Calculus

Joshua Nauman

Contents

I Calculus I **7**

0.1 Introduction . . . 9
0.2 In Progress Note (v2) . . . 9
0.3 Important Algebra Rules and Facts . . . 10
0.4 Important Trigonometric Relations . . . 10
0.5 Important Geometric Facts . . . 12

1 Limits **13**

1.1 What is a Limit? . . . 13
1.2 Limits, Graphically . . . 13
1.3 Limits, Algebraically . . . 14
1.4 Limits at Infinity . . . 15
1.5 Does Not Exist . . . 16
1.6 The Squeeze Theorem . . . 17
1.7 Special Limits . . . 17
1.8 Continuity and Discontinuity . . . 20
1.8.1 Removable Discontinuity . . . 20
1.8.2 Jump Discontinuity . . . 22
1.8.3 Infinite Discontinuity . . . 23
1.8.4 Oscillating Discontinuity . . . 24
1.8.5 Continuity at a Point . . . 25
1.8.6 Examples . . . 26
1.9 Intermediate Value Theorem . . . 28
1.10 Chapter 1 Summary and Exercises . . . 29
1.10.1 Chapter 1 Exercises . . . 29

2 Derivatives **31**

2.1 What is a Derivative? . . . 31
2.2 Limit Definition of the Derivative . . . 31
2.3 Differentiability . . . 33
2.4 Derivative Shortcuts . . . 33
2.4.1 Properties of the Derivative . . . 34
2.4.2 Power Rule . . . 34
2.4.3 Product Rule . . . 35
2.4.4 Chain Rule . . . 36
2.4.5 Quotient Rule . . . 37
2.4.6 Known Derivatives . . . 38
2.4.7 Implicit Differentiation . . . 41

2.4.8 Logarithmic Differentiation . . . 41
2.4.9 Piecewise Derivatives . . . 44
2.5 Examples . . . 44
2.6 Chapter 2 Summary and Exercises . . . 44

3 Derivative Applications **45**

4 Integrals **47**
4.1 What is an Integral? . . . 47
4.2 Approximating Integrals . . . 48
4.2.1 Left and Right Rectangular Approximation Methods . . . 48
4.2.2 Midpoint Rectangular Approximation Method . . . 52
4.2.3 Trapezoid Rule . . . 55
4.2.4 Net vs. Total Area with Geometry and Symmetry . . . 58
4.3 Properties of Integrals . . . 60
4.4 The Fundamental Theorem of Calculus . . . 62
4.5 Indefinite Integrals and the Constant of Integration . . . 66
4.6 Basic Methods for Integration . . . 67
4.6.1 Reverse Power Rule . . . 67
4.6.2 Known Integrals . . . 70
4.6.3 Split The Fraction . . . 73
4.7 Substitution Method . . . 74
4.7.1 Converting to u bounds . . . 77
4.8 Net vs. Total Area with Integrals . . . 78
4.9 Tips, Tricks, and Creativity . . . 79
4.9.1 A Different Kind of u-substitution . . . 79
4.9.2 Creativity . . . 80
4.10 Extra Examples . . . 81
4.11 Chapter 4 Summary and Exercises . . . 81

5 Applications of Integrals **83**
5.1 Initial Value Problems . . . 83
5.2 Average Value of a Function . . . 83
5.3 1 Dimensional Particle Motion Revisited . . . 84
5.4 Integrals as Accumulation of Change . . . 85
5.5 Area Between Curves . . . 87
5.6 Volumes of Revolution . . . 88
5.6.1 The Disc Method . . . 90
5.6.2 The Cylindrical Shell Method . . . 92
5.7 Chapter 5 Summary and Exercises . . . 94

6 Introduction to Differential Equations **95**

II Calculus II **97**

7 More on Integrals **99**

8 Series **101**

9 Introduction to the Calculus of Parametric Equations, Polar Equations, and Vector-Valued Functions **103**

III Calculus III **105**

IV Proofs, Answers to Exercises, and more **107**

Part I

Calculus I

0.1 Introduction

I started writing this book as a junior in high school in my second semester of AP Calculus BC with the goal of writing a textbook from the perspective of a student so that better analogies could be made to increase understanding, since I remember what it was like not long ago to not understand.

Calculus is the mathematical study of change. In this text, we outline the basics of the three main subjects of calculus: Differential Calculus (Calculus I), Integral & Series Calculus (Calculus II), and Multivariable & Vector Calculus (Calculus III). We attempt to do this in a fluid and intuitive manner that is easy for those of various mathematical backgrounds to understand. This book is intended as an introductory text to calculus and is best used when supplemented with lectures on the main topics of each chapter.

At the end of each chapter, following the chapter summary, there will be exercises split up by which section of the chapter they should be assigned with. Solutions to odd-numbered exercises are provided at the end of the book. Also at the end of the book will be some of the more involved derivations of different derivative and integral rules, like the chain rule, derivative of inverse trig functions, integral of $\frac{1}{x^2+1}$ (without trig substitution), and more.

0.2 In Progress Note (v2)

This textbook is very much a work in progress. In its current stage, no formatting has been finalized (in fact, in most areas, it *will* be changed). Exercises will be added in future versions. In this version, chapter 5 on some of the applications of integrals has been added. Small refinements have been made to other sections, particularly to the section on limits, though there are many more to come in the near future.

In the next version, I hope to add more examples and a few exercises for the currently existing sections. I would also like to formalize my write-up of chapter 3 and roughly complete that section here. Moreover, I plan on making major modifications to Chapter 1 on limits for the reason that, while most of the explanations are okay for an intuitive understanding, they lie on very shaky ground mathematically. I would like to be slightly more precise (while still avoiding epsilons and deltas). I will also finish the main content of chapters 2 (derivatives) and 5 (applications of integrals) by finishing the sections on differentiability, implicit differentiation, and initial value problems, along with adding numerous more examples. My formatting largely varies by chapter as I am trying to discover which formatting I like the best. It should again be noted that no formatting is finalized, and that much of the formatting is currently in a very prototypical state.

A rough draft of chapter 3 is complete, and chapter 6's rough draft is about halfway done. Suggestions and corrections are welcome and encouraged, both of current content and suggestions for future content (examples, exercises, formatting tips, etc.). Contact information can be found at my website.

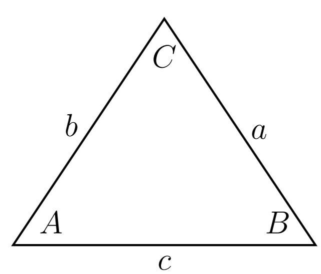
0.3 Important Algebra Rules and Facts

Here, we list the results of algebra that are most important to the study of introductory calculus.

1. For an equation in the form $ax^2+bx+c=0$,
$$x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}.$$
2. $a^b \cdot a^c = a^{b+c}$.
3. $\dfrac{a^b}{a^c} = a^{b-c}$.
4. $(a^b)^c = a^{bc}$.
5. $a^{-b} = \dfrac{1}{a^b}$ unless $a = 0$.
6. $a^0 = 1$ whenever $a \neq 0$.
7. $a^{\frac{m}{n}} = \sqrt[n]{a^m}$.
8. $\ln(e^k) = k$.
9. $e^{\ln(k)} = k$.
10. $\ln(a) = \log_e(a)$.
11. $\log_a(a^k) = k$.
12. $a^{\log_a(k)} = k$
13. $\log_a(a) = 1$.
14. $\log_a(1) = 0$.
15. $\log_a(x)$ is undefined whenever $x \leq 0$ or when $a = 1$ or $a \leq 0$.
16. $\log_a(bc) = \log_a(b) + \log_a(c)$.
17. $\log_a(b^c) = c\log_a(b)$. This is only true if the exponent is inside the log. $\log_a(b)^c \neq c\log_a(b)$.
18. $\log_a\left(\dfrac{b}{c}\right) = \log_a(b) - \log_a(c)$.
19. $\log_a(b) = \dfrac{\log_c(b)}{\log_c(a)}$. You can choose c.

0.4 Important Trigonometric Relations

For any formula that references A, B, C, a, b, c, refer to the following triangle. This is a generic triangle. These rules apply to all triangles, even ones that don't look like this one. What is important is that side a is opposite of angle A, side b is opposite of angle B, and side c is opposite of angle C.



1. $\tan(x) = \dfrac{\sin(x)}{\cos(x)}$.
2. $\csc(x) = \dfrac{1}{\sin(x)}$.
3. $\sec(x) = \dfrac{1}{\cos(x)}$.
4. $\cot(x) = \dfrac{1}{\tan(x)} = \dfrac{\cos(x)}{\sin(x)}$.
5. $\sin^2(x) + \cos^2(x) = 1$.

6. $\sin(2x) = 2\sin(x)\cos(x)$.

7. $\cos(2x) = 2\cos^2(x) - 1 = 1 - 2\sin^2(x) = \cos^2(x) - \sin^2(x)$.

8. $\sin(\alpha \pm \beta) = \sin(\alpha)\cos(\beta) \pm \cos(\alpha)\sin(\beta)$.

9. $\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) \mp \sin(\alpha)\sin(\beta)$.

10. The Law of Sines: $\dfrac{\sin A}{a} = \dfrac{\sin B}{b} = \dfrac{\sin C}{c}$.

11. The Law of Cosines:

    - $c^2 = a^2 + b^2 - 2ab\cos C$.
    - $b^2 = a^2 + c^2 - 2ac\cos B$.
    - $a^2 = b^2 + c^2 - 2bc\cos A$.

The unit Circle:

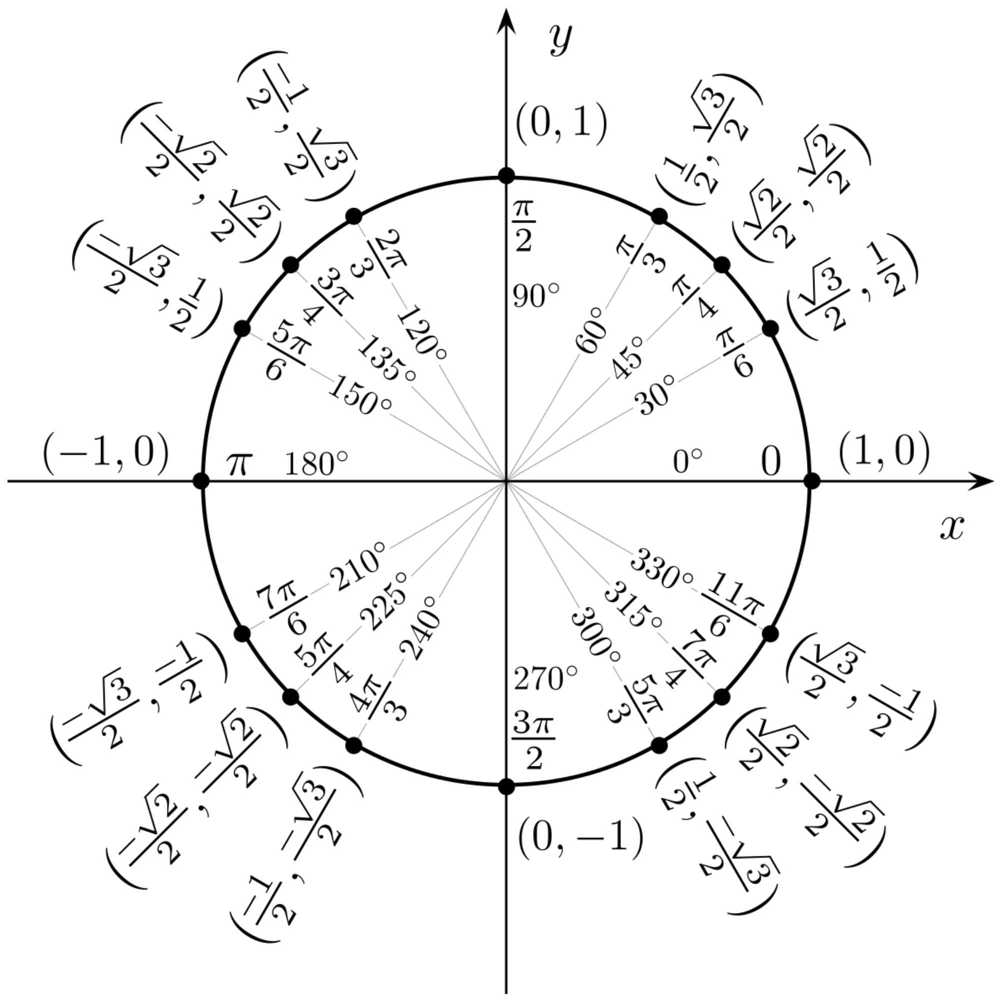
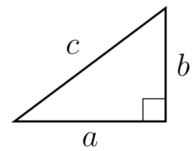


Figure 1: The Unit Circle.

0.5 Important Geometric Facts



1. The Pythagorean Theorem: $a^2 + b^2 = c^2$.
2. Area of a circle: πr^2.
3. Circumference of a circle: $2\pi r$.
4. Volume of a sphere: $\frac{4}{3}\pi r^3$.
5. Surface area of a sphere: $4\pi r^2$.
6. Volume of a cylinder: $\pi r^2 h$.
7. Surface area of a cylinder: $2\pi rh + 2\pi r^2$.
8. Volume of a cone: $\frac{1}{3}\pi r^2 h$.

Chapter 1

Limits

1.1 What is a Limit?

The limit is one of the most fundamental ideas of calculus and all of mathematics. This simple definition of the limit is the value that a function is approaching when the function's input approaches some value. Let's take a look at an example.

$$\lim_{x \to 3} x^2 = ?$$

This problem is asking "as x gets really, really close to 3, what does x^2 get really really close to?" For most problems, a graph of the function is helpful. The graph of the function x^2 is shown below, with a red point indicating the function value when $x = 3$.

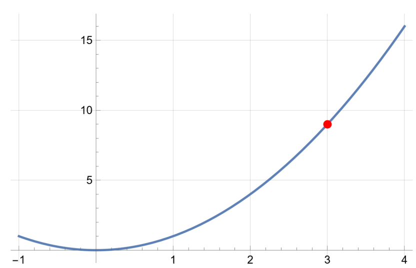


Figure 1.1: Plot of $f(x) = x^2$ with a point at $(3, f(3))$.

So, as we get really really close to the point when $x = 3$ from each side, what function value are we approaching? From the graph, it appears that the function x^2 is approaching the value of 9. Let's confirm this with an algebraic limit: $\lim_{x \to 3} x^2 = 3^2 = 9$. That's not everything about limits, however—they're not just a more complex way to tell you a function value. This section will explain the basics of limits.

1.2 Limits, Graphically

If possible, when calculating a limit, it's a good idea to look at the graph of the function you're taking the limit of. It's much easier to see this way what value a given function is approaching, if any. Let's take a look at a piecewise function—a function comprised of multiple functions, depending on what the value of x is. Consider the following piecewise function and its graph, shown below.

$$f(x) = \begin{cases} \frac{3}{2}x & x < 4 \\ 2 & 4 \leq x \end{cases}$$

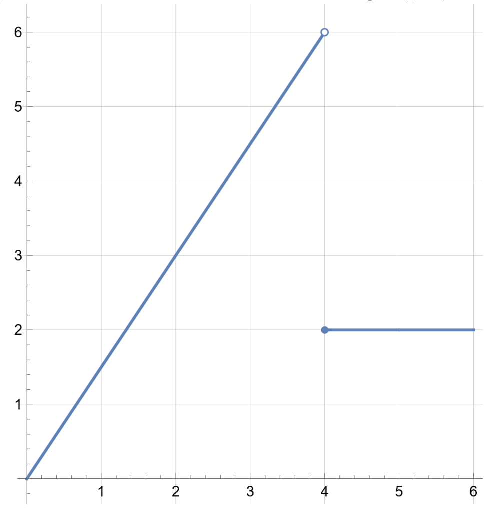


Figure 1.2: A graph of the piecewise function $f(x)$.

As you can see, the function is defined for all x. At $x = 4$, the function value is 2 because, if you look at the constraints on x, you can see that when $x = 4$, the bottom equation of the piecewise function defines the function value.

Example 1. What is the limit as x approaches 4 from the left?

Solution. To determine this, we first must decide which part of the piecewise function we are analyzing. Since we are going from the left of $x = 4$, we will use the top equation of the piecewise because it is the part of the function that is defined when x is less than 4. As we move from the left towards the open circle on the graph, we see that we are approaching the y value of 6. □

Example 2. What is the limit as x approaches 4 from the right?

Solution. First, like above, we must decide which part of the piecewise function to use. In this case, since the bottom equation of the piecewise function is the part that is defined when x is greater than (or equal to) 4, we will use the bottom equation. As we move from right towards the closed circle on the graph, we see that we are approaching the y value of 2. □

What is the limit as x approaches 4 for this function? To answer this, we must compare the limits on the right and left-hand sides of $x = 4$. We will talk more about this soon.

1.3 Limits, Algebraically

As nice as it is, you won't always be able to use a calculator to graph the function you are taking the limit of. Even for simple functions like the ones previously discussed that you can probably graph without a calculator, it is important to understand how to take a limit of a function without looking at a graph.

In the example above, we analyzed the limits of the piecewise function $f(x)$, where $f(x) = \begin{cases} \frac{3}{2}x & x < 4 \\ 2 & 4 \leq x \end{cases}$. Let's compute these same limits, but, this time, without a graph. The notation for a limit as x is approaching some value n from the left for a given function $g(x)$ is $\lim_{x\to n^-} g(x)$. Similarly, the notation when x is approaching n from the right for a given function $g(x)$ is $\lim_{x\to n^+} g(x)$.

Let's start by computing $\lim_{x\to 4^-} f(x)$. Since we are approaching 4 from the left, we will use the top equation of the piecewise function because it is the equation that determines what is to the left of the function at $x = 4$. The first tool to solving any limit is always plugging in the value that x is approaching, whether it is a limit from the left, right, both sides, a piecewise function, or some other situation. Plugging in 4 for x in the top equation, we see that $\lim_{x\to 4^-} f(x) = 6$——the same answer we got when computing this limit graphically. Now let's take a look at $\lim_{x\to 4^+} f(x)$. Since we are approaching 4 from the right, we will use the bottom equation of the piecewise function as it is the equation that determines what is to the right of the function at $x = 4$. Again, let's just plug in x=4 to the bottom equation of $f(x)$. Doing this, we find that $\lim_{x\to 4^+} f(x) = 2$.

1.4 Limits at Infinity

What value does a function approach as x approaches infinity ($x \to \infty$)? What about when x approaches negative infinity ($x \to -\infty$)? Let's take a look at a few examples.

Example 3. Evaluate $\lim_{x\to\infty} \frac{1}{x^2}$.

Solution. Let's think about fractions for a moment. What happens to the value of a fraction when the denominator grows while the numerator remains constant? The value of the fraction decreases. So, as x grows infinitely, the value of the fraction approaches 0. Any constant number divided by an infinitely large value is equal to 0, so we say $\lim_{x\to\infty} \frac{1}{x^2} = 0$. One may also see this by plugging in increasingly large values of x. Doing so:

$$\begin{aligned} x = 1: &\qquad \frac{1}{1^2} = 1 \\ x = 10: &\qquad \frac{1}{10^2} = 0.01 \\ x = 100: &\qquad \frac{1}{100^2} = 0.0001 \\ x = 1000: &\qquad \frac{1}{1000^2} = 0.000001 \\ x = 10000: &\qquad \frac{1}{10000^2} = 0.00000001. \end{aligned}$$

As x increases to infinity, $\frac{1}{x^2}$ decreases to 0. □

Example 4. Evaluate $\lim_{x\to\infty} e^x$.

Solution. What value does this function approach? To start, note that e is a constant approximately equal to 2.71828. Any positive constant greater than one raised to a large power will result in a very large positive number. So, as x grows infinitely, this function approaches infinity, so we say $\lim_{x\to\infty} e^x = \infty$. Again, one may see this by plugging in increasingly large values of x. □

Example 5. Evaluate $\lim_{x\to\infty} 3^{-\frac{1}{x}}$.

Solution. What value will this interesting function approach as x approaches infinity? Recall our first example: any constant divided by x as x grows infinitely is equal to 0. So, this limit simplifies to 3^0. Any non-zero number raised to the power of 0 is equal to 1, so this function approaches 1 as x grows infinitely, that is, $\lim_{x\to\infty} 3^{-\frac{1}{x}} = 1$. Again, one may see this too by plugging in increasingly large values of x. □

Example 6. Evaluate $\lim_{x\to\infty} -3x^2$.

What value will this simple function approach as x tends towards infinity? If we omit the negative sign, we see that as x grows infinitely, the function $3x^2$ approaches no finite number, it just keeps getting bigger as x keeps getting bigger. This means that $\lim_{x\to\infty} 3x^2 = \infty$. If we "re-insert" the negative sign, we see that $\lim_{x\to\infty} -3x^2 = -\infty$. Once again, this may be seen by plugging in increasingly large values of x.

Example 7. Evaluate $\lim_{x\to-\infty} -3x^2$.

Solution. After our last example, one may be tempted to say that, since we are approaching negative infinity, this will just be the opposite of our last answer, that this limit is equal to positive infinity. However, that is not correct. For this problem, remember that any negative number squared is a positive number. Since x is being squared, this problem is actually identical to the last one, and $\lim_{x\to-\infty} -3x^2 = -\infty$. This can be seen by plugging in negative numbers of increasing size;

$$\begin{aligned} x = 1 : &\qquad -3(1)^2 = -3 \\ x = 10 : &\qquad -3(10)^2 = -300 \\ x = 100 : &\qquad -3(100)^2 = -30000 \\ x = 1000 : &\qquad -3(1000)^2 = -3000000 \\ x = 10000 : &\qquad -3(10000)^2 = -300000000. \end{aligned}$$

As x decreases to negative infinity, so too does $-3x^2$. □

1.5 Does Not Exist

What is the $\lim_{x\to 0} \frac{1}{x}$? or $\lim_{x\to\infty} \sin(x)$? What about $\lim_{x\to 4} \begin{cases} \frac{3}{2}x & x < 4 \\ 2 & 4 \le x \end{cases}$? Let's explore these intriguing questions.

Let's begin with $\lim_{x\to 0} \frac{1}{x}$. If we plug in 0 for x, we see that this function approaches the value of $\frac{1}{0}$ when x is approaching 0. Does this mean that $\lim_{x\to 0} \frac{1}{x} = \infty$ because a constant divided by an infinitely small number is infinity? Not exactly. Until this point, the following has not been emphasized because it was not necessary in any of the problems we have done so far. To determine this limit, as with any other limit, we must ensure that the left and right-handed limits are equal, that is that $\lim_{x\to 0^-} \frac{1}{x} = \lim_{x\to 0^+} \frac{1}{x}$. First, let's compute the left-hand limit. When x is an infinitely small number to the left of 0, it is an infinitely small negative number. This means that $\lim_{x\to 0^-} \frac{1}{x} = -\infty$.

As for the right-handed limit, 1 divided by an infinitely small positive number (since to the right of 0 is positive) will be infinity, that is, $\lim_{x \to 0^+} \frac{1}{x} = \infty$. Since the left-hand limit approaches negative infinity and the right-hand limit approaches positive infinity, the limits on each side are not equal. This means that there is not a single value that this function is approaching when x approaches 0, so, we say that $\lim_{x \to 0} \frac{1}{x}$ Does Not Exist, or "DNE" for short.

Now, let's take a look at that second limit, $\lim_{x \to \infty} \sin(x)$. What value does the function $\sin(x)$ approach as x grows infinitely? Note that the sine function can output any value between (and including) -1 and 1. In other words, the range of the sine function is $[-1, 1]$. So which of these values does the sine function approach as x approaches infinity? Does it approach a different value between those two values? No, this function doesn't approach any value. The sine function, no matter how large x grows, never approaches a single value. It continues infinitely oscillating between -1 and 1. For this reason, we say that $\lim_{x \to \infty} \sin(x) = \text{DNE}$.

What about that third limit? This is a familiar piecewise function, one that we explored when we discussed limits graphically and algebraically on pages 10 and 11. However, when we previously discussed this function $f(x) = \begin{cases} \frac{3}{2}x & x < 4 \\ 2 & 4 \leq x \end{cases}$, we only discussed the left-hand and right-hand limits separately. We found that $\lim_{x \to 4^-} f(x) = 6$ and that $\lim_{x \to 4} f(x) = 2$. So, what is $\lim_{x \to 4} \begin{cases} \frac{3}{2}x & x < 4 \\ 2 & 4 \leq x \end{cases}$? To determine any limit that approaches a number from both sides, we can simply compare the left-hand and right-hand limits. Because $\lim_{x \to 4^-} f(x) \neq \lim_{x \to 4^+} f(x)$, we can deduce that $\lim_{x \to 4} f(x) = \text{DNE}$.

1.6 The Squeeze Theorem

The **Squeeze Theorem**, also known as the **Sandwich Theorem**, is a simple, intuitive theorem. The squeeze theorem is defined with three functions: $f(x)$, $g(x)$, and $h(x)$. It states that if $f(x) \leq g(x) \leq h(x)$, where $f(x) = h(x)$, $g(x)$ must also be equal to $f(x)$ and $h(x)$. This also means that, even if $f(x) \neq h(x)$, if at some point $x = a$ $f(a) = h(a)$, then $g(a) = f(a)$ and $g(a) = h(a)$. This makes sense—suppose that $f(3) = 2 = h(3)$. The inequality for $g(x)$ is now $2 \leq g(3) \leq 2$. The only way that this inequality can be true is if $g(3)$ also equals 2. This is a very simple but extremely important theorem. It is used in many proofs and in computing many limits. This theorem will come in handy when we compute the limits in the following section. The name of this theorem comes from the fact that the function $g(x)$ is being squeezed (or sandwiched) between the two functions $f(x)$ and $h(x)$, and is forced to be equal to them.

1.7 Special Limits

There are a few very special, very common limits you should know. The three most important are:

Special Limits

$$\lim_{x\to 0}\frac{\sin(x)}{x}=1 \qquad \lim_{x\to 0}\frac{1-\cos(x)}{x}=0 \qquad \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^{x}=e\approx 2.71828$$

Before we verify these algebraically, we present a graph of each special limit.

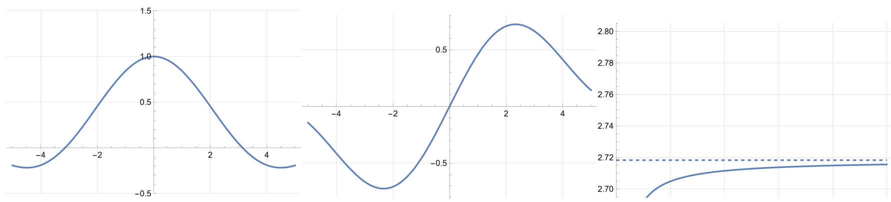


Figure 1.3: Graphs of $\frac{\sin(x)}{x}$, $\frac{1-\cos(x)}{x}$, and $\left(1+\frac{1}{x}\right)^{x}$, respectively. The third plot includes a dashed line representing the number e.

We now go about showing that $\lim_{x\to 0}\frac{\sin(x)}{x}=1$.

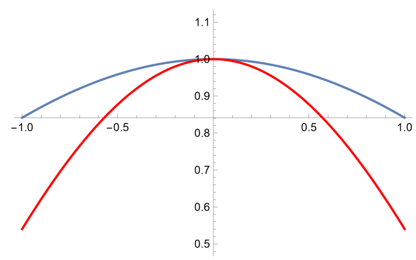


Figure 1.4: A plot of $\frac{\sin(x)}{x}$, shown in blue, and $\cos(x)$, shown in red.

As you can see from Figure 1.4, for x near 0, $\cos(x)\le\frac{\sin(x)}{x}$. Also, note that $\frac{\sin(x)}{x}\le 1$ for x near 0 (it is actually less than or equal to 1 for any x, but we only need it to be less than 1 near $x=0$). With this, we can say that

$$\cos(x)\le\frac{\sin(x)}{x}\le 1 \text{ for } x \text{ near } 0.$$

Then, taking the limit as x approaches 0 of all 3 parts of the equation,

$$\begin{array}{ccccc}
\lim_{x\to 0}\big(\cos(x)\big) & \le & \lim_{x\to 0}\frac{\sin(x)}{x} & \le & \lim_{x\to 0}(1)\\
\cos(0) & \le & \lim_{x\to 0}\frac{\sin(x)}{x} & \le & 1\\
1 & \le & \lim_{x\to 0}\frac{\sin(x)}{x} & \le & 1.
\end{array}$$

Since the leftmost and rightmost parts of the inequality are equal $(1 = 1)$, $\lim_{x \to 0} \frac{\sin(x)}{x} = 1$ by the Squeeze Theorem. A direct result of this is also that $\lim_{x \to 0} \frac{\sin(ax)}{ax} = 1$. This is because we can make a substitution, $y = ax$. As $x \to 0$, $y \to 0$ as well. That means we can write our limit as $\lim_{y \to 0} \frac{\sin(y)}{y} = 1$.

We will now show that $\lim_{x \to 0} \frac{1 - \cos(x)}{x} = 0$. For this, we start with one of the well-known cosine double-angle identities, $\cos(2\theta) = 1 - 2\sin^2(\theta)$. If we subtract 1 from each side, we have $-1 + \cos(2\theta) = -\sin^2(\theta)$. Now, multiplying each side by -1, we are left with $1 - \cos(2\theta) = \sin^2(\theta)$. Then, if we let $x = 2\theta$, dividing each side of this new equation by two we get that $\theta = \frac{x}{2}$. Now, substituting into our trig identity, we have

$$1 - \cos(x) = 2\sin^2\left(\frac{x}{2}\right).$$

With this, our limit has now become $\lim_{x \to 0} \frac{2\sin^2\left(\frac{x}{2}\right)}{x}$. Now, we can split this up into a product of two sine functions:

$$\begin{aligned} \lim_{x \to 0} \frac{2\sin^2\left(\frac{x}{2}\right)}{x} &= \lim_{x \to 0} \left[\left(\sin\left(\frac{x}{2}\right) \right) \left(\frac{2\sin\left(\frac{x}{2}\right)}{x} \right) \right] \\ &= \lim_{x \to 0} \left[\left(\sin\left(\frac{x}{2}\right) \right) \left(\frac{\sin\left(\frac{x}{2}\right)}{\frac{x}{2}} \right) \right]. \end{aligned}$$

Then, since $\lim_{x \to 0} \frac{\sin(ax)}{ax} = 1$, taking $a = \frac{1}{2}$, we have

$$\begin{aligned} \lim_{x \to 0} \frac{2\sin^2\left(\frac{x}{2}\right)}{x} &= \lim_{x \to 0} \left[\left(\sin\left(\frac{x}{2}\right) \right) (1) \right] \\ &= \lim_{x \to 0} \sin\left(\frac{x}{2}\right). \end{aligned}$$

For this, we can just plug in $x = 0$, as is our usual strategy for limits. Doing so, $\sin\left(\frac{0}{2}\right) = 0$, so we have that $\lim_{x \to 0} \frac{1 - \cos(x)}{x} = 0$.

We will save the final special limit as an example later on. We will show that $\lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^x = e$ in Chapter 3, Derivative Applications.

1.8 Continuity and Discontinuity

An important thing to understand in any calculus class is continuity. Continuity, as it may sound, deals with whether or not a function is continuous. Basically, if you can graph a function on a piece of paper without ever lifting your pencil, it is continuous. Here are some examples of functions that are NOT continuous:

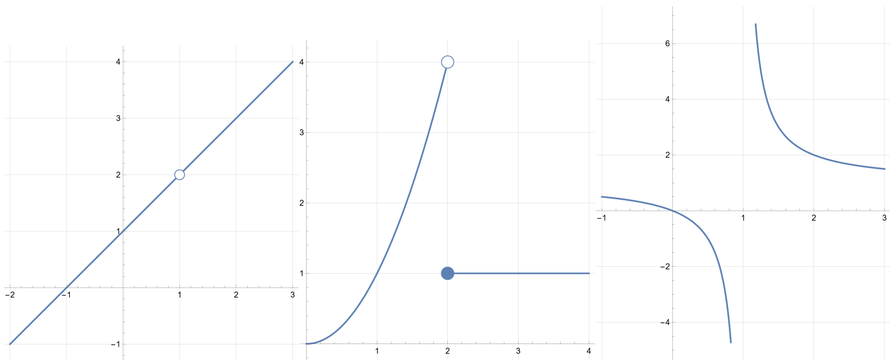


Figure 1.5: Example graphs of discontinuous functions (functions that are not continuous).

How can we classify these 3 types of discontinuities?

1.8.1 Removable Discontinuity

The first type of discontinuity is called a **removable discontinuity**, often informally referred to as a hole. As shown in the first graph above, the graph of the function $f(x) = \dfrac{x^2-1}{x-1}$, at a hole, the value of the function is undefined. Does this mean that $\lim\limits_{x\to 1} \dfrac{x^2-1}{x-1} = \text{DNE}$? No. Algebraically, every function with a hole will break down into $\frac{0}{0}$ at the x value of the hole. This is the reason that the hole exists—you cannot divide 0 by 0, even with limits (it's not necessarily infinity or 0, as we will see). What if we could find a way to get rid of this spot where the function breaks down into $\frac{0}{0}$ without changing the function at any other x value? There is. It's part of the reason this type of discontinuity is referred to as a *removable* discontinuity. To do this, we have to understand what is making the function 0 in the numerator at the same time that the denominator is 0. Let's begin by factoring the numerator of the function. The factored version of this function looks like this: $\frac{(x-1)(x+1)}{(x-1)}$. Is there anything we can "remove" from the function? Yes. Notice the term that is repeated between the numerator and the denominator, $(x-1)$. If we cancel them out, we find that the function is now $x+1$. This is a much easier limit to compute. We now know that $\lim\limits_{x\to 1} x+1 = 2$. Does this seem plausible? Let's take another look at the graph of $\dfrac{x^2-1}{x-1}$.

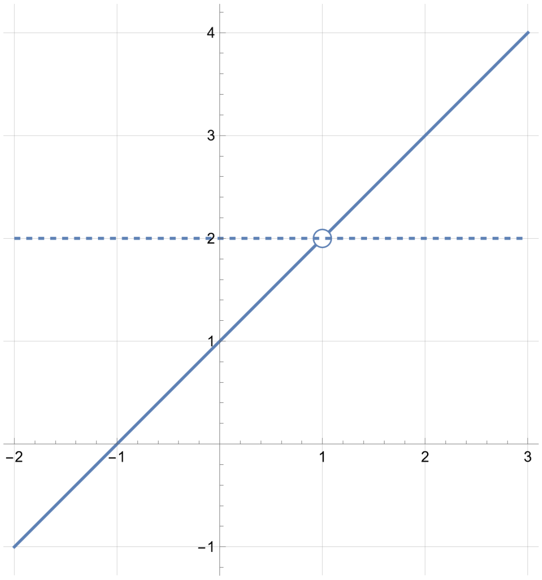


Figure 1.6: Plot of $f(x) = \dfrac{x^2 - 1}{x + 1}$ with a dashed line representing the line $y = 2$.

The dashed line you see is the y value of 2. As you can see, this line perfectly intersects with the function $\dfrac{x^2 - 1}{x - 1}$ at the point $x = 1$. So, graphically, we can confirm that $\lim_{x \to 1} \dfrac{x^2 - 1}{x - 1} = 2$. So, as x approaches 1, the function approaches 2. In general, our limit strategy is now as follows:

1. For a limit $\lim_{x \to a} f(x)$ where $f(x)$ is not a piecewise function, start by plugging in a to $f(x)$. If $f(a)$ is a number, you are done.

2. If after plugging a into $f(x)$, you get an expression of the form $\frac{0}{0}$, factor the numerator and denominator of the fraction as much as possible. Once it is clear what is causing the numerator to be 0 and the denominator to be 0, cancel out these terms to simplify the fraction. Then, plug in a into your new expression. If you have a number, you are done.

1.8.2 Jump Discontinuity

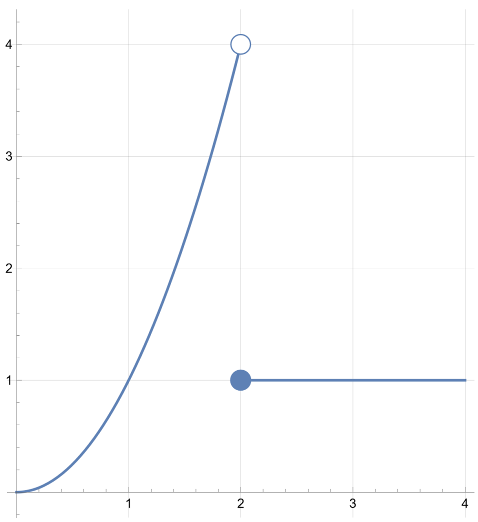


Figure 1.7: Plot of the piecewise function $f(x) = \begin{cases} x^2 & 0 \le x < 2 \\ 1 & 2 \le x \le 4 \end{cases}$.

What about that second picture? What is that discontinuity known as? Does the limit exist for that function at the x value of 2? Let's find out. First, let's take a look at the function itself. This piecewise function is defined as follows:

$$f(x) = \begin{cases} x^2 & 0 \le x < 2 \\ 1 & 2 \le x \le 4 \end{cases}.$$

You can see that when $x = 2$, the function value appears to be jumping from the curve $y = x^2$ to the line $y = 1$. In mathematics, this is referred to as a *jump discontinuity*, or, more commonly, a "jump." As you can probably tell, the function is defined when $x = 2$. You simply plug in 2 to the bottom equation and get that $f(2) = 1$. Does this mean that $\lim_{x \to 2} f(x) = 1$? No, the jump discontinuity helps to highlight something important about limits: **limits are not just another way of expressing the function value**.

As we discussed earlier, limits show the value the function is *approaching*. Because this limit does not specify whether x is approaching 2 from the left or the right, we have to take the limit from both sides and see if they are equal. The $\lim_{x \to 2^-} f(x) = \lim_{x \to 2} x^2 = 4$ and the $\lim_{x \to 2^+} f(x) = \lim_{x \to 2}(1) = 1$. Because clearly $4 \neq 1$, it can be said that $\lim_{x \to 2} f(x) = \text{DNE}$.

In general, our limit strategy for $\lim_{x \to a} f(x)$ is now as follows:

1. If $f(x)$ is a piecewise function (if $f(x)$ is not, proceed to step 2), check that it is continuous. If it is continuous, proceed to step 2. If it is not continuous (it is discontinuous), classify the type of discontinuity. If it is a hole, proceed to step 3. If it is a jump, does the jump occur at $x = a$? If no, proceed to step 2. If the jump occurs at $x = a$ and the limit is not a one-sided limit, the limit does not exist.

2. Start by plugging in a to $f(x)$. If $f(a)$ is a number, you are done. If not, continue to step 3.

3. If after plugging a into $f(x)$, you get an expression of the form $\frac{0}{0}$, factor the numerator and denominator of the fraction as much as possible. Once it is clear what is causing the numerator to be 0 and the denominator to be 0, cancel out these terms to simplify the fraction. Then, plug in a into your new expression. If you have a number, you are done.

1.8.3 Infinite Discontinuity

Now, let's talk about the third type of discontinuity. First, let's take another look at the graph (left).

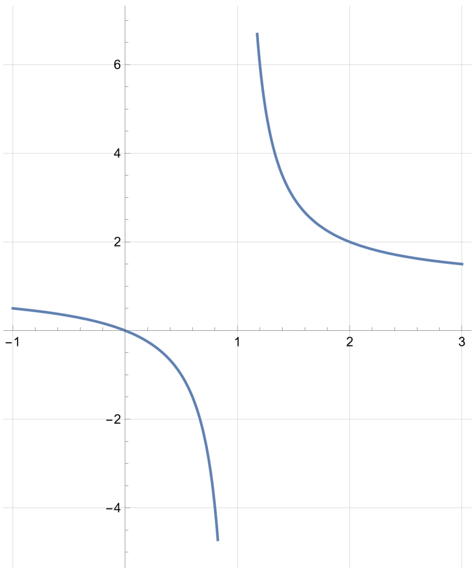


Figure 1.8: Plot of the function $f(x) = \dfrac{x}{x-1}$.

This is the graph of the function $\dfrac{x}{x-1}$. The only interesting thing in this function occurs when $x = 1$, as you can see from the graph. At $x = 1$, there is a *vertical asymptote*, something you may remember from algebra, a vertical line the function is asymptotic to (approaches). We refer to this as an *infinite discontinuity*, since the ends are approaching infinity and the function is discontinuous because of it. Let's compute the $\lim\limits_{x\to 1} \dfrac{x}{x-1}$. Here, plugging in $x = 1$, we get $\frac{1}{0}$. What does this mean? Well, taking another look at the graph, we can see that this limit does not exist—the two sides of $x = 1$ do not approach the same value. Mathematically, $\lim\limits_{x\to 1^-} \dfrac{x}{x-1} = \lim\limits_{x\to 1^+} \dfrac{x}{x-1}$. However, this isn't always the case, even with this type of discontinuity. Let's take a look at another example.

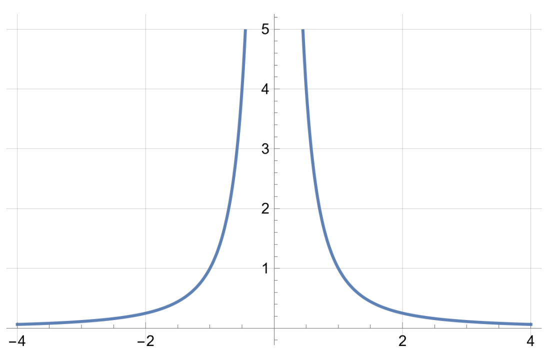


Figure 1.9: A plot of the function $\dfrac{1}{x^2}$.

This is actually the graph of a function whose limit as $x \to 0$ we've already computed in section 1.4, Limits With Infinity. However, when computing that limit, we didn't consider the limit from each side. We start with the process listed at the end of 1.8.2.

1. $f(x)$ is not a piecewise function, so we proceed to step 2.

2. Plugging in $x = 0$ to $f(x)$, we get $f(0) = \frac{1}{0^2} = \frac{1}{0}$. Since this is not a number, we proceed to step 3.

3. Step 3 only helps us if we have something in the form $\frac{0}{0}$, so where can we go from here?

Now, we must define step 4. If $\lim_{x \to a} f(x) = \dfrac{1}{0}$, check whether or not $\lim_{x \to a^-} f(x) = \lim_{x \to a^+} f(x)$. If these limits are equal, that is the limit of the function. If these limits are unequal, the limit does not exist. We now use this final step to evaluate our limit. First, we check the limit as $x \to 0^-$, the limit as x approaches 0 from the left. As $x \to 0^-$, x is becoming an extremely small negative number. Because x is in the denominator of the fraction, as x gets smaller, the value of the fraction increases. So, if x is as small as possible (infinitely close to 0), the fraction is as big as possible (infinity). Now, we just have to decide whether this is a positive or negative infinity. As you can see from the graph, as $x \to 0^-$, the function grows in the positive direction, so $\lim_{x \to 0^-} \dfrac{1}{x^2} = \infty$. But why positive infinity? Why not negative infinity? This is because the denominator is squared, so the negative, in a sense, goes a way. If x was being raised to an odd power, this would limit would approach negative infinity. For example, $\lim_{x \to 0^-} \dfrac{1}{x^5} = -\infty$.

Now, we check the other side of this limit, $\lim_{x \to 0^+} \dfrac{1}{x^2}$. We use the same reasoning as previously. Since the denominator of this fraction is getting endlessly smaller and the numerator is not changing, the value of the fraction is getting endlessly bigger, that is, it is approaching infinity. Since x is positive (since it is a number just to the right of 0) and a positive number squared is also positive, $\lim_{x \to 0^+} \dfrac{1}{x^2} = \infty$. Since our limits on each side are equal, we say that $\lim_{x \to 0} \dfrac{1}{x^2} = \infty$.

We now present our updated general limit strategy:

1. If $f(x)$ is a piecewise function (if $f(x)$ is not, proceed to step 2), check that it is continuous. If it is continuous, proceed to step 2. If it is not continuous (it is discontinuous), classify the type of discontinuity. If it is a hole, proceed to step 3. If it is an infinite discontinuity, proceed to step 4. If it is a jump, does the jump occur at $x = a$? If no, proceed to step 2. If the jump occurs at $x = a$ and the limit is not a one-sided limit, the limit does not exist.

2. Start by plugging in a to $f(x)$. If $f(a)$ is a number, you are done. If it is instead of the form $\frac{0}{0}$, proceed to step 3. If it is in the form of $\frac{n}{0}$ for some number n, proceed to step 4.

3. Factor the numerator and denominator of the fraction as much as possible. Once it is clear what is causing the numerator to be 0 and the denominator to be 0, cancel out these terms to simplify the fraction. Then, return to step 2.

4. Check the left and right limits. If they are equal, the limit exists and is equal to those limits. If they are unequal, the limit does not exist.

1.8.4 Oscillating Discontinuity

We now introduce our final type of discontinuity, the *Oscillating Discontinuity*. This is occasionally also referred to as an *essential discontinuity*, as it cannot be removed from the function (it is essential to the definition of the function). Before we take a look at an example, we outline two important characteristics of an oscillating discontinuity. These will be explained in our example. Those are:

1. The oscillations are bounded.

2. At least one side of the limit does not exist.

Now, for an example. Consider the function $f(x) = \sin(\frac{1}{x})$, shown below.

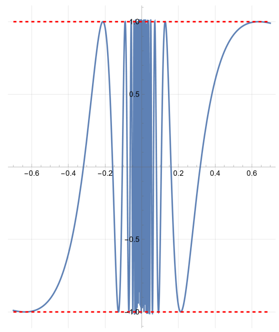


Figure 1.10: A plot of the function $\sin(\frac{1}{x})$.

You can probably see where the name *oscillating* discontinuity comes from—the function oscillates rapidly between -1 and 1 when x is close to 0. As you can see, the function stays nicely bounded between the dashed lines $y = -1$ and $y = 1$ (as is expected, since sine only outputs values between -1 and 1 when it takes real numbers as inputs). We now only have left to ensure that the limit doesn't exist from at least one side of the discontinuity (which occurs at 0, since the only thing undefined in this function is $\sin(\frac{1}{0})$). We will check both sides of the limit, however:

$$\lim_{x \to 0^-} \sin\left(\frac{1}{x}\right) = \sin(-\infty) = \text{DNE}$$
$$\lim_{x \to 0^+} \sin\left(\frac{1}{x}\right) = \sin(\infty) = \text{DNE}.$$

Since the limit does not exist from either side, this is an oscillating discontinuity.

Note: We used some improper notation above. Since infinity is not a number, it is meaningless in an equation without a limit present. The $\sin(\infty)$ is a meaningless statement. What we should really say is that as $x \to 0^+$, $\frac{1}{x} \to \infty$, so, if we take $y = \frac{1}{x}$, then, as $x \to 0^+$, $y \to \infty$, and we can write $\lim_{x \to 0^+} \sin\left(\frac{1}{x}\right) = \lim_{y \to \infty} \sin(y) = \text{DNE}$. For $x \to 0^-$, we follow a very similar process. We will present our general limit strategy after we define Continuity at a Point.

1.8.5 Continuity at a Point

For a given function $f(x)$ to be continuous at a point (a, b), $f(a)$ must be defined and equal to b, and $\lim_{x \to a} f(x)$ must exist and be equal to $f(a)$. This is just a fancy way of saying that, for a function $f(x)$ to be continuous at a point, the limit to that point must exist and be equal to the function value at the point. The limit and the function value being equal ensures there is not a jump discontinuity, and the function being defined at the point ensures there is not a removable discontinuity, infinite discontinuity, or oscillating discontinuity.

We now give a final summary of our general limit strategy.

General Limit Strategy

For a limit $\lim_{x \to a} f(x)$,

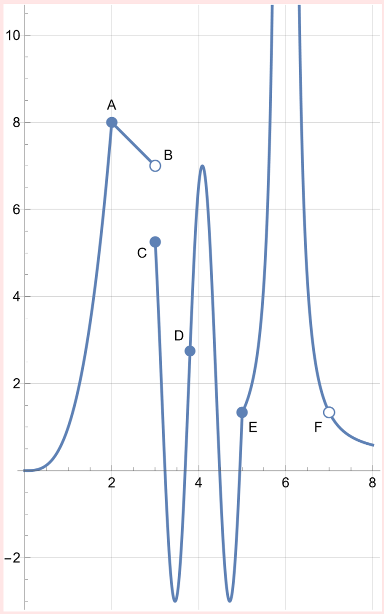
1. Is $f(x)$ a piecewise function? If no, go to step 2. If yes, skip to step 3.
2. Plug in $x = a$ to $f(x)$. If $f(a)$ is a number, you are done. If not, go to step 4.
3. Is there a jump discontinuity at the point $x = a$? If yes, the limit does not exist. If no, plug a in for x. If $f(a)$ is a number, you are done. Otherwise, proceed to step 4.
4. Classify the type of discontinuity.
   a. If it is an oscillating discontinuity, the limit does not exist.
   b. If it is a removable discontinuity, remove it by factoring the numerator and denominator and canceling out like terms. Then, return to step 2.
   c. If it is an infinite discontinuity, check the limit from each side. If they are equal, that is the limit. If they are unequal, the limit does not exist.

1.8.6 Examples

We will now work through some examples. You should try each example on your own, and then read through the solution. If you get stuck, give yourself a hint by reading a part of the solution.

Example 1 - Classifying Discontinuities

Determine whether or not there is a discontinuity at each point labeled A through F on the plot below. If there is, classify the type of discontinuity.



Solution:
There are two conditions for continuity: the limit must exist AND the function must be defined.

First, notice that, at points A, C, D, and E, the circles are closed (filled in), so, by convention, the function is defined at those points. Also notice that B has the same x value as point C, so the function is still defined for the x value of B. Since F is an open circle and there is no other closed circle with the same x value, the function is not defined at F.

We now analyze the limit of each point. At point A, notice that the two functions on either side (a cubic on the left and a linear function on the right) meet. Since they meet at this point, the limit from the left and right will be the same, and thus the limit will exist. As x approaches the x value of the points B and C, from the left, the limit is B, and, from the right, the limit is C. Thus, the limit does not exist. At point D, notice that the function runs right through the point. This is a surefire way to know that the limit exists at point D. This same logic applies to points E and F (and also A), and so the limits exist at those points as well.

Conclusion:
Since the limit exists and the function is defined at points A, D, and E, the function is continuous at those points. At points B and C, the limit does not exist, so there is a discontinuity. Since you have to "jump" from B to get to C (and also to get from C to B), this is a jump discontinuity. At point F, the function is not defined, so the function is discontinuous at F. This is a removable discontinuity (a hole) because it does not fit our definition of any of the other discontinuities.

Note: There is also an infinite discontinuity at $x = 6$, but that does not affect any of the other points A through F. Their continuity does not depend on what is going on elsewhere in the function.

1.9 Intermediate Value Theorem

The **Intermediate Value Theorem**, or IVT for short, is a simple theorem that states that for any function $f(x)$ that is continuous on the closed interval $[a, b]$, the function takes on all values between $f(a)$ and $f(b)$. The only reason this must be a closed interval is so that the function is defined at the endpoints. Otherwise, how could $f(a)$ and $f(b)$ be defined? A more complete definition is as follows:

Intermediate Value Theorem (IVT)

Suppose $f(x)$ is a function continuous on the closed interval $[a, b]$. Then, for some y where $f(a) \leq y \leq f(b)$ or $f(b) \leq y \leq f(a)$, there exists a value c in the interval $[a, b]$ such that $f(c) = y$.

We will first illustrate this with a general picture, then with an example.

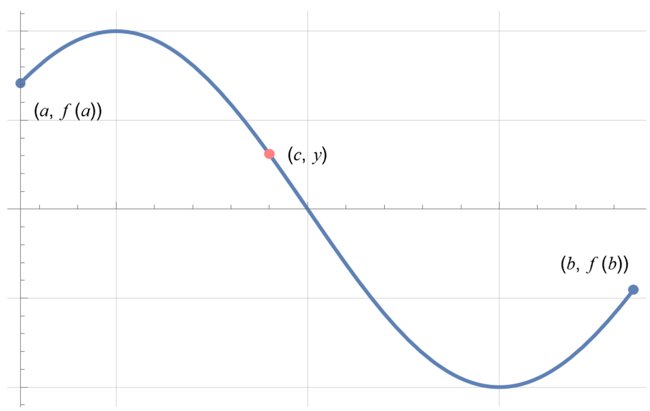


Figure 1.11: A plot of a generic function with endpoints $(a, f(a))$ and $(b, f(b))$, where $a \leq c \leq b$ and y between $f(a)$ and $f(b)$.

Here's an example. Suppose you have some function $f(x)$ that is continuous on the closed interval $[0, 3]$. If $f(0) = 3$ and $f(3) = 7$, must there exist some value c that is $0 \leq c \leq 3$ such that $f(c) = 4$? The answer to this question would be:

The function $f(x)$ is continuous on the interval $[0, 3]$, $f(0) = 3$, and $f(3) = 7$. Since 4 is between 3 and 7, the IVT applies, and it follows that there exists some value c on the interval $[0, 3]$ such that $f(c) = 4$.

1.10 Chapter 1 Summary and Exercises

Chapter Summary

- A limit tells us what a function is approaching as x approaching a given number (or infinity).
- We can analyze limits graphically by looking at what y value the function approaches as x approaches some number.
- We can compute limits algebraically with the process outlined at the end of section 1.8.5.
- As $x \to \infty$, the end behavior of a function is determined by the fastest growing term. Lower ordered polynomial terms can be disregarded.
- If the left and right-handed limits to not agree or if the limit does not approach a single value or infinity (such as with $\sin x$), the limit Does Not Exist.
- If $f(x) \leq g(x) \leq h(x)$ and $f(x) = h(x)$, then $f(x) = g(x) = h(x)$ by the Squeeze Theorem.
- There are three special limits you should know:

$$\lim_{x\to 0}\frac{\sin(x)}{x} = 1 \qquad \lim_{x\to 0}\frac{1-\cos(x)}{x} = 0 \qquad \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e \approx 2.71828$$

- In general, if you can draw a function without lifting your pencil, it is continuous.
- The 4 types of discontinuities are removable, jump, infinite, and oscillating discontinuities.
- A function $f(x)$ is continuous at a point a if $\lim_{x\to a} f(x) = f(a)$.
- In short, the Intermediate Value Theorem states that if a function is continuous on a closed interval $[a, b]$, then for every y between $f(a)$ and $f(b)$, there exists a $x = c$ in the interval $[a, b]$ such that $f(c) = y$.

1.10.1 Chapter 1 Exercises

Chapter 2

Derivatives

2.1 What is a Derivative?

For functions of the form $f(x) = mx + b$, we can easily tell what the slope at any point x is. However, for functions like x^2, $\sin(x)$, e^x and $\log_3(x)$, we cannot. This is where the derivative comes in. The derivative is a way of defining the slope at any given point. For example, on the graph of the function x^2, the slope at the point $x = 3$ has a value of 6 (we'll talk about why in just a moment). This means that the tangent line[1] at the point $x = 3$ has a slope of 6. So, the tangent line at the point would be $y - 9 = 6(x - 3)$. This will make more sense in the upcoming sections.

2.2 Limit Definition of the Derivative

We begin with the definition of rate of change from Algebra I, the change in y, Δy divided by the change in x, Δx, giving us the change in y per unit change in x:

$$\textit{Rate of change} = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}.$$

Now, let the distance between x_2 and x_1 be h, that is, $\Delta x = x_2 - x_1 = h$, as shown in the picture (left).

[1] A tangent line is a line that only touches the function at one point instead of crossing through the function.

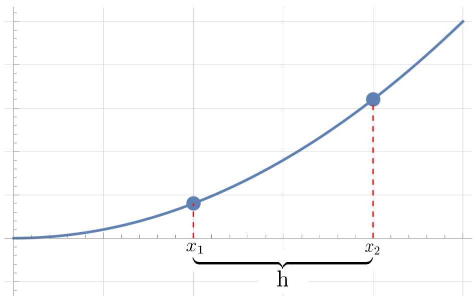


Figure 2.1: A plot of a function with two points x_1 and x_2 labeled with distance h between them.

With this, the new formula for the rate of change becomes

$$\textit{Rate of change} = \frac{y_2 - y_1}{h}.$$

In function notation, $y_2 = f(x_2)$ and $y_1 = f(x_1)$. With this, we have the formula

$$\textit{Rate of change} = \frac{f(x_2) - f(x_1)}{h}.$$

Now, remember that we defined $h = x_2 - x_1$. This means that, adding x_1 to each side, $x_1 + h = x_2$. Making this substitution,

$$\textit{Rate of change} = \frac{f(x_1 + h) - f(x_1)}{h}.$$

Remember that the goal of the derivative is to define the rate of change at a single point (called instantaneous rate of change). To do this, the distance between our two points, h, must be 0. However, if we make this replacement, we are left with $\frac{f(x_1+0)-f(x_1)}{0} = \frac{0}{0}$. This is undefined, but, using limits, we can find what value this slope function approaches as $h \to 0$. With this, we define the slope at a point x_1 for a function $f(x)$, denoted $f'(x_1)$[2]:

$$f'(x_1) = \lim_{h \to 0} \frac{f(x_1 + h) - f(x_1)}{h}.$$

We can generalize this to find an expression for the slope at any point by instead evaluating f' at a generic x. This results in the definition of the derivative[3], as follows.

Definition of the Derivative

For a function $f(x)$ satisfying certain properties (to be discussed shortly), the derivative of $f(x)$ is

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

We can verify this using a basic function in the format $y = mx + b$. Take the function $f(x) = 3x + 6$, for example. From our knowledge of linear equations, we know that this function has a slope of 3. We will verify this with our new formula:

$$f'(x) = \lim_{h \to 0} \frac{(3(x + h) + 6) - (3x + 6)}{h} = \lim_{h \to 0} \frac{\cancel{3x} + 3h + \cancel{6} - \cancel{3x} - \cancel{6}}{h} = \lim_{h \to 0} \frac{3\cancel{h}}{\cancel{h}} = 3.$$

From this, we can determine that the slope at any given point of the function $f(x) = 3x + 6$ is 3.

[2] You will often see the derivative denoted as $f'(x)$, read "f-prime of x." The mathematical function that represents the derivative is $\frac{\mathrm{d}}{\mathrm{d}x}$, read "dee-dee x," meaning take the derivative with respect to x. The derivative of a function y is often denoted y', read "y prime."

[3] This definition of the derivative is often referred to as the "First Principles" definition or "Limit Definition" of the Derivative.

We can be more general with this and show that the slope of $f(x) = mx + b$ is m:

$$f'(x) = \lim_{h\to 0} \frac{m(x+h) + b - (mx+b)}{h} = \lim_{h\to 0} \frac{\cancel{mx} + mh + \cancel{b} - \cancel{mx} - \cancel{b}}{h} = \lim_{h\to 0} \frac{m\cancel{h}}{\cancel{h}} = m.$$

And thus, the derivative of a linear function $f(x) = mx + b$ is m.

What if we test this using a function we don't know the slope of? Take the function $f(x) = x^2$, for example. Using our limit definition of the derivative and the fact that $(a+b)^2 = a^2 = 2ab + b^2$,

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

This means that the slope at any given point on the graph of x^2 is 2 multiplied by the value of x at that point. Back to the example at the beginning of this chapter, the derivative of $f(x) = x^2$ at the point $x = 3$ is $f'(3) = 2(3) = 6$. How can we verify this? By graphing the tangent line[4] with a graph of the original function, we can verify whether or not this is true. The graph of the functions can be found below.

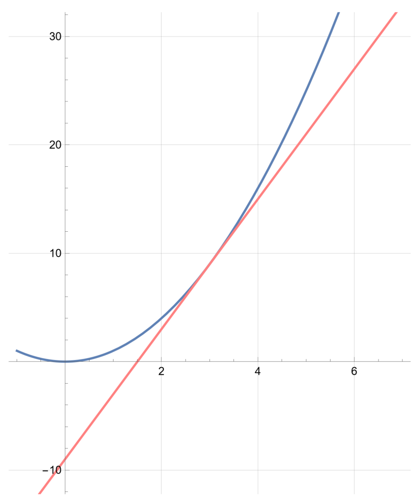


Figure 2.2: The function $f(x) = x^2$, shown in blue, and its tangent line at the point $x = 3$, shown in red.

2.3 Differentiability

2.4 Derivative Shortcuts

Of course, doing the limit definition of a derivative every time you want to take the derivative of something would be very tedious. Luckily, there are shortcuts. How these shortcuts are derived will not be shown immediately, but, as we discover more calculus techniques, derivations will be shown for many of these tricks.

[4]We can graph a tangent line using "point-slope form", $y - y_1 = m(x - x_1)$, for a slope m and point (x_1, y_1). In function notation with derivatives, we write $f(x) - f(a) = f'(a)(x - a)$ for the tangent line to a function f at a point $a, f(a)$.

2.4.1 Properties of the Derivative

There are two very important properties of the derivative. These will make taking derivatives much easier, and there are still more tricks to come.

1. For differentiable functions $f(x)$ and $g(x)$, $\frac{\mathrm{d}}{\mathrm{d}x}\Big[f(x)+g(x)\Big] = \frac{\mathrm{d}}{\mathrm{d}x}f(x) + \frac{\mathrm{d}}{\mathrm{d}x}g(x)$.
2. For a differentiable function $f(x)$ and constant n, $\frac{\mathrm{d}}{\mathrm{d}x}\Big[n \cdot f(x)\Big] = n \cdot \frac{\mathrm{d}}{\mathrm{d}x}f(x)$.

We will use these properties and provide examples in the following sections.

2.4.2 Power Rule

Power Rule for Differentiation

For a function $f(x)$ of the form x^n, the derivative is $f'(x) = nx^{n-1}$. Using property 2 of the derivative, for a function $f(x) = ax^n$ where a is a constant, we have that $\frac{\mathrm{d}}{\mathrm{d}x}f(x) = a \cdot \frac{\mathrm{d}}{\mathrm{d}x}\Big[x^n\Big] = anx^{n-1}$.

This will be proven in Section 2.4.8. We will now perform some examples of this rule. Starting simply, consider the function $f(x) = x^2$. We took the derivative of this function in section 2.2 and got $2x$, so we expect to see $f'(x) = 2x$. We will now use the power rule to differentiate x^2. In this equation, $n = 2$, so $f'(x) = 2x^{2-1} = 2x$, as expected. We will now attempt a more difficult example.

Given $g(x) = 2x^3 - \sqrt[3]{x} + \frac{1}{x^2} - x^{3\pi^2}$, find $g'(x)$.

This probably looks very overwhelming. There are a lot of different components to this function, and many are not in a form where it is easy to apply the power rule. We must stay calm, and apply our properties of derivatives and rules of algebra, then we will have a simple derivative left at the end. First, we apply property 1 of the derivative:

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big[2x^3 - \sqrt[3]{x} + \frac{1}{x^2} - x^{3\pi^2}\Big] = \frac{\mathrm{d}}{\mathrm{d}x}\Big[2x^3\Big] + \frac{\mathrm{d}}{\mathrm{d}x}\Big[-\sqrt[3]{x}\Big] + \frac{\mathrm{d}}{\mathrm{d}x}\Big[\frac{1}{x^2}\Big] + \frac{\mathrm{d}}{\mathrm{d}x}\Big[-x^{3\pi^2}\Big].$$

We will now evaluate each of these derivatives separately, and then combine them again at the end. Firstly, $\frac{\mathrm{d}}{\mathrm{d}x}\Big[2x^3\Big] = 2\frac{\mathrm{d}}{\mathrm{d}x}\Big[x^3\Big]$ by property 2 of the derivative. We now have $n = 3$, so $2\frac{\mathrm{d}}{\mathrm{d}x}\Big[x^3\Big] = 2(3)x^{3-1} = 6x^2$. For our second derivative, we use the second property of derivatives and write $\frac{\mathrm{d}}{\mathrm{d}x}\Big[-\sqrt[3]{x}\Big] = -\frac{\mathrm{d}}{\mathrm{d}x}\Big[\sqrt[3]{x}\Big]$. We then use the fact that $\sqrt[b]{x^a} = x^{\frac{a}{b}}$. Then, we have

$$-\frac{\mathrm{d}}{\mathrm{d}x}\Big[\sqrt[3]{x}\Big] = -\frac{\mathrm{d}}{\mathrm{d}x}\Big[x^{\frac{1}{3}}\Big] \implies n = \frac{1}{3}, \text{ so: } -\frac{\mathrm{d}}{\mathrm{d}x}\Big[x^{\frac{1}{3}}\Big] = -\frac{1}{3}x^{\frac{1}{3}-1} = -\frac{1}{3}x^{-\frac{2}{3}}.$$

For the third derivative, there is no coefficient, so we do not have to apply either property of the derivative. Instead, we apply the rule of algebra that $\frac{1}{x^a} = x^{-a}$. With this, we can rewrite $\frac{1}{x^2} = x^{-2}$. With this, we have $n = -2$, and it follows that

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big[\frac{1}{x^2}\Big] = \frac{\mathrm{d}}{\mathrm{d}x}\Big[x^{-2}\Big] = -2x^{-2-1} = -2x^{-3}.$$

Now, for our final derivative of this example, we apply the second property of derivatives, and have that $\frac{\mathrm{d}}{\mathrm{d}x}\left[-x^{3\pi^2}\right] = -\frac{\mathrm{d}}{\mathrm{d}x}\left[x^{3\pi^2}\right]$. While this looks very complex, notice that since 3, π, and 2 are all numbers, $3\pi^2$ is also a number, so we can apply the power rule with $n = 3\pi^2$:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[-x^{3\pi^2}\right] = -\frac{\mathrm{d}}{\mathrm{d}x}\left[x^{3\pi^2}\right] = -3\pi^2 x^{3\pi^2-1}.$$

We now combine all of these derivatives (by the first property of the derivative) and see that (using some rules of algebra to rewrite our expression), that

$$g'(x) = 6x^2 - \frac{1}{3}x^{-\frac{2}{3}} - 2x^{-3} - 3\pi^2 x^{3\pi^2-1} = 6x^2 - \frac{1}{3\sqrt[3]{x^2}} - \frac{2}{x^3} - 3\pi^2 x^{3\pi^2-1}.$$

If any of those algebraic manipulations are unclear, review section 0.3.

For our second and final example of this section, we will consider the function $f(x) = \frac{x^3 + x^2 + 1}{x}$. This may look like something we don't yet know how to solve. If you look back at section 0.3, you won't see any trick for this either. So what can we do? Well, we first must remember how to add fractions. When we have two fractions with the same denominator, we can simply add the numerator. Here, we will do that in reverse:

$$\frac{x^3 + x^2 + 1}{x} = \frac{x^3}{x} + \frac{x^2}{x} + \frac{1}{x} = x^2 + x + \frac{1}{x}.$$

Using property 1 of the derivative, we consider these as 3 separate derivatives. For x^2, we have computed this derivative multiple times: $\frac{\mathrm{d}}{\mathrm{d}x}x^2 = 2x$. For x, this is a linear function $y = mx + b$ with $m = 1$ and $b = 0$, so its slope (derivative) is 1. This agrees with the power rule: $x = x^1$, so we have $n = 1$. Then, $\frac{\mathrm{d}}{\mathrm{d}x}x = 1 \cdot x^{1-1} = 1 \cdot x^0 = 1$. We have one final derivative to consider: $\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{1}{x}\right]$. Here, we can use the algebra rule that $\frac{1}{x^a} = x^{-a}$. Here, $a = 1$, so $\frac{1}{x} = x^{-1}$. Then, $\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{1}{x}\right] = \frac{\mathrm{d}}{\mathrm{d}x}x^{-1} = -x^{-1-1} = -x^{-2} = -\frac{1}{x^2}$. This gives us our answer,

$$\frac{\mathrm{d}}{\mathrm{d}x}f(x) = 2x + 1 - \frac{1}{x^2}.$$

2.4.3 Product Rule

Product Rule for Differentiation

For a function $F(x) = f(x) \cdot g(x)$ where f and g are both differentiable,

$$F'(x) = f(x)g'(x) + g(x)f'(x).$$

This rule looks more difficult than it is. Essentially, it says that if you have two functions being multiplied and want to take the derivative, write down the first function and multiply it by the derivative of the second function, then add to that the second function multiplied by the derivative of the first function (or in the opposite order). Before we get into any examples, we present a proof of this relation.

Proof. Let $F(x) = f(x)g(x)$ with $f(x)$ and $g(x)$ both differentiable. By the definition of the derivative,

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

Observe that $F(x+h) = f(x+h)g(x+h)$. Then,

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

Now, note that $f(x+h)g(x) - f(x+h)g(x) = 0$, so adding this term to the numerator will not change the result. Doing so,

$$\begin{aligned}
F'(x) &= \lim_{h\to 0} \frac{f(x+h)g(x+h) + -f(x+h)g(x) + f(x+h)g(x) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+\overset{0}{\cancel{h}})\Big(g(x+h) - g(x)\Big) + g(x)\Big(f(x+h) - f(x)\Big)}{h} \\
&= \lim_{h\to 0} \frac{f(x)\Big(g(x+h) - g(x)\Big)}{h} + \frac{g(x)\Big(f(x+h) - f(x)\Big)}{h} \\
&= f(x)\lim_{h\to 0} \frac{g(x+h) - g(x)}{h} + g(x)\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \\
F'(x) &= f(x)g'(x) + g(x)f'(x).
\end{aligned}$$

□

2.4.4 Chain Rule

Chain Rule for Differentiation

For a function $F(x) = (f \circ g)(x) = f\Big[g(x)\Big]$ with differentiable f and g,

$$F'(x) = f'\Big[g(x)\Big]g'(x).$$

This is also commonly written

$$\frac{\mathrm{d}}{\mathrm{d}x} f(u) = f'(u)\frac{\mathrm{d}u}{\mathrm{d}x}.$$

This is one of the most important rules in calculus and is essential to success in calculus. It states that if you have a function composed of another function, you can simply take the derivative of the "outer" function evaluated at the "inner" function and multiply it by the derivative of this inner function. Proofs of the chain rule are generally beyond the scope of calculus classes, however, one is included at the back of the book. This is easier to see with examples. We will do one now, a few at the end of this section, and many at the end of this chapter.

Consider the function $h(x) = (x^4 - 5x^3 + 2x^2 + 3x - 7)^3$. Find $h'(1)$.

It is possible to take this derivative using only the power rule, or, if you're feeling especially adventurous, with only the limit definition of the derivative. Doing either of these options would involve an extensive amount of algebra, and doing the limit definition would involve expanding an $(x+h)^{12}$ term! Clearly, these are not good options for this function as they would take quite some time to compute by hand. Instead, we can use the chain rule. Notice that the function $h(x)$ can be thought of as a composition of two functions, $f(x) = x^3$ and $g(x) = x^4 - 5x^3 + 2x^2 + 3x - 7$. Then,

$$h(x) = f(g(x)) = f(x^4 - 5x^3 + 2x^2 + 3x - 7) = (x^4 - 5x^3 + 2x^2 + 3x - 7)^3.$$

This means we can use the chain rule. To do so, we need to know what $f'(x)$ and $g'(x)$ are. Let's compute these derivatives. For $f(x) = x^3$, this is simply a power rule with $n = 3$: $f'(x) = 3x^{3-1} = 3x^2$. As for $g(x)$, we split this into 5 power rules:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x} &\quad x^4 = 4x^3 \\
\frac{\mathrm{d}}{\mathrm{d}x} &\quad \left[-5x^3\right] = -5\frac{\mathrm{d}}{\mathrm{d}x}x^3 = -15x^2 \\
\frac{\mathrm{d}}{\mathrm{d}x} &\quad 2x^2 = 2\frac{\mathrm{d}}{\mathrm{d}x}x^2 = 4x \\
\frac{\mathrm{d}}{\mathrm{d}x} &\quad 3x = 3\frac{\mathrm{d}}{\mathrm{d}x}x = 3 \\
\frac{\mathrm{d}}{\mathrm{d}x} &\quad \left[-7\right] = -7\frac{\mathrm{d}}{\mathrm{d}x}x^0 = 0
\end{aligned}$$

So we have that $g'(x) = 4x^3 - 15x^2 + 4x + 3$ (by property 1 of the derivative). We can now put this all together:

$$\begin{aligned}
h'(x) &= f'\left[g(x)\right]g'(x) = 3\left[g(x)\right]^2\left(4x^3 - 15x^2 + 4x + 3\right) \\
h'(x) &= 3(x^4 - 5x^3 + 2x^2 + 3x - 7)^2(4x^3 - 15x^2 + 4x + 3) \\
h'(1) &= 3\left(1 - 5 + 2 + 3 - 7\right)^2\left(4 - 15 + 4 + 3\right) = -432
\end{aligned}$$

So the slope of the line tangent to $h(x)$ at $x = 1$ is $h'(1) = -432$. In other words, the instantaneous rate of change of $h(x)$ at $x = 1$ is $h'(1) = -432$.

2.4.5 Quotient Rule

Quotient Rule for Differentiation

For a function $F(x) = \dfrac{f(x)}{g(x)}$ where f and g are both differentiable,

$$F'(x) = \frac{g(x)f'(x) - f(x)g'(x)}{g(x)^2}.$$

Again, this looks quite complicated, but it is actually a very nice rule that can speed up the process of taking derivatives. A way to remember this formula is by saying to yourself "low d high minus high d low all over low squared," where "d" represents taking the derivative, "low" represents the denominator, and "high" represents the numerator. We will now go about proving this using the product, power, and chain rules (this can be proven with only the limit definition of the derivative, but that is far more difficult).

Proof. Let $F(x) = \dfrac{f(x)}{g(x)}$ for differentiable functions $f(x)$ and $g(x)$. Then, using the fact that $\dfrac{1}{a} = a^{-1}$, we can rewrite $F(x)$ as $f(x) \cdot g(x)^{-1}$. Now, we can apply the product rule:

$$F'(x) = f(x) \cdot \left(g(x)^{-1}\right)' + g(x)^{-1} \cdot f'(x).$$

From the power rule and chain rule, we know that $\left(g(x)^{-1}\right)' = -g(x)^{-2} \cdot g'(x)$. With this, we have

$$\begin{aligned} F'(x) &= \frac{f(x) \cdot -g'(x)}{g(x)^2} + \frac{f'(x)}{g(x)} = \frac{f(x) \cdot -g'(x)}{g(x)^2} + \frac{f'(x) \cdot g(x)}{g(x)^2} \\ &= \frac{-f(x)g'(x)}{g(x)^2} + \frac{f'(x)g(x)}{g(x)^2} \\ &= \frac{g(x)f'(x) - f(x)g'(x)}{g(x)^2}. \end{aligned}$$

□

An example now follows. Let $p(x) = \dfrac{x^2 + x + 1}{x^4 + x^3}$. Find $p'(x)$. In this example, $f(x) = x^2 + x + 1$ and $g(x) = x^4 + x^3$. Using the power rule, we can find that:

$$f'(x) = 2x + 1$$
$$g'(x) = 4x^3 + 3x^2$$

Now we simply plug in our $f, f', g,$ and g' into the quotient rule:

$$\begin{aligned} p'(x) &= \frac{(x^4 + x^3)(2x + 1) - (x^2 + x + 1)(4x^3 + 3x^2)}{(x^4 + x^3)^2} \\ p'(x) &= \frac{(x^2 + x)(2x + 1) - (x^2 + x + 1)(4x + 3)}{(x^3 + x^2)^2}. \end{aligned}$$

2.4.6 Known Derivatives

There are 16 "must-know" derivatives. These will come in handy often and will be a pain to derive each time you have to use them. We will present derivations for some of them here, and the rest in the back of the book. A table will be included at the end of this section.

We begin with the derivative of $\sin(x)$. From the limit definition of the derivative,

$$\frac{\mathrm{d}}{\mathrm{d}x} \sin(x) = \lim_{h \to 0} \frac{\sin(x + h) - \sin(x)}{h}.$$

Then, applying $\sin(\alpha \pm \beta) = \sin(\alpha)\cos(\beta) \pm \cos(\alpha)\sin(\beta)$, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x} \sin(x) &= \lim_{h \to 0} \frac{\sin(x)\cancelto{1}{\cos(h)} + \cos(x)\sin(h) - \sin(x)}{h} \\ &= \lim_{h \to 0} \frac{\cancel{\sin(x)} + \cos(x)\sin(h) \cancel{- \sin(x)}}{h} \\ &= \cos(x) \lim_{h \to 0} \frac{\sin(h)}{h} = \cos(x). \end{aligned}$$

In the last step, we applied one of the special limits from 1.7. The method for differentiating cosine is similar, so it will not be shown here (it is at the back of the book). Note that $\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) = -\sin(x)$.

Using these two derivatives, we now have enough to differentiate $\tan(x)$ using the quotient rule. Since $\tan(x) = \frac{\sin(x)}{\cos(x)}$, we have $f(x) = \sin(x)$ and $f'(x) = \cos(x)$ along with $g(x) = \cos(x)$ and $g'(x) = -\sin(x)$. Using the quotient rule, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan(x) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\sin(x)}{\cos(x)} = \frac{\cos^2(x)+\sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)} = \sec^2(x).$$

We will now differentiate $\sec(x)$. Note that $\sec(x) = \frac{1}{\cos(x)}$. This means we can use the quotient rule with $f(x) = 1, f'(x) = 0$ and $g(x) = \cos(x), g'(x) = -\sin(x)$. Plugging this into our formula,

$$\frac{\mathrm{d}}{\mathrm{d}x}\sec(x) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{\cos(x)} = \frac{0-(-\sin(x))}{\cos^2(x)} = \frac{1}{\cos(x)}\frac{\sin(x)}{\cos(x)} = \sec(x)\tan(x).$$

The derivatives for cosecant and cotangent are performed in a similar manner. Those are:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\cot(x) &= -\csc^2(x)\\ \frac{\mathrm{d}}{\mathrm{d}x}\csc(x) &= -\csc(x)\cot(x).\end{aligned}$$

We now move away from trig derivatives for a moment. What is the derivative of $\ln x$? We start with the limit definition of the derivative:

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln x = \lim_{h\to 0}\frac{\ln(x+h)-\ln(x)}{h} = \lim_{h\to 0}\frac{\ln\left(\frac{x+h}{x}\right)}{h} = \lim_{h\to 0}\frac{1}{h}\ln\left(1+\frac{h}{x}\right) = \lim_{h\to 0}\ln\left(\left(1+\frac{h}{x}\right)^{\frac{1}{h}}\right)$$

Since the entire limit is inside the log, we can move the notation for the limit inside the natural logarithm, as such

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln x = \ln\lim_{h\to 0}\left(1+\frac{h}{x}\right)^{\frac{1}{h}}.$$

Now, note that

$$e^y = \lim_{n\to 0}(1+ny)^{\frac{1}{n}} \text{ which implies that } e^{\frac{1}{x}} = \lim_{n\to 0}\left(1+\frac{n}{x}\right)^{\frac{1}{n}}.$$

Letting $n = h$, we see that

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln x = \ln\left(e^{\frac{1}{x}}\right) = \frac{1}{x}.$$

Here, we used a variation of the third special limit from section 1.7. We will consider these limits more in the next Chapter. Using this result and the log change of base formula, you should verify that $\frac{\mathrm{d}}{\mathrm{d}x}\log_a(x) = \frac{1}{x\ln(a)}$.

We will now take the derivative of e^x. To begin, remember that, from section 1.7, $e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$. There is another common definition for the number e which is the one we will need for finding the derivative of e^x. To find this formula, consider letting $n = \frac{1}{m}$. Then, multiplying each side by $\frac{m}{n}$, we are left with $m = \frac{1}{n}$. We now wish to substitute this into our limit, but, notice that if we were to do so, there would no longer be an n in the problem, which is the variable that is involved in the limit. To resolve this issue, notice from the equation $m = \frac{1}{n}$ that as $n \to \infty$, $m \to 0$. We now have everything we need to make some replacements and get an alternate definition of the number e:

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \lim_{m\to 0}(1+m)^{\frac{1}{m}}.$$

We will now set this aside for a moment and utilize the limit definition of the derivative,

$$\frac{\mathrm{d}}{\mathrm{d}x}e^x = \lim_{h\to 0}\frac{e^{x+h}-e^x}{h} = \lim_{h\to 0}\frac{e^x\left(e^h-1\right)}{h} = e^x\lim_{h\to 0}\frac{e^h-1}{h}.$$

We now proceed by making a substitution. Let $m = e^h - 1$. Then, solving for e^h, we have $e^h = m+1$, so, taking the natural logarithm of each side, $h = \ln(m+1)$. Then, since h is going to 0, we have $\ln(m+1) \to 0$. The only way for this to be true is if m is also going to 0, since the only time $\ln(a) = 0$ is when $a = 1$. If we now put m into our equation, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}e^x = e^x\lim_{m\to 0}\frac{m}{\ln(1+m)} = e^x\lim_{m\to 0}\frac{1}{\frac{1}{m}\ln(1+m)} = e^x\lim_{m\to 0}\frac{1}{\ln\left[(1+m)^{\frac{1}{m}}\right]}.$$

As in the last problem, since the limiting variable (m) is now entirely inside the natural log, we can bring in the limit and say that

$$\frac{\mathrm{d}}{\mathrm{d}x}e^x = e^x\frac{1}{\ln\left[\lim\limits_{m\to 0}(1+m)^{\frac{1}{m}}\right]}.$$

But we already showed that that limit is e, so we are left with

$$\frac{\mathrm{d}}{\mathrm{d}x}e^x = e^x\frac{1}{\ln(e)} = e^x.$$

Now, try to show that $\frac{\mathrm{d}}{\mathrm{d}x}a^x = a^x\ln(a)$. If you get stuck, try rewriting a^x as $e^{\ln(a^x)}$.

If some of these derivations were unclear, whether in this section or in one of the preceding sections (product and quotient rules), that's okay. What matters for now is that you understand how and when to use these rules.

We will now summarize these results in the table below. Note that there are 6 more derivatives written. These will be verified in Chapter 7.

Function	Derivative	Function	Derivative
$\sin(x)$	$\cos(x)$	$\cos(x)$	$-\sin(x)$
$\tan(x)$	$\sec^2(x)$	$\cot(x)$	$-\csc^2(x)$
$\sec(x)$	$\sec(x)\tan(x)$	$\csc(x)$	$-\csc(x)\cot(x)$
$\ln(x)$	$\frac{1}{x}$	$\log_a(x)$	$\frac{1}{x\ln(a)}$
e^x	e^x	a^x	$a^x\ln(a)$
$\sin^{-1}(x)$	$\frac{1}{\sqrt{1-x^2}}$	$\cos^{-1}(x)$	$\frac{-1}{\sqrt{1-x^2}}$
$\sec^{-1}(x)$	$\frac{1}{\lvert x\rvert\sqrt{x^2-1}}$	$\csc^{-1}(x)$	$\frac{-1}{\lvert x\rvert\sqrt{x^2-1}}$
$\tan^{-1}(x)$	$\frac{1}{x^2+1}$	$\cot^{-1}(x)$	$\frac{-1}{x^2+1}$

Table 2.1: A table of derivatives you should know. Note that $\sin^{-1}(x) = \arcsin(x)$, the inverse of sine (similar for the other trig functions), not the reciprocal $\frac{1}{\sin(x)}$, which will be denoted as $\Big(\sin(x)\Big)^{-1}$ or simply $\csc(x)$.

2.4.7 Implicit Differentiation

2.4.8 Logarithmic Differentiation

We've covered methods to differentiate almost every type of common function, but there are still a couple of questions left open. How do we take the derivative of something like $y = x^x$? How do we know that $\frac{\mathrm{d}}{\mathrm{d}x}x^n = nx^{n-1}$? This is our motivation for the method of *logarithmic differentiation.*

In general, logarithmic differentiation is most useful whenever we want to differentiate something of the form $y = f(x)^{g(x)}$. If we take the natural log[5] of each side, we are left with

$$\begin{aligned}\ln(y) &= \ln\Big[f(x)^{g(x)}\Big]\\ \ln(y) &= g(x)\cdot\ln[f(x)].\end{aligned}$$

[5]You could also use any other type of log, but the natural log is the "nicest" since it has the simplest derivative of all the types of logs.

Where we used the fact that $\ln(a^b) = b\ln(a)$ in the second line. From here, we now take the derivative using the product rule and chain rule:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\ln(y) &= \frac{\mathrm{d}}{\mathrm{d}x}\Big[g(x)\cdot\ln[f(x)]\Big] \\
\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} &= g(x)\cdot\frac{1}{f(x)}f'(x) + \ln[f(x)]\cdot g'(x) \\
\frac{\mathrm{d}y}{\mathrm{d}x} &= y\bigg(g(x)\cdot\frac{1}{f(x)}f'(x) + \ln[f(x)]\cdot g'(x)\bigg) \\
\frac{\mathrm{d}y}{\mathrm{d}x} &= f(x)^{g(x)}\bigg(g(x)\cdot\frac{1}{f(x)}f'(x) + \ln[f(x)]\cdot g'(x)\bigg) \\
\frac{\mathrm{d}}{\mathrm{d}x}f(x)^{g(x)} = \frac{\mathrm{d}y}{\mathrm{d}x} &= f(x)^{g(x)}\left(\frac{g(x)f'(x)}{f(x)} + \ln[f(x)]\cdot g'(x)\right). \qquad (2.1)
\end{aligned}$$

Do not worry, this is not a formula you will ever be required to remember. If you can write down the first step (taking the log of each side), you will be fine. After that, it is just product and chain rules. This will be easier to see with an example.

Let $y = x^x$. Find y'.

We begin by taking the natural log of each side. This leaves us with $\ln(y) = \ln\big[x^x\big] = x\ln(x)$. We now take the derivative,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\ln(y) &= \frac{\mathrm{d}}{\mathrm{d}x}\Big[x\ln(x)\Big] \\
\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} &= x\cdot\frac{1}{x} + \ln(x) \\
\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} &= 1 + \ln(x) \\
\frac{\mathrm{d}y}{\mathrm{d}x} &= y\big(1 + \ln(x)\big) \\
y' = \frac{\mathrm{d}y}{\mathrm{d}x} &= x^x\Big(1 + \ln(x)\Big).
\end{aligned}$$

And we have thus differentiated x^x. You should verify that this result agrees with (2.1).

Let's now return to the power rule, $\frac{\mathrm{d}}{\mathrm{d}x}x^n = nx^{n-1}$. How do we know this works?

Proof. First, let's start by letting $y = x^n$. Then, we take the natural log of each side, yielding $\ln(y) = \ln\big(x^n\big) = n\ln(x)$. Now, we differentiate:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\ln(y) &= \frac{\mathrm{d}}{\mathrm{d}x}\Big[n\ln(x)\Big] = n\frac{\mathrm{d}}{\mathrm{d}x}\ln(x) \\
\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{n}{x} \\
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{n}{x}y = \frac{n}{x}x^n = nx^{n-1}.
\end{aligned}$$

And thus the power rule is true. □

However, you may raise some concerns with this proof. We derived this using the logarithm, which is only defined (in the real numbers) when the argument is greater than 0, and you're exactly right. However, this breakdown can be avoided for some powers n, specifically for integers n. If n is an even number, then $(-x)^n = x^n$, and for n an odd number, we have $(-x)^n = -x^n$. Since we will now be differentiating x^n where $x \geq 0$ in either of these cases, our derivation holds. If n is some other non-integer, this (generally) does not hold for $x < 0$. This makes sense, however. Functions of the form x^n where n is not an integer are generally undefined in the real numbers for $x < 0$, but there are some exceptions. Whenever $n = \dfrac{p}{q}$ for integers p, q where q is an odd number, x^n is still defined for $x < 0$, because $(-x)^{1/a} = (-1)^{1/a}x^{1/a} = -x^{1/a}$ for odd a.

Logarithmic Differentiation is an extremely powerful tool. It isn't just helpful for taking the derivative of $f(x)^{g(x)}$, but can also be used to verify the product and quotient rules, or even to simplify problems that would have been quite complex using the product/quotient rule. We will not use logarithmic differentiation to verify the product or quotient rule since we have already verified each, but, if you wish to do so, it is fairly straightforward. Instead, we will provide an example where logarithmic differentiation can be used in place of other rules to simplify derivatives.

If $y = \dfrac{x^4}{(3x-4)\sqrt{x^3-7}}$, find $\dfrac{\mathrm{d}y}{\mathrm{d}x}$.

We *can* do this with only the product and quotient rules (and power and chain rules of course), but it will be much easier using logarithmic differentiation instead. We begin as usual, taking the natural log of each side. We then use properties of logs to simplify the right-hand side:

$$\begin{aligned}
\ln(y) &= \ln\left(\frac{x^4}{(3x-4)\sqrt{x^3-7}}\right)\\
\ln(y) &= \ln\left(x^4\right) - \ln\left((3x-4)\sqrt{x^3-7}\right)\\
\ln(y) &= 4\ln(x) - \left(\ln(3x-4) + \ln\left(\sqrt{x^3-7}\right)\right).\\
\ln(y) &= 4\ln(x) - \ln(3x-4) - \frac{1}{2}\ln\left(x^3-7\right).
\end{aligned}$$

We now take the derivative of each side using the chain rule:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\ln(y) &= \frac{\mathrm{d}}{\mathrm{d}x}\left[4\ln(x) - \ln(3x-4) - \frac{1}{2}\ln\left(x^3-7\right)\right]\\
\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{4}{x} - \frac{3}{3x-4} - \frac{3x^2}{2(x^3-7)}\\
\frac{\mathrm{d}y}{\mathrm{d}x} &= y\left(\frac{4}{x} - \frac{3}{3x-4} - \frac{3x^2}{2(x^3-7)}\right)\\
\frac{\mathrm{d}y}{\mathrm{d}x} &= \left(\frac{x^4}{(3x-4)\sqrt{x^3-7}}\right)\left(\frac{4}{x} - \frac{3}{3x-4} - \frac{3x^2}{2(x^3-7)}\right).
\end{aligned}$$

Which is much easier than what we would have had to do to differentiate using the quotient and product rules.

2.4.9 Piecewise Derivatives

2.5 Examples

2.6 Chapter 2 Summary and Exercises

Chapter 3

Derivative Applications

Chapter 4

Integrals

4.1 What is an Integral?

There are two main types of integrals—definite integrals and indefinite integrals. This doesn't mean you have two new things to learn, however. When doing an integral, a definite integral is done the same way as an indefinite integral, but you evaluate the output at two x and find the difference between them, similar to evaluating a derivative at a given x. The definite integral is the first type of integral we will discuss here. It is a measure of the net area between a curve and the x-axis.

To better understand what this means, consider the function $f(x) = \sin(x)$. How much area is between this function and the x-axis on the interval $[-\pi, \pi]$? We write this question as $\int_{-\pi}^{\pi} \sin(x)\mathrm{d}x$. Let's break down what each part of this expression means. The symbol $\int$ is the *integral*. The *lower bound* of the integral is $-\pi$, and let's us know where our interval starts. The *upper bound* of the integral is π, and lets us know where the interval over which we are trying to find the area between the function and the x-axis ends. The dx at the end is very important and has quite a bit of meaning. When dealing with definite integrals, it specifies which axis we want to find the area between the function with. For example, if we had $\int_1^3 3\ \mathrm{d}y$, that would mean we are trying to find the area between the line $x = 3$ and the y-axis between the y values of 1 and 3. To give a better idea of what $\int_{-\pi}^{\pi} \sin(x)\mathrm{d}x$ is asking us, we present a picture with the area that we are trying to measure shaded in:

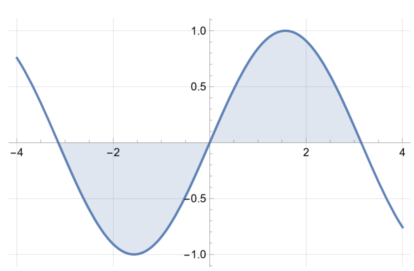


Figure 4.1: A plot of the curve $f(x) = \sin(x)$ with the area between $f(x)$ and the x-axis shaded when x is in the interval $[-\pi, \pi]$.

4.2 Approximating Integrals

The definition of the definite integral is:

$$\int_a^b f(x)\mathrm{d}x = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x \text{ where } \Delta x = \frac{b-a}{n}.$$

This is called the **Riemann Integral**. Essentially, this just says that we can split up a region into n rectangles and calculate the area of each rectangle by multiplying its height, $f(x_i)$, by its width, $\frac{b-a}{n}$, which is the length of the interval we are integrating over divided by the number of rectangles, then take the sum (which is what the symbol $\sum$ means) of all of these rectangles. As n approaches infinity, we are adding up so many rectangles that the sum of all of these areas is equal to the exact area under the curve $f(x)$.

Of course, calculating integrals this way would be pretty difficult (although it can be done, in fact, it's how many calculators calculate definite integrals). That's why mathematicians have come up with ways to approximate the value of integrals. We will discuss 4 ways to approximate an integral, and then discuss ways to get the exact values.

4.2.1 Left and Right Rectangular Approximation Methods

The Left Rectangular Approximation Method (LRAM) and Right Rectangular Approximation Method (RRAM) are the two methods of approximation we will discuss first. LRAM and RRAM are also known as Left Riemann Sum and Right Riemann Sum, respectively. Before we define either, we will visually show examples of each.

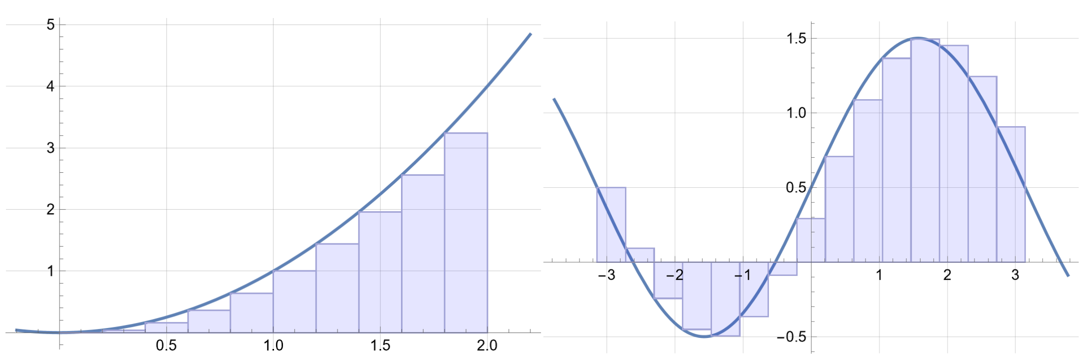


Figure 4.2: Graphs of x^2 and $\sin(x) + 0.5$ with their Left Riemann Sum shown from 0 to 2 with 10 rectangles and $-\pi$ to π with 15 rectangles, respectively.

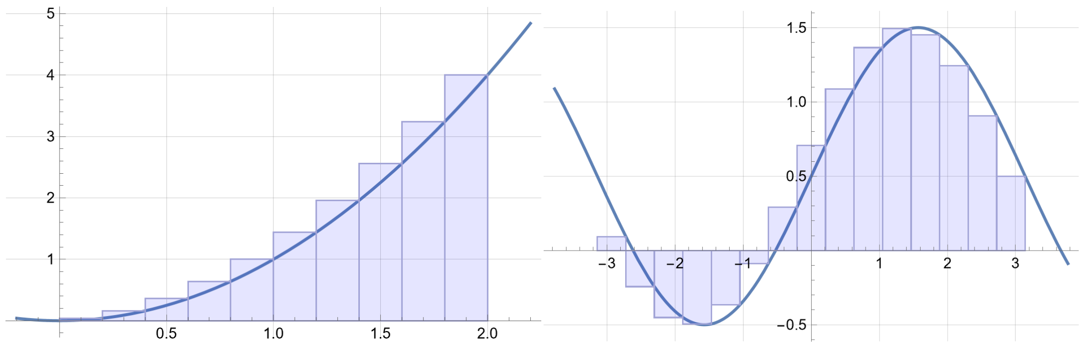


Figure 4.3: Graphs of x^2 and $\sin(x) + 0.5$ with their Right Riemann Sum shown from 0 to 2 with 10 rectangles and $-\pi$ to π with 15 rectangles, respectively.

As you can see, each method involves a set number of rectangles with a uniform width. The only thing that varies is the height, which is determined by the function value. In the Left Riemann Sum, the height is determined by the function value at the leftmost point in each rectangle. This causes a certain amount of error. For LRAM, when a function is positive and increasing, as seen in the graphs above, the area measured by the LRAM method will be an under-approximation. When the is negative and decreasing, it will also be an under-approximation. When the function is positive and decreasing or when it is negative and increasing, LRAM over-estimates the area. This is easiest to recall not by sheer memorization, but by drawing out a sample curve with a couple of LRAM rectangles shown and determining if LRAM over or under approximates in a given instance.

Similarly to LRAM, the Right Riemann Sum consists of rectangles; however, the height of each rectangle is determined by its rightmost x value. As you can see from the above graphs, when a function is increasing and positive or decreasing and negative, RRAM overestimates the area under a curve. When the function is positive and decreasing or negative and increasing, RRAM underestimates the area under a curve.
From the images above, you can probably see that these methods introduce quite a bit of error. Thus, a natural question arises: How can we minimize this area? Let's take a moment to remember how we defined the derivative. We said that for some differentiable function $f(x)$,

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

The reason this limit naturally appears in this formula is that to get the slope at a specific point, we must make the distance between two points, h, 0. This is a similar idea to what we would like to do here. If we make each rectangle infinitely thin while also ensuring that we have an infinite number of them in a given interval, then the sum of the area of all of our rectangles will tell us exactly the area under a curve. Before we can do so, let us explain a bit more about Riemann Sums. If you look at either graph of x^2 shown above, you will count 10 rectangles split along a distance of 2 units. Thus, the width of each rectangle is $\frac{2}{10} = 0.2$ units, which you can confirm if you look closely at either graph. For the plots of sine, you count 15 rectangles. The domain of each plot of sine is $[-\pi, \pi]$, so the total distance from one endpoint to the other is $\pi - (-\pi) = 2\pi$. This means that the width of each rectangle is $\frac{2\pi}{15} \approx 0.418879$. From here we define a Riemann Sum.

Riemann Sum

For a continuous function $f(x)$, the net area between $f(x)$ and the x-axis on an interval $[a, b]$ can be approximated with n rectangles as

$$\sum_{i=1}^{n} f(x_i)\Delta x.$$

Where Δx is the width of each rectangle, $\Delta x = \frac{b-a}{n}$.

If we want this width to be minimized, we need our number of rectangles, n, to grow drastically, $n \to \infty$. This is where the definition of the **Riemann Integral** given at the beginning of this section comes from.

Riemann Integral

For a continuous function $f(x)$, the net area between $f(x)$ and the x-axis on an interval $[a, b]$ is

$$\int_a^b f(x)\mathrm{d}x = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$

Where Δx is the width of each rectangle, $\Delta x = \frac{b-a}{n}$.

We will now perform some examples using Left and Right Rectangular Approximation Methods.

Consider the function $f(x) = x^2$. Estimate the area between the x-axis and $f(x)$ over the interval $[4, 20]$ using LRAM with 4 rectangles.

We begin by determining the width of each rectangle. The right end-point of our interval is 20, and the left end-point is 4. We have 4 rectangles, so the width of each rectangle is

$$\frac{b-a}{n} = \frac{20-4}{4} = 4.$$

Since this is a *Left* Riemann Sum, we will use the leftmost point in each rectangle to determine our function value. This means our points will be $x = 4, 8, 12, 16$. Finding the areas of each rectangle and taking the sum:

$$4 \cdot f(4) + 4 \cdot f(8) + 4 \cdot f(12) + 4 \cdot f(16) = 2496.$$

This is an *under-approximation* because x^2 is both positive and increasing on the interval $[4, 20]$. A graphical depiction is below.

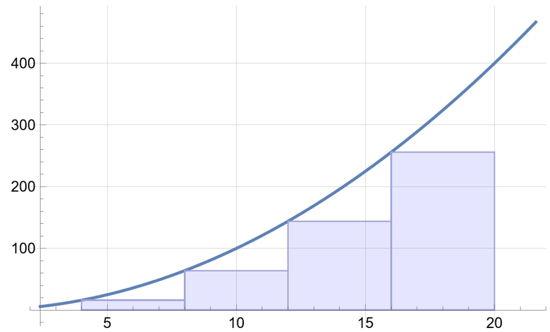


Figure 4.4: The Left Riemann Sum of the function $f(x) = x^2$ from 4 to 20.

Let's now repeat this problem with a Right Riemann Sum with 4 rectangles on the same interval to see how our answer differs. We expect a much higher number since x^2 is positive and increasing on the interval $[4, 20]$, which will cause an over-approximation by RRAM. Since this is a Right Riemann Sum, we use the right endpoints of each rectangle: $x = 8, 12, 16, 20$. Finding the areas of each rectangle and taking the sum:

$$4 \cdot f(8) + 4 \cdot f(12) + 4 \cdot f(16) + 4 \cdot f(20) = 3456.$$

Much higher than LRAM, as expected. Note that the actual amount of area underneath x^2 on the interval $[4, 20]$ is $2645.333\ldots$, so LRAM served as a much better approximation in this scenario. A plot of the Right Riemann Sum is included below.

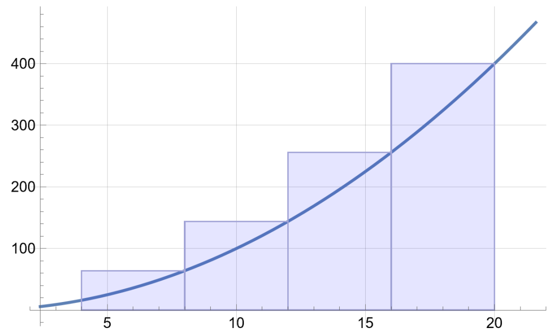


Figure 4.5: The Right Riemann Sum of the function $f(x) = x^2$ from 4 to 20.

We now present a second example, this time with a different format.

x	0	0.5	1	1.5	2	2.5	3
$f(x)$	1	2	3	5	8	13	21

Table 4.1: A table of x and $f(x)$ values for a unknown function.

Information: This function is positive and increasing on the interval [0, 3].
Questions:

A. Using a Left Riemann Sum and 3 rectangles of equal width, approximate $\int_0^3 f(x)\mathrm{d}x$. Is this an over or under-approximation?

B. Using a Right Riemann Sum and 6 rectangles of equal width, approximate $\int_0^3 f(x)\mathrm{d}x$. Is this an over or underestimation?

Answers:

A. First, we must decide how to create our intervals. Since we want these intervals to be of equal length (meaning the rectangles have equal width), the only option is defining our rectangles as having width 1, with the first rectangle running from 0 to 1, the second from 1 to 2, and the third from 2 to 3. Now, we have to calculate the area of our 3 rectangles. For the rectangle from 0 to 1, since this is a Left Riemann Sum, we use $f(0)$ as our rectangle height. Similarly, for the rectangle running from 1 to 2, we use $f(1)$, and, for the rectangle running from 2 to 3, we use $f(2)$. Since the width of our rectangles is 1, their areas are just their function values (since $1 \cdot f(a) = f(a)$. Then:

$$Estimation = f(0) + f(1) + f(2) = 12.$$

This is an underestimate because this is a Left Riemann Sum and the function is both positive and increasing.

B. For this second problem, we proceed similarly. To make 6 rectangles of equal length, we only have one option for the width of each rectangle: $\frac{b-a}{n} = \frac{3-0}{6} = 0.5$. So, the width of each rectangle (and length of each interval) must be 0.5. Then, our rectangles are defined from 0 to 0.5, 0.5 to 1, 1 to 1.5, 1.5 to 2, 2 to 2.5, and 2.5 to 3. Since this is a Right Riemann Sum, we use the right point of each interval to determine the height of our rectangles:

$$Estimation = 0.5\Big(f(0.5) + f(1) + f(1.5) + f(2) + f(2.5) + f(3)\Big) = 0.5 \cdot 52 = 26.$$

This is an overestimate because this is a Right Riemann Sum and the function is both positive and increasing.

4.2.2 Midpoint Rectangular Approximation Method

The Midpoint Rectangular Approximation Method (MRAM), or Midpoint Riemann Sum, is similar to LRAM and RRAM. As you may have guessed, instead, this method uses the midpoint of each interval to determine the height of each rectangle. This (usually) results in a more accurate approximation of area. Below are plots of MRAM for a quadratic and a sine curve, with 10 and 15 rectangles, respectively. The area under the quadratic is approximated on the interval $[0, 2]$, and the sine curve $[-\pi, \pi]$.

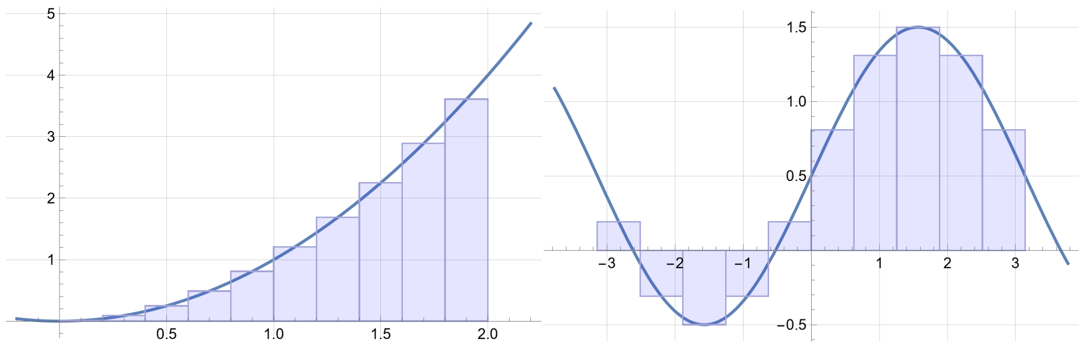


Figure 4.6: Graphs of x^2 and $\sin(x) + 0.5$ with their Midpoint Riemann Sum shown from 0 to 2 with 10 rectangles and $-\pi$ to π with 15 rectangles, respectively.

Note that both of these estimations look much more reasonable than either of the LRAM or RRAM equivalents given at he beginning of section 4.2.1. We will now continue with our examples from the previous section.

Consider the function $f(x) = x^2$. Estimate the area between the x-axis and $f(x)$ over the interval $[4, 20]$ using MRAM with 4 rectangles.

Again, the width of each rectangle will be $\frac{b-a}{4} = \frac{20-4}{4} = 4$. So our intervals will be of length 4. This gives us rectangles from 4 to 8, 8 to 12, 12 to 16, and 16 to 20. Because we are doing a Midpoint Riemann Sum, our $f(x_i)$ is determined by the x value in the middle of each rectangle interval. This yields $\frac{4+8}{2} = 6$, $\frac{8+12}{2} = 10$, $\frac{12+16}{2} = 14$, $\frac{16+20}{2} = 18$. Now, calculating the area of each rectangle and taking the sum:

$$4 \cdot \Big(f(6) + f(10) + f(14) + f(18)\Big) = 2624.$$

This estimate is very close to the actual value of $2645.333\ldots$, so this was a very good choice of approximation method for this problem. Let's now continue with our table problem.

x	0	0.5	1	1.5	2	2.5	3
$f(x)$	1	2	3	5	8	13	21

Table 4.2: A repeat of Table 4.1 - a table of x and $f(x)$ values for a unknown function.

Information: $f(x)$ is positive, increasing, and concave up on the interval $[0, 3]$.
Question:

C. Using a Midpoint Riemann Sum and 3 rectangles of equal width, approximate $\displaystyle\int_0^3 f(x)\mathrm{d}x$.

Answer:

C. First we must determine the length of each interval. Since our upper bound is 3, our lower bound is 0, and we are using 3 rectangles, the length of each interval (or width of each rectangle) is

$\frac{3-0}{3} = 1$. This means our intervals for each rectangle will be 0 to 1, 1 to 2, and 2 to 3. The midpoint of each rectangle is 0.5, 1.5, and 2.5, respectively. Using this, we calculate the area of each rectangle and take the sum:

$$1 \cdot \Big(f(0.5) + f(1.5) + f(2.5)\Big) = 20.$$

Is this an over or under-approximation? To determine this, we will take a look at a few extreme examples.

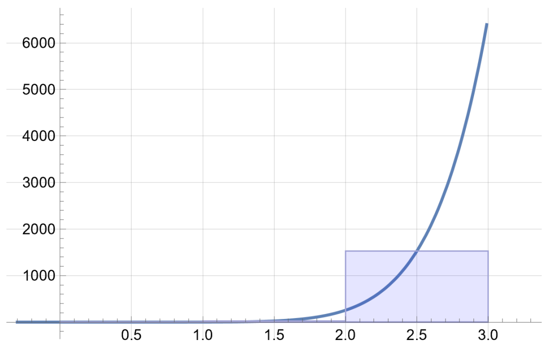


Figure 4.7: A graph of the area under $f(x) = x^8$ approximated by MRAM with 3 rectangles on the interval $[0, 3]$.

In this graph where function is positive, increasing, and concave up, it is clear that MRAM underestimates the value of the function. Is this true in general? Yes, it is. To understand this a bit more intuitively, think about what it means for a function be concave up. This means that the second derivative is positive, which in turn means the first derivative is increasing (so the function is increasing at a faster and faster rate). Because of this, the area above the second half of each rectangle (and below the curve) will be greater than the area above the curve and inside the rectangle to the left of the midpoint.

As an additional way of thinking about why concavity plays a role in this, think about a linear function. Take for example $y = 3x$. Shown below is the Midpoint Riemann Sum for this function over the interval $[0, 5]$ with one rectangle.

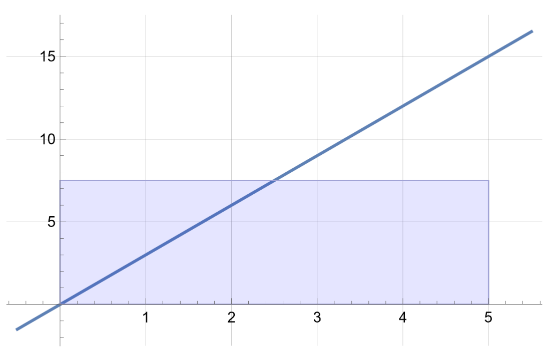


Figure 4.8: Midpoint Riemann Sum for $y = 3x$ on the interval $[0, 5]$ with 1 rectangle.

Note that the function $y = 3x$ has no concavity because $y'' = 0$. This suggests that MRAM neither over nor underestimates the area under this curve. By looking at the graph, this does in fact appear to be true, but we should verify with geometry. First, let's find the area of the rectangle. Our interval is $[0, 5]$ and we have one rectangle, so this rectangle's width must be 5. The midpoint is $\frac{5+0}{2} = 2.5$. Then the area of the rectangle must be $5 \cdot (3 \cdot 2.5) = 37.5$. As for the area under $y = 3x$ on the interval $[0, 5]$, notice this is a triangle with base 5 and height $3 \cdot 5 = 15$. Then, the area of this triangle is $\frac{1}{2} 5 \cdot 15 = 37.5$, so MRAM neither over nor under-approximated the area under this curve. This means that the area of the part of the MRAM rectangle above the line $y = 3x$ must be equal to the area under $y = 3x$ and above the MRAM rectangle on this interval. If we instead let the slope increase slightly after this midpoint, as illustrated in the next picture, then the area under the curve will increase but MRAM will still estimate a total area of 37.5.

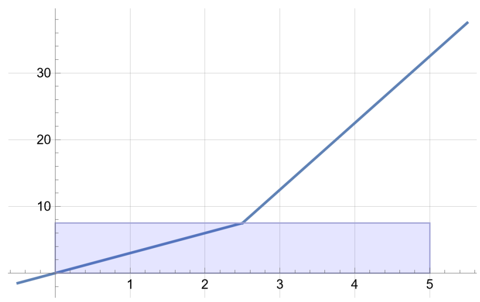


Figure 4.9: Midpoint Riemann Sum for $f(x) = \begin{cases} 3x & x \leq 2.5 \\ 10x - 17.5 & x > 2.5 \end{cases}$ on the interval $[0, 5]$ with 1 rectangle.

From this, we can extrapolate that (for a positive function) **if the slope of the function is larger after the midpoint than before** (an effect of $f''(x) > 0$), **then MRAM will underestimate the area.** If the function is positive and concave down, MRAM will overestimate the area. If the function is negative and concave up, MRAM will overestimate the area, and if the function is negative and concave down, MRAM will underestimate the area.

Now, returning to our question as to whether we under or over-approximated the area in part C. of our table question, since the function was positive and concave up (this was given in the information for the question) on the interval we approximated, MRAM underestimated the area under $f(x)$.

4.2.3 Trapezoid Rule

The Trapezoid Rule is the final method of numerical approximation we will discuss. As you may have guessed from the name, this method will involve the use of trapezoids, and so it is quite different from the 3 other methods we have discussed up to this point. However, the idea is still very similar. As with the previous 3 methods, we will begin by showing a visual example for the functions $f(x) = x^2$ and $f(x) = \sin(x) + 0.5$:

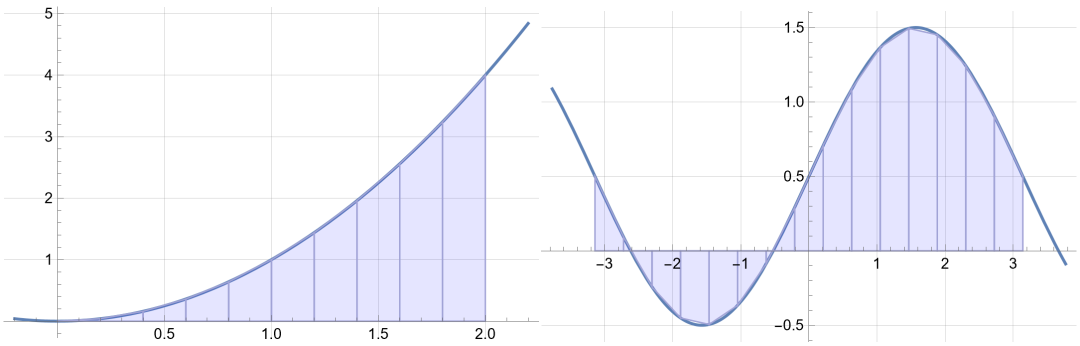


Figure 4.10: Graphs of x^2 and $\sin(x) + 0.5$ with their Trapezoidal Riemann Sum shown from 0 to 2 with 10 trapezoids and $-\pi$ to π with 15 trapezoids, respectively.

From the figures, it appears that the Trapezoid Rule is the most exact of the 4 rules (and that it is nearly perfect), however, MRAM is (typically) the best method for approximation.

Before we discuss more about the Trapezoid Rule, we will review how to find the area of a trapezoid. We begin with a generic trapezoid,

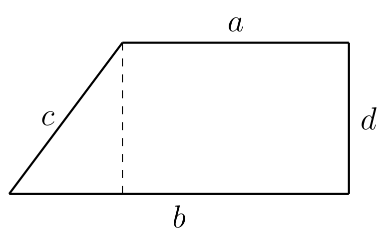


Figure 4.11: A right trapezoid with major base b, minor base a, height d along the side of the trapezoid with the right angles, and length c fourth side. A dashed line cuts the trapezoid into a rectangle and a triangle.

To find the area of this trapezoid, we will split this into two shapes—a rectangle with width a and height d and a triangle with width $b - a$ and height d. The "slice" is shown in the picture above, represented by a dashed line. These are two shapes we easily know how to find the area of. We will depict the area of the trapezoid as A_{trap}, area of the triangle as A_{tri}, and the area of the rectangle as A_{rect}. Since the trapezoid is split into these two shapes,

$$A_{trap} = A_{tri} + A_{rect}.$$

The area of the rectangle is $A_{rect} = ad$ and the area of the triangle is $A_{tri} = \frac{1}{2}(b - a)d$. Then,

$$\begin{aligned} A_{trap} &= ad + \frac{1}{2}(b-a)d \\ &= \frac{d}{2}(2a + b - a) \\ A_{trap} &= \frac{a+b}{2}d. \end{aligned}$$

One can also derive this formula by thinking of the trapezoid as a distorted rectangle. Simply take the average of the major and minor base ($\frac{a+b}{2}$) to find the equivalent length of the rectangle, then multiply by the height, d. Now we turn this trapezoid on its side:

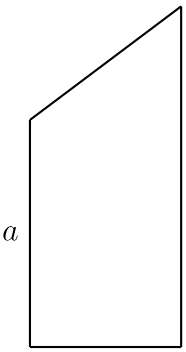


Figure 4.12: A right trapezoid with width d, left height a, and right height b.

This is exactly what the trapezoids we will make in doing the trapezoid rule will look like (reference Figure 4.10). To find the heights a and b, note that these are just the heights of the function at the endpoints of the interval. So, if our trapezoid interval is $[0, 3]$, our trapezoid area formula would be $\frac{f(0)+f(3)}{2}(3)$. Let's now use the trapezoid rule to continue with our examples from the preceding sections.

Consider the function $f(x) = x^2$. Estimate the area between the x-axis and $f(x)$ over the interval $[4, 20]$ using the Trapezoid Rule with 4 trapezoids.

As usual, we begin by determining the length of each interval. In this case, the interval lengths will represent the width of the trapezoids, not the width of rectangles. Since we want 4 trapezoids of equal length and our interval is $[4, 20]$, the width of each trapezoid is $\frac{20-4}{4} = 4$. This yields trapezoid intervals of $[4, 8], [8, 12], [12, 16]$ and $[16, 20]$. Then, factoring out the $\frac{1}{2}$ and the width, 4, from each expression for area of a trapezoid, we have

$$4\frac{1}{2}\Big(f(4)+f(8)+f(8)+f(12)+f(12)+f(16)+f(16)+f(20)\Big)= 2688.$$

An excellent approximation, albeit slightly worse than that of MRAM. We now consider our second table example with the Trapezoid rule.

x	0	0.5	1	1.5	2	2.5	3
$f(x)$	1	2	3	5	8	13	21

Table 4.3: A repeat of Table 4.1 - a table of x and $f(x)$ values for a unknown function.

Information: $f(x)$ is positive, increasing, and concave up on the interval $[0, 3]$.
Question:

D. Using the Trapezoid Rule and 2 trapezoids of equal width, approximate $\int_0^3 f(x)\mathrm{d}x$.

Answer:

D. Since we will be using two trapezoids of equal width, our interval length will be 1.5, making our intervals $[0, 1.5]$ and $[1.5, 3]$. Then, the total approximation for the area between the curve and the x-axis is the sum of the areas of the trapezoids defined by these intervals,

$$\frac{f(0) + f(1.5)}{2}(1.5) + \frac{f(1.5) + f(3)}{2}(1.5) = 24.$$

Is this an over or underestimation? It is much easier to tell intuitively whether or not the Trapezoid Rule will overestimate when compared to the Midpoint Riemann Sum. Take a look at the following plots of the $\sin(x)$ with a Trapezoid Rule with 4 trapezoids and x^3 with a Trapezoid Rule with 2 trapezoids.

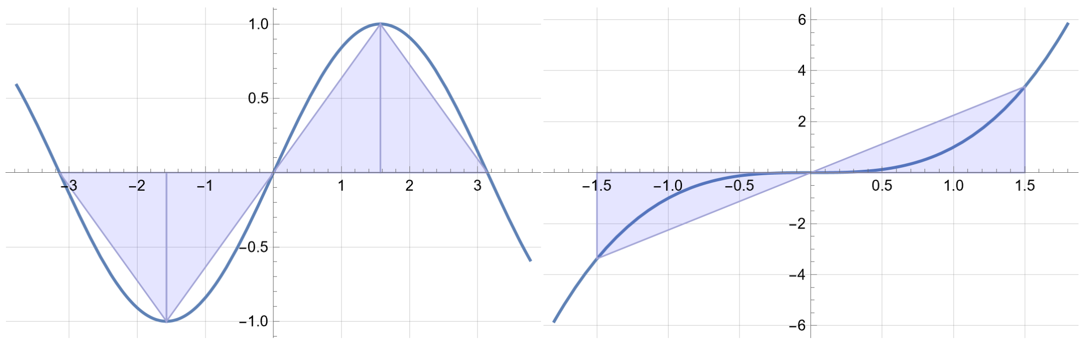


Figure 4.13: A plot of the function $f(x) = \sin(x)$ with the area between $-\pi$ and π approximated using a Trapezoid Rule with 4 trapezoids of equal width. On the right is a plot of the function $f(x) = x^3$ with the area between -1.5 and 1.5 approximated using a Trapezoid Rule with 2 trapezoids of equal width.

In this instance, it so happened that each trapezoid turned out to be a triangle because each trapezoid had one endpoint that lay on the x-axis. However, that does not affect our analysis of the approximation. From the first graph, we deduce that **the Trapezoid Rule underestimates whenever the function is concave down and positive or concave up and negative**. From the second graph, we notice that **whenever a function is negative and concave down or positive and concave up, the Trapezoid Rule will overestimate the area.**

We can now return to our question about part D. of the table problem, and we deduce that, because the function is positive and concave up on the interval $[0, 3]$, the Trapezoid Rule provided an over-approximation for the area between the curve and the x-axis. Since the Midpoint Riemann Sum underestimated 20 and the Trapezoid Rule overestimated 24, the true area under this curve is between 20 and 24.

4.2.4 Net vs. Total Area with Geometry and Symmetry

We now must make an important distinction. Up until now, we have thrown around the phrases *net area* and *total area*. We will now define each and give some examples using geometry.

To find *net area*, we treat area under the x-axis as "negative" area. This can be understood more easily with an example.

Find the net area between the x-axis as the curve $x^2 + y^2 = 3$.

Recall that an equation of the form $x^2 + y^2 = r^2$ is a circle of radius r centered at the origin. Thus the equation above is a circle of radius 3 centered at the origin. A graph is included below.

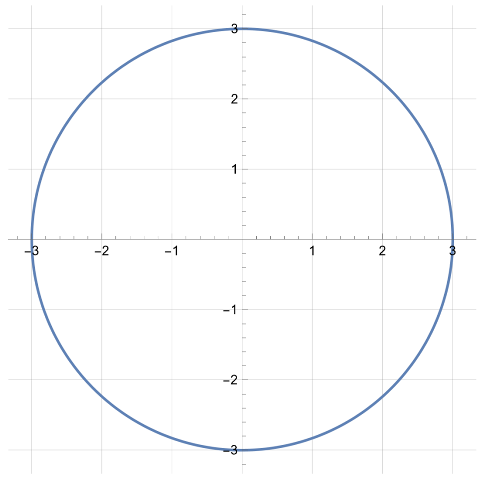


Figure 4.14: A circle of radius 3 centered at the origin.

Remember, we want to treat all area below the x-axis as negative, and all that above as positive. Since half the circle is plotted beneath the x-axis and half above, the net area will be 0. We can verify this mathematically using the formula for the area of a semicircle (half the area of a circle). Doing so for the positive area,

$$\frac{1}{2}\pi(3)^2 = \frac{9\pi}{2}.$$

For the area beneath the x-axis, we have

$$-\frac{1}{2}\pi(3)^2 = -\frac{9\pi}{2}.$$

To find the net area, we take the sum,

$$\frac{9\pi}{2} + \frac{-9\pi}{2} = 0.$$

And thus the net area between the curve and the x-axis is 0. We now present a second example.

Find the net area between the function $y = x$ and the x-axis on the interval $[-1, 2]$. Below is a graphical depiction of this with the positive area shown in blue and the negative area shown in red.

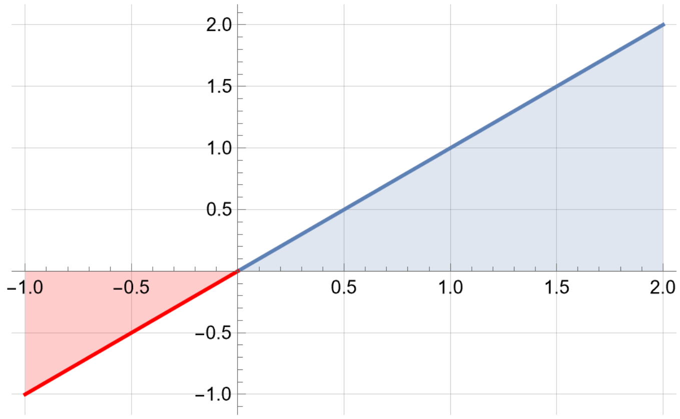


Figure 4.15: A graph of $y = x$ on the interval $[-1, 2]$ with negative area shown in red and positive area shown in blue.

To find the net area, we must subtract the area under the x-axis from that above. That is, we must calculate *blue area* − *red area*. Let's calculate each area. For the red area, we have a triangle with height 1 and base 1. This means the area is $\frac{1}{2} \cdot 1 \cdot 1 = \frac{1}{2}$. For the blue triangle, the base is 2 and the height is also 2. This means the area of the blue triangle is $\frac{1}{2} \cdot 2 \cdot 2 = 2$. Then, subtracting the red area from the blue area, we have that the net area between the function $y = x$ over the interval $[-1, 2]$ is $2 - \frac{1}{2} = 1.5$.

We now move to a discussion of total area. The *total area* between a curve and the x-axis can be found by treating all area as positive. So, in our previous example with a circle, the total area is the sum of the area above and below the x-axis without treating any of it as negative. Doing this yields

$$\frac{9\pi}{2} + \frac{9\pi}{2} = 9\pi.$$

This makes sense because the only area between the circle and the x-axis is the area of the circle itself, which is given by πr^2: $\pi(3)^2 = 9\pi$. As for our second example, now we simply add the red and blue areas instead of subtracting one from the other. Doing so, we get $2 + \frac{1}{2} = 2.5$. So the total area beneath the function $y = x$ and the x-axis is 2.5.

4.3 Properties of Integrals

Here, we will list a few properties of integrals with explanations. These will become more clear in the coming sections as we use them to evaluate integrals, but it is necessary that we list them first so that we can use them.

1. $\displaystyle\int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x.$

   While this may look complicated, it really isn't when we break it down. Remember that the integral $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ measures the area between the function $f(x)$ and the x-axis over the interval $[a, b]$ and that $\displaystyle\int_b^c f(x)\,\mathrm{d}x$ measures the area between the function $f(x)$ and the x-axis over the interval $[b, c]$. It stands to reason then, that the area between $f(x)$ and the x-axis on the interval $[a, c]$ would just be the sum of these two areas, which is what this rule says. This rule can be used similarly in reverse. For example, since $0 \leq 3 \leq 5$,

$$\int_0^5 x^2 \mathrm{d}x = \int_0^3 x^2 \mathrm{d}x + \int_3^5 x^2 \mathrm{d}x.$$

2. $\displaystyle\int f(x) + g(x)\ \mathrm{d}x = \int f(x)\ \mathrm{d}x + \int g(x)\ \mathrm{d}x.$

   Here, we displayed the integrals without any bounds. We will discuss what this means in Section 4.5. However, this same rule applies with bounds:

$$\int_a^b f(x) + g(x)\ \mathrm{d}x = \int_a^b f(x)\ \mathrm{d}x + \int_a^b g(x)\ \mathrm{d}x.$$

   This may seem difficult, but it is actually the exact same property as we are used to with derivatives:

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big[f(x) + g(x)\Big] = \frac{\mathrm{d}}{\mathrm{d}x}f(x) + \frac{\mathrm{d}}{\mathrm{d}x}g(x).$$

   Basically, this rule just says that to integrate a function with multiple terms, you can integrate each term separately then add them back together again, just as we did with derivatives. This makes sense in terms of area. Consider the function $y = x + 1$. The graph of this function from 0 to 3 is shown below.

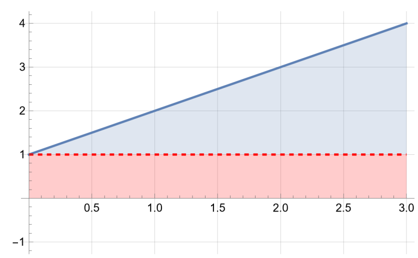


Figure 4.16: A graph of the function $y = x + 1$ with the triangular area shaded in blue and rectangular area shaded in red.

   As you can see, the area beneath this curve can be split into a rectangle and a triangle. The rectangle is a result of the $+1$ term and the triangle is a result of the x term. This suggests that we can split up the integrals to find these areas, and this is indeed true for the sum of any two functions.

3. $\displaystyle\int_a^b f(x)\ \mathrm{d}x = -\int_b^a f(x)\ \mathrm{d}x.$

   This rule will make more sense in the coming sections once we have the tools to do some examples, but, for now, you can think of this rule as helping us define a interval $[b, a]$ where $b > a$. Since it doesn't make sense to start our interval at the larger number, this rule tells us that instead we can swap the bounds and make the answer negative. This can be compared to derivatives in the following way. Consider the function $y = 3x + 4,\ y' = 3$. The derivative tells us that for every 1 unit increase (moving from smaller to larger) in x, we have a 3 unit change in y. If x moves from a larger value to a smaller value, then for each 1 unit change in x, there will be a -3 unit change in y.

4. $\int_a^a f(x)\ \mathrm{d}x = 0.$

   To understand this rule, remember that the integral from a to b measures the area between the x-axis and a given function by multiplying each function value by $\frac{b-a}{n}$ as $n \to \infty$. However, if our interval is from a to a, this width of each rectangle will be $\frac{a-a}{n} = 0$, so our integral will simply give us the function value at a (the height of the rectangle) multiplied by 0 (the width of the rectangle), giving us that $\int_a^a f(x)\ \mathrm{d}x = 0$.

5. $\int a \cdot f(x)\ \mathrm{d}x = a \int f(x)\ \mathrm{d}x.$

   Again, this rule was presented without any bounds, but works if there are bounds on the integral as well. To think about this intuitively, consider $\int_b^c a \cdot f(x)\ \mathrm{d}x$. The area under $a \cdot f(x)$ over the interval $[b, c]$ is the same as the between the function $f(x)$ and the x-axis multiplied by a. This is because the constant term just multiplies each function value that we are adding by 3, and we can instead factor this out of each term and multiply it at the end. In Riemann Integral definition notation, this is

$$\lim_{n\to\infty} \sum_{i=1}^{n} a \cdot f(x_i)\Delta x = a \cdot \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$

The rest of this chapter will consist of examples using these properties and defining and using other integration methods.

4.4 The Fundamental Theorem of Calculus

As you can no doubt tell from the name, the Fundamental Theorem of Calculus (FTC) is extremely important to the study of Calculus. There are two important parts of this theorem, often referred to as FTC 1 and FTC 2. The first part of the Fundamental Theorem of Calculus relates differentiation to integration. It establishes that integration and differentiation are inverses of one another. We now present the definition.

Fundamental Theorem of Calculus (1)

Let $f(x)$ be a continuous function on the interval $[a, b]$. Then, let $F(x)$ be a function defined such that

$$F(x) = \int_a^x f(t)\mathrm{d}t.$$

Then, $F(x)$ is continuous on $[a, b]$, differentiable on (a, b), and $F'(x) = f(x)$. A direct result of this theorem is that

$$\frac{\mathrm{d}}{\mathrm{d}x} \int_a^x f(t)\ \mathrm{d}t = f(x).$$

We have presented the definition of FTC 1 as it is typically stated, but it can be a little bit misleading. If the upper bound is something other than the variable you are differentiating with respect to, for example, $\dfrac{\mathrm{d}}{\mathrm{d}x}\displaystyle\int_a^{x^2} f(t)\,\mathrm{d}t$, you must use the chain rule. In this instance,

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^{x^2} f(t)\,\mathrm{d}t = f(x^2)\frac{\mathrm{d}}{\mathrm{d}x}\left[x^2\right] = 2x\cdot f(x^2).$$

So, the FTC 1 can be rewritten as:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^{g(x)} f(t)\,\mathrm{d}t = f(g(x))g'(x). \tag{4.1}$$

This will be easier to understand with a problem. Consider the function $h(x)$ such that $h(x) = \displaystyle\int_3^{\sin(x)} t^3\mathrm{d}t$. Find $h'\left(\frac{\pi}{4}\right)$.

Let's start by comparing this to equation (4.1). In this problem, $g(x) = \sin(x)$ and $f(t) = t^3$. We know that our expression for the derivative of $h(x)$, $h'(x)$ will involve $g'(x)$, so we compute that now. By Table 2.1, $\dfrac{\mathrm{d}}{\mathrm{d}x}\sin(x) = \cos(x)$, so $g'(x) = \cos(x)$. This gives us that

$$h'(x) = \frac{\mathrm{d}}{\mathrm{d}x}h(x) = \frac{\mathrm{d}}{\mathrm{d}x}\int_3^{\sin(x)} t^3\mathrm{d}t = \big(\sin(x)\big)^3\cos(x).$$

We now simply must evaluate this at $x = \frac{\pi}{4}$. From the unit circle, Figure 1, $\sin\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) = \dfrac{\sqrt{2}}{2}$. Plugging this in,

$$h'\left(\frac{\pi}{4}\right) = \left(\sin\left(\frac{\pi}{4}\right)\right)^3\cos\left(\frac{\pi}{4}\right) = \left(\frac{\sqrt{2}}{2}\right)^3\cdot\frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{4}\cdot\frac{\sqrt{2}}{2} = \frac{1}{4}.$$

Let's take a look at a second example. Consider the function $F(x) = \displaystyle\int_{-\tan(x)}^{\ln(x)^2} 3t\,\mathrm{d}t$. Find an expression for $F'(x)$.

We begin by using our second property of integrals from Section 4.3 and split this integral into two so that each integral only has one variable bound. We can choose any constant to be between these two bounds, since, according to the definition of FTC 1, if the lower bound is any number a, we can use the theorem. So, we split up our integral for $F(x)$ into two separate integrals:

$$F(x) = \int_{-\tan(x)}^{\ln(x)^2} 3t\,\mathrm{d}t = \int_{-\tan(x)}^{a} 3t\,\mathrm{d}t + \int_a^{\ln(x)^2} 3t\,\mathrm{d}t.$$

Now, notice that the left integral has a lower bound with a variable and an upper bound of a constant. Because the FTC 1 says the upper bound must contain the function, we use property 3 of integrals and flip the bounds, making our integral negative:

$$\int_{-\tan(x)}^{a} 3t\,\mathrm{d}t = -\int_a^{-\tan(x)} 3t\,\mathrm{d}t.$$

We can now take the derivative of each side to find $F'(x)$:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}F(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left[-\int_a^{-\tan(x)} 3t\ \mathrm{d}t + \int_a^{\ln(x)^2} 3t\ \mathrm{d}t\right] \\ F'(x) &= -\frac{\mathrm{d}}{\mathrm{d}x}\int_a^{-\tan(x)} 3t\ \mathrm{d}t + \frac{\mathrm{d}}{\mathrm{d}x}\int_a^{\ln(x)^2} 3t\ \mathrm{d}t. \end{aligned}$$

We now apply the fundamental theorem of calculus, remembering to apply the chain rule. Since $\frac{\mathrm{d}}{\mathrm{d}x}\Big[-\tan(x)\Big] = -\sec^2(x)$ and $\frac{\mathrm{d}}{\mathrm{d}x}\Big[\ln(x)^2\Big] = \frac{2\ln(x)}{x}$, we have

$$\begin{aligned} F'(x) &= -3\big(-\tan(x)\big)\cdot -\sec^2(x) + 3\big(\ln(x)^2\big)\cdot\frac{2\ln(x)}{x} \\ F'(x) &= -3\tan(x)\sec^2(x) + \frac{6\ln(x)^3}{x}. \end{aligned}$$

We now move on to part 2 of the Fundamental Theorem of Calculus. Its statement is as follows.

Fundamental Theorem of Calculus (2)

Let $f(x)$ be a continuous function on the interval $[a, b]$ and let $F(x)$ be a function such that $F'(x) = f(x)$. Then,

$$\int_a^b f(x)\ \mathrm{d}x = F(x)\Big|_a^b = F(b) - F(a).$$

This definition is truly fundamental in the use of integrals as it actually allows us to begin evaluating integrals, as we will do in the coming sections. With this theorem now defined, we can actually use it to verify some of our previous claims about integrals, which will help to give you a better understanding of how this theorem is used and how some of the properties of integrals agree with this theorem. Consider the two following integrals,

A. $\displaystyle\int_a^b f(x)\ \mathrm{d}x.$

B. $\displaystyle\int_b^a f(x)\ \mathrm{d}x.$

First, let's evaluate each using the FTC 2. Let $F(x)$ be a function with derivative $f(x)$. Then,

A. $\displaystyle\int_a^b f(x)\ \mathrm{d}x = F(b) - F(a).$

B. $\displaystyle\int_b^a f(x)\ \mathrm{d}x = F(a) - F(b) = -\big(F(b) - F(a)\big).$

This verifies property 3 of integrals. We can also use this to easily verify properties 1, 4, and 5, as we will do now.

Consider the sum of integrals $\int_a^b f(x)\ \mathrm{d}x + \int_b^c f(x)\ \mathrm{d}x$. Let $F(x)$ be a function satisfying $F'(x) = f(x)$. Then, by part 2 of the Fundamental Theorem of Calculus, we have that

$$\int_a^b f(x)\,\mathrm{d}x \quad = \quad F(b) - F(a)$$

and

$$\int_b^c f(x)\,\mathrm{d}x \quad = \quad F(c) - F(b).$$

Taking the sum, we have $F(b)-F(a)+F(c)-F(b) = F(c)-F(a) = \int_a^c f(x)\ \mathrm{d}x$, verifying property 1 of integrals.

To verify property 4, consider the integral $\int_a^a f(x)\ \mathrm{d}x$ and let $F'(x) = f(x)$. Then, $\int_a^a f(x)\ \mathrm{d}x = F(a) - F(a) = 0$, verifying property 4.

To verify property 5, consider the function $F(x)$ such that $F'(x) = f(x)$. Then by the properties of derivatives, $\frac{\mathrm{d}}{\mathrm{d}x}\Big[a \cdot F(x)\Big] = a \cdot f'(x)$. Then, $\int_b^c a \cdot f(x)\ \mathrm{d}x = a \cdot F(c) - a \cdot F(b) = a\Big(F(c) - F(b)\Big) = a\int_b^c f(x)\ \mathrm{d}x$, which verifies property 5 of integration.

We can even use this to verify FTC 1. Let $F'(x) = f(x)$. Now, consider the integral $\int_a^{g(x)} f(t)\ \mathrm{d}t$. By FTC 2, this integral is $F(g(x)) - F(a)$. Note $F(a)$ is some number, so its derivative is 0. This means that, by the chain rule and the fact that $F'(x) = f(x)$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big[F(g(x)) - F(a)\Big] = \frac{\mathrm{d}}{\mathrm{d}x}\Big[F(g(x))\Big] = F'(g(x))g'(x) = f(g(x))g'(x).$$

This is exactly as stated in the definition of part 1 of the Fundamental Theorem of Calculus. Let's try an example.

Find the area under one (completely positive) half-cycle of $\sin(x)$.

First, let's find the bounds for this interval. We must start by remembering that $\sin(x)$ completes one full cycle every x increases by 2π. Since we want the interval to find the area under one completely positive interval, the interval must start at some integer multiple of 2π (anywhere where a new period starts: $\ldots, -2\pi, 0, 2\pi, 4\pi, \ldots$). Naturally, we will choose 0, as this is the simplest value to work with. Thus, we set up our integral as

$$\int_0^\pi \sin(x)\ \mathrm{d}x.$$

We want to find a function whose derivative is $\sin(x)$. What function do we know that satisfies a similar property? We know that $\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) = -\sin(x)$. If we multiply each side here by -1, we have $\frac{\mathrm{d}}{\mathrm{d}x}\Big[-\cos(x)\Big] = \sin(x)$. Thus,

$$\int_0^{\pi} \sin(x)\ \mathrm{d}x = -\cos(x)\Big|_0^{\pi} = -\cos(\pi) - (-\cos(0)) = 1 + 1 = 2.$$

And thus the area under one half cycle of a sine wave is 2. We will present more concrete examples using this theorem in the coming sections.

4.5 Indefinite Integrals and the Constant of Integration

Before we specifically define the indefinite integral, we will first talk about the "antiderivative," the inverse of the derivative. The goal of the antiderivative is, given a function $f(x)$, to find a function $F(x)$ such that $F'(x) = f(x)$. Before we define any methods for integration, we are still able to integrate some basic functions. Take for example $f(x) = 2x$. The question $\int 2x\ \mathrm{d}x$ says "Give me a function whose derivative is $2x$." As you may recall from Chapter 2, we excessively found a function whose derivative was $2x$. That function was $y = x^2$, so x^2 is an antiderivative of $2x$.

Another example is $\int \cos(x)\ \mathrm{d}x$. This asks us to find a function whose derivative is $\cos(x)$. From Table 2.1, we of course know that a function satisfying this equation is $\sin(x)$.

You may have noticed in the paragraph above that we spoke very loosely when saying that a function was the antiderivative of another, saying things like "an antiderivative" or "a function." This is because there are infinitely many solutions to these problems. Take again the example $\int 2x\ \mathrm{d}x$. We already said that an answer to this problem is x^2 since $\frac{\mathrm{d}}{\mathrm{d}x}x^2 = 2x$. However, consider $y = x^2 + 1$. Then $y' = 2x$, and thus $x^2 + 1$ is also an answer to this integral. In fact, since x^2 plus any constant term will have the same derivative (the derivative of a constant is 0), the answer to the integral $\int x^2\ \mathrm{d}x$ is $x^2 + C$ where C is some constant number. We call C the *Constant of Integration* since it is the constant that results from an integration of a function. This is the distinction between antiderivatives and indefinite integrals that is often overlooked. The *Indefinite Integral* tells us all of the possible antiderivatives of a function, while the antiderivative tells us one specific function. Whenever you are asked to integrate a function (and there are not bounds on the integral), you should include the $+C$ so that your solution encompasses all possible solutions. It is natural to ask now, why don't we care about this constant of integration when there are bounds on the integral? This is because if we have some function, $f(x)$, and integrate it to find its antiderivative, $F(x)+C$, and now want to find the area under the curve $f(x)$ on the interval $[a, b]$ we will plug in a and b for x in our expression of the antiderivative and find the difference (this is the fundamental theorem of calculus). Since C represents a constant for one function only, $C - C = 0$. Knowing this, when we evaluate our expression to find the area under $f(x)$ from a to b, we see that:

$$F(b) + C - (F(a) + C) = F(b) + C - F(a) - C = F(b) - F(a),$$

exactly as the fundamental theorem of calculus says. We will now get into some concrete examples.

4.6 Basic Methods for Integration

In this section, we will derive some basic methods for integration, use the properties of the integral outlined in Sections 4.3 and 4.4.

4.6.1 Reverse Power Rule

Just as with derivatives, we have a general method to integrate any power function, that is, anything of the form x^n where n is some number. We begin with the definition.

The Power Rule for Integration

$$\int x^n \, \mathrm{d}x = \frac{x^{n+1}}{n+1} + C.$$

Where C is some arbitrary constant.

We will now show that this statement is true.

Proof. Consider the polynomial x^a. Then, by the power rule for differentiation,

$$\frac{\mathrm{d}}{\mathrm{d}x} x^a = a x^{a-1}.$$

Integrating both sides with respect to x,

$$\int \frac{\mathrm{d}}{\mathrm{d}x} x^a \, \mathrm{d}x = \int a x^{a-1} \, \mathrm{d}x.$$

Since integration is the inverse of differentiation,

$$x^a + C_1 = \int a x^{a-1} \, \mathrm{d}x.$$

Where C_1 is some arbitrary constant. Now, we divide each side by a. Since C_1 is an arbitrary number, $\frac{C_1}{a}$ will be some other arbitrary number, C.

$$\frac{x^a}{a} + C = \int x^{a-1} \, \mathrm{d}x.$$

Now, let $n = a - 1$. Then, adding 1 to each side, $n + 1 = a$. We now substitute this into the equation above,

$$\frac{x^{n+1}}{n+1} + C = \int x^n \, \mathrm{d}x. \tag{4.2}$$

Which is exactly what we wanted to show. □

Let's do an example. By geometry, we know that the area between the function $f(x) = 2x$ and the x-axis over the interval $[0, 3]$ is $\frac{1}{2}bh = \frac{1}{2}(3)(6) = 9$. We will confirm this using integration to find the area under the curve $f(x) = 2x$. First, we set up the integral:

$$\int_0^3 2x \, \mathrm{d}x.$$

Using property 5 of integrals, we can rewrite this as $2\int_0^3 x\ \mathrm{d}x$. Notice that this is of the form x^n with $n = 1$. Then, using our formula of the power rule for integration and FTC 2,

$$2\int_0^3 x\ \mathrm{d}x = 2\left(\frac{x^2}{2}\right)\Bigg|_0^3 = 2\left(\frac{3^2}{2} - \frac{0^2}{2}\right) = 9.$$

So calculus agrees with geometry. Let's try a harder example.

Consider the function $f(x) = 6x^2 - \frac{1}{3\sqrt[3]{x^2}} - \frac{2}{x^3} - 3\pi^2 x^{3\pi^2-1}$. Evaluate $\int f(x)\ \mathrm{d}x$. A lot is going on with this problem, but if we break it down, it will become much less intimidating. First, we remember that the integral of a sum of functions is the sum of the integrals of the functions (this is a restatement of the second property of integrals). Then, we pull the coefficients out of each integral (property 5 of integrals) and rewrite the terms using properties of exponents. This gives us the following problem instead.

$$\begin{aligned}\int f(x)\ \mathrm{d}x &= \int 6x^2\ \mathrm{d}x + \int \frac{-1}{3\sqrt[3]{x^2}}\ \mathrm{d}x + \int \frac{-2}{x^3}\ \mathrm{d}x + \int -3\pi^2 x^{3\pi^2-1}\ \mathrm{d}x \\ &= 6\int x^2\ \mathrm{d}x - \frac{1}{3}\int x^{\frac{-2}{3}}\ \mathrm{d}x - 2\int x^{-3}\ \mathrm{d}x - 3\pi^2\int x^{3\pi^2-1}\ \mathrm{d}x.\end{aligned}$$

Now, we consider each integral separately. For $\int x^2\ \mathrm{d}x$, we simply add 1 to the exponent then divide by the new exponent and we have our integral (not forgetting the $+C$, of course):

$$6\int x^2\ \mathrm{d}x = 6\frac{x^{2+1}}{2+1} + C_1 = 2x^3 + C_1.$$

Where C_1 is some arbitrary constant. Taking a look at our next integral, we simply add one to the power $(\frac{-2}{3})$ and divide by this new exponent:

$$-\frac{1}{3}\int x^{\frac{-2}{3}}\ \mathrm{d}x = -\frac{1}{3}\left(\frac{x^{-\frac{2}{3}+1}}{-\frac{2}{3}+1}\right) + C_2 = -\frac{1}{3}\frac{x^{\frac{1}{3}}}{\frac{1}{3}} + C_2 = -x^{\frac{1}{3}} + C_2 = -\sqrt[3]{x} + C_2.$$

Where C_2 is some arbitrary constant. We proceed in the same way for our third integral, adding one to the exponent and dividing by this new power, remembering to add the arbitrary constant on to the end.

$$-2\int x^{-3}\ \mathrm{d}x = -2\frac{x^{-2}}{-2} + C_3 = \frac{1}{x^2} + C_3.$$

Again, C_3 represents an arbitrary constant, not necessarily the same as C_1 or C_2. We proceed in the same fashion for the fourth integral. Since $(3\pi^2 - 1) + 1 = 3\pi^2$,

$$-3\pi^2\int x^{3\pi^2-1}\ \mathrm{d}x = -x^{3\pi^2} + C_4.$$

Where C_4 is again an arbitrary constant. We then take the sum of these four separate answers. Since C_1, C_2, C_3, C_4 are arbitrary, their sum is also arbitrary, $C_1 + C_2 + C_3 + C_4 = C$. With this, we have that

$$\int f(x)\ \mathrm{d}x = 2x^3 - \sqrt[3]{x} + \frac{1}{x^2} - x^{3\pi^2} + C.$$

We can confirm this by taking the derivative. We split this up into two derivatives:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[2x^3 - \sqrt[3]{x} + \frac{1}{x^2} - x^{3\pi^2} + C\right] = \frac{\mathrm{d}}{\mathrm{d}x}\left[2x^3 - \sqrt[3]{x} + \frac{1}{x^2} - x^{3\pi^2}\right] + \frac{\mathrm{d}}{\mathrm{d}x}[C] = \frac{\mathrm{d}}{\mathrm{d}x}\left[2x^3 - \sqrt[3]{x} + \frac{1}{x^2} - x^{3\pi^2}\right].$$

We differentiated this in Section 2.4.2 and got $f(x)$, so we are assured our answer is correct. It is a good idea, especially while integration is new to you, to take derivatives of your integrals to ensure you get the same function back.

The Exception

This rule is pretty unstoppable, but, unlike with the power rule for derivatives, there is one type of function for which it falls apart. While we have already seen that it works just fine for something like $\int \frac{1}{x^3}\,\mathrm{d}x$, but this isn't true for a very similar integral, $\int \frac{1}{x}\,\mathrm{d}x$. Let's assume that this power rule works for this function, apply it, and see what happens. Rewriting $\frac{1}{x}$,

$$\int \frac{1}{x}\,\mathrm{d}x = \int x^{-1}\,\mathrm{d}x.$$

So far, so good. Remember our general formula, $\int x^n\,\mathrm{d}x = \frac{x^{n+1}}{n+1} + C$. Here, $n = -1$, so $n + 1 = 0$. Then, if the power rule holds,

$$\int \frac{1}{x}\,\mathrm{d}x = \int x^{-1}\,\mathrm{d}x = \frac{x^0}{0} + C.$$

Something must be wrong here since we have a 0 in the denominator, which we know cannot happen without the help of limits (we cannot divide by 0). So what is the integral of $\frac{1}{x}$ then if we can't use the power rule? Well, remember that the integral is the opposite of the derivative and that the derivative of the natural logarithm is $\frac{1}{x}$. So,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\ln(x) &= \frac{1}{x} \\ \int \frac{\mathrm{d}}{\mathrm{d}x}\ln(x)\,\mathrm{d}x &= \int \frac{1}{x}\,\mathrm{d}x \\ \ln(x) = \int \frac{1}{x}\,\mathrm{d}x. \end{aligned} \tag{4.3}$$

Right? Almost. This is the right spirit, but this is not quite right. Note that for a function f, its antiderivative, F, must be continuous everywhere f is continuous. The function $\frac{1}{x}$ is continuous everywhere except at 0. How can we make this true for the natural logarithm? To do this, we can simply take the absolute value of x, giving $\int \frac{1}{x}\,\mathrm{d}x = \ln|x| + C$. Since $\ln|x|$ is simply the result of reflecting $\ln(x)$ across the y-axis while also leaving behind a copy of $\ln(x)$ to the right of the y-axis, as shown below.

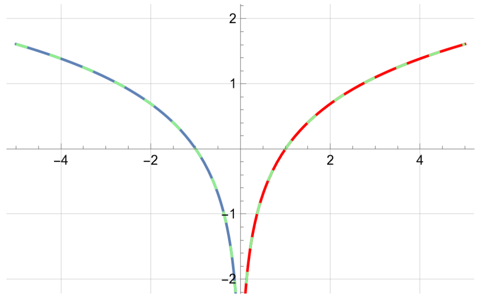


Figure 4.17: A plot of $\ln(x)$, shown in red, its reflection across the y-axis, $\ln(-x)$, shown in blue, and $\ln|x|$, shown by green dashes.

As we can see in the graph above, the slope at a point $a > 0$ will be the opposite (-1 times) the slope at a point $-a$. If we plug both values into the function that represents the derivative of $\ln|x|$, we see that this is true, getting $\frac{1}{a}$ and $\frac{1}{-a} = -\frac{1}{a}$, so we have confirmed that the function $\frac{1}{x}$ represents the slope at any point of the function $\ln|x|$. This can also be shown using limits, just as we did in Section 2.4.6 to show that $\frac{\mathrm{d}}{\mathrm{d}x}\ln(x) = \frac{1}{x}$. Also, because this graph is just a reflection, the area between the function and the x-axis between two points a and b where $a, b > 0$ will be the same as that between $-b$ and $-a$, so this also confirms that the area under the function $\frac{1}{x}$ can be measured with the function $\ln|x|$, confirming that it is the antiderivative.

4.6.2 Known Integrals

In the previous section, we integrated $\frac{1}{x}$ using only our knowledge that integration is the inverse of differentiation. We did not use any fancy new integration technique. This method of guessing a function whose derivative is the function we are trying to integrate can, in theory, be applied to any integrable function. However, this is generally not a very efficient method.

While, in general, this is not a "good" method for integrating functions, it can be used to easily integrate functions such as $\frac{1}{x^2+1}$ and $\cos(x)$ since we know functions with those derivatives, namely $\tan^{-1}(x)$ and $\sin(x)$. We will now use this method to re-present Table 2.1 of known derivatives as a table of known integrals.

Let us begin with the function $f(x) = \cos(x)$. To integrate $\cos(x)$, we need to recall what function has a derivative equal to $\cos(x)$. From Table 2.1, we know that the function that satisfies this is $\sin(x)$, so we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\sin(x) &= \cos(x) \\ \int \frac{\mathrm{d}}{\mathrm{d}x}\sin(x)\,\mathrm{d}x &= \int \cos(x)\,\mathrm{d}x \\ \sin(x) + C &= \int \cos(x)\,\mathrm{d}x. \end{aligned} \tag{4.4}$$

This method can be used for more complex-looking integrals as well, such as $\int \frac{-1}{|x|\sqrt{x^2-1}}\,\mathrm{d}x$. To integrate this crazy-looking function, we reference Table 2.1 and note that $\frac{\mathrm{d}}{\mathrm{d}x}\csc^{-1}(x) =$

$\dfrac{-1}{|x|\sqrt{x^2-1}}$. Then, using the same method we used to integrate $\cos(x)$ and $\ln(x)$,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\csc^{-1}(x) &= \frac{-1}{|x|\sqrt{x^2-1}} \\ \int \frac{\mathrm{d}}{\mathrm{d}x}\csc^{-1}(x)\,\mathrm{d}x &= \int \frac{-1}{|x|\sqrt{x^2-1}}\,\mathrm{d}x \\ \csc^{-1}(x) + C &= \int \frac{-1}{|x|\sqrt{x^2-1}}\,\mathrm{d}x. \end{aligned} \tag{4.5}$$

However, also notice that $\dfrac{\mathrm{d}}{\mathrm{d}x}\sec^{-1}(x) = \dfrac{1}{|x|\sqrt{x^2-1}}$ which implies that $\dfrac{\mathrm{d}}{\mathrm{d}x}\left[-\sec^{-1}(x)\right] = \dfrac{-1}{|x|\sqrt{x^2-1}}$. Then, by the same logic above, $\displaystyle\int \frac{-1}{|x|\sqrt{x^2-1}}\,\mathrm{d}x = -\sec^{-1}(x)\,\mathrm{d}x + C$. Is something here wrong? No, nothing here is wrong. While this may seem contradictory, remember, the integral does not assign a single function as the antiderivative of a function, it assigns an infinite family of functions, off only by some constant. In the case of the inverse secant and cosecant functions, they are equal when this constant is $\frac{\pi}{2}$. That is, $-\sec^{-1}(x) + \frac{\pi}{2} = \csc^{-1}(x)$. A graph of this is shown below.

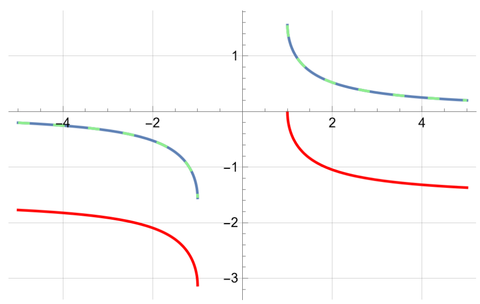


Figure 4.18: A plot of $\csc^{-1}(x)$, shown in blue, $-\sec^{-1}(x)$, shown in red, and $-\sec^{-1}(x) + \dfrac{\pi}{2}$, shown in dashed green.

We will continue with 2 more examples of this method. First, we will integrate a^x. Since we do not know a function whose derivative is a^x, we instead think of one that is only off by a constant. The derivative of a^x is $a^x\ln(a)$. That is,

$$\frac{\mathrm{d}}{\mathrm{d}x}a^x = a^x\ln(a).$$

If we divide each side by $\ln(a)$, we are left with a function whose derivative is a^x,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{a^x}{\ln(a)}\right] = a^x.$$

We now integrate both sides,

$$\begin{aligned} \int \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{a^x}{\ln(a)}\right]\mathrm{d}x &= \int a^x\,\mathrm{d}x \\ \frac{a^x}{\ln(a)} + C &= \int a^x\,\mathrm{d}x. \end{aligned} \tag{4.6}$$

And we have thus integrated a^x using only our knowledge of derivatives and the fundamental theorem of calculus.

We will now continue with one last example. How do we integrate the function $\frac{1}{x\ln(a)}$? We begin by noticing that this is the derivative of the function $\log_a(x)$, which gives us the following chain of logic:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\log_a(x) &= \frac{1}{x\ln(a)} \\ \int \frac{\mathrm{d}}{\mathrm{d}x}\log_a(x)\,\mathrm{d}x &= \int \frac{1}{x\ln(a)}\,\mathrm{d}x \\ \log_a(x) + C &= = \int \frac{1}{x\ln(a)}\mathrm{d}x. \end{aligned} \tag{4.7}$$

Unfortunately, this is wrong for the same reason that $\int \frac{1}{x}\,\mathrm{d}x = \ln|x| + C$ instead of $\ln(x) + C$. We can easily fix this in the same fashion as before, however, by adding absolute value bars around x. While this method works completely for computing this integral, we actually have an easier method we can do. Notice that, by property 5 of integrals,

$$\int \frac{1}{x\ln(a)}\,\mathrm{d}x = \frac{1}{\ln(a)}\int \frac{1}{x}\,\mathrm{d}x.$$

Then, since we know that $\int \frac{1}{x}\,\mathrm{d}x = \ln|x| + C$, we have that

$$\int \frac{1}{x\ln(a)}\,\mathrm{d}x = \frac{1}{\ln(a)}\int \frac{1}{x}\,\mathrm{d}x = \frac{1}{\ln(a)}\ln|x| + C.$$

This result is equivalent to $\log_a|x| + C$ by the change of base formula for Logarithms, found in 0.3.

With this same line of reasoning for each cell, we re-present Table 2.1 as follows.

Function	Integral	Function	Integral
$\cos(x)\ \mathrm{d}x$	$\sin(x) + C$	$\sin(x)\ \mathrm{d}x$	$-\cos(x) + C$
$\sec^2(x)\ \mathrm{d}x$	$\tan(x) + C$	$\csc^2(x)\ \mathrm{d}x$	$-\cot(x) + C$
$\sec(x)\tan(x)\ \mathrm{d}x$	$\sec(x) + C$	$\csc(x)\cot(x)\ \mathrm{d}x$	$-\csc(x) + C$
$\frac{1}{x}\ \mathrm{d}x$	$\ln\lvert x\rvert + C$	$\frac{1}{x\ln(a)}\ \mathrm{d}x$	$\log_a \lvert x\rvert + C$
$e^x\ \mathrm{d}x$	$e^x + C$	$a^x\ \mathrm{d}x$	$\frac{a^x}{\ln(a)} + C$
$\frac{1}{\sqrt{1-x^2}}\ \mathrm{d}x$	$\sin^{-1}(x) + C$	$\frac{-1}{\sqrt{1-x^2}}\ \mathrm{d}x$	$\cos^{-1}(x) + C$
$\frac{1}{\lvert x\rvert\sqrt{x^2-1}}\ \mathrm{d}x$	$\sec^{-1}(x) + C$	$\frac{-1}{\lvert x\rvert\sqrt{x^2-1}}\ \mathrm{d}x$	$\csc^{-1}(x) + C$
$\frac{1}{x^2+1}\ \mathrm{d}x$	$\tan^{-1}(x) + C$	$\frac{-1}{x^2+1}\ \mathrm{d}x$	$\cot^{-1}(x) + C$

Table 4.4: A table of integrals you should know. If you have Table 2.1 memorized, this should be very easy to remember—these are just the opposites of the known derivatives.

4.6.3 Split The Fraction

We now continue with a "new" method of integration. This integration method is more of an algebraic trick than a method of calculus, but it is still an important method of integration that may make integrals that look tricky actually become much more clear. In algebraic form, the rule is simply

$$\frac{a+b}{c} = \frac{a}{c} + \frac{b}{c}.$$

This is just the inverse of the rule for adding fractions. When we have two fractions with the same denominator and want to add them, we just add the numerators, which is exactly as we did above, just reversed. This rule can be applied in the same way for larger fractions. For example,

$$\frac{a+b+c+d}{x} = \frac{a}{x} + \frac{b}{x} + \frac{c}{x} + \frac{d}{x}.$$

We will now go through an example where this method is helpful. Consider the function $f(x) = \frac{x^3 - 2x + 3}{x^2}$. Find the integral of $f(x)$.

Without the above method, this would be a very difficult integral. However, using the above method, we can simply split up this fraction into three separate fractions then use property 2 of integrals and integrate each fraction separately:

$$\int f(x)\,\mathrm{d}x = \int \frac{x^3-2x+3}{x^2}\,\mathrm{d}x = \int \frac{x^3}{x^2} - \frac{2x}{x^2} + \frac{3}{x^2}\,\mathrm{d}x = \int x\,\mathrm{d}x + \int \frac{-2}{x}\,\mathrm{d}x + \int \frac{3}{x^2}\,\mathrm{d}x.$$

We now consider each integral separately. The first integral is a simple power rule. Here, the exponent is 1, so

$$\int x\,\mathrm{d}x = \frac{x^2}{2} + C_1.$$

For the second integrand, we must remember first that, by property 5 of integration, we can pull the constant term -2 out of the integral. Then, we must recall that the integral of $\frac{1}{x}$ is $\ln|x| + C$, so we have

$$\int \frac{-2}{x}\,\mathrm{d}x = -2\int \frac{1}{x}\,\mathrm{d}x = -2\ln|x| + C_2.$$

For the third integral, we again use property 5 of integrals to pull the 3 out of the integral so that we are left with $3\int \frac{1}{x^2}\,\mathrm{d}x$. Then, using the algebraic rule that $a^{-b} = \frac{1}{a^b}$, we have a simple power rule:

$$3\int \frac{1}{x^2}\,\mathrm{d}x = 3\int x^{-2}\,\mathrm{d}x = 3\frac{x^{-1}}{-1} + C_3 = \frac{-3}{x} + C_3.$$

Now, taking the sum of our three solutions, we have

$$\int \frac{x^3-2x+3}{x^2}\,\mathrm{d}x = \frac{x^2}{2} - 2\ln|x| - \frac{3}{x} + C,$$

where we used the fact that the sum of three arbitrary constants C_1, C_2, C_3 will give us another arbitrary constant C.

4.7 Substitution Method

Before we discuss how to perform a u-substitution (referred to by most as a "u-sub"), we first discuss the idea behind a u-substitution. Consider a function of the form $f'(g(x))g'(x)$. We know that the antiderivative of this function will be $f(g(x))$. How do we know this? Recall the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}x}f(g(x)) = f'(g(x))g'(x).$$

Then, integrating both sides,

$$\begin{aligned} \int \frac{\mathrm{d}}{\mathrm{d}x}f(g(x))\,\mathrm{d}x &= \int f'(g(x))g'(x)\,\mathrm{d}x \\ f(g(x)) + C &= \int f'(g(x))g'(x)\,\mathrm{d}x. \end{aligned}$$

This is what we hope to achieve when performing a u-substitution–an inverse of the chain rule. How do we do this? What do we substitute? To allow ourselves to compute an integral of this form, we let $u = g(x)$, the function who's derivative is also present in the expression we are trying to integrate. Since this effectively "removes" the x's from the problem (we now are trying to integrate $f'(u)u'$), we must change the $\mathrm{d}x$ so that we are integrating with respect to u without changing the result of the integral. To do this, we can find an expression for $\mathrm{d}x$ in terms of u by taking the derivative of $u = g(x)$ with respect to x. By the chain rule,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}u &= \frac{\mathrm{d}}{\mathrm{d}x}g(x) \\ 1 \cdot \frac{\mathrm{d}u}{\cancel{\mathrm{d}x}} &= g'(x)\frac{\mathrm{d}x}{\cancel{\mathrm{d}x}} \\ \mathrm{d}u &= g'(x)\ \mathrm{d}x \\ \frac{\mathrm{d}u}{g'(x)} &= \mathrm{d}x. \end{aligned}$$

If we make this replacement into our new integrand, we have

$$\int f'(g(x))g'(x)\ \mathrm{d}x = \int f'(u)\cancel{g'(x)}\frac{\mathrm{d}u}{\cancel{g'(x)}} = \int f'(u)\ \mathrm{d}u = f(u) + C = f(g(x)) + C.$$

Alternatively, instead of solving for $\mathrm{d}x$, we can simply replace $g'(x)\ \mathrm{d}x$ with $\mathrm{d}u$, since we found that $g'(x)\ \mathrm{d}x = \mathrm{d}u$, as we will do in the following example.

Consider the function $h(x) = 2x\cos(x^2)$. What is the integral of $h(x)$? We begin by noticing that x^2 is contained in another function and that its derivative, $2x$, is sitting outside. This means that we can make a substitution where $u = x^2$ and $\mathrm{d}u = 2x\ \mathrm{d}x$. We can then make these substitutions,

$$\int 2x\cos(x^2)\ \mathrm{d}x = \int \cos(x^2)2x\ \mathrm{d}x = \int \cos(u)\ \mathrm{d}u = \sin(u) + C = \sin(x^2) + C,$$

where C is the constant of integration. We verify this by taking the derivative:

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big[\sin\left(x^2\right) + C\Big] = \frac{\mathrm{d}}{\mathrm{d}x}\sin(x^2) + \frac{\mathrm{d}}{\mathrm{d}x}C = 2x\cos(x^2) + 0.$$

So our result was correct. Let's try another. Find all $f(x)$ satisfying $f(x) = \displaystyle\int \frac{1}{4x^2+2}\ \mathrm{d}x$. Remember, we only know how to solve integrals of the form $\frac{1}{u^2+1}$ (this is the derivative inverse tangent), so our goal is to rewrite the given integral in such a form. We start by recognizing that the constant term must be a 1, so we can begin by factoring out a 2 from the denominator,

$$\int \frac{1}{4x^2+2}\ \mathrm{d}x = \int \frac{1}{2(2x^2+1)}\ \mathrm{d}x = \frac{1}{2}\int \frac{1}{2x^2+1}\ \mathrm{d}x.$$

Now, remember we want the denominator to have some variable squared by itself, not a variable squared times a constant (in this case, the variable is x and the constant is 2). We can effectively remove this constant by moving it inside the square by taking its square root:

$$2x^2 = (\sqrt{2}x)^2.$$

This is valid because $(ab)^2 = a^2b^2$, so $(\sqrt{2}x)^2 = \sqrt{2}^2x^2 = 2x^2$. Using this, we can rewrite the integral as $\displaystyle\frac{1}{2}\int \frac{1}{(\sqrt{2}x)^2+1}\ \mathrm{d}x$. Now, the integral is nearly in the form where we can easily integrate it. We can get it into the form $\frac{1}{u^2+1}$ by making a substitution where $u = \sqrt{2}x$ and $\mathrm{d}u = \sqrt{2}\,\mathrm{d}x$ so $\mathrm{d}x = \frac{\mathrm{d}u}{\sqrt{2}}$. Making this substitution,

$$\frac{1}{2}\int \frac{1}{(\sqrt{2}x)^2+1}\,\mathrm{d}x = \frac{1}{2}\int \frac{1}{u^2+1}\frac{\mathrm{d}u}{\sqrt{2}} = \frac{1}{2\sqrt{2}}\int \frac{1}{u^2+1}\,\mathrm{d}u$$
$$= \frac{1}{2\sqrt{2}}\tan^{-1}(u) + C = \frac{1}{2\sqrt{2}}\tan^{-1}(\sqrt{2}x) + C.$$

You should differentiate this result and ensure that it is in fact the integral of $\frac{1}{4x^2+2}$.

Let's take a look at another example, $\displaystyle\int e^x\sqrt{e^x+1}\,\mathrm{d}x$. Here, notice that the derivative of $e^x + 1$, e^x, is right outside the square root, so we can "remove" it by making a substitution $u = e^x + 1$ and $\mathrm{d}u = e^x\mathrm{d}x$. Then,

$$\int e^x\sqrt{e^x+1}\,\mathrm{d}x = \int \sqrt{u}\,\mathrm{d}u = \frac{u^{\frac{1}{2}+1}}{\frac{1}{2}+1} + C = \frac{2}{3}\sqrt{u^3} + C = \frac{2}{3}\sqrt{(e^x+1)^3} + C.$$

Again, you should verify this result by taking its derivative.

Consider the example $\displaystyle\int \frac{(\ln x)^2}{x}\,\mathrm{d}x$. Here, we must notice that the derivative of $\ln x$, $\frac{1}{x}$, is present in this problem. This may be easier to see if we write our integral as

$$\int \frac{(\ln x)^2}{x}\,\mathrm{d}x = \int \frac{1}{x}(\ln x)^2\,\mathrm{d}x.$$

Then, we more clearly see that we can make the substitution where we take $u = \ln x$ and thus $\mathrm{d}u = \dfrac{1}{x}\mathrm{d}x$:

$$\int \frac{1}{x}(\ln x)^2\mathrm{d}x = \int u^2\,\mathrm{d}u = \frac{u^3}{3} + C = \frac{(\ln x)^3}{3} + C.$$

Again, differentiating this result will help in your understanding and is generally good practice to ensure no mistakes were made.

For our final example, consider the integral $\displaystyle\int \frac{x+1}{\sqrt{x-1}}\,\mathrm{d}x$. In this problem, there isn't an obvious place to start—as we have discussed u-substitution up to this point, there isn't a clear choice of u. However, remember that we know how to easily integrate anything of the form x^n, so we want to try to get this function into that form using a substitution. We can do this by changing the form of the square root from having two arguments to just one by making the substitution $u = x - 1$ and $\mathrm{d}u = \mathrm{d}x$. Then we have

$$\int \frac{x+1}{\sqrt{x-1}}\,\mathrm{d}x = \int \frac{x+1}{\sqrt{u}}\mathrm{d}u.$$

However, before we continue, we must ensure that the integral is completely in terms of u, that is, we must eliminate all x's from the integral. Since we said $u = x - 1$, we know $u + 1 = x$ which means that the numerator, $x + 1$ can be written as $x + 1 = (u + 1) + 1 = u + 2$. This gives us

$$\int \frac{x+1}{\sqrt{x-1}}\,\mathrm{d}x = \int \frac{u+2}{\sqrt{u}}\,\mathrm{d}u = \int u^{-\frac{1}{2}}(u+2)\,\mathrm{d}u = \int u^{\frac{1}{2}} + 2u^{-\frac{1}{2}}\,\mathrm{d}u =$$
$$\frac{2}{3}u^{\frac{3}{2}} + 4u^{\frac{1}{2}} + C = \frac{2}{3}\sqrt{(x-1)^3} + 4\sqrt{x-1} + C.$$

4.7.1 Converting to u bounds

We now move on to a more subtle concept. When we evaluate a definite integral using the method of u-substitution, the bounds change along with the argument of the integral. To illustrate this, it is best to start with an example.

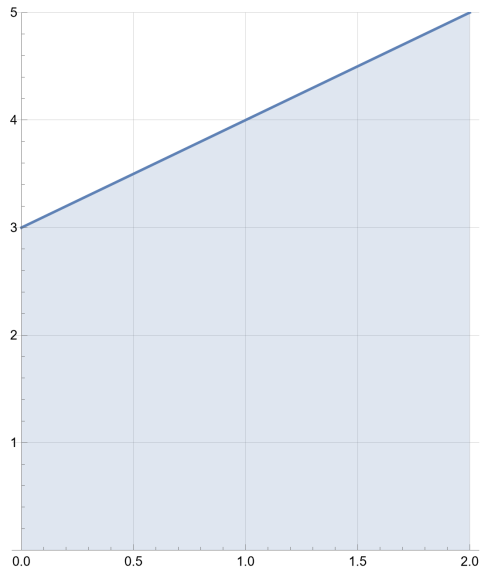


Figure 4.19: A plot of the function $y = x + 3$ over the interval $[0, 2]$.

Using geometry, we can easily see that the area of this shaded region is 8. Let us now use an integral by u-substitution to see if calculus agrees with geometry. Letting $u = x + 3$, we have $\mathrm{d}u = \mathrm{d}x$. Then, integrating, we have:

$$\int_0^2 x + 3 \; \mathrm{d}x \overset{\star}{=} \int_0^2 u \; \mathrm{d}u = \left. \frac{u^2}{2} \right|_0^2 = 2 - 0 = 2.$$

So where did we go wrong? The answer to that question, is that the bounds $[0, 2]$ are the bounds on the variable x, not on u, thus, the error occurs at the starred equals sign. If the lower bound in the x world is $x = 0$ and we know that $u = x + 3$, then the lower bound in the u world must be $u = 0 + 3 = 3$. Similarly, the upper bound with respect to u is $u = 2 + 3 = 5$. If we apply this to this problem, we have:

$$\int_0^2 x + 3 \; \mathrm{d}x = \int_3^5 u \; \mathrm{d}u = \left. \frac{u^2}{2} \right|_3^5 = \frac{25}{2} - \frac{9}{2} = \frac{16}{2} = 8,$$

which agrees with our geometrical analysis. This technique of changing the bounds on the integral becomes especially important in Multivariable Calculus, but can often be avoided in the setting of Calculus I. An alternative to the above method is to switch the integral back to the x variable before applying any bounds. This way, you can skip changing the bounds to the u variable. To do this, first take the indefinite integral of the function (in this case $x + 3$). Using a u-substitution, we get that

$$\int x + 3 \; \mathrm{d}x = \frac{(x+3)^2}{2} + C.$$

If we now apply the original bounds of $[0, 2]$ to this expression, we have

$$\frac{(x+3)^2}{2} + C \Bigg|_0^2 = \frac{25}{2} + \cancel{C} - \left(\frac{9}{2} + \cancel{C}\right) = 8,$$

so calculus once again agrees with geometry. For simplicity, the C is often left out of this type of calculation.

4.8 Net vs. Total Area with Integrals

To begin our discussion on net and total area, we return to an example from section 4.2.4, Net vs. Total Area with Geometry and Symmetry. Below is a graphical depiction of the function $f(x) = x$ over the interval $[-1, 2]$.

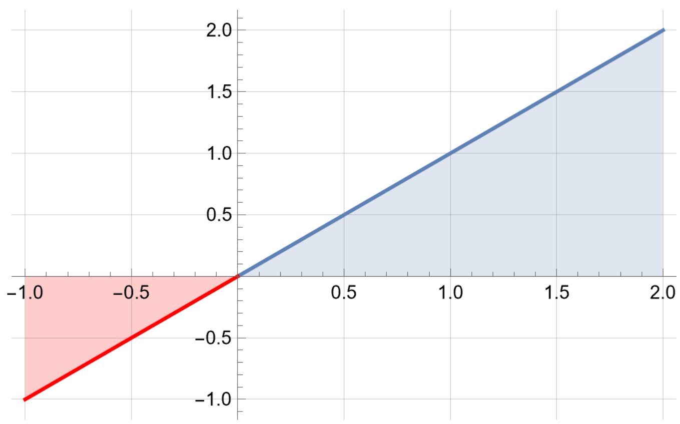


Figure 4.20: A graph of $f(x) = x$ on the interval $[-1, 2]$ with negative area shown in red and positive area shown in blue.

Using geometry, we found the *net* area to be 1.5 and the *total* area to be 2.5. Let's begin by evaluating the integral of $f(x) = x$ over the given interval to see which answer our integral agrees with. We have

$$\int_{-1}^{2} x \, dx = \frac{x^2}{2}\Bigg|_{-1}^{2} = 2 - \frac{1}{2} = 1.5.$$

So it seems that **integrals measure net area**. Further evidence of this claim can be seen by looking at the definition of the definite integral from section 4.2 on Approximating Integrals. Since, as with the LRAM, RRAM, and MRAM approximation methods, we add up rectangles between the function and the x-axis (see figures 4.2, 4.3, and 4.6).

This begs the question, how do we measure the *total* area between a function and the x-axis? To measure total area, we must treat all area as positive. To do this, we can simply make the argument of the integral always positive using absolute value. So, to measure the total area between the above function, $f(x)$, and the x-axis, we must evaluate the integral $\int_{-1}^{2} |x| \, dx$. So how can we evaluate this integral? We've never integrated anything involving absolute value before. The answer is to utilize properties of integrals to turn this into two regular integrals. We do this at every place the value of the part inside the absolute value bars changes from positive to negative (or from negative to positive). In this case, the expression inside the absolute value bars, x, only changes sign at $x = 0$. The first property of integrals stated in section 4.3 allows us to rewrite our integral as follows:

$$\int_{-1}^{2} |x| \, \mathrm{d}x = \int_{-1}^{0} |x| \, \mathrm{d}x + \int_{0}^{2} |x| \, \mathrm{d}x.$$

Now, if we omit the absolute value for a moment, the first of the two integrals on the right-hand side measures the area beneath the x-axis and above the function, while the second measures the area beneath the function and above the x-axis. This means the first integral's value (again omitting the absolute value) will be negative, and the second will be positive. Since the second integral will be positive even without the absolute value, we can remove the absolute value bars entirely. As for the first integral, we can move the absolute value bars to the outside of the integral so that we have only $\left|\int_{-1}^{0} x \, \mathrm{d}x\right|$. This leaves us with

$$\int_{-1}^{2} |x| \, \mathrm{d}x = \left|\int_{-1}^{0} x \, \mathrm{d}x\right| + \int_{0}^{2} x \, \mathrm{d}x = 2.5,$$

as you should verify using the power rule. It is important to note that you can *only* pull absolute value bars out of the integral if the function inside the absolute value does not change signs. For example, $\int_{-\pi}^{\pi} |\sin(x)| \, \mathrm{d}x \neq \left|\int_{-\pi}^{\pi} \sin(x) \, \mathrm{d}x\right|$.

4.9 Tips, Tricks, and Creativity

After completing many integrals, you will begin to notice patterns that will make integration much easier. This section hopes to give an overview of these various tips and tricks, as well as discuss some of the more creative ways of integrating various functions.

4.9.1 A Different Kind of u-substitution

Earlier, we discussed u-substitutions as a method of integration when the derivative of a function in the integrand is also in the integrand. However, u-substitutions are often used in other integrals as well. Take the following integral, for example:

$$\int x\sqrt{x+1} \, \mathrm{d}x.$$

Here, there is no clear option to set as u. As we have presented u-sub previously, this likely doesn't even seem like a time to employ the substitution technique. However, this is a great time to employ it. If we take $u = x + 1$, then $\mathrm{d}u = \mathrm{d}x$. We can also solve this equation for x, yielding $u - 1 = x$. Then, making some substitutions, we have

$$\int (u-1)\sqrt{u} \, \mathrm{d}u = \int u^{\frac{3}{2}} - u^{\frac{1}{2}} \, \mathrm{d}u = \frac{2}{5}u^{\frac{5}{2}} - \frac{2}{3}u^{\frac{3}{2}} + C.$$

Then, substituting back, we have

$$\int x\sqrt{x+1} \, \mathrm{d}x = \frac{2}{5}(x+1)^{\frac{5}{2}} - \frac{2}{3}(x+1)^{\frac{3}{2}} + C.$$

Another, albeit slightly more difficult example, is the following integral:

$$\int \frac{\ln x}{x(\ln(x)+1)^2}\,\mathrm{d}x.$$

Here, you may jump to think that $u = \ln(x)$ is a good substitution. Let's test this theory. If we take $u = \ln(x)$, then $\mathrm{d}u = \frac{1}{x}\,\mathrm{d}x$. Substituting, we have

$$\int \frac{u}{(u+1)^2}\,\mathrm{d}u.$$

At this point, we could perform a second u-substitution in a similar fashion as we did above. This would require a second substitution (of $t = u + 1$, or with some other letter than t), which could get quite confusing. We would then have to substitute back after integrating to have everything in terms of u then again in terms of t. However, this integral can be computed much more simply, with just one u-sub. We said a second u-sub of $t = u + 1$ would be necessary at this point, but, since $u = \ln(x)$, this is just $t = \ln(x) + 1$. This suggests that a more efficient substitution to begin with would have been $u = \ln(x) + 1$. If we work with this substitution as opposed to our earlier one, we still have $\mathrm{d}u = \frac{1}{x}\,\mathrm{d}x$, but we also have $u - 1 = \ln(x)$. Then, substituting, we have

$$\int \frac{u-1}{u^2}\,\mathrm{d}u = \int \frac{u}{u^2} - \frac{1}{u^2}\,\mathrm{d}u = \int \frac{1}{u}\,\mathrm{d}u - \int \frac{1}{u^2}\,\mathrm{d}u.$$

Then, integrating, we have

$$\int \frac{1}{u}\,\mathrm{d}u - \int \frac{1}{u^2}\,\mathrm{d}u = \ln(u) + \frac{1}{u} + C = \ln(\ln(x)+1) + \frac{1}{\ln(x)+1} + C.$$

Learning to notice clever substitutions like these will save you much time.

4.9.2 Creativity

There are times when you have to be quite creative to solve a given integral. Consider the following integral:

$$\int \frac{x^2}{x^2+1}\,\mathrm{d}x.$$

Here, there is no clear first step. This is where you must get creative. We begin by noting that, if the numerator had a $+\,1$ term (like the denominator), this would just be the integral of $\frac{x^2+1}{x^2+1} = 1$. So then, how can we get this $+\,1$ term in the numerator? We can do this by adding and subtracting one, effectively adding zero and not changing the integrand:

$$\int \frac{x^2}{x^2+1}\,\mathrm{d}x = \int \frac{x^2+1-1}{x^2+1}\,\mathrm{d}x.$$

We can now split this fraction into two, one with our desired $\frac{x^2+1}{x^2+1}$, allowing us to complete the integral:

$$\int \frac{x^2+1-1}{x^2+1}\,\mathrm{d}x = \int \frac{x^2+1}{x^2+1}\,\mathrm{d}x - \int \frac{1}{x^2+1}\,\mathrm{d}x = x - \tan^{-1}(x) + C.$$

Adding zero, as we have done above, has already shown to be extremely powerful. We will now take a look at another integral that requires a similar, but unique, technique. Consider the following trigonometric integral,

$$\int \sec(x)\,\mathrm{d}x.$$

This integral went unsolved for many years, but can actually be solved fairly simply, if you get creative enough. When rationalizing the denominator of a fraction like $\frac{1}{\sqrt{2}}$, we typically use a trick to move the irrational part, $\sqrt{2}$, to the numerator. This trick is to multiply by $\frac{\sqrt{2}}{\sqrt{2}} = 1$. Multiplying by one can be a powerful technique, and it is what will allow us to solve this integral. If we multiply the numerator and denominator of the fraction by the expression $\sec(x)+\tan(x)$, we will get something far more manageable. Doing so,

$$\int \sec(x)\frac{(\sec(x)+\tan(x))}{(\sec(x)+\tan(x))}\,\mathrm{d}x = \int \frac{\sec^2(x)+\sec(x)\tan(x)}{\sec(x)+\tan(x)}\,\mathrm{d}x.$$

Now, take a look at the denominator. The derivative of $\sec(x)$ is $\sec(x)\tan(x)$, and the derivative of $\tan(x)$ is $\sec^2(x)$. Both of these terms are present in the numerator. Adding them together, $\frac{\mathrm{d}}{\mathrm{d}x}\Big(\sec(x)+\tan(x)\Big) = \sec(x)\tan(x)+\sec^2(x)$, exactly what is in the numerator. Thus, a u-substitution is called for, with $u = \sec(x)+\tan(x)$ and $\mathrm{d}u = (\sec(x)\tan(x)+\sec^2(x))\,\mathrm{d}x$. Making this substitution, we have

$$\int \sec(x)\,\mathrm{d}x = \int \frac{\sec^2(x)+\sec(x)\tan(x)}{\sec(x)+\tan(x)}\,\mathrm{d}x = \int \frac{1}{u}\,\mathrm{d}u = \ln|u| + C.$$

Now, substituting back, we finally have that

$$\int \sec(x)\,\mathrm{d}x = \ln|\sec(x)+\tan(x)| + C.$$

The actual integration here was not hard, but there was some trickery that is difficult to notice. In future chapters, we will discuss other methods to tackle this specific integral.

4.10 Extra Examples

4.11 Chapter 4 Summary and Exercises

Chapter 5

Applications of Integrals

5.1 Initial Value Problems

5.2 Average Value of a Function

A simple, yet important application of integrals is finding the average value of a function. We can estimate the average value of a function $f(x)$ on $[a, b]$ by sampling points in the interval $[a, b]$. If we sample n points, the approximate average is given by

$$\text{Average value of } f(x) \text{ on } [a, b] \approx \frac{1}{n} \sum_{i=1}^{n} x_i.$$

If we sample every point, that is, we take the limit as $n \to \infty$, we see that the average value of $f(x)$ is given by an integral. We can see this more clearly by setting $\Delta x = \frac{b-a}{n}$:

$$\text{Average value of } f(x) \text{ on } [a, b] = \lim_{n\to\infty} \sum_{i=1}^{n} x_i \frac{b-a}{(b-a)n} = \frac{1}{b-a} \lim_{n\to\infty} \sum_{i=1}^{n} x_i \Delta x.$$

The right-hand-side here is just the limit of a Riemann sum, exactly an integral. We summarize our results below.

Average Value of a Function

Suppose $f(x)$ is integrable on $[a, b]$. Then the average value of $f(x)$ on this interval is given by

$$\text{Average value of } f(x) \text{ on } [a, b] = \frac{1}{b-a} \int_a^b f(x)\, \mathrm{d}x.$$

Curiously, the average value of a function $f(x)$ on $[a, b]$ is just the area between $f(x)$ and the x-axis divided by the length of the interval. One may expect this, however. If you multiply the average height of an object by its length, you get its area (this is exactly the idea in finding the area of a triangle or a trapezoid).

Example 8. Find the average value of $f(x) = x^3 - 2x - 1$ on the interval $[-1, 3]$.

Solution. The length of this interval is $3-(-1)=4$. Then the average value of the function is

$$\begin{aligned}\frac{1}{4}\int_{-1}^{3} x^3-2x-1\ \mathrm{d}x &= \frac{1}{4}\left(\frac{x^4}{4}-x^2-x\right)\Bigg|_{-1}^{3}\\ &= \frac{1}{4}\left(\frac{81}{4}-9+3\right)-\frac{1}{4}\left(\frac{1}{4}-1+1\right)\\ &= \frac{7}{2},\end{aligned}$$

so the average value of $f(x)$ over $[-1,3]$ is $\frac{7}{2}$. □

5.3 1 Dimensional Particle Motion Revisited

Exactly as we used derivatives to find the velocity and acceleration of a particle, we can learn about a particle's motion using integrals. Since integrals are functionally the opposite of derivatives, the functions of position, velocity, and acceleration are related to one another via integrals in the opposite manner that they are with derivatives. Thus, given the velocity function of a particle, we can find its position function, and, given the acceleration function of a particle, we can find its velocity function. However, there is an important caveat.

Suppose that the position of a particle is modeled by the position function $s(t)=\frac{t^3}{3}-2t+1$. Its corresponding velocity function would then be $s'(t)=v(t)=t^2-2$, and its acceleration function would be $v'(t)=a(t)=2t$. However, if we integrate the acceleration, we get t^2+C. Integrating the velocity, we get $\frac{t^3}{3}-2t+C$. Each of these calculations is off by a constant. If we knew that $s(0)=1$ and $v(0)=-2$, we would have been able to get the correct velocity and position functions from the acceleration functions. Thus, we need an extra condition, an initial value. Without an initial value, we cannot find a specific velocity or position function using integrals.

Example 9. The acceleration of a particular particle is given by $a(t)=3t-6$. Also, the velocity at time $t=2$ seconds is $v(2)=3$ meters per second, and that the position at time 0 is $s(0)=4$ meters. Find the position at time $t=2$ seconds.

Solution. Since we have the acceleration, we can integrate it and use the initial condition for the velocity of the particle to find the velocity function. Doing so,

$$v(t)=\int a(t)\ \mathrm{d}t=\int 3t-6\ \mathrm{d}t=\frac{3t^2}{2}-6t+C.$$

We can solve for C by plugging in $t=2$ since we know the value of $v(t)$ at that time. If we do so, we see that

$$3=v(2)=\frac{12}{2}-12+C=-6+C,$$

so $C=9$ meters per second. This means that $v(t)=\frac{3t^2}{2}-6t+9$. To find the position, we integrate velocity and apply the initial condition:

$$s(t)=\int v(t)\ \mathrm{d}t=\int \frac{3t^2}{2}-6t+9\ \mathrm{d}t=\frac{t^3}{3}-3t^2+9t+C.$$

Using our initial condition,

$$4 = s(0) = C,$$

so $C = 4$ meters. Thus $s(t) = \frac{t^3}{2} - 3t^2 + 9t + 4$. Plugging in $t = 2$,

$$s(2) = \frac{2^3}{2} - 3(2)^2 - 9(2) + 4 = -22.$$

The position of the particle at time $t = 2$ seconds is -22 meters. □

Example 10. Find the average velocity of a particle whose position, in meters, is modeled by the function $s(t) = e^{2t} - t^3 + 2$ between time $t = 0$ and $t = 3$ seconds.

Solution. The velocity of the particle is measured by the derivative of position,

$$v(t) = s'(t) = 2e^{2t} - 3t^2.$$

To find the average velocity on $[0, 3]$, we find the average value of $v(t)$ on $[0, 3]$. This gives us

$$\begin{aligned} \frac{1}{3-0}\int_0^3 2e^{2t} - 3t^2 \, \mathrm{d}t &= \frac{1}{3}(e^{2t} - t^3)\Big|_0^3 \\ &= \frac{1}{3}(e^6 - 27) - \frac{1}{3}(e^0 - 0) \\ &\approx 125.143, \end{aligned}$$

so the average velocity of this particle on the interval $[0, 3]$ is approximately $125.143\,\frac{\mathrm{m}}{\mathrm{s}}$. □

5.4 Integrals as Accumulation of Change

A natural question one may wonder is how we can interpret integrals in ways other than as the area between a function and an axis. A pleasing answer is that integrals can be used to measure change. This is best seen with examples.

Example 11. Suppose that a particle is oscillating up and down and that its vertical position at time t seconds is $s(t) = \sin(t) + 2$ nanometers, as below. What is the *net distance* traveled by this particle on $[0, 2\pi]$? On $[0, \frac{3\pi}{2}]$? What is the *total distance* traveled by this particle on $[0, 2\pi]$? On $[0, \frac{3\pi}{2}]$?

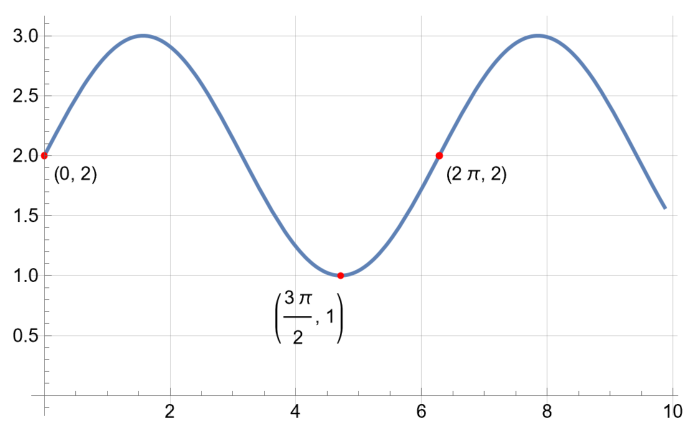


Figure 5.1: A plot of the function $f(x) = \sin(x) + 2$ with the points $(0, 2)$, $(\frac{3\pi}{4}, 1)$, and $(2\pi, 2)$ labeled.

Solution. *Net distance* measures the change from starting position to ending position. To calculate it, we just subtract the initial position from the ending position. On $[0, 2\pi]$, the net distance traveled is then $s(2\pi) - s(0) = 2 - 2 = 0$ nanometers. On $[0, \frac{3\pi}{2}]$, the net distance traveled is $s(\frac{3\pi}{2}) - s(0) = 1 - 2 = -1$ nanometers. Note that we also could have calculated each of these as

$$\int_0^{2\pi} s'(t)\ dt = \int_0^{2\pi} v(t)\ dt$$

and

$$\int_0^{3\pi/2} s'(t)\ dt = \int_0^{3\pi/2} v(t)\ dt,$$

respectively. So integrals measure net change, and thus **the area under the curve of $f'(x)$ on $[a, b]$ is exactly the net change in $f(x)$ on $[a, b]$.** This is another way to state the Fundamental Theorem of Calculus (2).

Total distance measures how far the particle travels, no matter the direction. Total distance is often a more useful measurement than net distance. For example, if you were to go for a run, starting at the same place you stopped, your net distance traveled would be 0, no matter how many miles you actually ran (your total distance). To measure total distance, we want to treat all distance as positive. Thus, instead of measuring this distance by integrating $s'(t) = v(t)$, we integrate $|s'(t)| = |v(t)|$ to treat everything as positive. Then the total distance traveled by the particle on $[0, 2\pi]$ is

$$\int_0^{2\pi} |s'(t)|\ dt = \int_0^{2\pi} |\cos(t)|\ dt.$$

To evaluate this integral, we consider the plot below.

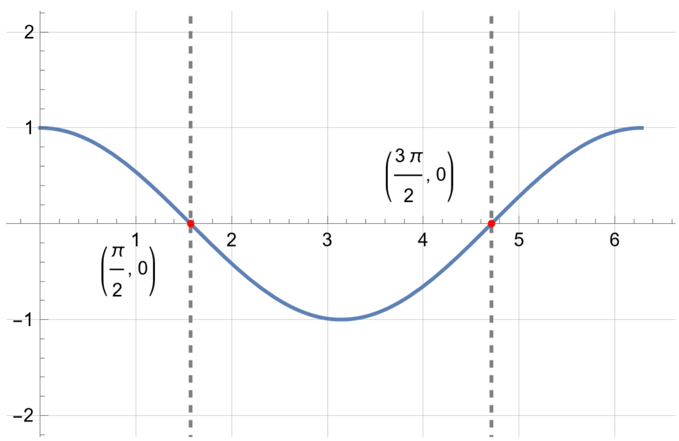


Figure 5.2: A plot of the function $\cos(x)$ on $[0, 2\pi]$ with the x-intercepts labeled.

Notice that on the intervals $[0, \frac{\pi}{2}]$ and $[\frac{3\pi}{2}, 2\pi]$, the function is positive and on $[\frac{\pi}{2}, \frac{3\pi}{2}]$, negative. Thus we split up the integral into these three parts, omitting the absolute value on the positive intervals:

$$\int_0^{2\pi} |\cos(t)|\ dt = \int_0^{\pi/2} \cos(t)\ dt + \left| \int_{\pi/2}^{3\pi/2} \cos(t)\ dt \right| + \int_{3\pi/2}^{2\pi} \cos(t)\ dt = 4.$$

Thus the total distance traveled by this particle is 4 nanometers. For $[0, \frac{3\pi}{2}]$, we see that the total distance traveled by the particle is

$$\int_0^{3\pi/2} |\cos(t)|\ dt = \int_0^{\pi/2} \cos(t)\ dt + \left| \int_{\pi/2}^{3\pi/2} \cos(t)\ dt \right| = 2 \text{ nanometers}.$$

□

5.5 Area Between Curves

Using integration, we can find the area of irregularly shaped regions. We do this by defining the region with multiple functions, then finding the area between those functions, or *curves*.

Example 12. Find the area of the shaded region below. The curve in blue is $f(x) = x$, and the curve in red is $g(x) = (x-3)^2 + 1$.

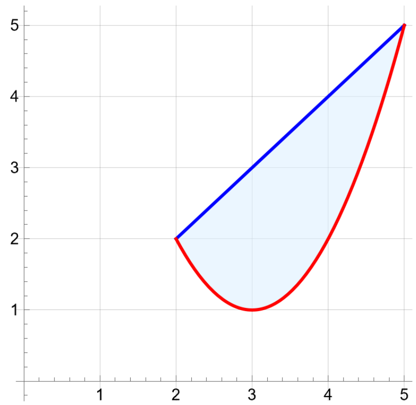


Figure 5.3: A plot of $f(x)$ and $g(x)$, with the area between them shaded.

Solution. To solve a problem like this, we must remember what an integral solves for. The following is a visual representation of how we will solve this problem.

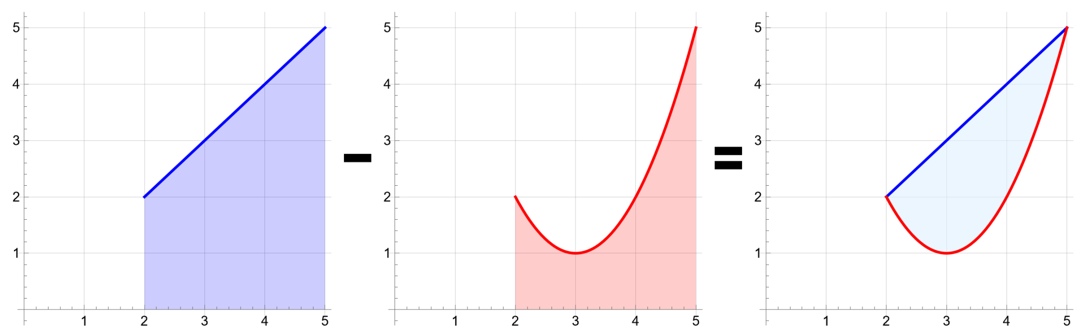


Figure 5.4: A graphical depiction of the area under $f(x)$ minus the area under $g(x)$.

To do this algebraically, we must set up and evaluate both integrals:

$$\begin{aligned} \int_2^5 f(x)\ \mathrm{d}x - \int_2^5 g(x)\ \mathrm{d}x &= \int_2^5 f(x) - g(x)\ \mathrm{d}x \\ &= \int_2^5 x - (x-3)^2 - 1\ \mathrm{d}x \\ &= \left(\frac{x^2}{2} - \frac{(x-3)^3}{3} - x\right)\Bigg|_2^5 \\ &= \frac{21}{2} - 3 - 3 = \frac{9}{2}. \end{aligned}$$

So the area bounded by these two curves is $\frac{9}{2}$. □

Example 13. Find the area of the shaded region below. The curve in blue is $x = f(y) = (y-5)^2+1$, and the curve in red is is $x = g(y) = -(y-5)^2 + 9$, each defined on the interval $[3, 7]$.

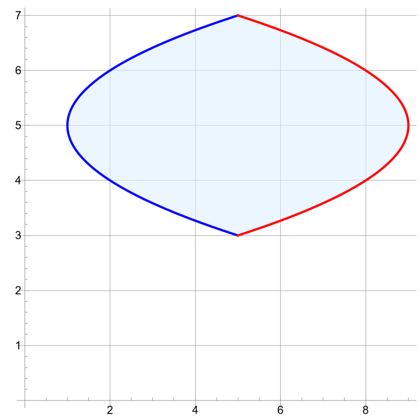


Figure 5.5: A plot of $f(x)$ and $g(x)$, with the area between them shaded.

Solution. While similar to our first example, there is a slight difference. While it could be done using the same method as in our first example[1], that would be a long computation and would allow errors in computation to occur with ease. Instead of solving for y and integrating with respect to x, we will integrate with respect to y. We again start by visualizing this:

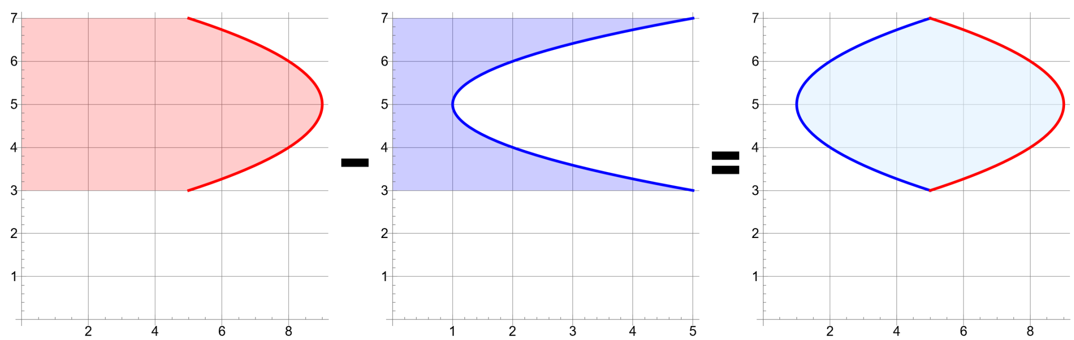


Figure 5.6: A graphical depiction of the area between $f(x)$ and the y-axis minus the area between $g(x)$ and the y-axis.

Again, to find the numerical value of this area, we set up both integrals, and subtract them:

$$\int_3^7 -(y-5)^2 + 9 \text{ dy} - \int_3^7 (y-5)^2 + 1 \text{ dy} = \frac{64}{3}.$$

We leave the computation of these integrals to the readers. As an addition exercise, try computing this area using the method of the prior example (integrating with respect to x instead of y). □

5.6 Volumes of Revolution

Not only can integrals be used to measure area, they can also measure volume. Using integration to find the volume of various solids is a large topic in Calculus 3, and here we will see a very small

[1]To do this, one would solve for $x = f(y)$ and $x = g(y)$ for y, then solve the resulting 4 integrals. The 4 integrals instead of two are a result of having a $\pm$ when solving for y since these are quadratic equations.

sample of that. Before getting into the specifics, we will first get an idea of what it means to revolve a curve about an axis and how to visualize it.

Consider the curve below.

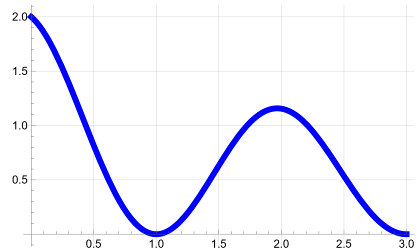


Figure 5.7: A plot of the curve $f(x) = \frac{\cos(\pi x)+1}{\sqrt{x+1}}$

If we embed this curve into the 3 dimensional world, one with an x, y, and z-axis, we would then have the following plot:

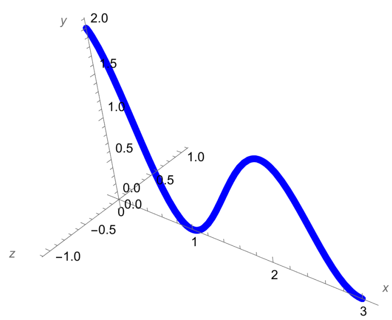


Figure 5.8: A plot of the two dimensional curve $f(x)$ embedded into three dimensional space.

If we were to take this curve and revolve it around the x-axis, we would wind up with the 3D-shape below, called a *surface*.

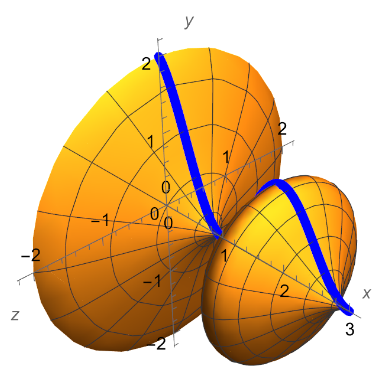


Figure 5.9: A 3D-plot of the curve $f(x)$ revolved about the x-axis.

We may instead want to rotate this curve about the y-axis, resulting in the following plot.
In general, there's nothing "special" about rotating around the x or y-axis. We might rotate about some other axis, called the *axis of rotation*, such as the vertical line $x = 5$, the horizontal line

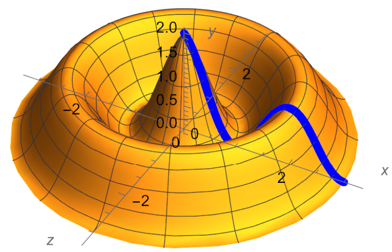


Figure 5.10: A 3D-plot of the curve $f(x)$ revolved about the y-axis.

$y = -1$, or even a more complex axis like $y = x + 12$. The axis of rotation can even be inside the curve (in our above example, such as the line $y = 1$), though the interpretation as a physical 3D solid is then slightly more difficult. Here, we will only cover rotations about simple axes, horizontal and vertical lines.

5.6.1 The Disc Method

To find the volume under a surface like the ones above, one option is to use the *disc method*. To do this, we imagine slices of the surface, all of which are discs (circles), as depicted below.

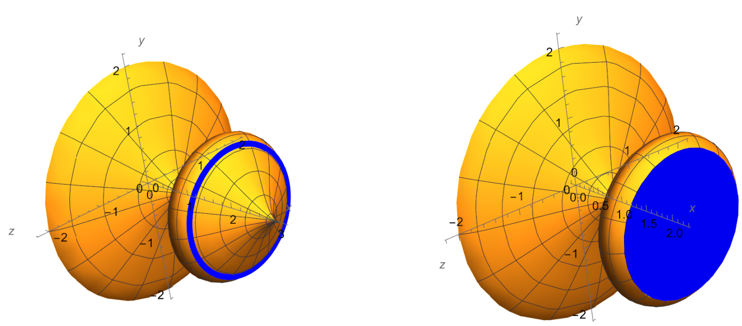


Figure 5.11: A disc (slice) of the revolved $f(x)$ when $x = 2.2$.

Since each slice is a circle, its area is given by πr^2, where r is the radius. Here, for a given circle at $x = a$, its radius is exactly $f(a)$, the distance from the x-axis to the curve at a. Now, recall that, to find the volume of a cylinder with radius r and height h, we compute $\pi r^2 h$. This formula can be viewed as finding first the area of any circular cross-section, πr^2, then multiplying by the height h, and adding up all these little cross-sections. Thus, we might view the formula for the volume of a cylinder as an integral,

$$\text{Volume of a Cylinder} = \pi r^2 h = \int_0^h \pi r^2 \, \mathrm{d}x.$$

If the height started at $x = a$ and ended at $x = b$ so that $h = b - a$, we might instead write

$$\text{Volume of a Cylinder} = \pi r^2 (b - a) = \int_a^b \pi r^2 \, \mathrm{d}x.$$

We use the exact same method here, except now the "cylinder" is of variable radius, $f(x)$. Taking $r = f(x)$, we see that the volume of a curve $f(x)$ on $[a, b]$ revolved around the x-axis is as follows.

Discs Method

The volume of the resulting solid after rotating an integrable curve $f(x)$ on $[a, b]$ about an axis can be found using the *discs method*,

$$\text{Volume} = \int_a^b \pi (f(x))^2 \, dx = \pi \int_a^b f(x)^2 \, dx.$$

Note that this is not the most general form of the formula, and some set up will still be required when completing most problems (see Example 15).

Example 14. Find the volume of the resulting solid when the curve $f(x) = \frac{\cos(\pi x)+1}{\sqrt{x+1}}$ on $[0, 3]$ is revolved around the x-axis.

Solution. This is the curve we have been working with throughout this section. For plots, see Figures 5.9 and 5.11. Evaluating the integral, we see that the volume of this solid is

$$\pi \int_0^3 \left(\frac{\cos(\pi x) + 1}{\sqrt{x+1}} \right)^2 \, dx \approx 7.06703.$$

□

Example 15. Find the volume of the resulting solid when the area enclosed by the curves $f(x) = x^2$ and $g(x) = \sqrt{x}$ is rotated about the line $x = 2$.

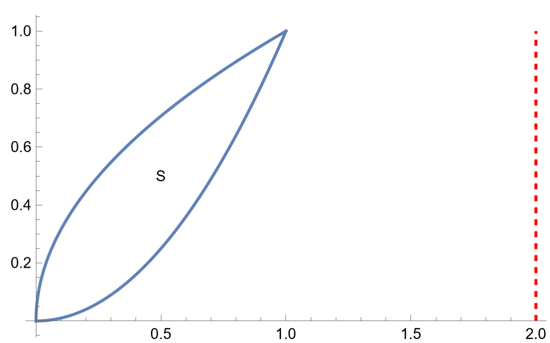


Figure 5.12: A plot of the region S bounded by the curves $f(x) = x^2$ and $g(x) = \sqrt{x}$ and with the axis of rotation, $x = 2$.

Solution. To solve this problem, we will use an extension of the discs method called the *washers method* (think of a physical washer). We will find the volume of the resulting solid when $f(x)$ is rotated about the line $x = 2$, then subtract from it the volume of the resulting solid when $g(x)$ is rotated about the line $x = 2$ (exactly the same way that we would find the enclosed area between the two curves, this time with volume of their resulting surfaces).

Since we're rotating about a vertical line, we want to integrate with respect to y here. We solve the equations for x:

$$f(x) = y = x^2 \implies x = \sqrt{y},$$
$$g(x) = y = \sqrt{x} \implies x = y^2.$$

We need not worry about the $\pm$ since we are only considering positive x. We have in mind the following picture:

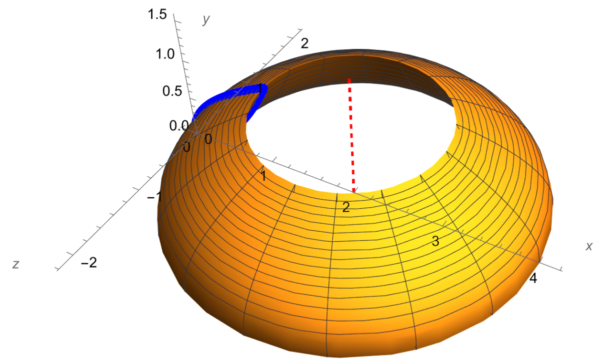


Figure 5.13: A plot of the region S rotated about the line $x = 2$.

The distance between the leftmost curve ($y = f(x) = \sqrt{x}$) and $x = 2$ is given by

$$R = 2 - y^2$$

and the distance between the rightmost curve ($y = g(x) = x^2$) and $x = 2$ is given by

$$r = 2 - \sqrt{y}.$$

To understand why the "$2 -$ " is there, note that the distance between each x value and the red dashed line is $2 - x$. We now find the volume of the resulting solid. Our "big radius" is $R = 2 - y^2$ and our "small radius" is $r = 2 - \sqrt{y}$, so:

$$\begin{aligned}
\text{Volume} &= \pi \int_0^1 R^2 - r^2 \, dy \\
&= \pi \int_0^1 (2 - y^2)^2 - (2 - \sqrt{y})^2 \, dy \\
&= \pi \int_0^1 4 + y^4 - 4y^2 - 4 - y + 4\sqrt{y} \, dy \\
&= \pi \left(4 + \frac{1}{5} - \frac{4}{3} - 4 - \frac{1}{2} + \frac{8}{3}\right) = \frac{31\pi}{30}.
\end{aligned}$$

□

5.6.2 The Cylindrical Shell Method

The *cylindrical shell method*, or *shell method* for short, is another way to measure the volume of these 3-dimensional solids that are the product of revolutions. Here, instead of thinking of circular cross sectional areas, we think of the solid as an infinite number of cylinders next to one another.

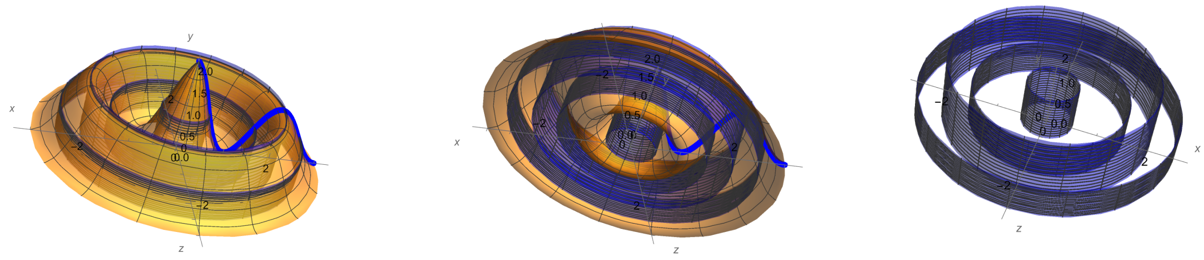


Figure 5.14: Plots of $f(x) = \frac{\cos(\pi x)+1}{\sqrt{x+1}}$ revolved about the y-axis with cylinders with radii $x = 0.5, 1.5, 2, 2.5$ and corresponding heights $f(0.5), f(1.5), f(2), f(2.5)$ centered at the origin from above and below. A plot of the cylinders without $f(x)$.

The plot below helps to illustrate this.
To find the volume of the resulting shape of the revolution of a curve $f(x)$ on $[a, b]$ about the y-axis, we add up the volume of each cylindrical shell. Each cylinder has a radius x and height $f(x)$, along with a small thickness Δx. The corresponding area of the base of a shell is then $2\pi x\Delta x$. We multiply by the height of the cylinder, $f(x)$, to get a volume of each cylindrical shell as $2\pi x f(x)\Delta x$. Roughly speaking, if we make each cylinder infinitely thin and add up all the volumes, we then get the exact volume of the solid, as follows.

Cylindrical Shells Method

The volume of the resulting solid after rotating an integrable curve $f(x)$ on $[a, b]$ about an axis can be found using the *cylindrical shells method*,

$$\text{Volume} = \int_a^b 2\pi x f(x)\ \mathrm{d}x = 2\pi \int_a^b x f(x)\ \mathrm{d}x.$$

Example 16. Find the volume of the resulting solid when the curve $f(x) = \frac{\cos(\pi x)+1}{\sqrt{x+1}}$ on $[0, 3]$ is revolved around the y-axis.

Solution. This is the curve we have been working with throughout this section. For plots, see Figures 5.10 and 5.14. Evaluating the integral, we see that the volume of this solid is

$$2\pi \int_0^3 \frac{x\cos(\pi x) + x}{\sqrt{x+1}}\ \mathrm{d}x \approx 15.9985.$$

□

Example 17. Let $f(x) = x^2$. When the region bounded above by $f(x)$, below by the x-axis, and on the left and right by the lines $x = 1$ and $x = 2$ is revolved around the y-axis, find the volume of the resulting region.

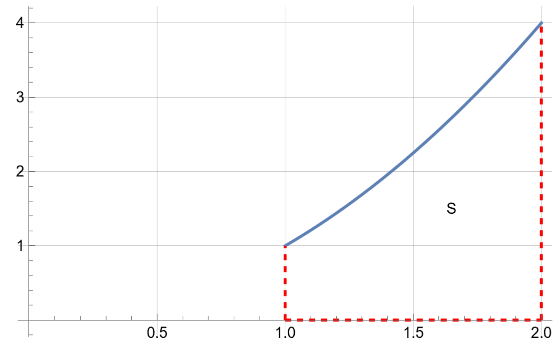


Figure 5.15: A plot of the region S bounded by the curves $f(x) = x^2$, $x = 1$, $x = 2$, and the x-axis.

Solution. We have in mind the following picture:

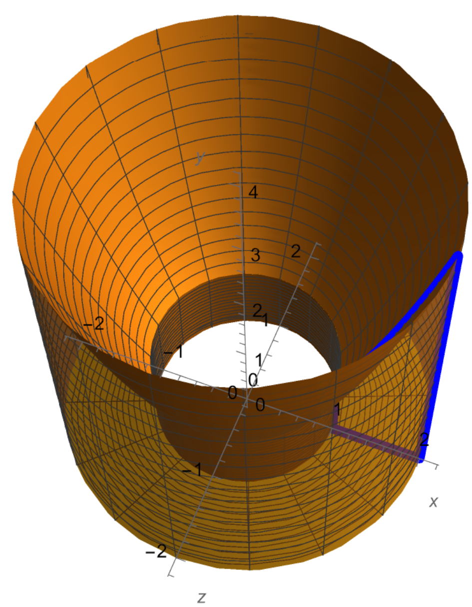


Figure 5.16: A plot of the resulting solid when the region S is revolved around the y-axis.

We now evaluate the integral to get the volume.

$$\begin{aligned} \text{Volume} &= 2\pi \int_1^2 x f(x)\ \mathrm{d}x = 2\pi \int_1^2 x(x^2)\ \mathrm{d}x \\ &= 2\pi \int_1^2 x^3\ \mathrm{d}x = 2\pi \frac{x^4}{4}\bigg|_1^2 \\ &= 2\pi\left(4 - \frac{1}{4}\right) = \frac{15}{2}\pi \approx 23.5619. \end{aligned}$$

□

5.7 Chapter 5 Summary and Exercises

Chapter 6

Introduction to Differential Equations

Part II

Calculus II

Chapter 7

More on Integrals

Chapter 8

Series

Chapter 9

Introduction to the Calculus of Parametric Equations, Polar Equations, and Vector-Valued Functions

Part III

Calculus III

Part IV

Proofs, Answers to Exercises, and more