A Supplement to ASU's MAT 370: Examples and Applications

Jackson Carpenter

Contents

1 Introduction

Graduating as a mathematics major at ASU usually boils down to one final hurdle: passing MAT 371, also know as Advanced Calculus I. Classes beyond Advanced Calculus I are more or less 'à la carte' and can be taken depending on the interests of the individual student. However, for most students, MAT 371 is the last math class they need to take. In fact, for almost all mathematics majors at ASU (including Computational Mathematical Sciences, Actuarial Science, Statistics, Math for Secondary Education, and the Mathematics (BA) track), MAT 371 is the last 'necessary course' with the MAT prefix on major maps. In addition, students from other disciplines (physics, business, economics, computer sciences, etc.) might enroll in MAT 371 for personal interest or for applications in their main field. What this all means is that MAT 371 classes can be filled with students who are entering their first truly advanced mathematical course or those who would like it to be their last. MAT 371, as most students first foray into real analysis and analysis in general, can also be notoriously difficult. From ASU data on students enrolled in MAT 371 over the last 4 semesters (Fall 2018 - Spring 2020), some 57% of students who take the class finish with a grade C or lower. Of these students 42% withdraw from the course and 10% receive a failing grade. This sector of students represent those who leave MAT 371 with negative connotations of real analysis and perhaps advanced, abstract mathematics in general.

For this thesis, I want to make a supplement for a course similar to MAT 371 offered at ASU: MAT 370, or Intermediate Analysis. The scope of MAT 370, in comparison with MAT 371, is more aligned with what I plan to present. For examples, topics such as sequences of functions and series in general are covered in MAT 371, but will not be discussed in this thesis. I want this thesis to alleviate some of the difficulties of MAT 370 and cater to the needs of both non-experts, struggling students, and those who might come from a different major or background. This supplement aims to be exactly that, a supplement; students can use this treatment of MAT 370 material to both enrich and review what they see in class, or if nothing else, offer them a different proof or presentation of concepts. I will still be giving definitions and proving theorems, but I want to focus on content that might not normally be covered in a regular 75 minute class period: developing the motivation behind proofs and theorems, providing worked examples and counterexamples, and applying results to other fields when appropriate.

For preparatory work, I have researched both qualities of effective mathematics textbooks and qualities of effective teaching strategies for classes similar to MAT 370. There are of course many textbooks written on the subject of Intermediate Calculus, or Introductory Real Analysis, and I'm sure every teacher at ASU has their own set of notes for teaching the class or their favorite reference material. I will lay out in the next section what I will do to make this supplement unique for the audience I am hoping to benefit, but I hope that professors and students alike will find this work helpful during what can sometimes be a traumatic semester of MAT 370.

2 Preliminary Research

Before beginning the supplement itself, it would be remiss to not acknowledge some of the extensive research done on this subject. Beyond examining what has been done with respect to teaching courses like MAT 370, it is necessary to examine the qualities of effective textbooks and other learning materials as well. In this way, the supplement will aim to implement both effective conceptual and pedagogical notions with regards to introductory real analysis, but also to be an appealing textbook.

2.1 Textbook Creation

Much of the literature on textbook analysis is in the scope of elementary and high school education. However, this will still be useful for the purpose of this supplement given the intended audience. Non-experts and those who struggle in MAT 370 might not prefer or benefit from a standard "definition-theorem-proof" layout in a textbook [1]. Howson, in a personal perspective on the development of mathematics textbooks, presents a list of some of the attributes that a textbook reviewer might consider [2]:

- Mathematical coherence
- Clarity and accuracy of explanations
- The range, quantity and quality of the exercises
- The connections with real-life and with other curricular subjects
- The physical attractiveness of the texts: format, type, colour, illustrations
- Some signs of originality in material

Much of the information that Howson gives is in the context of elementary and high school education, but the meaning and intent for university-level students still hold. The first two attributes are quite straightforward and hopefully a standard for any mathematics text. The next two attributes, however, are of the utmost importance for this supplement. The quality of exercises for this supplement will be focused on both variety and using different techniques to get students more used to the semantics of formal mathematical language. This variety will help students choose their own strategy instead of seeing the same type of exercise with the same repeated technique. This type of "blocked" practice from the context of elementary education has a negative effect on students [3] and instills within them a sense of algorithmic problem-solving in which every exercise can be solved the same way. Instead, this supplement will incorporate the "interleaved" practice model [3]: working examples that do not use the same technique although they deal with the same content. This interleaved model, although seen in the context of elementary school textbooks, is especially conducive to a text of supplement of this level. Seeing the connections between these analysis concepts (sequences, continuity, and differentiability, for example) allows for worked examples that call back to previous content while being different in their own right.

The final two attributes of the list are the most surprising to be applied to university level content. Often times, textbook chapters appear to be long blocks of text after each other, with an occasional theorem or proof block to separate. This supplement will not only utilize color to make it more visually appealing and original, but to emphasize important concepts and differentiate what is meant to be proved and what has been proved [2]. Using different colors for definitions, theorems, lemmas, corollaries, etc. will help prepare the reader for what is coming.

The implementation of diagrams and mathematical imagery is yet another way to include color that has many more benefits to student understanding. For one, as discussed in Pinto and Weber, mathematical imagery is effective at linking the formal definitions and concepts to intuitive and semantic notions of analysis [4, 1]. Pinto highlights the story of a gifted mathematics student who is able, for example, to translate effortlessly between the formal definition and a graphical representation of its meaning [4]. While this situation is not a relatable one to many students in MAT 370, this supplement will help guide students to at least see the inner workings behind definitions and concepts. Furthermore, some concepts in real analysis lend themselves nicely to a more semantic and visual description such as topological notions (interiors, closures, etc.) as seen in [1].

Beyond visual elements of this supplement and content of the above list, one final attribute this supplement will focus on is the use of appropriate mathematical language. Raman, in her analysis on the presentation of the definition of continuity in different levels of textbooks, found that these textbooks send different epistemological messages to students [5]. The three calculus books, denoted *Pre-calculus*, *Calculus*, and *Analysis*, provided different levels of formality with the definition of continuity. The level of this supplement falls somewhere between *Calculus* and *Analysis* given that the former only gives the "$\lim_{x \to x_0} f(x) = f(x_0)$" definition while the latter gives one in terms of abstract metric spaces. This study describes that the *Analysis* definition lacks motivation although it is abstract, formal, and is referenced in later chapters of the book. Additionally, there are no informal connections given for the formal definition in this text. On the other hand, *Calculus* sends the message that formal definitions are merely made to "make an easy task cumbersome" and need not be remembered. This supplement will apply this situation as a model: formal definitions will always be given, but not without motivations and appropriate informal connections in the form of visuals or semantic discussion.

2.2 Introductory Real Analysis Pedagogy

Much has been written on the teaching of classes like MAT 370, or Introductory Real Analysis (IRA), but we focus on some examples in early stages of the course and how students take to these different learning methods. We begin with Roh's study of an ϵ-strip activity on the development of students' intuition based on prior learning [6]. Concepts discussed in the article include "primary" and "secondary intuition". Primary intuition in an IRA context includes all student knowledge of content that they may have learned in previous calculus classes. Secondary intuition, however, is that which is learned and has become evident "through systematic instruction." The standard, formal ϵ-N definition of convergence falls into the second category when fully understood. The difficulty here is filling the gaps between students' primary and secondary intuition. As discussed by Raman, textbooks may leave students with certain opinions about mathematical formality that are hard to sway [5]. When a newer, more formal definition arises to clash with a semantic, intuitive one that students already believe about convergence, there is inner conflict. Roh uses the classic example of the ϵ-strip activity for students to construct the formal definition of convergence through self-discovery [6]. The results of this study showed the development that students' secondary intuitions underwent during the ϵ-strip activity. Some of these new realizations included the arbitrary nature of ϵ and the dependence of N on ϵ (not the other way around). Given that sequences and their convergence lay at the heart of IRA content, we will employ a discussion similar to the ϵ-strip activity in the section on sequences.

Another pivotal part of garnering understanding in an IRA course that can be applied to a written supplement is the development of IRA heuristics as discussed by Weber in his analysis of an IRA professor's lectures [1]. This professor, Dr. T, wanted to provide his students with a *mathematical toolbox*, a "set of techniques and heuristics" that can be used in future analysis courses. One of the major topics emphasized by Dr. T is the importance of understanding inequalities, especially in the context of proving results about sequences and limits. From previous personal experience and reflection, tricks like the classic $\epsilon/2$ argument come to mind and are essential for growing an intuition or 'feel' for the type of proofs that are quite standard in an IRA course. Explicitly mentioning these techniques in the supplement will allow students to understand where they are coming from and why they are necessary, instead of simply seeing a fraction in a string of inequalities.

An interesting application of the importance of taking an IRA course concerns Math Education majors who might not see the benefit in succeeding in an abstract course like MAT 370. Wasserman and Weber find in their study on the pedagogical applications of an IRA course for secondary mathematics teachers that many of the proof concepts play a role in understanding math at the high school level. One example they discussed was that of rounding error when computing an answer via calculator [7]. Many students often don't wait to simplify their answers of long computation to the end and work with decimal estimates (like $\sin(12°) \approx 0.208$). With IRA experience, secondary math teachers could help their students understand why pre-rounding can have detrimental results in the end. For example, the concepts of limits and especially something like $|x_n - x| < \epsilon$ is essentially an approximation with precision ϵ. But understanding the way in which $x_n y_n \to xy$ and $x_n + y_n \to x + y$ (i.e. algebraic limit laws) can give secondary math teachers insight into how rounding errors may quite literally multiply. I would love to share applications like these in the manuscript to entice those in different fields to not give up on their IRA course.

3 The Real Number System

MAT 370, or Intermediate Calculus, is another name for the types of classes concerned with Introductory **Real** Analysis. Thus, before jumping into the analysis, we need to feel confident about working with the mathematical objects that make up the foundation of this course, the real numbers themselves. We all know and love the natural numbers, ($\mathbb{N} = \{1, 2, 3, \ldots\}$), the integers ($\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$) and the rationals ($\mathbb{Q} = \{\frac{m}{n} : m, n \in \mathbb{Z}, n \neq 0\}$), each a superset of the previous ($\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$). However, the real numbers, which we will denote as $\mathbb{R}$ includes special numbers like $\sqrt{2}$, π, e, etc. that make $\mathbb{R}$ an appropriate setting over which to do mathematical analysis that the other sets of numbers lack.

Definition 3.1: The Real Numbers, $\mathbb{R}$

We assume the existence of the set of numbers $\mathbb{R}$ that satisfy the following axioms. First, $\mathbb{R}$ is equipped with two binary operations, $+$ (addition) and $\cdot$ (multiplication) such that for all $x, y, z \in \mathbb{R}$:

1. (associative) $(x + y) + z = x + (y + z)$; $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.
2. (commutative) $x + y = y + x$; $x \cdot y = y \cdot x$.
3. (distributive) $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.
4. There exists a unique element $0 \in \mathbb{R}$ such that $x + 0 = x$ for all $x \in \mathbb{R}$.
5. For every $x \in \mathbb{R}$, there exists a unique $y \in \mathbb{R}$ such that $x + y = 0$.
6. There exists a unique element $1 \in \mathbb{R}$ such that $x \cdot 1 = x$ for all $x \in \mathbb{R}$.
7. For every nonzero $x \in \mathbb{R}$, there exists a unique $y \in \mathbb{R}$ such that $x \cdot y = 1$.

Many of these axioms probably seem fairly familiar from since the early days of studying mathematics. Everything we are used to doing with real numbers is still valid. For notation's sake, from associativity axiom we will not write parentheses between sums or products of more than two real numbers and omit the $\cdot$ in general as well: $x + y + z$ and xyz will be standard. Also, the unique additive and multiplicative (if $x \neq 0$) inverses of x will be denoted $-x$ and $\frac{1}{x} = x^{-1}$, respectively. Beyond just operations with real numbers, we will equip $\mathbb{R}$ with an order $<$, the standard "less than" or "greater than" comparison that we are familiar with as well.

Definition 3.2: The Standard Ordering $<$ on $\mathbb{R}$

The real number system $\mathbb{R}$ is equipped with an order relation $<$ such that for all $x, y, z \in \mathbb{R}$:

1. If $x < y$, then $x + z < y + z$.
2. If $x < y$ and $z > 0$, then $xz < yz$.
3. (transitive) If $x < y$ and $y < z$, then $x < z$.
4. (Trichotomy) For any $x, y \in \mathbb{R}$ exactly **one** of the following hold: $x < y$, $y < x$, or $x = y$.

To put these axioms to use, we prove some results about real numbers that are useful to know and not shocking by any means.

Example 3.3: Additional Properties of $\mathbb{R}$

The following properties of real numbers follow from the above axioms:

1. If $x + z = y + z$ then $x = y$.
2. $x \cdot 0 = 0$ for all $x \in \mathbb{R}$.
3. If $x < y$ then $-y < -x$.
4. $-x = (-1) \cdot x$ for all $x \in \mathbb{R}$.
5. $(-1) \cdot (-1) = 1$
6. $x^2 \geq 0$ for all $x \in \mathbb{R}$.
7. $0 < 1$.
8. If $0 < x < y$ then $0 < \frac{1}{y} < \frac{1}{x}$.

Proof.

1. Suppose that $x+z = y+z$. By Definition 3.1.5, z has an additive inverse that we add to both sides of $x+z = y+z$, $(x+z)+(-z) = (y+z)+(-z)$. By Definition 3.1.1, 3.1.5, and 3.1.4, we have that $(x+z)+(-z) = (y+z)+(-z)$. Then $x+(z+(-z)) = y+(z+(-z))$, which implies $x + 0 = y + 0$, which implies $x = y$.

2. By Definition 3.1.4, we know that for any $x \in \mathbb{R}$, $x \cdot 0 = x \cdot (0 + 0)$. The distributive axiom says that
$$x \cdot 0 + 0 = x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0$$
By the previous result, we have that $x \cdot 0 = 0$.

3. Supposing that $x < y$, we add the quantity $(-x) + (-y)$ to both sides by Definition 3.2.1: $x + ((-x) + (-y)) < y + ((-x) + (-y))$. Reassociating on the left side gives $x+((-x)+(-y)) = (x+(-x))+(-y) = 0+(-y) = -y$. Commuting and reassociating on the right sides gives $y + ((-x) + (-y)) = y + ((-y) + (-x)) = (y + (-y)) + (-x) = 0 + (-x) = -x$. Finally, it follows that $-y < -x$.

4. We prove that $x+(-1)\cdot x = 0$. This would imply that $(-1)\cdot x = -x$, the unique additive inverse of x. By the distributive axiom, we have that $x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1+(-1)) \cdot x = 0 \cdot x = 0$. The result follows. We make a special note that the additive inverse of 0 is $-0 = (-1) \cdot 0 = 0$.

5. As an application of (4), we know that $(-1)\cdot(-1) = -(-1)$. Further, since $-1+1 = 0$, it follows that the additive inverse of -1 is $-(-1) = 1$. Therefore, $(-1) \cdot (-1) = -(-1) = 1$. Essentially, two negative signs multiplied together cancel out.

6. If $x = 0$, then $x^2 = 0 \cdot 0 = 0$. If $x > 0$, then by Definition 3.2.2, $0 = 0 \cdot x < x^2$. Finally, if $x < 0$, then by result (3) we have $0 < -x$ and by Definition 3.2.2 once again, $-x^2 = (-1) \cdot x \cdot x = x \cdot (-x) < 0 \cdot (-x) = 0$. By result (3), $0 < x^2$.

7. We know from (6) that $1 \cdot 1 = 1 \geq 0$. We suppose by contradiction that $0 = 1$. Then for any nonzero $x \in \mathbb{R}$ (we assume that not all real numbers are 0), we would have that $0 = x \cdot 0 = x \cdot 1 = x$, a contradiction to x being nonzero. Therefore, 1 must be strictly larger than 0, $0 < 1$.

8. We show that if $x > 0$, then $\frac{1}{x} > 0$. Suppose by contradiction that $\frac{1}{x} \leq 0$. If $\frac{1}{x} = 0$, then $1 = x \cdot \frac{1}{x} = x \cdot 0 = 0$, contradicting result (7). If $\frac{1}{x} < 0$, then $0 < -\frac{1}{x}$ and $-1 = -\frac{1}{x} \cdot x > 0$, or rather, $1 < 0$, further contradicting (7). Thus, $\frac{1}{x}$ is positive if x is. Now to the property at hand. We multiply the inequality $0 < x < y$ by the number $\frac{1}{x} \cdot \frac{1}{y} > 0$. It follows that $0 \cdot \frac{1}{x} \cdot \frac{1}{y} < x \cdot \frac{1}{x} \cdot \frac{1}{y} < y \cdot \frac{1}{x} \cdot \frac{1}{y}$, or rather, that $0 < \frac{1}{y} < \frac{1}{x}$.

$\square$

Many of these results seem tedious, and they are, but proving them once if only to forget about the axiomatic mechanics behind the scenes is beneficial. Finally, we need a way to talk about the 'distance' between two real numbers x and y. For example, how 'far away' are 0 and 1? How 'close' are $-\frac{1}{4}$ and $-\frac{1}{3}$ (see Figure 1)? Regardless, we want our 'distance' to be either positive or 0 (negative distances do not make sense). Further, we expect that the only point 0 distance away from $x \in \mathbb{R}$ should be x itself. There are other properties that we like a distance between real numbers to satisfy, and they are all accomplished with the following definition of absolute value.

We define the absolute value of a real number x as follows.

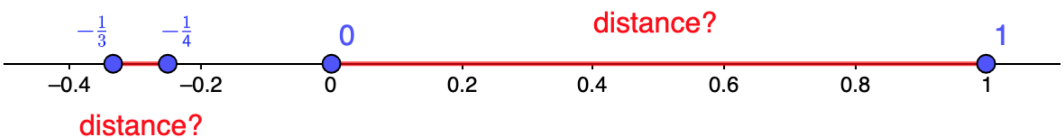


Figure 1: Distances

Definition 3.4: The Absolute Value

The absolute value of a real number x is given by the function

$$|x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases}.$$

Figure 2 shows a graph of $|x|$.

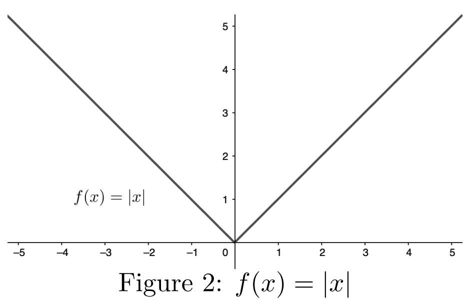


Figure 2: $f(x) = |x|$

The absolute value lends itself to computing distance between real numbers: the distance between $x \in \mathbb{R}$ and $y \in \mathbb{R}$ is $|x - y|$. For example, 0 and 1 are $|0 - 1| = |-1| = 1$ unit apart while $-\frac{1}{4}$ and $-\frac{1}{3}$ are $\left|-\frac{1}{4}-(-\frac{1}{3})\right| = \left|\frac{1}{12}\right| = \frac{1}{12}$ units apart. We now prove various properties of $|x|$ that will be extremely useful throughout MAT 370.

Theorem 3.5: Absolute Value Properties

For all $x, y \in \mathbb{R}$ the following hold:

1. $|x| = 0$ if and only if $x = 0$.
2. $|x| = |-x|$.
3. $|xy| = |x|\,|y|$.
4. For $M \geq 0$, $|x| \leq M$ if and only if $-M \leq x \leq M$.
5. $|x \pm y| \leq |x| + |y|$
6. $||x| - |y|| \leq |x \pm y|$.

Proof.

1. If $x = 0$, then by the definition of $|x|$, $|x| = 0$. For the converse, we argue by contraposition. If $x < 0$, $|x| = -x > 0$. If $x > 0$, $|x| = x > 0$. Thus, $x \neq 0$ implies $|x| \neq 0$.

2. If $x = 0$. $-x = 0$ and thus $0 = |x| = |-x| = 0$. If $x < 0$, $-x > 0$ and thus $|x| = -x = |-x|$. Similarly, if $x > 0$, $-x < 0$ and $|x| = x = -(-x) = |-x|$.

3. We check cases here as well. If $xy = 0$ then one of x or y equals 0. Thus, one of $|x|$ or $|y|$ equals 0 which implies that $|xy| = 0 = |x|\,|y|$. If $xy > 0$, x and y are both positive or both negative. In the first case, $|xy| = xy = |x|\,|y|$. In the second case, $|xy| = xy = (-x)(-y) = |x|\,|y|$. If $xy < 0$, then x and y have opposite signs. If $x < 0$ and $y > 0$ then $|xy| = -xy = (-x)y = |x|\,|y|$. The same holds for $x > 0$ and $y < 0$. All cases of $xy \in \mathbb{R}$ hold.

4. Let $M \geq 0$. Suppose that $|x| \leq M$. If $x \geq 0$, then $-M \leq 0 \leq x = |x| \leq M$. If $x < 0$, then $0 < -x$ and $-M \leq 0 < -x = |x| \leq M$. The forward direction holds in all cases. Suppose now that $-M \leq x \leq M$. The second inequality tells us that $x \leq M$, of course, and the first inequality tells us that $-x \leq M$. Thus, in all cases of the definition of $|x|$ we must have that $|x| \leq M$.

5. We apply result (4) with $M = |x| \geq 0$. Since $|x| \leq |x|$, it follows that $-|x| \leq x \leq |x|$. The same holds for y: $-|y| \leq y \leq |y|$. Adding these inequalities, we have that $-(|x| + |y|) \leq x + y \leq |x| + |y|$. By result (4), it follows that $|x + y| \leq |x| + |y|$. For $|x - y| \leq |x| + |y|$, we note that $-|y| \leq -y \leq |y|$ as well and we go through the same process.

6. The trick here is to add a 0 and apply result (5). We know that

$$|x| = |x + y - y| \leq |x + y| + |-y| = |x + y| + |y|\,. \tag{1}$$

This implies that $|x| - |y| \leq |x + y|$. We now do the same calculation but switch x and y:

$$|y| = |x + y - x| \leq |x + y| + |-x| = |x + y| + |x|\,, \tag{2}$$

which means that $|y| - |x| \leq |x + y|$ or rather, that $-|x + y| \leq |x| - |y|$. By result (4), $-|x + y| \leq |x| - |y| \leq |x + y|$ implies $||x| - |y|| \leq |x + y|$. For $||x| - |y|| \leq |x - y|$, we simply note that in equations (1) and (2), we could have written $|x + y - y| \leq |x - y| + |y|$ and $|x + y - x| \leq |x - y| + |x|$.

$\square$

These properties are all essential of a 'distance function' like $|x|$. Results (5) and (6) are often referred to as the Triangle Inequality and the Reverse Triangle Inequality, respectively. Geometrically, in the plane $\mathbb{R}^2$, the triangle inequality can be seen as "the length of one side of a triangle cannot exceed the sum of the lengths of the other two sides". Here, the reverse triangle inequality is of course the fact that "the length of one side cannot be smaller than the difference of the other two lengths" (see Figures 3 and 4).

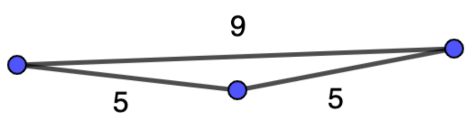


Figure 3: The Triangle Inequality

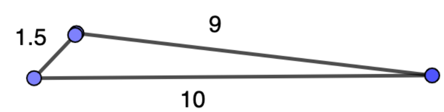


Figure 4: The Reverse Triangle Inequality

3.1 The Completeness Axiom

What makes $\mathbb{R}$ truly special in comparison to $\mathbb{N}$, $\mathbb{Z}$, or even $\mathbb{Q}$ is that it is 'complete'. In other words, $\mathbb{R}$ does not have any 'gaps' between its elements. These gaps in $\mathbb{N}$ and $\mathbb{Z}$ are quite obvious: think about all the numbers in between 1 and 2! However, even $\mathbb{Q}$ has its gaps that are detrimental enough to require a 'larger' space like $\mathbb{R}$ to work over. In order to begin to talk about the completeness of $\mathbb{R}$, we introduce the notion of bounded subsets of $\mathbb{R}$.

Definition 3.6: Boundedness

Let $E \subseteq \mathbb{R}$. E is said to be **bounded above** if and only if there exists an $M \in \mathbb{R}$ such that $x \leq M$ for all $x \in E$. We call M an **upper bound** of E.
This subset E is said to be **bounded below** if and only if there exists an $M' \in \mathbb{R}$ such that $M' \leq x$ for all $x \in E$. We call M' a **lower bound** of E.
We say that E is **bounded** if and only if E is bounded above and bounded below.
An **unbounded** set E is a set that is not bounded.

Since bounded sets are so important and this definition is based off of other definitions, we give some equivalent characterizations of boundedness.

Lemma 3.7: Boundedness Characterizations

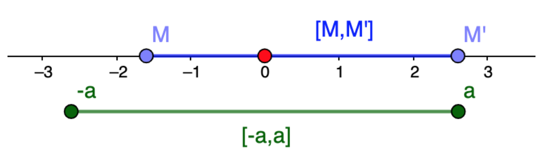
Let $E \subseteq \mathbb{R}$. The following are equivalent:

1. E is bounded.
2. There exist $M, M' \in \mathbb{R}$ such that $E \subseteq [M, M']$.
3. There exists $M \in \mathbb{R}$ such that $|x| \leq M$ for all $x \in E$.

Proof. That (1) and (2) are equivalent follows from the definition of bounded and that $[M, M'] = \{x \in \mathbb{R} : M \leq x \leq M'\}$.

For the equivalence of (2) and (3), we start with the converse. Suppose that there exists $M \in \mathbb{R}$ that $|x| \leq M$ for all $x \in E$. By Theorem 3.5.4, we have that for all $x \in E$, $-M \leq x \leq M$, or rather, $E \subseteq [-M, M]$. For the forward direction, we suppose that there exist $M, M' \in \mathbb{R}$ such that $E \subseteq [M, M']$. We need one real number a such that $[-a, a]$ contains the entirety of $[M, M']$. So, we pick whichever value of M and M' is furthest from 0, $a := \max\{|M|, |M'|\}$ (see Figure 5). Thus, for all $x \in E$, we have that $x \in [M, M'] \subseteq [-a, a]$. In other words, $-a \leq x \leq a$, or rather, $|x| \leq a$, by Theorem 3.5.4. $\square$

$a = \max\{|M|, |M'|\} = |M'|$



$[M, M'] \subseteq [-a, a]$

Figure 5: Constructing $[-a, a]$

This trick of picking a maximum or minimum of finitely many values is used frequently in MAT 370, so get used to it! We look at some quick examples of bounded and unbounded sets:

Example 3.8: Bounded and Unbounded Sets

1. Any interval of real numbers (a, b), $(a, b]$, $[a, b)$ $[a, b]$ is bounded.
2. The natural numbers are bounded below but not bounded. For example, $n \geq 1$ for all $n \in \mathbb{N}$. We will examine the unboundedness of $\mathbb{N}$ later.
3. The set $E = \{q \in \mathbb{Q} : q^2 < 2\}$ is bounded since $E \subseteq [-2, 2]$. It also holds that $E \subseteq [-1, 1.5]$ and even $E \subseteq [0, \sqrt{2}]$.

From the last part of the previous example, we see that a bounded set E can be contained in many such sets. This begs the question: is there a 'smallest' set that bounds E? Is there a smallest number M such that $|x| \leq M$ for all $x \in E$? Similarly, is there a greatest lower bound for E? We give a definition for such special upper and lower bounds of a subset of $\mathbb{R}$.

Definition 3.9: Suprema and Infima

Let $E \subseteq \mathbb{R}$. We define the **supremum** of E, notated $\sup(E)$, as the real number α that satisfies

i. α is an upper bound of E

ii. α is the least upper bound of E: for any other upper bound M of E, $\alpha \leq M$.

Similarly, the **infimum** of E, denoted $\inf(E)$ is the real number β that satisfies

i. β is a lower bound of E

ii. β is the greatest lower bound of E: for any other lower bound M' of E, $M' \leq \beta$.

Note that throughout the definition we have used the article 'the' when talking about these suprema and infima. We confirm that these special values of a set E are indeed unique:

Theorem 3.10: Uniqueness of $\sup(E)$ and $\inf(E)$

Let $E \subseteq R$, $\alpha = \sup(E)$, and $\beta = \inf(E)$. Then α and β are both unique.

Proof. Suppose, for the case of the supremum, that α_1 and α_2 are both suprema of E. By definition, they are both upper bounds of E and both the least upper bound of E. Thus,

$\alpha_1 \leq \alpha_2$ and $\alpha_2 \leq \alpha_1$, which implies $\alpha_1 = \alpha_2$. The same argument can be given for infima. $\square$

Beyond just being unique, these suprema and infima behave nicely with sets that we know have maxima and minima. Nowhere in the definitions of suprema and infima does it say that they have to be elements of the set E themselves. However, if E has a maximum or minimum, then the supremum and infimum fall in the set:

Theorem 3.11: Maxima and Minima are Suprema and Infima

If $E \subseteq R$ has a maximum or minimum, then $\sup(E) = \max(E)$ and $\inf(E) = \min(E)$.

Proof. For the supremum case, we check that $\max(E)$ satisfies the definition of supremum. $\max(E) \geq x$ for all $x \in E$ by definition, so it is an upper bound. Let M be any upper bound of E. Since $max(E) \in E$, then $\max(E) \leq M$ and thus $\max(E) = \sup(E)$. The same argument can be carried out for infima. $\square$

We give one final tool for working with suprema and infima, an equivalence that defines suprema and infima from the 'inside' of a set. It might be easier to prove things about suprema and infima this way:

Theorem 3.12: Suprema and Infima Equivalences

Let $E \subseteq \mathbb{R}$. Then $\alpha = \sup(E)$ if and only if

i. α is an upper bound of E

ii. If $x < \alpha$, then there exists $y \in E$ such that $x < y$.

Similarly, $\beta = \inf(E)$ if and only if

i. β is a lower bound of E

ii. If $x > \beta$, there exists a $y \in E$ such that $y < x$.

Proof. To prove the equivalences, we only check (ii) for each, since (i) is the same everywhere. For the supremum case, If α is the supremum, it follows that if x is an upper bound of E, then $\alpha \leq x$. We examine the contrapositive of this statement. Equivalently, this means that if $x < \alpha$, then M is not an upper bound of E. Thus, if $x < \alpha$, there exists an $y \in E$ such that $x < y$, what we wanted to show. Sorry for the abuse of notation from what we have been using to denote upper bounds, but this makes the equivalence a bit clearer to see. Again, a symmetric argument can be given for infima. $\square$

Now is the time for the Completeness Axiom, which has to do with boundedness and suprema of subsets of real numbers. This is what makes $\mathbb{R}$ so special in comparison to some of the subsets we have mentioned previously, and we take it to be an axiom, something we hold to be intrinsically true about the real number system:

Definition 3.13: The Completeness Axiom

Every nonempty, bounded above subset of $\mathbb{R}$ has a supremum.

Let's unpack the power of this statement. If we know that a subset E of the real numbers is bounded above by some number M, no matter how 'far away' from E it is, we are guaranteed the existence of another real upper bound that is as 'close' as we want it to be to E. Let's look at some examples of this:

Example 3.14: The Completeness Axiom, examples

1. $\sup((0,1)) = 1$.
2. $\sup(\{x \in \mathbb{R} : 0 < x < 5 \text{ and } \cos(x) = 0\}) = \frac{3\pi}{2}$.
3. $\sup(\{\frac{1}{n} : n \in \mathbb{N}\}) = 1$.

Proof. All of these sets are nonempty and bounded above (with upper bounds 1, 5, and 1, respectively), so we are guaranteed the existence of their suprema.

1. We check that 1 satisfies condition (ii) given in Theorem 3.12. Suppose that $x < 1$. It follows that the average of these two numbers, $\frac{x+1}{2}$ is strictly between them: $x < \frac{x+1}{2} < 1$. Thus, this $\frac{x+1}{2}$ is our element in $(0,1)$ that satisfies condition (ii) and hence $\sup((0,1)) = 1$.
2. Since $\cos(x) = 0$ if and only if x is a multiple of $\frac{\pi}{2}$, then the set whose seprumum we want to find is actually $\{\frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$, a finite set. Thus, its supremum is the maximum value: $\frac{3\pi}{2}$.
3. Again, this is not too bad for the supremum case, but this set looks like $\{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$. Thus, 1 is the maximum value of this set and hence its supremum as well.

$\square$

There are various applications of the Completeness Axiom that we can discuss now to be able to work with infima, natural numbers, and rational numbers more efficiently. However, in general, and later in the course, the Completeness Axiom will be the anchor to which many 'deep theorems' of Real Analysis are linked. Let's begin with infima:

Theorem 3.15: The Completeness Axiom (for infima)

Every nonempty, bounded below subset of $\mathbb{R}$ has an infimum.

Proof.

Let E be a nonempty set that is bounded below, say by some $M' \in \mathbb{R}$ (i.e. $M' \leq x$ for all $x \in E$). We need to get our hands on a set that is bounded **above** to be able to apply the Completeness Axiom. We consider the set $-E = \{-x : x \in E\}$ (see Figure 6; by **no means** does E need to be an interval of any kind, but it is good for illustration).

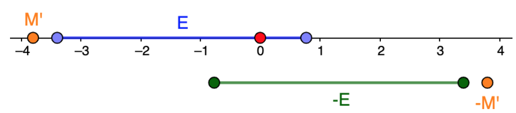


Figure 6: E and $-E$

Since $M' \leq x$ for all $x \in E$, then $-x \leq -M'$ for all $x \in E$, or rather $-M'$ is an upper bound for all elements of $-E$. So, by the Completeness Axiom, $\sup(-E)$ exists and we call it α. We check now that $-\alpha$ satisfies the conditions to be the infimum of E. Since α is an upper bound of $-E$, $-\alpha$ is a lower bound of E, using the same technique as with M' and $-M'$. Now, for any other lower bound M of E, we need to show that $E \leq -\alpha$. Well, since M is a lower bound of E, $-M$ is an upper bound of $-E$ and by definition of supremum, $\alpha \leq -M$. Negating this inequality gives our desired $M \leq -\alpha$. It follows that E has an infimum, namely, $-\alpha$. $\square$

The result of this proof can be summarized nicely using this definition of the negation of a set, $-E$: $-\inf(E) = \sup(-E)$ and $-\sup(E) = \inf(-E)$. In a way, negating suprema and infima equates to flipping the suprema and infima and 'negating' the set itself. We now look at a critical example that handles the relationship between set inclusion, suprema and infima. It is a good example of applying the definitions of suprema and infima.

Example 3.16: Suprema, Infima, and Subsets

Suppose that A and B are bounded subsets of $\mathbb{R}$ such that $A \subseteq B$. Then

$$\inf(B) \leq \inf(A) \leq \sup(A) \leq \sup(B).$$

Proof. Since both A and B are bounded, we know that the numbers $\inf(A)$, $\sup(A)$, $\inf(B)$, and $\sup(B)$ all exist by the Completeness Axiom and Theorem 3.15. Since $\inf(A)$ is the greatest lower bound of A it suffices to show that $\inf(B)$ is a lower bound of A. Similarly, it suffices to show that $\sup(B)$ is an upper bound for A. Let $a \in A$. Then since $A \subseteq B$, $a \in B$ and thus $\inf(B) \leq a \leq \sup(B)$. Thus $\inf(B)$ and $\sup(B)$ are indeed lower and upper bounds for A, respectively. Therefore, by the definition of the supremum and infimum of A, $\inf(B) \leq \inf(A)$ and $\sup(A) \leq \sup(B)$. For the final, middle inequality we know by definition that for all $a \in A$, $\inf(A) \leq a \leq \sup(A)$, so the whole chain of inequalities is satisfied. $\square$

The natural numbers will be used frequently in MAT 370 for approximations, which we will get to, so we need the following essential result, as an application of the Completeness Axiom.

Theorem 3.17: The Archimedean Property

The natural numbers are not bounded above. For all $x \in \mathbb{R}$, there exists $N \in \mathbb{N}$ such that $x < N$. Also, for all $x > 0$, there exists $N \in \mathbb{N}$ such that $\frac{1}{N} < x$.

Proof. We suppose, by contradiction, that $\mathbb{N}$ is a bounded above subset of $\mathbb{R}$. Thus, $\sup(\mathbb{N})$ exists, and we call it $M \in \mathbb{R}$. Thus, for all $n \in \mathbb{N}$ $n \leq M$. From condition (ii) of Theorem 3.12, we know that $M - 1 < M$ and thus there exists a $k \in \mathbb{N}$ such that $M - 1 < k$. But this implies that $k + 1 \in \mathbb{N}$ and $M < k + 1$, contradicting the fact that M is an upper bound. $\mathbb{N}$ must not be bounded above.

The second claim in the statement of the theorem follows immediately from negating the definition of bounded above. For the third, suppose that $x > 0$ so that $0 < \frac{1}{x} \in \mathbb{R}$. Then, there exists $N \in \mathbb{N}$ such that $\frac{1}{x} < N$, or rather, $\frac{1}{N} < x$. $\square$

With our tools we have so far, we return to Example 3.14 to compute the infima of the given sets.

Example 3.18: The Completeness Axiom, examples

1. $\inf((0,1)) = 0$.
2. $\inf(\{x \in \mathbb{R} : 0 < x < 5 \text{ and } \cos(x) = 0\}) = \frac{\pi}{2}$.
3. $\inf(\{\frac{1}{n} : n \in \mathbb{N}\}) = 0$.

Proof.

1. As above, we know that 0 is a lower bound for $(0,1)$. To prove condition (ii) suppose that $x > 0$. Taking the average analogously, $0 < \frac{x}{2} < x$ and we have our element in the set that is less than x as well: $\frac{x}{2}$.
2. Given our finite set in the previous example, the set has a minimum of $\frac{\pi}{2}$ and thus a infimum of $\frac{\pi}{2}$.

3. We know that every natural number satisfies $n > 0$ which implies $\frac{1}{n} > 0$. Hence 0 is a lower bound for this set. Now, let $x > 0$. By Theorem 3.17, there exists $N \in \mathbb{N}$ such that $\frac{1}{N} < x$. Thus, this $\frac{1}{N}$ satisfies condition (ii) of Theorem 3.12.

□

Two famous corollaries of The Archimedean Property of $\mathbb{R}$ show how the rationals and irrationals are intertwined respectively and as subsets of $\mathbb{R}$. They give us a method of approximate real numbers by rationals.

Corollary 3.19: Density of the Rationals

For any two real numbers $x < y$, there exists a $q \in \mathbb{Q}$ such that $x < q < y$.

Proof. Let x and y be two distinct real numbers such that $x < y$. We break into cases based on where x and y are located

Case 1: ($x < 0 < y$). The rational number that we need for the statement of theorem is already given: 0. We are done with this case.

Case 2: ($0 < x < y$). Here both x and y are strictly positive. We consider the positive value $y - x > 0$ and, from the Archimedean Property, there exists an $N \in \mathbb{N}$ such that $\frac{1}{N} < y - x$. This $\frac{1}{N}$ is going to be our 'step size' to fall into the interval (x, y) (see Figure 7).

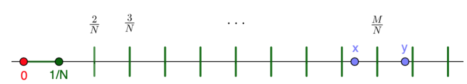


Figure 7: Steps to fall inside (x, y)

As for the number of steps it takes to get to (x, y), we consider the set $\{\frac{m}{N} : m \in \mathbb{N}\}$ of multiples of our step size $\frac{1}{N}$ and its further subset $S = \{\frac{m}{N} : m \in \mathbb{N} \text{ and } x < \frac{m}{N} < y\}$. Are we guaranteed that S is nonempty, that one of the steps falls into (x, y)? Well, if none of them did, we must have that two adjacent steps were farther apart than the distance between y and x, $|y - x| = y - x$. Thus, for some $N_0 \in \mathbb{N}$ we must have that $\frac{N_0+1}{N} - \frac{N_0}{N} \geq y - x$. This is a contradiction as $\frac{1}{N} = \frac{N_0+1}{N} - \frac{N_0}{N} \geq y - x > \frac{1}{N}$. Thus, by our choice of $\frac{1}{N}$, we are guaranteed the existence of at least one step in S. We pick $M = \min(S)$ since every nonempty subset of $\mathbb{N}$ has a least element. Finally, it follows that $\frac{M}{N} \in \mathbb{Q}$ and $x < \frac{M}{N} < y$.

Case 3: ($x < y < 0$) Here, x and y are negative and we multiply the whole inequality by -1: $0 < -y < -x$. Now, Case 2 can be applied in a similar way to get $\frac{M}{N} \in \mathbb{Q}$ such that $-y < \frac{M}{N} < -x$, which implies $x < -\frac{M}{N} < y$.

□

Corollary 3.20: Density of the Irrationals

For any two rational numbers $r_1 < r_2$, there exists a $\xi \in \mathbb{Q}^c$ such that $r_1 < \xi < r_2$.

Proof. We know $\sqrt{2}$ to be irrational and $q\sqrt{2}$ to be irrational for any $q \in \mathbb{Q}$. Thus, for $\frac{r_1}{\sqrt{2}}, \frac{r_2}{\sqrt{2}} \in \mathbb{R}$ satisfying $\frac{r_1}{\sqrt{2}} < \frac{r_2}{\sqrt{2}}$, we know there exists a $q \in \mathbb{Q}$ such that $\frac{r_1}{\sqrt{2}} < q < \frac{r_2}{\sqrt{2}}$, or rather, $r_1 < q\sqrt{2} < r_2$. This $q\sqrt{2}$ is our desired $\xi \in \mathbb{Q}^c$. $\square$

Before studying sequences of real numbers in Chapter 5, we take a quick look at the (standard) topology on $\mathbb{R}$, which will lead to visual, less technical interpretations of some of the results of Chapter 5.

3.2 Exercises

1. Show that if $0 < x < y$, then $0 < x^n < y^n$ for all $n \in \mathbb{N}$. Deduce a similar rule for negative numbers $x < y < 0$.

2. For $a \in \mathbb{R}$ and $b \neq 0$, show that
$$\left|\frac{a}{b}\right| = \frac{|a|}{|b|}.$$

3. Prove that for any two distinct $x, y \in \mathbb{R}$ there exists an infinity of real numbers between x and y.

4. Suppose that E is a nonempty subset of $\mathbb{R}$. Show that if $x \in \mathbb{R}$ is not an upper bound of E, then there exists $y \in E$ such that $y > x$ and y is also not an upper bound of E.

5. Suppose that $\mathbb{R}$ is partitioned into two disjoint, nonempty sets L and U, i.e. $\mathbb{R} = L \cup U$ with $L \cap U = \emptyset$, such that every element in L is less than or equal to every element of U. Show that there exists a unique element $x \in \mathbb{R}$ such that $l \leq x \leq u$ for all $l \in L$ and $u \in U$.

6. Suppose A and B are bounded, nonempty subsets of $\mathbb{R}$. We define $A + B := \{a + b : a \in a, b \in B\}$. Show that $\sup(A + B)$ exists and that
$$\sup(A + B) = \sup(A) + \sup(B).$$
Prove the analogous result relating for infima.

7. For A and B nonempty bounded subsets of positive real numbers, define $AB = \{ab : a \in A, b \in B\}$. Show that $\sup(AB)$ and $\inf(AB)$ exists and that
$$\sup(AB) = \sup(A)\sup(B).$$

and

$$\inf(AB) = \inf(A)\inf(B)$$

Can anything be said if A or B contain negative real numbers or 0?

4 Basic Topology on $\mathbb{R}$

Topology is the study of the structure of certain 'open' subsets of a given set X. For our purposes, we are only concerned with the set $\mathbb{R}$. This section will mainly focus on the 'standard' topology on $\mathbb{R}$ and offer definitions and examples to provide a more visual and somewhat less technical treatment of results to come. We want the following properties of open sets to be true:

1. $\mathbb{R}$ and $\emptyset$ are 'open'
2. Any union (infinite or not) of 'open' sets is 'open'.
3. Any finite intersection of 'open' sets is 'open'.

With our set $\mathbb{R}$, we start with a collection of subsets of $\mathbb{R}$ that will behave nicely with these notions of 'open' sets we will use in subsequent sections.

Definition 4.1: Open Intervals in $\mathbb{R}$

For $a, b \in \mathbb{R}$, with $a < b$, intervals of the form $(a, b) : \{x \in \mathbb{R} : a < x < b\}$ is called an **open interval** of $\mathbb{R}$.

These open intervals will be the building blocks of the more general open sets in $\mathbb{R}$. For any set $U \subseteq \mathbb{R}$, we give the following definition to check when it is open:

Definition 4.2: Open Sets in $\mathbb{R}$

A subset U of $\mathbb{R}$ is said to be **open** if and only if for every $x \in U$, there exists an open interval (a, b) such that $x \in (a, b)$ and $(a, b) \subseteq U$.

Essentially what this definition is saying is that a set U is open in $\mathbb{R}$ is it can be approximated from the inside by these 'building block' sets, the open intervals. To give a more visual explanation of this definition, we may consider an analogous result for $\mathbb{R}^2$, the real plane. Any set U in the plane is considered open if it can be approximated from the inside by the 'building block' sets that would be open circles in this case (see Figure 8). Every point in U can be contained in a circle that is further contained in U. For our purpose, the real line

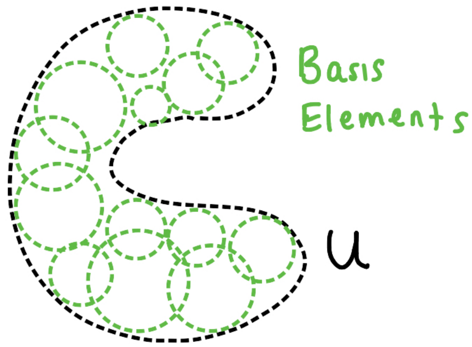


Figure 8: Approximating a general open set U by open circles

does not have the same geometric flare of the real plane but the example of Figure 8 works the same for open sets in $\mathbb{R}$.

With all this talk of open sets, it begs the question: is there such a notion of closed sets? We give the following definition:

Definition 4.3: Closed Sets of $\mathbb{R}$

A subset E of real numbers is closed if and only if its complement E^c is open, as defined in Definition 4.2.

This definition of closed subsets is not entirely elementary. Being closed is not the opposite of being open. In fact, in other topologies on $\mathbb{R}$, i.e. not the one we are concerned with given by open intervals, there exists subsets that are **both** open and closed, and are called **clopen** sets. There are, however, subsets of $\mathbb{R}$ that are **neither** open **nor** closed in the topology given by open intervals. Lets look at some examples of these different types of sets in $\mathbb{R}$.

Example 4.4: Closed Intervals in $\mathbb{R}$

Every **closed** interval of the form $[a, b]$ for $a, b \in \mathbb{R}$ and $a < b$ is in fact closed, as we expect.

Proof. We show that the complement of $[a, b]$ is open. This complement satisfies

$$[a, b]^c = (-\infty, a) \cup (b, \infty)$$

We want to show that both $(-\infty, a)$ and (b, ∞) are open. For the first set, we pick a natural number N such that $-a < N$, which gives $-N < a$. For all naturals $m \geq N$, we have that

$$(-\infty, a) = \bigcup_{m \geq N} (-m, a).$$

See Figure 9 for a visual of $[a, b]^c$ and the reasoning behind $(-\infty, a)$ being open. For the second set, we go through a similar procedure: pick an $N' \in \mathbb{N}$ such that $b < N'$. Thus,

$$(b, \infty) = \bigcup_{m \geq N'} (b, m)$$

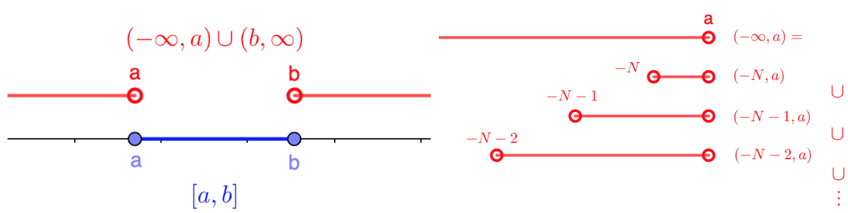


Figure 9: $[a, b]^c$ and showing $(-\infty, a)$ is open

Finally, it follows that

$$[a, b]^c = (-\infty, a) \cup (b, \infty) = \left(\bigcup_{m \geq N} (-m, a) \right) \cup \left(\bigcup_{m \geq N'} (b, m) \right),$$

so that the complement is the union of open intervals and hence open by our properties that open sets follow. By definition, it follows that $[a, b]$ is closed. $\square$

Example 4.5: Subsets neither Open nor Closed in $\mathbb{R}$

Intervals of the form

$$\{(a, b] : a, b \in \mathbb{R}, a < b\}$$

are neither open nor closed in $\mathbb{R}$.

Proof. Here is some good practice in negating the definition of being open to show that sets of the form $(a, b]$ are not: A set U is not open in the standard topology if and only if there exists an $x_0 \in U$ such that for all open intervals (c, d), $x \notin (c, d)$ or $(c, d) \not\subseteq U$. To prove a statement like this, we negate one of the conclusions and assume it as a hypothesis. So, we need to find a problematic $x_0 \in (a, b]$ such that for every open interval (c, d) with $x_0 \in (c, d)$, it follows that $(c, d) \not\subseteq (a, b]$. To prove this, set $x_0 = b$ and let (c, d) be any open set containing b. Therefore, $b < d$ and it follows that $\frac{b+d}{2} \in (c, d)$ but not in $(a, b]$. It follows that $(c,d) \nsubseteq (a,b]$. Thus $(a,b]$ is not open. To show that $(a,b]$ is not closed, it suffices to show that its complement

$$(a,b]^c = (-\infty,a] \cup (b,\infty)$$

is not open. The set on the right here is open, as above. However, the set $(-\infty,a]$ is not open with the problematic point a. Therefore, $(a,b]$ is neither open nor closed. A similar argument can be made for sets of the form $[a,b)$. $\square$

For the final definitions in this section that will assist us later on, we introduce the notion of neighborhoods and limit points:

Definition 4.6: Neighborhoods of Real Numbers

For a given $x \in \mathbb{R}$, a subset N of $\mathbb{R}$ is said to be a **neighborhood** of x if and only if there there exists an $\epsilon > 0$ such that the open interval centered at x satisfies $(x-\epsilon, x+\epsilon) \subseteq N$.

Definition 4.7: Limit Points of a subset E

Suppose $E \subseteq \mathbb{R}$. A real number $x \in \mathbb{R}$ is a **limit point** of E if and only if every neighborhood of x contains infinitely many points of E. We will define the set of all limit points of a set E as

$$E' := \{x \in \mathbb{R} : x \text{ is a limit point of } E\}.$$

It is **crucial** to notice here that limit points of E need not be elements of E.

It is clear that these two definitions are intertwined. Let's discuss the subtleties concerning neighborhoods. At first glance, neighborhoods of some $x \in \mathbb{R}$ seem to be the same as the open sets containing x. However, neighborhoods of x are a bit more general. Clearly every open set containing x is a neighborhood of x: If U is open and $x \in U$, then there exists some open interval (a,b) such that $x \in (a,b) \subseteq U$. As in Lemma 3.7 and Figure 5, we use a similar strategy to set $\epsilon = \min\{x-a, b-x\} > 0$. Thus, $x \in (x-\epsilon, x+\epsilon) \subseteq (a,b) \subseteq U$. and it follows that U is a neighborhood of x. However, we note that for the point $\frac{1}{2}$, the closed interval $[0,1]$ is still a neighborhood of $\frac{1}{2}$: $(\frac{1}{2}-\frac{1}{4}, \frac{1}{2}+\frac{1}{4}) = (\frac{1}{4}, \frac{3}{4}) \subseteq [0,1]$. However, $[0,1]$ is not a neighborhood of 0 or 1. Any open interval (a,b) of 0 or 1 will spill out of $[0,1]$ (i.e. is not contained in $[0,1]$). Further note that by definition even if our neighborhood of x is not open, we can always find an open set that we can consider our interval: $(x-\epsilon, x+\epsilon)$.

Therefore, to prove statements about limit points, and for arbitrary neighborhoods of these points, it suffices to consider only the open intervals containing the limit points themselves. Still though, proving something about infinitely many points of a subset E is still cumbersome. We give an alternate characterization for showing a real number is a limit point of a set E:

Theorem 4.8: Alternate Limit Point Characterization

A point $x \in \mathbb{R}$ is a limit point of $E \subseteq \mathbb{R}$ if and only if every neighborhood of x contains a point $y \in E$ such that $y \neq x$.

Proof. For the simpler forward direction, suppose $x \in \mathbb{R}$ is a limit point of E and let N be an arbitrary neighborhood of x. By the definition of a limit point, there exists an infinite number of points of E that are also members of this neighborhood N. Clearly, we may choose some y from this infinite number of points that is not x.

For the converse direction, we suppose that every neighborhood of x contains a point $y \in E$ such that $y \neq x$. Let N be an arbitrary neighborhood of x. We need to find or construct an infinite number of points from E that are also members of this neighborhood N. To begin, we are guaranteed the existence of some $y_1 \neq x$ such that $y_1 \in N \cap E$ by hypothesis. To continue we consider the set $N \setminus \{y_1\}$. Is this set a neighborhood of x?

Well, since N is a neighborhood of x, there exists an $\epsilon > 0$ such that $(x-\epsilon, x+\epsilon) \subseteq N$. If it happens that $y_1 \notin (x-\epsilon, x+\epsilon)$, then $(x-\epsilon, x+\epsilon) \subseteq N \setminus \{y_1\}$ and $N \setminus \{y_1\}$ is a neighborhood of x as well. If $y_1 \in (x-\epsilon, x+\epsilon)$, then using the same technique as above in Lemma 3.7, we may find an $\epsilon' = \min\{\epsilon, |x-y_1|\}$ such that $(x-\epsilon', x+\epsilon')$ does not contain y_1 and thus $(x-\epsilon', x+\epsilon') \subseteq N \setminus \{y_1\}$. In either case $N \setminus \{y_1\}$ is a neighborhood of x. Therefore, we may pick a $y_2 \in E \cap (N \setminus \{y_1\})$ such that $y_2 \neq x$. We are also guaranteed that $y_1 \neq y_2$.

Going through this inductive process to find the y_n, we will not run into problems and are thus guaranteed an infinite list (since each y_n is distinct by construction) of elements of $E \cap N$: $y_1, y_2, y_3, \ldots$ for all $n \in \mathbb{N}$. This completes our equivalence. $\square$

These notions of neighborhoods and limit points will be extremely helpful in our discussion of sequences of real numbers, which are the foundation for real analysis and approximation. If this section seemed a bit abstract now, it will be seen in a different context in the following chapter and especially the chapter on continuity as well.

4.1 Exercises

1. Show that for an arbitrary collection $\{K_\alpha\}$ of closed subsets of $\mathbb{R}$, $\bigcap_\alpha K_\alpha$ is closed. Further, for a finite collection $\{K_i\}_1^n$ of closed subsets of $\mathbb{R}$, show that $\bigcup_1^n K_i$ is closed.

2. Show that a subset E of $\mathbb{R}$ is closed if and only if E contains all of its limit points, i.e. $E' \subseteq E$.

3. Show that open intervals of the form

$$\{(p,q) : p,q \in \mathbb{Q}, p < q\}$$

may be used instead of open intervals with general real endpoints. In other words, we may further approximate general open sets in $\mathbb{R}$ by intervals with **rational** endpoints.

4. Show that every finite subset of $\mathbb{R}$ has no limit points. In other words, if $E \subseteq \mathbb{R}$ is finite, then $E' = \emptyset$. Deduce that every finite subset of $\mathbb{R}$ is closed.

5. Show that a and b are limit points of the open interval (a,b) with $a < b$. Deduce that $(a,b)' = [a,b]$.

6. Prove that $\mathbb{N}' = \emptyset$ and $\mathbb{Z}' = \emptyset$, but that $\mathbb{Q}' = \mathbb{R}$.

7. Let $E \subseteq \mathbb{R}$. We define the **closure** of E, notated $\overline{E}$, to be the intersection of all closed sets that contain E:

$$\overline{E} = \bigcap \{K \subseteq \mathbb{R} : K \text{ is closed and } E \subseteq K\}.$$

In other words, $\overline{E}$ is the 'smallest' closed set containing E.

   a) Show that $\overline{E}$ is indeed closed.
   b) Show that a subset E is closed if and only if $E = \overline{E}$.
   c) Show that $x \in \overline{E}$ if and only if every neighborhood of x contains some element $y \in E$. Deduce that $E' \subseteq \overline{E}$.
   d) Show that $\overline{E} = E \cup E'$. This means that to find the closure of E, you simply need to add to E all of its limit points.
   e) Suppose that E is a bounded, nonempty subset of $\mathbb{R}$. Prove then, that both $\sup(E) \in \overline{E}$ and $\inf(E) \in \overline{E}$.

5 Real-valued Sequences

Much of the results from the standard calculus class, such as continuity, differentiation, and even integration are developed with the use of 'limits'. For example, it might be familiar that a function f is continuous at $x = a$ if $\lim\limits_{x \to a} f(x) = f(a)$ or that f is differentiable at $x = a$ if the limit $\lim\limits_{x \to a} \frac{f(x)-f(a)}{x-a}$ exists. But what does all of this actually mean? What does it mean for x to approach a ($x \to a$)? What does it mean for a limit to exist? The tool for dealing with these limits is sequences, and in our case, sequences of real numbers.

Definition 5.1: Real-valued Sequences

A **sequence** of real numbers is a function f from $\mathbb{N}$ to $\mathbb{R}$. In this way, for each $n \in \mathbb{N}$, $f(n) = x_n$ for $x_n \in \mathbb{R}$.
The common notation we will use, however, gets rid of f altogether and denotes the sequence as

$$(x_n)_{n\in\mathbb{N}} = (x_1, x_2, x_3, \ldots)$$

In general, sequences can be functions from the natural numbers into any set X. We can examine sequences of functions, sequences of matrices, or even sequences of U.S. cities if we want:

$$\begin{aligned}&(f_1, f_2, f_3, \ldots)\\ &\left(\left(\begin{smallmatrix}1 & 4\\ 6 & 2\end{smallmatrix}\right), \left(\begin{smallmatrix}4 & -1\\ 0 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right), \ldots\right)\\ &(\text{Annapolis}, \text{Salt Lake City}, \text{Tempe}, \ldots)\end{aligned}$$

Of course, we will only be focused on sequences of real numbers as stated above. It is important to note in any case that a sequence also assumes the order of the natural numbers, we cannot simply switch around elements of these 'lists'. Unlike sets, sequences are changed when the order of elements is changes.

The primary question of interest for sequences of real numbers is: where do they 'go'? A perhaps familiar result is that for the sequence $(\frac{1}{n})_{n\in\mathbb{N}} = (1, \frac{1}{2}, \frac{1}{3}, \ldots)$,

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

Although $\frac{1}{n}$ is never 0 for any $n \in \mathbb{N}$, it seems that as n gets larger, the terms of the sequence get closer and closer to 0, arbitrarily so. Three important notions arise in the discussion of this specific sequence:

- The $\frac{1}{n}$ elements get '**close**' to 0. The distance from $\frac{1}{n}$ to 0, $\left|\frac{1}{n} - 0\right|$ appears to get small.
- This distance $\left|\frac{1}{n} - 0\right|$ is **arbitrarily** small. For any positive number $r > 0$ (no matter how small) we can find an $N \in \mathbb{N}$ such that $\left|\frac{1}{N} - 0\right| = \frac{1}{N} < r$ using the Archimedean Property (Theorem 3.17).
- This N **depends on** our arbitrary small number r, and not the other way around.

Although we have just considered the sequence $(\frac{1}{n})_{n\in\mathbb{N}}$ 'going' to the number 0, we incorporate these three points in the standard definition of an arbitrary sequence converging to a point in $\mathbb{R}$.

Definition 5.2: Sequence Convergence

A sequence $(x_n)_{n\in\mathbb{N}}$ **converges to** $a \in \mathbb{R}$ if and only if the following is true: for all $\epsilon > 0$, there exists a natural number $N \in \mathbb{N}$ such that for all $n \geq N$,

$$|x_n - a| < \epsilon.$$

We say that $(x_n)_{n\in\mathbb{N}}$ is **convergent** if and only if there exists an $a \in \mathbb{R}$ such that $(x_n)_{n\in\mathbb{N}}$ converges to a. In this case, we say that $\lim_{n\to\infty} x_n = a$ and that a is the **limit** of the sequence $(x_n)_{n\in\mathbb{N}}$.

We emphasize some of the subtleties of this definition:

- In order to talk about convergence of a sequence, we need to know the point or have a guess about the point a to which it converges. If we do not have a guess at its limit, we cannot prove that $(x_n)_{n\in\mathbb{N}}$ converges.
- Do not worry about the variable ϵ. It is only the standard notation for a small, positive real number. We could have used any other letter of symbol if we wanted.
- Again, we need to assume that ϵ can be any positive real number first, and then find some way of getting N to make the absolute value inequality work, not the other way around.
- We used a definite article 'the' when defining a limit of $(x_n)_{n\in\mathbb{N}}$, implying that a is the only limit of the sequence if it is convergent. We prove this claim now.

Theorem 5.3: Unique Limits

If $(x_n)_{n\in\mathbb{N}}$ converges to $a \in \mathbb{R}$, then a is unique.

Proof. For uniqueness proofs in general, we suppose that two real numbers satisfy the given property and show they must be equal. So suppose that $(x_n)_{n\in\mathbb{N}}$ converges to the real numbers a and b. We apply Exercise 1 of this section's exercises to $|a-b|$. Let $\epsilon > 0$ be arbitrary to show that $|a-b| < \epsilon$. Since $\epsilon > 0$, $\frac{\epsilon}{2} > 0$ as well and we use $\frac{\epsilon}{2}$ in Definition 5.2. Then we know there exists a natural number N_a such that $|x_n - a| < \frac{\epsilon}{2}$ for all $n \geq N_a$ and there exists a natural number N_b such that $|x_n - b| < \frac{\epsilon}{2}$ for all $n \geq N_b$. We define

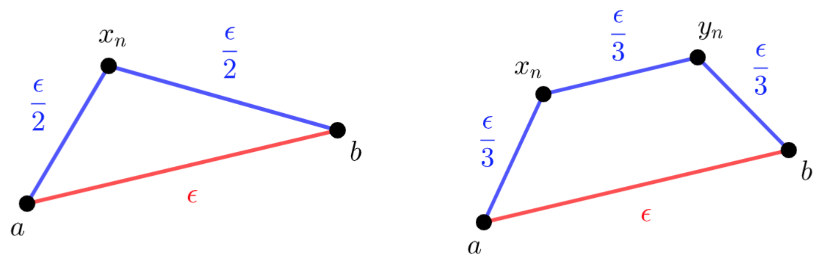


Figure 10: An $\frac{\epsilon}{2}$ or $\frac{\epsilon}{3}$ argument

$N := \max\{N_a, N_b\}$. Then, for all $n \geq N$, it follows that

$$\begin{aligned} |a-b| &= |a - x_n + x_n - b| \\ &\leq |x_n - a| + |x_n - b| \quad \text{(Triangle Inequality)} \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Therefore, since ϵ was arbitrary, by Exercise 1, we have that $a - b = 0$ or rather that $a = b$. It follows that the limit of a sequence is unique. $\square$

For this proof, we used the common '$\frac{\epsilon}{2}$ argument', an indispensable tool in MAT 370. Since we knew that $(x_n)_{n\in\mathbb{N}}$ converged to a and b by hypothesis, we were able to **use** any positive number given Definition 5.2, rather than prove it using an arbitrary positive number. Further, when we gained our natural numbers N_a and N_b guaranteed by Definition 5.2, we cannot be sure they are equal. However, since both conditions depended on all $n \geq N_a$ and all $n \geq N_b$, **both** would be satisfied simultaneously when $n \geq \max\{N_a, N_b\}$ since $\max\{N_a, N_b\}$ is greater than or equal to both N_a and N_b. This is another important trick of MAT 370: choosing a new natural number that is the maximum of a finite number of natural numbers. We will use it and the '$\frac{\epsilon}{2}$ argument' often. Finally, from Theorem 3.5.5, we were able to split $|a - b|$ into two quantities that we knew were less than $\frac{\epsilon}{2}$ and thus their sum would be less than ϵ. This is exactly the '$\frac{\epsilon}{2}$ argument'. We leave a visual in Figure 10 to clarify what is going on and even show that it can be done for 3 or more arbitrarily close points.

To appeal to our work done in the previous chapter, we provide an alternate characterization of convergence that has to do with neighborhoods of the limit point and leads to a more visual understanding of sequence convergence:

Theorem 5.4: Alternate Convergence Characterization

A sequence $(x_n)_{n\in\mathbb{N}}$ converges to $a \in \mathbb{R}$ if and only if for every neighborhood U of a, only finitely many elements of the sequence fall outside U.

Proof. For the forward direction, we suppose that $(x_n)_{n\in\mathbb{N}}$ converges to a as in Definition 5.2 and let U be an arbitrary neighborhood of a. Then we know there exists an $\epsilon > 0$ such that $(a-\epsilon, a+\epsilon) \subseteq U$ from Definition 4.6. Since $\epsilon > 0$, we know from Definition 5.2 that there exists an $N \in \mathbb{N}$ such that for all $n \geq N$, $|x_n - a| < \epsilon$. The key here is noting that $|x_n - a| < \epsilon$ if and only if $a - \epsilon < x_n < a + \epsilon$ if and only if $x_n \in (a-\epsilon, a+\epsilon)$. Therefore, for all $n \geq N$ $x_n \in U$ since $(a-\epsilon, a+\epsilon) \subseteq U$. The finitely many $\{x_1, x_2, \ldots, x_{N-1}\}$ may or may not also lie in U, but we know that, at most, only finitely many will be outside of U. This is exactly what we needed to show for the forward direction.

For the converse direction, we suppose the statement about neighborhoods and let $\epsilon > 0$ be arbitrary as in Definition 5.2. Well, we know that $(a-\epsilon, a+\epsilon)$ is a neighborhood of a since it is an open interval containing a and thus, by hypothesis, there exist a finite number of x_n that are not in $(a-\epsilon, a+\epsilon)$. We call these elements $\{x_{n_1}, x_{n_2}, \ldots, x_{n_k}\}$. We do not know which natural numbers they correspond to, but we know there are a finite number k of them. For Definition 5.2, we need one natural number above which everything works. As above, we choose $N = \max\{n_1, n_2, \ldots, n_k\} + 1$ since there are only finitely many natural numbers to choose from. Thus, since N is the max of all these faulty n_i plus 1, $x_n \in (a-\epsilon, a+\epsilon)$ for all $n \geq N$. As in the forward direction, for all $n \geq N$,

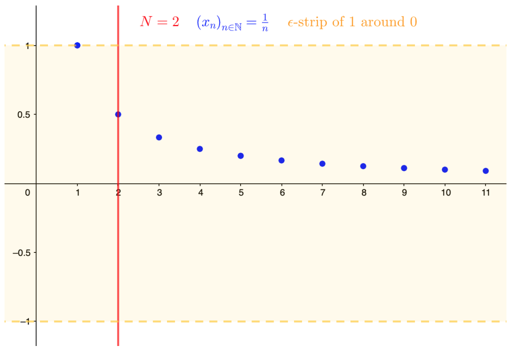
$$x_n \in (a-\epsilon, a+\epsilon), \text{ which implies } |x_n - a| < \epsilon,$$

exactly what we wanted to show for Definition 5.2. $\square$

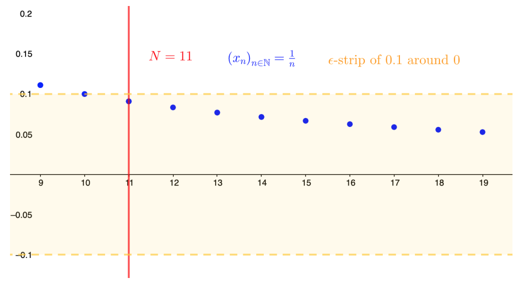
This Theorem gives us a different interpretation of convergence that may be easier to visualize graphically. If the sequence $(x_n)_{n\in\mathbb{N}}$ converges to $a \in \mathbb{R}$, we have that for every $\epsilon > 0$, the neighborhood $(a-\epsilon, a+\epsilon)$ contains all but finitely x_n members of the sequence. As in [6], we call these open interval neighborhoods **ϵ-strips of** a. As we have hinted at, $\lim\limits_{n\to\infty} \frac{1}{n} = 0$ (see Exercise 2), and thus, for any ϵ-strip $(-\epsilon, \epsilon)$ around 0, the strip contains all but finitely many of the $x_n = \frac{1}{n}$. We illustrate this in Figure 11a with $\epsilon = 1$, in Figure 11b with $\epsilon = 0.1$, and in Figure 11c with $\epsilon = 0.03$.

For each given ϵ, the N in red is the smallest natural number such that $x_n = \frac{1}{n}$ is contained in the yellow ϵ-strip for all $n \geq N$. The finitely many elements are the elements of the sequence that come before N, (there are only $N-1$ of them). It is important again to note that the N depends on the ϵ given. In this case with $x_n = \frac{1}{n}$, the smaller the ϵ, the 'further' we need to go out to find the appropriate $N \in \mathbb{N}$ that satisfies what we want.

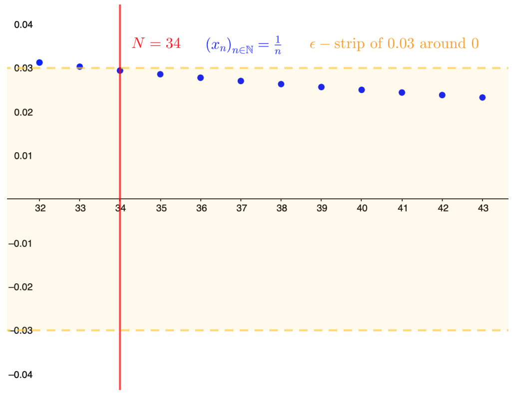
With hopefully a better understanding of the mechanics of convergence sequences, we prove a lemma about convergent sequences that will help us with later results.



(a) ϵ-strip of 1 around 0



(b) ϵ-strip of 0.1 around 0



(c) ϵ-strip of 0.03 around 0

Figure 11: Some ϵ-strips of 0 and the sequence $x_n = \frac{1}{n}$

Lemma 5.5: Convergent Sequences are Bounded

a) If the sequence $(a_n)_{n\in\mathbb{N}}$ converges to a, then the set $\{a_n : n \in \mathbb{N}\}$ is a bounded subset of $\mathbb{R}$.

b) If the sequence $(b_n)_{n\in\mathbb{N}}$ converges to b, and, with $b_n \neq 0$ for all $n \in \mathbb{N}$ and $b \neq 0$, then the set $\{\frac{1}{b_n} : n \in \mathbb{N}\}$ is a bounded subset of $\mathbb{R}$.

Proof.

a) Suppose $\lim\limits_{n\to\infty} a_n = a$. Since this is our hypothesis, we can claim something about any positive number. We choose our favorite, $\epsilon = 1$. Thus, there exists an $N \in \mathbb{N}$ such that $|a_n - a| < 1$ for all $n \geq N$. Using the reverse triangle inequality from Theorem 3.5, we have that $|a_n| - |a| \leq |a_n - a| < 1$, or rather, $|a_n| < |a| + 1$ for such $n \geq N$. So for all $n \geq N$, we have the upper bound $|a| + 1 > 0$ of a_n. To deal with the finitely many other values of n, we simply choose $M = \max\{|a_1|, |a_2|, \ldots, |a_{N-1}|, |a| + 1]\}$ Therefore, $M > 0$ as well, and it follows that for all $n \in \mathbb{N}$, $|a_n| \leq M$. Hence the set of all sequence elements $\{a_n : n \in \mathbb{N}\}$ is bounded.

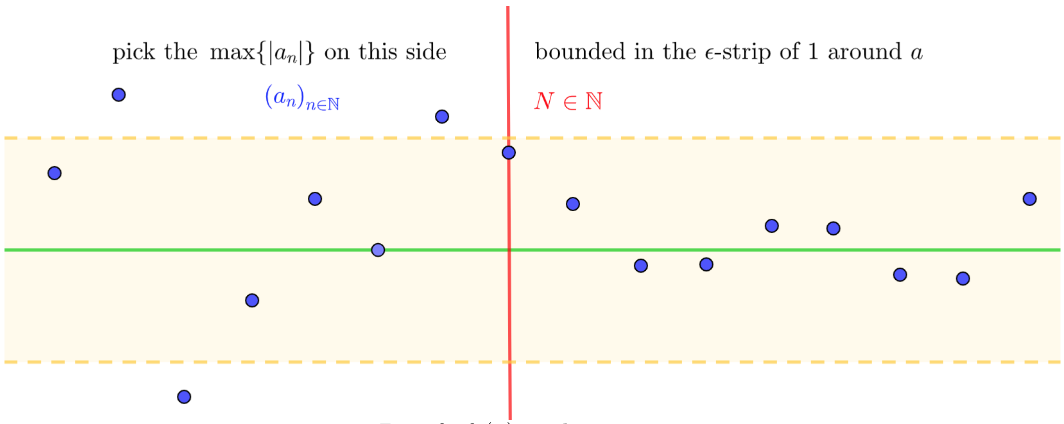


Figure 12: Proof of (a) with an ϵ-strip argument

b) For this part, we again know that $\lim_{n\to\infty} b_n = b$ but that all of the b_n and b are nonzero. We use a special positive number for this part, $\epsilon = \frac{|b|}{2}$. Thus, there exists an $N \in \mathbb{N}$ such that $|b_n - b| < \frac{|b|}{2}$ for all $n \geq N$. Using the same techniques this implies that

$$\begin{aligned} |b_n| &= |(b_n - b) + b| \\ &\geq |b| - |b_n - b| \\ &> |b| - \tfrac{|b|}{2} = \tfrac{|b|}{2}. \end{aligned}$$

Therefore, for all $n \geq N$, we have that $|b_n| > \frac{|b|}{2}$, or rather, $\frac{1}{|b_n|} < \frac{2}{|b|}$, and we have a bound on $\frac{1}{b_n}$ for all $n \geq N$. Again, we choose $M = \max\{\frac{1}{|b_1|}, \frac{1}{|b_2|}, \ldots, \frac{1}{|b_{N-1}|}, \frac{2}{|b|}\}$ and it follows as above that $\left|\frac{1}{b_n}\right| = \frac{1}{|b_n|} \leq M$ for all $n \in \mathbb{N}$. Hence, $\{\frac{1}{b_n} : n \in \mathbb{N}\}$ is bounded.

$\square$

Essentially, this lemma tells us that convergent sequences cannot 'blow up' or be unbounded. In the case of part (b), the reciprocal of sequence elements cannot be unbounded either. As a side-note, this gives us a way to prove that sequences do not converge, from the contrapositve of these statements. If a sequence is not bounded, then it is not convergent.

We now present one of the more important results concerned with real-valued sequences which allows us to find limits quickly:

Theorem 5.6: Algebra of Limits

Suppose that for sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$, that $\lim\limits_{n\to\infty} a_n = a$ and $\lim\limits_{n\to\infty} b_n = b$. Then

1. $\lim\limits_{n\to\infty} \alpha a_n = \alpha a$ for any $\alpha \in \mathbb{R}$,
2. $\lim\limits_{n\to\infty} (a_n + b_n) = a + b$,
3. $\lim\limits_{n\to\infty} a_n b_n = ab$, and
4. If $b_n \neq 0$ for all $n \in \mathbb{N}$ and $b \neq 0$, then $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$.

Proof. □

For each part of the theorem, we technically have new sequences we need to work with in the definition of sequence convergence, but we will be using the fact that $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ converge, of course.

1. Let $\alpha \in \mathbb{R}$. If $\alpha = 0$, then the sequence $(\alpha a_n)_{n\in\mathbb{N}}$ is a constant sequence of zeroes, which converges to $\alpha a = 0$. For the remainder of this part, $\alpha \neq 0$. Let $\epsilon > 0$. Since $\alpha \neq 0$, $|\alpha| > 0$ and hence $\frac{\epsilon}{|\alpha|} > 0$ as well. We use this positive number in Definition 5.2 with $(a_n)_{n\in\mathbb{N}}$ converging to a. There exists an $N \in \mathbb{N}$ such that for all $n \geq N$, $|a_n - a| < \frac{\epsilon}{|\alpha|}$. It follows that for such $n \geq N$,
$$\begin{aligned} |\alpha a_n - \alpha a| &= |\alpha(a_n - a)| \\ &= |\alpha|\,|a_n - a| \\ &< |\alpha| \cdot \frac{\epsilon}{|\alpha|} = \epsilon. \end{aligned}$$
Thus $\lim\limits_{n\to\infty} \alpha a_n = \alpha a$.

2. Let $\epsilon > 0$. As above, there exist some $N_a, N_b \in \mathbb{N}$ such that $|a_n - a| < \frac{\epsilon}{2}$ for all $n \geq N_a$ and $|b_n - b| < \frac{\epsilon}{2}$ for all $n \geq N_b$ (get ready for an $\frac{\epsilon}{2}$ argument). Thus, with $N = \max\{N_a, N_b\}$, we have that for all $n \geq N$,
$$\begin{aligned} |(a_n + b_n) - (a + b)| &= |(a_n - a) + (b_n - b)| \\ &\leq |a_n - a| + |b_n - b| \\ &< \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon. \end{aligned}$$
Thus $\lim\limits_{n\to\infty} (a_n + b_n) = a + b$.

3. We first note that if $a = 0$, then $\lim\limits_{n\to\infty} a_n b_n = 0 = ab$ by Exercise 7. This follows from the fact that $(b_n)_{n\in\mathbb{N}}$ is bounded since it is convergent. For the remainder of

the proof, we suppose that $a \neq 0$. Again, by Lemma 5.5, $(b_n)_{n\in\mathbb{N}}$ is bounded so that there exists an $M > 0$ such that $|b_n| \leq M$ for all $n \in \mathbb{N}$. Thus, we have the positive numbers $\frac{\epsilon}{2M} > 0$ and $\frac{\epsilon}{2|a|} > 0$ that we have picked so everything works out nicely. The positive number $\frac{\epsilon}{2M}$ will be used with $(a_n)_{n\in\mathbb{N}}$: there exists an $N_a \in \mathbb{N}$ such that $|a_n - a| < \frac{\epsilon}{2M}$ for all $n \geq N_a$. Similarly, $\frac{\epsilon}{2|a|}$ will be used with $(b_n)_{n\in\mathbb{N}}$: there exists an $N_b \in \mathbb{N}$ such that $|b_n - b| < \frac{\epsilon}{2|a|}$ for all $n \geq N_b$. We pick $N = \max\{N_a, N_b\}$, so that for all $n \geq N$, we have that

$$\begin{aligned}
|a_nb_n - ab| &= \left|a_nb_n \underbrace{-b_na + b_na}_{\text{special } 0} -ab\right| \\
&= |b_n(a_n - a) + a(b_n - b)| \\
&\leq |b_n|\,|a_n - a| + |a|\,|b_n - b| \\
&\leq M\,|a_n - a| + |a|\,|b_n - b| \\
&< M\tfrac{\epsilon}{2M} + |a|\,\tfrac{\epsilon}{2|a|} = \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Therefore, $\lim\limits_{n\to\infty} a_nb_n = ab$, as desired.

4. Again, if $a = 0$, $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$ as above by Exercise 7. We suppose again that $a \neq 0$ for the remainder of the proof. By Lemma 5.5, there exists an $M > 0$ such that $\frac{1}{|b_n|} \leq M$ for all $n \in \mathbb{N}$. This time, our positive numbers are a bit more extravagant and the algebra in the chain of inequalities is a bit more complex. We have that $\frac{\epsilon}{2M} > 0$ and thus, there exists an $N_a \in \mathbb{N}$ such that $|a_n - a| < \frac{\epsilon}{2M}$ for all $N \geq N_a$. Similarly, since $\frac{\epsilon|b|}{2M|a|} > 0$ (given $a, b \neq 0$), there exists an $N_b \in \mathbb{N}$ such that $|b_n - b| < \frac{\epsilon|b|}{2M|a|}$ for all $n \geq N_b$. Therefore, for all $n \geq \max\{N_a, N_b\}$,

$$\begin{aligned}
\left|\frac{a_n}{b_n} - \frac{a}{b}\right| &= \left|\frac{a_nb - b_na}{b_nb}\right| \\
&= \left|\frac{a_nb - ab + ab - b_na}{b_nb}\right| \\
&= \frac{|b(a_n - a) - a(b_n - b)|}{|b_nb|} \\
&\leq \frac{|b|\,|a_n - a| + |a|\,|b_n - b|}{|b_nb|} \\
&= \frac{1}{|b_n|}\,|a_n - a| + \frac{|a|}{|b_n|\,|b|}\,|b_n - b| \\
&\leq M\,|a_n - a| + \frac{M\,|a|}{|b|}\,|b_n - b| \\
&< M \cdot \frac{\epsilon}{2M} + \frac{M\,|a|}{|b|} \cdot \frac{\epsilon\,|b|}{2M\,|a|} = \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Once again, our desired limit is proven.

These last two proofs of parts (3) and (4) may seem a bit confusing when considering the choice of the positive numbers $\frac{\epsilon}{2M}$ or $\frac{\epsilon|b|}{2M|a|}$ that made everything work out. Another common trick in a class like MAT 370 is to work out the scratch-work of the algebra **beforehand** to see what is needed to get everything to be less than ϵ. That is what we did in the proofs; we just rewrote them to make it look like we got the special positive numbers out of nowhere like magic. With these results on the algebra of limits try out Exercise 6 to practice with some rational sequences!

This theorem on the algebra of limits is really only useful when we know that the sequences we are adding, multiplying, or dividing were originally convergent. A 'deep' result in MAT 370, one that appeals to the Completeness Axiom of $\mathbb{R}$, is a theorem that tells us when certain sequences are convergent.

Definition 5.7: Monotone Sequences

A sequence $(x_n)_{n\in\mathbb{N}}$ is said to be **monotone increasing** if and only if $x_n \leq x_{n+1}$ for all $n \in \mathbb{N}$. Similarly, a sequence $(y_n)_{n\in\mathbb{N}}$ is said to be **monotone decreasing** if and only if $y_n \geq y_{n+1}$ for all $n \in \mathbb{N}$. The term **monotone** is used in general if a sequence is monotone increasing or decreasing.

In addition to this notion of monotonicity, there is another quality that a monotone sequence needs that is necessary to guarantee convergence. This quality is boundedness. We give a visualization. Suppose we have a monotone sequence that is bounded above (see Figure 13). Then, informally, the sequence elements are 'trapped' and will continue to move upward to the upper bound. It turns out that in cases like this, the limit of the sequence is the supremum of the sequence elements, which we prove here:

Theorem 5.8: The Monotone Convergence Theorem

Suppose $(x_n)_{n\in\mathbb{N}}$ is a monotone increasing sequence that is bounded above. Then, $\lim\limits_{n\to\infty} x_n$ exists and is equal to $\sup\{x_n : n \in \mathbb{N}\}$.

Proof. Since the sequence is bounded above by hypothesis, the set $\{x_n : n \in \mathbb{N}\}$ is bounded above and nonempty and thus by the Completeness Axiom, $\sup\{x_n : n \in \mathbb{N}\}$ exists and we denote it by $\alpha \in \mathbb{R}$. We let $\epsilon > 0$ to show that $\lim\limits_{n\to\infty} x_n = \alpha$. Since $\alpha - \epsilon < \alpha$ and α is the supremum, there exists an $N \in \mathbb{N}$ such that $\alpha - \epsilon < x_N$ by Theorem 3.12. For all $n \geq N$, we also have that $\alpha - \epsilon < x_n$ since $(x_n)_{n\in\mathbb{N}}$ is monotone increasing. By the definition of α as an upper bound, we know that $x_n \leq \alpha < \alpha + \epsilon$ for all $n \in \mathbb{N}$ and hence for all $n \geq N$.

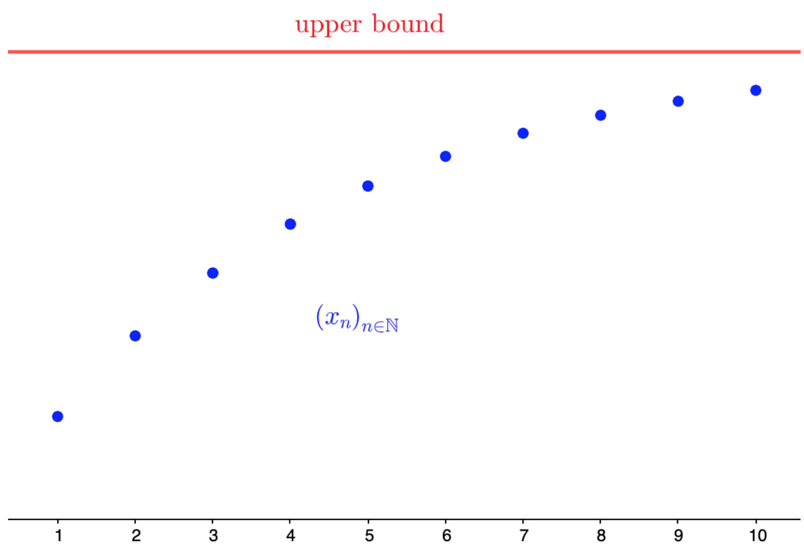


Figure 13: A monotone, bounded above sequence

Thus, for $n \geq N$, we have that $\alpha - \epsilon < x_n < \alpha + \epsilon$,which implies $|x_n - \alpha| < \epsilon$. Therefore, $\lim_{n\to\infty} x_n = \alpha$. □

Using the Monotone Convergence Theorem, the convergence of many different sequences can be proven. For example, many sequences defined recursively, i.e. $x_{n+1} = f(x_n)$, can be shown to be monotone and bounded and thus convergent. See the exercises in this section for some examples of this.

One last simple result in this section on sequences is appropriately named the Squeeze Theorem. Imagine you have two sequences $(a_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$ that both converge to $x \in \mathbb{R}$ (see Figure 14). If another sequence $(b_n)_{n\in\mathbb{N}}$ satisfies $a_n \leq b_n \leq c_n$ for all $n \in \mathbb{N}$, then it seems reasonable that $(b_n)_{n\in\mathbb{N}}$ would also converge to x. We give a more rigorous proof of the result here:

Theorem 5.9: The Squeeze Theorem

Suppose that for the sequences $(a_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$, $\lim_{n\to\infty} a_n = x = \lim_{n\to\infty} c_n$ for some $x \in \mathbb{R}$. If $(b_n)_{n\in\mathbb{N}}$ is another sequence that satisfies $a_n \leq b_n \leq c_n$ for all $n \in \mathbb{N}$, then $\lim_{n\to\infty} b_n = x$ as well.

Proof. We let $\epsilon > 0$. By our hypothesis, we know that $a_n - x \leq b_n - x \leq c_n - x$. Since $(a_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$ are convergent, there exist $N_a, N_c \in \mathbb{N}$ such that $-\epsilon < a_n - x < \epsilon$ for all $n \geq N_a$ and $-\epsilon < c_n - x < \epsilon$ for all $n \geq N_c$. Thus, with $N = \max\{N_a, N_c\}$, when

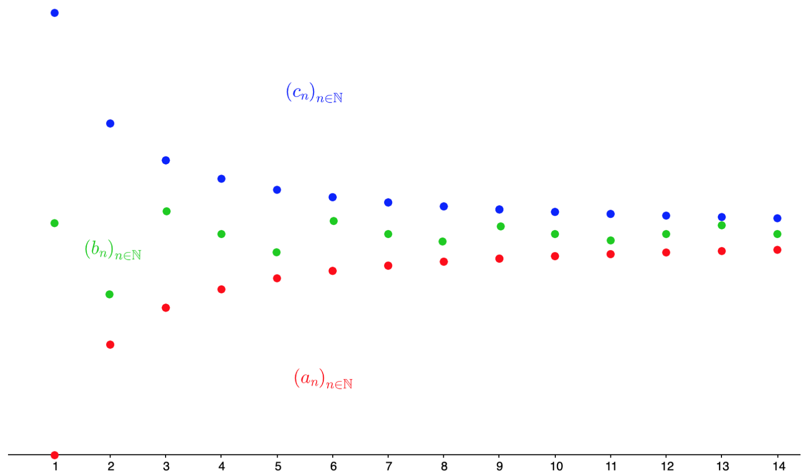


Figure 14: A Squeeze Theorem Visualization

$n \geq N$, both inequalities are satisfied simultaneously, and thus

$$-\epsilon < a_n - x \leq b_n - x \leq c_n - x < \epsilon,$$

which implies $|b_n - x| < \epsilon$. Therefore, $\lim_{n\to\infty} b_n = x$, as desired. $\square$

5.1 Subsequences

Another tool at our disposal in MAT 370 is the notion of **subsequences** of a sequence $(x_n)_{n\in\mathbb{N}}$. If you imagine the sequence $(x_n)_{n\in\mathbb{N}}$ as a list $(x_1, x_2, x_3, \ldots)$, a subsequence is any other list of elements that we select where the indices are increasing, i.e.

$$\begin{aligned}
&(x_1, x_3, x_5, x_7, \ldots), \\
&(x_4, x_{35}, x_{102}, x_{7,589}, \ldots), \text{ or even,} \\
&(x_1, x_2, x_3, x_4, \ldots) = (x_n)_{n\in\mathbb{N}}.
\end{aligned}$$

These would all be subsequences of $(x_n)_{n\in\mathbb{N}}$. Note that each of these subsequence is a sequence in and of itself, and thus we can discuss whether or not it converges just as we may talk about $(x_n)_{n\in\mathbb{N}}$ converging or not. We give our official definition of subsequences here:

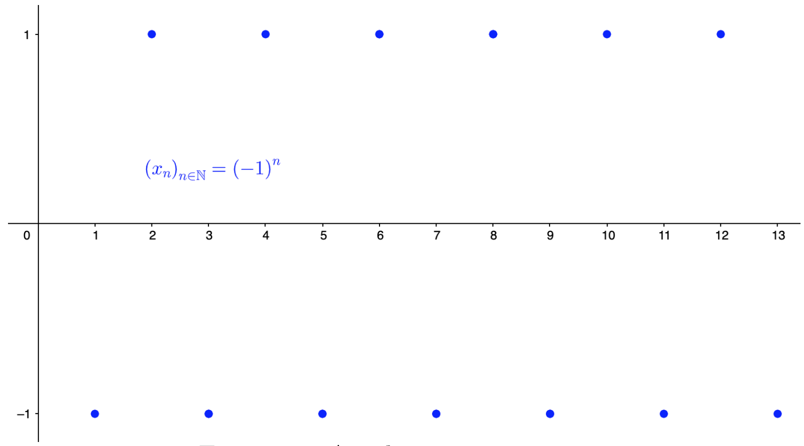


Figure 15: An alternating sequence

Definition 5.10: Subsequences

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence and let $(n_k)_{k\in\mathbb{N}}$ be a strictly increasing sequence of natural numbers, i.e. $n_1 < n_2 < n_3 < \cdots$. The sequence where we only look at the sequence elements at the given n_k indices, $(x_{n_k})_{k\in\mathbb{N}}$, is called a **subsequence** of $(x_n)_{n\in\mathbb{N}}$.

We give a simple example of a sequence that does not converge yet has convergent subsequences.

Example 5.11: An Alternating Sequence

Consider the sequence $(x_n)_{n\in\mathbb{N}}$ given by $x_n = (-1)^n$ for all $n \in \mathbb{N}$ (see Figure 15). Show that $(x_n)_{n\in\mathbb{N}}$ does not converge to 1 or -1, but has subsequences that do.

Proof. To deal with $(x_n)_{n\in\mathbb{N}}$ not converging to 1, we pick a bad ϵ_0, which in this case we pick to be $\epsilon_0 = 1$ and let $n \in \mathbb{N}$ (all of this is coming from negating Definition 5.2). If n is even, then $x_n = 1$ and we pick deal with $n+1$. It follows that $|x_{n+1} - 1| = |-1-1| = 2 > \epsilon_0$. If n is odd to begin with, then $|x_n - 1| = |-1-1| = 2 > \epsilon_0$ as well. It follows that $(x_n)_{n\in\mathbb{N}}$ does not converge to 1 and a similar argument follows for why it does not converge to -1 either.

For the convergent subsequences, if we take the subsequence of all even indices, we have that this sequences is a constant sequence of ones and thus converges to 1. Alternatively, if we take the sequence of all odd indices, we have that this sequences is a constant sequence of negative ones and thus converges to -1. $\square$

This example is a good indicator to an essential theorem relating sequence convergence and subsequence convergence:

Theorem 5.12: Subsequence and Sequence Convergence

A sequence $(x_n)_{n\in\mathbb{N}}$ converges to $x \in \mathbb{R}$ if and only if every subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ converges to x.

Proof. The converse direction is easy: if every subsequence converges to x, then $(x_n)_{n\in\mathbb{N}}$ converges to x since it is a subsequence of itself, as noted in the introduction to this section on subsequences.

For the forward direction, we suppose $\lim_{n\to\infty} x_n = x$, pick an arbitrary subsequence $(x_{n_k})_{k\in\mathbb{N}}$, and let $\epsilon > 0$ to show convergence. Since $\lim_{n\to\infty} x_n = x$, there exists an $N \in \mathbb{N}$ such that $|x_n - x| < \epsilon$ for all $n \geq N$. A general fact about sequences is that $n_k \geq k$ for the all subindices of the subsequence. Since the n_k are increasing, the first subindex $n_1 \geq 1$ since it is a natural number, but the second subindex $n_2 \geq 2$ since it is strictly greater than n_1. It of course can equal 2, so we cannot guarantee strict inequalities. However, we do not need them. For all $k \geq N$, we have that $n_k \geq k \geq N$ as well and thus,

$$|x_{n_k} - x| < \epsilon,$$

since these n_k are greater than or equal to N as well. □

We give a major result of this theorem as a corollary even through it follows immediately from being the contrapositive of the theorem:

Corollary 5.13: $(x_n)_{n\in\mathbb{N}}$ Not Converging

A sequence $(x_n)_{n\in\mathbb{N}}$ is not convergent if and only if there exist two subsequences that converge to different numbers.

In general, it is hard to prove that a sequence $(x_n)_{n\in\mathbb{N}}$ does not converge from just negating the definition: something needs to be checked for all real numbers!. However, if you can find two convergent subsequences that converge to different numbers, then you are done. We could have used this corollary to prove Example 5.11 quickly.

The 'deep' result of this section is attributed to the Bohemian mathematician Bernard Bolzano who proved it as a lemma to the Intermediate Value Theorem in 1817. It was later proved by German mathematician Karl Weierstrass in its own right some 50 years later. The theorem itself is connected to the 'complete' nature of $\mathbb{R}$, as described in Chapter 2. We have already seen in an example that the bounded sequence $x_n = (-1)^n$ has a convergent subsequence. In fact, the Bolzano-Weierstrass Theorem guarantees this fact for any arbitrary bounded sequence:

Theorem 5.14: The Bolzano-Weierstrass Theorem

Every bounded sequence $(x_n)_{n\in\mathbb{N}}$ has some convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$.

We first prove a lemma about arbitrary sequences that will allow us to apply the Monotone Convergence Theorem.

Lemma 5.15: Monotone Subsequences

Every sequence has a monotone subsequence.

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be an arbitrary sequence of real numbers. For this proof, we call an element x_m of $(x_n)_{n\in\mathbb{N}}$ a **peak** if $x_m \geq x_n$ for all $n > m$. We now break into two cases to construct monotone subsequences of $(x_n)_{n\in\mathbb{N}}$

Case 1: $(x_n)_{n\in\mathbb{N}}$ has infinitely many peaks. Set $n_1 \in \mathbb{N}$ as the smallest natural number such that x_{n_1} is a peak. Then $x_{n_1} \geq x_n$ for all $n \geq n_1$. Next, pick n_2 to be the smallest natural number not equal to n_1 such that x_{n_2} is a peak. It follows that $n_1 < n_2$ so that $x_{n_1} \geq x_{n_2}$ and $x_{n_2} \geq x_n$ for all $n \geq n_2$. Suppose that for an arbitrary $k \in \mathbb{N}$, $x_{n_1}, x_{n_2}, \ldots, x_{n_k}$ are peaks chosen as above, $x_{n_1} \geq x_{n_2} \geq \cdots x_{n_k}$. Since there are infinitely many peaks, we are able to pick a natural number n_{k+1} larger than n_k such that $x_{n_{k+1}}$ is a peak and $x_{n_k} \geq x_{n_{k+1}}$. By induction, this gives us a monotone decreasing subsequence $(x_{n_k})_{k\in\mathbb{N}}$.

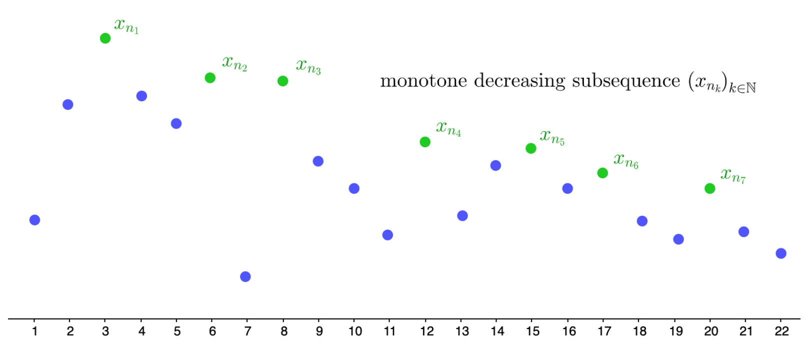


Figure 16: A monotone decreasing subsequence

Case 2: $(x_n)_{n\in\mathbb{N}}$ has only finitely many peaks. Suppose our finitely many peaks are labeled as $x_{n_1}, x_{n_2}, \ldots, x_{n_k}$ with $n_1 < n_2 < \cdots < n_k$. We set $m_1 := n_k + 1$. so that x_n is not a peak for all $n \geq N$. Starting with x_{m_1}, we have that there exists an $m_2 > m_1$ such that $x_{m_1} < x_{m_2}$, since x_{m_1} is not a peak. Further, since x_{m_2} is not a peak, there exists an $m_3 > m_2$ such that $x_{m_2} < x_{m_3}$. We continue inductively to gain a subsequence $(x_{m_k})_{k \in \mathbb{N}}$ such that $x_{m_1} < x_{m_2} < x_{m_3} < \cdots$, i.e. is strictly increasing.

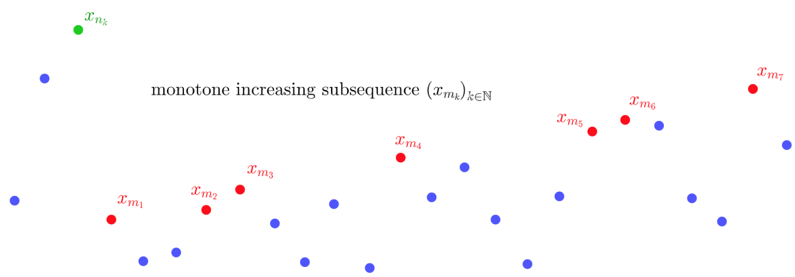


Figure 17: A monotone increasing subsequence

□

We now return to the proof of the Bolzano-Weierstrass Theorem:

Proof. Let $(x_n)_{n \in \mathbb{N}}$ be an arbitrary bounded sequence of real numbers. Then by Lemma 5.15, there exists some monotone subsequence $(x_{n_k})_{k \in \mathbb{N}}$ of $(x_n)_{n \in \mathbb{N}}$. Since the original sequence $(x_n)_{n \in \mathbb{N}}$ is bounded, $(x_{n_k})_{k \in \mathbb{N}}$ is as well. By the Monotone Convergence Theorem, $(x_{n_k})_{k \in \mathbb{N}}$ is convergent and the result follows. □

5.2 Cauchy Sequences

The last type of real-valued sequences we will investigate in this chapter are called Cauchy sequences. Unlike convergent sequences, Cauchy sequences need not have a point that they converge to. However, for a Cauchy sequence $(x_n)_{n \in \mathbb{N}}$, the x_n terms get arbitrarily close to **each other** as n approaches infinity. The terms of a convergent sequence get arbitrarily close to the limit as n approaches infinity. Here is the official definition of a Cauchy sequence:

Definition 5.16: Cauchy Sequences

A sequence is said to be **Cauchy** if and only if for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$|x_m - x_n| < \epsilon$$

for all $m, n \geq N$.

The condition here that $|x_m - x_n| < \epsilon$ for all $m, n \geq N$ is crucial yet subtle. This implies that any two elements past the special x_N are less than ϵ apart. Of course, if a sequence is Cauchy, then subsequent terms x_n and x_{n+1} will be less than ϵ apart for all $n \geq N$. However, the converse is not true. Showing that $|x_n - x_{n+1}| < \epsilon$ for all $n \geq N$ does not imply that $(x_n)_{n\in\mathbb{N}}$ is Cauchy, in general (see Exercise 15 for an example of this).

We give two lemmas that will help us show that being Cauchy is in fact no different than being convergent when talking about sequences of real numbers.

Lemma 5.17: Convergent Implies Cauchy

Every convergent sequence is Cauchy.

Proof. Suppose $(x_n)_{n\in\mathbb{N}}$ is convergent such that $\lim_{n\to\infty} x_n = x$. To show Cauchy, we let $\epsilon > 0$. Then, there exists a $N \in \mathbb{N}$ such that $|x_n - x| < \frac{\epsilon}{2}$ (get ready for an $\frac{\epsilon}{2}$ argument). Thus, for any two $m, n \geq N$, we have that

$$\begin{aligned} |x_m - x_n| &= |(x_m - x + x - x_n| \\ &= |(x_m - x) - (x_n - x)| \\ &\leq |x_m - x| + |x_n - x| \\ &< \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon. \end{aligned}$$

Thus, $(x_n)_{n\in\mathbb{N}}$ is Cauchy. □

Just as with convergent sequences, it can be shown that Cauchy sequences form bounded subsets of $\mathbb{R}$ as well.

Lemma 5.18: Cauchy Sequences are Bounded

For any Cauchy sequence $(x_n)_{n\in\mathbb{N}}$, the set $\{x_n : n \in \mathbb{N}\}$ is a bounded subset of $\mathbb{R}$.

Proof. We use a similar technique as with Lemma 5.5. Since we know that $(x_n)_{n\in\mathbb{N}}$ is Cauchy, we pick our favorite positive number $\epsilon = 1$ to start building an upper bound. Thus, there exists an $N \in \mathbb{N}$ such that $|x_m - x_n| < 1$ for all $m, n \geq N$. We have that for all $n \geq N$,

$$\begin{aligned} |x_n| &= |x_n - x_N + x_N| \\ &\leq |x_n - x_N| + |x_N| \\ &< 1 + |x_N|. \end{aligned}$$

Since this x_N is fixed, $1 + |x_N| > 0$ is still a valid upper bound for the x_n, $n \geq N$, for the finitely many $n \leq N - 1$, we again pick $M = \max\{|x_1|, |x_2|, \ldots, |x_{N-1}|, 1 + |X_N|\} > 0$. Thus $|x_n| \leq M$ for all $n \in \mathbb{N}$ and it follows that $\{x_n : n \in \mathbb{N}\}$ is bounded. □

The 'deep' result of this section on sequences is that a sequence is Cauchy if and only if it is convergent. Thus, if you can get your hands on a Cauchy sequence of real numbers $(x_n)_{n\in\mathbb{N}}$, which might be easier to do in practical situations, then you are guaranteed the existence of an $x \in \mathbb{R}$ such that $\lim_{n\to\infty} x_n = x$.

Theorem 5.19: Cauchy-Convergent Equivalence

A sequence of real numbers is Cauchy if and only if it is convergent.

Proof. The converse direction has already been proven with Lemma 5.17.

For the forward direction, suppose that $(x_n)_{n\in\mathbb{N}}$ is Cauchy. By Lemma 5.18, $(x_n)_{n\in\mathbb{N}}$ is bounded, and thus by the Bolzano-Weierstrass Theorem, there exists a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$. By Exercise 16, which we leave to the reader to prove, it follows that entire $(x_n)_{n\in\mathbb{N}}$ is convergent as well. □

5.3 Exercises

1. Show that $|x| < \epsilon$ for all $\epsilon > 0$ if and only if $x = 0$.

2. Prove that $\lim_{n\to\infty} \frac{1}{n} = 0$. For the sequence $x_n = \frac{1}{n}$, find the smallest $N \in \mathbb{N}$ such that $x_n = \frac{1}{n} \in (-\epsilon, \epsilon)$ for all $n \geq N$, with $\epsilon = 0.005$, $\epsilon = 0.0006$, and $\epsilon = 0.00007$.

3. Prove that the sequence $x_n = \frac{n^2+n+1}{n+1}$ is not convergent.

4. a) State what it means for a sequence $(x_n)_{n\in\mathbb{N}}$ to not converge to a point $a \in \mathbb{R}$ by negating Definition 5.2 of sequence convergence.

   b) For the sequence given by

   $$x_n = \begin{cases} 1 & n \text{ is a multiple of } 3 \\ \frac{1}{n} & \text{else} \end{cases}$$

   show that $(x_n)_{n\in\mathbb{N}}$ is bounded, but does not converge to 0. Hence, unbounded sequences are not the only sequences that do not converge.

5. Suppose that two convergent sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ satisfy $x_n \leq y_n$ for all $n \in \mathbb{N}$. Then $\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n$.

6. Prove that

   a) $\lim_{n\to\infty} \frac{1}{n^k} = 0$ for any $k \in \mathbb{N}$.

   b) $\lim_{n\to\infty} \dfrac{a_1n^2 + b_1n + c_1}{a_2n^2 + b_2n + c_2} = \dfrac{a_1}{a_2}$.

c) $\lim_{n\to\infty} \dfrac{P(n)}{Q(n)} = \dfrac{\text{leading coefficient of } P(n)}{\text{leading coefficient of } Q(n)}$, where $P(n)$ and $Q(n)$ are two polynomials in n, both having degree $k \in \mathbb{N}$.

d) $\lim_{n\to\infty} \dfrac{a_1 n + b_1}{a_2 n^2 + b_2 n + c_2} = 0$.

e) $\lim_{n\to\infty} \dfrac{P(n)}{Q(n)} = 0$, where $P(n)$ and $Q(n)$ are two polynomials in n such that the degree of $P(n)$ is less than the degree of $Q(n)$.

7. Suppose that for the sequence $(a_n)_{n\in\mathbb{N}}$, $\lim_{n\to\infty} a_n = 0$. If $(b_n)_{n\in\mathbb{N}}$ is any bounded sequence, show that $\lim_{n\to\infty} a_n b_n = 0$ as well.

8. Prove the Monotone Convergence Theorem for sequences that are monotone decreasing and bounded below.

9. Using the Monotone Convergence Theorem, prove that the following sequences defined recursively are convergent and find their limit using the algebra of limits:

   a) $x_1 = 2$ and $x_{n+1} = 2 - \frac{1}{x_n}$ for all $n \in \mathbb{N}$.

   b) $x_1 = \sqrt{2}$ and $x_{n+1} = \sqrt{2 + x_n}$ for all $n \in \mathbb{N}$.

   c) $x_1 = 2$ and $x_{n+1} = \frac{1}{3 - x_n}$ for all $n \in \mathbb{N}$. *Note*: This sequence $(x_n)_{n\in\mathbb{N}}$ is a sequence of rational numbers whose limit is not a rational number. The Completeness Axiom for the **real numbers** is critical.

10. Using the Monotone Convergence Theorem, show that for $0 \le a < 1$, $\lim_{n\to\infty} a^n = 0$.

11. Using the Squeeze Theorem, prove that $\lim_{n\to\infty} \frac{\sin(n)}{n} = 0$.

12. Show that the sequence $(x_n)_{n\in\mathbb{N}}$ given by $x_n = \sin(\frac{\pi n}{2})$ is not convergent, but does have subsequences that converge to 0, 1, and -1.

13. Suppose that the sequence $(x_n)_{n\in\mathbb{N}}$ is unbounded. Show that there exists some subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ that satisfies $\lim_{k\to\infty} \frac{1}{x_{n_k}} = 0$.

14. Prove that the Bolzano-Weierstrass Theorem is logically equivalent to the following statement:

    every bounded, infinite subset of $\mathbb{R}$ has a limit point.

15. Consider the sequence $(x_n)_{n\in\mathbb{N}}$ given by the pattern

$$(0, \tfrac{1}{2}, 1, \tfrac{2}{3}, \tfrac{1}{3}, 0, \tfrac{1}{4}, \tfrac{2}{4}, \tfrac{3}{4}, 1, \tfrac{4}{5}, \tfrac{3}{5}, \ldots) \quad \text{(see Figure 18)}.$$

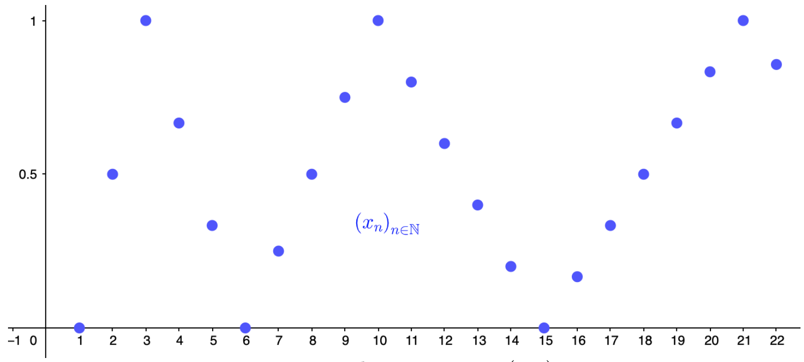


Figure 18: The sequence $(x_n)_{n\in\mathbb{N}}$

Prove that $(x_n)_{n\in\mathbb{N}}$ satisfies the following property:

$$\text{for all } \epsilon > 0, \text{ there exists an } N \in \mathbb{N} \text{ such that } |x_n - x_{n+1}| < \epsilon \text{ for all } n \geq N,$$

but that $(x_n)_{n\in\mathbb{N}}$ is not Cauchy.

16. If a sequence $(x_n)_{n\in\mathbb{N}}$ is Cauchy and has some convergent subsequence, show that $(x_n)_{n\in\mathbb{N}}$ is convergent.

17. Prove that a sequence of real numbers is Cauchy if and only if every subsequence is Cauchy.

18. Show that instead of the Completeness Axiom, we can prove the Archimedean Property by supposing the Monotone Convergence Theorem instead.

6 Continuity

Continuous functions are perhaps the bread and butter of calculus classes and for the results we know and love. These functions behave as we expect they should: one can draw a continuous function without lifting their pen from the page. This more intrinsic definition of continuity needs to be made mathematically rigorous in order to prove functions are continuous and to prove things about continuous functions.

Definition 6.1: Continuity

Let $f : D \to \mathbb{R}$, where $D \subseteq \mathbb{R}$. We say that f is **continuous at** $x_0 \in D$ if and only if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x \in D$,

$$|x - x_0| < \delta \quad \text{implies} \quad |f(x) - f(x_0)| < \epsilon.$$

We say f **continuous** if and only if it is continuous at every $x \in D$.

What this definition is telling us is that a function is continuous if the outputs are arbitrarily close if we can restrict the inputs to a certain interval, $(x_0 - \delta, x_0 + \delta)$. We prove the continuiuty of an important function and by negating Definition 6.1, we can begin to look at non examples of continuity:

Example 6.2: Absolute Value is Continuous

The function $f(x) = |x|$ for all $x \in \mathbb{R}$ is continuous.

Proof. To show continuity via Definition 6.1, we let $c \in \mathbb{R}$ and $\epsilon > 0$ be arbitrary. It turns out that with the absolute value function $\delta = \epsilon > 0$ works. Thus, for all $x \in \mathbb{R}$ such that $|x - c| < \delta$ we have that

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

The first inequality is the reverse triangle inequality from Theorem 3.5. □

Example 6.3: Non-continuous Functions

Show that

1. $f(x) = \begin{cases} 0 & x \leq 1 \\ 1 & x > 1 \end{cases}$ is not continuous at 1.
2. (Dirichlet's function) $f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$ is not continuous at any $c \in \mathbb{R}$.
3. $f(x) = \begin{cases} x & x \in \mathbb{Q} \\ -x & x \notin \mathbb{Q} \end{cases}$ is continuous at 0 and nowhere else.

Proof. 1. In the negation of Definition 6.1, our faulty ϵ_0 is $\epsilon_0 = \frac{1}{2}$. We let $\delta > 0$. Then with the point $1 + \frac{\delta}{2}$, $\left|1 + \frac{\delta}{2} - 1\right| = \frac{\delta}{2} < \delta$ but $\left|f(1 + \frac{\delta}{2}) - f(1)\right| = |1 - 0| = 1 \geq \epsilon_0$. For continuity elsewhere, note that over the intervals $(1, \infty)$ and $(-\infty, 1$, f is constant (either 1 or 0, respectively) and hence is continuous by Exercise 1.

2. First, we suppose that $q \in \mathbb{Q}$ so that $f(q) = 1$. Setting $\epsilon_0 = \frac{1}{2}$, we know that for all $\delta > 0$, there exists an irrational number $\xi_\delta \in (q - \delta, q + \delta)$ by the density of the irrationals. It follows that $|\xi_\delta - q| < \delta$, but $|f(\xi_\delta) - f(q)| = |0 - 1| = 1 \geq \epsilon_2$. Thus, f is not continuous on all of $\mathbb{Q}$ since q was arbitrary. The proof is similar for the irrational numbers, and by the density of the rationals, f is not continuous on all of $\mathbb{R} \setminus \mathbb{Q}$. Therefore, f is not continuous on all of $\mathbb{R}$.

3. We first show that f is continuous at 0. Let $\epsilon > 0$. Now, we let $x \in \mathbb{R}$ and suppose that $|x - 0| = |x| < \epsilon$. Regardless of x, we know that $|f(x) - f(0)| = |f(x)| = |x|$. Therefore, $|f(x) - f(0)| = |x| < \epsilon$ and f is continuous at 0. We show that f is not continuous on $(0, \infty)$; the negative side has a symmetric argument. Let $q \in \mathbb{Q}$ with $q > 0$. It turns out that for the negation of not continuous, $q > 0$ is also our faulty ϵ value. We let $\delta > 0$. As above, there exists a positive irrational number $\xi_\delta \in (q - \delta, q + \delta)$, or rather, $|\xi_\delta - q| < \delta$. However, $|f(\xi_\delta) - f(q)| = |-\xi_\delta - q| = |\xi_\delta + q| = \xi_\delta + q > q$. The argument is analogous if we had started with an arbitrary positive irrational ξ instead of q and the cases when $q < 0$ or $\xi < 0$ are similar as well.

$\square$

Even without being continuous, we can still talk about a function 'approaching' a point even it if might not be defined on that point. We now define what it means for the limit $\lim_{x \to c} f(x)$ to exist.

Definition 6.4: Limit of a Function

Let $f : D \to \mathbb{R}$ and let c be a limit point of D. We say that

$$\lim_{x \to c} f(x) = L$$

if and only if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x \in D$,

$$0 < |x - c| < \delta \quad \text{implies} \quad |f(x) - L| < \epsilon.$$

We discuss some of the subtleties of this definition. First, why do we only consider approaching a limit point? The natural numbers $\mathbb{Z}$ as a subset of $\mathbb{R}$ have no limit points. For a function $f : \mathbb{Z} \to \mathbb{R}$, what would we want $\lim_{m \to 3} f(m) = L$ to satisfy? As in Theorem 5.3, we hope that a function should only approach one value. Since there are no limit points of $\mathbb{Z}$, 3 is not a limit point and thus there exists a neighborhood $(3 - \epsilon_0, 3 + \epsilon_0)$ such that 3 is the only integer in $(3 - \epsilon_0, 3 + \epsilon_0)$. Thus, for any $L \in \mathbb{R}$ and any $\epsilon > 0$, that $0 < |m - 3| < \epsilon_0$ implies $f(x) - L < \epsilon$ is vacuously true: there are no $m \in \mathbb{Z}$ such that $0 < |m - 3| < \epsilon_0$.

It follows that in this case, f approaches every real number as n approaches 3. This is certainly not desirable and we restrict our discussion of function limits to limit points of D.

If this new notation of function limits is unfamiliar, no fear. We prove a lemma that relates function limits of Definition 6.4 with sequences of real numbers that we have seen in Chapter 5.

Lemma 6.5: Function Limits with Sequences

Let $f : D \to \mathbb{R}$ and c be a limit point of D. The following statements are equivalent:

1. $\lim_{x \to c} f(x) = L$.
2. For all sequences $(x_n)_{n\in\mathbb{N}} \subseteq D$ such that $x_n \neq c$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} x_n = c$, $\lim_{n\to\infty} f(x_n) = L$.

Proof. We first prove that (1) implies (2). Suppose $\lim_{x\to c} f(x) = L$ as in Definition 6.4 and let $(x_n)_{n\in\mathbb{N}}$ be an arbitrary sequence of D such that $x_n \neq c$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} x_n = c$. To prove (2), let $\epsilon > 0$. By Definition 6.4, there exists a $\delta > 0$ such that $0 < |x - c| < \delta$ implies $|f(x) - L| < \epsilon$ for all $x \in D$. Also, since $\lim_{n\to\infty} x_n = c$, there exists an $N \in \mathbb{N}$ such that $|x_n - c| < \delta$ for all $n \geq N$. Since $x_n \neq c$ for all $n \in \mathbb{N}$, $0 < |x_n - c| < \delta$ for all $n \geq N$. Thus, by Definition 6.4, it follows that $|f(x_n) - L| < \epsilon$ and thus $\lim_{n\to\infty} f(x_n) = L$.

For (2) implies (1), we prove by contradiction. We suppose (2) and that $\lim_{x\to c} f(x) \neq L$ by negating Definition 6.4. Then there exists an $\epsilon_0 > 0$ such that for all $\delta > 0$, there exists an $x_\delta \in D$ such that $0 < |x_\delta - c| < \delta$ but $|f(x_\delta) - L| \geq \epsilon_0$. For this proof, we need to contradict something about sequences and thus use $\delta = \frac{1}{n}$. Thus, for every $n \in \mathbb{N}$, there exists an $x_n \in D$ such that $0 < |x_n - c| < \frac{1}{n}$ but $|f(x_n) - L)| \geq \epsilon_0$. By construction, this $(x_n)_{n\in\mathbb{N}}$ is a sequence of D such that $x_n \neq c$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} x_n = c$. However, $\lim_{n\to\infty} f(x_n) \neq L$ since for $\epsilon_0 > 0$ and for all $n \in \mathbb{N}$, $|f(x_n) - L| \geq \epsilon_0$. This is exactly the negation of Definition 5.2 and hence is a contradiction to (2). $\square$

Another result that is helpful to have is similar to Theorem 5.6. We expect that function limits behave as sequences should, and in fact, they do.

Lemma 6.6: Algebra of Function Limits

Let $f, g : D \to \mathbb{R}$ with $D \subseteq \mathbb{R}$ and c a cluster point of D. Further suppose that $\lim_{x\to c} f(x) = L$ and $\lim_{x\to c} g(x) = M$. Then

1. $\lim_{x\to c} \alpha f(x) = \alpha L$ for any $\alpha \in \mathbb{R}$.
2. $\lim_{x\to c}(f(x) + g(x)) = L + M$.
3. $\lim_{x\to c} f(x)g(x) = LM$.
4. $\lim_{x\to c} \frac{f(x)}{g(x)} = \frac{L}{M}$ if $g(x) \neq 0$ for all $x \in D$ and $M \neq 0$.

Proof. We walk through the proof of the second result, but all other results follow symmetrically. We wish to show that $\lim_{x\to c}(f(x) + g(x)) = L + M = \lim_{x\to c} f(x) + \lim_{x\to c} g(x)$. By Lemma 6.5, it suffices to show for all sequences $(x_n)_{n\in\mathbb{N}}$ of D with $\lim_{n\to\infty} x_n = c$ and $x_n \neq c$ for all $n \in \mathbb{N}$ that $\lim_{n\to\infty}(f(x_n) + g(x_n)) = L + M$. So let $(x_n)_{n\in\mathbb{N}}$ be such a sequence. The result follows almost immediately since we already know that $f(x)$ approaches L and $g(x)$ approaches M. Thus, $\lim_{n\to\infty} f(x_n) = L$ and $\lim_{n\to\infty} g(x_n) = M$. Now, since we can consider the sequences $(f(x_n))_{n\in\mathbb{N}}$ and $(g(x_n))_{n\in\mathbb{N}}$,

$$\lim_{n\to\infty}(f(x_n) + g(x_n)) = \lim_{n\to\infty} f(x_n) + \lim_{n\to\infty} g(x_n) = L + M.$$

by Theorem 5.6. The results from (1), (3), (4) can be proved in the same way. □

Just as in Lemma 6.5, there is an equivalence for dealing with the continuity of a function at a point in terms of real-valued sequences. Further, we examine the standard definition of continuity seen in most calculus classes: that f is continuous at c if $\lim_{x\to c} f(x) = f(c)$. We show that our definition of continuity is equivalent to this one.

Theorem 6.7: Continuity Equivalences

Let $f : D \to \mathbb{R}$ and suppose that c is both an element of D as well as a limit point of D. The following statements are equivalent:

1. f is continuous at c.
2. For all sequences $(x_n)_{n\in\mathbb{N}} \subseteq D$ such that $\lim_{n\to\infty} x_n = c$, $\lim_{n\to\infty} f(x_n) = f(c)$.
3. $\lim_{x\to c} f(x) = f(c)$.

Proof. We have that (2) is equivalent to (3) by Lemma 6.5, replacing L with $f(c)$. The restriction for $x_n \neq c$ for all $n \in \mathbb{N}$ is not a problem since $c \in D$. If $x_N = c$ for some $N \in \mathbb{N}$, then $f(x_N) = f(c)$, which implies $|f(x_N) - f(c)| < \epsilon$ for all $\epsilon > 0$.

It is clear from Definitions 6.1 and 6.4 that (3) implies (1). We finish by showing that (1) implies (2). Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of D such that $\lim_{n\to\infty} x_n = c$ and let $\epsilon > 0$. By (1), there exists a $\delta > 0$ such that for all $x \in D$, $|x - c| < \delta$ implies $|f(x) - f(c)| < \epsilon$. Since $\lim_{n\to\infty} x_n = c$, there exists an $N \in \mathbb{N}$ such that $|x_n - c| < \delta$ for all $n \geq N$. Thus, for all such $n \geq N$, it follows that $|f(x_n) - f(c)| < \epsilon$, which implies $\lim_{n\to\infty} f(x_n) = f(c)$. □

A different way of viewing this idea of continuity is noting that from (2) above, we may write that for any sequence $(x_n)_{n\in\mathbb{N}}$ in D such that $\lim_{n\to\infty} x_n = c$, then

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right).$$

Essentially, limits can 'pass through' continuous functions.

Now for some of the most important theorems of this section: some tools for constructing continuous functions from functions that we know are continuous. These are the algebra of contiuous functions, that is, adding, multiplying, and dividing continuous functions produces continuous functions. Also, the composition of two continuous functions is continuous as well.

Theorem 6.8: Algebra of Continuous Functions

Suppose $f, g : D \to \mathbb{R}$, $c \in D$, and c is a limit point of D. Suppose f and g are continuous at c. Then,

1. $\alpha f + \beta g$ is continuous at c, for any $\alpha, \beta \in \mathbb{R}$.
2. fg is continuous at c.
3. If $g(x) \neq 0$ for all $x \in D$, then $\frac{f}{g}$ is continuous at c.

Proof. For this proof, we need only use Theorem 6.7 and the Algebra of Limits from Chapter 5. We walk through the process with (1). Using Theorem 6.7, we suppose $(x_n)_{n\in\mathbb{N}}$ is a sequence of D with $\lim_{n\to\infty} x_n = c$. Since both f and g are continuous at c, then $\lim_{n\to\infty} f(x_n) = f(c)$ and $\lim_{n\to\infty} g(x_n) = g(c)$. By the Algebra of Limits, we have that

$$\begin{aligned}
\lim_{n\to\infty} (\alpha f + \beta g)(x_n) &= \lim_{n\to\infty} (\alpha f(x_n) + \beta g(x_n)) \\
&= \alpha \lim_{n\to\infty} f(x_n) + \beta \lim_{n\to\infty} g(x_n) \\
&= \alpha f(c) + \beta g(c) \\
&= (\alpha f + \beta g)(c)
\end{aligned}$$

The proofs of (2) and (3) follow exactly the same argument, utilizing the Algebra of Limits that we know exist from continuity. □

Theorem 6.9: Compositions of Continuous Functions

Let $f : A \to B$ and $g : B \to \mathbb{R}$ for $A, B \subseteq \mathbb{R}$. Suppose f is continuous at $c \in A$ and g is continuous at $f(c) \in B$. Then the composition $g \circ f : A \to \mathbb{R}$ given by $(g \circ f)(x) = g(f(x))$ for all $x \in A$ is continuous at c.

Proof. This result follows by using Definition 6.1 a couple of times. Let $\epsilon > 0$. Since g is continuous at $f(c)$, then there exists a $\delta > 0$ such that for all $y \in B$, $|y - f(c)| < \delta$ implies $|g(y) - g(f(c))| < \epsilon$. Since f is continuous at c and $\delta > 0$, there exists an $\eta > 0$ such that for all $x \in A$, $|x - c| < \eta$ implies $|f(x) - f(c)| < \delta$. We emphasize that this $\eta > 0$ is what we will be using to prove that $g \circ f$ is continuous. For this $\eta > 0$ and all $x \in A$, it follows that by the two implications we have established, $|x - c| < \eta$ implies $|g(f(x)) - g(f(c))| < \epsilon$. This is exactly what we needed to show for $g \circ f$ to be continuous at c by Definition 6.1. $\square$

These two results help us prove that many types of functions are continuous. For some examples, see the exercises!

6.1 The Intermediate and Extreme Value Theorems

In this subsection, we prove some of the most important properties of continuous functions: the Intermediate and Extreme Value theorems. The results of this section depend heavily on domains of the form $[a, b]$, closed, bounded intervals of $\mathbb{R}$. These closed, bounded intervals are special in that they are **compact**, a topological property of much important in mathematics. Going through compactness is a bit past the scope of MAT 370, but just know that these compact intervals play a key role in the results of this section.

Definition 6.10: The Intermediate Value Property

Let $f : [a, b] \to \mathbb{R}$. We say that f **satisfies the Intermediate Value Property** if and only if for all y between $f(a)$ and $f(b)$, there exists an $x \in (a, b)$ such that $f(x) = y$.

Functions that satisfy the Intermediate Value Property need not be continuous, consider the following example: $f : [0, 2] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x & x \in [0, 1) \\ x - 1 & x \in [1, 2]. \end{cases}$$

This function f is not continuous at 1 but does satisfy the Intermediate Value Property: every y-value between $f(0) = 0$ and $f(2) = 1$ is hit by some $c \in (0, 2)$.

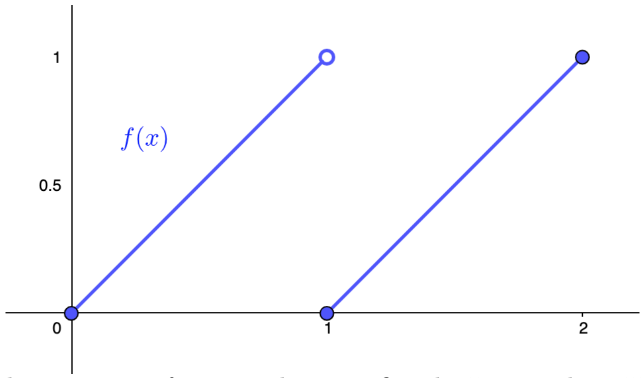


Figure 19: A discontinuous function that satisfies the Intermediate Value Property

In order to prove our first result, the Intermediate Value Theorem, we prove a lemma that is actually useful for proving the existence of roots of a function f: x-values such that $f(x) = 0$.

Lemma 6.11: Finding Roots of a Function

Suppose $f : [a, b] \to \mathbb{R}$ is continuous with $f(a) < 0 < f(b)$. Then there exists $c \in (a, b)$ such that $f(c) = 0$.

Proof. We begin by defining the set $N := \{x \in [a, b] : f(x) \leq 0\}$, which is nonempty since $a \in N$ and bounded above by b. By the Completeness Axiom, $\sup(N)$ exists and we call it c. We have that $c \in [a, b]$. We suppose by contradiction that $c = b$. In this case, there exists a sequence $(x_n)_{n\in\mathbb{N}}$ of elements of N such that $\lim\limits_{n\to\infty} x_n = c$ since c is the supremum of N. Since f is continuous we have that $\lim\limits_{n\to\infty} f(x_n) = f(c) = f(b)$. This is the contradiction since $f(x_n) \leq 0$ for all $n \in \mathbb{N}$ implies $\lim\limits_{n\to\infty} f(x_n) \leq 0$ (see Exercise 5 from the previous chapter) but $f(b) > 0$.

If $c = a$, then $N = \{a\}$, further contradicting the continuity of f. Since f is continuous at a and $f(a) < 0$, then $-\frac{f(a)}{2} > 0$ and there exists a $\delta > 0$ such that for all $x \in [a, b]$ with $|x - a| < \delta$, then $|f(x) - f(a)| < -\frac{f(a)}{2}$. We have that $a + \frac{\delta}{2} \in [a, a + \delta)$ so that $\left|f(a + \frac{\delta}{2}) - f(a)\right| < -\frac{f(a)}{2}$, which implies $f(a + \frac{\delta}{2}) < \frac{f(a)}{2} < 0$. Essentially, by continuiuty, of f we have found another element $a + \frac{\delta}{2} \in N$ that contradicts $N = \{a\}$. It follows that $c \in (a, b)$.

We claim that $f(c) = 0$. From Exercise 1 in the previous chapter, we show that $|f(c)| < \epsilon$ for all $\epsilon > 0$. We let $\epsilon > 0$ and note that since f is continuous at $c \in [a, b]$, there exists a $\delta > 0$ such that for all $x \in [a, b]$, $|x - c| < \delta$ implies $f(x) - \epsilon < f(c) < f(x) + \epsilon$.

By Theorem 3.12, we have that $c - \delta < c$ and thus there exists a $y \in N$ such that $c - \delta < y$. It follows that $c - \delta < y \leq c < c + \delta$ and thus $|y - c| < \delta$. By one side of our continuity implication, $f(c) < f(y) + \epsilon \leq \epsilon$ for this $y \in N$.

For the other inequality, we note that since c is the supremum of N, $c + \frac{\delta}{2}$ is not a member of N and thus $f(c + \frac{\delta}{2}) > 0$. Further, since $c + \frac{\delta}{2} \in (c, c + \delta)$, we apply the other side of our continuity implication: $f(c) > f(c + \frac{\delta}{2}) - \epsilon > -\epsilon$. By these two established inequalities, we have that $|f(c)| < \epsilon$ and our desired result $f(c) = 0$ is satisfied. $\square$

This lemma is a key tool for proving that certain functional equations have solutions. Check out the Exercises of this section for some examples of this.

Theorem 6.12: The Bolzano Intermediate Value Theorem

Let $f : [a, b] \to \mathbb{R}$ is continuous. Then f satisfies the Intermediate Value Property.

Proof. We let y be an arbitrary point between $f(a)$ and $f(b)$. We suppose that $f(a) < y < f(b)$. We consider the function $g : [a, b] \to \mathbb{R}$ given by $g(x) = f(x) - y$ for all $x \in [a, b]$. This g is continuous by the algebra of continuous functions since f is and y is a constant. Further, we have that $g(a) = f(a) - y < 0$ and $g(b) = f(b) - y > 0$. By Lemma 6.11, we have that there exists a $c \in (a, b)$ such that $g(c) = 0$. Thus, this $c \in (a, b)$ satisfies $f(c) = y$. and f satisfies the Intermediate Value Property.

In the case that $f(a) > y > f(b)$, this is proven given the result of Exercise 6. $\square$

To summarize this result, we have that the continuous image of a closed, bounded interval containing $[\min\{f(a), f(b)\}, \max\{f(a), f(b)\}]$ (see Figure 20).

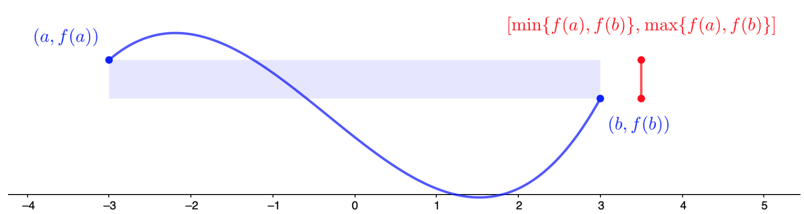


Figure 20: A visualization of the Intermediate Value Theorem

We can take this result one step further with the next theorem of this section, the Extreme Value Theorem. Not only does the continuous image of a closed, bounded interval contain a closed, bounded interval; in fact, such a continuous image is a closed, bounded interval. There exists $c, d \in [a, b]$ such that $f(c) \leq f(x) \leq f(d)$ for all $x \in [a, b]$.

Theorem 6.13: The Extreme Value Theorem

Suppose $f : [a, b] \to \mathbb{R}$ is continuous. Then there exists $x_m, x_M \in [a, b]$ such that $\sup\{f(x) : x \in [a, b]\} = f(x_M) = \max\{f(x) : x \in [a, b]\}$ and $\inf\{f(x) : x \in [a, b]\} = f(x_m) = \min\{f(x) : x \in [a, b]\}$

Proof. We know that the domain $[a, b]$ is bounded, but since f is continuous, we may show that f is bounded, or rather that $\{f(x) : x \in [a, b]\}$ is a bounded subset of $\mathbb{R}$. We suppose by contradiction that f is unbounded. In this case, for every $n \in \mathbb{N}$, there exists an $x_n \in [a, b]$ such that $|f(x_n)| > n$. Since $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence, by the Bolzano-Weierstrass Theorem, there exists a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ with $\lim_{k\to\infty} x_{n_k} = x_0$. Since f is continuous, we have that $\lim_{k\to\infty} f(x_{n_k}) = f(x_0)$. However, since this sequence $(f(x_{n_k})_{k\in\mathbb{N}}$ satisfies $|f(x_{n_k})| \geq n_k \geq k$, it is unbounded and does not converge by Lemma 5.5. It follows that f is bounded.

Therefore, $E := \{f(x) : x \in [a, b]\}$ is both nonempty and bounded. By the Completeness Axiom, both $\sup(E) = M$ and $\inf(E) = m$ exist. We prove that in the supremum case, there exists some x_M in $[a, b]$ such that $f(x_M) = M$. By Theorem 3.12, we once again know that since M is the supremum of E, there exists a sequence $(x_n)_{n\in\mathbb{N}}$ of elements of E such that $M - \frac{1}{n} < f(x_n) \leq M$ for all $n \in \mathbb{N}$. Since $(x_n)_{n\in\mathbb{N}}$ is bounded, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} x_{n_k}$ exists and we call this limit x_M. By the Squeeze Theorem and the continuity of f, $\lim_{k\to\infty} f(x_{n_k}) = f(x_M) = M$ and the result follows. The proof for x_m and the infimum is analogous; use the fact that there exists some sequence with $m \leq x_n < m + \frac{1}{n}$ for all $n \in \mathbb{N}$.

The fact that these $f(x_M)$ and $f(x_m)$ are the max and min of E, respectively follows since $f(x_M)$ and $f(x_m)$ are indeed elements of E. $\square$

To speak on this result, we can update Figure 20. The continuous image of a closed, bounded interval is equal to a closed, bounded interval, namely $[f(x_m), f(x_M)]$.

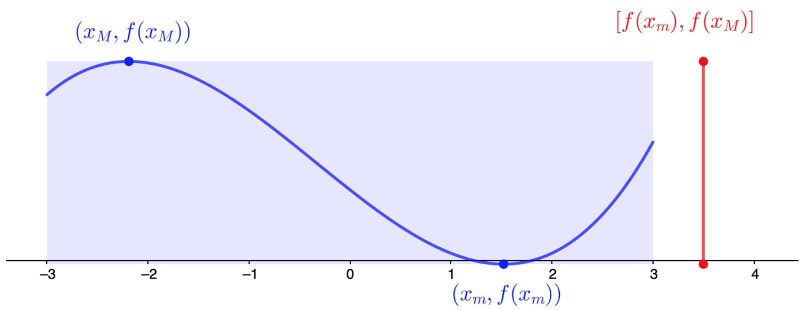


Figure 21: A visualization of the Extreme Value Theorem

6.2 Uniform Continuity

For the final subsection of this chapter, we introduce a stronger notion of continuity. To prove $f : D \to \mathbb{R}$ is continuous on D, the proof following Definition 6.1 would begin with an arbitrary $x \in D$ and an arbitrary $\epsilon > 0$. Thus, the choice of $\delta > 0$ may depend on both ϵ and x, the location at which we are examining the continuity of f. This stronger notion of continuity, uniform continuity, requires the $\delta > 0$ to only depend on ϵ. In other words, the $\delta > 0$ must work for all $x \in D$.

Definition 6.14: Uniform Continuity

We say that $f : D \to \mathbb{R}$ is **uniformly continuous on** D if and only if for all $\epsilon > 0$ there exists a $\delta > 0$ such that for all $x, y \in D$,

$$|x - y| < \delta \quad \text{implies} \quad |f(x) - f(y)| < \epsilon$$

We expand on this definition. Although it does seem quite similar to Definition 6.1, this definition requires something more from a function. In Definition 6.1, we are only concerned with continuity at a point $x \in D$. Thus, the δ chosen may depend on ϵ and x: $\delta(\epsilon, x)$. However, for a function to be uniformly continuous on D, the δ value will depend only on ϵ and must work for all $x, y \in D$: $\delta(\epsilon)$.

With continuity, we have a technique from Theorem 6.7 to show that a function $f : D \to \mathbb{R}$ is not continuous at $c \in D$: if there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in D such that $\lim\limits_{n\to\infty} x_n = c$ but $\lim\limits_{n\to\infty} f(x_n) \neq f(c)$, then f is not continuous. There is an analogous result for showing f is not uniformly continuous on D. However, since uniform continuity is a more demanding property, so is this technique to show f is not uniformly continuous:

Lemma 6.15: Proving *not* Uniformly Continuous

Let $f : D \to \mathbb{R}$. Then f is not uniformly continuous if and only if there exists sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ of D and an $\epsilon_0 > 0$ such that $\lim\limits_{n\to\infty} |x_n - y_n| = 0$ but $|f(x_n) - f(y_n)| \geq \epsilon_0$ for all $n \in \mathbb{N}$.

Proof. We start with the converse direction, i.e. proving the negation of Definition 6.14. We know from out lemma hypothesis that there exists an $\epsilon_0 > 0$ such that $|f(x_n) - f(y_n)| \geq \epsilon_0$ for all $n \in \mathbb{N}$. We let $\delta > 0$. Since $\lim\limits_{n\to\infty} |x_n - y_n| = 0$, there exists an $N \in \mathbb{N}$ such that $|x_n - y_n| < \delta$ for all $n \geq N$. Therefore, we take out problematic $x, y \in D$ in the negation of Definition 6.14 to be x_N and y_N. Thus, $|x_N - y_N| < \delta$ but $|f(x_N) - f(y_N)| \geq \epsilon_0$. It follows that f is not uniformly continuous on D.

For the forward direction, we suppose that f is not uniformly continuous as in Definition 6.14. Thus, for every $n \in \mathbb{N}$, $\delta_n = \frac{1}{n} > 0$ so that there exists $x_n, y_n \in D$ such that $|x_n - y_n| < \frac{1}{n}$ but $|f(x_n) - f(y_n)| \geq \epsilon_0$. It follows by the Squeeze Theorem that for the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ of D, $\lim\limits_{n\to\infty} |x_n - y_n| = 0$. We already have that $|f(x_n) - f(y_n)| \geq \epsilon_0$ is satisfied for all $n \in \mathbb{N}$ and so our lemma hypothesis is satisfied. □

Usually, functions that are not uniformly continuous on their domains have a 'problematic' point. The trick to apply Lemma 6.15 is to create two distinct sequences that converge to said point but whose functions values stay separated. We illustrate this with a classic example. See the exercises for some more practice.

Example 6.16: $f(x) = \frac{1}{x}$ is not Uniformly Continuous

Consider the function $f(x) = \frac{1}{x}$ defined for all $x > 0$. Using Lemma 6.15, we prove that f is not uniformly continuous on $(0, \infty)$.

Proof. The function f gets steep around 0 so we consider the two sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ of $(0, \infty)$ given by $x_n = \frac{1}{n}$ and $y_n = \frac{1}{n+1}$ for all $n \in \mathbb{N}$. We have that for all $n \in \mathbb{N}$,

$$0 \leq |x_n - y_n| = \left|\tfrac{1}{n} - \tfrac{1}{n+1}\right| = \left|\tfrac{1}{n(n+1)}\right| = \tfrac{1}{n^2+n} < \tfrac{1}{n}.$$

By Exercise 5 in the previous chapter and the Squeeze Theorem, $\lim\limits_{n\to\infty} |x_n - y_n| = 0$. However, we chose these sequences so that for all $n \in \mathbb{N}$,

$$|f(x_n) - f(y_n)| = |n - (n+1)| = 1.$$

Thus, with $\epsilon_0 = 1$, it follows that f is not continuous on $(0, \infty)$. If we restrict the domain of f to remove the problematic point 0, then f will be uniformly continuous. See the exercises of this section. □

The final result of this section has a similar feel to Theorem 5.19. Over closed, bounded intervals of $\mathbb{R}$ it turns out that continuity and uniform continuity are one and the same. From Exercise 8, we know that every uniformly continuous function is continuous. This additional hypothesis of a closed, bounded interval domain is enough to get the other direction:

Theorem 6.17: Continuity on a Closed Interval

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

We suppose by contradiction that f is not uniformly continuous. Then by Lemma 6.15, there exist sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ of $[a,b]$ and an $\epsilon_0 > 0$ such that $\lim_{n\to\infty} |x_n - y_n| = 0$ but $|f(x_n) - f(y_n)| \geq \epsilon_0$ for all $n \in \mathbb{N}$. By the Bolzano-Weierstrass Theorem, there exists a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ with $\lim_{k\to\infty} x_{n_k} = c$. We claim that for the same sub-indices, $\lim_{k\to\infty} y_{n_k} = c$ as well. Let $\epsilon > 0$. There exists $K', N \in \mathbb{N}$ such that $|x_{n_k} - c| < \frac{\epsilon}{2}$ for all $k \geq K'$ and $|x_n - y_n| < \frac{\epsilon}{2}$ for all $n \geq N$. Thus, with $K := \max\{K', N\}$, we have that for all $k \geq K$,

$$|y_{n_k} - c| \leq |y_{n_k} - x_{n_k}| + |x_{n_k} - c| < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon.$$

Since the absolute value function is continuous and f is continuous, it follows that

$$\lim_{k\to\infty} |f(x_{n_k}) - f(y_{n_k})| = |f(x) - f(x)| \geq \epsilon_0 > 0.$$

Of course, this is our contradiction, since $|f(x) - f(x)| = 0$.

6.3 Exercises

1. Show that both any constant function $f(x) = c$ for $c \in \mathbb{R}$ and the identity function given $f(x) = x$ for all $x \in \mathbb{R}$ are continuous, using Definition 6.1.

2. (Thomae's function) We define the function (see Figure 22) $f : [0,1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} \frac{1}{q} & x = \frac{p}{q} \text{ in lowest terms, } p \in \mathbb{Z},\ q \in \mathbb{N} \\ 0 & x \text{ irrational} \end{cases}$$

   Show that f is continuous at every irrational number and is not continuous at every rational number.

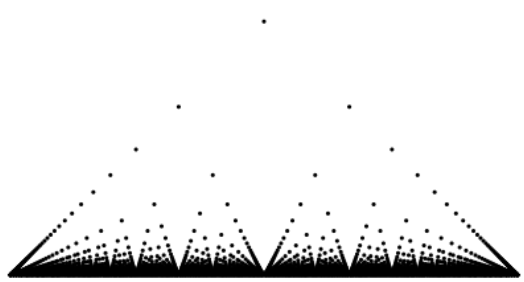


Figure 22: Thomae's function

3. a) Show that $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^n$ is continuous for every $n \in \mathbb{N}$.

   b) Prove that any polynomial $P(x) = \sum_{k=0}^{n} c_k x^k$ is continuous as a function on $\mathbb{R}$.

c) Prove that any rational function

$$g(x) = \frac{\sum_{k=0}^{n} c_k x^k}{\sum_{k=0}^{m} d_k x^k}$$

is continuous as a function on $\mathbb{R} \setminus \{x \in \mathbb{R} : \sum_{k=0}^{m} d_k x^k = 0\}$.

4. a) Show that $x \mapsto \sin(x) : \mathbb{R} \to \mathbb{R}$ is continuous.

   b) Show that $x \mapsto \cos(x) : \mathbb{R} \to \mathbb{R}$ is continuous.

   c) Show that $x \mapsto \tan(x)$ is continuous on the domain $\mathbb{R} \setminus \left\{ \frac{(2k+1)\pi}{2} : k \in \mathbb{Z} \right\}$.

   d) Deduce that all other trigonometric functions $\csc(x), \sec(x), \cot(x)$ are continuous on their respective domains.

5. Consider the function $f : [0, \frac{2}{\pi}] \to \mathbb{R}]$ given by

$$f(x) = \begin{cases} \sin(\frac{1}{x}) & x \in (0, \frac{1}{\pi}] \\ -1 & x = 0 \end{cases}$$

(see Figure 23). Show that f is not continuous at 0 but does satisfy the Intermediate Value Property.

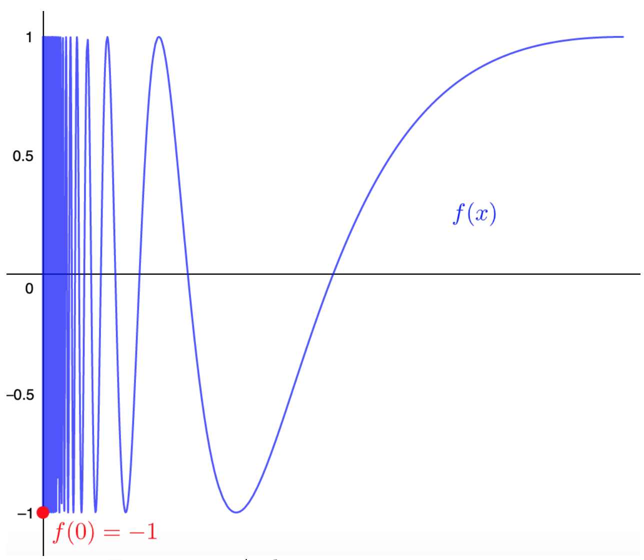


Figure 23: A discontinuous sine wave

6. Extend Lemma 6.11 by proving the result in the case that $f(a) > 0 > f(b)$.

7. a) Show that $\cos(x) = x$ has a solution in the interval $[0, \frac{\pi}{2}]$.

b) Show that $x^5 - x^4 + x^3 - x^2 + x$ has a root in the interval $[-1, 1]$.

c) Show that any polynomial of odd degree has at least one real root.

8. Using Definitions 6.1 and 6.14, show that every function $f : D \to \mathbb{R}$ that is uniformly continuous on D is continuous on D.

9. As stated in Example 6.16, show that $f(x) = \frac{1}{x}$ defined for all $x > 1$ is uniformly continuous.

10. Using Lemma 6.15, show that the $f(x) = \sin(\frac{1}{x})$ defined on $(0, \infty)$ is continuous but not uniformly continuous.

11. a) Show that if $f : A \to B$ is uniformly continuous on A and $g : B \to \mathbb{R}$ is uniformly continuous on B, then $g \circ f : A \to \mathbb{R}$ is uniformly continuous on A.

    b Suppose that $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ are uniformly continuous on D. Prove that $\alpha f + \beta g : D \to \mathbb{R}$ is uniformly continuous on D but that fg and $\frac{f}{g}$ are not in general.

7 Differentiation

The last three chapters of this supplement will cover some of the classic notions of calculus: differentiation and integration, and explore the relationship between them. To begin with differentiation, we consider a function $f : D \to \mathbb{R}$, with $D \subseteq \mathbb{R}$. For c a limit point of D, we will define the derivative of f at c, $f'(c)$, much like in an ordinary calculus class.

Definition 7.1: Differentiability at a Point

Let $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}$ and $c \in D$, with c a limit point of D. We say that f is **differentiable at** c if and only if the limit

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

exists. In this case, we call this limit $f'(c)$ or $\frac{df}{dx}|_{x=c}$.

Right off the bat, we may introduce a result from the previous to deal with sequences instead of function limits:

Corollary 7.2: Sequential Characterization of a Derivative

A function $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}$ is differentiable at a limit point c of D if and only if for every sequence $(x_n)_{n\in\mathbb{N}}$ of D with $\lim_{n\to\infty} x_n = c$ and $x_n \neq c$ for all $n \in \mathbb{N}$, it follows that

$$\lim_{n\to\infty} \frac{f(x_n) - f(c)}{x_n - c}$$

exists.

Proof. This is a direct application of Lemma 6.5 with $g(x) = \frac{f(x)-f(c)}{x-c}$. Then $\lim_{x\to c} g(x) = f'(c)$ exists if and only if for all such sequences,

$$\lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} \frac{f(x_n) - f(c)}{x_n - c} = f'(c)$$

exists. □

This leads to our first example of a function that is not differentiable at a point: the absolute value function.

Example 7.3: $|x|$ not Differentiable at 0

The absolute value function is not differentiable at 0.

Proof. The trick here is to use Corollary 7.2. If $f(x) = |x|$ were differentiable at 0 then for every sequence $(x_n)_{n\in\mathbb{N}}$ of nonzero points converging to 0, we would have that $\lim_{n\to\infty} \frac{|x_n|-|0|}{x_n-0} = f'(0)$ since the limit exists and is unique. However, consider the two sequences $x_n = \frac{1}{n}$ and $y_n = -\frac{1}{n}$ for all $n \in \mathbb{N}$. We see that

$$\lim_{n\to\infty} \frac{|x_n| - |0|}{x_n - 0} = \lim_{n\to\infty} \frac{\frac{1}{n}}{\frac{1}{n}} = 1,$$

but

$$\lim_{n\to\infty} \frac{|y_n| - |0|}{y_n - 0} = \lim_{n\to\infty} \frac{\frac{1}{n}}{-\frac{1}{n}} = -1.$$

Therefore, as these two limits differ, $|x|$ cannot be differentiable at 0. □

One of the first results linking differentiability and continuity is that being differentiable at a point is stronger than being continuous at a point:

Theorem 7.4: Differentiable Implies Continuous

If $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}$, is differentiable at $c \in D$, c a limit point of D, then f is continuous at c.

Proof. We show that $\lim_{x \to c} f(x) = f(c)$. Since f is differentiable at c, $\lim_{x \to c} \frac{f(x)-f(c)}{x-c}$ exists. By the trick of multiplying by $1 = \frac{x-c}{x-c}$,

$$\begin{aligned}
\lim_{x \to c}(f(x) - f(c)) &= \lim_{x \to c}(x-c)\frac{f(x)-f(c)}{x-c} \\
&= \lim_{x \to c}(x-c) \cdot \lim_{x \to c}\frac{f(x)-f(c)}{x-c} \\
&= 0 \cdot f'(c) = 0.
\end{aligned}$$

Since $\lim_{x \to c}(f(x) - f(c)) = \lim_{x \to c} f(x) - f(c) = 0$, it follows that $\lim_{x \to c} f(x) = f(c)$ and that f is continuous at c. $\square$

Much like with continuous functions in Theorem 6.8, there is an algebra for functions differentiable at a point c. However, the formulas for the derivative at a point c are not as simple.

Theorem 7.5: Algebra of Differentiable Functions

Suppose $f, g : D \to \mathbb{R}$ are both differentiable at a limit point $c \in D \subseteq \mathbb{R}$. Then

1. $f + g$ is differentiable at c with $(f+g)'(c) = f'(c) + g'(c)$,
2. fg is differentiable at c with $(fg)'(c) = f(c)g'(c) + f'(c)g(c)$,
3. if $g(x) \neq 0$ for all $x \in D$, then $\frac{f}{g}$ is differentiable at c with

$$\left(\frac{f}{g}\right)'(c) = \frac{g(c)f'(c) - f(c)g'(c)}{(g(c))^2}.$$

Proof. This proof will feel much like the proof of Theorem 5.6, with trickier algebra as we go through the different function operations (multiplication and division).

1. To show $f + g$ is differentiable at c, we note that

$$\begin{aligned}
\lim_{x \to c}\frac{(f+g)(x) - (f+g)(c)}{x-c} &= \lim_{x \to c}\frac{f(x) + g(x) - (f(c) + g(c))}{x-c} \\
&= \lim_{x \to c}\left(\frac{f(x)-f(c)}{x-c} + \frac{g(x)-g(c)}{x-c}\right) \\
&= \lim_{x \to c}\frac{f(x)-f(c)}{x-c} + \lim_{x \to c}\frac{g(x)-g(c)}{x-c} = f'(c) + g'(c)
\end{aligned}$$

We were able to split up the limit in the final line by Lemma 6.6 and since $f'(c)$ and $g'(c)$ both exist. So this limit involving $f + g$ exists and we have a formula for $(f+g)'(c)$.

2. For fg, we have that

$$\begin{aligned}\lim_{x\to c}\frac{(fg)(x)-(fg)(c)}{x-c} &= \lim_{x\to c}\frac{f(x)g(x)-f(c)g(c)}{x-c}\\ &= \lim_{x\to c}\frac{f(x)g(x)-f(x)g(c)+f(x)g(c)-f(c)g(c)}{x-c}\\ &= \lim_{x\to c}\left(f(x)\frac{g(x)-g(c)}{x-c}+g(c)\frac{f(x)-f(c)}{x-c}\right)\\ &= \lim_{x\to c}f(x)\lim_{x\to c}\frac{g(x)-g(c)}{x-c}+g(c)\lim_{x\to c}\frac{f(x)-f(c)}{x-c}\\ &= f(c)g'(c)+f'(c)g(c).\end{aligned}$$

   Again, the limit exists through algebraic manipulation and limit algebra. For the last line, since f is differentiable and hence continuous at c, $\lim\limits_{x\to c} f(x) = f(c)$.

3. For $\frac{f}{g}$, we compute:

$$\begin{aligned}\lim_{x\to c}\frac{(\frac{f}{g})(x)-(\frac{f}{g})(c)}{x-c} &= \lim_{x\to c}\frac{\frac{f(x)}{g(x)}-\frac{f(c)}{g(c)}}{x-c}\\ &= \lim_{x\to c}\frac{f(x)g(c)-f(c)g(x)}{g(x)g(c)(x-c)}\\ &= \lim_{x\to c}\frac{f(x)g(c)-f(x)g(x)+f(x)g(x)-f(c)g(x)}{g(x)g(c)(x-c)}\\ &= \lim_{x\to c}\left(-\frac{f(x)}{g(x)g(c)}\frac{g(x)-g(c)}{x-c}+\frac{g(x)}{g(x)g(c)}\frac{f(x)-f(c)}{x-c}\right)\\ &= -\frac{f(c)g'(c)}{(g(c))^2}+\frac{f'(c)}{g(c)}\\ &= \frac{g(c)f'(c)-f(c)g'(c)}{(g(c))^2}.\end{aligned}$$

□

In addition to this algebra of differentiable functions, there is an analog to Theorem 6.9: the Chain Rule. The composition of two differentiable functions is also differentiable, with its own formula as well.

Theorem 7.6: The Chain Rule

Let $f : D_1 \to \mathbb{R}$ and $g : D_2 \to \mathbb{R}$ with $f(D_1) \subseteq D_2$ ($g(f(x))$ is defined for all $x \in D_1$). If f is differentiable at $c \in D_1$, c a limit point of D_1 and g is differentiable at $f(c)$, with $f(c)$ a limit point of D_2, then $g \circ f : D_1 \to \mathbb{R}$ is differentiable at c and

$$(g \circ f)'(c) = g'(f(c))f'(c).$$

Proof. To show $g \circ f$ is differentiable at c, we first define an auxiliary function $h : D_2 \to \mathbb{R}$ given by

$$h(y) = \begin{cases} \frac{g(y)-g(f(c))}{y-f(c)} & y \neq f(c) \\ g'(f(c)) & y = f(c) \end{cases}.$$

We notice this about h: since g is differentiable at $f(c)$, $\lim_{y \to f(c)} h(y) = g'(f(c)) = h(f(c))$ so that h is continuous at $f(c)$. Thus, since f is continuous at c, $h \circ f$ is continuous at c, by Theorem 6.9. Therefore, with $y = f(x)$ in the definition of h,

$$\begin{aligned} \lim_{x \to c} \frac{g(f(x)) - g(f(c))}{x - c} &= \lim_{x \to c} \frac{h(f(x))(f(x) - f(c))}{x - c} \\ &= \lim_{x \to c} h(f(x)) \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \\ &= h(f(c))f'(c) = g'(f(c))f'(c), \end{aligned}$$

just as desired. □

7.1 The Mean Value Theorem

With the previous two theorems, we have the tools to construct continuous and differentiable functions from the sums, products, and compositions of functions we know to be continuous and differentiable. For this subsection, we focus on major results if we are given differentiable function to begin with. The first result is a deep theorem that makes sense graphically and has a more or less simple proof. We first introduce a definition to help us describe special points of a function rigorously.

Definition 7.7: Local Extrema

Let $f : (a, b) \to \mathbb{R}$. We say x_0 is a **local maximum** (resp. **minimum**) of f if and only if there exists a $\delta > 0$ such that for all $x \in (x_0 - \delta, x_0 + \delta)$,

$$f(x_0) \geq f(x) \quad (f(x_0) \leq f(x)).$$

In general, a **local extremum** of f refers to either a local maximum or minimum of f.

These local extrema are called local for a reason. They are not 'global' maxima or minima, as in the Extreme Value Theorem, where $f(x_M) \geq f(x)$ for **all** $x \in [a, b]$, for example. In Figure 24, x_0 and x_1 are local extrema, but as you can see at the left and right ends of the function, there are other points that reach higher and lower than $f(x_0)$ and $f(x_1)$, respectively. With this notion of local maxima and minima of function, we prove

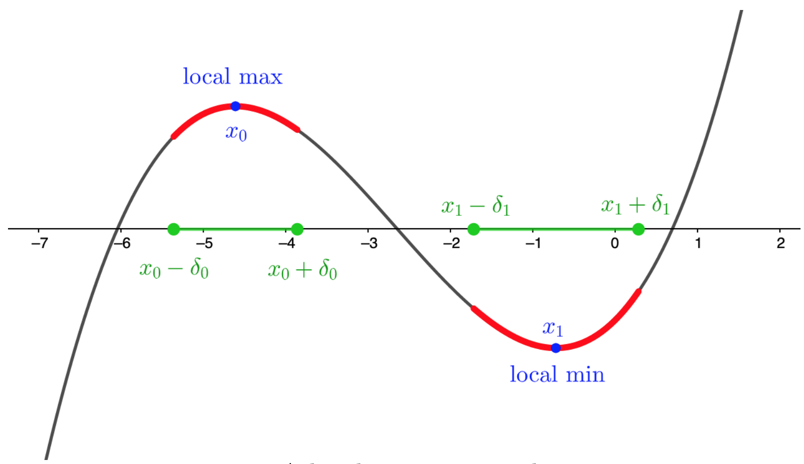


Figure 24: A local maximum and minimum

a lemma about the derivative at local extrema that will be familiar from previous calculus classes:

Lemma 7.8: Local Extrema Have Zero Derivative

Suppose $f : (a, b) \to \mathbb{R}$ is differentiable on (a, b) and $x_0 \in (a, b)$ is a local extremum of f. Then $f'(x_0) = 0$.

Proof. We prove the case where x_0 is a local maximum. The work for a local minimum is analogous, but with a minus sign thrown in. Since x_0 is a local maximum, there exists a $\delta_0 > 0$ such that $f(x) \leq f(x_0)$ for all $x \in (x_0 - \delta_0, x_0 + \delta_0)$. It might be the case that this interval $(x_0 - \delta_0, x_0 + \delta_0)$ is not a subset of the domain (a, b). If so, we simply define a new $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq (a, b) \cap (x_0 - \delta_0, x_0 + \delta_0)$. For example, $\delta := \min\{x_0 - a, b - x_0, \delta_0\} > 0$ would work. We construct a sequence in $(x_0, x_0 + \delta)$ and one in $(x_0 - \delta, x_0)$ as follows:

$$x_n = x_0 - \frac{\delta}{2^n} \quad \text{and} \quad y_n = x_0 + \frac{\delta}{2^n}, \text{ for all } n \in \mathbb{N}.$$

We note that by construction, $\lim_{n\to\infty} x_n = x_0 = \lim_{n\to\infty} y_n$ and that $x_n \neq x_0 \neq y_n$ for all $n \in \mathbb{N}$. Since f is differentiable at x_0,

$$\lim_{n\to\infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} = f'(x_0) = \lim_{n\to\infty} \frac{f(y_n) - f(x_0)}{y_n - x_0},$$

by Corollary 7.2. The key here is that $f(x_n) - f(x_0) \leq 0$ and $f(y_n) - f(x_0) \leq 0$ for all $n \in \mathbb{N}$ since both $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are sequences of $(x_0 - \delta, x_0 + \delta)$. Finally, by construction,

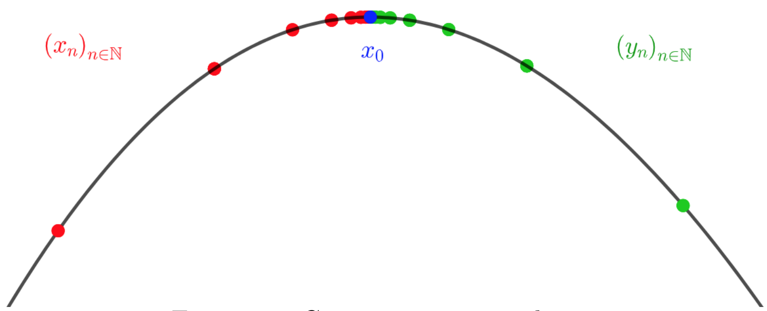


Figure 25: Constructing x_n and y_n

$x_n - x_0 = -\frac{\delta}{2^n} < 0$ and $y_n - x_0 = \frac{\delta}{2^n} > 0$ for all $n \in \mathbb{N}$. Thus,

$$\frac{f(x_n) - f(x_0)}{x_n - x_0} \geq 0 \quad \text{and} \quad \frac{f(y_n) - f(x_0)}{y_n - x_0} \leq 0$$

for all $n \in \mathbb{N}$. We may now conclude that

$$0 \leq \lim_{n\to\infty} \frac{f(y_n) - f(x_0)}{y_n - x_0} = f'(x_0) = \lim_{n\to\infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} \leq 0,$$

which implies that $f'(x_0) = 0$, as desired. □

The next step to the Mean Value Theorem is Rolle's Theorem, attributed to French mathematician Michel Rolle in 1691. Moving from Lemma 7.8, all we need to do is assume that f is continuous on $[a, b]$ to guarantee the existence of a point c such that $f'(c) = 0$.

Theorem 7.9: Rolle's Theorem

Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b) = 0$, then there exists a point $c \in (a, b)$ such that $f'(c) = 0$.

Proof. We first deal with the case when $f(x) = 0$ for all $x \in [a, b]$. It follows that $f'(x) = 0$ for all $x \in [a, b]$, so pick $c = \frac{b+a}{2} \in (a, b)$. It follows that $f'(c) = 0$.

For the remainder of the proof, we assume that f is not identically 0. By the Extreme Value Theorem, there exists $x_m, x_M \in [a, b]$ such that $f(x_m) \leq f(x) \leq f(x_M)$ for all $x \in [a, b]$. Since f is nonzero somewhere, one of x_m, x_M is a member of (a, b), which we will call x_0. This extremum is a global extremum and hence a local extremum. It follows that $f'(x_0) = 0$ and our proof is complete. □

The Mean Value Theorem, perhaps more famous, is nothing more than Rolle's Theorem with a linear translation. However, the Mean Value Theorem has a special geometric interpretation. For a function f continuous on $[a, b]$ and differentiable on (a, b), there exists a point $c \in (a, b)$ such that the slope of the secant line connecting the endpoints equals the slope, or derivative at c, $f'(c)$.

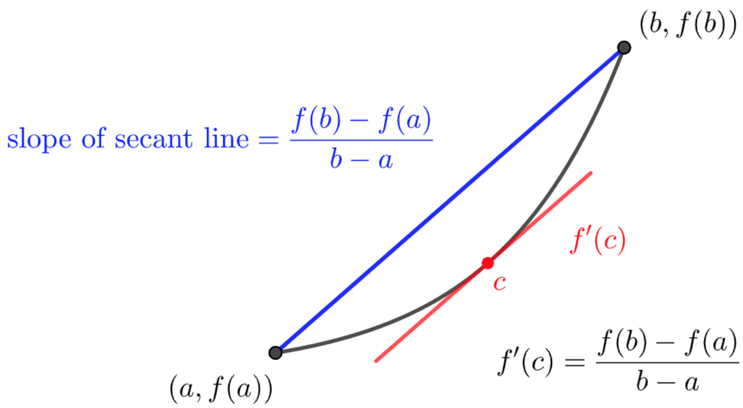


Figure 26: The geometry of the Mean Value Theorem

Theorem 7.10: The Mean Value Theorem

Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Then there exists a point $c \in (a, b)$ such that

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

Proof. In order to use Rolle's Theorem, we need a continuous and differentiable function such that $g(a) = g(b) = 0$. After some tinkering around with f we define $g : [a, b] \to \mathbb{R}$ by

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

By the algebra of continuous and differentiable functions, g is also continuous on $[a, b]$ and differentiable on (a, b). By construction,

$$g(a) = f(a) - f(a) - \frac{f(b) - f(a)}{b - a}(a - a) = 0$$

and

$$g(b) = f(b) - f(a) - \frac{f(b) - f(a)}{b - a}(b - a) = 0.$$

Therefore, by Rolle's Theorem, there exists a $c \in (a, b)$ such that $g'(c) = 0$. Computing the derivative of g in terms of f, we have that

$$g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a} = 0, \text{ or rather, } f'(c) = \frac{f(b) - f(a)}{b - a}.$$

□

Another Mean Value Theorem, attributed to Augustin-Louis Cauchy, is similar to the Mean Value Theorem, but adds another continuous, differentiable function into the mix:

Theorem 7.11: The Cauchy Mean Value Theorem

Suppose $f, g : [a, b] \to \mathbb{R}$ are continuous on $[a, b]$ and differentiable on (a, b). If $g'(x) \neq 0$ for all $x \in (a, b)$, then there exists a $c \in (a, b)$ such that

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

Proof. As in the Mean Value Theorem, we define another function $h : [a, b] \to \mathbb{R}$ given by

$$h(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)} g(x)$$

for all $x \in [a, b]$. How do we know that we can divide by $g(b) - g(a)$? Well, if $g(b) - g(a) = 0$ or rather $g(a) = g(b)$, then by Rolle's Theorem, there would exist a $c \in (a, b)$ such that $g'(c) = 0$, contradicting our hypothesis on $g'(x)$. Thus, $g(a) \neq g(b)$. We know that h is continuous on $[a, b]$ and differentiable on (a, b) by the algebra of continuous and differentiable functions, with derivative

$$h'(x) = f'(x) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(x)$$

for all $x \in (a, b)$. We know note that

$$h(a) = f(a) - \frac{f(b) - f(a)}{g(b) - g(a)} g(a) = \frac{f(a)g(b) - f(b)g(a)}{g(b) - g(a)}$$

and

$$h(b) = f(b) - \frac{f(b) - f(a)}{g(b) - g(a)} g(b) = \frac{f(a)g(b) - f(b)g(a)}{g(b) - g(a)}$$

so that $h(a) = h(b)$. By Rolle's Theorem, there exists a $c \in (a, b)$ such that $h'(c) = 0$, or rather, that

$$f'(c) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(c) = 0, \text{ or rather, } \frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

□

This quotient of function endpoints being equal to the quotient of function derivatives is exactly what we need to prove l'Hôpital's Rule, attributed to French mathematician Guillaume de l'Hôpital and first introduced by Swiss mathematician Johann Bernoulli. It is essential for evaluating limits of functions with 'indeterminate forms' such as $\frac{0}{0}$ or $\frac{\pm\infty}{\pm\infty}$.

Corollary 7.12: l'Hôpital's Rule

Suppose $f, g : [a, b] \to \mathbb{R}$ are continuous on $[a, b]$ and differentiable on (a, b). For $c \in [a, b]$, if

1. $g'(x) \neq 0$ for all $x \in (a, b)$,
2. $f(c) = g(c) = 0$, and
3. $\lim\limits_{x\to c} \frac{f'(x)}{g'(x)}$ exists,

then $\lim\limits_{x\to c} \frac{f(x)}{g(x)}$ exists and

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}$$

Proof. We begin this proof with a sequence $(x_n)_{n\in\mathbb{N}}$ of $[a, b]$ such that $x_n \neq c$ for all $n \in \mathbb{N}$ and $\lim\limits_{n\to\infty} x_n = c$. For every $n \in \mathbb{N}$, by the Cauchy Mean Value Theorem, there exists a c_n between x_n and c such that $\frac{f(x_n)-f(c)}{g(x_n)-g(c)} = \frac{f'(c_n)}{g'(c_n)}$. Actually, these c_n may be chosen as in Lemma 7.8. Since c_n is between x_n and c for all $n \in \mathbb{N}$, the sequence $(c_n)_{n\in\mathbb{N}}$ satisfies $c_n \neq c$ for all $n \in \mathbb{N}$ and $\lim\limits_{n\to\infty} c_n = c$ as well. By our hypothesis, $\lim\limits_{n\to\infty} \frac{f'(c_n)}{g'(c_n)}$ exists and thus $\lim\limits_{n\to\infty} \frac{f(x_n)-f(c)}{g(x_n)-g(c)} = \lim\limits_{n\to\infty} \frac{f(x_n)}{g(x_n)} = \lim\limits_{n\to\infty} \frac{f'(c_n)}{g'(c_n)}$. By Lemma 6.5, $\lim\limits_{x\to c} \frac{f(x)}{g(x)}$ exists with $\lim\limits_{x\to c} \frac{f(x)}{g(x)} = \lim\limits_{x\to c} \frac{f'(x)}{g'(x)}$. □

7.2 Exercises

1. Show that for a function $f : (a, b) \to \mathbb{R}$, differentiable at $c \in (a, b)$,

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

2. Show that any constant function $f(x) = \alpha$, $\alpha \in \mathbb{R}$, is differentiable on $\mathbb{R}$ with $f'(x) = 0$. Also, show that the identity function $g(x) = x$ is differentiable on $\mathbb{R}$ with $g'(x) = 1$.

3. a) Using the algebra of differentiable functions and induction, show that $f(x) = x^n$ is differentiable on $\mathbb{R}$ for all $n \in \mathbb{N}$ and $f'(x) = nx^{n-1}$.

   b) Prove the previous result for negative integers.

4. a) Show that $f(x) = \sin(x)$ is differentiable on $\mathbb{R}$ and $f'(x) = \cos(x)$.
   b) Show that $f(x) = \cos(x)$ is differentiable on $\mathbb{R}$ with $f'(x) = -\sin(x)$.
   c) Prove differentiability (over their respective domains), and calculate derivatives for the other four trigonometric functions: $\tan(x)$, $\csc(x)$, $\sec(x)$, and $\cot(x)$.
5. Consider the function $f(x) = x^2 \sin(\frac{1}{x})$ defined for all $x \neq 0$ with $f(0) = 0$. Show from Definition 7.1 that f is differentiable at 0 but that f' is not continuous at 0. This is a critical example of the fact that being differentiable at a point is much weaker than being **continuously differentiable**, i.e. f' is continuous.
6. Suppose that $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Show that
   a) if $|f'(x)| \leq M$ for all $x \in (a, b)$, then f is uniformly continuous on $[a, b]$;
   b) if $f'(x) \neq 0$ for all $x \in (a, b)$, then f is injective;
   c) if $f'(x) \geq 0$ (resp. $f'(x) \leq 0$) for all $x \in (a, b)$, then f is monotone increasing (decreasing);
   d) if $f'(x) = 0$ for all $x \in (a, b)$, then f is a constant function.
7. Suppose that $f, g : [a, b] \to \mathbb{R}$ are continuous on $[a, b]$, differentiable on (a, b), and satisfy $f'(x) = g'(x)$ for all $x \in (a, b)$. Show that $f(x) = g(x) + C$, for some $C \in \mathbb{R}$.

8 Integration

Unlike the previous chapter on differentiation, this chapter on integration will feature much more technical notation and many more definitions. Just to build up the theory of integration (Riemann or Darboux) requires substantial preparation. But, with figures along the way, the material will be presented in an organized and productive manner. Before discussing anything about functions, we define the notion of partitions of closed intervals on the real line.

Definition 8.1: Partitions

Let $[a, b]$ be a closed, bounded interval of $\mathbb{R}$.

1. A **partition** P **of** $[a, b]$ is a finite set of points $\{x_0, x_1, \ldots, x_n\}$ such that

$$a = x_0 < x_1 < \cdots < x_n = b.$$

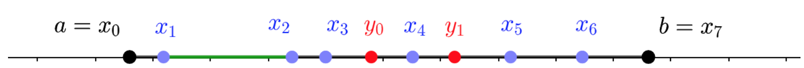
2. A partition Q of $[a, b]$ is said to be **finer than** a partition P of $[a, b]$ if and only if $P \subseteq Q$.

3. The **mesh** of a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ is defined by

$$\|P\| := \max_{i \in \{1,2,\ldots,n\}} \{x_i - x_{i-1}\}.$$

$P = \{x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ A parition of 8 points

$Q = \{x_0, x_1, x_2, x_3, y_0, x_4, y_1, x_5, x_6, x_7\} \supseteq P$ A finer parition



$||P|| = ||Q||$ The mesh of P and Q

Figure 27: Partitions visually

All of the previous definitions only have to deal with the x-axis, the independent variable. What we discuss next are functions. In calculus classes, the integral is usually described as the ‘area under the curve’. Before defining the integral, we start with the rectangles that are the building blocks for the definition. We give two definitions differing in how the heights of the rectangles are chosen.

Definition 8.2: Riemann Sums

Let $f : [a, b] \to \mathbb{R}$ be a bounded function and let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$. We define a **Riemann sum** of f and P as

$$S(P, f) = \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}), \text{ where } t_i \in [x_{i-1}, x_i].$$

Definition 8.3: Darboux Sums

Let $f : [a, b] \to \mathbb{R}$ be a bounded function and let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$. We define the **upper Darboux sum** of f and P as

$$U(P, f) = \sum_{i=1}^{n} M_i(f)(x_i - x_{i-1}), \text{ where } M_i(f) = \sup\{f(x) : x \in [x_{i-1}, x_i]\}.$$

Analogously, we define the **lower Darboux sum** of f and P as

$$L(P, f) = \sum_{i=1}^{n} m_i(f)(x_i - x_{i-1}), \text{ where } m_i(f) = \inf\{f(x) : x \in [x_{i-1}, x_i]\}.$$

These three sums may seem quite different, but essentially, the only thing that differentiates them is the 'height' chosen over each subinterval $[x_{i-1}, x_i]$. For Riemann sums, any point in each subinterval can be chosen and plugged into the function to get the height. For the upper and lower Darboux sums, the heights are fixed: you will either be using the supremum or infimum, respectively, over each function range on the subinterval. Since we assume our function f is bounded to begin with, $f([x_{i-1}, x_i])$ is bounded for each $i = 1, 2, \ldots, n$ and thus the supremum and infimum over such a set exist.

Before Definitions 8.2 and 8.3 were presented, we discussed rectangles as motivation for the approximation of an integral, which is exactly what is going on in each definition. The 'base' of each rectangle is the length of the subinterval $[x_{i-1}, x_i]$ and the height depends on the type of sum you are concerned with as described in the previous paragraph (see Figure 28). Thus, to get an approximation for the 'area under the curve', or the integral of f, you take the finite sum of the area of each rectangle, base $*$ height.

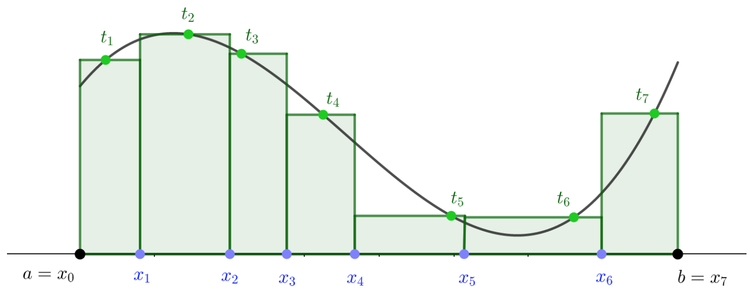
We know prove a lemma about the relation between upper and lower Darboux sums for partitions P and Q with $P \subseteq Q$. We also give an inequality relating Darboux sums and Riemann sums of a given partition P

Lemma 8.4: Comparing Sums

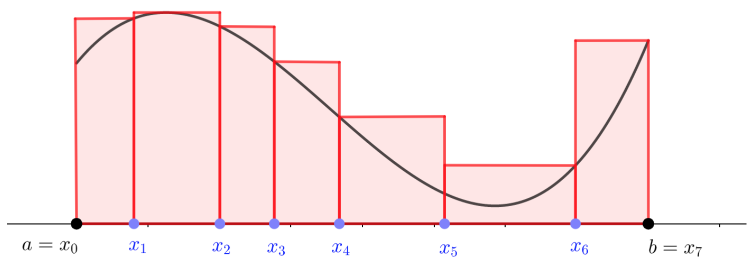
Let $f : [a, b] \to \mathbb{R}$ be bounded and P and Q be partitions of $[a, b]$ such that $P \subseteq Q$. Then

1. $L(P, f) \leq L(Q, f) \leq U(Q, f) \leq U(P, f)$ and
2. $L(P, f) \leq S(P, f) \leq U(P, f)$.

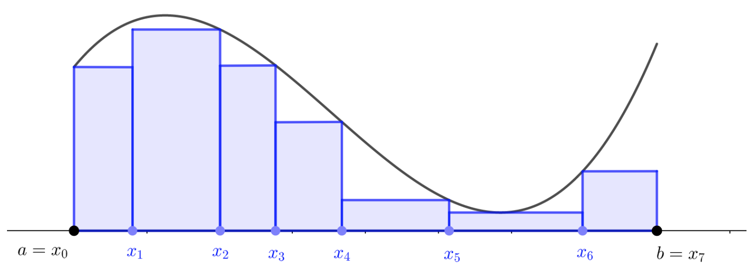
Proof.



(a) A Riemann sum



(b) An upper Darboux sum



(c) A lower Darboux sum

Figure 28: Some examples of rectangular sums for a given function and partition

1. If $P = Q = \{x_0, x_1, \ldots, x_n\}$, then the result follows almost immediately. We just need to show that $L(P, f) \leq U(P, f)$. Since $m_i(f) \leq M_i(f)$ for all $i \in \{1, \ldots, n\}$,

$$L(P, f) = \sum_{i=1}^{n} m_i(f)(x_i - x_{i-1}) \leq \sum_{i=1}^{n} M_i(f)(x_i - x_{i-1}) = U(P, f).$$

We now suppose that $P \subsetneq Q$. Since both P and Q are finite, Q can be written as P with a finite number of extra points thrown in. We prove our result in the case that Q only has one extra point c, say

$$P = \{x_0, x_1, \ldots, x_n\} \quad \text{and} \quad Q = \{x_0, c, x_1, \ldots, x_n\}.$$

If Q has more than one extra point, we just repeat the process finitely many times. Note that in the case of one extra point, the location of c does not matter; the subinterval $[x_{i-1}, x_i]$ that contains c is the only one affected in the summation. In this proof, we say that $c \in [x_0, x_1]$, the first subinterval. Outside of this subinterval, the summation over P and Q agree.

Over the first subinterval, we have that since both $[x_0, c]$ and $[c, x_1]$ are subsets of $[x_0, x_1]$, the images of these sets under f satisfy $f([x_0, c]) \subseteq f([x_0, x_1])$ and $f([c, x_1]) \subseteq f([x_0, x_1])$. Thus, by Example 3.16,

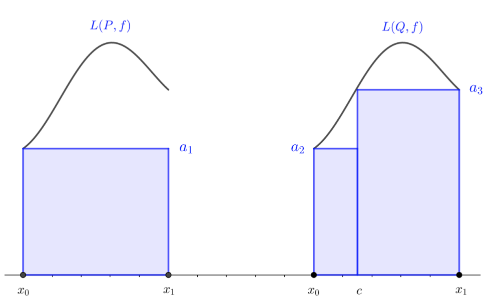
$$\inf(f([x_0, x_1])) \leq \inf(f([x_0, c])) \leq \sup(f([x_0, c])) \leq \sup(f([x_0, x_1])).$$

The same follows for $f([c, x_1])$ in the middle two inequalities as well. For some ease of notation we denote these infimum and supremum values by

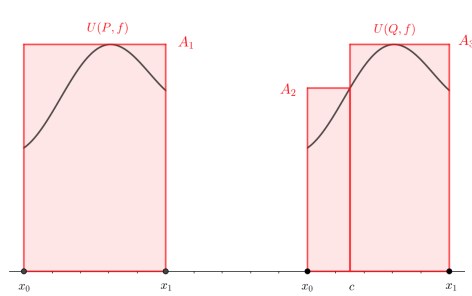
$$\begin{array}{lll} a_1 := \inf(f([x_0, x_1])) & a_2 := \inf(f([x_0, c])) & a_3 := \inf(f([c, x_1])) \\ A_1 := \sup(f([x_0, x_1])) & A_2 := \sup(f([x_0, c])) & A_3 := \sup(f([c, x_1])) \end{array}$$

(see Figure 29). The above chains of inequalities may thus be rewritten as

$$a_1 \leq a_2 \leq A_2 \leq A_1 \quad \text{and} \quad a_1 \leq a_3 \leq A_3 \leq A_1$$



(a) Lower Darboux sums



(b) Upper Darboux sums

Figure 29: Darboux sums over the split subinterval $[x_0, x_1]$

Now, for P, the area of the lower and upper rectangles over the first interval $[x_0, x_1]$ are $a_1(x_1 - x_0)$ and $A_1(x_1 - x_0)$, respectively. For Q, the first subinterval is split in two with lower area $a_2(c - x_0) + a_3(x_1 - c)$ and upper area $A_2(c - x_0) + A_3(x_1 - c)$. The key here is the chain of supremum and infimum inequalities from above. We know that

$$a_2(c - x_0) + a_3(x_1 - c) \geq a_1(c - x_0) + a_1(x_1 - c) = a_1(x_1 - x_0)$$

and

$$A_2(c - x_0) + A_3(x_1 - c) \leq A_1(c - x_0) + A_1(x_1 - c) = A_1(x_1 - x_0).$$

Therefore, over $[x_0, x_1]$, the lower area for P is less than or equal to the lower area for Q; also, the upper area for P is greater than or equal to the upper area for Q. Given that the other rectangles are equal over $[x_1, x_n]$, it follows that

$$L(P, f) \leq L(Q, f) \leq U(Q, f) \leq U(P, f).$$

2. We note that for the partition $P = \{x_0, x_1, \ldots, x_n\}$, $m_i(f) \leq f(t_i) \leq M_i(f)$, for all $t_i \in [x_{i-1}, x_i]$ and all $i = 1, 2, \ldots, n$. Therefore,

$$\sum_{i=1}^{n} m_i(f)(x_i - x_{i-1}) \leq \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) \leq \sum_{i=1}^{n} M_i(f)(x_i - x_{i-1}),$$

or rather,

$$L(P, f) \leq S(P, f) \leq U(P, f).$$

□

From this lemma, we can make a general claim about lower and upper Darboux sums over arbitrary partitions of $[a, b]$.

Corollary 8.5: Lower and Upper Darboux Sums in general

Let $f : [a, b] \to \mathbb{R}$ be bounded and P and Q be any two partitions of $[a, b]$. Then $L(P, f) \leq U(Q, f)$.

Proof. For two arbitrary partitions P and Q of $[a, b]$, $P \cup Q$ is a partition of $[a, b]$ that is finer than both P and Q. Thus, by the previous lemma,

$$L(P, f) \leq L(P \cup Q, f) \leq L(P \cup Q, f) \leq U(Q, f)$$

and we have our desired result. □

We now introduce some more machinery in the form of Riemann-integrable and Darboux-integrable functions and the Riemann and Darboux integrals. Each of these definitions include a 'limiting process' whether that be taking a supremum or taking a limit:

Definition 8.6: Riemann-Integrability

Let $f : [a, b] \to \mathbb{R}$ be bounded. We say f is **Riemann-integrable** if and only if the limit

$$\lim_{\|P\| \to 0} S(P, f)$$

exists. We say that f has **Riemann integral** equal to $I \in \mathbb{R}$ if and only if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for every partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ with $\|P\| < \delta$ and for any arbitrary choice of $t_i \in [x_{i-1}, x_i]$ (as per Definition 8.2),

$$|S(P, f) - I| < \epsilon.$$

Definition 8.7: Darboux-Integrability

Let $f : [a, b] \to \mathbb{R}$ be bounded. We define the **upper Darboux integral** of f as

$$\overline{\int_a^b} f := \inf\{U(P, f) : P \text{ a partition of } [a, b]\}.$$

Similarly, we define the **lower Darboux integral** of f as

$$\underline{\int_a^b} f = \sup\{L(P, f) : P \text{ a partition of } [a, b]\}.$$

Now, we say that f is **Darboux-integrable** if and only if $\overline{\int_a^b} f = \underline{\int_a^b} f$. In this case, we define the **Darboux integral** of f to be this common value and denote it by $\int_a^b f$.

These two definitions are clunky when it comes to showing a function is integrable. To show f is Riemann-integrable, we already need to have a guess at the integral that the limit in Definition 8.6 converges to. Also, dealing with this limit means defining sufficiently fine partitions (as $\|P\|$ approaches 0) to prove something with an arbitrary $\epsilon > 0$. In short, there are many quantifiers to deal with.

On the other hand, the Darboux definition tasks you with computing the infimum and supremum over all partitions P of $[a, b]$ (there are a lot to check...) and showing that these two values are equal.

We note here that the supremum and infimum defining the lower and upper Darboux integrals always exist. Let P be an arbitrary partition of $[a, b]$, thus satisfying $\{a, b\} \subseteq P$. Also, since f is bounded, there exists an $M \geq 0$ such that $|f(x)| \leq M$ for all $x \in [a, b]$.

From Lemma 8.4, we have that

$$-M(b-a) \leq L(\{a,b\},f) \leq L(P,f) \leq U(P,f) \leq U(\{a,b\},f) \leq M(b-a).$$

Therefore, the sets $\{U(P,f) : P \text{ a partition of } [a,b]\}$ and $\{L(P,f) : P \text{ a partition of } [a,b]\}$ are both bounded above and below (and nonempty) so that the supremum and infimum of each of them exist.

We now argue that the Darboux definition is going to be much more beneficial and we prove some results to help us eschew the Riemann definition altogether.

Lemma 8.8: Upper and Lower Darboux Integrals in general

Let $f : [a,b] \to \mathbb{R}$ be bounded. Then

$$\underline{\int_a^b} f \leq \overline{\int_a^b} f.$$

Proof. Let P and Q be two arbitrary partitions of $[a,b]$. From Corollary 8.5, $L(P,f) \leq U(Q,f)$. Thus, since $U(Q,f)$ does not depend on P, $U(Q,f)$ is an upper bound for all lower Darboux sums over $[a,b]$. Hence by the definition of supremum,

$$\underline{\int_a^b} f \leq U(Q,f).$$

Now, $\underline{\int_a^b} f$ is definitely not in terms of Q so that $\underline{\int_a^b} f$ is a lower bound for all upper Darboux sums over $[a,b]$. Therefore, by the definition of infimum

$$\underline{\int_a^b} f \leq \overline{\int_a^b} f,$$

as desired. $\square$

We now present an alternate characterization for proving Darboux-integrability. This alternate characterization is an invaluable tool for proving functions are Darboux-integrable. As will be seen, this definition only requires the existence of one partition P; there is no need to prove the equality suprema and infima or to check all partitions P with $\|P\| < \delta$ for some $\delta > 0$.

Theorem 8.9: Alternate Darboux-Integrability Characterization

Let $f : [a, b] \to \mathbb{R}$ be bounded. Then f is Darboux-integrable if and only if for all $\epsilon > 0$, there exists a partition P of $[a, b]$ such that

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

Proof. We start with the forward direction by supposing f is Darboux-integrable as in Definition 8.7 and let $\epsilon > 0$. By the definition of infimum, there exists a partition P_1 of $[a, b]$ such that $U(P_1, f) < \overline{\int_a^b} f + \frac{\epsilon}{2}$. In addition, there exists a partition P_2 of $[a, b]$ such that $L(P_2, f) > \underline{\int_a^b} f - \frac{\epsilon}{2}$, by the definition of supremum. Our partition for the proof of this direction will be $P := P_1 \cup P_2$ so that

$$L(P_2, f) \leq L(P, f) \leq U(P, f) \leq U(P_1, f).$$

Therefore, since $\underline{\int_a^b} f = \overline{\int_a^b} f$ by hypothesis,

$$\begin{aligned} U(P, f) - L(P, f) &\leq U(P_1, f) - L(P_2, f) \\ &< \overline{\int_a^b} f + \tfrac{\epsilon}{2} - \underline{\int_a^b} f + \tfrac{\epsilon}{2} = \epsilon. \end{aligned}$$

This is exactly the forward direction. For the converse direction, we show that $\overline{\int_a^b} f - \underline{\int_a^b} f < \epsilon$ for all $\epsilon > 0$, which would imply that $\underline{\int_a^b} f = \overline{\int_a^b} f$. Let $\epsilon > 0$. By the hypothesis of this direction there exists a partition P of $[a, b]$ such that $U(P, f) - L(P, f) < \epsilon$. By Definition 8.7 and Lemma 8.8, we know that

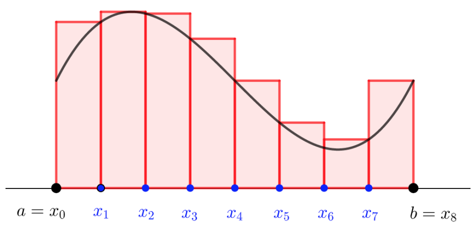
$$L(P, f) \leq \underline{\int_a^b} f \leq \overline{\int_a^b} f \leq U(P, f).$$

Therefore, we have that

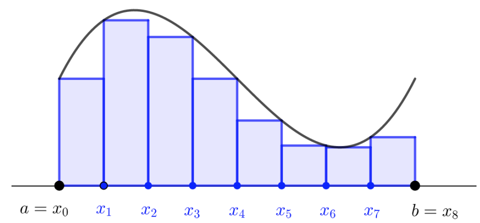
$$\overline{\int_a^b} f - \underline{\int_a^b} f \leq U(P, f) - L(P, f) < \epsilon.$$

The result follows for the converse direction and the equivalence is established. □

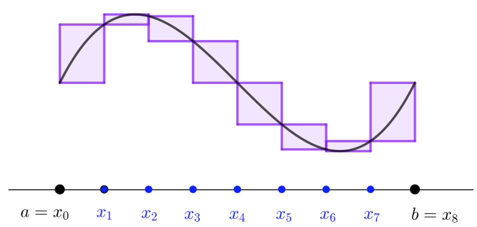
With this handy alternate characterization for Darboux-integrability, we no longer need to take suprema to show a function is Darboux-integrable. For an arbitrary $\epsilon > 0$, we need only find a partition such that the upper and lower Darboux sums with respect to that partition are less than ϵ apart. Here is a visual interpretation of this theorem, showing the lower and upper Darboux sums as before, along with the purple difference, which is what we are looking to be arbitrarily small:



(a) Upper Darboux sum of $f(x)$



(b) Lower Darboux sum of $f(x)$



(c) A visualization of $U(P, f) - L(P, f)$

Figure 30: An alternate Darboux-integrability characterization

In order for f to be Darboux-integrable, for any given $\epsilon > 0$, we must find a partition $P = \{x_0, \ldots, x_n\}$ such that the purple area above is less than ϵ. We give an example to show a function is Darboux-integrable using this alternate characterization:

Example 8.10: A Darboux-integrable Function

We define $f : [0, 1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1 & x \in [0, 1) \\ 0 & x = 1 \end{cases} .$$

Show that f is Darboux-integrable and that $\int_a^b f = 1$.

Proof. Let $\epsilon > 0$. We need to find a partition P of $[0, 1]$ such that $U(P, f) - L(P, f) < \epsilon$. The trick with these discontinuous functions is to partition the interval $[0, 1]$ in such a way so that the discontinuous point is contained in an arbitrarily small subinterval. For example, we will prove Darboux-integrability with the partition $P = \{0, 1 - \frac{\epsilon}{2}, 1\}$. Over $[0, 1 - \frac{\epsilon}{2}$, the supremum and infimum are both 1. Over $[1 - \frac{\epsilon}{2}, 1]$, though, the supremum is 1 and the infimum is 0. Now, we compute:

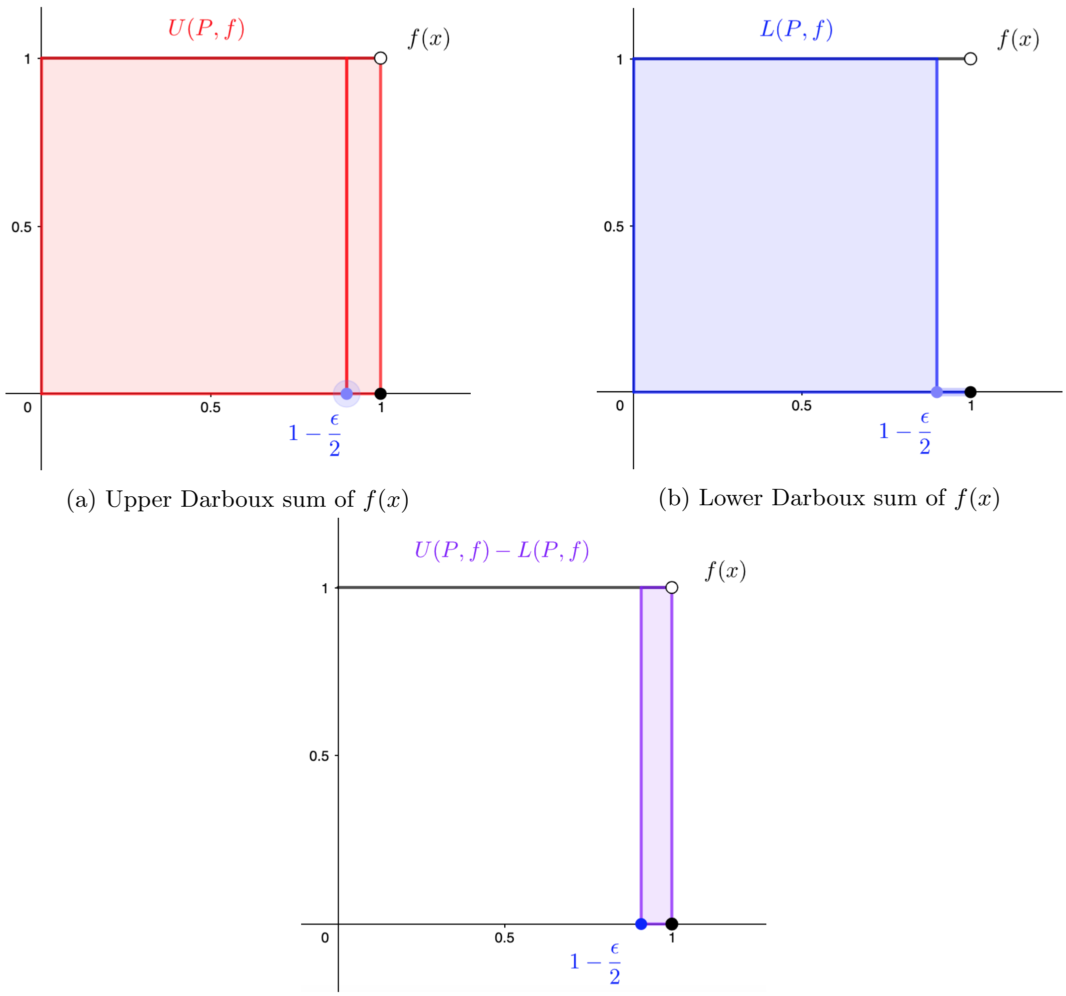
$$U(P, f) = 1 \cdot (1 - \tfrac{\epsilon}{2} - 0) + 1 \cdot (1 - (1 - \tfrac{\epsilon}{e})) = 1$$

and

$$L(P,f) = 1 \cdot (1 - \tfrac{\epsilon}{2} - 0) + 0 \cdot (1 - (1 - \tfrac{\epsilon}{e})) = 1 - \tfrac{\epsilon}{2}$$

so that

$$U(P,f) - L(P,f) = 1 - (1 - \tfrac{\epsilon}{2}) = \tfrac{\epsilon}{2} < \epsilon.$$



(a) Upper Darboux sum of $f(x)$

(b) Lower Darboux sum of $f(x)$

(c) Calculating $U(P,f) - L(P,f)$

Figure 31: Darboux sums of $f(x)$

It follows that f is Darboux-integrable. Now, to compute the integral, we note that

$$1 - \tfrac{\epsilon}{2} = L(P,f) \leq \int_a^b f \leq U(P,f) = 1,$$

as in the proof of Theorem 8.9. Thus, $-\epsilon < -\frac{\epsilon}{2} \leq \int_a^b f - 1 \leq 0 < \epsilon$. Therefore, $\left| \int_a^b f - 1 \right| < \epsilon$.

Since $\epsilon > 0$ was arbitrary to begin with, it follows that $\int_a^b f = 1$, as we wanted to show. $\square$

The major result of this subsection is showing that being Darboux-integrable is equivalent to being Riemann-integrable. In addition, when a function is integrable, the Riemann and Darboux integrals coincide.

Theorem 8.11: Darboux-integrable if and only if Riemann-integrable

A bounded function $f : [a, b] \to \mathbb{R}$ is Darboux-integrable if and only if it is Riemann-integrable. In this case, the Riemann and Darboux integrals as defined in Definitions 8.6 and 8.7, respectively, are equal.

Proof. In lieu of a rigorous proof that may go beyond the necessities of this course, we give a sketch of the argument for both directions.

For the forward direction, suppose that f is Darboux-integrable. For any partition P of $[a, b]$, we know that

$$L(P, f) \leq S(P, f) \leq U(P, f)$$

from Lemma 8.4. Thus, the Riemann sum is always sandwiched between the lower and upper Darboux sums. Taking the limit as $\|P\| \to 0$ tells us that $\lim\limits_{\|P\|\to 0} S(P, f)$ is sandwiched between $\lim\limits_{\|P\|\to 0} L(P, f)$ and $\lim\limits_{\|P\|\to 0} U(P, f)$. It can be shown (although the proof is a bit technical) that

$$\lim_{\|P\|\to 0} L(P, f) = \underline{\int_a^b} f \quad \text{and} \quad \lim_{\|P\|\to 0} U(P, f) = \overline{\int_a^b} f.$$

Therefore, by the same notion as the Squeeze Theorem, $\lim\limits_{\|P\|\to 0} S(P, f)$ exists and must be equal to the common value $\underline{\int_a^b} f = \overline{\int_a^b} f = \int_a^b f$ (since f is Darboux-integrable).

For the converse direction, suppose f is Riemann integrable and let $\epsilon > 0$. There exists a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ and two choices of $t_i^1, t_i^2 \in [x_{i-1}, x_i]$ (call these tagged partitions of $[a, b]$ P_1 and P_2, respectively) such that

1. $U(P, f) < S(P_1, f) + \frac{\epsilon}{6}$,
2. $L(P, f) > S(P_2, f) + \frac{\epsilon}{6}$, and
3. $|S(P_1, f) - S(P_2, f)| < \frac{2\epsilon}{3}$ (we get this since f is Riemann-integrable).

Therefore,

$$\begin{aligned} U(P, f) - L(P, f) &< S(P_1, f) - S(P_2, f) + \tfrac{\epsilon}{3} \\ &\leq |S(P_1, f) - S(P_2, f)| + \tfrac{\epsilon}{3} \\ &< \tfrac{2\epsilon}{3} + \tfrac{\epsilon}{3} = \epsilon. \end{aligned}$$

By Theorem 8.9, f is Darboux-integrable. Now, f is Darboux-integrable if and only if it is Riemann-integrable. Thus, the Riemann integral is always seen as the common value of the upper and lower Darboux integrals as shown above. $\square$

After this proof, we will only say that a function $f : [a, b] \to \mathbb{R}$ is **integrable** or that $f \in R([a, b])$. The integral of f will be denoted by $D(f) = R(f) = \int_a^b f$. For all intents and purposes, we will continue to use proof techniques related to Darboux-integrability given the utility of Theorem 8.9.

8.1 Classes of Integrable Functions

Now that we have rigorously defined the integral of f, we need to start building our catalog of functions that are integral. We start by proving that we can construct integrable functions from functions that we know to be integrable.

Theorem 8.12: Constructing Integrable Functions

Suppose that $f, g \in R([a, b])$ and $c \in \mathbb{R}$. Then

a) $cf \in R([a, b])$ with $\int_a^b cf = c\int_a^b f$, and

b) $f + g \in R([a, b])$ with $\int_a^b (f + g) = \int_a^b f + \int_a^b g$.

Proof.

a) We let $\epsilon > 0$ to make use of Theorem 8.9. If $c = 0$, then cf is the constant zero function. By a problem in the Exercises, $\int_a^b cf = \int_a^b 0 = 0 = c\int_a^b f$, as desired. Suppose that $c > 0$. Since $f \in R([a, b])$ there exists a partition P of $[a, b]$ such that $U(P, f) - L(P, f) < \frac{\epsilon}{c}$. We note that for the function cf and partition $P = \{x_0, x_1, \ldots, x_n\}$, $m_i(cf) = \sup\{cf(x) : x \in [x_{i-1}, x_i]\} = c\sup\{f(x) : x \in [x_{i-1}, x_i]\} = cm_i(f)$. The same holds for the $M_i(cf)$. Therefore,

$$\begin{aligned} U(P, cf) - L(P, cf) &= \sum_{i=1}^{n} (M_i(cf) - m_i(cf))(x_i - x_{i-1}) \\ &= c\sum_{i=1}^{n} (M_i(f) - m_i(f))(x_i - x_{i-1}) \\ &= c(U(P, f) - L(P, f)) < \epsilon. \end{aligned}$$

To show equality of the integrals, we notice that

$$\begin{aligned}\int_a^b cf &= \sup\{L(P, cf) : P \text{ a partition of } [a,b]\} \\ &= \sup\{cL(P, f) : P \text{ a partition of } [a,b]\} \\ &= c\sup\{L(P, f) : P \text{ a partition of } [a,b]\} \\ &= c\int_a^b f.\end{aligned}$$

We give a summary for the analogous proof of $c < 0$. For any $\epsilon > 0$, we get a partition P of $[a,b]$ such that $U(P, f) - L(P, f) < \frac{\epsilon}{-c}$. The trick here is that for every subinterval of P, $m_i(cf) = cM_i(f)$ and $M_i(cf) = cm_i(f)$ since $c < 0$. We can then compute $U(P, cf) - L(P, cf) = -c(U(P, f) - L(P, f)) < \epsilon$. The integral calculation will also go through two sign flips since c is negative. Ultimately, the end result is not affected:

$$\int_a^b cf = \sup\{L(P, cf)\} = \sup\{cU(P, f)\} = c\inf\{U(P, f)\} = c\int_a^b f.$$

b) We again let $\epsilon > 0$. Since f and g are integrable, there exists partitions P_f and P_g of $[a,b]$ such that

$$U(P_f, f) - L(P_f, f) < \tfrac{\epsilon}{2} \quad \text{and} \quad U(P_g, g) - L(P_g, g) < \tfrac{\epsilon}{2}.$$

We define the partition $P = P_f \cup P_g = \{x_0, x_1 \dots, x_n\}$, which also satisfies both of the above inequalities as a common refinement of P_f and P_g. Over every subinterval $[x_{i-1}, x_i]$ of P, we know that

$$M_i(f+g) \le M_i(f) + M_i(g) \quad \text{and} \quad m_i(f+g) \ge m_i(f) + m_i(g).$$

Therefore,

$$\begin{aligned}U(P, f+g) - L(P, f+g) &= \sum_{i=1}^{n}(M_i(f+g) - m_i(f+g))(x_i - x_{i-1}) \\ &\le \sum_{i=1}^{n}(M_i(f) - m_i(f) + M_i(g) - m_i(g))(x_i - x_{i-1}) \\ &= U(P, f) - L(P, f) + U(P, g) - L(P, g) \\ &< \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon.\end{aligned}$$

It follows that $f + g$ is Darboux integrable. Let P be any partition of $[a,b]$. We compute the Darboux integral in two ways. First,

$$\int_a^b (f+g) \le U(P, f+g) \le U(P,f) + U(P,g),$$

which implies that for the infimum over all partitions P of $[a,b]$

$$\int_a^b (f+g) \le \inf\{U(P,f)+U(P,g)\} = \inf\{U(P,f)\} + \inf\{U(P,g)\} = \int_a^b f + \int_a^b g.$$

Similarly,

$$\int_a^b (f+g) \ge L(P, f+g) \ge L(P,f) + L(P,g),$$

which implies that for the supremum over all partitions P of $[a,b]$

$$\int_a^b (f+g) \ge \sup\{L(P,f)+L(P,g)\} = \sup\{L(P,f)\} + \sup\{L(P,g)\} = \int_a^b f + \int_a^b g.$$

It follows that with both inequalities satisfied,

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g,$$

as desired. □

Sadly, there is no general formula for the integral of a product of integrable functions like there is for sums or scalar multiples of integrable functions. However, it can be proven that if $f, g \in R([a,b])$, then $fg \in R([a,b])$ as well: check the Exercises.

We now prove some order properties of the integral. If we know one integrable functions is greater than another of their domain, we argue that their integrals respect that order. Furthermore, we gain a result about the absolute value of an integrable function and the comparison between the resultant integrals.

Theorem 8.13: Order Properties of the Integral

Suppose $f, g \in R([a,b])$ with $f(x) \le g(x)$ for all $x \in [a,b]$. Then

a) $\int_a^b f \le \int_a^b g$, and

b) $|f| \in R([a,b])$ with $\left|\int_a^b f\right| \le \int_a^b |f|$.

Proof.

a) Let $P = \{x_0, x_1, \ldots, x_n\}$ be any partition of $[a,b]$. Since $f(x) \le g(x)$ for all $x \in [a,b]$, $m_i(f) \le m_i(g)$ and $M_i(f) \le M_i(g)$. These follow since for all $x \in [x_{i-1}, x_i]$, $m_i(f) \le$

$f(x) \leq g(x)$ and $f(x) \leq g(x) \leq M_i(g)$. Therefore,

$$L(P,f) = \sum_i^N m_i(f)(x_i - x_{i-1}) \leq \sum_i^N m_i(g)(x_i - x_{i-1}) = L(P,g).$$

The same holds for $U(P,f)$ and $U(P,g)$. Now, we see that

$$\begin{aligned}
\int_a^b f &= \sup\{L(P,f) : P \text{ a partition of } [a,b]\} \\
&\leq \sup\{L(P,g) : P \text{ a partition of } [a,b]\} \\
&= \int_a^b g.
\end{aligned}$$

b) To show $|f| \in R([a,b])$, let $\epsilon > 0$. There exists a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a,b]$ such that $U(P,f) - L(P,f) < \epsilon$. We prove that $M_i(f) - m_i(f) = \sup\{|f(x) - f(y)| : x, y \in [x_{i-1}, x_i]\}$. Since f is bounded, this supremum exists. We have that

$$\begin{aligned}
M_i(f) - m_i(f) &= \sup\{f(x) : x \in [x_{i-1}, x_i]\} - \inf\{f(y) : y \in [x_{i-1}, x_i]\} \\
&= \sup\{f(x) : x \in [x_{i-1}, x_i]\} + \sup\{-f(y) : y \in [x_{i-1}, x_i]\} \\
&= \sup\{f(x) - f(y) : x, y \in [x_{i-1}, x_i]\} \\
&\leq \sup\{|f(x) - f(y)| : x, y \in [x_{i-1}, x_i]\}.
\end{aligned}$$

For the other inequality, we notice that for any $x, y \in [x_{i-1}, x_i]$,

$$f(x) - f(y) \leq M_i(f) - m_i(f) \quad \text{and} \quad f(y) - f(x) \leq M_i(f) - m_i(f)$$

by symmetry so that $|f(x) - f(y)| \leq M_i(f) - m_i(f)$. Thus,

$$\sup\{|f(x) - f(y)| : x, y \in [x_{i-1}, x_i]\} \leq M_i(f) - m_i(f)$$

and equality holds. Therefore,

$$\begin{aligned}
M_i(|f|) - m_i(|f|) &= \sup\{||f(x)| - |f(y)|| : x, y \in [x_{i-1}, x_i]\} \\
&\leq \sup\{|f(x) - f(y)| : x, y \in [x_{i-1}, x_i]\} \\
&= M_i(f) - m_i(f)
\end{aligned}$$

by the Reverse Triangle Inequality. Finally,

$$\begin{aligned}U(P,|f|)-L(P,|f|)&=\sum_{i=1}^{n}(M_i(|f|)-m_i(|f|))(x_i-x_{i-1})\\&\le\sum_{i=1}^{n}(M_i(f)-m_i(f))(x_i-x_{i-1})\\&=U(P,f)-L(P,f)<\epsilon,\end{aligned}$$

so that $|f|\in R([a,b])$.

The trick to compare the integrals is to notice that $-|f(x)|\le f(x)\le |f(x)|$ for all $x\in[a,b]$. Thus, from the result above,

$$-\int_a^b |f| = \int_a^b (-|f|) \le \int_a^b f \le \int_a^b |f|.$$

It follows that $\left|\int_a^b f\right| \le \int_a^b |f|$.

□

To conclude this chapter, we prove that two types of functions will always be integrable. Continuous functions being integrable seems straightforward; the 'area under the curve' of a continuous function seems to be well defined since there are no 'jumps' in such a function. On the other hand, monotone (increasing or decreasing) does not make much sense at first. Just being monotone is not much of a restriction. However, both continuous functions and monotone functions will always be integrable.

Theorem 8.14: Monotone Implies Integrable

Suppose $f:[a,b]\to\mathbb{R}$ is monotone on $[a,b]$. Then $f\in R([a.b])$.

Proof. We consider the case when f is monotone increasing, i.e. $x\le y$ implies $f(x)\le f(y)$. We know that f is bounded since $f(a)\le f(x)\le f(b)$ for all $x\in[a,b]$. Let $\epsilon>0$. If $f(a)=f(b)$, then f is constant and thus integrable. So suppose $f(a)<f(b)$. There exists an $N\in\mathbb{N}$ such that $\frac{1}{N}<\frac{\epsilon}{(b-a)(f(b)-f(a))}$ and consider the partition $P=\{x_0,x_1,\ldots,x_N\}$ of $[a,b]$ given by

$$x_i=a+i\frac{b-a}{N},\ \text{for } i=0,1,\ldots,N.$$

Then $\|P\|=\frac{b-a}{N}<\frac{\epsilon}{f(b)-f(a)}$. We now discuss $m_i(f)$ and $M_i(f)$. Since f is increasing on $[a,b]$, we know that for every subinterval $[x_{i-1},x_i]$, $M_i(f)=f(x_i)$ and $m_i(f)=f(x_{i-1})$. This follows because $f(x_i)$ is an upper bound for $f(x)$ on $[x_{i-1},x_i]$ and is also an element of the set $\{f(x) : x \in [x_{i-1}, x_i]\}$. Thus, $f(x_i)$ is the maximum of the set and thus the supremum of the set. The same can be said for $f(x_{i-1})$ as the minimum and infimum. We now compute:

$$\begin{aligned} U(P,f) - L(P,f) &= \sum_{i=1}^{N} (M_i(f) - m_i(f)) \|P\| \\ &= \sum_{i=1}^{N} (f(x_i) - f(x_{i-1})) \|P\| \\ &= (f(b) - f(a)) \|P\| < \epsilon. \end{aligned}$$

Therefore, f is integrable. The case when f is monotone decreasing is analogous, but with a negative sign thrown in. □

Theorem 8.15: Continuous Implies Integrable

Suppose $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$. Then $f \in R([a.b])$.

Proof. We know many things about a continuous function f on a closed, bounded interval, as seen in Chapter 6. For starters, from the Extreme Value Theorem, there exists $x_m, x_M \in [a,b]$ such that $f(x_m) \leq f(x) \leq f(x_M)$, so that f is bounded.

Let $\epsilon > 0$. Since f is continuous on $[a,b]$, it is uniformly continuous, by Theorem 6.17. Thus, there exists a $\delta > 0$ such that $|f(x) - f(y)| < \frac{\epsilon}{b-a}$ whenever $|x - y| < \delta$. Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a,b]$ such that $\|P\| < \delta$. This partition P can be constructed as in the proof of the previous theorem, with subintervals of equal length $\frac{\delta}{2}$.

For every subinterval $[x_{i-1}, x_i]$ of P, we know that since f is continuous, there exists $x_i', y_i' \in [x_{i-1}, x_i]$ such that $f(x_i') = M_i(f)$ and $f(y_i') = m_i(f)$. This also follows by the Extreme Value Theorem. Note that since $x_i', y_i' \in [x_{i-1}, x_i]$, it follows that $|x_i' - y_i'| \leq x_i - x_{i-1} \leq \|P\| < \delta$. By our uniformly continuous condition, $|f(x_i') - f(y_i')| = f(x_i') - f(y_i') <$

$\frac{\epsilon}{b-a}$. We now compute:

$$\begin{aligned} U(P,f) - L(P,f) &= \sum_{i=1}^{n} (M_i(f) - m_i(f))(x_i - x_{i-1}) \\ &= \sum_{i=1}^{n} (f(x_i') - f(y_i'))(x_i - x_{i-1}) \\ &< \sum_{i=1}^{n} \frac{\epsilon}{b-a}(x_i - x_{i-1}) \\ &= \frac{\epsilon}{b-a} \sum_{i=1}^{n} (x_i - x_{i-1}) \\ &= \frac{\epsilon}{b-a} \cdot (b-a) = \epsilon. \end{aligned}$$

Therefore, f is integrable. □

8.2 Exercises

1. Show that the function $f : [0,2] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 1 & x \in [0,1) \\ 0 & x = 1 \\ 1 & x \in (1,2] \end{cases}$$

   is Darboux-integrable and that $\int_a^b f = 2$.

2. Show that the function $g : [0,1] \to \mathbb{R}$ given by

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

   is Darboux-integrable and that $\int_a^b f = 1$.

3. Prove that the function $f(x) : [0,1] \to \mathbb{R}$ given by

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

   is not Darboux-integrable.

4. Let $f : [a,b] \to \mathbb{R}$ be bounded and let $c \in \mathbb{R}$. Show that for any partition P of $[a,b]$,

   a) $U(P, f+c) = U(P,f) + U(P,c)$,

b) $L(P, f+c) = L(P, f) + L(P, c)$, and

c) $U(P, c) = L(P, c) = c(b-a)$.

5. Let $f : [a, b] \to \mathbb{R}$ be given bu $f(x) = c$ for all $x \in [a, b]$. Show that f is both Riemann and Darboux-integrable and that

$$\lim_{\|P\| \to 0} S(P, c) = \underline{\int_a^b} c = \overline{\int_a^b} c = \int_a^b c = c(b-a).$$

6. For a bounded function $f : [a, b] \to \mathbb{R}$ and any $c \in \mathbb{R}$, show that

$$\underline{\int_a^b} (f+c) = \underline{\int_a^b} f + c(b-a) \quad \text{and} \quad \overline{\int_a^b} (f+c) = \overline{\int_a^b} f + c(b-a).$$

7. The aim of this exercise is to prove that the product of integrable functions is integrable.

   a) Show that if $f : [a, b] \to \mathbb{R}$ is integrable, then f^2 is integrable.

   b) Suppose that both $f, g : [a, b] \to \mathbb{R}$ are integrable. Then fg is integrable.

8. Let $f : [a, b] \to \mathbb{R}$ be bounded and $c \in (a, b)$. Show that $f \in R([a, b])$ if and only if $f \in R([a, c])$ and $f \in R([c, b])$, with

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

9. Suppose that $f \in R([a, b])$ and that $f = g$ except at the finite number of points $\{x_1, x_2, \ldots, x_m\} \subseteq [a, b]$. Show that $g \in R([a, b])$ and that $\int_a^b f = \int_a^b g$.

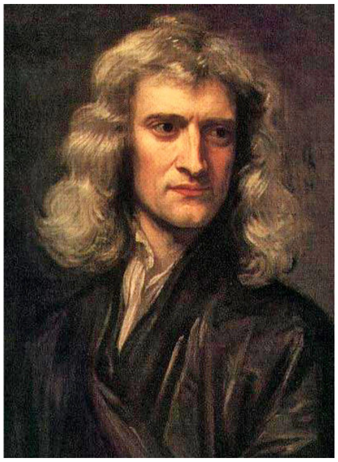
10. Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$. Show there exists a $c \in (a, b)$ such that

$$f(c) = \frac{1}{b-a} \int_a^b f.$$

9 The Fundamental Theorem of Calculus

For the culminating section of this supplement, and for the culminating topic of a course like MAT 370, we will examine the relationship between differentiation and integration. Although the past two chapters have seemed relatively separated and don't appear to be connected, their connection is one of the most profound and famous results of differential and integral calculus.

We give a bit of history on the development of the discovery and proof of the Fundamental Theorem of Calculus. Of course, some of the big names in its discovery were English mathematician Isaac Newton and German mathematician Gottfried Leibniz, two leaders in the development of calculus.



Sir Isaac Newton

Gottfried Leibniz

However, there are other names to be mentioned in relation to this result at the heart of calculus. Scottish mathematician James Gregory is claimed to have published the first statement and proof of the Fundamental Theorem of Calculus in the 17th century. Also, English mathematician Isaac Barrow, a mentor of Isaac Newton, has stakes in the discovery of the theorem as well, although Newton developed much of the theory of calculus itself.

James Gregory

Isaac Barrow

With all of the tools we have developed throughout this supplement and especially in the preceding two chapters, we can now present and prove the statement of the Fundamental

Theorem of Calculus. However, before going into the Fundamental Theorem itself, we make a quick detour to talk about **accumulation functions**.

Consider a bounded function $f : [a.b] \to \mathbb{R}$ such that $f \in R([a, b])$. Then, we may define a new function $F : [a, b] \to \mathbb{R}$ given by

$$F(x) = \int_a^x f = \int_a^x f(t)\, dt.$$

This function F scans from a to x and returns the area under the curve of f over the interval $[a, x]$. Thus, $F(a) = 0$, for example, since no area is under f considering the interval $[a, a]$. As x goes from a to b, F accumulates the area under f. In some case, we will use the 'dummy variable' t to distinguish between the independent variable of F and the integration variable t. We illustrate this with a geometric example:

Example 9.1: A Geometric Accumulation Function

Let $f : [0, 2] \to \mathbb{R}$ be given by $f(x) = \begin{cases} 0 & 0 \le x < 1 \\ 1 & 1 \le x \le 2 \end{cases}$. Confirm f is integrable and compute $F(x) = \int_0^x f$.

Proof. That f is integrable follows from Theorem 8.14 and the fact that f is monotone (increasing). Thus, $F(x) = \int_0^x f$ is well defined on $[0, 2]$. We compute F considering the geometry of the problem and finding the area under the curve.

Suppose $0 \le x < 1$. Then the area under the curve of f is $F(x) = 0$. At $x = 1$. $F(1) = \int_0^1 f = 0$ as well given that on this interval $f = 0$ at all but one point. By Exercise 9 from the previous chapter, the integral $F(1) = \int_0^1 f$ equals 0. Now, suppose $1 < x \le 2$. The area under the graph of f is a rectangle of height 1 and base $x - 1$. Thus, the area under the curve for such x values is $F(x) = 1 \cdot (x - 1) = x - 1$. We may now say that

$$F(x) = \int_0^x f = \begin{cases} 0 & 0 \le x \le 1 \\ x - 1 & 1 < x \le 2 \end{cases}.$$

□

The power of this example may not be apparent at the moment, but after looking at the graphs of these two functions f and F, we may see what integration does to integrable functions:

Essentially, integrating from $f(x)$ to $F(x)$ took a discontinuous function and made it a continuous one. This is a nice result for this example but in fact, it holds true for every integrable function:

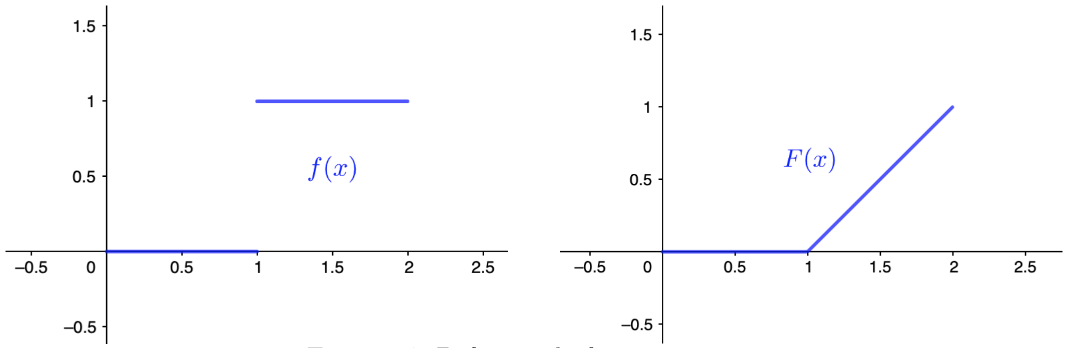


Figure 34: Before and after integration

Theorem 9.2: Accumulation Functions are Continuous

Let $f : [a, b] \to \mathbb{R}$ be bounded and suppose that f is integrable. Then, for the function $F : [a, b] \to \mathbb{R}$ given by

$$F(x) = \int_a^x f,$$

F is continuous on $[a, b]$.

Proof. We prove continuity using Definition 6.1. Let $\epsilon > 0$. Since f is bounded, there exists an $M \geq 0$ such that $|f(x)| \leq M$ for all $x \in [a, b]$. If $M = 0$, then f is 0 everywhere and hence F is 0 everywhere, and thus continuous. We suppose for the rest of the proof that $M > 0$. At $x_0 = a$, $F(a) = 0$. In this case, we set $\delta = \frac{\epsilon}{M}$ and let $x \in [a, b]$ such that $|x - a| < \delta$. We have that

$$\begin{aligned} |F(x) - F(a)| &= \left| \int_a^x f \right| \\ &\leq \int_a^x |f|, \quad \text{(Theorem 8.13)} \\ &\leq \int_a^x M = M(x - a) < M\delta = \epsilon. \end{aligned}$$

Thus, F is continuous at a. Now, we let $x_0 \in (a, b]$. Again, we pick $\delta = \frac{\epsilon}{M}$ and let $x \in [a, b]$ such that $|x - x_0| < \delta$. Whether $x < x_0$ or $x > x_0$ does not matter since $\int_x^{x_0} f = -\int_{x_0}^x f$ from Exercise 8 in the previous chapter. Thus, these two integrals will have the same absolute value. Without loss of generality, we suppose that $x > x_0$. It follows that

$$\begin{aligned} |F(x) - F(x_0)| &= \left| \int_a^x f - \int_a^{x_0} f \right| \\ &= \left| \int_{x_0}^x f \right| \quad \text{(Exercise 8, Chapter 8)} \\ &\leq \int_{x_0}^x |f| \leq \int_{x_0}^x M = M(x - x_0) < M\delta = \epsilon. \end{aligned}$$

Therefore F is continuous on $(a, b]$ and hence continuous on all of $[a, b]$. □

Now that we are more familiar with these types of accumulation functions and what they do, we ask ourselves: what if we integrate a function with a condition stronger than integrability? What happens if we require our function f to be continuous, not just integrable? This is exactly what the first part of the Fundamental Theorem of Calculus tells us:

Theorem 9.3: Fundamental Theorem of Calculus, part one

Suppose $f : [a, b] \to \mathbb{R}$ is continuous. With the function $F : [a, b] \to \mathbb{R}$ given by

$$F(x) = \int_a^x f,$$

we have that F is differentiable on $[a, b]$ and $F' = f$.

Proof. We prove this theorem using Definitions 7.1 and 6.4. Let $\epsilon > 0$. We need to use the fact that f is continuous on $[a, b]$ and thus uniformly continuous on $[a, b]$ (Theorem 6.17). Thus, by Definition 6.14 there exists a $\delta > 0$ such that $|f(x) - f(y)| < \frac{\epsilon}{2}$ whenever $x, y \in [a, b]$ with $|x - y| < \delta$.

Now, we let $x_0 \in [a, b]$ to prove differentiability. Further, we let $x \in [a, b]$ such that $0 < |x - x_0| < \delta$. We note here that for any t between x and x_0, it follows that $|t - x_0| < \delta$ as well. Thus, $|f(t) - f(x_0)| < \frac{\epsilon}{2}$ by uniform continuity. We now compute, with the help of the dummy variable t:

$$\begin{aligned}
\left|\frac{F(x)-F(x_0)}{x-x_0}-f(x_0)\right| &= \left|\frac{\int_a^x f(t)\,dt-\int_a^{x_0} f(t)\,dt}{x-x_0}-\frac{f(x_0)(x-x_0)}{x-x_0}\right| \\
&= \left|\frac{\int_{x_0}^x f(t)\,dt}{x-x_0}-\frac{\int_{x_0}^x f(x_0)\,dt}{x-x_0}\right| \\
&= \frac{\left|\int_{x_0}^x (f(t)-f(x_0))\,dt\right|}{|x-x_0|} \\
&\leq \frac{\left|\int_{x_0}^x |f(t)-f(x_0)|\,dt\right|}{|x-x_0|} \quad (t \text{ is between } x \text{ and } x_0 \text{ here}) \\
&\leq \frac{\left|\int_{x_0}^x \frac{\epsilon}{2}\,dt\right|}{|x-x_0|} = \frac{\frac{\epsilon}{2}|x-x_0|}{|x-x_0|} = \frac{\epsilon}{2} < \epsilon.
\end{aligned}$$

Therefore, F is differentiable on $[a,b]$ and from Definition 6.4, we know that the limit $F'(x_0)$ is $f(x_0)$ for all $x_0 \in [a,b]$, i.e. $F' = f$. $\square$

So, when we require our function f to be continuous, we not only get that the integral accumulation function is continuous, but also that it is differentiable. Further, taking the derivative of F gives us back f. In other symbolic notation, with t as a 'dummy variable', this part of the Fundamental Theorem tells us that

$$\frac{d}{dx}\left(\int_a^x f(t)\,dt\right) = f(x).$$

the rate of change of the area accumulated under the curve at a point x is equal to the function at x itself. In a way, differentiation and integration are inverses operations; they cancel each other out in this order (taking the derivative of an integral).

For the other part of the Fundamental Theorem, or rather the other direction, suppose we are given an **antiderivative** of f, a differentiable function F such that $F' = f$. What can we say about integrating this derivative F'? This is what the second part of the Fundamental Theorem has to say:

Theorem 9.4: Fundamental Theorem of Calculus, part two

Let $f : [a,b] \to \mathbb{R}$ be a continuous function and suppose that $F : [a,b] \to \mathbb{R}$ satisfies $F' = f$. Then,

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

Proof. For this proof, we actually rely on the proof of part one of the Fundamental Theorem of Calculus. We define the function $G : [a, b] \to \mathbb{R}$ by

$$G(x) = \int_a^x f.$$

By Theorem 9.3, we know that $G' = f$ so that G is an antiderivative of f as well. Thus, $G' = F'$ and by Exercise 7 in Chapter 7, we know that $G(x) = F(x) + C$ for some $C \in \mathbb{R}$. To find this C, we note that $G(a) = 0$. Thus, $0 = F(a) + C$, or rather, $C = -F(a)$. Finally, plugging b into G, we get that

$$G(b) = \int_a^b f = F(b) - F(a),$$

as desired. $\square$

The strength of this part of the Fundamental Theorem is that it allows us to compute the integral of f if we can find any antiderivative F of f. We simply plug in the endpoints to F and subtract! In addition, to echo what we said after the proof of the first part, this second part tells us, in different symbolic notation, that

$$\int_a^x f'(t)\,dt = f(x) - f(a).$$

Integrating the rate of change function returns the original function itself, up to some constant $f(a)$.

9.1 Integration Techniques

Now that we have the Fundamental Theorem of Calculus at our disposal, we prove some of the classic techniques to compute integrals: u-substitution and integration by parts. The first technique, known as u-substitution or perhaps change of variables, allows us to integrate certain compositions of functions with ease. If our function is of the form $f(g(x)) \cdot g'(x)$, i.e. a function g is composed inside f and the derivative of g is multiplied, we may compute the integral in a different, simpler way:

Theorem 9.5: u-substitution for Integration

Suppose $\varphi : [a, b] \to I$, $I \subseteq R$ an interval, is a differentiable function with continuous derivative φ' and $f : I \to \mathbb{R}$ is continuous. Then

$$\int_a^b (f \circ \varphi) \cdot \varphi' = \int_{\varphi(a)}^{\varphi(b)} f.$$

Proof. We first confirm that both of these integrals exist. Since f is continuous, $\int_{\varphi(a)}^{\varphi(b)} f$ makes sense. On the other hand, since both φ and φ' are continuous, $(f\circ\varphi)\cdot\varphi'$ is continuous; the integral $\int_a^b (f\circ\varphi)\cdot\varphi'$ makes sense as well. We need to show that these two integrals are equal.

We know by Theorem 9.3 that since f is continuous, an antiderivative of f exists: namely, $F(x) = \int_a^x f$ with $F' = f$. We consider the function $F\circ\varphi$, which is differentiable as well, by Theorem 7.6. Applying the formula for the chain rule, we have that

$$(F\circ\varphi)' = (F'\circ\varphi)\cdot\varphi' = (f\circ\varphi)\cdot\varphi'.$$

Therefore, by Theorem 9.4 now,

$$\begin{aligned}\int_a^b (f\circ\varphi)\cdot\varphi' &= \int_a^b (F\circ\varphi)' \\ &= F(\varphi(b)) - F(\varphi(a)) \\ &= \int_a^{\varphi(b)} f - \int_a^{\varphi(a)} f \\ &= \int_{\varphi(a)}^{\varphi(b)} f,\end{aligned}$$

exactly the equality we desired. $\square$

This proof took advantage of both parts of the Fundamental Theorem of Calculus, integrating a derivative and taking the derivative of an integral. In the proof, φ was our differentiable function composed inside f. In many calculus classes, this function φ is usually written instead as u, where u is some differentiable function of x. Hence, this technique usually goes by the name u-substitution.

For the other famous technique, however, we will only need the second part. If we are given an integrand of the form $f(x)g'(x)$, where f and g are differentiable functions, then we may compute $\int_a^b fg'$ in terms of $\int_a^b f'g$. This is integration by parts:

Theorem 9.6: Integration by Parts

Suppose $f, g : [a,b] \to \mathbb{R}$ are two differentiable functions with continuous derivatives f' and g'. Then

$$\int_a^b fg' = fg\big|_a^b - \int_a^b f'g, \qquad \text{where } fg\big|_a^b = f(b)g(b) - f(a)g(a).$$

Proof. Again, both of the integrals exist since f, g, f', and g' are all continuous. For the proof of Theorem 9.5 we took advantage of the Chain Rule for derivatives. For this proof, since there are products with no compositions of functions, we will be first using the Product Rule for derivatives, Theorem 7.5. We know that $(fg)' = fg' + fg'$, or rather, that $fg' = (fg)' - f'g$. Integrating both sides, we have that by Theorem 9.4,

$$\begin{aligned}\int_a^b fg' &= \int_a^b ((fg)' - f'g)\\ &= \int_a^b (fg)' - \int_a^b f'g\\ &= f(b)g(b) - f(a)g(a) - \int_a^b f'g\\ &= fg|_a^b - \int_a^b f'g,\end{aligned}$$

as desired. $\square$

These will be the last results of this supplement, and hopefully some helpful proofs of content seen and used in Calculus II. From the building blocks of the field of real numbers $\mathbb{R}$, we were able to build up the foundations to rigorously discuss sequences, continuous functions, differentiation, and integration. I hope that this text has not only helped elucidate why the results from Calculus classes are mathematically sound, but made MAT 370 go by a bit more easily!

9.2 Exercises

1. For the following functions $f : [a, b] \to \mathbb{R}$, geometrically compute the accumulation functions $F(x) = \int_a^x f$:

   a) $f : [0, 3] \to \mathbb{R}$, $f = \begin{cases} 0 & 0 \leq x < 1 \\ 1 & 1 \leq x \leq 2 \\ 0 & 2 < x \leq 3 \end{cases}$,

   b) $f : [0, 2] \to \mathbb{R}$, $f = \begin{cases} x & 0 \leq x < 1 \\ 0 & 1 \leq x \leq 2 \end{cases}$,

   c) $f : [0, 2] \to \mathbb{R}$, $f = \begin{cases} x & 0 \leq x < 1 \\ 2 - x & 1 \leq x \leq 2 \end{cases}$,

   d) $f : [0, 2] \to \mathbb{R}$, $f = \begin{cases} x & 0 \leq x < 1 \\ x - 2 & 1 \leq x \leq 2 \end{cases}$,

   e) $f : [0, 3] \to \mathbb{R}$, $f = \begin{cases} x & 0 \leq x < 1 \\ 1 & 1 \leq x \leq 2 \\ 3 - x & 2 < x \leq 3 \end{cases}$.

2. Use the Fundamental Theorem of Calculus, part one to show that for a continuous function $f : [a, b] \to \mathbb{R}$ and $a < c < b$,

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

3. Suppose we restrict the absolute value function to the domain $|\cdot| : [-1, 1] \to \mathbb{R}$. The absolute value is not differentiable at 0, but it is continuous everywhere on $[-1, 1]$. Using the Fundamental Theorem of Calculus, part one, compute an antiderivative $F(x)$ of $|x|$, i.e. a function such that $F'(x) = |x|$ for all $x \in [-1, 1]$.

4. Use the Fundamental Theorem of Calculus, part one to compute the following derivatives:

   a) $\frac{d}{dx} \int_a^{2x} e^{t^2} \, dt$,
   b) $\frac{d}{dx} \int_{-3x^2}^{b} \sin^2(t) \, dt$,
   c) $\frac{d}{dx} \int_a^{\cos x} \ln(t^3) \, dt$,
   d) $\frac{d}{dx} \int_{-e^x}^{e^x} \cos t \, dt$.

5. Use u-substitution to compute

$$\int_0^{\sqrt{\pi}/2} x \cos(x^2) \, dx.$$

6. Use u-substitution to compute

$$\int_{\pi/4}^{3\pi/4} \csc x \, dx.$$

7. Compute $\int_1^e \ln x \, dx$ using integration by parts.

8. Compute $\int_0^1 e^x \sin(x) \, dx$ using integration by parts.

Appendices

A Hints to Exercises

A.1 Chapter 3: The Real Number System

1. For $0 < x < y$, prove by induction. For $x < y < 0$, consider the behavior of x^n for $x < 0$ at even and odd values of n.

2. Work through the cases when $\frac{a}{b} = 0$, $\frac{a}{b} > 0$, and $\frac{a}{b} < 0$. For example, if $\frac{a}{b} > 0$, then a and b have the same sign.

3. How does the average of two distinct real numbers x and y compare to x and y? Can this process of taking an average be repeated?

4. For x not an upper bound of E, there exists $y \in E$ such that $y > x$, by negating Definition 3.6. If y turns out to be an upper bound of E, can we define another element greater than x that is not?

5. Since U is nonempty and L is bounded above by every element of U, we may use the Completeness Axiom: we know that $x = \sup(L)$ exists. Prove that $l \leq x \leq u$ for all $l \in L$ and $u \in U$ using Definition 3.9. Could the same result be proved using Theorem 3.15?

6. Show that $\sup(A) + \sup(B)$ is an upper bound of $A + B$. Since $\sup(A + B)$ is the least upper bound, $\sup(A + B) \leq \sup(A) + \sup(B)$. For the other inequality, we know that for any $a \in A$ and $b \in B$, $a + b \leq \sup(A + B)$. Solve this inequality for a. This gives a new upper bound for A. How does $\sup(A)$ compare to this new upper bound? Solve for b in the resulting inequality to include $\sup(B)$.

7. That $\sup(AB) \leq \sup(A)\sup(B)$ is the same argument as in the previous hint. For the other inequality, we know that $ab \leq \sup(AB)$ for all $a \in A$ and $b \in B$. Solving for a here gives a new upper bound of A with $\sup(A) \leq \frac{1}{b}\sup(AB)$ for all $b \in B$. Solve for b here to get a new inequality with $\sup(B)$. Why can we divide and multiply by $\sup(A)$ and $\sup(B)$?

A.2 Chapter 4: Basic Topology on $\mathbb{R}$

1. Apply De Morgan's laws for sets.

2. For the forward direction, argue by contradiction that some limit point x_0 of E is an element of E^c, which is open. Thus, on open interval may be found that contains x_0 but does not intersect E. For the converse direction, show that the complement of E is open, using Definition 4.2.

3. Use the density of the rationals from Corollary 3.19.

4. Show that singleton sets $\{x\}$ are closed in $\mathbb{R}$. Then, use the result from Exercise 1 of this chapter.

5. For an arbitrary $\epsilon > 0$, show that $(a-\epsilon, a+\epsilon)$ and $(b-\epsilon, b+\epsilon)$, respective neighborhoods of a and b. intersect (a, b).

6. For the first two statements, consider any $n \in \mathbb{N}$ or $m \in \mathbb{Z}$. What can you say about the neighborhoods $(n - \frac{1}{2}, n + \frac{1}{2})$ and $(m - \frac{1}{2}, m + \frac{1}{2})$ or n and m? Where do they intersect $\mathbb{N}$ and $\mathbb{Z}$? For the last result, use the density of rationals, Corollary 3.19.

7. a) Use Exercise 1 of this chapter.

   b) If E is closed, then E **is** one of closed sets that contains E. For the other direction use part (a) above.

   c) Prove that $x \in \overline{E}$ if and only if there exists a neighborhood of x that contains no element of E, using the above definition of $\overline{E}$. By definition,

$$\begin{aligned}(\overline{E})^c &= \bigcup \{K^c : K \text{ is closed and } E \subseteq K\} \\ &= \bigcup \{U : U \text{ is open and } U \subseteq E^c\}\end{aligned}$$

   How does $x \in (\overline{E})^c$ imply that there exist a neighborhood that contains no element of E?

   d) For one inclusion, show that $E \subseteq \overline{E}$ and apply result (c). For the other, Suppose that $x \in \overline{E} \cap E^c$ to show that $x \in E'$. Use part (c) and Theorem 4.8 to show x is a limit point of E.

   e) Apply Theorem 3.12 to result (c). For an arbitrary $\epsilon > 0$, consider the neighborhoods $(\sup(E) - \epsilon, \sup(E) + \epsilon)$ and $(\inf(E) - \epsilon, \inf(E) + \epsilon)$ of $\sup(E)$ and $\inf(E)$, respectively. Can we always find elements of E in these neighborhoods?

A.3 Chapter 5: Real-valued Sequences

1. For the forward direction, prove by contrapositive. If $x \neq 0$, pick an $\epsilon > 0$ such that $|x| \geq \epsilon$.

3. Use the contrapositive of Lemma 5.5.

4. b) Pick $\epsilon_0 = 1$. For any $N \in \mathbb{N}$ we need to find an $n \geq N$ such that $|x_n - 0| = x_n \geq 1$. If N is a multiple of 3, N suffices. If N is not a multiple of 3, which n could we pick?

5. Call $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$. Suppose by contradiction that $x > y$. Consider $\epsilon_0 = x - y$. Pick an $N \in \mathbb{N}$ such that $|x_n - x| < \frac{\epsilon_0}{4}$ and $|y_n - y| < \frac{\epsilon_0}{4}$ for all $n \geq N$ (why can this N be picked?). Show that for such $n \geq N$, $x_n > y_n$, a contradiction. For example, begin by noting

$$x_n - y_n = (x_n - x) + (x - y) + (y - y_n) \geq x - y - |x_n - x| - |y_n - y| .$$

6. Prove that

   a) Apply Theorem 5.6 and use induction on k.

   b) Multiply numerator and denominator by $\frac{1}{n^2}$ and use part (a) along with Theorem 5.6.

   c) Same argument as with part (b), only now multiply by $\frac{1}{n^k}$.

   d) Repeat part (b).

   e) Suppose $P(n)$ has degree k and $Q(n)$ has degree l with $k < l$. Repeat part (c), multiplying by $\frac{1}{n^l}$.

7. Use a similar technique as in the proof of Theorem 5.6.1. We know that for all $n \in \mathbb{N}$, $|b_n| \leq M$ for some $M \geq 0$. If $M = 0$, $b_n = 0$ for all $n \in \mathbb{N}$ and the result follows. Let $\epsilon > 0$. For $M > 0$, pick an $N \in \mathbb{N}$ such that $|a_n - 0| = |a_n| < \frac{\epsilon}{M}$ for all $n \geq N$. Show that for such $n \geq N$, $|a_n b_n - 0| < \epsilon$.

8. Adapt the proof of Theorem 5.8 while working with infima instead.

9. a) Use induction to show that x_n is a decreasing sequence, i.e. $x_n \geq x_{n+1}$ for all $n \in \mathbb{N}$. For a lower bound, use induction to show that $x_n > 1$ for all $n \in \mathbb{N}$. With limit $\lim\limits_{n\to\infty} x_n = x$, we know x is a solution to $x^2 - 2x + 1 = 0$.

   b) Prove this sequence is increasing with an upper bound of 2. The limit x satisfies $x^2 - x - 2 = 0$. Why can't $x = -1$?

   c) Show that x_n is decreasing with a lower bound of 0. The limit x satisfies $x^2 - 3x + 1 = 0$.

10. If $a = 0$, the result is trivial. If $0 < a < 1$, show that the sequence $x_n = a^n$ is decreasing and bounded below by 0. To show that the limit x is 0, note that $x = \lim\limits_{n\to\infty} a^{n+1} = a \lim\limits_{n\to\infty} a^n = ax$. Therefore, the limit satisfies $x(a-1) = 0$. What does our hypothesis on a require?

11. Use the fact that for all $n \in \mathbb{N}$, $-1 \leq \sin(n) \leq 1$.

13. Since $(x_n)_{n\in\mathbb{N}}$ is unbounded, then for every $k \in \mathbb{N}$, there exists an $n_k \in \mathbb{N}$ such that $|x_{n_k}| > k$, or rather that $\frac{1}{|x_{n_k}|} < \frac{1}{k}$. Can these n_k be chosen so that $n_1 < n_2 < n_3 < \cdots$? Use the Squeeze Theorem for the inequality $0 < \frac{1}{|x_{n_k}|} < \frac{1}{k}$ to prove the result.

14. For the forward direction, given a bounded infinite subset $E \subseteq \mathbb{R}$, we may create a sequence of distinct elements of E, say $(x_n)_{n\in\mathbb{N}}$. The Bolzano-Weierstrass Theorem says that this sequence $(x_n)_{n\in\mathbb{N}}$ has a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ with limit x. Show that x is a limit point of E.

For the converse direction, suppose we have a bounded sequence of real numbers $(x_n)_{n\in\mathbb{N}}$. If $(x_n)_{n\in\mathbb{N}}$ consists of a finite number of distinct elements $\{x_{n_1}, \ldots, x_{n_m}\}$, then there has to be an infinite repetition of one of the x_{n_i}. Our convergent subsequence is this constant subsequence of x_{n_i} terms. If x_n has infinite distinct elements, then by our hypothesis, $E = \{x_n : n \in \mathbb{N}\}$ is an infinite bounded subset of $\mathbb{R}$ and thus has a limit point x. Show that there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ that converges to x.

15. The distance between subsequent terms of this sequence satisfies $|x_1 - x_2| = \frac{1}{2}$, $|x_3 - x_4| = \frac{1}{3}$, $|x_6 - x_7| = \frac{1}{4}$, and so on. Given an $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $\frac{1}{N} < \epsilon$. Can you figure out how far you need to go in the sequence so that $|x_n - x_{n+1}| \leq \frac{1}{N} < \epsilon$?

    To show x_n is not Cauchy, try the problematic $\epsilon_0 = 1$. For any $N \in \mathbb{N}$, can you find $m, n \geq N$ such that $|x_m - x_n| \geq 1$?

16. Suppose that $(x_n)_{n\in\mathbb{N}}$ is Cauchy with a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ that converges to x. Pick an $N \in \mathbb{N}$ such that $|x_n - x_m| < \frac{\epsilon}{2}$ and $|x_{n_k} - x| < \frac{\epsilon}{2}$ for all $n, m, k \geq N$. Then for $n \geq N$, $|x_n - x| \leq |x_n - x_{n_k}| + |x_{n_k} - x| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. Why do we get that $|x_n - x_{n_k}| < \frac{\epsilon}{2}$ as well?

17. You will be relying on Theorem 5.19.

18. Suppose by contradiction that $\mathbb{N}$ is bounded above by some $x_0 \in \mathbb{R}$, i.e. $n \leq x_0$ for all $n \in \mathbb{N}$. Consider the sequence $(n)_{n\in\mathbb{N}}$. This sequence is bounded above and increasing, so by Theorem 5.8, $\lim\limits_{n\to\infty} n = x$, for some $x \in \mathbb{R}$. What is the contradiction here? Use Definition 5.2 with $\epsilon = \frac{1}{2}$.

A.4 Chapter 6: Continuity

2. For an irrational number $\xi \in [0, 1]$, we use the fact that as rationals approach ξ, their denominators grow larger and larger without bound. Thus, for any $\epsilon > 0$, pick an $N \in \mathbb{N}$ such that $\frac{1}{N} < \epsilon$. It is a bit technical, but we may also pick a $\delta > 0$ such that for rationals $\frac{p}{q}$, $\left|\frac{p}{q} - \xi\right| < \delta$ implies $\frac{1}{q} < \frac{1}{N}$. Finish the proof by showing that for all $x \in [0, 1]$ with $|x - \xi| < \delta$. $|f(x) - f(\xi)| = |f(x)| < \epsilon$. Work in cases, when x is rational and irrational.

   To show that f is discontinuous at every rational $\frac{p}{q}$, use the density of the irrationals, Corollary 3.20.

3. This is an application of Theorem 6.8.

4. a) Use the facts that $\sin(a) - \sin(b) = 2\cos(\frac{a+b}{2})\sin(\frac{a-b}{2})$ for any $a, b \in \mathbb{R}$ and $|\sin(x)| \leq |x|$ for all $x \in \mathbb{R}$.

b) Use the identity $\cos(x) = \sin(x + \frac{\pi}{2})$.

c) Apply Theorem 6.8.

d) Apply Theorem 6.8.

5. Consider the sequence $x_n = \frac{2}{(4n+1)\pi}$ so that $\lim_{n\to\infty} x_n = 0$. What does the sequence $(f(x_n))_{n\in\mathbb{N}}$ converge to? Why does this prove f is not continuous at 0, by Theorem 6.7?

6. Define $g(x) = -f(x)$ and apply the lemma.

7. a) Consider the function $f(x) = \cos(x) - x$. Apply Lemma 6.11.

   c) Polynomials $p(x)$ of odd degree extend to positive and negatives as x moves far from 0 in either directions. Essentially, we can find always find an $M \geq 0$ such that $p(-M)$ and $p(M)$ are on different sides of the y-axis. Use this to deduce the result.

9. Try $\delta = \epsilon$.

10. Mimic the technique used to solve Exercise 5 of this chapter.

11. a) Follow the proof of Theorem 6.9.

    b) For some counterexamples, consider $D = \mathbb{R}$ and $f(x) = g(x) = x$ for the product. With $D = (0, \infty)$, try $f(x) = 1$ and $g(x) = x$ for the quotient.

A.5 Chapter 7: Differentiation

1. Use the substitution $x = c + h$ and show that $\lim_{h\to 0} x = c$. Rewrite the limit on the right-hand side.

3. a) Note that $f(x) = x^n = x^{n-1} \cdot x$. Use the product rule to find the derivative: $f'(x) = (x^{n-1} \cdot x) = (n-1)x^{n-2} \cdot x + x^{n-1}$.

   b) Use the quotient rule from Theorem 7.5 and the fact that $x^{-n} = \frac{1}{x^n}$ for $n \in \mathbb{N}$.

4. a) Use the alternate presentation of a derivative from Exercise 2 above. Then, use trigonometric sum identities and the fact that
$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1.$$

   b) Use the identity $\cos(x) = \sin(\frac{\pi}{2} - x)$.

   c) Apply Theorem 7.5.

6. Apply the Mean Value Theorem.

7. Consider the function $h(x) = f(x) - g(x)$ and apply part (d) of the previous exercise.

A.6 Chapter 8: Integration

1. Imitate the proof of Example 8.10.

2. Imitate the proof of Example 8.10.

3. Negate the alternate characterization in Theorem 8.9. Show that for any partition P of $[0,1]$, the m_i and M_i values as in Definition 8.3 are always equal to 0 and 1, respectively.

4. For the function $g(x) = f(x) + c$, how do $m_i(f)$ and $M_i(f)$ compare to $m_i(g)$ and $M_i(g)$, with notation from Definition 8.3?

5. For Darboux-integrability, use part (c) of the previous exercise. For Riemann-integrability, $f(t_i) = c$ for any $t_i \in [x_i, x_{i+1}]$, for any partition $P = \{x_1, \dots, x_n\}$ of $[a,b]$.

6. Apply Exercises 4 and 5.

7. a) Since $|f(x)| \leq M$ for some $M \geq 0$, $\left|f(x)^2 - f(y)^2\right| \leq 2M\,|f(x) - f(y)|$ for any $x, y \in [a,b]$.

   b) Use the identity $fg = \frac{1}{4}((f+g)^2 - (f-g)^2)$.

8. Let $\epsilon > 0$. For the forward direction, pick a partition that satisfies Theorem 8.9 for ϵ. If $c \in P$, then split P into partitions P_1 of $[a,c]$ and P_2 of $[c,b]$. How do $U(P_1, f) - L(P_1, f)$ and $U(P_2, f) - L(P_2, f)$ compare to $U(P, f) - L(P, f)$? If $c \neq P$, throw it in to start with a partition finer than P and split into P_1 and P_2 as above.

   For the converse direction, pick two respective partitions of $[a,c]$ and $[c,b]$ satisfying Theorem 8.9 for $\frac{\epsilon}{2}$. Take the union P of these two partitions (which will be a partition of $[a,b]$) and show that $U(P, f) - L(P, f) < \epsilon$

9. Show that the function $h(x) = f(x) - g(x)$, which is 0 except at the finitely many points, is integrable. Set $M = \max\{|h(x_i)| : i = 1, \dots, m\}$ and construct a partition P that isolates the problematic points, say in subintervals $[x_i - \frac{\epsilon}{4Mm}, x_i + \frac{\epsilon}{4Mm}]$. Over every other subinterval, $h = 0$. Compute $U(P, h)$ and $L(P, h)$. Conclude with $g = h + f \in R([a,b])$. To compute $\int_a^b g$, show that $\int_a^b h = 0$.

10. You will be using the Intermediate Value Theorem and the fact that $m \leq f(x) \leq M$ implies $m(b-a) \leq \int_a^b f \leq M(b-a)$.

A.7 Chapter 9: The Fundamental Theorem of Calculus

1. Imitate the proof of Example 9.1.

2. If f is continuous, then it admits an antiderivative $F(x) = \int_a^x f$ by Theorem 9.3. Furthermore, $F(b) - F(a) = \int_a^b f$. Write the analogous equations for $\int_a^c f$ and $\int_c^b f$.
3. Use the same geometric techniques as in Example 9.1.
4. Don't forget the Chain Rule for derivatives.
5. Try the substitution $u = x^2$.
6. Multiply by $1 = \frac{\csc x + \cos x}{\csc x + \cos x}$ and use the trig derivative formulas derived in Exercise 5 of Chapter 7.
7. There is still a product of two functions in this integral. How might you switch derivatives?
8. Denote the desired integral by $I := \int_0^1 e^x \sin(x)\, dx$ and go through the process of integration by parts twice. You should end up with an equation in I.

References

[1] K. Weber, "Traditional instruction in advanced mathematics courses: a case study of one professor's lectures and proofs in an introductory real analysis course," *The Journal of Mathematical Behavior*, vol. 23, no. 2, pp. 115–133, 2004.

[2] G. Howson, "The development of mathematics textbooks: Historical reflections from a personal perspective," *ZDM*, vol. 45, no. 5, pp. 647–658, 2013.

[3] D. Rohrer, R. F. Dedrick, and M. K. Hartwig, "The scarcity of interleaved practice in mathematics textbooks," *Educational Psychology Review*, pp. 1–11, 2020.

[4] M. Pinto and D. Tall, "Building formal mathematics on visual imagery: A case study and a theory," *For the Learning of Mathematics*, vol. 22, no. 1, pp. 2–10, 2002.

[5] M. Raman, "Epistemological messages conveyed by three high-school and college mathematics textbooks," *The Journal of Mathematical Behavior*, vol. 23, no. 4, pp. 389–404, 2004.

[6] K. H. Roh and Y. H. Lee, "Designing tasks of introductory real analysis to bridge a gap between students' intuition and mathematical rigor: the case of the convergence of a sequence," *International Journal of Research in Undergraduate Mathematics Education*, vol. 3, no. 1, pp. 34–68, 2017.

[7] N. H. Wasserman and K. Weber, "Pedagogical applications from real analysis for secondary mathematics teachers," *For the Learning of Mathematics*, vol. 37, no. 3, pp. 14–18, 2017.