provided the limit exists. Similarly, if $p \in I$ satisfies $(-\infty, p) \cap I \neq \emptyset$, then the left derivative of $f$ at $p$, denoted $f_{-}^{\prime}(p)$, is given by

$$
\begin{equation*}
f_{-}^{\prime}(p)=\lim _{h \rightarrow 0^{-}} \frac{f(p+h)-f(p)}{h} \tag{48}
\end{equation*}
$$

provided the limit exists. NOTE: if $I=[a, b]$, the right derivative applies to $p \in[a, b)$, but not for $p=b$, because $I \cap(b, \infty)=\emptyset$. Similar for left derivative.
(b) If $p \in \operatorname{Int}(I)$, then $\exists f^{\prime}(p) \Longleftrightarrow \exists f_{+}^{\prime}(p), \exists f_{-}^{\prime}(p), f_{+}^{\prime}(p)=f_{-}^{\prime}(p)$.

If $p \in I$ is the left (right) endpoint of $I$, then $f^{\prime}(p)$ exists if and only if $f_{+}^{\prime}(p)\left(f_{-}^{\prime}(p)\right)$ exists. In this case, $f^{\prime}(p)=f_{+}^{\prime}(p)\left(f_{-}^{\prime}(p)\right)$.
(c) The right derivative of $f$ at $p$ is $f_{+}^{\prime}(p)$, given in (47), and requires that $p \in I$; while $f^{\prime}(p+)$ refers to the right limit of the derivative; i.e., $f^{\prime}(p+)=\lim _{x \rightarrow p^{+}} f^{\prime}(x)$, and can exist if $p \notin I$. The distinction between $f_{-}^{\prime}(p)$ and $f^{\prime}(p-)$ is similar.
2. For differentiable functions $f(t), g(t)$, heuristically, it appears we can write

$$
\begin{equation*}
[d f / d g](t)=[(d f / d t) /(d g / d t)](t)=f^{\prime}(t) / g^{\prime}(t) \tag{49}
\end{equation*}
$$

What is this, rigorously?
Before proceeding, what in fact does $d f / d g$ mean? It is the rate of change of $f(t)$, not with respect to an infinitesimal change in $t$, but rather with respect to an infinitesimal change in $g(t)$. Take for example $f(t)=t^{2}$ and $g(t)=t^{3}$. Clearly, $f^{\prime}(t)=2 t$ and $g^{\prime}(t)=3 t^{2}$. In terms of $g, f(t)=\left(t^{3}\right)^{2 / 3}$, and

$$
\frac{d f}{d g}=\frac{d f}{d\left(t^{3}\right)}=\frac{d}{d t^{3}}\left(t^{2}\right)=\frac{d}{d t^{3}}\left(t^{3}\right)^{2 / 3}=\frac{2}{3}\left(t^{3}\right)^{-1 / 3}=\frac{2}{3} t^{-1} .
$$

And indeed, $f^{\prime}(t) / g^{\prime}(t)=(2 t) /\left(3 t^{2}\right)=(2 / 3) t^{-1}$.
ANS: This is the Cauchy Mean Value Theorem (e.g., Stoll, Thm 5.2.8): If $f, g$ are continuous realvalued functions on $[a, b]$ that are differentiable on $(a, b)$, then

$$
\begin{equation*}
\exists c \in(a, b) \text { such that }[f(b)-f(a)] g^{\prime}(c)=[g(b)-g(a)] f^{\prime}(c) \tag{50}
\end{equation*}
$$

Indeed, as shown subsequently,

$$
\begin{equation*}
\text { if } g^{\prime}(x) \neq 0 \text { for all } x \in(a, b) \text {, then } g(a) \neq g(b) \tag{51}
\end{equation*}
$$

and (50) can be written as

$$
\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f^{\prime}(c)}{g^{\prime}(c)}
$$

which, when multiplied by $1=(b-a) /(b-a)$ gives (49). To prove (51), observe that this is the contrapositive of Rolle's Theorem:
(Stoll, Thm 5.2.5) Suppose $f$ is a continuous real-valued function on $[a, b]$ with $f(a)=f(b)$, and that $f$ is differentiable on $(a, b)$. Then there exists $c \in(a, b)$ such that $f^{\prime}(c)=0$.

From the contrapositive, if $\nexists c$ such that $f^{\prime}(c)=0$, then $f(a) \neq f(b)$.
3. (Ghorpade and Limaye, Coro 4.23; Stoll, Thm 5.2.9(e)) Let $I$ be an interval containing more than one point, and let $f: I \rightarrow \mathbb{R}$ be any function. Prove: $f$ is a constant function on $I$ if and only if $f^{\prime}$ exists and is identically zero on $I$.
ANS: If $f$ is a constant function on $I$, then it is obvious that $f^{\prime}$ exists on $I$ and $f^{\prime}(x)=0$ for all $x \in I$. Conversely, suppose $f^{\prime}$ exists and vanishes identically on $I$. Let $x_{1}, x_{2} \in I$ with $x_{1}<x_{2}$. Then $\left[x_{1}, x_{2}\right] \subseteq I$, and by applying the MVT to the restriction of $f$ to $\left[x_{1}, x_{2}\right]$, we obtain

$$
f\left(x_{2}\right)-f\left(x_{1}\right)=f^{\prime}(c)\left(x_{2}-x_{1}\right) \text { for some } c \in\left(x_{1}, x_{2}\right) .
$$

Since $f^{\prime}(c)=0$, we obtain $f\left(x_{1}\right)=f\left(x_{2}\right)$. Since $x_{1}, x_{2}$ are arbitrary elements of $I$ with $x_{1}<x_{2}$, it follows that $f$ is a constant function on $I$.
4. (Stoll, Remark, p. 199) Every book (e.g., Stoll, Thm 5.2.9) proves the usual calculus results about derivatives, e.g.: Suppose $f: I \rightarrow \mathbb{R}$ is differentiable on the interval $I$. If $f^{\prime}(x) \geq 0$ for all $x \in I$, then $f$ is monotone increasing on $I$; etc.. Paraphrasing, Stoll's important remark regards the case in which $f^{\prime}$ is not continuous, and states:

It needs to be emphasized that if the derivative of a function $f$ is positive at a point $c$, then this does not imply that $f$ is increasing on an interval containing $c$; it could be non-monotone on any interval containing $c$. If $f^{\prime}(c)>0$, the only conclusion that can be reached is that

$$
\begin{equation*}
\exists \delta>0 \text { such that } \forall x \in(c-\delta, c), f(x)<f(c) \text { and } \forall x \in(c, c+\delta), f(x)>f(c) . \tag{52}
\end{equation*}
$$

This does not mean that $f$ is increasing on $(c-\delta, c+\delta)$.
However, [recalling the result in exercise 10b,] if $f^{\prime}(c)>0$ and $f^{\prime}$ is continuous at $c$, then $\exists \delta>0$ such that, $\forall x \in(c-\delta, c+\delta), f^{\prime}(x)>0$. Thus $f$ is increasing on $(c-\delta, c+\delta)$.

While perhaps a bit tricky to visualize because $f^{\prime}$ is not continuous, imagine, for example, a differentiable (and thus continuous) function that is (pathologically) oscillatory for $x \in(c-\delta, c+\delta)$ but such that (52) is satisfied, i.e., all its values for $x \in(c-\delta, c)$ lie below $f(c)$, and all its values are $x \in(c, c+\delta)$ lie above $f(c)$.
(a) Prove the standard result statement mentioned above: Suppose $f: I \rightarrow \mathbb{R}$ is differentiable on the interval $I$. If $f^{\prime}(x) \geq 0$ for all $x \in I$, then $f$ is monotone increasing on $I$.
ANS: Suppose $x_{1}, x_{2} \in I$ with $x_{1}<x_{2}$. By the MVT applied to $f$ on $\left[x_{1}, x_{2}\right], f\left(x_{2}\right)-f\left(x_{1}\right)=$ $f^{\prime}(c)\left(x_{2}-x_{1}\right)$ for some $c \in\left(x_{1}, x_{2}\right)$. If $f^{\prime}(c) \geq 0$, then $f\left(x_{2}\right) \geq f\left(x_{1}\right)$. Thus, if $f^{\prime}(x) \geq 0$ for all $x \in I$, we have $f\left(x_{2}\right) \geq f\left(x_{1}\right)$ for all $x_{1}, x_{2} \in I$ with $x_{1}<x_{2}$. Thus $f$ is monotone increasing on $I$.
(b) Prove (52).

ANS:
Proof I (Marc): Let $R=R_{x, c}=(f(x)-f(c)) /(x-c)$, and take $\epsilon>0$. If $f^{\prime}(c)$ exists, then $\exists \delta_{\epsilon}>0$ such that, if $0<|x-c|<\delta_{\epsilon}$, then $\left|R-f^{\prime}(c)\right|<\epsilon$. If $f^{\prime}(c)>0$, we can choose $\epsilon$ small enough, say $\epsilon_{0}$, so that $0<f^{\prime}(c)-\epsilon_{0}<R<f^{\prime}(c)+\epsilon_{0}$, i.e., $R>0$. Choose $\delta=\delta_{\epsilon_{0}}$. For $c-\delta<x<c$, the denominator of $R$ is negative, so the numerator must be negative, i.e., for this choice of $\delta, c-\delta<x<c, f(x)<f(c)$. Likewise, for $c<x<c+\delta$, the denominator of $R$ is positive, so the numerator must be positive, i.e., $f(x)>f(c)$. Thus, $\exists \delta>0$ such that (52) holds.

Proof II (Ralf): By hypothesis, $f^{\prime}(c)$ exists and $f^{\prime}(c)>0$. Let $\epsilon=f^{\prime}(c)>0$. Then by existence of the derivative, $\exists \delta>0$ such that

$$
\begin{equation*}
\left|\frac{f(x)-f(c)}{x-c}-f^{\prime}(c)\right|<\epsilon, \quad \forall x \in I \text { with } 0<d(x, c)<\delta . \tag{53}
\end{equation*}
$$

This implies that

$$
\begin{equation*}
\underbrace{f^{\prime}(c)-\epsilon}_{=0}<\frac{f(x)-f(c)}{x-c}, \forall x \in I \text { with } 0<d(x, c)<\delta . \tag{54}
\end{equation*}
$$

Note that, whenever $x>c$ (with $0<x-c<\delta$ ), the denominator is positive, so that $f(x)>f(c)$. Similarly, whenever $c>x$ (with $0<c-x<\delta$ ), we must have $f(c)>f(x)$. We have therefore proven the existence of a $\delta$ as required.
5. (Pons, Thm 5.4.5) Let $f$ and $g$ be functions, differentiable on $(0, \infty)$ and continuous on $[0, \infty)$. Prove: If $f^{\prime}(x) \leq g^{\prime}(x)$ for every $x \in(0, \infty)$ and $f(0)=g(0)$, then $f(x) \leq g(x)$ for every $x \in[0, \infty)$.
ANS: Let $h=g-f$, so $h^{\prime}(x)=g^{\prime}(x)-f^{\prime}(x) \geq 0$ for all $x \in(0, \infty)$. Fix $x_{0} \in(0, \infty)$ and apply the MVT to $h \in\left(0, x_{0}\right)$, showing $\exists c \in\left(0, x_{0}\right)$ such that

$$
h^{\prime}(c)=\frac{h\left(x_{0}\right)-h(0)}{x_{0}-0}
$$

The quotient and the denominator are nonnegative; thus it must be the case that $h\left(x_{0}\right)-h(0) \geq 0$. Substituting for $f$ and $g$,

$$
0 \leq h\left(x_{0}\right)-h(0)=g\left(x_{0}\right)-f\left(x_{0}\right)-(g(0)-f(0))=g\left(x_{0}\right)-f\left(x_{0}\right),
$$

implying $f\left(x_{0}\right) \leq g\left(x_{0}\right)$.
6. (Stoll, p. 203, \#5(a); Sasane, Example 4.20) Prove: For all $x>0, \sqrt{1+x}<1+\frac{1}{2} x$.

ANS:
Proof I (Sasane): Consider $f:[0, \infty) \rightarrow \mathbb{R}$ defined by $f(x)=\sqrt{1+x}$ for $x \geq 0$. Then $f$ is continuous on $[0, \infty)$, differentiable on $[0, \infty)$ and

$$
f^{\prime}(x)=\frac{1}{2 \sqrt{1+x}}, \quad x \geq 0
$$

Fix $x>0$. Applying the MVT to $f$ on [ $0, x$ ], it follows that, for some $c$ satisfying $0<c<x$,

$$
\frac{\sqrt{1+x}-1}{x}=\frac{f(x)-f(0)}{x-0}=f^{\prime}(c)=\frac{1}{2 \sqrt{1+c}}<\frac{1}{2 \sqrt{1+0}}=\frac{1}{2}
$$

and so $\sqrt{1+x}<1+\frac{1}{2} x$. Stoll adds that the result also holds for $x>-1$, which is easy to see.
Proof II (Ralf) Let $f(x)=\sqrt{1+x}$ and $g(x)=1+\frac{1}{2} x$. Define $h(x)=g(x)-f(x)$. Clearly, $h(0)=0$ and $h^{\prime}(x)=\frac{1}{2}-\frac{1}{2 \sqrt{1+x}}>0$, for all $x>0$. As $h$ is differentiable (and thus also continuous) on $[0, \infty$ ), we can apply the MVT: Fix $x>0$. Applying the MVT to $f$ on $[0, x]$, it follows that, for some $c$ satisfying $0<c<x$,

$$
h^{\prime}(c)=\frac{h(x)-h(0)}{x} .
$$

As $h^{\prime}(c)>0$ and $x>0$, it must be that $h(x)>h(0)$, which implies $g(x)>f(x)$, so that

$$
\sqrt{1+x}<1+\frac{1}{2} x, \quad \forall x \in(0, \infty)
$$

as required.
7. Some books (e.g., Sasane, p. 163), show a proof of the Cauchy Mean Value Theorem via matrix determinants. In particular, let $f, g$ be continuous on $[a, b]$ and differentiable on $(a, b)$, and define, for $x \in[a, b]$,

$$
\varphi(x)=\operatorname{det}\left[\begin{array}{ccc}
f(x) & g(x) & 1 \\
f(a) & g(a) & 1 \\
f(b) & g(b) & 1
\end{array}\right]=f(x)(g(a)-g(b))-g(x)(f(a)-f(b))+f(a) g(b)-f(b) g(a)
$$

Use this to prove the result.
ANS: $\varphi$ is continuous on $[a, b]$, differentiable on $(a, b)$ and, for $x \in(a, b)$,

$$
\varphi^{\prime}(x)=f^{\prime}(x)(g(a)-g(b))-g^{\prime}(x)(f(a)-f(b))
$$

As $\varphi(a)=\varphi(b)=0$, Rolle's Theorem implies $\exists c \in(a, b)$ such that $\varphi^{\prime}(c)=0$, that is,

$$
f^{\prime}(c)[g(a)-g(b)]-g^{\prime}(c)[f(a)-f(b)]=0
$$

which, multiplying by -1 , is the Cauchy MVT.
8. (Jacob and Evans, Coro 8.12) Suppose that $f:[a, b] \rightarrow \mathbb{R}$ fulfils the assumptions of the MVT. Further suppose that $m \leq f^{\prime}(\eta) \leq M$ for all $\eta \in(a, b)$.
Prove: $\forall x, y \in(a, b), y \leq x, m(x-y) \leq f(x)-f(y) \leq M(x-y)$.
ANS: Apply the MVT to the restricted function $\left.f\right|_{[y, x]}$, so $\exists \xi \in(y, x)$ such that $f(x)-f(y)=$ $f^{\prime}(\xi)(x-y)$. As $f^{\prime}(\xi) \geq m$ and $x-y \geq 0$, this implies $m(x-y) \leq f(x)-f(y)$. Further, as $f^{\prime}(\xi) \leq M$ and $x-y \geq 0, f(x)-f(y) \leq M(x-y)$.
9. Can a function be strictly increasing but such that its derivative is not always positive?

ANS: Take $f(x)=x^{3}$, for which $f^{\prime}(0)=0$.
10. (Stoll, Example 5.1.3(g) and p. 191, \#9) Let $g$ be defined by

$$
g(x)=\left\{\begin{array}{cl}
x^{2} \sin \frac{1}{x}, & x \neq 0 \\
0, & x=0
\end{array}\right.
$$

The function is shown in Figure 7.
(a) Prove that $g$ is differentiable at 0 and that $g^{\prime}(0)=0$.


Figure 7: From Ghorpade and Limaye, page 264.
(b) Show that $g^{\prime}(x)$ is not continuous at 0 .

ANS: (Ghorpade and Limaye, Example 7.21) $g$ is differentiable at 0:

$$
g^{\prime}(0)=\lim _{x \rightarrow 0} \frac{g(x)-g(0)}{x-0}=\lim _{x \rightarrow 0} x \sin \frac{1}{x}=0
$$

But $g^{\prime}$ is not continuous at 0 , because $\lim _{x \rightarrow 0} g^{\prime}(x)$ does not exist. This follows because $\lim _{x \rightarrow 0} \cos (1 / x)$ does not exist.
11. (Darboux' Intermediate Value Theorem for Derivatives) Here we take a deep dive into the subtle aspects of this intriguing result. We first state some relevant results that we will need; then the theorem and its short proof; some comments; and then we come the exercises.

- (Stoll, Def 5.2.1) Suppose $E \subset \mathbb{R}$ and $f$ is a real-valued function with domain $E$. The function $f$ has a local maximum at a point $p \in E$ if $\exists \delta>0$ such that, $\forall x \in E \cap N_{\delta}(p), f(x) \leq f(p)$. The function $f$ has an absolute maximum at $p \in E$ if $f(x) \leq f(p)$ for all $x \in E$.
Similarly, $f$ has a local minimum at a point $q \in E$ if $\exists \delta>0$ such that, $\forall x \in E \cap N_{\delta}(q)$, $f(x) \geq f(q)$; and $f$ has an absolute minimum at $q \in E$ if $f(x) \geq f(q)$ for all $x \in E$.
- (Stoll, Thm 5.2.2): Let $f$ be a real-valued function defined on an interval $I$, and suppose $f$ has either a local minimum or local maximum at $p \in \operatorname{Int}(I)$. If $f$ is differentiable at $p$, then $f^{\prime}(p)=0$.
- (Stoll, Remark following Thm 5.2.9): This is exercise 4; go quickly review it.
(Stoll, Thm 5.2.13, Darboux' IVT for Derivatives): Suppose $I \subset \mathbb{R}$ is an interval and $f: I \rightarrow \mathbb{R}$ is differentiable on $I$. Then given $a, b \in I$ with $a<b$ and a real number $\lambda$ between $f^{\prime}(a)$ and $f^{\prime}(b)$, $\exists c \in(a, b)$ such that $f^{\prime}(c)=\lambda$.
Proof: Define $g$ by $g(x)=f(x)-\lambda x$. Then $g$ is differentiable on $I$ with $g^{\prime}(x)=f^{\prime}(x)-\lambda$. Suppose $f^{\prime}(a)<\lambda<f^{\prime}(b)$. Then $g^{\prime}(a)<0$ and $g^{\prime}(b)>0$. As in the remark following Theorem 5.2.9, since $g^{\prime}(a)<0$ there exists an $x_{1}>a$ such that $g\left(x_{1}\right)<g(a)$. Also, since $g^{\prime}(b)>0$, there exists an $x_{2}<b$ such that $g\left(x_{2}\right)<g(b)$. As a consequence, $g$ has an absolute minimum [Stoll, Def 5.2.1] at some point $c \in(a, b)$. But then $g^{\prime}(c)=f^{\prime}(c)-\lambda=0$ [note this crucially uses Stoll, Thm 5.2.2], i.e., $f^{\prime}(c)=\lambda$.
- Stoll, p. 201, writes: The remarkable aspect of this theorem is that the hypothesis does not require continuity of the derivative. If the derivative were continuous, then the result would follow from the usual Intermediate Value Theorem (Stoll, Thm 4.2.11) applied to $f^{\prime}$.
- As emphasized, the whole point of the theorem is that continuity is not required. One might then envision the graph for $f^{\prime}$ having a jump discontinuity (see, e.g., Stoll, Figure 4.9, page 165), but then the intermediate value property (IVP) cannot apply. Indeed, TBB, ${ }^{13}$ page 445 say "The result does imply, however, that $f^{\prime}$ cannot have jump discontinuities and cannot have removable discontinuities." Corropborating this, Wikipedia states that the only possible discontinuity in the derivative is an "essential discontinuity": At least one of the left and right side limits at the point do not exist; i.e., a derivative cannot have a jump discontinuity or a removable discontinuity.
- In the proof, it appears $f^{\prime}(a)$ cannot equal $f^{\prime}(b)$, i.e., strict $f^{\prime}(a)<\lambda<f^{\prime}(b)$. Indeed, Terell, p. 133 , imposes this in the statement of the theorem.
- Stoll, p. 201, writes:

The theorem is often used in calculus to determine where a function is increasing or decreasing. Suppose it has been determined that the derivative $f^{\prime}$ is zero at $c_{1}$ and $c_{2}$ with $c_{1}<c_{2}$, and that $f^{\prime}(x) \neq 0$ for all $x \in\left(c_{1}, c_{2}\right)$. Then by the theorem, it suffices to check the sign of the derivative at a single point in the interval $\left(c_{1}, c_{2}\right)$ to determine whether $f^{\prime}$ is positive or negative on the whole interval $\left(c_{1}, c_{2}\right)$. Theorem 5.2.9 [that is, the collection of usual calculus results, starting with: If $f^{\prime}(x) \geq 0$ for all $x \in I$, then $f$ is monotone increasing on $I$ ] then allows us to determine whether $f$ is increasing or decreasing on $\left(c_{1}, c_{2}\right)$.

Regarding the above "Then by the theorem", one might think to take end points $c_{1}$ and $c_{2}$, so that $0=f^{\prime}\left(c_{1}\right)<\lambda<f^{\prime}\left(c_{2}\right)=0$, which makes no sense. Instead, for any two values say $a, b \in\left(c_{1}, c_{2}\right)$, and knowing there is no $c \in\left(c_{1}, c_{2}\right)$, and thus no $c \in(a, b)$ such that $f^{\prime}(c)=0=\lambda$, the theorem implies, via the IVP, that $f^{\prime}$ cannot have both a positive and negative value in $(a, b)$; for otherwise, $\exists x_{0}$ such that $f^{\prime}\left(x_{0}\right)=0$, but this possibility was excluded.

- Darboux' IVT for Derivatives plays a crucial role in Newton's method, detailed on Stoll, p. 214, where he writes: "If in addition $f$ is differentiable on $(a, b)$ with $f^{\prime}(x) \neq 0$ for all $x \in(a, b)$, then $f$ is either strictly increasing or decreasing on $[a, b]$ ". This is precisely Darboux. However, one might argue that, in most situations, $f^{\prime}(x)$ is continuous, and the usual IVT, applied to $f^{\prime}(x)$ applies.

We now turn to the exercises.
(a) Pons, p. 198, gives this example (having changed $f$ to $g$, to avoid possible confusion):

The beauty of Darboux's Theorem has been discussed, but we now consider how subtle the result truly is. Take the function $g$ to be defined by

$$
g(x)= \begin{cases}1, & x \geq 0 \\ -1, & x<0\end{cases}
$$

Is there a function $f$ that has this function $g$ as its derivative? The answer is no, because $g$ does not attain the values between 1 and -1 . However, at first glance you should be reminded of the derivative of the absolute value function. The problem is the pesky fact that $g$ has 0 in its domain whereas the absolute value function is not differentiable at 0 .

[^11]The assignment is to precisely detail the explanation for what Pons states.
ANS: Recall Darboux' IVT for Derivatives: Let the domain of $f$ be an interval $I \subset \mathbb{R}$. Let $a, b \in I$, with $a<b$. The theorem says:
$f: I \rightarrow \mathbb{R}$ is differentiable on $I \Longrightarrow$
For $\lambda$ between $f^{\prime}(a)$ and $f^{\prime}(b), \exists c \in(a, b)$ such that $f^{\prime}(c)=\lambda$.

- (Marc) Use the contrapositive of the theorem: The contrapositive is not satisfied for $a<0$ and $b>0$, so $f$ is not differentiable on $I$.
- (Ralf) Let $a<0$ and $b \geq 0$. Assume for contradiction that $g$ were the derivative of some function. Since $g(a)=-1<g(b)=1$, then by Darboux' theorem, for any $\lambda \in(-1,1)$, there must exist $c \in(a, b)$ such that $g(c)=\lambda$. Since this is clearly not the case as $g(x)$ only takes on two discrete values -1 and $1, g$ cannot be the the derivative of any function.
(b) (Stoll, p. 191, \#10) Let $f$ be defined by

$$
f(x)= \begin{cases}x^{2}+2, & x \leq 2 \\ m x+c, & x>2\end{cases}
$$

i. For what values of $m$ and $c$ is $f$ continuous at 2 ?
ii. For what values of $m$ and $c$ is $f$ differentiable at 2 ?
iii. Building further on this question from Stoll, take $m=5, c=-4$. Apply the Darboux' IVT for Derivatives theorem to determine an interval $I$ such that $f$ is not differentiable on $I$.

ANS: For (i), we require $6=2^{2}+2=2 m+c \rightarrow\{m, c\}: 2 m+c=6$.
For (ii), taking first derivatives, we require $\left.(2 x=m)\right|_{x=2} \rightarrow m=4 \rightarrow c=-2$.
For (iii), $\forall x \leq 2, f^{\prime}(x) \leq 4$; while $\forall x>2, f^{\prime}(x)=5$. From the contrapositive of Darboux, $f^{\prime}(x)$ is not differentiable on any interval that includes the point 2.
(c) State and prove the theorem that relates $f^{\prime}\left(c^{+}\right):=\lim _{x \rightarrow c^{+}} f^{\prime}(x)$ and $f_{+}^{\prime}(c)$; and $f^{\prime}\left(c^{-}\right):=$ $\lim _{x \rightarrow c^{-}} f^{\prime}(x)$ and $f_{-}^{\prime}(c)$.
ANS: First recall the definitions of $f_{+}^{\prime}(p)$ and $f_{-}^{\prime}(p)$, as given in (47) and (48), respectively. Stoll, Thm 5.2.11, states:

Suppose $f:[a, b) \rightarrow \mathbb{R}$ is continuous on $[a, b)$ and differentiable on $(a, b)$. If $f^{\prime}\left(a^{+}\right)=$ $\lim _{x \rightarrow a^{+}} f^{\prime}(x)$ exists, then $f_{+}^{\prime}(a)$ exists and $f_{+}^{\prime}(a)=f^{\prime}\left(a^{+}\right)$.
Trivially, we have the counterpart:
Suppose $f:(a, b] \rightarrow \mathbb{R}$ is continuous on $(a, b]$ and differentiable on $(a, b)$. If $f^{\prime}\left(b^{-}\right)=$ $\lim _{x \rightarrow b^{-}} f^{\prime}(x)$ exists, then $f_{-}^{\prime}(b)$ exists and $f_{-}^{\prime}(b)=f^{\prime}\left(b^{-}\right)$.
Stoll gives the proof of the first statement as follows.
Proof: Let $L=\lim _{x \rightarrow a^{+}} f^{\prime}(x)$, that is assumed to exist. Given $\epsilon>0, \exists \delta>0$ such that, $\forall x: a<x<a+\delta,\left|f^{\prime}(x)-L\right|<\epsilon$.
Suppose $0<h<\delta$ is such that $a+h<b$. Since $f$ is continuous on $[a, a+h]$ and differentiable on $(a, a+h)$, by the MVT, $f(a+h)-f(a)=f^{\prime}\left(\zeta_{h}\right) h$ for some $\zeta_{h} \in(a, a+h)$. Therefore, $\forall h, 0<h<\delta$,

$$
\left|\frac{f(a+h)-f(a)}{h}-L\right|=\left|f^{\prime}\left(\zeta_{h}\right)-L\right|<\epsilon .
$$

Thus $f_{+}^{\prime}(a)=L$. The proof that $f_{-}^{\prime}(b)=f^{\prime}\left(b^{-}\right)$is analogous, and stated for completeness: Let $L=\lim _{x \rightarrow b^{-}} f^{\prime}(x)$. Given $\epsilon>0, \exists \delta>0$ such that, $\forall x: b-\delta<x<b,\left|f^{\prime}(x)-L\right|<\epsilon$.

Suppose $0<h<\delta$ is such that $b-h>a$. Since $f$ is continuous on $[b-h, b]$ and differentiable on $(b-h, b)$, by the MVT, $f(b)-f(b-h)=f^{\prime}\left(\zeta_{h}\right) h$ for some $\zeta_{h} \in(b-h, b)$. Therefore, $\forall h, 0<h<\delta$,

$$
\left|\frac{f(b)-f(b-h)}{h}-L\right|=\left|f^{\prime}\left(\zeta_{h}\right)-L\right|<\epsilon .
$$

Thus $f_{-}^{\prime}(b)=L$.
(d) Prove: In order for the Darboux' Intermediate Value Theorem for Derivatives to be applicable, the derivative cannot have jump/removable discontinuities. In fact, this applies in general to the existence of the derivative for function $f$ in some interval $I$.
ANS: (Ralf) Inspired by https://joelshapiro.org/Pubvit/Downloads/vandenberge_derivatives.pdf
Let $f:(a, b) \rightarrow \mathbb{R}$ be a function continuous at $c \in(a, b)$ and differentiable everywhere except perhaps at $c$. Then $f:(a, c] \rightarrow \mathbb{R}$ is continuous on $(a, c]$ and differentiable on $(a, c)$. Similarly, $f:[c, b) \rightarrow \mathbb{R}$ is continuous on $[c, b)$ and differentiable on $(c, b)$. We can therefore apply Stoll, Thm 5.2.11, to each case:

$$
\begin{aligned}
& \text { If } \lim _{x \rightarrow c^{+}} f^{\prime}(x) \text { exists, then } \lim _{x \rightarrow c^{+}} f^{\prime}(x)=f_{+}^{\prime}(c) \text {, and } \\
& \text { if } \lim _{x \rightarrow c^{-}} f^{\prime}(x) \text { exists, then } \lim _{x \rightarrow c^{-}} f^{\prime}(x)=f_{-}^{\prime}(c) \text {. }
\end{aligned}
$$

Note that if both $\lim _{x \rightarrow c^{+}} f^{\prime}(x)$ and $\lim _{x \rightarrow c^{-}} f^{\prime}(x)$ exist, but $f_{-}^{\prime}(c) \neq f_{+}^{\prime}(c)$, then $f^{\prime}(c)$ does not exist, so the function is not differentiable on the whole interval $(a, b)$. This means, first, that the required condition for Darboux is not fulfilled, and thus, for Darboux to apply, it rules out that a derivative can contain a jump discontinuity. However, notice it applies more generally: A derivative for $f$ cannot exist on an interval $I$ if $f^{\prime}$ has a jump discontinuity at one or more points in $I$.
Therefore, if $f$ is differentiable on $(a, b)$, it must be the case that: If both $\lim _{x \rightarrow c^{+}} f^{\prime}(x)$ and $\lim _{x \rightarrow c^{-}} f^{\prime}(x)$ exist, then $\lim _{x \rightarrow c^{+}} f^{\prime}(x)=f_{-}^{\prime}(c)=\lim _{x \rightarrow c^{+}} f^{\prime}(x)=f_{+}^{\prime}(c)$. This rules out the case of a removable discontinuity. (Recall: A function $f: I \rightarrow \mathbb{R}$ has a *removable* discontinuity at $p$ if $\lim _{x \rightarrow p} f(x)$ exists but does not equal $f(p)$ or $f(p)$ is not defined.)
The only possible discontinuity that can exist is if at least one of $\lim _{x \rightarrow c^{+}} f^{\prime}(x)$ and $\lim _{x \rightarrow c^{-}} f^{\prime}(x)$ fails to exist.
12. (Stoll, Coro 4.2.14) As a corollary of the Intermediate Value Theorem (IVT), Stoll gives: If $f$ : $[0,1] \rightarrow[0,1]$ is continuous, then there exists $y \in[0,1]$ such that $f(y)=y$. The short proof is as follows. Let $g(x)=f(x)-x$. Then $g(0)=f(0) \geq 0$ and $g(1)=f(1)-1 \leq 0$. Thus, $\exists y \in[0,1]$ such that $g(y)=0$; i.e., $f(y)=y$.
NOTE 1: We have seen this already in exercise 23 . Value $y$ is called a fixed point of $f$. Indeed, this is stated in Stoll, p. 162, \#13; and detailed extensively in, e.g., Ghorpade and Limaye, Sec. 5.4, in which the Picard method and Newton-Raphson are presented.
NOTE 2: The IVT requires $f(a)<f(b)$ or $f(a)>f(b)$. In the case here with function $g$, consider when this is not the case, i.e., imagine $g(0)=g(1)$, which means $f(1)-f(0)=1$, i.e., given the domain and range, $f(1)=1, f(0)=0$, which is the extreme case of the above constraints $f(0) \geq 0$ and $f(1) \leq 1$. Imagine the graph of $f(x)=x^{2}$ or $\sqrt{x}$ : then $f(y)=y$ for both endpoints $y=a$ and $y=b$, and nowhere in between. If $g(0)>0$ and $g(1)<0$, then $f(y)=y$ is satisfied somewhere in $(0,1)$, for possibly more than one value.
(a) Devise an example such that $f(1)=0, f(0)=1$, and for which there is one, and only one fixed point.

ANS: In this case, $g(0)=f(0)=1>0$ and $g(1)=f(1)-1=-1<1$, so the IVT requirement is met, and there must exist at least one solution to $f(y)=y \in(0,1)$. For example, $f(x)=1-x^{2}$. Then, imagining the graph, clearly there is a value satisfying the IVT and $f(y)=y$. For $f(x)=1-x^{2}, y^{2}+y-1=0$ has solution (in $\left.[0,1]\right) \sqrt{5} / 2-1 / 2 \approx 0.618$.
(b) Devise an example for which there are infinite solutions.

ANS: (Ghorpade and Limaye, Examples 5.16(i)) For example, $f:[0,1] \rightarrow[0,1]$ defined by $f(x)=x$, where every point of $[0,1]$ is a fixed point of $f$.
(c) Devise an example to show that closure of $I=[0,1]$ is necessary.

ANS: (Ghorpade and Limaye, Examples $5.16(i i)$ ) For example, if $f:[0,1) \rightarrow \mathbb{R}$ is defined by $f(x):=(1+x) / 2$, then $f$ maps $[0,1)$ into itself, and $f$ is continuous. But $f$ has no fixed point in $[0,1)$. Indeed, $(1+x) / 2=x$ only when $x=1$.
(d) Devise an example to show that it is necessary for $f$ to be defined on a bounded subset of $\mathbb{R}$.

ANS: (Ghorpade and Limaye, Examples $5.16($ iii) ) For example, if $f:[1, \infty) \rightarrow \mathbb{R}$ is defined by $f(x):=x+(1 / x)$, then $f$ maps $[1, \infty)$ into itself, and $f$ is continuous. But clearly, $f$ has no fixed point in $[1, \infty)$.
(e) Devise an example to show that the condition that $f$ is defined on an interval in $\mathbb{R}$ is necessary. ANS: (Ghorpade and Limaye, Examples 5.16(iv)) For example, if $D=[-2,-1] \cup[1,2]$ and $f$ : $D \rightarrow \mathbb{R}$ is defined by $f(x):=-x$, then $f$ maps $D$ into itself, and $f$ is continuous. But $f$ has no fixed point in $D$.
13. (Stoll, p. 162, \#8; Petrovic, p. 95, \#3.8.17) Suppose $E$ is a subset of a metric space $X$ and $f: E \rightarrow \mathbb{R}$ is uniformly continuous. If $\left\{x_{n}\right\}$ is a Cauchy sequence in $E$, prove that $\left\{f\left(x_{n}\right)\right\}$ is a Cauchy sequence.
ANS: Let $\epsilon>0$. As $f$ is uniformly continuous on $E, \exists \delta_{0}$ such that, if $x, y \in E$ and $|x-y|<\delta$, then $|f(x)-f(y)|<\epsilon$. As $\left\{x_{n}\right\}$ is a Cauchy sequence in $E, \exists N \in \mathbb{N}$ such that, if $m \geq n \geq N$, then $\left|x_{m}-x_{n}\right|<\delta$. It follows that if $m \geq n \geq N$, then $\left|f\left(x_{m}\right)-f\left(x_{n}\right)\right|<\epsilon$.
14. (Stoll, p. 176, \#15) Let $I \subset \mathbb{R}$ be an interval and let $f$ be a one-to-one continuous real-valued function on $I$. Prove that $f$ is strictly monotone on $I$.
ANS: We show two proofs; mine and from Loya.

- Let $f$ be continuous and one-to-one on an interval $I=[a, b]$. Then it is strictly monotone on $I$.

Proof: $f$ one-to-one $\Longrightarrow f(a) \neq f(b)$; we assume $f(a)<f(b)$, and show $f$ is strictly increasing. The case $f(a)>f(b)$ and $f$ is strictly decreasing is similar. We need to show:

$$
f(a)<f(b) \Longrightarrow \forall x \in(a, b), \quad f(a)<f(x)<f(b)
$$

- Assume $\exists x \in(a, b)$ such that $f(x) \leq f(a)<f(b)$. Then by the IVT (which uses the continuity assumption), $\exists y \in[x, b]$ with $f(y)=f(a)$, implying that $f$ is not one-to-one.
- Assume $\exists x \in(a, b)$ such that $f(a)<f(b) \leq f(x)$. Then by the IVT, $\exists y \in[a, x]$ with $f(y)=$ $f(b)$, again implying that $f$ is not one-to-one.
- (Loya, Thm 4.28) A one-to-one continuous function $f: \mathbb{R} \longrightarrow \mathbb{R}$ is strictly monotone, its range is an interval, and it has a continuous strictly monotone inverse (with the same monotonicity as $f$ ).
NOTE: The second part of Loya's proof refers to his Thm 4.27. This is stated and proved in exercise 33c.
Proof: Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a one-to-one continuous function. We shall prove that $f$ is strictly monotone. Fix points $x_{0}<y_{0}$. Then $f\left(x_{0}\right) \neq f\left(y_{0}\right)$, so either $f\left(x_{0}\right)<f\left(y_{0}\right)$ or $f\left(x_{0}\right)>f\left(y_{0}\right)$.

For concreteness, assume that $f\left(x_{0}\right)<f\left(y_{0}\right)$; the other case $f\left(x_{0}\right)>f\left(y_{0}\right)$ can be dealt with analogously. We claim that $f$ is strictly increasing. Indeed, if this were not the case, then there would exist points $x_{1}<y_{1}$ such that $f\left(y_{1}\right)<f\left(x_{1}\right)$. Now consider the function $g:[0,1] \rightarrow \mathbb{R}$ defined by

$$
g(t)=f\left(t y_{0}+(1-t) y_{1}\right)-f\left(t x_{0}+(1-t) x_{1}\right)
$$

Since $f$ is continuous, $g$ is continuous, and

$$
g(0)=f\left(y_{1}\right)-f\left(x_{1}\right)<0 \quad \text { and } \quad g(1)=f\left(y_{0}\right)-f\left(x_{0}\right)>0 .
$$

Hence by the IVT, there is a $c \in[0,1]$ such that $g(c)=0$. This implies that $f(a)=f(b)$, where $a=c x_{0}+(1-c) x_{1}$ and $b=c y_{0}+(1-c) y_{1}$. Since $f$ is one-to-one, we must have $a=b$; however, this is impossible, since $x_{0}<y_{0}$ and $x_{1}<y_{1}$ implies $a<b$. This contradiction shows that $f$ must be strictly monotone.
Now let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a continuous strictly monotone function and let $I=f(\mathbb{R})$. By [Loya, Thm 4.27], we know that $I$ is an interval too. We shall prove that $f^{-1}: I \rightarrow \mathbb{R}$ is also a strictly monotone function; then Theorem 4.27 implies that $f^{-1}$ is continuous. Now suppose, for instance, that $f$ is strictly increasing; we shall prove that $f^{-1}$ is also strictly increasing. If $x<y$ both belong $I$, then $x=f(\xi)$ and $y=f(\eta)$ for some $\xi$ and $\eta$. Since $f$ is strictly increasing it must be that $\xi<\eta$, and hence, since $\xi=f^{-1}(x)$ and $\eta=f^{-1}(y)$, we have $f^{-1}(x)<f^{-1}(y)$. Thus, $f^{-1}$ is strictly increasing, and our proof is complete.
15. (Stoll, p. 176, \#12)

If $m \in \mathbb{Z} \backslash\{0\}, n \in \mathbb{N}$, prove that $f(x)=x^{m / n}$ is continuous on $(0, \infty)$. Note that $m=0$ implies $f(x)=x^{0}=1$ is continuous.
ANS: (Loya, Example 4.32) Note that for every $n \in \mathbb{N}$, the function $f(x)=x^{n}$ is strictly increasing on $[0, \infty)$. Therefore, $f^{-1}(x)=x^{1 / n}$ is continuous. In particular, for every $m \in \mathbb{N}, g(x)=x^{m / n}=$ $\left(x^{1 / n}\right)^{m}$ is continuous on $[0, \infty)$, being a composition of the continuous functions $f^{-1}$ and the $m$ th power. Similarly, the function $x \mapsto x^{m / n}$ for $m \in \mathbb{Z}$ with $m<0$ is continuous on ( $0, \infty$ ). Therefore, for every $r \in \mathbb{Q}, x \mapsto x^{r}$ is continuous on $[0, \infty)$ if $r>0$ and on $(0, \infty)$ if $r<0$.
16. (Stoll, p. 205, \#22) Let $f(x)=x^{2}, g(x)=x^{3}, x \in[-1,1]$.
(a) Find $c \in(-1,1)$ such that the conclusion of the Cauchy Mean Value Theorem (Stoll, Thm 5.2.8) holds.

ANS: The assumptions of the theorem are satisfied, i.e., $f, g$ are continuous real-valued functions on $[a, b]$ that are differentiable on $(a, b)$. With $a=-1, b=1$, we need to solve for

$$
0=[f(b)-f(a)] g^{\prime}(c)=[g(b)-g(a)] f^{\prime}(c)=[1-(-1)] \times 2 c,
$$

so $c=0$.
(b) Show that there does not exist any $c \in(-1,1)$ for which $\frac{f(1)-f(-1)}{g(1)-g(-1)}=\frac{f^{\prime}(c)}{g^{\prime}(c)}$.

ANS: We need to rule out the case $c=0$, as $g^{\prime}(0)=0$. We then need to attempt to solve

$$
0=\frac{f(1)-f(-1)}{g(1)-g(-1)}=\frac{f^{\prime}(c)}{g^{\prime}(c)}=\frac{2 c}{3 c^{2}}=(2 / 3) c^{-1}
$$

which has no solution in $(-1,1) \backslash\{0\}$.
17. (Stoll, p. 204, \#13) Give an example of a uniformly continuous function on $[0,1]$ that is differentiable on $(0,1)$ but for which $f^{\prime}$ is not bounded on $(0,1)$.
ANS: Take $f(x)=\sqrt{x}$, noting $f^{\prime}(0+)$.
18. (Stoll, p. 204, $\# 11(\mathrm{a}))$ Suppose $f$ is differentiable on an interval $I$. Prove that $f^{\prime}$ is bounded on $I$ if and only if there exists a constant $M$ such that $|f(x)-f(y)| \leq M|x-y|$ for all $x, y \in I$.
Note: There is the related question from Petrovic, \#4.4.27: Suppose that $f$ has a bounded derivative on ( $a, b$ ). Prove that $f$ is uniformly continuous on $(a, b)$.
ANS: (Chen) Stoll's question:
$(\Longrightarrow)$ Take any $x, y \in I$. If $x=y$, then $0 \leq 0$ holds vacuously. Assume, w.l.o.g., $x<y$. Suppose that $f^{\prime}$ is bounded on $I$, i.e., $\exists M>0$ such that, $\forall c \in I,\left|f^{\prime}(c)\right| \leq M$. Since $f$ is differentiable on $I, f$ is continuous on $[x, y]$ and differentiable on $(x, y)$. Thus, the MVT applies to $f$ (but not necessarily $f^{\prime}$, which was not assumed continuous), so that $\exists x_{0} \in(x, y)$ such that $f(y)-f(x)=f^{\prime}\left(x_{0}\right)(y-x)$. As $\left|f^{\prime}\left(x_{0}\right)\right| \leq M$ and $y-x>0$,

$$
-M(y-x) \leq f(y)-f(x)=f^{\prime}\left(x_{0}\right)(y-x) \leq M(y-x) .
$$

That is, $|f(y)-f(x)| \leq M(y-x)=M|y-x|$.
$(\Longleftarrow)$ Suppose $\forall x, y \in I,|f(y)-f(x)| \leq M|y-x|$, where $M>0$. Fix an arbitrary $x \in I$. Then

$$
\forall y \in I, y \neq x, \quad\left|\frac{f(y)-f(x)}{y-x}\right| \leq M, \quad \text { or } \quad-M \leq f^{\prime}(x)=\lim _{y \rightarrow x} \frac{f(y)-f(x)}{y-x} \leq M .
$$

The limit exists because $f$ is differentiable on $I$. That is, $\forall x \in I,\left|f^{\prime}(x)\right| \leq M$, i.e., $f^{\prime}$ is bounded on $I$.
19. (l'Hospital's rule)
(a) State l'Hospital's rule (Stoll, Thm 5.3.2)

ANS: Suppose $f, g$ are real-valued differentiable functions on $(a, b)$, with $g^{\prime}(x) \neq 0$ for all $x \in(a, b)$, where $-\infty \leq a<b \leq \infty$. Suppose

$$
\lim _{x \rightarrow a^{+}} \frac{f^{\prime}(x)}{g^{\prime}(x)}=L, \quad \text { where } \quad L \in \mathbb{R} \cup\{-\infty, \infty\}
$$

If (i) $\lim _{x \rightarrow a^{+}} f(x)=0$ and $\lim _{x \rightarrow a^{+}} g(x)=0$, or (ii) $\lim _{x \rightarrow a^{+}} g(x)= \pm \infty$, then $\lim _{x \rightarrow a^{+}} \frac{f(x)}{g(x)}=L$. NOTE: In (i), there are two cases: $a$ is finite; and $a$ is $-\infty$. There are separate proofs for each. The same two cases hold for (ii), but Stoll gives the proof that applies to both cases simultaneously.
(b) (Stoll, p. 212, \#2) Suppose $f, g$ are differentiable on $(a, b), x_{0} \in(a, b)$, and $g^{\prime}\left(x_{0}\right) \neq 0$. If $f\left(x_{0}\right)=g\left(x_{0}\right)=0$, prove that

$$
\lim _{x \rightarrow x_{0}} \frac{f(x)}{g(x)}=\frac{f^{\prime}\left(x_{0}\right)}{g^{\prime}\left(x_{0}\right)}
$$

(Hint: apply the definition of the derivative.)
Note: This is a very useful exercise because it provides a very simple proof of l'Hospital's rule in this simplified case.
ANS: For $x \neq x_{0}$, write

$$
\begin{equation*}
\frac{f(x)-0}{g(x)-0}=\frac{f(x)-f\left(x_{0}\right)}{x-x_{0}} \frac{x-x_{0}}{g(x)-g\left(x_{0}\right)}=\frac{f(x)-f\left(x_{0}\right)}{x-x_{0}} / \frac{g(x)-g\left(x_{0}\right)}{x-x_{0}} . \tag{55}
\end{equation*}
$$

Next recall the definition of the derivative; and (Stoll, Thm 4.1.6(c)): Suppose $E$ is a subset of a metric space $X, f, g: E \rightarrow \mathbb{R}$, and $p$ is a limit point of $E$. If $\lim _{x \rightarrow p} f(x)=A$ and $\lim _{x \rightarrow p} g(x)=B$, then $\lim _{x \rightarrow p} \frac{f(x)}{g(x)}=\frac{A}{B}$, provided $B \neq 0$.
The result follows by taking the limit, $x \rightarrow x_{0}$, in (55).
(c) Evaluate $\lim _{x \rightarrow 0+} x \log x$ and $\lim _{x \rightarrow \infty} x \log (1+1 / x)$, these being common and important $0 \cdot \infty$ examples.
ANS:

$$
\begin{aligned}
\lim _{x \rightarrow 0+} x \log x & =\lim _{x \rightarrow 0+} \frac{\log x}{1 / x}=\lim _{x \rightarrow 0+} \frac{1 / x}{-1 / x^{2}}=-\lim _{x \rightarrow 0+} x=0 . \\
\lim _{x \rightarrow \infty} x \log (1+1 / x) & =\lim _{x \rightarrow \infty} \frac{\log (1+1 / x)}{1 / x}=\lim _{x \rightarrow \infty} \frac{[1 /(1+1 / x)]\left(-1 / x^{2}\right)}{-1 / x^{2}} \\
& =\lim _{x \rightarrow \infty} \frac{1}{1+1 / x}=1 .
\end{aligned}
$$

20. Let $f:(0, \infty) \rightarrow \mathbb{R}$ be given by $f(x)=x^{1 / x}$. What are $\lim _{x \rightarrow \infty} f(x)$ and $\lim _{x \rightarrow 0} f(x)$ ? Derive the extrema of $f$.
ANS: From exercise 5, we know, for $n \in \mathbb{N}$, $\lim _{n \rightarrow \infty} n^{1 / n}=1$. As such, we conjecture from a continuity argument that $\lim _{x \rightarrow \infty} f(x)=1$. We can be more rigorous: Write $f(x)=\exp \left(\log x^{1 / x}\right)$ and note $\lim _{x \rightarrow \infty} \log x^{1 / x}=\lim _{x \rightarrow \infty}(\log x) / x$ is the $\infty / \infty$ form, so l'Hospital gives $\lim _{x \rightarrow \infty} \log x^{1 / x}=$ $\lim _{x \rightarrow \infty} 1 / x=0$, so, from continuity of $\exp , \lim _{x \rightarrow \infty} f(x)=\exp \left(\lim _{x \rightarrow \infty} \log x^{1 / x}\right)=1$.
For the limit to zero,

$$
\lim _{x \rightarrow 0} f(x)=\lim _{n \rightarrow \infty} f(1 / n)=\lim _{n \rightarrow \infty} 1 / n^{n}=0
$$

Next,

$$
f^{\prime}(x)=\exp (\ln (x) / x) \frac{x(1 / x)-\ln (x)}{x^{2}}=\frac{x^{1 / x}(1-\ln (x))}{x^{2}}
$$

and setting to zero yields gives $\ln (x)=1$, or $x=e$. Function $f$ is continuous, with limits at zero and infinity given by 0 and 1 , respectively, and there is only one extremum point, with $x=e>1$, so $x=e$ is the single maximum of $f$.
21. (Stoll, p. 213, \#7) Let $f(x)=x^{2} \sin (1 / x)$ and $g(x)=\sin x$. Show that $\lim _{x \rightarrow 0}[f(x) / g(x)]$ exists but that $\lim _{x \rightarrow 0}\left[f^{\prime}(x) / g^{\prime}(x)\right]$ does not exist.
ANS: Write

$$
\lim _{x \rightarrow 0} \frac{f(x)}{g(x)}=\lim _{x \rightarrow 0} \frac{x \sin (1 / x)}{\sin (x) / x}=\frac{\lim _{x \rightarrow 0} x \sin (1 / x)}{\lim _{x \rightarrow 0} \sin (x) / x}=\frac{0}{1}=0
$$

from Stoll, EXAMPLES 4.1.10 (c) and (d). Next,

$$
\begin{aligned}
& f^{\prime}(x)=x^{2} \cos (1 / x)\left(-x^{-2}\right)+2 \sin (1 / x) x=-\cos (1 / x)+2 \sin (1 / x) x \\
& g^{\prime}(x)=\cos (x)
\end{aligned}
$$

with $\lim _{x \rightarrow 0} \cos (x)=1$, but $\lim _{x \rightarrow 0} f^{\prime}(x)$ does not exist, because of the cosine term.
22. (Stoll, p. 213, \#9a) Let $f(x)=(\sin x) / x$ for $x \neq 0$, and $f(0)=1$ Show that $f^{\prime}(0)$ exists, and determine its value.
ANS: Stoll's provided solution just says $f^{\prime}(0)=0$. From Stoll, Def 5.1.1, the Newton quotient is (multiply by $x / x) \lim _{x \rightarrow 0}[f(x)-1] / x=\lim _{x \rightarrow 0}[\sin (x)-x] / x^{2}$. Apply l'Hospital to get $\lim _{x \rightarrow 0}[\cos (x)-$ $1] /(2 x)$, apply again to get $-\lim _{x \rightarrow 0} \sin (x) / 2=0$.
23. (Application of Stoll, Lemma 5.4.3; Trench, Example 2.5.7) First recall Stoll, Lemma 5.4.3:

Suppose $f:[a, b] \rightarrow \mathbb{R}$ is such that $f$ and $f^{\prime}$ are continuous on $[a, b]$ and $f^{\prime \prime}(x)$ exists for all $x \in(a, b)$. Let $x_{0} \in[a, b]$. Then for any $x \in[a, b]$, there exists a real number $\zeta$ between $x_{0}$ and $x$ such that

$$
\begin{equation*}
f(x)=f\left(x_{0}\right)+f^{\prime}\left(x_{0}\right)\left(x-x_{0}\right)+\frac{1}{2} f^{\prime \prime}(\zeta)\left(x-x_{0}\right)^{2} . \tag{56}
\end{equation*}
$$

Replacing $x$ by $x_{0}+h$, this can also be written (and used below) as

$$
\begin{equation*}
\forall x_{0} \in[a, b], \quad \exists c \in\left(x_{0}, x_{0}+h\right): f\left(x_{0}+h\right)=f\left(x_{0}\right)+h f^{\prime}\left(x_{0}\right)+\frac{h^{2}}{2} f^{\prime \prime}(c) \tag{57}
\end{equation*}
$$

Our first task is to extend (56) and, in particular, (57), to the next order. This is in fact yet another special case of a Taylor series expansion, proven in general in all books. We will require that $f^{(3)}(x)$ exists for all $x \in(a, b)$, resulting in

$$
\begin{equation*}
\forall x_{0} \in[a, b], \quad \exists c \in\left(x_{0}, x_{0}+h\right): f\left(x_{0}+h\right)=f\left(x_{0}\right)+h f^{\prime}\left(x_{0}\right)+\frac{h^{2}}{2} f^{\prime \prime}\left(x_{0}\right)+\frac{h^{3}}{6} f^{\prime \prime \prime}(c) . \tag{58}
\end{equation*}
$$

In numerical analysis, forward differences are used to approximate derivatives. If $h>0$, the first and second forward differences with spacing $h$ are defined by $\Delta f(x)=f(x+h)-f(x)$ and

$$
\begin{align*}
\Delta^{2} f(x)=\Delta[\Delta f(x)] & =\Delta f(x+h)-\Delta f(x)  \tag{59}\\
& =f(x+2 h)-2 f(x+h)+f(x)
\end{align*}
$$

Though we will not use it, the reader can attempt by induction to show (Trench, p. 111, \#18(b)), with $\Delta^{0} f(x)=f(x)$,

$$
\Delta^{n} f(x)=\sum_{m=0}^{n}(-1)^{n-m}\binom{n}{m} f(x+m h), \quad n \geq 1
$$

Here is the exercise: For an appropriately defined function $f: \mathbb{R} \rightarrow \mathbb{R}$, derive upper bounds for the magnitudes of the errors in the approximations

$$
\begin{equation*}
f^{\prime}\left(x_{0}\right) \approx \frac{\Delta f\left(x_{0}\right)}{h} \tag{60}
\end{equation*}
$$

and

$$
\begin{equation*}
f^{\prime \prime}\left(x_{0}\right) \approx \frac{\Delta^{2} f\left(x_{0}\right)}{h^{2}} \tag{61}
\end{equation*}
$$

The general statement is (Trench, p. 112, \#22) If $f^{(n+1)}$ is bounded on an open interval containing $x_{0}$ and $x_{0}+n h$, then

$$
\left|\frac{\Delta^{n} f\left(x_{0}\right)}{h^{n}}-f^{(n)}\left(x_{0}\right)\right| \leq A_{n} M_{n+1} h,
$$

where $A_{n}$ is a constant independent of $f$ and

$$
M_{n+1}=\sup _{x_{0}<c<x_{0}+n h}\left|f^{(n+1)}(c)\right|
$$

ANS: If $f^{\prime \prime}$ exists on an open interval containing $x_{0}$ and $x_{0}+h$, we can use (57) to estimate the error in (60) by writing

$$
\begin{equation*}
f\left(x_{0}+h\right)=f\left(x_{0}\right)+f^{\prime}\left(x_{0}\right) h+\frac{f^{\prime \prime}(c) h^{2}}{2} \tag{62}
\end{equation*}
$$

where $x_{0}<c<x_{0}+h$. We can rewrite (62) as

$$
\frac{f\left(x_{0}+h\right)-f\left(x_{0}\right)}{h}-f^{\prime}\left(x_{0}\right)=\frac{f^{\prime \prime}(c) h}{2}, \text { i.e., } \quad \frac{\Delta f\left(x_{0}\right)}{h}-f^{\prime}\left(x_{0}\right)=\frac{f^{\prime \prime}(c) h}{2}
$$

Therefore, with $M_{2}$ an upper bound for $\left|f^{\prime \prime}\right|$ on $\left(x_{0}, x_{0}+h\right)$,

$$
\left|\frac{\Delta f\left(x_{0}\right)}{h}-f^{\prime}\left(x_{0}\right)\right| \leq \frac{M_{2} h}{2}
$$

If $f^{\prime \prime \prime}$ exists on an open interval containing $x_{0}$ and $x_{0}+2 h$, we can use (58) to get expressions for $f\left(x_{0}+h\right)$ and $f\left(x_{0}+2 h\right)$, and estimate the error in (61) by writing

$$
f\left(x_{0}+h\right)=f\left(x_{0}\right)+h f^{\prime}\left(x_{0}\right)+\frac{h^{2}}{2} f^{\prime \prime}\left(x_{0}\right)+\frac{h^{3}}{6} f^{\prime \prime \prime}\left(c_{0}\right)
$$

and

$$
f\left(x_{0}+2 h\right)=f\left(x_{0}\right)+2 h f^{\prime}\left(x_{0}\right)+2 h^{2} f^{\prime \prime}\left(x_{0}\right)+\frac{4 h^{3}}{3} f^{\prime \prime \prime}\left(c_{1}\right)
$$

where $x_{0}<c_{0}<x_{0}+h$ and $x_{0}<c_{1}<x_{0}+2 h$. These two equations imply that

$$
f\left(x_{0}+2 h\right)-2 f\left(x_{0}+h\right)+f\left(x_{0}\right)=h^{2} f^{\prime \prime}\left(x_{0}\right)+\left[\frac{4}{3} f^{\prime \prime \prime}\left(c_{1}\right)-\frac{1}{3} f^{\prime \prime \prime}\left(c_{0}\right)\right] h^{3},
$$

which can be rewritten as

$$
\frac{\Delta^{2} f\left(x_{0}\right)}{h^{2}}-f^{\prime \prime}\left(x_{0}\right)=\left[\frac{4}{3} f^{\prime \prime \prime}\left(c_{1}\right)-\frac{1}{3} f^{\prime \prime \prime}\left(c_{0}\right)\right] h
$$

because of (59). Therefore, with $M_{3}$ an upper bound for $\left|f^{\prime \prime \prime}\right|$ on $\left(x_{0}, x_{0}+2 h\right)$,

$$
\left|\frac{\Delta^{2} f\left(x_{0}\right)}{h^{2}}-f^{\prime \prime}\left(x_{0}\right)\right| \leq \frac{5 M_{3} h}{3}
$$

24. (Convex and Concave Functions) We investigate the relations between convexity and continuity and derivatives. Convexity is an extremely important property in optimization. A few analysis books cover the basics of this (not Stoll, and only very lightly in Fitzpatrick). There are many books dedicated to convexity and optimization.
Most students will have learned the following: Given a twice differentiable function $f: \mathbb{R} \rightarrow \mathbb{R}, f$ is convex if $f^{\prime \prime}(x) \geq 0$ for all $x \in \mathbb{R}$. Likewise, $f$ is concave if $f^{\prime \prime}(x) \leq 0$ for all $x \in \mathbb{R}$. For example, $f(x)=a x^{2}+b x+c$ is convex if $a \geq 0$, and is concave if $a \leq 0$.
These definitions are too specific, requiring $f$ to be twice differentiable. The following gives the more general definition and some basic results. I mostly draw from the solid presentation in Ghorpade and Limaye. Figure 8 shows an illustration of convex and concave functions.
Geometrically, a function is convex if the line segment joining any two points on its graph lies on or above the graph. A function is concave if any such line segment lies on or below the graph. Another geometrically obvious fact is that each tangent line of the function lies entirely below the graph of the function. More specifically, Let $f:(a, b) \rightarrow \mathbb{R}$ be a convex function. Then, for every point $c \in(a, b)$, one can prove there exists a line $L$ in $\mathbb{R}^{2}$ with the following properties:
(a) $L$ passes through the point $(c, f(c))$.
(b) The graph of $f$ lies entirely above $L$.


Figure 8: From Ghorpade and Limaye, page 25

Any line satisfying the above is referred to as a tangent line for $f$ at $c$. Note that $f$ does not need to be differentiable. If not, then the slope of a tangent line may not be uniquely determined. As an example, consider $f:[0,1] \rightarrow \mathbb{R}$, with $f(x)=x / 2$ for $x \in[0,1 / 2]$; and $f(x)=x-1 / 2$, for $x \in(1 / 2,1]$. Another canonical example is $f(x)=|x|$.
This gives rise to a very simple proof that
every convex function is continuous.
As in https://e.math.cornell.edu/people/belk/measuretheory/Inequalities.pdf:
Let $f:(a, b) \rightarrow \mathbb{R}$ be a convex function, and let $c \in(a, b)$. Let $L$ be a linear function whose graph is a tangent line for $f$ at $c$, and let $P$ be a piecewise linear function consisting of two chords to the graph of $f$ meeting at $c$. See Figure 9. Then $L \leq f \leq P$ in a neighborhood of $c$, and $L(c)=f(c)=P(c)$. As $L$ and $P$ are continuous at $c$, it follows from the Squeeze Theorem and the sequential definition of continuity that $f$ is also continuous at $c$.


Figure 9: Convex function $f$ is continuous at each point in an open interval of its domain. Taken from https://e.math. cornell.edu/people/belk/measuretheory/Inequalities.pdf.

Analytically, for $x_{1}<x<x_{2}$, we think in terms of the slope of the line from $x_{1}$ to $x$, compared to the slope of the line from $x_{1}$ to $x_{2}$. For convex, the latter should be larger than the former. Specifically, let $D \subseteq \mathbb{R}$ be such that $D$ contains an interval $I$, and let $f: D \rightarrow \mathbb{R}$ be a function. We say that $f$ is convex on $I$ if

$$
\begin{equation*}
x_{1}, x_{2}, x \in I, x_{1}<x<x_{2} \Longrightarrow f(x)-f\left(x_{1}\right) \leq \frac{f\left(x_{2}\right)-f\left(x_{1}\right)}{x_{2}-x_{1}}\left(x-x_{1}\right), \tag{64}
\end{equation*}
$$

and $f$ is concave on $I$ if

$$
\begin{equation*}
x_{1}, x_{2}, x \in I, x_{1}<x<x_{2} \Longrightarrow f(x)-f\left(x_{1}\right) \geq \frac{f\left(x_{2}\right)-f\left(x_{1}\right)}{x_{2}-x_{1}}\left(x-x_{1}\right) . \tag{65}
\end{equation*}
$$

An alternative way, and the one more commonly seen in the literature, to formulate the definitions of convexity and concavity is as follows. First note that, for all $x_{1}, x_{2} \in \mathbb{R}$ with $x_{1}<x_{2}$, the points $x$ between $x_{1}$ and $x_{2}$ are of the form $(1-t) x_{1}+t x_{2}$ for some $t \in(0,1)$; in fact, $t$ and $x$ determine each other uniquely, since

$$
x=(1-t) x_{1}+t x_{2} \Longleftrightarrow t=\frac{x-x_{1}}{x_{2}-x_{1}} .
$$

Substituting this into the previous definition, we see that $f$ is convex on $I$ if (and only if)

$$
\forall x_{1}, x_{2} \in I, \quad x_{1}<x_{2}, \quad \forall t \in(0,1), \quad f\left((1-t) x_{1}+t x_{2}\right) \leq(1-t) f\left(x_{1}\right)+t f\left(x_{2}\right)
$$

The roles of $t$ and $1-t$ can be readily reversed, and with this in view, one need not assume that $x_{1}<x_{2}$. Thus, $f$ is convex on $I$ if (and only if)

$$
\begin{equation*}
f\left(t x_{1}+(1-t) x_{2}\right) \leq t f\left(x_{1}\right)+(1-t) f\left(x_{2}\right) \quad \text { for all } x_{1}, x_{2} \in I \text { and } t \in(0,1) \tag{66}
\end{equation*}
$$

Similarly, $f$ is concave on $I$ if (and only if)

$$
\begin{equation*}
f\left(t x_{1}+(1-t) x_{2}\right) \geq t f\left(x_{1}\right)+(1-t) f\left(x_{2}\right) \quad \text { for all } x_{1}, x_{2} \in I \text { and } t \in(0,1) \tag{67}
\end{equation*}
$$

We state without proof the following result: (Ghorpade and Limaye, p. 102, \#3.34) A function $f:(a, b) \rightarrow \mathbb{R}$ is convex if and only if it is continuous on $(a, b)$ and satisfies

$$
\begin{equation*}
f\left(\frac{x_{1}+x_{2}}{2}\right) \leq \frac{f\left(x_{1}\right)+f\left(x_{2}\right)}{2}, \quad \forall x_{1}, x_{2} \in(a, b) \tag{68}
\end{equation*}
$$

We will in fact prove one direction of this below, and for a more general linear combination of $x_{i}$. The result is well-known, and very important; it is called Jensen's inequality.
We turn now to the exercise, the first three of which are from Stoll's "Miscellaneous Exercises" for his chapter 5 , which means they are more difficult and delve into further topics beyond the main syllabus.
((a), (c), (d), from Stoll, p. 221, \#3) A function $f$ is convex (or concave up) on the interval $(a, b)$ if for any $x, y \in(a, b)$, and $0<t<1, f(t x+(1-t) y) \leq t f(x)+(1-t) f(y)$.
(a) Let $f$ be convex on $(a, b)$. Prove that $f$ is continuous on $(a, b)$. [Without appealing to geometric arguments as above; so purely analytic.] [Also holds for $f$ concave.]
ANS: (Ghorpade and Limaye, Prop 3.15) Let $I$ be an open interval in $\mathbb{R}$ and let $f: I \rightarrow \mathbb{R}$ be convex on $I$ or concave on $I$.
First, suppose $f$ is convex. Let $c \in I$. Then there is $r>0$ such that $[c-r, c+r] \subseteq I$. Let $M:=\max \{f(c-r), f(c+r)\}$. For each $x \in[c-r, c+r]$, there is $t \in[0,1]$ such that $x=(1-t)(c-r)+t(c+r)$, and, hence, from (66),

$$
\begin{equation*}
f(x) \leq(1-t) f(c-r)+t f(c+r) \leq(1-t) M+t M=M \tag{69}
\end{equation*}
$$

Given any $\epsilon>0$ with $\epsilon \leq 1$, and $x \in \mathbb{R}$, we claim that

$$
|x-c| \leq r \epsilon \Longrightarrow x \in I \text { and }|f(x)-f(c)| \leq \epsilon(M-f(c)) .
$$

Suppose $|x-c| \leq r \epsilon$. Then $x \in[c-r, c+r]$, since $\epsilon \leq 1$, and so $x \in I$. Define

$$
y:=c+\frac{x-c}{\epsilon} \quad \text { and } \quad z:=c-\frac{x-c}{\epsilon} .
$$

Then $|y-c|=|z-c|=|x-c| / \epsilon \leq r$, and so $y, z \in[c-r, c+r]$. Moreover,

$$
x=(1-\epsilon) c+\epsilon y \quad \text { and } \quad c=\frac{1}{1+\epsilon} x+\frac{\epsilon}{1+\epsilon} z .
$$

Since $f$ is convex and $0<\epsilon \leq 1$, we see that

$$
\begin{equation*}
f(x) \leq(1-\epsilon) f(c)+\epsilon f(y), \quad \text { that is, } \quad f(x)-f(c) \leq \epsilon(f(y)-f(c)) \tag{70}
\end{equation*}
$$

Recall $y \in[c-r, c+r]$ and, for each $y \in[c-r, c+r]$, (69) implies $f(y) \leq M$. Thus, (70) implies that $f(x)-f(c) \leq \epsilon(M-f(c))$. Also, as $f$ is convex and $x, y, z \in[c-r, c+r]$,

$$
f(c) \leq \frac{1}{1+\epsilon} f(x)+\frac{\epsilon}{1+\epsilon} f(z), \quad \text { that is, } \quad(1+\epsilon) f(c) \leq f(x)+\epsilon f(z)
$$

The last inequality implies that $f(c)-f(x) \leq \epsilon(f(z)-f(c)) \leq \epsilon(M-f(c))$. It follows that $|f(x)-f(c)| \leq \epsilon(M-f(c))$, and thus the claim is established. The result of continuity of $f$ at $c$ follows from the $\delta-\epsilon$ definition of continuity. If $f$ is concave, it suffices to apply the result just proved to $-f$.
(b) Prove at least part (i) of the following theorem, which is taken (along with the proof of (i)) from Ghorpade and Limaye, Prop 4.33.

Let $I$ be an interval containing more than one point, and let $f: I \rightarrow \mathbb{R}$ be a differentiable function. Then
(i) $f^{\prime}$ is monotonically increasing on $I \Longleftrightarrow f$ is convex on $I$.
(ii) $f^{\prime}$ is monotonically decreasing on $I \Longleftrightarrow f$ is concave on $I$.
(iii) $f^{\prime}$ is strictly increasing on $I \Longleftrightarrow f$ is strictly convex on $I$.
(iv) $f^{\prime}$ is strictly decreasing on $I \Longleftrightarrow f$ is strictly concave on $I$.

ANS: Proof of (i). First, assume that $f^{\prime}$ is monotonically increasing on $I$. Let $x_{1}, x_{2}, x \in I$ be such that $x_{1}<x<x_{2}$. By the MVT, there are $c_{1} \in\left(x_{1}, x\right)$ and $c_{2} \in\left(x, x_{2}\right)$ satisfying

$$
f(x)-f\left(x_{1}\right)=f^{\prime}\left(c_{1}\right)\left(x-x_{1}\right) \quad \text { and } \quad f\left(x_{2}\right)-f(x)=f^{\prime}\left(c_{2}\right)\left(x_{2}-x\right)
$$

Now $c_{1}<c_{2}$ and $f^{\prime}$ is monotonically increasing on $I$, and so

$$
\frac{f(x)-f\left(x_{1}\right)}{x-x_{1}}=f^{\prime}\left(c_{1}\right) \leq f^{\prime}\left(c_{2}\right)=\frac{f\left(x_{2}\right)-f(x)}{x_{2}-x} .
$$

Collecting only the terms involving $f(x)$ on the left side, we obtain

$$
f(x)\left(\frac{1}{x-x_{1}}+\frac{1}{x_{2}-x}\right) \leq \frac{f\left(x_{1}\right)}{x-x_{1}}+\frac{f\left(x_{2}\right)}{x_{2}-x} .
$$

Multiplying throughout by $\left(x-x_{1}\right)\left(x_{2}-x\right) /\left(x_{2}-x_{1}\right)$, we see that

$$
f(x) \leq \frac{f\left(x_{1}\right)\left(x_{2}-x\right)+f\left(x_{2}\right)\left(x-x_{1}\right)}{x_{2}-x_{1}}=f\left(x_{1}\right)+\frac{f\left(x_{2}\right)-f\left(x_{1}\right)}{x_{2}-x_{1}}\left(x-x_{1}\right)
$$

where the last equality follows by writing $x_{2}-x=\left(x_{2}-x_{1}\right)-\left(x-x_{1}\right)$. Thus, recalling (64), $f$ is convex on $I$.
Conversely, assume that $f$ is convex on $I$. Let $x_{1}, x_{2}, x \in I$ be such that $x_{1}<x<x_{2}$. Then

$$
\begin{aligned}
f(x) & \leq f\left(x_{1}\right)+\frac{f\left(x_{2}\right)-f\left(x_{1}\right)}{x_{2}-x_{1}}\left(x-x_{1}\right)=f\left(x_{1}\right)+\frac{f\left(x_{2}\right)-f\left(x_{1}\right)}{x_{2}-x_{1}}\left[\left(x_{2}-x_{1}\right)-\left(x_{2}-x\right)\right] \\
& =f\left(x_{1}\right)+\left[f\left(x_{2}\right)-f\left(x_{1}\right)\right]-\frac{f\left(x_{2}\right)-f\left(x_{1}\right)}{x_{2}-x_{1}}\left(x_{2}-x\right)=f\left(x_{2}\right)-\frac{f\left(x_{2}\right)-f\left(x_{1}\right)}{x_{2}-x_{1}}\left(x_{2}-x\right) .
\end{aligned}
$$

As a consequence, the slopes of chords are increasing, that is,

$$
\frac{f(x)-f\left(x_{1}\right)}{x-x_{1}} \leq \frac{f\left(x_{2}\right)-f\left(x_{1}\right)}{x_{2}-x_{1}} \leq \frac{f\left(x_{2}\right)-f(x)}{x_{2}-x} .
$$

Taking limits as $x \rightarrow x_{1}^{+}$and $x \rightarrow x_{2}^{-}$, we obtain

$$
f^{\prime}\left(x_{1}\right) \leq \frac{f\left(x_{2}\right)-f\left(x_{1}\right)}{x_{2}-x_{1}} \leq f^{\prime}\left(x_{2}\right)
$$

Thus, $f^{\prime}$ is monotonically increasing on $I$.
(c) Suppose $f^{\prime \prime}(x)$ exists for all $x \in(a, b)$. Prove:

$$
\begin{equation*}
f \text { convex on }(a, b) \Longleftrightarrow \forall x \in(a, b), \quad f^{\prime \prime}(x) \geq 0 \tag{71}
\end{equation*}
$$

ANS: Just apply exercise $24 \mathrm{~b}(\mathrm{i})$ and exercise 4 a .
(d) If $f$ is convex on $(a, b)$, prove that $f_{+}^{\prime}(p)$ and $f_{-}^{\prime}(p)$ exist for every $p \in(a, b)$. Hint: Use the following result (Ghorpade and Limaye, p. 40, \#1.63):

Let $I$ be an interval containing more than one point and let $f: I \rightarrow \mathbb{R}$ be a function. Define $\phi\left(x_{1}, x_{2}\right):=\left(f\left(x_{1}\right)-f\left(x_{2}\right)\right) /\left(x_{1}-x_{2}\right)$ for $x_{1}, x_{2} \in I$ with $x_{1} \neq x_{2}$. Show that $f$ is convex on $I$ if and only if $\phi$ is a monotonically increasing function of $x_{1}$, that is,

$$
\begin{equation*}
\forall x_{1}, x_{2} \in I, \quad x_{1}<x_{2}, \quad \forall x \in I \backslash\left\{x_{1}, x_{2}\right\}, \quad \phi\left(x_{1}, x\right) \leq \phi\left(x_{2}, x\right) . \tag{72}
\end{equation*}
$$

The proof of this is asked for in the next exercise.
Finally, show by example that a convex function on $(a, b)$ need not be differentiable on $(a, b)$.
ANS: The case of interest is when $f$ is not differentiable at some $p \in(a, b)$. From (63) and exercise 24a, we know $f$ is continuous at $p$, i.e., $f(p)=f_{+}(p)=f_{-}(p)$. From (47) and (48), we need to show the existence of

$$
f_{+}^{\prime}(p)=\lim _{h \rightarrow 0^{+}} \frac{f(p+h)-f(p)}{h} \quad \text { and } \quad f_{-}^{\prime}(p)=\lim _{h \rightarrow 0^{-}} \frac{f(p+h)-f(p)}{h} .
$$

Recall that a limit of function $f: D \rightarrow \mathbb{R}$ as $x \rightarrow c$ exists, denoted $f(x) \rightarrow \ell$ as $x \rightarrow c$, or $\lim _{x \rightarrow c} f(x)=\ell$, if there exists $\ell \in \mathbb{R}$ such that, for any sequence $\left\{x_{n}\right\} \in D \backslash\{c\}$ with $x_{n} \rightarrow c$, $f\left(x_{n}\right) \rightarrow \ell$. Consider $f_{+}^{\prime}(p)$ and let $x_{n}=p+1 / n$, for $n \in\left\{n_{0}, n_{0}+1, \ldots\right\}$, where $n_{0}$ is the smallest value of $n \in \mathbb{N}$ such that $p+1 / n<b$ (which we know exists, by invoking the wellordering principle and the Archimedean Property). Let $h, k \in \mathbb{N}$ such that $n_{k}>n_{h} \geq n_{0}$, so $p<x_{n_{k}}<x_{n_{h}}$. From (64),

$$
\begin{equation*}
\frac{f\left(x_{n_{k}}\right)-f(p)}{x_{n_{k}}-p} \leq \frac{f\left(x_{n_{h}}\right)-f(p)}{x_{n_{h}}-p} \tag{73}
\end{equation*}
$$

The result now follows from (72). The proof for $f_{-}^{\prime}(p)$ is similar, or possibly could be elicited from that of $f_{+}^{\prime}(p)$ and some clever "symmetry" argument, defining some function $g$ in terms of $f$.
For the example: For any $a>0$, let $I=[-a, a]$, and $f: I \rightarrow \mathbb{R}$ defined by $f(x)=|x|$. Function $f$ is clearly convex, but not differentiable at the interior point $0 \in(-a, a)$. Similarly, $-f$ is concave on $I$, but not differentiable at 0 .
(e) Prove (72).

ANS: (Ralf)
Montonically increasing if convex: Suppose that $f$ is convex. Assume for contradiction that $\phi\left(x_{1}, x\right)$ is not monotonically increasing in $x_{1}$. That is, we can find $x_{1}, x_{2} \in I$ with $x_{1}<x_{2}$ and $x \in I \backslash\left\{x_{1}, x_{2}\right\}$ such that $\phi\left(x_{1}, x\right)>\phi\left(x_{2}, x\right)$. Assume $x_{1}<x<x_{2} .{ }^{14}$ Note that, since $x_{1}<x<x_{2}, \exists t \in(0,1)$ such that $t x_{1}+(1-t) x_{2}=x$. Then

$$
\begin{aligned}
\frac{f\left(x_{1}\right)-f\left(t x_{1}+(1-t) x_{2}\right)}{x_{1}-\left(t x_{1}+(1-t) x_{2}\right)} & >\frac{f\left(x_{2}\right)-f\left(t x_{1}+(1-t) x_{2}\right)}{x_{2}-\left(t x_{1}+(1-t) x_{2}\right)} \\
\Leftrightarrow \frac{f\left(x_{1}\right)-f\left(t x_{1}+(1-t) x_{2}\right)}{(1-t)\left(x_{1}-x_{2}\right)} & >\frac{f\left(x_{2}\right)-f\left(t x_{1}+(1-t) x_{2}\right)}{-t\left(x_{1}-x_{2}\right)} \\
\Leftrightarrow-t\left[f\left(x_{1}\right)-f\left(t x_{1}+(1-t) x_{2}\right)\right] & >(1-t)\left[f\left(x_{2}\right)-f\left(t x_{1}+(1-t) x_{2}\right)\right] \\
\Leftrightarrow f\left(t x_{1}+(1-t) x_{2}\right) & >(1-t) f\left(x_{2}\right)+t f\left(x_{1}\right)
\end{aligned}
$$

Note that going from the second to third row we multiply by $-t(1-t)\left(x_{1}-x_{2}\right)>0$, therefore the inequalities do not switch. The last line contradicts that $f$ is convex. We therefore conclude that $\phi\left(x_{1}, x\right)$ is monotonically increasing in $x_{1}$.

Convex if montonically increasing: Conversely, suppose that $\phi\left(x_{1}, x\right) \leq \phi\left(x_{2}, x\right)$ for any $x_{1}, x_{2} \in I$ and $x \in I \backslash\left\{x_{1}, x_{2}\right\}$. Now assume for contradiction that $f$ is not convex, i.e., $\exists t \in(0,1)$ and $x_{1}, x_{2} \in I\left(\text { with } x_{1} \neq x_{2}\right)^{15}$ such that $f\left(t x_{1}+(1-t) x_{2}\right)>t f\left(x_{1}\right)+(1-t) f\left(x_{2}\right)$. Assume w.l.o.g. that $x_{1}<x_{2}$. Let $x=t x_{1}+(1-t) x_{2}$, then clearly $x \in I$. Then

$$
\begin{aligned}
& f\left(t x_{1}+(1-t) x_{2}\right)>t f\left(x_{1}\right)+(1-t) f\left(x_{2}\right) \\
& \Leftrightarrow f\left(t x_{1}+(1-t) x_{2}\right)-f\left(x_{2}\right)>t\left(f\left(x_{1}\right)-f\left(x_{2}\right)\right) \\
& \Leftrightarrow \frac{f\left(t x_{1}+(1-t) x_{2}\right)-f\left(x_{2}\right)}{t x_{1}+(1-t) x_{2}-x_{2}}<\frac{t\left(f\left(x_{1}\right)-f\left(x_{2}\right)\right)}{t x_{1}+(1-t) x_{2}-x_{2}} \\
& \Leftrightarrow \frac{f\left(t x_{1}+(1-t) x_{2}\right)-f\left(x_{2}\right)}{t x_{1}+(1-t) x_{2}-x_{2}}<\frac{f\left(x_{1}\right)-f\left(x_{2}\right)}{x_{1}-x_{2}} \\
& \Leftrightarrow \phi\left(x, x_{2}\right)<\phi\left(x_{1}, x_{2}\right)
\end{aligned}
$$

The inequality flips in the third line when we divide both sides by $\left(t x_{1}+(1-t) x_{2}-x_{2}\right)<0$. The last line contradicts our assumption that $\phi\left(x_{1}, x\right)$ is monotonically increasing in $x_{1}$, because $x_{1}<x$ but $\phi\left(x, x_{2}\right)<\phi\left(x_{1}, x_{2}\right)$. Therefore, we conclude that $f$ is convex.
(f) (Young's inequality) Let $a, b \in \mathbb{R}_{\geq 0}$ and $p, q \in(1, \infty)$ such that $1 / p+1 / q=1$. Prove:

$$
\begin{equation*}
a b \leq \frac{a^{p}}{p}+\frac{b^{q}}{q} . \tag{74}
\end{equation*}
$$

ANS: (Nair, Lemma 5.2.3) Function $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ defined by $\varphi(x)=e^{x}, \quad x \in \mathbb{R}$ is convex, i.e., for every $x, y \in \mathbb{R}$ and $0<\lambda<1, \varphi(\lambda x+(1-\lambda) y) \leq \lambda \varphi(x)+(1-\lambda) \varphi(y)$. Taking $\lambda=1 / p$ we have $1-\lambda=1 / q$ and

$$
e^{x / p+y / q} \leq \frac{e^{x}}{p}+\frac{e^{y}}{q}
$$

Now, taking $x>0$ and $y>0$ such that $a=e^{x / p}$ and $b=e^{y / q}$, that is, $x=\ln \left(a^{p}\right)$ and $y=\ln \left(b^{q}\right)$, we obtain (74).

[^12](g) (Jensen's inequality, Finite Version) Let $\varphi:(a, b) \rightarrow \mathbb{R}$ be a convex function, where $-\infty \leq a<$ $b \leq \infty$, and let $x_{1}, \ldots, x_{n} \in(a, b)$. Then
\[

$$
\begin{equation*}
\varphi\left(\lambda_{1} x_{1}+\cdots+\lambda_{n} x_{n}\right) \leq \lambda_{1} \varphi\left(x_{1}\right)+\cdots+\lambda_{n} \varphi\left(x_{n}\right) \tag{75}
\end{equation*}
$$

\]

for any $\lambda_{1}, \ldots, \lambda_{n} \in[0,1]$ satisfying $\lambda_{1}+\cdots+\lambda_{n}=1$. Prove (75).
NOTE: A more general version that subsumes this case is Jensen's inequality for the Lebesgue integral. An excellent presentation can be found in M. Thamban Nair's Measure and Integration: A First Course, Thm 5.2.5. Nair subsequently also shows that Young's inequality is a special case of Jensen.
ANS: (https://e.math.cornell.edu/people/belk/measuretheory/Inequalities.pdf) Let $c=\lambda_{1} x_{1}+\cdots+\lambda_{n} x_{n}$, and let $L$ be a linear function whose graph is a tangent line for $\varphi$ at $c$. Since $\lambda_{1}+\cdots+\lambda_{n}=1$, we know that $L\left(\lambda_{1} x_{1}+\cdots+\lambda_{n} x_{n}\right)=\lambda_{1} L\left(x_{1}\right)+\cdots+\lambda_{n} L\left(x_{n}\right)$. As $L \leq \varphi$ and $L(c)=\varphi(c)$, we conclude that

$$
\begin{aligned}
\varphi(c)=L(c) & =L\left(\lambda_{1} x_{1}+\cdots+\lambda_{n} x_{n}\right) \\
& =\lambda_{1} L\left(x_{1}\right)+\cdots+\lambda_{n} L\left(x_{n}\right) \leq \lambda_{1} \varphi\left(x_{1}\right)+\cdots+\lambda_{n} \varphi\left(x_{n}\right) .
\end{aligned}
$$

(h) Recall the AM-GM inequality (6). A more general version (allowing for unequal weights) is: (AM-GM Inequality) Let $x_{1}, \ldots, x_{n}>0$, and let $\lambda_{1}, \ldots, \lambda_{n} \in[0,1]$ so that $\lambda_{1}+\cdots+\lambda_{n}=1$. Then

$$
\begin{equation*}
x_{1}^{\lambda_{1}} \cdots x_{n}^{\lambda_{n}} \leq \lambda_{1} x_{1}+\cdots+\lambda_{n} x_{n} . \tag{76}
\end{equation*}
$$

Prove (76) using convexity.
ANS: This theorem is equivalent to the convexity of the exponential function. Specifically, from (75),

$$
\exp \left\{\lambda_{1} t_{1}+\cdots \lambda_{n} t_{n}\right\} \leq \lambda_{1} e^{t_{1}}+\cdots+\lambda_{n} e^{t_{n}}, \quad \forall t_{1}, \ldots, t_{n} \in \mathbb{R}
$$

Substituting $x_{i}=e^{t_{i}}$ gives the desired result.

### 4.3 Stoll Chapter 6 (Integration, FTC)

1. (Stoll, p. 238, \#6) If $f, g \in \mathcal{R}[a, b]$ with $f(x) \leq g(x)$ for all $x \in[a, b]$, prove that $\int_{a}^{b} f \leq \int_{a}^{b} g$.

ANS: Let $\mathcal{P}=\left\{x_{0}, x_{1}, \ldots, x_{n}\right\}$ be a partition of $[a, b]$. The key observation is, as $f(x) \leq g(x)$ for all $x \in[a, b]$, and recalling (5),

$$
\forall i \in\{1, \ldots, n\}, \sup \left\{f(t): t \in\left[x_{i-1}, x_{i}\right]\right\} \leq \sup \left\{g(t): t \in\left[x_{i-1}, x_{i}\right]\right\}
$$

As a consequence $\mathcal{U}(\mathcal{P}, f) \leq \mathcal{U}(\mathcal{P}, g)$ for all partitions $\mathcal{P}$ of $[a, b]$. Taking the infimum over $\mathcal{P}$ gives $\overline{\int_{a}^{b}} f \leq \overline{\int_{a}^{b}} g$. The result now follows from the fact that $f, g \in \mathcal{R}[a, b]$.
2. (Stoll, p. 238, \#7)
(a) Suppose $f$ is continuous on $[a, b]$ with $f(x) \geq 0$ for all $x \in[a, b]$. If $\int_{a}^{b} f=0$, prove that $f(x)=0$ for all $x \in[a, b]$.
ANS: Here are three ideas for the proof.
i. Recall the result (Stoll, p. 157, \#21; see also exercise 10b):

Let $E \subset \mathbb{R}$ and let $f$ be a real-valued function on $E$ that is continuous at $p \in E$. If $f(p)>0$, prove that there exists an $\alpha>0$ and a $\delta>0$ such that $f(x) \geq \alpha$ for all $x \in N_{\delta}(p) \cap E$.
This, along with continuity of $f$ imply that the integral cannot be zero. The contrapositive (crucially under the assumption that $f$ is continuous) implies the desired result.
ii. In general, if $\int_{a}^{b} f=0$ and $f \geq 0$, then $f$ can differ from the identically zero function only on a set of measure zero. Choose partition $P$ of $[a, b]$ such that, in each of the $n$ intervals, $\left[x_{i-1}, x_{i}\right]$, there is at most one value of $x$ such that $f(x)>0$. This is possible because the set of such points has measure zero. For any particular $i$, the interval has nonzero length, say $\delta_{0}$. Then, "almost all" $f\left(\left[x_{i-1}, x_{i}\right]\right)$ are zero (at most one is not, by construction), and we want to show that, in fact, literally all of them are. Function $f$ is continuous on a closed interval, so it is uniformly continuous. This means, for any $\epsilon>0, \exists \delta_{1}$ such that $|f(x)-f(y)|<\epsilon$, when $|x-y|<\delta_{1}$. Let $\delta=\min \left(\delta_{0}, \delta_{1}\right)$, so that, $\forall x, y \in\left[x_{i-1}, x_{i}\right]$, at least one of $f(x), f(y)$ is zero, say $f(x)$, and, as $f \geq 0,0 \leq f(y)<\epsilon$. As $\epsilon>0$ is arbitrary, $f(y)=0$, i.e., $f([a, b])=0$.
iii. As $f(x) \geq 0$, and $\int_{a}^{b} f=0$, for any closed interval $I \subset[a, b], \int_{I} f$ must be 0 also (because if it were positive, some other interval must result in a negative value to offset it so that the total integral is 0 , but this cannot happen because the function is non-negative). So, for all intervals $I \subset[a, b]$, the Mean Value Theorem for Integrals (Stoll, Thm 6.3.6) implies $\exists c \in I$ such that $f(c)=0$. Now by the Nested Intervals Theorem, there is a single point remaining in a decreasing sequence of closed intervals, so that no value of $x$ such that $f(x)>0$ can exist.
NOTE: (Manfred Stoll, after my inquiry with him about my above ideas): This idea has merit. Assume $\exists c$ such that $f(c)>0$. now choose decreasing intervals whose intersection is $c$ and apply the IVT on each of the intervals.
NOTE: (Chen) Since $f(x) \geq 0$, for any closed interval $I \subset[a, b], 0 \leq \int_{I} f \leq \int_{a}^{b} f=0$, i.e., $\int_{I} f=0$. Take an arbitrary $c \in[a, b]$. Since $c$ is a limit point of $[a, b]$, we can construct a set of nested closed intervals $I_{n}=\left[a_{n}, b_{n}\right] \subset[a, b]$ with $a_{n}<b_{n}$ such that $a_{n} \leq a_{n+1} \leq$ $c \leq b_{n+1} \leq b_{n}$ and $\lim _{n \rightarrow \infty} a_{n}=c=\lim _{n \rightarrow \infty} b_{n}$. By the Mean Value Theorem for Integrals (Stoll, Thm 6.3.6), $\exists c_{n} \in\left[a_{n}, b_{n}\right]$ such that $f\left(c_{n}\right)=0$. Then, by the Squeeze Theorem, $\lim _{n \rightarrow \infty} c_{n}=c$. By continuity of $f$, we get $f(c)=f\left(\lim _{n \rightarrow \infty} c_{n}\right)=\lim _{n \rightarrow \infty} f\left(c_{n}\right)=0$. Note: in order to get $f\left(c_{n}\right)=0$ from MVT, we need $\forall n, a_{n}<b_{n}$.
(b) Show by example that the conclusion in part (a) may be false if $f$ is not continuous.

ANS: See, e.g., Stoll, Example 6.1.14(a); or, more simply, take a finite or countably infinite set of points on the interval and, starting with a function $f(x)=0$, set the function equal to a positive value (recall $f(x) \geq 0$ ) for these points.
3. (Stoll, p. 239, \#9) Suppose $f$ is a nonnegative Riemann integrable function on $[a, b]$ satisfying $f(r)=0$ for all $r \in \mathbb{Q} \cap[a, b]$. Prove that $\int_{a}^{b} f=0$.
ANS: Here are four proofs. The first shows, for any $\epsilon^{\prime}>0, \int_{a}^{b} f=\inf \{\mathcal{U}(\mathcal{P}, f)\}<\epsilon^{\prime}$. The second proof demonstrates $\int_{a}^{b} f=\sup \{\mathcal{L}(\mathcal{P}, f)\}=0$. The third proof extends these results to show that the result holds for function $g$ with the same properties of $f$, except without the nonnegativity constraint. The fourth proof directly shows $\int_{a}^{b} g=0$, where $g$ is not necessarily nonnegative, by using Riemann's definition of the integral.

- (Marc): As $f$ is Riemann integrable, this cannot be the Dirichlet function. We only learned so far that $f$ needs to be bounded, and (Stoll, Thm 6.1.8) can be continuous, or monotone. Let $\mathbb{I}=\mathbb{R} \backslash \mathbb{Q}$; and, for short, write $\mathbb{Q}^{\prime}=\mathbb{Q} \cap[a, b] ;$ and $\mathbb{I}^{\prime}=\mathbb{I} \cap[a, b]$. We have

$$
\int_{a}^{b} f=\int_{\mathbb{Q}^{\prime}} f+\int_{\mathbb{I}^{\prime}} f
$$

Irrespective of the value of $f$ on $\mathbb{Q}^{\prime}, \int_{\mathbb{Q}^{\prime}} f$ is zero, because $\mathbb{Q}^{\prime}$ is of measure zero. However, it is crucial for this question that $f(r)=0$ for all $r \in \mathbb{Q}^{\prime}$. Recall one of the definitions of a dense set, e.g., exercise 3: A set $S$ of real numbers is said to be dense in $\mathbb{R}$ provided that every interval $I=(a, b)$, where $a<b$, contains a member of $S$. And recall that $\mathbb{Q}$, and thus $\mathbb{Q}^{\prime}$, is dense in $\mathbb{R}$. Consider a point $p \in \mathbb{I}^{\prime}$. Then, $\forall \delta>0$, interval $(p-\delta, p+\delta)$ contains a rational number (in fact, countably infinitely many).

We recall the basic definitions associated with the Riemann integral for $f$ a bounded real-valued function on $[a, b]$. Given a partition $\mathcal{P}=\left\{x_{0}, x_{1}, \ldots, x_{n}\right\}$ of $[a, b]$, for each $i=1,2, \ldots, n$, let

$$
m_{i}=\inf \left\{f(t): x_{i-1} \leq t \leq x_{i}\right\}, \quad M_{i}=\sup \left\{f(t): x_{i-1} \leq t \leq x_{i}\right\}
$$

The upper and lower sums $\mathcal{U}(\mathcal{P}, f)$ and $\mathcal{L}(\mathcal{P}, f)$ for $\mathcal{P}$ and $f$ are denoted

$$
\mathcal{U}(\mathcal{P}, f)=\sum_{i=1}^{n} M_{i} \Delta x_{i}, \quad \mathcal{L}(\mathcal{P}, f)=\sum_{i=1}^{n} m_{i} \Delta x_{i}
$$

For any bounded real function, the upper and lower integrals of $f$ are respectively defined as

$$
\overline{\int_{a}^{b}} f=\inf \{\mathcal{U}(\mathcal{P}, f)\}, \quad \underline{\int_{a}^{b}} f=\sup \{\mathcal{L}(\mathcal{P}, f)\}
$$

where the inf and sup are over all partitions $\mathcal{P}$ of $[a, b]$. The Riemann integral exists when these latter two quantities are equal, denoted $\int_{a}^{b} f$.
Finally, (Stoll, Thm 6.1.7): A bounded real-valued function $f$ is Riemann integrable on $[a, b]$ if and only if for every $\epsilon>0$, there exists a partition $\mathcal{P}$ of $[a, b]$ such that $\mathcal{U}(\mathcal{P}, f)-\mathcal{L}(\mathcal{P}, f)<\epsilon$. Furthermore, if $\mathcal{P}$ is a partition of $[a, b]$ for which this holds, then the inequality also holds for all refinements of $\mathcal{P}$.

Given that $f$ is Riemann integrable and nonnegative, and such that, $\forall r \in \mathbb{Q}^{\prime}, f(r)=0$, we have $m_{i}=0$ for $i=1, \ldots, n$, because $\mathbb{Q}^{\prime}$ is dense in $\mathbb{R}$, so that $\mathcal{L}(\mathcal{P}, f)=0$. Thus, $\forall \epsilon>0$,

$$
\mathcal{U}(\mathcal{P}, f)-\mathcal{L}(\mathcal{P}, f)=\mathcal{U}(\mathcal{P}, f)=\sum_{i=1}^{n} M_{i} \Delta x_{i} \leq\|P\| \sum_{i=1}^{n} M_{i}<\epsilon
$$

where $\|P\|=\max \left\{\Delta x_{j}: j=1,2, \ldots, n\right\}$. Thus, $\int_{a}^{b} f=\inf \{\mathcal{U}(\mathcal{P}, f)\}<\epsilon /\|P\|$.

- (Ralf): Let $\mathcal{P}=\left\{x_{0}, x_{1}, \ldots, x_{n}\right\}$ be a partition of $[a, b]$. The key observation is, as $\mathbb{Q}$ is dense in $\mathbb{R}$, that for every $i \in\{1, \ldots, n\}, \exists r \in \mathbb{Q}$ in $\Delta x_{i}$. Therefore, (further noting that the function is nonnegative),

$$
\forall i \in\{1, \ldots, n\}, \quad \inf \left\{f(t): t \in\left[x_{i-1}, x_{i}\right]\right\}=0
$$

As a consequence, $\mathcal{L}(\mathcal{P}, f)=0$ for any partitions $\mathcal{P}$ of $[a, b]$ so that

$$
\underline{\int_{a}^{b}} f=\sup \{\mathcal{L}(\mathcal{P}, f): \mathcal{P} \text { is a partition of }[a, b]\}=0
$$

As $f$ is Riemann integrable, we require $\overline{\int_{a}^{b}} f=\underline{\int_{a}^{b}} f=\int_{a}^{b} f$ and consequently, $\int_{a}^{b} f=0$.

- (Chen, extension to not necessarily nonnegative, using above results): Suppose that $g$ is Riemann integrable on $[a, b]$, satisfying $g(r)=0, \forall r \in \mathbb{Q} \cap[a, b]$. By Stoll, Thm 6.2.2, $|g|$ is also Riemann integrable on $[a, b]$ and $\left|\int_{a}^{b} g\right| \leq \int_{a}^{b}|g|$. Note that $|g|$ is now nonnegative, Riemann integrable, satisfying $|g(r)|=0, \forall r \in \mathbb{Q} \cap[a, b]$. Thus, from the above results for the nonnegative case, $\int_{a}^{b}|g|=0$. Then, $0 \leq\left|\int_{a}^{b} g\right| \leq \int_{a}^{b}|g|=0$, i.e., $\int_{a}^{b} g=0$.
- (Chen, independent proof of the result not imposing nonnegativity): Suppose $g$ is a Riemann integrable function on $[a, b]$ satisfying $f(r)=0$ for all $r \in \mathbb{Q} \cap[a, b]$. Prove that $\int_{a}^{b} g=0$.
Proof: Recall Stoll, Def 6.2.4, Def 6.2.5 and Thm 6.2.6. Let $\mathcal{P}=\left\{x_{0}, x_{1}, \ldots, x_{n}\right\}$ be a partition of $[a, b]$. By density of $\mathbb{Q}$, there exists a rational number $r_{i} \in\left[x_{i-1}, x_{i}\right]$ for each $i=1,2, \ldots, n$. As $g$ is Riemann integrable on $[a, b], \int_{a}^{b} g=\lim _{\|\mathcal{P}\| \rightarrow 0} \sum_{i=1}^{n} g\left(r_{i}\right) \Delta x_{i}=\lim _{\|\mathcal{P}\| \rightarrow 0} \sum_{i=1}^{n} 0 \cdot \Delta x_{i}=0$.

4. (Bounded and Total Variation) A function $f$ is said to be of bounded variation on $[a, b]$ if there is a number $K$ such that

$$
\begin{equation*}
\sum_{j=1}^{n}\left|f\left(a_{j}\right)-f\left(a_{j-1}\right)\right| \leq K, \quad \text { for any } a=a_{0}<a_{1}<\cdots<a_{n}=b \tag{77}
\end{equation*}
$$

The smallest number with this property is the total variation of $f$ on $[a, b]$.
NOTE: This is commonly expressed as follows. The total variation of a real-valued (or more generally complex-valued) function $f$, defined on an interval $[a, b] \subset \mathbb{R}$ is the quantity

$$
V_{a}^{b}(f)=\sup _{\mathcal{P}} \sum_{i=0}^{n_{P}-1}\left|f\left(x_{i+1}\right)-f\left(x_{i}\right)\right|,
$$

where the supremum runs over the set of all partitions $\mathcal{P}=\left\{P=\left\{x_{0}, \ldots, x_{n_{P}}\right\} \mid P\right.$ is a partition of $[a, b]\}$ of the given interval.
The concept arises in various mathematical contexts, including measure theory and stochastic calculus. An interesting further result (from https://en.wikipedia.org/wiki/Total_variation) is this:

The total variation of a differentiable function $f$, defined on an interval $[a, b] \subset \mathbb{R}$, has the following expression if $f^{\prime}$ is Riemann integrable:

$$
V_{a}^{b}(f)=\int_{a}^{b}\left|f^{\prime}(x)\right| \mathrm{d} x
$$

If $f$ is differentiable and monotonic, then the above simplifies to $V_{a}^{b}(f)=|f(a)-f(b)|$. For any differentiable function $f$, we can decompose the domain interval [ $a, b$ ], into subintervals $\left[a, a_{1}\right],\left[a_{1}, a_{2}\right], \ldots,\left[a_{N}, b\right]$ (with $a<a_{1}<a_{2}<\cdots<a_{N}<b$ ) in which $f$ is locally monotonic, then the total variation of $f$ over $[a, b]$ can be written as the sum of local variations on those subintervals:

$$
\begin{aligned}
V_{a}^{b}(f) & =V_{a}^{a_{1}}(f)+V_{a_{1}}^{a_{2}}(f)+\cdots+V_{a_{N}}^{b}(f) \\
& =\left|f(a)-f\left(a_{1}\right)\right|+\left|f\left(a_{1}\right)-f\left(a_{2}\right)\right|+\cdots+\left|f\left(a_{N}\right)-f(b)\right| .
\end{aligned}
$$

(a) (Trench, p. 135, \#7) Prove: If $f$ is of bounded variation on $[a, b]$, then $f$ is bounded on $[a, b]$. ANS: Let $V$ be the total variation of $f$ on $[a, b]$. Then, as

$$
\begin{array}{r}
f(x)=\frac{f(a)+f(b)}{2}+\frac{(f(x)-f(a))+(f(x)-f(b))}{2}, \quad \forall a<x<b, \\
|f(x)| \leq \frac{|f(a)+f(b)|}{2}+\frac{|f(a)-f(x)|+|f(x)-f(b)|)}{2} \leq \frac{|f(a)+f(b)|+V}{2} .
\end{array}
$$

(b) (Trench, p. 135, \#7) Prove: If $f$ is of bounded variation on $[a, b]$, then $f$ is integrable on $[a, b]$. Hint: Use:

- Trench, p. 114) If $f$ is defined on $[a, b]$, then a sum

$$
\sigma=\sum_{j=1}^{n} f\left(c_{j}\right)\left(x_{j}-x_{j-1}\right), \quad \text { where } \quad x_{j-1} \leq c_{j} \leq x_{j}, \quad 1 \leq j \leq n
$$

is a Riemann sum of $f$ over the partition $P=\left\{x_{0}, x_{1}, \ldots, x_{n}\right\}$. Note: As $c_{j}$ can be chosen arbitrarily in $\left[x_{j}, x_{j-1}\right]$, there are infinitely many Riemann sums for a given function $f$ over a given partition $P$.

- Trench, Thm 3.1.4: Let $f$ be bounded on $[a, b]$, and let $P$ be a partition of $[a, b]$. Then
i. The upper sum $S(P)=\sum_{j=1}^{n} M_{j}\left(x_{j}-x_{j-1}\right)$ of $f$ over $P$ is the supremum of the set of all Riemann sums of $f$ over $P$.
ii. The lower sum $s(P)=\sum_{j=1}^{n} m_{j}\left(x_{j}-x_{j-1}\right)$ of $f$ over $P$ is the infimum of the set of all Riemann sums of $f$ over $P$.
- Trench, Thm 3.2.7: If $f$ is bounded on $[a, b]$, then $f$ is integrable on $[a, b]$ if and only if, for each $\epsilon>0$, there is a partition $P$ of $[a, b]$ for which $S(P)-s(P)<\epsilon$.
ANS: (From Trench, but corrected and augmented by Ralf)
Note by part (a), as $f$ is a function of bounded variation, $f$ is bounded.
Let $P=\left\{x_{0}, x_{1}, \ldots, x_{n}\right\}$ be a partition of $[a, b]$ and $\epsilon>0$. From [Trench, Thm 3.1.4], we can choose $c_{1}, \ldots, c_{n}$ and $c_{1}^{\prime}, \ldots, c_{n}^{\prime}$ so that $x_{j-1} \leq c_{j}, c_{j}^{\prime} \leq x_{j}$, and

$$
\text { (A) }\left|S(P)-\sum_{j=1}^{n} f\left(c_{j}\right)\left(x_{j}-x_{j-1}\right)\right|<\epsilon / 2, \quad(B)\left|s(P)-\sum_{j=1}^{n} f\left(c_{j}^{\prime}\right)\left(x_{j}-x_{j-1}\right)\right|<\epsilon / 2 .
$$

Adding and subtracting the term $\sum_{j=1}^{n}\left(f\left(c_{j}\right)-f\left(c_{j}^{\prime}\right)\right)\left(x_{j}-x_{j-1}\right)$ to $S(P)-s(P)$ :

$$
\begin{aligned}
S(P)-s(P) & =S(P)-s(P) \\
& -\sum_{j=1}^{n} f\left(c_{j}\right)\left(x_{j}-x_{j-1}\right) \\
& +\sum_{j=1}^{n}\left(f\left(c_{j}\right)-f\left(c_{j}^{\prime}\right)\right)\left(x_{j}-x_{j-1}\right) \\
& +\sum_{j=1}^{n} f\left(c_{j}^{\prime}\right)\left(x_{j}-x_{j-1}\right)
\end{aligned}
$$

Reordering the terms and applying the triangle inequality:

$$
\begin{aligned}
S(P)-s(P) & \leq \underbrace{\left|S(P)-\sum_{j=1}^{n} f\left(c_{j}\right)\left(x_{j}-x_{j-1}\right)\right|}_{A} \\
& +\left|\sum_{j=1}^{n}\left(f\left(c_{j}\right)-f\left(c_{j}^{\prime}\right)\right)\left(x_{j}-x_{j-1}\right)\right| \\
& +\underbrace{\left|\sum_{j=1}^{n} f\left(c_{j}^{\prime}\right)\left(x_{j}-x_{j-1}\right)-s(P)\right|}_{B}
\end{aligned}
$$

Recall that $\|P\|$ is the norm of the partition, defined as $\|P\|=\max \left\{\Delta x_{j}: j=1,2, \ldots, n\right\}$. Using (A) and (B), if $\|P\|<\epsilon / K$, where $K$ is defined in (77),

$$
S(P)-s(P) \leq \epsilon+\sum_{j=1}^{n}\left|f\left(c_{j}\right)-f\left(c_{j}^{\prime}\right)\right|\left(x_{j}-x_{j-1}\right) \leq \epsilon+K\|P\|<2 \epsilon
$$

[Trench, Thm 3.2.7] then implies that $f$ is integrable on $[a, b]$.
5. (Differentiating Under the Integral) This is work in progress.

ANS:


[^0]:    ${ }^{1}$ Robert B. Ash, Real Variables, with Basic Metric Space Topology, 2007.

[^1]:    ${ }^{2}$ Former Professor of Economics and of Sociology at the University of Chicago, and recipient of the 1992 Nobel Memorial Prize in Economic Sciences.

[^2]:    ${ }^{3}$ Matthew A. Pons, Real Analysis for the Undergraduate: With an Invitation to Functional Analysis, 2014. From the several chapters I have read, the book is extremely attractive.

[^3]:    ${ }^{4}$ Professor, School of Mathematics, Georgia Institute of Technology. Chris is the author of several outstanding books, notably an undergraduate Calculus textbook, but also, and of more interest to us, an accessible book on metric spaces, and a first-year graduate (thus accessible) book on measure theory. See https://heil.math.gatech.edu/.

[^4]:    ${ }^{5}$ Argue by induction
    ${ }^{6}$ Convergent sequence in $A$ exists since $A$ is bounded.

[^5]:    ${ }^{7}$ See, e.g., Heil (Metrics), Thm 3.2.5 and Sec. 3.6; Jacob and Evans (Vol. II), p. 37; Gopal et al. (An Introduction to Metric Spaces), Sec. 2.10.

[^6]:    ${ }^{8}$ One might be tempted to argue: As $E$ is bounded, and all subsets of $X$ are closed, Heine-Borel (HB) implies $E$ is compact. But HB is not applicable here. To generalise it to any metric space (and therefore also the real numbers with the non-usual metric) we require that the metric space is what is called totally bounded, i.e., for any $\epsilon>0$, we need to cover $X$ with finitely many balls of a fixed radius $\epsilon$. See, e.g., Jacob and Evans, A Course in Analysis, Vol II, Def 3.15. Clearly this fails because, for $\epsilon<1$, each ball contains only one point and we thus require infinitely many balls to cover $[a, b]$. Totally bounded is discussed in, e.g., Heil, Metrics, Norms, Inner Products, and Operator Theory, Sec. 2.8; other books on metric spaces; and those on functional analysis.

[^7]:    ${ }^{9}$ This rules out cases where the rhs is of the form $0 \cdot \infty$, as mentioned by Stoll. Petrovic does not mention this requirement. Can this be proven without it?

[^8]:    ${ }^{10}$ For the curious, a nice example is given on https://math.stackexchange.com/questions/234216, as follows: Suppose $\mu(X)=\infty$. We want the integral of non-negative simple functions to satisfy $\int \sum_{i=1}^{n} c_{i} \chi_{E_{i}} d \mu=\sum_{i=1}^{n} c_{i} \mu\left(E_{i}\right)$. In particular for the 0 function:

    $$
    0=\int 0 d \mu=\int 0 \cdot \chi_{X} d \mu=0 \cdot \int \chi_{X} d \mu=0 \cdot \mu(X)=0 \cdot \infty
    $$

[^9]:    ${ }^{11}$ Arthur P. Mattuck, Introduction to Analysis, 1999.

[^10]:    ${ }^{12}$ Hossein Giv, Mathematical Analysis and Its Inherent Nature, 2016.

[^11]:    ${ }^{13}$ Brian S. Thomson, Judith B. Bruckner, and Andrew M. Bruckner, Elementary Real Analysis, Second Edition, 2008.

[^12]:    ${ }^{14}$ The other potential cases, $x<x_{1}<x_{2}$ and $x_{1}<x_{2}<x$, similarly lead to a contradiction of convexity of $f$ by expressing the "middle" value as a mix of the largest and smallest.
    ${ }^{15}$ The case $x_{1}=x_{2}$ trivially fulfills the requirements of convexity for any function.

