Calculus I

Daniel Reinholz

May 6, 2009

OpenText License (Version 1.0).

1. Definitions

   1.1. "Creator" means each individual or entity that first made a Work available under this License.

   1.2. "Contributor" means each individual or entity that contributes to the creation of any Derivative Work made available under the terms of this License.

   1.3. "Work" means any creation eligible for copyright protection pursuant to the Copyright Act, United States Code, Title 17, Section 102 (17 U.S.C. § 102).

   1.4. "Derivative Work" has the same meaning given to that term by 17 U.S.C. § 101.

   1.5. "License" means this document.

   1.6. "OpenText" means any Work or Derivative Work made available under the terms of this License.

   1.7. "Larger Work" means any Work combining an OpenText with a Work or Derivative Work not governed by the terms of this License.

   1.8. "Editable Form" means the common form of Works in which modifications are made, typically by use of computer text editing software, and includes but is not limited to the Work in electronic format as a TeX or LaTeX project file, an Oasis OpenDocument project file, or any other format now in existence or later developed in which modifications can be readily made using computer software or any other means.

   1.9. "You" or "Your" means an individual or a legal entity exercising rights under, and complying with all of the terms of, this License. For legal entities, "You" includes any entity which controls, is controlled by, or is under common control with You. For purposes of this definition, "control" means (a) the power, direct or indirect, to cause the direction or management of such entity, whether by contract or otherwise, or (b) ownership of more than fifty percent (50%) of the outstanding shares or beneficial ownership of such entity.

   1.10. "Distribute" means to make an OpenText or a Larger Work, or any portion of either, available, directly or indirectly, to anyone other than You.

2. License Grant.

   2.1. Expressly conditioned on your compliance with the terms of this License, the Creator and Contributors grant You a world-wide, royalty-free, non-exclusive license to use, modify, display, perform, and Distribute the OpenText, with or without modifications, or as part of a Larger Work.

3. Your obligations under this License.

   3.1. You must include a copyright notice bearing the names of the Creator and each subsequent Contributor to any OpenText upon which any Derivative Work you create or distribute is based.

3.2. Any OpenText that You distribute or otherwise make available in any medium now in existence or hereafter developed must also be made available in Editable Form and that Editable Form must also be distributed under the terms of this License. You must include a copy of this License with every copy of the OpenText and must either inform recipients of the OpenText as to how they can obtain the OpenText in Editable Form in a reasonable manner or You must make the OpenText publicly available in Editable Form.

3.3. You may create or distribute a Larger Work, provided that you distribute the OpenText-derived portions of the Larger Work exclusively under the terms of this License and include a copy of this License with every copy of the Larger Work. Any portions of the Larger Work that are not to be governed by the terms of this License must be conspicuously marked so as to avoid any confusion as to which portions of the Larger Work are governed by the terms of this License and which portions are not governed by the terms of this License. In the absence of a conspicuous marking to the contrary, any portion of the Larger Work in question shall be governed by the terms of this License.

3.4. Exhibit 1 to this License contains sample language for incoporating this License into your OpenText or Larger Work and informing downstream users of their rights under this License. This sample language is provided as an example; you may fulfill your obligations under this License using any other method reasonably calculated to give a similar level of notice to downstream users.

3.5. You expressly warrant that you hold sufficient intellectual property rights in any contribution or modification you make to the OpenText to offer that contribution or modification under the terms of this License or under terms of your own choosing, and agree to indemnify, defend, and hold harmless the Creator and Contributors for any claims asserted against or any liability incurred by the Creator or Contributors directly or indirectly resulting from the creation or distribution of your contribution or modification either as part of an OpenText or as part of a Larger Work.

4. Disclaimer of Warranty and Limitation of Liability.

4.1. THIS OpenText IS PROVIDED "AS IS," WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, WARRANTIES THAT THE OpenText IS FREE OF DEFECTS, MERCHANTABLE, FIT FOR A PARTICULAR PURPOSE OR NON-INFRINGING. THE ENTIRE RISK AS TO USE OR DISTRIBUTION OF THE OpenText IS WITH YOU. THIS DISCLAIMER OF WARRANTY CONSTITUTES AN ESSENTIAL PART OF THIS LICENSE. NO USE OF ANY OpenText IS AUTHORIZED HEREUNDER EXCEPT UNDER THIS DISCLAIMER.

4.2. IN NO EVENT SHALL THE Creator OR Contributors BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OR DISTRIBUTION OF AN OpenText OR THIS License, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

5. New Versions of this License.

   5.1. Daniel L. Reinholz is the promulgator and copyright holder of this License and may publish revised or new versions of this License at any time. Each version of this License shall be given a distinguishing version number. If an OpenText has been distributed under a particular version of this License, you may continue to use it under the terms of that version of this License or under the terms of any subsequent version of this License published by Daniel L. Reinholz. Except as provided in section 5.2 of this License, only Daniel L. Reinholz has the right to modify the terms of this License or to create new versions of this License.

   5.2. If You want to create a new license for Your original product, You may create and use a modified version of this License so long as You rename the license and remove any references to the name of the promulgator and make it clear that Your license contains terms which are different from this License. In no event are You permitted to use the label "OpenText" on a modified version of this License without the express written permission of the promulgator.

6. Termination.

   6.1. Expressly conditioned on your compliance with the terms of this License, the copyright license granted herein shall remain in full force and effect for the duration of the applicable copyright unless terminated as provided in sections 6.2. or 6.3. of this License.

   6.2. This License and the rights granted hereunder will terminate automatically if You fail to comply with any of the terms herein and fail to cure such breach within 30 days of becoming aware of the breach. Provisions which, by their nature, must remain in effect beyond the termination of this License shall survive.

   6.3. If you assert a copyright infringement or any other claim against the Creator or Contributors in connection with the creation, use, or distribution of an OpenText or this License, this License shall automatically terminate unless you dismiss, with prejudice, your claim and agree in writing to waive any future claims against the Creator or Contributors on any theory of liability and in any jurisdiction arising in connection with the creation, use, or distribution of the OpenText or this License.

   6.4. In the event of termination, the rights granted to any third-parties under this License with respect to any OpenText or OpenText-derived portions of any Larger Work created or distributed by You shall expressly survive termination of Your rights granted under this License.

7. Purpose and Scope.

   7.1. The purpose of this License is to promote the development of high-quality educational materials and to make the same readily available to students and educators at little or no cost. Its conditions are necessary to enable the Creator and Contributors to benefit from the work of other Contributors and downstream users. The Creator and Contributors maintain significant economic rights in OpenTexts even when they distribute an OpenText free of charge or make it publicly available, and this License is intended to protect those rights in a manner that promotes collaboration among the Creator and Contributors.

7.2. The terms of this License define its scope and are conditions of the copyright license granted to You hereunder. You acknowledge that any use, modification, or distribution of an OpenText that is outside the scope of this License will subject You to liability for copyright infringement in a cause of action brought by the Creator or any Contributor.

8. Miscellaneous Provisions.

8.1. This License represents the complete agreement concerning the subject matter hereof.

8.2. If any provision of this License is held to be unenforceable, such provision shall be reformed only to the extent necessary to make it enforceable, and the remainder of this License shall remain in full force and effect.

8.3. This License shall be governed by the law of the State of California, excluding conflict-of-law provisions. Exclusive jurisdiction and venue over any litigation relating to this License or any OpenText or Larger Work shall lie in the judicial district (state and county in the United States) where the Creator or Contributor who is a party to the litigation resides, and you hereby consent to the personal jurisdiction of, and venue in, that court.

8.4. The application of the United Nations Convention on Contracts for the International Sale of Goods is expressly excluded.

8.5. Any law or regulation which provides that the language of a contract shall be construed against the drafter shall not apply to this License.

8.6. This License shall not be construed as creating any agency, partnership, joint venture, or any other form of legal association or privity between You, the Creator, or any Contributor to an OpenText or a Larger Work. Under no legal theory and under no circumstances shall You, the Creator, or any Contributor be held jointly and severally liable for Your infringement of the intellectual property rights of a third party, nor shall You, the Creator, or any Contributor be held jointly and severally liable for another Contributor or the Creator's infringement of the intellectual property rights of a third party.

8.7. You may translate this License into other languages provided that you also include an unmodified copy of this License in its original English text in any OpenText or Larger Work that You distribute. You agree and certify that You have requested that this License be provided in English, and acknowledge that the terms of this License in its original English text are controlling.

8.8. Failure by the Creator or any Contributor to enforce any provision of this License shall not be deemed a waiver of future enforcement of that or any other provision.

Exhibit 1

Sample notice to be included with an OpenText:

This work is an OpenText, and use is subject to the terms below:

[insert full text of the OpenText License here]

Electronic versions of this work in editable computer word processing document format are available for the cost of media or for free download upon written request made to Michel Ardan, 777 Leviathan Avenue, Nantes, CA 92314.

Sample notice to be included with a Larger Work:

Copyright © 2009, Daniel L. Reinholz
Copyright © 2010, Pierre Aronnax
Copyright © 2010, Axel Lidenbrock
Copyright © 2015, Michel Ardan
Copyright © 2016, Roberto Bellarmino

Portions of this work are an OpenText, and their use is subject to the terms of the OpenText License version 1.0 (or any later version). You can find the full text of the OpenText License here:

[insert full text of the OpenText License here]

The remainder of this work is Copyright © Roberto Bellarmino, all rights reserved. OpenText portions of this work are denoted by the Arial Font, whereas those portions Copyright © Roberto Bellarmino, all rights reserved are denoted by the Courier New Font. Copyright © Roberto Bellarmino, all rights reserved images are denoted by the caption "Copyright © Roberto Bellarmino, all rights reserved." All other images are licensed in accordance with the OpenText license. Electronic versions of the OpenText portions of this work in editable computer word processing document format are available for the cost of media or for free download upon written request made to Roberto Bellarmino, 666 Inquisition Boulevard, Rome, CA 92159.

Suggested Course Outline

In this text I have tried to include all of the standard topics covered in a first-semester calculus course, even the ones I feel would be better left out. Following is a suggestion of how I would structure a course using this text, assuming that a semester consists of 40 lectures, 50 minutes each.

Semester I: Limits, Differentiation, Integration

1. **Functions:** A general review of functions, and motivation of limits through instantaneous velocity (2 lectures total).

2. **Limits:** All topics in this chapter, with 2 lectures for the definition and 1 lecture for all others (11 lectures total).

3. **Differentiation:** Cover the definition through trigonometric functions, with 1 lecture for each topic, except for 2 lectures for the differentiation rules. Cover extreme values and all topics beyond with 1 lecture (14 lectures total).

4. **Integration:** Cover all topics in this chapter up partial fractions, skip the sections on trigonometric integrals, and cover numerical integration through improper integrals, with 1 lecture each (13 lectures total).

In the above syllabus I have omitted many of the traditional *applications* of calculus, such as related rates, calculating volumes, etc. The hope is to provide students with a strong foundation in calculus, so that they will be better served in other courses where they will actually need to apply calculus; most students will never use any of the standard calculus applications outside of an introductory calculus course. I have also favored to omit some of the more obscure techniques of integration, which I do not feel are particularly useful.

The above course outline is ambitious. Depending on the structure of the course and nature of students enrolled, it may be worthwhile to slow down the pace, excluding some topics. In my opinion, order of magnitude analysis could be done without, requiring a brief introduction to the topic at the time l'Hôpital's rule is presented. If absolutely necessary, the lectures on sequences through the bisection method could be omitted, but I feel these are useful topics that are unfortunately seldomly presented, and really drive home the point that calculus is about approximation.

Lastly, the recommendation of one lecture per topic is tenuous at best. For instance, in the beginning of the discussion of differentiation, the lecture on the definition is probably considerably longer than a single lecture, but the follow up lecture on differentials is very short. Thus, an instructor would probably find it most reasonable to present the definition in a day and a half, and spend only half of a day with the short discussion on differentials.

Contents

Suggested Course Outline **i**

1 Mathematical Theory **1**

1.1 What Mathematics Is, and Why It Is Worth Studying . . . 2
1.2 What Calculus Is . . . 5
1.3 How Mathematics is Developed . . . 6
1.4 How to Succeed in Studying Calculus . . . 9

2 Functions **11**

2.1 Sets . . . 12
2.2 Definition of a Function . . . 14
2.3 Inverse Functions . . . 19
2.4 Exponential and Power Functions . . . 24
2.5 Trigonometric Functions . . . 29
2.6 Dimensions and Dimensional Analysis . . . 34
2.7 Rates of Change . . . 36

3 Limits **38**

3.1 Definition of a Limit . . . 39
3.2 Properties of Limits . . . 49
3.3 One-sided Limits . . . 55
3.4 Infinite Limits . . . 60
3.5 Finding Infinite Limits: Order of Magnitude Analysis . . . 65
3.6 Continuity . . . 70
3.7 Sequences . . . 74
3.8 Irrational Numbers . . . 77
3.9 The Bisection Method . . . 83
3.10 First-Order Approximations . . . 85

4 Differentiation **91**

4.1 Definition of the Derivative . . . 92
4.2 Differentials and Infinitesimals . . . 99
4.3 Properties of Differentiation . . . 101
4.4 The Chain Rule . . . 107
4.5 Derivatives of Trigonometric Functions . . . 113

4.6 Derivatives of Inverse Functions . . . 119
4.7 Differentiating Implicitly Defined Functions . . . 122
4.8 Related Rates . . . 126
4.9 Extreme Values . . . 131
4.10 Mean Value Theorem . . . 136
4.11 Concavity and Curve Sketching . . . 140
4.12 Optimization . . . 145
4.13 l'Hôpital's Rule . . . 155
4.14 Newton's Method . . . 158
4.15 Euler's Method . . . 161

5 Integration **165**
5.1 Indefinite Integrals . . . 166
5.2 Area . . . 170
5.3 The Fundamental Theorem of Calculus . . . 177
5.4 Properties of Definite Integrals . . . 180
5.5 Average Values . . . 183
5.6 Integration by Substitution . . . 186
5.7 Integration By Parts . . . 192
5.8 Partial Fractions . . . 198
5.9 Trigonometric Integrals . . . 203
5.10 Trigonometric Substitution . . . 206
5.11 Numerical Integration . . . 209
5.12 Numerical Integration - Composite Rules . . . 219
5.13 Improper Integrals - Infinite Limits of Integration . . . 224
5.14 Improper Integrals - Infinite Integrands . . . 231
5.15 Area Between Curves . . . 234
5.16 Finding Volumes Using Slabs . . . 237
5.17 Finding Volumes Using Shells . . . 241
5.18 Lengths of Plane Curves . . . 244
5.19 Work . . . 248

Chapter 1

Mathematical Theory

For most students calculus marks their entry into the world of higher-level mathematics. In many ways higher-level mathematics is a drastic departure from the lower-level mathematics that most students are familiar with. For this reason, it is worthwhile to understand the differences one is likely to encounter. In this chapter we begin with the notion of mathematics as a language, and why it is important to have a strong grasp of this language. Having considered what mathematics is from a broad perspective, we discuss the essence of the branch of mathematics known as analysis (or calculus). Mathematics, more than a collection of facts, is about structures. A key characteristic of higher-level mathematics is looking at how those structures are developed. As a result, the presentation of material in a calculus course can be quite different from courses in lower-level mathematics. For this reason, this chapter ends with some advice for students on how to succeed in calculus.

Contents

1.1 What Mathematics Is, and Why It Is Worth Studying 2

1.2 What Calculus Is 5

1.3 How Mathematics is Developed 6

1.4 How to Succeed in Studying Calculus 9

1.1 What Mathematics Is, and Why It Is Worth Studying

Simply put, mathematics is a language. Just like any other language, it has its own symbols to represent ideas, and logical rules for combining and using those symbols. In order to create a better picture of this language, it is helpful to compare it to another language, English, the language of this text. These two languages are very different in both the ways they are used, and their form. English is a language used to qualify, or describe qualitatively, some idea, situation, or place. On the other hand, mathematics is a language used to quantify, or describe quantitatively, an idea or phenomenon. When individuals make quantitative descriptions using English, they are expressing mathematical ideas, casted in the form of the English language. Similarly, mathematics relies on English (or another suitable language) in order to communicate ideas more easily between humans, who do not primarily think in terms of mathematics.

Because English (or another suitable language) is used for qualitative descriptions of the world, and mathematics is used to make quantitative descriptions of the world, the two languages serve fundamentally different purposes. These languages are useful in answering different types of questions, and can be very different in the ways in which they answer those questions. Consider some basic descriptions of reality: the weather is hot today; that woman runs very quickly; my hat is the color blue. All of these descriptions are qualitative, giving a basic idea of the qualities of some person, place, or thing. Consider these descriptions in contrast: the weather is 95° Fahrenheit (35° Celsius) today; that woman is running at 20 mph (about 9 m/s); the wavelength of light scattered most by that hat is 460 nm. These are quantitative descriptions, giving the amount or magnitude of some quantity, be it temperature, speed, or wavelength of light. Of these two descriptions, the latter contains much more information; rather than simply saying fast, hot, or blue, it says how fast, how hot, how blue. The downside to the latter description is that it requires more information to be understood. The qualitative description gives a general idea, whereas the quantitative description gives a precise description.

Although the quantitative descriptions are more precise, that does not imply they are more useful (the usefulness of a description depends on its context). Knowing that a hat scatters a wavelength of 460 nm is useless to someone who doesn't have a great deal of familiarity with the visible spectrum of light. Furthermore, in many situations it may be difficult or impossible to give a description with that level of precision. If a person is taking a walk on a hot summer day, what is the probability that they will be carrying a thermometer? If they aren't carrying such a device, saying that the day is hot is probably more than sufficient for casual conversation. Nevertheless, if the individual was interested in calculating the thermal expansion of the sidewalk beneath him or herself, such a qualitative description would not be very helpful - it does not contain the information required to calculate by how much the sidewalk has expanded. Thus, qualitative and quantitative descriptions are both useful, but depending on the situation, one is usually more useful than the other. It is for those situations where a quantitive description is absolutely necessary that mathematics becomes essential.

It is very important to emphasize that a quantitative description is not strictly superior to a qualitative one. As above, the description of color in terms of wavelength of light would not only be confusing in a common-day scenario, it would also be very difficult to make; how many individuals are equipped with the ability to determine a given wavelength of visible light using only the naked eye? Moreover, there are many questions that simply cannot be answered or qualities that cannot be described quantitatively. Emotion is a prime example of this. Think of love and hate for instance. It is not as though there is a natural scale upon which one can assign a quantity for one's love or hate for another individual. Imagine saying to a lover, I love you 450 points of love, or to an enemy, I hate you 7000 points of hate. It simply doesn't make sense. We would be much better served by a qualitative description: I love you more than I ever knew possible, or, I absolutely loathe you, even looking at your face disgusts me[1]. Even more than the extreme cases of love and hate, all emotions - pain, sadness, motivation, etc. - are best described qualitatively. Nevertheless, as humans we often make attempts to artificially quantify emotions, which is best evinced by pain scales used in the medical industry. After coming into the office, a patient is confronted by a chart ranking 1 to 10, perhaps with a scale of faces ranging from happy to very sad underneath. A nurse or physician asks the patient to rate his or her pain on this chart. In doing so the patient is faced with a number of difficult, if not impossible questions to answer: *how does my pain compare to four weeks ago when I last filled out this chart? How did I even feel four weeks ago? What would a 10 on the pain scale feel like? Death?* At very best the patient can make an approximation, and pick a high number for a lot of pain, and a low number for a little bit of pain. There is undoubtedly a lot of error in this procedure, but it is much easier to compare quantitative data than it is to compare qualitative data, so there is a strong desire to quantify the unquantifiable, even at the cost of accuracy. In this situation, a paragraph of testimony from the patient would give a much more accurate description of the pain and problem, with the small caveat that it takes more time to read through.

Hopefully the above discussion makes it more clear that both qualitative and quantitative descriptions of reality are useful, and that they serve different purposes. As such, both English (or your language of choice) and mathematics are useful, but useful for answering different types of questions. With the disclaimer that the following statement is a bit of a generalization, and slightly inaccurate, I feel that we can largely separate the sciences into two domains - social sciences and physical sciences. Social sciences are primarily concerned with the nature of humans, and human interactions. Thus, the social sciences consist of subjects such as anthropology, psychology, political science, etc. On the other hand, physical sciences are concerned primarily with the nature of physical reality, including subjects such as physics, chemistry, astronomy, etc. We can also extend this group to contain engineering, which is essentially the application of discoveries of the physical sciences in order to design and construct tools to fulfill specific purposes. This description of engineering, however, is incomplete, because in many situations it is engineers who contribute to our understanding of the physical world, and thus make contributions to the physical sciences. Although the social sciences use mathematics (and sometimes quite effectively), the objects of their study are often unquantifiable, or do not easily or naturally lend themselves to being quantified. These objects of study include human emotions and motivations. The way in which mathematics is applied to their study is usually statistically, in making predictions of these very complicated issues. On the other hand, the role mathematics plays in the physical sciences is very fundamental - it is the language in which the laws of nature present themselves.[2] Through exploring and solving these mathematical relationships, we are able to uncover some of the mystery of the underlying mechanisms that govern the nature of our physical reality.

Although the above explanations are oversimplifications, they still illustrate that there is a difference between these two realms of scientific inquiry. They also provide us with a step in the direction of answering the question, why study mathematics. For physical scientists, mathematics is the natural language with which to describe the key object of study - physical reality. Just as

[1] It is probably best that one not take these examples too seriously. Trying to learn how to woo a potential lover, or spite an enemy from a mathematics text is definitely not the most efficient use of time, and certainly not the answer to the question of why study mathematics.

[2] This statement is somewhat inaccurate. Scientific theories are models of how reality works, and these models are expressed mathematically, but we can never be certain that a scientific model explains what reality *actually is*. In fact, we can generally be pretty certain that our best scientific theories only provide approximations, but as we will see throughout the study of calculus, approximations can be extremely useful.

it is difficult to write a compelling essay without a strong command of qualitative language, it is difficult to understand the nature of physical reality without a strong command of quantitative language - mathematics. More specifically, many physical phenomena manifest themselves in the language of calculus. Thus, an understanding of calculus is essential to an individual who wants a precise description of physical reality. This level of precision is exactly what is required in the field of engineering in order to manipulate the objects of physical reality.

1.2 What Calculus Is

Calculus (also referred to as analysis) is the mathematics of approximation. The most fundamental tool of calculus, the limit, is about approximation. Suppose we are trying to measure the *exact* height of an object, using a ruler with a scale of centimeters. Using this ruler we are only able to approximate the exact height of the object, because of the size of its scale. Nevertheless, if we were able to find another ruler with a finer scale, say milimeters, we could improve our approximation of the exact height of the object. Suppose that by some means or another (such as finding better and better rulers) we could continually refine our approximation of the exact height of the object. The logical conclusion of this process, the quantity that we approach as we continue to refine our approximations, is called a limit. Thus, for this example the limit of this process is the exact height of the object.

There are fundamentally two approaches two analysis - approximation in theory and in practice. In theory, a limit is an exact value that we are able to approximate as accurately as we desire, simply by following a given process to its logical conclusion. In practice, however, we may not be able to follow the limiting process to its logical conclusion. In the above example, there are limitations to how accurate we can make rulers, so even if the given object has an exact height, we can only approximate what it actually is. The accuracy of the approximation required depends on the situation. The study of issues related to how well we can approximate a given quantity in practice is called numerical analysis. Limitations of computational devices are central to this field.

The two cornerstones of introductory calculus - differentiation and integration - both follow from limiting processes. Through differentiation we approximate (or define) instantaneous rate of change in terms of average rates of change, in the limit as the period of time over which the change occurs approaches 0. The derivative also provides us with a means of finding (or defining) a tangent line - the line which best approximates a curve in the vicinity of the point of tangency. Similarly, integration is a limiting process, through which we find areas (and many other things) through approximation, and following through a limit to make the approximation as accurate as desired. The study of differentiation and integration leads naturally into the study of differential equations - equations that describe how functions change. For this reason calculus is sometimes described as the mathematics of change, because differentiation and integration are principal topics of calculus. If we define calculus solely as the study of differentiation and integration, then perhaps this description is apt. However, the scope of the material taught in most calculus courses extends far beyond these two topics. In this context, the notion of approximation is much more fundamental than the study of change.

Approximation is manifest in numerous other places in analysis. The real numbers are exactly the set of numbers that can be approximated as accurately as desired by rational numbers. In essence, we can only access irrational numbers by approximating them with rational numbers - sequences are the tool used to do so. We are able to extend the notion of sequences to series, through which we can *add* an infinite set of numbers together. We are able to do so by approximating the infinite sum by finite sums, and following through a limiting process. We can even go further to approximate functions with other functions (such as polynomials), and follow a limiting process to make the approximation as accurate as desired. The list of examples goes on, but hopefully the point has been made. Calculus is about approximation. If you want to understand calculus, then you need to understand approximation.

1.3 How Mathematics is Developed

For most students calculus marks their entry into the world of higher-level mathematics. In many ways this is foreign and unknown territory, and at times seems to bear very little resemblance to the lower-level mathematics that we have been taught for so many years. For this reason (among others), learning calculus can seem to be a very daunting task. It is apt to classify calculus as the gateway to higher-level mathematics, at least at the university level, because the study of calculus carries with it the development of calculus. Although different from the historical development of calculus, topics are generally studied in a way that the key concepts are deduced from previous concepts, with an introductory emphasis on proof. Because this aspect of mathematics is very unfamiliar for most students entering a calculus classroom, it makes sense to begin our journey with an introduction to how mathematics is developed, and the field of formal logic and proof.

Mathematical theory, like any other theory, begins with axioms. Axioms are basic facts or truths that are assumed, not proven, to be true. The necessity of axioms follows from the way arguments are constructed. In order to create a logically valid argument, there must be some basic truths that can be called upon to build the argument. From these basic facts a theory is developed. Since axioms are simply assumed to be true, not shown to be true, it is important that they are self-evident, and not subject to contention. This is because the validity of the entire theory is incumbent upon the validity of the axioms. Were the axioms not valid, then the theory that follows from them would be invalid, because it would be based on false assumptions.

The reader is probably already familiar with some of the most basic axioms of mathematics. For our purposes let us consider the field axioms, which define rules for the basic operations of addition and subtraction.

Definition 1.3.1. Field

If $\mathbb{F}$ is a set with two operations, addition and multiplication, satisfying the following the following properties (the field axioms), we call $\mathbb{F}$ a field. For all x, y, z in $\mathbb{F}$:

Property	Addition	Multiplication
Commutativity	$x + y = y + x$	$xy = yx$
Associativity	$(x+y)+z = x+(y+z)$	$(xy)z = x(yz)$
Distributivity	$x(y + z) = xy + xz$	
Identities	$x + 0 = x$	$x \cdot 1 = x$
Inverses	$x + (-x) = 0$	$x \cdot (1/x) = 1$

At the moment lets not focus on what a field is, other than to say that the real numbers are a field, and are assumed to satisfy the above properties.

As was said previously, axioms are merely assumptions, so why should we choose these assumptions to underly mathematical theory? The reason is simple - these properties are consistent with reality. Because we want to use mathematics to describe reality, these are reasonable axioms to choose. There is no reason we couldn't choose another set of axioms, but if they weren't consistent with reality, then the theory built upon them would not be very useful for describing physical situations.

These axioms (among others) provide a basic foundation for mathematical theory. In order to develop the theory, we use deductive logic, to prove that certain results necessarily follow from our chosen axioms, and from other results we have already proven. A mathematical claim consists of a hypothesis and a conclusion. If a claim is true, then anything which satisfies the hypothesis also satisfies the conclusion. Claims can be proven using deductive logic or disproven by giving a counterexample - any object that satisfies the hypotheses of the claim but not the conclusion, which shows the claim must be false[3].

Using the above field axioms we can already begin to prove some basic results of algebra. For instance, we can prove the well-known fact that any real number multiplied by zero is zero. This may seem like an obvious fact, but how do you know it to be true? It is likely that at some point in your education you were told it was true, and over time shown that it seemed to work consistently with other rules of algebra, so you finally accepted it to be true. However, without logical proof, we really cannot be certain that this *self-evident* fact is true. First we begin with a preliminary result we will use to prove our claim of interest.

Claim 1.3.1. *If* $x + y = x$ *then* $y = 0$.

Proof. Suppose that x and y are real numbers. From the field axioms it follows that

$$y = 0 + y = (-x + x) + y = -x + (x + y) = -x + x = 0.$$

$\square$

Claim 1.3.2. $0 \cdot x = 0$.

Proof. Suppose that x is a real number. It follows that

$$0 \cdot x + 0 \cdot x = x \cdot 0 + x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 = 0 \cdot x.$$

From the previous result it follows that $0 \cdot x = 0$. $\square$

The above illustration probably seems very arcane. To begin, the claim is written explicitly, and the word proof is written preceding the deductive argument. Each equality that is written in the argument follows directly from either one of the field axioms or a previous result. Finally, the proof is concluded with a small square, to signify that the argument is complete, and the claim has been proven true (assuming the logic is correct). Once proven, major results are called theorems rather than claims, and minor results are called lemmas[4]. Finally, definitions play an intricate role in mathematics. By defining certain classes of objects, one can prove results about a given class of objects, whereby the simple information that a given object falls under a certain class immediately bestows one a wealth of information about that object. Through the construction of appropriate definitions and proof of theorems, mathematical theory is developed.

One is justified in feeling that this process is rigorously formal. The pertinent question is whether or not such rigor is required. The answer depends on the context. In terms of mathematical theory, the above level of rigor is required. After all, the self-evident facts we often take for granted are not really self-evident; if they were, there would be no need for them to be taught. Through deductive reasoning we are often able to prove claims that our intuition would have told us are untrue, and disprove claims that seem very reasonable, and our intuition tells us should be true. Through this process we continually refine our mathematical intuition, developing mathematical structure simultaneously. This structure provides insight both into the nature of other mathematical structures, and into the structure of physical reality, as scientific theories are built upon mathematical structures.

For individuals who do not wish to pursue the development of mathematics, but merely *use* mathematics, the above level of rigor is probably not required. For this reason, our emphasis will

[3]There are also claims that cannot be proven. This is Gödel's first incompleteness theorem.

[4]A lemma generally does not have much applicability in its own right - the primary purpose of a lemma is to prove a stronger theorem.

not be on constructing proofs, but in understanding their role and purpose, as well as being able to apply them. For students particularly interested in mathematics this will provide the basis for the study of higher-level mathematics, and for those more interested in the sciences it will provide them with the foundation to properly use and interpret mathematical results.

A major consideration to understanding mathematical logic is that logic is directional. What this means is that just because an object satisfies the conclusion of a claim, it does not necessarily satisfy its hypotheses. For instance, knowing that an animal is a Jack Russell Terrier implies that it is a dog. However, knowing that an animal is a dog is not enough to know that it is a Jack Russel Terrier. Some authors would say that if an animal is a Jack Russel Terrier then it is *necessarily* a dog. Nevertheless, knowing that an animal is a dog is not *sufficient* to conclude that it is a Jack Russell Terrier. Although logic is directional, in some instances we encounter conditions that both imply each other. That is, we have a hypothesis which implies the conclusion, and the conclusion also implies the hypothesis. In such a case the two conditions are equivalent, because each one implies the other, so it is impossible to have one condition without the other. An example of two conditions which are equivalent are as follows: if one can legally purchase alcohol in the USA, then one is at least 21 years old. In this case the implication works conversely, in that if one is at least 21 years old then one can legally purchase alcohol in the USA.

1.4 How to Succeed in Studying Calculus

It cannot be overemphasized how different the study of higher-level mathematics is from lower-level mathematics. The primary focus of calculus is concepts, which forces students to write much more than would be expected in a mathematics course. Precision in language is of absolute importance; for instance, students should recognize what the differences between the two phrases *arbitrarily close* and *arbitrary and close* are. Such subtleties of language render rote memorization of definitions and theorems useless, forcing students to be able to reconstruct the definitions and theorems themselves, using their internal understanding of the concepts. Due to the conceptual nature of calculus, memorizing algorithmic processes is not a very effective means of solving problems. The best advice that can be offered to students is to really focus on grasping the key concepts of calculus - limit, derivative, integral - as well as the language associated with them.

In addition to preparing for the conceptual focus of calculus, students should not underestimate the amount of effort required to succeed in calculus. In the USA, about only 60% of students successfully complete calculus. This means that for most students, a considerable amount of effort will be required to succeed. As far as how to focus one's efforts, here are a few essential strategies:

- Attend and actively participate (including asking questions) in every class possible.
- Complete all assignments in the course.
- Seek help immediately if a given concept is unclear or confusing.

The mentality of *figuring things out later* does not work in calculus, where nearly all concepts build on one another, and they build very quickly. A student struggling with current concepts lacks the foundation required to understand subsequent concepts, and as such will only struggle more with them. Students should always seek assistance from an appropriate source - the course text, class notes, additional practice through examples, fellow students, the course instructor, etc. - until they are confident in their understanding of the current lesson/concept before moving onto the next.

Simply attending lecture is not enough - focus and active participation are required. Even losing focus for a few minutes can be enough for a student to miss a key idea or concept which makes the rest of the lecture very difficult to understand. Similarly, it is important to ask questions during lecture to clarify confusing points, so that you are able to follow along as the lecture progresses. Unfortunately, most students are hesitant to ask questions in class out of fear of the perceptions of other students. In truth, most of your classmates probably have very little interest in you, and at the very least it is unlikely you will ever see them again once you complete the course. By asking questions you can greatly improve the value of lecture, and your instructor will appreciate your interest and engagement. Your instructor has a lot more experience with mathematics than you do. Utilize this expertise by actively participating and asking questions in class, asking questions after class, and seeking additional help outside of class in terms of office hours.

Even if you are intrinsically motivated to learn calculus, the constraints of time and life may make it so that your primary goal is simply passing the course with a certain grade. The best strategy for achieving a high grade is understanding the concepts of calculus, and how to apply them to solve problems. These are the skills required to perform well on exams, which are the source of majority of your grade. Even if you achieve perfect scores on all class assignments, if you are unable to perform during an exam you will not receive a very high grade in the class. For this reason you should view assignments as an opportunity to practice and develop problem-solving skills, not as a means of improving your grade. Collaborate with other students and utilize the different perspectives they can offer in order to supplement your own understanding, not as a replacement for it. If you work in an environment where another student does the majority of the work for you, you will not adequately develop problem-solving skills. Even if you are able to increase your assignment scores in this way, you will fail to develop the requisite problem-solving skills to perform well on exams, and as a result receive a much lower grade in the course.

Lastly, as a student you need to be able to internally assess your understanding of the course content. You should be able to predict reasonably well the grade you will receive on an assignment; if your grades are not consistent with your expectations, speak with your instructor. Either there was an error in grading, or you misassessed your understanding of the material. In the first situation the instructor can easily enough fix the grade, and in the second he or she can help guide you to correct your understanding. View your grades on assignments as a means of getting feedback on the work you have done in the course. When assignments with low grades are returned to you, do not see it as a punishment, and definitely do not simply file the assignment away and get ready to do the next one. Use the low grade as feedback that you did not adequately understand the topic at hand, and make all efforts required to remedy that lack of understanding.

Chapter 2

Functions

Functions represent relationships between quantities. Thus, functions are key to the study of the physical sciences, in which we want to represent the relationships between physical quantities, such as length, temperature, etc. In this chapter we will explore functions from the perspective of higher-level mathematics. In this context a function is a very general and abstract entity. Thus, while students may be familiar with some of the material in this chapter, the way it is presented will probably seem unfamilar. This should serve as an introduction to the perspective from which we will view mathematics throughout this text. Since the three central concepts of calculus - limit, derivative, integral - all build on functions, it is important for students to have a strong grasp of the function concept before moving along further in study.

Contents

2.1 Sets 12
2.2 Definition of a Function 14
2.3 Inverse Functions 19
2.4 Exponential and Power Functions 24
2.5 Trigonometric Functions 29
2.6 Dimensions and Dimensional Analysis 34
2.7 Rates of Change 36

2.1 Sets

Sets are ubiquitous in higher-level mathematics. For our purposes, the place sets will play the largest role is in that functions are defined on sets, and the properties of a given function depend on the sets over which it is defined.

Definition 2.1.1 (Set)**.** A set is a collection of distinct objects, called elements or members.

Conventionally, sets are denoted by capital letters, such as X, Y, etc. In order to specify or define a set, we use curly braces $\{\}$, where the contents of the curly braces describe which elements belong to the set. One way to do this is to simply list all of the members explicitly. For instance,

$$M = \{2, 4, 6\}$$

defines a set M containing the elements 2, 4, and 6. Thus, the set M contains three elements. If we were to write

$$N = \{2, 4, 2, 6, 4\},$$

it would be equivalent to the definition above, because each element of a set is only counted once (elements are defined as being distinct). Thus,

$$N = M,$$

which means that the elements of M are exactly the elements of N. If we wanted to define a set that contained all even positive integers, we could specify it as

$$P = \{2, 4, 6, \ldots\}$$

where the '. . . ' indicate that the pattern continues on forever. We could also simply write

$$P = \{\text{even positive integers}\}.$$

Finally, if we want to express that an object is a member of some set, we use the notation $2 \in P$, to say that *2 belongs to P* or *2 is a member of P*.

There is one more way to define sets which can be particularly useful. Let us define the set Q by

$$Q = \{z \in P \mid z < 7\},$$

which literally reads *Q is the set of elements z in P such that z is less than 7.* Thus,

$$Q = \{2, 4, 6\} = M = N.$$

The strength of this notation lies that we can specify rather complicated sets easily, provided that we have a reference set. It turns out that there are a number of sets that are used so commonly there are special symbols to represent them.

1. The Natural Numbers: $\mathbb{N} = \{1, 2, 3, \ldots\}$
2. The Integers: $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$
3. The Rational Numbers: $\mathbb{Q} = \{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\}$
4. The Real Numbers: $\mathbb{R}$

5. The Complex Numbers: $\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}\}$

At the moment we refrain from giving a more explicit definition of the real numbers, because it is not an issue we are prepared to deal with yet. We could say that the set of real numbers is the set of rational numbers plus the set of irrational numbers, which contains esoteric examples such as π, e, and $\sqrt{2}$, but we still don't really know how to deal with irrational numbers yet. After we have developed more mathematical theory we will be able to view the set of real numbers as the set of numbers we can approximate with rational numbers. This is an issue we will touch on later when we look at sequences of numbers.

Lastly, when we think of sets there is a notion of containment. That is, a set can be a *subset* of another set, if all of the members of the later set are contained in the former set.

Definition 2.1.2 (Subset)**.** A set M is a *subset* of a set N, written $M \subset N$, if for each $m \in M$, we also have $m \in N$. In other words, each element of M is also an element of N.

A simple chain of subsets is given by the commonly used sets of numbers above, namely

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

In the context of the real numbers, we will often consider special subsets, called intervals. An interval is defined by two numbers, a left endpoint, and a right endpoint. An interval is any set that contains all numbers greater than its left endpoint, and less than its right endpoint; ie. intervals contain the points between their endpoints. These points are called the *interior points* of an interval. An interval may or may not contain its endpoints. Intervals are denoted as follows:

$$(a, b), \quad [a, b], \quad [a, b), \quad \text{and} \quad (a, b].$$

A parenthesis (or) denotes that the endpoint is not included in the interval, whereas a bracket [or] denotes that the endpoint is included in the interval. Thus, (a, b) includes neither endpoint, $[a, b]$ includes both endpoints, $[a, b)$ includes the left but not the right endpoint, and $(a, b]$ includes the right but not the left endpoint. An interval of the form (a, b) is called an *open interval* and an interval of the form $[a, b]$ is called a *closed interval*. These names stem from the topological properties of intervals over the real numbers, which is not a topic we will focus on here.

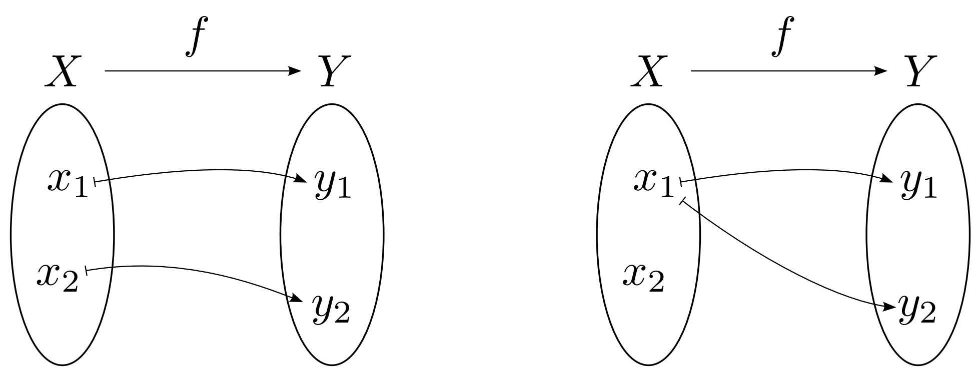
2.2 Definition of a Function

In mathematics, we are interested in relations between different quantities, both in order to represent unknown quantities in terms of known quantities, and to understand the behavior of one quantity with respect to another. There are many situations in which it may be very difficult or impossible to measure the value of a given quantity directly. For instance, it is a simple task to measure the height of an object, but it is not as simple to measure its gravitational potential energy. Nevertheless, if we can find a relation between height and gravitational potential energy, then finding an object's gravitational potential energy becomes as simple as measuring its height.

Relations are also useful for understanding how quantities depend on each other. Particularly, in calculus we will be interested in how one quantity changes with respect to another. For instance, we may know the position of an object with respect to time, and want to find its velocity and acceleration. Conversely, we may know acceleration or velocity and be interested in finding the other two quantities. Thus, we will be interested in relating a given quantity and its rate of change with respect to time (or another variable of interest, such as height). The specific type of relation we are interested in for both of these scenarios is called a function.

Definition 2.2.1 (Function)**.** A function f is a relation between sets X and Y, where each element of X is associated with a unique element $y = f(x) \in Y$. The set X is called the *domain* of f, and the set Y is called the *codomain* of f. This relation is denoted by $f : X \to Y$, $x \mapsto y = f(x)$.

By this definition, a function is any description that relates two sets of values, where each input is associated with only one output. Such a description can be algebraic, graphical, verbal, etc. The domain X represents the set of inputs, and the codomain Y represents the set of possible outputs. In addition to being considered an association between two sets, functions can also be thought of as *maps* between sets, where each member of the domain is *mapped* to a member of the codomain. This language is very natural when combined with the pictorial representation of a function in figure 2.1.



(a) Diagram of a function.

(b) Not a function.

Figure 2.1: The diagrammatic representation of a map makes it easy to tell whether or not it is a function.

Here we have the representation of two sets, and a sample of some of their elements (or members). Here we see that f literally *maps* one element to another, where the arrow shows to which element of the codomain Y each element of the domain X is mapped to. In figure 2.1(a) the map is of x_1 to y_1 and x_2 to y_2, so the map is of a function. However, the mapping in figure 2.1(b) maps x_1 to both y_1 and y_2, so there is a single input with multiple outputs - not a function.

Strictly speaking, a function should have an output for each member of the domain, but we will also consider *partial functions* where some members of the domain have no output. An example of such is the function

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = 1/x,$$

which is not defined for $x = 0$, because division by 0 is not permitted. Thus, strictly speaking we should write

$$f : \{x \in \mathbb{R} \mid x \neq 0\} \to \mathbb{R}$$

to represent the function f, but rather than encumber our notation we will simply work with partial functions, and henceforth refer to them as just functions.

The domain of a function is an extremely important part of a function's definition. Consider the function

$$f : \mathbb{N} \to \mathbb{R}, \quad f(x) = 1/x,$$

which is defined by the same rule as the function above, but on a domain of natural numbers rather than real numbers. The behavior of this new function is very different - it never exceeds a maximum value of $1/2$, and moreover we don't have to worry about division by 0 on this new domain.

The codomain Y represents the set of possible outputs of a function f. This means there is at least some dependence of the codomain on the choice of domain for f, as the set of possible outputs large enough for one domain may not be large enough for another domain. Since Y is the set of possible outputs, there is no requirement that each member (or element) of Y have a member of X associated with it (in general Y may be much larger than the set of actual outputs). For this reason one can always increase the size of the codomain, and arrive at a new valid codomain. Thus, there is no largest codomain. Nevertheless, there is a smallest codomain, in which every member of the codomain has a member of the domain associated with it. This set is called the *range* or *image* of f, denoted by $f(X)$, which is the set of all values that f associates a member of X with. Written explicitly,

$$f(X) = \{y \in Y \mid y = f(x) \text{ for some } x \in X\}.$$

In this way we are actually viewing the function f as a map between sets rather than just points. Finally, it is worth noting that the notation f is used to represent the function or map f, whereas $f(x)$ is used to represent the value of the function when it is evaluated at the element x.

We will usually consider functions defined on sets of numbers, such as the real numbers $\mathbb{R}$, rational numbers $\mathbb{Q}$, integers $\mathbb{Z}$, and natural numbers $\mathbb{N}$. Nevertheless, there is no reason that we cannot choose much more exotic sets, such as shapes, colors, or names. We might have a function that maps type of polygon (input) to the number of sides it has (output),

$$f : \{\text{polygons}\} \to \mathbb{N}.$$

Here the domain X consists of members such as square and triangle, and the range is natural numbers greater than or equal to 3. Thus,

$$f(\{\text{polygons}\}) = \{x \in \mathbb{N} | x \geq 3\}.$$

If we were to turn this mapping around, and instead think of a mapping from number of sides to type of polygon, it would no longer be a function, as a single input would have multiple outputs (4 sides would map to square, rhombus, etc.).

We can also think of functions where neither the set of inputs or outputs consists of numbers. For instance, we could have a function which maps square to circle, circle to triangle, and triangle to square. For this function, the domain and range are the same, the set of the shapes square, circle, and triangle. This a function, because each input is mapped to only a single output.

The sky is the limit in defining functions, but when we define a function we must be careful that the function is actually well-defined (that it is defined in a mathematically rigorous and sensible way, and that it does in fact only have a single output for each input). Let us attempt to define a function that maps the natural numbers to the real numbers, so that for each number $n \in \mathbb{N}$, the output is the smallest real number that is larger that n. This function is not well-defined, because there is no smallest real number larger than n. For any number larger than n, say y, we can always find a number between n and y[1]. The fact that this *function* is not well-defined means it's not really a function at all.

Example 1. Consider the following descriptions of possible functions:

1. An object's gravitational potential energy with respect to height above the earth.

2. $f(x) = x^2$

3. $x^2 + y^2 = 1$, where x is the input and y is the output.

4. A raindrop counter. At the beginning of a storm, the raindrop counter starts with a value of zero. Each time a raindrop hits the roof of a specific building, the counter increases by one. The raindrop counter tells how many drops of rain have hit the roof after a given amount of time.

5. Suppose we know the position of a jogger with respect to time. We devise an average velocity meter, which for any instant in time gives the runner's average velocity.

6. Suppose we have a polynomial function $f(t)$. We devise an area meter, which for any given x gives us the area between $f(t)$ and the t-axis, over the interval $[0, x]$.

Which of the above descriptions could represent functions if given an appropriate domain and codomain? If possible, give an example of such a domain and codomain.

Solution The major questions to ask about the above relations is whether or not they are well-defined, and if each input is assigned to only one output over an appropriately chosen domain.

1. For each height we have a unique gravitational potential energy, so this description is well-defined and each input has at most one output. A possible domain and codomain would be the nonnegative real numbers, as we are considering heights above the earth, and do not have negative potential energies associated with such heights.

2. This relation is well-defined and each input is assigned to only one output. The domain could be chosen as the real numbers $\mathbb{R}$, which would have a corresponding range of nonnegative real numbers.

3. This relation is well-defined - the equation for a circle of radius 1. If we assign a domain of real numbers then it is not a function, because each x value has two possible y values, one above and the other below the x-axis. However, if we restrict the domain to the set

$$\{-1, 1\},$$

[1]This fact will be extremely important in calculus, when we want to make approximations; namely, for any real number we can always choose another real number that is as close to it as we want.

then we do have a function, as each of these inputs are associated with only a single output - zero. Thus, the range here is

$$\{0\}.$$

4. At a given instant in time there is a unique value for how many raindrops have already fallen, so this description can belong to a function. Since the input is time, nonnegative real numbers is a reasonable domain (avoiding negative time), and the natural numbers could be chosen for the codomain (saying that there is no such thing as a fraction of a raindrop).

5. At each position there is a unique average velocity, so this is the description of a function. Since we are considering time for the domain, we could use the nonnegative real numbers as the domain, with a corresponding codomain of real numbers.

6. Although esoteric, this does describe a function. The domain is real numbers, and the codomain should be nonnegative reals, given that area should not be negative. Later on we will define a notion of signed area (permitting negative output values), and such a function will be called the definite integral of $f(t)$.

Most of the above functions are formulated in an abstract sense, so we do not actually know their values, only that they would define valid functions. When we have a graphical representation of a function, say in the Cartesian coordinate plane, then we can easily determine whether or not the graph could represent a function - we use the vertical line test. For the graph of a function in the coordinate plane, any vertical line can only cross the function at at most one point. If it were to intersect the graph at multiple points, it would mean that a single input has multiple outputs, so the graph would not be of a function. For example, think of the graph of a circle. Since we can find vertical lines that cross the circle in two places, we know that it is not the graph of a function.

If we have a number of functions at our disposal, we can combine them to create new functions. For functions f and g and constants α and β we can create many new functions, such as

$$\alpha f,\ f+g,\ \alpha f+\beta g,\ fg,\ \frac{f}{g},\ \text{etc.}$$

The above functions are constructed using arithmetic operations: $+, -, \times, \div$. Note that f/g is only defined for all values x where $g(x) \neq 0$, as division by 0 is not permitted. There is one additional process through which we can combine functions, which is known as *composition*, written $f \circ g$. In the composition $f \circ g$ the output of the function g is used as the input of the function f, so $f \circ g(x) = f(g(x))$. If $f : Y \to Z$, $g : X \to Y$, then $f \circ g : X \to Z$ (see figure 2.2.) Graphically, we can think of composition as allowing us to create maps between *distant* sets, using multiple maps.

Properties of functions related to the way they change do not follow from the formal definition of a function we have presented above. The notion of a changing function only makes sense for certain domains. Namely, we need to have a notion of distance, so that we can say as the input changed by some distance, the output correspondingly changed by some distance. The way in which we determine distance is called a *metric*, and a set combined with a metric is called a *metric space*. An example of a metric space would be any of $\mathbb{R}, \mathbb{Q}, \mathbb{Z}, \mathbb{N}$ combined with a metric

$$d(x, y) = |x - y|,$$

which means that the distance between two points is given by the absolute value or magnitude of their difference. The notion of distance does not make sense for all domains. Consider the function defined on the domain of polygons. Even if we were able to define a notion of distance for this domain, it would be extremely artificial, and thus not very useful.

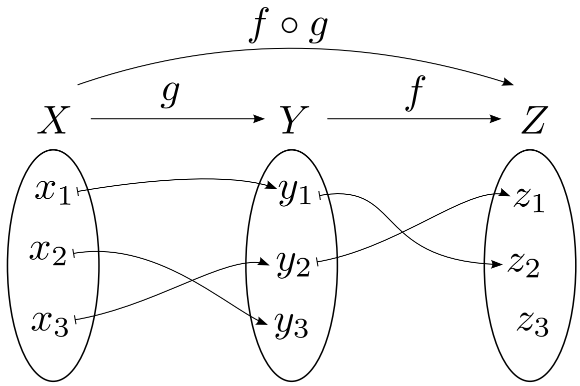


Figure 2.2: The composition $f \circ g$ of f and g maps X to Z.

As stated previously, calculus (or analysis) is the mathematics of approximation. Approximation requires a notion of distance, so that we can make sense of the idea of how close our approximation is. Thus, as we study calculus we will be interested in looking at functions defined on metric spaces, so that the notion of approximation makes sense. We will focus primarily on the real numbers, with the metric $d(x, y) = |x - y|$. From the notion of approximation we will able to establish properties of changing functions.

2.3 Inverse Functions

The inverse of a function f is another function f^{-1}, which essentially undoes the action of f. We can think of using inverse functions in order to solve algebraic equations, where the goal is to isolate a given variable. If one knows that

$$x^2 = 2,$$

then in order to solve for x, one wants to apply the inverse operation of squaring, namely the square root. This operation undoes the action of the square, isolating x, so we see that

$$x = \sqrt{2}$$

is a solution to the above equation. In terms of functions, we know that for each input to a function, there is exactly one output. Thus,

$$x = y \quad \Longrightarrow \quad f(x) = f(y),$$

which says that for any function f, if $x = y$ then $f(x) = f(y)$. In the above example, we know that $x^2 = 2$, so $f(x^2) = f(2)$, for any function f defined at 2. In this case we can choose $f(x) = \sqrt{x}$, so that

$$f(x^2) = f(2) \quad \Longrightarrow \quad \sqrt{x^2} = \sqrt{2},$$

which provides a solution to the above equation. We find the other solution to the above equation using the function

$$g(x) = -\sqrt{x}.$$

Now suppose that we have an arbitrary function f. There are two natural questions to ask: does there exist an inverse function f^{-1}, and if so, how do we find it? In order to guide our discussion we will first need to define what we mean by inverse function.

Definition 2.3.1 (Inverse Function)**.** A function f^{-1} is said to be the *inverse* of a function f if:

- $f^{-1}(f(x)) = x$ for all x in the domain of f
- $f(f^{-1}(x)) = x$ for all x in the domain of f^{-1}

By this definition we don't just require that f^{-1} undoes the action of f, but also that f undoes the action of f^{-1}. Starting from the definition of a function $f : X \to Y$, for each element $x \in X$, f maps x to an element $y \in Y$. In order to undo the action of f, we want to take these elements in Y and map them back to the original elements in X. Thus, we want

$$f^{-1} : Y \to X, \quad f^{-1}(y) = x,$$

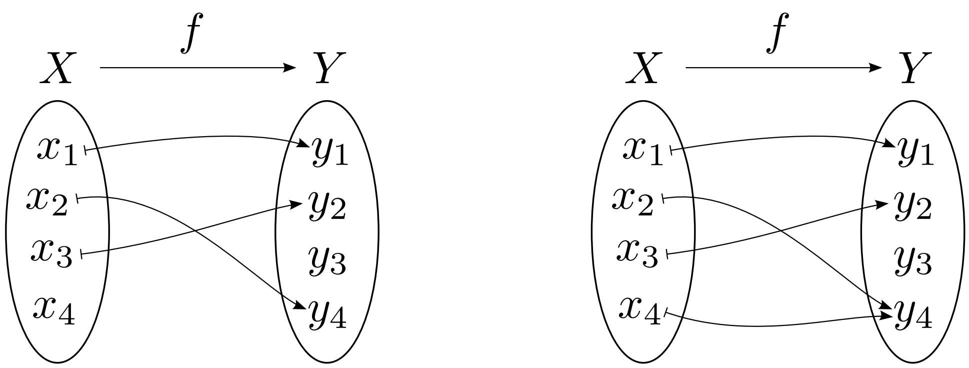
where we initially had $f(x) = y$. We immediately run into a problem; what if multiple elements of X are mapped to the same element of Y? For instance, we could have

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = x^2,$$

so that both -1 and 1 are mapped to 1. When we try to construct an inverse, we will have 1 mapping back to both 1 and -1. The problem is, now we have two outputs associated with a single input, so our inverse map is no longer a function. Thus, if our original function f maps multiple inputs to the same output, then the inverse map will not be a function, so we say that as function, f is noninvertible. In order to have an invertible function, we require the condition that each member of the range of f is mapped to by at most one element of the domain. Such a map is called an *injection*, or *one-to-one function*.

Definition 2.3.2 (One-to-One Function)**.** A function $f : X \to Y$ is called *one-to-one* if each element $y \in Y$ is associated with at most one element $x \in X$. In other words,

$$f(x) = f(y) \implies x = y.$$



(a) A one-to-one function.

(b) A non-one-to-one function.

Figure 2.3: A one-to-one function maps at most one element to each element of the codomain.

As stated above, we need to have a one-to-one function f in order for its inverse f^{-1} to exist. The members of the codomain Y for which we know f^{-1} are exactly those that are mapped to by some member of X, because it is this member of X that defines the inverse map. For example, consider

$$f : \{x \in \mathbb{R} \mid x \geq 0\} \to \mathbb{R}, f(x) = x^2.$$

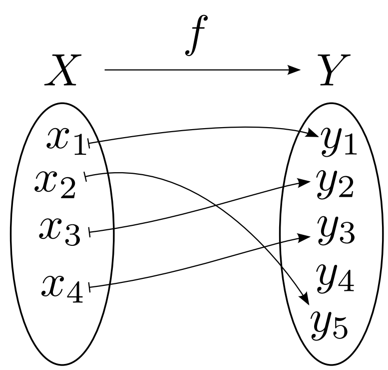
Here $\{x \in \mathbb{R} \mid x \geq 0\}$ is the set of nonnegative real numbers. We do not want to include the negative real numbers in the domain because each of them is mapped to the same number as a positive real number (both x and $-x$ are mapped to x^2). When we remove the negative reals, the square root function provides us with a map that undoes the action of f. However, this map is still not an inverse function, because it does not map the entire codomain $\mathbb{R}$ back to the real numbers (so f will not undo the action of f^{-1} for all x in the domain of f^{-1}); the square root is only defined for nonnegative reals - which are exactly the range $f(\{x \in \mathbb{R} | x \geq 0\})$. Thus, in order to find the inverse map we need to look at the function

$$f : \{x \in \mathbb{R} \mid x \geq 0\} \to \{x \in \mathbb{R} \mid x \geq 0\}, \quad f(x) = x^2$$

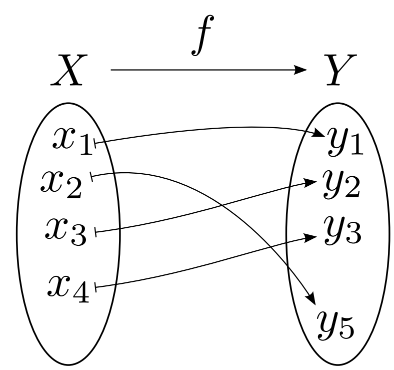
which is simply the above function with the restriction of the codomain to the range of f. When we make this restriction we remove all of the extra elements from the codomain that we didn't know how to map back to domain of f (which is the range of f^{-1}). We could try to redefine f^{-1} to map these extra elements to some other members of the domain of f, but then we would no longer have the desired property that f undoes the action of f^{-1} (because f^{-1} would map multiple elements to the same place, and thus not be one-to-one). A function with the property we require here is called a *surjection*, or *onto function*, which means that for every member of the codomain there is an element in the domain mapped to it by f.

Definition 2.3.3 (Onto Function)**.** A function $f : X \to Y$ is called *onto* if each element $y \in Y$ is associated with at least one element $x \in X$. In other words, for all $y \in Y$, there is an $x \in X$ so that $f(x) = y$.

It is very easy to make any function f into an onto function - we simply choose the codomain to be the range of f. Consider the function in figure 2.4(a). We can make this function onto simply by removing the extra element y_4 from the codomain, arriving at the onto function in figure 2.4(b). In doing so we are redefining the codomain to be exactly the range $f(X)$, just as we've done previously.



(a) A non-onto function.



(b) An onto function.

Figure 2.4: Making a function onto.

For a one-to-one function each element of the codomain is associated with at most one element of the domain, and for an onto function each element of the codomain is associated with at least one element of the domain. Thus, for a one-to-one and onto function, each element of the co-domain is associated with exactly one element of the domain, and conversely. Such functions are called *bijections*.

Definition 2.3.4 (Bijective Function)**.** A function $f : X \to Y$ is called a *bijection* if each element $y \in Y$ is associated with exactly one element $x \in X$.

A bijective function is exactly what we need for invertibility. If $f : X \to Y$ is a bijection, there is a unique correspondence between elements of X and Y, so the choice for an inverse is very natural - simply map the members of each of these pairs to one another. Bijective functions are invertible, and in fact, are the only functions that are invertible. Furthermore, because of the unique correspondence between members of X and Y, the choice for f^{-1} is unique.

Theorem 2.3.1 (Invertible Function)**.** *A function $f : X \to Y$ is invertible if and only if f is a bijection. If f is invertible, its inverse f^{-1} is unique.*

Here we have a statement which includes *if and only if* sometimes written *iff* or $\iff$. This means being a bijection implies a function is invertible, and being invertible implies a function is a bijection; ie. being a bijection and being invertible are equivalent. With the above result we can determine exactly which functions are invertible, and also using the unique correspondence between the domain and codomain (which is also the range in this case), we can determine what the inverse function is (see figure 2.5).

Since it is usually not feasible to evaluate a function for every one of its inputs, it is helpful to have an alternative means for determining whether or not a function is invertible. Particularly, this is of interest for functions $f : \mathbb{R} \to \mathbb{R}$, because we would need to evaluate the function f at an infinite number of inputs to determine if it is invertible. In order to get around this problem, we need to find some classification for functions, and then show that all functions that fall under that class are invertible. The functions we will be interested in are *monotone functions*.

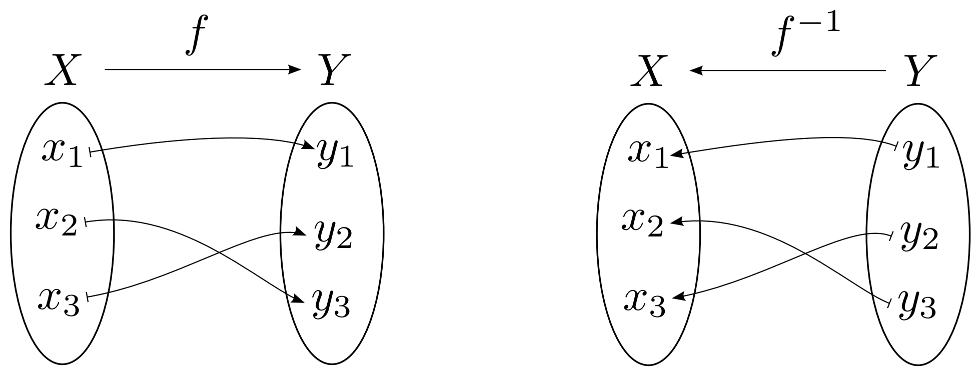


Figure 2.5: A bijective function f and its inverse.

Definition 2.3.5 (Monotone Function)**.** A *monotone function* is a function that is either increasing or decreasing.

- A function is *increasing* if $x < y \implies f(x) \leq f(y)$.
- A function is *decreasing* if $x < y \implies f(x) \geq f(y)$.

The above definitions may not be completely intuitive at first. An increasing function is one where when the input to the function increases, the output does not decrease. This means that as the input gets larger, the output may increase or stay constant, but it can never decrease. Similarly, a decreasing function is a function that never increases. As inputs get larger the function either gets smaller or stays constant. This property is not sufficient to have invertibility, because if the function stays constant for a period of time, we will have a number of inputs mapped to the same output. Instead, what we really need is a function that is never constant and never overlaps itself. One class of such functions are *strictly monotone* functions.

Definition 2.3.6 (Strictly Monotone Function)**.** A *strictly monotone function* is a function that is either strictly increasing or strictly decreasing.

- A function is *strictly increasing* if $x < y \implies f(x) < f(y)$.
- A function is *strictly decreasing* if $x < y \implies f(x) > f(y)$.

The only difference in the above definitions is that $\leq$ becomes $<$ and $\geq$ becomes $>$. Thus, the difference between a monotone and strictly monotone function is that a strictly monotone function can never be constant, although a monotone function can. A strictly monontone function must always get larger, or always get smaller. From this it is not difficult to show that a strictly monotone functions is one-to-one, and simply by restricting the co-domain to the range of the function, we can make it onto.

Theorem 2.3.2 (Invertibility of Strictly Monotone Functions)**.** *Let $f : X \to f(X), X \subset \mathbb{R}$. If f is strictly monotone, then f is invertible.*

Proof. In order to show that f is invertible, we must show that it is both one-to-one and onto; ie. f is a bijection. Because f is a mapping to its range, it follows automatically that f is onto. Now consider the case where f is strictly increasing. Then we have

$$x < y \implies f(x) < f(y) \quad \text{for } x, y \in X.$$

Consider two elements $x, y \in X, x \neq y$. Since $x \neq y$, either $x < y$ or $x > y$. Analogous arguments hold in both cases, so assume $x < y$. Since f is strictly increasing, it follows that $f(x) < f(y)$, so $f(x) \neq f(y)$. Thus, distinct inputs produce distinct outputs, so f is one-to-one. Therefore f is a bijection, so by a previous result, f is invertible. $\square$

2.4 Exponential and Power Functions

When we are working with natural numbers, we can think of exponentiation as short-hand notation for multiplication. If we let n be a nonnegative integer, we can intuitively think of a^n as a multiplied by itself n times. For instance,

$$a^3 = a \cdot a \cdot a.$$

This is similar to how multiplication of natural numbers can be viewed as a short-hand notation for addition. For instance,

$$3 \cdot 2 = 3 + 3 = 2 + 2 + 2.$$

Nevertheless, when we extend multiplication to the integers or rational numbers, multiplication stops resembling addition, becoming an operation with its own unique properties. Just imagine trying to represent

$$4 \cdot \frac{1}{2}$$

in terms of addition - it doesn't make sense. Analogously, we can begin with this intuitive notion of exponentiation as stemming from multiplication, but we will see that in extending exponentiation to larger sets of numbers, it will become a distinct entity from multiplication. In order to begin, we will slightly modify our intuitive notion of exponentiation, and work on the nonnegative integers rather than natural numbers (in order to include 0). We can think of a^n as 1 multiplied by a, n times. When we start with this definition we have that

$$a^0 = 1 \quad \text{and} \quad a^1 = 1 \cdot a = a.$$

From this intuitive view it is not difficult to deduce that

$$a^n \cdot a^m = a^{n+m}.$$

For instance, if we let $n = 2$ and $m = 3$, then we are multiplying $a \cdot a$ and $a \cdot a \cdot a$, so the result is five terms of a multiplied together, or $a^5 = a^{2+3}$. A similar argument works for any other values of n or m. Now what if we have something of the form a^m/a^n? Letting $n = 2$ and $m = 3$ again, we have

$$\frac{a^m}{a^n} = \frac{a^3}{a^2} = \frac{a \cdot a \cdot a}{a \cdot a} = a = a^{3-2}.$$

Once again we can make the same argument for any values of n and m. Thus,

$$\frac{a^m}{a^n} = a^{m-n},$$

where we are assuming $m > n$. According to the above rules, if we multiply two base a exponential functions together, the exponents add, and if we divide, the exponent in the denominator is subtracted. If we write the above expression slightly differently, we see

$$\frac{a^m}{a^n} = a^m \cdot \frac{1}{a^n} = a^{m-n},$$

where we now have written the expression as a product rather than a quotient. When we multiply any base a exponential by $1/a^n$, $1/a^n$ behaves similar to a base a exponential (even though it isn't one) except that n is subtracted from the exponent rather than added to it. From another viewpoint, when we multiply by $1/a^n$, we *add* $-n$ to the exponent. In accordance with our rule for multiplying exponentials, it is natural to *define* negative exponents, so that

$$a^{-n} = \frac{1}{a^n}.$$

This definition of negative exponents is natural because it is completely consistent with our previous rule for multiplying positive exponentials (if we multiply by a negative exponent, we add the negative exponent). Now we have the rule

$$a^n \cdot a^m = a^{n+m} \quad \text{for } n, m \in \mathbb{Z}.$$

We have gained more flexibility by extending exponentiation to all integers, but at the cost of our initial intuitive understanding; the notion of multiplying a number by itself a negative number of times is purely nonsense.

It's possible to extend the notion of exponentiation even further to rational numbers, if we notice that

$$(a^n)^m = a^{nm} \quad \text{for } n, m \in \mathbb{Z}.$$

Just for another simple illustration, think of $(a^2)^3$. Here we have

$$(a^2)^3 = (a \cdot a)^3 = (a \cdot a) \cdot (a \cdot a) \cdot (a \cdot a) = a^6 = a^{2 \cdot 3}.$$

We can extend exponentiation to rational numbers by requiring that rational exponents abide by this same rule. In other words, we should have

$$(a^{1/n})^n = a^{n/n} = a \quad \text{and} \quad (a^n)^{1/n} = a^{n/n} = a.$$

In this way we can *define* $a^{1/n}$ as the number such that when it is raised to the n^{th} power, the result is a. This extended definition of exponentiation is also consistent with our previous rules (although this may be a little difficult to verify).

Since we've already made it to the rational numbers, a reasonable goal would be to try and achieve real exponents. This is indeed possible, but we will need the aid of limits in order to do so, so we will return to this issue at a later time (just as we avoided defining the real numbers earlier). Nevertheless, for the moment we will take for granted that real exponentials work (as do complex exponentials), and that we have the following rules of exponentiation.

Theorem 2.4.1 (Rules of Exponentiation)**.** *Let $a, b, x, y \in \mathbb{R}$. It follows that:*

1. $a^x \cdot a^y = a^{x+y}$
2. $(a^x)^y = a^{xy}$
3. $a^{-x} = \frac{1}{a^x}$
4. $(ab)^x = a^x b^x$
5. $a^1 = a$
6. $a^0 = 1$

Even though we have defined rational and real exponentials, in most cases we do not actually know how to evaluate them yet. For instance, consider

$$2^{1/2} = \sqrt{2}.$$

By definition, $\sqrt{2}$ is the number that when squared is 2. However, right now we have no idea what this number actually is. We will see later on that the answer to this question lies in approximation, so we will not be able to answer this question without the aid of calculus.

Setting the above difficulties aside we can use exponentiation to define polynomial and exponential functions, which will be some of the most basic functions available to us.

Definition 2.4.1 (Power Function)**.** A *power function* is a function of the form $f(x) = x^a$, where $a \in \mathbb{R}$.

Thus, a power function is a function where the *base* of the exponential varies as an input. Very basic examples of power functions include $f(x) = x$ and $f(x) = x^2$. Note that $f(x) = x$ maps the real numbers to the real numbers, where $f(x) = x^2$ maps the real numbers to the nonnegative real numbers. Some power functions are only defined as maps on the real numbers for a domain of nonnegative real numbers, such as

$$f(x) = x^{1/2} = \sqrt{x}.$$

For this function a negative input is not defined, because complex numbers are required to make sense of the square root of a negative number.

Using power functions as our most basic building blocks, we arrive at polynomial functions. One of the most basic ways in which we combine functions is in a *linear combination*.

Definition 2.4.2 (Linear Combination)**.** A *linear combination* of the functions $\{f_1, f_2, \ldots, f_n\}$ is a function

$$f(x) = c_1 f_1(x) + c_2 f_2(x) + \ldots + c_n f_n(x),$$

where $c_1, c_2, \ldots, c_n \in \mathbb{R}$.

Definition 2.4.3 (Polynomial)**.** A *polynomial* is a function that can be written as a linear combination of power functions raised to nonnegative integer exponents. Thus, polynomials take on the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n,$$

where $a_i \in \mathbb{R}$.

Polynomials provide us with a large class of simple functions to work with. In constrast to an arbitrary function, a polynomial is very well-behaved, and as a result has a number of useful properties. As we delve further into the study of calculus we will use polynomials to approximate other functions, and eventually be able to represent much more complicated functions as infinite sums of power functions. Such representations are called *power series*.

In contrast to power functions, *exponential functions* are functions where the exponent varies as an input.

Definition 2.4.4 (Exponential Function)**.** An *exponential function* is a function

$$f : \mathbb{R} \to \mathbb{R}^+ \text{ (positive real numbers)}, \quad f(x) = a^x,\ a \in \{x \in \mathbb{R} \mid x > 0, x \neq 1\}.$$

Note that when we are talking about exponential functions we are only interested in exponentials with base $a > 0$. We are not interested in $a = 1$, because it is simply a constant function. Since this constant function behaves differently from the rest of the exponential functions we will deal with, we simply exclude it from the list of exponential functions. All of the exponential functions have a domain of $\mathbb{R}$ and a range of $\mathbb{R}^+$ (positive real numbers). This means that the output of an exponential function is always positive. In fact, exponential functions are strictly increasing, which means for each exponential there is a corresponding inverse function (see theorem 2.3.2). These inverse functions are called *logarithms*.

Definition 2.4.5 (Logarithmic Function)**.** A *logarithmic function* is a function

$$f : \mathbb{R}^+ \to \mathbb{R}, \quad f(x) = \log_a(x),\ a \in \{x \in \mathbb{R} \mid x > 0, x \neq 1\},$$

where $\log_a(x)$ is the inverse function of a^x.

Since exponentials and logarithms are inverses, we have

$$\log_a(a^x) = x = a^{\log_a(x)}$$

for all $a > 0, a \neq 1$. By virtue of this inverse relationship, logarithms inherit a number of useful properties from exponentials. For instance, by using the relationship

$$a^x a^y = a^{x+y}$$

we can deduce

$$\log_a(a^x a^y) = \log_a(a^{x+y}) = x + y = \log_a(a^x) + \log_a(a^y).$$

Above we simply use the property that $\log_a(a^x) = x$, in order to move from step 2 to 3, and step 3 to 4.

We can also deduce a rule for the logarithm of a product, noting that the exponential function a^x has a range of the entire positive real numbers $\mathbb{R}^+$. In other words, for any $x \in \mathbb{R}^+$ there is some $b \in \mathbb{R}$ so that

$$x = a^b.$$

Similarly, for any positive y, we can write $y = a^c$, for some c. As a result,

$$\log_a(xy) = \log_a(a^b a^c) = \log_a(a^b) + \log_a(a^c) = \log_a(x) + \log_a(y).$$

Thus, the logarithm of a product of two numbers is the sum of the logarithms. The full table of properties of logarithms follows.

Theorem 2.4.2 (Rules of Logarithms)**.** *Let $a, x, y \in \mathbb{R}^+$. It follows that:*

1. $\log_a(xy) = \log_a(x) + \log_a(y)$
2. $\log_a(x^y) = y \log_a(x)$
3. $\log_a(1) = 0$
4. $\log_a(\frac{x}{y}) = \log_a(x) - \log_a(y)$
5. $\log_a(\frac{1}{y}) = -\log_a(y)$

Property 2 above follows if we write $m = \log_a(x)$, where $m \in \mathbb{R}$. Using the inverse relationship of exponentials and logarithms we also know $x = a^m$. Thus,

$$x^y = (a^m)^y = a^{my}.$$

Taking the base a logarithm of each side of the above equation,

$$\log_a(x^y) = \log_a(a^{my}) = my = ym = y \log_a(x).$$

The third property is simply that $a^0 = 1$ for all a. Properties 4 and 5 are written only for convenience - they follow immediately from the previous three properties. Can you see why?

Given that for any value of $a > 0, a \neq 1$ we have both an exponential and corresponding logarithmic function, we have access to a plethora of functions through exponentials and logarithms. However, as we noted previously, we are currently unable to evaluate exponentials for all but a very small set of numbers. Similarly, we have difficulty in actually finding the values of logarithmic functions. Thus, even though we have defined this large class of functions, and have found that they have a number of useful properties, we still cannot actually evaluate them in most situations.

It turns out that the solution to this problem lies in polynomial functions. Given the difficulty of evaluating exponential functions, we can instead turn to approximating their values. Using the power of calculus (through the limit) we will actually be able to represent exponential functions as a sum of an infinite number of power functions (with natural-number exponents). Since it is relatively easy to evaluate power functions, this will give us a means of accessing these much more elusive exponential functions. We will deal with logarithmic functions in a slightly different way, but calculus will once again be essential.

Although exponential and logarithmic functions define infinite classes of functions, we are really only interested in a single function from each of these classes. The exponential function we will be interested in is the base e exponential, where e is a specific irrational number, defined by the limit

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

Because this function is used so much more often than other exponential functions, it is often referred to as *the* exponential function. Do not worry about understanding the above notation - as we delve into the study of limits and sequences it will begin to make sense. Corresponding to the base e exponential, we are interested in $\log_e(x)$, the *natural logarithm*, which is often written as $\ln(x)$. The graphs of these two functions are given in figure 2.6.

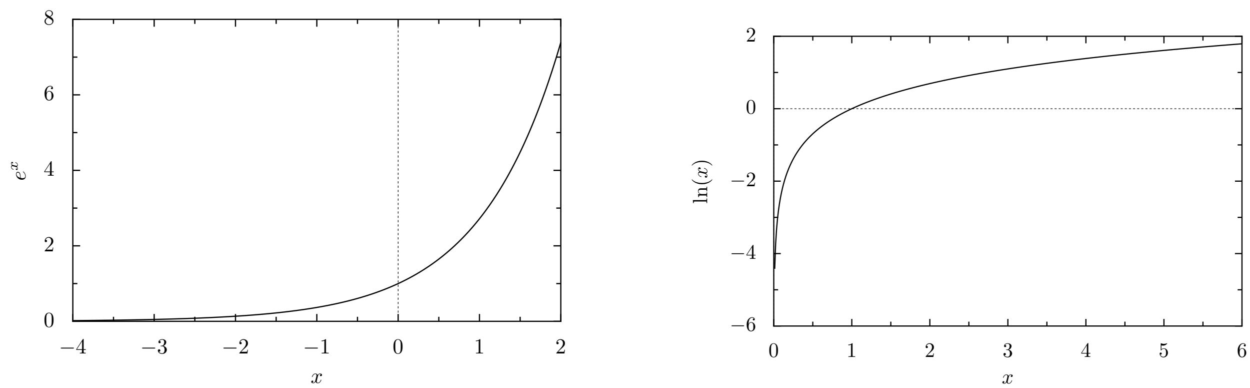


Figure 2.6: Exponential and Logarithmic Functions

While the exponential function and natural logarithm are really the only two functions of interest for us, the base 10 logarithm is sometimes encountered as well. The base 10 logarithm often written $\log(x)$ for short, which was widely used to simplify calculations before the advent of computers. The base 10 logarithmic is a part of the definition of a decibel, so it is encountered in fields such as telecommunications and acoustics.

2.5 Trigonometric Functions

Almost all there is to know about trigonometric functions can be derived from a single picture - the unit circle (a circle centered at the origin with a radius of 1). For this reason we will spend a considerable amount of our time studying trigonometric functions focusing on the unit circle.

First we should note that it is customary to use radians rather than degrees in the study of calculus. In the measure of radians $2\pi = 360°$, which means that the angle rotated in completing a circle is 2π. This corresponds exactly with the circumference of 2π of the unit circle. If we draw half of the unit circle, the length of the arc will be π (half of the circumference), and so will the angle rotated. Thus, on the unit circle, the length of an arc is equal to the angle of the arc. This is an immediate benefit of radians over degrees. There are other benefit of radians over degrees, but they will not become evident until we study differentiation.

Many times when angles are measured physically they are measured in degrees, so to deal with them mathematically we will want to convert between degrees and radians. To do so we use the following conversion factors

$$1 = \frac{180°}{\pi \text{ radians}} \quad \text{and} \quad 1 = \frac{\pi \text{ radians}}{180°}$$

which are derived from the fact that $2\pi = 360°$. Using this relationship we can see for instance, that that $60° = \frac{\pi}{3}$radians, and that $\frac{\pi}{4}$radians $= 45°$.

The two basic trigonometric functions sine and cosine are defined using coordinates along the unit circle. More specifically, if we draw a ray with angle θ with respect to the x-axis, $\sin(\theta)$ and $\cos(\theta)$ are given by the point of intersection of the ray and the unit circle (see figure 2.7), with

$$(x, y) = (\cos\theta, \sin\theta).$$

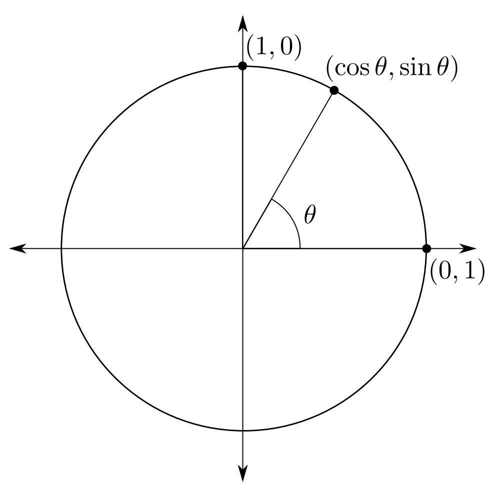


Figure 2.7: Sine and cosine defined on the unit circle.

This single relationship defines sine and cosine as functions with a domain of all of the real numbers. Positive angles are defined by counterclockwise revolutions, and negative angles are defined by clockwise revolutions. Angles larger than 2π are defined by a number of complete revolutions on the unit circle, plus a specific position on the unit circle. For instance, an angle of $7\pi/3$ corresponds to one full revolution around the unit circle 2π plus $\pi/3$, which ends up in the first quadrant. Negative angles smaller than -2π are defined analogously.

Using the above relationships we can sketch graphs of the sine and cosine functions. The easiest way to do so is to find the values of $\sin(\theta)$ and $\cos(\theta)$ for a number of specific values of θ, and then *connect the dots*. The easiest value to begin with is $\theta = 0$, which corresponds to the point $(1, 0)$ on the unit circle, so

$$\cos(0) = 1 \quad \text{and} \quad \sin(0) = 0.$$

Next we choose the point $\theta = \pi/2$, which corresponds to a right angle, touching the point $(0, 1)$. Here we have

$$\cos(\pi/2) = 0 \quad \text{and} \quad \sin(\pi/2) = 1.$$

Similar analysis works for the points $\pi, 3\pi/2, 2\pi$, etc. In order to sketch a reasonable graph we need many more points that this. We can pick up a number of points using the relationships between sides of a triangle. Namely, we will construct triangles with the angles $\pi/6, \pi/4, \pi/3$. In the first quadrant, we have the triangles shown in figure 2.8.

For the other quadrants we will use the same process, except that the reference triangles will be in different places.

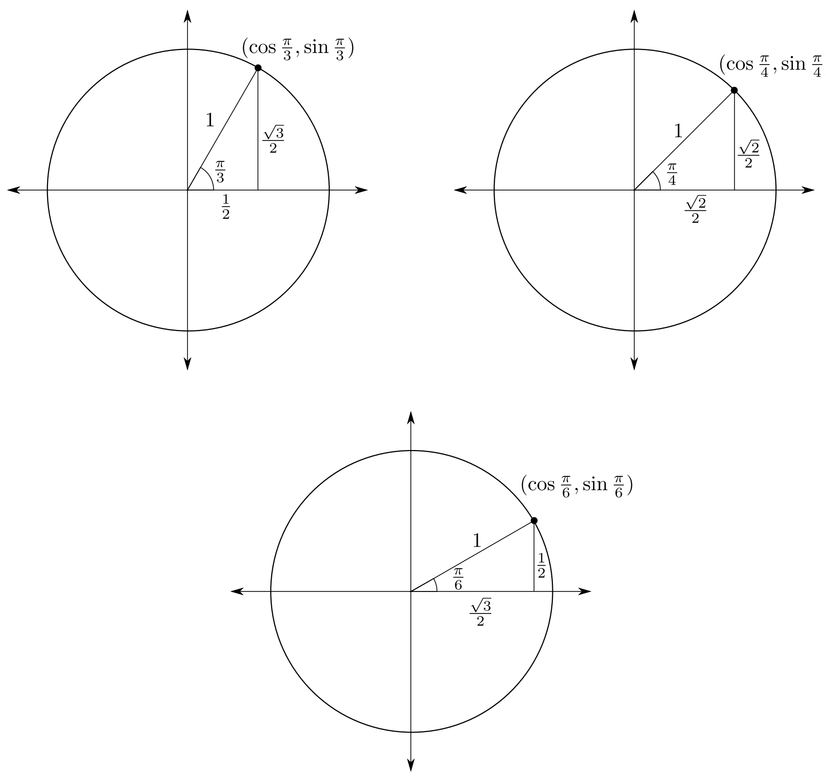


Figure 2.8: Deducing Values for Sine and Cosine Using Familiar Triangles

These particular values $\pi/6, \pi/4, \pi/3$ correspond to $30°, 45°, 60°$, which belong to both the 30,60,90 and 45,45,90 right triangles. The sides of these triangles are well-known, stemming from the study of geometry. The sides of these triangles are

$$\frac{1}{2}. \frac{\sqrt{3}}{2}, 1 \quad \text{and} \quad \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 1$$

respectively. Using these relationships, and correlating the x and y coordinates of the triangles

θ (radians)	θ (degrees)	$\sin\theta$	$\cos\theta$
0	0	0	1
$\pi/6$	$30°$	$1/2$	$\sqrt{3}/2$
$\pi/4$	$45°$	$\sqrt{2}/2$	$\sqrt{2}/2$
$\pi/3$	$60°$	$\sqrt{3}/2$	$1/2$
$\pi/2$	$90°$	1	0

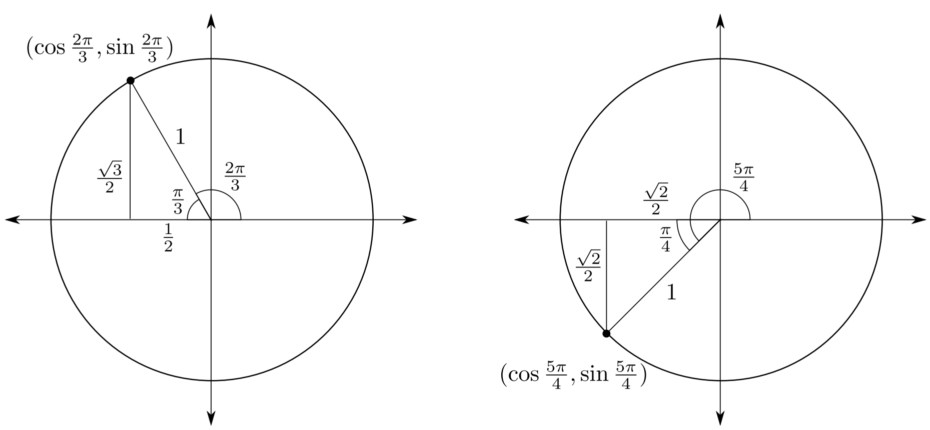
Table 2.1: Values of Sine and Cosine in the First Quadrant.

with the sine and cosine functions, we construct a table of values for the first quadrant (see table 2.1).

To find the values of sine and cosine in the other quadrants we construct the same triangles, but position them differently. The magnitude of sine and cosine are the same throughout each of the four quadrants, the only thing that changes are the signs. In the first two quadrants sine is positive (because it is above the x-axis) and in the third and fourth quadrants it is negative (below the x-axis). Cosine is positive in the first and fourth quadrants (right of the y-axis), and negative in the other two (left of the y-axis). For example,

$$\cos(2\pi/3) = -\frac{1}{2} \quad \text{and} \quad \sin(3\pi/2) = -\frac{\sqrt{2}}{2}$$

which are both found by constructing the same triangles as above only in the second and third quadrants, respectively (see figure 2.9).



(a) Cosine is negative in the 2^{nd} quadrant

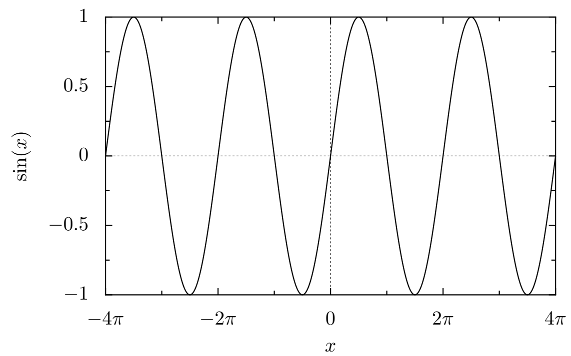
(b) Sine (and cosine) are negative in the 3^{rd} quadrant

Figure 2.9: Values of Trigonometric Functions in Other Quadrants

Using this method we can plot a reasonable sample of points for both the sine and cosine function, and *connecting the dots* we generate the graphs in figure 2.10.

Since after revolving 2π along the unit circle we return to our starting point, the sine and cosine functions repeat every 2π radians. Functions with this type of repeating behavior are said to be *periodic*. Mathematically,

$$\sin(x) = \sin(x + 2n\pi) \quad \text{and} \quad \cos(x) = \cos(x + 2n\pi),\ n \in \mathbb{Z}.$$



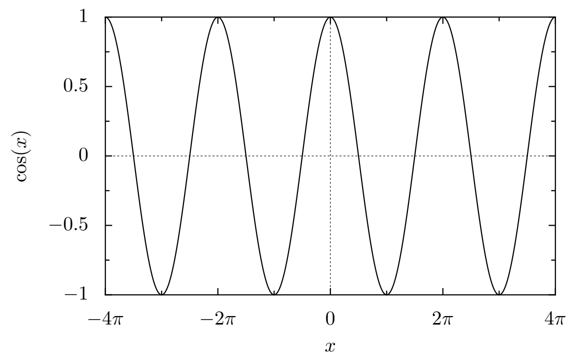


Figure 2.10: Fundamental Trigonometric Functions

Since the sine and cosine functions repeat every 2π radians, it follows that they repeat every $2n\pi$ radians, which just corresponds to rotating around the unit circle multiple times, reaching the same point.

We can also see that the sine and cosine graphs look nearly identical, except that one is a horizontal shift of the other. Thus,

$$\sin(x + \pi/2) = \cos(x).$$

Finally, we note that since the values of cosine and sine correspond to the x and y coordinates of the unit circle, satisfying the equation

$$x^2 + y^2 = 1,$$

so we have that

$$(\cos x)^2 + (\sin x)^2 = 1,$$

which is often written

$$\cos^2 x + \sin^2 x = 1.$$

In the way it is written above, the notation means

$$\cos^2 x = \cos x \cdot \cos x,$$

simply written out of stylistic preference. This identity is called the Pythagorean identity, because if we draw any of the right triangles in the unit circle as we did above, this is just a statement of the Pythagorean theorem (the sum of the squares of the two sides equal the square of the hypotenuse, which is in this case 1). Similarly, we obtain all of the rules of right triangle trigonometry simply by viewing any right triangle as similar to a triangle in the unit circle. Recall that for similar triangles, while the lengths of the sides are different, the lengths of the ratios of sides are the same. In other words, the ratio of the height of a similar triangle over its hypotenuse will be same as the ratio of the height of the triangle in the unit circle, over its hypotenuse of one. Since the height of such a triangle is just $\sin(\theta)$, we have that

$$\frac{\text{height}}{\text{hypotenuse}} = \sin(\theta).$$

Similarly we find that

$$\frac{\text{base}}{\text{hypotenuse}} = \cos(\theta).$$

Another trigonometric function encountered frequently in right-triangle trigonometry is the tangent function. Recall that the other four trigonometric functions are all simply combinations of sine and cosine. They are defined as

$$\tan(x) = \frac{\sin(x)}{\cos(x)}, \quad \cot(x) = \frac{\cos(x)}{\sin(x)}, \quad \sec(x) = \frac{1}{\cos(x)}, \quad \text{and} \quad \csc(x) = \frac{1}{\sin(x)}.$$

Once again viewing a right triangle as similar to one of our right triangles with unit hypotenuse, we see that

$$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)} = \frac{\text{height}}{\text{base}}.$$

When using these relationships it is important to be cognizant of the location of the angle θ, as *height* and *base* are relative concepts. For most of our purposes sine and cosine will be sufficient, so we will not focus much on the other trigonometric functions.

2.6 Dimensions and Dimensional Analysis

When we deal with physical systems, we need standard conventions for measuring things. The universe does not come equipped with a standard system of units, so a measurement such as a length of 10 is meaningless. A measurement of 10 forces one to ask the question: *10 what?* In order to answer this question, we develop standard systems of units. Units are defined relative to some known quantity. For instance, the meter is defined as the distance traveled by light in a vacuum during a time of 1/299758458 seconds. When this level of precision is not required, one can simply refer to a meter stick or other physical object which is *close* to this length. Having such a reference, we can now assign meaning to a distance of 10 m, which is simply 10 times this standard distance of reference. In this way, one can see that units are somewhat arbitrary, in that there is no reason we could not choose a difference reference distance for our standard of length (which is one reason there are multiple systems of units).

Even though systems of units are somewhat arbitrary, it does not mean that all systems of units are equal. The SI system of units by far the most widely system of units but other systems of units may be better suited for some applications, such as atomic units, which can simplify calculations at the atomic level (because the SI units are defined relative to such larger quantities). One key benefit to the SI system of units is that all multiples of a given unit are represented using convenient prefixes, such as kilo, nano, etc. This allows one to avoid tedious conversions between the same type of unit, unlike certain other systems of units, notably the U.S. customary units, which are based on the antiquated Imperial units, which are no longer used in England. In this sense the U.S. customary units can be considered inferior to the SI system of units, because they require numerous extra calculations to convert between units such as inches and miles, and the relationships between different units of the same dimension are extremely arbitrary, and difficult to remember. For these reasons, we will use the SI system of units, which is the international standard in science.

In the SI system of units, all units are derived as combinations from a set of base units, consisting of: meter (m), kilogram (kg), second (s), ampere (A), kelvin (K), mole (mol), and candelas (cd). For instance, Joules (J), the unit of energy, are defined by

$$J = m^2 \cdot kg/s^2.$$

The fact that all SI units are defined in terms of these base units gives us access to a very powerful tool - dimensional analysis.

Definition 2.6.1 (Dimension)**.** A *dimension* is a fundamental aspect of nature, such as length, time, energy, etc. Physical quantities of a certain dimension are represented relative to some reference, using units. For instance, in the SI system of units, length is represented in meters, time in seconds, and energy in Joules.

The idea of dimensional analysis is that in any physical relationship the dimensions must be consistent. This can equivalently be implemented using units, whereby after all units have be converted to some standard (for instance, all units could be written in terms of the base units of the SI system), the units of the relationship must be consistent. For a basic example, let's suppose someone asserts the velocity of an atom in space is given by the height of the atom above the earth times the width of the atom. Using dimensional analysis we can immediately refute this claim, by noting that the units of velocity are meters per second (m/s), whereas the units of width times height are meters squared (m^2). Since the units (and thus dimensions) of the relationship are inconsistent, it is impossible for the relationship to be true.

It's worth emphasizing that simply because the dimensions or units of a relationship match up, the relationship is not necessarily true. Suppose that a individual asserts that his velocity is equal to his height divided by the amount of time he can hold his breath. If we look at the units on each side we have m/s, but this relationship is clearly nonsense. Thus, dimensional analysis is not a tool to be used in order to verify whether or not a relationship is correct, but it can be used to verify if a relationship is incorrect.

The idea of dimensional analysis is extremely simple, but also extremely useful. If one is making a calculation to try and find the energy of a system, but ends up with units of force, then one immediately knows there was an error in the calculation. Noting that

$$J = m^2 \cdot kg/s^2 \quad \text{and} \quad N = m \cdot kg/s^2,$$

we see that

$$J = N \cdot m,$$

so the error in the units is a factor of meters. This information can guide an individual in looking for where the error in calculation occurs. In this way, one can think of dimensional analysis as a *sanity check* of one's work, to see if the result is even sensible.

Another useful tool for checking one's work is *order of magnitude analysis*. Roughly speaking, the order of magnitude of a quantity is the number of digits that appear in it. Thus, the order of magnitude of five thousand is four. It is often convenient to represent order of magnitude using scientific notation, so we would say the order of magnitude of 5000 is 10^4. The basic idea of order of magnitude analysis is to have some idea of what approximate order of magnitude a physical quantity should have. For instance, if one is looking at the velocity of a person running, one should recognize that 100 m/s is probably not a very reasonable answer, as it is much faster than a human can run. This is a very simple example, but one might also use it in calculating the wavelength of a laser. If one was calculating the wavelength of an ultraviolet laser, and the calculated wavelength was in the visible spectrum, one should know that something is awry.

2.7 Rates of Change

By definition, a function is simply a relation between the points or members of two sets. In this setting, the output of a function at a single point may be completely unrelated to the output of the function at other points. However, if a function represents a physical quantity, then it's reasonable to believe the outputs of the function for different points in the domain are related (at least in the realm of classical mechanics). One of the basic goals of calculus is finding and understanding these relationships, in order to understand the evolution of a system.

Let us suppose we know the position of an object as a function of time, at all times. We know that there must be at least some relationship between the positions of the object for closely related times. In other words, the position of the object at a given time is definitely related to the position of the object 1 second later. We may not be able to find one position given the other, but we know the distance traveled can only be so great, as macroscopic objects do not teleport. A very natural and interesting question that arises is as follows: given the position of an object at all times, can we find its velocity? Roughly speaking, an object's velocity (or speed) is a measure of how fast it is moving, or changing positions. Noting that velocity (or speed) is simply the rate of change of position with respect to time, we certainly *should* be able to find it.

Definition 2.7.1 (Average Rate of Change)**.** The *average rate of change* of a function f over an interval $[x_0, x_1]$ is the ratio of its change in output $(f(x_1)-f(x_0))$ over change in input $(x_1-x_0 = h)$. Mathematically,

$$\text{Average Rate of Change} = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{f(x_0 + h) - f(x_0)}{h}$$

The above definition holds in a more general context, but for the moment we're interested in a changing position.

Definition 2.7.2 (Velocity and Speed)**.** The *velocity* of an object is the rate of change of its position with respect to time. The *speed* of an object is the magnitude of its velocity.

According to the above definition, velocity describes how fast an object is moving, and in which direction, whereas speed simply denotes how fast an object is moving. Since velocity is the rate of change of position with respect to time, *average velocity* is the average rate of change of position with respect to time.

How should we interpret average velocity? Let $x(t)$ denote the position of an object D, and t represent time. The average velocity of an object over an interval $[a, b]$ in time is given by

$$v_{ab} = \frac{x(b) - x(a)}{b - a} = \frac{x(a + h) - x(a)}{h},$$

where $h = b - a$. If another object E is moving with a constant velocity v_{ab}, then over the interval of time $[a, b]$, E will travel a distance of

$$v_{ab} \cdot h = \frac{x(a + h) - x(a)}{h} \cdot h = x(a + h) - x(a) = x(b) - x(a),$$

which is exactly the distance that D travels over the interval $[a, b]$. In other words, the average velocity of an object over a given interval of time is the unique number such that for any other object, if the second object travels at a constant velocity of the average velocity for that interval of time, it will travel the same distance as the original object. The key thing to note here is that object D is not necessarily traveling at a constant velocity - it could be speeding up and slowing down.

In the above discussion we are implicitly referencing some mysterious *nonaverage* velocity, which changes throughout an interval of time. In most physical situations, we are interested in this other velocity, the one that it is changing throughout an interval of time. This velocity is called an object's *instantaneous velocity*. If we think about two objects colliding, we are interested in how fast the objects are moving right when they hit, not over an interval of time before they hit. Knowing the average velocity of the objects for 5 seconds before they collide doesn't tell us much, because we don't know whether the objects are moving much faster or slower than average when they collide. Nevertheless, the average velocity does tell us *something*.

If over some small period of time the objects have a given average velocity, we know that the instantaneous velocities within that interval of time are somewhat close to the average, because of the physical limits on how rapidly the velocity of an object can change. Stated more precisely, the difference between the average and instantaneous velocities is limited, because objects have finite accelerations. Thus, the average velocity over a small interval provides an approximation of the instantaneous velocity of points inside the interval.

In order to find the instantaneous velocity of an object at a point, we begin by finding the average velocity of the object over an interval including that point. This provides us with an approximation of the instantaneous velocity. By decreasing the size of the interval over which we calculate the average velocity, we will get another approximation. It is possible that this approximation may be worse than the original one, but on a whole, as we decrease the length of time over which we calculate average velocity, we should increase the accuracy of our approximation. We should be able to make the approximation as accurate as we like, simply by calculating the average velocity over small enough intervals. Given that

$$v_{ab} = \frac{x(a+h) - x(a)}{h},$$

we want to make h as close to 0 as possible, in order to increase the accuracy of our approximation. The instantaneous velocity is the value *approached* by these successive approximations, as h approaches 0.

We can think of the above process in two different but equivalent ways. As described above, the instantaneous velocity is a property of the object that already exists, that we are able to find by using average velocities. Alternatively, we can simply define it as the velocity we find by looking at the average velocity over shorter and shorter intervals. Which viewpoint you choose to hold is simply a matter of preference.

Although we have outlined a process for finding instantaneous velocity, there is still one more question we need to answer. Practically speaking, how do we find instantaneous velocity? There are fundamentally two different approaches we can take here. From a theoretical perspective, if we follow this process of approximation to its logical conclusion, we should be able to find the exact value of the instantaneous velocity. However, at the present time we do not have the tools to follow this process to its logical conclusion. The tool we need to do so is called a limit, which will be the fundamental concept of study in calculus. From a practical perspective, there are physical limits to how accurate we can make our approximations. These limitations may result from computers, measuring devices, etc. The study of such limitations and how to deal with them falls primarily in the realm of numerical analysis, so they will not be our primary focus of study.

Chapter 3

Limits

Calculus would be aptly described as the mathematics of approximation. Essentially every idea in calculus utilizes the notion of approximation, in order to solve problems that would either be very difficult or impossible to solve otherwise. The notion of approximation is realized through the limit. For this reason, we will begin our study of calculus with limits.

Contents

3.1 Definition of a Limit 39

3.2 Properties of Limits 49

3.3 One-sided Limits 55

3.4 Infinite Limits 60

3.5 Finding Infinite Limits: Order of Magnitude Analysis 65

3.6 Continuity 70

3.7 Sequences 74

3.8 Irrational Numbers 77

3.9 The Bisection Method 83

3.10 First-Order Approximations 85

3.1 Definition of a Limit

Suppose you have decided to volunteer (or are being paid to do so, if you need money as a motivator) to help with a juvenile sports team (either baseball or cricket - your sport of choice). Because of your adept mathematical and reasoning skills, your task is to help out with technical and strategical issues. During practice the coach of the team asks you to setup the pitching machine to 27 m/s (about 60 mph). Let the pitching machine be represented by the function $f(x) = x^3, x \neq 3$, where the domain of this function is the dial that sets the speed, and the speed of the pitch is the output of the function. Now you have a problem here, because the dial is broken, insofar that it skips over the value $f(3) = 27$, which is exactly the value you need! Rather than disappointing the coach and the team, you try to find a solution to this problem.

Realizing that by sight it is pretty difficult to judge the speed of a soaring projectile, you decide to set the knob to 3.1, presuming that the coach and team members won't be able to tell the difference (30 m/s is close to 27 m/s after all). This might work at first, but the coach might bring out an instrument to measure the speed of the pitch. Also realizing that there is a limited accuracy to this device, you set the dial to 3.001 (so the output is about 27.03). If the coach brought out a more accurate instrument, you could just crank the dial closer and closer to 3 (granted that there would be physical limitations after some point).

The idea here is that even though you cannot make the machine pitch at *exactly* 27 m/s, you can approximate 27 m/s as accurately as you would like by turning the nob close enough (but not equal to) 3. This is the basic idea of the limit - approximation. In this case we would write

$$\lim_{x \to 3} f(x) = 27$$

to mean that for this function f, we can approximate the output of 27 as accurately as we like by considering x values approaching (but not equal to) 3.

Another useful observation is that when we look at the graph of $x^3, x \neq 3$, in a sense we can see what the value of the function *ought to be* at $x = 3$, even though the function is not defined there. In this sense, the limit as $x \to 3$ is the value that the function *ought to have* at $x = 3$, regardless of whether or not it actually has the value at that point. We gain one advantage by thinking of a limit in this way. Intuitively, it only makes sense that a function ought to take on a single value at a given point (after all, if it were to take on multiple values it would not be a function). Due to this observation, we demand that a limit be unique, and if it is not possible to find a unique value so that a function *ought to have* that value at a given point x_0, then we would say the limit as $x \to x_0$ does not exist.

A key aspect of mathematics is precision of language. Having developed some notion of a limit, we need to formulate a definition that encapsulates the properties a limit *should* have. Let's examine a number of possible definitions, where they fall short, and how they lead us to the *actual* definition of a limit. Remember that we could define a limit differently, but then it would have different properties. The definition we want is the one that will give the limit the properties we want it to have (which may not yet be obvious at this time). In the following discussion, we will move from the most coarse, to most refined definition of the limit.

Proposed Definition 3.1.1 (Limit)**.** When we write $\lim_{x \to x_0} f(x) = L$ we mean that as the input values x approach the point x_0, the output values of the function $f(x)$ approach L.

This is in a sense the most crude understanding of what a limit is. It is based on a perception that for a reasonable function, it seems like there is a value that the function *ought to have* at the point x_0, which is the limit L - whether or not the function is actually defined so that $f(x_0) = L$.

However, there are some major problems with this definition. First of all, this definition does not indicate the fact that the value of the function at the point x_0, $f(x_0)$, has no bearing on whether or not the limit exists, or what it exists as. Second, the terms *approach* and *approaches* are much too vague for us to really work with them with any amount of precision. It is reasonable to assume that when we say approach, we mean *get closer to*. Thus, this definition says as the input values x *get closer to* x_0, the function values $f(x)$ *get closer to* the limit value L. However, there is no mention of *how close* the function values need to get to the limit. Consider the example

$$f(x) = -x^2,$$

which is a parabola opening downward, with its vertex at the origin of the Cartesian coordinate plane. From the above discussion about the value the function *ought to have*, we know that

$$\lim_{x \to 0}(-x^2) = 0,$$

yet based on our current definition, we could just as well conclude that

$$\lim_{x \to 0}(-x^2) = 5,$$

or any other positive number. As $x \to 0$, clearly $f(x)$ gets closer to 5, which means that the limit is 5, according to the above definition. In this way we have an infinite number of different values for the limit, yet we want the limit to be unique. This is a serious problem with our definition, because there is nothing about the value 5 (or any other positive number) that makes it seem it *ought to be* the value of $f(0)$. We will need to keep searching for a better definition.

Proposed Definition 3.1.2 (Limit)**.** When we write $\lim_{x \to x_0} f(x) = L$ we mean that we can force the output values $f(x)$ to be arbitrarily close to, but not necessarily equal to, the limit value L, by evaluating f at x arbitrarily close to, but not equal to x_0.

This definition is a definite improvement upon the previous one. First, it illustrates the fact that the value of $f(x_0)$ has no bearing on the existence or value of the limit, because we don't care about the point $x = x_0$. Second, it clears up the problem with the previous example, insofar that

$$\lim_{x \to 0}(-x^2) \neq 5,$$

or any other positive value that one might decide to choose (the function values don't become arbitrarily close to 5). Nevertheless, this definition still does not guarantee uniqueness of a limit. For instance, consider the function

$$f(x) = \frac{|x|}{x},$$

which is 1 for $x > 0$ and -1 for $x < 0$. Using this definition we can conclude both

$$\lim_{x \to 0} f(x) = 1 \quad \text{and} \quad \lim_{x \to 0} f(x) = 1,$$

since we can choose x values arbitrarily close but not equal to 0, so that $f(x)$ is both arbitrarily close (in fact equal to) 1 or -1. Although this definition is an improvement, it is still insufficient, as it does not guarantee uniqueness of a limit.

Proposed Definition 3.1.3 (Limit)**.** When we write $\lim_{x \to x_0} f(x) = L$ we mean that we can force the output values $f(x)$ to be arbitrarily close to, but not necessarily equal to, the unique limit value L, by choosing x arbitrarily close to, but not equal to x_0.

This is exactly the same definition as above, with a single addition - the word unique. Now we require that the limit value L be uniquely related to the function f and the point x_0 in the way described by the definition - it cannot be related to another x_1 in the same way. This is more promising, as we can conclude that $\lim_{x\to 0}|x|/x$ does not exist, as there is no unique value that satisfies the relationship (there are two).

This is a perfectly good definition of a limit, as it encapsulates all of the properties we have decided a limit ought to have. Nevertheless, we will work with another definition instead of this one, in order to simplify the language used (avoiding the term *arbitrarily close*), and to emphasize that limits are about approximation.

Definition 3.1.4 (Limit)**.** When we write $\lim_{x\to x_0} f(x) = L$ we mean that L is the unique number that can be approximated as accurately as desired by evaluating the function f using x-values close enough to, but not equal to x_0.

This definition is equivalent to the one above, insofar that they classify limits in the same way. Here we replace the term *arbitrarily close* with *approximated as accurately as desired*, but the two phrases have essentially the same meaning. For any distance away from L, we can force the function values to be within that distance; ie. approximate the limit value L with any degree of accuracy. When we use the words *close enough to*, we have the exact same meaning as *sufficiently close* or *arbitrarily close to.*

At this point we could be satisfied with our definition of the limit, and agree to let this be our definition of the limit. However, if one does a little bit of reading he or she will probably not encounter this definition of the limit. The reason for this lies in the way mathematics is developed. Although we have a perfectly good definition of the limit, this definition is not ideal for making deductive arguments. Instead, we favor to recast this verbal definition in terms of variables, to arrive at the *formal* definition of a limit, which is found in most calculus texts.

Definition 3.1.5 (Formal ε-δ Definition of the Limit)**.** Let f be defined on an open interval containing x_0 (except possibly at x_0), and $L \in \mathbb{R}$. We say that L is the limit of f as x approaches x_0, written

$$\lim_{x\to x_0} f(x) = L$$

if for every number $\varepsilon > 0$, there exists a corresponding number $\delta > 0$ such that for all x with $0 < |x - x_0| < \delta$ we have

$$|f(x) - L| < \varepsilon.$$

Written in English, this definition reads: for any error tolerance of approximation ($\varepsilon > 0$), we can restrict the values $f(x)$ to be within that error tolerance of L ($|f(x) - L| < \varepsilon$) by restricting the values of x to be sufficiently close but not equal to x_0 (by choosing $\delta > 0$ and considering x with $0 < |x - x_0| < \delta$). If we are able to meet any error tolerance of approximation, it follows that we are able to approximate our function as accurately as we like. Here the introduction of δ makes precise the notion of sufficiently close - we only care about the points that fall within the given δ-interval around x_0. We require $0 < |x - x_0|$ in order to exclude the point x_0 from the interval of interest, because the limit does not depend on the function value at that single point.

The way this definition classifies a limit as existing or not existing differs from our previous definition. Consider

$$\lim_{x\to 0} \frac{|x|}{x}.$$

According to our previous definiton the limit does not exist because the values we can approximate are not unique (we can approximate both 1 and -1). For the ε-δ definition of the limit, the limit does not exist because if we try and put an error tolerance of $\varepsilon < 2$ around either of the limit values $L = -1$ or $L = 1$, any δ-interval around $x_0 = 0$ will contain function values that fall outside of the required error tolerance. In our previous definition we needed to explicitly state uniqueness in order to properly classify limits; this is not required with the ε-δ definition. Uniqueness follows from the way we have formulated the definition.

Our two definitions of the limit generally classify the same limits as existing or not existing, except for in very esoteric examples. Consider

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

This is a function that looks like $1/x$ when the input is a rational number, and 0 when the input is an irrational number. According to the first definition

$$\lim_{x \to 0} f(x) = 0,$$

because 0 is the unique number we can approximate as accurately as desired. However, according to the ε-δ definition of the limit the limit does not exist, because for any error tolerance around 0, the function values on the rational numbers fall outside of the error tolerance. For an example such as this, our guiding principle of what value the function *ought to have* doesn't really make much sense anymore. Thus, it is not entirely clear whether or not we would want a definition of the limit that classifies this limit as existing. Either way, the ε-δ definition of the limit is the definition we will work with, so for this function the limit as $x \to 0$ does not exist.

The formal ε-δ definition of the limit requires that the function be defined on an interval around the point x_0, which ensures that x_0 is a reasonable point to even consider the notion of a limit. Consider a point in the domain of f with no other points around it. In such a case we could construct a small enough δ-interval so that it would only contain x_0, and the limit would trivially exist, even though it would be completely unrelated to the notion of the limit we are interested in.

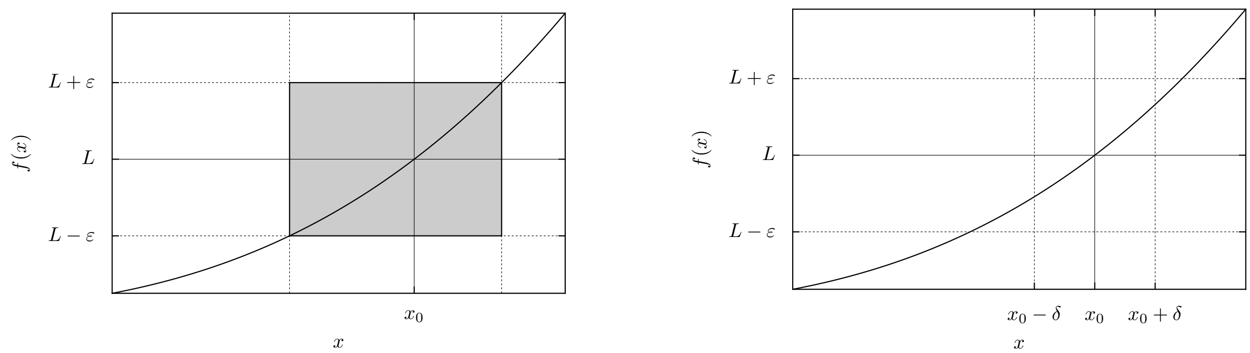
In figure 3.1 we have a graphical representation of the ε-δ definition of the limit. The point x_0 is the point we are looking at the limit at, and the point L is the actual value for the limit. Here $L+\varepsilon$ and $L-\varepsilon$ signify the bounds of the error tolerance - any function value must be within these two bounds in order to satisfy the error tolerance. In the first figure the vertical lines denote the positions on the x-axis where the function values transition between satisfying and not satisfying the given error tolerance. Thus, for any choice of x in between them, the function value satisfies the error tolerance, signified by the shaded region. In the second figure a specific δ has been chosen so that all function values for points within the δ-interval satisfy the error tolerance. Note that this is just a sample interval, and that any smaller (and some larger) intervals would work as well.

In order to make sense of the notion of approximation, we need to have a notion of distance. When we are looking at real numbers, the distance between two numbers x and y is simply $|x - y|$, the magnitude of their difference. Thus, both $|f(x) - L|$ and $|x - x_0|$ should be interpreted as the distance between points. When we are working with distances, the fundamental result we will build from is the triangle inequality.

Theorem 3.1.1 (Triangle Inequality)**.** *For $x, y \in \mathbb{R}$ we have $|x + y| \leq |x| + |y|$.*

Proof.

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



(a) The shaded region represents points where the function satisfies the error tolerance.

(b) One possible choice for δ satisfying the given ε.

Figure 3.1: Graphical interpretation of a limit.

The result follows upon taking square roots of both sides of the equation. $\square$

As stated before, it follows directly from the definition that a limit is unique.

Theorem 3.1.2 (Uniqueness of Limits). *Let $f : \mathbb{R} \to \mathbb{R}$. Suppose $\lim_{x\to x_0} f(x) = L$ and $\lim_{x\to x_0} g(x) = M$, where $L, M \in \mathbb{R}$. It follows that $L = M$.*

Proof. Fix $\varepsilon > 0$, arbitrary. Since $\lim_{x\to x_0} f(x) = L$ and $\lim_{x\to x_0} g(x) = M$ there exist δ_L and δ_M so that for x with $|x - x_0| < \delta_L$ we have

$$|f(x) - L| < \varepsilon/2$$

and for x with $|x - x_0| < \delta_M$ we have

$$|f(x) - M| < \varepsilon/2.$$

Choose $\delta > 0$ so that $\delta < \delta_L$ and $\delta < \delta_M$. Consider x with $|x - x_0| < \delta$. It follows that

$$|L-M| = |L+f(x)-f(x)-M| = |L-f(x)+f(x)-M| \leq |f(x)-L|+|f(x)-M| < \varepsilon/2+\varepsilon/2 = \varepsilon.$$

Thus, for any $\varepsilon > 0$, $|L - M| < \varepsilon$. It follows that $L = M$. Suppose $L \neq M$. This would imply that $|L - M|$ is a positive number, but then it could not be smaller than any arbitrary positive number ε. $\square$

Although proof is not our emphasis here, it is worth taking a minute to look at how the above proof works. To begin, we *fix* an arbitrary ε. By fixing the ε we are saying that we want to look at an arbitrary value, but one that does not change throughout the course of the proof. If we can show that we meet this arbitrary error tolerance, then it easily follows we can meet *any* error tolerance, because the proof doesn't depend on ε having a specific value - the proof works no matter what value is substituted for ε.

Next, we know that $\lim_{x\to x_0} = L$, which means that for any arbitrary $\epsilon > 0$, we can find a δ so that

$$|x - x_0| < \delta \implies |f(x) - L| < \epsilon.$$

Namely, we know that if we let $\epsilon = \varepsilon/2$, which is our fixed value, we can find such a δ. Since we want to work with multiple δ-intervals, we call this one δ_L. Using the exact same logic we find a

δ_M. Now we choose a δ smaller than both δ_L and δ_M. This means that any x in this δ-interval will fall in both the δ_L and δ_M intervals, so it will satisfy the corresponding inequalities in $f(x)$.

Now we want to show $L = M$, and one way of doing so is showing that $|L - M|$ is less than any positive number. If this is true, then the only possible conclusion is $|L - M| = 0$. Since we are working with approximations of arbitrary accuracy, this is going to be the best way to proceed, as trying to show directly that $L = M$ will not allow us to use the information we already have. Now our goal is to write $|L-M|$ in terms of quantities we already know are small, and if we *overestimate* in order to do so, it is no problem. The trick here is to add 0, or $f(x) - f(x)$, which lets us use the triangle inequality to overestimate by known quantities, eventually allowing us to reach the desired conclusion. Although these techniques may seem a bit unfamiliar, they are very common and very standard to proving results involving limits. With time they will feel much more familiar.

Example 1. Given

$$f(x) = 2x + 1,$$

is it possible to find a δ-interval around $x_0 = 2$ so that for x with $0 < |x - 2| < \delta$ it follows that

$$|f(x) - 5| < \varepsilon,$$

if $\varepsilon = 0.5$? Using this information can you conclude whether or not

$$\lim_{x\to 2} f(x) = 5$$

is a true statement?

Solution We will begin with the inequality we want to satisfy, and try to deduce a restriction on x in order to satisfy it.

$$\begin{aligned} |f(x) - 5| &< \varepsilon \\ |2x + 1 - 5| &< 0.5 \\ |2x - 4| &< 0.5 \\ 2|x - 2| &< 0.5 \\ |x - 2| &< 0.25 \end{aligned}$$

Since the inequality $|f(x) - 5| < \varepsilon$ is true for $|x - 2| < 0.25$, it follows that $\delta = 0.25$ will define an interval sufficiently small. Note that any smaller value for δ will also work, so we could just as well let $\delta = 0.1$. Either way, we conclude that it is possible to find such a δ-interval. Although

$$\lim_{x\to 2} f(x) = 5$$

is in fact a true statement, being able to find δ for a single error tolerance (as we have just done) is insufficient information to make this conclusion; we would need to find the width of the δ-interval for *every* ε. In order to do so we would need to work with an arbitrary ε.

Example 2. Given

$$f(x) = \sin(\frac{1}{x}),$$

is it possible to find a δ-interval around $x_0 = 0$ so that for x with $0 < |x| < \delta$ it follows that

$$|f(x) - 0| < \varepsilon,$$

if $\varepsilon = 0.1$? Using this information can you conclude whether or not

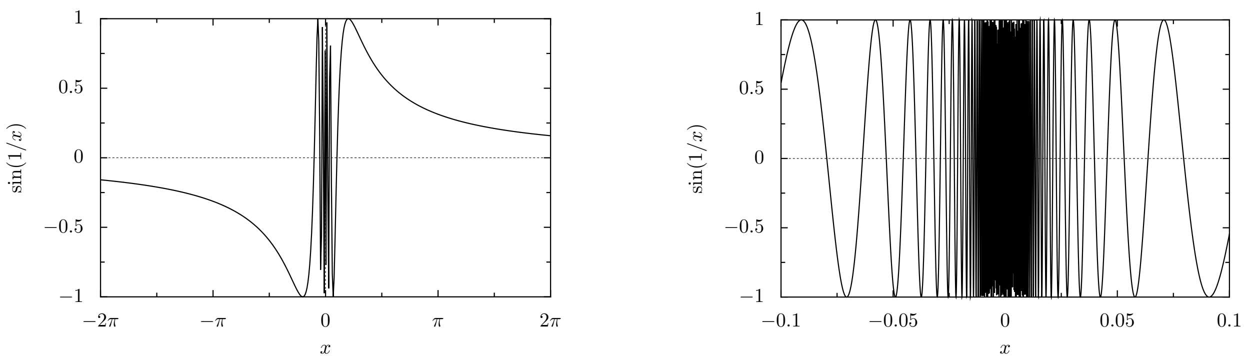
$$\lim_{x \to 0} f(x) = 0$$

is a true statement?

Solution If we approach this problem algebraically, we run into difficulty trying to manipulate the inequality

$$|\sin(\frac{1}{x}) - 0| < 0.1$$

in order to reach a restriction for x. Instead, let us consider the problem graphically (see figure 3.2).



(a) Zoomed out view.

(b) Zoomed in view.

Figure 3.2: Graph of the *sewing machine* function.

Figure 3.2 shows that the function values oscillate between 1 and -1, and that oscillations become more rapid as x becomes closer to 0. Because these oscillations continue indefinitely close to $x = 0$, no matter how small of a δ-interval we define around $x = 0$ we will be able to find a point x within it such that $f(x) = 1$, which is outside of the error tolerance around 0. Recalling that

$$\sin(x) = 1 \quad \text{when} \quad x = \frac{\pi}{2} + 2n\pi, \quad n \in \mathbb{Z}$$

we find that

$$\sin(\frac{1}{x}) = 1 \quad \text{when} \quad x = \frac{1}{\frac{\pi}{2} + 2n\pi}, \quad n \in \mathbb{Z}.$$

No matter how small of a δ-interval we consider, we will always find a point x as defined above, by considering sufficiently large values of n. Thus, in any interval around $x_0 = 0$ there will be values of x such that

$$|\sin(\frac{1}{x}) - 0| = 1 > 0.1.$$

We conclude that it is impossible to find a δ-interval so that the inequality $|\sin(\frac{1}{x}) - 0| < 0.1$ is satisfied for all x in the interval. Using this information we can conclude that

$$\lim_{x \to 0} \sin(\frac{1}{x}) \neq 0,$$

because we have found a single error tolerance $\varepsilon = 0.1$ for which we cannot find a corresponding δ-interval to satisfy the definition of the limit. In fact, it is not just that the limit does not equal 0, but that there is no value L that satisfies the limit (so the limit does not exist), which can be shown using a technique similar to the one we used above; we would simply fix a given error tolerance, such as $\varepsilon = 0.1$, and show that in any δ-interval we can find two points where the function values are further than ε apart. It is probably easiest to choose points where the output values are 1 and -1, but other choices would work as well.

Now let's consider some examples where we want to prove the existence or nonexistence of a limit for a given function. In order to do so, we need to work with an arbitrary ε. The standard means of doing so is *fixing* an arbitrary ε, and then performing the analysis with that ε. For this fixed ε, we need to find a δ so that if we have any x with $0 < |x - x_0| < \delta$ it follows that $|f(x) - L| < \varepsilon$, meaning that the values of the function within this δ-interval are within the error tolerance ε of the limit L. In general, δ will depend on ε (except for in some cases like $f(x) = c$).

Example 3. Prove that $\lim_{x\to 1} f(x) = 3$ if $f(x) = 2x + 1$.

Solution Fix $\varepsilon > 0$, arbitrary. We must find a corresponding $\delta > 0$ such that for $0 < |x-1| < \delta$ we have $|f(x) - 3| < \varepsilon$. We find that

$$|f(x) - 3| = |(2x + 1) - 3| = |2x - 2| = 2|x - 1|$$

and

$$2|x - 1| < \varepsilon \quad \iff \quad |x - 1| < \frac{\varepsilon}{2}.$$

If we let $\delta = \dfrac{\varepsilon}{2}$, then for an arbitrary x with $0 < |x - 1| < \delta$ it follows

$$|f(x) - 3| = 2|x - 1| < 2 \cdot \frac{\varepsilon}{2} = \varepsilon.$$

Thus, from the definition of the limit, it follows that

$$\lim_{x\to 1} f(x) = 3.$$

Example 4. Prove that $\lim_{x\to 3} f(x) = 9$ for

$$f(x) = \begin{cases} x^2 & x \neq 3 \\ 3 & x = 3 \end{cases}$$

Solution Fix $\varepsilon > 0$, arbitrary. We must find a corresponding $\delta > 0$ such that for $0 < |x-3| < \delta$ we have $|f(x) - 9| < \varepsilon$. Since $0 < |x - 3|$, it follows $x \neq 3$, so $f(x) = x^2$. Thus,

$$|f(x) - 9| = |x^2 - 9| = |(x + 3) \cdot (x - 3)| = |x + 3| \cdot |x - 3|.$$

We can control the value of $|x - 3|$ directly with δ, but we will have to control $|x + 3|$ indirectly. If we knew that δ would be less than 1, then we'd know from $0 < |x - 3| < \delta$ that $3 < x < 5$. Now the expression

$$|f(x) - 9| = |x + 3| \cdot |x - 3| < \varepsilon$$

consists of two terms we can control. If $\delta < 1$, then the maximum value of x is 5, so the maximum value of $|x + 3|$ is 8. It follows, granted $\delta < 1$, that if $|x - 3| < \varepsilon/8$, then $|f(x) - 9| < \varepsilon$. Thus, we choose $\delta = \min(1, \varepsilon/8)$.

Consider arbitrary x with $0 < |x - 3| < \delta$. For such an x, it follows

$$|f(x) - 9| = |x + 3| \cdot |x - 3| < 8 \cdot \frac{\varepsilon}{8} = \varepsilon.$$

Example 5. Prove that $\lim_{x \to 0} \frac{1}{x} \neq 3$.

Solution When we want to prove a limit is equal to a certain value, we need to work with an arbitrary ε, in order to show that our function values are close to the limit value for any error tolerance. In contrast, in order to show a limit does not equal a certain value, we need to find a single value for ε, so that for any δ-interval there is at least one point x within the interval with $|f(x) - L| > \varepsilon$, meaning that $f(x)$ is outside of the error tolerance. This shows that it is impossible to restrict the function values to be as close as we want to the limit value by choosing input values close enough to x_0, by exhibiting a *bad error tolerance*, that we cannot meet.

This particular function $f(x)$ grows without bound as $x \to 0$ from the right and decreases without bound as $x \to 0$ from the left, so clearly the limit does not exist (note that we are not trying to show that the limit does not exist, only that it does not have a specific value). Now it is just our job to come up with a value for ε, and show that no matter how small someone chooses δ to be, they cannot meet this error tolerance. One possibility is to simply guess a very small number, such as $\varepsilon = 10^{-10}$ and check if it will work. This strategy is not ideal however, because it complicates the mechanics of the problem. Since our function in this case grows without bound, we know it will exceed any error tolerance. Thus, we might as well choose a simple value for the error tolerance, and see if we can make progress from there. Let $\varepsilon = 1$.

Now we consider $\delta > 0$, arbitrary. At this point we need to use a little trick. We want to specify an x value within the δ-interval around 0, but we can't simply specify a point like $x = 0.2$, because we don't know how large δ is. Instead, we need to specify x in terms of δ. If $x < 0$, then we know that $f(x) < 0$, so we let

$$x = -\frac{\delta}{2}$$

which ensures both that $x < 0$, and that it falls in our δ-interval. For this x we have

$$0 < |x - 0| = \frac{\delta}{2} < \delta$$

with

$$|\frac{1}{x} - 3| = |\frac{1}{-\delta/2} - 3| = 3 + \frac{2}{\delta} > 1.$$

This proves that

$$\lim_{x \to 0} \frac{1}{x} \neq 3,$$

because we have shown that no matter how close to $x_0 = 0$ one creates a δ-interval, we can find points within where the function does not meet the error tolerance of 1 around the limit value $L = 3$.

Example 6. Prove that $\lim_{x \to 0} \frac{1}{|x|} \neq 10$.

Solution We need to find a value for ε so that for any $\delta > 0$ we can find an x with $0 < |x - 0| < \delta$ and $|f(x) - 10| > \varepsilon$. Looking at this function graphically we see this must be possible, but because we do not have the same separation of positive and negative values for $f(x)$ as we cross the y-axis, the mechanics of this problem are a little bit more difficult. For the sake of illustration, let us solve this problem using $\varepsilon = 0.5$, rather than $\varepsilon = 1$.

Consider $\delta > 0$, arbitrary. Suppose we try to use the same trick as last time and define

$$x = \frac{\delta}{2}.$$

Since δ is arbitrary, we don't know what value it might have. If it happens that $\delta = 0.2$, making $x = 0.1$, we would find

$$|f(x) - 10| = |\frac{1}{0.1} - 10| = 0 < \varepsilon = 0.5,$$

so we would have failed in our task to show the limit value is not 10. There is nothing special about choosing

$$x = \frac{\delta}{2},$$

and in fact we will run into the same problem no matter how we define x directly in terms of δ; if $x = k \cdot \delta$, then we will run into problems if it happens that $\delta = 0.1/k$. No matter how we define x, there is a corresponding range of δ values that will cause us trouble. However, if we can guarantee that x will not be in this problematic range of values, and $0 < |x| < \delta$ will be satisfied, then we can overcome this hurdle. We will need to restrict x close to 0 in order to accomplish this, but how small must we make x be? This depends on our choice of ε. Since we let $\varepsilon = 0.5$, in order to have

$$|f(x) - 10| = |\frac{1}{|x|} - 10| > 0.5$$

we will need to have

$$|x| < \frac{1}{10.5}.$$

After considering arbitrary $\delta > 0$, let us set $x = \min(\frac{1}{10.5}, \frac{\delta}{2})$. For such an x we find

$$|f(x) - 10| = |\frac{1}{|x|} - 10| > 0.5 = \varepsilon,$$

which proves

$$\lim_{x \to 0} \frac{1}{|x|} \neq 10.$$

3.2 Properties of Limits

Just as polynomials are linear combinations of power functions, most functions can be written as a combination of simpler functions. For instance, the function

$$f(x) = 2 \cdot x + 5$$

can be viewed as a combination of the functions

$$g(x) = 2, \quad h(x) = x, \quad \text{and} \quad i(x) = 5,$$

wherein

$$f(x) = g(x) \cdot h(x) + i(x).$$

Ideally, we should be able to infer something about a limit involving f using information of limits involving g, h, and i. If we are able to find rules for combining limits of functions, then we can replace the task of proving very complicated limits with that of using knowledge of basic functions, and the corresponding rules for combining limits. For functions we can construct from functions with known limits, this will allow us to avoid working with the ε-δ definition of the limit directly, which is often extremely difficult. As we increase the library of functions for which we know their limits, the more limits we will be able to evaluate while avoiding formal proof. We begin by establishing limits for two very basic functions.

Claim 3.2.1. $\lim\limits_{x \to x_0} c = c$

Proof. Fix $\varepsilon > 0$. For any $x \in \mathbb{R}$, we have $f(x) = c$. Therefore, for any $\delta > 0$, and x with $|x - x_0| < \delta$ we have

$$|f(x) - c| = |c - c| = 0 < \varepsilon.$$

□

Claim 3.2.2. $\lim\limits_{x \to x_0} x = x_0$

Proof. Fix $\varepsilon > 0$. Set $\delta = \varepsilon$. For any x with $|x - x_0| < \delta$ we have

$$|f(x) - x_0| = |x - x_0| < \delta = \varepsilon.$$

□

Now that we've proven limits for the most basic building blocks, we need to look at rules for combining them. Note that these results follow from the way in which we have defined the limit; thus, they are inescapable. By considering this notion of the limit, it must have the following properties.

Theorem 3.2.3 (Rules for Combining Limits)**.** *Suppose* $\lim_{x \to x_0} f(x) = L$ *and* $\lim_{x \to x_0} g(x) = M$ *where* $L, M \in \mathbb{R}$. *The following results are true.*

1. *Constant Multiple Rule:* $\lim_{x \to x_0} k \cdot f(x) = k \cdot L, \quad k \in \mathbb{R}$
2. *Sum Rule:* $\lim_{x \to x_0} [f(x) + g(x)] = L + M$
3. *Product Rule:* $\lim_{x \to x_0} [f(x) \cdot g(x)] = L \cdot M$

4. Quotient Rule: $\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{L}{M}, \quad M \neq 0$

The proofs of the first of these two results are well within our grasp. For the first we can fix an ε and choose δ so that

$$|f(x) - L| < \varepsilon/|k|.$$

For the second we just need to choose a δ so that

$$|f(x) - L| < \varepsilon/2 \quad \text{and} \quad |g(x) - M| < \varepsilon/2.$$

Once we've done so we simply apply the triangle inequality. The proof of the product rule requires the notion of boundnedness (or local boundedness more precisely), which we are not yet ready to tackle. Finally, the quotient rule requires assurity that our function stays sufficiently far away from 0 (which it must for the limit to exist as nonzero), but we will also avoid that proof at the moment. Using these rules in conjunction with the limits of $f(x) = x$ and $f(x) = c$ at any point we are now equipped to deal with the limit of *any* polynomial at *any* point.

Example 1. Find $\lim\limits_{x\to 1}(3x^2 + x - 4)$

Solution First we apply the sum rule to spread the limit to each of the individual terms, and then we can evaluate limit of the second two terms immediately. For the first, we use the constant product rule to separate 3 from x^2, and finally the product rule (using $x \cdot x$) to evaluate the limit for x^2. Thus,

$$\lim_{x\to 1}(3x^2 + x - 4) = 3\lim_{x\to 1} x + \left(\lim_{x\to 1} x\right)\cdot\left(\lim_{x\to 1} x\right) + \lim_{x\to 1}(-4) = 3\cdot 1 + 1\cdot 1 - 4 = -2$$

Example 2. Find $\lim_{x\to x_0} f(x)$ for all $x_0 \in [0,4]$, for f defined below. If the limit does not exist at any x_0, state why.

$$f(x) = \begin{cases} x & 0 \leq x < 1 \\ 0 & x = 1 \\ 1 & 1 < x \leq 2 \\ 3 - x & 2 < x < 3 \\ 1 & 3 \leq x \leq 4, x \neq 3.5 \\ 2 & x = 3.5 \end{cases}$$

Solution The first step is to graph this piecewise function, in order to obtain a better idea of what it looks like.

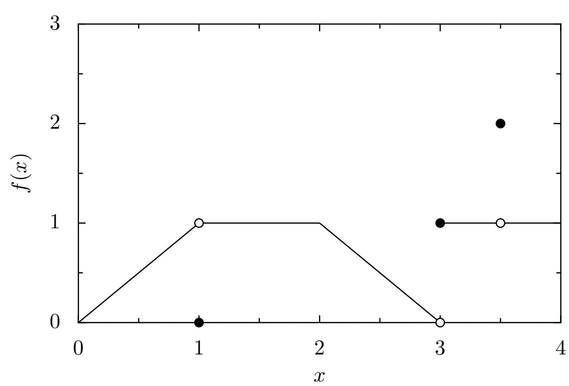


Figure 3.3: Graph of the piecewise function f.

Using the rules for combining limit and our results about $f(x) = x$ and $f(x) = c$ we can already determine that in the middle of an interval where $f(x)$ is defined as a line the limit will exist. Thus, we need only to concern ourselves with the junctions of these intervals, and the points where there are holes in the graph. Since the value of the function at a given point does not alter its limit, we can immediately conclude that the two holes in the graph are irrelevant, meaning that the limit exists the same as if the holes were filled. At a junction where the function lines up (ie. there are no jumps in the graph) the limit exists, as both sides are approaching the same value, in the fashion required by the limit. Thus, the limit exists at all points other than $x = 3$ where there is an actual jump in the graph of the function.

Polynomial functions are nicely behaved, in the sense that the limit value of a polynomial is the same as the function value at that point. However, rational functions (a ratio of two polynomials) do not always exhibit such nice behavior. Consider a rational function such as

$$\frac{x^2 - 1}{x - 1}.$$

If we choose $x_0 \neq 1$, then we can simply apply the quotient rule in order to find the limit as $x \to x_0$ (evaluating the limits of both the numerator and denominator separately). However, when we are looking at a point where the limit of the denominator is 0, such as when $x = 1$ for the above function, we cannot apply the quotient rule. Nevertheless, in some situations we may be able to rewrite the rational function so that it has a nonzero limit at the point of interest. Upon doing so we are once again able to apply the quotient rule. In general, we use one of the following techniques to rewrite a rational function.

1. *Factoring and Canceling a Common Factor*: If both the numerator and denominator are polynomial functions, look to see if the numerator is also zero at the same point as the denominator, say $x = x_0$. If so, then one can factor both the numerator and denominator, pulling out a factor of $(x - x_0)$. Once these factors are canceled, one checks to see if the quotient rule can be applied, repeating one of these techniques of necessary.

2. *Rationalizing the Numerator or Denominator*: Some rational functions involve square roots, so trying to directly find a factor is not the best approach. Instead, one multiplies the rational function by a convenient choice of 1, in order to remove the square root. After doing so, one looks for factors that can be divided out from the numerator and denominator. The idea lies in the fact that $a^2 - b^2 = (a + b)(a - b)$. If we have a term such as $\sqrt{x+1} + 2$, then if we multiply by $\sqrt{x+1} - 2$, we get a result of $(x + 1) - 4$, which does not have a square root in it. The term that we can multiply a term involving a square root by in order to remove the square root is called its *conjugate*. When we need to find a convenient choice for 1, we simply use the conjugate of the factor that is giving us trouble, divided by itself. Thus, if $\sqrt{x+1} - 2$ was a factor in the denominator giving us trouble, we would choose to multiply by

$$1 = \frac{\sqrt{x+1} + 2}{\sqrt{x+1} + 2}.$$

3. *l'Hôpital's Rule*: l'Hôpital's rule relies on differentiation, so it will not be discussed until section 4.13.

Example 3. Find $\lim\limits_{x \to 3} \dfrac{x^2 - 1}{x}$

Solution Since we are dealing with a rational function, the first thing to do is look at the limit of the denominator. Since it is nonzero, we are able to apply the quotient rule and evaluate the limits of the numerator and denominator separately. We find that

$$\lim_{x\to 3} \frac{x^2-1}{x} = \frac{\lim_{x\to 3}(x^2-1)}{\lim_{x\to 3} x} = \frac{8}{3}.$$

Example 4. Find $\lim\limits_{x\to 2} \dfrac{3x^2-5x-2}{x-2}$

Solution Since $\lim_{x\to 2}(x-2) = 0$, we cannot apply the quotient rule. Nevertheless, when we evaluate the numerator at $x = 2$ we see that $12-10-2 = 0$, so the numerator also has a 0 at $x = 2$. In order for a polynomial to have a 0 at some point x_0, it must contain a factor of $(x - x_0)$. Thus, we know there is a factor of $(x-2)$ in the numerator. Now we need to find another first-order term that will factor the original polynomial; ie, we need to find a and b so that

$$(3x^2-5x-2) = (a+b)(x-2).$$

Since we know $a \cdot x = 3x^2$ it follows $a = 3x$, and similarly $b = 1$. Thus,

$$\lim_{x\to 2} \frac{3x^2-5x-2}{x-2} = \lim_{x\to 2} \frac{(3x+1)(x-2)}{x-2} = \lim_{x\to 2}(3x+1) = 7.$$

We were able to *cancel* the factors of $(x-2)$ because when we are considering the limit we are looking only at points where $x \neq 2$ (it would be a mistake conclude that $0/0 = 1$). In other words,

$$\frac{(3x+1)(x-2)}{x-2} = 3x+1$$

only for $x \neq 2$.

Example 5. Find $\lim\limits_{x\to 0} \dfrac{\sqrt{x^2+100}-10}{x^2}$.

Solution We cannot use the quotient rule, because $\lim_{x\to 0} x^2 = 0$. Nevertheless, if we notice that

$$\lim_{x\to 0}(\sqrt{x^2+100}-10) = 0,$$

we might suspect that we can rewrite this rational function into a form in which we can evaluate the limit. There is no way for us to directly factor either the numerator or denominator, but if we can rewrite the expression without the square roots we may be able to. The technique here is to multiply by 1, where 1 is conveniently chosen to be the conjugate of the numerator over itself (which will remove the square root from the numerator).

$$\begin{aligned}
\frac{\sqrt{x^2+100}-10}{x^2} &= \frac{\sqrt{x^2+100}-10}{x^2} \cdot \frac{\sqrt{x^2+100}+10}{\sqrt{x^2+100}+10} = \frac{x^2+100-100}{x^2(\sqrt{x^2+100}+10)} \\
&= \frac{x^2}{x^2(\sqrt{x^2+100}+10)} = \frac{1}{\sqrt{x^2+100}+10}.
\end{aligned}$$

Having rewritten our function in a way so that the limit of the denominator is no longer 0, we can now apply the quotient rule. Taking the limits of the numerator and denominator we find that

$$\lim_{x\to 0} \frac{\sqrt{x^2+100}+10}{x^2} = \lim_{x\to 0} \frac{1}{\sqrt{x^2+100}+10} = \frac{1}{\sqrt{0^2+100}+10} = \frac{1}{20}.$$

Finally, there is one more useful theorem to discuss, which can be used to calculate some more obscure limits, and even some limits where we do not have an explicit expression for the function at hand.

Theorem 3.2.4 (The Sandwich Theorem). *Suppose $g(x) \leq f(x) \leq h(x)$ for all x in an open interval about x_0, not necessarily containing x_0. If*

$$\lim_{x \to x_0} g(x) = \lim_{x \to x_0} h(x) = L,$$

then

$$\lim_{x \to x_0} f(x) = L.$$

As the name suggests, the sandwich theorem can be applied to a function f sandwiched between between two other functions g and h. If the sandwiching functions g and h both have the same limit at the point of interest, then the function sandwiched between them *must* have the same limit, because there is nowhere else for it to go. One can imagine the function f to be a piece of lettuce, and the sandwiching functions as two pieces of bread. As randomly as the lettuce may move around, it is stuck between the pieces of bread. If we press tightly at some point on both sides of our sandwich, then at that point we know exactly where the lettuce must be, even if we don't know much about where it is at other points.

Example 6. Find $\lim\limits_{x \to 0} x \sin 1/x$.

Solution This is our favorite sewing machine function $\sin(1/x)$ multiplied (or modulated) by x. We know that the sewing machine behaves very badly $x \to 0$, oscillating more and more rapidly, but when we modulate it with x, the amplitude (or height) of the oscillations become smaller and smaller as $x \to 0$. In this way the function actually settles on a limit value of 0. Now we just need to show it.

To apply the sandwich theorem we need to find two functions - one that is always larger than the function and one that is always smaller - with the same limit at the point of interest. For an amplitude-modulated sinuousoid (sine or cosine) the choice should be directly related to the function modulating the amplitude, because the magnitude of sine or cosine will never exceed 1. Thus, we choose the functions $|x|$ and $-|x|$, because

$$-|x| \leq x \sin(1/x) \leq |x|, \quad \text{for all } x.$$

Now we have sandwiched our sinuousoid between two functions, which have that

$$\lim_{x \to 0} |x| = \lim_{x \to 0}(-|x|) = 0.$$

By the sandwich theorem it follows that

$$\lim_{x \to 0} x \sin 1/x = 0.$$

Example 7. Find $\lim_{x \to 0} f(x)$ if for every $x \neq 0$

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

Solution Since

$$\lim_{x \to 0} -x^2 = 0 = \lim_{x \to 0} x^2$$

it follows from the sandwich theorem that

$$\lim_{x\to 0} f(x) = 0.$$

The above analysis shows, for instance, that

$$\lim_{x\to 0} x^2 \cos^2(\frac{1}{x}) = 0.$$

3.3 One-sided Limits

When we consider

$$\lim_{x\to x_0} f(x),$$

we need to consider the values of the function f on an interval surrounding both sides of x_0. This may be impossible if x_0 is an endpoint of the domain of f, or may at the very least be problematic if f is defined piecewise, with different definitions on each side of x_0 (making it difficult to algebraically manipulate f). In order to overcome these difficulties, we can define one-sided limits, where we only consider x values on one side of x_0. When we write $x \to 2$ it means x approaches 2. Similarly, $x \nearrow 2$ means x *increases to* 2 and $x \searrow 2$ means x *decreases to* 2. On a number line, if $x \nearrow 2$ (is increasing to 2), then x approaches 2 from the left side, and if $x \searrow 2$ (is decreasing to 2), then x approaches 2 from the right side.

Definition 3.3.1 (Right-Hand Limit)**.** We say that L is the *right-hand limit* of $f(x)$ at x_0, written

$$\lim_{x\searrow x_0} f(x) = L$$

if for every number $\varepsilon > 0$, there exists a corresponding number $\delta > 0$ such that for all x with $x_0 < x < x_0 + \delta$ we have

$$|f(x) - L| < \varepsilon.$$

Definition 3.3.2 (Left-Hand Limit)**.** We say that L is the *left-hand limit* of $f(x)$ at x_0, written

$$\lim_{x\nearrow x_0} f(x) = L$$

if for every number $\varepsilon > 0$, there exists a corresponding number $\delta > 0$ such that for all x with $x_0 - \delta < x < x_0$ we have

$$|f(x) - L| < \varepsilon.$$

The above definitions are identical to ε-δ definition of the limit except that now we only require that the δ-interval extend on one side of x_0. Because these definitions are otherwise identical to the ε-δ definition, it follows that all of the rules we previously established for combining two-sided limits (and the sandwich theorem) still hold for one-sided limits. There is one additional result for relating one-sided and two-sided limits.

Theorem 3.3.1 (One-sided and Two-sided Limits)**.** *A function f has a limit L at x_0 if and only if both its left-hand and right-hand limits at x_0 are L.*

The if and only if in this theorem states an equivalence - having a two-sided limit is the same thing as having equal one-sided limits. Although this may seem like a very simple fact, it is indeed useful for showing that limits do not exist. Rather than trying to show all possible limit values do not work, we simply need to show that the one-sided limits are different. If the one-sided limits exist but disagree, then it is impossible for the function to approach a single value as $x \to x_0$, which implies that the two-sided limit does not exist.

Example 1. Prove that $\displaystyle\lim_{x\to 0} \frac{|x|}{x}$ does not exist.

Solution Recall that

$$\frac{|x|}{x} = \begin{cases} -1 & x < 0 \\ 1 & x > 0 \end{cases}$$

In order to prove the limit does not exist, we simply need to show that the one-sided limits as $x \nearrow 0$ and $x \searrow 0$ either disagree, or that one does not exist. Looking from just the left or right side of the point $x_0 = 0$, we have two constant functions. Since the limit of a constant is just that constant, it follows that

$$\lim_{x \searrow 0} \frac{|x|}{x} = 1 \quad \text{and} \quad \lim_{x \nearrow 0} \frac{|x|}{x} = -1.$$

It follows that $\lim_{x\to 0} |x|/x$ does not exist. Note that we have not just shown that the function does not have a specific limit value as $x \to 0$; we have shown that it does not have any limit value.

Example 2. Evaluate $\lim\limits_{x\to 0} f(x)$ and $\lim\limits_{x\to 2} f(x)$ for

$$f(x) = \begin{cases} x^2\cos(1/x) & x < 0 \\ \frac{3x^2-5x-2}{x-2} & 0 \le x < 2 \\ \frac{14(\sqrt{x^2+12}-4)}{x-2} & 2 < x. \end{cases}$$

Solution Since we are dealing with a piecewise function, we will need to look at the one-sided limits on both sides of the points of interest in order to evaluate the two-sided limits. Beginning with $x \to 0$, we find

$$\lim_{x \nearrow 0} f(x) = \lim_{x \nearrow 0} x^2 \cos(1/x) = 0$$

by the sandwich theorem, using sandwiching functions x^2 and $-x^2$. Looking at the other side we find

$$\lim_{x \searrow 0} f(x) = \lim_{x \searrow 0} \frac{3x^2-5x-2}{x-2} = \frac{-2}{-2} = 1.$$

Since the one-sided limits do not agree, we conclude that $\lim_{x\to 0} f(x)$ does not exist. Next we consider the limit as $x \to 2$. Doing so we find

$$\lim_{x \nearrow 2} f(x) = \lim_{x \nearrow 2} \frac{3x^2-5x-2}{x-2},$$

which we cannot evaluate directly using the quotient rule. Noting that the numerator has a root at $x = 2$, we know that

$$3x^2 - 5x - 2 = (a+b)(x-2)$$

for some a and b. By inspection we can see that $a = 3x$ and $b = 1$. It follows that

$$\lim_{x \nearrow 2} f(x) = \lim_{x \nearrow 2} \frac{(3x+1)(x-2)}{x-2} = \lim_{x \nearrow 2} (3x+1) = 7.$$

Considering the limit as $x \searrow 2$ we find

$$\begin{aligned} \lim_{x \searrow 2} f(x) &= \lim_{x \searrow 2} \frac{14(\sqrt{x^2+12}-4)}{x-2} = \lim_{x \searrow 2} \frac{14(\sqrt{x^2+12}-4)}{x-2} \cdot \frac{\sqrt{x^2+12}+4}{\sqrt{x^2+12}+4} \\ &= \lim_{x \searrow 2} \frac{14(x^2+12-16)}{(x-2)(\sqrt{x^2+12}+4)} = \lim_{x \searrow 2} \frac{14(x^2-4)}{(x-2)(\sqrt{x^2+12}+4)} \\ &= \lim_{x \searrow 2} \frac{14(x+2)(x-2)}{(x-2)(\sqrt{x^2+12}+4)} = \lim_{x \searrow 2} \frac{14(x+2)}{\sqrt{x^2+12}+4} = \frac{14 \cdot 4}{4+4} = 7. \end{aligned}$$

Seeing that both the left-hand and right-hand limits agree, we conclude that $\lim_{x\to 2} f(x) = 7$.

Example 3. Find $\lim_{x\searrow 0} \dfrac{1-\sqrt{x+1}}{\sqrt{x}}$.

Solution First note that we need to consider a one-sided limit because $\sqrt{x}$ is undefined for $x<0$. We cannot use the quotient rule, because $\lim_{x\searrow 0}\sqrt{x}=0$. Nevertheless, noticing that

$$\lim_{x\searrow 0}(1-\sqrt{x+1})=0,$$

we suspect that we can rewrite this rational function into a form in which we can evaluate the limit. Multiplying by a convenient choice of 1 (the conjugate of the numerator divided by itself) we find

$$\frac{1-\sqrt{x+1}}{\sqrt{x}}=\frac{1-\sqrt{x+1}}{\sqrt{x}}\cdot\frac{1+\sqrt{x+1}}{1+\sqrt{x+1}}=\frac{-x}{\sqrt{x}\cdot(1+\sqrt{x+1})}=\frac{-\sqrt{x}}{1+\sqrt{x+1}}$$

With this simplified expression, we can now apply the quotient rule, and taking the limits of the numerator and denominator, we find that

$$\lim_{x\searrow 0}\frac{1-\sqrt{x+1}}{\sqrt{x}}=\lim_{x\searrow 0}\frac{-\sqrt{x}}{1+\sqrt{x+1}}=\frac{0}{1}=0.$$

At this point there is a very useful result we can establish using one-sided limits. Consider

$$\lim_{x\to 0}\frac{\sin(x)}{x}.$$

We cannot apply the quotient rule because the limit of the denominator is 0. However, since the limit of the numerator is zero as well, we suspect that the limit still might exist. We will see in a moment that this limit is 1, which in essence states that for small values of x, $\sin(x)$ behaves like x. In many physical situations we can use this approximation to simplify the problem at hand. We will also see that this result has profound implications in the study of differentiation.

Claim 3.3.2. $\lim_{x\to 0}\dfrac{\sin(x)}{x}=1$, *($x$ in radians)*

Proof. Both $\sin(x)$ and x are odd functions, which means they are antisymmetric about the y-axis. In other words,

$$\sin(-x)=-\sin(x)\quad\text{and}\quad(-x)=-(x),$$

If we let $\mathrm{sinc}(x)=\sin(x)/x$, then

$$\mathrm{sinc}(-x)=\frac{\sin(-x)}{-x}=\frac{-\sin(x)}{-x}=\frac{\sin(x)}{x}=\mathrm{sinc}(x).$$

This means that $\mathrm{sinc}(x)$ is symmetric about the y-axis, so if the right-hand limit exists, then the left-hand limit must also exist and be equal. Thus, if we can show the right-hand limit exists it will follow that the two-sided limit exists.

Since we are looking as $x\searrow 0$, we only need to concern ourselves with $0<x<\pi/2$, which means we can view $\sin(x)$ as the height of a right triangle in the first quadrant of the unit circle. Inside the unit circle we construct the triangle with this height, and a base of 1. Outside of the circle we construct a triangle with base 1, and height $\tan(x)$. Between these two triangles we have a semicircular area (see figure 3.4).

The areas of the triangles are just given by $1/2\cdot\text{base}\cdot\text{height}$. To find the area of the semicircle we note that it is simply some percentage of the total area of the circle, which is given by $\pi\cdot 1^2=\pi$.

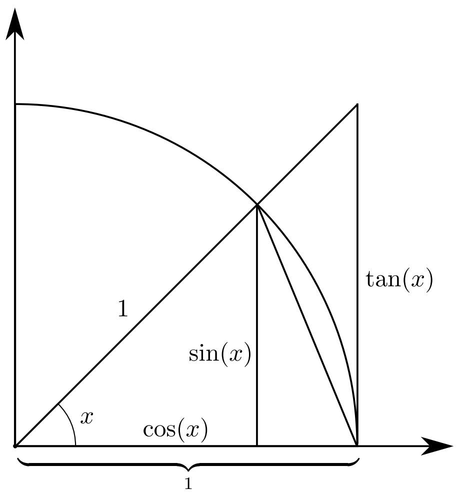


Figure 3.4: Graph of the areas in claim 3.3.2.

The percentage of the area is given by the percentage of the total angle, $x/2\pi$, times the total area (an angle of $\pi/4$ would correspond to $1/8^{th}$ of the total area). Thus, the area of the semicircle is given by

$$x/2\pi \cdot \pi = x/2.$$

Looking at the nesting of the areas in figure 3.4 we arrive at the inequalities

$$\frac{1}{2}\sin(x) < \frac{1}{2}x < \frac{1}{2}\tan x.$$

We can multiply all terms by 2 to remove the fraction 1/2, and divide by $\sin(x)$, which is positive (so it does not switch the order of the inequality), finding that

$$1 < \frac{x}{\sin(x)} < \frac{1}{\cos(x)}.$$

We take the reciprocal of this chain of inequalities, which reverses the direction, and we find

$$1 > \frac{\sin(x)}{x} > \cos(x).$$

Since $\lim_{x\searrow 0}\cos(x) = 1$, it follows using the sandwich theorem (for one-sided limits) that

$$\lim_{x\searrow 0}\frac{\sin(x)}{x} = 1,$$

and because the left-hand limit equals the right-hand limit, we have that

$$\lim_{x\to 0}\sin(x) = 1.$$

$\square$

Example 4. Show that $\lim_{x\to 0} \frac{\cos(x)-1}{x} = 0$.

Solution Since the limit of the denominator is zero we cannot use the quotient rule. Given that we've just proven (through a reasonable amount of work) a similar limit, we might try to rewrite cosine in terms of sine, using an appropriate trigonometric identity. In order to remove the 1 from the numerator, we use the half-angle identity

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

We find that

$$\lim_{x\to 0} \frac{\cos(x)-1}{x} = \lim_{x\to 0} \frac{-2\sin^2(x/2)}{x} = -\lim_{x\to 0} \frac{\sin^2(x/2)}{(x/2)}.$$

From here we are getting much closer to something of $\sin(x)/x$. The only problem is we have $\sin(x/2)$. However, we also have $(x/2)$ in the denominator. If we let $h = (x/2)$, then we are looking at

$$\lim_{x\to 0} \frac{\sin^2(h)}{(h)} = \lim_{2h\to 0} \frac{\sin(h)}{(h)} \cdot \sin(h).$$

Here we have written everything in terms of h, and have gotten very close to a familiar limit. We just need to notice that as $2h \to 0$, we have $h \to 0$, so we find

$$\lim_{2h\to 0} \frac{\sin(h)}{(h)} \cdot \sin(h) = 1 \cdot 0 = 0.$$

Thus,

$$\lim_{x\to 0} \frac{\cos(x)-1}{x} = 0.$$

3.4 Infinite Limits

The concept of infinity (or symbolically, ∞) plays an important role in calculus. This concept is related to the boundedness of a function.

Definition 3.4.1 (Bounded Function)**.** A function $f : X \to \mathbb{R}$ is *bounded* if there exists $M \in \mathbb{R}$ so that for all $x \in X$,

$$|f(x)| \leq M.$$

We say f is *bounded above* if there exists $M \in \mathbb{R}$ so that for all $x \in X$,

$$f(x) \leq M.$$

We say f is *bounded below* if there exists $M \in \mathbb{R}$ so that for all $x \in X$,

$$f(x) \geq M.$$

In each of these cases M is said to be a bound, upper bound, or lower bound, respectively. If no such M exists, then a function is said to be unbounded, unbounded above, or unbounded below, respectively.

The basic idea behind boundedness is creating a wall (either a ceiling or floor) that the function cannot pass. If a function is bounded above then we can create a ceiling above the function, so that the value of the function is always below the ceiling. For a function that is bounded below, we can construct a floor below it, so that the value of the function is always above the floor. For a bounded function, we can construct both a ceiling above it and a floor below it. The smallest number that works as an upper bound is called the *least upper bound* or *supremum*, and the largest number that works as a lower bound is called the *greatest lower bound* or *infemum*. Other than remarking that these concepts are extremely important in analysis, we will not focus on them here.

The notion of infinity is contrary to that of boundedness; infinity represents unboundedness. The exact definition of infinity is contextual, but in general ∞ represents something increasing without bound (which means there is no upper bound for the function), and $-\infty$ represents something decreasing without bound (which means there is no lower bound for the function). Before proceeding further, we need to emphasize that $\pm\infty$ are not numbers. Since ∞ and $-\infty$ are not numbers, performing arithmetic operations with them is nonsensical.

Claim 3.4.1 (∞ as a number is nonsense)**.** *Suppose there exists a real number ∞, which is larger than all other real numbers. It follows that $0 = \infty$, which is nonsense.*

Proof. Let us begin by multiplying ∞ by 2. Since there is no number larger than ∞, it is clear that the result must be ∞. Thus,

$$2 \cdot \infty = \infty.$$

It follows that

$$2 \cdot \infty - \infty = \infty - \infty = 0.$$

because for any number x, $x - x = 0$. We can also write that

$$2 \cdot \infty - \infty = (2 - 1)\infty = \infty.$$

Finally, we conclude that

$$0 = \infty,$$

by equating the above expressions. Thus, the number which is largest in magnitude is equal to the number which is smallest in magnitude, which is nonsense. $\square$

We were able to draw such a nonsensical conclusion in the above illustration because we manipulated ∞ as a number, which it is not. Since $\pm\infty$ are not numbers, we will never encounter functions that have a value of $\pm\infty$ at some point. The question then remains, when do we encounter $\pm\infty$? In order to proceed further, we need to be more specific. There are two primary contexts in which the notion of infinity will be useful, and both are related to limits.

Consider the function

$$f(x) = 1/x.$$

If we look as $x \searrow 0$, the value of the function becomes larger and larger (increasing without bound), because the fraction is positive, and the magnitude of its denominator becomes smaller and smaller. On the other side, as $x \nearrow 0$, the values of the function become smaller and smaller (decreasing without bound), because the denominator is negative, and its magnitude is approaching 0. It is not possible for this function to have a (finite) limit from either side, as on the right it becomes larger than every finite number, and on the left it becomes smaller than any finite number. Not only does this function not have one-sided limits as $x \to 0$, the limits do not exist for very specific reasons - either because the function is increasing or decreasing without bound. In order to represent this using the concept of infinity, we write

$$\lim_{x \nearrow 0} \frac{1}{x} = -\infty \quad \text{and} \quad \lim_{x \searrow 0} \frac{1}{x} = \infty.$$

In general, when we write

$$\lim_{x \to x_0} f(x) = \infty$$

we are not saying that the limit exists; we are saying is that *the limit does not exist*, and it does not exist because the values of the function $f(x)$ grow without bound as $x \to x_0$. Similarly,

$$\lim_{x \to x_0} f(x) = -\infty$$

means that the limit as $x \to x_0$ of $f(x)$ does not exist because the function values decrease without bound (become arbitrarily large in magnitude, and negative in sign). Throughout this discussion one-sided limits will be particularly useful to us, because it is quite common that a function may approach ∞ from one side of a point, and $-\infty$ from the other side of the point, just as the function $f(x) = 1/x$ does.

Definition 3.4.2 (Infinite Limits)**.** We write

$$\lim_{x \to x_0} f(x) = \infty$$

if for every number $B > 0$ there exists a corresponding $\delta > 0$ such that for all x with $0 < |x - x_0| < \delta$ we have

$$f(x) > B.$$

Similarly, we write

$$\lim_{x \to x_0} f(x) = -\infty$$

if for every number $B > 0$ there exists a corresponding $\delta > 0$ such that for all x with $0 < |x - x_0| < \delta$ we have

$$f(x) < -B.$$

It is important to emphasize that the rules we previously had for combining limits do not apply to limits as $f(x) \to \pm\infty$, because these limits do not actually exist. If we try to apply the rules for combining finite limits, we will run into many nonsensical possibilities, such as $\infty - \infty$, ∞/∞, etc., which are meaningless, because ∞ is not a symbol to be manipulated as a number.

Limits where $f(x) \to \pm\infty$ have an important physical interpretation in terms of frequency response. For this reason, we will write f as $f(\omega)$, using ω to denote we have a function of frequency. The frequency response of a system gives the output of the system with respect to the frequency of the input signal. In general, a system responds differently to inputs coming in at different frequencies. The types of systems we are interesting in are *oscillating* systems, those in which some value is moving repetitively between a minimum and maximum. Oscillating systems are important in the fields of physics, mechanics, and electromagnetics. Examples are the simple harmonic oscillator (mass on a spring or pendulum), electrical circuits, electromagnetic waves, etc. Places where the (ideal or undamped) frequency response of a system approaches $\pm\infty$ (or in a real physical situation, becomes very large) are called *resonant frequencies*, and the phenomenon of the output of the system becoming very large in reaction to an input at a certain frequency is call *resonance*.

A simple physical example related to frequency response is pushing a child on a swing. If one pushes the swing every time the child moves all the way back, the swing will start moving rapidly, elevating the child higher and higher. If we look at the separation between each push, we can think of the pushes occurring with a certain frequency. Now what happens if we change that frequency, by pushing more frequently, or less frequently? If we start pushing the child in the middle of the motion of the swing, our input forces will act destructively rather than constructively, and not be very effective at swinging the child (just think about what would happen if you pushed the child in the middle of a swing, rather than at the end of it). Here the resonant frequency is the one at which we need to push the child in order for the swinging to occur effectively.

In the above example there are other factors to consider, such as dampening in the chains of the swing, and the physical setup of the swing. These impose practical limitations on the speed and height at which the swing could be pushed. For this reason the response of the system at the resonant frequency is not an *infinite* output, merely an output much larger than if the input force were at a different frequency. Only in an *ideal* case (such as a model which ignores certain physical phenomena) will the frequency response of the system ever be $\pm\infty$. However, if in a model we have an infinite frequency response, we know that in practice we will have a large response, unless there are significant dampening factors to control the resonance.

In electrical circuits resonance can be used for tuning signals. A circuit is setup so that its frequency response is much higher at a given frequency than others, so effectively only that frequency (and the narrow band of frequencies around it) is transmitted (a simple RLC circuit tunes signals in this way). Another example of resonance is in a laser cavity, which consists of two mirrors and a gain medium in between them. As light oscillates between the two mirrors, some frequencies undergo constructive interference, whereas others undergo destructive interference, which frequencies of light depending on the length of the optical cavity. The result is that certain frequencies are amplified greatly - the resonant frequencies - and others are not amplified. As a result, a beam of light with a very narrow frequency spectrum can be generated, which is called a laser. If we look at the frequency response of the laser cavity, we can see which frequencies resonate, and thus determine what frequency lasers can be generated with it. Of course there are other issues to consider, such as which frequencies are amplified by the gain medium, how to get enough energy into the cavity, and how much power can be sustained by the cavity before the mirrors are destroyed.

Just as resonance can be exploited for positive effects, it is also a hazard to avoid. In mechanical systems there is a worry about the system vibrating at the resonant frequency, until the forces become so great that the machine is destroyed. For this reason it is important to include dampening to curb the *infinite* (ideal) responses of the system, until they are controlled to the point where they are not self-destructive. A famed example of resonance as a hazard is the Tacoma Narrows bridge failure of 1940. Although this is not an example of forced resonance, it is a related phenomenon.

In addition to considering the behavior of a physical system as its output grows considerably large, it is also interesting to look at the behavior of a system after a consider amount of time has passed. The definition of considerable depends on the nature of the system, because it is not physically possible (or necessarily useful) for us to observe a system for an arbitrarily large amount of time. Unless a system is chaotic, over time it should fall into some sort of steady state behavior. This is what leads us to the second usage of the concept of infinity.

We can look at the behavior of a function (representing a physical system) after a considerable amount of time has passed by looking at what happens as the input (time) grows without bound; ie. we consider the limit as $t \to \infty$. If the independent variable is not time, we might look at limits as $x \to \pm\infty$. Even though a concept of *infinite* time is meaningless physically, a limit as $t \to \infty$ describes what *would* happen in a system, were it possible to let it continue on indefinitely. We call the behavior of a function as its input approaches $\pm\infty$ the asymptotic behavior of the function.

Definition 3.4.3 (Definition: Limits as $x \to \infty$ or $x \to -\infty$)**.** We write

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

if for every number $\varepsilon > 0$ there exists a corresponding number M such that for all x with $x > M$ we have

$$|f(x) - L| < \varepsilon.$$

Similarly, we write

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

if for every number $\varepsilon > 0$ there exists a corresponding number N such that for all x with $x < N$ we have

$$|f(x) - L| < \varepsilon.$$

The method for showing a limit exists as $x \to \pm\infty$ is very similar to showing that a limit exists at a point. The only difference is that rather than creating a δ interval around the point of interest to restrict the function values within the error tolerance of the limit value, we need to find a value M or N so that if the input values x are larger/smaller than M or N, the function values are within the error tolerance of the limit. Due to the remarkable similarities between these notions of the limit, it may not be surprising that all of the properties we stated previously for combining limits also hold for limits at $\pm\infty$. We will not focus on proofs of limits as $x \to \pm\infty$.

A function may exhibit many possible behaviors in the limits as its argument $x \to \pm\infty$. The function can either approach $\pm\infty$, meaning that it increases or decreases without bound, it can approach a constant value, or it might not approach any value. For instance,

$$\lim_{x\to\infty} x = \infty,$$

because of the output of a function is its input, so as the input grows without bound, so does the output. If we have a constant function, $f(x) = c$, then the input approaching ∞ is irrelevant, so we find

$$\lim_{x\to\infty} c = c.$$

Finally, we may encounter functions that oscillate with time, and never settle down to a single value, or increase/decrease without bound. Here we do not use specific notation to represent that the limit does not exist, we simply write

$$\lim_{x\to\infty} \cos(x) \text{ does not exist.}$$

Nevertheless, if we modulate the cosine function with a decreasing amplitude, then we will have a function that approaches 0.

$$\lim_{x\to\infty} x^{-2} \cos(x) = 0.$$

For any of the above limits, if we multiply the function by -1, the sign will simply change; if a function increases without bound, the negative of that function will decrease without bound. Thus,

$$\lim_{x\to\infty} -x = -\infty.$$

Similarly, if we look in the limit as $x \to -\infty$, the only thing we might have to worry about is a negative sign - otherwise the analysis is the same.

In the next section we will consider how to analyze functions and determine whether or not they approach $\pm\infty$ in certain limits.

3.5 Finding Infinite Limits: Order of Magnitude Analysis

In order to work with infinite limits it is helpful to introduce the set of *affinely extended real numbers*, which contains all of the real numbers as well as $\pm\infty$. Using set notation, we write

$$\mathbb{R}^* = \mathbb{R} \cup \{\infty, -\infty\}.$$

One way in which we encounter limits where

$$\lim_{x\to x_0} f(x) = \infty$$

is when $x_0 \in \mathbb{R}$, and f contains a division by 0 at the point x_0. In such a situation $f(x_0)$ is often undefined, but because the denominator of $f(x)$ approaches 0 as $x \to x_0$, the output of $f(x)$ grows without bound as $x \to x_0$.

Example 1. Find $\displaystyle\lim_{x\to -4} \frac{1}{(x+4)^4}$.

Solution In analyzing this function, the first thing to notice is that we have a fourth power in the denominator, $(x+4)^4$. Since $(x+4)$ is raised to an even power, the denominator is always non-negative (it is 0 when $x = -4$ and positive otherwise). Since the numerator is always positive as well, it means that this function is always non-negative. As $x \nearrow -4$ and $x \searrow 4$ the denominator approaches 0, so correspondingly the values of the function grow without bound (so the limit does not exist). We write

$$\lim_{x\to -4} \frac{1}{(x+4)^4} = \infty$$

to signify the function values grow without bound as $x \to -4$ from both sides.

Example 2. Find $\displaystyle\lim_{x\to 3} \frac{3-x}{(x-3)^4}$.

Solution The first thing for us to do is simplify this fraction as much as possible. If we factor -1 from the numerator, we find that

$$\lim_{x\to 3} \frac{3-x}{(x-3)^4} = \lim_{x\to 3} \frac{-(x-3)}{(x-3)^4} = \lim_{x\to 3} \frac{-1}{(x-3)^3}.$$

Note that above we were able to cancel the factor of $(x-3)$ because in looking at the limit as $x \to 3$, we know that $x \neq 3$, so the factor will never be 0 (we, of course, cannot divide by 0). In this case, unlike the previous one, we have an odd power in the denominator; that is, $(x-3)$ is raised to the third power. Because of this, we will have

$$(x-3)^3 < 0 \text{ when } x < 3 \quad \text{and} \quad (x-3)^3 > 0 \text{ when } x > 3.$$

Since the sign of the denominator is different depending on what side $x \to 3$ from, we should look at the one-sided limits rather than trying to calculate a two-sided limit directly (because they will have different signs, so unless they are both 0, the two-sided limit will not exist). As $x \nearrow 3$ the denominator is negative, so the entire fraction is positive (because there is a -1 in the numerator). As $x \searrow 3$ the denominator is always positive, so the entire function is negative. In both cases the denominator approaches 0, so we find that

$$\lim_{x\nearrow 3} \frac{3-x}{(x-3)^4} = \lim_{x\nearrow 3} \frac{-1}{(x-3)^3} = \infty \quad \text{and} \quad \lim_{x\searrow 3} \frac{3-x}{(x-3)^4} = \lim_{x\searrow 3} \frac{-1}{(x-3)^3} = -\infty.$$

Since our function has no consistent behavior as $x \to 3$, all we can say is that the limit does not exist.

Example 3. Find $\lim_{x\to 3} \dfrac{4x^2-9x-9}{x-3}$.

Solution As $x \to 3$, the denominator approaches 0, but so does the numerator. For any polynomial to have a zero at 3 means it has a factor of $(x-3)$. In fact,

$$4x^2-9x-9=(x-3)(4x+3).$$

Now we can cancel a factor of $(x-3)$ from numerator and denominator, to find that

$$\lim_{x\to 3}\frac{4x^2-9x-9}{x-2}=\lim_{x\to 3}(4x+3)=15.$$

Suppose that we have two functions, f and g, with

$$\lim_{x\to x_0}|f(x)|=\lim_{x\to x_0}|g(x)|=\infty,$$

where $x_0 \in \mathbb{R}^*$. What happens when we look at

$$\lim_{x\to x_0}\frac{f(x)}{g(x)}?$$

In this situation we cannot apply the quotient rule for limits, because

$$\lim_{x\to x_0} f(x) \quad \text{and} \quad \lim_{x\to x_0} g(x)$$

do not actually exist (recall that writing a limit is ∞ means that the limit doesn't exist precisely because the function values grow without bound). Nevertheless, just as when we were faced with a limit of a quotient where both the limits of the numerator and denominator were 0, we suspect that it is possible for the limit of this quotient to be finite. For instance, if we let

$$f(x)=g(x)=x,$$

then we have two functions both approaching ∞ as $x \to \infty$, yet the limit of their ratio is 1. Before delving into this problem in more generality, we will note that there is a simple trick for evaluating this type of limit when we are dealing with rational functions (a ratio of polynomials). If we divide by the highest power of x in the denominator we can evaluate the limit.

Example 4. Find $\lim_{x\to\infty} \dfrac{2x^2-1}{5x^2-x}$.

Solution As before, we want to multiply by a convenient choice of 1 in order to rewrite the expression into something we can manipulate. Since the highest power of x in the denominator is 2 (ie, x^2) we multiply by

$$\frac{1/x^2}{1/x^2},$$

in order to remove all terms in the denominator approaching $\pm\infty$, reducing the limit into something we can analyze (because considering something of the form ∞/∞ is meaningless). Doing so we find

$$\lim_{x\to\infty}\frac{2x^2-1}{5x^2-x}=\lim_{x\to\infty}\frac{2x^2-1}{5x^2-x}\cdot\frac{1/x^2}{1/x^2}=\lim_{x\to\infty}\frac{2-1/x^2}{5-1/x}=\frac{2-0}{5-0}=\frac{2}{5}.$$

Example 5. Find $\lim_{x\to\infty} \dfrac{2x-1}{5x^2-x}$.

Solution Using the same trick as before we find

$$\lim_{x\to\infty} \frac{2x-1}{5x^2-x} = \lim_{x\to\infty} \frac{2x-1}{5x^2-x} \cdot \frac{1/x^2}{1/x^2} = \lim_{x\to\infty} \frac{2/x-1/x^2}{5-1/x} = \frac{0-0}{5-0} = 0.$$

Example 6. Find $\lim_{x\to-\infty} \dfrac{2x^4-x^3}{5x^3-x}$.

Solution Yet again we perform the same trick, finding that

$$\lim_{x\to-\infty} \frac{2x^4-x^3}{5x^3-x} = \lim_{x\to-\infty} \frac{2x^4-x^3}{5x^3-x} \cdot \frac{1/x^3}{1/x^3} = \lim_{x\to-\infty} \frac{2x-1}{5-1/x^2} = -\infty.$$

In the above examples, it was the relative order of magnitude of the polynomials (the highest power of x) in the numerator and denominator that determined the limit. This is true in general. If the numerator has a higher order of magnitude than the denominator, then the rational function will approach $\pm\infty$. If the numerator has the same order of magnitude as the denominator, then then limit will be the ratio of the coefficients of the highest order terms in the numerator and denominator. Finally, if the numerator has a lower order of magnitude than the denominator, the limit will approach 0.

The key to evaluating limits of $f(x)/g(x)$ in the more general case is extending this idea of order of magnitude to functions other than polynomials. We will call this order of magnitude analysis of growth and decay. The basic idea is that when we look at functions that approach ∞, we can distinguish functions based on the *rate* at which they approach ∞. For a simple example, think of the functions $f(x) = x$ and $g(x) = x^2$. As x gets larger, both functions grow without bound, but $g(x)$ grows much more rapidly than $f(x)$. This distinction is not difficult to make because they functions are so familiar. The goal is to extend this classification to many more functions, in order to make evaluating limits much easier. We formalize the idea of rate of growth with the following definition.

Definition 3.5.1 (Relative Rates of Growth)**.** Let $f, g : \mathbb{R} \to \mathbb{R}$, $x_0 \in \mathbb{R}^*$, and

$$\lim_{x\to x_0} |f(x)| = \lim_{x\to x_0} |g(x)| = \infty.$$

We say f approaches ∞ *on a higher order of magnitude* than g if

$$\lim_{x\to x_0} \left|\frac{f(x)}{g(x)}\right| = \infty.$$

We say f approaches ∞ *on a lower order of magnitude* than g if

$$\lim_{x\to x_0} \left|\frac{f(x)}{g(x)}\right| = 0.$$

We say that f and g approach ∞ *on the same order of magnitude* if

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = a, \quad a \in \mathbb{R}, a \neq 0.$$

Now our goal is to try to put familiar functions into each of these categories. In accordance with the above examples, we already have distinctions for polynomial functions. In general, l'Hôpital's rule is extremely useful for making these classifications, but we do not yet have the tool of differentiation at our disposal. At the moment let us restrict ourselves to three classes of functions: exponentials, power functions, and logarithms.

Theorem 3.5.1 (Relative Rates of Growth)**.** *As $x \to \infty$, the magnitude of exponential, power, and logarithmic functions approach ∞. Moreover, the order of magnitude with which they approach ∞ is given in the following order, with exponential functions approaching the fastest, and logarithmic functions approaching the slowest.*

1. *Exponentials of the form a^x, where $a > 0, a \neq 1$. If $a > b$, then a^x approaches ∞ faster than b^x as $x \to \infty$.*

2. *Power functions, x^n, $n \in \mathbb{N}$, If $n > m$, then x^n approaches ∞ faster than x^m.*

3. *Logarithmic Functions, $\log_a(x)$, $a > 0, a \neq 1$. All logarithmic functions approach ∞ at the same rate, because $\log_a(x)/\log_b(x) = \ln(b)/\ln(a)$.*

If any of the above functions is multiplied by a nonzero constant, it does not change the relative order of magnitude at which the magnitude of the function approaches ∞.

Using this information, we can immediately evaluate limits such as

$$\lim_{x\to\infty} \frac{e^x}{\ln(x)} = \infty \quad \text{and} \quad \lim_{x\to\infty} \frac{\ln(x)}{e^x} = 0.$$

There is one additional fact we need in order to really utilize this tool. The key thing is that when we look at sums and differences of these functions, it is only the term that grows that fastest that matters - this is called the *dominant term*.

Definition 3.5.2 (Dominant Term)**.** Suppose f can be written as a linear combination of functions $\{f_1, f_2, \ldots, f_n\}$, and

$$\lim_{x\to x_0} |f(x)| = \infty,$$

where $x_0 \in \mathbb{R}^*$. The *dominant term* of f is the function f_i which approaches ∞ on the highest order of magnitude, denoted by $\hat{f}$.

Recall that a linear combination of functions is simply a sum of constant multiples of the functions. For instance, a polynomial is simply a linear combination of power functions.

Theorem 3.5.2 (Dominant Terms)**.** *Consider $f(x)$ and $g(x)$, with*

$$\lim_{x\to x_0} |f(x)| = \lim_{x\to x_0} |g(x)| = \infty.$$

If $\hat{f}$ and $\hat{g}$ are the dominant terms of f and g respectively, then

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{\hat{f}(x)}{\hat{g}(x)}.$$

Now for any given limit we simply need to look at the dominant terms. For instance,

$$\lim_{x\to\infty} \frac{e^x - 12x^2 + x}{x^4 + 1} = \lim_{x\to\infty} \frac{e^x}{x^4} = \infty.$$

We can also apply this same idea to limits in which both the numerator and denominator approach 0.

Definition 3.5.3 (Relative Rates of Decay)**.** Let $f, g : \mathbb{R} \to \mathbb{R}$, $x_0 \in \mathbb{R}^*$, and

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0.$$

We say f approaches 0 *on a higher order of magnitude* than g if

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0.$$

We say f approaches 0 *on a lower order of magnitude* than g if

$$\lim_{x \to x_0} \left| \frac{f(x)}{g(x)} \right| = \infty.$$

We say that f and g approach 0 *on the same order of magnitude* if

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = a, \quad a \in \mathbb{R}, a \neq 0.$$

The most simple case is to look at the reciprocals of exponential and power functions. The relative rates of decay of such functions follow immediately from the relative rates at which they grow. For instance, consider the functions e^{-x} and x^{-2}. As $x \to \infty$, e^x grows faster than x^2, so it follows that its reciprocal decays faster, because in viewing e^{-x} as $1/e^x$ and x^{-2} as $1/x^2$, we see that the magnitude of the denominator of the first fraction is growing the fastest. Thus, the magnitude of the overall fraction is decaying the fastest. Another way to evaluate limits involving these reciprocal functions is simply to rewrite them in terms of the growing functions.

Example 7. Evaluate $\lim\limits_{x \to \infty} \dfrac{e^{-x} + 1}{x^{-2}}$.

Solution In order evaluate this limit we simply rewrite the constituent functions in terms of their reciprocals.

$$\frac{e^{-x} + 1}{x^{-2}} = \frac{e^{-x}}{x^{-2}} + \frac{1}{x^{-2}} = \frac{x^2}{e^x} + x^2$$

Now we simply evaluate the limit, using our knowledge of the relative rates of growth of the given functions, and see that

$$\lim_{x \to \infty} (\frac{x^2}{e^x} + x^2) = \infty.$$

In some more complicated cases we can also evaluate infinite limits by using substitution.

Example 8. Find $\lim\limits_{x \to \infty} \sin(\dfrac{1}{x})$

Solution Here we will use the fact that as $x \to \infty$ we have $1/x \to 0$, and introduce the variable $t = 1/x$. Thus,

$$\lim_{x \to \infty} \sin(\frac{1}{x}) = \lim_{t \to 0^+} \sin(t) = 0$$

The reason we need to write this as a one-sided limit is because $x \to \infty$ from only one side.

3.6 Continuity

In this section we will explore a very useful class of functions - continuous functions.

Definition 3.6.1 (Continuity at a Point)**.** A function $f : X \to \mathbb{R}$, $X \subset \mathbb{R}$ is continuous at a point $x_0 \in X$ if for every $\varepsilon > 0$, there exists $\delta > 0$ so that for all $x \in X$ with $|x - x_0| < \delta$ we have

$$|f(x) - f(x_0)| < \varepsilon.$$

This definition is written in terms of ε and δ, just like the formal definition of a limit. Essentially, this definition of continuity is the same as a function having a limit at a given point, except that here the limit value is the value of the function. However, this definition includes functions mapping $X \to \mathbb{R}$, in which these notions are not strictly equivalent[1]. This issue is not really related to our goals at the moment, so let's restrict ourselves to functions defined on intervals, where these notions *are* equivalent.

Theorem 3.6.1 (Continuity at a Point)**.** *Suppose $f : I \to \mathbb{R}$, I an interval. Let $x_0 \in I$. There are three situations:*

1. *x_0 is an interior point of I. f is continuous at x_0 iff $\lim\limits_{x \to x_0} f(x) = f(x_0)$.*
2. *x_0 is a left endpoint of I. f is continuous at x_0 iff $\lim\limits_{x \searrow x_0} f(x) = f(x_0)$.*
3. *x_0 is a right endpoint of I. f is continuous at x_0 iff $\lim\limits_{x \nearrow x_0} f(x) = f(x_0)$.*

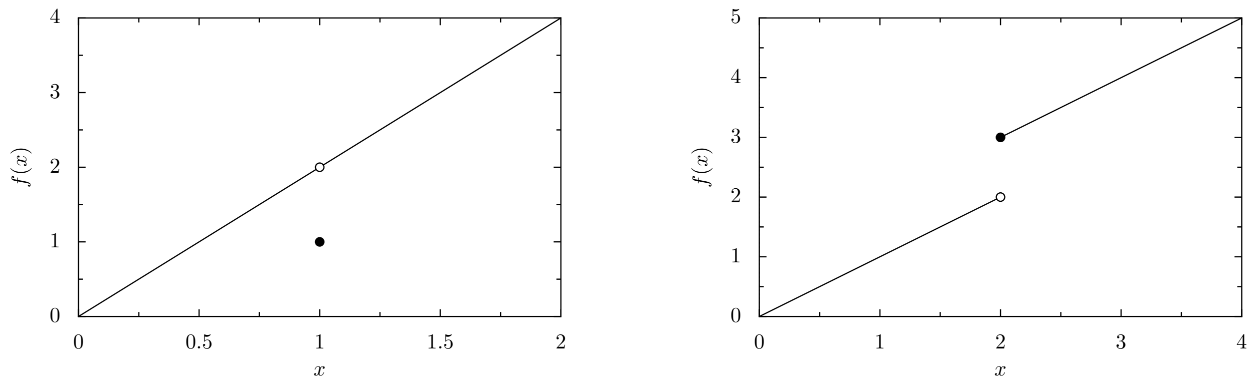
This theorem states that for functions defined on intervals, continuity is equivalent to having a limit that is equal to the value of the function, if we are careful to use one-sided limits for endpoints. Although continuity is a local property, defined in terms of points, we are really more interested in a global notion of continuity - functions for which every point in their domain they are continuous.

Definition 3.6.2 (Continuity)**.** A function f is *continuous* if it is continuous at every point in its domain.

In essence, continuity is related to the notion of how fast a function can change. If a function is continuous at a point, it means that we can restrict nearby function values to be as close as we want, simply by restricting ourselves to near enough input values. This property most clearly manifests itself over intervals. If a function is continuous on an interval, then it cannot contain any holes or jumps within the interval; if it contained any holes or jumps in an interval, then there would be some points where we cannot restrict how much the function changes by looking at a small enough interval, because no matter how small the interval, the function would change by the height of the jump or the distance between the removed point and the rest of the function. Stated in other words, a function is *discontinuous* at any point where it has a jump or hole (see figure 3.5).

From the perspective of limits, it is clear to see that for a hole the limit value does not equal the function value. For a finite jump, the one-sided limits are different, so the limit does not exist, and thus the function cannot be continuous there. The fact that a function that is continuous on an interval cannot contain any holes or jumps has an immediate consequence, called the *intermediate value property*.

[1]Consider a function $f : \mathbb{N} \to \mathbb{R}$. Such a function is continuous at all points in its domain, if we simply choose $\delta \leq 1$. However, this notion is completely unrelated to limits, as it doesn't make sense to look at a limit on the domain of natural numbers - there are no x values *approaching* any point in the domain to consider.



(a) Discontinuity at a hole.

(b) Discontinuity at a jump.

Figure 3.5: Example discontinuities.

Theorem 3.6.2 (The Intermediate Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous. For any value y_0 between $f(a)$ and $f(b)$, there exists some $c \in [a, b]$ such that $f(c) = y_0$.*

We use the statement between $f(a)$ and $f(b)$ because it is not clear which one is larger than the other, if in fact $f(a) \neq f(b)$. If we have $f(a) = f(b)$ then the above statement tells us nothing, because there are no values between $f(a)$ and $f(b)$, and we already know our function takes on the value $f(a) = f(b)$ at the endpoints of the interval.

Why is this property true? Suppose that there was some value between $f(a)$ and $f(b)$ skipped by the function. How could f skip such a value? Either by having that value removed at a point of the function, or by having a jump. Either of these methods would require the function have a discontinuity, which it doesn't, by hypothesis.

The intermediate value property is particularly useful for finding roots of a function, insofar that if we have a function $f(x)$, and we know that for some a we have $f(a) < 0$ and for some b we have $f(b) > 0$, then there is some c between a and b where $f(c) = 0$.

Continuity also has implications for the invertibility of a function. Consider a continuous function $f : I \to \mathbb{R}$. Since f is a continuous function defined on an interval, it cannot have any holes or jumps. In order for f to be invertible then, it follows that f must be strictly monotone (always increasing or always decreasing). Clearly if f were constant for some period of time it would not be invertible. However, we can go further. If f is increasing over some subinterval (or decreasing for that matter) unless it continues to increase, it will begin to repeat output values, because it cannot jump over them. For instance, suppose $f(x)$ increases from 1 to 5 over some subinterval of I, and then later on has a value of 2. The only way it could get back down to 2 is by moving through all of the values 5 to 2, by the intermediate value theorem. Such a function would have multiple inputs with the same output, and hence not be invertible. Thus, although strict monotonicity is normally sufficient (but not necessary) for invertibility, in the case of continuous functions (defined on intervals) it also becomes necessary.

Theorem 3.6.3 (Invertibility of Continuous Functions)**.** *Let $f : I \to \mathbb{R}$ be continuous. It follows f is invertible if and only if f is strictly monotone.*

Just as we had rules for combining limits, we can also combine continuous functions to create continuous functions.

Theorem 3.6.4 (Properties of Continuous Functions)**.** *Suppose $g(x)$ is continuous at x_0*

1. *If f is continuous at x_0, then $f+g$ is continuous at x_0.*

2. *If f is continuous at x_0, then $f \cdot g$ is continuous at x_0.*

3. *If f is continuous at x_0, $g(x_0) \neq 0$, then f/g is continuous at x_0.*

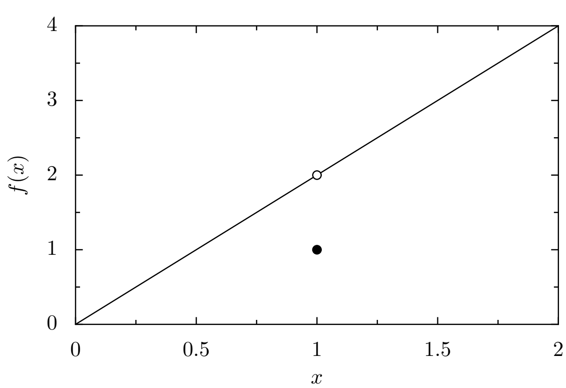
4. *If f is continuous at $g(x_0)$, then $f \circ g$ is continuous at x_0.*

As before, these rules will only be useful if we have some knowledge about basic continuous functions. We will be able to use these functions as building blocks to verify the continuity of more complex functions. We already know that for any polynomial function, the limit is the same as the value of the polynomial at every point, so it follows that all polynomial functions are continuous on $\mathbb{R}$. Sine, cosine, and the exponential function are also continuous on all of $\mathbb{R}$. In fact, most functions that we encounter are continuous, as long as we restrict them to the proper domain. For instance, $\ln(x)$ is continuous for $x > 0$, which is the entire domain over which it is defined. It follows from the continuity of these basic functions, and the above properties, that much more complicated functions such as

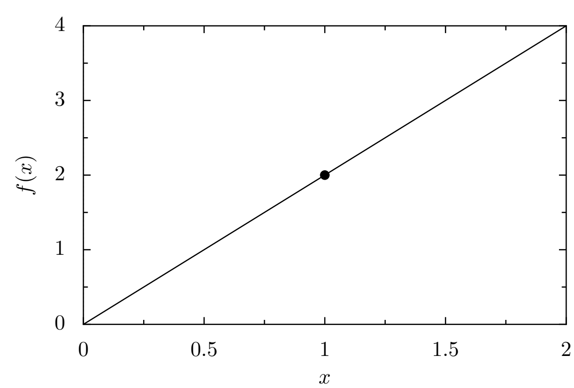
$$f(x) = \frac{x\sin(x)}{x^2+2}$$

are continuous (when combining functions in this way one must be careful to avoid division by 0, which we do here because $(x^2+2) > 0$ for all x).

At this point let us turn our attention to discontinuities, the points where functions fail to be continuous. In this discussion, unless otherwise stated we will be looking at functions defined on $\mathbb{R}$. There are two classes of discontinuities - removable and nonremovable. If we have a *removable* discontinuity, it means that we can get remove the discontinuity just by redefining the function at the point of discontinuity. This corresponds to a function which is not continuous at a given point, but that *ought to be*. Think about a hole removed from a function and put in a different place. All we need to do is fill up that hole, and the discontinuity will be removed (see figure 3.6). Thus, holes in a function are an example of removable discontinuities.



(a) Removable discontinuity.



(b) Discontinuity removed.

Figure 3.6: Removing a removable discontinuity.

In contrast, nonremovable discontinuities are places where a function is broken so badly that there is no *simple* way to fix the function. Think about the function $|x|/x$. It is discontinuous at $x = 0$, because it is not defined. Nevertheless, no matter what value we try to assign to the function at that point, it will still be discontinuous from one side or the other. In this case there is a jump in the function that cannot be removed simply by redefining a single point. In order to remove this discontinuity from the function, we would need to redefine it over an entire interval, significantly changing the characteristics of the function. Any discontinuity we cannot remove by redefining the function at a single point is called a *nonremovable* discontinuity. In general, if a function is discontinuous because the function value is not equal to the limit, we simply redefine the function value to the limit value and remove the discontinuity. If there is a place in the function where a limit does not exist, then simply redefining a single value will not make the limit exist, so the discontinuity is nonremovable.

Example 1. Is $\sin(1/x)$ continuous over the real numbers? If not, does it have any discontinuities which can be redefined to make it continuous over the real numbers?

Solution Since there is a division by 0 at $x = 0$, this function is not defined at $x = 0$, so it is discontinuous over a domain of real numbers (a limit value cannot be equal to a function value that does not exist). As $x \to 0$ the function oscillates faster and faster between 1 and -1, so it does not approach any single value. Thus, there is no way that we could define $f(0)$ in order to make the function continuous over the domain of real numbers. However, if the domain is restricted to the real numbers minus 0 this function is continuous.

Example 2. Is $x \cdot \sin(1/x)$ continuous over the real numbers? If not, does it have any discontinuities which can be redefined to make it continuous over the real numbers?

Solution Since this function is undefined at $x = 0$, it is discontinuous over a domain of real numbers. However, as $x \to 0$ despite the fact that the sine function oscillates faster and faster between 1 and -1, it is multiplied by x, so it approaches 0. Thus, if we define $f(0) = 0$, then it will be a continuous function over the domain of reals.

Example 3. Is $\sin(x)/x$ continuous over the real numbers? If not, does it have any discontinuities which can be redefined to make it continuous over the real numbers?

Solution Once again this function is undefined at $x = 0$, and thus is discontinuous there. However, we have previously shown that

$$\lim_{x \to 0} \frac{\sin(x)}{x} = 1,$$

so if we define $f(0) = 1$, then we have a continuous function over the real numbers.

Example 4. Where is $\dfrac{x^2 + x - 6}{x^2 - 4}$ discontinuous? Are these discontinuities removable?

Solution We have a quotient of two continuous functions, so this rational function is continuous wherever we do not have division by 0. Since there is division by 0 when $x = \pm 2$, we have two discontinuities. Are either of them removable? In order to find out we need to look at the limits as $x \to 2$ and $x \to -2$. If we look as $x \to 2$ we find

$$\lim_{x \to 2} \frac{x^2 + x - 6}{x^2 - 4} = \lim_{x \to 2} \frac{(x+3)(x-2)}{(x+2)(x-2)} = \lim_{x \to 2} \frac{(x+3)}{(x+2)} = \frac{5}{4}.$$

By redefining $f(2) = 5/4$, we remove this discontinuity. Looking as $x \to -2$ we find

$$\lim_{x \searrow -2} \frac{x^2 + x - 6}{x^2 - 4} = \infty,$$

so we find that this discontinuity is nonremovable (no matter how we redefine the function this limit will not exist).

3.7 Sequences

Sequences are an extremely powerful, and deceptively simple tool in calculus. At the base of it, a sequence is just a function.

Definition 3.7.1 (Sequence)**.** A function $f : \mathbb{N} \to \mathbb{R}$ is called a *sequence* of real numbers.

Given that we've already spent all of this time working with functions that have the real numbers as their domain, it almost seems a step back to be working with a function on the natural numbers. What do the natural numbers have to offer that the real numbers don't? After all, the natural numbers are only a subset (and a very small one at that) of real numbers. However, when we move from the natural numbers up the rationals or real numbers, we lose something very important, a notion of order. This is not to say that the rational and real numbers are not ordered, as they certainly are. Given two (or any number for that matter) different real numbers, we can order them from smallest to largest. However, given two different real numbers we can always also find another real number in between. In some situations this is wonderful - it is from this fact that we are able to approximate numbers to an arbitrary degree of accuracy. However, in some situations it is also detrimental, and that is why we move back to the natural numbers for sequences.

The benefit we gain from working on the natural numbers is that given a natural number there is a logical *next* natural number - the given number plus 1. Although a sequence is defined simply as a function on the natural numbers, that's generally not the way it is visualized or thought of. We generally think of each natural number as denoting a term in the sequence, and that as we move up to higher numbers, we are moving farther in the sequence. Because we are working on the natural numbers, after each term in the sequence, there is a logical *next* term - after the 3^{rd} comes the for 4^{th}.

We generally don't look at sequences where the terms just continue along randomly. As we move further out in the sequence, we'd like for the terms to progress towards something. This is where the idea of the limit comes into play. Just as for a function on the real numbers, we look at the limit as the input variable $n \to \infty$. If this limit exists, we say the sequence *converges.* If it does not exist, we say the sequence *diverges.* When a sequence converges, the terms at the end or *tail* of the sequence *pile up* around some value, which is the limit of the sequence. When the terms don't *pile up* around a given value, we have a sequence that diverges.

In accordance with this alternative visualization of a sequence, we generally do not use function notation for a sequence. Instead, we use the notation $\{a_n\}_n$, to denote a sequence $a_1, a_2, \ldots$, where the i^{th} term is a_i. For example, the sequence

$$\{a_n\}_n = \{1/n\}_n$$

is a sequence with $a_1 = 1$, $a_2 = 1/2$, $a_3 = 1/3$ and so on. We could also write the same sequence as

$$\{a_n\}_n = \{1, 1/2, 1/3, 1/4, \ldots\}_n.$$

There is no standard convention for how many terms one should write, simply that one should write enough terms so that the pattern of the sequence is evident. Note that just like a regular function, a sequence need not have a specific pattern. We could construct a sequence where the terms are numbered as days, starting from 100 million years ago, where the output is the mean temperature of that day. Note that a sequence's terms must continue on forever, but if the universe lasted forever, then in theory we could continue the terms of such a sequence indefinitely.

Although we stated above that the convergence of a sequence is exactly the same as whether or not a limit as $n \to \infty$ exists, it is helpful to restate convergence in the notation of a sequence.

Definition 3.7.2 (Convergent Sequence)**.** A sequence $\{a_n\}_n$ is said to *converge* to a limit L if for every $\varepsilon > 0$ there exists an N so that for all $n \geq N$ we have

$$|a_n - L| < \varepsilon.$$

A sequence is *convergent* if it converges to a limit L, for some L. If a sequence is not convergent, it is said to be *divergent*.

According to the above definition, in order for a sequence to converge, for any error tolerance we need to be able to find a point far enough out in the sequence where all of the later terms are within the error tolerance of the limit.

Example 1. Does the sequence $\{1/n\}_n$ converge? If so, to what value?

Solution Simply in terms of real numbers, we know that

$$\lim_{x\to\infty} \frac{1}{x} = 0.$$

Similarly, as n gets larger, the terms $1/n$ become arbitrarily close to 0. Thus, this sequence converges to a limit 0, and we write

$$\{1/n\}_n \to 0.$$

Example 2. Does the sequence $\{1, -1, 1, -1, \ldots\}_n$ converge? If so, to what value?

Solution This sequence simply alternates between the two terms 1 and -1. No matter how far out in the sequence we go, it does not settle on a single value, so the limit does not exist. Thus, the sequence diverges.

Since sequences are really just functions with a domain of the natural numbers, all of the rules we had for combining limits of functions hold equally well for sequences.

Theorem 3.7.1 (Rules for Combining Sequences)**.** *Suppose* $\{a_n\}_n \to L$ *and* $\{b_n\}_n \to M$*, where* $L, M \in \mathbb{R}$*. The following results are true.*

1. *Constant Multiple Rule:* $\{k \cdot a_n\}_n \to k \cdot L, \quad k \in \mathbb{R}$
2. *Sum Rule:* $\{a_n + b_n\}_n \to L + M$
3. *Product Rule:* $\{a_n \cdot b_n\}_n \to L \cdot M$
4. *Quotient Rule:* $\{a_n/b_n\}_n \to \frac{L}{M}, \quad M \neq 0, b_i \neq 0$ *for all* i

It turns out that all of the terms in the sequence *piling up* around a single value (convergence of a sequence) is equivalent to another condition - that all of terms in the *tail* of the sequence become arbitrarily close to each other. If all of the terms of the sequence begin to pile together like this, it follows then that there must be something they are piling up around (at least in a complete space, which the real numbers are). A sequence with this property is called a *Cauchy sequence*, and this equivalence is called the *Cauchy criterion for convergence*.

Definition 3.7.3 (Cauchy Sequence)**.** A sequence $\{a_n\}_n$ is called *Cauchy* if for every ε there exists an N so that for all $n, m \geq N$, we have

$$|a_n - a_m| < \varepsilon.$$

Theorem 3.7.2 (Cauchy Criterion for Convergence)**.** *A sequence* $\{a_n\}_n$ *of real numbers converges iff it is a Cauchy Sequence.*

Rather than trying to find an explicit limit, or show that a sequence is Cauchy, it is sometimes easier to show a sequences satisfies different properties, which imply that it must converge. We start by looking at monotone sequences, which are simply monotone functions defined on the natural numbers.

Definition 3.7.4 (Monotone Sequence)**.** A *monotone sequence* is a sequence that is either increasing or decreasing.

- A sequence is *increasing* if $a_{n+1} \geq a_n$ for all n.
- A sequence is *decreasing* if $a_{n+1} \leq a_n$ for all n.

Written in English, an increasing sequence is one in which each term is no smaller than the terms before it. A decreasing sequence is one in which each term is no larger than the terms before it. Monotonicity in itself is not enough to imply that a sequence converges. Take the sequence $\{n\}_n$. It is an increasing sequence, but it diverges to ∞. What if we have a monotone function that is bounded? Let's consider an increasing sequence which is bounded above. For each term in the sequence, it must be larger than the terms before it, but it must also be smaller than its *least upper bound* (the smallest number M that serves as an upper bound). As the terms of the sequence continue to grow (or at least not decrease) these two bounds become closer and closer together, until there is essential nowhere for the terms to go but bunch up below the least upper bound. Thus, such a sequence converges to its least upper bound. Similarly, a decreasing bounded sequence converges to its greatest lower bound.

Theorem 3.7.3 (Convergence of Monotone Sequences)**.** *A monotone sequence* $\{a_n\}_n$ *converges iff it is bounded. If* $\{a_n\}_n$ *is increasing, it converges to its least upper bound. If* $\{a_n\}_n$ *is decreasing, it converges to its greatest lower bound.*

For monotone sequences, it is usually *much* easier to show they are bounded than to show they converge. The above result then implies that such a sequence converges. We will use this important fact later on with the bisection method.

There is a lot of theory around sequences, which we will not build up at this point. In the next section we will use sequences in order to finally define the irrational and real numbers. Later on we will use sequences in order to evaluate infinite sums, and we will look at sequences of functions, rather than just numbers. This will allow us to rewrite a number of common functions in terms of power series, or *infinite degree polynomials.*

3.8 Irrational Numbers

With the tool of sequences at our disposal, we are finally able to answer the question as to what the real numbers are. Since we are working from a perspective where we do not yet know what the real numbers are, it would be wrong for us to begin by working with real numbers (so far we've taken the existence and usage of real numbers on faith, an unfortunate but sometimes necessary part of learning mathematics). Thus, our tool will be *sequences of rational numbers*.

Definition 3.8.1 (Sequence of Rational Numbers)**.** A function $f : \mathbb{N} \to \mathbb{Q}$ is called a *sequence* of rational numbers.

The only difference between this definition and that of a sequence of real numbers is that now our range is rational numbers - a set of numbers we are comfortable with and understand. A key observation for all rational numbers is that when we write them in terms of a decimal expansion, we either have a finite decimal expansion, or an infinite expansion with a repeating pattern. For instance,

$$\frac{1}{5} = 0.2,$$

a finite pattern, and

$$\frac{8}{11} = 0.72727272\ldots,$$

an infinite repeating decimal pattern. The fact of the matter is, for *any* rational number, we will run into one of the above types of decimal expansions.

The problem is, even the ancient Greeks knew, that in a sense, the rational numbers are just not *enough* numbers. One of the simplest examples lies in the length of the diagonal of a square. If a square has sides of length 1, a nice simple number, the length of the diagonal is $\sqrt{2}$. How can we determine whether or not this is a rational number? Well, if $\sqrt{2}$ is a rational number, then we should be able to write it with either a finite or infinite repeating decimal expansion. What if we aren't able to find such an expansion. Does it then follow that $\sqrt{2}$ is irrational (not rational)? To show $\sqrt{2}$ is irrational, we need to be certain that its decimal expansion is infinite, and has no repeating pattern. Suppose we look to 100,000 terms, and see no pattern. Can we be convinced? Isn't it very possible that after a very long time the decimal expansion finally repeats, and we just haven't gone out far enough to see that? In fact, no matter how far we go out in a decimal expansion, we can never be certain it doesn't repeat. Thus, a finite number of terms in the decimal expansion of a number will *never* be enough to show that it is irrational. Fortunately, hope is not lost, because it *is* possible to show that a number cannot be rational - we just need a different approach.

Claim 3.8.1. *$\sqrt{2}$ is an irrational number.*

Proof. Suppose there exists a rational number p so that

$$p^2 = 2.$$

It follows that we can write $p = m/n$, where m, n are integers, and not both are even; if both m, n are even, then we can divide a factor of 2 in the numerator and denominator, to find new integers whose ratio represents p. This is a process we could continue indefinitely until both the numerator and denominator did not contain a factor of 2. Let's proceed supposing we have gone through this process, and found an appropriate m and n. From the first equation we can write

$$m^2 = 2n^2,$$

which implies that m^2 is even, because it contains a factor of 2. This implies now that m is even, because if m were odd then m^2 would also be odd. Given that m is even, it must contain a factor of 2, so therefore m^2 contains a factor of $2^2 = 4$. Since then the left side of the equation is divisible by 4, so must the right, which implies the term n^2 is even, because it must contain a factor of 2. Using the same logic as before, n must be even. We have now shown that both n and m are even, containing a factor of 2 each, but this is contrary to our choice of m and n, which was such that only one of them contained a factor of 2. It follows that there is no rational number p so that

$$p^2 = 2.$$

$\square$

From the above argument we know that $\sqrt{2}$ is not a rational number. Thus, it is impossible to write it using either a finite or infinite repeating decimal pattern. If we can't express $\sqrt{2}$ in this way, then how could we express it - with an infinite, nonrepeating decimal expansion? What would such an expansion be anyway? Since it goes on forever, and never repeats, there is no possible way for us to write it down. We can't write it down explicitly anyway.

Sequences are the key to solving this problem. For any decimal expansion, we can look at a sequence of rational numbers that converges to that decimal expansion. Take the example above of 8/11. We could just simply write a constant sequence

$$\{8/11\}_n$$

which obviously converges to the desired value. Alternatively, we could look at a sequence of numbers where each term gets closer and closer to the desired value, by making each term pick up an additional decimal place. We would write a sequence of the form

$$\{0.7, 0.72, 0.727, 0.7272, 0.72727, \ldots\}.$$

Now we have a sequence that does not actually contain the term 8/11, but converges to 8/11, because if we look in the limit as $n \to \infty$, we will pick up all of the infinite decimal expansion. In the same way, we can construct a sequence of rational numbers that converges to any conceivable decimal expansion, simply by writing the sequence with one additional decimal place each time. In this way, for any decimal expansion (be it infinite and nonrepeating), there exists a corresponding sequence that *converges* to it. The term *converges* is being used very loosely here, because in order for a sequence to converge, it must have a limit that is in its range, yet we can construct sequences of decimal expansions where the *number* that we are approximating with the sequence of rational numbers is in fact not a rational number. This is the sense in which the rational numbers are lacking - we have Cauchy sequences of rational numbers that don't converge, because there is no rational number for them to converge to. In a sense, by the nature of a Cauchy sequence it *should* converge, so in this way the rational numbers are lacking the points where the sequences should converge to. This is how we define the real numbers.

Definition 3.8.2. The real numbers $\mathbb{R}$ are the set of all limits of Cauchy sequences of rational numbers.

By this definition the real numbers contain all of the rational numbers (just by looking at constant sequences), plus all of the *gaps* in between them, where our Cauchy sequences *should* converge (these are the numbers with infinite, nonrepeating decimal patterns). If we construct a Cauchy sequence of rational numbers, then it has a limit, represented by some real number. By definition, any real number can be represented as the limit of at least one Cauchy sequence of rational numbers (but we can easily construct an infinite number of such sequences). Fortunately, the set of real numbers turns out to be *enough*. For any Cauchy sequence of real numbers, the limit is also a real number. This property is called *completeness* in that for every Cauchy sequence of members of the set, there is a corresponding limit within the set.

To show something is a real number, we need to show there is a Cauchy sequence of rational numbers converging to it, generally by finding a recursive method for constructing such a sequence. In this way, even if we are unable to explicitly write out a decimal expansion for a real number, we can write it out with as much accuracy as desired (in approximating by a rational number), simply by following the algorithm far enough. One such method of doing so for $\sqrt{2}$ (and many other numbers) is the bisection method, which we will explore in the next section. For other irrational numbers, there are other methods. For now, let us suffice ourselves to say that there are methods of doing so, and we will look at the exact methods a bit later.

The above discussion gives a means of performing arithmetic operations with real numbers. Just imagine trying to add two numbers with infinite decimal expansions together - where does one start adding from? It should be evident we can't simply work with real numbers in the same way we work with rationals. The answer to performing arithmetic operations with real numbers is to simply perform the associated operations with sequences converging to them (where the new sequence will converge to the result). For instance, in order to to find the real number π^2, we first start with a sequence $\{a_n\}_n \to \pi$. We find then that

$$\{a_n \cdot a_n\}_n \to \pi^2,$$

which gives meaning to the symbol π^2. In order to work with π^2 numerically, we simply go far enough in this sequence, and use an approximation for it by a rational number. We need to be a little bit more careful when we try to work with exponentiation. Suppose we are interested in finding $2^{\sqrt{2}}$. We start with a sequence $\{b_n\}_n \to \sqrt{2}$, and then we look at the sequence

$$\{2^{b_n}\}_n \to 2^{\sqrt{2}}.$$

The problem here is that along the way we might (and probably will) run into more irrational numbers. Since we're now dealing with a sequence of irrational numbers, we will need to look at sequences converging to each term. This is quite a mess we've gotten ourselves into! In order to avoid this mess, here is one possibility. We could start with a decimal expansion for $\sqrt{2}$, say 8 decimal places long,

$$\sqrt{2} \approx 1.41421356.$$

Now we could just look at

$$2^{1.41421356}$$

and find a single sequence converging to it, using some algorithm (like the bisection method). While in theory we can work with exponentials in this way, we will find a much simpler means of doing so in working with power series.

It's worthwhile to take a moment to look at two of the most commonly used irrational numbers at this point, e and π.

Claim 3.8.2. *The sequence $\{(1 + 1/n)^n\}_n$ converges, and its limit is defined as e.*

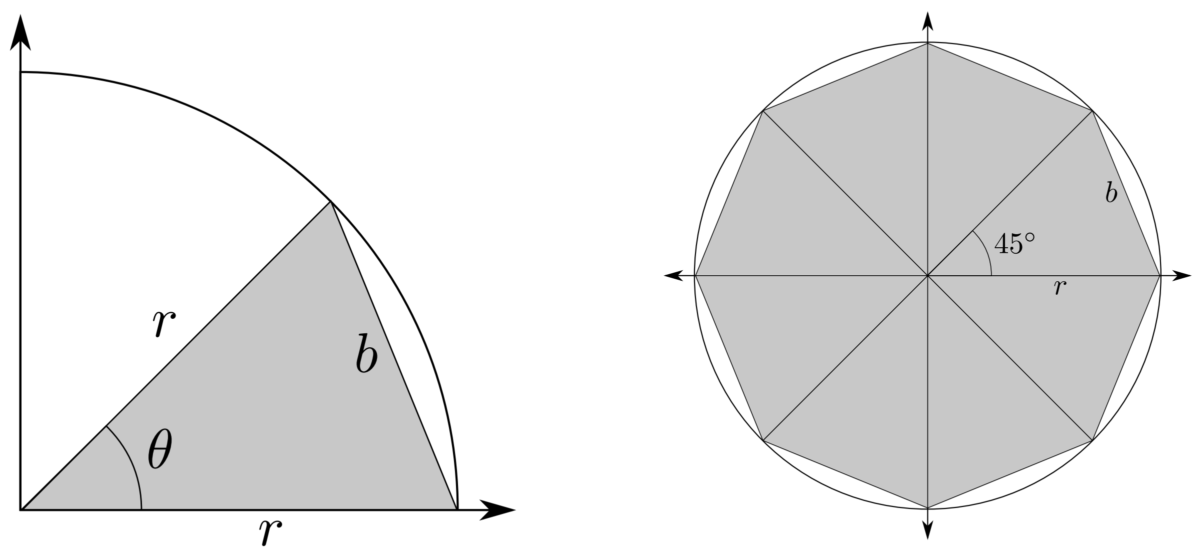
At the moment we'd have a bit of difficulty proving this proposition, but it is done by showing that $\{(1 + 1/n)^n\}_n$ is an increasing sequence, which is bounded above. The least upper bound of the sequence is e.

In order to gain insight into the number π, we will look at approximating a circle r by a number of inscribed isosceles triangles. Each of these triangles will have two sides of length r, and will be used to approximate a portion of the circle that has a certain arc length. The radius r and this arc length will determine all of the other properties of the given triangle. For an arc of angle θ, we can use the law of cosines to find the base of the triangle (see figure 3.7), given by

$$b = \sqrt{r^2 + r^2 - 2r^2 \cos(\theta)} = r\sqrt{2(1 - \cos(\theta))}.$$

Now that we have an expression for the base of one of these triangles, we can begin the iterative process. Let's begin with triangles that have a 45° angle. Inside of a given circle, we can inscribe 8 of these triangles. Thus, using 8 of these triangles (see figure 3.7), the circumference is approximated as $8b$, or

$$8 \cdot r\sqrt{2(1 - \cos(45°))}.$$



(a) Finding the base of an approximating triangle.

(b) Approximating the circumference of a circle.

Figure 3.7: Using triangles to approximate the circumference of a circle.

Now if we divide by the diameter of the circle, we find the ratio of circumference to diameter for the first approximation, which gives us the first term of our sequence,

$$a_1 = \frac{8 \cdot r\sqrt{2(1 - \cos(45°))}}{2r} = 4\sqrt{2(1 - \cos(45°))} \approx 3.06147,$$

where we use the well-known fact that $\cos(45°) = \sqrt{2}/2$. Interestingly enough, we note that r does not appear in this expression - at least for the first approximation, the ratio of circumference to diameter is constant. Given how the approximations are constructed, it's not difficult to see that none of these approximations will depend on the radius r. It follows then in the limiting case as $n \to \infty$ this will still hold true, so we conclude our first major result.

Claim 3.8.3. *The ratio of a circle's circumference to its diameter is a constant, denoted by π.*

Now as we move into further approximations, our target will be this elusive value π, given by the ratio of circumference to diameter. In order to make better approximations, we simply increase the number of triangles involved, decreasing the arc each triangle is approximating. In this way, we can see that our approximations will get closer and closer to the actual circumference of the circle, so our sequence will approach π. We know that this sequence will converge, because it is increasing, and it is bounded by 4 (just look at the circle as inscribed in a square, and note that the ratio of the perimeter of the square to the diameter of the circle is 4, where the perimeter of the square is clearly larger than the circumference of the circle). In each of our approximations we need to calculate cosine, so we need to be somewhat careful in the angles we choose to use. From the half-angle identity for cosine, we have

$$\cos(\theta/2) = \pm\sqrt{(1-\cos(\theta))/2},$$

but we know that we will always use the positive half, because we are looking at an angle inside a triangle. Now we can use our knowledge of $\cos(45°)$ to find

$$\cos((45/2)°) = \cos(22.5°) = \frac{\sqrt{\sqrt{2}+2}}{2}.$$

We can go through this process recursively, in order to double the number of triangles we use with each step of approximation. In this way, a general term of our sequence can be written

$$\{a_n\}_n = \frac{8 \cdot 2^{n-1} \cdot r\sqrt{2(1-\cos((45/2^{n-1})°))}}{2r} = 4 \cdot 2^{n-1}\sqrt{2(1-\cos((45/(2^{n-1}))°))}.$$

Here we have $8 \cdot 2^{n-1}$ because the number of triangles doubles every step, and on the first step $2^0 = 1$ so we have 8 triangles. Similarly, we have $(45/2^{n-1})$ as the angle of each triangle is halved each step, with a beginning angle of 45°. Since we know that this sequence converges, we don't actually need to evaluate each term of the sequence. We can simply move out to a very far term in the sequence and we'll get a decent approximation for π. For instance,

$$a_6 = 4 \cdot 2^5\sqrt{2(1-\cos((45/(2^5))°))} \approx 3.14151$$

and

$$a_{10} = 4 \cdot 2^9\sqrt{2(1-\cos((45/(2^9))°))} \approx 3.1415923.$$

One should be careful in continuing this algorithm out much further on a calculator, because round-off error will become a problem when we are asking the calculator to deal with numbers as small as $1-\cos(\theta)$ when $\cos(\theta)$ is becoming very near 1. Nevertheless, through this method we could show that the ratio of a circle's circumference and its diameter is constant, and we made a reasonable approximation at its value.

Interestingly enough, we can take this analysis one step further. Let's look at the area of a circle using this same process. To do so we need to find the area of one of our approximating triangles. To find the area of one of these triangles we first need to find the height (see figure 3.8).

If we cut the base of the triangle in half, then we have a right triangle, and using the Pythagorean theorem we find that

$$h = \sqrt{r^2-(b/2)^2} = \sqrt{r^2-r^2(1-\cos(\theta))/2} = \frac{r}{2}\sqrt{2(1-\cos(\theta))}.$$

In the above analysis some steps are skipped, because algebraic manipulation is not our focus. We now find that the area of one of our triangles is given by

$$A = \frac{1}{2}b \cdot h = \frac{1}{2}r\sqrt{2(1-\cos(\theta))} \cdot \frac{r}{2}\sqrt{2(1-\cos(\theta))} = \frac{r^2}{2}\sqrt{1-\cos(\theta)}\sqrt{1+\cos(\theta)}.$$

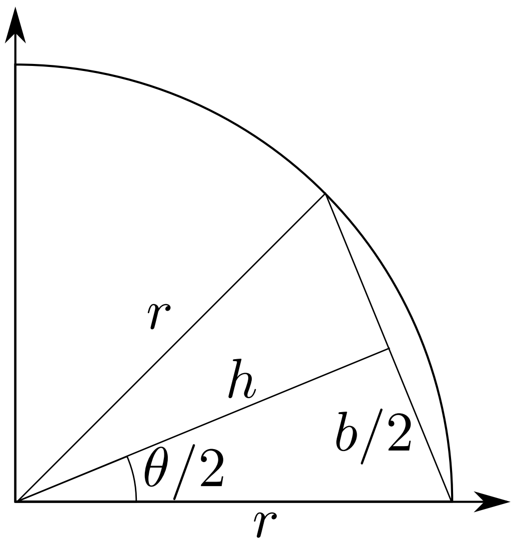


Figure 3.8: Finding the area of an approximating triangle.

With the way we have two square root terms multiplied together, we might expect that they will simplify into something more familiar. If we square these terms we find

$$(1+\cos(\theta))(1-\cos(\theta)) = 1-\cos^2(\theta) = \sin^2(\theta),$$

so it follows that

$$\sqrt{1-\cos(\theta)}\sqrt{1+\cos(\theta)} = \sin(\theta).$$

It follows now that the area of one of these triangles is given by

$$A = \frac{r^2}{2}\sin(\theta).$$

We proceed as before in order to make a sequence of approximations, defining

$$b_n = 8\cdot 2^{n-1}\cdot\frac{r^2}{2}\cdot\sin((45/2^{n-1})^\circ) = 4\cdot 2^{n-1}\cdot r^2\cdot\sin((45/2^{n-1})^\circ).$$

This shows that the area of a circle is proportional to the square of its radius (which we already *knew*), and also gives us a new sequence approximating π,

$$\{c_n\}_n = \{4\cdot 2^{n-1}\cdot\sin((45/2^{n-1})^\circ)\}_n \nearrow \pi.$$

3.9 The Bisection Method

The bisection method is an algorithmic procedure for finding a root of a given equation - a place where it is 0. This is useful because we can rewrite any algebraic equation in terms of finding a root. For our purposes we will use the bisection method to find sequences converging to $\sqrt{2}$, thus providing meaning to this elusive irrational number. $\sqrt{2}$ is defined as the number that when squared it equals 2. Thus, it is a solution to the equation

$$f(x) = x^2 - 2 = 0.$$

We know there are two solutions to this equation, which are $\pm\sqrt{2}$. In order to use the bisection method, we need to find two starting guesses for a solution, one which is below the actual value, and one that is above it. Starting with very simple and inaccurate guesses, we know that $\sqrt{2}$ lies between 1 and 2. If we plug 1 into the equation

$$f(x) = x^2 - 2$$

the result is -1, which is less than 0. If we plug 2 into the above equation the result is 2, which is greater than 0. Now that we know a place where the function is less than 0 and greater than 0, we know that somewhere in between 1 and 2 it must be 0, by the intermediate value theorem. It is worth noting that since the bisection method relies upon the intermediate value property it can only be applied to continuous functions.

From these two starting points we will generate two sequences, one that decreases to the root of the equation, and one that increases to it. Since both of these sequences will be bounded and monotone, it follows that they must converge, and their limit will be a root of the equation. As the name bisection suggests, at each step we will cut our interval in half, in order to generate these two sequences. Rather than giving the algorithm in generality this point, we will simply work through it as a demonstration. The first term of our first sequence is $a_1 = 1$, and the first term of the other sequence is $b_1 = 2$. If we let x_0 represent the root we will find through the bisection method, then we will have

$$\{a_n\}_n \nearrow x_0 \quad \text{and} \quad \{b_n\}_n \searrow x_0.$$

Without further ado, let us construct the required sequences.

In the first step our interval is $[1, 2]$. Now we cut this interval in half, adding the midpoint of 1.5. Next we evaluate our function at this point, and see

$$f(1.5) = 1.5^2 - 2 = 0.25 > 0.$$

Since $f(1) < 0$ and $f(1.5) > 0$, we know that the function has a root between these two points (by the intermediate value theorem). Thus, we set the next terms of our sequences as

$$a_2 = 1 \quad \text{and} \quad b_2 = 1.5.$$

To find the next terms we go through the same process. We cut our new interval $[1, 1.5]$ in half, adding in a point at 1.25. We evaluate the function here, and find

$$f(1.25) = 1.25^2 - 2 = -0.4375 < 0.$$

It follows then that our point of interest lies in the interval $[1.25, 1.5]$. We set the next terms of the sequence as

$$a_3 = 1.25 \quad \text{and} \quad b_3 = 1.5.$$

We can continue in this way indefinitely, until we find smaller and smaller intervals surrounding the point of interest. Both $\{a_n\}_n$ and $\{b_n\}_n$ converge to $\sqrt{2}$, so we can approximate $\sqrt{2}$ as accurately as required simply by moving far enough in these two sequences.

We can go through this process for many other functions, and even functions that have multiple roots. One difficulty to note however, is that we don't know which root we will converge to using the bisection method. If there are multiple roots between our two starting points, we are not sure which one we will get. For this reason it helps to have some starting knowledge to choose reasonable starting points. The general procedure for the bisection method is outlined below.

Algorithm 3.9.1 (Bisection Method)**.** The bisection method is used to solve equations of the form $f(x) = 0$, where f is a continuous function.

1. Find values a and b with $a < b$ for which $f(a)$ and $f(b)$ have opposite signs; if either $f(a)$ or $f(b)$ is zero then you have already found a root and are done. If not, set $a_1 = a$ and $b_1 = b$.
2. Given a_n and b_n look at the point $c_n = (a_n + b_n)/2$.
   - If $f(c_n) = 0$, you have found a root, and are done.
   - If $f(a_n)f(c_n) < 0$, then $f(a_n)$ and $f(c_n)$ have opposite signs, so set $a_{n+1} = a_n$ and $b_{n+1} = c_n$.
   - Otherwise, $f(b_n)f(c_n) < 0$, so set $a_{n+1} = c_n$ and $b_{n+1} = b_n$.
3. Repeat step 2 until an approximation of the desired accuracy is found

In the above algorithm we use the trick that if $a \cdot b < 0$, then a and b must have opposite signs. If both a and b were positive then $a \cdot b$ would be positive, and similarly if both a and b were negative, then the negative signs would cancel in the multiplication. Thus, the only way the product of two numbers can be negative is if one of the numbers is positive and the other is negative; ie. they have opposite signs.

The above formula is written inductively, which means that it gives you a method for completing the first step, and a means to compute the next step given the previous. An algorithm of this form is well suited to implementation on a computer, and is not really meant to be performed by hand. The bisection method is not the most efficient method for finding roots, but since the root always remains in between the two sequences, we are guaranteed to find a root. This makes the bisection method particularly appealing from a theoretical standpoint, even if in practice we might use a faster converging method (that gives us a better approximation after fewer terms). Another method we will consider later on is Newton's method, which will converge much more quickly than the bisection method (but doesn't always converge).

If we look in the error of our approximation by either of the two sequences converging to the root, for the first term the maximum error is $b - a$, which would occur if the root lies exceedingly close to one of the end points, but we use the other endpoint for our approximation. Since on the second step our interval is cut in half, the maximum error becomes $(b - a)/2$. In general, the maximum error for the n^{th} term of either sequence is

$$\frac{b-a}{2^{n-1}}.$$

In order to use the bisection algorithm to approximate a given irrational number, one simply needs to rewrite the problem into the problem of finding a root. For instance, to approximate $a^{m/n}$, one uses the equation

$$f(x) = x^{n/m} - a = 0.$$

In doing so one must then have a means of evaluating something to the $(n/m)^{th}$ power.

3.10 First-Order Approximations

When we are faced with a function that is too difficult to work with directly, sometimes we can instead work with a simpler function that approximates the function we are interested in. Even though the resulting solutions will only be approximations, approximate solutions can often provide a lot of insight into a problem. In fact, in many situations exact solutions may be impossible to find.

The problem of approximating functions is closely linked to the problem of sampling. For a function defined on the real numbers, we can find the value of the function at any point. However, when we measure a physical quantity, say temperature, it doesn't come packaged as an explicitly-written mathematical function. To construct a function of time we must measure or sample the given quantity at a number of instants in time. This leads immediately to the following question - how often should one measure the quantity? Every hour, minute, or second? Even if we measure the temperature every second, there will be an infinite number of times between each second in which we do not know the temperature; our function of temperature with respect to time looks like a number of points with large gaps in between. Such a function is called a discrete-time function (or just discrete function), usually defined at regular intervals, in contrast to a continuous-time function defined for all possible values of time.

Since no matter how often we sample a physical quantity there will still be gaps between our measurements, anytime we record data from a physical situation for mathematical analysis we are forced to work with discrete functions. Anytime we want to know the value of the function at a point other than a sampling point, we are forced guess what the value would have been, had we actually measured it at that time. Trying to determine the value of the function at unknown points using known, surrounding points, is called interpolation. Through interpolation we can construct a continuous-time function that approximates our discrete-time function. One of the simplest ways of interpolating between data points is to simply draw a line segment connecting each two successive points. Such a segment is called a chord, and its extension to a line is called a secant line.

Definition 3.10.1 (Secant Line)**.** A line that intersects a function f at the points (x_0, y_0) and (x_1, y_1) is called the *secant line* intersecting f at (x_0, y_0) and (x_1, y_1).

Once we've found a secant line we can find the chord between two points simply by restricting the domain of the line to be between the two points. Recall that for any two points in the Cartesian plane, there is a unique line that passes through both of them. Thus, just knowing the values of two points we can find the secant line passing through them. We generally define such a line using the point-slope form.

Definition 3.10.2 (Point-Slope Form)**.** The equation for the unique line passing through the point (x_0, y_0) with slope m written in *point-slope form* is given by

$$y - y_0 = m(x - x_0).$$

Now let's suppose that we know the two points (x_0, y_0) and (x_1, y_1) for our discrete function. In order to find the slope of the line passing between them we simply substitute the point (x_1, y_1) into the point slope equation, finding that

$$y_1 - y_0 = m(x_1 - x_0),$$

or

$$m = \frac{y_1 - y_0}{x_1 - x_0},$$

which is simply the ratio of change in output over change in input. Having found the slope we find that the equation for the secant line will be given by

$$y - y_0 = \frac{y_1 - y_0}{x_1 - x_0} \cdot (x - x_0).$$

Having found the secant line intersecting (x_0, y_0) and (x_1, y_1) we can now define our approximating function so that it coincides with the secant line on the interval $[x_0, x_1]$. Continuing in the same fashion for each other pair of points we can replace our discrete-time function with a continuous-time approximation, defined piecewise by a number of first-order polynomials.

Let us return to the problem of approximating a complicated function, such as an exponential function (of any base). As we've defined it, such a function is evaluated at a given point by finding the solution to a certain equation (for instance, $a^{1/n}$ is a root to the equation $x^n - a = 0$). We've previously seen that we can find such solutions using the bisection method, but doing so takes some work. Thus, exponential functions are a reasonable candidate for approximation, as they difficult to actually evaluate except at a very limited number of points.

For the sake of illustration, let's consider a base 2 exponential. We know that $2^0 = 1$ and $2^1 = 2$, so we can use the same method as previously to approximate the values of the function between 0 and 1. The slope of the secant between these two points is given by

$$m = \frac{2-1}{1-0} = 1,$$

so we find that

$$2^x \approx x + 1, \qquad \text{for } 0 \leq x \leq 1.$$

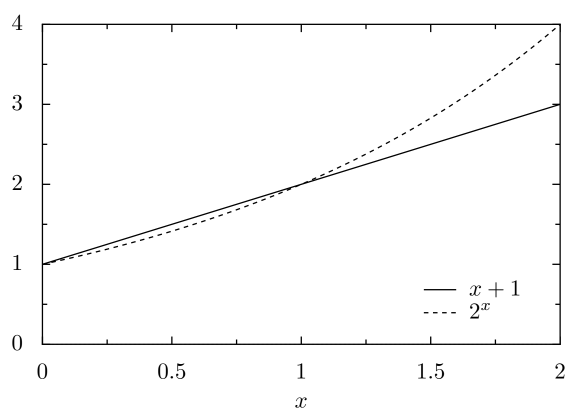


Figure 3.9: Graph of the function 2^x and the secant line through (0,1) and (1,2) approximating it.

This provides us, for instance, with the estimate that $2^{0.5} \approx 1.5$. For such a quick and simple method this is a reasonable approximation to the actual value of

$$2^{0.5} = \sqrt{2} = 1.414\ldots.$$

Nevertheless, in most situations we require a higher degree of accuracy. Before looking at higher-order approximations (which we won't for some time), let's further investigate first-order approximations. Above we decided to use the secant line to make an approximation because it is very simple to calculate, and seems to be an obvious first choice. However, that by no means implies it is the *best* choice. To proceed in answering this question we'll have to further investigate the properties of functions.

Let us take a look at the *smooth*, continuous function $f(x) = x^2$, at the point $x = 1$ (these choices are quite arbitrary - it just happens that this function has the property of interest). If we restrict our view of this function to very small intervals surrounding the point $x = 2$, we begin to notice something in the behavior of x^2 (the easiest way to do this is by looking on a small window of a graphing calculator or computer - see figure 3.10). The more we *zoom in* the more it appears that the function we are looking at is a line. If we look at other points we'll notice that the function has the same behavior. To classify this behavior we say that, locally, x^2 behaves like a first-order polynomial. This means that over a small enough interval, the behavior of the function closely resembles that of a line. A special name is given to the line that best resembles the function in a small neighborhood of the point of interest - the *tangent line*. It is best to disclaim that a tangent line only touches a function at a single point or cannot cross a function. These misunderstandings are often used to intuitively *motivate* the tangent line, but both are untrue.

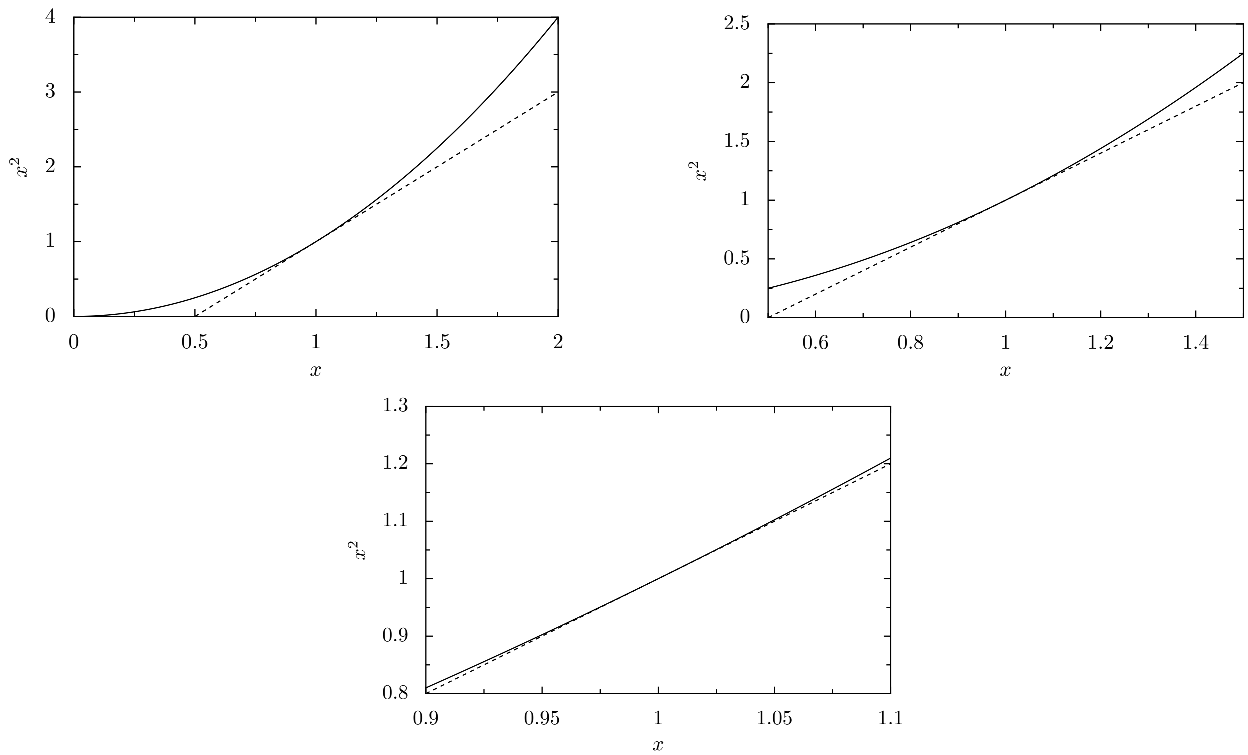


Figure 3.10: In the near vicinity of $x = 1$ the behavior of x^2 resembles that of a line. The dotted line represents the line that x^2 begins to behave like.

We are immediately brought to the question of how to find a tangent line. In order to find such a line we return to the secant lines that we dealt with earlier. Over a large interval, our secant lines give a pretty poor approximation, because there is a lot of distance between the two points of the function connected by the secant line - in such a large space the function can change a lot. However, as we decrease the distance between the points of intersection of a secant line, the secant line becomes a better approximation, at least in the limited region between the two points of intersection, and the points nearby them. The more we decrease the distance between the two points a secant line intersects, the better the approximation becomes in the nearby region (see figure 3.11). If we look in the limit as the distance between the two points of intersection approaches 0, then we approach the line that provides the best first-order approximation in that region - the tangent line.

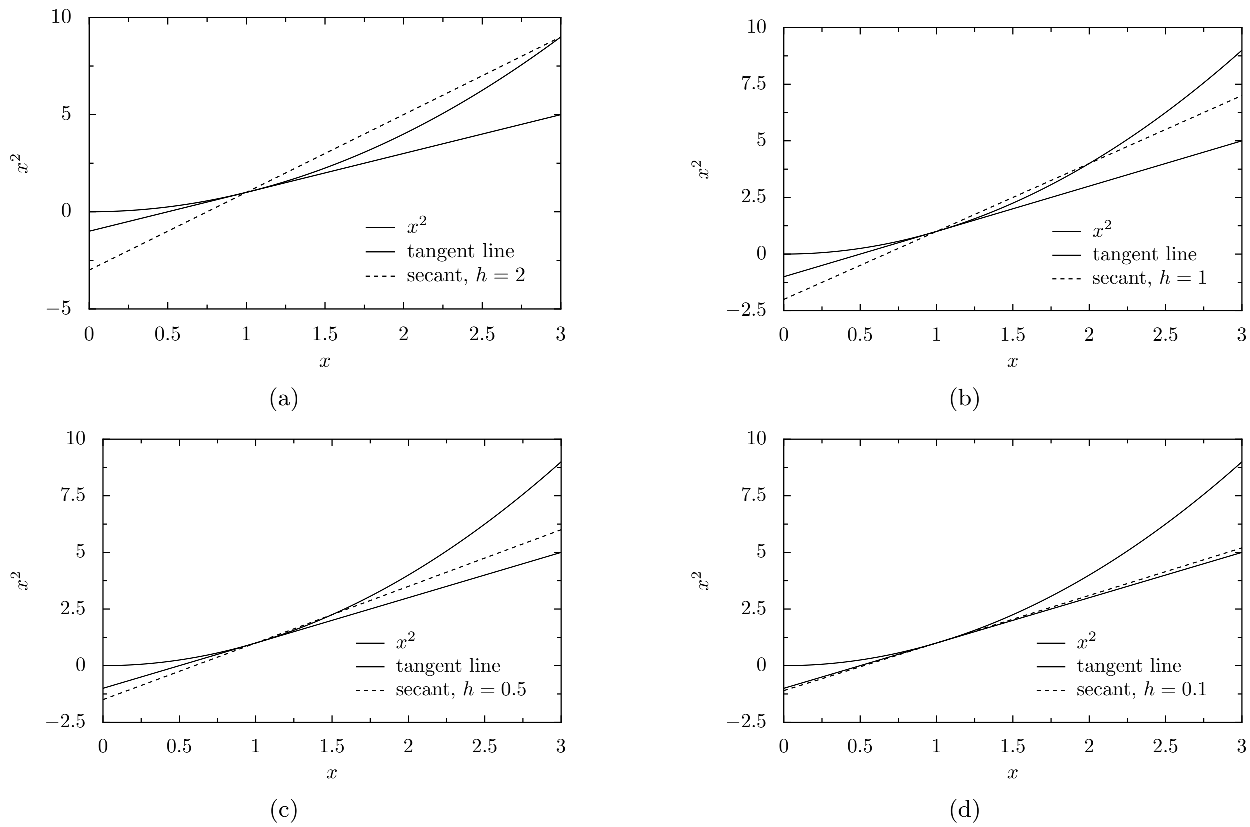


Figure 3.11: As the distance h between the two points the secant line intersects becomes smaller, the secant line more closely represents the tangent line.

Let's perform the above argument mathematically in order to find a mathematical representation for the tangent line. Let's consider a function f and a point x_0. Since we want the behavior of f to be resembled by that of the tangent line near x_0, we begin by requiring that the tangent line intersect the point $(x_0, f(x_0))$. This is called the *point of tangency*, where the tangent line has the same value as the function it is tangent to. Now in order to find the point-slope form for the equation of the tangent line, we just need to find the tangent line's slope. As stated above, we find the slope of the tangent line through a limiting process governed by secant lines. Let's consider the secant line passing through the points $(x_0, f(x_0))$ and $(x_0 + h, f(x_0 + h))$, where $h \in \mathbb{R}$, $h \neq 0$. This slope is given by

$$\frac{f(x_0 + h) - f(x_0)}{h}.$$

In order to find the slope of the tangent line, we simply look at the limit as $h \to 0$. In other words, the slope of the tangent line is given by

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

This quantity plays such an important role in calculus that it is given its own name.

Definition 3.10.3 (Derivative at a Point)**.** The *derivative* of a function f at the point x_0 is denoted by $f'(x_0)$, where

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

If the above limit exists, f is said to be *differentiable* at x_0.

Through the above discussion we see that in order to find the line tangent to a function at a given point, we need to evaluate a certain limit. Even though we used the notion of a tangent line to motivate the above definition, we will actually formally *define* the tangent line in terms of it.

Definition 3.10.4 (Tangent Line)**.** Let f be differentiable at x_0. The line tangent to f at the point x_0 is the unique line passing through the point $(x_0, f(x_0))$ with slope $f'(x_0)$.

A function only has a tangent line at a point if it is differentiable at that point. It follows that a function does not behave like a line in the vicinity of points at which it is not differentiable - it instead exhibits some more complicated behavior. There are a number of ways in which a function can fail to be differentiable at a given point. A function is not differentiable at any place it has a:

1. *corner.* Consider $f(x) = |x|$ which has a corner at $x = 0$. If we look at the secant lines in the limit as $h \to 0$, we see that as $h \nearrow 0$ the slope of the secant lines approaches -1, and that as $h \searrow 0$ the slope of the secant lines approaches 1. Since these limits do not agree, the derivative and thus tangent line do not exist at $x = 0$ (a similar analysis will hold for a corner of any function). More intuitively, in the vicinity of $x = 0$, the function $|x|$ does not behave like a line.

2. *a cusp.* Consider the function $f(x) = \sqrt{|x|}$ at the point $x = 0$ (where it has a cusp). If we look at the secant lines as $h \searrow 0$, their slopes approach ∞, and as $h \nearrow 0$ their slopes approach $-\infty$. Thus, the limit that defines the slope of the tangent line doesn't exist, to the function is not differentiable at $x = 0$ (and so it has no tangent line there).

3. *vertical tangent.* A vertical tangent line exists when a function locally behaves like a line, but the line it behaves like is vertical. As defined above, a vertical tangent line is not really considered a tangent line, because it occurs at a point where a function is not differentiable. Similar to the above example, a vertical tangent line occurs if the limit of the slopes of the secant lines as $h \to 0$ approaches either ∞ or $-\infty$ from both sides. An example of function with a vertical tangent is $(2 - x)^{1/5}$.

4. *a discontinuity.* If we have a point discontinuity the secant lines will behave like when we have a cusp. A function with a jump discontinuity will have different behavior on both sides of the point of interest. For instance, $f(x) = |x|/x$, $f(0) = 1$ has a jump discontinuity at $x = 0$. The secant lines approach a horizontal tangent from the right, and vertical from the left. Thus, the derivative does not exist at this point (nor does a tangent line).

There is one subtlety in the above statements that we will repeat again for emphasis. Above it is said that a function is not differentiable at any place it has a discontinuity. The contrapositive of this statement is that at any point a function is differentiable, it must be continuous. Thus, if we know that a function is differentiable then we immediately know it is continuous, or *differentiability implies continuity.* This subtle fact can come in very handy at times.

Although we did not explicitly say so at the time, the argument we made above with secant and tangent lines is essentially the same argument we made in motivating the concept of the limit, using average and instantaneous velocity. In fact, the slope of a secant line is the same as the average rate of change of a function over the interval defined by those two points. Similarly, the slope of the line tangent to a point is the same as the instantaneous rate of change of the function at that point. Thus, the derivative of a function at a point has many meanings. The derivative of a function f at a point x_0 can be interpreted as:

1. The instantaneous rate of change of f at the point x_0.
2. The slope of the line given by looking in the limit as $h \to 0$ of secant lines passing through $(x_0, f(x_0))$ and $(x_0 + h, f(x_0 + h))$.
3. The slope of the line that best approximates the function f in the vicinity of x_0, the tangent line at x_0.
4. The slope of the function f at the point x_0.

From here on we will begin the study of differential calculus, where our primary object of study is the derivative.

Chapter 4

Differentiation

When we study differentiation, we see the reason why calculus is often described as the mathematics of change. That is exactly what differentiation is about - how functions change.

Contents

4.1 Definition of the Derivative . . . 92
4.2 Differentials and Infinitesimals . . . 99
4.3 Properties of Differentiation . . . 101
4.4 The Chain Rule . . . 107
4.5 Derivatives of Trigonometric Functions . . . 113
4.6 Derivatives of Inverse Functions . . . 119
4.7 Differentiating Implicitly Defined Functions . . . 122
4.8 Related Rates . . . 126
4.9 Extreme Values . . . 131
4.10 Mean Value Theorem . . . 136
4.11 Concavity and Curve Sketching . . . 140
4.12 Optimization . . . 145
4.13 l'Hôpital's Rule . . . 155
4.14 Newton's Method . . . 158
4.15 Euler's Method . . . 161

4.1 Definition of the Derivative

As previously defined, the derivative of a function at a point is a number that is equal to the slope of the line tangent to a function at that point. It will be extremely useful for us to extend this notion to define a new function, which provides the derivative of a function at all points for which its derivative at a point is defined.

Definition 4.1.1 (Derivative)**.** The *derivative* of a function f is a function denoted by f', where

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

The domain of f' is all x for which f is defined and the above limit exists.

Note that the domain of a function's derivative will always be smaller than that of the original function. In finding a function's derivative we say that we *differentiate* the function. The reason it is so useful to consider the derivative of a function as a new function in its own right is because many physical laws manifest themselves in terms of derivatives. One of the most common of these physical laws is Newton's second law of motion, which states: from a given inertial frame the instantaneous rate of change, or derivative, of an object's momentum (with respect to time) is equal to the net force acting on it. Symbolically, Newton's second law is written

$$F_{net} = \frac{dp}{dt},$$

where F_{net} represents force, p momentum, and d/dt represents the derivative with respect to time. An equation that involves a function's derivatives, like of this form, is called a differential equation. Differential equations are extremely important precisely because they can be used to describe so many physical phenomena. This is perhaps the primary reason why the study of calculus is so important to physical scientists.

When we have a differential equation, we know something about a function's derivative(s), but what we really want to know is the function itself. For instance, if we have a differential equation of the form

$$\frac{dF}{dt} = f(t),$$

it is our goal to find a function $F(t)$ so that $F'(t) = f(t)$. In other words, we are interested in reversing the process of differentiation - we want to antidifferentiate $f(t)$.

Definition 4.1.2 (Antiderivative)**.** A function F is an antiderivative of f if

$$\frac{dF}{dx} = F'(x) = f(x),$$

for all x in the domain of f.

In general, a function can depend on any number of input variables. For instance, the pressure of an ideal gas depends on the volume of the gas, the number of moles of the gas, as well as the temperature of the gas. When we differentiate such a function, we need to be specific with respect to which input variable we are differentiating, as the pressure of the gas will change differently depending on which input variable is changing. For instance, the pressure will change differently depending on whether it is temperature or pressure that is changing as an input. Correspondingly, we may have multiple different derivatives for a given function, depending on which input variable is changing.

In situations where it may not be clear which input variable are differentiating with respect to, we will use the notation of the differentiation operator. In analogy to how functions define relations between sets of numbers, operators define relations between sets of functions: given a function as an input, an operator returns a new function as an output. Accordingly, we define the differentiation operator, which when applied to a function returns the derivative of that function as an output. With this operator we can explicitly reference which input variable we are differentiating with respect to. To differentiate with respect to x the differentiation operator is written

$$\frac{d}{dx}$$

(pronounced 'dee dee x'). In describing the action of this operator we write

$$\frac{d}{dx}f(x) = f'(x),$$

meaning that when the differentiation operator is applied to a function, it returns the derivative of that function (with respect to x) as an output. For this reason, it is common to see the notation

$$\frac{df}{dx}$$

to represent the derivative $f'(x)$. If we have a function f that doesn't depend on an input variable y, then we would find

$$\frac{df}{dy} = 0,$$

as there is no change in the output of f as y changes (because a change in y has no influence over f). This is a subtle concept one should look out for, especially when dealing with functions of multiple variables.

In addition to providing extra clarity with respect to which variable one is differentiating, the operator perspective provides another useful insight into differentiation. Suppose we have that

$$f(x) = x^2 + 1.$$

The above equation states that $f(x)$ and x^2+1 are equal, so just like when we apply an arithmetic operator to both sides of the equation, both sides of the equation should remain equal when we apply the differentiation operator. In other words, if we apply the differentiation operator to both sides of an equation, the result remains true upon differentiation, or

$$\frac{d}{dx}f(x) = \frac{d}{dx}(x^2+1)$$

so

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

Thus, anytime we want to find the derivative of some function, we can just think of applying the differentiation operator to both sides of the equation (even if we still need to evaluate the derivative on the right-hand side). This may not seem like a very useful insight, but we can just as well think of differentiating an equation of the form

$$x^2 + f(x) = 2x,$$

yielding

$$\frac{d}{dx}(x^2 + f(x)) = \frac{d}{dx}2x.$$

Although we are not yet able to evaluate something like the equation above, we will be able to exploit this idea in order to differentiate implicitly-defined functions. Thus, in addition to clarifying which input variable we are differentiating with respect to, the operator perspective provides a great deal of flexibility.

Before proceeding there is one more matter of notation to discuss. When we differentiate the function f twice, we write

$$f''(x),$$

which is called the second derivative of f with respect to x. In general, we write the n^{th} derivative as

$$f^{(n)}(x),$$

where the parenthesis around the n are used to represent differentiation, in contrast to exponentiation ($f^n(x)$ would be written to represent the n^{th} power of f). When we find the derivative of a function beyond the first, we call it a higher-order derivative. In a similar vein, we can define higher-order differentiation operators, so that an operator which returns the n^{th} derivative of a function is written

$$\frac{d^n}{dx^n} \quad \text{and} \quad \frac{d^n}{dx^n}f(x) = f^{(n)}(x).$$

Example 1. Find the derivative of an arbitrary first-order function $f(x) = mx + b$.

Solution In order to solve this problem we work with the definition of the derivative,

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

In other words, the derivative or instantaneous rate of change of a line is simply the slope of the line. This also happens to be the average rate of change of a line over any interval. It is worth noting that when we want to calculate derivatives the definition will always provide us with a limit with a denominator approaching 0. Thus, we will never be able to use the quotient rule for limits directly, but will need to rewrite the limit in a form we can evaluate it (using algebraic simplification).

The above example illustrates a very important aspect of differentiation. Suppose we have that

$$f(x) = 2x + 1 \quad \text{and} \quad g(x) = 2x + 5.$$

Upon differentiation, we find that

$$f'(x) = 2 \quad \text{and} \quad g'(x) = 2,$$

or that both f and g have exactly the same derivative. In general, anytime we differentiate a function we lose any constant that might have been added to it. Since two different functions can have the same derivative, we see that when we differentiate a function, we lose some information. Namely, the derivative of a function tells us exactly how the function changes, but it doesn't tell us where the function begins from. Thus, in order to reverse the process of differentiation, we will need more information about a function than just its derivative.

Example 2. Find a solution to the differential equation

$$\frac{dx}{dt} = \alpha, \quad \alpha \in \mathbb{R}.$$

Solution We need to find a function $x(t)$ so that $x'(t) = \alpha$. Using our knowledge of derivatives, we know that $x(t) = \alpha t$ is one such function, so it serves as a solution to the differential equation.

In the above example we saw that $x(t) = \alpha t$ was a solution to the differential equation because it's derivative was α. However, any line with slope α will just as well serve as a solution, because it also has a derivative of α. Thus, we end up with an entire family of solutions to the differential equation, all of the form

$$x(t) = \alpha t + c,$$

where $c \in \mathbb{R}$. If we are dealing with a physical situation, then an entire family of solutions is completely unacceptable. Suppose that $x(t)$ describes the position of an object. Based on physical intuition we must have a single solution, not an entire family of solutions. In order to arrive at a single solution, however, we need some additional information to determine the arbitrary constant c. Namely, we need to know the position of the object at a single instant in time. Intuitively, if we know the position of an object at some time, and how it changes for all time (given the derivative), we should be able to find its position for all subsequent times, which in fact we can. A problem in which we have a differential equation and some additional information to determine any arbitrary constants is called an initial value problem.

Example 3. Solve the initial value problem

$$\frac{dx}{dt} = \alpha, \quad \alpha \in \mathbb{R},\ x(0) = 1.$$

Solution We have already seen that a general solution to the differential equation is of the form

$$x(t) = \alpha t + c.$$

Now we need to use the additional information provided by the initial condition $x(0) = 1$. Evaluating the general solution at $t = 0$ we find

$$1 = x(0) = \alpha \cdot 0 + c,$$

which implies that

$$c = 1.$$

Thus, our solution to the initial value problem is given by

$$x(t) = \alpha t + 1.$$

In the above example we used the additional information provided, an *initial condition*, in order to find a unique solution. The information is called an initial condition because it often corresponds to knowledge of the function at the time $t = 0$. Nevertheless, a measurement of x at any other time t would work just as well for determining c. This general class of problems is called initial value problems, because they rely upon solving a differential equation using an initial value for the function. Such problems are solved by finding a general solution, and picking out a specific member of the family using initial conditions.

If we want to consider the derivative of a function at an endpoint of the interval over which it is defined, we calculate the derivative using a one-sided limit, because the function is not defined beyond the endpoint. Consider $\sqrt{x}$, which is not defined for $x < 0$. In order to look at a derivative at $x = 0$, we use the limit

$$\lim_{h \searrow 0} \frac{\sqrt{0+h} - \sqrt{0}}{h} = \lim_{h \searrow 0} \frac{1}{\sqrt{h}} = \infty,$$

which implies $\sqrt{x}$ is not differentiable at $x = 0$.

Example 4. Let the position of a car be given by

$$x(t) = \begin{cases} \frac{t^2}{2} & 0 \le t < 1 \\ t - \frac{1}{2} & 1 \le t < 2 \\ \frac{t}{t-1} - \frac{1}{2} & 2 \le t < 4 \\ \frac{16}{9t} + \frac{7}{18} & 4 \le t. \end{cases}$$

Find the velocity of the car as a function of t. Are there any places where this velocity function does not make physical sense?

Solution On the interiors of these intervals (all points except for the endpoints) we can find the derivative using a two-sided limit. On the boundary points between intervals we will need to consider the one-sided limits on each side, and see if they are equal (which will tell us whether or not the two-sided limit exists). Starting on the first subinterval,

$$v(t) = \lim_{h\to 0} \frac{(t+h)^2/2 - t^2/2}{h} = \frac{1}{2}\lim_{h\to 0} \frac{(t+h)^2 - t^2}{h} = \frac{1}{2}2t = t,$$

which holds for $0 \le t < 1$. We can include $t = 0$ in this interval because the function is not defined for negative time, so there is no need to look at the left-hand limit there. On the second interval $1 < t < 2$ we have

$$v(t) = 1,$$

because the derivative of a line is just its slope. To find $v(1)$ we need to look at the one-sided limits

$$\lim_{h\nearrow 0} \frac{(1+h)^2/2 - 1^2/2}{h} = 1 \quad \text{and} \quad \lim_{h\searrow 0} \frac{(1+h) - 1/2 - (1-1/2)}{h} = 1.$$

Since the limits on both sides agree, it follows that $v(1) = 1$. Note that it was clear by inspection that the one-sided limits would match (based on calculating the derivative on the interior of the interval), but as a matter of technicality we calculated the one-sided limits.

To look at the derivative on the third interval it is enough to consider the derivative of $t/(t-1)$, since the constant term will disappear when we differentiate.

$$\begin{aligned} v(t) &= \lim_{h\to 0} \frac{\frac{t+h}{t+h-1} - \frac{t}{t-1}}{h} = \lim_{h\to 0} \frac{(t+h)(t-1) - t(t+h-1)}{h(t+h-1)(t-1)} \\ &= \lim_{h\to 0} \frac{t^2 + th - t - h - t^2 - th + t}{h(t+h-1)(t-1)} = \lim_{h\to 0} \frac{-h}{h(t+h-1)(t-1)} = \frac{-1}{(t-1)^2}. \end{aligned}$$

Evaluating

$$\lim_{t\to 2} \frac{-1}{(t-1)^2} = -1$$

we see that the derivatives do not match on both sides of $t = 2$, so this portion of the derivative only holds for $2 < t < 4$ (and the derivative does not exist for $t = 2$). Looking to the fourth interval, once again ignoring the constant

$$v(t) = \lim_{h\to 0} \frac{\frac{16}{9(t+h)} - \frac{16}{9t}}{h} = \frac{16}{9}\lim_{h\to 0} \frac{t - (t+h)}{ht(t+h)} = \frac{16}{9}\lim_{h\to 0} \frac{-h}{ht(t+h)} = -\frac{16}{9t^2}.$$

Since both of the one-sided derivatives match as $t \to 4$, $v(4) = -1/9$. Thus, we have $v(t)$ for all points on $[0, \infty)$, except for at $t = 2$, where the derivative does not exist. Physically, an

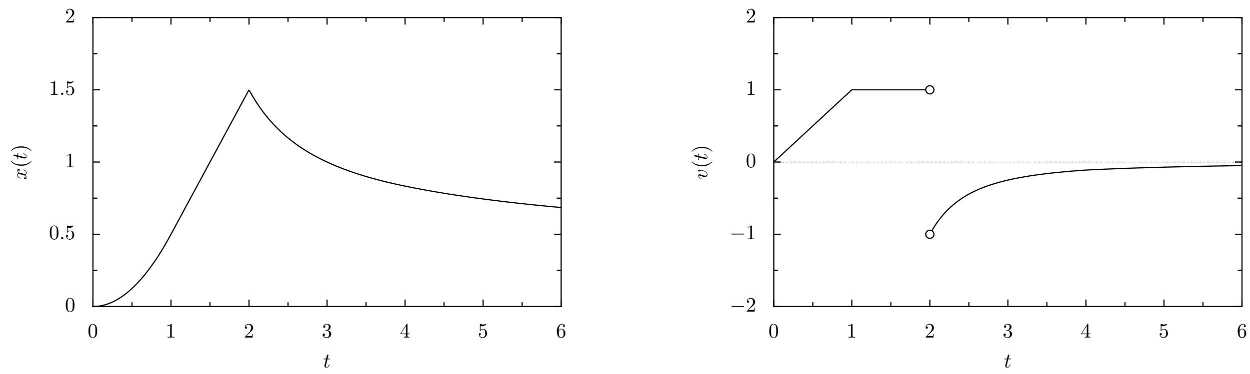


Figure 4.1: Position and velocity of the car.

instantaneous change or jump in velocity corresponds to some type of 'infinite' acceleration; such an acceleration is clearly not physically realizable, and corresponds to the point in the graph of the function $x(t)$ where there is a corner.

Interpreting this figure we see that initially the velocity of the car is increasing, and then the car continues to move along at a constant velocity. After some strange event at $t = 2$ the car is instantly (and nonphysically) shifted into a negative acceleration, where it obtains a negative velocity and begins moving the other way. In this situation a positive velocity corresponds to traveling right from some reference point, and a negative velocity to traveling left.

Essentially, the derivative of a function tells us two things about how the function is changing - in which direction a function is changing, and how fast it is changing. The sign of the derivative tells us the direction in which a function is changing. A positive derivative corresponds to an increasing function, a negative derivative to a decreasing function, and a zero derivative to a function that is not changing. On the other hand, the magnitude of the derivative tells us how rapidly the function is changing. A large magnitude corresponds to rapid change, and a small magnitude corresponds to gradual change. Thus, a large positive derivative corresponds to a rapid increase, and a small negative derivative corresponds to a gradual decrease.

By mentally superimposing tangent lines onto a given function, we can gain quick intuitive insight into the behavior of a function. The basic idea is to look at the function at a point, and draw what would *seem* to be the tangent line, using an intuitive understanding of the best first-order approximation. From this mental picture one can quickly estimate the slope of the tangent line, and thus how fast the function is changing, as well as whether or not it is increasing or decreasing. This type of technique is an important tool for quickly grasping information about a function, just using a qualitative picture, even if exact values are unavailable.

Example 5. Given the function shown in figure 4.2, of population vs time, answer the following questions. Where does the function seem to be increasing the slowest. The fastest? How fast is it changing at these times?

Solution The first thing to observe from our function is that it is always increasing; as we move from left to right the function values always get larger. We can also confirm this by visualizing any tangent line, and realizing that it has a positive slope. In order to estimate the area where we are increasing the fastest, it is probably easiest to use a ruler or other straight edge to trace out tangent lines. Doing so we find the slope is actually the steepest in the very beginning (near $t = 0$), and begins to gradually soften. As we continue along the slope reaches its minimum just after $t = 3$, before it begins to steepen again. Thus, the population seems to be growing fastest around times

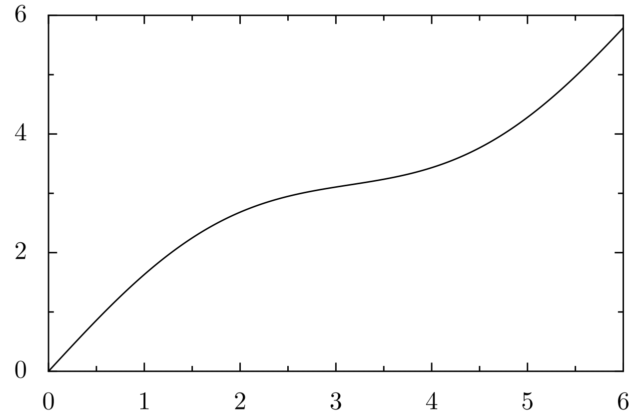


Figure 4.2: Graph of animal population vs time.

$t = 0$ and $t = 6$ (although it is actually growing slightly faster at $t = 0$, even if it is difficult to tell from this rough estimate), and growing slowest around $t = 3$. At $t = 0$ the derivative is about 1, and at $t \approx 3$ the derivative is about $1/4$.

Example 6. Consider the velocity of an object given in figure 4.3. When is the object moving forward; backward? When does the object move at its greatest speed? When is its acceleration positive, negative, zero? When is its acceleration the greatest? When does it stand still for more than an instant?

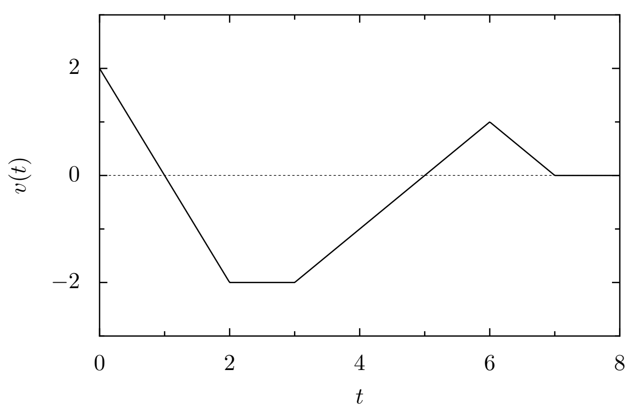


Figure 4.3: Velocity of an object vs time.

Solution The object is moving forward when it has a positive velocity, so it is moving forward on [0,1) and (5,7). Moving backwards corresponds to a negative velocity, so it is moving backwards from (1,5). The object is moving at its greatest speed when the magnitude of its velocity is the largest, which corresponds to $t = 0$ and $t \in [2, 3]$. The acceleration is positive when the slope of velocity is positive, so the object has a positive acceleration (speeding up) on (3,6). Similarly, it is slowing down on [0,2) and (6,7). The object has its greatest acceleration when the slope of its velocity is the steepest, which corresponds to [0,2). The only time the object stays still for more than a moment is for $t > 7$. It is worth noting that there is a slightly inaccuracy in this graph, insofar that corners of the object's velocity do not correspond to physical accelerations, as the object's acceleration is not defined at these points. To be strictly correct rounded corners should be used, but are not for the sake of simplicity.

4.2 Differentials and Infinitesimals

Before we proceed further in studying differentiation, it is worthwhile to acknowledge the formal mathematical path we are going to take. When calculus was originally developed (primarily by Newton and Leibniz) it was built on the notion of infinitesimals, rather than limits. Just as infinity represents something larger in magnitude than any real number, an infinitesimal represents something smaller in magnitude than any nonzero real number. It follows that zero is the only infinitesimal real number, which is not particularly helpful. Nevertheless, it is possible to extend the real numbers into the hyperreal numbers, in which there are nonzero infinitesimals. This allows for calculus to be developed using infintesimals rather than limits. We will focus our attention on nonzero infinitesimals, as they will provide us with the required behavior to develop calculus. Namely, we can perform multiplication and division with (nonzero) infinitesimals and arrive at finite results. In this way we can perform the fundamental operations of differentiation and integration using infinitesimals rather than limits.

From the infinitesimal point of view we think of the derivative of a function at a point as the change of the function over an infinitesimal interval containing that point. The notation

$$\frac{\Delta f}{\Delta x}$$

is used to represent the average rate of change of f over an interval, so Δf represents a finite change in f, and Δx a finite change in x. In contrast, we use the notation

$$\frac{df}{dx}$$

to represent instantaneous rate of change at a point, or average rate of change over an interval of infinitesimal length. Here df represents an infinitesimal change in f, and dx represents an infinitesimal change in x. These quantities df and dx are called *differentials*, which just means that they represent infinitesimal changes. Although these infinitesimals are *smaller* than any nonzero real number, by looking at a ratio of them, we are able to arrive at a finite result. The mechanics of doing so are very similar to finding a derivative using limits.

Example 1. Find the derivative of $f(x) = x^2$ using infinitesimals.

Solution In order to find the derivative of f we need to look at the average rate of change over an interval of infinitesimal length. Thus, we consider

$$\begin{aligned} f'(x) &= \frac{df}{dx} = \frac{f(x+dx) - f(x)}{dx} = \frac{(x+dx)^2 - x^2}{dx} \\ &= \frac{x^2 + 2x(dx) + (dx)^2 - x^2}{dx} = \frac{2x(dx) + (dx)^2}{dx} = 2x + dx \end{aligned}$$

Now in order to actually complete the process of finding the derivative, we take what is called the *standard part* of the above expression, which basically amounts to neglecting any infinitesimal contributions. In other words, we let the $dx \to 0$, and find that

$$f'(x) = 2x,$$

which is exactly the result we find using limits.

In the above example the process for finding the derivative is essentially the same as using limits. We begin with a fraction where the denominator is approaching 0 (or is infinitesimal), and need to use algebraic manipulation in order to find the finite contributions. The infinitesimal approach simply provides us a different means of looking at derivatives, so we will refer to it from time to time to build intuition.

The differentiation operator is also based on idea of an infinitesimal. We can view the notation

$$\frac{d}{dx},$$

as saying: compute an infinitesimal change in a function and divide it by an infinitesimal change in its input variable.

Given that our notation relies heavily on the infinitesimal approach, one might ask the question why we deal with limits. Historically, it was not until many years after the initial development of calculus that a mathematically rigorous notion of infinitesimal was developed. Although it was acknowledged that infinitesimals *worked*, and could be used to solve problems, there was also a lot of skepticism of the validity of the methods.

It was not until later when Karl Weierstrass provided the notion of limits and continuity that calculus had a rigorous mathematical foundation. These notions of limits and continuity are the same as the ones we have already considered up to now. With this new rigorous base there was a shift in the perceptions of calculus, and up to today the infinitesimal approach is largely ignored or neglected, with little appeal as to how our notation has developed.

Interestingly enough, long after the use limits became more wide spread, a rigorous mathematical base was provided for working with calculus in terms of infinitesimals. Differential geometry and the hyperreal number system both provide means of working with infintesimals rigorously. Because of this, we can feel reasonably comfortable using the intuitive insights that infinitesimals provide us, rather than ignoring their prevalent place in our notation.

Since there is now a rigorous base to calculus with infinitesimals, the natural question to ask is why do we still use limits. The reason we use limits is because it turns out that they are a much broader, more fundamental concept than infinitesimals. From limits we build up the notion of approximation, which can be used in many ways. As seen earlier, we used limits of sequences in order to provide meaning to irrational numbers. As we continue to develop our theory we will encounter more cases where limits provide us with an indispensable tool. By building up our theory using limits, we are able to use the same tool again and again. Notably, we will use limits in order to work in-depth with objects such as sequences and infinite sums of functions. However, we will not come to that place for quite some time. For the moment we'll work in the world of differential and integral calculus where infinitesimals would work just fine. Rather than shunning one particular view of calculus over another, we'll use both perspectives complementarily, noting that two lenses will provide no less light than one in looking at any problem.

4.3 Properties of Differentiation

From our previous work we've seen that it can be quite a task to calculate the derivative of an arbitrary function. Just working with a second-order polynomial things get pretty complicated - imagine computing the derivative of a fifth-order polynomial. It is the same now as it was with limits. Rather than trying to calculate every possible limit individually, we sought rules upon which we could combine known limits, in order to save ourselves a great deal of work. We have already noticed some such rules (and used them) up to this point, but here we will formally acknowledge and develop these rules a bit further.

The first thing that we noticed is that the derivative of a line is just the slope of the line. The immediately corollary is that the derivative of a constant is 0 (because a constant is just a horizontal line). There are two more basic differentiation results which we will simply state at this point - we will not gain a whole lot by trying to prove them or justify them as true. These are results for the derivatives of power and exponential functions.

Theorem 4.3.1 (Basic Derivatives). *We have the following results for derivatives.*

1. $\dfrac{d}{dx}c = 0, \quad c \in \mathbb{R}$
2. $\dfrac{d}{dx}x^p = p \cdot x^{p-1}, \quad p \in \mathbb{R}, p \neq 0$
3. $\dfrac{d}{dx}e^x = e^x$

We have already seen the derivative of a power function for $p = 1$ and $p = 2$, but the above result generalizes to all other p as well. The last of the above derivatives sheds some light on the elusive number e. The exponential function e^x is such that its derivative is itself - the instantaneous rate of change of e^x at all points is equal to its current value. This is quite a remarkable property for a function to have, and interestingly enough the exponential function underlies many physical processes, such as growth and decay. At the moment we don't have the tools to show why this is true, but we'll use this useful fact, adding one more function to our arsenal of differentiable functions.

Example 1. Find the derivative of x^2.

Solution We apply the power rule to find

$$\frac{d}{dx}x^2 = 2x^{2-1} = 2x.$$

Example 2. Find the derivative of $\sqrt{x}$.

Solution First we rewrite $\sqrt{x} = x^{-1/2}$ and then apply the power rule.

$$\frac{d}{dx}x^{1/2} = \frac{1}{2}x^{1/2-1} = \frac{1}{2}x^{-1/2} = \frac{1}{2\sqrt{x}}$$

Since the square root function is not real valued for $x < 0$, the function is not defined for $x < 0$, and its derivative is not defined for $x \leq 0$, as division by 0 is forbidden.

Example 3. Find the derivative of x^{-1}.

Solution Since $x^{-1} = \frac{1}{x}$ is not defined for $x = 0$, neither is its derivative. Nevertheless, we can apply the power rule to find the value of the derivative at all other points, to find

$$\frac{d}{dx}x^{-1} = -x^{-2}.$$

Having looked at derivatives of these basic building blocks, we'd like to generalize our results to combinations of these functions. We very easily gain results for linear combinations of such functions.

Theorem 4.3.2 (Linearity of Differentiation)**.** *Suppose f and g are differentiable functions, $\alpha \in \mathbb{R}$. It follows that*

1. $\dfrac{d}{dx}(\alpha \cdot f(x)) = \alpha \cdot \dfrac{d}{dx} f(x). = \alpha f'(x)$
2. $\dfrac{d}{dx}(f+g)(x) = \dfrac{d}{dx} f(x) + \dfrac{d}{dx} g(x) = f'(x) + g'(x).$

Proof. From the definition of the derivative,

1. $\dfrac{d}{dx}(\alpha \cdot f(x)) = \lim\limits_{h\to 0} \dfrac{\alpha f(x+h) - \alpha f(x)}{h} = \alpha \cdot \lim\limits_{h\to 0} \dfrac{f(x+h) - f(x)}{h} = \alpha \cdot \dfrac{d}{dx} f(x).$
2. $$\begin{aligned} \frac{d}{dx}(f+g)(x) &= \lim_{h\to 0} \frac{(f+g)(x+h) - (f+g)(x)}{h} = \lim_{h\to 0} \frac{f(x+h) - f(x) + g(x+h) - g(x)}{h} \\ &= \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} + \lim_{h\to 0} \frac{g(x+h) - g(x)}{h} = \frac{d}{dx} f(x) + \frac{d}{dx} g(x). \end{aligned}$$

□

The above properties are called linearity, insofar that differentiation respects linear combinations of functions. If we differentiate a linear combination of functions, the result is the same linear combination of their derivatives. The above properties might also be referred to as the constant product and sum rules for differentiation. Since all polynomial functions are simply linear combination of power functions, these rules give us the means to differentiate any polynomial, without going through the arduous task of starting from the definition.

Example 4. For a general object, surface area and volume are related by

$$S = cV^{2/3}.$$

where c is a constant of proportionality, which depends on the given object. Find the derivative of surface area with respect to volume, and describe how the proportionality between surface area and volume changes with changing volume.

Solution We must be careful in that we cannot immediate apply the power rule in this situation, as we have a power function multiplied by a constant. However, by linearity of differentiation, we can factor out the the constant, and then apply the product rule. We find

$$\frac{dS}{dV} = c\frac{d}{dV} V^{2/3} = \frac{2}{3} c V^{-1/3}.$$

The rate at which the surface area changes is proportionate to $\frac{1}{\sqrt[3]{V}}$. Thus, as volume increases, the surface area increases more and more slowly; as an object gets large, the ratio of surface area to volume decreases.

Example 5. Find the derivative of $2e^x + x^4$.

Solution Here we will use linearity of differentiation, in order to see

$$\frac{d}{dx}(2e^x + x^4) = 2 \cdot \frac{d}{dx} e^x + \frac{d}{dx} x^4 = 2 \cdot e^x + 4x^3.$$

Example 6. Find the derivative of $ax^3 + bx^2 + cx + d$, where $a, b, c, d \in \mathbb{R}$.

Solution First we apply the sum rule, then the constant product rule, and finally the power rule to find that

$$\frac{d}{dx}(ax^3+bx^2+cx+d) = \frac{d}{dx}ax^3 + \frac{d}{dx}bx^2 + \frac{d}{dx}cx + \frac{d}{dx}d = a\frac{dx^3}{dx} + b\frac{dx^2}{dx} + c\frac{dx}{dx} + d = 3ax^2 + 2bx + c$$

Example 7. Solve the differential equation

$$\frac{dP}{dt} = t + e^t,$$

given that $P(1) = 2$.

Solution To find a solution we need to perform *educated guesswork* for each of the above terms. We know that if we differentiate $t^2/2$ the result is t, and likewise if we differentiate e^t we will have e^t. We also need to add an arbitrary constant (that is lost in the process of differentiation). Thus,

$$P(t) = \frac{t^2}{2} + e^t + c.$$

Now to find a unique solution, we must use the condition given to us about $P(1)$. Doing so we find

$$P(1) = \frac{1}{2} + e + c = 2,$$

which implies that

$$c = \frac{3}{2} - e.$$

This gives us the unique solution

$$P(t) = \frac{t^2}{2} + e^t + \frac{3}{2} - e = \frac{t^2+3}{2} + e(e^{t-1} - 1).$$

Example 8. Consider an object falling near the surface of the earth. Neglecting wind resistance, the only force acting on the object is gravity, which is given by approximately 9.8 m/s times the mass of the object (and the direction is down, which gives us a minus sign). From Newton's second law (in a nonrelativistic situation) we find that

$$\frac{F}{m} = \frac{d^2x}{dt^2},$$

which in this situations implies

$$\frac{d^2x}{dt^2} = -9.8 \text{ m/s}^2.$$

Supposing that the object has an initial position of $x(0) = x_0$ and initial velocity of $v(0) = v_0$, find the position and velocity of the object for all times $t \geq 0$.

Solution Noting that velocity is the first derivative of position, we can rewrite the differential equation in terms of the first derivative of v, rather than the second derivative of x. Doing so we find

$$\frac{dv}{dt} = -9.8.$$

Using our knowledge of differentiation (and an educated guess) we can see that the family of functions of the form

$$v(t) = -9.8t + c$$

where $c \in \mathbb{R}$ all provide solutions to the differential equation. However, an object can only have one velocity, not an entire family of velocities, so we need use the initial condition $v(0)$ to pick out the correct solution from this family of solutions. Evaluating the general solution at $t = 0$ we find

$$v(0) = -9.8 \cdot 0 + c = v_0.$$

This implies $c = v_0$, and $v(t) = -9.8t + v_0$. We can write this as another differential equation in order to find the position of the object, $x(t)$. We find that

$$\frac{dx}{dt} = -9.8t + v_0,$$

and using a little more educated guesswork we find

$$x(t) = -4.9t^2 + v_0 t + c$$

where c is another arbitrary constant. Now we use our other initial condition, $x(0) = x_0$. Evaluating $x(t)$ at $t = 0$ we find

$$x(0) = -4.9 \cdot 0 + v_0 \cdot 0 + c = x_0$$

which implies $c = x_0$. The overall solution is that we find the position of a falling object is given by

$$x(t) = -4.9t^2 + v_0 t + x_0.$$

In addition to finding the derivatives of linear combination of known functions, we can also find the derivative of products of known functions. One would probably guess immediately that the result is simply the product of the individual derivatives. However, we can verify immediately that this is not true. Think about multiplying the functions $f(x) = x$ and $g(x) = x$ together. Each of these individual derivatives are 1, yet the derivative of the product is $2x$. Thus, the process of finding the derivative of a product of functions is not as simple as just finding the product of the derivatives. In order to unravel this mystery we'll need to look at things a bit differently.

Let's suppose that we have two functions f and g. One way of interpreting the product $f(x) \cdot g(x)$ is as an area of a square, where the value of each of these functions represents the length of one of the sides. In order to find the derivative of this function, we're going to need to look at

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

If we think about modifying these functions by a little bit, to consider $f(x+h)$ and $g(x+h)$ then we have a new square, with these new respective lengths. For the sake of illustration, let us suppose that $f(x+h)$ and $g(x+h)$ are larger than $f(x)$ and $g(x)$ (this is not required, but we would need to draw a number of pictures for each different situation otherwise). When we do so, we can view the original square as residing inside this new, larger square we have created (see figure 4.4).

Our next task is to look at this new, larger area, in terms of the original area. Using the square $f(x)g(x)$ as our guide, we break up this larger square into four pieces. When we write the total area in terms of these four squares, we find that

$$\begin{aligned} f(x+h)g(x+h) &= f(x)g(x) + \big(g(x+h) - g(x)\big)f(x) \\ &\quad + \big(f(x+h) - f(x)\big)g(x) + \big(f(x+h) - f(x)\big)\big(g(x+h) - g(x)\big), \end{aligned}$$

or

$$\begin{aligned} f(x+h)g(x+h) - f(x)g(x) &= f(x)g(x) + \big(g(x+h) - g(x)\big)f(x) + \big(f(x+h) - f(x)\big)g(x) \\ &\quad + \big(f(x+h) - f(x)\big)\big(g(x+h) - g(x)\big). \end{aligned}$$

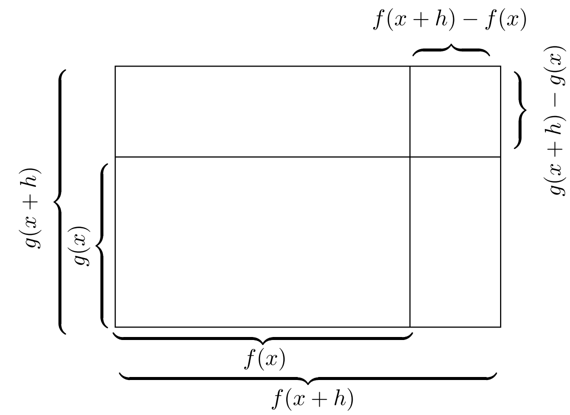


Figure 4.4: Illustration of the areas involved in the derivative of a product of functions.

If we look at the left-hand side of the above equation, this is exactly the form that shows up in the definition of the derivative of the product. We simply need to divide by h and look in the limit as $h \to 0$. We find then that

$$\begin{aligned}
\frac{d}{dx}(f \cdot g)(x) &= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} \frac{(g(x+h)-g(x))f(x)}{h} + \lim_{h\to 0} \frac{(f(x+h)-f(x))g(x)}{h} \\
&\quad + \lim_{h\to 0} \frac{(f(x+h)-f(x))(g(x+h)-g(x))}{h} \\
&= g'(x)f(x) + f'(x)g(x) + \lim_{h\to 0} \frac{(f(x+h)-f(x))(g(x+h)-g(x))}{h}.
\end{aligned}$$

Now we still need to evaluate this final limit, corresponding to the little corner of the big square. Intuitively, as we let $h \to 0$, both sides of this square are shrinking on the order of h, so this little square shrinks much faster than the others, and becomes negligible. Alternatively, we can think of multiplying this term by a convenient choice of 1, h/h, and since both $f'(x)$ and $g'(x)$ exist, it follows that

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

Thus, this final term vanishes. In summary, we have the following result.

Theorem 4.3.3 (Product Rule for Derivatives)**.** *Suppose that f and g are functions differentiable at x. It follows that*

$$\frac{d}{dx}(f \cdot g)(x) = f(x)\frac{d}{dx}g(x) + g(x)\frac{d}{dx}f(x) = f(x)g'(x) + f'(x)g(x).$$

Example 9. Find the derivative of $p(x) = (x-1)(x+1)$.

Solution Let $f(x) = x - 1$, and $g(x) = x + 1$. Since these are both linear functions with slope 1, it follows that $f'(x) = 1 = g'(x)$. Now we can apply the product rule

$$p'(x) = f(x)g'(x) + g(x)f'(x) = (x-1) \cdot 1 + (x+1) \cdot 1 = 2x.$$

In this situation, we can also multiply the two polynomials out, and find that

$$p(x) = x^2 - 1$$

Applying the sum and product rules as we did in the last section we find

$$p'(x) = 2x + 0 = 2x$$

which is consistent with the product rule.

Example 10. Suppose the volume of a plant is described by

$$V(t) = 100 + 12t,$$

where t is measured in days, and V is measured in cm^3. Let the density of the plant decrease according to

$$\rho(t) = 0.8 - 0.05t,$$

where ρ is measured in grams per cm^3. Describe the change of plant's mass over time.

Solution First recall that mass is related to volume and density by

$$M(t) = \rho(t)V(t)$$

Now we can apply the product rule to find the rate of change of the mass

$$M'(t) = \rho(t)V'(t)+\rho'(t)V(t) = (0.8-0.05t)12+(100+12t)(-0.05) = 9.6-0.6t-5-0.6t = 4.6-1.2t$$

Looking at the rate of change of mass as a function of time, we can see that initially the mass is increasing, but the rate it is increasing at decreases with time. Eventually, when $t > \frac{23}{6}$, the derivative becomes negative, so the mass begins to decrease.

In addition to using limits to evaluate derivatives, we can go through the opposite process. If we can recognize a limit simply as the derivative of a known function, then we can save a lot of work in evaluating the limit.

Example 11. Evaluate $\lim\limits_{h\to 0} \dfrac{2(x+h)^2 - 2x^2}{h}$.

Solution One way to work through this problem is to just blindly begin evaluating the limit. However, if we recognize the form of this limit is simply the derivative of $2x^2$, then we can use the power rule and save ourselves all of the work. We find that

$$\lim_{h\to 0} \frac{2(x+h)^2 - 2x^2}{h} = \frac{d}{dx}2x^2 = 4x.$$

4.4 The Chain Rule

A basic illustration of the chain rules comes in thinking about runners in a race. Suppose two brothers, Mark and Brian, hold an annual race to see who is the fastest. Last year Mark won the race, so this year he will be considered the pace setter. Let M and B be functions that represent the position of each of the brothers at any given time. Since Mark is the pace setter, his position only depends on time, so we write $M(t)$. In contrast, we measure Brian's velocity (and thus position) relative to Mark's, so we write $B(M(t))$. If Brian is running $2/3$ as fast as Mark, then Brian's velocity is simply

$$\frac{dB}{dt} = \frac{2}{3} \cdot \frac{dM}{dt},$$

meaning that Brian's position changes at a rate $2/3$ of that of Mark's. Knowing Mark's velocity (or instantaneous rate of change of position with time) we can find Brian's velocity just by multiplying by $2/3$, the rate at which Brian's position is changing relative to Mark's. Thus, if Mark is running at 10 mph, it follows Brian is running at $20/3 \approx 6.67$ mph. Written symbolically, we observe

$$\frac{dB}{dt} = \frac{dB}{dM} \cdot \frac{dM}{dt}.$$

This is called the chain rule. It states that if we know the rate of change of Brian's position with respect to Mark's and we know the rate of change of Mark's position with respect to time, we can find the rate of change of Brian's position with respect to time, simply by multiplying them together. If we add a third runner into the race, Kevin, who is running twice as fast as Brian, then his velocity is given by

$$\frac{dK}{dt} = 2 \cdot \frac{dB}{dT} = 2 \cdot \frac{dB}{dM} \cdot \frac{dM}{dt}.$$

Noting that 2 is just the rate of change of Kevin's position with respect to Brian's, we find that

$$\frac{dK}{dt} = \frac{dK}{dt} \cdot \frac{dB}{dM} \cdot \frac{dM}{dt}.$$

In this way we can *chain together* any number of functions, and find the overall rate of change by looking at the relative rates of change at each step. Note that here we have the composition $K(B(M(t)))$, so we first differentiate K with respect to B, B with respect to M, and finally M with respect to t.

Theorem 4.4.1 (The Chain Rule). *Suppose f is differentiable at $u(x)$ and u is differentiable at x. It follows that the composition $f \circ u$ is differentiable at x, and*

$$\frac{d}{dx} f(u(x)) = \frac{df}{du} \cdot \frac{du}{dx} = f'(u(x)) \cdot u'(x).$$

In the composition $f(u(x))$ we call f the outer function and u the inner function. The basic mechanism of the chain rule is as follows: differentiate the outer function holding the inner function as a constant, then multiply the result by the derivative of the inner function. If there is a composition of more than two functions, the above process is simply repeated as many times as necessary.

Example 1. Find the derivative of $(2x + 1)^4$ with respect to x.

Solution One means of finding this derivative would be to expand the polynomial and use the power rule. However, this would be rather time consuming, and we can find this derivative much more easily using the chain rule. Note that the above function is a composition of the functions u^4 and $u = 2x + 1$. The chain rule tells us that the derivative is equal to

$$\frac{d}{dx}(2x+1)^4 = \frac{d}{du}u^4 \cdot \frac{d}{dx}(2x+1) = 4u^3 \cdot 2 = 8(2x+1)^3.$$

Example 2. Find the derivative of $\frac{1}{1+y^2}$ (with respect to y).

Solution In order to solve this problem it is helpful to first rewrite the function as $(1+y^2)^{-1}$. Then we can see there is a composition of two functions u^{-1} and $u = 1 + y^2$. Applying the chain rule we find

$$\frac{d}{dy}(1+y^2)^{-1} = -(1+y^2)^{-2} \cdot 2y = \frac{-2y}{(1+y^2)^2}.$$

Example 3. Find the derivative of $e^{\alpha t}$ (with respect to t), $\alpha \in \mathbb{R}$.

Solution The above function is a composition of two functions, e^u and $u = \alpha t$. Thus, we can apply the chain rule. We take the derivative of the outer function (which is e^u), evaluate the result at the inner function ($u = \alpha t$), differentiate the inner function (yielding α), and then multiply the results.

$$\frac{d}{dt}e^{\alpha t} = e^{\alpha t} \cdot \alpha = \alpha e^{\alpha t}.$$

Thus, to find the derivative of an exponential function where the argument is multiplied by a constant, simply multiply the exponential function by that constant. This is consistent with the derivative of the exponential function, where that constant is simply a 1.

When a quantity satisfies the differential equation

$$\frac{dN}{dt} = \alpha N$$

it is said to be either growing or decaying exponentially, depending on whether or not $\alpha > 0$ or $\alpha < 0$. From the above example we can see that a solution to the above differential equation is given by

$$N(t) = e^{\alpha t}.$$

It turns out that if we multiply this solution by a constant we still have a solution to the differential equation, so the general solution is given by

$$N(t) = N_0 e^{\alpha t},$$

where N_0 is a constant. We would need to have some initial condition in order to solve for N_0.

Example 4. Solve the initial value problem

$$\frac{dN}{dt} = -2t, \quad N(0) = 3.$$

Solution We know the general solution of this equation is given by

$$N(t) = N_0 e^{-2t},$$

which can be verified through differentiation. Using the initial condition $N(0) = 2$ we find

$$3 = N(0) = N_0 e^0 = N_0,$$

so the solution to the initial value problem is given by

$$N(t) = 3e^{-2t}.$$

Example 5. Find the derivative of e^{x^2}.

Solution Once again we can apply the chain rule, to the composition of e^u and $u = x^2$. We take the derivative of the outer function (which is e^u), evaluate the result at the inner function ($u = x^2$), differentiate the inner function (yielding $2x$), and then multiply the results.

$$\frac{d}{dx}e^{x^2} = e^{x^2} \cdot 2x.$$

Example 6. Find the derivative of a^t (with respect to t), where $a \in \mathbb{R}^+$.

Solution Using the chain rule we can find the derivative of a base a exponential using the derivative of the base e exponential. Write $a = e^{\ln(a)}$, which can be done as the exponential function and natural logarithm are inverses (as long as a is a positive real number, as we have not defined the natural logarithm for negative numbers). Now we take the derivative using the chain rule, finding

$$\frac{d}{dt}a^t = \frac{d}{dt}e^{\ln(a)t} = e^{\ln(a)t} \cdot \ln(a) = \ln(a) \cdot a^t.$$

Example 7. Find the derivative of $e^{e^{2x+1}}$.

Solution In this case we're actually looking at a composition of three functions we know how to differentiate. How can we apply the chain rule in this case? Let us choose our two functions as e^u and $u = e^{2x+1}$. The chain rule tells us

$$\frac{d}{dx}e^{e^{2x+1}} = \frac{d}{du}e^u \cdot \frac{d}{dx}e^{2x+1} = e^{e^{2x+1}} \cdot \frac{d}{dx}e^{2x+1}.$$

Now we need to find the derivative of e^{2x+1} to complete the problem. Choosing e^u and $u = 2x + 1$ we find

$$\frac{d}{dx}e^{2x+1} = e^{2x+1} \cdot 2 = 2e^{2x+1}.$$

Substituting this into our first equation we find

$$\frac{d}{dx}e^{e^{2x+1}} = 2e^{e^{2x+1}} \cdot e^{2x+1}.$$

The above example illustrates an extremely useful observation we made earlier. If we have a composition of three functions, to find the derivative we simply the multiply the derivatives of the three component functions (at each respective step) together. Thus, we find that

$$\frac{d}{dx}f(g(h(x))) = \frac{df}{dg}\frac{dg}{dh}\frac{dh}{dx},$$

where we evaluate df/dg at $g(h(x))$ and dg/dh at $h(x)$. We can easily extend the formula above to a composition of as many functions as we like.

Example 8. Find the derivative of $(x^2 + 1)^3 + 2$.

Solution There is nothing specific about the chain rule that tells us *how* we must choose our inner and outer functions; the only requirement is we choose functions that we know how to differentiate (even if we need to use the chain rule again to do so). Let us choose $u^3 + 2$ and $u = x^2 + 1$ as our functions, because we know how to differentiate both of them. We find

$$\frac{d}{dx}((x^2 + 1)^3 + 2) = 3(x^2 + 1)^2 \cdot 2x.$$

Choosing those two functions as our *chain* of functions worked well because we knew how to differentiate both of them. Nevertheless, another *chain* of functions would work just as well. Let us choose u^3+2, $u=v+1$ and $v=x^2$. Then we find

$$\frac{d}{dx}((x^2+1)^3+2)=3(x^2+1)^2\cdot 1\cdot 2x=3(x^2+1)^2\cdot 2x,$$

where $du/dv=1$ and $dv/dx=2x$. This result is exactly the same as our previous result, as it must be. The lesson to be learned here is that the most important thing to consider with the chain rule is choosing functions that are easy to differentiate. As long as the composition or *chain* of functions chosen to represent the original function is equivalent, it doesn't matter how many functions are chosen.

Example 9. Find the derivative of $y=(1+a^{2t})^{-4}$ with respect to t.

Solution Letting our representative functions be u^{-4}, $u=1+a^v$ and $v=2t$ we find

$$\frac{dy}{dt}=-4(1+a^{2t})^{-5}\cdot\ln(a)a^{2t}\cdot 2=-\frac{8\ln(a)a^{2t}}{(1+a^{2t})^5}.$$

Using the chain rule in conjunction with the product rule we can derive a formula for directly taking the derivative of a quotient of functions with known derivatives. We begin by writing

$$u(x)/v(x)=u(x)\cdot(v(x))^{-1},$$

where $(v(x))^{-1}$ is raised the negative first power, not an inverse function. Differentiating we find

$$\begin{aligned}\frac{d}{dx}(u(x)\cdot(v(x))^{-1}) &= u'(x)\cdot(v(x))^{-1}+u(x)\cdot(-(v(x))^{-2}\cdot v'(x)\\ &= \frac{(v(x))^2}{(v(x))^2}\cdot\Big(u'(x)\cdot(v(x))^{-1}+u(x)\cdot(-(v(x))^{-2}\cdot v'(x)\Big)\\ &= \frac{u'(x)v(x)-u(x)v'(x)}{(v(x))^2}.\end{aligned}$$

This result is called the quotient rule.

Theorem 4.4.2 (Quotient Rule for Derivatives)**.** *Suppose u and v are both differentiable at x, and $v(x)\neq 0$. It follows that*

$$\frac{d}{dx}\Big(\frac{u(x)}{v(x)}\Big)=\frac{v(x)\frac{d}{dx}u(x)-u(x)\frac{d}{dx}v(x)}{[v(x)]^2}=\frac{v(x)u'(x)-u(x)v'(x)}{[v(x)]^2}.$$

This rule can be remembered as "low 'd' high minus high 'd' low over low squared." 'High' is the numerator of the quotient, where 'low' is the denominator. Finally, 'd' represents the derivative of. When one is faced with the task of differentiating a quotient, one can either use the quotient rule directly, or simply use the product and chain rules in conjunction, rather than working with (and remembering) the quotient rule.

Example 10. Find the derivative of $\frac{x^2+1}{x-1}$.

Solution Let $u(x)=x^2+1$, and $v(x)=x-1$. It follows that $u'(x)=2x$ and $v'(x)=1$. Now we can use the quotient rule

$$\frac{d}{dx}\frac{u(x)}{v(x)}=\frac{v(x)u'(x)-u(x)v'(x)}{v(x)^2}=\frac{(x-1)\cdot 2x-(x^2+1)\cdot 1}{(x-1)^2}=\frac{2x^2-2x-x^2-1}{(x-1)^2}=\frac{x^2-2x-1}{(x-1)^2}.$$

Example 11. The set of Hill functions is a family of functions used to describe stimulus response in the study of biology. Hill functions take on the general form

$$h(x) = \frac{x^n}{1 + x^n}$$

where $n \in \mathbf{Z}^+$ (is a positive integer $1, 2, \ldots$). Calculate the derivative of the Hill functions for $n = 1$ and $n = 2$.

Solution We will begin with $n = 1$. Let $u(x) = x$ and $v(x) = 1 + x$. It follows that $u'(x) = v'(x) = 1$. Using the quotient rule we find

$$\frac{dh}{dx} = \frac{(1 + x) \cdot 1 - x \cdot 1}{(x + 1)^2} = \frac{1}{(x + 1)^2}$$

This derivative is always positive for $x > 0$, and $h'(0) = 1$. Now let $n = 2$, so $u(x) = x^2$ and $v(x) = 1 + x^2$. It follows that $u'(x) = v'(x) = 2x$. From the quotient rule we find

$$\frac{dh}{dx} = \frac{(1 + x^2) \cdot 2x - x^2 \cdot 2x}{(x^2 + 1)^2} = \frac{2x}{(x^2 + 1)^2}.$$

This derivative is also always positive for $x > 0$, but has the value $h'(0) = 0$.

Given two functions that are differentiable, the chain rule tells us how to find the derivative of their composition. Nevertheless, it may be possible to compose functions which are not differentiable and arrive at a result which is differentiable. Let us consider the functions

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

and

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

These in and of themselves seem to be very esoteric functions, so let's take a minute to think about what they really are. For the function f, whenever the input is a rational number, we get an output of 1. Otherwise, for an irrational number, the output is 0. This function is so badly broken up that we cannot draw it, but we can note that it is discontinuous everywhere (because between any two rationals is an irrational, and between any two irrationals is a rational). The second function u behaves similarly to f, except that it is 0 for rationals and 1 for irrationals. It is equally badly behaved.

Nevertheless, when we look at the composition $f(u(x))$, something remarkable happens. No matter the input, u outputs a rational number (either 0 or 1), so the output of $f(u(x)) = 1$ for all x. This function is continuous everywhere, as well as differentiable, as $(f(u(x)))'(x) = 0$ for all x. There is no inconsistency here with the chain rule, because the component functions do not satisfy the hypothesis of the chain rule; the chain rule tells us nothing about the derivative of a composition of functions which are not differentiable.

The above discussion elucidates the fact that the chain rule

$$\frac{df}{dx} = \frac{df}{du} \cdot \frac{du}{dx}$$

is not verified by simply canceling the differentials du on the right-hand side of the equation. If we could simply *cancel* differentials then this would be a trivial identity. Similarly, if we could *cancel* the differentials then we should be able add them in, using whatever intermediate functions we desire. However, we just saw that it is possible for the derivative df/dx to exist, even if both df/du and du/dx do not. Thus, it would be a grave mistake to add in or cancel additional differentials and conclude that the product of two nonexistent terms exists.

4.5 Derivatives of Trigonometric Functions

In order to find the derivative of the sine function, we begin with the definition of the derivative. Let $f(x) = \sin(x)$. Then,

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{\sin(x+h) - \sin x}{h} = \lim_{h\to 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \lim_{h\to 0} \Big(\frac{\sin x(\cos h - 1)}{h} + \frac{\cos x \sin h}{h} \Big) = \sin x \cdot \lim_{h\to 0} \frac{\cos h - 1}{h} + \cos(x) \cdot \lim_{h\to 0} \frac{\sin h}{h} \end{aligned}$$

We have already seen that the limit

$$\lim_{h\to 0} \frac{\sin h}{h} = 1,$$

and we can actually use this limit to find the value of the other one using the half angle formula $\cos h = 1 - 2\sin^2(h/2)$. We find

$$\lim_{h\to 0} \frac{\cos h - 1}{h} = \lim_{h\to 0} -\frac{2\sin^2(h/2)}{h} = -\lim_{\theta\to 0} \frac{\sin\theta}{\theta} \sin\theta = -1 \cdot 0 = 0,$$

where in the above analysis we substituted $h/2 = \theta$ to evaluate the limit. Using these two limits, we find that

$$f'(x) = \sin x \cdot 0 + \cos x \cdot 1 = \cos x.$$

This is a very interesting result. When we look at the rate of change of the function $\sin(x)$ we find that it is the other sinusoidal function $\cos(x)$. In fact,

$$\cos(x) = \sin(x + \frac{\pi}{2}),$$

or in other words, the instantaneous rate of change of $\sin(x)$ at any point x is simply the value of $\sin(x + \pi/2)$, the value of the same function $\pi/2$ to the right of the point of interest. Thus, $\sin(x)$ is a function which has rate of change directly related to itself. From a global perspective we see that the derivative of the sine function is a $\pi/2$ horizontal left-shift of itself.

We could go through similar analysis to find the derivative of the cosine function, but by noting that cosine is just a horizontal shift of the sine function, it follows that its derivative will just be a horizontal shift of the derivative of the sine function. Namely, we shift by another $\pi/2$ radians, and find

$$\frac{d}{dx}\cos(x) = \sin(x + \pi) = -\sin(x).$$

It is useful to look at the graph of a function and its derivative together, to see just how much information is contained in the derivative. Notice that at the peaks of the sine function, the cosine function is 0. This is because the line tangent to the sine function is horizontal at these points. At the peaks of the cosine function (the derivative of sine) the sine function crosses the x-axis - these are the points where the sine function has the greatest slope, or is changing the most rapidly. The graphs of these trigonometric functions also give us a clue as to which derivative contains the negative sign. At $x = 0$, $\sin(x)$ is increasing, and $\cos(x)$ is positive, so it makes sense that the derivative is $+\cos(x)$. On the other hand, just after $x = 0$, $\cos(x)$ is decreasing, and $\sin(x)$ is positive, so the derivative must be $-\sin(x)$.

Example 1. Find all derivatives of $\sin(x)$.

Solution Since we know $\cos(x)$ is the derivative of $\sin(x)$, if we can complete the above task, then we will also have all derivatives of $\cos(x)$.

$$\frac{d}{dx}\sin(x) = \cos(x)$$

gives us the first derivative of the sine function.

$$\frac{d^2}{dx^2}\sin(x) = \frac{d}{dx}\cos(x) = -\sin(x)$$

gives us the second derivative. Also

$$\frac{d^3}{dx^3}\sin(x) = \frac{d}{dx}(-\sin(x)) = -\frac{d}{dx}\sin(x) = -\cos(x).$$

Finally,

$$\frac{d^4}{dx^4}\sin(x) = -\frac{d}{dx}\cos(x) = \sin(x).$$

Now we can see that the fourth derivative of $\sin(x)$ is $\sin(x)$, so we can easily enough find any derivative of the sine function as follows. Suppose we want to find the n^{th} derivative of sine. All we need to do is divide n by 4, and look at the remainder r. If we take the r^{th} derivative of sine, it will be exactly the same as taking the n^{th} derivative, as every four derivatives will simply return us to the original result of the sine function. Applying this principle, we find that the 17^{th} derivative of the sine function is equal to the 1^{st} derivative, so

$$\frac{d^{17}}{dx^{17}}\sin(x) = \frac{d}{dx}\sin(x) = \cos(x)$$

The derivatives of $\cos(x)$ have the same behavior, repeating every cycle of 4. The n^{th} derivative of cosine is the $(n+1)^{th}$ derivative of sine, as cosine is the first derivative of sine.

Rather than mechanically calculating derivatives as above, the graphical perspective provides a much more elegant solution to this problem. If we recognize that differentiating sine results in shifting the function $\pi/2$ to the left, then differentiating twice should shift it π to the left, and so on. Since the sine function is 2π-periodic, it then follows that we must have our derivatives repeat in cycles of 4, because everytime we differentiate 4 times we have shifted the function by 2π, returning it to its original position.

Knowledge of the derivatives of sine and cosine allows us to find the derivatives of all other trigonometric functions using differentiation rules. Recall the definitions of the other trigonometric functions are as follows:

$$\tan(x) = \frac{\sin(x)}{\cos(x)} \quad \cot(x) = \frac{\cos(x)}{\sin(x)} \quad \sec(x) = \frac{1}{\cos(x)} \quad \csc(x) = \frac{1}{\sin(x)}$$

In order to find the derivatives of these functions we will also need to use a few forms of the Pythagorean identity

$$\sin^2(\theta) + \cos^2(\theta) = 1.$$

From here we can divide both sides of the equation by $\cos^2(\theta)$ to find

$$1 + \tan^2(\theta) = \sec^2(\theta).$$

If we divide both sides of the equation by $\sin^2(\theta)$ we find

$$1 + \cot^2(\theta) = \csc^2(\theta).$$

Keeping these identities in mind, we will look at the derivatives of the other trigonometric functions.

Example 2. Find the derivatives of $\tan(x)$, $\cot(x)$, $\csc(x)$, and $\sec(x)$.

Solution We can find all of the above derivatives using the quotient rule and the derivatives of sine and cosine. Starting with tangent

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

and moving to cotangent

$$\frac{d}{dx}\cot(x) = \frac{d}{dx}\frac{\cos(x)}{\sin(x)} = \frac{\sin(x)(-\cos(x)) - \cos(x)\sin(x)}{\sin^2(x)} = \frac{-\sin^2(x) - \cos^2(x)}{\sin^2(x)} = \frac{-1}{\sin^2(x)} = -\csc^2(x)$$

next cosecant

$$\frac{d}{dx}\csc(x) = \frac{d}{dx}\frac{1}{\sin(x)} = \frac{\sin(x)\cdot 0 - 1\cdot\cos(x)}{\sin(x)^2} = -\frac{\cos(x)}{\sin(x)}\cdot\frac{1}{\sin(x)} = -\cot(x)\csc(x)$$

and finally secant

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

Example 3. Find the derivative of $x\cdot\cos(x) + x^2$.

Solution To find this derivative, we will utilize the sum and product rules.

$$\frac{d}{dx}(x\cdot\cos(x) + x^2) = \cos(x) + x\cdot(-\sin(x)) + 2x = \cos(x) + x(2 - \sin(x))$$

Example 4. Find the derivative of the sinc function, which is commonly encountered in signal processing. Note that

$$\operatorname{sinc}(x) = \frac{\sin(x)}{x}.$$

Solution We will use the quotient rule and our knowledge that the derivative of the sine function is cosine to find

$$\frac{d}{dx}\operatorname{sinc}(x) = \frac{x\cos(x) - \sin(x)}{x^2}.$$

Example 5. Find the derivative of $\sin(1/x)$.

Solution Here we have $\sin(u)$ and $u = 1/x$. Applying the chain rule

$$\frac{d}{dx}\sin(1/x) = \cos(1/x)\cdot(-x^{-2}) = -\frac{\cos(1/x)}{x^2}.$$

Example 6. Find the derivative of $x\sin(1/x)$

Solution We have already used the chain rule to find the derivative of $\sin(1/x)$, so now we just need to use the product rule

$$\frac{d}{dx}x\sin(1/x) = \sin(1/x) + x\cdot(-\frac{\cos(1/x)}{x^2}) = \sin(1/x) - \frac{\cos(1/x)}{x}.$$

Example 7. Find the derivative of $\cos(x)\tan(x)$.

Solution Applying the product rule

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

of course we should find the above result, because

$$\cos(x)\tan(x) = \cos(x)\frac{\sin(x)}{\cos(x)} = \sin(x)$$

and the derivative of the sine function is the cosine function. It is useful to check if a product or quotient of trigonometric functions can be simplified; after all, all of the trigonometric functions are defined directly in terms of sine and cosine.

We have found that the derivatives of the trigonometric functions exist at all points in their domain. For instance, $\tan(x)$ is differentiable for all $x \in \mathbb{R}$ with $x \neq \pi/2 + 2n\pi$ (the points where cosine is 0). It follows immediately that $\tan(x)$ must be continuous at all of these points (because a discontinuity would preclude differentiability). Using this information, we can easily evaluate limits involving trigonometric functions.

Example 8. Evaluate $\lim_{x\to 0} \frac{\sqrt{2+\sec(x)}}{\cos(\pi - \tan(x))}$

Solution Since we are looking at sums, quotients, and a composition of functions which are continuous at $x = 0$, we can simply plug in $x = 0$ to evaluate the limit

$$\lim_{x\to 0} \frac{\sqrt{2+\sec(x)}}{\cos(\pi - \tan(x))} = \frac{\sqrt{2+\sec(0)}}{\cos(\pi - \tan(0))} = \frac{\sqrt{2+1}}{\cos(\pi - 0)} = \frac{\sqrt{3}}{-1} = -\sqrt{3}$$

Example 9. Evaluate $\lim_{x\to 1} x\sec(\pi x/2)$

Solution In this example,

$$\sec(\pi/2) = \frac{1}{\cos(\pi/2)}$$

which is not defined. Thus, we cannot simply plug in for the limit. If we look as $x \to 1$, we see that $\cos(\pi x/2) \to 0$, which implies the fraction goes to $\pm\infty$, depending on which side we look at the limit from. Thus, the limit does not exist.

The sine and cosine functions occur in countless applications throughout the physical sciences. In electric circuits, AC currents are described by sine and cosine functions, the so-called sinusoidal functions. Another important example is the simple harmonic oscillator. Harmonic oscillators are frequently encountered in physics, because numerous phenomena behave in an identical mathematical fashion, such as a weight on a spring, the motion of a pendulum, and LC circuits.

We will focus on the mechanical simple harmonic oscillator - a weight on a frictionless spring. Imagine a spring which is protruding sideways, with a weight resting on a frictionless track. Such a spring has an equilibrium position, at which it is not too compressed or stretched. When brought out of equilibrium the spring exerts a restoring force, which is given by

$$F = -kx.$$

Here k is the spring's spring constant (a parameter which describes the stiffness of the spring) and x is the displacement from equilibrium. We can see from the negative sign that the force acts opposite to direction of the displacement, tending to bring the system back to equilibrium, hence the name restoring force. The simple harmonic oscillator is the simplest of oscillating systems we can consider, because it is an idealization which ignores the dampening force of friction. Since the restoring force is the only force considered, Newton's second law becomes

$$m\frac{d^2x}{dt^2} = -kx.$$

In order to solve this differential equation we can first divide by m, to arrive at

$$\frac{d^2x}{dt^2} = -\frac{k}{m}x.$$

As we have done previously, inspection and a clever choice of a function will be our method to solve this differential equation. As is often useful, let us first try and infer something about the form of the solution, before working out the exact details. The above differential equation tells us that we need to find a function whose second derivative is some constant times the original function itself. The only functions we have encountered so far that satisfy this property are the exponential function and the sine and cosine functions. In order to have the derivative of one of these functions be a constant times itself, we will need to have some composition, so that the inside of the function will yield a constant multiple when we differentiate using the chain rule. Since we are differentiating twice, this constant multiple will *pop out* twice.

Let's suppose that our solution is of the form

$$e^{\alpha t},$$

for some α. If we differentiate twice we will arrive at

$$\alpha^2 e^{\alpha t}.$$

In order for such a function to satisfy the above differential equation, we must have that

$$\alpha^2 = -\frac{k}{m}.$$

This will only occur of α is an imaginary number, which we are not prepared to deal with at this time[1]. Since we're not looking for an exponential solution, we will try to find a sinusoidal solution. Consider a function of the form

$$\cos(\omega t).$$

When we differentiate twice we arrive at

$$-\omega^2 \cos(\omega t),$$

so we see this function satisfies the differential equation when

$$\omega^2 = \frac{k}{m}.$$

Thus, we find

$$\cos(\frac{k}{m}t)$$

to be a solution. Similarly, we can find that

$$\sin(\sqrt{\frac{k}{m}}t)$$

[1]This is not entirely true - we could simply treat the complex exponential function like a regular exponential and we would more or less be fine. The fact that the complex exponential provides us with a solution is related to Euler's identity, that $e^{ix} = \cos(x) + i\sin(x)$.

is a solution. Since differentiation is linear, if we multiply either of these solutions by a constant it still remains a solution. We also find that a sum of the above two solutions is a solution (which can easily be verified through differentiation), so the general solution is given by

$$A\cos(\sqrt{\frac{k}{m}}t) + B\sin(\sqrt{\frac{k}{m}}t).$$

From this general solution we can match the given initial conditions of an initial value problem to find the position of the weight for all subsequent times.

Example 10. Let the position of a mass on a spring be given by $x(t) = A\cos(\omega t)$. Find the velocity and acceleration. What can be said about the motion of the simple harmonic oscillator?

Solution To find the velocity and acceleration we differentiate twice, finding

$$v(t) = \frac{dx}{dt} = -A\omega\sin(\omega t),$$

and

$$a(t) = \frac{dv}{dt} = -A\omega^2\cos(\omega t).$$

We make the following observations

1. Because the position of the mass is given by $A\cos(\omega t)$, we see that the mass oscillates between -A and A, where we call A the amplitude of the oscillations.

2. The acceleration is exactly opposite to the position of the object. This is consistent with the notion of the restoring force. When the mass is out of equilibrium, there is a force exerted by the spring (resulting in an acceleration) which pulls or pushes the mass back to the equilibrium position. The acceleration is 0 (and so is the force exerted by the spring) only when the position of the mass is equilibrium.

3. The peaks of the velocity function correspond to the zeroes of the position and acceleration functions, and conversely, the zeroes of the velocity function correspond to the peaks of the position and acceleration functions. This result is of particular physical significance, and corresponds to conservation of energy in a physical system. There are two forms of energy in the system - kinetic energy in the motion of the spring, and potential energy in the coils of the spring. When the spring is in its equilibrium position, the velocity is at its maximum - there is no potential energy in the spring, and all energy is kinetic in the motion of the spring. When the position of the mass is at a peak, the velocity is zero - all the energy is stored potential within the coils of the spring, and there is no kinetic energy because the velocity is zero. This is characteristic of an oscillating system - oscillations occur in correspondence with the exchange of energy between two different forms.

   Compare this situation to an oscillating LC circuit, where energy is exchanged between a capacitor and inductor. In an LC circuit the peaks and zeroes of the oscillations correspond to where either all of the energy is stored in the charged capacitor, or in the magnetic field generated by the current flowing through the inductor.

4.6 Derivatives of Inverse Functions

The chain rule not only provides us with a means of differentiating compositions, but inverse functions as well, through the clever usage of compositions. If we have a function f and its inverse f^{-1}, by definition

$$(f^{-1} \circ f)(x) = (f \circ f^{-1})(x) = x.$$

Now if two sides of an expression are equal, it follows that they must remain equal after differentiation (after all, how can the same function have two different derivatives?). Using this observation along with the chain rule, we find that

$$(f \circ f^{-1})'(x) = f'(f^{-1}(x))(f^{-1}(x))' = 1.$$

where the right hand side comes from the fact $dx/dx = 1$. Solving this equation we find

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}.$$

This amazing result is enough to allow us to find the derivative of a function we don't know, if we can find the derivative of its inverse function. Through this technique we can find the derivatives of logarithmic functions, and later will be able to use it to find the derivatives of inverse trigonometric functions.

Theorem 4.6.1 (Derivative of Inverse Functions)**.** *Let f be a function differentiable at f, with inverse function f^{-1}. It follows that $(f^{-1})'$ exists, and*

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}.$$

Example 1. Find the derivative of $\ln(x)$

Solution If $f^{-1}(x) = \ln(x)$ then $f(x) = f'(x) = e^x$, so using the above result we find

$$\frac{d}{dx}\ln(x) = \frac{1}{e^{\ln(x)}} = \frac{1}{x},$$

because the exponential undoes the action of the logarithm (leaving us with x) in the above equation.

Example 2. Find the derivative of $\log_a(x)$.

Solution If $f^{-1}(x) = \log_a(x)$ then $f(x) = a^x$ and $f'(x) = \ln(a)a^x$. Using the above result we find

$$\frac{d}{dx}\log_a(x) = \frac{1}{\ln(a)a^{\log_a(x)}} = \frac{1}{\ln(a)x},$$

because the base a exponential undoes the action of the logarithm (leaving us with x) in the above equation.

In addition to providing us with a means of differentiating logarithmic functions, we can use this idea to work with inverse trigonometric functions. Given that all of the trigonometric functions are periodic, they are not one-to-one, and thus noninvertible. In order to find inverses for the trigonometric functions we need to restrict their domains. We do so according to the following table.

Function	Domain	Range
$\sin(x)$	$[-\pi/2, \pi/2]$	$[-1, 1]$
$\cos(x)$	$[0, \pi]$	$[-1, 1]$
$\tan(x)$	$(-\pi/2, \pi/2)$	$(-\infty, \infty)$
$\cot(x)$	$(0, \pi)$	$(-\infty, \infty)$
$\sec(x)$	$[0, \pi/2) \cup (\pi/2, \pi]$	$(-\infty, -1] \cup [1, \infty)$
$\csc(x)$	$[-\pi/2, 0) \cup (0, \pi/2]$	$(-\infty, -1] \cup [1, \infty)$

On these restricted domains we have one-to-one functions, so they have inverses. Correspondingly, we construct the following table of inverse functions.

Function	Domain	Range
$\sin^{-1}(x)$	$[-1, 1]$	$[-\pi/2, \pi/2]$
$\cos^{-1}(x)$	$[-1, 1]$	$[0, \pi]$
$\tan^{-1}(x)$	$(-\infty, \infty)$	$(-\pi/2, \pi/2)$
$\cot^{-1}(x)$	$(-\infty, \infty)$	$(0, \pi)$
$\sec^{-1}(x)$	$(-\infty, -1] \cup [1, \infty)$	$[0, \pi/2) \cup (\pi/2, \pi]$
$\csc^{-1}(x)$	$(-\infty, -1] \cup [1, \infty)$	$[-\pi/2, 0) \cup (0, \pi/2]$

When working with trigonometric functions the input can be thought of an angle, with an output that corresponds to a corresponding ratio of sides on a right triangle, or combination of coordinates on the unit circle. Thus, when we are looking at inverse trigonometric functions our input is a length or ratio of sides, and the output is the corresponding angle.

Working with the chain rule we can find the derivatives of the inverse trigonometric functions. Recall that for a differentiable function f, we have

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}.$$

Let us begin with the sine function, so $f(x) = \sin(x)$ and $f^{-1}(x) = \sin^{-1}(x)$. Applying the above result we have

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

Above we use the identity $\cos^2(x) + \sin^2(x) = 1$, so $\cos(x) = \sqrt{1 - \sin^2(x)}$, noting that cosine is positive over the interval $[-\pi/2, \pi/2]$ where the inverse sine is defined.

We can go through a similar process to find the derivative of the inverse tangent. Here $f(x) = \tan(x)$ and $f^{-1}(x) = \tan^{-1}(x)$. We find that

$$(\tan^{-1}(x))' = \frac{1}{\sec^2(\tan^{-1}(x))} = \frac{1}{1 + \tan^2(\tan^{-1}(x))} = \frac{1}{1 + x^2}.$$

Here we are using the identity $1 + \tan^2(x) = \sec^2(x)$.

We can also proceed to find the derivative of the inverse secant. Here $f(x) = \sec(x)$ and $f^{-1}(x) = \sec^{-1}(x)$. We find that

$$(\sec^{-1}(x))' = \frac{1}{\sec(\sec^{-1}(x))\tan(\sec^{-1}(x))} = \frac{1}{x\tan(\sec^{-1}(x))}.$$

Once again we will use the identity $1+\tan^2(x) = \sec^2(x)$, so $\tan(x) = \pm\sqrt{\sec^2(x) - 1}$. Here we need to be careful about the $\pm$ sign. For $x \in (-\infty, -1]$ we have $\sec^{-1}(x) \in (\pi/2, \pi]$, which corresponds to a negative tangent. For $x \in [1, \infty)$ we have $\sec^{-1}(x) \in [0, \pi/2)$, which corresponds to a positive tangent. Keeping track of these negative signs we find

$$(\sec^{-1}(x))' = \begin{cases} 1/(x\sqrt{\sec^2(\sec^{-1}(x)) - 1}) = 1/(x\sqrt{x^2-1}) & \text{for } x > 1 \\ -1/(x\sqrt{\sec^2(\sec^{-1}(x)) - 1}) = -1/(x\sqrt{x^2-1}) & \text{for } x < -1 \end{cases}$$

We can combine the above expressions into a single expression using the absolute value of x, noting that the derivative is positive for $x > 1$ and negative for $x < -1$. In conclusion, we find

$$(\sec^{-1}(x))' = \frac{1}{|x|\sqrt{x^2-1}}.$$

It is possible to find the derivatives of the other inverse trigonometric identities in the same way as we have done above, but fortunately there's a much simpler way, using a handful of identities.

$$\begin{aligned} \cos^{-1}(x) &= \pi/2 - \sin^{-1}(x) \\ \cot^{-1}(x) &= \pi/2 - \tan^{-1}(x) \\ \csc^{-1}(x) &= \pi/2 - \sec^{-1}(x) \end{aligned}$$

Using any of these identities when we differentiate the constant $\pi/2$, and we will simply pick up a negative sign in front of the derivative we already know. We summarize our results following.

Theorem 4.6.2 (Derivatives of Inverse Trigonometric Functions)**.** *We have the following results for derivatives.*

1. $\dfrac{d}{dx}\sin^{-1}(x) = \dfrac{1}{\sqrt{1-x^2}}, \quad |x| < 1$
2. $\dfrac{d}{dx}\cos^{-1}(x) = -\dfrac{1}{\sqrt{1-x^2}}, \quad |x| < 1$
3. $\dfrac{d}{dx}\tan^{-1}(x) = \dfrac{1}{1+x^2}$
4. $\dfrac{d}{dx}\cot^{-1}(x) = -\dfrac{1}{1+x^2}$
5. $\dfrac{d}{dx}\sec^{-1}(x) = \dfrac{1}{|x|\sqrt{x^2-1}}, \quad |x| > 1$
6. $\dfrac{d}{dx}\csc^{-1}(x) = -\dfrac{1}{|x|\sqrt{x^2-1}}, \quad |x| > 1$

4.7 Differentiating Implicitly Defined Functions

Thus far, the functions we have been concerned with have been defined explicitly. A function is defined explicitly if *the output is given directly in terms of the input.* For instance, for the function $f = 4x^2$ the value of f is given explicitly or directly in terms of the input. Just by knowing the input we can immediately find the output. A second type of function that is also useful for us to consider is an implicitly-defined function. A function is defined implicitly if *the output cannot be found directly from the input.* For instance, $\sqrt{f} = 2x$ is an implicitly-defined function, because for each positive x value there is a corresponding f value, but we cannot find it directly from the function. We first need to square both sides, and then we find the values of f directly from the explicitly-defined function $f = 4x^2$.

In general, it is not possible to rewrite an implicitly-defined function as an explicitly-defined function. Consider the function f given by $\sin(f) + f = x$. For a given x value, there *may* be a corresponding output value f which makes this a true statement. In this way a function f would be defined for all such x where there is a solution. For instance, we have $f(0) = 0$, because setting $x = f = 0$ makes the above equation a true statement. As we have defined them, exponential functions are an entire class of implicitly-defined functions. For instance, f given by $f = 2^x$ is defined as the number that when raised to the $1/x$ power gives 2, even if we don't know how to find that number.

A function is a relation that provides us with at most output for a single input. More generally, we can consider relations that provide us with multiple outputs for a single input, such as $x^2+y^2 = 1$, the equation of the unit circle. We can think of this more general relation as consisting of multiple functions, in this case $y = \pm\sqrt{1-x^2}$.

In general, we can describe an arbitrary curve (like a circle) as a combination of some number of functions. It follows that we should be able to find lines tangent to arbitrary curves, simply by differentiating the function that describes the appropriate portion of the curve. In some situations, like a circle, we can find explicit expressions for the multiple functions our curve consists of. However, in general we will not be able to do so. Nevertheless, because a more complicated curve still consists of functions, we *should* be able to find its derivative at a particular point.

This is a situation in which the power of the operator perspective of differentiation shines through. Just like when we multiply by a constant, add a constant, etc, differentiation should not change the truth of a given equation (if we are differentiating the same or *equal* functions, the result must also be equal). In this way, we can look at the function

$$f = x^2$$

and think about finding the derivative of f with respect to x by differentiating both sides of the equation. When we do so, we find

$$\frac{d}{dx}f = \frac{d}{dx}x^2,$$

or

$$\frac{df}{dx} = 2x,$$

which is a familiar result.

When we differentiate implicitly-defined functions it is essential to keep in mind that f and x are not independent. In the above expression, f has an explicit dependence on x, but x also has an implicit dependence on f. If we change either of these two variables, the other must correspondingly change to make the equation remain true. Now let's suppose we were interested in looking at

$$\frac{d}{df}f = \frac{d}{df}x^2,$$

which would yield

$$1 = \frac{d}{df}x^2.$$

It is clear that the right-hand side of the equation must be nonzero, because otherwise the equation would not be a true statement. However, it is not clear exactly how x^2 changes with respect to f, so we cannot directly evaluate the derivative on the right-hand side. Nevertheless, we can view x as a some unknown function of f, and apply the chain rule. The chain rule tells us to multiply the rate of change of x^2 with respect to x by the rate of change of our intermediate function $x(f)$ with respect to f. In other words,

$$\frac{d}{df}x^2 = \frac{d}{dx}x^2 \cdot \frac{d}{df}x = 2x \cdot \frac{dx}{df}.$$

In the above expression we are left with df/dx, which is an unknown quantity. However, if we equate this expression with the left-hand side of our equation we find

$$1 = 2x\frac{dx}{df},$$

which allows us to solve for dx/df, finding

$$\frac{dx}{df} = \frac{1}{2x}.$$

In essence, when differentiating implicitly-defined functions we need to take into account the fact that there is dependence between our two variables of interest. The chain rule gives us a means of taking this dependence into account, by allowing us to differentiate through an *invisible* intermediate function. Once we differentiate in this way we are left with the unknown derivative, which we are then able to solve for.

Note that our derivative depends on dx/df depends on x. If we look at the graph of the curve $x^2 = f$ we can see the reason for this. This curve consists of the two functions, $\pm\sqrt{f}$. Now suppose that we are interested in finding the line tangent to this curve at some point $f \neq 0$. If the derivative depended solely on f, we would have no way of distinguishing whether or not we were on the function $\sqrt{f}$ or $-\sqrt{f}$, yet clearly the lines tangent to these two functions are different. Thus, we need the derivative to depend on the dependent variable x, because we need information about x to distinguish on which portion of the curve we are on, as different portions of the curve with the same f value have different derivatives.

Example 1. Find df/dx for $\sin(f) + f = x$.

Solution We proceed by differentiating both sides of the given equation.

$$\begin{aligned}
\sin(f) + f &= x \\
\frac{d}{dx}(\sin(f) + f) &= \frac{dx}{dx} \\
\cos(f)\frac{df}{dx} + \frac{df}{dx} &= 1 \\
\frac{df}{dx}(\cos(f) + 1) &= 1 \\
\frac{df}{dx} &= \frac{1}{\cos(f) + 1}.
\end{aligned}$$

Before we proceed, we'd like to introduce the notion of a line that is normal to a curve. The words normal, orthogonal, and perpendicular all have the same meaning, but are generally used in different fields. A line normal to a curve is a line that crosses the line tangent to the curve at a 90 degree angle. The slope of this line is the negative reciprocal of the slope of the tangent line. Thus, if the slope of the tangent line is m, the slope of the normal will be $-1/m$. In some fields, such as geometric optics, the line normal to a curve (or lens) is a particular interest.

Example 2. Find the equation for the lines tangent and normal at the point $(2, 4)$ to the folium of Descartes (see figure 4.5), given by

$$x^3 + y^3 - 9xy = 0.$$

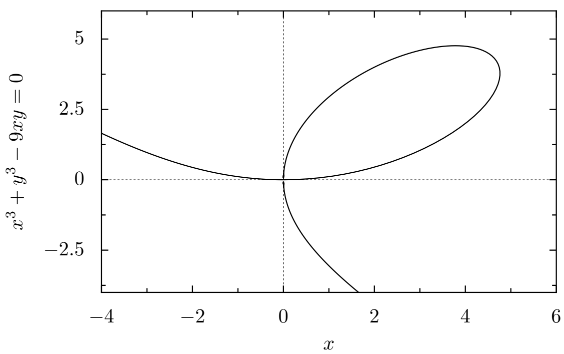


Figure 4.5: The Folium of Descartes.

Solution The first thing we will need to do is find the slope of the curve at the point (2,4).

$$\begin{aligned}
x^3 + y^3 - 9xy &= 0 \\
3x^2 + 3y^2\frac{dy}{dx} - 9y - 9x\frac{dy}{dx} &= 0 \\
(3y^2 - 9x)\frac{dy}{dx} &= 9y - 3x^2 \\
\frac{dy}{dx} = \frac{9y - 3x^2}{3y^2 - 9x} &= \frac{3y - x^2}{y^2 - 3x}.
\end{aligned}$$

Evaluating the derivative at the point (2,4) we find

$$\left.\frac{dy}{dx}\right|_{(2,4)} = \left.\frac{3y - x^2}{y^2 - 3x}\right|_{(2,4)} = \frac{12 - 4}{16 - 6} = \frac{4}{5}.$$

Now that we have the slope of the tangent line at the point of interest, we use the point-slope form to find

$$y_t = \frac{4}{5}(x - 2) + 4 = \frac{4}{5}x + \frac{12}{5}.$$

Now as far as the line normal to the curve at this point is concerned, we need to find the line perpendicular to the tangent line. This line will cross through the same point, but the slope will be the negative reciprocal of the slope of the tangent line. It follows that

$$y_n = -\frac{5}{4}(x - 2) + 4 = -\frac{5}{4}x + \frac{13}{2}.$$

Example 3. Find d^2y/dx^2 for $2x^3 - 3y^2 = 8$.

Solution We begin by finding the first derivative

$$\begin{aligned} \frac{d}{dx}(2x^3 - 3y^2) &= \frac{d}{dx}8 \\ 6x^2 - 6yy' &= 0 \\ x^2 - yy' &= 0 \\ y' &= \frac{x^2}{y}. \end{aligned}$$

Now the second derivative

$$y'' = \frac{d}{dx}\frac{x^2}{y} = \frac{2xy - x^2y'}{y^2} = \frac{2x}{y} - \frac{x^2y'}{y^2} = \frac{2x}{y} - \frac{x^4}{y^3}.$$

Finally, we can use implicit differentiation to find the derivative of inverse functions.

Example 4. Find the derivative of $y = \ln(x)$ using implicit differentiation.

Solution Presuming that we don't know the derivative of $\ln(x)$, we would rewrite this equation as $e^y = x$ using the inverse function. Now we can use implicit differentiation (because we know how to differentiate both sides of the equation) to find

$$e^y \frac{dy}{dx} = 1,$$

so

$$\frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{e^{\ln(x)}} = \frac{1}{x},$$

which is the familiar result.

Example 5. Find the derivative of $y = \sin^{-1}(x)$ using implicit differentiation.

Solution Writing $\sin(y) = x$ and differentiating we find

$$\cos(y)\frac{dy}{dx} = 1,$$

so that

$$\frac{dy}{dx} = \frac{1}{\cos(y)} = \frac{1}{\cos(\sin^{-1}(x))}.$$

To simplify, we recall the useful identity $\sin^2(x)+\cos^2(x) = 1$ which leads to $\cos(x) = \sqrt{1 - (\sin(x))^2}$. When we substitute into the above equation, we will apply sine to its inverse (yielding x), and find

$$\frac{d}{dx}\sin^{-1}(x) = \frac{1}{\cos(\sin^{-1}(x))} = \frac{1}{\sqrt{1 - (\sin(\sin^{-1} x))^2}} = \frac{1}{\sqrt{1 - x^2}}.$$

4.8 Related Rates

In many physical situations we have a relationship between multiple quantities, and we know the rate at which one of the quantities is changing. Oftentimes we can use this relationship as a convenient means of measuring the unknown rate of change of one of the other quantities, which may be very difficult to measure directly. Such a situation is called a related rates problem. The key to solving related rates problems is using the known relationship between the quantities to find a relationship between their rates of change, by differentiating the relationship between the quantities themselves.

Example 1. Suppose that a spherical balloon is being inflated, with the radius of the balloon increasing at .1 m/s. Find the rate of change of the volume of the balloon with respect to time.

Solution The first step to solving this problem is assigning variables to the quantities of interest - the radius and volume of the balloon. Let us refer to the radius as r and the volume as V. Now we need to find a relationship between these quantities. For a sphere, we know that

$$V = \frac{4}{3}\pi r^3.$$

Since the balloon is being pumped up, clearly both V and r are changing with time. We are given $dr/dt = 0.1$ m/s, and our goal is to find the time rate of change of the volume, dV/dt. We do so by differentiating the relationship between V and r with respect to time, applying the differentiation operator d/dt to both sides of our relationship between V and r. Doing so we find

$$\begin{aligned}
\frac{d}{dt}V &= \frac{d}{dt}\left(\frac{4}{3}\pi r^3\right) \\
\frac{dV}{dt} &= \frac{4}{3}\pi \cdot \frac{d}{dt}r^3 \\
\frac{dV}{dt} &= \frac{4}{3}\pi \cdot 3r^2\frac{dr}{dt} \\
\frac{dV}{dt} &= 4\pi r^2\frac{dr}{dt}
\end{aligned}$$

The above expression gives us a relationship between the time rate of change of volume and radius of the balloon. In this particular situation we are given a specific value for dr/dt, so we find that

$$\frac{dV}{dt} = 4\pi r^2 \cdot 0.1 \text{ m} = 0.4\pi r^2 \text{ m/s}.$$

Note that in the above expression we would need to substitute with the radius of the balloon in meters in order to have the units of dV/dt be given in m^3/s.

Example 2. Consider a right-circular cylindrical tank of water of constant radius r with a tap at its bottom from which water is flowing at a constant rate of 3000 L/min. At what rate is the height of the water in the tank decreasing?

Solution The first step is to identify the variables of interest. If we let h denote the height of the water in the tank, then what we are interested in finding is dh/dt. From the problem statement we are given the rate of change of the volume of the water, in L/min, which we will denote as dV/dt. For a right-cylinder, volume is given by the product of the base area and height of the cylinder. Since we have a right-circular cylinder, volume, radius, and height are related by the expression

$$V = \pi r^2 h.$$

We differentiate both sides of the equation with respect to t to find

$$\frac{dV}{dt} = \pi r^2 \frac{dh}{dt}.$$

Here it is important to note that only V and h are changing with time, and that r is simply a constant, so we do not need to worry about using the product rule. Solving for dh/dt we find

$$\frac{dh}{dt} = \frac{1}{\pi r^2}\frac{dV}{dt}.$$

In this situation we need to be a little bit careful with units. We must note that $1L = (10\text{cm})^3$, so we have that

$$\frac{dV}{dt} = -3000\frac{(10\text{cm})^3}{\text{min}} = -3 \cdot 10^6 \frac{\text{cm}^3}{\text{min}}.$$

Now in order to have the units match on both sides of the expression we need give r in cm, so h will be expressed in cm, and dh/dt will be expressed in cm/min. Finally, we arrive at the expression

$$\frac{dh}{dt} = -\frac{3 \cdot 10^6}{\pi r^2}\frac{\text{cm}^3}{\text{min}}.$$

Note that smaller the radius is, the faster the height drops. This is because with a small radius 3000 L of water leaving per minute would correspond to a volume of water with a large height, whereas when r is larger the volume would have a relatively smaller height.

The above example illustrates the importance of keeping track of units when solving related rates problems. In general we can associate the units as a part of the variables in the relationship between the quantities of interest, and differentiate to relate the rates of change without paying regards to units. However, once we want to substitute numerical values for different quantities in the relationship we need to make sure our units match.

In more complicated situations it can be very helpful to draw a picture of the situation at hand. Keeping this in mind the generally strategy for solving a related rates problem is to identify the quantities of interest and assign appropriate variables names. Having done so one draws a picture of the physical situation, and identifies the relationship between the previously defined variables, and whatever additional constants are a part of the situation. These relationships generally stem from a geometric or physical consideration. Finally one differentiates both sides of this relationship to relate the rates of change of the variables of interest, and solves for the unknown rate of change in terms of known quantities.

Example 3. Consider a rising hot-air balloon with a completely vertical ascent. The height of the balloon is being tracked by a rangefinder 500ft from the liftoff point. At the moment the rangefinder's elevation angle is $\pi/4$, the angle is increasing at the rate of 0.14 rad/min. How fast is the balloon rising at that instant?

Solution First we identify our variables of interest - the angle the range finder makes with the ground θ, and the height of the balloon y. In order to relate these quantities we must consider the physical situation. The rangefinder is on the ground a distance 500ft from the liftoff point from the balloon, and since the the ascent of the balloon is completely vertical, the balloon is some height y above the liftoff point. These two distances form two sides of a right triangle, with a hypotenuse completed by the laser of the rangefinder. This is a right triangle with a base of 500 ft, height of y ft, and with angle θ adjacent to the base. Given this right triangle we can relate the angle on the rangefinder and the height of the balloon. We find that

$$\tan(\theta) = \frac{y}{500 \text{ ft}}.$$

Differentiating both sides of the equation with respect to t, we find

$$\sec^2(\theta)\frac{d\theta}{dt} = \frac{1}{500 \text{ ft}}\frac{dy}{dt}.$$

Thus,

$$\frac{dy}{dt} = 500\sec^2(\theta)\frac{d\theta}{dt} \text{ ft}.$$

We know that $d\theta/dt|_{\theta=\pi/4} = 0.14$ rad/min, so evaluating this expression at this angle of interest we find

$$\left.\frac{dy}{dt}\right|_{\theta=\pi/4} = 500\sec^2(\pi/4)\cdot 0.14\frac{ft}{min} = 70(\sqrt{2})^2\frac{\text{ft}}{\text{min}} = 140\frac{\text{ft}}{\text{min}}.$$

Example 4. Suppose a police cruiser is chasing a speeding car. The police cruiser is heading south approaching an intersection, and the speeding car has already passed the intersection and is heading east. When the police cruiser is 0.6 mi north of the intersection, the car is 0.8 mi to the east, and the police radar determines that the distance between them and the car is increasing at 20mph. If the cruiser is moving at 60mph when this measurement is made, what is the speed of the car?

Solution Let y denote the position of the police cruiser, and x the position of car. It is convenient to define our origin at the intersection, so that the position of the cruiser will fall entirely on the vertical axis, and the position of the car on the horizontal axis. Let s be the distance between the cruiser and the speeding car. In this case we have three variables of interest, which happen to be related as the three sides of a right triangle. Our triangle has base x, height y, and hypotenuse s, so using the Pythagorean theorem, it follows

$$s^2 = x^2 + y^2.$$

Differentiating with respect to time we find

$$2s\frac{ds}{dt} = 2x\frac{dx}{dt} + 2y\frac{dy}{dt}.$$

In the above expression both s and dx/dt are unknowns. Fortunately, using the Pythagorean theorem again we can rewrite s in terms of x and y, finding

$$s = \sqrt{x^2 + y^2}.$$

Solving for dx/dt, the time rate of change of the speeding car we find

$$\frac{dx}{dt} = \frac{1}{x}\Big(s\frac{ds}{dt} - y\frac{dy}{dt}\Big) = \frac{1}{x}\Big(\sqrt{x^2+y^2}\frac{ds}{dt} - y\frac{dy}{dt}\Big).$$

This expression relates the rate of change of the police car, the distance between the two cars, and the speeding car. To find the speed of the chased car we must use the numerical data given in the problem statement. Here we must be a careful in order to make sure that our interpretation of these data are consistent with the coordinate system we have setup. Namely, since the police cruiser is moving downwards, we find that $dy/dt = -60$ mph, a negative rate of change. Since the distance between the two cars is increasing we find $ds/dt = 20$ mph. Now we can substitute the known quantities into the above expression finding

$$\frac{dx}{dt} = \frac{1}{0.8}\Big(\sqrt{0.8^2+0.6^2}\cdot 20 - 0.6\cdot(-60)\Big) \text{ mph} = 70 \text{ mph}.$$

Example 5. Consider an oil well located near the middle of a calm, shallow lake leaking oil onto the surface of the lake at the rate of 16 cubic feet per minute. Since oil and water don't mix, as the oil leaks onto the surface of the lake it spreads out into a uniform layer, called an oil slick, that floats on top of the water. The slick is circular, centered at the leak, and 0.02 ft thick. Find the rate at which the radius of the slick is increasing when the radius is both 500 ft and 700 ft. Why would you expect the rates of change of the radii to be different, even though the oil is being added at a constant rate?

Solution The variables of interest are the volume V and the radius r of the oil slick. Since the surface of the oil slick is circular, and it has a height of 0.02 ft, we are dealing with a very thin, right-circular cylinder. The relationship between the volume and radius of the cylinder is given by

$$V = \pi r^2 h = 0.02\pi r^2 \text{ ft}.$$

Differentiating both sides of the equation with respect to t we find

$$\frac{dV}{dt} = 0.04\pi r \frac{dr}{dt} \text{ ft}.$$

We know that $dV/dt = 16 \text{ ft}^3/\text{min}$, and want to find dr/dt. Solving for dr/dt we find that

$$\frac{dr}{dt} = \frac{1}{0.04\pi r}\frac{dV}{dt}\frac{1}{\text{ft}} = \frac{400}{\pi r} \text{ ft}^2.$$

Substituting the appropriate values for r we find

$$\left.\frac{dr}{dt}\right|_{r=500} = \frac{400}{500\pi} \approx 0.255 \text{ ft} \quad \text{and} \quad \left.\frac{dr}{dt}\right|_{r=700} = \frac{400}{700\pi} \approx 0.182 \text{ ft}.$$

One way of interpreting the increasing size of the oil slick is that the existing portion of the slick does not change, but concentric hollow discs are being added around the initial circular surface. From this perspective, the larger r is, the larger the volume of such a disc is, whereby the radius r needs to increase by less in order for a constant volume of oil to be added to the slick.

Example 6. Suppose we have two right circular cones, cone A and cone B. The height of cone A and the diameter of cone B both change at a rate of 4 cm/s, while the diameter of cone A and the height of cone B are both constant. At a particular instant, both cones have the same shape: $h = d = 10$ cm where h is height and d is diameter. Find the rates of change of the volume of the two cones at this time. Why would you expect the volume of the cones to be changing at different rates?

Solution Before proceeding it is helpful to visualize the situation. Since the height of cone A is increasing, cone A is becoming a tall and skinny cone. Since cone B becomes wider and wider, it starts to looks short and fat. Of interest here are the height h, diameter D (using capital D as d would create a notation issue), and volume V for the two cones. For a right-circular cone, the volume is given by

$$V = \frac{1}{3}\pi r^2 h.$$

Since our given the diameter rather than radius, we rewrite this equation

$$V = \frac{1}{3}\pi (D/2)^2 h = \frac{\pi}{12} D^2 h.$$

Before proceeding we need to recognize that we in fact have two related rates problems here - one for cone A and another for cone B. We will use the subscripts a and b to distinguish between the two cones. For cone A, D will be a constant and h will vary with time, while for cone B, h will be a constant and D will vary with time. Thus, we must differentiate this expression separately for each cone, because different quantities will depend on time. At the particular instant of interest, we know $h_a = h_b = D_a = D_b = 10$ cm. We also know that $dh_a/dt = 4$ cm/s and $dD_b/dt = 4$ cm/s. For cone A, differentiating with respect to t, we find the relationship

$$\frac{dV_a}{dt} = \frac{\pi}{12} D_a^2 \frac{dh_a}{dt} = \frac{\pi}{12} \cdot 10^2 \cdot 4 \frac{\mathrm{cm}^3}{\mathrm{s}} = \frac{100\pi}{3} \frac{\mathrm{cm}^3}{\mathrm{s}} \approx 104.72 \frac{\mathrm{cm}^3}{\mathrm{s}}.$$

Now for cone B, differentiating with respect to t, we find the relationship

$$\frac{dV_b}{dt} = \frac{\pi}{12} h_b \cdot 2D_b \cdot \frac{dD_b}{dt} = \frac{\pi}{6} 10 \cdot 10 \cdot 4 \frac{\mathrm{cm}^3}{\mathrm{s}} = \frac{200\pi}{3} \frac{\mathrm{cm}^3}{\mathrm{s}} \approx 209.44 \frac{\mathrm{cm}^3}{\mathrm{s}}.$$

We expect that the two cones will have different rates of change, because in one case, the volume is related to the first power of a changing quantity, and in the second it is related to the square of the changing quantity.

4.9 Extreme Values

The extreme values, or extrema, of a function are places where a function reaches its minimum and maximum values. Extreme values come into use in both design scenarios, as well as visualizing functions. Suppose we are designing the packaging for a product. In such a situation it would be beneficial to minimize the waste used in production, in order to conserve resources. If we have a function that describes the amount of waste used in production, it would be our goal to find its minimum, and use that to guide our production process. These types of design problems are called optimization problems, where we either want to maximize a benefit, or minimize a loss. Solving an optimization problem comes down to finding the extreme values of a function modeling the situation, and choosing a suitable operating point.

We will also see extreme values play a vital role in sketching functions, where we can use a very limited amount of information about a function in order to gain a lot of information about its behavior. In order to graph, and reason visually about functions, we will need to: identify the function's extrema, find intervals over which the function is increasing and decreasing, and determine the concavity of the function. We can accomplish the first of these two tasks using the first derivative, and the last using the second derivative. We will begin with finding the extrema, or extreme values of a given function.

We can identify two types of extrema - local and global. Global extrema are the largest and smallest values that a function takes on over its entire domain, and local extrema are extrema which occur in a specific neighborhood of the function. In both the local and global cases, it is important to be cognizant of the domain over which the function is defined. That which is an extremum on one domain may very well not be over a new domain, and vice versa. Before delving further, let us give the formal definitions of these various extrema.

Definition 4.9.1 (Global Extrema)**.** The function $f : D \to \mathbb{R}$ has a *global maximum* at $c \in D$ if

$$f(x) \leq f(c) \quad \text{for all } x \in D.$$

We say that f has a *global minimum* at $c \in D$ if

$$f(x) \geq f(c) \quad \text{for all } x \in D.$$

Definition 4.9.2 (Local Extrema)**.** The function $f : D \to \mathbb{R}$ has a *local maximum* at $c \in D$ if

$$f(x) \leq f(c) \quad \text{for all } x \text{ in some open interval centered at } c.$$

We say that f has a *local minimum* at $c \in D$ if

$$f(x) \geq f(c) \quad \text{for all } x \text{ in some open interval centered at } c.$$

If c is an interior point of D we require that the interval extend on both sides of c; if c is at an endpoint of the domain, we only require that the above inequalities hold on the portion of the interval that is in the domain D.

Note that in the above definitions for local extrema, the interval we find may be quite small. Nevertheless, if we can find a single interval over which the above inequalities hold, then they will also hold for any smaller interval. Since any interval of finite length contains an infinite number of points, there is no way we can test every point to look for extrema. Instead, we will need to be more clever with the places we look for extrema. We need to emphasize that we are only interested in identifying the extrema of continuous functions. If we are faced with a discontinuous function, then it may very well jump to an exceedingly high or low value at any strange point, which would make tracking down extrema a very difficult endeavor.

As stated earlier, the first derivative is the tool we will need to use to find global and local extrema. Recall that the first derivative tells us whether a function is increasing, decreasing or neither, and if it is changing, how quickly. Here we are not so interested in how quickly the function is changing, but whether or not it is changing, and in which direction; we gain this information from the sign of the first derivative: if the derivative is positive, the function is increasing, if it is negative, the function is decreasing, and if it is zero the function is not changing. If the derivative has the same sign over an entire interval, then we can say that the function is either increasing, decreasing, or not changing over that interval. Formally, for a function to be increasing on an interval means that ($x_1 < x_2 \implies f(x_1) < f(x_2)$), and for the function to be decreasing ($x_1 < x_2 \implies f(x_1) > f(x_2)$).

A careful observer may already be able to deduce how we can use the first derivative to find extrema. If we want to identify a local maximum at a point c, we want the function to increase to the left of c, and to decrease to the right of c. This will put c at a peak in the function. If the function is increasing to the left of c, the derivative must be positive to the left of c, and if the function is decreasing to the right of c, the derivative must be negative to the right of c. When we say that the function is increasing to the left of c, we mean there is some interval (b, c) so that the derivative is positive on (b, c), and when we say the function is decreasing to the right of c, we mean there is some interval (c, d) so that the derivative is negative on (c, d).

It is only possible for a derivative to go from positive to negative in one of two ways: either the derivative must cross 0 (which it would if the derivative is a continuous function), or the derivative must be discontinuous, so that it can jump from positive to negative (for instance in the corner of $|x|/x$ at $x = 0$). The point at which the derivative should cross zero (or fail to exist) is c, the local maximum. Similar analysis would indicate that the derivative should be zero (or not exist) at a local minimum, but with the derivative negative to the left, and positive to the right. In conclusion, we can only have a local extremum at a point c if the sign of the derivative changes around c (either from positive to negative - a maximum, or from negative to positive - a minimum). And the only way the derivative can change signs around c is if $f'(c) = 0$ or $f'(c)$ does not exist. Such points are called *critical points*.

Definition 4.9.3 (Critical Point)**.** An interior point of the domain of a function f where f' is either zero or undefined is a *critical point* of f.

Apropos to the previous discussion, the only place in the interior of a function's domain that it can have a local extremum is at a critical point (but a critical point doesn't guarantee there is a local extremum). The only points left to consider are the endpoints of the domain. Thus, if we want to find extrema, we need to look at critical points, and the endpoints of the domain. Now we have transformed the problem of evaluating the function at an infinite number of points, to evaluating it at a few, select points - much easier. Once we have found the local extrema, it is an easy task to find the global extrema. If a function has a global maximum on a given domain, it will correspond to the local maximum with the largest value. Similarly, a global minimum corresponds to the local minimum with the smallest value.

Just as we would expect from the previous discussion, if a function f has an extremum at an interior point c of its domain, and $f'(c)$ is defined, then $f'(c) = 0$. However, the converse of this statement is untrue - the fact that $f'(c) = 0$ does not guarantee that c is a local extremum. For instance, we could have a function where the derivative is positive, becomes zero at a point c, and then becomes positive again (x^3 is an example of this, at $x = 0$). Since the function is increasing to the right of c, it is clear that c cannot be a local maximum, and since the function is increasing to the left of c, c also cannot be a local minimum. For any function that isn't a constant, we only have a local extremum at c when the derivative changes signs around c; if the derivative doesn't change signs, then we do not have a local extremum. In conclusion, a zero derivative doesn't guarantee a local extremum, but it is a very good place to try and identify one. We may also find extrema at points where the derivative is undefined (because the derivative could jump and change signs at such a discontinuity).

It turns out that a continuous function defined on a closed interval always achieves its global maximum and minimum.

Theorem 4.9.1 (Extreme Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous. It follows that f attains both a global minimum value m and a global maximum value M in the interval $[a, b]$. That is, there exist $x_1, x_2 \in [a, b]$ so that $f(x_1) = m$ and $f(x_2) = M$, and $m \leq f(x) \leq M$ for all $x \in [a, b]$.*

In contrast, it is possible for a function to not have any extreme values on a given domain. Consider the function

$$f : [0, 1) \to \mathbb{R},\ f(x) = x.$$

The smallest value is $f(0) = 0$, which corresponds to the global minimum. However, there is no global maximum, as no matter how close a point is to $x = 1$, one can choose a point even closer that yields a larger function value. If the function was defined on $0 < x \leq 1$, then it would instead have a global maximum but no global minimum. If we defined this function on $(0, 1)$, then it would have no extreme values. In contrast, if we define the function on the closed interval $[0, 1]$ then it has both a global maximum and minimum, as required by the extreme value theorem.

Example 1. Find all minima and maxima of $f(x) = xe^{-x}$ on the domain $[0, 4]$.

Solution To find local extrema, we need to look at the behavior of the first derivative around critical points, as well as at endpoints. We find

$$f'(x) = 1 \cdot e^{-x} + x \cdot (-e^{-x}) = (1 - x)e^{-x}.$$

Solving for $f'(x) = 0$ we find $0 = (1-x)e^{-x}$. The right-hand side is 0 only if one of the two terms is zero, and since e^{-x} is always nonnegative, we find that $x = 1$ is the only critical point. According to our previous discussion, we know that the derivative of a function can only change signs around a critical point, so if we partition the interval $[0, 4]$ into subintervals, we know that the sign of the derivative must be constant on $[0, 1)$, as well as on $(1, 4]$. If the signs of the derivatives differ on these two subintervals we will know that we have found an extremum.

For $x < 1$ we have $1 - x > 0$, and for $x > 1$ we have $1 - x < 0$, so this critical point is a local maximum. On the endpoints we find $f'(0) = 1 > 0$, so the left endpoint is a local minimum (as the function is increasing to the right of it). Similarly, we find $f'(4) = -3e^{-4} \approx -0.055 < 0$, so the right endpoint is also local minimum (as the function is decreasing to the left of it).

To find the global minima and maxima, we need to compare the function values at all critical points, as well as the endpoints of the domain. We find that $f(0) = 0$, $f(1) = e^{-1} \approx 0.367$, and $f(4) = 4e^{-4} \approx 0.073$. Thus, $f(0)$ is both a local and global minimum, and $f(1)$ is both a local and global maximum. It is important to once again note that the global and local maxima are domain specific - that is, the points at which they occur may be different for different domains of the function (which they certainly are in this case). The graph of this function and its derivative are given in figure 4.6.

Example 2. Find the global minima and maxima of $f(x) = x^3 - x^2$ on the interval $[-1, 1]$.

Solution The first step is to identify the critical points. To do so, we compute

$$f'(x) = 3x^2 - 2x.$$

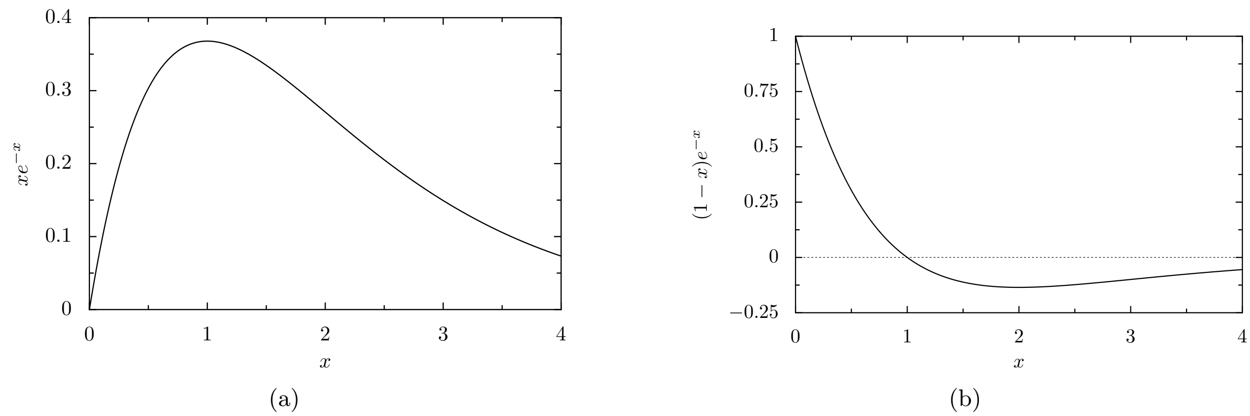


Figure 4.6: xe^{-x} and its derivative.

Setting the derivative equal to 0 we find $0 = 3x^2 - 2x = x(3x-2)$, so we have critical points at $x = 0$ and $x = \frac{2}{3}$. Comparing the values of the critical points and endpoints, we find that: $f(-1) = -2$, $f(0) = 0$, $f(\frac{2}{3}) \approx -0.15$, and $f(1) = 0$. Thus, the global minimum is $f(-1)$, while both $f(0)$ and $f(1)$ are global maxima.

Example 3. Find the global minima and maxima of $f(x) = x^3 - x^2$ on the interval $[-2, 2]$.

Solution We have already found the critical points, so we simply need to compare their values to those of the new endpoints. $f(-2) = -12$ and $f(2) = 4$, so clearly $f(-2)$ is the global minimum, and $f(2)$ is the global maximum. Note that on the new domain there is a single global maximum, rather than two (see figure 4.7).

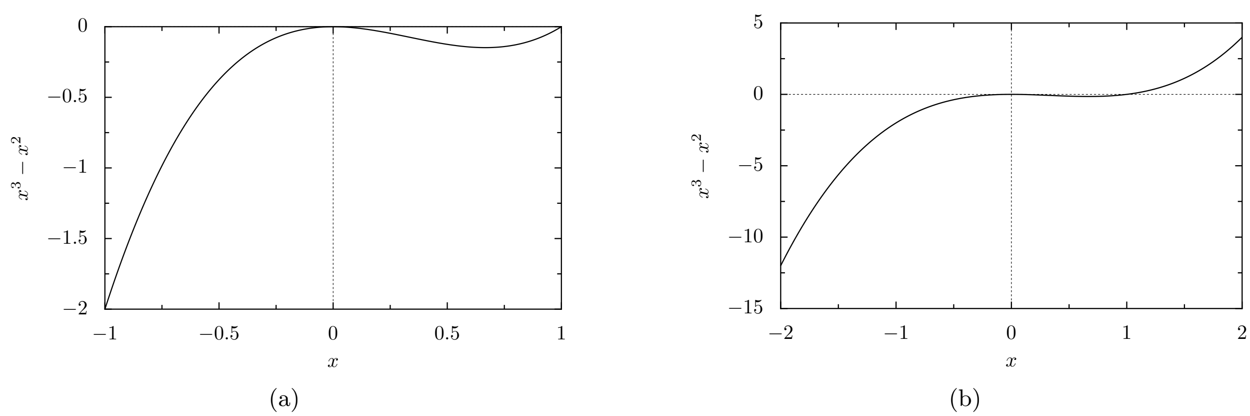


Figure 4.7: The extreme values of $x^3 - x^2$ depend on the domain.

Example 4. Find the global maxima and minima of $f(x) = |x|$ over the domain $-2 \le x \le 3$.

Solution In order to differentiate this function, it is useful to define it piecewise as follows

$$f(x) = \begin{cases} x & x \ge 0 \\ -x & x < 0 \end{cases}$$

We can see that for $x > 0$ we have $f'(x) = 1$, and for $x < 0$ we have $f'(x) = -1$, so we don't obtain any critical points there. However, the derivative is not defined at $x = 0$ (because it is a corner), so we have a critical point at $x = 0$ (and it turns out to be a local minimum, because the derivative changes signs from negative to positive as we cross it). The values of the endpoints and this critical point are $f(-2) = 2$, $f(0) = 0$, and $f(3) = 3$. Thus, $x = 0$ is the global minimum, and $x = 3$ is the global maximum.

4.10 Mean Value Theorem

Before discussing the mean value theorem, it is worthwhile to note an interesting result about the nature of derivatives. When we differentiate a function $f : I \to \mathbb{R}$, there is no guarantee that the result is a continuous function. For instance, consider

$$f(x) = \begin{cases} x^2 \sin(1/x) & x \neq 0 \\ 0 & x = 0. \end{cases}$$

For $x \neq 0$, we find that

$$f'(x) = 2x \sin(1/x) + x^2 \cos(1/x) \cdot (-x^{-2}) = 2x \sin(1/x) - \cos(1/x).$$

In order to find $f'(0)$ we will need to appeal directly to the definition of the derivative, and find

$$f'(0) = \lim_{h \to 0} \frac{f(h) - f(0)}{h} = \lim_{h \to 0} h \sin(1/h) = 0.$$

Thus, we have a function that is differentiable everywhere, but its derivative is not continuous, as it has no limit as $x \to 0$ (the cosine term oscillates too rapidly). This example may seem a bit artificial, but it illustrates the fact that the derivative of a function may not behave nearly as nicely as the function itself. Even though we are not guaranteed continuity of a derivative, we are guaranteed one useful property of derivatives.

Theorem 4.10.1 (Intermediate Value Property of Derivatives)**.** *Let $f : [a, b] \to \mathbb{R}$ be differentiable on (a, b). For any value y_0 between $f'(a)$ and $f'(b)$, there exists some $c \in [a, b]$ such that $f'(c) = y_0$.*

This theorem simply states that any derivative has the intermediate value property. Thus, even though a derivative may not be continuous, it cannot also just be *any* function. For instance, the above theorem implies that $|x|/x$ cannot be the derivative of a function that is differentiable over an interval containing 0, because over any interval containing $x = 0$ it does not have the intermediate value property (it skips the values between -1 and 1). We know in fact that the derivative of $f(x) = |x|$ is $f'(x) = |x|/x$, except for at $x = 0$, where the derivative $f'(x)$ does not exist. It is the way that $|x|/x$ jumps at $x = 0$ that makes it impossible for it to be the derivative of any function that is differentiable over an entire interval, which $f(x) = |x|$ is not. We will keep this intuition in mind while considering the mean value theorem.

The mean value theorem provides us with a means of relating a function's instantaneous rate of change inside an interval with its average (or mean) rate of change over an interval. The power of the mean value theorem does not truly shine in introductory calculus, as its main applications are proving other results. Nevertheless, we will discuss this very useful theorem here, and some of its implications.

Theorem 4.10.2 (The Mean Value Theorem)**.** *Let f be continuous on $[a, b]$, and differentiable on the interval's interior (a, b). It follows that there exists at least one point $c \in (a, b)$ at which*

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

Verbally, the mean value theorem states: for a continuous and differentiable function f on an interval, we can find a point inside the interval at which the instantaneous rate of change of f is equal to the average rate of change of f over that interval. In other words, if we draw the secant line between the points $(a, f(a))$ and $(b, f(b))$, we can find a point $c \in (a, b)$ so that the slope of the tangent line at c, $f'(c)$, is the same as the slope of the secant line through the endpoints of the interval. If we think of the average rate of change over an interval as the *mean* rate of change, then the name of this theorem makes sense. A graphical representation of the mean value theorem is given in figure 4.8.

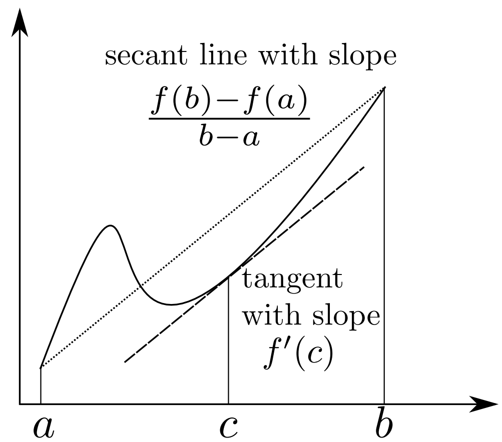


Figure 4.8: Geometrically, the mean value theorem states that there is some point $c \in [a, b]$ where the tangent line is parallel to the secant line through $(a, f(a))$ and $(b, f(b))$.

In order for a function to have a given average rate of change, it must either be constantly changing at that rate, or at some instants in time change faster, and at others change slower. In this way the mean value theorem follows from the intermediate value property of derivatives, as if the derivative is larger than the average rate of change at some time, and smaller at another time, there must be a time in between when it is exactly equal.

It is important to emphasize that the mean value theorem does not apply for functions which are not continuous on the closed interval $[a, b]$, or not differentiable on the open interval (a, b). Consider the function

$$f(x) = \begin{cases} 0 & x = 0 \\ 0.5 & 0 < x < 1 \\ 1 & x = 1 \end{cases}$$

If we look at the average rate of change on $[0, 1]$ we find

$$\frac{f(1) - f(0)}{1 - 0} = 1,$$

and the secant line through the points $(0, f(0))$ and $(1, f(1))$ is just the line of slope one that passes through the origin. What about the derivative of this function? It turns out $f'(x)$ does not exist at $x = 0$ and $x = 1$, because the function is not continuous at those points. However, for $0 < x < 1$ the function is continuous, and just a constant, so the derivative $f'(x) = 0$. It can be seen that there are no points in this interval for which the slope of the tangent line is 1, because for all points in $(0, 1)$ the tangent line is horizontal, having slope 0. We cannot apply the mean value theorem to this function because it is not continuous on $[0, 1]$, even though it is differentiable on $(0, 1)$. It is noteworthy that we could still apply the mean value theorem on any closed interval contained in $(0, 1)$, and it would tell us that somewhere in the interval the slope of the tangent line is 0 (which is not particularly useful, as we already know that the tangent line is 0 everywhere in the interval).

Now let us consider a function that is continuous on a closed interval, but not differentiable within the open interval. Let us consider $f(x) = |x|$ on the interval $[-1, 1]$. If we look at the average rate of change over the interval we find

$$\frac{f(1) - f(-1)}{1 - (-1)} = \frac{0}{2} = 0,$$

yet there are no points $c \in (-1, 1)$ such that $f'(c) = 0$. For $x < 0$ we have $f'(x) = -1$ and for $x > 0$ we have $f'(x) = 1$ (of course $f'(0)$ does not exist). If the function $|x|$ were differentiable at $x = 0$, then in a sense, we would have to have $f'(0) = 0$, as this would be the point where the derivative would change from negative to positive. However, in this case we don't have any points with a zero derivative, and because the derivative is discontinuous at $x = 0$, it is possible for us to jump from a positive to negative derivative.

If we apply the mean value theorem on an interval for which $f(a) = f(b)$, it follows that there exists a point $c \in (a, b)$ so that

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

This means that if we have any two points a and b at which our function of interest crosses the x-axis, we can find a point in the interval which is a critical point (at which we could look for a local extremum). This special case of the mean value theorem is called Rolle's theorem, and is used in the proof of the mean value theorem, If we think about the function

$$f(x) = \frac{x^3}{3} - 3x$$

which is 0 at $-3, 0, 3$, then we know it has a critical point somewhere in the intervals $(-3, 0)$ and $(0, 3)$. Thus, we know it has at least two critical points, but it could have more.

The mean value theorem is also useful in the classification of critical points. Consider

$$f : [-3, 3] \to \mathbb{R},$$

where $f(x)$ is defined as above. Differentiating f we find that

$$f'(x) = x^2 - 3,$$

and that $f'(x) = 0$ when $x = \pm\sqrt{3}$, meaning that $\pm\sqrt{3}$ are critical points. Now we evaluate the function at each of the critical points and endpoints, yielding $f(-3) = 0$, $f(-\sqrt{3}) = 2\sqrt{3}$, $f(\sqrt{3}) = -2\sqrt{3}$, and $f(3) = 0$. We can of these points as partitioning our larger interval into subintervals, over which we want to consider the sign of the derivative. Since the sign of the derivative cannot change over these subintervals, knowing the sign of the derivative at a single point in the interval tells us the sign of the derivative over the entire subinterval.

If we apply the mean value theorem to the endpoints of one of these subintervals, say $[\sqrt{3}, 3]$, we find that there exists some $c \in (\sqrt{3}, 3)$ so that

$$f'(c) = \frac{f(3) - f(\sqrt{3})}{3 - \sqrt{3}} = \frac{0 - (-2\sqrt{3})}{3 - \sqrt{3}} > 0.$$

Since the derivative is positive at some point in the subinterval, and the sign of the derivative cannot change over the subinterval, it follows that the derivative is positive over the entire subinterval. Note that the denominator will always be positive when we apply the mean value theorem, and since we only care about the sign of the derivative, we only need to determine the relative magnitude of the endpoints to see if the derivative will be positive or negative. Doing so we can see that $f(\sqrt{3}) = -2\sqrt{3} < 2\sqrt{3} = f(-\sqrt{3})$ which implies the derivative is negative on the subinterval $(-\sqrt{3}, \sqrt{3})$, and finally $f(-\sqrt{3}) > f(3)$ implies that the derivative is positive over the subinterval $(-3, -\sqrt{3})$. It follows that $f(-\sqrt{3})$ is a local maximum and $f(\sqrt{3})$ is a local minimum.

The mean value theorem allows us to transform a bounded derivative into a bound for a function defined on an interval. Let's suppose that we have a function f which is continuous on the interval $[a, b]$, differentiable on (a, b), and that $|f'(x)| < B$ for all $x \in (a, b)$ (this means that the derivative of f is bounded). This bound on the derivative tells us that the function can only change so fast. In fact, the maximum rate at which it can change is B. Intuitively, if the function can only change so fast, then it should not be able to become too much larger or smaller than $f(a)$ (or $f(b)$ for that matter) if we don't move too far from $x = a$. Let us consider a point $x \in (a, b)$. Applying the mean value theorem on the interval $[a, x]$ we find

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

for some $c \in (a, x)$. Rearranging terms it follows

$$f(x) = f'(c)(x - a) + f(a).$$

Although this looks like the equation for a line, remember that x is a fixed value, and for any value of x, we might actually find a different value for c. Normally we would find $f'(c)$ using information about $f(x)$ and $f(a)$, but here we want to gain some information about $f(x)$. This means that we don't know exactly what $f'(c)$ is, but we do know that $f'(c) < B$ and $f'(c) > -B$. It then follows that

$$-B(x - a) + f(a) \leq f(x) \leq B(x - a) + f(a),$$

which is an equation that holds for all $x \in (a, b)$. In fact, if we consider the above equation in the cases of equality, we get the equations of two lines. When $x = a$, the value of both inequalities is $f(a)$. From there, the lines increase (or decrease) with slope B (or $-B$), which describe bounds on the function $f(x)$. At any given point x, the function $f(x)$ must be between the two lines, which implies the function is bounded. The only way the function would be equal to one of the two lines, is if the derivative was a constant of B or $-B$ up to that point, meaning that the magnitude of the derivative was equal to the bound, which is the largest possible magnitude of the derivative (or may not even be achievable, if the bound is larger than any value the derivative ever achieves). Because in this region the rate of change of the function is bounded, it is only possible for the function to change by so much, which forces it to be within these bounds. This is an important concept for visualizing functions, and is useful in many proofs.

4.11 Concavity and Curve Sketching

The standard technique for sketching a function is plotting points and connecting the dots. However, there are issues to consider such as which points to plot and how to connect them. Consider a sine function. If all of the points we choose happen to be at the peaks of the function, say $\pi/2$, $5\pi/2$, etc., then upon connecting our dots we will just get a straight line, which is not like a sine function at all. Similarly, if all of the points we plot are at alternating peaks, $\pi/2$, $3\pi/2$, $5\pi/2$, etc., when we connect the dots we will get a jagged sawtooth curve - also inaccurate. Although this plot is more accurate than the original one, it is still missing a very important feature - curvature. Curvature is also called concavity.

Over a given interval, a function can either be concave up, concave down, or have no concavity. A function is concave up when it is curved upwards, like a bowl, concave down when it is curved downward like a hill, and is neither if it has no curvature. Referencing the sine function, it is concave up over the interval $[\pi, 2\pi]$, where is it shaped like a bowl, and concave down over the interval $[0, \pi]$, where is it shaped like a hill. There are no intervals over which the sine function has no concavity or curvature. For an example of a function without concavity, consider a straight line.

Concavity, or curvature of a function is related to a changing steepness of the curve. Thus, in order to determine the concavity of a function we need to know something about the magnitude of its first derivative, but we can't simply evaluate the first derivative at all points within an interval, as there are too many points to do so. We'll have to use a different method. Consider the function x^3 (see figure 4.9).

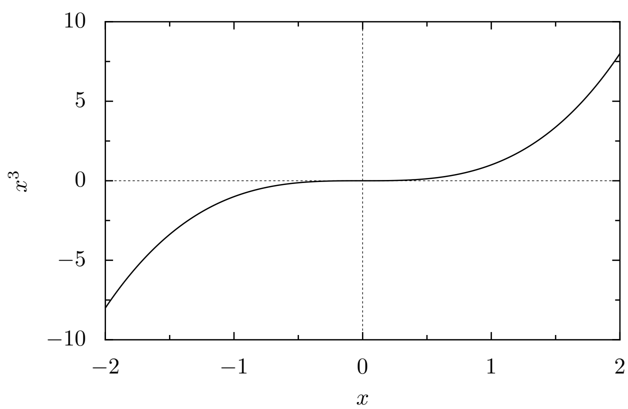


Figure 4.9: The graph of a function that is both concave up and down.

First note that this function is concave down for $x < 0$ and concave up for $x > 0$. Also notice that $f'(x) = 3x^2$, so the function $f(x) = x^3$ is increasing at all points other than $x = 0$ (because $3x^2 > 0$ for all $x \neq 0$). As $x \nearrow 0$, we see that the function begins to increase more slowly, corresponding to a decreasing first derivative, and as x becomes larger than 0, the rate at which the function increases increases, corresponding to an increasing first derivative. The changing magnitude of the first derivative lends curvature to the function.

Definition 4.11.1 (Concavity)**.** We say a function f is *concave up* on an interval I if f' is increasing on I. Similarly, we say a function f is *concave down* on an interval I if f' is decreasing on I.

From the above definition we have translated a visual property of how a function is curved into a mathematical property, related to a changing first derivative. Rather than trying to evaluate the first derivative of a function at many points over an interval, we look at the second derivative, which tells us the rate of change of the first derivative.

Theorem 4.11.1 (Second Derivative test for Concavity). *Let f be a twice-differentiable function.*

1. *If $f'' > 0$ on I, then f is concave up on I.*

2. *If $f'' < 0$ on I, then f is concave down on I.*

In addition to identifying the intervals over which a function is concave up and down, we are interested in identifying the points where concavity can possibly change. A point where the concavity of a function changes is called an *inflection point.* Using similar arguments as those made with critical points, we can see that if either $f''(c) = 0$ or $f''(c)$ does not exist, then c is a possible inflection point. If the second derivative changes sign around one of these points (from positive to negative, or negative to positive), then the point is an inflection point, which is a point where the function $f(x)$ changes concavity. If the second derivative does not change sign (ie. it goes from positive to zero to positive), then it is not an inflection point ($x = 0$ with $f(x) = x^4$ is an example of this).

Before we move onto using concavity as a part of curve sketching, we note that using a function's concavity can be a helpful tool for classifying its extrema. Consider a function f with $f'(c) = 0$. Now suppose $f'' > 0$ on an interval around the critical point c, which implies the first derivative is continuous and increasing on this interval. It follows that the first derivative must begin negative in order to increase to 0, and become positive after increasing beyond 0. In other words, such a point is a local minimum. There is a similar result for a negative second derivative, which we summarize in the following theorem.

Theorem 4.11.2 (Second Derivative Test for Local Extrema). *Suppose $f'(c) = 0$, and f'' is continuous on an open interval around $x = c$.*

1. *If $f''(c) < 0$, then f has a local maximum at $x = c$*

2. *If $f''(c) > 0$, then f has a local minimum at $x = c$*

3. *If $f''(c) = 0$, then the test is inconclusive. f may have a local maximum, minimum, or neither at $x = c$.*

Why do we require that f'' be continuous in the hypothesis of the above theorem? The reason is as follows. If f'' is continuous, then the fact that $f''(c) > 0$ or $f''(c) < 0$ implies that $f'' > 0$ or $f'' < 0$ on some interval around $x = c$, because the function is continuous so it can't jump. It's noteworthy that in many cases we cannot apply this theorem. First, we cannot use it for critical points at which the first derivative does not exist. Second, it does not give us any information for when the second derivative is 0 as well (such as x^3, x^4, etc). Finally, it is not useful for classifying endpoints. Despite these restrictions it is sometimes useful, for the simple fact that when we want to graph a function, we need to find the second derivative anyway.

Now equipped with our knowledge of concavity we can move on to sketching functions. The basic procedure will to be to start with the first derivative to identify the critical points of the function. This will partition the domain of our function into subintervals over which it is increasing, decreasing, or neither. Next we will want to identify the inflection points, dividing the domain further into subintervals with different concavities. Finally, we evaluate the function at the critical points, inflection points, and end points, *connecting the dots* using our knowledge of the way the function changes.

Example 1. Sketch $f(x) = x^3$ on $[-2, 2]$ using the first and second derivatives.

Solution We are already familiar with the graph of this function, so this will be a bit of a warmup. By repeated applications of the power rule, we find that

$$f'(x) = 3x^2 \quad \text{and} \quad f''(x) = 6x.$$

Since $f'(0) = 0$ and $f''(0) = 0$, we have that $x = 0$ is a critical point and a possible inflection point, but we cannot use the second derivative test because $f''(0) = 0$. Evaluating the function at the critical points and endpoints, we find that $f(-2) = -8$, $f(0) = 0$, and $f(2) = 8$. Since the first derivative cannot change signs on the subinterval $(-2, 0)$, and $f(0) > f(-2)$, it follows that the first derivative is positive on $(-2, 0)$. Likewise, the first derivative is positive on $(0, 2)$, so $f(0)$ is not an extreme value. Considering the second derivative, we see that for $x < 0$ we have $f''(x) < 0$, so $f(x)$ is concave down. For $x > 0$ we have $f''(x) > 0$, so $f(x)$ is concave up. Since the second derivative exists and changes signs around 0, $x = 0$ is an inflection point. Now we can plot the endpoints and the critical point, and use our knowledge of the second derivative to connect the dots.

Example 2. Sketch $f(x) = x^2$ on $[-3, 3]$ using the first and second derivatives.

Solution By repeated applications of the power rule, we find that

$$f'(x) = 2x \quad \text{and} \quad f''(x) = 2.$$

Thus, we have a critical point at $x = 0$, and evaluating the function at the critical points and endpoints we find $f(-3) = 9$, $f(0) = 0$, and $f(3) = 9$. It follows from the mean value theorem that the function is decreasing on $(-3, 0)$ and increasing on $(0, 3)$, which is confirmed by looking at the first derivative. It follows $f(0)$ is a local minimum. Also, using the second derivative test we see $f''(0) > 0$, which implies $f(0)$ is a local minimum. In fact, for all x, the second derivative $f''(x) > 0$, so the function $f(x)$ is always concave up. Now we plot the endpoints and critical points, connecting the dots with the knowledge that the function is always concave up.

Example 3. Find $f'(x)$ and $f''(x)$ if $f(x) = 2x$. Use this to analyze the familiar graph of the function.

Solution We find

$$f'(x) = 2 \quad \text{and} \quad f''(x) = 0.$$

For all x, the first derivative $f'(x) > 0$, so the function $f(x)$ is always increasing. Also, for all x, the second derivative is 0. This corresponds to a graph that does not have any concavity, which is the familiar line.

Example 4. Sketch $f(x) = \dfrac{x}{x-1}$ on $[-5, 5]$ using the first and second derivatives.

Solution The first thing to notice is that we have division by 0 at $x = 1$, so it is possible that our function could increase or decrease without bound as $x \to 1$. We find that

$$\lim_{x \nearrow 1} \frac{x}{x-1} = -\infty \quad \text{and} \quad \lim_{x \searrow 1} \frac{x}{x-1} = \infty.$$

In order to aid ourselves in sketching the graph of this function, we can draw a vertical dotted line at $x = 1$, called a vertical asymptote. This will provide us a guideline for sketching the fact that our function is increasing and decreasing without bound around this point of interest. Next, we need to calculate the first and second derivatives of f. Using the quotient rule we find

$$f'(x) = \frac{(x-1) \cdot 1 - x \cdot 1}{(x-1)^2} = \frac{-1}{(x-1)^2} = -(x-1)^{-2}.$$

Now we can use the power rule and chain rule, to find

$$f''(x) = 2 \cdot (x-1)^{-3} \cdot 1 = \frac{2}{(x-1)^3}.$$

Now we can use the above derivatives to analyze the behavior of the function. Since both the first and second derivative are fractions in which the numerator is constant, neither can equal zero. Thus, our only point of interest is $x = 1$, where the first and second derivative are not defined, because the function itself is not defined at $x = 0$. Since the denominator is nonnegative, the first derivative is negative for all $x \neq 1$. This indicates that for all $x \neq 1$ the function $f(x)$ is decreasing.

Looking at the second derivative we see that for $x < 1$ the function is concave down, as $f''(x) < 0$. For $x > 1$, $f''(x) > 0$, so the function is concave up. Even though the sign of the second derivative changes around $x = 1$, since the second derivative is not defined there, we do not call $x = 1$ an inflection point. Evaluating the function at the end points, we see $f(-5) = 5/6$ and $f(5) = -5/4$. Even with this much information, it is difficult to construct an accurate graph of this function. The simplest way to construct a more accurate sketch is to plot some additional points, such as $(0, f(0) = 0)$ and $(2, f(2) = 1)$ (see figure 4.10).

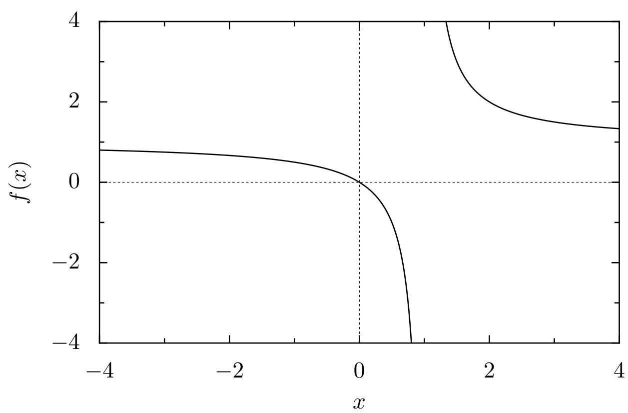


Figure 4.10: Graph of $f(x) = \frac{x}{x-1}$.

Example 5. Sketch $f(x) = 2x^3 - 7x^2 + 5$

Solution The first step is to find the first and second derivatives. We find that

$$f'(x) = 6x^2 - 14x \quad \text{and} \quad f''(x) = 12x - 14.$$

The first derivative is defined for all x, so in order to find the critical points we solve $f'(x) = 0$, finding

$$0 = 6x^2 - 14x = x(6x - 14).$$

Thus, we have critical points when $x = 0$ and $x = 14/6 = 7/3 = 2.\overline{3}$. These points divide $f'(x)$ into 3 intervals:

$$(-\infty, 0), \quad (0, 7/3), \quad \text{and} \quad (7/3, \infty).$$

Now we need to check the sign of the derivative over each of these subintervals (one point per interval is sufficient, as the derivative only changes sign at a critical point). By looking at $f'(x)$, we can see that for very large or small values of x, the positive quadratic term will dominate the negative first-order term, making the first derivative positive, so the function is increasing in the first and last interval. Now we can use a sample point of $x = 1$, which lies between the two critical points, to find that $f'(1) = 6-14 = -8 < 0$. Thus, in the middle interval the function is decreasing.

In order to search for possible inflection points, we set the second derivative equal to 0, and solve for x. We find that $0 = 12x - 14$, so $x = 14/12 = 7/6 = 1.1\overline{6}$. Now we must compare the sign of the second derivative for $x > 1.1\overline{6}$ and $x < 1.1\overline{6}$ to see if it an inflection point or not. We find that for $x < 1.1\overline{6}$ we have $f''(x) < 0$, so the function is concave down. For $x > 1.1\overline{6}$ we have $f''(x) > 0$, so the function is concave up. Thus, it is truly an inflection point, and we know the concavity of the function changes from down to up at this point.

The final step is to plot the function at the above points of interest: $(0, f(0))$, $(1.1\overline{6}, f(1.1\overline{6}))$, and $(2.\overline{3}, f(2.\overline{3}))$, which correspond to $(0, 5)$, $(1.16, -1.35)$, and $(2.3, -7.7)$. Once we have plotted these key points, we simply need to connect the dots, and be cognizant of the concavity and whether or not the function is increasing or decreasing. The final graph is pictured in figure 4.11.

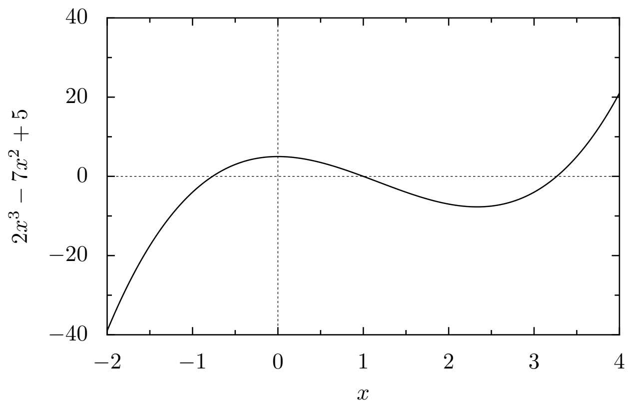


Figure 4.11: $2x^3 - 7x^2 + 5$.

4.12 Optimization

A solution to nearly any problem in life involves tradeoffs. In economics there is a tradeoff between selling more product at a lower price and selling less product at a higher price. This leads naturally to the question: what is the ideal level of production? A worker in a capitalist state has to balance time spent working to make more money, and having time available to spend that money. It is an interesting problem to try and find the ideal tradeoff in each of these situations.

In order to solve such problems, we need to find a point in the tradeoff where a desired quantity is maximized, or an undesired quantity is minimized. In life, we try to find a balance between work and free time in order to maximize our happiness. In economics, we look at maximizing profit. In manufacturing, we might want to minimize the amount of material we need to use. All of these problems are optimization problems, where the goal is to find an optimal solution, that maximizes or minimizes the quantity of interest. In order to solve optimization problems, we need to construct a mathematical model for the situation, and solve for the global minimum or maximum.

Example 1. Suppose you are trying to maximize your free time in order to get the most enjoyment out of life. In a given day you have 24 hours, but you must take out time for work and school, as well as eating and sleeping. Finally, if you do not sleep enough in a given day you will be tired and inefficient, wasting time. In order to try and find balance in your life, you have developed the following model for your free time

$$F(t) = 24 - I - t - \frac{100}{(t+1-s)^4}.$$

To use your model, you first need to set the parameters I and s, where I represents the amount of time you must spend working and in school, and s represents the minimum amount of time you need to sleep in order to be able to function. Once you have set these parameters, you must find a value of t for the amount of time you should spend sleeping and eating, in order to maximize your free time F.

Solution In order to maximize F, we need to look for a global maximum. The domain of this function must surely be smaller than $(s-1, 24-I)$, because these extremes represent the minimum amount of time we could reasonably spend sleeping, and the maximum amount of time we could possibly sleep in a day (if we slept the entire day). These extremes are unlikely to be ideal, so we should seek a maximum on the interior of the domain, by looking for critical points.

$$F'(t) = -1 + \frac{400}{(t+1-s)^5} = 0$$

which implies that

$$t = (400)^{1/5} + s - 1.$$

Looking at the second derivative we find

$$F''(t) = -\frac{2000}{(t+1-s)^6} < 0,$$

which means our function is concave down, and we have a maximum. For this value of t, we find that

$$F((400)^{1/5} + s - 1) = 24 - I - (400)^{1/5} - s + 1 - \frac{100}{((400)^{1/5} + s - 1 + 1 - s)^4} \approx 20.86 - I - s.$$

Thus, for an individual with a work load of $I = 10$ and a minimum sleep tolerance of $s = 7$, the ideal amount of time to spend sleeping and eating is

$$t = 9.3,$$

and such an individual has

$$F(9.3) = 20.86 - 10 - 7 = 3.86$$

hours of free time each day.

Example 2. The ideal dosage of a medication is a tradeoff between the efficacy of the drug and the severity of its side effects. If the side effects of a treatment are too severe, it will have to be discontinued, so it will fail to treat the patient. Similarly, if the dosage of the medicine is too low, its efficacy will not be sufficient to treat the patient. Suppose the probability of an effective treatment that accounts for this trade off is given by

$$P(x) = \frac{x^{1/2}}{1+x},$$

where x is the dosage of the medication given to the patient. What is the ideal dosage?

Solution To find the solution to this problem, we want to maximize $P(x)$, so we are looking for a global maximum. First we will need to find and classify the critical points, and finally compare them to the values of the endpoints.

$$P'(x) = \frac{(1+x)\frac{1}{2}x^{-1/2} - x^{1/2}}{(x+1)^2}$$

A fraction is only zero when its numerator is zero, so we need to solve

$$\begin{aligned} \frac{1}{2}(1+x)x^{-1/2} - x^{1/2} &= 0 \\ \frac{1}{2}(1+x) - x &= 0 \\ \frac{1}{2} &= x(1-\frac{1}{2}) \\ 1 &= x \end{aligned}$$

We find that $P(1) = 0.5$. We find that $P(0) = 0$, and since the sign of the derivative can only change at a critical point, and $P(1) > P(0)$, we find that $P'(x) > 0$ for $0 < x < 1$ (the derivative is not defined for $x = 0$). Similarly, to find the sign of the derivative for $x > 1$, we can choose a sample point, and the sign of the derivative of that point will give us the sign of the derivative for all $x > 1$. Choosing $x = 4$ (because the square root of 4 is easy to evaluate) we find

$$P'(4) = \frac{(1+4)\frac{1}{2}4^{-1/2} - 4^{1/2}}{(4+1)^2} = \frac{5/4) - 2}{25} = \frac{1.25 - 2}{25} < 0$$

Thus, for $x > 1$, the function is decreasing, so we find that x is a local maximum. Since there is no right endpoint, we conclude that $x = 1$ is the global maximum.

Example 3. An open-top box can be made by cutting out squares from a rectangular piece of material and folding up the resulting flaps. For a piece of material with length l and width w, what length for the sides of the cut-out squares will maximize the volume of the resulting box? For simplicity, assume $w < l$.

Solution In this problem we want to maximize the volume of the box. In order to do so, we first need to find an equation for the volume of the box, and then search for a global maximum. The volume of the box is given by multiplying the length, width, and height of the box together. Since we are removing squares of length x from the original piece of material, the length and width of the resulting box will be $l-2x$ and $w-2x$, respectively. The height of the box will simply be x. Thus, we find that

$$V(x) = x(l-2x)(w-2x) = 4x^3 - (2l+2w)x^2 + lwx$$

The domain of this function is $[0, w/2]$, because we can begin by cutting no material, or we can cut all the way through the width of the box. At both of these points we will clearly have 0 volume, and since this is a continuous function with positive volume on the interior of its domain, we know that we will have a global maximum. In order to find this global maximum we differentiate, yielding

$$V'(x) = 12x^2 - 4(l+w)x + lw$$

In order to find the roots of this equation, we apply the quadratic formula, which tells us that

$$x = \frac{4(l+w) \pm \sqrt{16(l+w)^2 - 48lw}}{24} = \frac{4(l+w \pm \sqrt{(l+w)^2 - 3lw})}{24} = \frac{l+w \pm \sqrt{l^2+w^2-lw}}{6}$$

How can we interpret these roots? Let us use the second derivative test to determine if they are minima or maxima.

$$V''(x) = 24x - 4(l+w)$$

For the first of these roots, we have

$$\begin{aligned} V''\left(\frac{l+w-\sqrt{l^2+w^2-lw}}{6}\right) &= 24\frac{l+w-\sqrt{l^2+w^2-lw}}{6} - 4(l+w) \\ &= 4(l+w) - 4\sqrt{l^2+w^2-lw} - 4(l+w) = -4\sqrt{l^2+w^2-lw} < 0 \end{aligned}$$

and for the second we have

$$\begin{aligned} V''\left(\frac{l+w+\sqrt{l^2+w^2-lw}}{6}\right) &= 24\frac{l+w+\sqrt{l^2+w^2-lw}}{6} - 4(l+w) \\ &= 4(l+w) + \sqrt{l^2+w^2-lw} - 4(l+w) = 4\sqrt{l^2+w^2-lw} > 0 \end{aligned}$$

It follows that the first is a local maximum, and the second is a local minimum. In the special case $l = w$, we have our maximum at

$$x = \frac{l+w-\sqrt{l^2+w^2-lw}}{6} = \frac{2w-\sqrt{2w^2-w^2}}{6} = \frac{2w-w}{6} = \frac{w}{6}$$

which tells us that we should use $1/6^{th}$ of the width of the original material for the width of each cut-out square. We see that for a square box $l = w$, the other critical point we found corresponds to

$$x = \frac{l+w+\sqrt{l^2+w^2-lw}}{6} = \frac{2w+\sqrt{2w^2-w^2}}{6} = \frac{2w+w}{6} = \frac{w}{2}$$

which would tell us to cut out the entire box. This is at the endpoint of the domain, and corresponds to 0 volume.

Example 4. Determine the dimensions (radius and height) of a 1 Liter (1000 cm^3) right circular cylindrical can that minimizes the amount of material used (ignore the thickness of the material, and waste).

Solution First we must note that we are constrained in volume to 1 Liter. Thus,

$$\pi r^2 h = 1000 \text{ cm}.$$

Since the thickness of the material is ignored, the amount of material used in our can will correspond to its surface area. In order to find the surface area we must separately add the surface area of the top and bottom, as well as cylindrical shell. The surface area of the shell can be found by thinking about unrolling the cylinder. Upon doing so, we will have a rectangle with width $2\pi r$ (the circumference of the circle) and height h, so it will have an area of $2\pi rh$. The top and bottoms are both circles, so they have area given by the familiar formula πr^2. Since there are two of them, we arrive at the final result

$$A = 2\pi r^2 + 2\pi rh.$$

Although this appears to be a function of two variables, by setting a given radius we already determine the height, due to the fact that we must have a volume of 1 Liter. Thus, we can rewrite surface area in terms of a single variable, using $h = 1000/\pi r^2$. We find

$$A = 2\pi r^2 + 2\pi r\frac{1000}{\pi r^2} = 2\pi r^2 + \frac{2000}{r}.$$

In order to find a minimum for the material used, we investigate the first derivative

$$\frac{dA}{dr} = 4\pi r - \frac{2000}{r^2}.$$

Stting the first derivative to 0 we find that

$$4\pi r = \frac{2000}{r^2},$$

or

$$r = \sqrt[3]{\frac{500}{\pi}}.$$

By simple reasoning we can be pretty confident that this will minimize the surface area, as a very tall and skinny or very short and fat cylinder will use a lot more material. We check this intuition with the second derivative

$$\frac{d^2A}{dr^2} = 4\pi + \frac{4000}{r^3} > 0,$$

which confirms that this is a minimum. If we use a little algebra, we can show that for this value of r we have

$$h = \frac{1000}{\pi r^2} = \frac{1000}{\pi}\cdot\Big(\frac{\pi}{500}\Big)^{2/3} = \Big(\frac{1000^3\pi^2}{\pi^3 500^2}\Big)^{1/3} = \Big(\frac{4000}{\pi}\Big)^{1/3} = \Big(\frac{2^3\cdot 500}{\pi}\Big)^{1/3}.$$

This implies

$$h = 2\Big(\frac{500}{\pi}\Big)^{1/3} = 2r \approx 10.84 \text{ cm}.$$

which gives the interesting result that the material is minimized when the diameter of the cylinder is the same as its height.

Example 5. An artisan uses mahogany to produce 5 cabinets each day. The delivery costs of the mahogany cost \$5000, regardless of how much wood is ordered, and the wood costs \$10 per day per unit stored, where one unit of wood is enough to construct a single cabinet. How much material should be ordered at a time, and how often, in order to minimize the costs of ordering and storing wood?

Solution The first thing to note is that there is no need for the artisan to order extra wood until she runs out of wood. Thus, if she orders wood every n days, she should order $5n$ units of wood. Let's suppose that our artisan is in the middle of a cycle. During a day she will withdraw her last 5 units of wood from storage, and then place $5n$ more from the delivery in storage, so her storage costs will be $5n$ for that day. The next day she will store $5(n-1)$, then $5(n-2)$, until she stores only 5 units. The next day after that she will withdraw the final 5 and be back up to storing $5n$ units from the new shipment. In summary, during an n day period of time she will pay storage costs for

$$5n + 5(n-1) + 5(n-1) + \ldots + 5 = 5(n + (n-1) + \ldots + 1) = 5\sum_{i=1}^{n} i.$$

Here we use the convenient sigma ($\sum$) notation, in order to succinctly represent the sum. To evaluate such a sum we evaluate the expression i for each of the values 1 to n, and sum them together. Thus, for instance

$$\sum_{i=1}^{3} i = 1 + 2 + 3 = 6.$$

In order to solve this problem, we need to figure out how to sum the first n natural numbers. There's a useful trick for seeing how to do so. We begin by writing the numbers 1 to n from left to right, then below write them from right to left, and sum the result.

$$\begin{array}{cccccc} & 1 & 2 & \ldots & (n-1) & n \\ + & n & (n-1) & \ldots & 2 & 1 \\ \hline & (n+1) & (n+1) & \ldots & (n+1) & (n+1) \end{array}$$

Here we have n terms all with a value of $(n+1)$, so we have $n \cdot (n+1)$. However, we have to be careful because we actually added all of the numbers 1 to n together twice, so we need to divide by 2. In the end we find

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

Finally, we can see that the cost for a single delivery cycle n days long is

$$5000 + 5\frac{n(n+1)}{2} \cdot 10 = 5000 + 25n(n+1).$$

To obtain the cost per day we need to divide by n, finding

$$c(n) = \frac{5000}{n} + 25(n+1) = \frac{5000}{n} + 25 + 25n.$$

To minimize the daily cost we consider the first derivative

$$c'(n) = -\frac{5000}{n^2} + 25 = 0.$$

This implies

$$n = \pm\sqrt{200} \approx \pm 14.14.$$

Since we cannot have negative days, we find $n = 14.14$. Looking at the second derivative

$$c''(n) = \frac{10000}{n^3} > 0,$$

so this is a minimum. Since we cannot have a delivery in a fraction of a day, we know our minimum should lie at either 14 or 15 days. We see that

$$c(14) = \frac{5000}{14} + 25(15) \approx 732.14 \quad \text{and} \quad c(15) = \frac{5000}{15} + 25(16) \approx 733.33.$$

Both of these costs are very similar, but it is ideal for the artisan to order wood every 14 days.

A powerful concept in geometric optics is Fermat's principle of least time. The principle states that light will travel between two points by taking the path that takes the least time. If light is traveling in a single medium, the path that will take the least time is the path that minimizes the distance traveled - a straight line. However, because the speed of light depends on the medium that light is traveling through, when light travels through multiple media it may no longer take a straight line path. The bending of light as it moves from one medium to another is known as refraction.

To undersatnd why the light should bend as it travels between two different media consider the following analogy. If you were on a beach trying to rescue someone in the water, you would want to reach him or her as quickly as possible. It is clear that you will be able to run faster on land than swim in water, but at some point you will need to enter the water to reach the person. Depending on the path you take you will spend more or less time on land and in water. In order to reach the person as soon as possible you need to balance taking a direct path, and spending as much time on land as possible. The most direct path might require a lot more swimming time, so it could be more efficient to take a less direct path, in order to spend more time running quickly on land.

Now let's consider the case of light, traveling from point A in medium 1 to point B in medium 2. Let c_1 denote the speed of light in the first medium, and c_2 the speed of light in the second medium. Let us set up a coordinate system, in which the x-axis resides on the junction of the two media (assume that the junction is flat). Let the coordinates of A be $(0, a)$, and the coordinates of B be $(d, -b)$. Let P be the point at which the light strikes the junction, given by coordinates $(x, 0)$. Define the angle of incidence as the angle θ_1 which is the angle between the ray of light and the normal from the junction as it exits media 1. Let θ_2 denote the angle from which the light enters media 2, with respect to the normal. See figure 4.12 for a graphical representation of this setup.

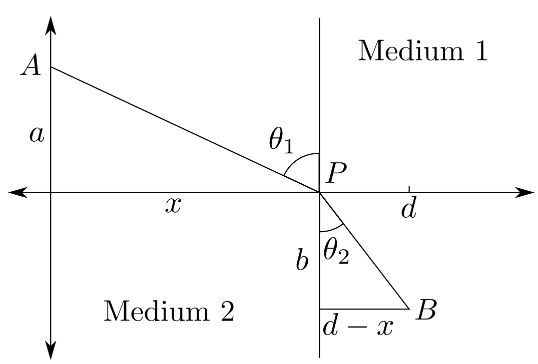


Figure 4.12: Refraction of light using Fermat's principle of least time.

The time it takes the light travel through a given medium is given by the distance it travels divided by the speed at which it travels. We find the relationships

$$t_1 = \frac{AP}{c_1} = \frac{\sqrt{a^2+x^2}}{c_1} \quad \text{and} \quad t_1 = \frac{PB}{c_2} = \frac{\sqrt{b^2+(d-x)^2}}{c_2},$$

where t_1 denotes the time it takes to travel through the first medium, and t_2 the time through the second medium. The total time is then

$$t = t_1 + t_2 = \frac{\sqrt{a^2+x^2}}{c_1} + \frac{\sqrt{b^2+(d-x)^2}}{c_2}$$

Now in this equation all quantities are fixed except for x. Thus, we want to find the value of x that minimizes the time it takes to travel. Taking the derivative we find

$$\begin{aligned}\frac{dt}{dx} &= \frac{1}{2}\cdot\frac{1}{c_1\sqrt{a^2+x^2}}\cdot 2x + \frac{1}{2}\cdot\frac{1}{c_2\sqrt{b^2+(d-x)^2}}\cdot 2(d-x)\cdot(-1)\\ &= \frac{x}{c_1\sqrt{a^2+x^2}} - \frac{d-x}{c_2\sqrt{b^2+(d-x)^2}}.\end{aligned}$$

We can greatly simplify this expression by rewriting it in terms of the angles of incidence and refraction, noting that

$$\sin(\theta_1) = \frac{x}{\sqrt{a^2+x^2}} \quad \text{and} \quad \sin(\theta_2) = \frac{d-x}{\sqrt{b^2+(d-x)^2}}.$$

Using these simplifications we find

$$\frac{dt}{dx} = \frac{\sin(\theta_1)}{c_1} - \frac{\sin(\theta_2)}{c_2}.$$

Now setting the derivative to 0 we find

$$\frac{\sin(\theta_1)}{c_1} = \frac{\sin(\theta_2)}{c_2}.$$

This relationship is called Snell's Law. We can verify this is a minimum by noticing that the derivative at the point $x = 0$ is negative and the derivative at $x = d$ is positive. It turns out that the speed of light in a vacuum is the constant $c \approx 3 \cdot 10^8$ m/s. We can rewrite c_1 and c_2 in terms of c by defining the indices of refraction so that $c_1 = c/\eta_1$ and $c_2 = c/\eta_2$. Then we can rewrite Snell's law using the indices of refraction as

$$\eta_1 \sin(\theta_1) = \eta_2 \sin(\theta_2).$$

Note that since the speed of light is largest in vacuum, in any other medium the index of refraction $\eta \geq 1$. The index of refraction for air is about 1, for water is 1.33, and for glass is about 1.5 (depending on the glass). Suppose that our first medium is water and the second is air. Then Snell's law tells us

$$1.33 \sin(\theta_1) = \sin(\theta_2).$$

We know that the maximum value for the sine function is 1, which occurs when $\theta_2 = \pi/2$. What happens then if

$$\theta_1 > 1/1.33 \approx .75.$$

For instance, let $\theta_1 = \pi/4$. In this case it will not be possible to find a value for θ_2 to satisfy the above equation. In such a situation, there is no refracted light; instead, the light is all reflected. This phenomenon is called total internal reflection. This is the principle used in optical waveguides, which is the principle that allows us to transmit light through fiber optic cables.

If we are dealing with a reflected beam, what should the path of the beam look like? Let us setup a situation with a mirror (but we could just as well use a junction like in the above example, which for all intents and purposes acts like a mirror), two points A and B, and apply Fermat's principle. In this case the path of least time would be just a straight line between the two points, but this does not include the mirror at all, so tells us nothing about what the behavior of light reflected off the mirror is. Instead, we need to stipulate that the light *must* strike the mirror before it reaches point B. In this case, we can apply the principle of least time to see what angle the light reflects off of the mirror at. This situation is nearly identical to the light passing through two media, except that now the point B is on the same side of the junction, so its coordinates are (d, b). If we replace the angle θ_2 with the new angle between the ray of light and the normal of the junction, we can call θ_2 the angle of reflection. In fact, this situation is now mathematically identical to the previous one, except that $c_1 = c_2$. Thus, we find that

$$\sin(\theta_1) = \sin(\theta_2),$$

and for $0 \leq \theta \leq \pi/2$ which is necessitated by the physical situation, it follows that

$$\theta_1 = \theta_2.$$

Thus, the angle of incidence is the same as the angle of reflection.

Example 6. Consider a bee foraging flowers in order to eat their nectar. The supply of nectar of a given flower depletes as a bee consumes it, so the rate at which it consumes nectar decreases with time. Nevertheless, as it takes time to travel between flowers, the bee does not want to leave too soon, as it will spend more time traveling than consuming nectar. Let t denote the time spent on each flower, $F(t)$ denote the amount of nectar consumed per flower, and τ denote the time required to travel between flowers. The rate at which nectar is consumed $R(t)$ is the amount of nectar consumed divided by the time time it takes to consume that nectar, so we find the relationship

$$R(t) = \frac{F(t)}{t + \tau}$$

What is the optimal time for a bee to spend at each flower, as a function of $F(t)$?

Solution Although we do not have a specific function for $F(t)$, and thus $R(t)$, there are some inferences we can make about it based on the physical situation. It is clear that for $t = 0$, $F(t) = 0$, and thus $R(t) = 0$, because if no time is spent on a flower, then no nectar can be consumed. We also know that at some point in time all of the nectar on a flower will be depleted, so staying longer will have no benefit. Thus, if we consider a large enough interval of time (say all possible times), we know that the ideal solution, a maximum for $R(t)$, is going to occur at a critical point. Let us calculate the derivative of $R(t)$, in order to find what value of $F(t)$ will yield a critical point.

$$R'(t) = \frac{(t + \tau)F'(t) - F(t)}{(t + \tau)^2}$$

Thus, we will find a critical point when

$$(t + \tau)F'(t) - F(t) = 0$$

or

$$F'(t) = \frac{F(t)}{t+\tau} = R(t)$$

If there is a single solution to this equation, we know it is a maximum, because $R(t)$ is increasing at first, and clearly is decreasing for very large t (particularly after the flower has been depleted).

The above result is called the marginal value theorem. The marginal theorem says that the optimal time to leave a flower is when the instantaneous rate of food consumption is equal to the average rate of food consumption ($F'(t) = R(t)$). This is an intuitive result, because if $F'(t) > R(t)$, the bee is consuming more than it could on average elsewhere, and if $F'(t) < R(t)$, it is consuming less than it could on average elsewhere, so it should seek a new flower. Thus, $F'(t) = R(t)$ is the ideal time to leave, before the consumption on that flower becomes inefficient.

Example 7. Find the ideal amount of time to spend at each flower if

$$F(t) = \frac{t}{t+0.5}.$$

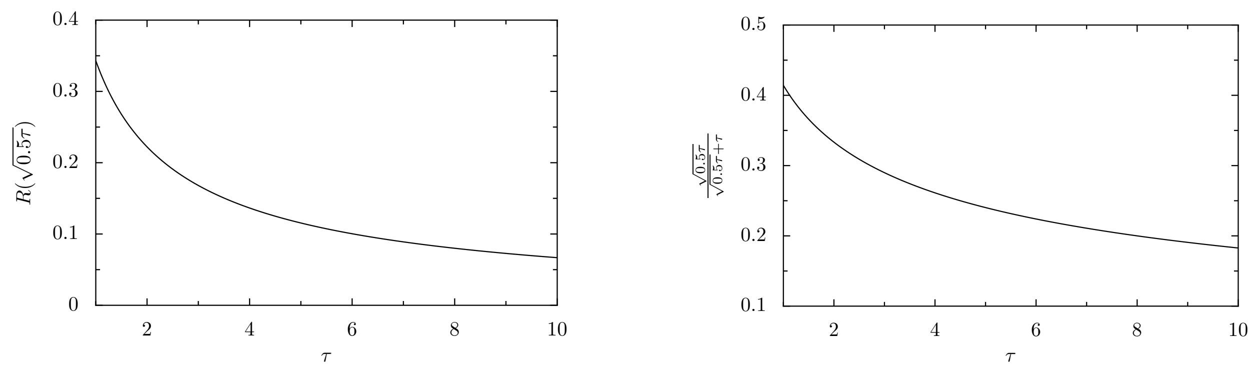
Solution To solve this problem, we will first calculate $F'(t)$, and then apply the marginal value theorem.

$$F'(t) = \frac{(t+0.5)-t}{(t+0.5)^2} = \frac{0.5}{(t+0.5)^2}.$$

Now we must solve the equation

$$\begin{aligned}
\frac{0.5}{(t+0.5)^2} &= R(t) = \frac{F(t)}{t+\tau} \\
\frac{0.5}{(t+0.5)^2} &= \frac{t}{(t+0.5)(t+\tau)} \\
\frac{0.5}{(t+0.5)} &= \frac{t}{(t+\tau)} \\
0.5(t+\tau) &= t(t+0.5) \\
0.5t+0.5\tau &= t^2+0.5t \\
0 &= t^2+0.5t-0.5t-0.5\tau \\
t^2 &= 0.5\tau \\
t &= \sqrt{0.5\tau}
\end{aligned}$$

Based on the above result we can see that as the time it takes to travel between flowers increases, the amount of time t that the bee spends on a single flower increases. Further, we can plot $R(\sqrt{0.5\tau})$ to see the rate at which nectar is collected as a function of the ideal time, and plot $t/(t+\tau) = \sqrt{0.5\tau}/(\sqrt{0.5\tau}+\tau)$ to see what the ideal fraction of time to spend on flowers as a function of τ (see figure 4.13).



(a) Average rate of consumption.

(b) Ideal amount of time to spend on each flower.

Figure 4.13: Behavior of the bee with respect to τ.

4.13 l'Hôpital's Rule

As we saw previously, when we consider a limit which is a quotient of two functions, it is possible for us to encounter indeterminate forms, where the limit of both the numerator and denominator individually are either 0 or ∞. In such a situation we cannot apply the quotient rule for limits, but if we were to blindly apply the rule we would have something that looks like

$$\frac{0}{0} \quad \text{or} \quad \frac{\infty}{\infty},$$

which are both meaningless expressions. As we saw before, whether or not such a limit exists depends on the rate at which the functions in the numerator and denominator approach 0 or ∞. Using this notion we were able to characterize the order of magnitude of growth and decay of functions. At the time we did not have access to l'Hôpital's rule, so we took such relationships for granted. However, with access to differentiation, we can actually evaluate limits of these types directly.

Theorem 4.13.1 (l'Hôpital's Rule)**.** *Suppose that f and g are differentiable on an open interval I containing $x_0 \in \mathbb{R}^* \equiv \mathbb{R} \cup \{\pm\infty\}$, and*

$$\lim_{x\to x_0} |f(x)| = \lim_{x\to x_0} |g(x)| = \infty$$

or

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0.$$

Further, suppose that $g'(x) \neq 0$ for all $x \neq x_0$ in I. It follows that

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)},$$

provided that the right limit exists, or is $\pm\infty$.

When we say that I contains ∞ or $-\infty$ we mean that I is either of the form $(-\infty, b)$ or (a, ∞)). This is useful if we have a limit of the form

$$\lim_{x\to\infty} \frac{e^x}{x} = \lim_{x\to\infty} \frac{\frac{d}{dx}e^x}{\frac{d}{dx}x} = \lim_{x\to\infty} \frac{e^x}{1} = \infty,$$

where $\lim_{x\to\infty} e^x = \lim_{x\to\infty} x = \infty$ allowed us to apply the theorem. This result tells us that e^x grows faster than x as $x \to \infty$.

As before, this allows us to classify the rates at which functions grow/decay as as $x \to x_0$, by seeing if the limit is 0, a nonzero constant, or $\pm\infty$. When the limit is 0, we say the denominator either grows on a higher order of magnitude, or decays on a lower order of magnitude than the numerator (depending on whether the functions are growing or decaying). If the limit is $\pm\infty$, then we say the denominator either grows on a lower order of magnitude, or decays on a higher order of magnitude. Finally, if the limit is a nonzero constant, we say the functions are growing or decaying on the same order of magnitude. This is a very useful and powerful tool when we want to think about the growth and decay of functions.

We can even apply this theorem to functions where the derivative is not defined at the point of interest but the limit still makes sense, such as a function like $\ln(x)$. The natural logarithm and its derivative are not defined for $x_0 = 0$, but we can still look at the limit as $x \to 0$. For instance,

$$\lim_{x\to 0} \frac{\ln(x)}{1/x} = \lim_{x\to 0} \frac{1/x}{-1/x^2} = \lim_{x\to 0} \frac{-x^2}{x} = 0.$$

which tells us that as $x \to 0$ we have $\ln(x) \to -\infty$ on a lower order of magnitude than $1/x \to \infty$. Finally, we can use l'Hôpital's rule multiple times successively. l'Hôpital's rule tells us that if a limit is in an indeterminate form, we can evaluate the limit by looking at the derivatives of the functions involved in the limit, provided that the derivatives exist. If we are faced with another indeterminate form, we can once again apply l'Hôpital's rule, provided that the second derivatives exist. In most cases, we should eventually find a limit that is not in an indeterminate form, which will tell us the limit of the original two functions. It should be emphasized that if a limit is not in an indeterminate form, then l'Hôpital's rule cannot be applied.

Example 1. Evaluate $\lim\limits_{x\to 0} \dfrac{e^{3x}-1}{x}$.

Solution By inspection, we can see that both the numerator and denominator are differentiable, and the limit is of the indeterminate form $0/0$ so it is okay to apply the rule. Thus, we find

$$\lim_{x\to 0} \frac{e^{3x}-1}{x} = \lim_{x\to 0} \frac{3e^{3x}}{1} = 3.$$

Example 2. Evaluate $\lim\limits_{x\to \infty} \dfrac{e^{3x}-1}{x}$.

Solution In this situation, we know that e^x approaches ∞ faster than x as $x \to \infty$. Thus, the limit should be ∞. Let us verify this using l'Hôpital's rule.

$$\lim_{x\to \infty} \frac{e^{3x}-1}{x} = \lim_{x\to \infty} \frac{3e^{3x}}{1} = \infty.$$

Example 3. Evaluate $\lim\limits_{x\to 0} \dfrac{e^{3x}}{x}$.

Solution In this case we do not have an indeterminate form, so we cannot apply l'Hôpital's rule. As $x \to 0$ the numerator approaches 1 while the denominator approaches 0. In this situation we can only conclude that

$$\lim_{x\to 0} \frac{e^{3x}}{x} \quad \text{does not exist.}$$

because looking at the one-sided limits we have

$$\lim_{x\nearrow 0} \frac{e^{3x}}{x} = -\infty \quad \text{and} \quad \lim_{x\searrow 0} \frac{e^{3x}}{x} = \infty.$$

Example 4. Evaluate $\lim\limits_{x\to 1} \dfrac{x^2-1}{x-1}$.

Solution First notice that this limit is in the indeterminate form $\frac{0}{0}$, so we can apply l'Hôpital's rule.

$$\lim_{x\to 1} \frac{x^2-1}{x-1} = \lim_{x\to 1} \frac{2x}{1} = 2$$

Alternatively, we can write $x^2-1 = (x+1)(x-1)$, and cancel a factor of $x-1$ from the numerator and denominator. Then, we find

$$\lim_{x\to 1} \frac{x^2-1}{x-1} = \lim_{x\to 1} x+1 = 2.$$

Example 5. Evaluate $\lim\limits_{x\to 0} \dfrac{3x-\sin(x)}{x}$.

Solution This limit is in the indeterminate form $\frac{0}{0}$, so we can apply l'Hôpital's rule, finding

$$\lim_{x\to 0} \frac{3x-\sin(x)}{x} = \lim_{x\to 0} \frac{3-\cos(x)}{1} = 2.$$

Example 6. Evaluate $\lim\limits_{x\to 0} \dfrac{\sqrt{1+x}-1-x/2}{x^2}$.

Solution Once again, this limit is in the indeterminate form $\frac{0}{0}$, so we can apply l'Hôpital's rule.

$$\lim_{x\to 0} \frac{\sqrt{1+x}-1-x/2}{x^2} = \lim_{x\to 0} \frac{(1/2)(1+x)^{-1/2}-1/2}{2x} = \lim_{x\to 0} \frac{-(1/4)(1+x)^{-3/2}}{2} = -\frac{1}{8}.$$

Example 7. Evaluate $\lim\limits_{x\to\infty} \dfrac{\ln(x)}{2\sqrt{x}}$.

Solution Now we have an indeterminate form $\frac{\infty}{\infty}$. We find

$$\lim_{x\to\infty} \frac{\ln(x)}{2\sqrt{x}} = \lim_{x\to\infty} \frac{1/x}{1/\sqrt{x}} = \lim_{x\to\infty} \frac{1}{\sqrt{x}} = 0.$$

It is also possible for us to encounter indeterminate forms other than $0/0$ and ∞/∞. For instance, we might be faced with forms such as $0\cdot\infty$ or $\infty-\infty$. In order to deal with such forms, we need to use some algebraic tricks to write them in the form of $0/0$ or ∞/∞, and then we can apply l'Hôpital's rule. Consider the following examples.

Example 8. Evaluate the limit $\lim\limits_{x\to 0}\left(\dfrac{1}{x}-\dfrac{1}{\sin(x)}\right)$.

Solution Here when we evaluate this limit we have two vanishing denominators, which leaves us with $\infty-\infty$, a meaningless result. If we multiply these fractions to a common denominator, we find

$$\lim_{x\to 0}\left(\frac{1}{x}-\frac{1}{\sin(x)}\right) = \lim_{x\to 0} \frac{x-\sin(x)}{x\sin(x)}$$

The leaves us in the form of $0/0$, so we can apply l'Hôpital's rule.

$$\lim_{x\to 0} \frac{x-\sin(x)}{x\sin(x)} = \lim_{x\to 0} \frac{1-\cos(x)}{\sin(x)+x\cos(x)} = \lim_{x\to 0} \frac{\sin(x)}{2\cos(x)-x\sin(x)} = \frac{0}{2} = 0.$$

Example 9. Evaluate the limit $\lim\limits_{x\to\pi/2} \tan(x)\cdot\ln(\sin(x))$.

Solution Here we have a limit of the form $\infty\cdot 0$, but by simply writing $\tan(x) = 1/\cot(x)$ we have

$$\lim_{x\to\pi/2} \tan(x)\cdot\ln(\sin(x)) = \lim_{x\to\pi/2} \frac{\ln(\sin(x))}{\cot(x)},$$

which is of the form $0/0$. Applying l'Hôpital's rule

$$\lim_{x\to\pi/2} \frac{\ln(\sin(x))}{\cot(x)} = \lim_{x\to\pi/2} \frac{\cos(x)/\sin(x)}{-\csc^2(x)} = \lim_{x\to\pi/2} (-\cos(x)\sin(x)) = 0.$$

4.14 Newton's Method

An unfortunate truth of mathematics is that the majority of problems cannot be solved exactly - instead we need to use numerical approximations. Let's consider the problem of solving an algebraic equation. For a second-order polynomial we can use the quadratic formula to find its roots, and for third and fourth order polynomials there are even more complicated methods to find the roots. However, for a fifth-order or higher polynomial there is no general method for finding roots. Right now we're only considering polynomials. What about equations involving trigonometric functions? We previously considered the bisection method, which is one numerical method we can use to solve such equations. Newton's method is another method for solving such equations, which can be very useful, because it can converge to a result much more quickly, or provide us with a better estimate after fewer iterations.

The basic idea of Newton's method is of first-order approximation. As stated previously, the line tangent to a curve at a given point is the best first-order approximation of the curve at that point. This means that very close to the point of tangency the tangent line is an excellent approximation to a function, but as we move further away from the point of tangency the accuracy of the approximation decreases. The key here is that the function behaves similar to the tangent line. For this reason, it is reasonable to believe that a tangent line will cross the x-axis at point somewhere similar to where the function itself crosses the x-axis (this is useful because we can reforge any equation into a question of where a specific function crosses the x-axis, or is 0).

If this were as far as we could go Newton's method would not be that useful. The trick comes in as follows. By finding the intercept of the tangent line we get a new x-coordinate, which is presumably closer to the 0 of our function than the original point we chose. Now we construct a tangent line at the new point and see where it crosses the x-axis. This is presumably once again closer to our real solution. By continuing this process we get a sequence of approximations to our solution, which hopefully converge to, or become increasingly closer to, the real solution. If we look in the limit as the number of iterations of this process tends to ∞, we should hopefully find the actual solution to the problem at hand. See figure 4.14 for an illustration of Newton's method.

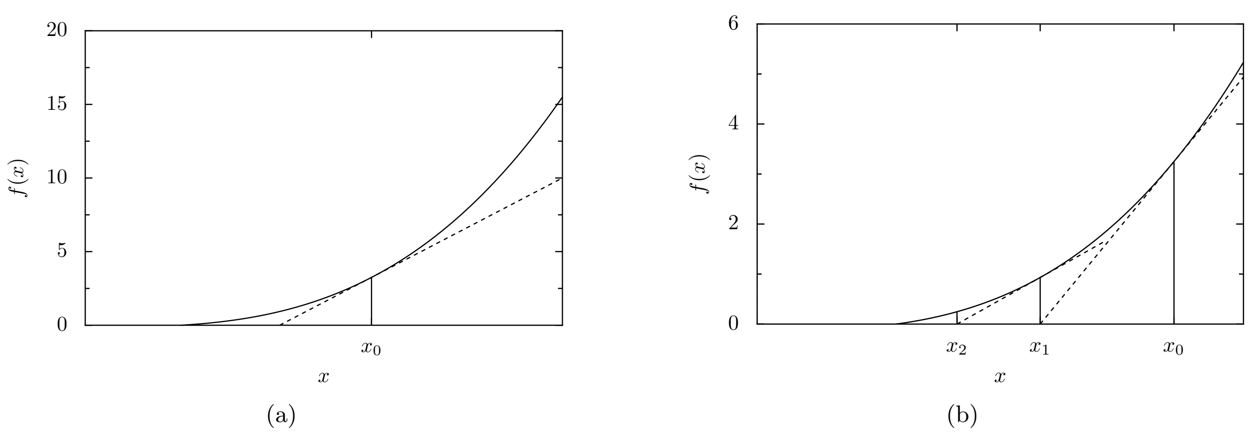


Figure 4.14: The initial guess generally provides a poor estimate of the actual root, but by iterating the process we hope to move closer and closer to the actual root.

In practice, we apply Newton's method in the following way. First, move all terms to one side of the equation in order to have an equation of the form $f(x) = 0$ (this transforms the problem of finding where two sides of an equation are equal to finding where a new function crosses the x-axis). Then, make an initial guess x_0 at a possible solution. Next, find the equation of the line tangent to the function $f(x)$ at the point x_0. Solve for the zero of the tangent line (which is a simple first-order equation), and use the result as a second guess for the solution to $f(x) = 0$. Continue iterating this process until solutions either converge or diverge.

Example 1. Solve of the equation $x^2 = 3$ using Newton's method.

Solution We already know that $\pm\sqrt{3}$ are solutions to this equation, but let's try and find them using Newton's method. The first step is to formulate the equation as $f(x) = x^2 - 3$, and set it equal to 0. Let us make an initial guess that $x = 2$ is a solution to the equation (even though we know it is a little high). We find $f'(x) = 2x$, so the equation of the tangent line is given by

$$\hat{f}_2(x) = f(2) + f'(2)(x-2) = (4-3) + 4(x-2) = 4x - 7.$$

If we set $\hat{f}_2(x) = 4x - 7 = 0$ we can see by inspection that the solution is $x = 7/4 = 1.75$, which is closer than 2 to $\sqrt{3} \approx 1.73205$. If we continue this process we can find

$$\hat{f}_{1.75}(x) = f(1.75) + f'(1.75)(x - 1.75) = (1.75^2 - 3) + 3.5(x - 1.75) = 3.5x - 6.0625.$$

Now when we set $\hat{f}_{1.75}(x) = 3.5x - 6.0625 = 0$ we find $x = 6.0625/3.5 \approx 1.73214$. After only two iterations we have found a solution that is relatively close to $\sqrt{3}$.

Since Newton's is an iterative process it is very useful to recast the process in a different form that simplifies calculations, and makes it much easier to implement Newton's method using a computer. Let us suppose we are on the n^{th} iteration of Newton's method, and we have found an x value of x_n. Setting the line tangent to $f(x)$ at x_n to 0 we get

$$0 = \hat{f}_{x_n} = f(x_n) + f'(x_n)(x - x_n),$$

where x is the solution to this equation that gives us our next value x. Solving this equation for x we find

$$\begin{aligned} f(x_n) + f'(x_n)(x - x_n) &= 0 \\ f'(x_n)(x - x_n) &= -f(x_n) \\ x - x_n &= -\frac{f(x_n)}{f'(x_n)} \\ x &= x_n - \frac{f(x_n)}{f'(x_n)} \end{aligned}$$

This will be x_{n+1}, the value we use in the next iteration. In conclusion, we have that

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

The above equation holds as long as $f'(x_n) \neq 0$ (if $f'(x_n) = 0$ then we have a horizontal tangent line, which does not have any zeros, unless it is zero everywhere). This simplifies the process of finding a numerical solution as follows. First, we set the equation $f(x) = 0$, and make an initial guess at a solution. Next, use the above formula to find a second guess, and continue iterating until the answers converge to a value up to the desired accuracy. In this way we construct a sequence $\{x_n\}_n$ that *hopefully* converges to a solution.

Example 2. Use Newton's method to find a nonzero solution to $x = \dfrac{\pi}{2}\sin(x)$.

Solution We can see by inspection that there are no possible solutions for $|x| > \pi/2$, because the sine function has an amplitude of $\pi/2$. Let us reframe the question to solving

$$f(x) = x - \frac{\pi}{2}\sin(x) = 0.$$

We will choose $x_0 = 1$ for an initial guess. Since

$$f'(x) = 1 - \frac{\pi}{2}\cos(x),$$

we find that

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 1 - \frac{1-(\pi/2)\sin(1)}{1-(\pi/2)\cos(1)} \approx 3.12683$$

Iterating again we find

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 3.12683 - \frac{3.12683 - \frac{\pi}{2}\sin(3.12683)}{1-\frac{\pi}{2}\cos(3.12683)} \approx 1.91948$$

Continuing these iterations, we find $x_3 = 1.63106$, $x_4 = 1.5734$, $x_5 = 1.5708$, and $x_6 = 1.570796$. At this point we can feel confident we've converged to a solution, which happens to be $x = \pi/2 \approx 1.570796$. We have found a solution that is accurate up to at least the first 6 decimal places (actually more) of the solution $\pi/2$.

For this particular equation, there are actually three solutions. Depending on where the initial point is chosen, Newton's method may find a different root ($x_0 = -1$ yields $x = -\pi/2$), or may not converge at all. Despite being very effective, Newton's method can be a bit of trial and error. Because of this, in practice one would use some method of determining a reasonable first guess, and then use Newton's method for increased accuracy.

Example 3. Use Newton's method to solve $e^x = x + 2$ with at least 3 decimal places of accuracy, beginning with $x_0 = 1$.

Solution First we set $f(x) = e^x - x - 2$ and differentiate to find $f'(x) = e^x - 1$. Now we iterate to find

$$\begin{aligned}
x_1 &= 1 - \frac{e^1 - 1 - 2}{e^1 - 1} \approx 1.16395 \\
x_2 &= 1.16395 - \frac{e^{1.13395} - 1.16395 - 2}{e^{1.16395} - 1} \approx 1.14642 \\
x_3 &= 1.14642 - \frac{e^{1.14642} - 1.14642 - 2}{e^{1.14642} - 1} \approx 1.14619.
\end{aligned}$$

In order to be confident we have achieved 3 decimal points of accuracy, we continue iterating until the first 3 decimal places do not change. To really be confident we might consider iterating one step further.

4.15 Euler's Method

Euler's method is a numerical method for solving initial value problems. Euler's method is based on the insight that some differential equations (which are the ones we can solve using Euler's method) provide us with the slope of the function (at all points), while an initial value provides us with a point on the function. Using this information we can approximate the function with a tangent line at the point given by the initial value. As we have seen, the tangent line is only a good approximation over a small interval. Thus, after moving a small interval, we will want to construct a new tangent line.

What happens when we move to the new point? Although we know the value of the derivative of the function at the new point, we do not know the exact value of the function. Thus, the best we can do is construct an approximate tangent line, using the actual slope of the function and an approximation of the value of the function at point of tangency; the approximate value of the function is given by the previous tangent line approximation. Continuing in this fashion, we construct a first-order piecewise approximation to the solution of the differential equation. We can also think of our approximation as a discrete function, which is defined for these approximated points. To make this discrete function a continuous function, we interpolate between each pair of these points with a line.

Before we begin we should emphasize the limitations of Euler's method are practical, not theoretical. In theory, we should be able to make the distance traveled along each tangent line as short as we like, making the approximation to the actual solution of the initial value problem as accurate as we like. The only reason we cannot do this in practice is because there is a limit to the number of computations we can perform. From a human perspective, it becomes unreasonable to look at more than 10-100 intervals. A computer might do well with millions or billions of intervals (but then there are other practical issues to deal with, such as roundoff error, which do limit the number of intervals we can work with). Even still, we would need to look in a limit as the length of the intervals approaches 0, which means the number of intervals approaches ∞. Due to these practical limitations more advanced numerical methods such as the Runge-Kutta are used to solve differential equations in practice, rather than Euler's method. Nevertheless, it is useful to consider Euler's method because these other numerical methods are conceptually very similiar.

Suppose we wish to solve the initial value problem

$$\frac{dm}{dt} = f(t), \quad m(t_0) = m_0.$$

Our goal is to construct a function $\hat{m}(t)$ which approximates the actual solution. First we set $\hat{m}(t_0) = m(t_0) = m_0$. We want to define $\hat{m}(t)$ on intervals of length Δt, so it will next be defined at $\hat{m}(t_0 + \Delta t)$, then $\hat{m}(t_0 + 2\Delta t)$, and so on. If we know $\hat{m}(t)$ we find $\hat{m}(t + \Delta t)$ by moving a distance Δt along the line passing through $(t, \hat{m}(t))$ with slope $m'(t)$, which is our approximate tangent line. In moving a distance Δt along a line with slope $m'(t)$ we increase the value of our function by $m'(t)\Delta t$, so we find that

$$\hat{m}(t + \Delta t) = \hat{m}(t) + m'(t)\Delta t.$$

The result is a discrete-time function, but it can be made into a continuous-time function through interpolation (see figure 4.15).

This new continuous-time function is piecewise defined on intervals, and for any point in the interval $[t_0, t_0 + \Delta t]$ the function is given by

$$\hat{m}(t_0 + t) = \hat{m}(t_0) + m'(t_0)t$$

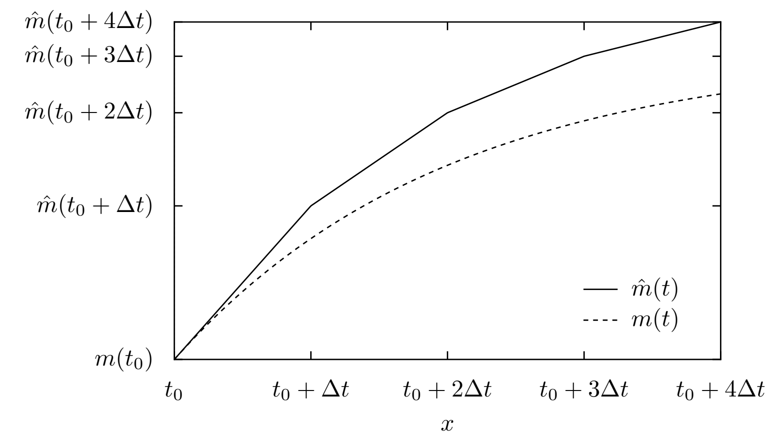


Figure 4.15: An illustration of Euler's method. The dotted line represents the actual (unknown) solution to the differential equation. The solid function is an illustration of the approximate solution obtained through Euler's method.

where $t \in [0, \Delta t]$ (this restriction upon t ensures that $t_0 + t \in [t_0, t_0 + \Delta t]$). This equation follows from the same observation as before, noting that the function value is increased by an amount $m'(t_0)t$ in moving a distance t along a line with slope $m'(t_0)$. In order to find the value of the function at a distant point we simply move along a number of approximate tangent lines, repeating this same process. In moving from t_0 to $t_0 + 4\Delta t$, for instance, we find that

$$\hat{m}(t + 4\Delta t) = \hat{m}(t_0) + m'(t_0)\Delta t + m'(t_0 + \Delta t)\Delta t + m'(t_0 + 2\Delta t)\Delta t + m'(t_0 + 3\Delta t)\Delta t.$$

In order to facilitate using Euler's method by hand it is often helpful to use a chart. This process is outlined in the following examples.

Example 1. Apply Euler's method to the differential equation

$$\frac{dV}{dt} = 2t$$

within initial condition $V(0) = 2$. Approximate the value of $V(1)$ using $\Delta t = 0.25$.

Solution We begin by setting $\hat{V}(0) = 2$. Next we construct the chart

t	$\hat{V}(t)$	$V'(t) = 2t$	$\hat{V}(t+\Delta t) = \hat{V}(t) + V'(t)\Delta t$
0	2	$2 \cdot 0 = 0$	$2 + 0 \cdot 0.25 = 2$
0.25	2	$2 \cdot 0.25 = 0.5$	$2 + 0.5 \cdot 0.25 = 2.125$
0.5	2.125	$2 \cdot 0.5 = 1$	$2.125 + 1 \cdot 0.25 = 2.375$
0.75	2.375	$2 \cdot 0.75 = 1.5$	$2.375 + 1.5 \cdot 0.25 = 2.75$
1	2.75		

So we find that $V(1) \approx 2.75$. Solving this initial value problem exactly we find $V(t) = t^2 + 2$. This yields $V(1) = 3$, so Euler's method provides a somewhat reasonable approximation, which would be greatly improved upon by decreasing the size of Δt. Note that the exact solution $t^2 + 2$ is concave up, which means that the value of the derivative is increasing along each interval. Thus, the value of the derivative at the beginning of an interval is smaller than the value of the derivative at all other points in the interval. The result is that Euler's method continually *underestimates* the actual value of the function.

Example 2. Apply Euler's method to the differential equation

$$\frac{dP}{dt} = e^{-t}$$

with the initial condition $P(0) = 0$. Approximate the value of $P(2)$.

Solution We begin by setting $\hat{P}(0) = 0$. We will use the time step $\Delta t = 0.5$. Next we construct the chart

t	$\hat{P}(t)$	$P'(t) = e^{-t}$	$\hat{P}(t+\Delta t) = \hat{P}(t) + P'(t)\Delta t$
0	0	$e^{-0} = 1$	$0 + 1 \cdot 0.5 = 0.5$
0.5	0.5	$e^{-0.5} \approx 0.60653$	$0.5 + 0.60653 \cdot 0.5 = 0.80326$
1	0.80326	$e^{-1} \approx 0.36788$	$0.80326 + 0.36788 \cdot 0.5 = 0.98721$
1.5	0.9872	$e^{-1.5} \approx 0.22313$	$0.9872 + .22313 \cdot 0.5 = 1.09877$
2	1.09877		

The actual solution to this problem is given by $P(t) = 1 - e^{-t}$, so we find the exact value $P(2) = 0.86466$. This is somewhat lower than our approximation. Note that the function $1 - e^{-t}$ is concave down, so the value of the derivative is decreasing, meaning the derivative at the beginning of an interval is always larger than at the end. For this reason, our Euler's method approximation continues to *overestimate* the actual solution.

The above examples may seem somewhat artificial, in that it was easy enough (or even possible) to find an exact solution to the differential equation. There are some situations where we can use an infinite series or other method in order to find an exact solution, but the truth is that very few differential equations can be solved exactly. Thus, numerical methods such as Euler's method are extremely important to solving practical problems. Consider the function

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

which corresponds to a Gaussian, or normal, distribution (such functions are ubiquitous in statistics, because they have a number of nice properties: the mean, median, and mode are the same, there is symmetry around the mean, and inflection points occurs one standard deviation away from the mean). This function *can* be integrated by using infinite series, but we currently don't have such a technique at our disposal. Instead, we can use Euler's method to approximate a solution.

Example 3. Apply Euler's method to the differential equation

$$\frac{df}{dt} = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

within initial condition $f(0) = 0.5$. Approximate the value of $f(1)$ using $\Delta t = 0.25$.

Solution We begin by setting $\hat{f}(0) = 0.5$. We will use the time step $\Delta t = 0.25$. Next we construct the chart

t	$\hat{f}(t)$	$f'(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\hat{f}(t+\Delta t) = \hat{f}(t) + f'(t)\Delta t$
0	0.5	0.399	$0.5 + 0.399 \cdot 0.25 = 0.59975$
0.25	0.59975	0.3866	$0.59975 + 0.3866 \cdot 0.25 = 0.6964$
0.5	0.6964	0.352	$0.6964 + 0.352 \cdot 0.25 = 0.7844$
0.75	0.7844	0.301137	$0.7844 + 0.301137 \cdot 0.25 = 0.859684$
1	0.859684		

Right now we can't verify this result, but with a little knowledge of statistics, we know that we should get a result of about 0.84. In this situation, Euler's method provides us with a very accurate estimate.

We can also used the idea of Euler's method when we do not have a differential equation, but simply data. This is particularly useful because anytime we measure physical data we only have values known over discrete-time intervals; ie. we do not have an exact function. Suppose we have measured the velocity of an object at various times, and want to estimate the position. We know position is given by the differential equation

$$\frac{dx}{dt} = v(t),$$

so we can perform Euler's method, but rather than evaluating $v(t)$ at the points of interest, we would just look up the value in a table.

Example 4. Suppose you have the following data for the velocity of an object. Find the distance it travels after 60 seconds.

Time (seconds)	Velocity (m/s)
0	0
10	12
20	18
30	15
40	12
50	8
60	0

Solution Here we can construct our table as normal, simply using the values we already know for velocity. Since we are looking at distance traveled, we will let $x(0)\hat{x}(0) = 0$. Our data comes in increments of 10, so we will let $\Delta t = 10$.

t	$\hat{x}(t)$	$v(t)$	$\hat{x}(t + \Delta t) = \hat{x}(t) + v(t)\Delta t$
0	0	0	$0 + 0 \cdot 10 = 0$
10	0	12	$0 + 12 \cdot 10 = 120$
20	120	18	$120 + 18 \cdot 10 = 300$
30	300	15	$300 + 15 \cdot 10 = 450$
40	450	12	$450 + 12 \cdot 10 = 570$
50	570	8	$570 + 8 \cdot 10 = 660$
60	660	0	

So we find that the object has traveled a total of 660 meters in 1 minute. Note that if we are only interested in the object's position at $t = 60$, we can simply use the shortcut

$$\hat{x}(60) = \hat{x}(0) + 10(v(10) + v(20) + v(30) + v(40) + v(50)).$$

Finally, if we had uneven intervals of time, we could use the formula

$$\hat{x}(t + \Delta t_1 + \Delta t_2) = \hat{x}(t) + v(t) \cdot \Delta t_1 + v(t + \Delta t_1) \cdot \Delta t_2.$$

Chapter 5

Integration

As useful as differentiation is, physical phenomena often present themselves in terms of derivatives of functions of interest, whereby we need to go backwards in order to find the original function. This is one key aspect to integration - antidifferentiation. At the same time, we will see that this process is also related to another useful idea - area.

Contents

5.1 Indefinite Integrals . . . 166
5.2 Area . . . 170
5.3 The Fundamental Theorem of Calculus . . . 177
5.4 Properties of Definite Integrals . . . 180
5.5 Average Values . . . 183
5.6 Integration by Substitution . . . 186
5.7 Integration By Parts . . . 192
5.8 Partial Fractions . . . 198
5.9 Trigonometric Integrals . . . 203
5.10 Trigonometric Substitution . . . 206
5.11 Numerical Integration . . . 209
5.12 Numerical Integration - Composite Rules . . . 219
5.13 Improper Integrals - Infinite Limits of Integration . . . 224
5.14 Improper Integrals - Infinite Integrands . . . 231
5.15 Area Between Curves . . . 234
5.16 Finding Volumes Using Slabs . . . 237
5.17 Finding Volumes Using Shells . . . 241
5.18 Lengths of Plane Curves . . . 244
5.19 Work . . . 248

5.1 Indefinite Integrals

Integration is a formalization of the process of antidifferentiation. When we differentiate a function, we lose some information about the function - although we know how it is changing we don't know where it started from. Thus, when we go through the reverse process of differentiation, we end up with a family of functions rather than just a single function. Because of this we need an initial value to find a unique solution to a differential equation. We define the indefinite integral in order to encompass the fact that we have a family of functions.

Definition 5.1.1 (Indefinite Integral)**.** We call the set of all antiderivatives of f the indefinite integral of f, denoted by

$$\int f(x)dx.$$

The symbol $\int$ is an integral sign. The function f is the called the integrand and x is the variable of integration. We say that dx is a differential of x.

If we know any antiderivative F to a given function f, then we simply find the indefinite integral as

$$\int f(x)dx = F(x) + c,$$

because the expression on the right-hand side represents all possible antiderivatives of $f(x)$, if we let c be an arbitrary constant.

When we find the indefinite integral of a function $f(x)$ we say that we integrate the integrand $f(x)$. Thus, the process of evaluating an integral is referred to as integration. Unfortunately, the notation we have introduced for the indefinite integral probably looks rather arcane. This notation comes from Leibniz, who developed calculus using the notion of an infinitesimal. In this notation the $\int$ is an elongated S, which represents summation. Intuitively, this notation says that integration is the summation of an infinite number of rectangles, each having height $f(x)$ (where x varies) and infinitesimal width dx. Such a summation gives exactly the area underneath the curve f. We will explore this connection in more detail later.

When it will not lead to confusion we will refer to the indefinite integral as simply the integral. The term indefinite integral is used to distinguish the process of indefinite and definite integration (which we will consider soon). Although both of these concepts are related to the area underneath a function, the indefinite integral is a function, whereas the definite integral is a constant, which is given by the area underneath a function over a set interval (defined by the limits of integration, which are not present in an indefinite integral).

Since the process of (indefinite) integration is an inverse to differentiation, we can derive many rules for integration using rules we already know for differentiation. For instance

$$\frac{d}{dx}x^{n+1} = (n+1)x^n$$

so we see that

$$\frac{d}{dx}\Big(\frac{x^{n+1}}{n+1}\Big) = x^n, \quad n \neq -1.$$

This leads us to the product rule for integrals.

Theorem 5.1.1 (Power Rule for Integrals)**.**

$$\int x^n dx = \frac{x^{n+1}}{n+1} + c$$

for every $n \neq -1$ ($n \in \mathbb{R} \setminus \{-1\}$).

Example 1. Evaluate the indefinite integral of x^2.

Solution In this case we use the product rule, to see that

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

Example 2. Evaluate the indefinite integral of $\sqrt{x}$.

Solution Once we rewrite $\sqrt{x} = x^{1/2}$ we see that

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

Example 3. Evaluate the indefinite integral of x^{-3}.

Solution We find that

$$\int x^{-3} = \frac{x^{-3+1}}{-3+1} + c = -\frac{1}{2}x^{-2} + c$$

Example 4. Evaluate $\int dt$.

Solution Rewriting the integrand we find

$$\int dt = \int t^0 dt = \frac{t^{0+1}}{0+1} + c = t + c.$$

Just like differentiation, integration is a linear operation. What this means is that integration satisfies the following two properties.

Theorem 5.1.2 (Linearity of Integration)**.** *Let f and g be Riemann integrable, $a \in \mathbb{R}$. It follows*

1. $\int \alpha f(x)dx = \alpha \int f(x)dx.$
2. $\int (f+g)(x)dx = \int f(x)dx + \int g(x)dx.$

Example 5. Evaluate the indefinite integral of $-2x^{-3} + 4x^{1/2}$.

Solution We can use the constant product rule and sum rules in conjunction to find

$$\int -2x^{-3} + 4x^{1/2} = -2\int x^{-3} + 4\int x^{1/2} = -2(-\frac{1}{2}x^{-2}) + 4(\frac{2}{3}x^{3/2}) + c = x^{-2} + \frac{8}{3}x^{3/2} + c$$

In the above analysis we do not write the result as $-2c + 4d$, because we can easily enough choose c and d so that we have any value for the arbitrary constant. Thus, it is cleaner to just replace $-2c + 4d$ with a single c, as both are capable of representing any arbitrary constant.

Corresponding to other differentiation rules we have learned, we can evaluate a number of other integrals, such as

$$\int e^x dx = e^x + c$$

$$\int \cos(x)dx = \sin(x) + c$$

$$\int \sin(x)dx = -\cos(x) + c$$

Note that the negative sign corresponds to the integral of the sine function, just as the negative sign corresponds to the derivative of cosine. This list is far from exhaustive, as we have results for other trigonometric functions as well as the inverse trigonometric functions. Rather than repeating all of these results here we simply note that the corresponding rule for integration can be read directly from the result for differentiation.

We'd also like to note

$$\int \frac{1}{x} dx = \ln|x| + c.$$

In the above formula we use the magnitude of x, $|x|$, because although $\frac{1}{x}$ is defined for negative x, $\ln(x)$ is not. In fact, the natural logarithm is usually *defined* in terms of the above integral, and then can be shown to be the inverse of the exponential function. However, since we did not have access to integration at the time we introduced the natural logarithm, we defined it as the inverse function, and now see that it satisfies the above integral. For $x > 0$

$$\frac{d}{dx} \ln|x| = \frac{d}{dx} \ln(x) = \frac{1}{x}.$$

If we consider $x < 0$ then

$$\frac{d}{dx} \ln|x| = \frac{d}{dx} \ln(-x) = \frac{1}{-x} \cdot (-1) = \frac{1}{x}.$$

We cannot evaluate the above derivative at $x = 0$, because $\ln(x)$ is undefined for $x = 0$.

Example 6. Evaluate $\displaystyle\int \Big(\frac{1}{x} + 2e^x + 3\cos(x)\Big) dx$

Solution We can evaluate the above integral using the sum and constant product rules in conjunction with the integrals we have just discussed.

$$\int \Big(\frac{1}{x} + 2e^x + 3\cos(x)\Big) dx = \int \frac{1}{x} dx + 2\int e^x dx + 3\int \cos(x) dx = \ln|x| + 2e^x + 3\sin(x) + c.$$

Example 7. Evaluate $\displaystyle\int 8\cos(8x) dx$.

Solution We don't know how to directly integrate something of this form. Nevertheless, we can find a solution with a bit of educated guesswork. If we recognize that in order find a cosine function with an argument of $8x$, we need to differentiate a sine function with an argument of $8x$, we consider

$$\frac{d}{dx} \sin(8x) = 8\cos(8x).$$

This turns out to be exactly our integrand, so we find that

$$\int 8\cos(8x) dx = \sin(8x) + c.$$

We can verify this result is correct by differentiating the right-hand side of the above equation.

Example 8. Evaluate $\displaystyle\int \sin(3x) dx$.

Solution Here the situation is similar, and we recognize that we will need to differentiate something of the form $-\cos(3x)$. However, when we do so we find

$$\frac{d}{dx}(-\cos(3x)) = 3\sin(3x).$$

This gives us an extra factor of 3 not present in the integrand. In order to account for this 3, we need to divide both sides of the equation by 3, so that

$$\frac{d}{dx}\frac{-\cos(3x)}{3} = \sin(3x).$$

Now we can see that

$$\int \sin(3x)dx = -\frac{\cos(3x)}{3} + c.$$

5.2 Area

In order to quantify the size of a 2-dimensional object, we use area. Since we measure area in square units, we can think of the area of an object as the number of such squares it fills up. Using this idea we can derive formulas for the area of a square, rectangle, triangle, etc. With a little bit of ingenuity we can also figure out the area of a circle. What about more complicated shapes, such as the area between a sine wave and the x-axis? (See figure 5.1.)

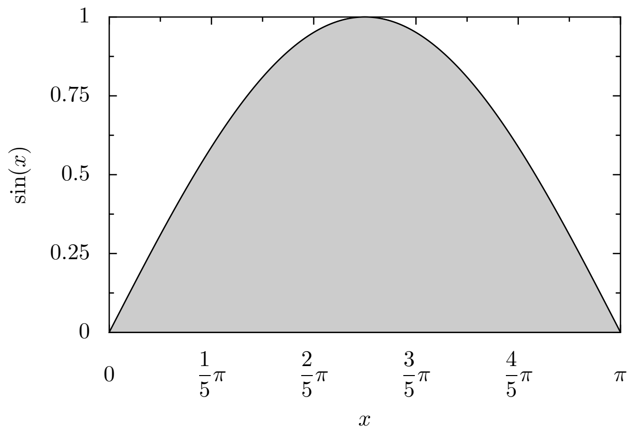


Figure 5.1: We need calculus to find the area of the shaded region.

A general technique for finding the area of a complicated shape is to break it up into smaller pieces which have known areas. Unfortunately, it is not possible to do this for most 2-dimensional objects, such as the sine wave in question. Instead, we can think about approximating the complicated shape with shapes that we can find the area of. The simplest and most practical 2-dimensional shape to use here is a rectangle, because it is much more flexible than using a square, but it is still very easy to calculate its area. If we only use a few rectangles, they will overlap our function of interest in many places, so it will be a rather crude approximation. However, if we use more rectangles, then we can get a better approximation (see figure 5.2).

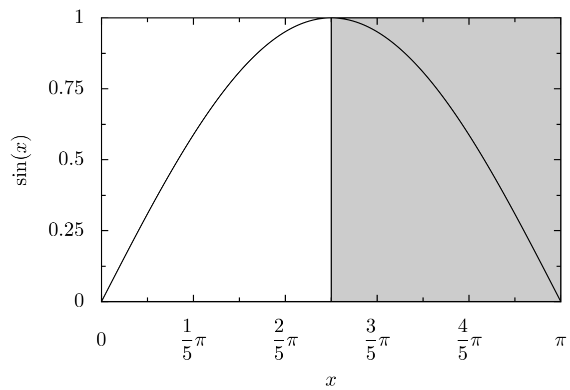
Just like with Euler's method, we can think of taking a limit as the width of these rectangles approaches 0, in order to find the exact area of the function or object of interest. The only real limitation we need to be concerned about is being able to compute the area of these rectangles. If we have billions of rectangles, then clearly we will need to use a computer to do the work, but still, computers have their own limitations.

Let us continue by trying to approximate the area between the above sine wave and the x-axis, over the interval $[0, \pi]$ (so we don't need to worry about what happens when we cross the x-axis). For simplicity (and scalability), let's divide the interval into equal subintervals, of length Δx. For convenience, we will use the value of the function at the left endpoint for the height of each rectangle. Finally, we must decide how many subintervals we want. We'll begin with just two, in order to illustrate the process. If we denote the approximated area function as $A(n)$, where n is the number of subintervals, we will find

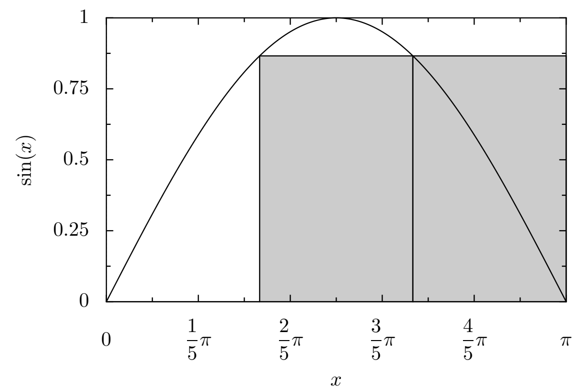
$$A(2) = \sin(0)\frac{\pi}{2} + \sin(\pi/2)\frac{\pi}{2} = \frac{\pi}{2} \approx 1.57.$$

For 3 subintervals we will find

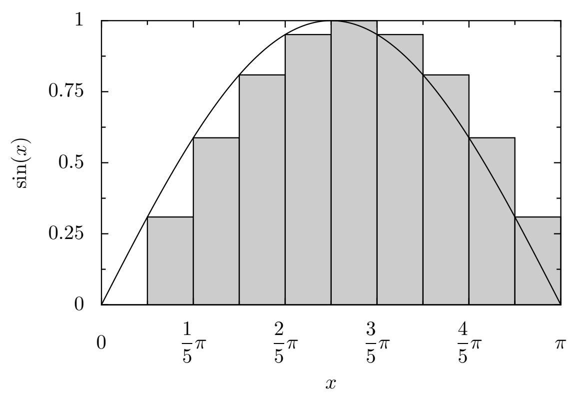
$$A(3) = \sin(0)\frac{\pi}{3} + \sin(\pi/3)\frac{\pi}{3} + \sin(2\pi/3)\frac{\pi}{3} = \frac{\pi(\sin(\pi/3) + \sin(2\pi/3))}{3} = \frac{\pi\sqrt{3}}{3} \approx 1.813.$$



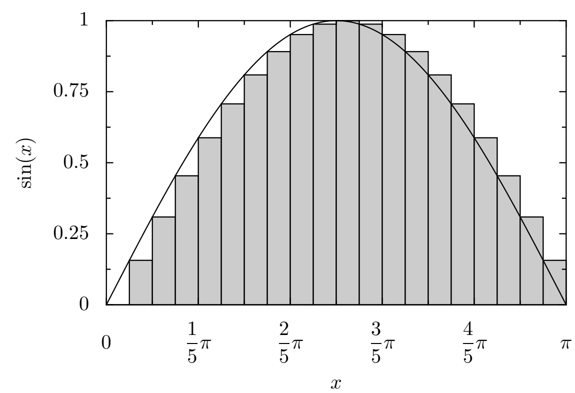
(a) 2 rectangles.



(b) 3 rectangles.



(c) 10 rectangles.



(d) 20 rectangles.

Figure 5.2: Approximating the area underneath a sine curve.

We could continue this way to find more accurate approximations, but more interesting than the specific case is a slightly more general problem. Let's think about finding the area underneath an arbitrary function $f(x)$ over the interval $[a, b]$. If we want n subintervals of equal length, then each will be of length

$$\Delta x = \frac{b-a}{n}.$$

The approximated area underneath the curve will be

$$A(n) = f(a)\Delta x + f(a+\Delta x)\Delta x + \ldots + f(a+(n-1)\Delta x)\Delta x + f(a+n\Delta x)\Delta x.$$

This is a very interesting result. When written in this way, the problem of approximating the area underneath a curve looks very familiar to something we've already done - approximating the solution to a differential equation. If we consider the differential equation

$$\frac{dF}{dx} = f(x),$$

where the value of $F(a)$ is known, we find that

$$F(b) = F(a) + f(a)\Delta(x) + f(a+\Delta x)\Delta x + \ldots + f(a+(n-1)\Delta x)\Delta x + f(a+n\Delta x)\Delta x,$$

and rearranging the terms we have the result

$$F(b) - F(a) = f(a)\Delta(x) + f(a+\Delta x)\Delta x + \ldots + f(a+(n-1)\Delta x)\Delta x + f(a+n\Delta x)\Delta x.$$

Above is only an approximation to $F(b)$, but we can it as accurate as we like, simply by using small enough intervals. We can see from the above expressions that the process of approximating a solution to a differential equation is the same as approximating the area underneath a curve (the connection is that the area of a rectangle with height $f(x)$ and base Δx is the same as the distance traveled in moving a distance Δx along a tangent line with slope $f(x)$).

Since we can make these approximations as accurate as we like, this means that the process of solving a differential equation is the same as finding the area underneath a curve (the curve we are finding the area underneath is the function representing the derivative in the differential equation). It follows that we can find the area underneath a curve through antidifferentiation, and conversely, solve differential equations by computing areas. In order to find the area underneath a function over an interval, we simply evaluate the antiderivative of the function at the endpoints and subtract the difference. Noting that

$$\int \sin(x)dx = -\cos(x) + c,$$

we find that the area underneath the sine curve from $[0, \pi]$ is

$$-\cos(\pi) - (-\cos(0)) = 1 + 1 = 2,$$

which was the value we began to approximate with rectangles.

We should note that when we find the area *underneath* a curve in this way, we are really finding a signed area. In places where the function is above the x-axis we have *positive* area, and in places where the function is below the x-axis, we have *negative* area. Using this reasoning, if we calculate the signed area under a sine wave over $[0, 2\pi]$, we get 0, because on $[0, \pi]$ the sine function is positive, and on $[\pi, 2\pi]$ the sine function is negative, and the negative portion is the mirror image of the positive portion of the function. To find the conventional area between a curve and the x-axis we would need to compute the signed area underneath the absolute value of the function of interest, so that all negative area becomes positive.

In general it is difficult if not impossible to find a closed-form expression for an antiderivative, so it is still worthwhile to investigate the problem of approximating areas in more detail. Before we proceed, it is worth pointing out that the reason for the notation of the indefinite integral should now be clear - in the process of integration, we want to sum up rectangles to approximate the area underneath a function, and look in the limit as the length of each rectangle approaches 0, so that they in a sense become *infinitesimal* in length.

We use the Greek letter Σ to represent summation in a succinct form. We can write a sample sum in the form

$$\sum_{i=0}^{3} x_i.$$

We call i the index of the this sum, and each x_i is a single term in the sum. There is nothing unique about the choice of i, and in general we can use whatever variable we like for the index of the sum. For every different value of the index, we have a corresponding term. We evaluate this sum by adding each of the terms together. Thus, this sum would be evaluated

$$\sum_{i=0}^{3} x_i = x_0 + x_1 + x_2 + x_3,$$

where each x_i is some value. If we let $x_0 = 1$, $x_1 = 4$, $x_2 = 3$, and $x_3 = 3$ then we would find

$$\sum_{i=0}^{3} x_i = x_0 + x_1 + x_2 + x_3 = 1 + 4 + 3 + 3 = 11.$$

It is noteworthy that for sums with a finite number of terms, it does not matter in which order the terms are added. If one is considering infinite sums however, the order *does* matter. It is also possible for the index to appear in the term of a sum as follows

$$\sum_{i=0}^{2} 1 + i^2 = (1 + 0) + (1 + 1) + (1 + 4) = 8.$$

In general we can have some arbitrary mix of both the index and other factors in each term. Finally, if we have a constant factor in every term of a sum, we can factor the constant from the sum, evaluate the sum, and multiply the result by that factor in the end. Thus

$$\sum_{i=0}^{2} 2 + 2i^2 = \sum_{i=0}^{2} 2(1 + i^2) = 2\sum_{i=0}^{2} 1 + i^2 = 2 \cdot 8 = 16.$$

Using this notation we can better define the problem of finding the area underneath a curve. To find the area underneath a function f over an interval $[a, b]$, we partition the interval $[a, b]$ into a number of subintervals. The endpoints of these subintervals are given by

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b,$$

where the $n+1$ points x_i define n subintervals. There is no reason in general that these subintervals need be of equal length, and in general we may not want them to be. Nevertheless, for simplicity let us begin with intervals of equal length. It then follows each subinterval will have a length of

$$\Delta x = \frac{b - a}{n}.$$

Finally, we need to define a height for the rectangle used to approximate the area of f over each subinterval. There is no reason we cannot choose any point within a given subinterval, but it is common, for the sake of simplicity, to choose such points uniformly. Most commonly one chooses to evaluate the function either at the left endpoint, right endpoint, or midpoint of an interval. For this illustration let us choose the left endpoint.

In summary, to approximate the area beneath a general function f over an interval $[a, b]$ using n subintervals of equal length, evaluating the function at left endpoints we have

$$I_l = \sum_{i=0}^{n-1} f(x_i)\Delta x.$$

Each term of this sum simply consists of the product of the height of the function at the left endpoint $f(x_i)$ and the width of the subinterval Δx. Looking at the way Δx is defined, we see that as the number of intervals n increases, the width of the intervals Δx correspondingly decreases. For I_l we are summing rectangles with a height equal to the value of the function at the left endpoint of each subinterval, so it is called a left-hand sum. If we look in the limit as $n \to \infty$, we find the area underneath the curve, assuming the limit exists.

Let's try and apply this machinery to finding the area underneath the curve $f(x) = x$, over $[0, b]$. We know that it should be

$$\frac{b^2}{2},$$

because the resulting figure is simply a triangle. We also know this because the antiderivative of x is $x^2/2$, and evaluating the antiderivative at the endpoints of the interval and subtracting yields $b^2/2 - 0 = b^2/2$. We should also be able to find the same result by approximating the area using rectangles, and looking in the limit as the length of the rectangles approaches 0. For n subintervals we find $\Delta x = b/n$, and the area is

$$I_l = \sum_{i=0}^{n-1} f(0 + i\Delta x)\Delta x = \sum_{i=0}^{n-1} \frac{ib}{n} \cdot \frac{b}{n} = \sum_{i=0}^{n-1} \frac{ib^2}{n^2} = \frac{b^2}{n^2} \sum_{i=0}^{n-1} i = \frac{b^2}{n^2} \frac{n(n-1)}{2} = \frac{b^2}{2}(1 - \frac{1}{n}),$$

where we use the fact that for the sum of the first n integers,

$$\sum_{i=0}^{n} i = \frac{n(n+1)}{2}.$$

In order to find the exact area underneath the curve, we look in the limit as $n \to \infty$ of the left-hand sum. Doing so

$$\lim_{n\to\infty} I_l = \lim_{n\to\infty} \frac{b^2}{2}(1 - \frac{1}{n}) = \frac{b^2}{2},$$

which is the familiar result.

Our left-hand sum is a special case of a Riemann sum, in which we made a number of simplifying assumptions. For a general Riemann sum the length of the subintervals need not all be the same. Rather than considering the limit as $n \to \infty$, we look in the limit as the widths of the subintervals approach 0. Furthermore, the point within each subinterval at which the function is evaluated can be anywhere. If all possible Riemann sums of this form converge to the same value for a given function (in the limit as the width of subintervals approaches 0), the function is said to be Riemann integrable. The Riemann integral is one type of definite integral, written

$$\int_a^b f(x)dx.$$

Here the definite integral represents the area underneath the function f on the interval $[a, b]$ (see figure 5.3). We call a the lower limit of integration and b the upper limit of integration.

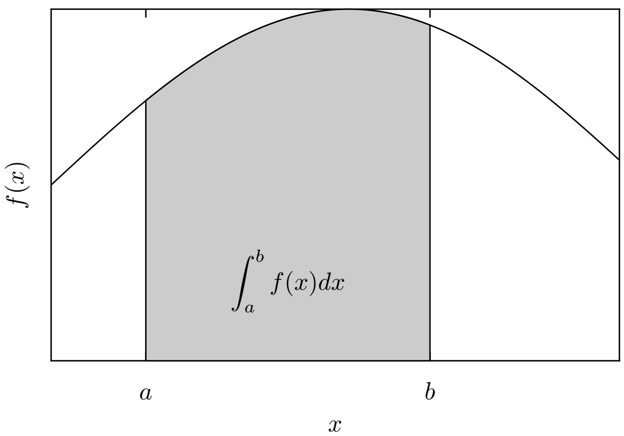


Figure 5.3: The definite integral is defined as the shaded area.

It is often easier to think about an equivalent notion of the integral - the Darboux integral. This notion is equivalent to the Riemann integral in that a function is Riemann integrable if and only if it is Darboux integrable, so both methods define the same definite integral for a function and a given interval. For the Darboux integral we once again consider partitions of arbitrary size, and look in the limit as the width of the partitions approaches 0. However, rather than considering any arbitrary point within each interval to evaluate the function, we are only interested in two sums - upper and lower sums. In each subinterval, upper sums overestimate the area with each rectangle, and the lower sums underestimate the area of with each rectangle. For an upper sum, the height of a rectangle on a given interval is given by the least upper bound of the function over than interval (the smallest number that is greater than or equal to the function on the interval); lower sums are similarly given by the greatest lower bound of the function over an interval. For a function to be Darboux integrable, we simply require that the upper and lower sums converge to the same limit (if these sums converge to the same limit, all other Riemann sums must as well, because they are sandwiched between these two extreme cases).

If the upper and lower Riemann sums don't converge, then we say a function is not Riemann (or Darboux) integrable, and in a sense we cannot define the area underneath the curve in this way. In truth, one must continue far into the study of mathematics and science to begin to see the shortcomings of the Riemann integral. For now we must suffice ourselves with the following result.

Theorem 5.2.1 (Integrability)**.** *Let $f : [a, b] \to \mathbb{R}$ be bounded, with a finite number of discontinuities. It follows that f is Riemann integrable.*

This theorem tells us immediately that an incredible number of functions are Riemann integrable, and it turns out that there are much more powerful theorems than above, but which are difficult to state without developing more mathematical theory. It's worth noting that here we require boundedness of the function - when we are faced with unbounded functions Riemann integrability breaks down. The reason for this is that for an unbounded function the notions of least upper bound and greatest lower bound are meaningless, so our definition of the integral becomes meaningless (later on we will be able to remedy this problem in some cases, using improper integrals). Given that so many functions are Riemann integrable, one might ask the question of what is not Riemann integrable?

One of the most simple examples is given as follows. Imagine a function that is 0 for rational numbers, and 1 for irrational numbers. Let us integrate over the interval $[0, 1]$. Since between every two rational numbers is an irrational, and between every two irrationals a rational, over every single interval we will an upper sum of 1, and a lower sum of 0, so the Riemann integral will not exist. In truth though, there are many many more irrationals than rationals in this interval, so we would expect (using intuition from higher-level mathematics) that the integral should be 1. In order to understand these issues better, one must focus on studying analysis, and eventually measure theory, in which a more powerful version of the integral can be constructed (the Lebesgue integral).

5.3 The Fundamental Theorem of Calculus

We recently observed the amazing link between antidifferentiation and the area underneath a curve - in order to find the area underneath a function f over some interval $[a, b]$, we simply look at the difference of the values of the antiderivative at the endpoints of the interval $F(b) - F(a)$. If we increase the value of b, then we begin to accumulate area underneath a larger portion of the function f. In a similar vein, we can think of defining a new function directly in terms of the area underneath a curve. Given some function $f(t)$ we can define (see figure 5.4)

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

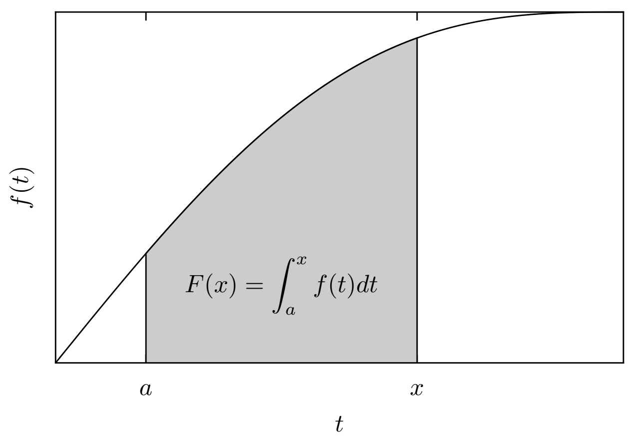


Figure 5.4: The area of the shaded region is the value of $F(x)$.

Here the variable t is called a dummy variable, or variable of integration. It could just as well be replaced with any other variable without changing the meaning of the integral - it just works as a placeholder for a varying input to the function f. The reason we use t rather than x is because x represents the end of the interval over which we are finding the area, so it doesn't make sense to use x to represent an arbitrary point inside the interval. The way we have defined this new function $F(x)$ is so that for any value b, $F(b)$ is the area under the function $f(x)$ over the interval $[a, b]$. It is noteworthy that we can generally choose any value for a. By choosing a different value for a, we simply begin accumulating area from a different point. We can just as well start from $c < a$, and define the function

$$G(x) = \int_c^x f(t)dt.$$

Then, the only difference between $F(x)$ and $G(x)$ will be a constant, and the constant will be exactly the area under $f(x)$ over the interval $[c, a]$, so

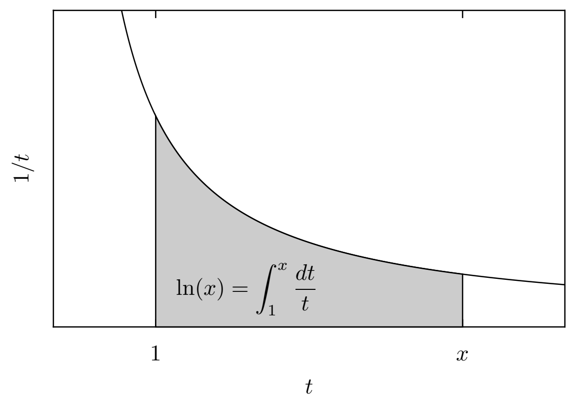
$$F(x) - G(x) = \int_c^a f(t)dt.$$

The fact that $F(x)$ and $G(x)$ differ by only a constant means they have the same derivative, namely $f(x)$. This makes sense because the rate at which we accumulate area should not depend on how much area we have already accumulated, which depends on where we start accumulating area from (the choice for a). This idea also relates to the constant of integration. These functions defined by accumulating area underneath the function $f(x)$ form a whole family of functions, all differing by a constant, where the constant depends on where we begin to accumulate area from. When we indefinitely integrate a function, we can think of the arbitrary constant relating to that we do not know where we began accumulating area from. This justifies the use the terms definite and indefinite. For a definite integral we know exactly over which interval we want to find the area. For an indefinite integral, we have not determined an interval over which to accumulate area.

This technique gives us an important means of defining new functions. In fact, the natural logarithm (see figure 5.5) is generally defined in this way, as

$$\ln(x) = \int_1^x \frac{dt}{t}.$$

We did not define the natural logarithm in this way because we didn't have access to this powerful idea, but this definition is completely consistent with all facts we previously established about the natural logarithm.



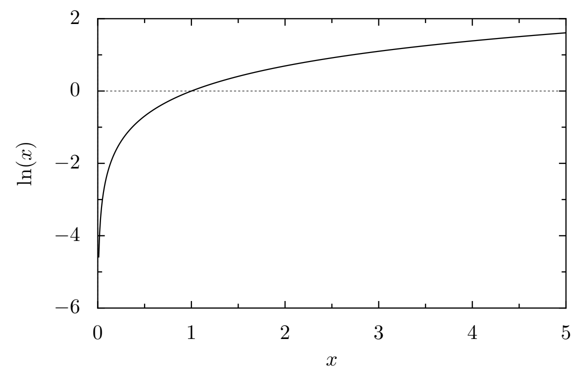


Figure 5.5: The natural logarithm is the area underneath the curve $1/x$ over the interval $[1, x]$.

More than just defining new functions using integrals, we already know an important piece of information about these functions - their derivatives. The derivative of one of these functions is simply given by the function we are finding the area under. Consider a function $f(x)$ with antiderivative $F(x)$. What happens when we differentiate? We need to think about the limit given by the definition of the derivative,

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

In the numerator, the difference is given by the area underneath the function $f(x)$ over an interval $[x, x+h]$. As $h \to 0$ the approximation of the area underneath the function by a single rectangle becomes as accurate as we would like. The height of this rectangle will just be $f(x)$, the value of the function at the left endpoint. It follows that

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

which verifies what we already knew to be true. This result is formalized by the Fundamental Theorem of Calculus.

Theorem 5.3.1 (Fundamental Theorem of Calculus: Part I)**.** *Let f be continuous on $[a, b]$. It follows the function $F(x) = \int_a^x f(t)dt$ is continuous on $[a.b]$ and differentiable on (a, b), with*

$$F'(x) = \frac{d}{dx} \int_a^x f(t)dt = f(x).$$

Example 1. Find $\frac{d}{dx}\int_a^x \cos(t)dt$.

Solution If we apply the fundamental theorem, we find

$$\frac{d}{dx}\int_a^x \cos(t)dt = \cos(x).$$

Example 2. Find $\frac{d}{dx}\int_a^{x^2} \cos(t)dt$

Solution Here we cannot directly apply the fundamental theorem, because the upper limit of integration is x^2, not x. Instead, we must use the chain rule, with

$$\int_a^u \cos(t)dt \quad \text{and} \quad u = x^2,$$

and we find

$$\frac{d}{dx}\int_a^{x^2} \cos(t)dt = \Big(\frac{d}{du}\int_a^u \cos(t)dt\Big)\cdot\frac{du}{dx} = \cos(u)\cdot\frac{du}{dx} = 2x\cos(x^2).$$

The second part of part of the fundamental theorem is something we have already discussed in detail - the fact that we can find the area underneath a curve using the antiderivative of the function.

Theorem 5.3.2 (Fundamental Theorem of Calculus: Part 2)**.** *Let f be continuous on $[a,b]$ and $F(x)$ an antiderivative of f on $[a,b]$. It follows that*

$$\int_a^b f(x)dx = F(x)\Big|_a^b = F(b) - F(a).$$

In the above theorem the notation

$$F(x)\Big|_a^b$$

is used to mean evaluate the function F at b and subtract by the function evaluated at a.

Example 3. Evaluate $\int_0^{10} x^2dx$.

Solution Using the fundamental theorem,

$$\int_0^{10} x^2dx = \frac{x^3}{3}\Big|_0^{10} = \frac{10^3}{3} - \frac{0}{3} = \frac{10^3}{3}.$$

Example 4. Evaluate $\int_1^4 (\frac{1}{x} + x)dx$.

Solution We find that

$$\int_1^4 (\frac{1}{x} + x)dx = (\ln|x| + \frac{x^2}{2})\Big|_1^4 = \ln(4) + 8 - (\ln(1) + \frac{1}{2}) = \ln(4) + 8 - \frac{1}{2} \approx 8.88.$$

Example 5. Let $n, m \in \mathbb{Z}$. Evaluate $\int_{2n\pi}^{2m\pi} \sin(x)dx$.

Solution We find that

$$\int_{2n\pi}^{2m\pi} \sin(x)dx = -\cos(x)\Big|_{2n\pi}^{2m\pi} = -\cos(2m\pi) + \cos(2n\pi) = -1 + 1 = 0.$$

5.4 Properties of Definite Integrals

We have seen that a definite integral represents the area underneath a function over a given interval. There are numerous useful properties of definite integrals worth studying, so that we can become adept at using and manipulating them. Suppose f and g are both Riemann integrable functions. In light of the fundamental theorem of calculus, and that

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

for F and antiderivative of f, we must have that

$$\int_b^a f(x)dx = F(a) - F(b) = -\int_a^b f(x)dx,$$

which in a sense states that by integrating in the opposite direction, or backwards, we pick up a negative sign. Let's also notice that if we integrate over an interval of 0 length, then we must find 0 area, so

$$\int_a^a f(x)dx = 0.$$

Just as we constant multiple and sum rules for indefinite integrals, they must hold for definite integrals, so

$$\int_a^b kf(x)dx = k\int_a^b f(x)dx,$$

and

$$\int_a^b (f(x) + g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

To illustrate the first property, suppose you are finding the area of a square. If you double the length of one of the sides, then the original square will fit twice inside the enlarged one, so the area of the enlarged square is twice that of the original square. The second property says that if we split an object into two pieces, the sum of the area should be the area of the sum. In a similar vein, it follows that

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx,$$

where here we are simply splitting the area in a different way (using a vertical rather than horizontal line). Finally, if $f(x) \geq g(x)$ on (a,b), then it must follow that there is more area under $f(x)$, so

$$\int_a^b f(x)dx \geq \int_a^b g(x)dx.$$

A special case of this property is looking at the rectangles formed by the minimum and maximum values of f (or least upper bounds and greatest lower bounds). The rectangle formed by the max of f must have a greater area than under f, which must have a greater area than under the minimum value of f. We summarize all of these results in the following theorem.

Theorem 5.4.1 (Properties of the Definite Integral)**.** *Let f and g be Riemann integrable functions, $k \in \mathbb{R}$. It follows that*

1. $\displaystyle\int_b^a f(x)dx = -\int_a^b f(x)dx$

2. $\int_a^a f(x)dx = 0$

3. $\int_a^b kf(x)dx = k\int_a^b f(x)dx$

4. $\int_a^b (f(x)+g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx$

5. $\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx$

6. $\int_a^b f(x)dx \geq \int_a^b g(x)dx, \quad \text{if } f(x) \geq g(x), x \in (a,b)$

Using the above rules, we can now integrate a wider variety of functions. For instance, we are now capable of dealing with piecewise-defined functions.

Example 1. Evaluate $\int_0^4 f(x)dx$ for

$$f(x) = \begin{cases} x & 0 \leq x < 2 \\ 4 & 2 \leq x < 3 \\ 3x^2 & 3 \leq x \end{cases}$$

Solution To solve this problem we need to split the integral into multiple pieces, according to where the function is defined differently.

$$\int_0^4 f(x)dx = \int_0^2 xdx + \int_2^3 4dx + \int_3^4 3x^2dx = \frac{x^2}{2}\Big|_0^2 + 4x\Big|_2^3 + x^3\Big|_3^4 = (2-0)+(12-8)+(64-9) = 61.$$

Example 2. Evaluate $\int_{-5}^5 |x|dx$.

Solution We do not know how to directly integrate $|x|$, but we can write it in terms of two functions we do now how to integrate. Using the fact that we can split the integral over the limits of integration, we find that

$$\begin{aligned} \int_{-5}^5 |x|dx &= \int_{-5}^0 |x|dx + \int_0^5 |x|dx = \int_{-5}^0 (-x)dx + \int_0^5 xdx \\ &= -\frac{x^2}{2}\Big|_{-5}^0 + \frac{x^2}{2}\Big|_0^5 = -\frac{0^2}{2} - \frac{-5^2}{2} + \frac{5^2}{2} - \frac{0^2}{2} = 2\cdot\frac{5^2}{2} = 5^2 = 25. \end{aligned}$$

This is really part of a more general phenomenon. When we have a function which is symmetric about the y-axis, and we integrate over an interval $[-a,a]$, we find a value that is twice the integral from $[0,a]$. For this reason we can see that

$$\int_{-\pi/2}^{\pi/2} \cos(x)dx = 2\int_0^{\pi/2} \cos(x)dx = 2\sin(x)\Big|_0^{\pi/2} = 2(\sin(\pi/2) - \sin(0)) = 2.$$

We have previously said that the definite integral finds the 'signed' area underneath a function. We know that functions like the sine and cosine function have certain symmetries about the x-axis that cause their definite integrals to be 0 over certain limits. If we want to find the amount of conventional area between such a function and the x-axis, we need to integrate the magnitude of the function.

Example 3. Find the conventional area between $\sin(x)$ and the x-axis over $[0, 2\pi]$.

Solution In order to solve this problem we integrate $|\sin(x)|$. However, we do not know how to integrate such a function, so we need to break it up into intervals where we know sine is positive and negative, so we can remove the absolute value sign.

$$\begin{aligned}\int_0^{2\pi} |\sin(x)| dx &= \int_0^{\pi} \sin(x) dx + \int_{\pi}^{2\pi} (-\sin(x)) dx = -\cos(x)\Big|_0^{\pi} + \cos(x)\Big|_{\pi}^{2\pi} \\ &= -\cos(\pi) + \cos(0) + \cos(2\pi) - \cos(\pi) = 1 + 1 + 1 + 1 = 4.\end{aligned}$$

5.5 Average Values

When we want to find the average or mean of a finite collection of objects we simply sum the values together, and divide by the number of objects in the collection. Thus, the average of the scores 45, 64, and 90 on a test is

$$\frac{45+64+90}{3} = \frac{199}{3} \approx 66.33.$$

What about a continuous function? Is there an intuitive notion of the average value of a continuous function? Let's think about a physical example. Suppose we know the instantaneous velocity of a person hiking in the mountains over some interval of time. What should the average velocity of the hiker be? It seems intuitive to define the average velocity as the total distance traveled divided by the time it took to travel this distance. In this way, another person who was traveling at the average velocity of the first hiker for the same amount of time would travel the same distance. We can find the distance traveled by calculating the integral of the hiker's velocity over the interval of time the hiker traveled. The time traveled will simply be the difference of the endpoints of the interval of time.

Mathematically, we say that the average value of a function f over an interval $[a, b]$ is given by

$$\frac{1}{b-a}\int_a^b f(x)dx$$

which is in complete analogy with the above thought experiment. The average value of a function over a given interval also has a useful graphical interpretation. If we integrate the average value over the interval $[a, b]$, we will find the same area as under the function f. Thus, the average value gives the height of a rectangular box with width $[a, b]$ that has the same area as the function f.

Example 1. Find the average value of x^3 over $[0, 3]$.

Solution Using the above formula we know the average value is

$$\frac{1}{3-0}\int_0^3 x^3dx = \frac{1}{3}\frac{x^3}{3}\Big|_0^3 = \frac{3^3}{9} - \frac{0}{9} = 3.$$

Example 2. Find the average value of $\sin(x)$ over $[0, 2\pi]$.

Solution We already know that

$$\int_0^{2\pi} \sin(x)dx = -\cos(2\pi) + \cos(0) = 0,$$

so the average value is 0.

The above result is a problem in the realm of electrical circuits, where AC currents and voltages are represented by sinusoidal functions. Over the appropriate intervals of 2π these functions have an average value of 0. Even if the average value over an interval is not 0, all segments of length 2π are canceled out when we find the average value. If we were to think of the average power moving through one of these systems, it would surely not be 0, simply because the average voltage and current are 0. Instead, we would use the root mean square or RMS value, in which we find the average value of the square of a function, and take the square root of the output.

$$\text{RMS value of } f(x) = \sqrt{\frac{1}{b-a}\int_a^b f^2(x)dx}.$$

Unfortunately, we do not currently have the tools to evaluate the required integrals to find the RMS values of sinusoidal functions, given by

$$\sqrt{\frac{1}{t_2-t_1}\int_{t_1}^{t_2} I^2\sin^2(t)dt}$$

where I represents the amplitude of the signal (and t_1 and t_2 occur at the end of a period of the waveform). Nevertheless, we can find the RMS value for other functions (although we cannot show it, it turns out that the value of the above expression is $I/\sqrt{2}$).

Example 3. Find the RMS value of $f(x) = x$ over $[-4, 4]$

Solution The first step in the process is to find the average value of $f^2(x) = x^2$.

$$\frac{1}{4-(-4)}\int_{-4}^{4} x^2dx = \frac{1}{8}\frac{x^3}{3}\Big|_{-4}^{4} = \frac{1}{24}(64-(-64)) = \frac{16}{3}$$

Now we take the square root of this value, to find

$$\text{RMS of } f(x) = \sqrt{\frac{16}{3}} = \frac{4}{\sqrt{3}}.$$

Example 4. Suppose you are the coach of a runner in a one-on-one race, who is facing another runner notorious for his running techniques. Your runner's best technique is to run at a constant pace throughout the whole race. If you know the opponent's running pattern is given by

$$v(t) = \begin{cases} 5 & 0 \le t < 1 \\ 4+t & 1 \le t < 3 \\ 15-t^2 & 3 \le t \le \sqrt{15} \end{cases}$$

based on previous races, at what constant pace should you advise your runner to run (velocity is in miles per hour)?

Solution In order to win the race, the constant pace will have to be slightly higher than the average velocity of the erratic runner. In order to find the average of this piecewise defined function we do not find the average of each piece. Instead, we find the total distance traveled, by integrating, and then divide the final result by the time it took to run.

$$\begin{aligned}\int_0^{\sqrt{15}} v(t)dt &= \int_0^1 5dt + \int_1^3(4+t)dt + \int_3^{\sqrt{15}}(15-t^2)dt = 5t\Big|_0^1 + (4t+\frac{t^2}{2})\Big|_1^3 + (15t-\frac{t^3}{3})\Big|_3^{\sqrt{15}} \\ &= (5-0) + (12+\frac{9}{2}-4-\frac{1}{2}) + (15\sqrt{15}-\frac{\sqrt{15}^3}{3}-45+\frac{27}{3}) \approx 19.73\end{aligned}$$

Now we divide by $\sqrt{15}$ to find the average speed, which is

$$\frac{19.73}{\sqrt{15}} \approx 5.094$$

Now in order to be sure our runner wins, we would like a constant pace slightly higher than this, so we might advise 5.25 miles per hour.

If we think about the erratic runner in the above race, it seems that at some point he must have been running at his average velocity. If he was always running lower than the average speed, then the average would be too high, and similarly if he was always running faster the average would be to slow. This means that at some points he was running faster and at some points running slower. Because in the physical world we cannot instantly accelerate, at least at some points he must have been running the average speed (even if only for a very short fraction of time). This result holds more generally, as long as we have continuous functions (so in this way we can think that a physical velocity must be continuous).

Theorem 5.5.1. *If f is continuous on $[a, b]$, then at some point c in $[a, b]$*

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

This theorem tells us (not too surprisingly) that if the integral of a continuous function over a finite interval is 0, then the value of the function must be 0 at some point in the interval. Compare this with the behavior of the sine function.

5.6 Integration by Substitution

Just as the chain rule is indispensable in differentiation, it will be extremely useful in integration. If we have two functions $f(u)$ and $u(x)$ which are both differentiable, the chain rule tells us that

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

From here it follows that

$$\int f'(u(x))u'(x)dx = f(u(x)) + c.$$

This rule tells us that if we can identify the integrand as the derivative of a composition of functions, we can evaluate the integral. However, working with the above formula directly is rather cumbersome. Nevertheless, with the aid of differentials, we can rewrite the above equation in a way that is much easier to work with.

Consider the differential du. One way of representing this infinitesimal change in u is by the difference $du = u(x+dx) - u(x)$. However, if we know u is differentiable, we can also represent the differential du in another way. Assuming that $u'(x)$ exists, moving an infinitesimal distance along the curve u is identical to moving an infinitesimal distance along the line tangent to u at the point x. This follows from the fact that a differentiable function behaves essentially like its tangent line in a small enough neighborhood around the point of tangency (and we are considering an interval of infinitesimal length). It follows that for a differentiable function $u(x)$, we can represent the differential du as

$$du = \frac{du}{dx} \cdot dx = u'(x)dx.$$

This allows us to rewrite

$$\int f'(u(x))u'(x)dx = \int f'(u)du.$$

Here we are changing the variable of integration from x to a new variable $u(x)$. This provides us with an integral written in terms of u, which we simply evaluate as normal, and replace $u = u(x)$ into the result, to find the antiderivative as a function of x. The key to integration by substitution is proper choice of u, in order to transform the integrand from an unfamiliar form to a familiar form.

Example 1. Evaluate $\int 2xe^{x^2}dx$.

Solution Although we don't know how to integrate $2xe^{x^2}$, we do know how to integrate e^u. Thus, our goal is to use substitution to change the integrand to the form of e^u. We can do this by defining

$$u = x^2.$$

Differentiating u we find

$$\frac{du}{dx} = 2x$$

which yields

$$du = 2xdx.$$

Now we replace any factors of x^2 with u in the above integral, and replace $2xdx$ with du. This yields

$$\int 2xe^{x^2}dx = \int e^u du.$$

We evaluate this as

$$\int e^u du = e^u + c = e^{x^2} + c.$$

Note that the last step is to replace u with x^2, to write the solution in terms of the original variable. Differentiating we find

$$\frac{d}{dx}(e^{x^2} + c) = e^{x^2} \cdot 2x,$$

which verifies that we have successfully evaluated the above integral.

It is important to note that substitution will not always work. If we were faced with an integral of the form

$$\int 2x^2 e^{x^2} dx,$$

substitution would not have simplified the problem (there is no useful choice for u). In fact, it is impossible to evaluate the above integral in terms of elementary functions. Above is an example of a function that is a perfect candidate for numerical integration, using a computer.

Example 2. Suppose fish size begins at $L(0) = 0$ (measured from fertilization) and follows the differential equation

$$\frac{dL}{dt} = 6.48e^{-0.09t} \text{ cm}$$

Find the length of the fish as a function of time.

Solution We can solve this pure-time differential equation using integration, but we will also have to apply the method of substitution. We would like to choose u such that our integrand is of the form e^u, which we know how to integrate. Define

$$u = -0.09t,$$

and note

$$du = -0.09dt \quad \text{or} \quad \frac{du}{-0.09} = dt.$$

Using substitution

$$\int 6.48e^{-0.09t}dt = \frac{6.48}{-0.09}\int e^u du = -72e^u + c = -72e^{-0.09t} + c.$$

Now we can use the initial condition to solve for an exact value for c.

$$L(0) = -72e^{-0.09\cdot 0} + c = -72 + c = 0,$$

which implies $c = 72$. Thus, our solution to the initial value problem is

$$L(t) = 72(1 - e^{-0.09t}) \text{ cm}.$$

As the fish matures it never stops growing, but the rate at which it grows continues to decrease. The length of the fish approaches but never reaches 72 cm.

Example 3. Evaluate $\displaystyle\int \frac{4x}{4x^2+2}dx$.

Solution Faced with a fraction of this form, it is our goal to use substitution to reduce it to the form $\frac{1}{u}$ which we can integrate to $\ln(|u|)$. If we include the second power of x in our definition of u, when we differentiate u with respect to x, we will be left with the first power of x, which will replace the x in the numerator nicely. Thus, we choose

$$u = 4x^2 + 2.$$

This yields

$$du = 8xdx.$$

We would like to identify factors of $4x^2 + 2$ and $8xdx$ in the integrand, but we notice that $8xdx$ does not appear anywhere. Thus, we cannot directly substitute in the term du. However, we do have a factor of $4xdx$, so if we note that

$$\frac{du}{2} = 4xdx,$$

we can substitute in $du/2$. Substituting into the original equation we find

$$\int \frac{4x}{4x^2+2}dx = \frac{1}{2}\int \frac{1}{u}du = \frac{1}{2}\ln|u| + c = \frac{1}{2}\ln|4x^2+2| + c = \frac{1}{2}\ln(4x^2+2) + c.$$

Example 4. Evaluate $\int \cos(t)\sin(t)dt$.

Solution In integrating this function, both $u = \sin(t)$ and $u = \cos(t)$ are reasonable choices for substitution. We'll consider both of them, to see how the process varies depending on the choice of u. Let us begin with

$$u = \cos(t),$$

which implies that

$$du = -\sin(t)dt \quad \text{or} \quad -du = \sin(t)dt.$$

We rewrite the above equation because there is no negative sign in the integrand which we can substitute in for. Substituting into the integral

$$\int \cos(t)\sin(t)dt = -\int udu = -\frac{u^2}{2} + c = -\frac{\cos(t)^2}{2} + c.$$

Alternatively, if we designate

$$u = \sin(t),$$

we find

$$du = \cos(t)dt.$$

Substituting into the integral

$$\int \cos(t)\sin(t)dt = \int udu = \frac{u^2}{2} + c = \frac{\sin(t)^2}{2} + c.$$

It is curious that we received two different results by using different representatives for substitution. In order for substitution to be a valid method, surely valid substitutions should lead to the same result. In fact, they do in this case, if we recognize that

$$\cos(t)^2 + \sin(t)^2 = 1 \quad \text{or} \quad \cos(t)^2 = 1 - \sin(t)^2.$$

Substituting into our first result

$$-\frac{\cos(t)^2}{2} + c = -\frac{1-\sin(t)^2}{2} + c = \frac{\sin(t)^2}{2} - \frac{1}{2} + c = \frac{\sin(t)^2}{2} + d.$$

Where d is an arbitrary constant. Absorbing the $-\frac{1}{2}$ into the arbitrary constant, we see that these two results are in fact consistent.

There are a few general strategies to keep in mind when solving a integral using substitution. Since we do not know how to directly integrate anything of the form $\sin(u)$, $\cos(u)$, e^u, or u^n where u is some function of x, in such an example we will almost surely need to substitute for u. The only additional thing we need to check is that when we differentiate u we find something that appears in the integrand as well (note that we may need to re-arrange constants in order to make this happen). Finally, we may need to break our integral into multiple integrals, separately performing substitution with each, in order to evaluate it.

Example 5. Evaluate $\int (3y^2 \cdot \sqrt{1+y^3} + \frac{2y}{\sqrt[3]{y^2+1}})dy$.

Solution In this case, we would like to substitute $u = 1 + y^3$ to evaluate the first term, but this will create quite a mess of the second term. Similarly, we'd like to substitute $u = 1 + y^2$ for the second term, which will make the first much more complicated. In order to avoid this problem, we simply break the integral into two, which we can do using the additivity of integrals. Thus,

$$\int (3y^2 \cdot \sqrt{1+y^3} + \frac{2y}{\sqrt[3]{y^2+1}})dy = \int 3y^2 \cdot \sqrt{1+y^3}dy + \int \frac{2y}{\sqrt[3]{y^2+1}}dy.$$

Now we can solve each of the integrals separately, using different substitutions. First, let

$$u = 1 + y^3,$$

so

$$du = 3y^2 dy.$$

This yields

$$\int 3y^2 \cdot \sqrt{1+y^3}dy = \int \sqrt{u}du = \frac{2}{3}u^{3/2} + c = \frac{2}{3}(1+y^3)^{3/2} + c.$$

For the second integral, let

$$u = y^2 + 1,$$

so

$$du = 2y,$$

which yields

$$\int \frac{2y}{\sqrt[3]{y^2+1}}dy = \int u^{-1/3}du = \frac{3}{2}u^{2/3} + c = \frac{3}{2}(1+y^2)^{2/3} + c,$$

with a final result of

$$\int (3y^2 \cdot \sqrt{1+y^3} + \frac{2y}{\sqrt[3]{y^2+1}})dy = \frac{2}{3}(1+y^3)^{3/2} + \frac{3}{2}(1+y^2)^{2/3} + c.$$

It's worth noting that we could have used a different substitution to integrate the second integral. Let

$$u = \sqrt[3]{y^2+1},$$

so

$$du = \frac{1}{3}(y^2+1)^{-2/3}2ydy,$$

which looks rather complicated. We cannot find a term of $(y^2+1)^{-2/3}$ in the integrand, so we will need to solve for part of this expression in terms of u. Rewriting the above expression as

$$3(y^2+1)^{1/3}du = \frac{2ydy}{\sqrt[3]{y^2+1}},$$

gives us the entire integrand on the right hand side, but we need to remove all terms including y on the left. Using the fact that

$$u = \sqrt[3]{y^2+1} = (y^2+1)^{1/3},$$

we find that

$$3udu = \frac{2ydy}{\sqrt[3]{y^2+1}}.$$

Now we can finally substitute to find

$$\int \frac{2y}{\sqrt[3]{y^2+1}}dy = \int 3udu = \frac{3}{2}u^2 + c = \frac{3}{2}(1+y^2)^{2/3} + c.$$

which is the same result as before.

The above example illustrates that it is possible to evaluate integrals using multiple different substitutions. Even though the choice of u and resulting integral of u may be very different, the final result must be the same. Furthermore, it illustrates how messy things can get with a poor choice for the substitution. A rule of thumb for choosing a function u is to choose the simplest function that will transform the integral into a solvable form. Note that our second choice for u above was more complicated, so it resulted in a simpler integral in u, but it required a lot more work to get there. Generally, it is worth dealing with a slightly more complicated (but still simple) integral of u, to avoid the headache of figuring out how to make the substitution.

In addition to evaluating indefinite integrals through substitution, we can evaluate definite integrals in the same way. One way of doing so is simply evaluating the indefinite integral to find an antiderivative, and applying the Fundamental Theorem of Calculus to evaluate the antiderivative at the appropriate limits of integration. Equivalently, we can modify the actual limits of integration, and simply evaluate the integral with respect to u over the new limits. We find that

$$\int_a^b f'(u(x))u'(x)dx = f(u(x))\Big|_{x=a}^{x=b} = f(u)\Big|_{u=u(a)}^{u=u(b)} = \int_{u(a)}^{u(b)} f'(u)du.$$

Since both of the methods work equally well, it is a matter of preference which to use. It is usually easier to substitute for the limits of integration, because it involves fewer steps to reach the final answer.

Example 6. Evaluate $\displaystyle\int_{-1}^{1} 3x^2\sqrt{x^3+1}dx$.

Solution We will demonstrate how both of the above methods work. First let us set

$$u = 1 + x^3,$$

so that

$$du = 3x^2dx,$$

and

$$\int 3x^2\sqrt{x^3+1}dx = \int \sqrt{u}du = \frac{2}{3}u^{3/2} + c = \frac{2}{3}(x^3+1)^{3/2} + c,$$

so it follows

$$\int_{-1}^{1} 3x^2\sqrt{x^3+1}dx = \frac{2}{3}(x^3+1)^{3/2}\Big|_{-1}^{1} = \frac{2}{3}((1^3+1)^{3/2} - (-1+1)^{3/2}) = \frac{2}{3}2^{3/2} = \frac{4\sqrt{2}}{3}.$$

Alternatively, we can perform the substitution, and note that $u(-1) = 0$ and $u(1) = 2$, so that

$$\int_{-1}^{1} 3x^2\sqrt{x^3+1}dx = \int_0^2 \sqrt{u}du = \frac{2}{3}u^{3/2}\Big|_0^2 = \frac{2}{3}2^{3/2} = \frac{4\sqrt{2}}{3}.$$

Example 7. Evaluate $\displaystyle\int_0^1 (1+2t)^2 dt.$

Solution Using substitution we let

$$u = 1 + 2t,$$

so

$$\frac{1}{2}du = dt.$$

Noting that $u(0) = 1$ and $u(1) = 3$ we find

$$\int_0^1 (1+2t)^2 dt = \int_1^3 u^2 du = \frac{u^3}{3}|_1^3 = \frac{27}{3} - \frac{1}{3} = \frac{26}{3}.$$

Example 8. Evaluate $\displaystyle\int_0^{\pi/6} \cos^{-3}(2x)\sin(2x)dx.$

Solution In this situation both $\cos(2x)$ and $\sin(2x)$ seem like reasonable choices for substitution. However, if we were to choose $u = \sin(2x)$, we would only end up with a single factor of $\cos(2x)$ when we differentiate, yet we need 3 factors of cosine in order to transform it into a workable form. Thus, we must choose to let

$$u = \cos(2x),$$

so that

$$du = -2\sin(2x)dx \quad \text{or} \quad \frac{du}{-2} = \sin(2x)dx.$$

Noting that $u(\pi/6) = \cos(\pi/3) = 1/2$ and $u(0) = \cos(0) = 1$ we find that

$$\int_0^{\pi/6} \cos^{-3}(2x)\sin(2x)dx = -\frac{1}{2}\int_1^{1/2} u^{-3}du = -\frac{1}{2}\cdot\frac{u^{-2}}{-2}\Big|_1^{1/2} = \frac{1}{4(1/2)^2} - \frac{1}{4(1)^2} = \frac{3}{4}.$$

Alternatively, we could write

$$\cos^{-3}(2x)\sin(2x) = \tan(2x)\sec^2(2x)$$

and use the substitution $u = \tan(2x)$ to solve the problem.

Example 9. Evaluate $\displaystyle\int_2^5 6.48e^{-0.09t}dt.$

Solution First we write

$$u = -0.09t,$$

so

$$du = -.09dt \quad \text{or} \quad -\frac{1}{0.09}du = dt.$$

We find $u(2) = -0.09\cdot 2 = -0.18$ and $u(5) = -0.09\cdot 5 = -0.45$. Substituting in

$$\int_2^5 6.48e^{-0.09t}dt = -\frac{6.48}{0.09}\int_{-0.18}^{-0.45} e^u du = -72e^u\Big|_{-0.18}^{-0.45} \approx -45.9 + 60.1 = 14.2.$$

5.7 Integration By Parts

We have already seen that like differentiation, integration is a linear operator, meaning that it treats linear combinations by working on each of the component functions (which means that integration *splits* over constant products and sums). Moving one step further we were able to use the chain rule in order to implement the powerful method of substitution. The final technique we had in differentiation was the product rule. Integration by parts is the integral analog to the product rule for differentiation.

Integration by parts is a method we can use to evaluate integrals of the form

$$\int f(x)g(x)dx,$$

where the term $f(x)$ is easy to differentiate, and $g(x)$ is easy to integrate, because integration by parts allows us to evaluate such an integral in terms of the derivative of $f(x)$ and the integral of $g(x)$. An example of an integral we can evaluate using integration by parts is

$$\int xe^x dx,$$

because x is easily differentiable, and e^x is easily integrable.

The product rule for derivatives states that

$$(u \cdot v)'(x) = u(x)v'(x) + v(x)u'(x).$$

If we integrate both sides of the equation with respect to x we find

$$\int (u \cdot v)'(x)dx = \int u(x)v'(x)dx + \int v(x)u'(x)dx,$$

which implies

$$u(x)v(x) = \int u(x)v'(x)dx + \int v(x)u'(x)dx.$$

Finally, we rewrite

$$\int u(x)v'(x)dx = u(x)v(x) - \int v(x)u'(x)dx.$$

On the left-hand side we have an integral involving the functions u and v', which is written in terms of an antiderivative v of v', and the derivative u' of u on the right. Thus, if we know how to differentiate u, and how to integrate v', then we can evaluate the above integral. The above formula is often rewritten in terms of differentials for brevity,

$$\int udv = uv - \int vdu,$$

where u and v are functions of x. The benefit of this form is that it is more compact and easier to remember than the previous.

Example 1. Evaluate $\int xe^x dx$.

Solution The first step to integrate is to choose which term will be used for $u(x)$ and which will be $v'(x)$. We make the choices

$$u = x, \quad \text{and} \quad dv = e^x dx,$$

as u is easy to differentiate and v' easy to integrate. Since we only need an antiderivative, when we integrate e^x we will set the arbitrary constant to 0 for convenience. Thus, we find that

$$du = dx, \quad \text{and} \quad v = e^x.$$

Using the formula for integration by parts

$$\int \underbrace{x}_{u} \underbrace{e^x dx}_{dv} = \underbrace{x}_{u} \underbrace{e^x}_{v} - \int \underbrace{e^x}_{v} \underbrace{dx}_{du} = xe^x - e^x + c.$$

Example 2. Evaluate $\displaystyle\int \ln(x) dx$.

Solution In this case we can write $\ln(x) = 1 \cdot \ln(x)$ and apply integration by parts. Since we don't know how to integrate $\ln(x)$, we will set

$$u = \ln(x), \quad \text{and} \quad dv = dx.$$

We find

$$du = \frac{1}{x} dx \quad \text{and} \quad v = x,$$

once again setting the arbitrary constant to 0 (which we will always do with integration by parts). Finally,

$$\int \ln(x) dx = x \cdot \ln(x) - \int x \frac{1}{x} dx = x \ln(x) - x + c.$$

Example 3. Evaluate $\displaystyle\int x \sec(x)^2 dx$.

Solution Although we know how to differentiate $\sec(x)^2$, it isn't likely to simplify the problem. The integral of $\sec(x)^2$ is $\tan(x)$, so it seems that the following definitions

$$u = x \quad \text{and} \quad dv = \sec(x)^2 dx,$$

will help us solve the problem. We find

$$du = dx \quad \text{and} \quad v = \tan(x).$$

This yields

$$\int x \sec(x)^2 dx = x \tan(x) - \int \tan(x) dx.$$

In order to evaluate the second integral we rewrite $\tan(x) = \frac{\sin(x)}{\cos(x)}$ and use substitution. Define

$$u = \cos(x),$$

which implies

$$du = -\sin(x) dx.$$

Now we can evaluate

$$\int \frac{\sin(x)}{\cos(x)} dx = -\int \frac{1}{u} du = -\ln|u| + c = -\ln|\cos(x)| + c.$$

Substituting into the original expression we find

$$\int x \sec(x)^2 dx = x \tan(x) - (-\ln|\cos(x)| + c) = x \tan(x) + \ln|\cos(x)| + c.$$

Example 4. Evaluate $\int x^2 \sin(x) dx$.

Solution Here we definitely do not want to integrate x^2, as it will only yield higher powers of x, complicated the problem. Due to the cyclical nature of the sine function, we can integrate or differentiate it as many times as required without complicating the problem. For this reason we will want to differentiate x^2 and integrate $\sin(x)$. Set

$$u = x^2 \quad \text{and} \quad dv = \sin(x)dx.$$

This leads to

$$du = 2xdx \quad \text{and} \quad v = -\cos(x).$$

Using integration by parts

$$\int x^2 \sin(x)dx = x^2 \cdot (-\cos(x)) - \int (-\cos(x)) \cdot 2xdx = -x^2 \cos(x) + 2 \int x \cos(x)dx.$$

In order to evaluate the integral on the right we'll need to use integration by parts once again, setting

$$v = x \quad \text{and} \quad du = \cos(x)dx,$$

which implies

$$dv = dx \quad \text{and} \quad u = \sin(x).$$

Now we can find

$$\int x \cos(x)dx = x \sin(x) - \int 1 \cdot \sin(x)dx = x \sin(x) - (-\cos(x)) + c = x \sin(x) + \cos(x) + c.$$

Substituting back into the original expression

$$\int x^2 \sin(x)dx = -x^2 \cos(x) + 2(x \sin(x) + \cos(x) + c) = (2 - x^2) \cos(x) + 2x \sin(x) + c.$$

Although we were able to work through the above example using integration by parts twice, the process can become a little bit cumbersome. Imagine an integral of the form

$$\int x^3 e^{-x} dx.$$

It's not hard to see how we could use the same process as above, but then we'd need to perform integration by parts three times. However, with a little bit of ingenuity we can greatly streamline the process. When we have one term we need to integrate repeatedly, and another we need to differentiate repeatedly, we can use tabular integration by parts. Essentially, we create a table of the respective derivatives and integrals, and combine the terms alternating between positive and negative signs. The easiest way to see how this works is through an example.

Example 5. Evaluate $\int x^3 e^x dx$.

Solution First we identify

$$u = x^3 \quad \text{and} \quad v' = e^x$$

as the functions we want to repeatedly integrate and differentiate. In order to solve the problem we now construct a table of the derivatives of u and the integrals of v'.

u **and its derivatives**	v' **and its integrals**
x^3	e^x
$3x^2$	e^x
$6x$	e^x
6	e^x
0	e^x

Now in order to solve the problem we simply combine the above terms in the proper fashion. We begin by taking the first member of the first column, and multiplying by the second term of the second column. From here we subtract the second member of the first column times the third member of the second column. Next we add the third term of the first column times the fourth term of the second column, and so, alternating between a positive and negative sign between each set of terms. In the end we find

$$\begin{aligned} \int x^3 e^x dx &= \Big(x^3 \cdot e^x\Big) - \Big(3x^2 \cdot e^x\Big) + \Big(6x \cdot e^x\Big) - \Big(6 \cdot e^x\Big) + c \\ &= (x^3 - 3x^2 + 6x - 6) \cdot e^x \end{aligned}$$

Example 6. Evaluate $\displaystyle\int x^2 \cos(x) dx$.

Solution In this situation we will identify

$$u = x^2 \quad \text{and} \quad v' = \cos(x).$$

We find that

u **and its derivatives**	v' **and its integrals**
x^2	$\cos(x)$
$2x$	$\sin(x)$
2	$-\cos(x)$
0	$-\sin(x)$

Thus, our end result is that

$$\begin{aligned} \int x^2 \cos(x) dx &= \Big(x^2 \cdot \sin(x)\Big) - \Big(2x \cdot (-\cos(x))\Big) + \Big(2 \cdot (-\sin(x))\Big) \\ &= (x^2 - 2) \cdot \sin(x) + 2x \cdot \cos(x) \end{aligned}$$

Comparing this to the previous example we working with sine rather than cosine, we see that tabular integration by parts is much more efficient when we need to integrate by parts multiple times.

In addition to using integration by parts for indefinite integrals, we can use integration by parts for definite integrals. Since indefinite integration by parts provides us an antiderivative for a given integral, all we need to do is apply the fundamental theorem of calculus, and we see that for a definite integral, we have

$$\begin{aligned} \int_a^b u(x)v'(x)dx &= \Big[u(x)v(x) - \int v(x)u'(x)dx\Big|_a^b \\ &= u(x)v(x)\Big|_a^b - \int_a^b v(x)u'(x)dx \end{aligned}$$

In order to interpret the above analysis, we can think of the indefinite integral

$$\int v(x)u'(x)dx$$

as accumulating area from some arbitrary starting point. If we evaluate this function at b and subtract by it evaluated at a, the result will be exactly the area accumulated over the interval (a,b). Formally, we state the following theorem.

Theorem 5.7.1 (Integration By Parts)**.** *Let u and v be differentiable functions, with continuous first derivatives. It follows that*

$$\int_a^b u(x)v'(x)dx = u(x)v(x)\Big|_a^b - \int_a^b u'(x)v(x)dx.$$

In terms of area, there is a nice graphical interpretation for integration by parts. Written in differential form, definite integration by parts reads

$$\int_{v(a)}^{v(b)} u(v)dv = u(b)v(b) - u(a)v(a) - \int_{u(a)}^{u(b)} v(u)du,$$

which expresses the functions u and v in terms of each other, using the intermediate variable x. Viewed from this lens, we arrive at figure 5.6, which relates the areas in the different portions of the integration by parts formula.

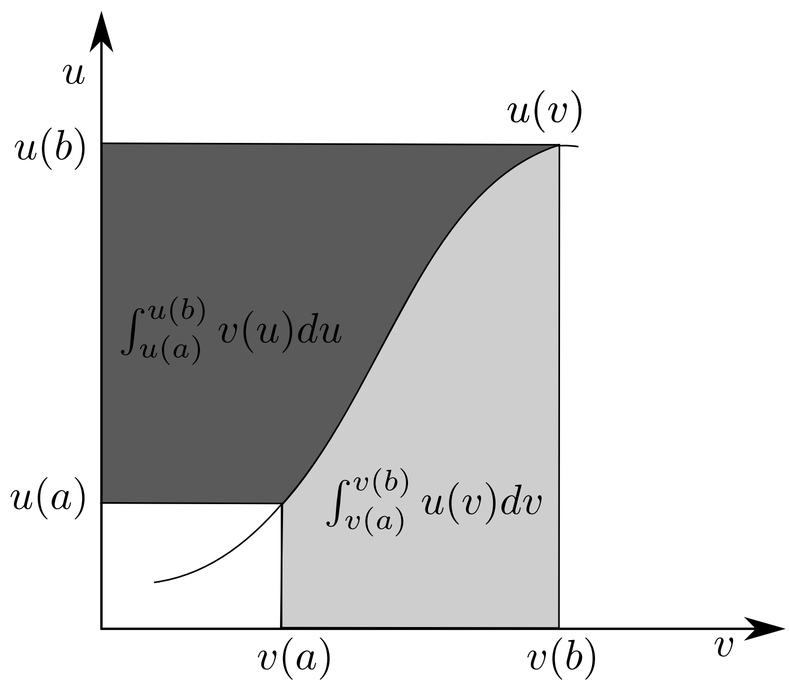


Figure 5.6: Graphical Interpretation of Integration By Parts.

Example 7. Evaluate $\displaystyle\int_1^2 t^6 \ln(t)dt$.

Solution One method would be simply to evaluate the indefinite integral, and evaluate at the appropriate limits. However, let's use this chance to showcase definition integration by parts. Since we can't integrate $\ln(t)$ directly, we'll need to make the following choices

$$u = \ln(t) \quad \text{and} \quad dv = t^6 dt$$

so that

$$du = \frac{1}{t}dt \quad \text{and} \quad \frac{1}{7}t^7.$$

We find that

$$\begin{aligned}
\int_1^2 t^6 \ln(t)dt &= \frac{1}{7}t^7 \ln(t)\Big|_1^2 - \int_1^2 \frac{1}{7}t^7 \cdot \frac{1}{t}dt = \frac{1}{7}(128\ln(2) - \ln(1)) - \frac{1}{7}\int_1^2 t^6 dt \\
&= \frac{128}{7}\ln(2) - \frac{1}{49}t^7\Big|_1^2 = \frac{128}{7}\ln(2) - \frac{127}{49} \approx 10.083.
\end{aligned}$$

Although in many previous examples we found ourselves differentiating the power function involved, since above we could not integrate $\ln(t)$, we instead had to integrate the power function. Related to integration by parts there is a little mnemonic ILATE, which specifies the priority in which someone should choose which function to differentiate and integrate. The higher on the list a function is, the higher of a priority choice for differentiation it is.

- Inverse Trigonometric Functions: $\tan^{-1}(x)$, $\cos^{-1}(x)$, etc.
- Logarithmic Functions: $\ln(x)$, $\log(x)$, etc.
- Algebraic Functions (polynomials): $x^2 + x$, x^3, etc.
- Trigonometric Functions: $\cos(x)$, $\tan(x)$, etc.
- Exponential Functions: e^x, a^x, etc.

The basic idea is that the higher up on the list a given function is, the more difficult it is to integrate, so it is the natural choice for differentiation in the process of integration by parts.

5.8 Partial Fractions

We already know how to evaluate the integral of any polynomial function, simply using linearity of integration. Unfortunately, the process becomes a bit more complicated when we move from polynomial to rational functions (quotients of polynomials). In the simplest cases, such as

$$\int \frac{6x+1}{3x^2+x}dx,$$

the numerator is simply the derivative of the denominator. Then we simply choose the substitution

$$u = 3x^2 + x \quad \text{and} \quad du = (6x+1)dx,$$

which gives us the result of

$$\int \frac{6x+1}{3x^2+x}dx = \int \frac{1}{u}du = \ln|u| + c = \ln|3x^2+1| + c.$$

However, in a more general context, a rational function is not the ratio of a function's derivative and the function itself. In such a situation we cannot simply use substitution to evaluate the integral. Instead, we will need to rewrite the rational function, in terms of *partial fractions*. In order to write a rational function in terms of partial fractions, we factor the denominator, and break it up into separate terms. For instance, we write

$$\frac{3x+1}{x^2-2x-1} = \frac{3x+1}{(x+1)(x-2)} = \frac{A}{x+1} + \frac{B}{x-2},$$

where A and B are as of yet undetermined constants. If we establish a common denominator on the right-hand side of the equation, the denominators will be the same on both sides, and we'll simply need to solve for A and B in the numerator. We will find that

$$3x+1 = A(x-2) + B(x+1),$$

or

$$3 = A + B \quad \text{and} \quad 1 = -2A + B.$$

We reach this set of two equations from the single equation above noting that two polynomials are equal if and only if all of their coefficients are equal. Subtracting the second equation from the first we find $A = 2/3$ and substituting A into the first we find $B = 7/3$. From here we find that

$$\frac{3x+1}{x^2-2x-1} = \frac{2/3}{x+1} + \frac{7/3}{x-2}.$$

Rewritten in this way the task of integrating the two functions on the right-hand side is simple enough. In theory we can rewrite any rational function in terms of partial fractions. In this section we will explore techniques for doing so.

The first thing to note is that before we can write a rational function in terms of partial fractions, we must have a proper fraction. What this means is that the degree of the polynomial in the numerator must be less than or equal to the degree of the polynomial in the denominator. Fortunately, it is possible to rewrite any improper fraction in terms of proper fractions. Polynomial division is the tool we use for the job. Suppose we have an improper fraction, such as

$$\frac{2x^3-4x^2-x-3}{x^2-2x-3}.$$

Using polynomial long division we find

$$
\begin{array}{rl}
 & 2x \\
x^2-2x-3 \big) & \overline{\;2x^3-4x^2 \quad -x-3} \\
 & -2x^3+4x^2+6x \\
\cline{2-2}
 & 5x
\end{array}
$$

We begin the above process by determining that we need to multiply the highest order term of the denominator by $2x^2$ in order to have the same term as in the numerator. From this starting point we multiply the denominator by $2x$, and subtract from the numerator. The result is $5x$, which is not divisible by x^2, so the process ends. We pull down the remaining term -3 to finish the remainder, which the final result that

$$\frac{2x^3-4x^2-x-3}{x^2-2x-3} = 2x + \frac{5x-3}{x^2-2x-3}.$$

Now that we've rewritten the improper fraction into a sum of proper fractions, we can rewrite it in terms of partial fractions and integrate as required.

Example 1. Write $\dfrac{x^4+x^2-1}{x^3+x}$ as a sum of proper fractions.

Solution Performing polynomial division we find

$$
\begin{array}{rl}
 & x \\
x^3+x \big) & \overline{\;x^4+x^2}-1 \\
 & -x^4-x^2 \\
\cline{2-2}
\end{array}
$$

so that

$$\frac{x^4+x^2-1}{x^3+x} = x - \frac{1}{x^3+x}.$$

With these preliminaries out of the way, we can move onto the techniques for rewriting rational functions in terms of partial fractions. Depending on the factors of the denominator the process varies slightly. We will need to distinguish between first-order and irreducible second-order factors, as well as account for repeated factors. For first-order factors we will proceed as we did in the beginning of the section. Finally, we will see that when we have a repeated factor in the denominator, such as $(x+1)^2$, we will need to include an additional term in the partial fraction decomposition.

Example 2. Evaluate $\displaystyle\int \frac{3x-1}{x^2-x-6}dx.$

Solution We begin by decomposing the integrand into partial fractions

$$\frac{3x-1}{x^2-x-6} = \frac{A}{x+2} + \frac{B}{x-3}.$$

On the right-hand side we multiply the term with A by $(x-3)/(x-3)$, and the term containing B by $(x+2)/(x+2)$. This establishes a common denominator on both sides of the equation, so in order to establish equality we only need to consider the numerators. Thus,

$$3x-1 = A(x-3) + B(x+2),$$

or

$$3 = A + B \quad \text{and} \quad -1 = -3A + 2B,$$

breaking up the above equation in terms of like powers of x. For a system of two simultaneous equations such as above, it is usually easiest to add or subtract a multiple of one equation from the other, so that one of the variables is canceled. If we multiply the first equation by 3 and add it to the second, we find

$$9 - 1 = 3B + 2B,$$

or $B = 8/5$. We can substitute this value for B into either of the above equations to find $A = 7/5$. Now we are ready to integrate, and we find

$$\int \frac{3x-1}{x^2-x-6}dx = \int \frac{7/5}{x+2}dx + \int \frac{8/5}{x-3}dx = \frac{7}{5}\ln|x+2| + \frac{8}{5}\ln|x-3| + c.$$

It's worth noting that in the above example we could deal with any number of factors. For instance, if we had a polynomial

$$\frac{x^2+1}{(x-1)(x-2)(x-3)}$$

we would break it up as

$$\frac{x^2+1}{(x-1)(x-2)(x-3)} = \frac{A}{x-1} + \frac{B}{x-2} + \frac{C}{x-3},$$

and we would need to solve a system of three simultaneous equations. Given that we have not studied efficient methods for solving large systems of equations (such as using augmented matrices and Gauss-Jordan elimination), we will try avoiding dealing with systems of this size.

Example 3. Evaluate $\displaystyle\int \frac{6x^2-3x+1}{(4x+1)(x^2+1)}dx.$

Solution In this situation we have a first-order factor $(4x+1)$ and a second-order factor (x^2+1), which does not factor over the real numbers. Here our partial fraction decomposition looks like

$$\frac{6x^2-3x+1}{(4x+1)(x^2+1)} = \frac{A}{4x+1} + \frac{Bx+C}{x^2+1}.$$

Now we multiply by the appropriate factors to create a common denominator, and solve the equation in the numerator, to find

$$6x^2 - 3x + 1 = A(x^2+1) + (Bx+C)(4x+1).$$

In order to solve this system of equations we can substitute using clever values for x. A good first choice is $x = -1/4$, because this will eliminate both B and C in the above equation, yielding

$$6/16 + 3/4 + 1 = 17/16 \cdot A,$$

so $A = 2$. If we let $x = 0$, we can eliminate B from the above equation, yielding

$$1 = 2 + C,$$

so $C = -1$. Finally if we let $x = 1$ we can substitute in the values of A and C to find

$$4 = 4 + (B-1) \cdot 5,$$

so $B = 1$. This leads to the partial fraction decomposition

$$\frac{6x^2 - 3x + 1}{(4x+1)(x^2+1)} = \frac{2}{4x+1} + \frac{x-1}{x^2+1} = \frac{2}{4x+1} + \frac{x}{x^2+1} - \frac{1}{x^2+1}.$$

Now we integrate termwise, and see that

$$\begin{aligned} \int \frac{6x^2 - 3x + 1}{(4x+1)(x^2+1)} dx &= \int \frac{2}{4x+1} dx + \int \frac{x}{x^2+1} dx - \int \frac{1}{x^2+1} dx \\ &= \frac{1}{2} \ln|4x+1| + \frac{1}{2} \ln(x^2+1) - \tan^{-1}(x) + c. \end{aligned}$$

Example 4. Evaluate $\displaystyle\int \frac{6x+7}{(x+2)^2} dx$.

Solution In this situation we are faced with a repeated factor in the denominator. This leads us to the following partial fraction decomposition

$$\frac{6x+7}{(x+2)^2} = \frac{A}{x+2} + \frac{B}{(x+2)^2}.$$

We find that

$$6x + 7 = A(x+2) + B,$$

which implies $A = 6$ and $B = -5$. Now we can integrate,

$$\int \frac{6x+7}{(x+2)^2} dx = \int \frac{6}{x+2} dx + \int \frac{-5}{(x+2)^2} dx = 6 \ln|x+2| + 5(x+2)^{-1} + c.$$

In the above example we have a repeated first-order factor, but the method is the same for a second-order factor. For instance, we would find the following partial fraction decomposition for a rational function with a repeated second-order factor

$$\frac{6x^2 - 15x + 22}{(x+3)(x^2+2)^2} = \frac{A}{x+3} + \frac{Bx+C}{x^2+2} + \frac{Dx+E}{(x^2+2)^2}.$$

Rather than dealing with a system of five equations, we simply note that the process is the same as in the examples above.

When we are dealing with a polynomial in the denominator that consists entirely of distinct, first-order factors, the Heaviside "cover-up" method is a quick means of finding a partial fraction decomposition. We will present this method through example.

Example 5. Expand $\displaystyle\frac{x^2+1}{(x-1)(x-2)(x-3)}$ into partial fractions.

Solution We begin by noting the expansion will be of the form

$$\frac{x^2+1}{(x-1)(x-2)(x-3)} = \frac{A}{x-1} + \frac{B}{x-2} + \frac{C}{x-3}.$$

Normally, we would begin by creating a common denominator for all terms. However, this becomes a bit cumbersome when dealing with three terms, because each of the unknowns A, B, and C will then be multiplied by two factors. Nevertheless, if we were to multiply both sides of the equation by $(x-1)$, we would find

$$\frac{x^2+1}{(x-2)(x-3)} = A + \frac{B}{x-2} \cdot (x-1) + \frac{C}{x-3} \cdot (x-1).$$

Now if we substitute for $x = 1$ the last two terms on the right-hand side will vanish. Doing so we find that

$$A = \frac{1^2 + 1}{(1-2)(1-3)} = 1.$$

Rather than actually going through the multiplication and rewriting the above expression, we note that this is the same result we would have arrived at if we were to just "cover-up" or ignore the term $(x - 1)$ on the left-hand side, and evaluate at $x = 1$. In a similar vein, we can find B by covering the factor $(x - 2)$, and evaluating the left-hand side at $x = 2$. Thus,

$$B = \frac{2^2 + 1}{(2-1)(2-3)} = -5.$$

In general, we find the value of the unknown with a factor of $(x-k)$ in its denominator by covering up the factor of $(x - k)$ on the left-hand side, and evaluating at $x = k$. For C we have $k = 3$, so

$$C = \frac{3^2 + 1}{(3-1)(3-2)} = 5.$$

This process works because when we are covering up the given factor it is equivalent to multiplying all terms by that factor and substituting in for x at the point where the factor is 0, which causes the other terms to vanish. It should be cautioned that this process only works when we have first-order factors.

5.9 Trigonometric Integrals

The key to evaluating most integrals involving trigonometric functions is usage of the appropriate trigonometric identity. In this section we will explore some of the general categories of trigonometric integrals one might encounter, and which trigonometric identities are used to solve them. More than anything this section should be used as a reference - the time required to compute most of these integrals by hand is generally unjustified.

The first category of trigonometric integrals to consider consists of products of sine and cosine functions, which occur in some engineering and physics applications. For instance, such integrals arise in looking at a one dimensional infinite potential well in quantum mechanics. In order to evaluate such integrals we use the following set of trigonometric identities

$$\begin{aligned}\sin(mx)\sin(nx) &= \frac{1}{2}[\cos((m-n)x) - \cos((m+n)x)]\\ \sin(mx)\cos(nx) &= \frac{1}{2}[\sin((m-n)x) + \sin((m+n)x)]\\ \cos(mx)\cos(nx) &= \frac{1}{2}[\cos((m-n)x) + \cos((m+n)x)]\end{aligned}$$

From the above identities integration becomes simple enough.

Example 1. Show that $\displaystyle\int_0^a \sin\left(\frac{m\pi x}{a}\right)\sin\left(\frac{n\pi x}{a}\right)dx = 0$, $m, n \in \mathbb{Z}$.

Solution We begin by using the appropriate trigonometric identity, namely that

$$\sin\left(\frac{m\pi x}{a}\right)\sin\left(\frac{n\pi x}{a}\right) = \frac{1}{2}\Big[\cos\left(\frac{(m-n)\pi x}{a}\right) - \cos\left(\frac{(m+n)\pi x}{a}\right)\Big].$$

Both of the terms on the right-hand side can be integrated using substitution, but the integration is easy enough to see by inspection (think of the constant $(m+n)\pi/a$ as a simple constant like 2). We find that

$$\begin{aligned}\int_0^a \sin\left(\frac{m\pi x}{a}\right)\sin\left(\frac{n\pi x}{a}\right)dx &= \frac{1}{2}\int_0^a \cos\left(\frac{(m-n)\pi x}{a}\right)dx - \frac{1}{2}\int_0^a \cos\left(\frac{(m+n)\pi x}{a}\right)dx\\ &= \frac{1}{2}\Big[\frac{a}{(m-n)\pi}\sin\left(\frac{(m-n)\pi x}{a}\right) - \frac{a}{(m+n)\pi}\sin\left(\frac{(m+n)\pi x}{a}\right)\Big|_0^a\\ &= 0.\end{aligned}$$

We arrive at the final result of 0 noting that the sine function vanishes at integer multiples of π.

Another class of trigonometric integrals we encounter involves powers of sine and cosine. If we have an integrand of the form $\sin^n(x)\cos^m(x)$ there are three cases to consider.

1. For m odd we write $m = 2k+1$, and use the identity $\sin^2(x) = 1 - \cos^2(x)$, yielding

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

   Now we use u-substitution, with $u = \cos(x)$, so $-du = \sin(x)dx$.

2. For m even but n odd we write $n = 2k+1$ and use the identity $\cos^2(x) = 1-\sin^2(x)$, yielding

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

   Now we use u-substitution, with $u = \sin(x)$, so $du = \cos(x)dx$.

3. For both m and n even we use

$$\sin^2(x) = \frac{1-\cos(2x)}{2} \quad \text{and} \quad \cos^2(x) = \frac{1+\cos(2x)}{2},$$

to reduce the integrand to lower powers of $\cos(2x)$.

Example 2. Evaluate $\displaystyle\int \sin^3(x)\cos^2(x)dx$.

Solution We start here by writing $\sin^3(x) = (1-\cos^2(x))\sin(x)$. Now we substitute into the integral,

$$\int \sin^3(x)\cos^2(x)dx = \int (1-\cos^2(x))\sin(x)\cos^2(x)dx = \int (1-\cos^2(x))\cos^2(x)\sin(x)dx.$$

At this point we use the substitution $u = \cos(x)$, so $-du = \sin(x)dx$. We find that

$$\int \sin^3(x)\cos^2(x)dx = \int (1-u^2)(u^2)(-du) = \int (u^4-u^2)du = \frac{u^5}{5} - \frac{u^2}{2} + c = \frac{\cos^5(x)}{5} - \frac{\cos^3(x)}{3} + c.$$

Example 3. Evaluate $\displaystyle\int \cos^5(x)dx$.

Solution Here we want to write $\cos^5(x) = (1-\sin^2(x))^2\cos(x)$. This will yield

$$\int \cos^5(x)dx = \int (1-\sin^2(x))^2\cos(x)dx.$$

If we let $u = \sin(x)$ then we have $du = \cos(x)dx$, and

$$\int (1-\sin^2(x))^2\cos(x)dx = \int (1-u^2)^2du = \int (1-2u^2+u^4)du = \sin(x) - \frac{2}{3}\sin^3(x) + \frac{1}{5}\sin^5(x) + c.$$

Example 4. Evaluate $\displaystyle\int \cos^2(x)\sin^2(x)dx$.

Solution First we write

$$\sin^2(x) = (1-\cos(2x))/2 \quad \text{and} \quad \cos^2(x) = (1+\cos(2x))/2.$$

Now we find

$$\int \cos^2(x)\sin^2(x)dx = \frac{1}{4}\int (1-\cos(2x))(1+\cos(2x))dx = \frac{1}{4}\int (1-\cos^2(2x))dx.$$

Once again we can use the power-reducing identity, this time writing

$$\cos^2(2x) = (1+\cos(4x))/2.$$

Substituting back into the integral we find

$$\int \cos^2(x)\sin^2(x)dx = \frac{1}{4}\int (1-(1+\cos(4x))/2)dx = \frac{1}{4}\int \Big(\frac{1}{2} - \frac{\cos(4x)}{2}\Big)dx = \frac{x}{8} - \frac{\sin(4x)}{32} + c.$$

This result should be sufficient, but at this point we can also use the double-angle identities

$$\sin(2x) = 2\sin(x)\cos(x) \quad \text{and} \quad \cos(2x) = 2\cos^2(x) - 1$$

in order to rewrite the solution in terms of x (rather than $4x$). Doing so we find

$$\begin{aligned}\frac{\sin(4x)}{32} &= \frac{2}{32}\sin(2x)\cos(2x) = \frac{2\cdot 2}{32}\sin(x)\cos(x)\cos(2x)\\ &= \frac{1}{8}\sin(x)\cos(x)\big(2\cos^2(x)-1\big) = \frac{\sin(x)\cos^3(x)}{4} - \frac{\sin(x)\cos(x)}{8}.\end{aligned}$$

This yields the final result that

$$\int \cos^2(x)\sin^2(x)dx = \frac{x}{8} - \frac{\sin(x)\cos^3(x)}{4} + \frac{\sin(x)\cos(x)}{8} + c.$$

Additionally, some integrals can be solved using trigonometric identities to remove square roots.

Example 5. Evaluate $\displaystyle\int_0^{\pi/4} \sqrt{1+\cos(4x)}dx$.

Solution Once again we use the power-reducing identity, but here we actually want to increase the power of cosine, to remove the square root. Noting that

$$1+\cos(2x) = 2\cos^2(x)$$

we find

$$1+\cos(4x) = 2\cos^2(2x).$$

Evaluating the integral we find

$$\int_0^{\pi/4} \sqrt{1+\cos(4x)}dx = \int_0^{\pi/4} \sqrt{2\cos^2(2x)}dx = \sqrt{2}\int_0^{\pi/4} |\cos(2x)|dx.$$

At this point it is important to note the limits of integration. Since cosine is nonnegative on $[0,\pi/2]$, we have that $\cos(2x)$ is nonnegative on $[0,\pi/4]$. Thus, we do not need to worry about the absolute value in the integrand. Proceeding further we find

$$\int_0^{\pi/4} \sqrt{1+\cos(4x)}dx = \sqrt{2}\int_0^{\pi/4} \cos(2x)dx = \sqrt{2}\cdot\frac{\sin(2x)}{2}\bigg|_0^{\pi/4} = \frac{\sqrt{2}}{2}.$$

There are also tricks for working with other trigonometric integrals, but we will not consider them here. For instance, integrals involving $\tan(x)$ and $\sec(x)$ can be evaluated using the Pythagorean identity along with integration by parts.

5.10 Trigonometric Substitution

When we are faced with integrals involving terms of the form $\sqrt{a^2 - x^2}$, $\sqrt{a^2 + x^2}$, or $\sqrt{x^2 - a^2}$ we can use trigonometric substitution in order to evaluate the integral. The three basic substitutions we use are $x = a\sin(\theta)$, $x = a\tan(\theta)$, and $x = a\sec(\theta)$, corresponding to the respective expressions above. See figure 5.7 for the triangles corresponding to these substitutions.

- With $x = a\sin(\theta)$,

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

- With $x = a\tan(\theta)$,

$$a^2 + x^2 = a^2 + a^2\tan(\theta) = a^2(1 + \tan^2(\theta)) = a^2\sec^2(\theta).$$

- With $x = a\sec(\theta)$,

$$x^2 - a^2 = a^2\sec^2(\theta) - a^2 = a^2(\sec^2(\theta) - 1) = a^2\tan^2(\theta).$$

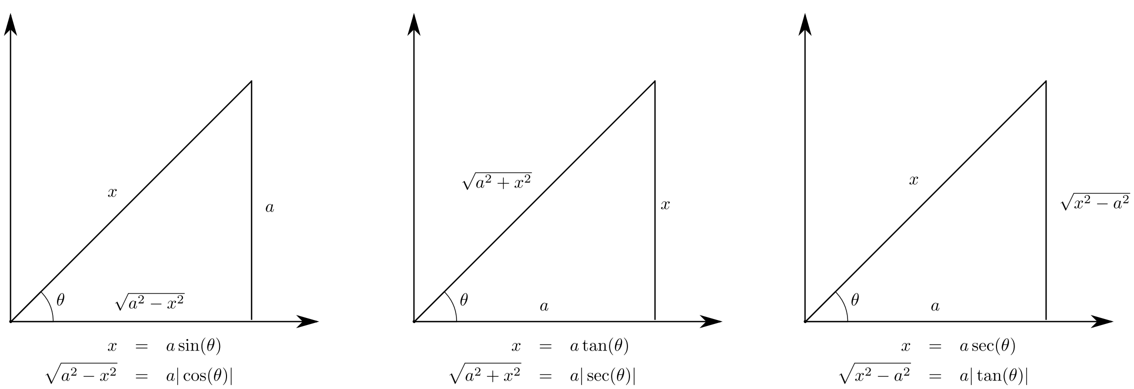


Figure 5.7: Triangles corresponding to trigonometric substitutions.

After making these substitutions, we are effectively able to remove the square root from the integrand. Nevertheless, we do need to be a little bit careful, because after making these substitutions we want to return to the original variable of x. This puts a restriction on θ, because the inverse trigonometric functions only exist for limited values of θ. In each of these three cases

$$\begin{array}{lll}
\theta = \sin^{-1}(x/a) & \text{requires} & -\pi/2 \le \theta \le \pi/2. \\
\theta = \tan^{-1}(x/a) & \text{requires} & -\pi/2 < \theta < \pi/2. \\
\theta = \sec^{-1}(x/a) & \text{requires} & \begin{cases} 0 \le \theta < \pi/2, & x/a \ge 1, \\ \pi/2 < \theta \le \pi, & x/a \le -1. \end{cases}
\end{array}$$

Let's further investigate the process of trigonometric substitution by considering some examples.

Example 1. Evaluate $\displaystyle\int \sqrt{a^2 - x^2}dx$.

Solution Here we use the substitution

$$x = a\sin(\theta), \quad -\pi/2 \le \theta \le \pi/2,$$

so $dx = a\cos(\theta)d\theta$ and $\sqrt{a^2-x^2} = a|\cos(\theta)| = a\cos(\theta)$, because cosine is positive over these limits. Now we find

$$\begin{aligned}\int \sqrt{a^2-x^2}dx &= \int a\cos(\theta)\cdot a\cos(\theta)d\theta = a^2\int \cos^2(\theta)d\theta \\ &= \frac{a^2}{2}\int (1+\cos(2\theta))d\theta = \frac{a^2}{2}\Big(\theta + \frac{1}{2}\sin(2\theta)\Big) + c = \frac{a^2}{2}(\theta + \sin(\theta)\cos(\theta)) + c.\end{aligned}$$

Now we need to use the inverse sine function, $\theta = \sin^{-1}(x/a)$, in order to rewrite our result in terms of x.

$$\cos(\theta) = \cos\Big(\sin^{-1}(x/a)\Big) = \frac{\sqrt{a^2-x^2}}{a},$$

using the first triangle in figure 5.7. Finally, we find that

$$\int \sqrt{a^2-x^2}dx = \frac{a^2}{2}\Big(\sin^{-1}(x/a) + \frac{x}{a}\cdot\frac{\sqrt{a^2-x^2}}{a}\Big) + c = \frac{a^2}{2}\sin^{-1}(x/a) + \frac{x}{2}\sqrt{a^2-x^2} + c.$$

Now if we recognize that $\sqrt{a^2-x^2}$ is just the equation of the upper half of a circle, we can evaluate this integral at the appropriate limits, calculating the area of the circle as

$$\int_{-a}^{a} \sqrt{a^2-x^2}dx = \Big[\frac{a^2}{2}\sin^{-1}(x/a) + \frac{x}{2}\sqrt{a^2-x^2}\Big|_{-a}^{a} = \frac{a^2}{2}\Big[\frac{\pi}{2} + \frac{\pi}{2}\Big] = \frac{\pi a^2}{2}.$$

Example 2. Evaluate $\displaystyle\int \frac{dx}{\sqrt{4+x^2}}$.

Solution Here we use the substitution

$$x = 2\tan(\theta), \quad -\pi/2 < \theta < \pi/2,$$

so $dx = 2\sec^2(\theta)d\theta$ and $\sqrt{4+x^2} = 2|\sec(\theta)| = 2\sec(\theta)$, because secant is positive over these limits. Now we find

$$\int \frac{dx}{\sqrt{4+x^2}} = \int \frac{2\sec^2(\theta)d\theta}{2\sec(\theta)} = \int \sec(\theta)d\theta.$$

Integrating secant here takes a bit of work, but with a clever use of substitution and multiplication by 1 we find

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

Now we let

$$u = \tan(x) + \sec(x) \quad \text{so} \quad du = (\sec^2(x) + \sec(x)\tan(x))dx,$$

and

$$\int \sec(x)dx = \int \frac{du}{u} = \ln|u| + c = \ln|\sec(x)+\tan(x)| + c.$$

With that out of the way we now have

$$\int \frac{dx}{\sqrt{4+x^2}} = \ln|\sec(\theta)+\tan(\theta)| + c,$$

which we need to rewrite in terms of x, using $\theta = \tan^{-1}(x/a)$. Using the second triangle in figure 5.7 we find

$$\ln|\sec(\theta)+\tan(\theta)| + c = \ln|\frac{\sqrt{4+x^2}}{2} + \frac{x}{2}| + c = \ln|\sqrt{4+x^2} + x| + d,$$

letting $d = c - \ln(2)$.

Example 3. Evaluate $\displaystyle\int_2^4 \frac{\sqrt{x^2-4}}{x}$.

Solution Here we use the substitution

$$x = 2\sec(\theta), \quad 0 \le \theta < \pi/2,$$

so $dx = 2\sec(\theta)\tan(\theta)d\theta$ and $\sqrt{x^2-4} = 2|\tan(\theta)| = 2\tan(\theta)$, because secant is positive over these limits. Now we can proceed with integration, finding

$$\begin{aligned} \int_2^4 \frac{\sqrt{x^2-4}}{x} &= \int_0^{\pi/3} \frac{2\tan(\theta)}{2\sec(\theta)} \cdot 2\sec(\theta)\tan(\theta)d\theta = 2\int_0^{\pi/3} \tan^2(\theta)d\theta \\ &= 2\int_0^{\pi/3} (\sec^2(\theta)-1)d\theta = 2\tan(\theta)\Big|_0^{\pi/3} = 2\sqrt{3} - 2\pi/3 \approx 1.37. \end{aligned}$$

5.11 Numerical Integration

The previous sections provide some methods for finding antiderivatives in very specialized cases, but only offer a little direction in evaluating an arbitrary integral. Not only is there difficulty in choosing which method to use to evaluate a given integral, the majority of integrals cannot be evaluated using any of the previous methods. A few examples are functions such as

$$\sin(x^2), \quad \sqrt{1-x^4}, \quad \sin(x)/x, \quad e^{x^2}.$$

Although some of these difficult integrals can be evaluated using power series, the problems we face above make numerical integration extremely important and useful. Just as we used Newton's method to solve equations we otherwise could not, we will use numerical integration. Furthermore, we can use numerical integration in situations where we only have sample data points, but no actual function. This is extremely useful for measured data, where we do not have a function to work with. Although numerical integration has a number of benefits, we can only use it to evaluate definite integrals, so it will never provide us with an antiderivative - it only provides us with a solution in a specific case. Having an antiderivative provides us with much more information than a single solution, so it is best to use numerical integration to complement our previous techniques of integration, not replace them.

It's worth noting that we already have experience with numerical integration - we used it in order to approximate the area underneath a curve. Recall that in order to calculate the area underneath a curve over a given interval, we divided the larger interval into a number of subintervals, and approximated our function (with a horizontal line) over each of the subintervals. We found the area on each of the subintervals and then added them up. Anytime we were using a finite number of subintervals, we were only approximating the actual value of the integral. Our goal here is to refine this basic process, in order to develop more accurate methods of approximation. Just as the limitations of Euler's method are practical and not theoretical, the limitations of using finite sums of rectangles to approximate integrals are practical and not theoretical. By further developing our techniques of numerical integration we want to remedy some of the practical problems with approximating area as we previously did. Namely, we want to achieve higher degrees of accuracy with less computation.

The basic idea of numerical integration (also called numerical quadrature) is to approximate the function or set of data points representing a function with another function that is simple to integrate. Polynomials are a natural choice, because they are simple to integrate and have a number of useful properties. Namely, for any set of data points (with only one output per input) we can a find polynomial that *interpolates* or intersects all of the given points.

Theorem 5.11.1 (Polynomial Interpolation)**.** *Given a set of $n+1$ data points $(x_i, f(x_i))$, where $x_i \neq x_j$ for $i \neq j$, there exists a unique polynomial p of degree $\leq n$ which interpolates all of the data points. That is,*

$$p(x_i) = f(x_i), \quad i = 0, 1, \ldots, n.$$

The above theorem states that we can find a polynomial to *fit* any set of data points that correspond to some function. An interpolating polynomial has the exact value of the function at the points of interpolation, and between the points of interpolation it provides an approximation. In general, by increasing the number of interpolation points, we should get a better approximation. Nevertheless, in doing so there is a tradeoff with the complexity of calculations that must be made. For our purposes we will only work with up to second-order interpolating polynomials. There are multiple *forms* or ways to represent this interpolating polynomial; we will begin with the *Lagrange form.*

Definition 5.11.1 (Lagrange Form)**.** Given $n+1$ data points $(x_i, f(x_i))$, where $x_i \neq x_j$ for $i \neq j$, the *Lagrange form* of the interpolating polynomial is given by

$$p(x) = \sum_{j=0}^{n} f(x_j) l_j(x),$$

where

$$l_j(x) = \prod_{i=0, i\neq j}^{n} \frac{x - x_i}{x_j - x_i}.$$

Just as $\sum$ is used to represent sums of terms, the symbol $\prod$ is used to represent products of terms. For reference,

$$\prod_{i=1}^{5} i = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$$

and

$$\prod_{i=0}^{2} (x - x_i) = (x - x_0)(x - x_1)(x - x_2).$$

For l_j above the way the product is written is such that the terms are multiplied for all i from 0 to n, with the exception of the i that equals j. For a few cases of interest the Lagrange form is given by

$$\begin{aligned}
p_0(x) &= f(x_0) \\
p_1(x) &= f(x_0) \cdot \frac{x - x_1}{x_0 - x_1} + f(x_1) \cdot \frac{x - x_0}{x_1 - x_0} \\
p_2(x) &= f(x_0) \cdot \left(\frac{x - x_1}{x_0 - x_1} \cdot \frac{x - x_2}{x_0 - x_2} \right) + f(x_1) \cdot \left(\frac{x - x_0}{x_1 - x_0} \cdot \frac{x - x_2}{x_1 - x_2} \right) + f(x_2) \cdot \left(\frac{x - x_0}{x_2 - x_0} \cdot \frac{x - x_1}{x_2 - x_1} \right)
\end{aligned}$$

Above the subscript corresponds to the order of the polynomial; $p_1(x)$ is a first-order interpolating polynomial. Although strictly speaking $p_0(x)$ isn't really given by the Lagrange form, it is included to emphasize that for a single data point $(x_0, f(x_0))$ the interpolating polynomial is just a constant (which is exactly what we used in approximating area with rectangles). It's easy enough to verify that these polynomials really do interpolate the data points $(x_i, f(x_i))$. For instance, consider $p_2(x)$. When we evaluate at x_0, then we find

$$\begin{aligned}
p_2(x_0) &= f(x_0) \cdot \cancelto{1}{\left(\frac{x_0 - x_1}{x_0 - x_1} \cdot \frac{x_0 - x_2}{x_0 - x_2} \right)} + f(x_1) \cdot \cancelto{0}{\left(\frac{x_0 - x_0}{x_1 - x_0} \cdot \frac{x_0 - x_2}{x_1 - x_2} \right)} \\
&\quad + f(x_2) \cdot \cancelto{0}{\left(\frac{x - x_0}{x_2 - x_0} \cdot \frac{x - x_1}{x_2 - x_1} \right)}.
\end{aligned}$$

Although it is easy to verify that the above form truly does provide us with the interpolating polynomial, the Lagrange form is not very practical to work with. We are much better served working with the *Newton form* of the interpolating polynomial.

Definition 5.11.2 (Newton Form)**.** Given $n+1$ data points $(x_i, f(x_i))$, where $x_i \neq x_j$ for $i \neq j$, the *Newton form* of the interpolating polynomial is given by

$$p(x) = \sum_{j=0}^{n} f[x_0, \ldots, x_j] n_j(x),$$

where

$$n_j(x) = \prod_{i=0}^{j-1}(x - x_i).$$

The notation $f[x_0, \ldots, x_j]$ represents the j^{th} divided difference of $f(x)$ at the points $x_0, x_1, \ldots, x_j$. Divided differences can be calculated through a recursive process, with the zeroth divided difference given by

$$f[x_0] = f(x_0)$$

and

$$f[x_0, \ldots, x_j] = \frac{f[x_1, \ldots, x_j] - f[x_0, \ldots, x_{j-1}]}{x_j - x_0}.$$

This gives us the first and second divided differences

$$\begin{aligned} f[x_0, x_1] &= \frac{f(x_1) - f(x_0)}{x_1 - x_0} \\ f[x_0, x_1, x_2] &= \frac{\frac{f(x_2)-f(x_1)}{x_2-x_1} - \frac{f(x_1)-f(x_0)}{x_1-x_0}}{x_2 - x_0}. \end{aligned}$$

This is enough information to calculate the first few interpolating polynomials. Rather than working through the calculations, they are simply listed as

$$\begin{aligned} p_0(x) &= f(x_0) \\ p_1(x) &= f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0} \cdot (x - x_0) \\ p_2(x) &= f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0} \cdot (x - x_0) + \frac{\frac{f(x_2)-f(x_1)}{x_2-x_1} - \frac{f(x_1)-f(x_0)}{x_1-x_0}}{x_2 - x_0} \cdot (x - x_0)(x - x_1) \end{aligned}$$

It is here that the benefit of the Newton form becomes much more clear. Anytime we want to add another interpolation point, assuming that we keep the other points the same, we maintain all of the previous terms of the interpolating polynomial. Thus, if we proceed using the Newton form of the interpolating polynomial, we should be able to reuse a lot of our previous work, as we increase the number of interpolation points.

Now that we've laid down the ground work of interpolating polynomials, we can focus on using them for numerical integration. Our goal is to approximate the value of the definite integral of a function f over an interval $[a, b]$. In order to do so we replace the function f with a polynomial p_n that interpolates (has the same values as) f at the points $x_0, x_1, \ldots, x_n$, and evaluate

$$\int_a^b f(x)dx \approx \int_a^b p_n(x)dx.$$

In general, we expect that using more interpolation points will result in a more accurate approximation. For the moment we'll replace our function f with a single interpolating polynomial over the entire interval $[a, b]$. We'll do this in order to develop the methods of numerical integration. Once we have developed these methods we can use the properties of the integral to write

$$\int_a^b f(x)dx = \int_a^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \ldots + \int_{x_{n-1}}^{b} f(x)dx,$$

and approximate each of the integrals on the right-hand side individually. This is the same idea we used in the past to approximate the area underneath a function with rectangles. Over each subinterval the area underneath the rectangle represents an approximation to the actual area (the definite integral). The only difference is now we will move beyond simply using rectangles, and use higher-order polynomials as well, in order to increase the accuracy of our approximation using fewer subintervals.

The first and simplest of the rules for numerical integration we present here is exactly the rule we've been discussing above - replacing our function of interest with a constant. When we do so the area we will be calculating will be that of a rectangle. Now we only need to determine the height of the rectangle. In general there is no reason to expect that that using the value of the function at the left or right endpoint will provide us with a better approximation, so we will simply choose to use the left endpoint (another person could just as well choose the right endpoint). Doing so we find that

$$\int_a^b f(x)dx \approx \int_a^b f(a)dx = (b-a)f(a).$$

We call this rule for numerical integration the *rectangle rule* (see figure 5.8(a)).

While one might have no reason to expect that the function value at the left or right endpoint will provide a better estimate in general, there is another point of interest we could work with - the midpoint $(a+b)/2$. It is reasonable to expect that the value of the function in the center of the interval will be a better representative of its overall value than one of its endpoints. Using the midpoint we find

$$\int_a^b f(x)dx \approx \int_a^b f\Big(\frac{a+b}{2}\Big)dx = (b-a)f\Big(\frac{a+b}{2}\Big).$$

We call this rule for numerical integration the *midpoint rule* (see figure 5.8(b)).

From here we proceed by moving away from zeroth-order polynomials (constants) to a first-order polynomial, a line. Now for a line we need to choose two points at which to match the function we are dealing with. The most reasonable choices seem to be the endpoints of the interval, a and b. By matching our function at these points we expect to have a better approximation than simply using a constant. Here the equation of our line (in Newton form) that matches $f(x)$ at the points a and b is given by

$$p_1(x) = f(a) + \frac{f(b)-f(a)}{b-a}\cdot(x-a).$$

We've already done the work for the first term with the rectangle rule, so we simply need to integrate the second term, finding

$$\begin{aligned}\int_a^b \frac{f(b)-f(a)}{b-a}\cdot(x-a)dx &= \frac{f(b)-f(a)}{b-a}\cdot(x^2/2-ax)\Big|_a^b \\ &= \frac{f(b)-f(a)}{b-a}\cdot(b^2/2-ab-a^2/2+a^2) \\ &= (f(b)-f(a))\cdot\frac{(b-a)^2}{2(b-a)} \\ &= \frac{(f(b)-f(a))(b-a)}{2}.\end{aligned}$$

Adding this to the result we found from the rectangle rule we find

$$(b-a)f(a) + \frac{(f(b)-f(a))(b-a)}{2} = \frac{1}{2}(b-a)(f(a)+f(b)),$$

so

$$\int_a^b f(x)dx \approx \frac{1}{2}(b-a)(f(a)+f(b)).$$

If both $f(a)$ and $f(b)$ have the same sign, then the area found by this rule corresponds to the area of a trapezoid underneath the function f. For this reason, we call this rule for numerical integration the *trapezoid rule*[1] (see figure 5.8(c)).

The final method we will consider here is using a second-order interpolating polynomial - a parabola. In this case we need to choose three interpolation points. The endpoints a and b are both natural choices, and for a third point we choose the midpoint $c = (a + b)/2$. Here our interpolating polynomial is given by

$$p_2(x) = f(a) + \frac{f(b) - f(a)}{b - a} \cdot (x - a) + \frac{\frac{f(c)-f(b)}{c-b} - \frac{f(b)-f(a)}{b-a}}{c - a} \cdot (x - b)(x - a).$$

From the trapezoid rule we've already integrated the first two of these terms. At this point we can replace $c = (a + b)/2$ and begin some simplification of the third term

$$\begin{aligned} \frac{\frac{f(c)-f(b)}{c-b} - \frac{f(b)-f(a)}{b-a}}{c - a} &= \frac{\frac{f(c)-f(b)}{\frac{a+b}{2}-b} - \frac{f(b)-f(a)}{b-a}}{\frac{a+b}{2} - a} \\ &= \frac{2\frac{f(c)-f(b)}{a-b} + \frac{f(b)-f(a)}{a-b}}{-\frac{a-b}{2}} \\ &= 2\frac{f(a) - 2f(c) + f(b)}{(b - a)^2}. \end{aligned}$$

Now we can integrate the other part of the third term, to find that

$$\begin{aligned} \int_a^b (x - a)(x - b)dx &= \int_a^b (x^2 - ax - bx - ab)dx \\ &= \frac{x^3}{3} - \frac{ax^2}{2} - \frac{bx^2}{2} - abx\Big|_a^b \\ &= \frac{b^3}{3} - \frac{ab^2}{2} - \frac{b^3}{2} - ab^2 - \frac{a^3}{3} + \frac{a^3}{2} + \frac{ba^2}{2} + a^2b \\ &= -\frac{b^3}{6} - \frac{3ab^2}{2} + \frac{3a^2b}{2} + \frac{a^3}{6} \\ &= \frac{-(b - a)^3}{6} \end{aligned}$$

Putting these two pieces together we find that the last term of our interpolating polynomial integrates to

$$2\frac{f(a) - 2f(c) + f(b)}{(b - a)^2} \cdot \frac{-(b - a)^3}{6} = \frac{b - a}{3}(-f(a) + 2f(c) - f(b)).$$

Finally, we can combine this with the result of the trapezoid rule, to find

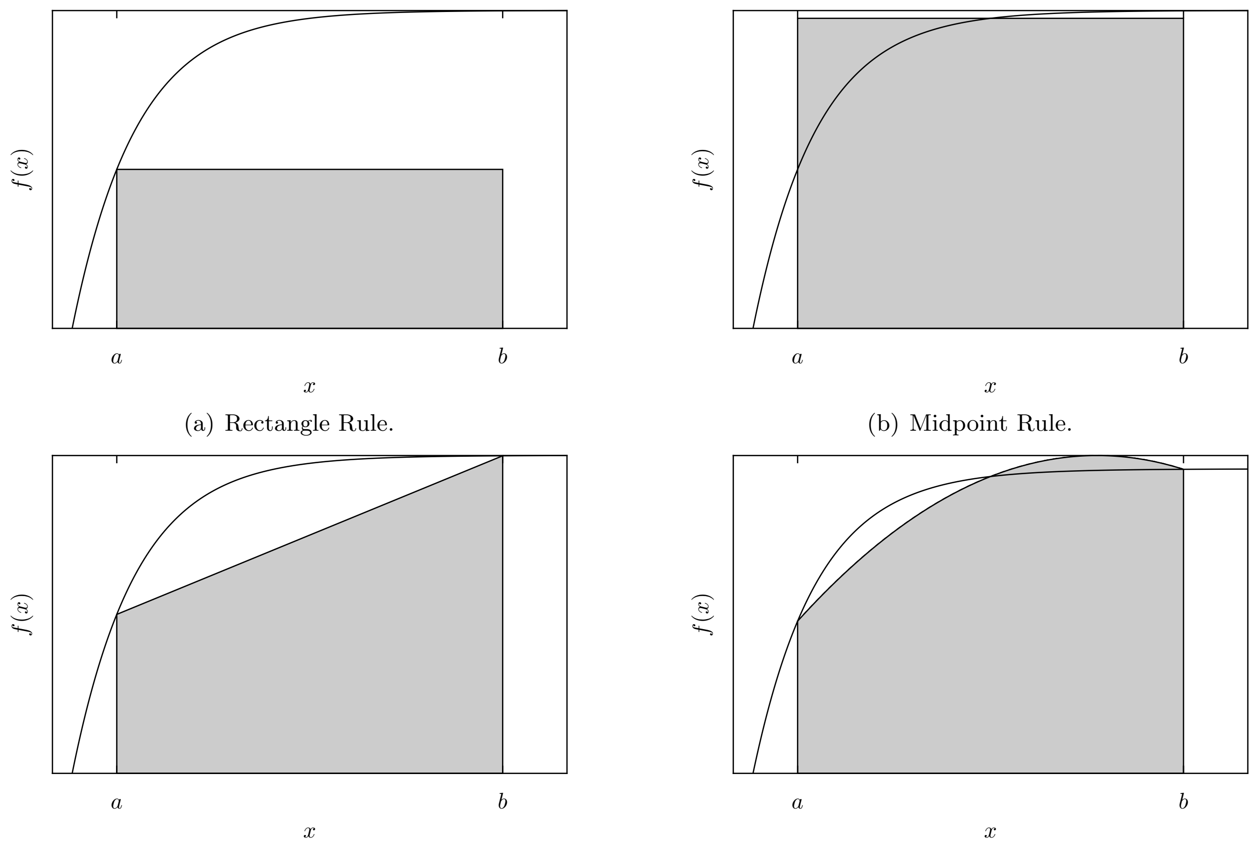
$$\frac{1}{2}(b - a)(f(a) + f(b)) + \frac{b - a}{3}(-f(a) + 2f(c) - f(b)) = \frac{b - a}{6}(f(a) + 4f(c) + f(b)).$$

Finally, substituting $c = (a + b)/2$ we arrive at the result

$$\int_a^b f(x)dx \approx \frac{b - a}{6}\Big(f(a) + 4f\Big(\frac{a + b}{2}\Big) + f(b)\Big).$$

[1] If $f(a)$ and $f(b)$ have opposite signs, then the area we are calculating actually corresponds to two triangles. This in no way invalidates this method of approximation, but it is interesting to note that the area calculated by the trapezoid rule isn't always a trapezoid.

This rule is called *Simpson's rule* (see figure 5.8(d)). It is also sometimes called the parabolic rule, because our interpolating polynomial is a parabola. In the special case where $f(a)$, $f(c)$, and $f(b)$ are all the same sign, we can think of the area given by Simpson's rule as the area underneath the approximating parabola.



(a) Rectangle Rule. (b) Midpoint Rule. (c) Trapezoid Rule. (d) Simpson's Rule.

Figure 5.8: Approximated areas found through numerical integration.

Rather than continuing further in this fashion, let us focus our attention to error analysis, in order to actually justify the claims we made about hoping to reduce the error in our approximations. We will do so by returning to the interpolating polynomial, and looking at the error which occurs at the points where the interpolating polynomial doesn't match the function of interest. Afterwards we will use that information in order to estimate the amount of error that results when we integrate the polynomial rather than our actual function.

Consider a function f on the interval $[a, b]$, with interpolating polynomial p_n interpolating the points $x_0, \ldots, x_n$. The *interpolation error* at any point $x \in [a, b]$ is given by

$$e_n(x) = f(x) - p_n(x).$$

Now let us consider a point $\bar{x}$ distinct from the other interpolation points, and consider the polynomial p_{n+1} which interpolates f at $x_0, \ldots, x_n$ and $\bar{x}$. The Newton form of this polynomial is given by

$$p_{n+1}(x) = p_n(x) + f[x_0, \ldots, x_n, \bar{x}] \prod_{i=0}^{n} (x - x_i).$$

Since p_{n+1} interpolates f at $\bar{x}$, the two functions agree at $\bar{x}$, and

$$f(\bar{x}) = p_{n+1}(\bar{x}) = p_n(\bar{x}) + f[x_0, \ldots, x_n, \bar{x}] \prod_{i=0}^{n} (\bar{x} - x_i).$$

Recall that $\bar{x}$ is simply an arbitrary point not interpolated by p_n. For such a point we have

$$e_n(\bar{x}) = f(\bar{x}) - p_n(\bar{x}) = f[x_0, \ldots, x_n, \bar{x}] \prod_{i=0}^{n} (\bar{x} - x_i).$$

The result here is that the error in our approximation takes on the form of the last term we would add by increasing the number of interpolation points by one. Unfortunately, this error relies on knowledge of $f(\bar{x})$. In some cases we may not know the value, but specifically in our situation we are interested in looking at the error over an interval, so we cannot and don't want to evaluate our function at every possible point $\bar{x}$. Fortunately, it turns out that the divided difference on the right-hand side is related to the $(n+1)^{st}$ derivative of f at a single point, which is enough for us to make estimates of the error.

Theorem 5.11.2. *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and n times differentiable on (a, b). For $n + 1$ distinct points $x_0, \ldots, x_n \in [a, b]$, there exists $\xi \in (a, b)$ such that*

$$f[x_0, \ldots, x_n] = \frac{f^{(n)}(\xi)}{n!}.$$

This theorem is a generalization of the mean value theorem; for $n = 1$ it is just a restatement of the mean value theorem. We can now apply this result to our error function $e_n(x)$ to arrive at another result.

Theorem 5.11.3. *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$, $n + 1$ times differentiable on (a, b), and p_n be the interpolating polynomial of the points $x_0, \ldots, x_n \in [a, b]$. For each $\bar{x} \in [a, b], \bar{x} \notin \{x_0, \ldots, x_n\}$ there exists a corresponding $\xi \in (a, b)$ such that*

$$e_n(\bar{x}) = f(\bar{x}) - p_n(\bar{x}) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^{n} (\bar{x} - x_i).$$

It may not yet seem like we've made much progress. At this point we've merely converted knowledge about the error in our interpolating polynomial in terms of some point in the middle of our interval. Since we don't even know what the point is, we won't be able to use this formula to calculate the exact error. However, if we have a bound on the derivative of the function over the entire interval, then we consequently have a bound on the error. To find the error in our numerical integration, we simply look at

$$\int_a^b e_n(x) dx.$$

There is one more minor issue we need to address at this point. We have found the error in the approximation for all points other than the interpolation points. While in theory we could think about breaking up the above integral at all of these places, it is not very practical. Fortunately, as long as the function f is well-behaved, we don't need to worry about this.

Theorem 5.11.4. *Let $f : [a,b] \to \mathbb{R}$ be $n+1$ times differentiable on $[a,b]$ with $f^{(n+1)}$ continuous on $[a,b]$. It follows that*

$$f[x_0, \ldots, x_n, x]$$

is continuous. Moreover, for all $x \in [a,b]$

$$f(x) = \sum_{j=0}^{n} f[x_0, \ldots, x_j] \prod_{i=0}^{j-1}(x - x_i) + f[x_0, \ldots, x_n, x] \prod_{i=0}^{n}(x - x_j).$$

We need to be a little careful here. Now we are looking at the divided difference as a function of x, where in the past we were simply looking at divided differences that resulted in constants. As a consequence of the above theorem, given that f has enough derivatives we can simply work with the error function $e_n(x)$ over the entire interval, without needing to worry about specific points causing problems. We also lose the restriction that the points $x_0, \ldots, x_n$ be distinct. Recall that we also need a certain number of derivatives in order to rewrite $e_n(x)$ in terms of the $(n+1)^{st}$ derivative of f. Thus, henceforth we will assume that f has the required derivatives. Before doing so we need to present one more result, the mean value theorem for integrals, which we will use to simplify some of the integrals we will be faced with.

Theorem 5.11.5 (Mean Value Theorem for Integrals)**.** *Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and $g : [a,b] \to \mathbb{R}$ be nonpositive or nonnegative. It follows that there exists $\xi \in [a,b]$ such that*

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx.$$

The requirement that g be nonpositive or nonnegative just means that $g(x)$ cannot change signs over the interval (a,b) (the value at the endpoints doesn't matter when it comes to integration). With this final result established we can move on to find the error in our approximation. Doing so we find our error is given by

$$\int_a^b e_n(x)dx = \int_a^b f[x_0, \ldots, x_n, x]\psi_n(x)dx,$$

where

$$\psi_n(x) = \prod_{i=0}^{n}(x - x_i).$$

If $\psi_n(x)$ does not change signs over (a,b), it follows from from the mean value theorem for integrals that for some $\eta \in [a,b]$

$$\int_a^b f[x_0, \ldots, x_n, x]\psi_n(x)dx = f[x_0, \ldots, x_n, \eta] \int_a^b \psi_n(x)dx.$$

It follows then from our previous results that for some $\xi \in [a,b]$

$$\int_a^b e_n(x)dx = \frac{f^{(n+1)}(\xi)}{(n+1)!} \int_a^b \psi_n(x)dx.$$

Even if $\psi_n(x)$ does change signs over (a,b), it is possible to make some simplifications. Particularly, if

$$\int_a^b \psi_n(x)dx = 0$$

then we can use the identity

$$f[x_0,\ldots,x_n,x] = f[x_0,\ldots,x_n,x_{n+1}] + f[x_0,\ldots,x_{n+1},x](x-x_{n+1})$$

to establish

$$\begin{aligned}
\int_a^b e_n(x)dx &= \int_a^b f[x_0,\ldots,x_n,x]\psi_n(x)dx \\
&= \int_a^b f[x_0,\ldots,x_n,x_{n+1}]\psi_n(x)dx + \int_a^b f[x_0,\ldots,x_{n+1},x](x-x_{n+1})\psi_n(x)dx \\
&= f[x_0,\ldots,x_n,x_{n+1}]\int_a^b \psi_n(x)dx + \int_a^b f[x_0,\ldots,x_{n+1},x](x-x_{n+1})\psi_n(x)dx \\
&= \int_a^b f[x_0,\ldots,x_{n+1},x]\psi_{n+1}(x)dx.
\end{aligned}$$

In the above simplification one needs to be careful to recognize that $f[x_0,\ldots,x_n,x_{n+1}]$ is simply a constant so it can be pulled out from the integrand. Doing so the term goes to zero by the assumption that

$$\int_a^b \psi_n(x)dx = 0.$$

Now if we can choose x_{n+1} in such a way that $\psi_{n+1}(x)$ does not change signs on (a,b) we find that for some $\xi \in [a,b]$ we have

$$\int_a^b e_n(x)dx = \frac{f^{(n+2)}(\xi)}{(n+2)!}\int_a^b \psi_{n+1}(x)dx,$$

using our previous result. Notice that the result of this process was that we increased the order of the derivative that appears. With this groundwork laid down we can finally look at analyzing the error in each of the previous cases.

For the rectangle rule $\psi_0(x) = x-a$, which is always positive on (a,b). Thus, we find the error using the first of the simplifications. We find that for some $\xi \in [a,b]$

$$E_R = f'(\xi)\int_a^b (x-a)dx = \frac{f'(\xi)(b-a)^2}{2}.$$

For the midpoint rule $\psi_0(x) = x-x_0$ where $x_0 = (a+b)/2$, which changes signs over (a,b). Thus, we cannot use the first simplification, but it turns out that

$$\int_a^b \psi_0(x)dx = \int_a^b (x-(a+b)/2)dx = 0,$$

so we can use the second of the simplifications. We simply need to choose to choose a value for x_1 so that $\psi_1(x) = (x-x_0)(x-x_1)$ does not change signs over (a,b). A natural choice is $x_1 = x_0$, as clearly $(x-x_0)^2$ does not change signs. Doing so we can use the second of the simplifications to calculate the error, and find that

$$E_M = \frac{1}{2}f''(\xi)\int_a^b \left(x - \frac{a+b}{2}\right)dx = \frac{f''(\xi)(b-a)^3}{24}.$$

For the trapezoid rule $\psi_1(x) = (x-a)(x-b)$ which does not change signs over (a, b). It follows that we can use the first simplification for the error term and doing so we find that

$$E_T = \frac{f''(\xi)}{2}\int_a^b (x-a)(x-b)dx = -\frac{f''(\xi)(b-a)^3}{12}.$$

For Simpson's rule we have $x_0 = a$, $x_1 = (a+b)/2$, and $x_2 = b$, with

$$\psi_2(x) = (x-x_0)(x-x_1)(x-x_2),$$

which changes signs over (a, b). Nevertheless, one can show that

$$\int_a^b \psi_2(x)dx = 0,$$

so we're once again reduced to the situation of needing to choose an appropriate value for x_3, in order to make it so that $\psi_3(x)$ does not change signs over (a, b). We will proceed similarly as before by letting $x_3 = x_1 = (a+b)/2$, which was the problematic term last time. When we do so we once again end up with a square, so

$$\psi_3(x) = (x-a)(x-(a+b)/2)^2(x-b)$$

which does not change signs over (a, b). Here we find that

$$E_S = \frac{f^{(4)}(\xi)}{24}\int_a^b \psi_3(x)dx = -\frac{f^{(4)}(\xi)[(b-a)/2]^5}{90}.$$

The reason for including the factor of 2 with $(b-a)$ is so that we can think of the term $(b-a)/2$ as representing not the length of the interval of integration, but the distance between points of interpolation. The reason for this will become clear when we look at composite rules in the next section.

It is worth emphasizing again that when using these methods we don't know where the point ξ lies, so we are usually not able to calculate the actual error. Nevertheless, if we have a bound on the appropriate derivative over the entire interval, then we know that the value of the derivative at evaluated at ξ must be less than the given bound, so in that way we are able to find a bound on the amount of error in our approximation. The bound only tells us how bad our approximation can get, but doesn't tell us our how close approximation actually gets. It is possible in some specific cases that a method such as the rectangle rule will provide a better estimate than Simpson's rule (in fact, it is very easy to come up with examples where any of the approximations will be exact), but in general Simpson's rule should be preferred because it has a much tighter bound on error.

Although we generally cannot find the exact error, there is a special case in which we can - when the error is 0. Namely, when the required derivative of the function f is identically 0 over the entire interval $[a, b]$, then it must also be 0 at ξ, so it follows that our approximation is exact. The simplest example functions we can look at are polynomials. For a general polynomial

$$q_n(x) = a_0 + a_1x + \ldots + a_nx^n$$

we have

$$\frac{d^n}{dx}q_n(x) = q_n^{(n+1)}(x) = 0.$$

It follows then that the midpoint and trapezoid rules are exact for first-order polynomials, and Simpson's rule is exact for third-order polynomials. With a bit of ingenuity one can derive rules for numerical integration that are exact for polynomials up to an arbitrary degree. These rules are called Gaussian rules, but are out of the scope of our current discussion.

5.12 Numerical Integration - Composite Rules

In the previous section we spent a considerable of effort to establish a number of rules for numerical integration. In general, a single interpolating polynomial approximating our function f over the entire interval $[a, b]$ will provide quite a poor approximation. In order to overcome this issue we simply need to partition our interval $[a, b]$ into a number of subintervals, with break points

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b,$$

and write

$$\int_a^b f(x)dx = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x)dx.$$

Having done so we approximate the integral over each subinterval using one of the previously established methods. In general there is no requirement on the spacing of the points chosen to break our larger interval into subintervals. With a bit of information about the function f and some ingenuity, one can increase one's accuracy by using more complicated partitioning schemes. However, for our current purposes, and the sake of simplicity, we will work with subintervals of equal length (see figure 5.9). Doing so we let

$$x_i = a + ih \quad i \in \{0, \ldots, n\},\ h = \frac{b-a}{n}.$$

Rather than handling the approximation and error corresponding to each rule of numerical integration separately as we did in the previous section, we can work with them together, noting that by adding the approximation and the error term together we find the exact value of the integral. In other words, over each subinterval we will find that

$$\int_{x_{i-1}}^{x_i} f(x)dx = \int_{x_{i-1}}^{x_i} p_n(x)dx + \int_{x_{i-1}}^{x_i} e_n(x)dx.$$

Since the integral over $[a, b]$ is just found as a sum of the integrals on the left-hand side of the above equation, we will find the approximation to be the sum of the approximation terms, and the error to be the sum of the error terms. Beginning with the rectangle rule we find

$$\int_{x_{i-1}}^{x_i} f(x)dx = (x_i - x_{i-1})f(x_{i-1}) + \frac{f'(\xi_i)(x_i - x_{i-1})^2}{2} = hf(x_{i-1}) + \frac{f'(\xi_i)h^2}{2}.$$

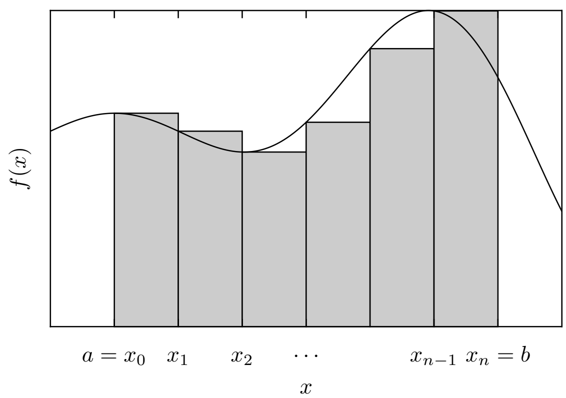
Summing over these terms we find

$$\int_a^b f(x)dx \approx h\sum_{i=1}^{n} f(a + (i-1)h) = \frac{b-a}{n}\sum_{i=1}^{n} f\Big(a + (i-1)\frac{b-a}{n}\Big),$$

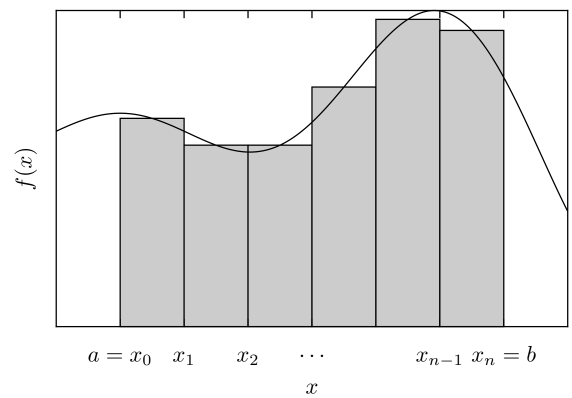
for the composite rectangle rule with n subintervals of length $h = (b-a)/n$. In order to simplify the error terms, we need to use the following result.

Theorem 5.12.1. *Let $f : [a, b] \to \mathbb{R}$ be continuous, $x_1, \ldots, x_n \in [a, b]$, and $g_1, \ldots, g_n \in \mathbb{R}$, where all $g_i \geq 0$ or $g_i \leq 0$. It follows that for some $\xi \in [a, b]$*

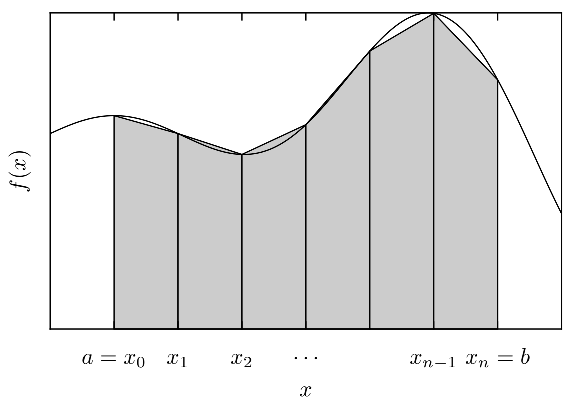
$$\sum_{i=1}^{n} f(x_i)g_i = f(\xi)\sum_{i=1}^{n} g_i.$$



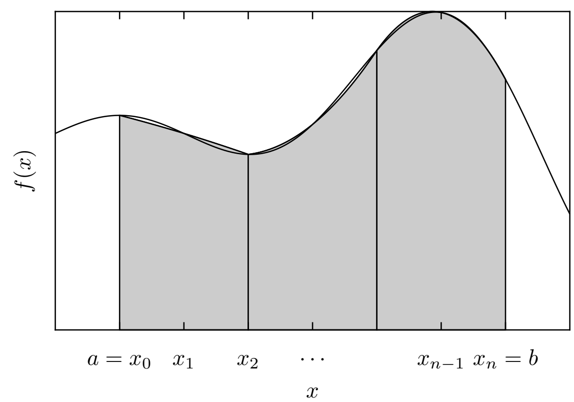
(a) Rectangle Rule.



(b) Midpoint Rule.



(c) Trapezoid Rule.



(d) Simpson's Rule.

Figure 5.9: Composite rules.

This result is essentially a finite case of the mean value theorem for integrals, and is useful to us here because it allows us to combine the terms involving the derivative of f over each of the subintervals into a single term over the interval $[a.b]$ (moving from a number of points $\xi_i \in [x_{i-1}, x_i]$ to a single point $\xi \in [a, b]$). As said before, the error is simply the sum of the previous error terms, so we find

$$E_{R,n} = \sum_{i=1}^{n} \frac{f'(\xi_i)h^2}{2} = f'(\xi)\sum_{i=1}^{n} \frac{h^2}{2} = \frac{f'(\xi)nh^2}{2} = \frac{f'(\xi)(b-a)^2}{2n}.$$

In a similar fashion we work with the midpoint rule, and find

$$\int_{x_{i-1}}^{x_i} f(x)dx = (x_i - x_{i-1})f\Big(\frac{x_{i-1}+x_i}{2}\Big) + \frac{f''(\xi_i)(b-a)^3}{24} = hf\Big(\frac{x_{i-1}+x_i}{2}\Big) + \frac{f''(\xi_i)h^3}{24}.$$

Summing over these terms we find

$$\int_a^b f(x)dx \approx \frac{b-a}{n} \cdot \sum_{i=1}^{n} f\Big(a + \Big(i - \frac{1}{2}\Big)\frac{b-a}{n}\Big),$$

for the composite midpoint rule. Once again we sum the error terms, and find that

$$E_{M,n} = \sum_{i=1}^{n} \frac{f''(\xi_i)h^3}{24} = \frac{f''(\xi)nh^3}{24} = \frac{f''(\xi)(b-a)^3}{24n^2}.$$

Moving along to the trapezoid rule we find

$$\int_{x_{i-1}}^{x_i} f(x)dx = \frac{1}{2}(x_i - x_{i-1})(f(x_{i-1}) + f(x_i)) - \frac{f''(\xi_i)(x_i - x_{i-1})^3}{12} = \frac{h}{2}(f(x_{i-1}) + f(x_i)) - \frac{f''(\xi_i)h^3}{12}.$$

It is slightly more complicated when we sum these terms, because $a = x_0$ and $b = x_n$ will both only be summed once, but all of the intermediate terms will be included twice in the overall sum. Keeping this in mind we find

$$\int_a^b f(x)dx \approx \frac{b-a}{2n}\sum_{i=1}^{n}(f(x_{i-1}) + f(x_i)) = \frac{b-a}{2n}\Big(f(a) + \sum_{i=1}^{n-1} f\Big(a + i\frac{b-a}{n}\Big) + f(b)\Big)$$

for the composite trapezoid rule. In this case the error is given by

$$E_{M,n} = \sum_{i=1}^{n}\Big(-\frac{f''(\xi_i)h^3}{12}\Big) = -\frac{f''(\xi)nh^3}{12} = -\frac{f''(\xi)(b-a)^3}{12n^2}.$$

Lastly, we work with Simpson's rule, finding

$$\int_{x_{i-1}}^{x_i} f(x)dx = \frac{h}{6}\Big(f(x_{i-1}) + 4f\Big(\frac{x_{i-1}+x_i}{2}\Big) + f(x_i)\Big) - \frac{f^{(4)}(\xi_i)(h/2)^5}{90}.$$

We have to be a little bit cautious here. Unlike the previous methods, we are actually using three points for each approximation with Simpson's rule. In order to have a fair comparison of Simpson's rule with the other methods we want the number of interpolation points to be the same. In order to keep the number of interpolation points consistent with the other composite rules, we think about looking at n subintervals, where each *subinterval* is of length

$$h = \frac{2(b-a)}{n}.$$

In this way each of our *subintervals* for Simpson's rule is twice the width of a subinterval for the other methods, but we break each of these *subintervals* in half into two subintervals when we interpolate at the midpoint. Making this adjustment imposes the condition that we must choose n to be even, because otherwise the last of our new larger *subintervals* would not coincide with the endpoint (it would overlap it by the width of one subinterval).

When we sum the approximation terms, the first and last term $f(a)$ and $f(b)$ will both only be counted once, but the intermediate terms are slightly more complex. Each term with a 4 in front of it will be counted only once, and the other terms will be counted twice. Thus, we get something of the form

$$f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \ldots + f(b).$$

Actually summing all the terms together the result is

$$\int_a^b f(x)dx \approx \frac{b-a}{3n} \cdot \Big(f(a) + 4\sum_{i=1}^{n/2} f\Big(a + (2i-1)\frac{b-a}{n}\Big) + 2\sum_{i=1}^{n/2-1} f\Big(a + 2i\frac{b-a}{n}\Big) + f(b)\Big).$$

Note here that although our sums only extend to $n/2$ because we are considering *subintervals* of twice the width. Nevertheless, the terms inside are multiplied by $2i$, so we do in fact cover all of our subintervals. Considering the error in the estimate we find

$$E_{S,n} = \sum_{i=1}^{n/2} \Big(- \frac{f^{(4)}(\xi_i)(h/2)^5}{90}\Big) = -\frac{f^{(4)}(\xi)n(h/2)^5}{180} = -\frac{f^{(4)}(\xi)(b-a)^5}{180n^4}.$$

Note that the way we redefined h above required us only to sum up to $n/2$, because each of our *subintervals* is twice the width of the subintervals we used to find the error in Simpson's rule. Here the reason for writing the error in terms of $h/2$ also becomes clear, because $h/2$ gives the distance between the interpolation points, which is really more important than the distance between subintervals.

Example 1. Estimate

$$\ln(2) = \int_1^2 \frac{1}{x}dx$$

using the composite rectangle, midpoint, trapezoid, and Simpson's rules, with $n = 10$ subintervals.

Solution Since our interval of interest is $[1, 2]$, and we have 6 subintervals, we will have the evaluation points $1, 1.1, 1.2, \ldots, 1.9, 2$. For the rectangle rule we find

$$\ln(2) \approx 0.1 \cdot \Big(\frac{1}{1} + \frac{1}{1.1} + \frac{1}{1.2} + \frac{1}{1.3} + \frac{1}{1.4} + \frac{1}{1.5} + \frac{1}{1.6} + \frac{1}{1.7} + \frac{1}{1.8} + \frac{1}{1.9}\Big) = 0.718771403.$$

For the midpoint rule we find

$$\ln(2) \approx 0.1 \cdot \Big(\frac{1}{1.05} + \frac{1}{1.15} + \frac{1}{1.25} + \frac{1}{1.35} + \frac{1}{1.45} + \frac{1}{1.55} + \frac{1}{1.65} + \frac{1}{1.75} + \frac{1}{1.85} + \frac{1}{1.95}\Big) = 0.69283536.$$

Here the trapezoid rule provides us with

$$\ln(2) \approx 0.05 \cdot \Big(\frac{1}{1} + 2 \cdot \Big(\frac{1}{1.1} + \frac{1}{1.2} + \frac{1}{1.3} + \frac{1}{1.4} + \frac{1}{1.5} + \frac{1}{1.6} + \frac{1}{1.7} + \frac{1}{1.8} + \frac{1}{1.9}\Big) + \frac{1}{2}\Big) = 0.693771403.$$

Finally, Simpson's rule gives the estimate of

$$\ln(2) \approx \frac{1}{30} \cdot \Big(\frac{1}{1} + \frac{4}{1.1} + \frac{2}{1.2} + \frac{4}{1.3} + \frac{2}{1.4} + \frac{4}{1.5} + \frac{2}{1.6} + \frac{4}{1.7} + \frac{2}{1.8} + \frac{4}{1.9} + \frac{1}{2}\Big) = 0.693150231.$$

For a point of reference, the actual value of ln(2) is given by

$$\ln(2) \approx 0.693147181.$$

In this case it turns out that the midpoint rule gives a better estimate than the trapezoid rule. Simpson's rule gives the best approximation, accurate up to 4 decimal places.

Example 2. Find an upper bound on the error that results from approximating ln(2) with $n = 10$ using the composite rectangle, midpoint, trapezoid, and Simpson's rules.

Solution In order to find the upper bound on the error we're going to need to take a number of derivatives. Letting $f(x) = 1/x$ we find

$$f'(x) = -\frac{1}{x^2}, \quad f''(x) = \frac{2}{x^3}, \quad f'''(x) = -\frac{6}{x^4}, \quad f^{(4)}(x) = \frac{24}{x^5}.$$

Here we're in luck because for all of these derivatives the maximum value (in magnitude) occurs at the left endpoint $x = 1$, saving us the work of looking for their maxima. For the rectangle rule

$$E_{R,10} \leq \frac{1}{20} \cdot \left|\frac{1}{1^2}\right| = 0.05.$$

For the midpoint rule

$$E_{M,10} \leq \frac{1}{2400} \cdot \left|\frac{2}{1^3}\right| = \frac{1}{1200} \approx 0.000833.$$

For the trapezoid rule

$$E_{T,10} \leq \frac{1}{1200} \cdot \left|\frac{2}{1^3}\right| = \frac{1}{600} \approx 0.00166.$$

For Simpson's rule

$$E_{S,10} \leq \frac{1}{1800000} \cdot \left|\frac{24}{1^5}\right| = \frac{4}{300000} \approx 1.33 \cdot 10^{-5}.$$

It's clear to see that Simpson's rule guarantees a much tighter bound on the error than the others in this case. Just for reference, we can calculate the actual error in each case, finding that for the rectangle rule

$$E_{R,10} \approx 0.693147181 - 0.718771403 = -0.02562.$$

For the midpoint rule

$$E_{M,10} \approx 0.693147181 - 0.69283536 = -0.000311821.$$

For the trapezoid rule

$$E_{T,10} \approx 0.693147181 - 0.693771403 = -0.000624222.$$

Finally, for Simpson's rule

$$E_{S,10} \approx 0.693147181 - 0.693150231 = -3.05 \cdot 10^{-6}.$$

In looking at the actual error we can see that it is consistent with the upper bounds of error.

5.13 Improper Integrals - Infinite Limits of Integration

By virtue of the way the Riemann integral is defined, there are two conditions we must be aware of when finding the area underneath a curve. The first is that the Riemann integral is only defined to find the area underneath a curve over a finite interval - some $[a, b]$. The second is that the function we wish to integrate must be bounded over the limits of integration. In dealing with an unbounded function the notion of the upper sum no longer makes sense, so we cannot Riemann integrate such functions. Fortunately, there are methods for getting around both of these problems, at least to an extent. In order to do so we extend the notion of the integral, and consider what are called improper integrals. In this section we will deal with infinite limits of integration, and we will deal with unbounded integrands in the next.

The motivation for finding the area underneath a curve of over a domain of infinite length arises primarily from statistics. When we think of performing an experiment, we can either have a discrete or continuous set of outcomes (possible results of the measurement). For instance, if we roll a six-sided die, the set of outcomes is

$$\{1, 2, 3, 4, 5, 6\},$$

which has a finite number of members. This is a discrete set of outcomes. In contrast, we can have an experiment where the set of outcomes is continuous (or so large that it is reasonable to model it as being continuous). Think about measuring the position of a particle, restricted to one-dimensional space for simplicity. In this situation we have an entire range of possible positions for the particle, so the set of possible outcomes is continuous (imagine trying to write down each possible position). In the discrete case it is possible to assign a probability to each individual event - it would be $(1/6)^{th}$ in the above case. However, in the case of a continuous set of outcomes the probability of a single event occurring is generally meaningless. In all likelihood, each individual outcome has a probability of 0 of occurring (because there are so many possibilities). This forces us to take a different approach for dealing with a continuous set of outcomes - probability distributions.

A probability distribution gives us a means of describing the probability of events occurring, given a continuous set of outcomes. In order to do so we need to look not at individual events (which generally all have probability 0 - except in some exceptional cases), but ranges of events. In order to find the probability of the result of an experiment falling under a certain range of possible outcomes, we consider the area underneath the probability distribution over that interval. For instance, in quantum mechanics the position of a particle is described by a wave function ψ, which gives the probability of finding a particle in different positions (in truth it is $|\psi|^2$ that is the probability distribution function, but this is merely a technical matter). By integrating this probability distribution over a given interval (for instance a 1 meter radius, for a one-dimensional particle) one computes the probability of finding the particle within that interval (a 1 meter radius in this case). In a sense, the area underneath any single point is going to be 0, but when we look at a range of values there is some area, so we are able to make at least some predictions about the outcome of an experiment.

Having some background of probability distributions we can now discuss why their significance to improper integrals. Since a probability distribution represents the entire set of outcomes, the area underneath the entire probability distribution had better be one, meaning that the result of our experiment lies somewhere in the set of possible outcomes. This condition is called the normalization of a probability distribution. In order to make sure our probability distribution is normalized, we need to be able to calculate the area underneath the entire distribution. If the domain of our probability distribution is the entire real numbers, then we need to be able to integrate over the entire real numbers. Our current definition of the Riemann integral does not allow for this; hence the motivation for improper integrals.

Unfortunately, we cannot simply discard the requirement of the Riemann integral of using an interval of finite length. However, we can work around this problem with a clever usage of limits. The idea is as follows. We begin by evaluating the integral over the limits $[a, b]$, and then consider a limit as $b \to \infty$. In this way we are always evaluating an integral over a finite interval, but we are approximating as accurately as we'd like what the area would be if we *could* integrate over an interval of infinite length.

Throughout this section we will need to impose the condition that our functions be bounded over any finite interval we wish to integrate over. This will allow us to use Riemann integration. This requirement does not demand however that our function be bounded over the entire real numbers. Thus, we can use improper integrals with infinite limits to evaluate integrands such as x, because even though

$$\lim_{x\to\infty} x = \infty$$

which is unbounded, if we look over any finite limits of integration, the function x will be bounded, and since we are really only integrating over finite limits at any given time, that is all we need. Although not explicitly stated, this is a requirement for the theorems in this section. For convenience, we usually do not write the limit explicitly.

Definition 5.13.1 (Improper integrals)**.** We use the following notation for improper integrals.

$$\begin{aligned}
\int_a^\infty f(x)dx &= \lim_{b\to\infty} \int_a^b f(x)dx. \\
\int_{-\infty}^b f(x)dx &= \lim_{a\to-\infty} \int_a^b f(x)dx. \\
\int_{-\infty}^\infty f(x)dx &= \lim_{a\to-\infty} \int_a^c f(x)dx + \lim_{b\to\infty} \int_c^b f(x)dx,\ c \in \mathbb{R}.
\end{aligned}$$

Consistent with our previous notions of limits, if the limit exists we say the integral converges, and if the limit does not exist we say the integral diverges. Let's consider some examples.

Example 1. Suppose chemical production is governed by the differential equation

$$\frac{dP}{dt} = e^{-\alpha t},\ \alpha > 0, \alpha \in \mathbb{R},$$

moles per second. If this experiment were allowed to run indefinitely, how much chemical would be produced?

Solution To find the amount of chemical produced as a function of time we need to solve for $P(t)$ (conveniently in this initial value problem our initial value is $P(0) = 0$). Since we're dealing with time our lower limit of integration will be 0. We find that

$$\int_0^\infty e^{-\alpha t}dt = \lim_{b\to\infty} \int_0^b e^{-\alpha t}dt.$$

Evaluating the integral over finite limits we find

$$P(b) = \int_0^b e^{-\alpha t}dt = -\frac{e^{-\alpha t}}{\alpha}\Big|_0^b = \frac{-e^{-\alpha b} - (-e^{-0})}{\alpha} = \frac{1 - e^{-\alpha b}}{\alpha}.$$

Now introducing the limit we find

$$\lim_{b\to\infty} P(b) = \int_0^\infty e^{-\alpha t}dt = \lim_{b\to\infty} \frac{1-e^{-\alpha b}}{\alpha} = \frac{1}{\alpha}.$$

Here the smaller α is, the slower the exponential function decays. Thus, as we increase α the amount of chemical produced decreases. Nevertheless, even if production continues indefinitely, we will always produce a finite amount of chemical.

Example 2. Suppose chemical production is governed by

$$\frac{dQ}{dt} = \frac{1}{1+t}$$

moles per second. How much chemical is generated if production continues indefinitely, beginning from $t = 0$?

Solution Once again, we write

$$\int_0^\infty \frac{1}{1+t}dt = \lim_{b\to\infty} \int_0^b \frac{1}{1+t}dt.$$

Using u substitution, with $u = 1+t$, so $du = dt$, we find

$$\int_0^b \frac{1}{1+t}dt = \int_1^{b+1} \frac{1}{u}du = \ln|u|\Big|_1^{b+1} = \ln(b+1) - \ln(1) = \ln(b+1).$$

Thus,

$$\int_0^\infty \frac{1}{1+t}dt = \lim_{b\to\infty} \ln(b+1) = \infty.$$

In this situation, if production is allowed to continue indefinitely, the amount produced grows without bound. In other words, this improper integral diverges.

What is it that differs between the integrands of these two integrals that causes one to converge and the other to diverge? In both of these cases the integrands are always positive over the limits of integration. Furthermore, they both approach 0 as $b \to \infty$. The difference is that e^{-t} decays much more quickly than $1/(1+t)$. Based on this observation, we should be able to generalize whether some simple functions will converge or diverge, based on the rate at which their integrands approach 0 (for positive functions).

We should emphasize that if the integrand does not decay to zero, then it is guaranteed the integral will diverge. As a simple example, think of

$$\lim_{b\to\infty} \int_0^b dx = \lim_{b\to\infty} x\Big|_0^b = \lim_{b\to\infty} b = \infty.$$

This shouldn't be surprising, because as we continue to pick up area underneath a rectangle of height one, the amount of area continues to increase without bound. We run into a similar situation with oscillating functions that also do not decay to 0. Consider

$$\int_0^\infty \cos(x)dx = \lim_{b\to\infty} \int_0^b \cos(x)dx = \lim_{b\to\infty} \sin(x)\Big|_0^b = \lim_{b\to\infty} \sin(b).$$

Here we have a limit that does not exist. If we were to break up the integral into any finite interval (no matter how large) the amount of area would be finite, because the positive and negative areas would cancel each other every 2π due to the nature of cosine. However, when we look as $b \to \infty$ we don't have any logical stopping point. Thus, the amount of area picked up keeps fluctuating between 0 and 2, not settling on a single value. For this reason the integral diverges. In conclusion, if the integrand of an improper integral does not decay as the function approaches ∞, it is impossible for the integral to converge, so there is no point in evaluating the integral, as we already know the result.

We have already seen that the integral over $(0, \infty)$ converges for any decaying exponential. It is also useful to consider functions of the form $1/t^p$. We will begin integrating at $t = 1$, to avoid division by 0. If $p = 1$, then we are considering the integral

$$\int_1^\infty \frac{1}{t} dt.$$

which we already observed diverges (because it is strictly larger than the integral we evaluated previously). For $p \neq 1$, we use the power rule so

$$\int_1^b \frac{1}{t^p} dt = \frac{t^{1-p}}{1-p}\Big|_1^b = \frac{b^{1-p}}{1-p} - \frac{1}{1-p}.$$

When we consider the limit as $b \to \infty$, we must consider whether $p > 1$ or $p < 1$. First recall that

$$\lim_{b\to\infty} b^p = \infty \quad \text{for } p > 0, \quad \text{and} \quad \lim_{b\to\infty} b^p = 0 \quad \text{for } p < 0.$$

For $p < 1$

$$\lim_{b\to\infty} \left(\frac{b^{1-p}}{1-p} - \frac{1}{1-p} \right) = \infty,$$

because $1 - p > 0$ so that the first term grows without bound. In the case that $p > 1$

$$\lim_{b\to\infty} \left(\frac{b^{1-p}}{1-p} - \frac{1}{1-p} \right) = \frac{1}{p-1},$$

because $1 - p < 0$ so that the first term goes to zero. In conclusion

$$\int_1^\infty \frac{1}{t^p} = \begin{cases} \infty & p \leq 1 \\ 1/(p-1) & p > 1. \end{cases}$$

It is noteworthy that for an improper integral, moving the lower limit of integration a finite amount will not alter the integral's convergence or divergence, as long as it does not introduce a portion of the function that is unbounded. This means that we can already gather a lot of information about the convergence and divergence of other improper integrals. For example,

$$\int_5^\infty \frac{1}{\sqrt{t}} dt = \int_1^\infty \frac{1}{\sqrt{t}} dt - \int_1^5 \frac{1}{\sqrt{t}} dt$$

using the summation property for integrals. We know that

$$\int_1^\infty \frac{1}{\sqrt{t}} dt$$

diverges, and that

$$\int_1^5 \frac{1}{\sqrt{t}} dt$$

is just some finite amount. If we subtract some finite amount from a diverging integral, the result will still be something that diverges. Thus, without any computation we can deduce that

$$\int_5^{\infty} \frac{1}{\sqrt{t}} dt$$

diverges. In a sense, this means that the function's behavior for small input values has no influence over the convergence of such an integral - convergence is related solely to the rate at which the function decays to zero as inputs grow larger.

Now suppose that we are faced with a more complicated integral, something like

$$\int_1^{\infty} \left(\frac{1}{t} + e^{-t}\right) dt.$$

We can write

$$\int_1^{\infty} \left(\frac{1}{t} + e^{-t}\right) dt = \int_1^{\infty} \frac{1}{t} dt + \int_1^{\infty} e^{-t} dt.$$

Since $\int_1^{\infty} e^{-t} dt$ is just a positive number, we can deduce that

$$\int_1^{\infty} \left(\frac{1}{t} + e^{-t}\right) dt > \int_1^{\infty} \frac{1}{t} dt = \infty.$$

Thus, we can conclude that our integral diverges, since it is larger than an integral diverging to ∞. A very similar idea to this one leads us to the comparison test.

Theorem 5.13.1 (Direct Comparison Test for Improper Integrals)**.** *Consider f and g with $0 \leq f(x) \leq g(x)$ for all $x \in (a, b)$, $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$. It follows that*

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx,$$

and hence

1. $\displaystyle\int_a^b f(x) dx$ *converges if* $\displaystyle\int_a^b g(x) dx$ *converges.*
2. $\displaystyle\int_a^b g(x) dx$ *diverges if* $\displaystyle\int_a^b f(x) dx$ *diverges.*

The above theorem is written to encapsulate all three types of improper integrals: starting from $-\infty$, ending at ∞, or over the entire real numbers. In essence, the above theorem states that if some positive function $f(x)$ is always less than or equal to another positive function $g(x)$, then its integral will be less, so if $g(x)$ converges, then $f(x)$ must converge as well. Similarly, if $f(x)$ diverges, and $g(x)$ is greater than or equal to it, then it must also diverge, as the integral will be greater. In the light of the previous discussion it is worth emphasizing that we really only need to use the comparison test to look at the tail of the functions of interest - for a bounded function its area underneath a finite interval is immaterial in determining its convergence.

Example 3. Determine whether or not $\displaystyle\int_1^{\infty} \frac{1}{t + e^t} dt$ converges or diverges.

Solution First we can note that because the integrand is always positive, the integral must be greater than 0. Also, for all $t > 0$ we have

$$\frac{1}{t + e^t} \leq \frac{1}{e^t}$$

so by the comparison test, we have that

$$0 \le \int_1^\infty \frac{1}{t+e^t} \le \int_1^\infty \frac{1}{e^t} = 1$$

Although we don't know to what exact value, we can conclude that this integral converges, and its value is between 0 and 1.

Example 4. Determine whether or not

$$\int_0^\infty \frac{1}{2x+2} dx$$

converges or diverges. If it converges, find to what value.

Solution This function looks like $\frac{1}{x}$, which is divergent, so we suspect that this integral should also diverge. However, it is not clear how to use the comparison test in this case, so let us rewrite

$$\int_0^\infty \frac{1}{2x+2} dx = \lim_{b\to\infty} \int_0^b \frac{1}{2x+2} dx.$$

Now, using substitution $u = 2x + 2$, so $du/2 = dx$, and

$$\int_0^b \frac{1}{2x+2} dx = \frac{1}{2} \int_2^{2b+2} \frac{1}{u} du.$$

Now that this integral is written in a more familiar form, we can see that it diverges, using the fact that

$$\int_1^\infty \frac{1}{u} du = \infty.$$

Example 5. Determine whether or not

$$\int_2^\infty \frac{1}{(x-1)^{1/2}} dx$$

converges or diverges. If it converges, find to what value.

Solution Using the comparison test,

$$\frac{1}{(x-1)^{1/2}} > \frac{1}{x^{1/2}}$$

where the latter of these terms is the integrand for a divergent integral under these limits. Thus, the integral in question diverges.

Example 6. Determine whether or not

$$\int_2^\infty \frac{2}{(3x-5)^2} dx$$

converges or diverges. If it converges, find to what value.

Solution This integrand has the form of $1/x^2$, so we suspect that it should converge. Let us use the substitution $u = 3x - 5$ so $du/3 = dx$. Evaluating the indefinite integral

$$\int \frac{2}{(3x-5)^2} dx = \frac{2}{3} \int u^{-2} du = -\frac{2}{3} \cdot \frac{1}{u} + c = -\frac{2}{3} \cdot \frac{1}{3x-5} + c.$$

Now we see that

$$\int_2^b \frac{2}{(3x-5)^2}dx = -\frac{2}{3}\cdot\frac{1}{3x-5}\Big|_2^b = -\frac{2}{3}\cdot\frac{1}{3b-5}+\frac{2}{3}.$$

Thus, we finally conclude

$$\int_2^\infty \frac{2}{(3x-5)^2}dx = \lim_{b\to\infty}\Big(-\frac{2}{3}\cdot\frac{1}{3b-5}+\frac{2}{3}\Big) = \frac{2}{3}.$$

In addition to using the direction comparison test to determine convergence and divergence of improper integrals, there is another useful result.

Theorem 5.13.2 (Limit Comparison Test for Improper Integrals)**.** *Let f and g be positive functions, If*

$$\lim_{x\to\infty}\frac{f(x)}{g(x)} = L,\ 0 < L < \infty,$$

then

$$\int_a^\infty f(x)dx \quad \textit{and} \quad \int_a^\infty g(x)dx$$

either both converge or diverge.

The limit comparison test will be particularly useful for cases where we have functions that behave like a function we know converges, but actually may be a little bit larger. For instance, If we want to evaluate an integral of the form

$$\int_1^\infty \frac{1}{x^2-1}dx,$$

we cannot use the direct comparison test, because

$$\frac{1}{x^2-1} > \frac{1}{x^2},$$

and the improper integral on the right actually converges. Thus, the direct comparison test tells us nothing because we only have an integrand larger than one we know converges. Nevertheless, using the limit comparison test

$$\lim_{x\to\infty}\frac{1/(x^2-1)}{1/x^2} = \lim_{x\to\infty}\frac{x^2}{x^2-1} = 1.$$

It then follows that both integrals converge, so

$$\int_1^\infty \frac{1}{x^2-1}dx < \infty.$$

There is one more result that can be useful as far as determining the convergence of improper integrals is concerned. It is a very simple and intuitive result, but we present it here for completeness.

Theorem 5.13.3 (Absolute Comparison Test for Improper Integrals)**.** *Let $a \in \mathbb{R}\cup\{-\infty\}$, $b \in \mathbb{R}\cup\{\infty\}$. It follows that*

$$\Big|\int_a^b f(x)dx\Big| \le \int_a^b |f(x)|dx,$$

so namely if the integral on the right converges, so does

$$\int_a^b f(x)dx.$$

5.14 Improper Integrals - Infinite Integrands

In the previous section we considered integrals where we wanted to look at infinite limits of integration. The second issue that can arise is when we want to integrate a function that is not bounded over the limits of integration. This is a subtle issue, but consider the integral

$$\int_{-1}^{2} \frac{1}{x^2} dx.$$

If we just blindly integrate, applying the power rule, we will find

$$\int_{-1}^{2} \frac{1}{x^2} dx = -\frac{1}{x}\Big|_{-1}^{2} = -\frac{1}{2} - 1 = -\frac{3}{2}.$$

However, note that $1/x^2$ is always positive over the limits of integration of $[-1, 2]$, so this clearly cannot be right. The problem here is that $1/x^2$ is not bounded over the interval $[-1, 2]$, so it is not Riemann integrable. Nevertheless, we may still be interested in defining some notion of area underneath a function, even if it is not Riemann integrable. In this way we need to extend the Riemann integral, to handle this second type of improper integral. Here the idea is going to be very similar as before. We'll want to look at our function over an interval where it is bounded, and use a limit to approximate as accurately as we'd like what the area would look like if we could integrate over the interval where it is unbounded.

To determine whether or not an function f is bounded over finite limits of integration $[a, b]$, we simply need to identify if there are any points $c \in [a, b]$ where

$$\lim_{x \nearrow c} |f(x)| = \infty \quad \text{or} \quad \lim_{x \searrow c} |f(x)| = \infty.$$

We look at one-sided limits because it is very possible for our function to be unbounded from one side and bounded from the other.

Definition 5.14.1 (Improper Integrals - Unbounded Integrand)**.** Let f be bounded on $[a, b]$, except for at most one $x_0 \in [a, b]$, where

$$\lim_{x \to x_0} |f(x)| = \infty.$$

We find the improper integral of $f(x)$ over $[a, b]$ as

$$\int_a^b f(x)dx = \lim_{c \nearrow x_0} \int_a^c f(x)dx + \lim_{d \searrow x_0} \int_d^b f(x)dx.$$

Note that the above definition only applies if we have a single point where the function of interest approaches ∞. If there are multiple such points, then we simply need to break the original limits of integration into a number of subintervals, wherein each subinterval there is only a single point where

$$\lim_{x \to x_0} |f(x)| = \infty.$$

It's also worth noting that above the definition is constructed to handle a point x_0 inside the interior of the interval, (a, b). However, if x_0 lies at either a or b, then one of the two integrals can simply be ignored, as the integration will take place over 0 limits, so no area will be accumulated by it.

Example 1. Evaluate $\int_0^1 x^{-1/2}dx$

Solution We find that

$$\int_0^1 x^{-1/2}dx = \lim_{d\searrow 0}\int_d^1 x^{-1/2}dx = \lim_{d\searrow 0} 2\sqrt{x}\Big|_d^1 = \lim_{d\searrow 0}(2-2\sqrt{d}) = 2.$$

Example 2. Evaluate $\int_0^b \frac{dt}{t}$.

Solution We find that

$$\int_0^b \frac{dt}{t} = \lim_{d\searrow 0}\int_d^b \frac{dt}{t} = \lim_{d\searrow 0} \ln|t|\Big|_d^b = \lim_{d\searrow 0}(\ln(b)-\ln(d)) = \infty.$$

Here $\ln(b)$ is just some number, which is completely dominated by the other term which grows without bound.

Just as with infinite limits of integration, we can look at the general class of functions $1/t^p$, and generalize our result for

$$\int_0^b \frac{dt}{t^p}.$$

We've seen that if $p = 1$ then this integral diverges. If $p \neq 1$ then

$$\int_0^b \frac{dt}{t^p} = \lim_{d\searrow 0}\int_d^b \frac{dt}{t^p} = \lim_{d\searrow 0} \frac{t^{1-p}}{1-p}\Big|_d^b = \lim_{d\searrow 0}\Big(\frac{b^{1-p}}{1-p} - \frac{d^{1-p}}{1-p}\Big).$$

If $p < 1$, the second term vanishes, so integral converges. If $p > 1$, the second term approaches $-\infty$, so the integral diverges. We can summarize these results as

$$\int_0^b \frac{dt}{t^p} = \begin{cases} \infty & p \geq 1 \\ \frac{b^{1-p}}{1-p} & p < 1 \end{cases}$$

It's worth noting that the situation here is opposite from when we considered $1/t^p$ over infinite limits of integration. In that case we had convergence for $p > 1$, and divergence for $p \leq 1$. Nevertheless, we can take this analysis one step further. Having evaluated this integral for a general bound b, we can consider the limit as $b \to \infty$. For $p \geq 1$ the result is clearly still ∞. For $p < 1$ we have

$$\int_0^\infty \frac{dt}{t^p} = \lim_{b\to\infty} \frac{b^{1-p}}{1-p} = \infty,$$

as well, because we are raising b to a positive power. In conclusion,

$$\int_0^\infty \frac{dt}{t^p} = \infty$$

for any value of p. Noteworthy however is the reason for divergence. When we have $p > 1$ the tail of the function decays rapidly enough that it does not cause problems, but it causes the function to grow much more quickly for small t, leading to an unbounded area. On the other hand, when we have $p < 1$ the beginning of the function grows slow enough to provide a finite area, but the tail of the function decays too slowly.

The direct comparison test we used in the previous section is based on the very simple observation that

$$|f(x)| \leq |g(x)| \text{ for } x \in (a,b) \quad \Longrightarrow \quad \int_a^b |f(x)|dx \leq \int_a^b |g(x)|dx.$$

We can use this very simple observation once again to make comparisons for integrals with unbounded integrands.

Example 3. Determine whether or not $\displaystyle\int_0^1 \frac{dt}{\sqrt{t}+\sqrt[3]{t}}$ converges or diverges.

Solution Here we will use the direct comparison test. For $0 \leq t \leq 1$ we have

$$\sqrt[3]{t} > \sqrt{t}$$

so that

$$\frac{1}{\sqrt{t}+\sqrt[3]{t}} < \frac{1}{\sqrt{t}+\sqrt{t}} = \frac{1}{2\sqrt{t}}.$$

Thus,

$$\int_0^1 \frac{dt}{\sqrt{t}+\sqrt[3]{t}} \leq \frac{1}{2}\int_0^1 \frac{dt}{\sqrt{t}} = \frac{1}{2}\cdot 2 = 1.$$

Finally, before closing let us return to the original integral we began with, from a more sophisticated perspective.

Example 4. Evaluate $\displaystyle\int_{-1}^2 \frac{dx}{x^2}$.

Solution Since the integrand is not bounded over these limits, we cannot just integrate blindly. Using a bit more care we find

$$\int_{-1}^2 \frac{dx}{x^2} = \lim_{c\nearrow 0}\int_{-1}^c \frac{dx}{x^2} + \lim_{d\searrow 0}\int_d^2 \frac{dx}{x^2}.$$

Evaluating the integrals (ignoring the limits for the time being) we find

$$\int_{-1}^c \frac{dx}{x^2} = -x^{-1}\Big|_{-1}^c = -\frac{1}{c} - 1.$$

and

$$\int_d^2 \frac{dx}{x^2} = -x^{-1}\Big|_d^2 = \frac{1}{d} - \frac{1}{2}.$$

Note that when we choose values for c and d that fall inside appropriate range of values, both of these results are positive, as they must be. Now we can combine these results and look at limits, yielding

$$\int_{-1}^2 \frac{1}{x^2}dx = \lim_{c\nearrow 0}\Big(-\frac{1}{c} - 1\Big) + \lim_{d\searrow 0}\Big(\frac{1}{d} - \frac{1}{2}\Big) = -\lim_{c\nearrow 0}\frac{1}{c} + \lim_{d\searrow 0}\frac{1}{d} - \frac{3}{2} = \infty.$$

It's interesting to note that the term of $-3/2$ is exactly the result we found when we *blindly* integrated this non-Riemann integrable function.

5.15 Area Between Curves

Using integration we can find the area between a function and the x-axis, which allows us to find the area of many different shapes. This method has one serious limitation, however - it can only be used to find the area of an object when one side of the object is flat. In general, we might want to find the area in an arbitrary shape. How can we propose doing so? We can define each edge of the object as a different function, and find the area between the two functions. In doing so, we will always be considering conventional area.

Suppose we want to find the area between $\sin(x)$ and $-\sin(x)$ over the interval $[0, \pi]$. In this case some of the space is above the x-axis, and some is below. If we want to find the conventional area between these two functions, then we must count the area above and below the x-axis both as positive. Using the fact that

$$\int_0^\pi \sin(x) = 2,$$

and recognizing the area below the x-axis is exactly the same, we find the total area to be 4. The next question is how to solve such a problem more generally, without needing to rely on symmetry considerations.

Suppose we want to find the area between two functions f and g over some interval of interest, where $f(x) \geq g(x)$. Let's consider transforming both f and g in an identical way, for instance, by adding 2 to both functions. In doing so both functions will shift up by 2, but the distance between the two functions will not be altered at all. Thus, the area between them remain the same. Keeping this in mind, we can go one step further, and think of deforming both functions in a way so that the bottom curve g is fit to the x-axis. Since both functions are being deformed identically, the area between them will not change. Now the area between f and g is defined by the distance between a deformed version of f and the x-axis, which is an area we know how to calculate. We simply need to integrate the deformed f over the interval of interest, and we will have the area between f and g.

It may appear difficult to find such a deformation, but it is actually quite a simple obstacle to overcome. If we subtract $g(x)$ from $g(x)$ at all points, the result will be a function that is identically 0 (which lies along the x-axis). We simply perform the same deformation to f, so the function we are interested in integrating is $f - g$. In conclusion, if we have two functions f and g and over some interval $[a, b]$, and $f(x) \geq g(x)$, the area between f and g is given by

$$A = \int_a^b (f(x) - g(x))dx.$$

When faced with a specific example of finding the area between two functions, we simply need to identify on which intervals one function is greater than or equal to the other, and evaluate the appropriate integrals. One additional subtlety to be wary of is finding the area *enclosed* between two curves. In this situation we will only be interested intervals that have endpoints where the functions f and g are equal, so that the area will form a closed region. In contrast, if we were trying to find the area enclosed between x and $-x$ the answer would be 0, because these functions only intersect at one point, not forming a closed shape.

Example 1. Find the area enclosed by $2 - x^2$ and $-x$.

Solution The first consideration we need to make is to find where the two curves intersect, so that we can determine which regions are enclosed by the two functions. To do so we set the two functions equal, and solve.

$$\begin{aligned} 2 - x^2 &= -x \\ x^2 - x - 2 &= 0 \\ (x+1)(x-2) &= 0. \end{aligned}$$

The solutions to the above equation are $x = -1$ and $x = 2$, so the region of interest is $[-1, 2]$. Next, we need to identify which function is greater on this interval. To do so, we only need to evaluate at a convenient sample point, because our functions are continuous, and only intersect at -1 and 2, so they cannot cross one another in this interval. Choosing the convenient point $x = 0$, we see that $2 - 0^2 > -0$, so the parabola is our top function. Now we subtract $-x$ from $2 - x^2$ and integrate, yielding

$$\int_{-1}^{2} (2 - x^2 + x)dx = 2x - \frac{x^3}{3} + \frac{x^2}{2}\Big|_{-1}^{2} = (4 - \frac{8}{3} + \frac{4}{2}) - (-2 + \frac{1}{3} + \frac{1}{2}) = \frac{9}{2}.$$

Example 2. Find the area enclosed between $\sqrt{x}$ from above and the x-axis and $y = x - 2$ from below.

Solution The first thing to recognize is that we can rewrite the lower bound of the x-axis and the line $y = x - 2$ as a single function, defined piecewise. At first the x-axis is above the line, so it acts as our lower bound, but after they intersect at $x = 2$, the line is higher and becomes our new lower bound. Thus, we can define a new function

$$f(x) = \begin{cases} 0 & 0 \le x \le 2 \\ x - 2 & x > 2 \end{cases}$$

and recognize that the original problem was really just to find the area between $\sqrt{x}$ and $f(x)$. Now we note that these two functions intersect at $x = 0$ (where $\sqrt{x}$ stems from the x-axis), and $x = 4$, because $\sqrt{4} = 2 = 4 - 2$. Finally, we integrate

$$\begin{aligned} A &= \int_0^4 (\sqrt{x} - f(x))dx = \int_0^4 \sqrt{x}dx - \int_0^4 f(x)dx = \int_0^4 \sqrt{x}dx - \int_0^2 0dx - \int_2^4 (x-2)dx \\ &= \frac{3}{2}x^{3/2}\Big|_0^4 + (\frac{x^2}{2} - 2x)\Big|_2^4 = \frac{10}{3}. \end{aligned}$$

Above we could have just as well split the integral from 0 to 2 and 2 to 4 for both functions, but we saved ourselves one integration of $\sqrt{x}$ by evaluating the integral as we did.

If we think about the previous problem is a slightly different way, we can much more easily find the area. Rather than thinking of both curves as functions of x, we can think of them as functions of y, and integrate along y. In doing so we avoid dealing with piecewise functions, which simplifies the calculation. To find the area between two functions of y, where f is to the right of g, we simply evaluate

$$A = \int_c^d (f(y) - g(y))dy,$$

where we are interested in the vertical region c to d.

Example 3. Find the area enclosed between $\sqrt{x}$ from above and the x-axis and $y = x - 2$ from below, by integrating with respect to y.

Solution The first thing is to write our functions in terms of y, so we have $x = y^2$, and $x = y + 2$. The functions still intersect at the same points $(0, 0)$ and $(4, 2)$, but now $y + 2 > y^2$. Rather than integrating from 0 to 4 along x, we integrate from 0 to 2 along y, and look at the difference of these two functions of y.

$$A = \int_0^2 (y + 2 - y^2)dy = \frac{y^2}{2} + 2y - \frac{y^3}{3}\Big|_0^2 = 2 + 4 - \frac{8}{3} = \frac{10}{3}.$$

In general, it is helpful to draw a picture of the situation to decide which variable to integrate along, and determine which function is above the other (or to the right of the other).

Example 4. Find the area between the curves $f(x) = -x^2 + 3x$ and $g(x) = 2x^3 - x^2 - 5x$.

Solution The first thing we need to do is find the points where these curves intersect, so that we can find the regions enclosed by them. Setting the two functions equal we find

$$\begin{aligned} f(x) &= g(x) \\ g(x) - f(x) &= 0 \\ 2x^3 - x^2 - 5x - (-x^2 + 3x) &= 0 \\ 2x^3 - 8x &= 0 \\ 2x(x^2 - 4) &= 0 \end{aligned}$$

From the above equation we find that the curves intersect at $x = 0$, and $x = \pm 2$. This defines two subintervals of interest, namely $[-2, 0]$ and $[0, 2]$. It is now our task to determine which of the curves is above the other on each of these subintervals. We do so as before by sampling a single point of convenience. For simplicitly, we'll consider the points -1 and 1 (as they should be relatively easy points to evaluate our functions at).

$$\begin{aligned} f(-1) &= -(-1)^2 + 3(-1) = -1 - 3 = -4 \\ g(-1) &= 2(-1)^3 - (-1)^2 - 5(-1) = -2 - 1 + 5 = 2 \\ f(1) &= -(1)^2 + 3(1) = -1 + 3 = 2 \\ g(1) &= 2(1)^3 - (1)^2 - 5(1) = 2 - 1 - 5 = -4 \end{aligned}$$

Thus, it follows that over $[-2, 0]$ we have f as our *top* function, and over $[0, 2]$ we have g as our top function. Noting this we find the area between the two curves is

$$\text{Area} = \int_{-2}^0 (g(x) - f(x))dx + \int_0^2 (f(x) - g(x))dx = \int_{-2}^0 (g(x) - f(x))dx - \int_0^2 (g(x) - f(x))dx.$$

Above we need to evaluate the same integral twice, so we first evaluate the indefinite integral to find an antiderivative, and then we will plug in the limits of integration afterwards, in order to save work. Noting that above we already found that $g(x) - f(x) = 2x^3 - 8x$, we find that

$$\int (g(x) - f(x))dx = \int (2x^3 - 8x)dx = \frac{x^4}{2} - 4x^2 + c.$$

Now we can evaluate the definite integrals, finding

$$\text{Area} = \Big(\frac{x^4}{2} - 4x^2\Big)\Big|_{-2}^0 - \Big(\frac{x^4}{2} - 4x^2\Big)\Big|_0^2 = \Big(0 - \Big(\frac{(-2)^4}{2} - 4(-2)^2\Big)\Big) - \Big(\frac{2^4}{2} - 4 \cdot 2^2 - 0\Big) = 8 - (-8) = 16.$$

5.16 Finding Volumes Using Slabs

While we have been successful in using integration to find the area of many objects, the notion of integration can also be extended to calculate volumes, which is our current goal. When we calculate the area of a region, we simply divide the region into a number of small pieces, each of which we can calculate its area. By summing all of these areas we find the total area. Similarly, we can divide a three-dimensional region into a number of small volumes, and sum the volumes of each of the smaller pieces. The result will give us the total volume of the region. Its worth noting that it does not matter how we cut the solid into smaller pieces; there are multiple methods for doing so, and the optimal method depends on the situation.

For now we are interested in approximating a volume using cylindrical slabs. In order to find the volume of a cylinder, we simply calculate the surface area of the base and multiply it by the height. The reason why this formula works is because the cross-section of the cylinder doesn't change. No matter where we think about slicing (perpendicular to the height), the two-dimensional figure will be identical to that of the base.

What if we have a more complicated object, so that the cross-sections vary as we move along its height? For instance, we could think of a perfect loaf of bread as our cylindrical object, and a lopsided loaf as our more complicated object. When we cut out slices of bread from the perfect loaf, they all look the same, but the slices from the lopsided loaf vary depending on where we cut from. However, if we cut the slices thin enough, and think of each slice as a new object in its own right, the cross-sections of the slice vary very little. That is, each of the individual slices is approximately like a cylinder, and the approximation improves the thinner we make the slices. Each slice will have a volume given by

$$A(x)dx,$$

where $A(x)$ is the area of the base, and dx represents the height of the very thin slice. If we add the volume of all of these slices together, we will get the total volume, so if our lopsided loaf extends from a to b, its volume will be given by

$$V = \int_a^b A(x)dx.$$

Example 1. A pyramid that is 3m in height has a square base with 3m on each side. The cross-section of the pyramid perpendicular to the altitude x m down from the vertex is a square x m on a side. Find the volume of the pyramid.

Solution In this case it is helpful to define a coordinate system so that the apex of the pyramid is at the origin, and the altitude of the pyramid is along the x-axis. In this way, the pyramid extends from 0 to 3 along the x-axis. A given cross-section is x by x m^2, so the volume of a given slab is x^2dx and the total volume is given by

$$V = \int_0^3 x^2 dx = \frac{x^3}{3}\bigg|_0^3 = 9 \text{ m}^3.$$

Example 2. Find the volume of a right circular cone with height h and radius r.

Solution Once again we will want to place the tip of the cone at the origin, and let its height extend along the x-axis, so the cone will begin at $x = 0$ and end at $x = h$. In order to find the area of a cross-section, we will need to use similar triangles. Since we know each cross-section is circular, we simply need to find the radius of a general cross-section. Moving out a distance x from the origin, let y denote the height of the triangle formed by the x-axis, the slant of the cone, and a line drawn at a right angle to the x-axis. This triangle is similar to the triangle given by the slant, the height h, and the radius r. Thus, we find that

$$\frac{y}{x} = \frac{r}{h},$$

or

$$y = \frac{r}{h}x,$$

where y will give us the radius of a general cross-section. Finding the area then

$$\pi y^2 = \pi \frac{r^2}{h^2}x^2.$$

Finally, we are ready to integrate, and find that

$$V = \int_0^h \pi \frac{r^2}{h^2}x^2 dx = \pi \frac{r^2}{h^2}\frac{x^3}{3}\Big|_0^h = \frac{\pi r^2 h}{3},$$

which is the familiar formula for the volume of a right-circular cone.

In light of the above example, we can think of generating a solid by considering a function f and rotating it about a line, for instance, the x-axis. The result will be a three-dimensional figure, with each of its cross-sections given by a circle. The function f and the line chosen will determine the radius of each cross-section, and the volume of a single slab will be given by

$$\pi R^2 dx,$$

where the radius R is a function of x or

$$\pi R^2 dy,$$

for R a function of y.

Example 3. The region between the curve $1+\sin(x)$, $0 \le x \le 4\pi$ and the x-axis is revolved around the x-axis to generate a solid. Find its volume.

Solution In this case the area of a cross section is given by

$$A(x) = \pi(1+\sin(x))^2 = \pi(1+2\sin(x)+\sin^2(x)) = \pi(\frac{3}{2}+2\sin(x)-\frac{\cos(2x)}{2})$$

noting that

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

and that we need to integrate from 0 to 4π, we find

$$\begin{aligned} V &= \pi\int_0^{4\pi}(\frac{3}{2}+2\sin(x)-\frac{\cos(2x)}{2})dx = \pi(\frac{3}{2}x-2\cos(x)-\frac{\sin(2x)}{4})\Big|_0^{4\pi} \\ &= \pi(6\pi-2-0-(0-2-0)) = 6\pi^2. \end{aligned}$$

Example 4. A sphere can be generated by rotating the circle

$$x^2+y^2=a^2$$

about the x-axis. Find the volume of the sphere.

Solution The radius of a typical cross-section is given by y, so

$$A(x) = \pi y^2 = \pi(a^2 - x^2).$$

In this case the sphere extends from $-a$ to a on the x-axis, so the volume is given by

$$V = \pi \int_{-a}^{a} (a^2 - x^2)dx = \pi(a^2 x - \frac{x^3}{3})\Big|_{-a}^{a} = \frac{4}{3}\pi a^3.$$

Example 5. Find the volume generated by revolving the region bounded by $y = \sqrt{x}$, $y = 1$, and $x = 4$ about the line $y = 1$. **Solution** Even though the revolution is not about the x-axis, we still simply find cross-sections in the same way as before, in order to find the area. First, note that along the x-axis we will need to integrate from 1 to 4. Next, see that a typical cross-section is given by

$$A(x) = \pi(\sqrt{x} - 1)^2 = \pi(x - 2\sqrt{x} + 1)$$

so

$$V = \int_1^4 \pi(x - 2\sqrt{x} + 1)dx = \pi(\frac{x^2}{2} - 2 \cdot \frac{2}{3}x^{3/2} + x)\Big|_1^4 = \frac{7\pi}{6}.$$

Example 6. Find the volume of the solid generated by revolving the region between the line $x = 2$ and the curve $x = 2/y + 2$, $1 \leq y \leq 4$, about the line $x = 2$.

Solution The only difference between this situation and the previous ones is now we will have cross-sections which are given as functions of y, and we will sum them vertically, rather than horizontally as we have been doing previously. The area of a cross-section is given by

$$A(y) = \pi(2/y + 2 - 2)^2 = \frac{4\pi}{y^2}.$$

Here our limits of integration are 1 to 4, so we find

$$V = 4\pi \int_1^4 \frac{dy}{y^2} = -\frac{4\pi}{y}\Big|_1^4 = -\pi + 4\pi = 3\pi.$$

What now, if we generated a solid using the above curve, but rather than rotating about $x = 2$, we rotate about the y-axis? The result will be a volume with a hollow inside. In order to find the area of a cross-section of this new solid, we simply find the area as if there were no hole, and then subtract the area from the inside. Thus, if the outside has a radius of R, and the inside has a radius of r, the cross-sectional area will be given by

$$A(x) = \pi(R^2 - r^2).$$

Example 7. Find the volume of the solid generated by revolving the region between the line $x = 2$ and the curve $x = 2/y + 2$, $1 \leq y \leq 4$, about the y-axis.

Solution The area of a cross-section is given by

$$A(y) = \pi(2/y + 2)^2 - \pi 2^2 = \pi(\frac{4}{y^2} + \frac{4}{y} + 4 - 4) = 4\pi(\frac{1}{y^2} + \frac{1}{y}).$$

Here our limits of integration are 1 to 4, so we find

$$V = 4\pi \int_1^4 (\frac{1}{y^2} + \frac{1}{y})dy = 4\pi(-\frac{1}{y} + \ln(|y|))\Big|_1^4 = 4\pi(-\frac{1}{4} + \ln(4) + 1 - \ln(1)) = 3\pi + 4\pi\ln(4) \approx 26.8455.$$

In a more general setting, the hollowed out inside of the function need not been a simple cylinder. Consider the following example.

Example 8. The region bounded by the curve $y = x^2 + 1$ and $y = -x + 3$ is revolved about the x-axis to generate a solid. Find the volume of the solid.

Solution First we need to find the intersection of these two curves to determine the region we are revolving.

$$\begin{aligned} x^2 + 1 &= -x + 3 \\ x^2 + x - 2 &= 0 \\ (x+2)(x-1) &= 0. \end{aligned}$$

so our points of intersection are $(-2, 5)$ and $(1, 2)$. Choosing a sample point of $x = 0$, we see that $-x + 3$ is the higher of these two curves, so it will define the outer radius of our cross-section. We find that

$$A(x) = \pi((-x+3)^2 - (x^2+1)^2) = \pi(8 - 6x - x^2 - x^4).$$

It follows then that

$$V = \pi \int_{-2}^{1} (8 - 6x - x^2 - x^4)dx = \pi(8x - 3x^2 - \frac{x^3}{3} - \frac{x^5}{5})\Big|_{-2}^{1} = \frac{117\pi}{5}.$$

5.17 Finding Volumes Using Shells

While we can find a number of volumes by chopping objects into cylindrical slabs, there are some situations in which it is difficult if not impossible to do so. Consider generating a solid of revolution with a hollow inside. Now imagine that a single curve was used to generate the solid. For instance, consider the region between the function $f(x) = 3x - x^2$ and the x-axis rotated around the line $x = -1$. Here we cannot exactly subtract an inner radius from an outer radius, because both are defined by the same curve. Nevertheless, we can still calculuate the volume of such a solid, by dividing it into cylindrical shells rather than cylindrical slabs.

Suppose that we have generated a three-dimensional solid by revolving the above curve around the line $x = -1$. Imagine viewing the solid from above, as if you were looking down from above on the x-z plane (here we let y be the direction that defines the height of our solid). Now draw concentric circles centered at $x = -1$, $z = 0$. Imagine taking a saw and cutting through our solid along each of these concentric circles. The result will be our solid cut into a number of cylindrical shells. Let us cut our concentric circles close enough so that the thickness of each shell is given by the infinitesimal width dx. At a point x the inside radius of our shell will be $1 + x$ (the distance between x and the axis of revolution), and the outer radius of the shell will be given by $1 + x + dx$. Imagine unrolling one of these shells into a nearly rectangular slab. In one dimension we will have the thickness dx, with a width given by $2\pi(1 + x)$ (the circumference of the circle defined by the inner radius). We can neglect any variation in the width, because it will be infinitesimal. Any variation in the height of the slab will also be infinitesimal, so we can simply use the height of the shell given by the function evaluated at the x value given by the inner radius. In conclusion, we find the volume of one of these infinitesimal shells (a slab when rolled out) to be given by

$$2\pi(1 + x)f(x)dx = 2\pi(1 + x)(3x - x^2)dx.$$

In order to find the volume of the entire solid we simply need to integrate (sum up) all of these infinitesimal shells. We integrate over $[0, 3]$ corresponding to where our parabola crosses the x-axis. We find that

$$V = \int_0^3 2\pi(1 + x)(3x - x^2)dx = 2\pi \int_0^3 (2x^2 + 3x - x^3)dx = 2\pi\Big[\frac{2}{3}x^3 + \frac{3}{2}x^2 - \frac{x^4}{4}\Big|_0^3 = \frac{45\pi}{2}.$$

The above process can be applied in a number of situations to calculate the volumes of solids of revolution. To find the volume of such a solid we imagine cutting the solid into a number of cylindrical shells and adding the volumes of the shells. The volume of each individual shell is given by the product of the circumference of its inner radius, height, and width. These are given by 2π times the distance between the x coordinate and the axis of rotation, the value of the rotated function evaluated at x, and finally the infinitesimal width dx. Once integrated over the appropriate limits (to include all of the sliced shells) the volume of the entire solid is found.

Example 1. The region bounded by $f(x) = \sqrt{x}$, the x-axis, and the line $x = 4$ is revolved around the y-axis to form a solid. Find the volume of the solid.

Solution The first step is to sketch the region of interest. After doing so we consider an arbitrary shell, and identify both its height and radius. The height of an arbitrary shell is given by $\sqrt{x}$, and its radius is simply x (because we revolve f around the y-axis). It follows that the volume of a representative shell is given by

$$2\pi x\sqrt{x}dx = 2\pi x^{3/2}dx.$$

Now we simply need to identify our limits of integration as $[0, 4]$ and calculate the volume. Doing so we find

$$V = \int_0^4 2\pi x^{3/2} dx = 2\pi \left[\frac{2}{5} x^{5/2}\right|_0^4 = \frac{128\pi}{5}.$$

Example 2. The region bounded by $f(x) = \sqrt{x}$, the x-axis, and the line $x = 4$ is revolved around the x-axis to form a solid. Find the volume of the solid.

Solution In this situation we can use cylindrical shells or slabs to calculate the volume. Using cylindrical slabs we find that the radius of an arbitrary slab is $\sqrt{x}$, so the area of a circular cross-section is given by

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

We need to integrate from $[0, 4]$ to pick up all of these cross-sections, so we find the volume to be given by

$$\int_0^4 \pi x dx = \frac{\pi x^2}{2}\Big|_0^4 = 8\pi.$$

As an alternative to using cylidrical slabs we can also use cylindrical shells. Doing so we find that the radius of a shell is given by y, with the height of a shell given by $4 - x = 4 - y^2$. Thus, the volume of a given cylindrical shell is

$$2\pi y(4 - y^2)dy = 2\pi(4y - y^3)dy.$$

Now we must be careful as we need to integrate along y rather than x. It follows that the volume is given by

$$2\pi \int_0^2 (4y - y^3)dy = 2\pi \left[2y^2 - \frac{y^4}{4}\right|_0^2 = 8\pi.$$

Example 3. A torus (doughnut) can be generated by revolving the region inside the circle

$$y^2 + (x - 3/2)^2 = (1/2)^2$$

around the y-axis. Set up an integral to find the volume of such a solid.

Solution The first thing to do is sketch the region of interest. We have a circle of radius 1/2 centered at $(3/2, 0)$. Now we slice the region into cylindrical shells. Each shell will have a radius of x, with a height given by $2y$. This height is $2y$ because the height of the region defined by the circle is twice that of the height of the region above the x-axis. Solving for y in terms of x we find

$$y = \sqrt{(1/2)^2 - (x - 3/2)^2},$$

so the height is given by

$$2y = 2\sqrt{(1/2)^2 - (x - 3/2)^2},$$

This leads to the integral

$$2 \cdot 2\pi \int_{1/2}^{3/2} x\sqrt{(1/2)^2 - (x - 3/2)^2}$$

for volume, where the 2 in front comes from the fact we are using twice the height. The above integral cannot be evaluated using a simple substitution, so in this case we would need to resort to numerical methods to approximate the volume of interest. Using familiar formulas for the volume of a torus we can find that the value of the above expression is

$$\frac{3}{4}\pi^2.$$

Example 4. Set up an integral to find the volume of solid generated by rotating the region between $y = \sqrt{3}$ and $x = 3 - y^2$ around the x-axis over $0 \leq x \leq 3$.

Solution Begin by sketching the region of interest. In this situation the height of a shell is given by $3 - (3 - y^2)$. The radius of such a shell is given by y. The volume integral is given by

$$\int_0^{\sqrt{3}} 2\pi y(3 - (3 - y^2))dy = \frac{9\pi}{2}.$$

Example 5. Consider the triangle with vertices $(1, 1)$, $(1, 2)$, and $(2, 2)$. Set up an integral for the volume of the region generated by rotating this triangle about the line $x = 10/3$ and the line $y = 1$.

Solution The first step is to note that this triangle is given by the region enclosed by the line $x = 1$, $y = 2$, and $y = x$. Rotating about $x = 10/3$ we find the radius of a cylindrical shell to be given by $10/3 - x$, and the corresponding height to be $2 - x$. Thus, the volume of this region is given by

$$2\pi \int_1^2 (10/3 - x)(2 - x)dx.$$

If we consider rotating about the line $y = 1$, we find the radius of a shell to be given by $y - 1$. The corresponding shell height is $y - 1$. It follows the volume is given by

$$2\pi \int_1^2 (y - 1)(y - 1)dy.$$

5.18 Lengths of Plane Curves

In addition to using integration to calculate area and volume, we can use integration to measure length. For the time being we will focus only on two-dimensional curves. For a general curve in a two-dimensional plane it is not clear exactly how to measure its length. In everyday physical situations one can place a string on top of the curve, and then measure the length of the string when it is straightened out, noting that the length of the string is the same whether it is wound up or not. Unfortunately, we have no means of running a string over an arbitrary curve, one that we might not even be able to sketch. Instead, we need to use the notion of approximation, and use a limit to make the approximation as accurate as we would like. The simplest means of approximating a curve is using straight line segments. As we increase the number of segments, they begin to hang closer and closer to the curve, and in the limit that the number of segments approaches infinity, we find the exact length of the curve.

The first issue to resolve is how to represent a general curve in a two-dimensional plane. Although we can use a function $f(x)$ to represent a curve, the number of curves we can represent using just functions is rather limited. Instead, we need to shift our focus to parameterized curves. Essentially, a parameterized curve consists of two time-dependent functions, $x(t)$ and $y(t)$, where one represents the x-coordinate of the curve at a given time, and the other the y-coordinate. The collection of points $(x(t), y(t))$ defines the curve itself. Since we are defining the curve in terms of the independent variable of time, we can actually think of the curve as representing the trajectory of some point, and as time increases the point moves along the curve.

In order to proceed in finding the lengths of two-dimensional curves we will need to impose some slight restrictions on the curves we are interested in dealing with. First of all, there is no guarantee that two arbitrary functions $x(t)$ and $y(t)$ will in any way define a nice curve. They may define some set of points filled with jumps, breaks, etc, that do very little to represent an actual curve. The first restriction we need to make is to consider plane curves.

Definition 5.18.1 (Plane Curve)**.** Let x and y be continuous functions on $[a, b]$. The set of points

$$\{((x(t), y(t)) \mid t \in [a, b]\}$$

defines a *plane curve.*

Here the restriction that x and y be continuous guarantees us that the curve we end up with will actually be connected. In order to proceed however, we still some additional restrictions.

Definition 5.18.2 (Smooth Curve)**.** A plane curve is called *smooth* if it is defined by functions x and y where x' and y' exist and are continuous on (a, b), and $x'(t)$ and $y'(t)$ are not simultaneously 0 on (a, b).

Here the requirement that the derivatives be continuous ensures that our curve will suffer no immediate changes in direction, and the requirement that the derivatives not be simultaneously 0 will guarantee that the curve does not stop or reverse over itself. With these restrictions in place, we are prepared to calculate the length of smooth curves in a two-dimensional plane.

To begin, we partition the interval $[a, b]$ into a number of subintervals,

$$a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b.$$

Over each of these subintervals we will replace the actual curve with a straight line segment. Using the formula for distance in a two-dimensional plane the length of the i^{th} segment will be given by

$$\sqrt{(x(t_i) - x(t_{i-1}))^2 + (y(t_i) - y(t_{i-1}))^2}.$$

Now we apply the mean value theorem (for derivatives), noting that there are some points $\hat{t_i}$ and $\tilde{t_i}$ in $[t_{i-1}, t_i]$ so that

$$\begin{aligned} x(t_i) - x(t_{i-1}) &= x'(\hat{t_i}) \cdot (t_i - t_{i-1}) = x'(\hat{t_i}) \cdot \Delta t_i \\ y(t_i) - y(t_{i-1}) &= y'(\tilde{t_i}) \cdot (t_i - t_{i-1}) = y'(\tilde{t_i}) \cdot \Delta t_i \end{aligned}$$

From here we find that the length of the i^{th} line segment is given by

$$\sqrt{(x'(\hat{t_i}) \cdot \Delta t_i)^2 + (y'(\tilde{t_i}) \cdot \Delta t_i)^2} = \sqrt{(x'(\hat{t_i}))^2 + (y'(\tilde{t_i}))^2}\Delta t_i.$$

Now to find the approximation of the length using n segments we simply sum up each of these contributions, to find

$$\sum_{i=1}^{n} \sqrt{(x'(\hat{t_i}))^2 + (y'(\tilde{t_i}))^2}\Delta t_i.$$

To make this approximation as accurate as we'd like, we need to consider the limit as the width of each of the subintervals approaches 0. As we do so the fact that above we have two different points $\hat{t_i}$ and $\tilde{t_i}$ becomes immaterial, because they are essentially forced to take on the same position as the width of a given interval shrinks to 0. We can see this using the mean value theorem over an interval $[t_{i-1}, t_{i-1} + \Delta t_i]$. As $\Delta t_i \to 0$ we have

$$\lim_{\Delta t_i \to 0} (x(t_i) - x(t_{i-1})) = \lim_{\Delta t_i \to 0} x'(\hat{t_i}) \cdot \Delta t_i = \lim_{\Delta t_i \to 0} x'(t_{i-1}) \cdot \Delta t_i,$$

because $x'(t)$ is a continuous function, and in the above limit there is no place for $\hat{t_i}$ to go except for t_{i-1}. An analogous result holds for $y'(\tilde{t_i})$. Thus, we arrive at the following result for the length of a smooth plane curve.

$$L = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2}dt = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}dt.$$

From this general result we can very easily move to the special cases of finding the length of a function. If we have a function $y(x)$, then we can simply treat x as the parameter, so we have the curve defined by $y(x)$ and $x(x) = x$. We find the length of such a curve by

$$L = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}dx.$$

Similarly, if we have a function $x(y)$, we can let y be the parameter, and

$$L = \int_a^b \sqrt{1 + \left(\frac{dx}{dy}\right)^2}dx.$$

Example 1. Find the circumference of the circle $x^2 + y^2 = a^2$.

Solution In order to solve this problem we first must note that the equation of the above circle is given in parameterized form by

$$x = a\cos(t), \quad y = a\sin(t), \ t \in [0, 2\pi].$$

From here we find that

$$\frac{dx}{dt} = -a\sin(t) \quad \text{and} \quad \frac{dy}{dt} = a\cos(t).$$

Now we can use the integral for the length of a curve to find

$$L = \int_0^{2\pi} \sqrt{a^2\sin^2(t) + a^2\cos^2(t)}dt = \int_0^{2\pi} a dt = at\Big|_0^{2\pi} = 2\pi a.$$

Example 2. Find the length of the curve $y = x^{3/2}$ from $0 \le x \le 2$.

Solution Here we have y as a function of x, so we can use one of the alternative forms for finding the length of the curve. We find

$$\frac{dy}{dx} = \frac{3}{2}x^{1/2},$$

so

$$\left(\frac{dy}{dx}\right)^2 = \frac{9}{4}x.$$

Now evaluating the appropriate integral

$$L = \int_0^2 \sqrt{1 + \frac{9}{4}x}dx,$$

which we solve through substitution, letting

$$u = 1 + \frac{9}{4}x,$$

and

$$du = \frac{9}{4}dx \quad \text{or} \quad \frac{4du}{9} = dx.$$

It follows $u(0) = 1$ and $u(2) = 1 + 9/2 = 11/2$. Thus,

$$L = \int_0^2 \sqrt{1 + \frac{9}{4}x}dx = \frac{4}{9}\int_1^{11/2} \sqrt{u}du = \frac{8}{27}u^{3/2}\Big|_1^{11/2} = \frac{8}{27} \cdot \left[\left(\frac{11}{2}\right)^{3/2} - 1\right] \approx 3.5255.$$

Example 3. Find the length of $y = (x/2)^{2/3}$ from $x = 0$ to $x = 2$.

Solution When we find

$$\frac{dy}{dx} = \frac{2}{3}\left(\frac{x}{2}\right)^{-1/3} \cdot \left(\frac{1}{2}\right) = \frac{1}{3}\left(\frac{2}{x}\right)^{1/3},$$

we notice that this function grows without bound as $x \to 0$. Rather than dealing with improper integrals, we can alternatively represent x as a function of y, and find

$$\begin{aligned} y &= \left(\frac{x}{2}\right)^{2/3} \\ y^{3/2} &= \frac{x}{2} \\ x &= 2y^{3/2}, \end{aligned}$$

which we have previously seen is differentiable. When $x = 0$ we have $y = 0$, and when $x = 2$ we have $y = 1$, so our limits of integration will be for y from 0 to 1. Calculating the derivative

$$\frac{dx}{dy} = 2\left(\frac{3}{2}\right)y^{1/2} = 3y^{1/2},$$

and

$$\left(\frac{dx}{dy}\right)^2 = 9y.$$

Thus,

$$L = \int_0^1 \sqrt{1 + 9y}dy = \frac{1}{9} \cdot \frac{2}{3}(1 + 9y)^{3/2}\Big|_0^1 = \frac{2}{27}(10\sqrt{10} - 1) \approx 2.27.$$

Example 4. Find the length of $x = (y^3/6) + 1/(2y)$ from $y = 2$ to $y = 3$.

Solution We begin by finding the derivative

$$\frac{dx}{dy} = \frac{y^2}{2} - \frac{1}{2y^2},$$

and

$$\left(\frac{dx}{dy}\right)^2 = \frac{1}{4}(y^4 - 2 + y^{-4}).$$

Thus,

$$\begin{aligned} L &= \int_2^3 \sqrt{1 + \frac{1}{4}(y^4 - 2 + y^{-4})}dy = \int_2^3 \sqrt{\frac{1}{4}(y^4 + 2 + y^{-4})}dy = \frac{1}{2}\int_2^3 \sqrt{(y^2 + y^{-2})^2}dy \\ &= \frac{1}{2}\int_2^3 (y^2 + y^{-2})dy = \frac{1}{2}\Big[\frac{y^3}{3} - \frac{1}{y}\Big|_2^3 = \frac{1}{2}\Big[\Big(\frac{27}{3} - \frac{1}{3}\Big) - \Big(\frac{8}{3} - \frac{1}{2}\Big)\Big] = \frac{13}{4}. \end{aligned}$$

The key to solving the above problem was the fact that

$$1 + \left(\frac{dx}{dy}\right)^2$$

was a perfect square, so it canceled out with the square root. However, most curves do not work out so nicely. Unfortunately, we cannot evaluate the vast majority of integrals that arise in finding arc length, and most of them cannot be evaluated by hand. This makes numerical methods such as Simpson's rule particularly appealing for evaluating such integrals.

5.19 Work

Contrary to everyday usage, the term work has a very specific meaning in physics. In physics, work is related to the transfer of energy by forces. There are essentially two complementary ways to conceive of work. According to the work-kinetic energy theorem, the net work done on a object is equal to the object's change in kinetic energy (energy associated with motion). Symbolically,

$$W_{net} = \Delta T = T_f - T_i,$$

where T_f represents the object's final kinetic energy, and T_i its initial kinetic energy. Complementarily, we find that the total work done on an object by *nonconservative* forces is equal to the object's change in total energy (kinetic energy T plus potential energy U).

$$W_{nonc.} = \Delta E = \Delta T + \Delta U = (T_f - T_i) + (U_f - U_i).$$

If we analyze the units of the above equations, we see that work must have units of energy, because the right-hand sides of the equations have units of energy. Since the units of work are energy, we can see that the amount of work done by a force on an object is simply the amount of energy transferred by the force into the object. In the SI system of units the units of energy are Joules, and

$$J = \text{Nm},$$

where Newtons (N) are a unit of force, and meters (m) are a unit of distance. Thus, it follows that the dimensions of work are a force times a distance. It turns out that a force only does work when it acts on an object over some distance. Think about a desk that is so heavy that it does not move no matter how hard you try to push it. Although you are exerting a force on the desk when you try to push it, the force you apply does not move the desk, so no energy is transfered to it.

If we restrict ourselves to one-dimensional motion, the work done by a constant force acting on an object is given by

$$W = Fd,$$

where F is the force and d is the displacement of the object (the amount it is moved by the force). This is consistent with our above definition of work, as the units of work are still Newton meters (Nm), which are better known as Joules (J), the unit of energy.

Gravity, electrostatic forces, and the restoring force exerted by a spring are all examples of conservative forces. When a conservative force does work on an object, the change in the object's energy is independent of the path taken, or equivalently, the amount of work done by a conservative force as an object moves along a closed path is zero. Consider an object in a gravity field. If the object is moved around but eventually returns to its initial position, the net work done by gravity will be zero, because gravity is a conservative force. Thus, if an object moves around in a gravity field with no forces other than gravity acting on it, by the work-kinetic energy theorem its change in kinetic energy will be zero, because the net work will be zero. Similarly, the change in total energy and thus potential energy will be zero because there are no nonconservative forces acting. It is for this reason that the (gravitational) potential energy of an object in a gravity field only depends on the position of the object, not how it got to that position.

Friction and applied mechanical forces are the most common examples of nonconservative forces. In contrast to conservative forces, work done by nonconservative forces does depend on the path taken. Imagine an object moving against the resistance of friction. Due to the force of friction energy from the object will be dissipated into heat. If the object is moved along some path so that it returns to its initial position, the amount of energy lost will depend on the path it took, because the more distance it traveled the more energy it will have lost to heat.

Let's consider the example of an individual throwing a ball. To throw the ball he or she exerts a nonconservative force on the ball, which transfers kinetic energy to the ball, sending it into motion through the air. Since energy is transferred to the ball by the force of throwing the ball, the individual throwing the ball is doing work on the ball. Once the individual has thrown the ball, there is no way for him or her to recover the energy exhausted in doing so. Thus, in terms of energy, the action of throwing the ball is nonreversible. Nonconservative forces, such as the applied force in throwing the ball, can be thought of as nonreversible forces. In contrast, work done by conservative forces is in a sense *recoverable*.

Think about lifting a box vertically to a height h. In order to lift the box we will need to exert a force opposite to the force of gravity, which is pushing the box downwards. How much force do we need to exert though? We need to exert at least as much force as gravity, in the opposite direction. In general, to calculate the amount of work required to move an object a given distance, we consider applying a force equal and opposite to the forces opposing the motion of the object. The idea is that we apply a force infinitesimally greater than the forces opposing the motion of object to set the object into motion, and then hold the applied force constantly equal to the forces opposing motion. In doing so the sum of the forces acting on the object will be zero, and consequently (by Newton's second law) its acceleration will be zero, so it will remain in motion with a constant, albeit infinitesimal, velocity[2]. Therefore, if we lift the box slowly, using a constant force about equal in magnitude to the weight of the box (the force due to gravity), the work done by lifting the box will be

$$W_l = Fd = mgh.$$

At the same time gravity is acting with a force of $-mg$, so

$$W_g = F_g d = -mgh,$$

and so no net work is done. This means that there is no change the object's kinetic energy. This is sensible, because after the object has been lifted to a height h it is at rest. If we consider only the nonconservative forces acting in this situation we can gain some additional information. The net work done by nonconservative forces is mgh, which corresponds to a change in total energy, but we already know there is no change in kinetic energy. Thus, this change in energy must be related to the potential energy of the box. After being lifted, the box has gravitational potential energy, related to its increased height. In this situation work is done by the person on the box, which increases the energy of the box, but the person doing work on the box expends energy in the exertion of lifting the box (thus the energy of the person is decreased). If the box is dropped, then it will begin to accelerate and have energy corresponding with motion. When it finally hits the ground energy will be transferred into heat and sound.

In a more general situation, we can consider the amount of work done on an object by a force that is not constant. Think of stretching out a spring. The further the spring is stretched, the more force (and thus work) required to move it an additional distance. In order to find the total work required to stretch the spring a given distance, we can think of subdividing the distance the spring is stretched into very small subintervals. As these subintervals become small enough, or of length dx, the force required to stretch the spring the distance dx will be relatively constant, because the resistance of the spring will not change over such a small distance. To find the total work required to stretch the spring a distance $b - a$, we simply sum the work required to stretch the spring each

[2]In a sense this is the minimum amount of energy that can be used to move the object a given distance. If we think about applying more force than this in order to move the object, the additional energy is transfered to something other than the object, for instance, heat.

infinitesimal distance. Thus, for a variable force, we find that the work done in moving an object from a to b is given by

$$W = \int_a^b F(x)dx.$$

Hooke's law says when a spring is stretched or compressed a distance x from its natural (uncompressed) state, it exerts a force of

$$F_s = -kx,$$

where k is a spring constant, related to the specific design of the spring. This force is a a restoring force, which tends to push the spring to equilibrium. Thus, in order to stretch or compress a spring a distance x, one must exert a force of

$$F_a = kx,$$

so that if this force is maintained it will balance with the restoring force, and the spring will remain at a position x.

Example 1. Find the work required to compress a spring from its natural length of 50 cm to a length of 40 cm if the spring constant is 20 N/m.

Solution It is often easiest to convert units into meters, so that we can be certain our units will work out properly. In this case we want to compress the spring by an amount of 10 cm, which is the same as 0.1 m. Thus, we will be integrating from 0 (no compression) to -0.1 (a compression of 10 cm). We find that

$$W = \int_0^{-0.1} kxdx = 20 \int_0^{-0.1} xdx = 10x^2\Big|_0^{-0.1} = 0.1 \text{ J}.$$

Example 2. Suppose a spring has a natural length of 1 m. A 24 Newton weight is placed on the end of the spring, which stretches it to a length of 1.8 m.

i. Find the spring constant k.

ii. How much work is required to stretch the spring 2m beyond its natural length?

iii. How far will a 45 N weight stretch the spring?

Solution Since we know $F = kd$, and we have a force of 24 Newtons, with $d = 0.8$, we find that

$$k = \frac{d}{F} = \frac{24}{0.8} = 30 \text{ N/m}.$$

Next, we apply Hooke's law and note that to stretch the spring 2 meters we will go from $x = 0$ to $x = 2$, so

$$W = 30 \int_0^2 xdx = 15x^2\Big|_0^2 = 60 \text{ J}.$$

Finally, we use the relationship $F = kd$ for $F = 45$ N to find that

$$d = \frac{F}{k} = \frac{45}{30} = 1.5 \text{ m},$$

which means that this weight will stretch the spring to a length of 2.5 meters (because it is stretched 1.5 m beyond its natural length of 1m).

Example 3. A 3 kg bucket is lifted from the ground into the air by pulling in 6 meters of rope at a constant speed. The rope weighs 0.12 kg/m. How much work is required to lift the bucket and rope?

Solution Since the weight of the bucket does not change, the work required to lift it is simply force times distance, which in this case is weight times distance. In order to find the weight of the object, we need to multiply its mass by the acceleration due to gravity, 9.8 m/s^2.

$$3 \text{ kg} \cdot 9.8 \text{ m/s}^2 \cdot 6 \text{ m} = 176.4 \text{ J}.$$

The weight of the rope remaining, however, varies with how far down the bucket is hanging. As more of the rope is pulled in, less of it remains hanging, and so the weight remaining decreases. As a function x of the height of the bucket, the weight of the rope is given by

$$0.12 \text{ kg/m} \cdot 9.8 \text{ m/s}^2 \cdot (6-x) \text{ m} = 1.17(6-x) \text{ N}.$$

Finally, we integrate from 0 to 6, finding that

$$W_{rope} = \int_0^6 1.17(6-x)dx = 1.17\Big[6x - \frac{x^2}{2}\Big|_0^6 = 1.17\Big(36-18\Big) = 21.06 \text{ J}.$$

Thus, the total work done is

$$176.4 + 21.06 = 197.46 \text{ J}.$$

Using calculus we can also calculate the work required to pump liquid up out of a given container. The method we use is to think about dividing the liquid into a number of cylindrical slabs, treat the slabs as a solid object with a given weight, and then look in the limit as the height of the slabs approaches zero. Let's suppose we have a liquid with a given weight per unit volume, ρ. The volume of a given slab will be given by

$$A(y)dy,$$

where $A(y)$ is the area of a typical cross-section. Thus, the weight of a single slab will be given by

$$\rho A(y)dy,$$

which is equal to the force required to lift the slab. Finally, we need to calculate how much distance we need to lift the slab. If the top of the container is at some height h, then we will need to lift the slab a distance

$$(h-y),$$

where y is the height of the slab itself. Summing up the work required to lift each individual slab leads to the integral

$$W = \int \rho A(y)(h-y)dy.$$

Example 4. Suppose a cylindrical tank of water has a radius of 1 meter, and a height of 10 meters. Further, suppose that this tank is filled to the brim with water. Use the fact that water has a density of 1000 kg/m^3 to determine how much work is required to pump all of the water up out of the tank.

Solution The first thing to do is find the weight of the water per unit volume, which we do by multiplying by 9.8, the constant of acceleration due to gravity on earth at sea level. We then find

$$9.8 \cdot 1000 = 980 \text{ N/m}^3.$$

The volume of a typical slab is given by

$$\pi 2^2 dy = 4\pi dy \text{ m}^3,$$

and the work required to lift a slab is equal to its weight, so we find

$$F(y) = 3920\pi \text{ N}.$$

Finally, we need to lift a slab from a height y to a height 10, so we need to lift each slab a distance $(10 - y)$. Integrating over the limits 0 to 10 (the portions of the tank where water is present - all of it), we find the total work as

$$W = \int_0^{10} 3920\pi(10 - y)dy = 3920\pi(10y - y^2/2)\Big|_0^{10} = 196000\pi \approx 615752.16 \text{ N}.$$

Example 5. Consider the cone generated by rotating the line $x = y/2$ around the y-axis. Suppose that this forms a tank with height 4m, and that the tank is filled up to 3m with olive oil, weighing 8953.98544 Newtons per meter cubed. At the top of the tank there is a pipe which transfers the liquid in this tank to another tank. How much work is required to pump all of the liquid out of this tank into the other one?

Solution We begin by finding the volume area of a cylindrical slab, given by

$$A(y)dy = \pi(y/2)^2 = \frac{\pi}{4}y^2 dy \text{ m}^3.$$

The force required to lift this slab is given by its weight

$$8953.98544 \cdot \frac{\pi}{4}y^2 dy = 2238.49636\pi y^2 dy \text{ N}.$$

In order to transfer the slab of water to the other tank we must lift it to the rim of the tank, so that it will go through the transfer pipe. This requires our force $F(y)$ to act through a distance of of the rim is $4 - y$, so integrating from 0 to 3 (the portion of the tank containing liquid), we find that

$$W = \int_0^3 2238.49636\pi y^2(4 - y)dy = 2238.49636\pi(\frac{4}{3}y^3 - \frac{y^4}{4})\Big|_0^3 = 110760.989 \text{ J}.$$