Preface

Calculus is notable in that any competent mathematician with at least a masters degree, and many with just a strong bachelors, should be fluent enough in the subject to passably teach the courses, since calculus and calculus-descendant studies form such important parts of their training. Perhaps consequently, there are almost as many opinions about how it should be taught as there are people teaching it. A fuzzy and somewhat artificial division into “traditional” and “reform” camps has been the rage for some fifteen years now, though neither seems able to define their own camp very well, let alone the other camp. Textbooks are often sold labeled as traditional or reform, with a growing number giving homage to both. Actually this is not difficult since the camps seem to really disagree mostly on emphasis. When given a choice, professors pick whatever textbook most closely resembles their own philosophy, and make up for the differences using the lectures. When not given a choice, many professors still give the nearly the same lectures since again, they understand calculus very thoroughly and have their own ideas about how to best make sense of it to their particular students. If it were not for the scale of such a project, in both writing and dealing with the actual publishing aspects, there would surely be many more—and more diverse—calculus textbooks available to reflect these opinions.

Into this mix I submit this textbook, hoping it will appeal to like minded instructors. It grew out of my own ideas about what was right, and what was lacking in the textbooks from which I learned and later taught. This text has been in the works conceptually since my own graduate school days, when I was privileged to twice teach summer Calculus I at Purdue University using the very ambitious text by Richard Hunt. When I later, as an assistant professor, found that his second edition would not be published, I searched in vain for an alternative that was of a similar spirit and could find none. Some seven years after teaching those courses at Purdue, and never being totally comfortable with the texts available in the market, I finally began putting my own ideas onto paper. In talking to colleagues over the years, I am led to believe several do share visions similar to my own. To them I offer this as at least a step in their direction. I hope that the approach is fresh and energizing to some of my fellow calculus instructors, high school as well as college, who have been looking for a textbook with some of the elements offered here.

Incidentally, the title of this textbook is not meant to exclude students who are in fields of study other than mathematics. Indeed, it is hoped that anyone who pursues calculus for whatever reason will do so as a “student of mathematics.” It is not uncommon for a gathering of individuals to contain some who may be considered “students of Shakespeare” but never formally

---

1This is not to say that this textbook is a clone of Richard Hunt’s. I only claim that his was my first inspiration, and none of the other available texts seemed to me as inspired with a vision as did his. I have heard Hunt’s strategy described as “sneaking in some real analysis.” This seems an apt description of his. I do the same, though to some extent I wonder who sneaked the real analysis out of the calculus. But my mission is not just to return to that, as the rest of the preface here will explain.
completed a Literature or related major. The phrase here simply means that such individuals care enough about the subject to take personal time to examine it thoughtfully and continuously, and to become respectably articulate in the subject, at least when among one’s peers. On the other hand, one can study the mechanics of, say, calculus without considering its more technical details or its conceptual content. To do so is akin to learning to quote Shakespeare’s plays without actually understanding the themes, or knowing the contexts. With calculus as with Shakespeare such lack of understanding can lead to trouble, in the form of embarrassment in the case of Shakespeare, and perhaps more catastrophic consequences in the case of calculus applied to real-world problems. The more of a “student” one is in a particular subject, the better trouble can be avoided and the more the subject can be enjoyed and enriching. Of course this textbook is intended to be thorough for those whose major field of study is Mathematics. However, it is hoped that Calculus for Students of Mathematics will inspire each reader—whose study may be any field—to become, for a while if not for a lifetime, a true “student of mathematics.”

What is different about this text?

At the risk of appearing trite, this text is actually meant to be read, perhaps even “curled up with and read.” Many texts are too sparse, and others too concise, in their explanations and it is often up to the lecturer to fill in the details or give alternate explanations more fit for student consumption. Indeed few students can today learn calculus on their own with one of the current calculus textbooks.

This text is an attempt to reverse the contemporary roles of mathematics textbook (quick guide) and professor (expander), so that the professor does not have to sprint through the details but can, in good conscience, give the highlights or supplement with his own particular insights, knowing that the students have a complete treatment in the textbook.

Much effort has been made for the text to be self-contained. A reasonably prepared and dedicated student should be able to learn enough calculus independently with this text to be able to solve all but perhaps the most challenging problems contained here. The text is naturally more verbose than most, and is peppered with cross references and footnotes. This will be a different style for many students, but one which is worth learning how to read.

Added depth

Concepts and examples are explored deeply and coherently, with an eye towards more advanced topics. Some insights students normally arrive at on their own are explained outright, but there is enough extra depth in the explanations and examples that students should have a more coherent overall understanding, and indeed may arrive at yet deeper insights on their own as they ponder, for instance, the applications of the principles they learn here.

Pedagogically Linear Order

This is as opposed to theoretically linear order. I have spent a great deal of thought on the order of topics, and have experimented with various orders extensively with my own classes. I have found that a few simple changes can make profound differences in the rate in which material is absorbed.


\[2\] Perhaps only in mathematics are the professor’s lectures traditionally more complete, for the key topics at least, than the text. This is especially true if we include the question and answer dialogues with the students. In contrast, imagine a biology or history professor giving quantitatively more details in lecture than contained in the readings!
While this text is more theoretical than most, it was written with an awareness that there is a momentum to learning. Too many starts and stops in the development can dissipate energy from a calculus class. For that reason, it is sometimes better to show the final, “working” theory than risk bogging down in the preliminary theorems, with or without proofs. For instance, many texts will develop the natural logarithm as a definite integral, show that it works like a logarithm should and therefore must be a logarithm of some kind, and then call some theorem on inverse functions—a topic often painfully developed in its own, barely motivated section—to finally derive the function $e^x$ and its algebraic and calculus properties. I am in good company in deferring the theoretical development—until the reader is well-practiced with both these functions—and then giving the axiomatic theoretical development for completeness. The speed in which the computational skills are developed is greater, and the theoretical development is better-appreciated.

I also develop all of the derivative rules in the same chapter (Chapter 4). A reasonable argument can be made that the exponential, logarithmic and arc-trigonometric functions should be introduced later, in between other calculus topics, so students can first further digest the earlier differentiation rules through applications. Though that is a standard pedagogical technique, instead I attempt to exploit the momentum of learning the differentiation rules so that they can be completely dispatched, and then reinforced through use in the chapters on applications. Of course the professor is welcome to break up the material, perhaps to have an exam after the first few sections if it seems appropriate for the particular class.

Similarly, after all the differentiation rules are developed, and a chapter has been devoted to applications of derivatives, I devote two chapters on indefinite integrals before using them in Riemann Sum-motivated applications. The first of these integration chapters exhausts all the functions introduced earlier, in substitution-type settings. The second is my advanced integration technique chapter which builds upon the momentum of the first integration chapter. After these two chapters comes the chapter on Riemann Sums and applications of definite integrals. This approach allows the text to maintain the momentum from the derivative chapters, uninterrupted by Riemann Sums until they can be immediately motivated by the applications, and the student should be able to handle any integral which might arise, since by then the student has accomplished a considerable amount of integration. While this approach is actually less “gentle” for the development of antidifferentiation techniques, it has less stops and starts, and should help the student retain those skills throughout the applications.

### Continuity before limits.

One reason I feel comfortable developing a topic completely—without interrupting to allow the reader to “sleep on it”—is that I front-load the text with rigor. Especially in the topics of limits and continuity, my path is perhaps not the quickest through these topics, but rather the path that will give the best hope for a comprehensive understanding. It is coincidentally also the most linear for the theoretical development.

In particular I put continuity before limits, defining both in their own rights, using $\varepsilon$-$\delta$ definition. I strongly believe that Calculus loses much rigor when we omit $\varepsilon$-$\delta$ (even if students do not always understand these proofs as much as we would like), and that this omission causes much ad hoc explanation in the rest of our limit discussions (which can then barely be called “developments”). However, I realize that this is not a real analysis text, and so I only require the student to give $\varepsilon$-$\delta$ proofs for the first continuity section where I think it is best motivated (for instance by reference to tolerances), after which theorems ensure we never need to use them again in the exercises. My section on continuity on intervals has a couple of intuitive topological theorems\(^3\) on the images of intervals under continuous functions, from which I can easily state

\(^3\)Topological proofs are omitted to avoid the need to define connectedness and compactness.
the Intermediate Value Theorem (IVT), and the Extreme Value Theorem (EVT), using the
former to give a method for solving polynomial and rational inequalities. I then have several
limit sections to take care of all the first semester techniques, including separate sections for
vertical and horizontal asymptote phenomena.

Compared to other texts, the extensiveness of this particular chapter is perhaps the most
innovative feature of the textbook. It is my sincere hope that it will help solve many of the
difficulties associated with teaching these two topics.

Symbolic logic included.

To help with the rigor and communication of ideas, I include an early introduction to symbolic
logic, which I then mix into the prose throughout the rest of the text. This is done for many
reasons. First, it adds clarity through precision of the arguments. Second, the symbols naturally
illustrate the logical “flow” of the arguments. Finally, it is my hope that this will be a hook
for many students who have had difficulty relating abstract mathematics to everyday life, since
the symbolic logic arguments have common sense appeal. Learning about logical equivalence
is particularly useful in calculus since many theorems are stated in one form, used in another
equivalent form, and possibly proved in still another form. Without some logical sophistication,
such a discussion can be very confusing for calculus students. In particular, the contrapositive
and the difference between implication and equivalence are stressed, as these can be problematic
throughout one’s college studies and beyond. Of course it is hoped that the discussion of logic will
help the dedicated student sharpen his or her own analytical skills in all disciplines, mathematical
or otherwise, where logical argument is required.

Studying symbolic logic has several other advantages. For instance, college calculus courses
are often populated by a mix of students who had some exposure to calculus in high school,
while the rest had none. This often leads to overconfidence in the former group and anxiety in
the latter. Beginning with symbolic logic evens the playing field at the start, and sends a clear
message to those who had calculus before that college calculus will be different, while giving both
the novice and the former high school calculus student an opportunity to build the momentum
to study calculus at a college level.

The logic also sets a tone for a generally more abstract text than most. I feel justified in
this since, after all, the underlying principles are abstract and understanding these is crucial
for proper application. In this spirit I include, for instance, the axiomatic definition of the real
numbers (though again, I am aware this is not supposed to be a real analysis text), in order that
correct algebraic operations can be discussed in more exalted language. The discussion includes
the least upper bound property so that, much later, convergence of sequences and series will
not need to be explained in an ad hoc manner. Using notation from logic, I give a somewhat
different review of algebra and trigonometry than what students may be used to, again to get
them thinking about these things from a more sophisticated and hopefully fresher perspective.

Applications.

On a visit to Singapore in 2005, I was twice asked casually why students should study
calculus. What was a bit shocking was that this question came from two young, successful
Singaporean adults who had actually studied calculus! Fifteen years earlier, as a newly minted
graduate teaching assistant I could rattle off what are probably standard answers: it is useful in
engineering, all sorts of sciences, economics, and so on because it allows for deep analysis and
computation, impossible without calculus, regarding among other things how quantities change
and how those changes accumulate. But by the summer of 2005, either from growing tired of
repeating myself or (I would like to think) a more mature understanding of the subject, my
instinct was to pause, look around and take in what I could see as calculus problems everywhere.
It happened that both times I was asked this, we were riding on public transportation, so I could
imagine applications of calculus in the mechanics of moving the bus or train I was on at the time,
in the rate of absorption of the sun’s rays on surfaces set at different angles from those rays,
in the centrifugal/centripetal forces generated by all manner of spinning objects (wheels, motor
internals), and plenty of other examples if I wanted to indulge my revelrie further. I then tried
to explain how eventually, as with all education, once one accumulates a kind of critical mass of
it, one starts to “see things” differently, and indeed more deeply. However these two very smart
individuals both apparently passed on opportunities to explore, even in their imaginations, the
analytic power of the calculus.

It is true that I hope the reader would appreciate that all ideas presented here have relevance
in either possible applications of the ideas themselves, or at least in their understanding of how
the world works. However, it is still somewhat up to the student to be open to the relevance of
whatever topic he or she is studying, and to use his or her own imagination as to relevance each
time a new idea is explored.

There is a kind of truism in dialogue form which has been promulgated by mathemtics
educators, which reads as follows: “You can lead a horse to water but you can’t make him
drink.” “Yes, but you can sure salt his oats!” My hope is that students will find novel (to them)
relevance to how they view the world in applications of even the most abstract treatment of a
topic, and that this exercise inspires students to rethink and enhance how they view mathematics
and its role in understanding the world.

For applications I stick more to physics examples, and only occasionally inject biological or
social scientific examples. I believe physics has the clearest connection to calculus, and offers the
best motivations for its study. In fact, I do not use the tangent line slope as my introductory
motivation for the derivative, but instead use velocity (vis-à-vis position). I believe velocity is
initially more intuitive to more students. The fact that the derivative is graphically the slope
of a tangent line is a very convenient device of course, and I exploit it extensively, but too
many students walk away uninspired from calculus thinking it is all about tangent lines, and not
instead about change (instantaneous and cumulative).

Other differences.

Also different is the fact that this text is in black and white, further reinforcing a more
abstract spirit. This may be more a matter of taste, but I believe there is a place for such a text
and that fancy, four-color illustrations can be distracting from the main themes, not to mention
far more expensive to produce, a fact not unnoticed by cash-strapped students.4

The entire textbook is typeset in \LaTeX by the author, using the \LaTeX book style, with
graphics handled by the \LaTeX pstricks package. Several other \LaTeX packages were also used,
mainly for modifying the format. No graphics were imported but are all generated using \LaTeX
code from these packages.

Acknowledgments

First I would like to thank anyone who reads any part of this book. I wrote it for you! Even
if you do not read it cover-to-cover, I very much appreciate your interest. And I would like very
much to hear back, regardless of your opinion.

4It has been pointed out that many middle school history textbooks use very sophisticated, ostensibly attractive
designs, and yet middle school students are not likely to be found under the blankets with a flashlight and their
history books. Contrast this with the sparsely illustrated Harry Potter books. Granted, fiction can be more
“fun,” but a good telling of the exploits of Julius Caesar might be more compelling than a chart or graph.
For helping to make this work possible, I am most grateful to my wife of sixteen years, Hung-Chieh Chang (a.k.a. Joy Dougherty), who never showed me any doubt in her mind that this work would eventually be finished, and who put up with the seemingly countless hours I was bonding with several computers to finish this book. Her opinions on things mathematical, pedagogical and artistic were invaluable.

I am also very grateful to Southwestern Oklahoma State University, particularly all in the Department of Mathematics, for their encouragement and support for this project. Few institutions would allow junior faculty so much freedom to undertake a calculus textbook. In particular I was given some release time and was allowed to use photocopied excerpts of the work-in-progress in my calculus courses. This was both risky and expensive for the department, and I sincerely hope my colleagues find that it was worth it.

This text is strongly influenced by James Phelan, my own high school calculus instructor. Though he also taught at a local college, he did not teach specifically to prepare us for the Advanced Placement test—as so many high school instructors are directed to do—but instead taught what he thought was a solid course. (With much credit due to him, I passed the AP test with 5/5 anyhow, which gave him much satisfaction and validation!) His mission was to make us literate in mathematics in as many ways as possible while teaching as much calculus as we could absorb as high school seniors. My desire to be a “student of mathematics,” that is, to acquire some mathematical sophistication but not necessarily to earn a degree in the field, came first from him.

The professor who, by example, convinced me to become a mathematics major was Shih-Chuan Cheng at Creighton University. The coherence and sophistication of his lecture notes first convinced me of the beauty of mathematics and probably constitute the single greatest influence on the style of this text. Other coursework under John Mordeson and James Carlson at Creighton convinced me that there is a style of learning mathematics which stresses depth and coherence first and breadth second, which is far from the “sink or swim” approach, and from which students can hold their own among their peers from top-tiered schools. In other words, if you have the depth, the breadth can come later.

I must also thank all who made desktop publishing of mathematics possible. In particular, those in the \LaTeX developers community who brought us not only \LaTeX but some amazing supplemental packages, particularly \texttt{pstricks}, \texttt{multicol}, \texttt{enumerate} and \texttt{caption}, and for Adobe for inventing Postscript and PDF standards and, as importantly, keeping them “open” so the \TeX community could exploit them for producing publication-quality mathematics with \TeX and \LaTeX. With these things I was able to produce textbook quality copy to give to my students in class, as well as online, when the text was still in preliminary form, and to present camera-ready copy to a publisher. This ability has been enormously helpful to the development of this text.

---

5When I was interviewing for my first position as an assistant professor, and was asked about my teaching philosophy, I explained that the students had been reading too much Stephen King mathematics. I wanted them to think they were reading Tolstoy! (A friend later told me, “Yes, and in Russian!”) I got the job.

6The opposite approach is that students will somehow acquire depth from knowing a breadth of topics. This is arguable, but for most students I personally believe the breadth first approach will breed more confusion and anxiety, and thus be less likely to produce understanding. However, a reasonable argument can be made that student understanding of mathematics more easily goes from particular to general. My counterargument is that students taking 15–18 hours of college courses may not have the time and mental energy to make all the connections (or “synthesize” in education-speak) we hope for. Still, I encourage the reader to have an open mind on the subject, and indeed I sometimes include “drill” exercises in the problems. But in the explanations and examples I attempt to consistently work from a “depth first” philosophy.
Introduction

The discovery of calculus was one of the most important and exciting achievements in the history of intellectual progress. Virtually every field which deals with quantities has benefited from calculus. It allowed Sir Isaac Newton to derive the laws of planetary motion, Albert Einstein to derive relativity, many an economist to model and analyze market variables, and countless other achievements. In particular, it is reasonable to estimate that physics would be centuries behind its present maturity were it not for the availability of calculus.

Despite the technicalities involved in the lofty fields already mentioned, the fundamental principles of calculus are quite accessible, especially now that the subject has been distilled into a more coherent form in the passing of centuries since its initial discoveries. Generations of researchers and authors have refined the presentations to be understandable to motivated college students with a variety of interests. Some larger universities have separate calculus courses specifically for majors in business, agriculture, forestry, and even English, as well as the mainstays of mathematics, engineering and science. Calculus is the marquee mathematical subject for many of these programs of study, particularly science and engineering. Its importance can not be overstated. Even the algebra-trigonometry courses at our institutions have been fashioned largely to groom students for eventual study of calculus.

So what is calculus? The short answer is that it is the field of mathematics which deals with change, both instantaneous and cumulative. Respectively, this means calculus is mainly—but certainly not exclusively—interested in solving the following two problems:

(1) given algebraic relationships among variables, compute their rates of change with respect to each other;

(2) given the rates of change of variables with respect to each other, find the algebraic relationships among the variables.

Indeed the second is simply the first in reverse. For a simple, though abstract example, consider the following questions:

(i) If we know the position \( s \) of an object at every time \( t \), can we know the velocity \( v \) of the object at every time \( t \)?

(ii) If we know the velocity \( v \) of an object at every time \( t \), can we know the position \( s \) of the object at every time \( t \)?

An important concept which will come through in the text, but which should already be intuitive, is that velocity is ultimately just the measure of how the position \( s \) is changing with time \( t \), so indeed (i) and (ii) are examples of (1) and (2) above. The answer to (i) is “yes,” and the answer to (ii) is “almost.” In fact, for (ii) we need some more information, like where the object is (i.e., \( s \)) for a particular time \( t \), and then we can usually “pin down” the position \( s \) for all time \( t \). So for instance, if we are given a starting point and time, and the velocity at every
time from then onwards, then we know where the object is at every time afterwards. In fact, we have even more freedom, for it is enough to know the velocity at all time, and then knowledge of the position at any time will determine the position for all time. But knowing velocity at all times is insufficient to finding position; we need one datum on the position to determine position for all time. In contrast, knowing position at all times is sufficient for knowing velocity at all times.

Problems of type (1) are part of the differential calculus, also known as calculus of derivatives. Problems of type (2) are part of integral calculus where, perhaps predictably, we will compute many antiderivatives. Problems of this second type tend to be more (sometimes much more) difficult than problems of the first type.

Before we even begin to work in the differential or integral calculus, we will need some preliminaries. We will begin with a chapter on symbolic logic so that we can employ that language throughout the text. Next we exercise that logic on some algebraic preliminaries. Our first preliminaries specific to calculus follow in the concepts of continuity and limits, which together form much of the theoretical foundation of the calculus, and so we will spend considerable effort on these. The bulk of our work is then contained in the chapters on differential calculus and integral calculus. A final major topic is series, which finishes our work here. This last topic will require much of its own foundational development, but is a very important aspect of the classical calculus. Within that study are the answers to questions such as how a calculator can find \( \sin 78^\circ \) to ten digits of accuracy (and how we could with pencil and paper as well, though it would require remarkable persistence!), but that is only a very small sample of the usefulness of that theory.

Throughout the text we will see other applications of limits, derivatives, antiderivatives and series, and we will explore as many of those as reasonable for a text of this scope. The development of the analytical tools is our main goal. The student well-versed in the mechanics of those tools will surely (and, it is hoped, easily) find numerous other uses for the methods developed here.
Reading this Book

The main body of the book is organized into Chapters 1, 2, 3, and so on. Chapters are then organized into sections, so for instance Chapter 1 is divided into Sections 1.1, 1.2, 1.3, etc. Most sections correspond to the amount of material a college professor should be able to introduce in a one or two hour lecture, not including time spent answering homework questions. Sections are themselves sometimes divided, so Section 5.2 (i.e., the second section of Chapter 5) may be divided into Subsections 5.2.1, 5.2.2, and so on. This is done to maintain a type of outline format, with each chapter, section and subsection given a title so that it is clear which topic (or subtopic, or sub-subtopic) is being developed at any particular location in the text. Because there are numerous tangential points and clarifications to be made, numbered footnotes—some quite lengthy—are employed extensively so that the regular flow of the text need not be interrupted.

While all this hierarchy and numbering may at first seem excessive, it has become standard practice, and does help to mark where a particular topic is developed. As mentioned in the preface, this book attempts to be a stand-in for an actual professor lecturing at a chalkboard in a calculus class. It is reasonable for students to expect the professor to write the main points on the board in an outline form. Textbook styles, however, differ from the usual lecture-notes outline form, which includes major headings (here chapters), Roman numerals (sections), upper-case Latin letters (subsections), Persian-Arabic numerals (definitions, examples, steps in the general explanation), and so on. But the basic idea of an outline, branching from general to specific, is the same. Of course it is not uncommon for the professor to stand away from the board from time to time and verbally elaborate on various points, or to mention relevant external topics which do not fit neatly into the flow of the outline and indeed may distract from the strictly-defined purposes of the course if included in the written outline. Nonetheless such points can do much to clarify the material, especially by importing relevance to outside topics. The footnotes provide the author of the textbook the same chance, to figuratively step away from the main flow of the text, clarify the discussion and connect it to the rest of the world.

This text also uses numerical labels for equations, theorems, corollaries, definitions, tables and figures. When a particular equation warrants, it is given a sequentially-assigned label based upon the chapter number for easy reference. For instance, Equation (7.23) would be the twenty-third such equation in Chapter 7. Theorems, definitions and figures are similarly numbered. (The exception is that each chapter’s footnote labels reset to 1, so that Chapter 1 has footnotes numbered 1, 2, 3, etc., as do Chapters 2, 3, and so on.) These are all standard, formal styles of labeling found in much of the technical literature. This calculus text presents perhaps an ideal opportunity for a first introduction to its extensive use.

The text, while not strictly linear, is mostly cumulative, with new topics constantly referring to earlier topics. Thus the material should be read in the order it is presented, with few exceptions possible (and then best chosen by an instructor).

It may not always be possible for the student to master each topic as it is presented. Noneth-

---

7This is an example of a footnote.
less, it is very important to work as hard as possible to become as proficient as possible in the topics as they are first encountered. This may sometimes require a very slow and deliberate approach to the exercises and explanations. However, it may occasionally be necessary to “table” a confusing topic, in order to return to it after seeing how it fits into the greater scheme. This is natural, and in fact even a topic seemingly “mastered” will undoubtedly benefit from a revisit. Still, strong efforts on all topics will continually pay returns as one’s intuition for calculus as a whole is nurtured.

There are numerous comments within the text explaining the importance of the various topics, and relating how difficult particular topics have historically proven to students. Again, all topics are important, but such comments are provided to indicate where particularly strong effort may be required. Comments regarding common mistakes are also common within the text.
We will often use Greek letters in this text, as is standard for technical writing. In calculus, \( \Delta, \delta, \varepsilon \) and \( \Sigma \) have particularly special roles, as do \( \theta \) and \( \phi \), among others, in trigonometry. Furthermore, we will also use other Greek letters when they are either appropriate stand-ins for their English/Latin counterparts, or when we want them to be conspicuous in mathematical expressions. Finally, a reasonably informed student in any technical discipline is expected to eventually know all of the letters of the Greek alphabet. For these reasons we include a table of Greek letters below.

<table>
<thead>
<tr>
<th>Greek Letter</th>
<th>Letter</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>( \alpha )</td>
<td>alpha</td>
</tr>
<tr>
<td>B</td>
<td>( \beta )</td>
<td>beta</td>
</tr>
<tr>
<td>\Gamma</td>
<td>( \gamma )</td>
<td>gamma</td>
</tr>
<tr>
<td>\Delta</td>
<td>( \delta )</td>
<td>delta</td>
</tr>
<tr>
<td>E</td>
<td>( \varepsilon )</td>
<td>epsilon</td>
</tr>
<tr>
<td>Z</td>
<td>( \zeta )</td>
<td>zeta</td>
</tr>
<tr>
<td>H</td>
<td>( \eta )</td>
<td>eta</td>
</tr>
<tr>
<td>\Theta</td>
<td>( \theta )</td>
<td>theta</td>
</tr>
<tr>
<td>I</td>
<td>( \iota )</td>
<td>iota</td>
</tr>
<tr>
<td>K</td>
<td>( \kappa )</td>
<td>kappa</td>
</tr>
<tr>
<td>\Lambda</td>
<td>( \lambda )</td>
<td>lambda</td>
</tr>
<tr>
<td>M</td>
<td>( \mu )</td>
<td>mu</td>
</tr>
<tr>
<td>N</td>
<td>( \nu )</td>
<td>nu</td>
</tr>
<tr>
<td>\Xi</td>
<td>( \xi )</td>
<td>xi</td>
</tr>
<tr>
<td>O</td>
<td>( \omicron )</td>
<td>omicron</td>
</tr>
<tr>
<td>\Pi</td>
<td>( \pi )</td>
<td>pi</td>
</tr>
<tr>
<td>P</td>
<td>( \rho )</td>
<td>rho</td>
</tr>
<tr>
<td>\Sigma</td>
<td>( \sigma )</td>
<td>sigma</td>
</tr>
<tr>
<td>T</td>
<td>( \tau )</td>
<td>tau</td>
</tr>
<tr>
<td>Y</td>
<td>( \upsilon )</td>
<td>upsilon</td>
</tr>
<tr>
<td>\Phi</td>
<td>( \phi )</td>
<td>phi</td>
</tr>
<tr>
<td>X</td>
<td>( \chi )</td>
<td>chi</td>
</tr>
<tr>
<td>\Psi</td>
<td>( \psi )</td>
<td>psi</td>
</tr>
<tr>
<td>\Omega</td>
<td>( \omega )</td>
<td>omega</td>
</tr>
</tbody>
</table>
# Contents

Preface .............................................. ii  
Introduction ..................................... viii  
Reading this Book ............................... x  
Table of Greek Letters ........................... xii

1 Mathematical Logic and Sets ............... 1  
  1.1 Logic Symbols and Truth Tables .......... 2  
    1.1.1 Lexicographical Listings of Possible Truth Values .......... 3  
    1.1.2 The Logic Operations ................... 4  
    1.1.3 Constructing Further Truth Tables ............... 9  
    1.1.4 Tautologies and Contradictions, A First Look .......... 11  
  1.2 Valid Logical Equivalences as Tautologies .......... 15  
    1.2.1 The Idea, and Definition, of Logical Equivalence ....... 15  
    1.2.2 Equivalences for Negations .................. 17  
    1.2.3 Equivalent Forms of the Implication .......... 18  
    1.2.4 Other Valid Equivalences ................. 19  
    1.2.5 Circuits and Logic ...................... 23  
    1.2.6 The Statements $T$ and $F$ ............ 26  
  1.3 Valid Implications and Arguments .......... 29  
    1.3.1 Valid Implications Defined .......... 29  
    1.3.2 Partial List of Valid Implications ....... 31  
    1.3.3 Fallacies and Valid Arguments .......... 32  
    1.3.4 Analyzing Arguments Without Truth Tables ....... 39  
  1.4 Quantifiers and Sets ...................... 45  
    1.4.1 Sets .................................. 45  
    1.4.2 Quantifiers ............................ 47  
    1.4.3 Statements with Multiple Quantifiers .......... 47  
    1.4.4 Detour: Uniqueness as an Independent Concept ...... 49  
    1.4.5 Negating Universally and Existentially Quantified Statements .......... 50  
    1.4.6 Negating Statements Containing Mixed Quantifiers ....... 51  
  1.5 Sets Proper ................................ 54  
    1.5.1 Subsets and Set Equality ................ 54  
    1.5.2 Intervals and Inequalities in $\mathbb{R}$ ........ 57  
    1.5.3 Most General Venn Diagrams .......... 59  
    1.5.4 Set Operations ......................... 60
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5.5</td>
<td>More on Subsets</td>
<td>64</td>
</tr>
<tr>
<td>1.6</td>
<td>Epilogue: Proofs Without Truth Tables</td>
<td>69</td>
</tr>
<tr>
<td>1.7</td>
<td>Logic Epilogue</td>
<td>73</td>
</tr>
<tr>
<td>2</td>
<td>Real Numbers, Algebra and Functions (Optional)</td>
<td>74</td>
</tr>
<tr>
<td>2.1</td>
<td>Real Numbers Defined</td>
<td>76</td>
</tr>
<tr>
<td>2.1.1</td>
<td>Axioms</td>
<td>78</td>
</tr>
<tr>
<td>2.2</td>
<td>Arithmetic with Real Numbers</td>
<td>86</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Subtraction is Addition of the Additive Inverse</td>
<td>86</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Division is Multiplication by the Reciprocal</td>
<td>87</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Addition &amp; Subtraction with Multiplication &amp; Division</td>
<td>88</td>
</tr>
<tr>
<td>2.2.4</td>
<td>Integer Exponents</td>
<td>90</td>
</tr>
<tr>
<td>2.2.5</td>
<td>Miscellaneous Manipulations</td>
<td>92</td>
</tr>
<tr>
<td>2.2.6</td>
<td>Order of Operations</td>
<td>93</td>
</tr>
<tr>
<td>2.2.7</td>
<td>Absolute Value, Size, and Distance on the Real Line</td>
<td>95</td>
</tr>
<tr>
<td>2.2.8</td>
<td>Binomial Expansion</td>
<td>99</td>
</tr>
<tr>
<td>2.3</td>
<td>Arithmetic of Units; Significant Digits</td>
<td>101</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Adding Units</td>
<td>101</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Multiplying Units: Simplest Cases</td>
<td>102</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Division of Units: Rates</td>
<td>103</td>
</tr>
<tr>
<td>2.3.4</td>
<td>Conversions</td>
<td>104</td>
</tr>
<tr>
<td>2.3.5</td>
<td>Tolerances</td>
<td>106</td>
</tr>
<tr>
<td>2.3.6</td>
<td>Significant Digits</td>
<td>107</td>
</tr>
<tr>
<td>2.4</td>
<td>Algebra</td>
<td>111</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Polynomial Equations: Using Theorem 2.1.8</td>
<td>113</td>
</tr>
<tr>
<td>2.4.2</td>
<td>More with Radicals</td>
<td>116</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Completing the Square</td>
<td>119</td>
</tr>
<tr>
<td>2.4.4</td>
<td>The Quadratic Formula</td>
<td>120</td>
</tr>
<tr>
<td>2.5</td>
<td>Functions-I</td>
<td>123</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Various Definitions of Functions; Notation</td>
<td>124</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Functions as Processors of Inputs</td>
<td>126</td>
</tr>
<tr>
<td>2.5.3</td>
<td>Graphs, Domain and Range</td>
<td>126</td>
</tr>
<tr>
<td>2.5.4</td>
<td>Functions By Formulas</td>
<td>127</td>
</tr>
<tr>
<td>2.6</td>
<td>Basic Functions of One Variable</td>
<td>129</td>
</tr>
<tr>
<td>2.6.1</td>
<td>“Linear” Functions</td>
<td>129</td>
</tr>
<tr>
<td>2.6.2</td>
<td>Simple Powers</td>
<td>131</td>
</tr>
<tr>
<td>2.6.3</td>
<td>Roots</td>
<td>133</td>
</tr>
<tr>
<td>2.6.4</td>
<td>Reciprocal Function</td>
<td>135</td>
</tr>
<tr>
<td>2.6.5</td>
<td>Absolute Value and Piecewise Defined Functions</td>
<td>136</td>
</tr>
<tr>
<td>2.6.6</td>
<td>Exponential Functions</td>
<td>137</td>
</tr>
<tr>
<td>2.6.7</td>
<td>Logarithmic Functions</td>
<td>140</td>
</tr>
<tr>
<td>2.6.8</td>
<td>Sine and Cosine Functions</td>
<td>141</td>
</tr>
<tr>
<td>2.6.9</td>
<td>Arcsine and Arccosine Functions</td>
<td>144</td>
</tr>
<tr>
<td>2.7</td>
<td>Functions-III: Simple Transformations of Functions</td>
<td>146</td>
</tr>
<tr>
<td>2.7.1</td>
<td>Diagramming Functions</td>
<td>146</td>
</tr>
<tr>
<td>2.7.2</td>
<td>Shift Transformations</td>
<td>147</td>
</tr>
<tr>
<td>2.7.3</td>
<td>Reflection Transformations</td>
<td>149</td>
</tr>
<tr>
<td>2.7.4</td>
<td>Rescaling: Contractions and Dilations</td>
<td>151</td>
</tr>
<tr>
<td>2.7.5</td>
<td>Multiple Simple Transformations of Basic Functions</td>
<td>152</td>
</tr>
<tr>
<td>2.7.6</td>
<td>Reflection Around Other Vertical Lines</td>
<td>154</td>
</tr>
</tbody>
</table>
## 2.8 Functions-IV: Combinations of Functions  
2.8.1 Adding Functions  
2.8.2 Multiplying Functions  
2.8.3 Compositions of Functions  

### 3 Continuity and Limits of Functions  
3.1 Definition of Continuity at a Point  
3.2 Continuity Theorems  
3.2.1 Basic Theorems  
3.2.2 One-Sided Continuity  
3.2.3 Essential versus Removable Discontinuities  
3.3 Continuity on Intervals  
3.4 Finite Limits at Points  
3.5 One-Sided Finite Limits  
3.6 Infinite Limits at Points  
3.7 Sandwich, Composition, Trigonometric Continuity  
3.7.1 Sandwich Theorem  
3.7.2 “Approaches” for Independent Versus Dependent Variables  
3.7.3 “Sandwiching” From One Side  
3.7.4 Limits with Compositions of Functions  
3.7.5 Continuity Considerations for Trigonometric Functions  
3.7.6 Proofs  
3.8 Limits “At Infinity”  
3.9 Further Limit Theorems and Trigonometric Limits  
3.9.1 Simple Limit Theorems  
3.9.2 A Trigonometric Limit  
3.9.3 Limits by Substitution  
3.9.4 Epilogue  
3.10 Limits of Sequences: A First Look  
3.10.1 Definitions and First Examples  
3.11 Sequences II  

## 4 The Derivative  
4.1 Derivative: Rate of Change, Velocity, Slope  
4.1.1 Slope of a Curve  
4.1.2 Marginal Cost  
4.1.3 Some Further Applications  
4.2 First Differentiation Rules; Leibniz Notation  
4.2.1 Positive Integer Power Rule  
4.2.2 Leibniz Notation  
4.2.3 Sum and Constant Derivative Rules  
4.2.4 Increasing and Decreasing Functions; Graphing Polynomials
## CONTENTS

4.2.5 Derivatives of Sine and Cosine ........................................... 317  
4.2.6 Absolute Values and Piece-wise Defined Functions ................. 321  

4.3 First Applications of Derivatives ........................................... 325  
4.3.1 Velocity: Vertical and Horizontal ..................................... 325  
4.3.2 Fluid Flow ........................................................................ 327  
4.3.3 Energy, Work and Power ..................................................... 328  
4.3.4 Electrical Current ............................................................. 329  
4.3.5 Ohm’s Law ....................................................................... 334  
4.3.6 Electromagnetic Inductance ................................................ 336  

4.4 Chain Rule I .......................................................................... 351  
4.4.1 Chain Rule in Prime Notation ............................................ 351  
4.4.2 Partial Proof of Chain Rule ............................................... 353  
4.4.3 Leibniz Notation and the Chain Rule ................................. 353  
4.4.4 Chain Rule Derivatives of Sine and Cosine ......................... 356  
4.4.5 Power Rule for Rational Powers ........................................ 359  
4.4.6 Proof Of Power Rule for Rational Powers ......................... 361  
4.4.7 Further Examples ............................................................... 362  

4.5 Product, Quotient and Other Trigonometric Rules ....................... 368  
4.5.1 Product Rule Stated and First Applied ............................... 368  
4.5.2 Product Rule Proof ............................................................ 372  
4.5.3 Quotient Rule .................................................................... 373  
4.5.4 Tangent, Cotangent, Secant and Cosecant Rules .................. 376  
4.5.5 Putting Rules Together—Carefully ..................................... 378  

4.6 Chain Rule II: Implicit Differentiation ................................... 383  
4.6.1 Review of Differentiation Rules ......................................... 383  
4.6.2 Implicit Functions and Their Derivatives ............................ 385  
4.6.3 A Mistake to Avoid ............................................................ 393  
4.6.4 An Application ................................................................. 395  

4.7 Arctrigonometric Functions and their Derivatives ....................... 398  
4.7.1 One-to-one Functions and Inverses, Briefly ....................... 398  
4.7.2 The Arctrigonometric Functions and Their Derivatives ....... 399  
4.7.3 The Arccosine Function and Its Derivative ....................... 405  
4.7.4 The Arcsecant and Its Derivative ...................................... 407  
4.7.5 The Arcsine Function and Its Derivative ......................... 409  
4.7.6 The Arcsecant Function and Its Derivative ...................... 412  
4.7.7 Geometric and Algebraic Arguments ................................. 414  

4.8 Exponential Functions .............................................................. 420  
4.8.1 Exponential Functions ......................................................... 420  
4.8.2 Derivative of a Special Exponential Function ...................... 423  
4.8.3 A Note on Differences Between Polynomials, Exponentials .... 427  

4.9 The Natural Logarithm I ........................................................... 429  
4.9.1 Algebra of Logarithms ......................................................... 429  
4.9.2 Graph of the Natural Logarithm Function ......................... 432  
4.9.3 Derivative of the Natural Logarithm .................................... 433  

4.10 The Natural Logarithm II: Further Results ............................. 440  
4.10.1 Logarithmic Differentiation ................................................. 440  
4.10.2 Bases Other Than $e$: Logarithms .................................... 444  
4.10.3 Bases other than $e$: Exponential Functions ...................... 445  
4.10.4 Derivative of $(f(x))^g(x)$ ............................................... 447  

4.11 Further Interpretations of the Derivative ................................. 452
### CONTENTS

4.11.1 Velocity and Acceleration ........................................ 452
4.11.2 Power .............................................................. 452
4.11.3 Fluid Flow and Volume ........................................... 452
4.11.4 Population Growth and Decay ..................................... 452
4.11.5 Consumption ....................................................... 454
4.11.6 Marginal Cost and Profit .......................................... 454

5 Using Derivatives to Analyze Functions; Applications 455
5.1 Higher-Order Derivatives and Graphing ............................... 456
  5.1.1 Higher-Order Derivatives (Derivatives of Derivatives) ............... 456
  5.1.2 Graphical Significance of the Second Derivative ....................... 457
5.2 Extrema on Closed Intervals .......................................... 477
  5.2.1 Argument for Main Theorem .................................. 477
  5.2.2 Main Theorem and Examples .................................. 480
  5.2.3 Applications (Max/Min Problems) ................................ 484
5.3 The Mean Value Theorem ............................................... 487
  5.3.1 Linear Interpolation and Average Rate of Change .................. 487
  5.3.2 Mean Value Theorem ........................................... 488
  5.3.3 Applications to Graphing Theorems ................................ 489
  5.3.4 Numerical Applications ....................................... 491
5.4 Differentials and the Linear Approximation Method ................. 493
  5.4.1 Some Approximation Problems .................................. 493
  5.4.2 Linear Approximation ........................................... 495
  5.4.3 Linear Approximations and Implicit Functions ..................... 499
  5.4.4 Differentials .................................................... 501
5.5 Newton’s Approximation Method .................................... 504
  5.5.1 Approximating Methods Applied to an Example ..................... 504
5.6 Chain Rule III: Related Rates ...................................... 512
  5.6.1 The Basic Idea; An Example .................................. 512
  5.6.2 Choosing the Form of the Relationship .......................... 513
  5.6.3 Examples: Finding a Relationship Among Variables ............... 517
  5.6.4 Further Examples ............................................... 519
  5.6.5 A Modeling Example ............................................ 520

6 Basic Integration 523
6.1 First Indefinite Integrals (Antiderivatives) .......................... 524
  6.1.1 Indefinite Integrals and Constants of Integration .................. 524
  6.1.2 Power Rule for Integrals ....................................... 526
  6.1.3 Finding C (Where Possible) .................................... 529
  6.1.4 First Trigonometric Rules ...................................... 532
  6.1.5 Integrals Yielding Inverse Trigonometric Functions ............... 534
  6.1.6 Integrals Yielding Exponential Functions ......................... 536
6.2 Summation (Sigma) Notation ......................................... 538
6.3 Riemann Sums and the Fundamental Theorem of Calculus .............. 544
  6.3.1 Absolute, Relative and Percent Errors .......................... 545
  6.3.2 A Physics Example ............................................... 546
  6.3.3 General Riemann Sums, FTC .................................... 547
  6.3.4 Computing Areas: Geometric Interpretation of FTC ............... 550
  6.3.5 Physics Application: Net Displacement versus Distance Traveling ... 554
  6.3.6 Infinitesimals .................................................... 554
# CONTENTS

6.4 Substitution With Power Rule .................................................. 557
  6.4.1 The Technique ............................................................... 558
  6.4.2 A Slight Twist on the Method .......................................... 563
  6.4.3 Other Miscellaneous Power Rule Substitutions ....................... 565

6.5 Second Trigonometric Rules ............................................... 568
  6.5.1 Antiderivatives of the Six Trigonometric Functions ............... 568
  6.5.2 Substitution and the Basic Trigonometric Functions, Part I ....... 569

6.6 Substitution with All Basic Forms ....................................... 576
  6.6.1 List of Basic Forms ..................................................... 577

6.7 Further Arctrigonometric Forms ......................................... 583
  6.7.1 Factoring to Achieve “1” ................................................ 583
  6.7.2 Completing the Square .................................................. 586

6.8 Hyperbolic Functions ...................................................... 588
  6.8.1 Hyperbolic Functions and Their Basic Intrinsic Structures ...... 588

7 Advanced Integration Techniques ........................................... 590

7.1 Substitution-Type Integration by Inspection ................................ 590
  7.1.1 The Method ................................................................. 591
  7.1.2 Limitations of the Method ............................................ 594

7.2 Integration By Parts ............................................................ 598
  7.2.1 The Idea by Examples ................................................... 598
  7.2.2 The Technique in Its Simpler Applications ......................... 599
  7.2.3 Repeated Use of Integration by Parts ............................... 603
  7.2.4 Integrals of Certain Other Functions ............................... 604
  7.2.5 An Indirect Method ..................................................... 605
  7.2.6 Miscellaneous Considerations ....................................... 608

7.3 Trigonometric Integrals ........................................................ 615
  7.3.1 Sample Problems ......................................................... 615
  7.3.2 Odd Powers of Sine or Cosine ....................................... 616
  7.3.3 Even Powers of Secant or Odd Powers of Tangent ................. 621
  7.3.4 Even Powers of Cosecant or Odd Powers of Cotangent .......... 623
  7.3.5 Even Powers of Sine and Cosine .................................... 625
  7.3.6 Miscellaneous Problems and Methods-I ........................... 628
  7.3.7 Miscellaneous Problems and Methods-II ........................... 632

7.4 Trigonometric Substitution .................................................. 638
  7.4.1 Introduction ............................................................... 638
  7.4.2 The General Approach .................................................. 643

7.5 Partial Fractions and Integration ........................................ 649
  7.5.1 Theory Behind the Forms of PFD’s (Optional) ..................... 649
  7.5.2 Partial Fraction Decompositions: The Rules ...................... 653
  7.5.3 Finding the Coefficients for PFD’s ................................. 654

7.6 Miscellaneous Methods ....................................................... 665

8 Applications of Definite Integrals I: General Arguments .............. 666

8.1 Riemann Sums and Approximations of Cumulative Quantities .......... 666
8.2 Other Complications ........................................................... 667
8.3 Approximation Methods for Definite Integrals .......................... 667
8.4 Physical Applications of Definite Integrals ............................. 667
  8.4.1 Geometric Applications ................................................ 667
9 Improper Integrals and Advanced Limit Techniques 669
  9.1 Some Asymptotics of Functions ........................................ 669
  9.2 L'Hôpital's Rule ......................................................... 674
  9.3 Other Indeterminate Forms ............................................. 683
    9.3.1 $\infty - \infty$, $0 \cdot \infty$ ......................... 683
    9.3.2 $1^\infty$, $0^0$, $\infty^0$ and similar forms. ............ 685
  9.4 Improper Integrals .................................................... 688
    9.4.1 First Examples ................................................ 688
    9.4.2 The Background and Method ................................ 689
    9.4.3 Infinite Areas do not Cancel ................................ 692
    9.4.4 More Complicated Integrals ................................ 694
    9.4.5 A Complication With Oscillatory and Other Integrals .... 697
10 Series of Constants 699
  10.1 Series and Partial Sums ............................................. 701
    10.1.1 Telescoping Series ............................................ 702
    10.1.2 Geometric Series .............................................. 705
    10.1.3 Convergence/Divergence in Geometric Series .............. 708
  10.2 NTTFD and Integral Test ............................................ 711
    10.2.1 First Divergence Test: NTTFD ............................... 713
    10.2.2 Integral Test Proper (IT) .................................. 715
  10.3 Comparison Tests .................................................... 720
    10.3.1 Direct Comparison Test (DCT) ................................ 721
    10.3.2 A Hierarchy of Functions and DCT ....................... 723
    10.3.3 Limit Comparison Test (LCT) ................................ 725
    10.3.4 Summary ....................................................... 729
    10.3.5 Elementary Series Theorems ................................ 730
  10.4 Alternating Series and Absolute Convergence .................... 732
    10.4.1 Alternating Series Test (AST) ................................ 732
    10.4.2 Absolutely and Conditionally Convergent Series ....... 735
    10.4.3 One Last Remark Concerning Absolute Convergence .... 737
  10.5 Ratio and Root Tests ................................................ 740
11 Taylor Series 747
  11.1 Taylor Polynomials: Examples and Derivation ...................... 748
  11.2 Accuracy of $P_N(x)$ ................................................. 769
  11.3 Taylor/Maclaurin Series ............................................ 776
    11.3.1 Checking Validity of Taylor Series ...................... 776
    11.3.2 Techniques for Writing Series using $\Sigma$-Notation .... 779
  11.4 Power Series, Interval of Convergence .......................... 782
    11.4.1 Definition of General Power Series ...................... 782
    11.4.2 Abel's Theorem ............................................... 782
    11.4.3 Finding the Interval and Radius of Convergence .......... 783
    11.4.4 Taylor/Power Series Connection ............................ 785
  11.5 Valid Manipulations with Taylor/Power Series .................... 788
    11.5.1 Algebraic Manipulations ..................................... 788
    11.5.2 Derivatives and Integrals .................................. 790
    11.5.3 The Binomial Series and an Application ................. 793
  11.6 Complications and the Role of Complex Numbers .................. 798
    11.6.1 Troubles stemming from continuity problems ............. 798
### 11.6 Complex Numbers
11.6.2 The Complex Plane ............................................. 801
11.6.3 The Complex Plane’s Role .................................. 805
11.6.4 Summary ...................................................... 807

### 12 Topics in Two-Dimensional Analytic Geometry 809
12.1 Conic Sections ................................................... 810
12.1.1 Conic Sections Defined Geometrically .................... 810
12.1.2 Simplest Equations of Conic Sections ..................... 811
12.1.3 Equations of Conic Sections from Definitions .......... 820
12.1.4 Eccentricity ................................................. 825
12.2 Parametric Curves ............................................... 826
12.3 Polar Coordinates ............................................... 826
12.4 Calculus in Polar Coordinates ................................. 826
12.5 Vectors in \(\mathbb{R}^2\); Scalar (Dot) Products .............. 827
12.6 Three-Dimensional Space ....................................... 831
12.7 Vectors in \(\mathbb{R}^3\); Vector (Cross) Products ............. 831
12.8 Lines and Planes in \(\mathbb{R}^3\) .................................. 831
In this chapter we introduce symbolic logic and set theory. These are not specific to calculus, but are shared among all branches of mathematics. There are various symbolic logic systems, and indeed mathematical logic is its own branch of mathematics, but here we look at that portion of mathematical logic which should be understood by any professional mathematician or advanced student. The set theory is a natural extension of logic, and provides further useful notation as well as some interesting insights of its own.

The importance of logic to mathematics cannot be overstated. No conjecture in mathematics is considered fact until it has been logically proven, and truly valid mathematical analysis is done only within the rigors of logic. Because of this dependence, mathematicians have carefully developed and formalized logic beyond some of the murkier “common sense” we learn from childhood, and given it the precision required to explore, manipulate and communicate mathematical ideas unambiguously. Part of that development is the codification of mathematical logic into symbols. With logic symbols and their rules for use, we can analyze and rewrite complicated logic statements much like we do with algebraic statements.

Symbolic logic is a powerful tool for analysis and communication, but we will not abandon written English altogether. In fact, most of our ideas will be expressed in sentences which mix English with mathematical expressions including symbolic logic. We will strive for a pleasant style of mixed prose, but we will always keep in mind the formal logic upon which we base our arguments, and resort to the symbolic logic when the logic-in-prose is complicated or can be illuminated by a symbolic representation.

Because we will use English phrases as well as symbolic logic, it is important that we clarify exactly what we mean by the English versions of our logic statements. Part of our effort in this chapter is devoted toward that end.

The symbolic language developed here is used throughout the text. It is descriptive and precise, and learning its correct use forces clarity in thinking and presentation. It is not common for a calculus textbook to include a study of logic, since authors have more than enough to accomplish in trying to offer a respectably complete treatment of the calculus itself. However, it

\footnote{In fact Bertrand Russell (1872–1970)—one of the greatest mathematicians of the twentieth century—argued successfully that mathematics and logic are exactly the same discipline. Indeed, they seem to be supersets of each other, implying they are the same set. It just happens that to many a lay person, mathematics may be associated only with numbers and computations while logic deals with argument. The field of geometry belies this categorization, but there are many other vast mathematical disciplines which are not so interested in our everyday number systems. These include graph theory (useful for network design and analysis), topology (used to study surfaces, relativity), and abstract algebra (used for instance in coding theory) to name a few. Indeed both mathematics and logic can be defined as interested in abstract, coherent structural systems. Thus, to a modern mathematician, logic-versus-mathematics may be considered a “distinction without a difference.”}
CHAPTER 1. MATHEMATICAL LOGIC AND SETS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Read</th>
<th>Example</th>
<th>Also Read</th>
</tr>
</thead>
<tbody>
<tr>
<td>∼</td>
<td>not</td>
<td>∼ P</td>
<td>P is not true</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P is false</td>
<td></td>
</tr>
<tr>
<td>∧</td>
<td>and</td>
<td>P ∧ Q</td>
<td>P and Q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>both P and Q are true</td>
<td></td>
</tr>
<tr>
<td>∨</td>
<td>or</td>
<td>P ∨ Q</td>
<td>P or Q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P is true or Q is true (or both)</td>
<td></td>
</tr>
<tr>
<td>−→</td>
<td>implies</td>
<td>P −→ Q</td>
<td>if P then Q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P only if Q</td>
<td></td>
</tr>
<tr>
<td>⇔</td>
<td>if and only if</td>
<td>P ⇔ Q</td>
<td>P bi-implies Q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P iff Q</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.1: Some basic logic notation.

is quite common for teachers and professors to insert some of the logic notation into the class lectures because of its usefulness for presenting and explaining calculus to students. Unfortunately a casual or “on the fly” introduction to these devices can cause as many problems as it solves. In this text we will instead commit early to developing and using the symbolic logic notation so we can take advantage of its correct use.

We begin with the first section (Section 1.1) devoted to the construction of truth tables, which ultimately define our first group of logic symbols. Subsequent sections in this chapter will explore valid logical equivalences (Section 1.2), valid implications and some general argument types (Section 1.3), quantifiers (Section 1.4) and sets (Section 1.5). An optional, final section (Section 1.6) considers further symbolic logic manipulations based upon those built up in the previous sections.

1.1 Logic Symbols and Truth Tables

The first logic symbols we develop in the text are listed in Table 1.1. In what follows we will explain their meanings and give their English versions, while also pointing out where casual English interpretations often differ, from each other as well as their formal meanings. It is useful to learn to read the symbols above as they would usually be said out loud. For instance, $P \land Q$ can be read, “$P$ and $Q$,” while $P \rightarrow Q$ is usually read “$P$ implies $Q$.” One reads $\neg P$ as “not $P$,” while more elaborate means for verbalizing, say, $\neg (P \lor Q)$ would include “it is not the case that $P$ or $Q$.” In fact, if $P$ is any statement, such as “it is raining,” then we can graft the words “it is the case that” and have a new statement with exactly the same meaning: “it is the case that it is raining.” This allows us more flexibility to read negations in a more natural order: $\neg P$ becomes “it is not the case that it is raining.”

Now we look again at the symbols in Table 1.1. The symbol $\neg$ is called a unary logic operation because it operates on one (albeit possibly compound) statement, say $P$. The symbols $\lor, \land, \rightarrow, \leftrightarrow$ are called connectives or binary logic operations, connecting two statements, such as $P, Q$. Both types will be developed in detail in this chapter.

---

2Occasionally some of these are verbalized using what amounts to their typographical descriptions, so for instance $P \land Q$ becomes “$P$ wedge $Q$,” while $P \lor Q$ becomes “$P$ vee $Q$.”
1.1. LOGIC SYMBOLS AND TRUTH TABLES

1.1.1 Lexicographical Listings of Possible Truth Values

In the next subsection we develop the logic operators listed previously in Table 1.1, page 2. These operators connect statements, such as \( P, Q \), etc., forming new, compound statements \( P \rightarrow Q \), \( \sim P \), \( P \land Q \), etc. In doing so, we analyze the truth or falsity of the compound statements based upon the truth or falsity of the underlying, component statements \( P, Q \), etc.\(^3\)

We always assume a particular statement can be either true or false, but not simultaneously both.\(^4\) We signify these possibilities by the truth values, \( T \) or \( F \), respectively. Note that for \( n \) independent statements \( P_1, \cdots, P_n \), there are \( 2^n \) different combinations of \( T \) and \( F \).\(^5\) Thus for a single statement \( P \), we have \( 2^1 = 2 \) truth value possibilities, \( T \) or \( F \). For two independent statements \( P \) and \( Q \), we have \( 2^2 = 4 \) possible combinations of truth values: \( TT, TF, FT, FF \), i.e., \( P \) and \( Q \) both true, \( P \) true and \( Q \) false, \( P \) false and \( Q \) true, or \( P \) and \( Q \) both false. For three statements \( P, Q, R \), the possibilities are \( 2^3 = 8 \)-fold. To list exhaustively all possible orders, we will employ a lexicographical order, as shown in Figure 1.1. If there are \( n \geq 2 \) independent statements, then for the first we write half \((2^{n-1})\) \( T \)'s and the same number of \( F \)'s. For the next statement we write half \((2^{n-2})\) \( T \)'s, and the same number of \( F \)'s, and then repeat. If there is a third, we simply alternate \( T \)'s with \( F \)'s twice as fast, i.e., \( 2^{n-3} \) \( T \)'s, as many \( F \)'s, and then repeat following that pattern until we fill out \( 2^n \) entries. The last statement’s entries are \( TFTF \cdots FT \), until \( 2^n \) entries are made. Figure 1.1 illustrates this pattern, for \( n = 1, 2, 3 \)

---

\(^{3}\)In our analyses, the component statements will consist of single letters \( P, Q \), and so on, and be allowed truth value \( T \) or \( F \). Compound statements are not necessarily allowed either truth value, but their truth values are determined by those of the underlying component statements. For instance, we will see \( P \lor (\sim P) \) can only have truth value \( T \), and \( P \land (\sim P) \) can have only \( F \), while \( P \rightarrow Q \) is sometimes \( T \), sometimes \( F \).

\(^{4}\)This is sometimes referred to as the “law of the excluded middle.” It is useful in future discussions since it is often easier to prove \( P \) is not false (i.e., \( (\sim P) \) is false) than to prove \( P \) is true.

\(^{5}\)This is a simple counting principle. For another example suppose we have four shirts and three pairs of pants, and we want to know how many different combinations of these we can wear, assuming we will wear exactly one shirt and one pair of pants. Since we can wear any of the shirts with any of the pants, the choices—for counting purposes—are independent. We have four choices of shirts, and for each of those we have three choices of pants. It is not difficult to see that we have \( 4 \times 3 = 12 \) possible combinations to choose from.

If we also include two choices of belts, and assume we will wear exactly one belt, then we have \( 4 \times 3 \times 2 = 24 \) possible combinations of shirt, pants, and belt.

---

Whole textbooks are written regarding this and other counting principles, but in this text we will only encounter a few. For undergraduates, some of these principles are often found embedded in probability courses or courses relying upon probability such as genetics, or in combinatorics which appears especially in computer science and electrical engineering.
1.1.2 The Logic Operations

The basic five logic operations we will use in this text are given in Table 1.2 for every possible truth value of underlying component statements. We say that the operation \( \sim \) takes one argument (not to be confused with the colloquial meaning of the term), that argument being \( P \) in the table above. The other operators \( \land, \lor, \rightarrow \) and \( \leftrightarrow \) each take two arguments, which for the table above we dub \( P \) and \( Q \).

We begin with the logical negation \( \sim \), which is a unary operation, i.e., acting on one (possibly compound) statement. For example consider the statement \( \sim P \), usually read “not \( P \).” This is the negation of the statement \( P \). Of course \( \sim P \) is not independent of \( P \), but its truth value is based upon that of \( P \); stating that \( \sim P \) is true is the same as stating that \( P \) is false, and stating that \( \sim P \) is false is the same as stating that \( P \) is true. We can completely describe the relationship between \( P \) and \( \sim P \) in the following truth table diagram:6

<table>
<thead>
<tr>
<th>( P )</th>
<th>( \sim P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

Table 1.2: The basic logic operations defined for all possible truth values of their arguments

We will always use double lines to separate the independent component statements \( P, Q, \) etc., from compound statements based upon them.
Perhaps a better English translation of \( \sim (\sim P) \) here would be, “it is not the case that I will not go to the store,” which clearly states that I will go to the store, i.e., \( P \). In the next section we will look at ways to calculate when two logic statements in fact mean the same thing, such as \( P \) and \( \sim (\sim P) \).

We next turn our attention to the binary operation \( \land \). This is called the logical conjunction, or just simply and: the statement \( P \land Q \) is usually read “\( P \) and \( Q \).” This compound statement \( P \land Q \) is true exactly when both \( P \) and \( Q \) are true, and false if a component statement is false. Thus its truth table is given by the following:

<table>
<thead>
<tr>
<th>( P )</th>
<th>( Q )</th>
<th>( P \land Q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

As an operation, \( \land \) returns T if both statements it connects have truth value T, and returns F otherwise, i.e., if either of the statements connected by \( \land \) is false.

**Example 1.1.1** Suppose we set \( P \) and \( Q \) to be the statements

\[
\begin{align*}
P : & \text{ I will eat pizza,} \\
Q : & \text{ I will drink soda.}
\end{align*}
\]

Connecting these with \( \land \) gives

\( P \land Q : \text{ I will eat pizza and I will drink soda.} \)

This is true exactly when I do both, eat pizza and drink soda, and is false if I fail to do one, or the other, or both.

Next we look at the binary operation \( \lor \), called the logical disjunction, or simply or. The statement \( P \lor Q \) is usually read “\( P \) or \( Q \).” For \( P \lor Q \) to be true we only need one of the underlying component statements to be true; for \( P \lor Q \) to be false we need both \( P \) and \( Q \) to be false. The truth table for \( P \lor Q \) is thus as follows:

<table>
<thead>
<tr>
<th>( P )</th>
<th>( Q )</th>
<th>( P \lor Q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

It is important to note that \( P \lor Q \) is not an exclusive or,\(^8\) so we still take \( P \lor Q \) to be true for the case that both \( P \) and \( Q \) are true. At times it is not interpreted this way in spoken English, but our standard for a statement being false, i.e., having truth value F, is that it is in fact contradicted. If we state “\( P \) or \( Q \),” to a logician we are only taken to be lying if both \( P \) and \( Q \) are false.

\(^7\)It will be taken for granted throughout the text that the reader has some familiarity with the use of parentheses ( ), brackets [ ], and similar devices for grouping quantities—logical, numerical, or otherwise—to be treated as single quantities. For instance, \( \sim (\sim P) \) means that the “outer” (or first) \( \sim \) will operate on the statement \( \sim P \), treated as a single, albeit “compound” statement. Thus we first find \( \sim P \), and then its logical negation is \( \sim (\sim P) \). This type of device is used throughout the chapter and the rest of the text.

\(^8\)The case where we have \( P \) or \( Q \) but not both is called an exclusive or. Computer scientists and electrical engineers know this as XOR. For our purposes the inclusive or \( \lor \) will suffice, and anyhow is much simpler to deal with computationally in symbolic logic manipulations, though XOR will appear in the exercises.
Example 1.1.2 For the $P$ and $Q$ from the previous example, we have

$$P \lor Q : \text{I will eat pizza or I will drink soda.}$$

Again, this is still true if I do both, eat pizza and drink soda, or just do one of these; it is sufficient that one be true, but it is not contradicted if both are true. Note that this is false exactly when both $P$ and $Q$ are false, i.e., for the case that I do not eat pizza and do not drink soda.

Sometimes in spoken English the above example of $P \lor Q$ would be considered false if I did both, eat the pizza and drink the soda. According to abstract logic, doing both does not technically make the speaker a liar. For many reasons, symbolic logic defines the operation $\lor$ to be inclusive, so that $P \lor Q$ is considered true in the case in which $P$ and $Q$ are both true.\(^9\)

Next we consider $\rightarrow$. Arguably the most common and therefore important logic statements in mathematics are of the form $P \rightarrow Q$, read “$P$ implies $Q$” or “if $P$ then $Q$.” These are also the most misunderstood by novice mathematics students, and so we will discuss them at length. As before, a truth table summarizes the action of this (binary) operation:

<table>
<thead>
<tr>
<th>$P$</th>
<th>$Q$</th>
<th>$P \rightarrow Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

Note that the only circumstance we take $P \rightarrow Q$ to be false is when $P$ is true, but $Q$ is false. As before, our standard for falsity is when the statement is actually contradicted, and that can be seen to be exactly when we have the truth of the antecedent $P$, but not of the consequent $Q$. In particular, if $P$ is false, then $P \rightarrow Q$ cannot be contradicted, so we take those two cases to be true, dubbing $P \rightarrow Q$ vacuously true for those two cases where $P$ is false.

In summary, the connection $\rightarrow$ returns T for all cases except when the first statement is true, but the second is false.

The importance of the implication extends beyond mathematics and into philosophy and other studies. Because of its ubiquity, logical implication has several syntaxes which all mean the same to a logician. It is interesting to compare the various phrases, but first we will look at an example in the same spirit as we had for $\sim$, $\land$ and $\lor$.

Example 1.1.3 For the $P, Q$ in the previous examples, we have

$$P \rightarrow Q : \text{If I will eat pizza then I will drink soda.}$$

It is useful to see when this is clearly false: when $P$ is true but $Q$ is false, which for these $P, Q$ would be the case that I eat pizza but do not drink soda. In fact, it is important that that is the only case in which we consider $P \rightarrow Q$ to be false. In particular, if $P$ is false, then $P \rightarrow Q$ is vacuously true. The idea is that if I do not eat pizza, then whether or not I drink soda I do not contradict the stating, “If I will eat pizza then I will drink soda.”

---

\(^9\)To see how English understanding is context-driven, consider the following situations. First, suppose a parent tells the child to “clean the bedroom or the garage” before dinner. If the child does both, the parent will likely take the request to be fulfilled. Next, suppose instead that parent tells the child to “take a cookie or a brownie” after dinner, and the child takes one of each. In this second context the parent may have a very different understanding of the child’s compliance to the parental instructions. To a logician (and perhaps to any self-respecting smart aleck) $\lor$ must be context-independent.
1.1. LOGIC SYMBOLS AND TRUTH TABLES

There are several English phrases which mean $P \rightarrow Q$. Below are five equivalent ways to write the corresponding English version of $P \rightarrow Q$ for the $P, Q$ in the examples. (That the fourth and fifth versions are equivalent will be proved in the next section.)

1. My eating pizza implies my drinking soda ($P$ implies $Q$).
2. If I will eat pizza then I will drink soda (if $P$ then $Q$).
3. I will eat pizza only if I will drink soda ($P$ only if $Q$).
4. I will drink soda or I will not eat pizza ($Q$ or not $P$).
5. If I will not drink soda, then I will not eat pizza (if not $Q$ then not $P$).

These five ways of stating $P \rightarrow Q$ might not all be immediately obvious, and so are worth reflection and eventual commitment to memory. Two other common—and rather elegant—ways of stating the same thing are given below in the abstract:

6. My drinking soda is necessary for my eating pizza ($Q$ is necessary for $P$).
7. My eating pizza is sufficient for my drinking soda ($P$ is sufficient for $Q$).

The kind of diction in 6 and 7 is very common in philosophical as well as mathematical discussions. We will return to implication after next discussing bi-implication, since a very common mistake for novice mathematics students is to confuse the two.

The bi-implication is denoted $P \iff Q$, and often read “$P$ if and only if $Q$.” This is sometimes also abbreviated “$P$ iff $Q$”. It states that $P$ implies $Q$ and $Q$ implies $P$ simultaneously. Thus truth of $P$ gives truth of $Q$, while truth of $Q$ would give truth of $P$. Furthermore, if $P$ is false, then so must be $Q$, because $Q$ being true would have forced $P$ to be true as well. Similarly $Q$ false would imply $P$ false (since if $P$ were instead true, so would be $Q$). The truth table for the bi-implication is the following:

<table>
<thead>
<tr>
<th>$P$</th>
<th>$Q$</th>
<th>$P \iff Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

An important, alternative way to describe the operation $\iff$ is to note that $P \iff Q$ is true exactly when $P$ and $Q$ have the same truth values (TT or FF). Thus the connective $\iff$ can be used to detect when the connected statements’ truth values match, and when they do not. This will be crucial in the next section.

**Example 1.1.4** Consider the statement $P \iff Q$ for our earlier $P$ and $Q$, for which we have

$P \iff Q$: I will eat pizza if and only if I will drink soda.

This is the idea that I can not have one without the other: if I have the pizza, I must also have the soda (“only if”), and I will have the pizza if I have the soda (“if”). This is false for the cases that I have one but not the other. Importantly it is not false if I have neither.

In fact a bi-implication $P \iff Q$ is well-named as such since it is actually the same as $(P \rightarrow Q) \land (Q \rightarrow P)$. (The proof of this fact is given in the next section.) Note that we can switch the order of statements connected by $\land$ (and), so we can instead write $(Q \rightarrow P) \land (P \rightarrow Q)$,
i.e., $Q \iff P$. In prose we can write “$P$ is necessary and sufficient for $Q$,” for $P \iff Q$, which
is then the same as “$Q$ is necessary and sufficient for $P$,” i.e., $Q \iff P$.

At this point we will make a few more observations concerning the differences between
the English and formal logic uses of terms in common. The cases below illustrate how casual English
users are often unclear about when “if,” “only if,” or “if and only if” are meant in both speaking
and listening. Again, mathematics requires absolute precision in these things.

The first difference involves the phrase “only if.” This is often misunderstood to mean “if
and only if” in everyday speech. When we combine the two words “only if,” the standard logic
meaning is not the same as “if” modified by the adverb “only.” Taken together, the words “only
if” have a different, but precise meaning in logic. Consider the following statements:

- You can drive that car only if there is gasoline in the tank.
- You can drive that car only if there is air in the tires.
- You can drive that car only if the ignition system is working.

Clearly it is not the case that you can drive that car if and only if there is gasoline in the
tank, since the gasoline is necessary but not sufficient for running the car; you also need the air,
ignition, etc., or the car still will not drive regardless of the state of the gasoline tank. Similarly
a father telling his teenaged child, “you can go out with your friends only if your homework
is finished” might justifiably find another reason to keep the child from joining the friends
even after the homework is done. (Sudden severe weather, inappropriate activities planned,
mechanical problems, and several of other reasons quickly come to mind.) Note that these are
all mathematical implications $\rightarrow$: that you can drive the car would imply there is gas, air and
ignition; that the child can go out implies that the homework is done. One has to be careful not
to read bi-implications (if and only if) into any of these statements which are only implications.

The other difference deals with another way to state implications: if/then. This is also often
misunderstood to mean if and only if. Consider the colloquial English statements:10

(a) If it stops raining, I’ll go to the store.
(b) If I win the lottery, I’ll buy a new car.

Unfortunately the “if” in statement (a) might be intended to mean “if and only if.” Thus by
stating (a) the speaker leads the listener to believe he will definitely go to the store if it stops
raining, but also that he will go to the store only if it stops raining (and thus will not go if it
does not stop raining). To the strict logician (a) is not violated in the case it does not stop
raining, but the speaker still goes to the store. Recall that in such a case (a) is vacuously true.

On the other hand, it seems somewhat more likely (b) is understood the same by the logician
and the casual user of English; though we are tempted to understand the speaker to mean if and
only if, upon reflection we would not consider him a liar for buying the car without having first
won the lottery.11

In both (a) and (b) the personalities and shared experiences of the speaker and listener will
likely play roles in what was meant by the speaker and what was understood by the listener. In
mathematics we cannot have this kind of subjectivity.

---

10 This is related by Steven Zucker, Ph.D. from Johns Hopkins University, writing in the appendix of

11 In everyday English, context may be important to our interpretations. For instance, if the first person were
asked, “Will you go to the store,” we might interpret (a) as an if and only if. For the second person, if he were
asked, “what will you do if you win the lottery,” then (b) might be interpreted as an “if,” while if he were instead
asked, “will you buy a new car,” this answer might be interpreted as an “only if.”
1.1. LOGIC SYMBOLS AND TRUTH TABLES

1.1.3 Constructing Further Truth Tables

Here we look at truth tables of more complicated compound statements. To do so, we first list the underlying component statements \( P, Q \), and so on in lexicographical order. We then proceed—working “inside-out” and step by step—to construct the resulting truth values of the desired compound statement for each possible truth value combinations of the component statements.

It will be necessary to recall the actions of each of the operations introduced earlier. These are completely summarized by their truth tables in the previous subsection, but we can summarize the actions in words:

\( \sim \) changes \( T \) to \( F \), and \( F \) to \( T \).

\( \land \) returns \( F \) unless both statements it connects are true, in which case it returns \( T \).

\( \lor \) returns \( T \) if either statements it connects are true, and \( F \) exactly when both statements are false.

\( \rightarrow \) returns \( T \) except when the first statement is true and the second false. In particular, if the first statement is false, then this returns \( T \) (vacuously).

\( \leftrightarrow \) returns \( T \) if truth values of both statements match, and \( F \) if they differ.

Example 1.1.5 Construct a truth table for \( \sim (P \rightarrow Q) \).

Solution: The underlying component statements are \( P \) and \( Q \), so we first list these, and then their possible truth value combinations in lexicographical order. In order to construct the resulting truth table values for \( \sim (P \rightarrow Q) \), we build this statement one step at a time with the operations, in an “inside-out” fashion. By this we mean that we write the truth table column for \( P \rightarrow Q \), and then apply the negation to get the truth table column for \( \sim (P \rightarrow Q) \):

\[
\begin{array}{c|c|c|c}
P & Q & P \rightarrow Q & \sim (P \rightarrow Q) \\
\hline
T & T & T & F \\
T & F & F & T \\
F & T & T & F \\
F & F & F & T \\
\end{array}
\]

This reflects a fact we had before: that we have \( \sim (P \rightarrow Q) \) true—i.e., we have \( P \rightarrow Q \) false—exactly when we have \( P \) true but \( Q \) false.

Note that in the example above the third column, which represents \( P \rightarrow Q \), essentially connects the statements represented by the first and second columns with the connective \( \rightarrow \), while the last column applied the operation \( \sim \) to the statement represented by that third column. Thus the example reads easily from left to right without interruption. It is not always possible (or easiest) to do so; often we will add a column connecting statements from previous columns which are some distance from where we want to place our new column, though our style here will always have our final column representing the desired compound statement.

Example 1.1.6 Compute the truth table for \( (P \lor Q) \rightarrow (P \land Q) \).

Solution: Our “inside-out” strategy is still the same. Here we list \( P \) and \( Q \), construct \( P \lor Q \) and \( P \land Q \) respectively, and the connect these with \( \rightarrow \):

\[
\begin{array}{c|c|c|c|c|c}
P & Q & P \lor Q & P \land Q & (P \lor Q) \rightarrow (P \land Q) \\
\hline
T & T & T & T & T \\
T & F & T & F & F \\
F & T & T & F & F \\
F & F & F & F & T \\
\end{array}
\]
Some texts refer to the \( \rightarrow \) above as the “major connective,” since ultimately the statement

\((P \lor Q) \rightarrow (P \land Q)\)

is an implication, albeit connecting two already-compounded statements \((P \lor Q)\) and \((P \land Q)\). Thus “major connective” can be seen as referring to the last operator whose action was computed in making the truth table for the statement as a whole. (In the previous example, \(\sim\) would be the major connective, though we do not refer to unitary operators as “connectives.”)

After constructing such a truth table step-by-step, it is also instructive to step back and examine the result. In particular, it is always useful to see which circumstances render the whole statement false, which here are the second and third combinations. In those, we have (the antecedent) \(P \lor Q\) true since one of the \(P\), \(Q\) is true, but (the consequent) \(P \land Q\) is not true, since \(P\) and \(Q\) are not both true.

**Example 1.1.7** Construct the truth table for \(P \rightarrow [\sim (Q \lor R)]\).

**Solution** Here we need \(2^3 = 8\) different combinations of truth values for the underlying component statements \(P\), \(Q\) and \(R\). Once we list these combinations in lexicographical order, we then compute \(Q \lor R\), \(\sim (Q \lor R)\), and then compute \(P \rightarrow [\sim (Q \lor R)]\), in essence computing the major connective \(\rightarrow\).

<table>
<thead>
<tr>
<th>(P)</th>
<th>(Q)</th>
<th>(R)</th>
<th>(Q \lor R)</th>
<th>(\sim (Q \lor R))</th>
<th>(P \rightarrow [\sim (Q \lor R)])</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T)</td>
<td>(T)</td>
<td>(T)</td>
<td>(T)</td>
<td>(F)</td>
<td>(F)</td>
</tr>
<tr>
<td>(T)</td>
<td>(T)</td>
<td>(F)</td>
<td>(T)</td>
<td>(F)</td>
<td>(F)</td>
</tr>
<tr>
<td>(T)</td>
<td>(F)</td>
<td>(T)</td>
<td>(T)</td>
<td>(F)</td>
<td>(F)</td>
</tr>
<tr>
<td>(T)</td>
<td>(F)</td>
<td>(F)</td>
<td>(F)</td>
<td>(T)</td>
<td>(T)</td>
</tr>
<tr>
<td>(F)</td>
<td>(T)</td>
<td>(T)</td>
<td>(T)</td>
<td>(F)</td>
<td>(T)</td>
</tr>
<tr>
<td>(F)</td>
<td>(T)</td>
<td>(F)</td>
<td>(T)</td>
<td>(F)</td>
<td>(T)</td>
</tr>
<tr>
<td>(F)</td>
<td>(F)</td>
<td>(T)</td>
<td>(T)</td>
<td>(T)</td>
<td>(T)</td>
</tr>
<tr>
<td>(F)</td>
<td>(F)</td>
<td>(F)</td>
<td>(F)</td>
<td>(T)</td>
<td>(T)</td>
</tr>
</tbody>
</table>

Note that the last four cases are true vacuously. The cases where this is false are when we have \(P\), but \(\sim (Q \lor R)\) is false, i.e., when we have \(P\) and \(Q \lor R\) true, i.e., when \(P\) is true and either \(Q\) or \(R\) is true.

**Example 1.1.8** Construct the truth table for \(P \land (Q \lor R)\).

**Solution:** Again we need \(2^3 = 8\) rows. The column for \(P\) is repeated, though this is not necessary or always desirable (see the above example), but here was done so that the last column represents the major connective operating on the two columns immediately preceding it.

<table>
<thead>
<tr>
<th>(P)</th>
<th>(Q)</th>
<th>(R)</th>
<th>(P \land (Q \lor R))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T)</td>
<td>(T)</td>
<td>(T)</td>
<td>(T)</td>
</tr>
<tr>
<td>(T)</td>
<td>(T)</td>
<td>(F)</td>
<td>(T)</td>
</tr>
<tr>
<td>(T)</td>
<td>(F)</td>
<td>(T)</td>
<td>(T)</td>
</tr>
<tr>
<td>(T)</td>
<td>(F)</td>
<td>(F)</td>
<td>(F)</td>
</tr>
<tr>
<td>(F)</td>
<td>(T)</td>
<td>(F)</td>
<td>(F)</td>
</tr>
<tr>
<td>(F)</td>
<td>(T)</td>
<td>(F)</td>
<td>(F)</td>
</tr>
<tr>
<td>(F)</td>
<td>(F)</td>
<td>(F)</td>
<td>(F)</td>
</tr>
</tbody>
</table>

In fact it is not difficult to spot which entries in the final column should have value \(T\), since what was required was that \(P\) be true, and at least one of the \(Q\) or \(R\) also be true. Later we will see that this has exactly the same truth value, in all circumstances, as \((P \land Q) \lor (P \land R)\), which is true if we have both \(P\) and \(Q\) true, or we have both \(P\) and \(R\) true. That these two compound statements, \(P \land (Q \lor R)\) and \((P \land Q) \lor (P \land R)\) basically state the same thing will be explored in Section 1.2, as will other “logical equivalences.”
1.1. Tautologies and Contradictions, A First Look

Two very important classes of compound statements are those which form tautologies, and those which form contradictions. As we will see throughout the text, the tautologies loom especially large in our study and use of logic. We will study both tautologies and contradictions further in the next section. Here we introduce the concepts and begin to develop an intuition for these types of statements. We begin with the definitions and most obvious examples.

**Definition 1.1.1** A compound statement formed by the component statements $P_1, P_2, \ldots, P_n$ is called a **tautology** iff its truth table column consists entirely of entries with truth value $T$ for each of the $2^n$ possible truth value combinations ($T$ and $F$) of the component statements.

**Definition 1.1.2** A compound statement formed by the component statements $P_1, P_2, \ldots, P_n$ is called a **contradiction** iff its truth table column consists entirely of entries with truth value $F$ for each of the $2^n$ possible truth value combinations ($T$ and $F$) of the component statements.

**Example 1.1.9** Consider the statement $P \lor (\sim P)$, which is a tautology:

<table>
<thead>
<tr>
<th>$P$</th>
<th>$\sim P$</th>
<th>$P \lor (\sim P)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

**Example 1.1.10** Next consider the statement $P \land (\sim P)$, which is a contradiction:

<table>
<thead>
<tr>
<th>$P$</th>
<th>$\sim P$</th>
<th>$P \land (\sim P)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
</tr>
</tbody>
</table>

We see that the statement $P \lor (\sim P)$ is always true, whereas $P \land (\sim P)$ is always false. There are other interesting tautologies, as well as other interesting contradictions. For the moment let us concentrate on the tautologies.

That the statement $P \lor (\sim P)$ is a tautology—especially when a particular example is examined—should be obvious when we consider what the statement says: $P$ is true or $\sim P$ is true. If $P$ is the statement that I will eat pizza, then we get the always true statement

$$P \lor (\sim P) : \text{I will eat pizza or I will not eat pizza.}$$

In some contexts, tautologies seem to provide no useful information. Indeed, there are times in formal speech that declaring a statement to be a tautology is meant to be demeaning. However, we will see that there are many nontrivial tautologies, and it can be quite useful to recognize complex statements which are always true. For the moment we will look at the most basic of tautologies. For instance, the next tautology is obvious if we can read and understand its symbolic representation.

**Example 1.1.11** $(P \land Q) \rightarrow P$ is a tautology:

<table>
<thead>
<tr>
<th>$P$</th>
<th>$Q$</th>
<th>$P \land Q$</th>
<th>$P$</th>
<th>$(P \land Q) \rightarrow P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

\[12\] In fact, we will essentially devote the next whole section to tautologies, though we use different terms there.
Note that the last three cases were true vacuously.

A simple English example shows how the above is a tautology. If we take $P$ as before, and $Q$ as the statement, “I will drink soda,” then $(P \land Q) \rightarrow P$ becomes, “If I will eat pizza and drink soda then I will eat pizza.” Looking at it abstractly, if we have both $P$ and $Q$ true, then we have $P$ true. Note that we cannot replace the implication $\rightarrow$ with a bi-implication $\leftrightarrow$.

In the next section we will be very much interested in tautologies in which the major connective is the bi-implication $\leftrightarrow$. In fact we will develop a variation of the notation for just those cases. The following tautology is one such example.

**Example 1.1.12** Show that $\sim (P \lor Q) \leftrightarrow ([\sim P] \land [\sim Q])$ is a tautology.\(^{13}\)

**Solution:**

<table>
<thead>
<tr>
<th>$P$</th>
<th>$Q$</th>
<th>$P \lor Q$</th>
<th>$\sim (P \lor Q)$</th>
<th>$\sim P$</th>
<th>$\sim Q$</th>
<th>$\sim P \land \sim Q$</th>
<th>$\sim (P \lor Q) \leftrightarrow ([\sim P] \land [\sim Q])$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
<td>$T$</td>
</tr>
<tr>
<td>$T$</td>
<td>$F$</td>
<td>$T$</td>
<td>$F$</td>
<td>$F$</td>
<td>$T$</td>
<td>$F$</td>
<td>$T$</td>
</tr>
<tr>
<td>$F$</td>
<td>$T$</td>
<td>$T$</td>
<td>$F$</td>
<td>$T$</td>
<td>$F$</td>
<td>$F$</td>
<td>$T$</td>
</tr>
<tr>
<td>$F$</td>
<td>$F$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
</tbody>
</table>

Later we will get into the habit of just pointing out how the two columns representing, say, $\sim (P \lor Q)$ and $(\sim P) \land (\sim Q)$ have the same truth values, so when connected by $\leftrightarrow$ we get a tautology. That will be more convenient, as the entire statement does not fit easily into a relatively narrow truth table column heading.

With reflection the various tautologies and contradictions become intuitive, and easy to identify. (For the above example consider the discussion of when $P \lor Q$ is false, as in Example 1.1.1, page 5.) However, not all things which appear to be contradictions are in fact contradictions. At the heart of the problem in such examples is usually the nature of the implication operation $\rightarrow$. Consider the following:

**Example 1.1.13** Write a truth table for $P \rightarrow (\sim P)$ to demonstrate that it is not a contradiction.

**Solution:** Note that there is only one independent statement $P$, so we need only $2^1 = 2$ rows.

<table>
<thead>
<tr>
<th>$P$</th>
<th>$\sim P$</th>
<th>$P \rightarrow (\sim P)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$F$</td>
<td>$F$</td>
</tr>
<tr>
<td>$F$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
</tbody>
</table>

Note that the second case is true vacuously. It is also interesting to note that the statement $P \rightarrow (\sim P)$ has the same truth values as $\sim P$, which can be interpreted as saying $\sim P$ is the same as $P \rightarrow (\sim P)$. That is worth pondering, but for the moment we will not elaborate.

It is perhaps easier to spot contradictions which do not involve the implication. For instance, $P \leftrightarrow (\sim P)$ is a contradiction, but the demonstration of that (by truth tables) is left to the reader. (Simply replace $\rightarrow$ with $\leftrightarrow$ in the above truth table.)

\(^{13}\)Some texts employ a strict hierarchy of “precedence” or “order of operations” on logic operations. It is akin to arithmetic, where $4 \cdot 5^2 + 3/5$ has us computing $5^2$, multiplying that by 4, separately computing $3/5$, and then adding our two results. With grouping symbols one might write $[4 \cdot (5^2)] + [3/5]$, but through conventions designed for convenience the grouping is understood in the original expression. For our example above, some texts will simply write

$\sim (P \lor Q) \leftrightarrow \sim P \land \sim Q$, understanding the precedence to be $\sim$, then $\land$ and $\lor$, and then $\rightarrow$ and $\leftrightarrow$ in the order of appearance. Thus we need the first parentheses to override the precedence of the $\sim$, else we would interpret $\sim P \lor Q$ to mean $(\sim P) \lor Q$. As this text is not for a course in logic per se, we will continue to use grouping symbols rather than spend the effort to develop and practice a procedure for precedence. (Besides, the authors find the clearly grouped statements easier and more pleasing to read and write.)
Exercises

1. A very useful way to learn the nuances of the logic operations is to consider when their compound statements are false. For each of the following compound statements, discuss all possible circumstances in which the given statement is false.

For example, $P \iff Q$ is false exactly when $P$ is true and $Q$ false, or $Q$ true and $P$ false.

(a) $\sim P$
(b) $P \land Q$
(c) $P \lor Q$
(d) $P \implies Q$
(e) $P \iff Q$
(f) $P \implies (\sim Q)$.

2. Repeat 1(a)–(f) above, except using truth tables for each to answer the question of when each statement is false. Compare and reconcile your answers to Exercise 1 above.

3. Consider the statement

$(\sim Q) \implies (\sim P)$.

(a) When is it false?
(b) Now consider $P \implies Q$. When is it false?
(c) Do you believe these two compound statements mean the same thing?
(d) Construct the truth table for the statement $(\sim Q) \implies (\sim P)$. Then revisit your answer to (c).

4. Construct the truth table for $P \text{ XOR } Q$.

(See Footnote 8, page 5.)

5. Construct the truth table for the statement $\sim (P \iff Q)$. Compare your answer to the previous exercise.

6. Construct truth tables for the following statements:

(a) $(\sim P) \iff (\sim Q)$ (Compare to $P \iff Q$.)

(b) $[P \lor (\sim Q)] \implies P$

(c) $(\sim [P \land (Q \lor R)])$

7. Find six other English statements which are equivalent to the statement,

“You can go out with your friends only if your homework is finished.”

(See page 7. Some of your answers may seem very formal.)

8. Construct truth tables for the following statements.

(a) $\sim (P \land Q)$
(b) $P \lor (Q \land R)$
(c) $P \lor (Q \lor R)$
(d) $(P \lor Q) \lor R$ (Compare to the previous statement.)
(e) $(P \implies Q) \land (Q \implies P)$

9. Decide which are tautologies, which are contradictions, and which are neither. Try to decide using intuition, and then check with truth tables.

(a) $P \implies P$
(b) $P \iff P$
(c) $P \lor (\sim P)$
(d) $P \land (\sim P)$
(e) $P \iff (\sim P)$
(f) $P \implies (\sim P)$
(g) $((P) \land (\sim P)) \implies Q$
(h) $(P \implies (\sim P)) \implies (\sim P)$
(i) $(P \land Q) \implies P$
(j) $(P \lor Q) \implies P$
(k) $P \implies (P \land Q)$
(l) $P \implies (P \lor Q)$

10. Some confuse implication $\implies$ with causation, interpreting $P \implies Q$ as “$P$ causes $Q$.” However, the implication is in fact weaker than the layman’s concept of causation. Answer the following:

(a) Show that \((P \rightarrow Q) \lor (Q \rightarrow P)\) is a tautology.

(b) Explain why, replacing \(\rightarrow\) with the phrase “causes” clearly does not give us a tautology.

(c) On the other hand, if \(P\) being true causes \(Q\) to be true, can we say \(P \rightarrow Q\) is true?

11. Write the lexicographical ordering of the possible truth value combinations for four statements \(P,Q,R,S\).
1.2 Valid Logical Equivalences as Tautologies

1.2.1 The Idea, and Definition, of Logical Equivalence

In lay terms, two statements are logically equivalent when they say the same thing, albeit perhaps in different ways. To a mathematician, two statements are called logically equivalent when they will always be simultaneously true or simultaneously false. To see that these notions are compatible, consider an example of a man named John N. Smith who lives alone at 12345 North Fictional Avenue in Miami, Florida, and has a United States Social Security number 987-65-4325. Of course there should be exactly one person with a given Social Security number. Hence, when we ask any person the questions, “are you John N. Smith of 12345 North Fictional Avenue in Miami, Florida?” and “is your U.S. Social Security number 987-65-4325?” we would be in essence asking the same question in both cases. Indeed, the answers to these two questions would always be both yes, or both no, so the statements “you are John N. Smith of 12345 North Fictional Avenue in Miami, Florida,” and “your U.S. Social Security number is 987-65-4325,” are logically equivalent. The notation we would use is the following:

\[
\text{you are John N. Smith of 12345 North Fictional Avenue in Miami, Florida} \quad \iff \quad \text{your U.S. Social Security number is 987-65-4325}.
\]

The motivation for the notation “\(\iff\)” will be explained shortly.

On a more abstract note, consider the statements \(\sim (P \lor Q)\) and \((\sim P) \land (\sim Q)\). Below we compute both of these compound statements’ truth values in one table:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>(P \lor Q)</th>
<th>(\sim (P \lor Q))</th>
<th>(\sim P)</th>
<th>(\sim Q)</th>
<th>((\sim P) \land (\sim Q))</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

We see that these two statements are both true or both false, under any of the \(2^2 = 4\) possible circumstances, those being the possible truth value combinations of the underlying, independent component statements \(P\) and \(Q\). Thus the statements \(\sim (P \lor Q)\) and \((\sim P) \land (\sim Q)\) are indeed logically equivalent in the sense of always having the same truth value. Having established this, we would write

\[
\sim (P \lor Q) \iff (\sim P) \land (\sim Q).
\]

Note that in logic, this symbol “\(\iff\)” is similar to the symbol “\(=\)” in algebra and elsewhere. There are a couple of ways it is read out loud, which we will consider momentarily. For now we take the occasion to list the formal definition of logical equivalence:

**Definition 1.2.1.** Given \(n\) independent statements \(P_1, \ldots, P_n\), and two statements \(R, S\) which are compound statements of the \(P_1, \ldots, P_n\), we say that \(R\) and \(S\) are **logically equivalent**, which we then denote \(R \iff S\), if and only if their truth table columns have the same entries for each of the \(2^n\) distinct combinations of truth values for the \(P_1, \ldots, P_n\). When \(R\) and \(S\) are logically equivalent, we will also call \(R \iff S\) a **valid equivalence**.

\(^{14}\)This is all, of course, fictitious.

\(^{15}\)Some texts use the symbol “\(\equiv\)” for logical equivalence. However there is another standard use for this symbol, and so we will reserve that symbol for that use later in the text.
Again, this is consistent with the idea that to say statements $R$ and $S$ are logically equivalent is to say that, under any circumstances, they are both true or both false, so that asking if $R$ is true is—functionally—exactly the same as asking if $S$ is true. (Recall our example of John N. Smith’s Social Security number.)

Note that if two statements’ truth values always match, then connecting them with $\leftrightarrow$ yields a tautology. Indeed, the bi-implication yields $T$ if the connected statements have the same truth value, and $F$ otherwise. Since two logically equivalent statements will have matching truth values in all cases, connecting with $\leftrightarrow$ will always yield $T$, and we will have a tautology. On the other hand, if connecting two statements with $\leftrightarrow$ forms a tautology, then the connected statements must have always-matching truth values, and thus be equivalent. This argument yields our first theorem:  

**Theorem 1.2.1** Suppose $R$ and $S$ are compound statements of $P_1 \cdots, P_n$. Then $R$ and $S$ are logically equivalent if and only if $R \leftrightarrow S$ is a tautology.

The theorem above gives us the motivation behind the notation $\iff$. Assuming $R$ and $S$ are compound statements built upon component statements $P_1 \cdots, P_n$, then 

$$R \iff S \text{ means that } R \leftrightarrow S \text{ is a tautology.}$$

(1.1) 

To be clear, when we write $R \leftrightarrow S$ we understand that this might have truth value $T$ or $F$, i.e., it might be true or false. However, when we write $R \iff S$, we mean that $R \leftrightarrow S$ is always true (i.e., a tautology), which partially explains why we call $R \iff S$ a *valid equivalence*.  

To prove $R \iff S$, we could (but usually will not) construct $R \leftrightarrow S$, and show that it is a tautology. We do so below to prove 

$$\sim (P \lor Q) \iff (\sim P) \land (\sim Q).$$

However, our preferred method will be as in the previous truth table, where we simply show that the truth table columns for $R$ and $S$ have the same entries at each horizontal level, i.e., for each truth value combination of the component statements. That approach saves space and reinforces our original notion of equivalence (matching truth values). However it is still important to understand the connection between $\leftrightarrow$ and $\iff$, as given in (1.1).

---

16 A *theorem* is a fact which has been proven to be true, particularly dealing with mathematics. We will state numerous theorems in this text. Most we will prove, though occasionally we will include a theorem which is too relevant to omit, but whose proof is too technical to include in an undergraduate calculus book. Such proofs are left to courses with titles such as mathematical (or real) analysis, topology, or advanced calculus.

Some theorems are also called lemmas (or, more archaically, lemmata) when they are mostly useful as steps in larger proofs of the more interesting results. Still others are called corollaries if they are themselves interesting, but follow with very few extra steps after the underlying theorem is proved.

17 Depending upon the author, both $R \leftrightarrow S$ and $R \iff S$ are sometimes verbalized “$R$ is equivalent to $S$,” or “$R$ if and only if $S$.” We distinguish the cases by using the term “equivalent” for the double-lined arrow, and “if and only if” for the single-lined arrow. To help avoid confusion, we emphasize this more restrictive use of “equivalences” (denoted with $\iff$) by calling them “valid equivalences.”
1.2. VALID LOGICAL EQUIVALENCES AS TAUTOLOGIES

1.2.2 Equivalences for Negations

Much of the intuition achieved from studying symbolic logic comes from examining various logical equivalences. Indeed we will make much use of these, for the theorems we use throughout the text are often stated in one form, and then used in a different, but logically equivalent form. When we prove a theorem, we may prove even a third, logically equivalent form.

The first logical equivalences we will look at here are the negations of the our basic operations. We already looked at the negations of \( \neg P \) and \( P \lor Q \). Below we also look at negations of \( P \land Q \), \( P \rightarrow Q \) and \( P \leftrightarrow Q \). Historically, (1.3) and (1.4) below are called De Morgan’s Laws, but each basic negation is important. We now list these negations.

\[
\begin{align*}
\neg (\neg P) & \iff P & (1.2) \\
\neg (P \lor Q) & \iff (\neg P) \land (\neg Q) & (1.3) \\
\neg (P \land Q) & \iff (\neg P) \lor (\neg Q) & (1.4) \\
\neg (P \rightarrow Q) & \iff P \land (\neg Q) & (1.5) \\
\neg (P \leftrightarrow Q) & \iff [P \land (\neg Q)] \lor [Q \land (\neg P)] & (1.6)
\end{align*}
\]

Fortunately, with a well chosen perspective these are intuitive. Recall that any statement \( R \) can also be read “\( R \) is true,” while the negation asserts the original statement is false. For example \( \neg R \) can be read as the statement “\( R \) is false,” or a similar wording (such as “it is not the case that \( R \)”).

Similarly the statement \( \neg (P \lor Q) \) is the same as “\( P \) or \( Q \)” is false.” With that it is not difficult to see that for \( \neg (P \lor Q) \) to be true requires both that \( P \) be false and \( Q \) be false.

For a specific example, consider our earlier \( P \) and \( Q \):

\[
\begin{align*}
P : & \text{ I will eat pizza} \\
Q : & \text{ I will drink soda} \\
P \lor Q : & \text{ I will eat pizza or I will drink soda} \\
\neg (P \lor Q) : & \text{ It is not the case that (either) I will eat pizza or I will drink soda} \\
(\neg P) \land (\neg Q) : & \text{ It is not the case that I will eat pizza, and it is not the case that I will drink soda}
\end{align*}
\]

That these last two statements essentially have the same content, as stated in (1.3), should be intuitive. An actual proof of (1.3) is best given by truth tables, and can be found on page 15.

Next we consider (1.5). This states that \( \neg (P \rightarrow Q) \iff P \land (\neg Q) \). Now we can read \( \neg (P \rightarrow Q) \) as “it is not the case that \( P \rightarrow Q \),” or “\( P \rightarrow Q \) is false.” Recall that there was only one case for which we considered \( P \rightarrow Q \) to be false, which was the case that \( P \) was true but \( Q \) was false, which itself can be translated to \( P \land (\neg Q) \). For our earlier example, the negation of the statement “if I eat pizza then I will drink soda” is the statement “I will eat pizza but (and) I will not drink soda.” While this discussion is correct and may be intuitive, the actual proof (1.5) is by truth table:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>( P \rightarrow Q )</th>
<th>( \neg (P \rightarrow Q) )</th>
<th>( P \land (\neg Q) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

the same
We leave the proof of (1.6) by truth tables to the exercises. Recall that \( P \leftrightarrow Q \) states that we have \( P \) true if and only if we also have \( Q \) true, which we further translated as the idea that we cannot have \( P \) true without \( Q \) true, and cannot have \( Q \) true without \( P \) true. Now \( \sim (P \leftrightarrow Q) \) is the statement that \( P \leftrightarrow Q \) is false, which means that \( P \) is true and \( Q \) false, or \( Q \) is true and \( P \) false, which taken together form the statement \( [P \land (\sim Q)] \lor [Q \land (\sim P)] \), as reflected in (1.6) above. For our example \( P \) and \( Q \) from before, \( P \leftrightarrow Q \) is the statement “I will at pizza if and only if I will drink soda,” the negation of which is “I will eat pizza and not drink soda, or I will drink soda and not eat pizza.”

Another intuitive way to look at these negations is to consider the question of exactly when is someone uttering the original statement lying? For instance, if someone states \( P \land Q \) (or some English equivalent), when are they lying? Since they stated “\( P \) and \( Q \),” it is not difficult to see they are lying exactly when at least one of the statements \( P,Q \) is false, i.e., when \( P \) is false or \( Q \) is false, i.e., when we can truthfully state \( (\sim P) \lor (\sim Q) \). That is the kind of thinking one should employ when examining (1.4), that is \( \sim (P \land Q) \iff (\sim P) \lor (\sim Q) \), intuitively.

### 1.2.3 Equivalent Forms of the Implication

In this subsection we examine two statements which are equivalent to \( P \rightarrow Q \). The first is more important conceptually, and the second is more important computationally. We list them both now before contemplating them further:

\[
P \rightarrow Q \iff (\sim Q) \rightarrow (\sim P)
\]

\[
P \rightarrow Q \iff (\sim P) \lor Q.
\]

We will combine the proofs into one truth table, where we compute \( P \rightarrow Q \), followed in turn by \( (\sim Q) \rightarrow (\sim P) \) and \( (\sim P) \lor Q \).

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( P )</td>
<td>( Q )</td>
<td>( P \rightarrow Q )</td>
<td>( \sim Q )</td>
<td>( \sim P )</td>
<td>( (\sim Q) \rightarrow (\sim P) )</td>
<td>( \sim P )</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
</tr>
</tbody>
</table>

The form (1.7) is important enough that it warrants a name:

**Definition 1.2.2** Given any implication \( P \rightarrow Q \), we call the (logically equivalent) statement \((\sim Q) \rightarrow (\sim P)\) its **contrapositive** (and vice-versa, see below).

In fact, note that the contrapositive of \((\sim Q) \rightarrow (\sim P)\) would be \([\sim (\sim P)] \rightarrow [\sim (\sim Q)]\), i.e., \( P \rightarrow Q \), so \( P \rightarrow Q \) and \((\sim Q) \rightarrow (\sim P)\) are contrapositives of each other.

We have proved that \( P \rightarrow Q \), its contrapositive \((\sim Q) \rightarrow (\sim P)\), and the other form \( (\sim P) \lor Q \) are equivalent using the truth table above, but developing the intuition that these should be equivalent can require some effort. Some examples can help to clarify this.

---

\(^{18}\)Note that here as always we use the inclusive “or,” so when we write “\( P \) is false or \( Q \) is false,” we include the case in which both \( P \) and \( Q \) are false. (See Footnote 8, page 5 for remarks on the exclusive “or.”)
1.2. VALID LOGICAL EQUIVALENCES AS TAUTOLOGIES

\[ P : \text{I will eat pizza} \]
\[ Q : \text{I will drink soda} \]
\[ P \rightarrow Q : \text{If I eat pizza, then I will drink soda} \]
\[ (\sim Q) \rightarrow (\sim P) : \text{If I do not drink soda, then I will not eat pizza} \]
\[ (\sim P) \lor Q : \text{I will not eat pizza, or I will drink soda.} \]

Perhaps more intuition can be found when \( Q \) is a more natural consequence of \( P \). Consider the following \( P, Q \) combination which might be used by parents communicating to their children.

\[ P : \text{you leave your room messy} \]
\[ Q : \text{you get spanked} \]
\[ P \rightarrow Q : \text{if you leave your room messy, then you get spanked} \]
\[ (\sim Q) \rightarrow (\sim P) : \text{if you do not get spanked, then you do (did) not leave your room messy} \]
\[ (\sim P) \lor Q : \text{you do not leave your room messy, or you get spanked.} \]

A mathematical example could look like the following (assuming \( x \) is a “real number,” as discussed later in this text):

\[ P : x = 10 \]
\[ Q : x^2 = 100 \]
\[ P \rightarrow Q : \text{if } x = 10, \text{ then } x^2 = 100 \]
\[ (\sim Q) \rightarrow (\sim P) : \text{if } x^2 \neq 100, \text{ then } x \neq 10 \]
\[ (\sim P) \lor Q : x \neq 10 \text{ or } x^2 = 100. \]

The contrapositive is very important because many theorems are given as implications, but are often used in their logically equivalent, contrapositive forms. However, it is equally important to avoid confusing \( P \rightarrow Q \) with either of the statements \( P \iff Q \) or \( Q \rightarrow P \). For instance, in the second example above, the child may get spanked without leaving the room messy, as there are quite possibly other infractions which would result in a spanking. Thus leaving the room messy does not follow from being spanked, and leaving the room messy is not necessarily connected with the spanking by an “if and only if.” In the last, algebraic example above, all the forms of the statement are true, but \( x^2 = 100 \) does not imply \( x = 10 \). Indeed, it is possible that \( x = -10 \). In fact, the correct bi-implication is \( x^2 = 100 \iff [(x = 10) \lor (x = -10)] \).

1.2.4 Other Valid Equivalences

While negations and equivalent alternatives to the implication are arguably the most important of our valid logical equivalences, there are several others. Some are rather trivial, such as

\[ P \land P \iff P \iff P \lor P. \]  \hspace{1cm} (1.9)

Also rather easy to see are the “commutativities” of \( \land, \lor \) and \( \iff \): \n
\[ P \land Q \iff Q \land P, \quad P \lor Q \iff Q \lor P, \quad P \iff Q \iff Q \iff P. \]  \hspace{1cm} (1.10)

There are also associative rules. The latter was in fact a topic in the previous exercises:

\[ P \land (Q \land R) \iff (P \land Q) \land R \]  \hspace{1cm} (1.11)
\[ P \lor (Q \lor R) \iff (P \lor Q) \lor R. \]  \hspace{1cm} (1.12)
However, it is not so clear when we mix together ∨ and ∧. In fact, these “distribute over each other” in the following ways:

\[ P \land (Q \lor R) \iff (P \land Q) \lor (P \land R), \quad (1.13) \]
\[ P \lor (Q \land R) \iff (P \lor Q) \land (P \lor R). \quad (1.14) \]

We prove the first of these distributive rules below, and leave the other for the exercises.

\[ \begin{array}{cccc|cccc}
 P & Q & R & Q \lor R & P \land (Q \lor R) & P \land Q & P \land R & (P \land Q) \lor (P \land R) \\
 T & T & T & T & T & T & T & T \\
 T & T & F & T & T & T & F & T \\
 T & F & T & T & T & F & T & T \\
 T & F & F & F & F & F & F & F \\
 F & T & T & T & F & F & F & F \\
 F & T & F & T & F & F & F & F \\
 F & F & T & T & F & F & F & F \\
 F & F & F & F & F & F & F & F \\
\end{array} \]

To show that this is reasonable, consider the following:

\[ P : \] I will eat pizza;
\[ Q : \] I will drink cola;
\[ R : \] I will drink lemon-lime soda.

Then our logically equivalent statements become

\[ P \land (Q \lor R) : \] I will eat pizza, and drink cola or lemon-lime soda;
\[ (P \land Q) \lor (P \land R) : \] I will eat pizza and drink cola, or
\[ \quad \] I will eat pizza and drink lemon-lime soda.

Table 1.3, page 22 gives these and some further valid equivalences. It is important to be able to read these and, through reflection and the exercises, to be able to see the reasonableness of each of these. Each can be proved using truth tables.

For instance we can prove that \( P \iff Q \iff (P \rightarrow Q) \land (Q \rightarrow P) \), justifying the choice of the double-arrow symbol \( \iff \):

\[ \begin{array}{cccc|cccc}
 P & Q & P \rightarrow Q & P \rightarrow Q & Q \rightarrow P & (P \rightarrow Q) \land (Q \rightarrow P) \\
 T & T & T & T & T & T \\
 T & F & F & F & T & F \\
 F & T & F & T & F & F \\
 F & F & T & T & T & T \\
\end{array} \]

This was discussed in Example 1.1.4 on page 7.

For another example of such a proof, we next demonstrate the following interesting equivalence:
1.2. **VALID LOGICAL EQUIVALENCES AS TAUTOLOGIES**

\[ P \rightarrow (Q \land R) \iff (P \rightarrow Q) \land (P \rightarrow R) \]

<table>
<thead>
<tr>
<th>P</th>
<th>Q</th>
<th>R</th>
<th>Q \land R</th>
<th>P \rightarrow (Q \land R)</th>
<th>P \rightarrow Q</th>
<th>P \rightarrow R</th>
<th>(P \rightarrow Q) \land (P \rightarrow R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

This should be somewhat intuitive: if \( P \) is to imply \( Q \land R \), that should be the same as \( P \) implying \( Q \) and \( P \) implying \( R \). This equivalence will be (1.33), page 22. According to (1.34) below it, we can replace \( \land \) with \( \lor \) and get another valid equivalence.

Still one must be careful about declaring two statements to be equivalent. These are all ultimately intuitive, but intuition must be informed. For instance, left to the exercises are some valid equivalences which may seem counter-intuitive. These are in fact left off of our Table 1.3 because they are somewhat obscure, but we include them here to illustrate that not all equivalences are transparent. Consider

\begin{align*}
(P \lor Q) \rightarrow R & \iff (P \rightarrow R) \land (Q \rightarrow R), \tag{1.15}
(P \land Q) \rightarrow R & \iff (P \rightarrow R) \lor (Q \rightarrow R). \tag{1.16}
\end{align*}

Upon reflection one can see how these are reasonable. For instance, we can look more closely at (1.15) with the following \( P, Q \) and \( R \):

\begin{align*}
P : & \text{I eat pizza,} \\
Q : & \text{I eat chicken,} \\
R : & \text{I drink cola.}
\end{align*}

Then the left and right sides of (1.15) become

\begin{align*}
(P \lor Q) \rightarrow R : & \text{If I eat pizza or chicken, then I drink cola} \\
(P \rightarrow R) \land (Q \rightarrow R) : & \text{If I eat pizza then I drink cola, and if I eat chicken then I drink cola.}
\end{align*}

In fact (1.16) is perhaps more difficult to see.

At the end of the chapter there will be an optional section for the reader interested in achieving a higher level of symbolic logic sophistication. That section is devoted to finding and proving valid equivalences (and implications as seen in the next section) without relying on truth tables. The technique centers on using a small number of established equivalences to rewrite compound statements into alternative, equivalent forms. With those techniques one can quickly prove (1.15) and (1.16), again without truth tables. It is akin to proving trigonometric identities, or the leap from memorizing single-digit multiplication tables and applying them to several-digit problems.

\[ ^{19} \text{It is a truism in mathematics and other fields that, while one part of learning is discovering what is true, another part is discovering what is not true, especially when the latter seems reasonable at first glance.} \]
CHAPTER 1. MATHEMATICAL LOGIC AND SETS

\[ P \land P \iff P \iff P \lor P \quad (1.17) \]
\[ \sim (\sim P) \iff P \quad (1.18) \]
\[ \sim (P \lor Q) \iff (\sim P) \land (\sim Q) \quad (1.19) \]
\[ \sim (P \land Q) \iff (\sim P) \lor (\sim Q) \quad (1.20) \]
\[ \sim (P \rightarrow Q) \iff P \land (\sim Q) \quad (1.21) \]
\[ \sim (P \leftrightarrow Q) \iff [P \land (\sim Q)] \lor [Q \land (\sim P)] \quad (1.22) \]
\[ P \lor Q \iff Q \lor P \quad (1.23) \]
\[ P \land Q \iff Q \land P \quad (1.24) \]
\[ P \lor (Q \land R) \iff (P \lor Q) \land R \quad (1.25) \]
\[ P \land (Q \land R) \iff (P \land Q) \land R \quad (1.26) \]
\[ P \lor (Q \land R) \iff (P \land Q) \lor (P \land R) \quad (1.27) \]
\[ P \lor (Q \land R) \iff (P \lor Q) \lor (P \land R) \quad (1.28) \]
\[ P \rightarrow Q \iff (\sim P) \lor Q \quad (1.29) \]
\[ P \rightarrow Q \iff (\sim Q) \rightarrow (\sim P) \quad (1.30) \]
\[ P \rightarrow Q \iff [P \land (\sim Q)] \quad (1.31) \]
\[ P \rightarrow Q \iff (\sim P) \rightarrow (\sim Q) \quad (1.32) \]
\[ P \rightarrow (Q \land R) \iff (P \rightarrow Q) \land (P \rightarrow R) \quad (1.33) \]
\[ P \rightarrow (Q \lor R) \iff (P \rightarrow Q) \lor (P \rightarrow R) \quad (1.34) \]
\[ (P \rightarrow Q) \land (Q \rightarrow P) \iff P \rightarrow Q \quad (1.35) \]
\[ (P \rightarrow Q) \land (Q \rightarrow R) \land (R \rightarrow P) \iff (P \rightarrow Q) \land (Q \land R) \quad (1.36) \]

Table 1.3: Table of common valid logical equivalence.

For a glance at the process, we can look at such a proof of the equivalence of the contrapositive: \( P \rightarrow Q \iff (\sim Q) \rightarrow (\sim P) \). To do so, we require (1.29), that \( P \rightarrow Q \iff (\sim P) \lor Q \).

The proof runs as follows:

\[ P \rightarrow Q \iff (\sim P) \lor Q \]
\[ \iff Q \lor (\sim P) \]
\[ \iff [\sim (\sim Q)] \lor (\sim P) \]
\[ \iff (\sim Q) \rightarrow (\sim P). \]

The first line used (1.29), the second commutativity (1.23), the third that \( Q \iff (\sim Q) \) (1.18), and the fourth used (1.29) again but with the part of “\( P \)” played by \( (\sim Q) \) and the part of “\( Q \)” played by \( (\sim P) \). This proof is not much more efficient than a truth table proof, but for (1.15) and (1.16) this technique of proofs without truth tables is much faster. However that technique assumes that the more primitive equivalences used in the proof are valid, and those are ultimately proved using truth tables. The extra section which develops such techniques, namely Section 1.6, is supplemental and not required reading for understanding sufficient symbolic logic to aid in developing the calculus. For that we need only up through Section 1.4.
1.2.5 Circuits and Logic

While we will not develop this next theory deeply, it is worthwhile to consider a short introduction. The idea is that we can model compound logic statements with electrical switching circuits.\(^{20}\) When current is allowed to flow across a switch, the switch is considered “on” when the statement it represents has truth value T and current can flow through the switch, and “off” and not allowing current to flow through when the truth value is F. We can decide if the compound circuit is “on” or “off” based upon whether or not current could flow from one end to the other, based on whether the compound statement has truth value T or F. The analysis can be complicated if the switches are not necessarily independent (\(P\) is “on” when \(\sim P\) is “off” for instance), but this approach is interesting nonetheless.

For example, the statement \(P \lor Q\) is represented by a parallel circuit:

\[
\begin{array}{c}
\text{in} \\
P \\
Q \\
\text{out}
\end{array}
\]

If either \(P\) or \(Q\) is on (T), then the current can flow from the “in” side to the “out” side of the circuit. On the other hand, we can represent \(P \land Q\) by a series circuit:

\[
\begin{array}{c}
\text{in} \\
P \\
Q \\
\text{out}
\end{array}
\]

Of course \(P \land Q\) is only true when both \(P\) and \(Q\) are true, and the circuit reflects this: current can flow exactly when both “switches” \(P\) and \(Q\) are “on.”

It is interesting to see diagrams of some equivalent compound statements, illustrated as circuits. For instance, (1.27), i.e., the distributive-type equivalence

\[
P \land (Q \lor R) \iff (P \land Q) \lor (P \land R)
\]

can be seen as the equivalence of the two circuits below:

\[
\begin{array}{c}
\text{in} \\
P \\
Q \\
R \\
\text{out}
\end{array}
\]

\[
\begin{array}{c}
\text{in} \\
P \\
Q \\
P \\
R \\
\text{out}
\end{array}
\]

\(^{20}\)Technically a circuit would allow current to flow from a source, through components and back to the source. Here we only show part of the possible path. We will encounter some complete circuits later in the text.
In both circuits, we must have $P$ “on,” and also either $Q$ or $R$ for current to flow. Note that in the second circuit, $P$ is represented in two places, so it is either “on” in both places, or “off” in both places. Situations such as these can complicate analyses of switching circuits but this one is relatively simple.

We can also represent negations of simple statements. To represent $\sim P$ we simply put “$\sim P$” into the circuit, where it is “on” if $\sim P$ is true, i.e., if $P$ is false. This allows us to construct circuits for the implication by using (1.29), i.e., that $P \rightarrow Q \iff (\sim P) \lor Q$:

![Diagram of circuit for $P \rightarrow Q$]

We see that the only time the circuit does not flow is when $P$ is true ($\sim P$ is false) and $Q$ is false, so this matches what we know of when $P \rightarrow Q$ is false. From another perspective, if $P$ is true, then the top part of the circuit won’t flow so $Q$ must be true, for the whole circuit to be “on,” or “true.”

When negating a whole circuit it gets even more complicated. In fact, it is arguably easier to look at the original circuit and simply note when current will not flow. For instance, we know $\sim (P \land Q) \iff (\sim P) \lor (\sim Q)$, so we can construct $P \land Q$:

![Diagram of circuit for $P \land Q$]

and note that it is off exactly when either $P$ is off or $Q$ is off. We then note that that is exactly when the circuit for $(\sim P) \lor (\sim Q)$ is on.

![Diagram of circuit for $\sim P \lor \sim Q$]

There are, in fact, electrical/mechanical means by which one can take a circuit and “negate” its truth value, for instance with relays or reverse-position switch levers, but that subject is more complicated than we wish to pursue here.

It is interesting to consider $P \iff Q$ as a circuit. It will be “on” if $P$ and $Q$ are both “on” or both “off,” and the circuit will be “off” if $P$ and $Q$ do not match. Such a circuit is actually used commonly, such as for a room with two light switches for the same light. To construct such a circuit we note that

$$P \iff Q \iff (P \rightarrow Q) \land (Q \rightarrow P) \iff [(\sim P) \lor Q] \land [(\sim Q) \lor P]$$

We will use the last form to draw our diagram:
The reader is invited to study the above diagram to be convinced it represents \( P \iff Q \), perhaps most easily in the sense that, “you can not have one (\( P \) or \( Q \)) without the other, but you can have neither.” While the above diagram does represent \( P \iff Q \) by the more easily diagrammed \( [(\sim P) \lor Q] \land [(\sim Q) \lor P] \), it also suggests another equivalence, since the circuits below seems to be functionally equivalent. In the first, we can add two more wires to replace the “center” wire, and also switch the \( \sim Q \) and \( P \), since \( (\sim Q) \lor P \) is the same as \( P \lor (\sim Q) \):

This circuit represents \( [(\sim P) \land (\sim Q)] \lor [P \land Q] \), and so we have (as the reader can check)

\[
P \iff Q \iff [(\sim P) \land (\sim Q)] \lor [P \land Q],
\]

which could be added to our previous Table 1.3, page 22 of valid equivalences. It is also consistent with a more colloquial way of expressing \( P \iff Q \), such as “neither or both.”

Incidentally, the circuit above is used in applications where we wish to have two switches within a room which can both change a light (or other device) from on to off or vice versa. When switch \( P \) is “on,” switch \( Q \) can turn the circuit on or off by matching \( P \) or being its negation. Similarly when \( P \) is “off.” Mechanically this is accomplished with “single pole, double throw (SPDT)” switches.

In the above, the switch \( P \) is in the “up” position when \( P \) is ‘true, and “down” when \( P \) is false. Similarly with \( Q \).
CHAPTER 1. MATHEMATICAL LOGIC AND SETS

Because there are many possible “mechanical” diagrams for switching circuits, reading and writing such circuits is its own skill. However, for many simpler cases there is a relatively easy connection to our symbolic logic.

1.2.6 The Statements \( T \) and \( F \)

Just as there is a need for zero in addition, we have use for a symbol representing a statement which is always true, and for another symbol representing a statement which is always false. For convenience, we will make the following definitions:

**Definition 1.2.3** Let \( T \) represent any compound statement which is a tautology, i.e., whose truth value is always \( T \). Similarly, let \( F \) represent any compound statement which is a contradiction, i.e., whose truth value is always \( F \).

We will assume there is a universal \( T \) and a universal \( F \), i.e., statements which are respectively true regardless of any other statements’ truth values, and false regardless of any other statements’ truth values. In doing so, we consider any tautology to be logically equivalent to \( T \), and any contradiction similarly equivalent to \( F \).\(^{21}\)

So, for any given \( P_1 \cdots P_n \), we have that \( T \) is exactly that statement whose column in the truth table consists entirely of \( T \)'s, and \( F \) is exactly that statement whose column in the truth table consists entirely of \( F \)'s. For example, we can write

\[
P \lor (\neg P) \iff T; \quad (1.38)
\]

\[
P \land (\neg P) \iff F. \quad (1.39)
\]

These are easily seen by observing the truth tables.

<table>
<thead>
<tr>
<th>( P )</th>
<th>( \neg P )</th>
<th>( P \lor (\neg P) )</th>
<th>( P \land (\neg P) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>( F )</td>
<td>( T )</td>
<td>( F )</td>
</tr>
<tr>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
<td>( F )</td>
</tr>
</tbody>
</table>

We see that \( P \lor (\neg P) \) is always true, and \( P \land (\neg P) \) is always false. Anything which is always true we will dub \( T \), and anything which is always false we will call \( F \). In the table above, the third column represents \( T \), and the last column represents \( F \).

From the definitions we can also eventually get the following.

\[
P \lor T \iff T \quad (1.40)
\]

\[
P \land T \iff P \quad (1.41)
\]

\[
P \lor F \iff P \quad (1.42)
\]

\[
P \land F \iff F. \quad (1.43)
\]

\(^{21}\)In fact it is not difficult to see that all tautologies are logically equivalent. Consider the tautologies \( P \lor (\neg P) \), \( (P \rightarrow Q) \rightarrow [\neg (Q) \rightarrow (\neg P)] \), and \( R \rightarrow R \). A truth table for all three must contain independent component statements \( P, Q, R \), and the abridged version of the table would look like

<table>
<thead>
<tr>
<th>( P )</th>
<th>( Q )</th>
<th>( R )</th>
<th>( P \lor (\neg P) )</th>
<th>( (P \rightarrow Q) \rightarrow [\neg (Q) \rightarrow (\neg P)] )</th>
<th>( R \rightarrow R )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
</tr>
<tr>
<td>( T )</td>
<td>( T )</td>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
</tr>
<tr>
<td>( T )</td>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
</tr>
<tr>
<td>( T )</td>
<td>( F )</td>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
</tr>
<tr>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
</tr>
<tr>
<td>( F )</td>
<td>( T )</td>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
</tr>
<tr>
<td>( F )</td>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
</tr>
<tr>
<td>( F )</td>
<td>( F )</td>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
</tr>
</tbody>
</table>

So when all possible underlying independent component statements are included, we see the truth table columns of these tautologies are indeed the same (all \( T \)'s!). Similarly all contradictions are equivalent.
To demonstrate how one would prove these, we prove here the first two, (1.40) and (1.41), using a truth table. Notice that all entries for $T$ are simply $T$:

\[
\begin{array}{c|c|c|c}
P & T & P \lor T & P \land T \\
T & T & T & T \\
F & T & T & F \\
\end{array}
\]

Equivalence (1.40) is demonstrated by the equivalence of the second and third columns, while (1.41) is shown by the equivalence of the first and fourth columns. The others are left as exercises.

These are also worth reflecting upon. Consider the equivalence $P \land T \iff P$. When we use $\land$ to connect $P$ to a statement which is always true, then the truth of the compound statement only depends upon the truth of $P$. There are similar explanations for the rest of (1.40)–(1.43).

Some other interesting equivalences involving these are the following:

\[
T \iff P \iff P \\
P \iff F \iff \sim P
\]

We leave the proofs of these for the exercises. These are in fact interesting to interpret. The first says that if a true statement implies $P$, that is the same as in fact having $P$. The second says that if $P$ implies a false statement, that is the same as having $\sim P$, i.e., as having $P$ false. Both types of reasoning are useful in mathematics and other disciplines.

If a statement contains only $T$ or $F$, then in fact that statement itself must be a tautology ($T$) or a contradiction ($F$). This is because there is only one possible combination of truth values. For instance, consider the statement $T \iff F$, which is a contradiction. One proof is in the table:

\[
\begin{array}{c|c|c}
T & F & T \iff F \\
T & F & F \\
\end{array}
\]

Since the component statement $T \iff F$ always has truth value $F$, it is a contradiction. Thus $T \iff F \iff F$.

\section*{Exercises}

Some of these were solved within the section. It is useful to attempt them here again, in the context of the other problems. Unless otherwise specified, all proofs should be via truth tables.

1. Prove (1.18): $\sim (\sim P) \iff P$.

2. Prove (1.32):

   \[ P \iff Q \iff (\sim P) \iff (\sim Q). \]

3. Prove the logical equivalence of the contrapositive (1.30):

   \[ P \iff (\sim Q) \iff (\sim P). \]

4. Prove (1.29): $P \iff Q \iff (\sim P) \lor Q$.

5. Show that $P \iff Q$ and $Q \iff P$ are not equivalent.

6. Prove De Morgan’s Laws (1.19) and (1.20), which are listed again below:

   (a) $\sim (P \lor Q) \iff (\sim P) \land (\sim Q)$.

   (b) $\sim (P \land Q) \iff (\sim P) \lor (\sim Q)$.

7. Use truth tables to prove the distributive-type laws (1.27) and (1.28):

   (a) $P \land (Q \lor R) \iff (P \land Q) \lor (P \land R)$.

   (b) $P \lor (Q \land R) \iff (P \lor Q) \land (P \lor R)$.

8. Repeat the previous problem but using circuit diagrams.
9. Prove (1.33): 

\[ P \rightarrow (Q \land R) \leftrightarrow (P \rightarrow Q) \land (P \rightarrow R). \]

10. Prove (1.34): 

\[ P \rightarrow (Q \lor R) \leftrightarrow (P \rightarrow Q) \lor (P \rightarrow R). \]

11. Prove (1.21): 

\[ \sim (P \rightarrow Q) \iff P \land (\sim Q). \]

12. Prove (1.35), which we write below as 

\[ P \iff (P \rightarrow Q) \iff (P \rightarrow R). \]

13. Prove (1.22): 

\[ \sim (P \rightarrow Q) \iff (P \land (\sim Q)) \lor (Q \land (\sim P)). \]


(a) Construct a truth table for 

\[ P \text{ XOR } Q. \]

(b) Compare to the previous problem. Can you make a conclusion?

(c) Find an expression for 

\[ P \text{ XOR } Q \text{ using } P, Q, \sim, \land \text{ and } \lor. \]

15. Prove that 

\[ (P \lor Q) \rightarrow (P \land Q) \text{ is equivalent to } P \rightarrow Q. \]

How would you explain in words why this is reasonable? (Perhaps you can think of a colloquial way to verbalize the statement so it will sound equivalent to 

\[ P \rightarrow Q. \]

16. Prove (1.31). How would you explain in words why this is reasonable?

17. Prove the following:

(a) (1.40): 

\[ P \lor T \iff T. \]

(b) (1.41): 

\[ P \land T \iff P. \]

(c) (1.42): 

\[ P \lor \mathcal{F} \iff P. \]

(d) (1.43): 

\[ P \land \mathcal{F} \iff \mathcal{F}. \]

18. For each of the following, find a simple, equivalent statement, using truth tables if necessary.

(a) \[ T \lor T \]

(b) \[ \mathcal{F} \lor \mathcal{F} \]

(c) \[ T \lor \mathcal{F} \]

(d) \[ \mathcal{T} \land \mathcal{T} \]

(e) \[ \mathcal{T} \land \mathcal{F} \]

(f) \[ T \land \mathcal{F} \]

(g) \[ T \rightarrow \mathcal{T} \]

(h) \[ \mathcal{F} \rightarrow \mathcal{F} \]

(i) \[ T \rightarrow \mathcal{F} \]

(j) \[ \mathcal{F} \rightarrow \mathcal{T} \]

19. Repeat the previous exercise for the following:

(a) \[ T \rightarrow P \]

(b) \[ \mathcal{F} \rightarrow P \]

(c) \[ P \rightarrow T \]

(d) \[ P \vdash \mathcal{F} \]

(e) \[ P \rightarrow (\sim P) \]

(f) \[ (P \rightarrow \sim P) \lor (\sim P \rightarrow \mathcal{F}) \]

20. Prove the associative rules (1.25) and (1.26), page 22.

21. Prove (1.36), page 22.

22. Show 

\[ (P \lor Q) \rightarrow R \iff [(P \rightarrow R) \land (Q \rightarrow R)]. \]

Try to explain why this makes sense.

23. Show 

\[ (P \land Q) \rightarrow R \iff [(P \rightarrow R) \lor (Q \rightarrow R)]. \]

(This is not so easily explained as is the previous exercise.)

24. There is a notion in logic theory regarding “strong” versus “weak” statements, the stronger ones claiming in a sense more information regarding the underlying statements such as 

\[ P, Q. \]

For instance, 

\[ P \land Q \]

is considered “stronger” than 

\[ P \lor Q, \]

because 

\[ P \land Q \]

tells us more about 

\[ P, Q \] (both are declared true) than 

\[ P \lor Q \] (at least one is true but both may be). Similarly 

\[ P \leftrightarrow Q \]

is stronger than 

\[ P \rightarrow Q. \]

Each of the following statements may appear “strong” but in fact give little interesting content regarding 

\[ P, Q. \]

Construct a truth table for each and use the truth tables to then explain why they are not terribly “interesting” statements to make about 

\[ P, Q. \]

(Hint: what are these equivalent to?)

(a) \[ (P \rightarrow Q) \lor (Q \rightarrow P) \]

(b) \[ (P \rightarrow Q) \rightarrow (P \leftrightarrow Q) \]
1.3 Valid Implications and Arguments

Most theorems in this text are in the form of implications, rather than the more rigid equivalences of the last section. Indeed, our theorems are usually of the form “hypothesis implies conclusion.” So we have need of an analog to our valid equivalences, namely a notion of valid implications.

1.3.1 Valid Implications Defined

Our definition of valid implications is similar to our previous definition of valid equivalences:

**Definition 1.3.1** Suppose that $R$ and $S$ are compound statements of some independent component statements $P_1, \cdots, P_n$. If $R \implies S$ is a tautology (always true), then we write

$$ R \implies S, \quad (1.46) $$

which we then call a valid logical implication.\(^{22}\)

**Example 1.3.1** Perhaps the simplest example is the following: $P \implies P$. This seems obvious enough on its face. It can be proved using a truth table (note the vacuous case):\(^{23}\)

<table>
<thead>
<tr>
<th>$P$</th>
<th>$P$</th>
<th>$P \implies P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

Thus we see that there are logical implications which are tautologies. A slightly more complicated—and very instructive—example is the following:

**Example 1.3.2** The following is a valid implication:

$$(P \land Q) \implies P. \quad (1.47)$$

To prove this, we will use a truth table to show that the following is a tautology:

$$(P \land Q) \implies P.$$

<table>
<thead>
<tr>
<th>$P$</th>
<th>$Q$</th>
<th>$P \land Q$</th>
<th>$P$</th>
<th>$(P \land Q) \implies P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

Notice that three of the four cases have the implication true vacuously.

Another intuitive example is fairly easy to interpret: if $P$ and $Q$ are true, then (of course) $P$ is true. Another intuitive example follows, basically stating that if we have bi-implication then we have implication.

---

22In this text we will differentiate between implications as statements, such as $P \implies Q$, which may be true or false, and valid implications which are declarations that a particular implication is always true. For example $R \implies S$ means $R \implies S$ is a tautology. (We similarly differentiated $\iff$ from $\implies$.)

23We could also show that $P \implies P$ is a tautology by way of previously proved results. For instance, with $P \implies Q \iff (\sim P) \lor Q$ ((1.29), page 22), with the part of $Q$ played by $P$, we have

$$ P \implies P \iff (\sim P) \lor P \iff T, $$

the second statement being in effect our original example of a tautology (Example 1.1.9, page 11).
Example 1.3.3 The following is a valid implication:

\[ P \iff Q \implies P \rightarrow Q. \] (1.48)

As before, we prove that replacing \( \implies \) with \( \rightarrow \) gives us a tautology.

<table>
<thead>
<tr>
<th>( P )</th>
<th>( Q )</th>
<th>( P \iff Q )</th>
<th>( P \rightarrow Q )</th>
<th>( (P \rightarrow Q) \implies (P \rightarrow Q) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

Also as before, we see the importance of the vacuous cases in the final implication. In fact, this is just an application of the previous implication (1.47), if we remember that \( P \iff Q \) is equivalent to \( (P \rightarrow Q) \land (Q \rightarrow P) \):

\[ \llbracket (P \rightarrow Q) \land (Q \rightarrow P) \rrbracket \implies \llbracket P \rightarrow Q \rrbracket. \]

The quotes indicate what roles in (1.47) are played by the parts of (1.48).

Another interesting valid implication is given next. (The reader should reflect on its apparent meaning.)

Example 1.3.4 \( (P \rightarrow Q) \land (Q \rightarrow R) \implies (P \rightarrow R) \).

Note that to prove this, we must show that the following statement is a tautology:

\[ [(P \rightarrow Q) \land (Q \rightarrow R)] \implies (P \rightarrow R). \]

<table>
<thead>
<tr>
<th>( P )</th>
<th>( Q )</th>
<th>( R )</th>
<th>( P \rightarrow Q )</th>
<th>( Q \rightarrow R )</th>
<th>( (P \rightarrow Q) \land (Q \rightarrow R) )</th>
<th>( P \rightarrow R )</th>
<th>( [(P \rightarrow Q) \land (Q \rightarrow R)] \implies (P \rightarrow R) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

At this stage in the development it is perhaps best to check such things using truth tables, however unwieldy they can be. Of course with practice comes intuition, informed memory and many shortcuts, but for now we will use this brute force method of truth tables to determine if an implication is valid.

For a simple algebraic perspective of the difference between a valid implication and a valid equivalence, consider the following list of algebraic facts:

\[ x = 5 \implies x^2 = 25, \]
\[ x = -5 \implies x^2 = 25, \]
\[ x^2 = 25 \iff (x = 5) \lor (x = -5). \]
1.3. VALID IMPLICATIONS AND ARGUMENTS

Knowing \( x = 5 \) is not \textit{equivalent} to knowing \( x^2 = 25 \). That is because there is an alternative explanation for \( x^2 = 25 \), namely that perhaps \( x = -5 \). But it is true that knowing \( x = 5 \) \textit{implies} knowing—at least in principle—that \( x^2 = 25 \) (just as knowing \( x = -5 \) implies knowing \( x^2 = 25 \)). If an equivalence is desired, a valid one is that knowing \( x^2 = 25 \) is equivalent to knowing that \( x \) must be either 5 or \(-5\).

Later in the text we will briefly focus on algebra in earnest, bringing our symbolic logic to bear on that topic. In algebra (and in calculus) it is often important to know when we have an equivalence and when we have only an implication. For some algebraic problems, the implication often means we need to check our answer while the equivalence means we do not. For an example of this phenomenon, consider

\[
\sqrt{x + 2} = x \iff x + 2 = x^2 \iff 0 = x^2 - x - 2 \iff 0 = (x - 2)(x + 1) \iff (x - 2 = 0) \lor (x + 1 = 0) \iff (x = 2) \lor (x = -1).
\]

We lost the equivalence at the first step, and so we can only conclude from the logic that

\[
\sqrt{x + 2} = x \implies (x = 2) \lor (x = -1).
\]

All this tells us is that \textit{if} there is a a number \( x \) so that \( \sqrt{x + 2} = x \), then the number must be either \( x = 2 \) or \( x = -1 \) (or perhaps both work; recall that we always interpret \( \lor \) inclusively). When we check \( x = 2 \) in the original equation, we get \( \sqrt{4} = 2 \), which is true. However, \( x = -1 \) gives \( \sqrt{-1} = -1 \), which is not true. Since we have now solved the original equation we can say that\(^{24}\)

\[
\sqrt{x + 2} = x \iff x = 2.
\]

For another example, consider how one can solve a linear equation:

\[
2x + 1 = 3 \iff 2x = 2 \iff x = 1.
\]

Here we subtracted 1 from both sides, and then divided by 2, neither of which break the logical equivalence. We do not have to check the answer (unless we believe our arithmetic or reasoning may be faulty). In Chapter 2 we will delve more deeply into algebra.

1.3.2 Partial List of Valid Implications

Table 1.4, page 32 lists some basic valid equivalences and implications. All can be proved using truth tables. However, it is important to learn to recognize validity without always resorting to truth tables. Each can be viewed in light of English examples. Still, it is the rigorous \textit{mathematical} framework which gives us the precise rules for rewriting and analyzing statements.\(^{25}\)

It is useful to see why (1.49)–(1.58) are not equivalences.\(^{26}\) For instance a little reflection should make clear that\(^{27}\) \( P \iff P \land Q \) (unless there is some underlying relationship between \( P \) and \( Q \) which is not stated), and so we cannot replace \( \iff \) with \( \iff \) in (1.49). Similarly, in (1.52) having \( P \rightarrow R \) in itself says nothing about \( Q \), so there is no reason to believe \( (P \rightarrow Q) \land (Q \rightarrow R) \) is implied by \( P \rightarrow R \).

Implicit in the above discussion is the fact that having \( \iff \) is the same as simultaneously having both \( \implies \) and \( \iff \). Put another way, \( R \iff S \) is the same as collectively having

\(^{24}\)Later, in Chapter 2 we will define \( \sqrt{z} \) to be only the nonnegative square root of \( z \), assuming \( z \geq 0 \) lest \( \sqrt{z} \) be undefined, at least as a real number.

\(^{25}\)Similarly we use the rules of algebra to rewrite and analyze equations in hopes of solving for the variables.

\(^{26}\)The exceptions in the table are (1.60) and (1.61), which are valid if we replace \( \iff \) with \( \iff \), as was discussed in the previous section. See (1.45), page 27 and the exercises of that section.

\(^{27}\)Note that it is common to negate a statement by including a “slash” through the main symbol, as in \( \neg (x = 3) \iff x \neq 3 \). What we mean by \( R \iff S \) is that it is not true that \( R \rightarrow S \) is a tautology.
\[ P \land Q \implies P \] (1.49)
\[ P \implies P \lor Q \] (1.50)
\[ P \iff Q \implies P \implies Q \] (1.51)
\[ (P \implies Q) \land (Q \implies R) \implies P \implies R \] (1.52)
\[ (P \iff Q) \land (Q \iff R) \implies (P \implies R) \] (1.53)
\[ (P \implies Q) \land (Q \implies R) \implies (P \iff R) \] (1.54)
\[ (P \iff Q) \land P \implies Q \] (1.55)
\[ (P \implies Q) \land (\sim Q) \implies \sim P \] (1.56)
\[ (P \lor Q) \land (\sim Q) \implies P \] (1.57)
\[ P \implies (\sim P) \implies \sim P \] (1.58)
\[ P \implies \mathcal{F} \implies \sim P \] (1.59)
\[ T \implies P \implies P \] (1.60)
\[ T \implies P \implies P \] (1.61)

Table 1.4: Table of Valid Logical Implications. If we replace \( \implies \) with \( \iff \) in each of the above (perhaps enclosing each side in brackets \([\ldots]\)), we would have tautologies.

Both \( R \implies S \) and \( S \implies R \). In fact it is important to note that all of the valid logical equivalences, for instance in Table 1.3, page 22 can be also considered to be combinations of two valid implications, one with \( \implies \), and the other with \( \iff \), replacing \( \iff \). We do not list them all here, but rather list the most commonly used implications which are not equivalences, except for the last three in the table.

1.3.3 Fallacies and Valid Arguments

The name fallacy is usually reserved for typical faults in arguments that we nevertheless find persuasive. Studying them is therefore a good defense against deception.

—Peter Suber, Department of Philosophy, Earlham College, Richmond, Indiana, 1996.

Here we look at some classical argument styles, some of which are valid, and some of which are invalid and therefore called fallacies (whether or not they may seem persuasive at first glance). The valid styles will mostly mirror the valid logical implications of Table 1.4.

A common method for diagramming simple arguments is to have a horizontal line separating the premises from the conclusions. Usually we will have multiple premises and a single conclusion. For style considerations, the conclusion is often announced with the symbol \( \therefore \), which is read “therefore.”

In this subsection we look at several of these arguments, both valid and fallacious. Many are classical, with classical names. We will see how to analyze arguments for validity. In all cases here, it will amount to determining if a related implication is valid.

---

28Note that \( R \iff S \) would be interpreted as \( S \implies R \). We will not make extensive use of “\( \iff \).”
29Valid argument forms are also called rules of inference.
30Premises are also called hypotheses. The singular forms are premise and hypothesis.
31In fact most texts use either the horizontal line or the symbol \( \therefore \), but not both. We use both to emphasize where the hypotheses end and the conclusion begins.
Example 1.3.5 Our first example we consider is the argument form which is classically known as modus ponens, or law of detachment. It is outlined as follows:

\[
\begin{align*}
P \rightarrow Q \\
P \\
\therefore Q
\end{align*}
\]

The idea is that if we assume \( P \rightarrow Q \) and \( P \) are true, then we must conclude that \( Q \) is also true. This is ultimately an implication. The key is that checking to see if this is valid is the same as checking to see if \((P \rightarrow Q) \wedge P \Rightarrow Q\), i.e., that \( [(P \rightarrow Q) \wedge P] \rightarrow Q \) is a tautology. We know this to be the case already, as this is just (1.56), though we should prove this by producing the relevant truth table to show that \( [(P \rightarrow Q) \wedge P] \rightarrow Q \) is indeed a tautology, i.e., has truth value T for all cases of truth values of \( P,Q \):

\[
\begin{array}{cccccccc}
P & Q & P \rightarrow Q & (P \rightarrow Q) \wedge P & Q & [(P \rightarrow Q) \wedge P] \rightarrow Q \\
T & T & T & T & T & T \\
T & F & F & F & F & T \\
F & T & T & F & T & T \\
F & F & T & F & T & T
\end{array}
\]

Thus to test the validity of an argument is to test whether or not the argument, written as an implication, is a tautology. This gives us a powerful, computational tool to analyze the classical argument styles. It also connects some of our symbolic logic to this style of diagramming arguments, so the intuition of these two flavors of logic can illuminate each other.

To repeat and emphasize the criterion for validity we list the following definition:

Definition 1.3.2 A valid argument is one which, when diagrammed as an implication, represents a tautology. In other words, if the premises are \( P_1, P_2, \cdots, P_m \) and the conclusion is \( Q \) (where \( P_1, P_2, \cdots, P_m \) and \( Q \) are compound statements based upon some underlying independent statements \( P_1, \cdots, P_n \)), then the argument is valid if and only if

\[
P_1 \wedge P_2 \wedge \cdots \wedge P_m \Rightarrow Q,
\]

i.e., if and only if \( [P_1 \wedge P_2 \wedge \cdots \wedge P_m] \rightarrow Q \) is a tautology. If not, then the argument is a called a fallacy.

Note that the validity of any argument does not depend upon the truth or falsity of the conclusion. Indeed the modus ponens argument in Example 1.3.5 is perfectly valid, regardless of whether or not \( Q \) is true. That is because we do not know—or even ask for purposes of discovering if the logic is valid—whether or not the premises are true. What we do know is that, if the premises are true, then so is the conclusion. In other words, the statement \( [(P \rightarrow Q) \wedge P] \rightarrow Q \) is always true. (If one or more of the premises are false, the implication is true vacuously.)

---

32 Modus ponens is short for modus ponendo ponens, which is Latin for “the way that affirms by affirming.” It is important enough that it has been extensively studied through the ages, and thus has many names, another being “affirmation argument.” As for “law of detachment,” it is pointed out in J.E. Rubin’s Mathematical Logic: Applications and Theory (Saunders, 1990), that the idea is that we can validly “detach” the consequent \( Q \) of the conditional \( P \rightarrow Q \) when we also assume the antecedent \( P \).
CHAPTER 1. MATHEMATICAL LOGIC AND SETS

One example often used to shed light on the law of detachment above, and other argument styles as well, uses the following choices for $P$ and $Q$.

\[ P : \text{It rained} \]
\[ Q : \text{The ground is wet} \]

The argument above could then be diagrammed again but using the words represented by $P, Q$:

If it rained, then the ground is wet.

\[
\text{It rained.} \]
\[ \therefore \text{The ground is wet.} \]

This is a perfectly valid argument, meaning that if we accept the premises we must accept the conclusion. In other words, the logic is flawless. That said, one need not necessarily accept the conclusion just because the argument is valid, since one can always debate the truthfulness of the premises. Again the key is that the logic here is valid, even if the premises may be faulty.  

Next we look at an example of an invalid argument, i.e., a fallacy. The following is called the fallacy of the converse:  

**Example 1.3.6** Show that the following argument is a fallacy:

\[
P \rightarrow Q \\
Q \therefore P \text{ (Invalid)}
\]

As before, we analyze the corresponding implication, in this case \([(P \rightarrow Q) \land Q] \rightarrow P\), with a truth table:

<table>
<thead>
<tr>
<th>$P$</th>
<th>$Q$</th>
<th>$P \rightarrow Q$</th>
<th>$(P \rightarrow Q) \land Q$</th>
<th>$P$</th>
<th>$[(P \rightarrow Q) \land Q] \rightarrow P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

It is always useful to review an invalid argument to see which conditions were problematic. In the third row of our truth table, $P \rightarrow Q$ is vacuously true, and $Q$ is true so the premises hold true, but the conclusion $P$ is false (which was why $P \rightarrow Q$ was vacuously true!). From a more common-sense standpoint, while $P \rightarrow Q$ is assumed, $P$ may not be the only condition which forces $Q$ to be true. (If it were, we would instead have $P \leftarrow \rightarrow Q$.) Consider again our previous choices for $P$ and $Q$:

If it rained, then the ground is wet.

\[
\text{The ground is wet.} \]
\[ \therefore \text{It rained. (Invalid)} \]

---

\[ ^{33} \]In mature philosophical discussions, the logic is rarely in question because the valid models of argument are well known. When a conclusion seems unacceptable or just questionable, it is usually the premises which then come under scrutiny.

\[ ^{34} \]The *converse* of an implication $R \rightarrow S$ is the statement $S \rightarrow R$ (or $R \rightleftharpoons S$ if we want to preserve the order). An implication and its converse are not logically equivalent, as a quick check of their truth tables would reveal. However, it is a common mistake to forget which direction an implication follows, or to just be careless and mistake an implication for a bi-implication. The "fallacy of the converse" refers to a state of mind where one mistakenly believes the converse true, based upon the assumption that the original implication is true. (Note that if we mistakenly replace $P \rightarrow Q$ with $Q \rightarrow P$ in Example 1.3.6, the new argument would be valid. In fact it is *modus ponens*.)
Even if the premises are correct in Example 1.3.6, the ground being wet does not guarantee that it rained. Perhaps it is wet from dew, or a sprinkler, or flooding from some other source. Here one can accept the premises, but the conclusion given above is not valid.

There is a subtle—perhaps difficult—general point in this subsection which bears repeating: the truth table associated with an argument reflects the validity or invalidity of the logic of the argument (i.e., the validity of the corresponding implication), regardless of the truthfulness of the premises. Indeed, note how the truth table for the valid form *modus ponens* of Example 1.3.5 (page 33) contains cases where the premises, \( P \rightarrow Q \) and \( P \), can have truth value \( F \) as well as \( T \).\(^{35}\)

For completeness, we mention that some use the adjective *sound* to describe an argument which is not only valid, but whose premises (and therefore conclusions) are in fact true. Of course in reality those are usually the arguments which we seek, but (arguably) one must first understand validity before probing the soundness of arguments, and so for this text, we are mostly interested in abstract, valid arguments, and worry about soundness only in context.

The next example is also very common. It is a valid form of argument often known by its Latin name *modus tollens*.\(^{36}\)

**Example 1.3.7** Analyze the following (modus tollens) argument.

\[
P \rightarrow Q \\
\sim Q \\
\therefore \sim P
\]

As before, we analyze the following associated implication (which we leave as a single-arrow implication until we establish it is a tautology):

\[
(P \rightarrow Q) \land (\sim Q) \rightarrow (\sim P).
\]

<table>
<thead>
<tr>
<th>( P )</th>
<th>( Q )</th>
<th>( P \rightarrow Q )</th>
<th>( \sim Q )</th>
<th>( (P \rightarrow Q) \land (\sim Q) )</th>
<th>( \sim P )</th>
<th>( (P \rightarrow Q) \land (\sim Q) \rightarrow (\sim P) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
<td>( F )</td>
<td>( F )</td>
<td>( F )</td>
<td>( T )</td>
</tr>
<tr>
<td>( T )</td>
<td>( F )</td>
<td>( F )</td>
<td>( T )</td>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
</tr>
<tr>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
</tr>
<tr>
<td>( F )</td>
<td>( F )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
<td>( T )</td>
</tr>
</tbody>
</table>

That the final column is all \( T \)'s thus establishes its validity. In fact, we see that the argument above is just a re-diagrammed version of (1.57), page 32.

A short chapter could be written just on the insights which can be found studying the above *modus tollens* argument. For instance, for a couple of reasons one could make the case that *modus tollens* and *modus ponens* are the same type of argument. We will see one of these reasons momentarily, but first we will look at *modus tollens* by itself. Inserting our previous \( P \) and \( Q \) into this form, we would have

\(^{35}\) Another, perhaps more subtle point here is the extensive role of the vacuous cases in the underlying implication of an argument. If the premises are not true, then the conclusion can not contradict them. This gives rise to some strange forms of argument indeed, for the premises can be self-contradictory and therefore, taken as a group of statements joined by \( \land \), can be equivalent to \( \mathcal{F} \). (Recall \( \mathcal{F} \rightarrow \mathcal{P} \) is a tautology.) However, the only practical uses of arguments come when we know the premises to be true, or we think they are false and demonstrate it by showing the valid conclusions they imply are demonstrably false. The latter use is often called *proof by contradiction* or *indirect proof*, but there are many structures which use the same idea. *Modus tollens* (Example 1.3.7) is one permutation of the idea behind indirect proof.

\(^{36}\) Short for *modus tollendo tollens*, Latin for “the way that denies by denying.” It is also called “denying the consequent,” which contrasts it to “affirming the consequent,” another name for the fallacy of the converse, Example 1.3.6, page 34. As the reader can deduce, most of these common arguments—valid or not—have many names, inspired by different contexts and considerations. Computationally they are simple enough, once seen as implications to be analyzed and found to be tautologies (in cases of valid implications) or non-tautologies (in cases of fallacies).
If it rained, then the ground is wet.

The ground is not wet.

∴ It did not rain.

The validity of the above argument should be intuitive. A common way of explaining it is that it must not have rained, because (first premise) if it had rained the ground would be wet, and (second premise) it is not wet. Of course that explanation is probably no simpler than just reading the argument as it stands.

Another way to look at it is to recall the equivalence of the implication to the contrapositive ((1.30), page 22 and elsewhere):

\[ P \rightarrow Q \iff (\sim Q) \rightarrow (\sim P). \]

Thus we can replace in the modus tollens argument the first premise, \( P \rightarrow Q \), with its (equivalent) contrapositive:

\[ (\sim Q) \rightarrow (\sim P) \]

\[ \sim Q \]

∴ \( \sim P \)

This is valid by modus ponens (Example 1.3.5, page 33), with the part of \( P \) there played by \( \sim Q \) here, and the part of \( Q \) by \( \sim P \). In fact the next valid argument form further unifies modus ponens and modus tollens, as will be explained below, though this next form is interesting in its own right.

**Example 1.3.8** Consider the following form of argument, called disjunctive syllogism, which is valid.\(^{37}\)

\[ P \lor Q \]

\[ \sim P \]

∴ \( Q \)

The proof of this is left as an exercise. To prove this one needs to show

\[ (P \lor Q) \land (\sim P) \Rightarrow Q, \]

that is, to show \([ (P \lor Q) \land (\sim P) ] \rightarrow Q \) to be a tautology. Of course the idea of this argument style is that when we assume “\( P \) or \( Q \)” to be true, and then further assume \( P \) is false (by assuming “\( \sim P \)” is true), we are forced to conclude \( Q \) must be true. Note that this is just a re-diagrammed version of (1.58), page 32 except with \( P \) and \( Q \) exchanging roles. For an example, we will use a different pair of statements \( P \) and \( Q \):

\[ P : I \text{ will eat pizza} \]

\[ Q : I \text{ will eat spaghetti} \]

The argument above becomes (after minor colloquial adjustment):

\[ I \text{ will eat pizza or spaghetti.} \]

\[ I \text{ will not eat pizza.} \]

∴ \( I \text{ will eat spaghetti.} \)

To see how this unifies both modus ponens and modus tollens as two manifestations of the same principle, recall the following (easily proved by a truth table):

\[ P \rightarrow Q \iff (\sim P) \lor Q. \]

This appeared as (1.29), page 22, for instance. Thus the modus ponens and modus tollens become, respectively,

\[^{37}\text{A lay person might call this a form of “process of elimination.”}\]
1.3. VALID IMPLICATIONS AND ARGUMENTS

\[
\begin{align*}
(\sim P) \lor Q \\
\sim (\sim P)
\end{align*}
\]

and

\[
\begin{align*}
(\sim P) \lor Q \\
\sim Q
\end{align*}
\]

\[\therefore Q,\]

\[\therefore \sim P.\]

Compare these to the original forms, respectively, to see they are the same:

\[
\begin{align*}
P &\rightarrow Q \\
P &
\end{align*}
\]

and

\[
\begin{align*}
P &\rightarrow Q \\
\sim Q
\end{align*}
\]

\[\therefore \sim P.\]

For another fallacy, consider the fallacy of the inverse:\textsuperscript{38}

**Example 1.3.9** Show that the following statement is a fallacy.

\[
P \rightarrow Q \\
\sim P
\]

\[\therefore \sim Q \text{ (Invalid)}\]

**Solution:** We check to see if the statement

\[
[(P \rightarrow Q) \land (\sim P)] \rightarrow (\sim Q)
\]

is a tautology (in which case we could replace the major operation \(\rightarrow\) with \(\Rightarrow\)).

<table>
<thead>
<tr>
<th>P</th>
<th>Q</th>
<th>P \rightarrow Q</th>
<th>\sim P</th>
<th>(P \rightarrow Q) \land (\sim P)</th>
<th>\sim Q</th>
<th>[(P \rightarrow Q) \land (\sim P)] \rightarrow (\sim Q)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

We note that the argument—as an implication—is false in the case \(P\) is false and \(Q\) is true. In fact, in that case we have both premises of the argument \((P \rightarrow Q)\) vacuously, and \(\sim P\) obviously), but not the conclusion. Let us return to our rain and wet grass statements from before:

If it rained then the grass is wet.

It did not rain.

\[\therefore \text{The grass is not wet.} \text{ (Invalid)}\]

For another example, consider the following obviously fallacious argument:

If you drink the hemlock then you will die.

You do not drink the hemlock.

\[\therefore \text{You will not die.} \text{ (Invalid)}\]

There are, of course, other reasons why you might die (or why the grass might be wet). The case we see in the truth table example which ruins the bid for the corresponding implication to be a tautology is that case in which you do not drink the hemlock, and still die, contradicting the conclusion but not the hypotheses.

\textsuperscript{38}The inverse of an implication \(R \rightarrow S\) is the statement \((\sim R) \rightarrow (\sim S)\). It is not equivalent to the original implication. In fact, it is equivalent to the converse (see Footnote 34, page 34), the proof of which is left to the exercises. A course on logic would emphasize these two statements which are related to the implication. However, they are a source of some confusion so we do not elaborate extensively here. It is much more important to realize that \(R \rightarrow S\) is equivalent to its contrapositive \((\sim S) \rightarrow (\sim R)\), and not equivalent to these other two related implications, namely the converse \(S \rightarrow R\) and inverse \((\sim R) \rightarrow (\sim S)\). These facts should become more self-evident as the material is studied and utilized.
CHAPTER 1. MATHEMATICAL LOGIC AND SETS

There are many other forms of argument, both valid and invalid. A rudimentary strategy for detecting if an argument is valid or a fallacy is the same: to look at the corresponding implication and see if it is a tautology. If so, the argument is valid and if not, the argument is a fallacy. In the next subsection we introduce a slightly more sophisticated method, in which we use previously established valid implications and equivalences to make shorter work of some complicated arguments, but for now we will continue to use the truth table test of the validity of the underlying implication.

Our next example was proved earlier in the form of a valid implication. As such it was the subject of Example 1.3.4, page 30, and was listed as (1.52), page 32. We will not rework the truth table here.

Example 1.3.10 The following is a valid form of argument:

\[
\begin{align*}
P \rightarrow Q \\
Q \rightarrow R \\
\therefore P \rightarrow R
\end{align*}
\]

For an example of an application, consider the following assignments for \( P, Q \) and \( R \):

- \( P : I \text{ am paid} \)
- \( Q : I \text{ will buy you a present} \)
- \( R : You \text{ will be happy} \)

The argument then becomes the following, the validity of which is reasonably clear:

If I am paid, then I will buy you a present.
If I will buy you a present, then you will be happy.
\[
\therefore \text{If I am paid, then you will be happy.}
\]

Next we look at an example which contains three underlying component statements, and three premises. Before doing so, we point out that we can compute the truth tables for \( P \land Q \land R \) and \( P \lor Q \lor R \) relatively quickly; the former is true whenever all three are true (and false if at least one is false), and the latter is true if any of the three are true (and false only if all three are false). The reason we can do this is that there is no ambiguity in computing, for instance, \( P \land (Q \land R) \) or \( (P \land Q) \land R \), as these are known to be equivalent (see page 22). Similarly for \( \lor \).\[^{39}\] To be clearer, we note the truth tables.

\[
\begin{array}{c|c|c|c|c|c}
P & Q & R & P \land Q \land R & P \lor Q \lor R \\
\hline
T & T & T & T & T \\
T & T & F & F & T \\
T & F & T & F & T \\
T & F & F & F & T \\
F & T & T & F & T \\
F & T & F & F & T \\
F & F & T & F & T \\
F & F & F & F & F \\
\end{array}
\]

\[^{39}\]This is similar to the arithmetic rules that \( A + B + C = A + (B + C) = (A + B) + C \). The first expression is not at first defined per se, but because the second and third are the same we allow for the first. Similarly \( A \cdot B \cdot C = A \cdot (B \cdot C) = (A \cdot B) \cdot C \). However, this does not extend to all operations, such as subtraction: Usually \( A - (B - C) \neq (A - B) - C \), so when we write \( A - B - C \) we have to choose one, and in fact we choose the latter. Similarly, we have to be careful with other logical operations besides \( \land \) and \( \lor \). For instance, \( P \rightarrow Q \rightarrow R \) would probably be interpreted \( (P \rightarrow Q) \rightarrow R \), but it would depend upon the author.
With this observation, we can more easily analyze arguments with more than two premises.

**Example 1.3.11** Consider the argument

\[
P \rightarrow (Q \lor R) \\
P \sim R \\
\therefore Q
\]

We need to see if the following conditional—which we dub ARG (for “argument”) for space considerations—is a tautology:

\[
\text{ARG} : \quad \{(P \rightarrow (Q \lor R)) \land P \land (\sim R)\} \rightarrow Q.
\]

<table>
<thead>
<tr>
<th>(P)</th>
<th>(Q)</th>
<th>(R)</th>
<th>(Q \lor R)</th>
<th>(P \rightarrow (Q \lor R))</th>
<th>(\sim R)</th>
<th>((P \rightarrow (Q \lor R)) \land P \land (\sim R))</th>
<th>(Q)</th>
<th>ARG</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

Note that “ARG” connects the two immediately preceding columns with the logical operation \(\rightarrow\). Also notice that most of the cases are true vacuously (as often happens when we have more premises to be met, connected by \(\land\)), and eventually we see that the argument is valid.

Reading more examples, and perhaps some trial and error, one’s intuition for what is valid and what is not should develop. With English examples one might be able to see that the above argument style is reasonable. For instance, consider the following.

\[
P : \text{I will eat pizza} \\
Q : \text{I will drink soda} \\
R : \text{I will drink beer}
\]

Then the premises become

\[
P \rightarrow (Q \lor R) : \text{If I eat pizza, then I will drink soda or beer.} \\
P : \text{I will eat pizza.} \\
\sim R : \text{I will not drink beer.}
\]

It is reasonable to believe that, after declaring that if I eat pizza then I will drink soda or beer, and that I indeed will eat pizza, but not drink the beer, then I must drink the soda.\(^{40}\)

### 1.3.4 Analyzing Arguments Without Truth Tables

In fact there is another approach for analyzing complicated arguments such as the one above. The strategy is manipulative, and there are two tools we can make use of. The first is the somewhat obvious fact that we can always replace one of the hypotheses with a logically equivalent

\(^{40}\)Or else I was lying when I recited my premises. The argument, anyhow, is valid.
CHAPTER 1. MATHEMATICAL LOGIC AND SETS

statement. That is because the truth table column entries are what matter in our computations there. But there is another strategy which is not quite so obvious. It relies upon the following logical equivalence, which is left as an exercise:

\[ P \rightarrow Q \iff P \leftrightarrow (P \land Q). \]  \hspace{1cm} (1.62)

For our purposes, it means that if our premises \( P \) imply \( Q \), then we can replace them with \( P \land Q \), in effect attaching \( Q \) to the list of hypotheses. Note also this is true if \( P \) is instead just a single hypothesis in the list of hypotheses, which are joined by “ands” \( \land \), which is commutative and associative so \( Q \) can be attached anywhere in the list as another hypothesis.

If one or more of the hypotheses taken collectively validly imply a statement \( Q \), then \( Q \) can be attached to the list of hypotheses, again due to the commutative and associative nature of \( \land \).

**Example 1.3.12** Let us re-examine the previous example, but this time we append some intermediate conclusions.

\[
\begin{align*}
P \rightarrow (Q \lor R) & \quad \therefore Q \lor R \\
\sim R & \quad \therefore Q
\end{align*}
\]

The first conclusion \( Q \lor R \) is implied by the first two premises by a simple modus ponens, and the second came from the third original hypothesis \( \sim R \) and our newly attached hypothesis \( (Q \lor R) \) by the way of the disjunctive syllogism (Example 1.3.8, page 36 and (1.57), page 32). It requires more creativity than a truth table verification, but is clearly less tedious.

**Example 1.3.13** Consider the following argument style:

\[
(P \lor Q) \rightarrow R \\
\sim R \\
\therefore \sim P.
\]

Here again we can work backwards. (A reasonable alternative style would skip some steps.)

\[
\begin{align*}
(P \lor Q) \rightarrow R & \quad \therefore \sim (P \lor Q) \\
\sim R & \quad \therefore \sim P
\end{align*}
\]

Note that in this latest example we could have appended our original hypotheses to include \( (\sim P) \land (\sim Q) \), which would be the same as appending \( \sim P \) and \( \sim Q \) separately.

This method is often quite useful for verifying validity, but it is not, by itself, a way to detect a fallacy. However, since common fallacies rest upon the fallacy of the inverse (page 37) or the fallacy of the converse (page 34), we can sometimes detect when an argument tempts us to agree with the conclusion because of such invalid, but common, reasoning.

**Example 1.3.14** Consider the following argument.

\[
(P \lor Q) \rightarrow R \\
\sim Q \\
R \\
\therefore P. \text{ (Invalid!)}
\]

What is tempting (but invalid) to do with this argument, is to reason that the first and third hypotheses imply \( P \lor Q \), and with the second reading \( \sim Q \), we would then (seemingly validly)
conclude $P$. However, we can not concluded $P$ from our premises. Indeed, $R$ does not imply anything about $P$ and $Q$, and if nothing else, a truth table will prove this style invalid. An errant diagram might look like the following:

\[
\begin{array}{c}
(P \lor Q) \rightarrow R \\
\neg Q \\
R \\
\hline
\therefore P
\end{array}
\] \quad \therefore (P \lor Q) \quad \text{(INVALID!)} \quad \therefore \neg Q \\
\therefore P
\]

An example in English might help to see that indeed this is invalid. If we define $P$, $Q$ and $R$ as below, it seems pretty clear this style is a fallacy.

$P : I$ shot you.
$Q : I$ stabbed you.
$R : You$ died.

This particular argument then reads:

If I shot you or stabbed you, then you died.
I did not stab you.
You died.
\[\therefore I$ shot you. \ (Invalid!\)

Clearly this is fallacious reasoning. The “alternative explanation” for the consequent, i.e., $Q$ where $P \rightarrow Q$, should always be looming when we look at implications and try to read them backwards. (In this case there is nothing in the premises that do not allow for other causes of your dying.)

From the above discussion, one might conclude that we can often intuitively detect the likelihood of an argument being invalid, but unless we can rewrite it as a known fallacy—or clearly see a case where the premises hold true and the conclusion does not—we would need to go back to our primitive but absolutely reliable method of constructing the truth table for the underlying implication to see if we have a tautology. If we do not have a tautology then the argument is a fallacy; if we do, then the argument is valid.

Still, this new method of proving validity for arguments can be very useful, especial in lieu of long truth tables, but of course it necessarily rests upon the styles we previously proved to be valid, or the valid equivalences or implications from before. (It also requires a bit of creativity in

---

41 Some texts have the reader check the validity of an argument by checking only the cases in which all premises are true, to see if the conclusions are also true for those cases. Indeed it is only the cases in which the premises are true and the conclusion is false which invalidate an argument style. However, that kind of analysis de-emphasizes the connection between valid argument styles and valid implications, and the role of tautologies and so we prefer to include these ideas, at the cost of looking at every case of truth values for the underlying statements $P, Q,$ and so on when analyzing an argument for validity.

42 By now the reader, having encountered abstract and applied implications on many abstract levels, should be aware of the reason to write both sides of semicolon in the sentence above, namely, “If we do not have a tautology then the argument is a fallacy; if we do, then the argument is valid.” To spell this out better, consider

\[
P : we$ have a tautology (in the form of the argument as an implication)
\]

\[
Q : the$ argument is valid
\]

Then the sentence in quotes reads $[(\neg P) \rightarrow (\neg Q)] \land [P \rightarrow Q]$, which is equivalent to $[Q \rightarrow P] \land [P \rightarrow Q]$, i.e., $P \rightarrow Q$. If we wrote only the first part, $(\neg P) \rightarrow (\neg Q)$, that alone would not declare the second part $P \rightarrow Q$, though many casual readers would assume that it would (in words if not symbols).
CHAPTER 1. MATHEMATICAL LOGIC AND SETS

using them.) The more known valid styles, equivalences and implications one has available, the larger the number of argument styles which can be proven without resorting to truth tables. Of course it was the truth tables which allowed us to prove the preliminary equivalences, implications and argument styles to be valid, so ultimately all these things rest upon the truth tables. They are the one device at our disposal that allowed logic to be rendered computational (and at a very fundamental level), rather than just intuitive.

Exercises

Some of these were proved previously in the text. The reader should attempt to prove these first for himself/herself without referring back to the text for the proofs.

1. Prove $P \rightarrow Q \iff P \iff (P \land Q)$. This is essentially (1.62), page 40.

2. Consider the statement $P \rightarrow (\sim P)$.

(a) Use a truth table to prove the validity of $P \rightarrow (\sim P) \implies (\sim P)$. Is this reasonable?

(b) If possible without truth tables, and using a short string of known equivalences, show that in fact $P \rightarrow (\sim P) \iff (\sim P)$.

(See (1.29), page 22 and (1.17), page 22.)

3. Use truth tables to show (1.49) and (1.50):

$$P \land Q \implies P,$$

$$P \implies P \lor Q.$$  (1.43)

4. Prove the following without using truth tables.

$$P \implies (P \lor Q)$$

by proving that $P \implies (P \lor Q)$ is a tautology. (Again, see (1.29), page 22.)

5. Prove the following two valid implications:

(a) (1.56): $(P \rightarrow Q) \land P \implies Q.$

(b) (1.57): $(P \rightarrow Q) \land (\sim Q) \implies \sim P.$

6. Prove the following using truth tables.

Note that there are $2^4 = 16$ different combinations of truth values for $P, Q, R$ and $S$.

$$ (P \rightarrow R) \land (Q \rightarrow S)$$

$$\implies (P \land Q) \rightarrow (R \land S), \quad (1.63)$$

$$ (P \rightarrow R) \land (Q \rightarrow S)$$

$$\implies (P \lor Q) \rightarrow (R \lor S) \quad (1.64)$$

These are useful because we often have a string of statements $A_1, \ldots, A_n$ connected entirely by $\land$ or entirely by $\lor$, and wish to replace them with $B_1, \ldots, B_n$ (or just some of them), where $A_1 \implies B_1, A_2 \implies B_2, \ldots, A_n \implies B_n$.  (1.44)

With (1.63) and (1.64) we can generalize and write

$$A_1 \land A_2 \land \cdots \land A_n$$

$$\implies B_1 \land B_2 \land \cdots \land B_n, \quad (1.65)$$

$$A_1 \lor A_2 \lor \cdots \lor A_n$$

$$\implies B_1 \lor B_2 \lor \cdots \lor B_n. \quad (1.66)$$

Valid implications (1.63) and (1.64) would be quite laborious to prove without truth tables.

---

43 Some would characterize $P \land Q \implies P \implies P \lor Q$ to be a progression from the strongest statement, $P \land Q$, to the weakest, $P \lor Q$ of the three. A similar form would be $P \land Q \implies P \implies P \lor R$.

44 Here the parts of $P \rightarrow R$ and $Q \rightarrow S$ are played by the $A_i \rightarrow B_i$, but the idea is the same. We could instead attach $T \iff (A_1 \rightarrow B_1) \land \cdots \land (A_n \rightarrow B_n)$ to the left-hand sides of (1.65) and (1.66) with the wedge operation, since $T \land U \iff U$. 
7. For each of the following, decide if you believe it is a valid argument or a fallacy. Then check by constructing the corresponding truth table.

(a) \[ P \rightarrow Q \quad Q \rightarrow P \quad \therefore P \leftrightarrow Q \]

(b) \[ P \quad P \rightarrow Q \quad \therefore Q \]

(c) \[ Q \quad P \rightarrow Q \quad \therefore P \]

(d) \[ Q \quad P \rightarrow Q \quad \therefore \sim P \]

(For this last case, see Footnote 35, page 35.)

8. For each of the following, decide if it is valid or a fallacy. For those marked “(Prove),” offer a truth table proof or a manipulation of the hypotheses to justify your answer.

(a) \[ (P \land Q) \rightarrow R \]

(b) \[ (P \land Q) \rightarrow R \]

(c) \[ (P \land Q) \rightarrow R \]

(d) \[ (P \land Q) \rightarrow R \]

(e) \[ P \rightarrow Q \quad Q \rightarrow R \quad \sim P \quad \therefore \sim R \]

(f) \[ P \rightarrow Q \quad Q \rightarrow R \quad R \rightarrow P \quad \therefore (P \leftrightarrow R) \]

(g) \[ P \lor Q \lor R \quad \sim P \quad \sim R \quad \therefore Q \quad (\text{Prove}) \]

(h) \[ P \rightarrow Q \quad Q \rightarrow (\sim R) \quad P \quad \therefore \sim Q \]

(i) \[ P \lor Q \lor R \quad \sim P \quad \sim R \quad \therefore Q \quad (\text{Prove}) \]

(j) \[ P \rightarrow (Q \land R) \quad \sim R \quad \therefore \sim P \quad (\text{Prove}) \]

(k) \[ (\sim S) \land (\sim U) \quad \therefore \sim R \]

(l) \[ (\sim P) \rightarrow (\sim Q) \quad \sim Q \quad \therefore \sim P \]

(m) \[ (P \land Q) \rightarrow R \quad S \rightarrow P \quad U \rightarrow Q \quad S \land U \quad \therefore R \]

(n) \[ (P \land Q) \rightarrow R \quad S \rightarrow P \quad U \rightarrow Q \quad (\sim S) \land (\sim U) \quad \therefore \sim R \]
(o) \[
(P \land Q) \rightarrow R \\
S \rightarrow P \\
U \rightarrow Q \\
\sim R
\]
\[
\therefore (\sim S) \lor (\sim U)
\]

(p) \[
\mathcal{F} \\
\therefore P
\]

(q) \[
P \\
\therefore T
\]

(r) \[
T \rightarrow P
\]
\[
\therefore P
\]

(s) \[
P \rightarrow \mathcal{F}
\]
\[
\therefore \sim P
\]

(t) \[
P \rightarrow R
\]
\[
R \rightarrow P
\]
\[
\therefore P \leftrightarrow R \text{ (Prove)}
\]

(u) \[
Q \leftrightarrow R
\]
\[
\therefore P \leftrightarrow R
\]
1.4 Quantifiers and Sets

In this section we introduce quantifiers, which form the last class of logic symbols we will consider in this text. To use quantifiers, we also need some notions and notation from set theory. This section introduces sets and quantifiers to the extent required for our study of calculus here. For the interested reader, Section 1.5 will extend this introduction, though even with that section we would be only just beginning to delve into these topics if studying them for their own sakes. Fortunately what we need of these topics for our study of calculus is contained in this section.

1.4.1 Sets

Put simply, a set is a collection of objects, which are then called elements or members of the set. We give sets names just as we do variables and statements. For an example of the notation, consider a set \( A \) defined by

\[ A = \{2, 3, 5, 7, 11, 13, 17\}. \]

We usually define a particular set by describing or listing the elements between “curly braces” \( \{ \} \) (so the reader understands it is indeed a set we are discussing). The defining of \( A \) above was accomplished by a complete listing, but some sets are too large for that to be possible, let alone practical. As an alternative, the set \( A \) above can also be written

\[ A = \{x \mid x \text{ is a prime number less than } 18\}. \]

The above equation is usually read, “\( A \) is the set of all \( x \) such that \( x \) is a prime number less than 18.” Here \( x \) is a “dummy variable,” used only briefly to describe the set.\(^{45}\) Sometimes it is convenient to simply write

\[ A = \{\text{prime numbers between } 2 \text{ and } 17, \text{ inclusive}\}. \]

(Usually “inclusive” is meant by default, so here we would include 2 and 17 as possible elements, if they also fit the rest of the description.) Of course there are often several ways of describing a list of items. For instance, we can replace “between 2 and 17, inclusive” with “less than 18,” as before.

Often an ellipsis “\( \cdots \)" is used when a pattern should be understood from a partial listing. This is particularly useful if a complete listing is either impractical or impossible. For instance, the set \( B \) of integers from 1 to 100 could be written

\[ B = \{1, 2, 3, \cdots, 100\}. \]

To note that an object is in a set, we use the symbol \( \in \). For instance we may write \( 5 \in B \), read “5 is an element of \( B \).” To indicate concisely that 5, 6, 7 and 8 are in \( B \), we can write \( 5, 6, 7, 8 \in B \).

Just as we have use for zero in addition, we also define the empty set, or null set as the set which has no elements. We denote that set \( \emptyset \). Note that \( x \in \emptyset \) is always false, i.e.,

\[ x \in \emptyset \iff \text{false}, \]

because it is impossible to find any element of any kind inside \( \emptyset \). We will revisit this set repeatedly in the optional, more advanced Section 1.5.

\(^{45}\)“Dummy variables” are also used to describe the actions of functions, as in \( f(x) = x^2 + 1 \). In this context, the function is considered to be the action of taking an input number, squaring it, and adding 1. The \( x \) is only there so we can easily trace the action on an arbitrary input. We will revisit functions later.
Of course for calculus we are mostly interested in sets of numbers. While not the most important, the following three sets will occur from time to time in this text:

- **Natural Numbers**\(^{46}\):
  \[ \mathbb{N} = \{1, 2, 3, 4, \cdots\} \]  
  (1.67)

- **Integers**:
  \[ \mathbb{Z} = \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\} \]  
  (1.68)

- **Rational Numbers**:
  \[ \mathbb{Q} = \left\{ \frac{p}{q} \bigg| (p, q \in \mathbb{Z}) \land (q \neq 0) \right\} \]  
  (1.69)

Here we again use the ellipsis to show that the established pattern continues forever in each of the cases \( \mathbb{N} \) and \( \mathbb{Z} \). The sets \( \mathbb{N} \), \( \mathbb{Z} \) and \( \mathbb{Q} \) are examples of **infinite** sets, i.e., sets that do not have a finite number of elements. The rational numbers are those which are **ratios** of integers, except that division by zero is not allowed, for reasons we will consider later.\(^{47}\)

For calculus the most important set is the set \( \mathbb{R} \) of **real numbers**, which cannot be defined by a simple listing or by a simple reference to \( \mathbb{N} \), \( \mathbb{Z} \) or \( \mathbb{Q} \). One intuitive way to describe the real numbers is to consider the horizontal **number line**, where geometric points on the line are represented by their **displacements** (meaning distances, but counted as positive if to the right and negative if to the left) from a fixed point, called the **origin** in this context. That fixed point is represented by the number 0, since the fixed point is a displacement of zero units from itself. In Figure 1.2 the number line representation of \( \mathbb{R} \) is shown. Hash marks at convenient intervals are often included. In this case, they are at the integers. The arrowheads indicate the number line is an actual line and thus infinite in both directions. The points \(-2.5\) and \(4.8\) on the graph are not integers, but are rational numbers, since they can be written \(-\frac{25}{10} = -\frac{5}{2}\), and \(\frac{48}{10} = 24/5\), respectively. The points \(\sqrt{2}\) and \(\pi\) are real, but not rational, and so are called **irrational**. To summarize,

**Definition 1.4.1** *The set of all real numbers is the set \( \mathbb{R} \) of all possible displacements, to the right or left, of a fixed point 0 on a line. If the displacement is to the right, the number is the positive distance from 0. If to the left, the number is the negative of the distance from 0.*\(^{48}\)

Thus
\[ \mathbb{R} = \{\text{displacements from 0 on the number line}\}. \]  
(1.70)

This is not a rigorous definition, not least because “right” and “left” require a fixed perspective. Even worse, the definition is really a kind of “circular reasoning,” since we are effectively defining

\(^{46}\)The natural numbers are also called **counting** numbers in some texts.

\(^{47}\)For a hint, think about what should be \( x = 1/0 \). If we multiply both sides by zero, we might think we get \( 0x = (1/0) \cdot 0 \), giving \( 0 = 1 \), which is absurd. In fact there was no such \( x \), so \( x = 1/0 \rightarrow 0 = 1 \), which is of the form \( P \rightarrow F \) which we may recall to be equivalent to \( \sim P \).

\(^{48}\)It should be noted that we have to choose a direction to call “right,” the other then being “left.” It will depend upon our perspective. When we look at the Cartesian Plane, the horizontal axis measures displacements as right (positive horizontal) or left (negative horizontal), and the vertical axis measures displacements as upward (positive vertical) or downward (negative vertical). In that context the **origin** is where the axes intersect.
the number line in terms of \( R \), and then defining \( R \) in terms of (displacements on) the number line. We will give a more rigorous definition in Chapter 2 for the interested reader. For now this should do, since the number line is a simple and intuitive image.

### 1.4.2 Quantifiers

The three quantifiers used by nearly every professional mathematician are as follow:

- **universal quantifier:** \( \forall \), read, “for all,” or “for every;”
- **existential quantifier:** \( \exists \), read, “there exists;”
- **uniqueness quantifier:** \( ! \), read, “unique.”

The first two are of equal importance, and far more important than the third which is usually only found after the second. Quantified statements are usually found in forms such as:

- \( (\forall x \in S) P(x) \), i.e., for all \( x \in S \), \( P(x) \) is true;
- \( (\exists x \in S) P(x) \), i.e., there exists an \( x \in S \) such that \( P(x) \) is true;
- \( (\exists! x \in S) P(x) \), i.e., there exists a unique (exactly one) \( x \in S \) such that \( P(x) \) is true.

Here \( S \) is a set and \( P(x) \) is some statement about \( x \). The meanings of these quickly become straightforward. For instance, consider

- \( (\forall x \in \mathbb{R})(x + x = 2x) \) : for all \( x \in \mathbb{R} \), \( x + x = 2x \);
- \( (\exists x \in \mathbb{R})(x + 2 = 2) \) : there exists (an) \( x \in \mathbb{R} \) such that \( x + 2 = 2 \);
- \( (\exists! x \in \mathbb{R})(x + 2 = 2) \) : there exists a unique \( x \in \mathbb{R} \) such that \( x + 2 = 2 \).

All three quantified statements above are true. In fact they are true under any circumstances, and can thus be considered tautologies. Unlike unquantified statements \( P, Q, R \), etc., from our first three sections, a quantified statement is either true always or false always, and is thus, for our purposes, equivalent to either \( T \) or \( F \). Each has to be analyzed on its face, based upon known mathematical principles; we do not have a brute-force mechanism analogous to truth tables to analyze these systematically.\(^{49}\) For a couple more short examples, consider the following cases from algebra which should be clear enough:

\[
(\forall x \in \mathbb{R})(0 \cdot x = 0) \iff T;
\]

\[
(\exists x \in \mathbb{R})(x^2 = -1) \iff F.
\]

The optional advanced section shows how we can still find equivalent or implied statements from quantified statements in many circumstances.

### 1.4.3 Statements with Multiple Quantifiers

Many of the interesting statements in mathematics contain more than one quantifier. To illustrate the mechanics of multiply quantified statements, we will first turn to a more worldly setting. Consider the following sets:

\[
M = \{\text{men}\},
\]

\[
W = \{\text{women}\}.
\]

\(^{49}\)This is part of what makes quantified statements interesting!
In other words, $M$ is the set of all men, and $W$ the set of all women. Consider the statement\(^{50}\)

$$\left( \forall m \in M \right) \left( \exists w \in W \right)[w \text{ loves } m]. \quad (1.71)$$

Set to English, (1.71) could be written, “for every man there exists a woman who loves him.”\(^{51}\) So if (1.71) is true, we can in principle arbitrarily choose a man $m$, and then know that there is a woman $w$ who loves him. It is important that the man $m$ was quantified first. A common syntax that would be used by a logician or mathematician would be to say here that, once our choice of a man is fixed, we can in principle find a woman who loves him. Note that (1.71) allows that different men may need different women to love them, and also that a given man may be loved by more than (but not less than) one woman.

Alternatively, consider the statement

$$\left( \exists w \in W \right) \left( \forall m \in M \right)[w \text{ loves } m]. \quad (1.72)$$

A reasonable English interpretation would be, “there exists a woman who loves every man.” Granted that is a summary, for the word-for-word English would read more like, “there exists a woman such that, for every man, she loves him.” This says something very different from (1.71), because that earlier statement does not assert that we can find a woman who, herself, loves every man, but that for each man there is a woman who loves him.\(^{52}\)

We can also consider the statement

$$\left( \forall m \in M \right) \left( \forall w \in W \right)[w \text{ loves } m]. \quad (1.73)$$

This can be read, “for every man and every woman, the woman loves the man.” In other words, every man is loved by every woman. In this case we can reverse the order of quantification:

$$\left( \forall w \in W \right) \left( \forall m \in M \right)[w \text{ loves } m]. \quad (1.74)$$

In fact, if the two quantifiers are the same type—both universal or both existential—then the order does not matter. Thus

$$\left( \forall m \in M \right) \left( \forall w \in W \right)[w \text{ loves } m] \iff \left( \forall w \in W \right) \left( \forall m \in M \right)[w \text{ loves } m],$$

$$\left( \exists m \in M \right) \left( \exists w \in W \right)[w \text{ loves } m] \iff \left( \exists w \in W \right) \left( \exists m \in M \right)[w \text{ loves } m].$$

In both representations in the existential statements, we are stating that there is at least one man and one woman such that she loves him. In fact that above equivalence is also valid if we replace $\exists$ with $\exists!$, though it would mean then that there is exactly one man and exactly one woman such that the woman loves the man, but we will not delve too deeply into uniqueness here.

Note that in cases where the sets are the same, we can combine two similar quantifications into one, as in

$$\left( \forall m \in M \right) \left( \forall w \in W \right)[x + y = y + x] \iff \left( \forall x, y \in \mathbb{R} \right)[x + y = y + x]. \quad (1.75)$$

Similarly with existence.

\(^{50}\)Note that this is of the form $(\forall m \in M)(\exists w \in W)P(m, w)$, that is, the statement $P$ says something about both $m$ and $w$. We will avoid a protracted discussion of the difference between statements regarding one variable object—as in $P(x)$ from our previous discussion—and statements which involve more than one as in $P(m, w)$ here. Statements of multiple (variable) quantities will recur in subsequent examples.

\(^{51}\)At times it seems appropriate to translate “$\forall$” as “for all,” and at other times it seems better to translate it as “for every.” Both mean the same.

\(^{52}\)We do not pretend to know the truth values of either (1.71) or (1.72).
1.4. QUANTIFIERS AND SETS

However, we repeat the point at the beginning of the subsection, which is that the order does matter if the types of quantification are different.

For another, short example which is algebraic in nature, consider

\[(\forall x \in \mathbb{R})(\exists K \in \mathbb{R})(x = 2K).\] (True.) (1.76)

This is read, “for every \(x \in \mathbb{R}\), there exists \(K \in \mathbb{R}\) such that \(x = 2K\).” That \(K = x/2\) exists (and is actually unique) makes this true, while it would be false if we were to reverse the order of quantification:

\[(\exists K \in \mathbb{R})(\forall x \in \mathbb{R})(x = 2K).\] (False.) (1.77)

Statement (1.77) claims (erroneously) that there exists \(K \in \mathbb{R}\) so that, for every \(x \in \mathbb{R}\), \(x = 2K\). That is impossible, because no value of \(K\) is half of every real number \(x\). For example the value of \(K\) which works for \(x = 4\) is not the same as the value of \(K\) which works for \(x = 100\).

1.4.4 Detour: Uniqueness as an Independent Concept

We will have occasional statements in the text which include uniqueness. However, most of those will not require us to rewrite the statements in ways which require actual manipulation of the uniqueness quantifier. Still, it is worth noting a couple of interesting points about this quantifier.

First we note that uniqueness can be formulated as a separate concept from existence, interestingly instead requiring the universal quantifier.

**Definition 1.4.2 Uniqueness** is the notion that if \(x_1, x_2 \in S\) satisfy the same particular statement \(P(\cdot)\), then they must in fact be the same object. That is, if \(x_1, x_2 \in S\) and \(P(x_1)\) and \(P(x_2)\) are true, then \(x_1 = x_2\). This may or may not be true, depending upon the set \(S\) and the statement \(P(\cdot)\).

Note that there is the vacuous case where nothing satisfies the statement \(P(\cdot)\), in which case the uniqueness of any such hypothetical object is proved but there is actually no existence. Consider the following, symbolic representation of the uniqueness of an object \(x\) which satisfies \(P(x)\):

\[
(\forall x, y \in S)[(P(x) \land P(y)) \rightarrow x = y].
\] (1.78)

Finally we note that a proof of a statement such as \((\exists! x \in S)P(x)\) is thus usually divided into two separate proofs:

1. **Existence:** \((\exists x \in S)P(x)\);
2. **Uniqueness:** \((\forall x, y \in S)[(P(x) \land P(y)) \rightarrow x = y]\).

For example, in the next chapter we rigorously, axiomatically define the set of real numbers \(\mathbb{R}\). One of the axioms\(^{54}\) defining the real numbers is the existence of an additive identity:

\[
(\exists z \in \mathbb{R})(\forall x \in \mathbb{R})(z + x = x).
\] (1.79)

---

\(^{53}\)The above statement indeed says that any two elements \(x, y \in S\) which both satisfy \(P\) must be the same. Note that we use a single arrow here, because the statement between the brackets \([\cdot]\) is not likely to be a tautology, but may be true for enough cases for the entire quantified statement to be true. Indeed, the symbols \(\Rightarrow\) and \(\iff\) belong between quantified statements, not inside them.

\(^{54}\)Recall that an axiom is an assumption, usually self-evident, from which we can logically argue towards theorems. Axioms are also known as postulates. If we attempt to argue only using "pure logic" (as a mathematician does when developing theorems, for instance), it eventually becomes clear that we still need to make some assumptions because one can not argue "from nothing." Indeed, some "starting points" from which to argue towards the conclusions are required. These are then called axioms.

The word “axiomatic” is often used colloquially to mean clearly evident and therefore not requiring proof. In
In fact it follows quickly that such a “$z$” must be unique, so we have

\[ (\exists! z \in \mathbb{R})(\forall x \in \mathbb{R})(z + x = x). \] (1.80)

To prove (1.80), we need to prove (1) existence, and (2) uniqueness. In this setting, the existence is an axiom so there is nothing to prove. We turn then to the uniqueness. A proof is best written in prose, but it is based upon proving that the following is true:

\[ (\forall z_1, z_2 \in \mathbb{R})[(z_1 \text{ an additive identity}) \land (z_2 \text{ an additive identity}) \rightarrow z_1 = z_2]. \]

Now we prove this. Suppose $z_1$ and $z_2$ are additive identities, i.e., they can stand in for $z$ in (1.79), which could also read $(\exists z \in \mathbb{R})(\forall x \in \mathbb{R})(x = z + x)$. Note the order there, where the identity $z$ (think “zero”) is placed on the left of $x$ in the equation $x = z + x$. So, assuming $z_1, z_2$ are additive identities, we have:

\[
\begin{align*}
z_1 &= z_2 + z_1 & \text{(since $z_2$ is an additive identity)} \\
 &= z_1 + z_2 & \text{(since addition is commutative—order is irrelevant)} \\
 &= z_2 & \text{(since $z_1$ is an additive identity)}.
\end{align*}
\]

This argument showed that if $z_1$ and $z_2$ are any real numbers which act as additive identities, then $z_1 = z_2$. In other words, if there are any additive identities, there must be only one. Of course, assuming its existence we call that unique real number zero. (It should be noted that the commutativity used above is another axiom of the real numbers. We will list fourteen in all.)

The distinction between existence and uniqueness of an object with some property $P$ is often summarized as follows:

1. Existence asserts that there is at least one such object.
2. Uniqueness asserts that there is at most one such object.

If both hold, then there is exactly one such object.

### 1.4.5 Negating Universally and Existentially Quantified Statements

For statements with a single universal or existential quantifier, we have the following negations.

\[ \sim [(\forall x \in S)P(x)] \iff (\exists x \in S)[\sim P(x)], \] (1.81)

\[ \sim [(\exists x \in S)P(x)] \iff (\forall x \in S)[\sim P(x)], \] (1.82)

The left side of (1.81) states that it is not the case that $P(x)$ is true for all $x \in S$; the right side states that there is an $x \in S$ for which $P(x)$ is false. We could ask when is it a lie that for all $x$, $P(x)$ is true? The answer is when there is an $x$ for which $P(x)$ is false, i.e., $\sim P(x)$ is true.

The left side of (1.82) states that it is not the case that there exists an $x \in S$ so that $P(x)$ is true; the right side says that $P(x)$ is false for all $x \in S$. When is it a lie that there is an $x$ making $P(x)$ true? When $P(x)$ is false for all $x$.

Thus when we negate such a statement as $(\forall x)P(x)$ or $(\exists x)P(x)$, we change $\forall$ to $\exists$ or vice-versa, and negate the statement after the quantifiers.

---

fact that is not always the case with mathematical axioms. Indeed, the postulates required for defining the real numbers seem rather strange at first. They were, in fact, developed to be a minimal number of assumptions required to give the real numbers their apparent properties which could be observed. In that case, it seemed we worked towards a foundation, after seeing the outer structure. A similar phenomenon can be seen in Einstein’s Special Relativity, where his two simple—yet at the time quite counterintuitive—axioms were able to completely replace a much larger set of postulates required to explain many of the electromagnetic phenomena discovered early during his time, and predict many new phenomena that were later observed.
1.4. QUANTIFIERS AND SETS

Example 1.4.1 Negate \((\forall x \in S)[P(x) \rightarrow Q(x)]\).

Solution: We will need (1.21), page 22, namely \(\sim (P \rightarrow Q) \iff P \land (\sim Q)\).

\[
\sim [(\forall x \in S)(P(x) \rightarrow Q(x))] \iff (\exists x \in S)[(\sim (P(x) \rightarrow Q(x))]
\iff (\exists x \in S)[P(x) \land (\sim (Q(x))].
\]

The above example should also be intuitive. To say that it is not the case that, for all \(x \in S\), \(P(x) \rightarrow Q(x)\) is to say there exists an \(x\) so that we do have \(P(x)\), but not the consequent \(Q(x)\).

Example 1.4.2 Negate \((\exists x \in S)[P(x) \land Q(x)]\).

Solution: Here we use \(\sim (P \land Q) \iff (\sim P) \lor (\sim Q)\), so we can write

\[
\sim [(\exists x \in S)(P(x) \land Q(x))] \iff (\forall x)[(\sim P(x)) \lor (\sim (Q(x))].
\]

This last example shows that if it is not the case that there exists an \(x \in S\) so that \(P(x)\) and \(Q(x)\) are both true, that is the same as saying that for all \(x\), either \(P(x)\) is false or \(Q(x)\) is false.

1.4.6 Negating Statements Containing Mixed Quantifiers

Here we simply apply (1.81) and (1.82) two or more times, as appropriate. For a typical case of a statement first quantified by \(\forall\), and then by \(\exists\), we note that we can group these as follows:55

\[
(\forall x \in R)(\exists y \in S)P(x,y) \iff (\forall x \in R)((\exists y \in S)P(x,y)].
\]

(Here “\(R\)” is another set, not to be confused with the set of real numbers \(\mathbb{R}\).) Thus

\[
\sim [(\forall x \in R)(\exists y \in S)P(x,y)] \iff \sim [(\forall x \in R)((\exists y \in S)P(x,y)]
\iff (\exists x \in R)\{\sim [(\exists y \in S)P(x,y)]\}
\iff (\exists x \in R)(\forall y \in S)(\sim P(x,y)].
\]

Ultimately we have, in turn, the \(\forall\)'s become \(\exists\)'s, the \(\exists\)'s become \(\forall\)'s, the variables are quantified in the same order as before, and finally the statement \(P\) is replaced by its negation \(\sim P\). The pattern would continue no matter how many universal and existential quantifiers arise. (The uniqueness quantifier is left for the exercises.) To summarize for the case of two quantifiers,

\[
\sim [(\forall x \in R)(\exists y \in S)P(x,y)] \iff (\exists x \in R)(\forall y \in S)(\sim P(x,y)] \quad (1.83)
\sim [(\exists x \in R)(\forall y \in S)P(x,y)] \iff (\forall x \in R)(\exists y \in S)(\sim P(x,y)]. \quad (1.84)
\]

Example 1.4.3 Consider the following statement, which is false:

\[
(\forall x \in \mathbb{R})(\exists y \in \mathbb{R})(xy = 1).
\]

One could say that the statement says every real number \(x\) has a real number reciprocal \(y\). This is false, but before that is explained, we compute the negation which must be true:

\[
\sim [(\forall x \in \mathbb{R})(\exists y \in \mathbb{R})(xy = 1)] \iff (\exists x \in \mathbb{R})(\forall y \in \mathbb{R})(xy \neq 1).\]

Indeed, there exists such an \(x\), namely \(x = 0\), such that \(xy \neq 1\) for all \(y\).56

---

55This may not be the most transparent fact, and indeed there are somewhat deep subtleties involved, but eventually this should be clear. The subtleties lie in the idea that once a variable is quantified, it is “fixed” for that part of the statement which follows it. For instance, that part \((\exists y \in S)P(x, y)\) treats \(x\) as if it were “constant.”

56Note that in using English, the quantification often follows after the variable quantified, as in Example 1.4.3 above. That can become quite confusing when statements get complicated. Indeed, much of the motivation of this section is so that we can use the notation to, in essence, diagram the logic of such statements, and analyze them to see if they may be false (by seeing if their negations ring true).
In the above, we borrowed one of the many convenient mathematical notations for the negations of various symbols. Some common negations follow:

\[ \sim (x = y) \iff x \neq y, \]
\[ \sim (x < y) \iff x \geq y, \]
\[ \sim (x \leq y) \iff x > y, \]
\[ \sim (x \in S) \iff x \notin S. \]

Of course we can negate both sides of any one of these and get, for example, \( x \in S \iff \sim (x \notin S). \) Reading one of these backwards, we can have \( \sim (x \geq y) \iff x < y. \)

**Exercises**

1. Consider the sets

\[ P = \{ \text{prisons} \}, \]
\[ M = \{ \text{methods of escape} \}. \]

For each of the following, write a short English version of the given statement.

(a) \( (\forall p \in P)(\exists m \in M)[m \text{ will get you out of } P] \]
(b) \( (\exists m \in M)(\forall p \in P)[m \text{ will get you out of } P] \]
(c) \( (\exists p \in P)(\forall m \in M)[m \text{ will get you out of } P] \]
(d) \( (\forall m \in M)(\exists p \in P)[m \text{ will get you out of } P] \]
(e) \( (\exists m \in M)(\exists p \in P)[m \text{ will get you out of } P] \]
(f) \( (\forall m \in M)(\forall p \in P)[m \text{ will get you out of } P] \]

2. For parts (a)–(d) above, write the negation of the statement both symbolically and in English.

3. Write negations for each of the following. If in the process a logical statement within the quantified statement is negated, write the negation in the clearest possible form. For instance, instead of writing \( \sim (P \to Q), \) write \( P \land (\sim Q). \) Similarly instead of writing \( \sim (x > y) \) write \( x \leq y. \)

(a) \( (\forall x \in R)[x \in S] \]
(b) \( (\forall x \in R)(\exists y \in S)[y \leq x] \]
(c) \( (\forall x,y \in R)(\exists r,t \in S)[rx + ty = 1] \)
   **Hint:** This can also be written \( (\forall x \in R)(\forall y \in R)(\exists r \in S)(\exists t \in S)[rx + ty = 1]. \)
(d) \( (\forall \varepsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x \in \mathbb{R}) \[ |x - a| < \delta \implies |f(x) - f(a)| < \varepsilon \] \)
   **Hint:** consider \( P : |x - a| < \delta \) and \( Q : |f(x) - f(a)| < \varepsilon. \) Then consider the usual negation of \( P \to Q, \) with these statements inserted literally, and then rewrite it in a more “understandable” way.
4. For each of the following, write the negation of the statement and decide which is true, the original statement or its negation.

(a) \((\exists x \in \mathbb{R})(x^2 < 0)\)
(b) \((\forall x \in \mathbb{R})(|−x| \neq x)\)
(c) \((\forall x \in \mathbb{R})(\exists y \in \mathbb{R})(y = 2x + 1)\)

5. Consider \((\forall c \in C)(\exists b \in B)(b \text{ would buy } c)\). Here \(C = \{\text{cars}\}\) and \(B = \{\text{buyers}\}\).

(a) Write in English what this statement says.
(b) Write in English what the negation of the statement should be.
(c) Write in symbolic logic what the negation of the statement should be.

6. Consider the statement \((\forall x, y \in \mathbb{R})(x < y \rightarrow x^2 < y^2)\).

(a) Write the negation of this statement.
(b) In fact it is the negation that is true. Can you explain why?

7. Using the fact that
\[(\exists! x \in S)P(x) \iff \exists x \in S P(x) \land \forall x, y \in S [(P(x) \land P(y)) \rightarrow x = y],\]

compute the form of the negation of the unique existence:
\[\sim [(\exists! x \in S)P(x)].\]
1.5 Sets Proper

In this section we introduce set theory in its own right. We also apply the earlier symbolic logic to the theory of sets (rather than vice-versa). We also approach set theory visually and intuitively, while simultaneously introducing all the set-theoretic notation we will use throughout the text. To begin we make the following definition:

**Definition 1.5.1** A set is a well-defined collection of objects.

By well-defined, we mean that once we define the set, the objects contained in the set are totally determined, and so any given object is either in the set or not in the set. We might also note that in a sense a set is defined (or determined) by its elements; sets which are different collections of elements are different sets, while sets with exactly the same elements are the same set. We can also define equality by means of quantifiers:

**Definition 1.5.2** Given two sets $A$ and $B$, we defined the statement $A = B$ as being equivalent to the statement $(\forall x)[(x \in A) \iff (x \in B)]$:

$$A = B \iff (\forall x)[(x \in A) \iff (x \in B)]. \quad (1.85)$$

If we allow ourselves to understand that $x$ is quantified universally (that is, we assume “$(\forall x)$” is understood) unless otherwise stated, we can write, instead of $A = B$, that $x \in A \iff x \in B$.

When we say a set is well-defined we also mean that once defined the set is fixed, and does not change. If elements can be listed in a table (finite or otherwise), then the order we list the elements is not relevant; sets are defined by exactly which objects are elements, and which are not. Moreover, it is also irrelevant if objects are listed more than once in the set, such as when we list $\mathbb{Q} = \{x \mid x = p/q, \ p, q \in \mathbb{Z}, \ q \neq 0\}$. In that definition $2 = 2/1 = 4/2 = 6/3$ is “listed” infinitely many times, but it is simply one element of the set of rational numbers $\mathbb{Q}$. While it actually is possible to “list” the elements of $\mathbb{Q}$ if we allow for the ellipsis (\cdots), it is more practical to describe the set, as we did, using some defining property of its elements (here they were ratios of integers, without dividing by zero), as long as it is exactly those elements in the set—no more and no fewer—which share that property. One usually uses a “dummy variable” such as $x$ and then describes what properties all such $x$ in the set should have. We could have just as easily used $z$ or any other variable.

1.5.1 Subsets and Set Equality

When all the elements of a set $A$ are also elements of another set $B$, we say $A$ is a subset of $B$. To express this in set notation, we write $A \subseteq B$. In this case we can also take another perspective, and say $B$ is a superset of $A$, written $B \supseteq A$. Both symbols represent types of set inclusions, i.e., they show one set is contained in another.

A useful graphical device which can illustrate the notion that $A \subseteq B$ and other set relations is the **Venn Diagram**, as in Figure 1.3. There we see a visual representation of what it means for $A \subseteq B$. The sets are represented by enclosed areas in which we imagine the elements reside. In each representation given in Figure 1.3, all the elements inside $A$ are also inside $B$.

---

57Note that not all sets can be listed in a table, even if it is infinitely long. We can list $\mathbb{N} = \{1, 2, 3, \ldots\}$ and $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \ldots\}$, and even with a little ingenuity list the elements of $\mathbb{Q}$, but we cannot do so with $\mathbb{R}$ or $\mathbb{R} - \mathbb{Q}$. Those sets which can be listed in a table are called countable, and the others uncountable. All sets with a finite number of elements are also countable. Of the others, some are countably infinite, and the others are uncountably infinite (or simply uncountable, as the “infinite” in “uncountably infinite” is redundant).

58Of course we would not use fixed elements of the set as “variables,” which they are not since each has a unique identity.
Using symbolic logic, we can define subsets, and the notation, as follows:

\[ A \subseteq B \iff (\forall x)(x \in A \implies x \in B). \quad (1.86) \]

The role of the implication which is the main feature of (1.86) should seem intuitive. Perhaps less intuitive are some of the statements which are therefore logically equivalent to (1.86):

\[ A \subseteq B \iff (\forall x)(x \in A \implies x \in B) \iff (\forall x)[(\sim (x \in A)) \lor (x \in B)] \iff (\forall x)[(x \notin A) \lor (x \in B)] , \]

which uses the fact that \( P \implies Q \iff (\sim P) \lor Q \), and

\[ A \subseteq B \iff (\forall x)[(\sim (x \in B)) \implies (\sim (x \in A))] \iff (\forall x)[(x \notin B) \implies (x \notin A)] \]

which uses the contrapositive \( P \implies Q \iff (\sim Q) \implies (\sim P) \). Note that we used the shorthand notation \( \sim (x \in A) \iff x \notin A \). With the definition (1.86) we can quickly see two more, technically interesting facts about subsets:

**Theorem 1.5.1** For any sets \( A \) and \( B \), the following hold true:

\[ A \subseteq A, \quad \text{and} \quad A = B \iff (A \subseteq B) \land (B \subseteq A). \quad (1.87) \quad (1.88) \]

Now we take a moment to remind ourselves of what is meant by *theorem*:

**Definition 1.5.3** A *theorem* is a statement which we know to be true because we have a proof of it. We can therefore accept it as a tautology.

A theorem’s scope may be very limited (the above theorem only applies to sets and subsets as we have defined them.) Furthermore, a theorem’s scope and “truth” depends upon the axiomatic system upon which it rests, such the definitions we gave our symbolic logic symbols (which might not have always been completely obvious to the novice, as in our definitions of “\( \lor \)” and “\( \rightarrow \)”). For another example there is Euclidean geometry, the theorems of which
rest upon Euclid’s Postulates (or axioms, or original assumptions), while other geometric systems begin with different postulates.

Nonetheless once we have the definitions and postulates one can say that a theorem is a statement which is always true (demonstrated by some form of proof), and in fact therefore equivalent to $T$ (introduced on page 26). We will use that fact in the proof of (1.87), but for (1.88) we will instead demonstrate the validity of the equivalence ($\iff$). For the first statement’s proof, we have

$$A \subseteq A \iff (\forall x)[(x \in A) \rightarrow (x \in A)] \iff T.$$  

Note that the above proof is based upon the fact that $P \rightarrow P$ is a tautology (i.e., equivalent to $T$). A glance at a Venn Diagram with a set $A$ can also convince one of this fact, that any set is a subset of itself. For the proof of (1.88) we offer the following:

$$A = B \iff (\forall x)[(x \in A) \leftrightarrow (x \in B)]$$

$$\iff (\forall x)[(x \in A) \rightarrow (x \in B)] \land (\forall x)[(x \in B) \rightarrow (x \in A)]$$

$$\iff [(\forall x)[(x \in A) \rightarrow (x \in B)] \land (\forall x)[(x \in B) \rightarrow (x \in A)]]$$

$$\iff (A \subseteq B) \land (B \subseteq A), \text{ q.e.d.}^{59}$$

A consideration of Venn diagrams also leads one to believe that for all the area in $A$ to be contained in $B$ and vice versa, it must be the case that $A = B$. That $A = B$ implies they are mutual subsets is perhaps easier to see.

Note that the above arguments can also be made with supersets instead of subsets, with $\supseteq$ replacing $\subseteq$ and $\leftarrow$ replacing $\rightarrow$.

One needs to be careful with quantifiers and symbolic logic, as is discussed later in Section 7.7, but in what we did above the $(\forall x)$ effectively went along for the ride.

Of course, Venn Diagrams can accommodate more than two sets. For example, we can illustrate the chain of set inclusions

$$N \subseteq Z \subseteq Q \subseteq R$$  \hspace{1cm} (1.89)$$

using a Venn Diagram, as in Figure 1.4. Note that this is a compact way of writing six different set inclusions: $N \subseteq Z$, $N \subseteq Q$, $N \subseteq R$, $Z \subseteq Q$, $Z \subseteq R$, and $Q \subseteq R$.

\hspace{1cm}^{59}\text{Latin, quod erat demonstrandum}, the traditional ending of a proof meaning that which was to be proved.
1.5. SETS PROPER

1.5.2 Intervals and Inequalities in $\mathbb{R}$

The number line, which we will henceforth dub the real line, has an inherent order in which the numbers are arranged. Suppose we have two numbers $a, b \in \mathbb{R}$. Then the order relation between $a$ and $b$ has three possibilities, each with its own notation:

1. $a$ is to the left of $b$, written $a < b$ and spoken “$a$ is less than $b$.”

2. $a$ is to the right of $b$, written $a > b$ and spoken “$a$ is greater than $b$.”

3. $a$ is at the same location as $b$, written $a = b$ and spoken “$a$ equals $b$.”

Figure 1.5 shows these three possibilities. Note that “less than” and “greater than” refer to relative positions on the real line, not how “large” or “small” the numbers are. For instance, $4 < 5$ but $-5 < -4$, though it is natural to consider $-5$ to be a “larger” number than $-4$. Similarly $-1000 < 1.60$. Of course if $a < b \iff b > a$. We have further notation which describes when $a$ is left of or at $b$, and when $a$ is right of or at $b$:

4. $a$ is at or left of $b$, written $a \leq b$ and spoken “$a$ is less than or equal to $b$.”

5. $a$ is at or right of $b$, written $a \geq b$ and spoken “$a$ is greater than or equal to $b$.”

Using inequalities, we can describe intervals in $\mathbb{R}$, which are exactly the connected subsets of $\mathbb{R}$, meaning those sets which can be represented by darkening the real line at only those points which are in the subset, and where doing so can be theoretically accomplished without lifting our pencils as we darken. In other words, these are “unbroken” subsets of $\mathbb{R}$. Later we will see that intervals are subsets of particular interest in calculus.

Intervals can be classified as finite or infinite (referring to their lengths), and open, closed or half-open (referring to their “endpoints”). The finite intervals are of three types: closed, open and half-open. Intervals of these types, with real endpoints $a$ and $b$, where $a < b$ (though the idea extends to work with $a \leq b$) are shown below respectively by graphical illustration, in interval notation, and using earlier set-theoretic notation:

---

60 Later we will refer to a number’s absolute size, in which context we will describe $-1000$ as “larger” than 1.
CHAPTER 1. MATHEMATICAL LOGIC AND SETS

open: \( (a, b) \) \( \{ x \in \mathbb{R} \mid a < x < b \} \)

closed: \( [a, b] \) \( \{ x \in \mathbb{R} \mid a \leq x \leq b \} \)

half-open: \( (a, b] \) \( \{ x \in \mathbb{R} \mid a < x \leq b \} \)

half-open: \( [a, b) \) \( \{ x \in \mathbb{R} \mid a \leq x < b \} \)

Note that \( a < x < b \) is short for \( (a < x) \land (x < b) \), i.e., \( (x > a) \land (x < b) \). The others are similar.

We will concentrate on the open and closed intervals in calculus. For the finite open interval above, we see that we do not include the endpoints \( a \) and \( b \) in the set, denoting this fact with parentheses in the interval notation and an “open” circle at each endpoint on the graph. What is crucial to calculus is that immediately surrounding any point \( x \in (a, b) \) are only other points still inside the interval; if we pick a point \( x \) anywhere in the interval \( (a, b) \), we see that just left and just right of \( x \) are only points in the interval. Indeed, we have to travel some distance—albeit possibly short—to leave the interval from a point \( x \in (a, b) \). Thus no point inside of \( (a, b) \) is on the boundary, and so each point in \( (a, b) \) is “safely” on the interior of the interval. This will be crucial to the concepts of continuity, limits and (especially) derivatives later in the text.

For a closed interval \( [a, b] \), we do include the endpoints \( a \) and \( b \), which are not surrounded by other points in the interval. For instance, immediately left of \( a \) is outside the interval \( [a, b] \), though immediately right of \( a \) is on the interior.\(^{61}\) We denote this fact with brackets in the interval notation, and a “closed” circle at each endpoint when we sketch the graph. Half-open (or half-closed) intervals are simple extensions of these ideas, as illustrated above.

For infinite intervals, we have either one or no endpoints. If there is an endpoint it is either not included in the interval or it is, the former giving an open interval and the latter a closed interval. An open interval which is infinite in one direction will be written \( (a, \infty) \) or \( (-\infty, a) \), depending upon the direction in which it is infinite. Here \( \infty \) (infinity) means that we can move along the interval to the right “forever,” and \(-\infty\) means we can move left without end. For infinite closed intervals the notation is similar: \( [a, \infty) \) and \( (-\infty, a] \). The whole real line is also considered an interval, which we denote \( \mathbb{R} = (-\infty, \infty) \).\(^{62}\) When an interval continues \textit{without bound} in a direction, we also darken the arrow in that direction. Thus we have the following:

\(^{61}\)For a closed interval \([a, b]\), later we will sometimes refer to the \textit{interior} of the interval, meaning all points whose immediate neighbors left and right are also in the interval. This means that the interior of \([a, b]\) is simply \((a, b]\).

\(^{62}\)For technical reasons which will be partially explained later, \( \mathbb{R} \) is considered to be both an open and a closed interval. Roughly, it is open because every point is interior, but closed because every point that can be approached as close as we want from the interior is contained in the interval. Those are the topological factors which characterize open and closed intervals as such. Topology as a subject is rarely taught before the junior level of college, or even graduate school, though advanced calculus usually includes some topology of \( \mathbb{R} \).
1.5. SETS PROPER

open: \( a \rightarrow (a, \infty) \) \( \{ x \in \mathbb{R} \mid x > a \} \)

open: \( a \rightarrow (-\infty, a) \) \( \{ x \in \mathbb{R} \mid x < a \} \)

closed: \( a \rightarrow [a, \infty) \) \( \{ x \in \mathbb{R} \mid x \geq a \} \)

closed: \( a \rightarrow (-\infty, a] \) \( \{ x \in \mathbb{R} \mid x \leq a \} \)

Note that we never use brackets to enclose an infinite “endpoint,” since \(-\infty, \infty\) are not actual boundaries but rather are concepts of unending continuance. Indeed, \(-\infty, \infty \notin \mathbb{R}\), i.e., they are not points on the real line, so they can not be boundaries of subsets of \( \mathbb{R} \); there are no elements “beyond” them.

1.5.3 Most General Venn Diagrams

Before we get to the title of this subsection, we will introduce a notion which we will have occasional use for, which is the concept of proper subset.

**Definition 1.5.4** If \((A \subseteq B) \land (A \neq B)\), we call \(A\) a **proper** subset of \(B\), and write \(A \subset B\).\(^{63}\)

Thus \(A \subset B\) means \(A\) is contained in \(B\), but \(A\) is not all of \(B\). Note that \(A \subset B \implies A \subseteq B\) (just as \(P \land Q \implies P\)). When we have that \(A\) is a subset of \(B\) and are not interested in emphasizing whether or not \(A \neq B\) (or are not sure if this is true), we will use the “inclusive” notation \(\subseteq\). In fact, the inclusive case is less complicated logically (just as \(P \lor Q\) is easier than \(P \lor Q\)) and so we will usually opt for it even when we do know that \(A \neq B\). We mention the exclusive case here mainly because it is useful in explaining the most general Venn Diagram for two sets \(A\) and \(B\).

Of course it is possible to have two sets, \(A\) and \(B\), where neither is a subset of the other. Then \(A\) and \(B\) may share some elements, or no elements. In fact, for any given sets \(A\) and \(B\), exactly one of the following will be true:

**case 1:** \(A = B\);

**case 2:** \(A \subset B\), i.e., \(A\) is a proper subset of \(B\);

**case 3:** \(B \subset A\), i.e., \(B\) is a proper subset of \(A\);

**case 4:** \(A\) and \(B\) share common elements, but neither is a subset of the other;

**case 5:** \(A\) and \(B\) have no common elements. In such a case the two sets are said to be disjoint.

Even if we do not know which of the five cases is correct, we can use a single illustration which covers all of these. That illustration is given in Figure 1.6, with the various regions labeled. (We will explain the meaning of \(U\) in the next subsection.) To see that this covers all cases, we take them in turn:

---

\(^{63}\)The notation has changed over the years. Many current texts use “\(\subset\)” the way we use “\(\subseteq\)” here. This is unfortunate, because the notations “\(\subseteq, \subset\)” here are strongly analogous to the notations \(\leq, <\) from arithmetic. One has to take care to know how notation is being used in a given context. (A few authors even use \(\subseteq, \not\subset\).)
Figure 1.6: Most general Venn diagram for two arbitrary sets $A$ and $B$. Here $U$ is some superset of both $A$ and $B$.

Figure 1.7: The most general Venn Diagram for three sets $A$, $B$ and $C$.

**case 1:** $A = B$: all elements of $A$ and $B$ are in Region IV; there are no elements in Regions II and III.

**case 2:** $A \subset B$: there are elements in Regions III and IV, and no elements in Region II.

**case 3:** $B \subset A$: there are elements in Regions II and IV, and no elements in Region III.

**case 4:** $A$ and $B$ share common elements, but neither is a subset of the other: there are elements in Region II, III and IV.

**case 5:** $A$ and $B$ have no common elements: there are no elements in Region IV.

Note that whether or not Region I has elements is irrelevant in the discussion above, though it will become important shortly.

The most general Venn diagram for three sets is given in Figure 1.7, though we will not exhaustively show this to be the most general. It is not important that the sets are represented by circles, but only that there are sufficiently many separate regions and that every case of an element being, or not being, in $A$, $B$ and $C$ is represented. Note that there are three sets for an element to be or not to be a member of, and so there are $2^3 = 8$ subregions needed.

### 1.5.4 Set Operations

When we are given two sets $A$ and $B$, it is natural to combine or compare their memberships with each other and the universe of all elements of interest. In particular, we form new sets called the union and intersection of $A$ and $B$, the difference of $A$ and $B$ (and of $B$ and $A$), and the complement of $A$ (and of $B$). The first three are straightforward, but the fourth requires
some clarification. Usually $A$ and $B$ contain only objects of a certain class like numbers, colors, etc. Thus we take elements of $A$ and $B$ from a specific universal set $U$ of objects rather than an all-encompassing universe of all objects. It is unlikely in mathematics that we would need, for instance, to mix numbers with persons and planets and verbs, so we find it convenient to limit our universe $U$ of considered objects. With that in mind (but without presently defining $U$), the notations for these new sets are as follow:

**Definition 1.5.5**

\[
A \cup B = \{ x \mid (x \in A) \lor (x \in B) \} \tag{1.90}
\]
\[
A \cap B = \{ x \mid (x \in A) \land (x \in B) \} \tag{1.91}
\]
\[
A - B = \{ x \mid (x \in A) \land (x \notin B) \} \tag{1.92}
\]
\[
A' = \{ x \in U \mid (x \notin A) \}. \tag{1.93}
\]

These are read “$A$ union $B$,” “$A$ intersect $B$,” “$A$ minus $B$,” and “$A$ complement,” respectively. Note that in the first three, we could have also written $\{ x \in U \mid \cdots \}$, but since $A, B \subseteq U$, there it is unnecessary. Also note that one could define the complement in the following way, though (1.93) is more convenient for symbolic logic computations:

\[
A' = \{ x \mid (x \in U) \land (x \notin A) \} = U - A. \tag{1.94}
\]

These operations are illustrated by the Venn diagrams of Figure 1.8, where we also construct $B'$ and $B - A$. Note the connection between the logical $\lor$ and $\land$, and the set-theoretical $\cup$ and $\cap$.

**Example 1.5.1** Find $A \cup B$, $A \cap B$, $A - B$ and $B - A$ if

\[
A = \{ 1, 2, 3, 4, 5, 6, 7 \}
\]
\[
B = \{ 5, 6, 7, 8, 9, 10 \}.
\]

\[\text{64} \text{The set-theoretical “−” could be interpreted as “$\land \sim \cdots \in$” and if we always assume we know what is the universal set, we can interpret the complement symbol “′” as “$\sim \cdots \in$.”}\]
Solution: Though not necessary (and often impossible), we will list these set elements in a table from which we can easily compare the membership.

\[
A = \{ 1, 2, 3, 4, 5, 6, 7 \} \quad \text{and} \quad B = \{ 5, 6, 7, 8, 9, 10 \}.
\]

Now we can compare the memberships using the operations defined earlier.

\[
A \cup B = \{ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 \},
\]
\[
A \cap B = \{ 5, 6, 7 \},
\]
\[
A - B = \{ 1, 2, 3, 4 \},
\]
\[
B - A = \{ 8, 9, 10 \}.
\]

The complements depend upon the identity of the assumed universal set. If in the above example we had \( U = \mathbb{N} \), then \( A' = \{ 8, 9, 10, 11, \cdots \} \) and \( B' = \{ 1, 2, 3, 4, 11, 12, 13, 14, 15 \cdots \} \). If instead we took \( U = \mathbb{Z} \) we have \( A' = \{ \cdots, -3, -2, -1, 0, 8, 9, 10, 11, \cdots \} \), for instance. (We leave \( B' \) to the interested reader.)

Just as it is important to have a zero element in \( \mathbb{R} \) for arithmetic and other purposes, it is also useful in set theory to define a set which contains no elements:

**Definition 1.5.6** The set with no elements is called the **empty set**,\(^{65}\) denoted \( \emptyset \).

One reason we need such a device is for cases of intersections of disjoint sets. If \( A = \{ 1, 2, 3 \} \) and \( B = \{ 4, 5, 6, 7, 8, 9, 10 \} \), then \( A \cup B = \{ 1, 2, 3, \cdots, 10 \} \), while \( A \cap B = \emptyset \). Notice that regardless of the set \( A \), we will always have \( A - A = \emptyset \), \( A - \emptyset = A \), \( A \cup \emptyset = A \), \( A \cap \emptyset = \emptyset \), and \( \emptyset \subseteq A \). The last statement is true because, after all, every element of \( \emptyset \) is also an element of \( A \).\(^{66}\) Note also that \( \emptyset' = U \) and \( U' = \emptyset \).

The set operations for two sets \( A \) and \( B \) can only give us finitely many combinations of the areas enumerated in Figure 1.6. In fact, since each such area is either included or not, there are \( 2^4 = 16 \) different diagram shadings possible for the general case as in Figure 1.6. The situation is more interesting if we have three sets \( A \), \( B \), and \( C \). Using Figure 1.7, we can prove several interesting set equalities. First we have some fairly obvious commutative laws (1.95), (1.96) and associative laws (1.97), (1.98):

\[
A \cup B = B \cup A \quad (1.95)
\]
\[
A \cap B = B \cap A \quad (1.96)
\]
\[
A \cup (B \cup C) = (A \cup B) \cup C \quad (1.97)
\]
\[
A \cap (B \cap C) = (A \cap B) \cap C \quad (1.98)
\]

Next are the following two **distributive laws**, which are the set-theory analogs to the logical equivalences (1.27) and (1.28), found on page 22.

\[
A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad (1.99)
\]
\[
A \cup (B \cap C) = (A \cup B) \cap (A \cup C). \quad (1.100)
\]

\(^{65}\) It is also called the **null set**. Some older texts use empty braces \( \emptyset = \{ \} \).

\(^{66}\) This is precisely because there are no elements of \( \emptyset \); the statement \( x \in \emptyset \rightarrow x \in A \) is vacuously true because \( x \in \emptyset \) is false, regardless of \( x \).
1.5. SETS PROPER

Figure 1.9: Venn Diagrams for Example 1.5.2 verifying one of the distributive laws, specifically \( A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \). It is especially important to note how one constructs the third box in each line from the first two.

**Example 1.5.2** We will show how to prove (1.99) using our previous symbolic logic, and then give a visual proof using Venn diagrams. Similar techniques can be used to prove (1.100). For the proof that \( A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \), we use definitions, and (1.27) from page 22 to get the following:

\[
\begin{align*}
  x \in A \cap (B \cup C) & \iff (x \in A) \land (x \in B \cup C) \\
 & \iff (x \in A) \land [(x \in B) \lor (x \in C)] \\
 & \iff [(x \in A) \land (x \in B)] \lor [(x \in A) \land (x \in C)] \\
 & \iff [x \in (A \cap B)] \lor [x \in (A \cap C)] \\
 & \iff x \in [(A \cap B) \cup (A \cap C)], \text{ q.e.d.}
\end{align*}
\]

We proved that \( (\forall x)[(x \in A \cap (B \cup C)) \iff (x \in (A \cap B) \cup (A \cap C))] \), which is the definition for the sets in question to be equal. The visual demonstration of \( A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \) is given in Figure 1.9, where we construct both sets of the equality in stages.

To construct the left-hand side of the equation, in the first box we color \( A \), then \( B \cup C \) in the second, and finally we take the area from the first, remove the area from the second, and are left with the difference \( A - (B \cup C) \). To construct the right-hand side of the equation, we color \( A - B \) and \( A - C \) in separate boxes. Then we color the intersection of these, which is the area colored in the previous two boxes. This gives us our Venn Diagram for \( (A - B) \cap (A - C) \). We see that the left- and right-hand sides are the same, and conclude the equality is valid.

The next two are distributive in nature also:

\[
\begin{align*}
  A - (B \cup C) & = (A - B) \cap (A - C) \quad (1.101) \\
  A - (B \cap C) & = (A - B) \cup (A - C) \quad (1.102)
\end{align*}
\]
Finally, if we replace \( A \) with \( U \), we get the set-theoretic version of de Morgan’s Laws:

\[
\begin{align*}
(B \cup C)' &= B' \cap C' \\
(B \cap C)' &= B' \cup C'.
\end{align*}
\] (1.103) (1.104)

Note that these are very much like our earlier de Morgan’s laws, and indeed use the previous versions (1.3) and (1.4), page 17 (also see page 22) in their proofs. For instance, assuming \( x \in U \) where \( U \) is fixed, we have

\[
x \in (B \cup C)' \iff \sim (x \in B \cup C)
\]

\[
\iff \sim ((x \in B) \lor (x \in C))
\]

\[
\iff [\sim (x \in B)] \land [\sim (x \in C)]
\]

\[
\iff [x \in B'] \land [x \in C']
\]

\[
\iff x \in B' \cap C', \text{ q.e.d.}
\]

That proves (1.103), and (1.104) has a similar proof. It is interesting to prove these using Venn Diagrams as well (see exercises).

**Example 1.5.3** Another example of how to prove these using logic and Venn diagrams is in order. We will prove (1.101) using both methods. First, with symbolic logic:

\[
x \in A - (B \cup C) \iff (x \in A) \land [\sim (x \in B \cup C)]
\]

\[
\iff (x \in A) \land [\sim ((x \in B) \lor (x \in C))]
\]

\[
\iff (x \in A) \land [(\sim (x \in B)) \land (\sim (x \in C))]
\]

\[
\iff (x \in A) \land (\sim (x \in B)) \land (\sim (x \in C))
\]

\[
\iff [(x \in A) \land (\sim (x \in B)) \land (\sim (x \in C))]
\]

\[
\iff (x \in A - B) \land (x \in A - C)
\]

\[
\iff x \in (A - B) \cap (A - C), \text{ q.e.d.}
\]

If we took the steps above in turn, we used the definition of set subtraction, the definition of union, (1.19), associative property of \( \land \), added a redundant \( (x \in A) \), regrouped, used the definition of set subtraction, and finally the definition of intersection.

Now we will see how we can use Venn diagrams to prove (1.101). As before, we will do this by constructing Venn Diagrams for the sets \( A - (B \cup C) \) and \( (A - B) \cap (A - C) \) separately, and verify that we get the same sets. We do this in Figure 1.10. (If it is not visually clear how we proceed from one diagram to the next “all at once,” a careful look at each of the \( 2^3 = 8 \) distinct regions can verify the constructions.)

### 1.5.5 More on Subsets

Before closing this section, a few more remarks should be included on the subject of subsets. Consider for instance the following:

**Example 1.5.4** Let \( A = \{1, 2\} \). List all subsets of \( A \).

**Solution:** As \( A = \{1, 2\} \) has two elements, it can have subsets which contain zero elements, one element, or two elements. The subsets are thus \( \emptyset, \{1\}, \{2\} \) and \( \{1, 2\} = A \).
1.5. SETS PROPER

It is common for novices studying sets to forget that $\emptyset \subseteq A$, and $A \subseteq A$, though by definition,

\[
x \in \emptyset \implies x \in A \quad \text{vacuously},
\]
\[
x \in A \implies x \in A \quad \text{trivially}.
\]

If one wanted only proper subsets of $A$, those would be $\emptyset, \{1\}, \{2\}$ (we omit the set $A$).

Note that with our set $A = \{1, 2\}$, we can reduce rephrase the question of which subset we might refer to, instead into a question of exactly which elements are in it, from the choices 1 and 2. In other words, given a subset $B \subseteq A$, which (if any) of the following are true: $1 \in B$, $2 \in B$.

From these statements we can construct a truth table-like structure to describe every possible subset of $A$:

\[
\begin{array}{ccc}
1 \in B & 2 \in B & \text{subset } B \\
T & T & \{1, 2\} = A \\
T & F & \{1\} \\
F & T & \{2\} \\
F & F & \emptyset
\end{array}
\]

Similarly, a question about subsets $B$ of $A = \{a, b, c\}$ can be placed in context of a truth table-like construct:

\[
A = \{a, b, c\}
\]
It would not be too difficult to list the elements of $A = \{1, 2, 3\}$ by listing subsets with zero, one, two and three elements separately, i.e., $\emptyset$, $\{a\}$, $\{b\}$, $\{c\}$, $\{a,b\}$, $\{a,c\}$, $\{b,c\}$, $\{a,b,c\}$, but if we were to need to list subsets of a set with significantly more elements, it might be easier to use the lexicographical order embedded in the truth table format to exhaust all the possibilities. The only disadvantage is that the order in which subsets are listed might not be quite as natural as the order we would likely find if we listed subsets with zero, one, two elements and so on.

<table>
<thead>
<tr>
<th>$a \in B$</th>
<th>$b \in B$</th>
<th>$c \in B$</th>
<th>subset $B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>${a,b,c} = A$</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>F</td>
<td>${a,b}$</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
<td>${a,c}$</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>${a}$</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>${b,c}$</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
<td>${b}$</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
<td>${c}$</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>$\emptyset$</td>
</tr>
</tbody>
</table>
1.5. SETS PROPER

Exercises

1. Draw all the $2^4 = 16$ possible shadings for Figure 1.6, page 60. Then use the sets $A$, $B$ and $U$, together with unions, intersections, complements and set differences ($\cup$, $\cap$, $'$, $-$) to write a corresponding expression for the each of the shaded areas. Note that Figure 1.8 illustrates six of them. Also note that there may be more than one way of representing a set. For example, $A' = U - A$.

2. Use symbolic logic and Venn Diagrams (as in Examples 1.5.2 and 1.5.3) to prove the other set equalities:

   (a) (1.97): $A \cup (B \cap C) = (A \cup B) \cap C$
   (b) (1.98): $A \cap (B \cap C) = (A \cap B) \cap C$
   (c) (1.102): $A - (B \cap C) = (A - B) \cup (A - C)$
   (d) (1.99): $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
   (e) (1.103): $(B \cup C)' = B' \cap C'$
   (f) (1.104): $(B \cap C)' = B' \cup C'$

3. Use Venn Diagrams to draw and determine a simpler way of writing the following sets:

   (a) $A - (B - A) =$
   (b) $A - (A - B) =$
   (c) $(A - B) \cap (B - A) =$

4. Answer each of the following.

   (a) If $A \subseteq B$, what is $A - B$?
   (b) If $A \subset B$, what can you say about $B - A$?
   (c) Referring to Figure 1.8, what are $U'$ and $\mathcal{O}'$?
   (d) Suppose $A \subseteq B$. How are $A'$ and $B'$ related?
   (e) Suppose $A - B = A$. What is $A \cap B$?
   (f) Suppose $A - B = B$. What is $B$? What is $A$?
   (g) Is it possible that $A \subseteq B$ and $B \subseteq A$?

5. The set of irrational numbers is the set

$$I = \{ x \in \mathbb{R} \mid x \text{ is not rational} \}.$$ 

Using previously-defined sets and set notation, find a concise definition of $I$.

6. Many books define the symmetric difference between two sets $A$ and $B$ by

$$A \triangle B = (A - B) \cup (B - A).$$

(a) Use a Venn Diagram to show that $A \triangle B = (A \cup B) - (A \cap B)$.
   (b) Is it true that $A \triangle B = B \triangle A$?
   (c) Use a Venn Diagram to show that $A \triangle (B \cap C) = (A \triangle B) \cup (A \triangle C)$.
   (d) Calculate $A \triangle A$, $A \triangle U$, and $A \triangle \mathcal{O}$.
   (e) If $A \subseteq B$, what is $A \triangle B$? What is $B \triangle A$?
   (f) Define $A \triangle B$ in a manner similar to the definitions (1.90)–(1.93). That is, replace the dots with a description of $x$ in the following:

$$A \triangle B = \{ x \mid \cdots \}.$$ 

7. Redraw the Venn Diagram of Figure 1.7 and label each of the eight disjoint areas I–VIII. Then use the sets $U$, $A$, $B$ and $C$, together with the operations $\cup$, $\cap$, $-$ and $'$, to find a definition of each of these sets I–VIII.

8. How many different shading combinations are there for the general Venn Diagram for 3 sets $A$, $B$ and $C$? (See Figure 1.7.) Speculate about how many combinations there are for 4 sets, 5 sets, and $n$ sets. Test your hypothesis for $n = 0$ and $n = 1$.

9. Show that $A \subseteq B \iff B' \subseteq A'$. (See the discussion immediately following (1.86), page 55.)
10. A useful concept in set theory is \textit{cardinality} of a set $S$, which we denote $n(S)$, which can be defined to be the number of elements in a set if the set is finite. Thus $n(\{1, 2, 3, 8, 9, 10\}) = 6$.

(a) Use a Venn Diagram to show that \[ n(S \cup T) = n(S) + n(T) - n(S \cap T) \] (1.106)

(b) Show that if $n(S \cup T) = n(S) + n(T)$, then $S \cap T = \emptyset$.

11. Given $U = \{1, 2, 3, \cdots, 12\}$, $A = \{1, 2, 3, 4, 5, 6\}$, $B = \{2, 3, 4, 6, 7, 8\}$, $C = \{3, 5, 7\}$, $D = \{3, 5, 7, 11\}$, find each of the following:

(a) $(A - B) \cup (B - A)$,
(b) $(A - D)'$,
(c) the number of subsets $C$,
(d) the number of subsets of $D$.
(e) Use the results from (c) and (d) to determine the number of subsets that a set of 5 elements should have.
1.6 Epilogue: Proofs Without Truth Tables

Truth tables are the most exhaustive tools for studying compound statements, because they simply check every case. This makes them useful tools to fall back upon if necessary, but as we learn and memorize more equivalences it is useful to put what we already showed to work in other proofs. For instance, recall the following equivalences, here written with our new notation:

1. \( P \implies Q \iff (\neg P) \lor Q \),
2. \( \neg [P \lor Q] \iff (\neg P) \land (\neg Q) \),
3. \( \neg (\neg P) \iff P \).

We will use these principles, in exactly the order above, to produce a very concise proof in the following example.

Example 1.6.1 Prove without truth tables that \( \neg (P \implies Q) \iff P \land (\neg Q) \).

Solution:

\[
\neg (P \implies Q) \iff \neg [(\neg P) \lor Q] \\
\iff [\neg (\neg P)] \land (\neg Q) \\
\iff P \land (\neg Q), \quad q.e.d.
\]

To show a statement is a tautology it is enough to show that it is logically equivalent to \( T \). We can use this fact to prove valid logical equivalences (i.e., show \( \iff \) gives a tautology) and valid logical implications (i.e., show \( \implies \) gives a tautology). There are often easier ways to prove valid equivalences, but for valid implications this method can be useful, as in the example below.

Example 1.6.2 We revisit Example 1.3.2, proving (1.47), page 29 using this notation. That is, we wish to show that \( (P \land Q) \implies P \) by showing

\[
(P \land Q) \implies P \iff T.
\]

The proof will run as follows:

\[
(P \land Q) \implies P \iff \neg (P \land Q) \lor P \\
\iff (\neg P) \lor (\neg Q) \lor P \\
\iff (\neg P) \lor P \lor (\neg Q) \\
\iff T \lor (\neg Q) \\
\iff T, \quad q.e.d.
\]

This kind of proof requires much practice, but also can be very satisfying when completely mastered. The highly interested student can work all of the exercises, and perhaps continue by attempting, without truth tables, proofs of the remaining valid equivalences and implications listed in the next subsection. We leave a couple of more difficult examples for later in the next (and final) subsection.

Example 1.6.3 Here we expand the simple statement \( P \iff Q \) using, respectively, the obvious equivalent statement, followed by (1.29), page 22, several applications of the distributive rule
(1.27), page 22 and finally (1.39), page 26.\(^{67}\)

\[
P \iff Q \iff (P \implies Q) \land (Q \implies P) \iff [\sim (P \lor Q) \land \sim (Q \lor P)]
\]

\[
\iff \left\{ \left[ (\sim P) \lor Q \right] \land \left[ (\sim Q) \lor P \right] \right\} \lor \left\{ \left[ (\sim P) \lor Q \right] \land P \right\}
\]

\[
\iff \left\{ \left[ (\sim P) \land (\sim Q) \right] \lor \left[ Q \land (\sim Q) \right] \right\} \lor \left\{ \left[ (\sim P) \lor Q \right] \lor \right\}
\]

\[
\iff \left\{ \left[ (\sim P) \land (\sim Q) \right] \lor \right\} \lor \left\{ P \land Q \right\}
\]

Though we had a good understanding of \(P \implies Q\) before—that it is true if and only if \(P\) and \(Q\) are both true or both false—we might not have been as quick to declare obvious what we just derived when read backwards:

\[
(P \land Q) \lor (\sim (P) \land (\sim Q)) \iff P \implies Q.
\]  \(\text{(1.107)}\)

**Example 1.6.4 (A Rather Sophisticated Proof)** We will prove (1.52), page 32 using its predecessors, particularly (1.49) and the distributive laws. The proof is not as straightforward as using truth tables, but it is worthwhile to follow along.

\[
(P \implies Q) \land (Q \implies R) \iff [\sim (P) \lor Q] \land [\sim Q \lor R]
\]

\[
\iff \left\{ \left[ (\sim P) \lor Q \right] \land \left[ (\sim Q) \lor R \right] \right\} \lor \left\{ \left[ (\sim P) \lor Q \right] \land R \right\}
\]

\[
\iff \left\{ \left[ (\sim P) \land (\sim Q) \right] \lor \left[ Q \land (\sim Q) \right] \right\} \lor \left\{ \left[ (\sim P) \lor Q \right] \lor \right\}
\]

\[
\iff \left\{ \left[ (\sim P) \land (\sim Q) \right] \lor \left\{ \left[ (\sim P) \lor Q \right] \lor \right\} \right\}
\]

\[
\iff \left\{ \left[ (\sim P) \land (\sim Q) \right] \lor \left\{ \left[ (\sim P) \lor Q \right] \lor \right\} \right\}
\]

\[
\iff \left\{ (A \lor B) \land \left\{ \left( (\sim P) \lor R \right) \land \left( (\sim Q) \lor R \right) \right\} \right\}
\]

\[
\iff \left\{ \left( (A \lor B) \lor \left( (\sim Q) \lor R \right) \right) \land \left\{ (\sim P) \lor R \right\} \right\}
\]

\[
\iff \left\{ \left( (A \lor B) \lor \left( (\sim Q) \lor R \right) \right) \land \left\{ (\sim P) \lor R \right\} \right\}
\]

\[
\iff P \implies R, \text{ q.e.d.}
\]

Here we used some labels and substitutions \(A, B, C, D\) for clarity and space considerations. The actual form of \(D\) did not matter. Our final step was of the form \(D \land (P \implies R) \implies P \implies R\), i.e., (1.49), page 32. Note also that it is enough to have \(\implies\) all the way down in such a proof.

\(^{67}\)We include the references to previous valid equivalences and implications in case they are needed, but the reader should attempt to read the proof first without resorting to the references. Indeed the reader might not feel the need to look up the previous results at all, if comfortable with each step.
since we could then follow the “arrows” from the first statement through to the last. However it is best to work with equivalences as long as possible, to avoid losing too much information in early steps where the goal is not so easily concluded.

Clearly this can be a messy process. Fortunately, we rarely need to expand things out so far. What is important is that we now have a language of logic and some very general valid equivalences and implications which

1. we can prove using truth tables—a process with an obvious road map—if necessary;

2. will become more and more natural for us to see as valid by inspection, and be prepared to use;

3. and (later) will make shorter, clearer work of our mathematical arguments.

Some further observations and insights are introduced in the exercises. Many are very important and will be referred to throughout the text.
Exercises

1. Show without truth tables the following. (Hint: There is a very easy way and a very long way.)

\[(P \rightarrow Q) \land ((\neg P) \rightarrow (\neg Q)) \iff P \iff Q.\] (1.108)

Does this make sense? (You should relate this to English statements of these forms.)
1.7 Logic Epilogue

Most mathematical statements begin with hypotheses, i.e., assumptions, and end in conclusions which follow from the assumptions through valid logic. Thus, any mathematics which is done correctly essentially states a tautology. The valid logic connecting assumptions to conclusions constitutes a proof that the assumptions imply the conclusions. Once a mathematical statement has been proven to be a tautology (i.e., always true), it is called a theorem.\footnote{There are also lemmas and corollaries, but these are also proved, and are technically theorems in their own rights. Often the term theorem refers to tautologies the author finds the most interesting, while lemmas are preliminary results (often where most of the work is done in proving the theorems) which lead to theorems, and corollaries are results which follow quickly from theorems.}

Because a theorem is a tautology, it is logically equivalent to $T$. This is quite useful. Recall

$$P \iff P \land T.$$  \hspace{1cm} (1.109)

The content of every theorem resides in $T$.

Figure 1.11: A model for mathematical reasoning. To show the conclusions follow from the assumptions, one usually attaches some theorems which are relevant to the assumptions. This is valid since every statement $P$ is equivalent to $P \land T$, and every theorem is equivalent to $T$. 

\[(\text{Assumptions}) \iff [\text{(Assumptions)} \land T] \implies \text{(Conclusions)} \downarrow \text{ Theorems}\]
Chapter 2

Real Numbers, Algebra and Functions (Optional)

It should be pointed out that in order to successfully study calculus, it is crucial to have a reasonably good background in college algebra, or at least a fairly rigorous knowledge of senior-level high school algebra. As we move through the textbook, trigonometry also becomes increasingly important. It is difficult but possible for a motivated student to learn enough trigonometry “along the way” to complete a calculus course, but if a student’s algebra is very weak that usually dooms the student to fail a study of calculus.

This chapter is still preliminary to the calculus. Its purpose is not to revisit every mathematical topic needed to study calculus. Rather, we will visit many of these topics in perhaps more sophisticated ways than they are usually learned. We need a more advanced mindset for the study of calculus than is needed in preliminary studies of algebra and trigonometry.

The chapter begins with a very sophisticated discussion of the formal definition of the real numbers. While this discussion will sometimes be at a somewhat advanced level, it is useful to recount the properties of $\mathbb{R}$ so we can better avoid common arithmetic and algebraic mistakes. A section on algebraic considerations follows, and then three sections concerning the functions encountered in the text.

Now it should be pointed out that many students who are not completely proficient in the prerequisite algebra or trigonometry finally discover many of the nuances of these two fields in the study of calculus. Technicalities of algebra and trigonometry that may have been ignored or de-emphasized in those courses can loom large in importance in rather common calculus problems, and students usually respond by sharpening their pre-calculus skills along the way. Put another way, applying algebra and trigonometry to the very rich topics of calculus makes them “come alive” for many students.

However, a “D” student in algebra is very unlikely to pass a calculus course. In fact at university, a “C” student in algebra will have much difficulty passing calculus, while for “B” or “A” students in algebra, it is often difficult to predict the level of success—or failure—in subsequent calculus courses. This is because calculus is so different from algebra or trigonometry, despite these being prerequisites.\footnote{Once into the calculus sequence, it is easier to predict future success. A “C” student in Calculus I has to work much harder to earn a “C” in Calculus II, and a “D” student in Calculus I is almost certain to fail Calculus II. In fact it is common to drop one letter grade from Calculus I to Calculus II, but most find multivariable calculus (Calculus III) easier than Calculus II. For our purposes Calculus I is, roughly, through Chapter 6.}

Almost all calculus textbooks begin with a short chapter devoted to review of some of the
algebra and trigonometry skills that are needed later in the textbook. We will have something similar here, but for a few reasons it will differ significantly from the usual approach.

- First, it is difficult or impossible to include every technicality which will occur later in the text, and still have a coherent review.

- Second, many students have seen some of the concepts in two, three, or even more classes and thus tend to “go through the motions” of solving the problems without thinking about the logic, and so a repeat of the same concepts in the same settings will probably do little good.\(^2\)

- Finally, the student reading along this textbook has some knowledge of formal symbolic logic, and this can better illuminate what may seem like mundane topics.

This chapter does therefore contain a review of algebra and trigonometry, but with a more sophisticated (and perhaps less comprehensive) approach. Much of the notation used here is familiar to senior-level mathematics majors, and much of it is even smuggled quietly into the lectures of calculus professors, but it is rarely found in calculus textbooks.

We begin with all of the axioms that define the real numbers \(\mathbb{R}\), in part to help us to see how we can manipulate and solve certain equations and inequalities. We then proceed to a section dealing with polynomial and radical methods, and then to functions, and finally trigonometric functions.

It is important for the student to read this chapter, but it is also important not to bog down here. This can be read quickly and revisited on occasion as one reads through the text. As mentioned before, the actual calculus begins in the next chapter.

For reference we mention again some special sets of numbers:

\[
\begin{align*}
\text{Natural Numbers:} & \quad \mathbb{N} = \{1, 2, 3, 4, \cdots\} \quad (2.1) \\
\text{Integers:} & \quad \mathbb{Z} = \{\cdots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \cdots\} \quad (2.2) \\
\text{Rational Numbers:} & \quad \mathbb{Q} = \{p/q \mid (p, q \in \mathbb{Z}) \land (q \neq 0)\} \quad (2.3) \\
\text{Real Numbers:} & \quad \mathbb{R} \text{ (to be defined)} \quad (2.4)
\end{align*}
\]

\(^2\)Some of the reviews found in the textbooks are actually quite good in summarizing the algebra and trigonometry principles needed in the course. However, particularly in college courses it is rare that a professor has the time or inclination to follow through with the complete review. Indeed, for whatever reasons the professor is usually impatient to get to the calculus, which for us starts with Chapter 3.

This resistance to starting with a standard algebra and trigonometry review is actually reasonable for pedagogical reasons. The tone set for the course by a review of high school algebra is arguably incompatible with a tone appropriate for the much more sophisticated study of calculus. In the author’s experience, it lowers the expectations of the students for load and level of the work required for the actual calculus, and thus can be very counter-productive. Indeed, for calculus it usually works better to communicate to the students that these skills are expected to be already possessed, putting students who may be somewhat weak “on notice” that they will have to work even harder, to resolve those weaknesses and not be “too comfortable” with their pre-calculus skills.
2.1 Real Numbers Defined

We will see that there is something special about the set \( \mathbb{R} \) of real numbers. This is not to say \( \mathbb{N}, \mathbb{Z} \) and \( \mathbb{Q} \) are not interesting. After all, \( \mathbb{N} \) has infinitely many elements, \( \mathbb{Z} \) is infinite in two directions and contains zero and negative numbers, and \( \mathbb{Q} \) allows us to divide any member by any other member except zero and the result will still be rational. In fact the leap from \( \mathbb{Z} \) to \( \mathbb{Q} \) is impressive for more than algebraic reasons since—unlike \( \mathbb{N} \) and \( \mathbb{Z} \)—the rational number set \( \mathbb{Q} \) is “infinitely crowded” in the sense that between any two rational numbers is another rational number. Indeed, given \( x_1, x_2 \in \mathbb{Q} \), where \( x_1 \neq x_2 \), we can find the number halfway between these, say \( x_3 \), and we will show that it too will be rational (a ratio of integers). To see this, we begin with the definition of \( x_1, x_2 \in \mathbb{Q} \): \( x_1 = \frac{p_1}{q_1}, x_2 = \frac{p_2}{q_2} \), where \( p_1, p_2, q_1, q_2 \in \mathbb{Z}, q_1, q_2 \neq 0 \). Then \( x_3 \) is given by

\[
x_3 = \frac{x_1 + x_2}{2} = \frac{1}{2} \left[ \frac{p_1}{q_1} + \frac{p_2}{q_2} \right] = \frac{p_1 q_2 + p_2 q_1}{2 q_1 q_2} \in \mathbb{Q}.
\]

That \( x_3 \in \mathbb{Q} \) is due to the fact that the numerator and denominator of \( x_3 \) are both integers (since sums and products of integers are also integers), and \( q_1 q_2 \neq 0 \) (see Theorem 2.1.8). It is not hard to see that we can repeat this construction to find a rational number halfway between \( x_3 \) and \( x_1 \) (or between \( x_3 \) and \( x_2 \)), and then find a rational number halfway between the new number and \( x_1 \), and so on ad infinitum\(^3\), and thus find infinitely many rational numbers between any such \( x_1, x_2 \in \mathbb{Q} \). Note that this is not the case with either \( \mathbb{N} \) or \( \mathbb{Z} \).

Despite \( \mathbb{Q} \)'s apparent density of elements, there are very important omissions in assuming all numbers are rational. For instance, it was known even to the ancient Greeks that numbers like \( \sqrt{2} \) exist, but are not rational. To see it exists, we need only consider an isosceles right triangle with small sides 1 and 1 as in Figure 2.1. (Recall that we know the hypotenuse \( c \) will be some number whose square is 2 because \( 1^2 + 1^2 = c^2 \).) So such a number does exist (in the sense that we can measure it as a length!).

That \( \sqrt{2} \notin \mathbb{Q} \) is not so obvious, but a proof is not difficult. The jist of a proof would go like this: Suppose \( \sqrt{2} \in \mathbb{Q} \). Then \( \sqrt{2} = p/q \), where \( p \) and \( q \) are integers. Squaring both sides we get \( 2 = p^2/q^2 \), upon which multiplying by \( q^2 \) gives \( 2q^2 = p^2 \). But if we were to do a prime factorization of both sides of this equation, the left-hand side would have an odd number of factors of 2, and the right-hand side an even number of factors of 2. This is a contradiction, so our assumption \( \sqrt{2} \in \mathbb{Q} \) must be false. (Recall \( P \rightarrow \mathcal{F} \iff (\sim P) \). Here \( P : \sqrt{2} \in \mathbb{Q} \).)

The argument that \( \sqrt{2} \notin \mathbb{Q} \) was not based on calculus, and it would be a long distraction to justify it rigorously here.\(^4\) Nonetheless we should take note that, despite its density, the set \( \mathbb{Q} \)

\(^3\)Latin, meaning without end (literally, to infinity).

\(^4\)It should still at least have the ring of truth to the reader familiar with elementary algebra and prime
2.1. REAL NUMBERS DEFINED

of rational numbers is not sufficient to make measures on a *continuum* like the real line. (Recall Figure 1.2.)

But we are not being careful in the above discussion. To remedy this, we will list thirteen axioms which have been used historically to define the real numbers. We will then derive some familiar properties of the real numbers, including their arithmetic, and show how to exploit these to solve algebraic problems.

Indeed the real numbers are special. Technically, they form the smallest ordered field\(^5\) which contains \(\mathbb{Q}\) and possesses the least upper bound property. Less formally \(\mathbb{R}\) is the set which can be identified member-by-point with the geometric line we regard as a continuum. And its elements are precisely those we must include in order to do calculus.

Before we go on, we should recognize that many students (and even some professionals!) perform calculations without knowing the underlying principles which justify their work. Much worse, students are often mistaken about what can and cannot be done, for example to both sides of an equation or inequality. It is important that we be always faithful to valid mathematical principles or our work will be incorrect.

At the same time, we should strive for clarity in our work, or else we will again be at much greater risk of error. We should be precise in our use of notation. We should avoid consolidating too many steps. With every manipulation, we should check that it is valid. If we are clear in our presentation, we can re-read (rather than re-do) our work to check for errors. Sometimes this requires us to re-examine, and possibly discard old algebraic habits, so let the reader beware!

In fairness, it should be noted that the next section is very technical. To exercise one’s ability to think abstractly, Section 2.1 should be re-read on occasion, for it is usually difficult to fully understand and appreciate upon the first reading. In fact, it is not likely to be transparent upon a first reading. If that is the case, it is useful to go on to other sections, and then revisit Section 2.1 in light of what is learned in later sections. The material in the later sections is likely familiar, but our logic studies should cast a new light upon the algebraic material there.

\(^5\)See comments after A8, page 81 for the definition of field.

\[\begin{align*}
3000 &= 3 \cdot 1000 = 3 \cdot 2 \cdot 500 = 3 \cdot 2 \cdot 2 \cdot 250 = 3 \cdot 2 \cdot 2 \cdot 2 \cdot 125 = 3 \cdot 2^3 \cdot 5 \cdot 25 \\
&= 2^3 \cdot 3^1 \cdot 5^3.
\end{align*}\]

So we have factored 3000 into a product of powers of prime numbers. A fact used often in elementary arithmetic is that such a factorization is unique, no matter what path we take to achieve a prime factorization (for instance, we could have begun by dividing by 5 first) so the number of factors of 2, for instance, is determined uniquely at the start, for the given natural number being factored. As we see in our example, there are exactly 3 factors of 2 in the prime factorization of 3000.
2.1.1 Axioms

In mathematics, we begin with fundamental statements which we assume to be true, and so we state them without proof. These are called **axioms**.\(^6\) In particular, we assume the following rules about the real numbers \(\mathbb{R}\), and these in turn are collectively the defining properties of \(\mathbb{R}\) with the familiar operations of addition (+) and multiplication(\(\cdot\)) (note that we often simply write \(xy\) instead of \(x \cdot y\)).

**A1. Closures** \(\mathbb{R}\) possesses two well-defined operations + and \(\cdot\) so that\(^7\)

\[
(\forall x, y \in \mathbb{R})(x + y \in \mathbb{R}) \land (xy \in \mathbb{R}).
\]

In other words, we do not leave the real numbers when we take two real numbers and add or multiply them. To state these separately, we say the real numbers are **closed under addition**, and **closed under multiplication**.

Note also that \(x + y\) is called the **sum** of \(x\) and \(y\), while \(xy\) is called the **product** of \(x\) and \(y\).

**A2. Commutative Properties of Addition and Multiplication**

\[
(\forall x, y \in \mathbb{R})(x + y = y + x) \land (xy = yx).
\]

These are just the familiar properties that order does not matter when we add, and does not matter when we multiply. Of course, if we do both in the same expression we must be more careful (see A8).

**A3. Associative Properties of Addition and Multiplication:**

\[
(\forall x, y, z \in \mathbb{R})\left\{(x + (y + z) = (x + y) + z) \land [x(yz) = (xy)z]\right\}
\]

Again, these should be familiar. The idea is that, in a sum we can group terms however we like, and that the same holds for a product. However, again we need to be cautious if we do both (see again A8).

**A4. Existence of an Additive Identity** \((\exists 0 \in \mathbb{R})(\forall x \in \mathbb{R})(0 + x = x)\).

At this point we are departing slightly from the natural numbers \(\mathbb{N}\), in that we require the presence of a zero element. Without it, the next axiom would be meaningless, and without that axiom, there would be no subtraction.

We should also note that the mere existence of such an identity implies its uniqueness:

**Theorem 2.1.1** \((\exists ! y \in \mathbb{R})(\forall x \in \mathbb{R})(x + y = x)\).

Note that existence is an axiom, so the real import of the theorem is that the additive identity element of \(\mathbb{R}\) is unique.

---
\(^6\)Also known as postulates. Borrowing an image from chemistry, we can identify axioms with the atomic elements we construct molecules (mathematical systems) from, with valid logic assuming the role of chemical bonds. Recall Section 1.7.

\(^7\)By well-defined, we mean that any time we combine the same two numbers in the same order, we get the same result. In other words, \(x + y\) and \(x \cdot y\) stay the same if \(x\) and \(y\) stay the same. Thus \(2 + 3\) gives the same result every time it is computed, and so does \(2 \cdot 3\).
2.1. REAL NUMBERS DEFINED

Proof: Existence is assumed in A4. To prove uniqueness, suppose 0, 0’ are both additive identities. Then

\[ 0 + 0’ = 0 \quad \text{since } 0’ \text{ is an additive identity, and} \]
\[ 0’ + 0 = 0’ \quad \text{since 0 is an additive identity}. \]

But then commutivity gives \( 0 = 0 + 0’ = 0’ + 0 = 0’ \), so 0 = 0’. Thus any two additive identities must be the same, q.e.d.

A5. Existence of Additive Inverses \((\forall x \in \mathbb{R})(\exists y \in \mathbb{R})(x + y = 0)\).

Now we can define subtraction in the familiar way (see Section 2.2: \(x - y = x + (-y)\)).

From this axiom and the closure under addition, we have closure under subtraction, for

\[ x, y \in \mathbb{R} \implies x, -y \in \mathbb{R} \implies x - y = x + (-y) \in \mathbb{R}, \]

where the first implication follows from A5 and the second from A1.

The uniqueness of the additive inverse now follows in a manner similar in spirit to the previous uniqueness theorem:

Theorem 2.1.2 \((\forall x \in \mathbb{R})(\exists! y \in \mathbb{R})(xy = x)\).

Proof: The existence is given by A5. For uniqueness, suppose \(y_1, y_2 \in \mathbb{R}\) are additive inverses of \(x\). Also let \(-x\) denote some fixed additive inverse of \(x\). Then

\[ x + y_1 = 0 = x + y_2 \]
\[ \implies x + y_1 = x + y_2 \]
\[ \implies x + y_1 - x = x + y_2 - x \quad \text{(by the well-defined nature of addition, A1)} \]
\[ \implies x - x + y_1 = x - x + y_2 \]
\[ \implies 0 + y_1 = 0 + y_2 \]
\[ \implies y_1 = y_2, \quad \text{q.e.d.} \]

A6. Existence of a Multiplicative Identity \((\exists 1 \in \mathbb{R})(\forall x \in \mathbb{R})(1 \cdot x = x)\).

Just as the presence of 0 \(\in \mathbb{R}\) was needed for subtraction, so is the presence of 1 \(\in \mathbb{R}\) necessary for division. Furthermore, just as existence implies uniqueness in A4 and A5, the same is true here. The proof of the following is left as an exercise (See Exercise 1).

Theorem 2.1.3 \((\exists! y \in \mathbb{R})(\forall x \in \mathbb{R})(xy = x)\).
With this axiom we are almost ready to construct \( \mathbb{N} \) and \( \mathbb{Z} \), except that—upon careful inspection—we notice that there is no guarantee that \( 1 \neq 0 \). If we knew for certain that \( 1, 0 \) were separate elements of \( \mathbb{R} \), then we could define \( 2 = 1 + 1 \in \mathbb{R}, 3 = 2 + 1 \in \mathbb{R}, \ldots \), allowing us to construct \( \mathbb{N} \), as well as defining \( -1 = 0 - 1 \in \mathbb{R}, -2 = -1 - 1 \in \mathbb{R}, \) etc., to finish the job of constructing \( \mathbb{Z} \) as well. Unfortunately, if \( 1 = 0 \) then these are all the same numbers, and so we would just be describing \( \{0\} \). Of course we expect \( \mathbb{R} \) to contain more than one element, or it would be of little use. So we must assume more.

**A7. Existence of Multiplicative Inverses**

\[
(\forall x \in \mathbb{R} - \{0\})(\exists x^{-1} \in \mathbb{R})(xx^{-1} = 1).
\]

Just as we had uniqueness of the additive inverse, so too can we prove uniqueness of the multiplicative inverse:

**Theorem 2.1.4** \((\forall x \in \mathbb{R} - \{0\})(\exists! y \in \mathbb{R})(xy = 1)\).

The proof is the same as for the additive inverse, except we replace \(+\) with \(\cdot\), \(0\) with \(1\) and \(-x\) with \(x^{-1}\). Note that this is all vacuously true if \(\mathbb{R} = \{0\}\), but we will eventually prove that \(1 \neq 0\), and then we will know \(\mathbb{R}\) is more than just \(\{0\}\).

At this point we can define division (and ratios) by defining, for all \(x,y \in \mathbb{R}, y \neq 0\) the ratio

\[
x/y = x \cdot y^{-1}.
\]

(2.5)

**A8. Distributivity of Multiplication over Addition** \((\forall x,y,z \in \mathbb{R})\)

\[
x(y + z) = xy + xz.
\]

(2.6)

Note that we carry out multiplications before additions on the right hand side of (2.6), i.e., \(xy + xz = (xy) + (xz)\). One intuitive way to think about (2.6) is that the left-hand side represents \(x “(y + z)’s,”\) while the right-hand side represents \(x “y’s”\) plus \(x “z’s.”\) Another way to look at it, at least for \(x,y,z\) all positive (which we define later), to compare the areas that both sides of (2.6) represent, as in Figure 2.2.
Axioms A1–A8 are algebraic in nature and are called field axioms; any set which has a structure conforming to A1–A8 is called a field. \( \mathbb{Q} \) is a field, as is \( \mathbb{R} \), though \( \mathbb{N} \) and \( \mathbb{Z} \) are not.

We now prove a theorem of arithmetic:

**Theorem 2.1.5** \((\forall x \in \mathbb{R})(x \cdot 0 = 0)\).

**Proof:** Suppose \( x \in \mathbb{R} \). Then by the nature of 0 and the distributive axiom we have

\[
\begin{align*}
  x \cdot 0 &= x \cdot (0 + 0) = x \cdot 0 + x \cdot 0 \\
  \implies x \cdot 0 &= x \cdot 0 + x \cdot 0 \\
  \implies x \cdot 0 - x \cdot 0 &= x \cdot 0 + x \cdot 0 - x \cdot 0 \\
  \implies 0 &= x \cdot 0, \quad \text{q.e.d.}
\end{align*}
\]

Next we have the axioms of order:

**A9. Trichotomy** \((\forall x, y \in \mathbb{R})[(x < y) \lor (y < x) \lor (x = y)]\). Furthermore, the three cases are mutually exclusive. Thus,

\[
\begin{align*}
  (\forall x, y \in \mathbb{R})[(x < y) \iff ((y < x) \lor (x = y))], \\
  (\forall x, y \in \mathbb{R})[(y < x) \iff ((x < y) \lor (x = y))], \\
  (\forall x, y \in \mathbb{R})[(x = y) \iff ((x < y) \lor (y < x))].
\end{align*}
\]

In other words, exactly one of these three cases will hold for a given \( x \) and \( y \).

**A10. Transitivity** \((\forall x, y, z \in \mathbb{R})[(x < y) \land (y < z) \implies (x < z)]\).

These first two axioms of order, A9 and A10, allow us to position elements of \( \mathbb{R} \) on a line in a consistent order from left to right. Two elements \( x \), \( y \) \( \in \mathbb{R} \) are placed so that \( x \) is left of, right of, or at the same position as \( y \), and the ranking is consistent in that if \( z \) is right of \( y \) (i.e., \( y < z \)), and \( y \) is right of \( x \) (\( x < y \)), then \( z \) is to the right of \( x \) (\( y < z \)) on the line.

Before moving to the next axiom, we make mention of a couple familiar definitions.

**Definition 2.1.1** For a given \( x, y \in \mathbb{R} \), define \( >, \leq, \geq \), positive, negative, nonpositive, nonnegative, less than, greater than, less than or equal to, and greater than or equal to as follow:

\[
\begin{align*}
  x > y &\iff y < x \\
  x \leq y &\iff (x < y) \lor (x = y) \\
  x \geq y &\iff (x > y) \lor (x = y).
\end{align*}
\]

Next, \( x \) is called positive iff \( x > 0 \), and negative iff \( x < 0 \). We also call \( x \) nonpositive iff \( x \leq 0 \), and nonnegative iff \( x \geq 0 \). For \( x < y \), we say \( x \) is less than \( y \), or equivalently, \( y \) is greater than \( x \). For \( x \leq y \), we say \( x \) is less than or equal to \( y \), or equivalently, \( y \) is greater than or equal to \( x \).

**A11. Translation Preserves Order** \((\forall x, y, z \in \mathbb{R})[(x < y) \implies x + z < y + z]\).

This is probably familiar, and is a very important property of \( \mathbb{R} \). Later we will take advantage of the fact that \( \mathbb{R} \) can be interpreted as a set of translations\(^{10}\) on the number line (which is also a representation of \( \mathbb{R} \)) from any fixed point.

\(^{10}\)The term translation motion, in this case to the left or right on the number line, which is introduced in Figure 2.3. It may seem strange to think of \( \mathbb{R} \) as translations along \( \mathbb{R} \), but it is really no different than contemplating transporting one vehicle on top of another vehicle. Even if they are exactly the same type of vehicle, it is not hard to imagine circumstances when such an arrangement would be useful.
CHAPTER 2. REAL NUMBERS, ALGEBRA AND FUNCTIONS (OPTIONAL)

-5 -4 -3 -2 -1 0 1 2 3 4 5

Figure 2.3: Real number line, with positive numbers to the left of zero, and negative numbers to the right of zero. See also Figure 1.2.

A12. Existence of Positive Numbers \((\exists x \in \mathbb{R})(x > 0)\).

After adding this assumption we know that \(\mathbb{R} \neq \{0\}\), since \(\mathbb{R}\) contains numbers which are greater than 0 and therefore cannot be equal to 0.

A13. Product of Positives is Positive \((\forall x, y \in \mathbb{R})(x > 0) \land (y > 0) \implies xy > 0\).

With these axioms we can prove a very simple but important theorem which helps us construct a geometrical interpretation of \(\mathbb{R}\):

Theorem 2.1.6 \(1 > 0\).

Proof: First we show that \(1 \neq 0\). Suppose \(1 = 0\). By A12, we can choose some real \(x > 0\). Then by Theorem 2.1.5 and A13, we have \(0 = 0 \cdot x = 1 \cdot x = x > 0\), which is a contradiction (since it implies \(0 > 0\)). We must conclude \(0 \neq 1\).

Since we showed above that \(\sim (0 = 1)\), by trichotomy we now need only show \(\sim (1 < 0)\) to conclude the desired result.

Suppose \(1 < 0\). By translation by \(-1\) (i.e., adding \(-1\) to both sides) we would get \(0 < -1\), i.e., \(-1 > 0\). But then we would have \(1 = (-1)(-1) > 0\) by A12, which gives \(1 > 0\), contradicting our assumption that \(1 < 0\). Hence the assumption is false, and \(\sim (1 < 0)\), so (with \(1 \neq 0\)) we conclude \(1 > 0\), q.e.d.

With this theorem and the axioms of order, we are led to a geometric interpretation of the real numbers, as we saw earlier (in Figure 1.2 of Section ??). In particular, with the theorem and the translation invariance of \(<\), we can add 1 or \(-1\) to both sides repeatedly to get

\[
\cdots \leftarrow -2 \leftarrow -1 \leftarrow -1 < 0 \leftarrow 0 < 1 \implies 1 < 2 \implies 2 < 3 \implies \cdots.
\]

Indeed, we are justified in representing the real numbers by a horizontal line with negative numbers to the left of zero, and positive numbers to the right, as in Figure 2.3.

Actually, at this point the axioms describe exactly the rational numbers \(\mathbb{Q}\) and not the real numbers \(\mathbb{R}\). The final axiom of the real numbers is at the heart of the idea of continuum, and is perhaps the deepest of the thirteen axioms. It is called the least upper bound property. First we begin with some definitions:

Definition 2.1.2 Given a set \(S \subseteq \mathbb{R}\),

1. Suppose \((\exists m \in \mathbb{R})(\forall s \in S)(m \leq s)\). Then we say \(S\) is bounded from below, and call \(m\) a lower bound for \(S\).

2. Suppose \((\exists M \in \mathbb{R})(\forall s \in S)(M \geq s)\). Then we say \(S\) is bounded from above, and call \(M\) an upper bound for \(S\).
3. If such \( m \) and \( M \) both exist, we say \( S \) is **bounded**. Otherwise we say \( S \) is **unbounded** (meaning from above, or below, or both).

If \( S \) has finitely many elements, it will be bounded, i.e., \((\forall s \in S)(m \leq s \leq M)\), where \( m \) is the least and \( M \) the greatest element of \( S \), respectively. On the other hand, \( S \) can have infinitely many elements and still be bounded. For instance, if \( S = \{ x \in \mathbb{R} \mid -2 < x < 5 \} \), then \( m = -2 \) and \( M = 5 \) will work. Of course so will \( m = -10 \), \( M = 1000 \). But there is something special about our first choices. There is no greater choice for \( m \) which will work, and there is no lesser choice for \( M \) which will work, then \( m = -2 \) and \( M = 5 \). We call \( m \) a **greatest lower bound** (glb) of \( S \), and \( M \) a **least upper bound** (lub) of \( S \).

**Definition 2.1.3** Given a set \( S \). Define the following, if they exist:

\[
M = \text{lub}(S) \iff ((\forall s \in S)(s \leq L)) \implies (M \leq L) \\
m = \text{glb}(S) \iff ((\forall s \in S)(s \leq l)) \implies (m \geq l).
\]

In other words, \( M \) is less than or equal to every upper bound for \( S \), and \( m \) is greater than or equal to every lower bound for \( S \). Remember that not all sets are bounded, and a set can be bounded from above and not below, or vice-versa. But if a set has a bound, then our final axiom applies.

**A14. Least Upper Bound Property** If \( S \subseteq \mathbb{R} \) has an upper bound, then it has a least upper bound. Restated,

\[
(\exists L \in \mathbb{R})(\forall s \in S)(s \leq L) \implies (\exists M \in \mathbb{R})(M = \text{lub}(S)). 
\]

(2.10)

This is equivalent to the existence of a lower bound for \( S \) implying existence of \( \text{glb}(S) \), a fact which is left as an exercise but stated below:

**Theorem 2.1.7** If \( S \subseteq \mathbb{R} \) has a lower bound, then it has a greatest lower bound.

As stated just above, Theorem 2.1.7 \( \iff \) A14.

Now we consider a simple example which demonstrates how the least upper bound property distinguishes the real numbers, \( \mathbb{R} \), from the rationals, \( \mathbb{Q} \). Consider the decimal expansion for \( \sqrt{2} \).

We can construct a set \( S \) which contains all the finite expansions:

\[
S = \{1, 1.4, 1.41, 1.414, 1.4142, 1.41421, 1.414213, 1.4142135, 1.41421356, \ldots \}.
\]

Now, \( S \subseteq \mathbb{Q} \) since all of its elements are in \( \mathbb{Q} \): \( 1 \in \mathbb{Q} \), \( 1.4 = \frac{14}{10} \in \mathbb{Q} \), \( 1.41 = \frac{141}{100} \in \mathbb{Q} \), etc. Furthermore, clearly \( S \) is bounded from above by 2. Of course we can find lesser upper bounds even with rational numbers, such as 1.42, or 1.415, or 1.4143, or 1.41422, etc. However, for any rational upper bound we find, we can find a better (i.e., lesser) rational upper bound. This leads us to conclude that \( \mathbb{Q} \) does not have the least upper bound property. If we instead look to \( \mathbb{R} \), we can see that \( \sqrt{2} \in \mathbb{R} \) is the least upper bound we are seeking for \( S \), i.e., \( \sqrt{2} = \text{lub}(S) \).

---

11 Some texts call \( m = \text{glb} S \) the **infimum** of \( S \), and write \( m = \text{inf} S \), and call \( M = \text{lub} S \) the **supremum** of \( S \), written \( M = \text{sup} S \).

12 Actually, we are **defining** \( \mathbb{R} \) to have numbers like \( \sqrt{2} \) by requiring \( \mathbb{R} \) to contain the least upper bounds of sets such as \( S \). It would be a distraction from our eventual goal—calculus—to be too rigorous in our axiomatic development here. It is really the role of junior or senior real analysis courses to define the real numbers from scratch. In such a course, one begins by defining \( \mathbb{N} \), then generalizes to \( \mathbb{Z} \) (requiring A1–A5), then generalizes further requiring A1–A13 to define \( \mathbb{Q} \), and finally adding A14, so that A1–A14 define \( \mathbb{R} \). At each step one adds another requirement on the abilities of the set and possibly requiring the set to be expanded to have all the new capabilities. In doing so we created the superset hierarchy

\[
\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}.
\]
For most of the remainder of this chapter we will be more concerned with the algebraic (or, more mundanely, the arithmetic) properties of \( \mathbb{R} \) which are contained in A1–A12. However, the calculus relies heavily on A13. In fact, there would be no calculus without it. Much of the time we can perform the mechanics of calculus without worrying about the least upper bound property, but there will be several occasions where it will be prominent in our mathematical arguments. For now we will content ourselves with some important algebraic propositions.

**Theorem 2.1.8 (No Real Zero Divisors)** If \( x, y \in \mathbb{R} \), then
\[
x \cdot y = 0 \iff (x = 0) \lor (y = 0).
\]

**Proof:** We prove the direction “\( \implies \)” by exhausting the cases: \( x \neq 0 \), \( y \neq 0 \), or \( x, y = 0 \).

1. Case \( x \neq 0 \):

\[
(x \cdot y = 0) \land (x \neq 0) \implies (x \cdot y = 0) \land [(\exists x^{-1})(x^{-1}x = 1)]
\]
\[
\implies x^{-1}xy = x^{-1}0
\]
\[
\implies 1y = 0 \implies y = 0.
\]

2. Case \( y \neq 0 \):

\[
(y \cdot x = 0) \land (y \neq 0) \implies (y \cdot x = 0) \land [(\exists y^{-1})(y^{-1}y = 1)]
\]
\[
\implies xy^{-1} = 0y^{-1}
\]
\[
\implies x \cdot 1 = 0 \implies x = 0.
\]

3. Case \( x, y = 0 \). Clearly \( x, y = 0 \implies (x = 0) \lor (y = 0) \), independent of the fact that \( xy = 0 \).

This proves the direction \( \implies \). Conversely, \( (x = 0) \lor (y = 0) \implies xy = 0 \), q.e.d.

This principle is crucial for many problems in algebra. For instance, solving \( x^2 - 9x + 14 = 0 \) is the same as solving \( (x - 7)(x - 2) = 0 \), which occurs if and only if \( (x - 7) = 0 \) \( \lor \ (x - 2) = 0 \), i.e., if and only if \( (x = 7) \lor (x = 2) \). We will make more use of this principle in the next section, and then throughout the rest of the book. Like many other principles, this one can be easily extended as follows:

**Corollary 2.1.1** For any \( x_1, x_2, \ldots, x_n \in \mathbb{R} \),
\[
x_1x_2 \cdots x_n = 0 \iff (x_1 = 0) \lor (x_2 = 0) \lor \cdots \lor (x_n = 0).
\]

We will go ahead and list the proof here to show what such a proof would look like. It is a matter of applying the previous theorem several times.

**Proof:**
\[
x_1x_2x_3 \cdots x_n = 0 \iff x_1(x_2x_3 \cdots x_n) = 0
\]
\[
\iff (x_1 = 0) \lor (x_2x_3 \cdots x_n = 0)
\]
\[
\iff (x_1 = 0) \lor (x_2 = 0) \lor (x_3 \cdots x_n = 0)
\]
\[
\vdots
\]
\[
\iff (x_1 = 0) \lor (x_2 = 0) \lor (x_3 = 0) \lor \cdots \lor (x_n = 0).
\]
2.1. REAL NUMBERS DEFINED

Exercises

1. (Uniqueness of multiplicative inverse) Show that the existence of a multiplicative inverse implies its uniqueness. In other words, prove Theorem 2.1.4. (The proof is similar to that of Theorem 2.1.1, page 78.)

2. Assume $a, b \neq 0$. Show that $(ab)^{-1} = a^{-1}b^{-1}$. (Hint: It is enough to show that $(ab)(a^{-1}b^{-1}) = 1$, since then $a^{-1}b^{-1}$ must then be the inverse of $(ab)$. By definition that inverse is $(ab)^{-1}$. Thus they must be equal.)

3. Assume $a, b \in \mathbb{R}$. Show that $-(a + b) = -a - b$. (This is similar to Exercise 2.)

4. Assume that $x, y \in \mathbb{R}$. Show the following implications, based on the fact that $T \iff ((x < y) \lor (x > y) \lor (x = y))$. See Definition 2.1.1, de Morgan’s laws and Equation (1.58). It might also be helpful to recall that $P \iff P \land T$.

   - $\sim (x \leq y) \implies x > y$
   - $\sim (x < y) \implies x \geq y$
   - $\sim (x = y) \implies (x < y) \lor (x > y)$
   - $(x \leq y) \land (x \geq y) \implies x = y$. (Actually, since in Definition 2.1.1, we had exclusive ors, the above are implications can be equivalences.)

   Hint: So the first computation should begin $\sim (x \leq y) \iff T \land \sim (x \leq y) \iff \cdots$. (a) Find an upper bound within 0.1 of lub(S).
   (b) Find an upper bound within 0.01 of lub(S).
   (c) Find an upper bound within 0.001 of lub(S).
   (d) Find glb(S). (See page 83.)
   (e) How would you write lub(S)?

5. Use A11, page 81 to show that $x < y \implies x - z < y - z$.

   (Hint: $z \in \mathbb{R} \implies (-z) \in \mathbb{R}$. Now what, by definition, is $x - z$?, $y - z$?)

6. Define

   $\mathbb{N} = \{1, 2, 3, \cdots\}$,
   $\mathbb{Z} = \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}$,
   $\mathcal{E} = \left\{ 2k \mid k \in \mathbb{Z} \right\}$,
   $\mathcal{O} = \left\{ 2k + 1 \mid k \in \mathbb{Z} \right\}$.

   In other words, $\mathbb{N}$ and $\mathbb{Z}$ are as before, while $\mathcal{E}$ is the set of all even integers and $\mathcal{O}$ is the set of all odd integers.

   (a) Is $\mathbb{N}$ closed under addition? subtraction? multiplication? division?
   (b) Is $\mathbb{Z}$ closed under addition? subtraction? multiplication? division?
   (c) Is $\mathcal{E}$ closed under addition? subtraction? multiplication? division?
   (d) Is $\mathcal{O}$ closed under addition? subtraction? multiplication? division?

7. Consider the set $S = \{.9, .99, .999, .9999, .99999, \cdots\}$.

   - (a) Find an upper bound within 0.1 of lub(S).
   - (b) Find an upper bound within 0.01 of lub(S).
   - (c) Find an upper bound within 0.001 of lub(S).
   - (d) Find glb(S). (See page 83.)
   - (e) How would you write lub(S)?

8. Find lub(S) if $S = \{.9, .99, .999, .9999, .99999, \cdots\}$.

9. Show that $\frac{a}{b} = \frac{c}{d} \implies ad = bc$. To do so, recall that $\frac{a}{b} = ab^{-1}$, and $\frac{c}{d} = cd^{-1}$, and multiply both sides of $\frac{a}{b} = \frac{c}{d}$ by the same expression, which gives an implication because of the fact that multiplication is well-defined.
CHAPTER 2. REAL NUMBERS, ALGEBRA AND FUNCTIONS (OPTIONAL)

2.2 Arithmetic with Real Numbers

There are several familiar manipulations allowed by the axioms of real numbers. When expressions become complicated, we need to be sure our methods of calculating, simplifying or recombining the terms of an expression are consistent with those axioms. In short, it is important that we are aware of exactly what we can do, and what we cannot in general do.

There are a few simple principles which we will employ often, and which are only small steps from our real number axioms. These are important arithmetic ally, and therefore algebraically, and so are necessary for an effective knowledge of calculus. We begin with subtraction and division.

2.2.1 Subtraction is Addition of the Additive Inverse

\[ x - y = x + (-y) \]  
\[(2.11)\]

This is quite useful because, though subtraction is not commutative or associative,\(^{13}\) addition is. Thus if we translate a difference into a sum, we can perform manipulations such as

\[ x - y = x + (-y) = (-y) + x = -y + x, \]

\[ x - y - z = x + (-y) + (-z) = x + (-z) + (-y) = x - z - y. \]

As long as we regard subtracting \(y\) as an abbreviation for adding \((-y)\), we have all the freedom afforded by addition. We just need to be careful in our implementation; the “sign” must always follow the variable.

Next we list several consequences of this approach to subtraction, and a few observations which follow from our real axioms.

- \(-y = (-1)y.\)

To prove this, we need only show that \(y + (-1)y = 0\), because then \((-1)y\) would act as the (unique) additive inverse of \(y\), that additive inverse being \(-y\) by definition. In light of the distributive axiom, this is easy:

\[ y + (-1)y = 1y + (-1)y = (1 + (-1))y = 0y = 0. \]

The last equality follows from Theorem 2.1.5.

- \(-(-y) = y.\) This is rather obvious when we analyze what this equation says: that the additive inverse of \((-y)\) is \(y\) which of course follows from the fact that \((-y) + y = 0.\)

It is worth pointing out that, by our definition of \((-y)\), this new equation and what we know of uniqueness of additive inverses, we have that \(y\) and \((-y)\) are inverses of each other.

- \(-(x + y) = -x - y.\) This was Exercise 3, page 85. Another proof of this is given below:

\[ -(x + y) = (-1)(x + y) = (-1)x + (-1)y = -x - y. \]

- \(-(x - y) = y - x.\) Restated, \(x - y\) and \(y - x\) are additive inverses of each other. One proof is:

\[ -(x - y) = -x - (-y) = -x + (-(-y)) = -x + y = y - x. \]

\(^{13}\)To see subtraction is neither commutative nor associative, notice the examples

\[ 2 - 3 \neq 3 - 2, \]

\[ 3 - (2 - 1) \neq (3 - 2) - 1. \]
2.2. ARITHMETIC WITH REAL NUMBERS

2.2.2 Division is Multiplication by the Reciprocal

\[
\frac{x}{y} = x \cdot y^{-1}
\]  
(2.12)

Here we have to assume that \(y \neq 0\). Of course, given a fraction \(x/y\), we call \(x\) the numerator and \(y\) the denominator of the fraction. The terms reciprocal and multiplicative inverse are interchangeable.

Equation (2.12) is as useful for multiplication and division as Equation (2.11) is for addition and subtraction. Many of the following have their analogs in Subsection 2.2.1.

\(a^{-1} = \frac{1}{a}\). This follows from (2.12), though is read right to left:

\[
\frac{1}{a} = 1 \cdot a^{-1} = a^{-1}.
\]

\(\frac{a}{a} = 1\). This is also obvious from Equation (2.12): \(\frac{a}{a} = a a^{-1} = 1\).

\((ab)^{-1} = a^{-1}b^{-1}\), assuming \(a, b \neq 0\) (and thus \(ab \neq 0\), see Theorem 2.1.8). To prove this, note that \((ab)^{-1}\) is that unique real number such that \((ab)^{-1}(ab) = 1\). We need only show that \((a^{-1}b^{-1})(ab) = 1\) as well, so that we have two expressions for the inverse of \(ab\), so they must be the same.

\[
(a^{-1}b^{-1})(ab) = a^{-1}b^{-1}ab = a^{-1}ab^{-1}b = 1 \cdot 1 = 1, \quad \text{q.e.d.}
\]

We used associativity and commutativity several times here.\(^{14}\)

\(\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}\). Thus, to multiply two fractions, we simply multiply their numerators and multiply their denominators to produce a new fraction. With Equation (2.12), such things become trivial to prove.

\[
\frac{a}{b} \cdot \frac{c}{d} = ab^{-1}cd^{-1} = acb^{-1}d^{-1} = (ac)(bd)^{-1} = \frac{ac}{bd}.
\]

\[(a/b)^{-1} = \frac{b}{a}\]. Here we have to assume that \(a, b \neq 0\).\(^{15}\) Note that this can also be written:

\[
\left(\frac{a}{b}\right)^{-1} = \frac{b}{a}.
\]

\(^{14}\)If we want to be careful in the extreme, we could take one step at a time and write something like

\( \left( a^{-1}b^{-1} \right)(ab) = ((a^{-1})(b^{-1}))(ab) = (a^{-1})( (b^{-1})(ab)) = (a^{-1})( (b^{-1})(ba)) = (a^{-1})( (b^{-1}b)a) = (a^{-1})(1a) = a^{-1}a = 1. \)

Since this can be unnecessarily tiresome, instead we will shorten such arguments with the knowledge that for a “pure product” of many terms (as for a “pure sum”) we can regroup and reorder however we like, ultimately because of such repeated applications of commutative and associative axioms.

\(^{15}\)In essence, we have to assume that nowhere in any expressions we are dividing by zero, i.e., trying to multiply by a multiplicative inverse of 0. This is because no such inverse exists. Indeed, since \(1 \neq 0\) (See Theorem 2.1.6), the existence of such a multiplicative inverse—say some \(z \in \mathbb{R}\)—would lead to a contradiction:

\[1 = z \cdot 0 = 0 \implies \mathcal{F}.\]

Thus

\[(\exists z \in \mathbb{R})(z \cdot 0 = 1).\]  
(2.13)
Another summary of this is the statement that the reciprocal of $a/b$ is $b/a$. The proof is left to the reader.

- $(a^{-1})^{-1} = a$. There are several ways to prove this. For instance,

$$
(a^{-1})^{-1} = \left(\frac{1}{a}\right)^{-1} = \frac{1}{\left(\frac{1}{a}\right)} = a \cdot 1 = a \cdot 1^{-1} = a \cdot 1 = a.
$$

This relied on the earlier items and the fact that 1 is its own multiplicative inverse (since $1 \cdot 1 = 1$). Alternatively, we could have gone back to the definition of multiplicative inverse, the fact that it is unique, and what this equation says in light of the definition: that $a$ is the multiplicative inverse of $a^{-1}$—which is true since $a \cdot a^{-1} = 1$.

This is worth a second (quick) look. Taken with the definition of $a^{-1}$, we conclude that $a$ and $a^{-1}$ are multiplicative inverses (i.e., reciprocals) of each other. It thus makes sense to say $a/b$ and $b/a$ are reciprocals of each other as well.

- $\frac{ab}{bc} = \frac{a}{c}$. This is the familiar “cancellation of common factors” often used to simplify fractions. Note that it requires $b, c \neq 0$. It is sometimes illustrated

$$
\frac{a \cdot \frac{1}{b}}{b \cdot c} = \frac{a}{c},
$$

though we must be careful to not be in the habit of “canceling” in cases where it is invalid.\(^{16}\)

To prove this is not difficult:

$$
\frac{ab}{bc} = (ab)(bc)^{-1} = abb^{-1}c^{-1} = a \left( \frac{b}{b^{-1}} \right) c^{-1} = a \cdot 1 \cdot c^{-1} = \frac{a}{c}.
$$

- $\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a}{b} \cdot \frac{d}{c}$. This is proven below, once we restate it.

The proof of this equation—the title of this subsection—is left as an exercise and follows from the previous equations. We restate it here as a numbered equation

$$
\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a}{b} \cdot \frac{d}{c}. \tag{2.14}
$$

This is what many think of when stating that division by a number (say, $c/d$) is simply the same as multiplication by the reciprocal of that number ($d/c$). There are many possible proofs based upon the previous equations.

### 2.2.3 Addition & Subtraction with Multiplication & Division

We cannot list exhaustively all the possible combinations of these operations, but we can show how the real axioms, further developed in the previous subsections, dictate how we can compute, simplify or recombine the terms of an expression. Some examples are presented in the items below.

---

\(^{16}\)The following manipulation is invalid but tempting, especially with more complicated expressions:

$$
\frac{5 + 3}{5 + 8} = \frac{3}{5} \Rightarrow \mathcal{F}.
$$

The problem is that addition and division are not related the same way that multiplication and division are. It is important to only cancel factors of products, if the factors are common to both the numerator and denominator (which must first both be written as products).
2.2. ARITHMETIC WITH REAL NUMBERS

- \(a(b - c) = ab - ac\). This is not difficult to see, when we recall that \(-c = (-1)c\).

\[a(b - c) = a[b + ((-1)c)] = ab + a((-1)c) = ab + (-1)(ac) = ab - ac.\]

Of course in practice we do not bother with all the intermediate steps, unless things are complicated and we want to be unusually careful.

- \((a + b)(c + d) = ac + ad + bc + bd\).\(^{17}\) This is just the distributive axiom twice:

\[(a + b)(c + d) = a(c + d) + b(c + d) = ac + ad + bc + bd.\]

An alternative method is to distribute the terms as below:

\[(a + b)(c + d) = (a + b)c + (a + b)d = ac + bc + ad + bd,\]

which is the same except for a rearrangement valid by the commutativity of addition.

- \(a(b + c + d) = ab + ac + ad\). Thus the distributive axiom is easily generalized. One form of proof is given below:

\[a(b + c + d) = a(b + (c + d)) = ab + a(c + d) = ab + (ac + ad) = ab + ac + ad.\]

Recall that in a sum of three terms, we do not need parentheses because the associativity axiom guarantees we get the same result regardless of grouping.

- \((a + b + c)(d + e + f) = ad + bd + cd + ae + be + ce + af + bf + cf\) (nine terms). This takes longer to prove but is straightforward.

\[(a + b + c)(d + e + f) = (a + b + c)d + (a + b + c)e + (a + b + c)f
\]

\[= ad + bd + cd + ae + be + ce + af + bf + cf.\]

With practice we see that the product requires all possible terms which are products of one term of \((a + b + c)\) and one from \((d + e + f)\). Perhaps a better way of stating this is that \(a\) must distribute across \((d + e + f)\), and so must \(b\) and \(c\). Again, we have to be careful.

- \((a + b)(a + b) = aa + 2ab + bb\). The proof of this is a simple exercise:

\[(a + b)(a + b) = a(a + b) + b(a + b) = aa + ab + ba + bb = aa + ab + ab + bb
\]

\[= aa + (1 + 1)ab + bb = aa + 2ab + bb.\]

Later we will have more compact ways of writing such things.\(^{18}\)

- \[\frac{a}{c} + \frac{b}{c} = \frac{a + b}{c}.\] This is the familiar manipulation allowed when the fractions have the same, i.e., common, denominators. It is easy to prove:

\[\frac{a}{c} + \frac{b}{c} = ac^{-1} + bc^{-1} = (a + b)c^{-1} = \frac{a + b}{c}.\]

\(^{17}\)This is a statement of what some call “FOIL,” meaning you multiply the “first” terms \((a\) and \(c)\) of each, then the “outer” terms \((a\) and \(d)\), then the “inner” terms \((b\) and \(c)\), and finally the “last” terms \((b\) and \(d)\) of the factors \((a + b)\) and \((c + d)\). This is a fine mnemonic device for such products, but does not so easily generalize to more complicated products.

\(^{18}\)When we have exponential notation we will write this more compactly as

\[(a + b)^2 = a^2 + 2ab + b^2.\]
Notice that we used the distributive axiom but in the reverse of the order it is written. To use the distributive axiom in such a way is referred to as factoring, i.e., writing a sum as a product (when possible). The second equal sign represents such a process, which we will employ throughout the text in different settings.

- \(-\frac{a}{b} = \frac{-a}{b} = \frac{a}{-b}\). Thus we can place the negative sign in any of three places. This is not difficult to see if we recall that the negative sign can be treated as a factor of \(-1\). We leave the details of a proof as an exercise.

- \(\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}\). The most obvious approach to proving this is to read it backwards. Perhaps more satisfying is a proof that reads this left to right while filling in some details. The usual approach then includes a basic mathematics technique which is to multiply by 1, which does not change the value of the terms, but can be done in such a way that 1 is written in a particularly useful form. For the problem at hand we use 1 = \(bb^{-1}\) and 1 = \(aa^{-1}\):

\[
\frac{a}{b} + \frac{c}{d} = ab^{-1} + cd^{-1} = ab^{-1}dd^{-1} + cd^{-1}bb^{-1} = adb^{-1}d^{-1} + bcb^{-1}d^{-1} = \frac{ad + bc}{bd}.
\]

This can be re-written (perhaps more clearly) as follows:

\[
\frac{a}{b} + \frac{c}{d} = \frac{a}{b} \cdot \frac{d}{d} + \frac{c}{d} \cdot \frac{b}{b} = \frac{ad + bc}{bd} = \frac{ad + bc}{bd}.
\]

2.2.4 Integer Exponents

We first define exponents for the positive integer case. For \(a \in \mathbb{R}\) and \(n \in \mathbb{N}\), define

\[
a^n = a \cdot a \cdot a \cdots a, \tag{2.15}
\]

Thus,

\[
a^1 = a, \quad a^2 = a \cdot a, \quad a^3 = a \cdot a \cdot a,
\]

and so on. Here \(a\) is called the base, \(n\) the exponent, and \(a^n\) is called the \(n\)th power of \(a\). There are several rules for this notation which follow quickly.

- \(a^na^m = a^{n+m}\). This is (as are many of these principles) really a simple counting exercise, and so we can easily observe the truth of this:

\[
a^n \cdot a^m = a^{n \cdot m} = a^{n + m}.
\]

\[\text{A common mistake in adding fractions is to just simply add the numerators and add the denominators. In that sense adding fractions is very different from multiplying fractions. It is very important that we take the distinction to heart:}
\]

\[
\frac{a}{b} + \frac{c}{d} = \frac{a+c}{b+d}; \quad \text{but} \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.
\]
2.2. ARITHMETIC WITH REAL NUMBERS

• \((ab)^n = a^n b^n\). This is also straightforward:

\[
(\underbrace{ab \cdots ab}_{n \text{ copies of } ab}) = \underbrace{a \cdots a}_{n \text{ copies of } a} \cdot \underbrace{b \cdots b}_{n \text{ copies of } b} = a^n b^n.
\]

• \((a^n)^m = a^{nm}\). This is also not difficult to see:

\[
(\underbrace{a^n \cdots a^n}_{m \text{ copies}})^m = \underbrace{a^m \cdot a^m \cdots a^m}_{n \text{ copies of } a} = a^{nm}.
\]

Next we define what it means to have a negative power of \(a\). In what follows, assume \(a \neq 0\). For the moment we will also assume \(n \in \mathbb{N}\) and define

\[
a^{-n} = \frac{1}{a^n}.
\]

(2.16)

Note that this is consistent with the earlier definition: \(a^{-1} = \frac{1}{a} = \frac{1}{a}\). Now we make some observations regarding positive and negative exponents.

• \(\frac{1}{a^{-n}} = a^n\). We leave the proof as an exercise. Note that together with (2.16), this shows that \(a^n\) and \(a^{-n}\) are reciprocals of each other. Next we have, if \(n \neq m\) (and for \(n = m\) later), that

\[
a^n/a^m = a^{n-m}.\]

If \(n > m\), then \(m\) factors will cancel in the numerator and denominator and leave \(n - m\) factors in the numerator (and only a factor of 1 left in the denominator). On the other hand, if \(m > n\), then \(n\) factors will cancel, leaving \(m - n > 0\) factors in the denominator and 1 in the numerator, i.e., \(1/a^{n-m} = a^{m-n}\), albeit a negative power of \(a\). Notice that this is formally consistent with the earlier observation—with positive exponents—that \(a^n a^m = a^{n+m}\).

Now we have definitions of \(a^n\) for \(n \in \mathbb{Z} - \{0\}\), with negative \(n\) for \(a > 0\). It is natural to ask next if we can make sense of \(a^0\), at least for positive \(a\). Since \(a^0 a^{-n} = \frac{a^n}{a^n} = 1\), it is natural to define:

• \(a^0 = 1\). Another way to see the reasonableness of this is to note that \(a^n/a^n = 1\), but from our earlier equations we would like \(a^n/a^n = a^{n-n} = a^0\), so this is consistent with the earlier equations, at least for \(a \neq 0\).

It will at times be be useful to make some very simple observations regarding powers of \((-1)\). Notice that \((-1)^1 = -1, (-1)^2 = 1, (-1)^3 = -1, (-1)^4 = 1\), etc. So even powers of \((-1)\) yield 1, while odd powers yield \(-1\). This also has implications for powers of \(-a\):

• \((-1)^{2n} = 1\).

• \((-1)^{2n-1} = -1\).

• \((-a)^{2n} = a^{2n}\).

• \((-a)^{2n-1} = -a^{2n-1}\). The last term is shorthand for \(-a^{2n-1}\). To prove the third one, for instance, we can note that \((-a)^{2n} = (-1 \cdot 1)^{2n} = (-1)^{2n} \cdot a^{2n} = 1 \cdot a^{2n} = a^{an}.
It is very important that students do not confuse contexts, in the sense that each of these exponential rules are about counting how many copies of a factor are multiplied or divided; this is a strictly multiplication and division context. If we mix addition and subtraction into this the computations become more complicated. For instance, it was already shown that \((a + b)^2 = a^2 + 2ab + b^2\), though we wrote it \((a + b)(a + b) = aa + 2ab + bb\) at the time. It is important to note, for instance, that (unless \(a = 0\) or \(b = 0\)) we will have \((a + b)^2 = a^2 + b^2\).

Though we have far from checked all the possible complications (some of which are left as exercises), it is not difficult to see that the following are consistent, as long as we do not allow for any division by, or negative powers of, zero:

\[
\begin{align*}
a^n a^m &= a^{n+m} & (2.17) \\
(ab)^n &= a^n b^n & (2.20)
\end{align*}
\]

\[
\begin{align*}
a^n &= a^{n-m} & (2.18) \\
a^{-n} &= \frac{1}{a^n} & (2.21)
\end{align*}
\]

\[
\begin{align*}
(a^n)^m &= a^{nm} & (2.19) \\
\left(\frac{a}{b}\right)^n &= \frac{a^n}{b^n} & (2.22)
\end{align*}
\]

\[
(-1)^{2n} = 1 & (2.24) \\
(-1)^{2n-1} = -1 & (2.25)
\]

2.2.5 Miscellaneous Manipulations

- \((a + b)^2 = a^2 + 2ab + b^2\). This was stated earlier, without exponential notation.\(^{20}\)

- \((a - b)^2 = a^2 - 2ab + b^2\). This can be proved by “multiplying out” the left hand side \((a - b)(a - b)\) using the distributive axiom, as in

\[
\begin{align*}
(a - b)(a - b) &= (a + (-b))(a + (-b)) = aa + a(-b) + (-b)a + (-b)(-b) \\
&= a^2 - ab - ab + b^2 = a^2 - 2ab + b^2,
\end{align*}
\]

or more elegantly by replacing \(b\) with \(-b\) in the first equation to give us the second:

\[
(a + (-b))^2 = a^2 + 2a(-b) + (-b)^2 = a^2 - 2ab + b^2.
\]

Note how the part of “\(b\)” in the first equation for \((a + b)^2\) is played by “\(-b\)” in this one. The next two are similarly related:

- \((a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3\).

- \((a - b)^3 = a^3 - 3a^2b + 3ab^2 - b^3\).

\(^{20}\)This we proved earlier, though students are encouraged to repeatedly prove this by writing \((a + b)(a + b)\) and performing the multiplication, yielding four terms, two of which are the same. This “formula” can be memorized, but it is easily enough rediscovered that it is arguably better to memorize it through computing it from first principles until the outcome can be easily anticipated, i.e., until the student can “see a few steps ahead.”
We will prove the first one, and leave the second for the exercises.

\[(a + b)^3 = (a + b)(a^2 + 2ab + b^2)\]
\[= a(a^2 + 2ab + b^2) + b(a^2 + 2ab + b^2)\]
\[= a^3 + 2a^2b + ab^2 + a^2b + 2ab^2 + b^3\]
\[= a^3 + 2a^2b + a^2b + ab^2 + 2ab^2 + b^3\]
\[= a^3 + (2 + 1)a^2b + (1 + 2)ab^2 + b^3\]
\[= a^3 + 3a^2b + 3ab^2 + b^3.\]

Of course one learns to consolidate steps in such a calculation. The next equations establish a pattern. These can be verified by multiplication.\(^{21}\)

- \((a - b)(a + b) = a^2 - b^2.\)
- \((a - b)(a^2 + ab + b^2) = a^3 - b^3.\)
- \((a - b)(a^3 + 2ab + b^2) = a^4 - b^4.\)
- \((a - b)(a^4 + 3a^3b + a^2b^2 + ab^3 + b^4) = a^5 - b^5.\)
- \((a - b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \cdots + ab^{n-2} + b^{n-1}) = a^n - b^n.\)

The first is how we factor a “difference of two squares,” the second a “difference of two cubes,” and so on. Actually, when we have an even power \(a^{2n} - b^{2n}\), we have a difference of two squares and can write \((a^n - b^n)(a^n + b^n)\) and proceed from there. For instance,

\[a^4 - b^4 = (a^2 - b^2)(a^2 + b^2) = (a - b)(a + b)(a^2 + b^2).\]

Note that for the odd exponents we can replace \(b\) with \(-b\) to get, for instance,

- \((a + b)(a^2 - ab + b^2) = a^3 + b^3.\)
- \((a + b)(a^4 - a^3b + a^2b^2 - ab^3 + b^4) = a^5 + b^5.\)

2.2.6 Order of Operations

Of this topic it is assumed the reader is already aware, but it bears repeating that there is an assumed order, or hierarchy of computations when we parse or evaluate a complicated expression. It is important that there be no ambiguity of meaning when we write, for instance, \(-5^2\). A novice may wonder if we are to raise \(-5\) the the second power, and thus get \((-5)^2 = 25\), or if we are to raise 5 to the second power and then negate the result, i.e., ultimately return \(-5^2 = -25\). One familiar with these computations will know that the latter is the conventional understanding, i.e., that \(-5^2 = -(5^2) = -25.\)

One of the most useful types of notation developed in the history of mathematics (and already used in this text) is that of the grouping symbols. These are usually parentheses (.), and less commonly brackets ([,]), and least commonly curly braces (,) which are also used for sets. Grouping symbols allow us to understand that an expression, which can be itself quite complicated, can be treated as a single quantity in some part of the computation. For instance, it is understood when we write \(5(3 - 7)\), that we mean 5 multiplies whatever is the value of \(3 - 7\). We will have a method for actually deriving, rather than verifying, these and similar equations later using polynomial long division and other methods.
3 − 7. Thus 5(3 − 7) = 5(−4) = −20. If one were to input “5 times 3 minus 7” into a calculator (without parentheses), it would return instead (5 · 3) − 7 = 15 − 7 = 8.

Things get more interesting when we consider 15 − 3 · 2. This is understood to be 15 − (3 · 2) = 15 − 6 = 9, as multiplication takes precedence over addition and subtraction (unless overridden by parentheses). Students should be aware that the least sophisticated calculators will often parse this incorrectly, as (15 − 3) · 2 = 12 · 2 = 24.

It is perfectly acceptable to have “nested parentheses,” or nested grouping symbols in general. Most authors use parentheses first, then brackets, or sometimes larger parentheses. Examples may look like

\[
3[(4 − 8)(6 − 10)] − 5[(3^2 + 9)(5/6)] = 3[(-4)(-4)] − 5[(18)(5/6)]
\]

\[
= (3 · 16) − (5 · 15)
\]

\[
= 48 − 75
\]

\[
= −27.
\]

Returning to our previous example of −5², we note that after explicit grouping, the next priority is given to exponents. Thus

\[
−5^2 = −(5^2) = −25.
\]

After groupings and exponents comes multiplication and division, in the order they appear from left to right. After all of these comes addition and subtraction, also in the order they appear. (It is often the case that not all such structures are simultaneously present, but the order of precedence remains the same.) Thus,

• \( 3 · 5 − 4 = (3 · 5) − 4 = 15 − 4 = 11, \) while
  \( 3(5 − 4) = 3 · 1 = 3, \)

• \( 4 · 6 − 5 · 3 = (4 · 6) − (5 · 3) = 24 − 15 = 9, \) while
  \( 4 · (6 − 5) = 4 · 1 = 4, \)

• \( 8 + 4/2 = 8 + (4/2) = 8 + 2 = 10, \) while
  \( (8 + 4)/2 = 12/2 = 6, \)

• \( 2 + 3^2 = 2 + 9 = 11, \) while
  \( (2 + 3)^2 = 5^2 = 25, \)

• \( 3^2 · 2^3 · 3 − 7 · 9 = [(9 · 8) · 3] − (7 · 9) = (72 · 3) − 63 = 216 − 63 = 153. \)

It is unusual, except for explanatory purposes, to insert parentheses in such cases, so the last case above would normally be written

\[
3^2 · 2^3 · 3 − 7 · 9 = 9 · 8 · 3 − 63 = 72 · 3 − 63 = 216 − 63 = 153.
\]

However, parentheses can help clarify a complicated expression.\(^{22}\)

\(^{22}\)It is not altogether uncommon for unneeded grouping symbols to be introduced, since one’s eyes can use them as visual cues for partitioning the various parts of a complicated expression into parts which are treated temporarily as single objects. We will see this occur often in our chapter on differentiation, namely Chapter 4.
2.2. ARITHMETIC WITH REAL NUMBERS

2.2.7 Absolute Value, Size, and Distance on the Real Line

We define the absolute value of \( x \in \mathbb{R} \) as follows:

\[
|x| = \begin{cases} 
  x & \text{if } x \geq 0, \\
  -x & \text{if } x < 0.
\end{cases}
\]  \hfill (2.26)

Thus \(|5| = 5\), and \(|-7| = -(−7) = 7\). Note that in all cases we have

\[
|x| \geq 0,
\]  \hfill (2.27)

since \( x \geq 0 \implies |x| = x \geq 0 \), while \( x < 0 \implies |x| = -x > 0 \). However, we must be careful to note that the absolute value does not mean that we ignore any “negative” signs: \(|x| = -x\) is a possibility, occurring if and only if \( x \leq 0 \). (See footnote on why we have “\( \leq 0\)” instead of “\(< 0\)” here.) Next note that \(|a - b|\) is the distance between \( a \) and \( b \) on the real number line:

\[
|a - b| = \begin{cases} 
  a - b & \text{if } a - b \geq 0 \\
  -(a - b) & \text{if } a - b < 0
\end{cases} = \begin{cases} 
  a - b & \text{if } a \geq b \\
  b - a & \text{if } a < b.
\end{cases}
\]

Indeed that is exactly what we expect to be the distance between \( a \) and \( b \): we subtract the lesser from the greater (right point minus left point). This is sometimes given as a definition of distance:

\[
(\text{distance between } a \text{ and } b) = |a - b|.
\]  \hfill (2.28)

Note that \(|a - b| = |b - a|\), as we would expect. Furthermore, note that \(|x| = |x - 0|\), which is the distance between \( x \) and 0.

Later in the text we will often refer to how large a number is, meaning its absolute size (absolute value), so \(-10,000\) is a “larger” number than 2, even though \(-10,000 < 2\).

For a few examples of this distance relationship, consider

\[
|5 - 3| = |2| = 2 \quad \text{(the distance between 5, \(-3\)),} \\
|-3 - 5| = |-8| = 8 \quad \text{(the distance between \(-5\), 5),} \\
|2 + 3| = |5| = 5 \quad \text{(the distance between 2 and \(-3\)).}
\]

These we calculated from the definition of absolute value. However, note that \(|5 - 3|\) is the distance between 5 and 3 on the line, while \(|-3 - 5|\) is the distance between \(-3\) and 5, and \(|2 + 3| = |2 - (−3)|\) is the distance between 2 and \(-3\) on the line.

**Example 2.2.1** Next consider the set \( S = \{ x \in \mathbb{R} \mid |x - 3| < 5 \} \).

*This is all the points \( x \) which are less than 5 units away from 3, since \(|x - 3| < 5\) can be read “the distance between \( x \) and 3 is less than 5.”* Note that as an interval, \( S = (-2, 8) \):

\[\begin{array}{c}
\hline
-10 & -8 & -6 & -4 & -2 & 0 & 2 & 4 & 6 & 8 & 10 \\
\hline
\end{array}\]

\[\begin{array}{cc}
\bullet & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ
\end{array}\]

\[\begin{array}{c}
\begin{array}{c}
5 \\
3 \\
5
\end{array}
\end{array}\]

\[\begin{array}{c}
\begin{array}{c}
\text{Example 2.2.1: Next consider the set } S = \{ x \in \mathbb{R} \mid |x - 3| < 5 \}.
\end{array}
\end{array}\]

\[\begin{array}{c}
\begin{array}{c}
\text{This is all the points } x \text{ which are less than 5 units away from 3, since } |x - 3| < 5 \text{ can be read “the distance between } x \text{ and 3 is less than 5.” Note that as an interval, } S = (-2, 8):\end{array}
\end{array}\]

Note that if \( x = 0 \), either definition for \(|x|\) works, so we could also define

\[
|x| = \begin{cases} 
  x & \text{if } x > 0, \\
  -x & \text{if } x \leq 0.
\end{cases}
\]

23Note that if \( x = 0 \), either definition for \(|x|\) works, so we could also define
It is useful to then notice the following equivalences, which are easy to see if we refer to the number line. These are also true if we replace \(<\) with \(\leq\) everywhere, and \(>\) with \(\geq\) everywhere.

\[
|x| < K \iff (x < K) \land (x > -K) \iff -K < x < K, \tag{2.29}
\]

\[
|x| > K \iff (x > K) \lor (x < -K). \tag{2.30}
\]

These are easy to see if \(K > 0\), but are also true (the first one vacuously, and the second trivially) if \(K < 0\). Using these principles, we could have read Example 2.2.1 as follows:

\[
|x - 3| < 5 \iff -5 < x - 3 < 5 \iff -5 + 3 < x - 3 + 3 < 5 + 3 \iff -2 < x < 8.
\]

Thus \(x \in S \iff -2 < x < 8 \iff x \in (-2, 8)\), as before. Thus \(S = (-2, 8)\), by the definition of set equivalence.

**Example 2.2.2** Describe as an interval, and graph the set \(S = \{ x \in \mathbb{R} \mid |x + 4| \leq 3 \}\).

**Solution**: We first use (2.29) to rewrite this

\[
x \in S \iff |x + 4| \leq 3 \iff -3 \leq x + 4 \leq 3 \iff -7 \leq x \leq 1
\]

\(\iff x \in [-7, 1]\).

Hence \(S = [-7, -1]\). On the other hand, we can also rewrite \(|x + 4| \leq 3 \iff |x - (-4)| \leq 3\), and so \(S\) is the set of all points less than or equal to 3 units distance from \(-4\):

Clever manipulations of expressions of the form \(|ax + b|\) can make them appear to be describing a distance, or a multiple of a distance. As we will see below, in particular we can write (for \(a \neq 0\)):

\[
|ax + b| = |a| \cdot \left| x - \left( -\frac{b}{a} \right) \right| = |a| \cdot \left( \text{the distance from } x \text{ to } -\frac{b}{a} \right). \tag{2.31}
\]

One should not be expected to memorize (2.31) but rather to derive the relevant version as it is yielded naturally from manipulations allowed by facts given below.

Indeed we should notice some properties of absolute values which are clear if we realize that, when casually multiplying or dividing positive and negative numbers, we can do so with the absolute values of the numbers (i.e., ignoring the signs) until the end, where we account for the signs. That is the gradeschool approach. Thus \(7 \cdot (-5) = -35\) because \(7 \cdot 5 = 35\), but we had an “extra” negative sign. That naive approach is perhaps the best approach when noticing the following (where instead we simply “make sure” the results are not negative):

\[
|ab| = |a| \cdot |b|, \tag{2.32}
\]

\[
\left| \frac{a}{b} \right| = \frac{|a|}{|b|}, \tag{2.33}
\]

\[
|a^n| = |a|^n, \text{ assuming } n \in \mathbb{Z}. \tag{2.34}
\]

It is by virtue of (2.32), perhaps read backwards, that we can derive (2.31).
Example 2.2.3 Consider the inequality $|2x + 5| \geq 4$. Find all $x$ satisfying this by

1. using either (2.29) or (2.30) as your starting point, and (separately)

2. using a distance argument.

Solution:

1. We will use the equivalence (2.30), page 96:

$$|2x + 5| \geq 4 \iff (2x + 5 \geq 4) \lor (2x + 5 \leq -4)$$
$$\iff (2x \geq -1) \lor (2x \leq -9)$$
$$\iff (x \geq -1/2) \lor (x \leq -9/2)$$
$$\iff x \in (-\infty, -9/2) \cup [-1/2, \infty).$$

2. Here we will use (2.32), page 96 and some cleverness to rewrite the inequality so that it resembles a statement about distances:

$$|2x + 5| \geq 4 \iff |2| \cdot |x - (-5/2)| \geq 4 \iff |x - (-5/2)| \geq 2.$$

This is satisfied by all points which are at least 2 units away from $-5/2$.

Either way, the solution is graphed below:

Absolute value inequalities can become quite complicated, but we are mostly interested in those which convey size or distance restrictions. An example of a size restriction is the following:

Example 2.2.4 Consider the inequality $|x| \leq 100$. Since this one is simpler than the previous examples, we can look at it a third way: as an expression regarding how large $x$ can be.

1. Interpreting $|x| \leq 100$ as a statement about “size.”

Clearly we do not want $x$ to be greater than 100, or it would violate $|x| \leq 100$. On the other hand, we do not want $x$ to be “more negative” than (that is, to the left of) $-100$ either. With a little thought we can see easily that it is the following set which satisfies $|x| \leq 100$.

From this we can write $x \in [-100, 100]$.

2. Interpreting $|x| \leq 100$ as a statement about distance.

We can do this by rewriting the inequality $|x - 0| \leq 100$, meaning $x$ is at most 100 units from zero. This again gives us $x \in [-100, 100]$ and our graph above.
3. Interpreting $|x| \leq 100$ using an equivalent statement, namely (2.29), page 96 but using $\leq$, giving us $|x| \leq 100 \iff -100 \leq x \leq 100$, which reads clearly on its own but which can also be rewritten $x \in [-100, 100]$.

For such a simple inequality $|x| \leq M$ one usually interprets it in the “size” context; an inequality of the form $|x-a| \leq d$ can easily be seen in the context of distances; an absolute value inequality of the form $|ax+b| \leq k$ may be rewritten as a distance problem, though it is also quite common (perhaps more so) to use an equivalent statement $-k \leq ax+b \leq k$ from which one solves for $x$.

Absolute value inequalities are also useful in the context of tolerances. Suppose we wish to construct a meter stick and have the length be accurate to within 0.0005 meters. When reporting indirectly expresses an acceptable range of values.

One way of reading the phenomenon of (2.35) is that within the absolute values expression that it is within 0.0005 of 1 (suppressing the units m for now):

\[ |a+b| \leq |a| + |b|. \quad (2.35) \]

Note that this last property gives the “=” case if $a$ and $b$ are the same “sign,” or if either of them is zero, while giving the $<$ case when $a$ and $b$ have opposite signs. It is called the triangle inequality for reasons that are more clear when dealing with vector arithmetic, which occurs in higher dimensions (whereas the real line $\mathbb{R}$ is one-dimensional). The reader can verify the reasonableness of this by considering various cases, such as the following.

- $|2+5| = |7| = 7$, while $|2| + |5| = 2 + 5 = 7$ (the same).
- $|-3+(-6)| = |-9| = 9$, while $|-3| + |-6| = 3 + 6 = 9$ (the same).
- $|-7+4| = |-3| = 3$, while $|-7| + |4| = 7 + 4 = 11$ (different).

One way of reading the phenomenon of (2.35) is that within the absolute values $|a+b|$ there is a possibility of full or partial “cancellation,” say, if $a, b$ are of opposite signs, while there is no such possibility in the expression $|a| + |b|$, because $|a|$ and $|b|$ cannot have opposite signs.

There are similar inequalities, some generalizing the triangle inequality and some which are again intuitive when we think of which cases allow (or force) full or partial cancellation. Consider for instance the following:

\[ |a+b+c| \leq |a| + |b| + |c|, \quad |a-b| \leq |a| + |b|, \quad ||a|-|b|| \leq |a-b|. \]

Note that the second case can also be written $|a-b| = |a+(-b)| \leq |a| + |-b| = |a| + |b|$, and so ultimately we do have $|a-b| \leq |a| + |b|$.

Note that the “±” in the quadratic formula means we have exactly two cases, one for “+” and “−”:

\[(a \neq 0) \land (ax^2 + bx + c = 0) \iff x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.\]

In contrast is the context of tolerance, where “±” precedes the allowable error from the target value, and thus indirectly expresses an acceptable range of values.
2.2.8 Binomial Expansion

The binomial expansion gives us a shortcut for writing out \((a + b)^n\). It would be somewhat of a distraction to prove the expansion, so its proof is left for the appendix. The expansion is given by the Binomial Theorem given below:

**Theorem 2.2.1 (Binomial Theorem):** Given \((a + b)^n\), where \(n \in \mathbb{N}\), the following holds\(^{25}\):

\[
(a + b)^n = a^n + \frac{n \cdot a^{n-1}b}{1 \cdot 2} + \frac{n(n-1) \cdot a^{n-2}b^2}{1 \cdot 2 \cdot 3} + \cdots + \frac{n \cdots 2 \cdot 1 \cdot b^n}{1 \cdot 2 \cdots (n-1)}.
\]

(2.36)

Notice that the powers are all \(a^k b^{n-k}\) (if we assume \(a^0\) is 1--of course we have to be careful, but the formula works fine with that assumption no matter what \(a, b\) we are given). Thus the powers of \(a\) and \(b\) always add to \(n\).

Notice also that we should get \((a + b)^n = (b + a)^n\). In other words, \(a\) and \(b\) are interchangeable in (2.36). Clearly the \(a^n\) and \(b^n\) terms are both multiplied by 1, and the \(a^{n-1}b\) and \(ab^{n-1}\) both have coefficients \(n\). As we work toward the middle of the left hand side of (2.36) we see other terms \(a^k b^{n-k}\), \(a^{n-k} b^k\) are multiplied by the same constant factors.

**Example 2.2.5** We now see a few expansions based on the binomial theorem.

\[
(a + b)^4 = a^4 + \frac{4}{1}a^3 b + \frac{4 \cdot 3}{1 \cdot 2}a^2 b^2 + \frac{4 \cdot 3 \cdot 2}{1 \cdot 2 \cdot 3}ab^3 + \frac{4 \cdot 3 \cdot 2 \cdot 1}{1 \cdot 2 \cdot 3 \cdot 4}b^4
\]

\[= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4.\]

\[
(a - b)^5 = (a + (-b))^5
\]

\[= a^5 + \frac{5}{1}a^4(-b) + \frac{5 \cdot 4}{1 \cdot 2}a^3(-b)^2 + \frac{5 \cdot 4 \cdot 3}{1 \cdot 2 \cdot 3}a^2(-b)^3 + \frac{5 \cdot 4 \cdot 3 \cdot 2}{1 \cdot 2 \cdot 3 \cdot 4}a(-b)^4 + \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}(-b)^5
\]

\[= a^5 - 5a^4b + 10a^3b^2 - 10a^2b^3 + 5ab^4 - b^5.\]

\[
(3x - 2y)^3 = (3x)^3 + \frac{3}{2}(3x)^2(-2y) + \frac{3 \cdot 2}{1 \cdot 2}(3x)(-2y)^2 + \frac{3 \cdot 2 \cdot 1}{1 \cdot 2 \cdot 3}(-2y)^3
\]

\[= 27x^3 - 36xy^2 - 8y^3,
\]

\[
(a^2 - b^3)^3 = (a^2)^3 + \frac{3}{2}(a^2)^2(-b^3) + \frac{3 \cdot 2}{1 \cdot 2}(a^2)(-b^3)^2 + \frac{3 \cdot 2 \cdot 1}{1 \cdot 2 \cdot 3}(-b^3)^3
\]

\[= a^8 - 4a^6b^3 + 6a^4b^6 - 4a^2b^9 + b^{12}.
\]

\(^{25}\)Of course this may terminate after only two terms if \(n = 1\), or three terms if \(n = 2\). The pattern listed here is for general \(n \in \mathbb{N}\).
Figure 2.4: Pascal’s Triangle. Except for the first and last 1 in each row, every entry is the sum of the two closest entries above. One application is the coefficients of the binomial expansion (2.36).

Note that if our $a$ and $b$ terms in the expansion of $(a + b)^n$ are replaced by different powers of variables, then we cannot necessarily expect the powers on each term to sum to the same number. The last example above, for instance, had terms with powers $a^8, a^6b^3, a^4b^6$ and so on, where the sums of the powers were, respectively, 8, 9, 10 and so on.

For the simpler case $(a + b)^n$ (from which we can derive the others) there is an interesting relationship between the numbers multiplying the terms $a^k b^{n-k}$ in (2.36), i.e., the binomial coefficients, and a mathematical curiosity known as Pascal’s Triangle, shown in Figure 2.4. The entries of the “triangle” are computed by summing the two entries directly above to the right and above left. The rows of Pascal’s Triangle correspond to the coefficients of $(a + b)^0$, $(a + b)^1$, $(a + b)^2$, and so on.

**Example 2.2.6** We can use Pascal’s Triangle to compute $(a + b)^7$. We know that the first term will be $a^7$, and the second will be $7a^6b$, so we find the row that begins with 1 and 7 and compute:

$$(a + b)^7 = a^7 + 7a^6b + 21a^5b^2 + 35a^4b^3 + 35a^3b^4 + 21a^2b^5 + 7ab^6 + b^7.$$
2.3 Arithmetic of Units; Significant Digits

A unit is a well-defined quantity of some type, from which quantities of the same type can be measured. For instance, once we define what we mean by a particular unit of length, such as a meter, we can measure other lengths in terms of (or colloquially “in units of”) meters, so we can determine a 2-meter length, half-meter length, 5.27-meter length, and so on.

This section describes how units fit into computations, and can be manipulated. Units of time may be seconds, minutes, hours, rotations of the earth, revolutions of the earth around the sun, periodic radiations of a Cesium-133 atom, etc. Units of length include feet, meters, miles, light-years, and so on. There are also electrical units, such as volts, ohms, coulombs, amps, watts, farads and henries (all representing very different types of quantities and not just different units for the same type of quantity), and magnetic units such as teslas. Angles can be measured in degrees, radians, and revolutions. As we encounter applied mathematical problems, other units become important, such as volumes (liters, gallons, cubic feet, pints, etc.) or even money ($ or US$, NT$, £, Y, P, etc.).

Another common name for units is dimensions, which we will see makes some sense, particularly when we look at the relationships among measurements of length, area and volume. However, the student should keep in mind that the terms units and dimensions both have multiple meanings, which are perhaps more loosely related to each other than we would prefer. The terms are used interchangeably in our context here.

The mathematical concepts and manipulations developed here are soon well-known to any first-year student of physics or chemistry. It is somewhat a pity that students of mathematics are often not so well versed in these same skills.

2.3.1 Adding Units

We learn rather early on to add units, such as 2 ft + 3 ft = 5 ft, as this is intuitive. Of course there is a geometric meaning here, where a length of 2 feet is placed on end where the (parallel) 3-foot length begins, or vice-versa, so the total length is 5 feet. However, note that we can, at least formally, use the distributive property to write

\[ 2 \text{ ft} + 3 \text{ ft} = (2 + 3)\text{ ft} = 5 \text{ ft}. \]

So the usual algebra rules applied, at least in this case.

Now it is also possible to add lengths with different dimensions, but the arithmetic is not so simple. For instance, one could certainly add a 2-foot length to an 18-inch length, and get 2 ft + 18 in, but we can not combine these arithmetically if the units are different. The usual approach is to then convert to one length or another, for instance using the fact that one foot is twelve inches:

\[ 2 \text{ ft} + 18 \text{ in} = 24 \text{ in} + 18 \text{ in} = 42 \text{ in}; \]
\[ 2 \text{ ft} + 18 \text{ in} = 2 \text{ ft} + 1.5 \text{ ft} = 3.5 \text{ ft}. \]

27By international standards, a second is defined to be 9,192,631,770 periods of a certain radiation of a particular Cesium isotope, namely $^{133}\text{Ce}$ at absolute zero temperature. We can physically determine the temperature with complete precision, and count the vibrations of the isotope, making a second completely determinable in theory, without reference to other units.

28Note how units are often abbreviated. While there are standards one learns, conceptually the abbreviations are not important. One writer may write out “feet” or “foot” while another may just use “f,” but the overwhelming majority will use “ft,” which has several advantages, among them that there is no need to distinguish between singular and plural. Some conventions are less obvious, such as “lbs” for pounds (a unit of force), or “oz” for ounces (a unit used either for volume, force or mass, depending upon the context).
Both quantities we computed represent the same length, but in different units.\footnote{There are other ways one writes feet and inches, such as \(1.5\) \text{ft} = 1.5' = 18'' = 18 in. In some contexts this is then written 1'6'', meaning 1 ft + 6 in. That is not so easy to manipulate arithmetically. A similar notation is used for angle measure in degrees-minutes-seconds, and with time in hours-minutes-seconds, but we will rarely mix units like that. For instance, for time we will leave the answer in seconds, or in minutes, or in hours, etc., but not in a combination of these unless we are curious as to such a form for our answer. Certainly the arithmetic we set up will not mix these different units of the same type of quantity.}

\subsection*{2.3.2 Multiplying Units: Simplest Cases}

Shortly after one is secure adding units of length to each other, it is natural to begin to consider area and volume units built from these. For instance, in abstract geometry, we define the area of a rectangle to be the product of its length and width, i.e., the product of two adjacent sides’ lengths (which is called length, the other being the width, is entirely up to the individual, and varies from text to text). Thus a rectangle with length 3 and width 5 would have area \(3 \cdot 5 = 15\). But what of the units? If both are in feet, and we assume multiplication is always commutative (which it is in calculus, until you define some “vector” products in multivariable calculus, much later), and we have

\[3\text{ft} \cdot 5\text{ft} = 15\text{ft} \cdot \text{ft} = 15(\text{ft})^2 = 15\text{ft}^2.\]

Note that there are conventions that arise over the years. For instance, we are to assume “\text{ft}” is one indivisible concept, and so \(15(\text{ft})^2\) means \(15\text{ft}^2\), and certainly not \(15\text{f}^4\text{t}^2\), which is meaningless in this context. Other conventions arise for convenience in other fields, and usually the meaning is obvious from the context.

Note also that it is easy to imagine a “foot squared,” or a “square foot.” It is a unit of area equal to that of a square with dimensions one foot by one foot. Similarly we can discuss a volume of solid which is a right parallelepiped measuring 3 ft long, 4 ft wide and 5 ft tall, having volume \(3\text{ft} \cdot 4\text{ft} \cdot 5\text{ft} = 60\text{ft}^3\), where one cubic foot is the volume of a cube with each edge length being 1 ft.

Now consider the area of a rectangle which is 4 feet by 3 yards. We have basically three immediate options for computing this area: we could simply multiply these two quantities; we could convert feet into yards and have an answer in square yards; or we could convert the yards into feet and have an answer in square feet. Each answer will represent the same area, but in different units.

\[4\text{ft} \cdot 3\text{yd} = 12\text{ft-yd}\]

\[4\text{ft} \cdot 3\text{yd} = \frac{4}{3}\text{yd} \cdot 3\text{yd} = 4\text{yd}^2\]

\[4\text{ft} \cdot 3\text{yd} = 4\text{ft} \cdot 9\text{ft} = 36\text{ft}^2.\]

Note first that there is a convention when expressing a product of two different units, in which a short “dash” is used where we might expect to see a “dot.” The typsetters and writers are careful to be sure it is short enough that it can be distinguished from the longer subtraction mark (“−” versus “−”) , but it takes some getting used to. The first answer would be read aloud as, “twelve foot-yards,” where the other two would be read, respectively, “four yards square(d)” (or “four square yards”), and “thirty-six feet squared” or similar.

Now it is rare to see a unit such as a foot-yard, but it is simply the amount of area equal to a rectangular area which is one foot long and one yard wide, for instance. (Occasionally we
see such mixed units in commerce, where for instance one might buy an area of fabric or carpet which comes off of a roll with a fixed width, and the customer officially pays by the length of the cut piece.)

Another place where units are multiplied is in dealing with torques. If an arm of some length is attached at one end to a pivot about which it can rotate in a circle within a plane, and a force is applied at the other end of the arm in a direction tangent to the circle, then the torque experienced at that central pivot is given by the product of the length of the arm. If the force is in pounds (lbs), and the length of the arm is in feet, then the torque is in ft-lbs (foot-pounds). If we prefer metric units, the force could be in newtons (n) and the length in meters (m), and we would have torque in newton-meters, i.e., n-m.

Oddly enough, energy and work can be given in the same units as torque, though they are completely distinct from the concept of torque. Energy is the ability to do work, and the amount of energy an object has is defined by the work it can accomplish, so energy and work have the same units. Work is defined as the product of a force and the displacement (or signed distance) of the motion in the direction at which the force acts. This will be clarified more later, but for a simple example we can consider a young man pushing a broken down car along a road, using perhaps 100 lbs of force as he pushes the car 40 ft in the direction of motion. In this action the young man contributed \((40 \text{ ft}) \cdot (100 \text{ lbs})\), or 4000 ft-lbs of energy to the car. (Most likely the energy was dissipated by internal friction of the car, and the flexion of the car-road interface.) We will see applications of work and energy later in the text.

### 2.3.3 Division of Units: Rates

While multiplication of units is particularly common in the integral calculus (starting with Chapter 8), rates actually appear more prevalently in the earlier parts of the text, being particularly important in the study of applied differential calculus (beginning in Chapter 4). However, both concepts appear throughout the rest of the text.

For our first example of a rate, note that in the United States the speed (rate) of travel in vehicles is still commonly computed in miles per hour, or mph. In fact, this is mathematically represented as \(\text{mi/hr}\), i.e., miles divided by hours. Indeed, the word “per” (sometimes rather carelessly read “for every”) always indicates a division. The reason for this is clear if we look at applications in which we multiply these compound units by other units in obvious applications. For instance, suppose a car travels at exactly 60 mph for exactly 1.5 hr. Then the total travel distance is given by

\[
\left( \frac{60 \text{ mi}}{\text{hr}} \right) \cdot (1.5 \text{ hr}) = \frac{90 \text{ mi} \cdot \text{hr}}{\text{hr}} = 90 \text{ mi}.
\]

This is the well-known elementary school idea that “distance equals rate times time.” Note, though, how the units resolve correctly when we allow “per” to mean division by the unit which follows.

While in calculus we usually avoid the elementary school “times” notation, \(\times\) for multiplication, it is fairly common in the physical sciences, so the above computation might be written\(^{30}\)

\[
\frac{60 \text{ mi}}{\text{hr}} \times 1.5 \text{ hr} = 90 \text{ mi}.
\]

\(^{30}\)Of course, one has to distinguish the multiplication operation \(\times\) with the variable \(x\). While this is easy enough with typesetting, some care has to be taken when writing by hand, particularly for other readers. One device used by lecturers in the physical sciences is to be sure there is extra space around, and careful vertical centering of, the operator but not so much with the variable. Also, using a more cursive style with the variable, and straighter “45°” strokes with the operator is standard in writing these. It is up to the writer to be clear, and not always assume the reader will understand from the context.
A very similar computation shows how a motorcycle, which can travel 60 miles per gallon of gasoline, will travel 90 miles if given 1.5 gallons of gasoline.

\[
\frac{60 \text{ mi}}{\text{gal}} \times 1.5 \text{ gal} = 90 \text{ mi}.
\]

Note that both 60 mi/hr and 60 mi/gal are rates, the former being a rate of miles per every hour, and the latter being a rate of miles per gallon. Rates are arguably as important as any other concept in the development of the calculus, as we will see later in the text.

Most physics textbooks prefer to use metric units, and so velocities in such texts tend to be given in units of meters per second, or m/sec (sometimes simply written m/s). Less often, velocities are given in feet per second, ft/sec, though other units are used when they seem illustrative.

Now consider an object which is dropped from a tall cliff (precipice), and encounters only negligible air resistance. Initially the object will have velocity zero, but after one second the earth’s gravity will have the object moving 32 ft/sec, and then subsequent velocities will be 64 ft/sec, 96 ft/sec, 128 ft/sec, etc., after seconds 2, 3 and 4, respectively, until the object impacts the ground or another object. Note that the velocity of the object increases by 32 ft/sec every second; the object is accelerated by 32 feet per second per second. When we carefully write this rate of acceleration, including the resultant units, we get

\[
\frac{32 \text{ ft/sec}}{\text{sec}} = \frac{32 \text{ ft}}{\text{sec}} \times \frac{1}{\text{sec}} = 32 \text{ ft sec}^2.
\]

If writing inline, one would write 32 ft/sec\(^2\). Phrases to verbalize this include “32 feet per second squared,” and “32 feet per second per second.”

Unlike previous squared units, it is not clear how we are to “visualize” what a second squared would be. In fact, it is perhaps best not to try, but instead give it meaning algebraically, and (perhaps separately) in context. Indeed, here it signifies that we are considering “a rate of (change of) a rate (of change),” in this case the rate in which a velocity (in ft/sec) is changing by the second. In that context it eventually makes sense, and certainly the algebra forces us to accept these units.\(^{31}\)

The possibilities for types of rates we can imagine are limitless. For instance, $/\text{gal}$ of some fluid, rounds/min for a machine gun, $/\text{student}$ costs, gal/mile for fuel efficiency in a tanker ship (as opposed to mile/gal, which is just the reciprocal), and so on.

### 2.3.4 Conversions

There are three insights we share here from our friends in the physical sciences.

- We can write 1 (one) in many different ways, some involving units.

\[^{31}\text{When dealing with the rate of energy energy produced or consumed, that rate being called power, and the unit being the watt, a careful look at its development shows that it can be broken down to more fundamental units, and eventually one gets}

\[
1 \text{ W} = 1 \text{ kg} \times \frac{\text{m}^2}{\text{sec}^3}.
\]

We see that it can become very confusing to attempt to visualize its significance from the fundamental units, which in part explains why we devise the (relatively simple to comprehend) unit of the watt. Here kg is kilogram, which is a unit of mass.

Other units of power include horsepower, calories/second, and BTU/hour, where BTU stands for “British thermal unit.” Any unit of energy per unit time is, technically, a unit of power. (A calorie is a unit of energy.) By the same logic, power times time would yield energy. For example, in the United States electricity from the electrical utilities is usually charged by the kilowatt-hour, i.e., kW-hr, which is a unit of energy.
2.3. ARITHMETIC OF UNITS; SIGNIFICANT DIGITS

For instance, 1 mi = 5280 ft, and so we can claim that
\[
1 = \frac{5280 \text{ ft}}{1 \text{ mi}}.
\]
This is because the numerator and denominator are, technically, the same (and nonzero), so the fraction must equal 1. Similarly, its reciprocal is also 1. By the same logic, 1 hr/60 min and 1000 m/1 km are both, ultimately, 1. Of course there are infinitely many other such possibilities.

- We can use these creative ways of writing 1 as multiplying factors to convert quantities expressed in one unit to another expression of the same quantity in another unit.

For instance, suppose we are riding in an aircraft which is reportedly at an altitude of 30,000 ft, and we wish to know this height in miles. This can be achieved by multiplying the altitude by 1—therefore not actually changing the length of the reported altitude—but choosing the expression for 1 strategically:
\[
30,000 \text{ ft} = 30,000 \text{ ft} \times \frac{1 \text{ mi}}{5280 \text{ ft}} \approx 5.68 \text{ mi}.
\]
In fact, a calculator will likely tell us that 30,000/5280 = 5.681818182, which is a ten-digit approximation of the exact answer, so we might guess that the exact answer is 5.681, i.e., 5.681818181818⋅⋅⋅, but to report this answer in an applied problem is, in fact, ridiculous upon reflection. After all, it is impossible—and probably meaningless—to discuss exact altitude, as the topography is not constant, and aircraft tend not to be so steady, so our actual height could not credibly be reported as exactly 30.0 ft. The recipient of the altitude report would likely be satisfied with an approximation such as 5.68 mi, or 5.7 mi, or even “approximately five and a half miles.” We will discuss this idea of “rounding” our answers in the subsection on significant digits below.

Returning to the idea of conversions, note that several conversions can be made at once. For instance, if an air gun is known to be able to launch its projectile at 400 ft/sec, and we are interested in what this would be in mi/hr, we can convert the ft and sec separately:
\[
400 \text{ ft} = 400 \text{ ft} \times \frac{1 \text{ mi}}{5280 \text{ ft}} \times \frac{60 \text{ sec}}{1 \text{ min}} \times \frac{60 \text{ min}}{1 \text{ hr}} \approx 270 \text{ mi/hr}.
\]
Again, a calculator would return 272.72, but it is intuitively unlikely that the velocity of the projectile is known so precisely, so we should not report our representation in mi/hr as though we know it with unlikely precision. (In future computations, we will simply use 1 hr=3600 sec.)

Note that the converted quantity’s unit must be of the same type as the original, in the sense that if the original represented a velocity, so must the converted quantity. For instance, we can not convert a length to an area, or a volume to a velocity. For this reason, sometimes we see employed a short-hand “dimensional analysis,” in which we write things such as velocity=length/time, acceleration=velocity/time=length/time², length=velocity × time.

- We can perform many conversions within formulas.

For instance, we know that a constant velocity allows us to use the gradeschool formula distance = rate × time. It is not in fact necessary that the units be compatible, though if we are to interpret the final result, the units should be simplified (which occurs automatically if they were “compatible” from the start). For instance, suppose we wish to travel 45 km/hr for 90 sec. Then the distance we travel will be
\[
\frac{45 \text{ km}}{\text{hr}} \times 90 \text{ sec} = 4050 \frac{\text{km-sec}}{\text{hr}}.
\]
The answer is completely correct, but unwieldy for interpretation, as we are led to wonder just what is a km-sec/hr. Instead we could have (1) converted seconds to hours from the start, (2) converted km/hr into km/sec from the start, or (3) perform the conversion after we multiplied our rate and time:

\[
\frac{45 \text{ km}}{\text{hr}} \times 90 \text{ sec} \times \frac{1 \text{ hr}}{3600 \text{ sec}} = 1.125 \text{ km}.
\]

Again we may feel more comfortable reporting the final answer as 1.13 km, 1.1 km or even 1 km, depending upon our confidence in the original velocity (and how constant it was for the 90 sec).

Note also that if we wished, we could have converted the final answer to express the distance in some other unit of length. Chemistry and Physics students especially learn this “conversion as we go” habit, beginning from the relevant formula, as it is often easier than looking ahead to the end calculation and attempting to convert the original, often disparate units we are given at the beginning in order to have complicated formulas output a simple answer. In fact, multiple early conversions require the original data to be stored or written down in their converted form, where converting the final answer, using the original numbers, can be done at once on the calculator, utilizing all of its internal precision.

### 2.3.5 Tolerances

In applied problems, it is important to communicate the level of precision to which a quantity is measured. Put another way, it is important to know the possible error, i.e., the difference between the measured or reported value and the true value. Of course if we knew the exact error, we would know the exact number we approximated, since we can use the equation

\[
\text{ERROR} = \text{ACTUAL} - \text{MEASURED}.
\]

In practice, we hope to find bounds on the error term. In fact we usually can only find a bound for the absolute error, which is

\[
\text{ABSOLUTE ERROR} = |\text{ERROR}| = |\text{ACTUAL} - \text{MEASURED}|.
\]

Thus,

\[
\text{ACTUAL} = \text{MEASURED} \pm \text{ABSOLUTE ERROR}.
\]

It is common for the term “error” to refer to absolute error.

Finding a bound for the error is an important part of science and engineering. Realizing that it is usually impossible to have an exact measurement of quantities whose possible values form a continuum, we content ourselves to finding good approximations and good bounds for the errors in those approximations.

There are two standard methods for communicating the (absolute) error in a measured quantity, namely (1) through tolerances, and (2) by the significant digits that are listed.

A tolerance is an upper bound on the (absolute) error in measuring a quantity. The term comes from the idea that we will “tolerate” an error of at most the given size. This is useful in design as well as measurement. For instance, one might give the designed gap between the top of a valve stem in a particular combustion engine, and the valve’s respective cam on the camshaft, as 0.003 in $\pm$ 0.0008 in, or $(0.003 \pm 0.0008)$ in. This means that the actual value of the gap’s length would be between 0.003 $-$ 0.0008 in and 0.003 $+$ 0.0008 in, putting the gap in the interval $[0.0022 \text{ in}, 0.0038 \text{ in}]$, meaning the engine would be running within design specifications for that gap if it is anywhere within that interval.

In electrical applications, components are often given tolerances as percentages, but it is the same idea. For instance, a resistor might be listed as 330 Ohms of resistance, usually written
330Ω, and if it is allowed a 10% tolerance, that is also a tolerance of 33Ω, so the actual value of the resistance should be within [297 Ω, 363 Ω].

Tolerances can be given in absolute terms, which describes the exact interval by its endpoints or its center point plus/minus the possible error, or tolerances can be given in relative terms, usually by percentages, where one often writes, for instance ±5% or claims a 5% tolerance. It is called a relative measure of error because 5% is considered relative to the size of the target value. Later we will define relative error to be the actual (absolute) error divided by the target value, but that is all for later.

2.3.6 Significant Digits

“Significant digits aren’t the only, and are not the best, way of keeping track of errors in experimental data, but they are often used in beginning science classes. Don’t get hung up on the rules for significant digits. They are only a crude tool for showing the approximate amount of error in a measurement or calculated result. Use the rules with good judgment, rather than following them rigidly at all times.”

—David Dice, Carlton

(This subsection is heavily footnoted because the exact understanding and best rules for significant digits are not uniform among physical scientists, and mathematically it is easy to see none of the usual guidelines given to students apply every time, but it is useful to be aware of the different meanings and methods as one encounters the discussions among these scientists.)

There is a more compact, if less precise way to communicate the error, and that is through significant digits. Unfortunately there are multiple interpretations of the meanings of significant digits. Fortunately there is agreement on how many significant digits a particular numerical expression of a quantity contains.

We will list a few interpretations, and in general assume the “intersection” of these, so that each interpretation will imply ours here, though ours will then be weaker. When knowing a more accurate or precise tolerance is necessary, one can resort to tolerances, but as we will see in Chapter 3 and beyond, that can be a difficult problem indeed. If we are unwilling or unable to find a precise tolerance, then we must at least not exaggerate our precision and accuracy.

When one writes a numerical representation of a measured or approximated quantity using any of the significant digits approaches, as we read the number’s written from left to right all but the last written digit are assumed to be correct, perhaps after the number is truncated or “rounded.” The interpretations differ on the meaning of that last written digit. A strict interpretation of significant digits would assume that the last written digit is correct to ±1; the weaker interpretation only assumes that the last digit is a “good guess,” not necessarily within 1 of the actual value if known to perfect precision. In both interpretations, we count the number of “significant digits” by counting the digits from the first nonzero digit on the left to the last nonzero digit on the right.

For instance, suppose we are asked to obtain approximately 150 g (150 grams) of some substance for a chemistry exercise. Further suppose that we are to write down our actual measurement of the mass we obtained for a future computation. Then the following table represents what we might report, and what is meant in terms of a strict interpretation of significant digits, followed by the weaker interpretation. (All quantities are in units of grams.)
Similarly, if we wrote some measurement as 123.456, we are claiming that the actual value is 123.456 ± .001. We would say that this represents six significant digits of precision. It has the disadvantage that the tolerance is always given as some power of 10 (such as 10⁻³ in this case), but many of the advantages of concision. For instance, perhaps we really know this number is correct to within ±.0008, but are satisfied to report it within ±.001. As we will see in Chapter 3, it is not always easy to get an exact tolerance, and we will often instead estimate the tolerance, being sure to never “under-estimate,” in the colloquial sense, the possible error.

There are some conventions for communicating the “significance” of digits. In our original example, note that 150 is assumed to have only two significant digits, the 1 and the 5. If there is a decimal written after the 0, it is also considered significant. It becomes confusing if one writes a number such as 150,000 because it is not clear how many (at least two) digits are significant. For that reason, the scientific notation is often used, where we write any nonzero number as \( n \times 10^m \), where \( 1 \leq n < 10 \) and \( m \in \{0, \pm 1, \pm 2, \pm 3, \cdots \} \). Here \( n \) is the it coefficient or mantissa (which unfortunately has other, very different meanings), and \( m \) the exponent or order of magnitude of the given number. For any nonzero number, it is not hard to show that there is only one way to express it in scientific notation exactly. However we are Thus we can write 150,000 as \( 1.5 \times 10^5 \) for two significant digits, and \( 1.50 \times 10^5 \) for three, and so on.

Note that we can also write 150,000 as 0,150,000 but the leading zero is not considered significant. Similarly, if we write 0.123, or 0.0103, or 0.000000120, we still have three significant digits, so the “leading zeroes” are not considered significant.

Science students are some loose guidelines for determining how many significant digits a computational result should contain, but they all have their exceptions so we point out these guidelines here but with the understanding that they are not rigid.

For instance, adding a number with three significant digits to another with only one significant digit can result in a number with one, two or three significant digits. Consider the following:

\[
123 + 20,000 \approx 20,000; \\
123 + 10 \approx 130; \\
123 + 1 \approx 124
\]

---

32In higher mathematics, the term “estimate” is used differently than it is in colloquial English. The colloquial meaning of “estimate” is instead expressed by professionals in the mathematical sciences as approximate, in both noun and verb forms. A mathematician using the term “estimate” means he or she is finding bounds for a quantity, either upper or lower bounds. For instance, if a quantity \( \xi \) is known to be less than 10, and a “good guess” is that it is close to 6, we have the estimate \( \xi < 10 \), and the approximation \( \xi \approx 6 \). The estimate is expressed as a truth that is provable, while the approximation is, in fact, subjective, but its voracity should be easily judged from the evidence given.

33In fact this is not quite true, as we can write, for instance, \( 1 = 0.9999999 \cdots \), and so \( 1 \times 10^0 = 9.99 \times 10^{-1} \), but this become moot in actual empirical sciences where measurements are never exact, and tolerances and thus significant digits become important. For practical purposes (involving tolerances), there is only one way to write a nonzero number in scientific notation. Since we never have infinitely many significant digits for a measured quantity, we will not find ourselves dealing with expressions like \( 9.99 \times 10^{-1} \) or, for that matter, \( 1.99 \times 10^0 \), though of course we have many times where we have, at least theoretically, the exact number 1.
The first line, if all things were exact, would yield 20,123, but we do not know our answer to such precision because the 20,000 term was only known to an accuracy ±10,000, as only the 2 was communicated to be significant. Similarly, when we added 123 to 10, we did not know the accuracy of the units digit in the 10, so we can not include it in the final answer as being significant. The final line is more likely to be considered at face value.

Often, if not usually, a textbook in the physical sciences would simply use = with the understanding that there is a tolerance expressed by the number of significant digits and the final digit’s placement. These textbooks usually do not use the approximation symbol ≈ (or its variations such as = or ≃) in these circumstances. However, this being a mathematics textbook, we will use the approximation when we finish with our (approximate) answer. In fact we will adopt the convention that we assume the input data is exact, i.e., solve the problem hypothetically, and then report the final answer to the number of significant digits actually warranted by the precision in the data, using ≈ in the final step. We will also not perform numerical, intermediate computations, but instead allow our calculators to assume the 10–12 digits of accuracy they are capable of, before truncating our final answer to a reasonable number of significant digits for the problem.

For instance, suppose we wish to calculate the area of a rectangle with length 6.25 cm and height 12.67 cm. The usual method for doing so is to allow the calculator to report the product as if these inputs are exact, and then do our rounding. Multiplication and division differ from addition and subtraction in that we usually leave our answer with minimum number of significant digits possessed by the factors. For this computation, we have

\[ 6.25 \text{ cm} \times 12.67 \text{ cm} = \frac{79.1875 \text{ cm}^2}{\text{calculator reports}} \approx 79.2 \text{ cm}^2. \]

In fact, if this were to be the final computation in our report, we might write simply

\[ 6.25 \text{ cm} \times 12.67 \text{ cm} \approx 79.2 \text{ cm}^2. \]

Since the first factor only had three significant digits, the final answer can not have more than three. However, we do not truncate the second factor first, because we would be needlessly discarding information about that second quantity if we replace it with 12.7, which assumes a tolerance of ±1, where the original, 12.67, assumes a tolerance of ±0.01, and is closer to the actual value than is 12.7. Indeed, if we round first, we get \(6.25 \times 12.7 = 79.375 \approx 79.4\), and our last digit differs from the previous computation by 2.

In general, it is better to assume the quantities are exact, and allow the calculator to use all of its digits of accuracy, rounding only its final output to reflect the uncertainty of the final result, that uncertainty being inherent in the original, usually measured, inputs.

---

34In fact, what we really have here in 123+1 is some \(x + y\), where \(x = 123 \pm 1\) and \(y = 1 \pm 1\), so we really have \(x + y = 123 + 1 \pm 2 = 124 \pm 2\), so the final digit can be argued to be not quite significant either. This is often glossed over in the literature, and we accept the answer 124 to be, more or less, having a tolerance of ±1 though, mathematically, it is more like ±2. There are several explanations, but in practice it is often left to the judgment of the one reporting the result, or the reader reviewing the computations.

There is even some disagreement regarding how to translate a number like 125 into a two significant digit truncation. Some argue to “round up” if the last digit is in \{5, 6, 7, 8, 9\} and “round down” if the last digit is in \{0, 1, 2, 3, 4\}, while some argue that 125 is half-way between 120 and 130 and there is no compelling reason to round up instead of down. Some will even use the odd or even nature of the preceding digit, in this case 2, to decide, reasoning that in a large pool of data this should the roundings up and down should be roughly equal in number.

In this textbook we will always round up if the last digit is 5 or greater, and down otherwise, and so 125 \(\approx 130\), if we wish to change the three significant digit number to a two-digit representation instead. This kind of truncation is often done in the last step of a series of computations, where the calculator outputs more digits than we can really call significant.
This can be a difficult lesson for students of physical sciences, because it is often easier to write out the intermediate computations in some rounded manner, but this introduces needless error. It is better to solve for the variable in question, input the numbers. This is true whether the operations involved are multiplication and division, addition and subtraction, trigonometric functions, exponential and logarithmic functions, or any combination of these or other functions. When intermediate results must be written down for future computations, it is better to write them in as many significant digits as possible, even if they are “reported” in a truncated manner.
2.4 Algebra

Algebra is a very wide field, but we will restrict ourselves to the familiar real-variable algebra, in which we have at least one unspecified quantity, say $x$, which we nonetheless attempt to say something about. Perhaps we know one equation the quantity is supposed to satisfy, and try to find others which are equivalent to, or implied by, the first equation. (Instead of an equation it could be an inequality we would like to derive other facts from.) In this section, we will delve into “solving” equations, i.e., finding all the values of $x$ for which a given equation is satisfied.

Our purpose here and in subsequent sections is not to simply review high school algebra. Instead we will look at the topics from the standpoint of our understanding of real numbers and logic. In fact, these sections are designed to slow down and make thoughtful the reasoning processes which too often come across to students as plays in a symbol manipulation game, with rules that seem as arbitrary and numerous as a tax code. The “rules” are not! They follow immediately from the properties of real numbers and logic. Indeed, we have already begun to build arithmetic on top of these, albeit keeping that discussion abstract so we could focus on the principles involved rather than special cases.

True, it is important to master the “mechanics” of algebra, since there are numerous strategies for solving the myriad of problems which present themselves, but it is equally important to understand the principles as they are invoked, and the limitations of those principles. Indeed, there are technicalities which the student must be aware of in order to avoid errors. Many of the “manipulations” can be done with seeming impunity as, for instance, $x$ is solved for. Nonetheless we should always be aware of the rules we are using, and whether or not some principle is being violated. It is a truism that most calculus mistakes students make are really algebra mistakes, so it is of utmost importance that we do not undermine our calculus with errors at the (more fundamental) level of algebra.

A general principle used in algebra is the following:

**Principle 2.4.1** Given an equality $x = y$ of real numbers, any algebraic operation which can be performed on one side can also be performed on the other, and equality will still hold. In other words, if $A(.)$ is some algebraic operation, and $A(x)$ is defined, then

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

Here $A(.)$ might represent adding a fixed number, multiplying or dividing by a fixed number, squaring what is inside $(.)$, taking the tangent $\tan(.)$, etc. Basically, we can do whatever we like algebraically\(^{35}\) to both sides of a true equation and still have a true equation, provided the new equation has meaning.\(^{36}\) However, we are only guaranteed implication and not logical

\(^{35}\)We cannot make typographical changes to both sides of an equation. For instance,

$$x + 3 = 4 + 9 \not\implies x - 3 = 4 - 9,$$

even though we made the same typographical change (switching the “+” before the last term to a “−”). We could write

$$x + 3 = 4 + 9 \implies -(x + 3) = -(4 + 9),$$
in which we negated (an algebraic operation) both sides.

This example may seem a bit ridiculous. For more subtle example compare the first line below (which is a very common mistake) with the second.

$$2 + 3 = 4 + 1 \implies 2^2 + 3^2 = 4^2 + 1^2 \quad \text{False}$$
$$2 + 3 = 4 + 1 \implies (2 + 3)^2 = (4 + 1)^2 \quad \text{True}.$$

\(^{36}\)We can not, for example, divide both sides by zero, or take the square root of both sides if the sides are negative, or take the reciprocals if a side is somewhere zero, etc.
equivalence for many operations $A$, and so some precision may be lost by applying $A$ to both sides. For example,

$$2x = 1 \implies 0 \cdot 2x = 0 \cdot 1 \implies 0 = 0,$$

which is true, but we can not read the solution $x = 1/2$ from the resulting statement. A better situation is when $A$ is “invertible,” which actually gives us logical equivalence.

**Definition 2.4.1** An operation $A$ is called invertible if and only if

$$A(x) = A(y) \implies x = y. \quad (2.37)$$

The other implication being obvious if $A$ is a legitimate operation, we can replace the definition with

$$A(x) = A(y) \iff x = y. \quad \text{or} \quad (2.38)$$

**Proof:** Assume $x, y, k \in \mathbb{R}$. Then

$$x = y \implies x + k = y + k \implies x + k + (-k) = y + k + (-k) \implies x = y.$$

Thus $x = y$ implies $x + k = y + k$, and vice versa. The second statement of the principle is exactly the first statement, where we substitute $-k$ for $k$.

What this principle says is that we can add any real number to both sides of an equation and get a logically equivalent statement, or subtract any real number from both sides, also without changing the logical truth value. A similar principle states that we can multiply or divide both sides by a nonzero real number and have a logically equivalent statement. In the following, the notation $\mathbb{R} - \{0\}$ represents the set of real numbers with the zero element deleted.

**Principle 2.4.3** For any $k \in \mathbb{R} - \{0\}$,

$$x = y \iff kx = ky \quad \text{or} \quad (2.40)$$

$$x = y \iff x/k = y/k. \quad (2.41)$$

The second statement is exactly the first, where we substitute $k^{-1}$ for $k$. The proof is very similar to the proof of the preceding principle. To see how these principles are used to solve equations, consider the following example:

**Example 2.4.1** Find $x$ such that $5x + 11 = 37$.

We will begin with the equation and find logically equivalent statements of increasing simplicity, until we can read the solution.

<table>
<thead>
<tr>
<th>Equation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$5x + 11 = 37$</td>
<td>original equation</td>
</tr>
<tr>
<td>$5x + 11 + (-11) = 37 + (-11)$</td>
<td>added $(-11)$ to both sides</td>
</tr>
<tr>
<td>$5x = 26$</td>
<td>simplified</td>
</tr>
<tr>
<td>$\frac{1}{5} \cdot 5x = \frac{1}{5} \cdot 26$</td>
<td>multiplied both sides by $\frac{1}{5}$</td>
</tr>
<tr>
<td>$1x = \frac{26}{5}$</td>
<td>simplified</td>
</tr>
<tr>
<td>$x = \frac{26}{5}$</td>
<td>done.</td>
</tr>
</tbody>
</table>
Here we borrowed the venerable BASIC computer language notation (') to signify a “comment” which explains what was done. A student well-versed in algebra would likely write only the first, third and last lines of this list. However, there is a constant need for vigilance. In the above we have logical equivalence at each step. We do not always have that luxury. For instance we have the following propositions:

**Principle 2.4.4** Given any real number \( x \),

\[
 x = k \quad \Rightarrow \quad x^2 = k^2. \tag{2.42}
\]

The converse “\( \Leftarrow \)” is not true, for \((-1)^2 = (1)^2\), but clearly \((-1) \neq 1\).

**Definition 2.4.2** For any nonnegative real number \( x \), define the square root of \( x \), written \( \sqrt{x} \), to be that nonnegative number \( k \) such that \( k^2 = x \). Thus \( k \) is the number satisfying the following statement:

\[
 \sqrt{x} = k \quad \iff \quad (k \geq 0) \land (k^2 = x). \tag{2.43}
\]

Notice first that the right hand side is automatically false if \( x < 0 \), because \( k^2 \) cannot be negative. Thus only nonnegative numbers \( x \) can possess square roots.\(^{37}\) Next, notice that when we discuss the square root of a number, we do not choose the “negative square root” of that number. In some contexts it is acceptable to say “\(-5\) is a square root of 25.” However, we would say “5 is the square root of 25,” and write \( 5 = \sqrt{25} \).

We will take for granted the existence of the square root \( \sqrt{x} \) of any \( x \geq 0 \) (as, for instance, the length of a side of a square with area \( x \)), and discuss the properties of this operation \( \sqrt{\cdot} \) as we need them. We will return to the topic of square roots after an interlude into polynomial equations.

### 2.4.1 Polynomial Equations: Using Theorem 2.1.8

Using different letters, we can write the theorem on zero divisors as

\[
 A \cdot B = 0 \quad \iff \quad (A = 0) \lor (B = 0).
\]

This can be quite useful for solving polynomial equations if we employ a small amount of cleverness.

**Example 2.4.2** Solve \( x^2 - 9x = 0 \).

**Solution:** Here we can simply factor the left-hand side (LHS), and the factors will be our “A” and “B” in our new form of the theorem.

\[
 \begin{align*}
 x^2 - 9x &= 0 & \text{given} \\
 x (x - 9) &= 0 & \text{factored LHS} \\
 \iff & \quad (x = 0) \lor (x = 9) & \text{no zero divisors} \\
 x & \in \{0,9\} & \text{add 9}
\end{align*}
\]

The above example was already in a form Polynomial=0, and so solving it was a matter of factoring the polynomial.\(^{38}\)

---

\(^{37}\)There is a superset of real numbers, called complex numbers and denoted by \( \mathbb{C} \), in which every nonzero element has two square roots. Though that set is surprisingly useful, we will not delve deeply at all into the structure of \( \mathbb{C} \) in this text.

\(^{38}\)While we will look at algebra in sophisticated ways here, we will not comprehensively review such skills as factoring (writing as a product) second-degree polynomials \( ax^2 + bx + c \). Some review of these skills will occur incidentally, but the reader is referred to almost any of the hundreds of algebra texts available for such skills.
Example 2.4.3 Solve the equation \( x^2 + 21x = 100 \).

Solution: We first make this a question about when a polynomial will equal zero, so we can use our theorem on no zero divisors.

\[

given
\]
\[
\Longleftrightarrow x + 21x - 100 = 0 \quad \text{'subtracted 100'}
\]
\[
\Longleftrightarrow (x + 25)(x - 1) = 0 \quad \text{'factored'}
\]
\[
\Longleftrightarrow (x + 25 = 0) \vee (x - 1 = 0) \quad \text{'no zero divisors'}
\]
\[
\Longleftrightarrow (x = -25) \vee (x = 1) \quad \text{'added -25, 1 respectively'}
\]
\[
\Longleftrightarrow x \in \{-25, 1\}.
\]

We will not continue to write justifications for every step, since the reader should be or become familiar enough with the logic as we proceed. We may instead write, for example, the following.

Example 2.4.4 Solve the polynomial equation \( x^5 - 5x^3 = 36x \).

Solution:

\[
x^5 - 5x^3 = 36x \\
\Longleftrightarrow x(x^4 - 5x^2 - 36) = 0 \\
\Longleftrightarrow x(x^2 - 9)(x^2 + 4) = 0 \\
\Longleftrightarrow (x + 3)(x - 3)(x^2 + 4) > 0 \\
\Longleftrightarrow x \in \{0, -3, 3\}.
\]

The above example used the fact that \( x^2 + 4 = 0 \) has no solution, and indeed \( x^2 + 4 \geq 4 > 0 \iff x^2 + 4 > 0 \) since \( x^2 \geq 0 \) regardless of \( x \in \mathbb{R} \). We could have included a line before the statement of the solution reading \( x = 0 \) or \( x + 3 = 0 \) or \( x - 3 = 0 \).

Example 2.4.5 Solve \((x + 4)(x - 3) = 18\).

Solution: We again need to rewrite this into the form Polynomial=0, in this case by multiplication first:

\[
(x + 4)(x - 3) = 18 \\
\Longleftrightarrow x^2 - 3x + 4x - 12 = 18 \\
\Longleftrightarrow x^2 + x - 30 = 0 \\
\Longleftrightarrow (x + 6)(x - 5) = 0 \\
\Longleftrightarrow (x = 6) \vee (x = 5 = 0) \\
\Longleftrightarrow x \in \{-6, 5\}.
\]

It should be pointed out that in each of the above polynomial equations we solved, assuming we execute each step correctly we do not in the end have to check our answers. This is because each step is logically equivalent to the previous step. In subsequent subsections we will encounter cases where we have \( \iff \) instead of \( \Longleftrightarrow \). In those cases we have to “check” the answers in the original equations, or one which is connected to the original equation through a series of equivalences, i.e., \( \Longleftrightarrow \)'s.
2.4. ALGEBRA

Example 2.4.6 Solve $x^3 - 6x^2 = 9x$.

Solution: We proceed as before:

$$x^3 - 6x^2 = -9x$$

\[\iff x^3 - 6x^2 + 9x = 0\]

\[\iff x(x^2 - 6x + 9) = 0\]

\[\iff x(x - 3)^2 = 0\]

\[\iff (x = 0) \lor (x - 3 = 0)\]

\[\iff x \in \{0, 3\}.\]

In the above example, we could have instead written $x(x - 3)(x - 3) = 0 \iff x \in \{0, 3, 3\}$, but writing 3 twice is redundant. That said, it is sometimes interesting to see how often the factor $x - a$ appears in a factorization of an expression. If $(x - a)^n$ appears as a factor in a polynomial but $(x - a)^{n+1}$ does not, we call $x = a$ a zero, or root, of multiplicity $n$ of the polynomial. Clearly it is called a zero because if we input $x = a$ into a polynomial with $x - a$ as a factor, then the polynomial will output zero; the polynomial will be of the form $(x - a)q(x)$ where $q(x)$ is another polynomial, and at $x = a$ we will have $(a - a)q(a) = 0 \cdot q(a) = 0$. It is often called a root because it is a solution of the equation Polynomial = 0. (Recall solutions of $x^2 = 5$ are called “square roots of 5.”) In the above example, $x = 0$ is a zero of multiplicity 1, where $x = 3$ is a zero of multiplicity 2.

At times we will be interested in the multiplicities of zeroes of polynomials, because if $x = a$ is a zero multiplicity $n$ of a polynomial, ultimately implying that the polynomial can be written $(x - a)^n q(x)$ where $q(a) \neq 0$ (see Theorem 2.4.2 below), then if $n$ is even the sign of the polynomial away from $x = a$ only depends upon the sign of $q(x)$, since $(x - a)^n$ will never be negative. If $n$ is odd, the polynomial changes signs as $x$ increases and passes through the value $x = a$, because the factor $(x - a)^n$ changes signs.

The following theorems should also be mentioned here:

Theorem 2.4.1 If $P(x)$ is a nontrivial polynomial (meaning $P(x)$ is not just the polynomial $0 = 0 + 0x + 0x^2 + \cdots$), then $P(a) = 0$ if and only if $x - a$ is a factor of $P(x)$:

$$(P(a) = 0) \iff (\exists q(x))[q(x) \text{ is a polynomial}] \land (P(x) = (x - a)q(x)]. \quad (2.44)$$

We will not prove this theorem, as it is best left for a course in algebra, though it should have a ring of truth to the reader. Using this, a corollary we will prove is the following:

Theorem 2.4.2 If $x = a$ is a zero of multiplicity $n$ of the polynomial $P(x)$, then $P(x) = (x - a)^n Q(x)$ where $Q(a) \neq 0$.

Proof: If $x = a$ is a zero of multiplicity $n$ of the polynomial $P(x)$, this means that $(x - a)^n$ is a factor of $P(x)$ but $(x - a)^{n+1}$ is not. Thus $P(x) = (x - a)^n Q(x)$.

Suppose (for purposes of proving otherwise) that $Q(a) = 0$. Then by Theorem 2.4.1, it would follow that $x - a$ is a factor of $Q(x)$, so there would exist another polynomial $Q_2(x)$ so that $Q(x) = (x - a)Q_2(x)$, and so $P(x) = (x - a)^n Q_2(x) = (x - a)(x - a)Q_2(x) = (x - a)^{n+1} Q_2(x)$, meaning $(x - a)^{n+1}$ is a factor of $P(x)$, but that is impossible because the multiplicity of $x = a$ is only $n$.

Since $[Q(a) = 0] \implies \mathcal{F}$, we must conclude $Q(a) \neq 0$, q.e.d.
2.4.2 More with Radicals

First we should point out the following:

**Principle 2.4.5** Suppose \( k \geq 0 \). Then

\[
x^2 = k \quad \iff \quad (x = \sqrt{k}) \lor (x = -\sqrt{k}).
\]

(2.45)

Instead of writing out that whole right hand side, we can simply write \( x = \pm \sqrt{k} \). For a proof, notice that if \( k \geq 0 \) then \( \sqrt{k} \) exists, and then

\[
x^2 = k \quad \iff \quad x^2 - k = 0 \quad \iff \quad (x - \sqrt{k})(x + \sqrt{k}) = 0 \quad \iff \quad (x - \sqrt{k} = 0) \lor (x + \sqrt{k} = 0) \quad \iff \quad (x = \sqrt{k}) \lor (x = -\sqrt{k}), \quad \text{q.e.d.}
\]

From the above proposition we can see a problem with squaring both sides of an equation: there is a loss of information about the sign (positive, negative or zero) of the number being squared. This can be summarized in the following logical statement, which ultimately says \( x = \sqrt{k} \implies x = \pm \sqrt{k} \) (see (??), \( P \implies P \lor Q \)):

\[
x = \sqrt{k} \implies x^2 = k \quad \iff \quad x = \pm \sqrt{k}.
\]

(2.46)

**Example 2.4.7** Consider the following:

\[
x = -5
\]

\[
\iff (x)^2 = (-5)^2 \quad \text{‘squared both sides}
\]

\[
x^2 = 25 \quad \text{‘simplified}
\]

\[
\iff (x = 5) \lor (x = -5) \quad \text{‘from (2.46)}
\]

\[
x = \pm 5 \quad \text{‘restatement}
\]

When we summarize, we get \( x = -5 \implies x = \pm 5 \), which is true, but we lost some information about \( x \), and introduced an extraneous solution, \( x = 5 \).

The loss of information came in the second line, where we had implication only, instead of equivalence. Thus the student must be aware that raising both sides of an equation to an even power may introduce extraneous solutions, and that these must be checked with the original equation for consistency. In fact, whenever we have only simple implication, we need only check the resulting statement with the original, and all true solutions will hold true, while extraneous ones will be detected false. In the next example, we see how this complication arises in a less obvious way, and how we deal with it by checking the possible solutions in the original equation to see if they are truly solutions.

**Example 2.4.8** Consider the following:

\[
\sqrt{x - 1} + 3 = x
\]

\[
\iff \sqrt{x - 1} = x - 3 \quad \text{‘added (\( -3 \)) to both sides}
\]

\[
\iff (\sqrt{x - 1})^2 = (x - 3)^2 \quad \text{‘squared both sides}
\]

\[
x - 1 = x^2 - 6x + 9 \quad \text{‘multiplied out the right}
\]

\[
\iff 0 = x^2 - 7x + 10 \quad \text{‘added (\( -1 \)) to both sides}
\]

\[
\iff 0 = (x - 2)(x - 5) \quad \text{‘factoring}
\]

\[
\iff (x - 2 = 0) \lor (x - 5 = 0) \quad \text{‘no zero divisors}
\]

\[
\iff x \in \{2, 5\} \quad \text{‘solving the above equations.}
\]
2.4. ALGEBRA

Check:

\[
\begin{array}{lcl}
  x = 2 : & \sqrt{2-1} + 3 & = 2 \\
        & \sqrt{1} + 3 & = 2 \\
        & 1 + 3 & = 2 \quad False.
\end{array}
\]

\[
\begin{array}{lcl}
  x = 5 : & \sqrt{5-1} + 3 & = 5 \\
        & \sqrt{4} + 3 & = 5 \\
        & 2 + 3 & = 5 \quad True.
\end{array}
\]

Conclude that the solution of the original equation is \( x = 5 \). (Note how the computations above yielded \( \sqrt{x-1} + 3 = x \iff x = 5 \iff (x = 5) \lor (x = 2) \), giving an extraneous value \( x = 2 \).)

We noticed that we did not have equivalence all the way down, so we could only follow the implication from end to end one way. It is true that the first statement implies the last, but the reverse is not clear. However, we do know that if there is a solution, it is contained in the last statement (along with any extraneous solutions). Thus we need only check our possible solutions in the original equation.

One further complication of the loss of information, coupled with the fact that \( \sqrt{\cdot} \) refers only to nonnegative square root, is the following (readers should verify (2.47)):

**Principle 2.4.6** The relationship between \((\cdot)^2\) and \(\sqrt{\cdot}\) is summarized in the following statements:

\[
\begin{align*}
\sqrt{x^2} &= |x| \quad \forall x \in \mathbb{R}, \quad (2.47) \\
(\sqrt{x})^2 &= x \quad \forall x \geq 0. \quad (2.48)
\end{align*}
\]

**Example 2.4.9** Solve the equation \( \sqrt{x^2 + 6x + 9} = 7 \).

\[
\begin{array}{lcl}
  \sqrt{x^2 + 6x + 9} & = 7 & \text{given} \\
  \iff \sqrt{(x+3)^2} & = 7 & \text{rewritten} \\
  \iff |x+3| & = 7 & \text{principle} \\
  \iff x + 3 & = \pm 7 & \text{absolute value fact} \\
  \iff x & = -3 \pm 7 & \text{subtracted 3} \\
  \iff x & \in \{-10, 4\} & \text{simplified, done.}
\end{array}
\]

This does not have to be checked, because we have logical equivalence all the way down. An alternative method might be simpler algebraically, but would not have equivalence:

\[
\begin{array}{lcl}
  \sqrt{x^2 + 6x + 9} & = 7 & \text{given} \\
  \implies x^2 + 6x + 9 & = 49 & \text{squared both sides} \\
  \iff x^2 + 6x - 40 & = 0 & \text{subtracted 49} \\
  \iff (x-4)(x+10) & = 0 & \text{factored} \\
  \iff (x-4 = 0) \lor (x+10 = 0) & \text{no zero divisors} \\
  \iff x & \in \{4, -10\} & \text{solved both equations.}
\end{array}
\]

A quick check shows that \( x = 4 \) and \( x = -10 \) in the original equation both give \( \sqrt{49} = 7 \), which is true. Thus both solution candidates solve the original equation, and we conclude \( x \in \{4, -10\} \) as before.

To show that the final check is crucial in such a case, consider the same example with a simple modification:

**Example 2.4.10** Solve the equation \( \sqrt{x^2 + 6x + 9} = -7 \).

Everything is the same as before, except for the \(-7\) in the first line:
\[ \sqrt{x^2 + 6x + 9} = -7 \quad \text{‘given} \]
\[ \implies x^2 + 6x + 9 = 49 \quad \text{‘squared both sides} \]
\[ \iff x^2 + 6x - 40 = 0 \quad \text{‘subtracted 49} \]
\[ \iff (x - 4)(x + 10) = 0 \quad \text{‘factored} \]
\[ \iff x = 0 \lor x + 10 = 0 \quad \text{‘no zero divisors} \]
\[ \iff x \in \{4, -10\} \quad \text{‘solved both equations.} \]

However, the first line reads \( \sqrt{49} = -7 \) for either of our solution candidates, and this is clearly false. There are no solutions of this equation. A trained student would likely notice this from the first line if he were to use the other method, since the first line gives immediately \( |x + 3| = -7 \) which is clearly false for all \( x \).

Another pitfall comes in multiplication by unknown quantities. One problem comes from the fact that there may be some values for the unknown which, when used for multiplying both sides of an equation, may be in effect multiplying both sides by zero. Once both sides are multiplied by zero, precise information about the quantities may be lost. Consider the following example:

**Example 2.4.11** Consider the following equation:
\[ x^2 - 9 = 0 \quad \text{‘given} \]
\[ \implies x \cdot (x^2 - 9) = x \cdot 0 \quad \text{‘multiplied by} \ x \]
\[ \iff x(x + 3)(x - 3) = 0 \quad \text{‘factored left, simplified right} \]
\[ \iff x = 0 \]
\[ \forall (x + 3) = 0 \]
\[ \forall (x - 3) = 0 \quad \text{‘no zero divisors} \]
\[ \iff x \in \{0, -3, 3\} \quad \text{‘but} \ x = 0 \text{ is wrong!} \]

We began with an equation which had solution \( x = -3, 3 \), and ended with the statement \( x = 0 \). The problem here occurred in the second line, where we multiplied both sides by \( x \), without knowing what \( x \) is. In particular, for one value of \( x \)—namely \( x = 0 \)—we multiplied both sides by zero. Here it is not so difficult to see what went wrong: for some reason we decided to multiply both sides by \( x \), which introduced the extraneous solution \( x = 0 \).

Though it seems unlikely we would arbitrarily multiply both side of an equation by, say, \( x \), there are times we may feel forced to. Consider the following:

**Example 2.4.12** Solve the following equation:
\[ \frac{x^2 + 9}{x} = \frac{9 - x}{x} \quad \text{‘given} \]
\[ \implies x^2 + 9 = 9 - x \quad \text{‘multiplied by} \ x \]
\[ \iff x^2 + x = 0 \quad \text{‘added} \ (-9) + x \text{ to both sides} \]
\[ \iff x(x + 1) = 0 \quad \text{‘factored} \]
\[ \iff (x = 0) \lor x + 1 = 0 \quad \text{‘no zero divisors} \]
\[ \iff x \in \{0, -1\} \quad \text{‘solved above equations.} \]

A quick check shows that \( x = 0 \) is not a solution (since we would be dividing by zero in the first line), while \( x = -1 \) is indeed a solution.

One could argue that, what we effectively did in that second line, was to multiply both sides by zero for a certain value of \( x \) (namely, when \( x = 0 \)). Multiplying by that particular value on both sides may introduce extraneous solutions, as indeed occurred in this case.

Another approach to the above problem is to note from the outset that \( x \neq 0 \), and then we can have equivalence all the way down. It would read something like the following:
\[
\frac{x^2 + 9}{x} = \frac{9 - x}{x}, \quad x \neq 0 \quad \text{‘given}
\]
\[
\iff \quad x^2 + 9 = 9 - x \land (x \neq 0) \quad \text{‘multiplied by } x
\]
\[
\iff \quad x^2 + x = 0 \land (x \neq 0) \quad \text{‘added } x - 9 \text{ to both sides}
\]
\[
\iff \quad x(x + 1) = 0 \land (x \neq 0) \quad \text{‘factored}
\]
\[
\iff \quad (x = 0) \lor (x + 1 = 0) \land (x \neq 0) \quad \text{‘no zero divisors}
\]
\[
\iff \quad x \in \{0, -1\} \land (x \neq 0) \quad \text{‘solved above equations}
\]
\[
\iff \quad x = -1 \quad \text{‘done.}
\]

2.4.3 Completing the Square

Definition 2.4.3 A \textit{quadratic polynomial} is an expression of the form
\[
ax^2 + bx + c, \quad \text{where } a, b, c \in \mathbb{R}, \ a \neq 0. \quad (2.49)
\]

In this section we will exploit the form of the perfect square
\[
(x + k)^2 = \frac{x^2 + 2kx + k^2}{x^2 + bx + k^2}
\]

If we assume that \(2k = b\), i.e., that \(b = k/2\), or just perform a brute-force computation, we can re-write this
\[
\left( x + \frac{b}{2} \right)^2 = x^2 + bx + \frac{b^2}{4}. \quad (2.50)
\]

Next we notice the solution to the following special type of equation (see (2.46)), where \(d \geq 0\):
\[
(x + k)^2 = d \iff x + k = \pm \sqrt{d} \iff x = -k \pm \sqrt{d}. \quad (2.51)
\]

Written with the same substitution as before gives
\[
\left( x + \frac{b}{2} \right)^2 = d \iff x + \frac{b}{2} = \pm \sqrt{d} \iff x = -\frac{b}{2} \pm \sqrt{d}. \quad (2.52)
\]

This provides a useful technique for solving quadratic equations for which factoring is not easily accomplished.

Example 2.4.13 Solve the equation \(x^2 + 4x = 9\).

Recognizing that \(b = 4\) here, we calculate \(b^2/4 = 16/4 = 4\), which “completes the square” on the left of the equation, and so we add that term to both sides:
\[
\begin{align*}
\quad & x^2 + 4x + \left( \frac{4}{2} \right)^2 = 9 + \left( \frac{4}{2} \right)^2 \\
\iff & x^2 + 4x + 4 = 9 + 4 \\
\iff & (x + 2)^2 = 13 \\
\iff & x + 2 = \pm \sqrt{13} \\
\iff & x = -2 \pm \sqrt{13}.
\end{align*}
\]
2.4.4 The Quadratic Formula

Theorem 2.4.3  Given a quadratic equation of the form

\[ ax^2 + bx + c = 0, \quad a \neq 0 \]  \hspace{1cm} (2.53)

- if \( b^2 - 4ac \geq 0 \), then the solution to the equation is given by

\[ x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}; \]  \hspace{1cm} (2.54)

- if \( b^2 - 4ac < 0 \), the equation (2.53) has no real solution.

The quadratic formula (2.54) helps us avoid factoring or completing the square. If the right hand side of (2.53) factors readily, then factoring is easier. Otherwise the quadratic formula is usually simpler than completing the square. This method is certainly amenable to programming. The proof relies on the method of completing the square, but once we have the result, we should and will quote it with impunity.

**Proof:** Assume that \( a \neq 0 \). Then we rewrite (2.53) as follows:

\[ ax^2 + bx + c = 0 \iff ax^2 + bx = -c \]

\[ \iff x^2 + \frac{b}{a}x = -\frac{c}{a} \]

\[ \iff x^2 + \frac{b}{a}x + \left(\frac{1}{2} \cdot \frac{b}{a}\right)^2 = -\frac{c}{a} + \left(\frac{1}{2} \cdot \frac{b}{a}\right)^2 \]

\[ \iff x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} = -\frac{c}{a} + \frac{b^2}{4a^2} \]

\[ \iff \left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - c}{4a^2} \]

\[ \iff x + \frac{b}{2a} = \pm \sqrt{\frac{b^2 - c}{4a^2}} \]

\[ \iff x = \frac{-b}{2a} \pm \sqrt{\frac{b^2 - 4ac}{4a^2}} \]

\[ \iff x = \frac{-b}{2a} \pm \frac{1}{2a} \sqrt{b^2 - 4ac} \]

\[ \iff x = -\frac{b}{2a} \pm \frac{1}{2a} \sqrt{b^2 - 4ac} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \]

The proof is essentially complete, but a couple of comments should be made. First, since we need the two cases \(+/-\), we get both quantities with or without the absolute values. If \( a > 0 \), we still have \( \pm \), whereas if \( a < 0 \), we have \( \mp \), giving the same two cases, so we can leave them written with \( \pm \). Also, we are looking for solutions \( x \in \mathbb{R} \), so if \( b^2 - 4ac < 0 \) in the last line, the final line is equivalent to \( \mathcal{F} \), meaning that so is (2.53), meaning there are no real solutions.

The quantity \( b^2 - 4ac \) is important enough to warrant a name:

\[ 39 \] If \( a = 0 \) the formula (2.53) does not work (since we would be dividing by zero), but in such a case we are really solving \( bx + c = 0 \), which is linear and can be easily solved without resorting to quadratic methods.
Definition 2.4.4 The quantity $b^2 - 4ac$ in Equation (2.53) is called the discriminant of the equation.

Given the form of the quadratic formula, in particular the term $\pm \frac{1}{2a} \sqrt{b^2 - 4ac}$, the following implications for the sign of the discriminant are clear:

Theorem 2.4.4 Given (2.53).

1. $b^2 - 4ac = 0 \iff (2.53)$ has exactly one real solution.
2. $b^2 - 4ac > 0 \iff (2.53)$ has exactly two real solutions.
3. $b^2 - 4ac < 0 \iff (2.53)$ has no real solutions.
Exercises

1. Prove Principle 2.4.3, page 112. You may wish to look at Principle 2.4.2.

2. Show by direct calculation that, if \( a \neq 0 \), then
\[
x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\]
satisfies \( ax^2 + bx + c = 0 \). (You need to check both values for \( x \).)

3. Show that, if \( a \neq 0 \), then
\[
a \left( x - \frac{-b + \sqrt{b^2 - 4ac}}{2a} \right) \left( x - \frac{-b - \sqrt{b^2 - 4ac}}{2a} \right) = ax^2 + bx + c.
\]
Notice that once this is established, Exercise 2 is trivial.
2.5 Functions-I

Behind every calculus problem looms at least one function. In fact one can credibly argue that the main goal of calculus is to provide (powerful) new tools for analyzing functions. As we will discuss momentarily, functions play a crucial role in our analysis of physical reality, and so ultimately calculus provides a (tremendous, as we will see) leap in our ability to analyze real-world problems.

The methods for analyzing functions can roughly, and imperfectly, into two classes:

(1) analyses of the internal “building blocks” of the function, and how they interact, and
(2) analyses of the emergent properties of these combinations of functions.

In this section we will concern ourselves more with (1), because much intuition can be gained from that approach. However, to illustrate our predictions we will also computer-generate the graphs of many of these functions, which is arguably a tool of the second approach (2). Eventually, however, the combinations of functions become sufficiently complicated that we find ourselves giving up—or just nearly so—the analysis of the inner workings of the functions and apply calculus principles to approach the functions from their emergent properties. Eventually the best analysis usually comes from a marriage of these two approaches.

It cannot be overstated that students should strive to develop skills needed to be able to spot trends in the functions without resorting to the calculus. Of particular importance is the ability to “parse” a given function, i.e., to examine the various simple components, and then “synthesize” information on how the various components would behave in isolation with how they are connected, interacting to form the final output.\(^{40}\)

However, with calculus we have many tools which are so powerful that one is tempted to conclude calculus should obsolete the original precalculus theory of functions. This is not the case. The tools are different and complementary. In fact, once the calculus is applied we often need the pre-calculus theory of functions to interpret the calculus computation.

It should be pointed out that a discussion of functions can proceed on a very abstract level, but for most of what we do here it will be sufficient to consider functions as processes which take an input and return an output in a deterministic way, meaning that when we feed a particular function the same input on two occasions, we get the same output both times.

A function would not be terribly interesting unless it could process many different inputs. The set of inputs that a function can process is called its domain. The domain is usually apparent from the description of the function, as we will see later. The set of all possible outputs is called the range of the function, and it is usually a more difficult set to compute.

Many physical phenomena are either functions or can be very well modeled by functions. Here are some examples of functional relations:

- The area of a circle is a function of its radius: \( A = \pi r^2 \); if the radius is known, the area follows.
- The voltage across an Alternating Current (AC) power outlet is a function of time, which we will see is oscillatory.
- In a right triangle, the ratio of lengths of the side opposite one of the nonright angles to the hypotenuse is a function (sine) of the angle’s measure.

\(^{40}\)Without actively considering how a particular function is put together with its component parts—each having its own peculiar behavior—to form a single process, we can usually apply formulas but we will not be able to fully interpret their outputs, and would remove a valuable layer of anticipations and consequent error correction from our computations.
Here are some examples that are well-approximated by functions, and would be functional relations in "ideal" circumstances.

- The gain of an audio amplifier is approximately a function of the position of the volume knob. This is true ideally, and assumes the other knobs, the qualities of the components in the circuits, the power supplied, etc., are all fixed into an unchanging position. In fact the gain might also change with the level of the input, so the input is usually restricted to some set of standard, acceptable levels.

- The range of an artillery projectile is approximately a function of the angle at which it is fired. This assumes each projectile is identical, and the artillery piece, the wind, temperature, humidity, etc., are all fixed.

- One's net monetary worth is a function of time, though it might not be perfectly known, and opinions may vary. It is also probably impossible to predict exactly for future values of the variable time.

Since the input can usually be varied, we usually consider it a variable. To name this particular variable we will usually use the obvious and descriptive term input variable, though most texts use the equally descriptive but more mathematically flavored term independent variable.

Because we can vary the input, we expect the output may also vary, and will so describe it as the output variable, or value of the function, or occasionally (here) use the more commonly found term dependent variable. This last term is quite descriptive as, by definition of function, the output variable completely depends upon the input variable.

In fact, often a function has many (independent) input variables, but this text is mostly concerned with single-variable calculus, for which we have one input variable (usually $x$) and one output variable (usually $y$).

### 2.5.1 Various Definitions of Functions; Notation

There are many ways to describe functions. To name a couple, we can look at them as mappings (which "map" the independent variable values to their respective dependent variable outputs), and they can be described as processes or "machines." We will include the abstract definition later, to be complete. Ideally it is best to consider functions in all these ways. However for our purposes we will concentrate on the notion of functions as machines; for our purposes, a function is defined by its action, by which it takes inputs from a set called its domain, and deterministically return outputs from another set called a target set, or if we know the exact set of possible outputs, we call this target set the function’s range (compare to the earlier description of range). By deterministically we mean that, for a given function, if we know its input then the output is determined; the same input cannot yield two different outputs.

As is the case for any mathematical object, it is customary to give a function a name. The most common name is $f$, for “function,” but other letters from various alphabets are common ($g$, $h$, $\phi$, $\Phi$, for some examples), and some particular functions have common names (sin, tan, log, etc.). See Figure 2.5, page 125.

---

41Like those in other “hard” sciences, professionals and developers in the mathematical sciences over the decades and centuries have settled upon some very refined definitions of important concepts. Particularly in mathematics, a definition must be absolutely precise, because we proceed with absolute logic from our definitions. Our definitions must also be robust, for many reasons, for instance so we can recognize the same phenomena in different settings. Some find these “refined” definitions very awkward at first. Part of the process of maturing in one’s understanding of a field is the realization that, ultimately, these definitions are in fact the most “natural” for what we require of them (and we require a lot!).
2.5. FUNCTIONS-I

Once we have a name for the function, we then usually give names to the input and output variables. The most common name for the input variable, also known as the independent variable, is \( x \) (though \( t \), \( \theta \), and others will be used on occasion). If \( f \) is the function, and the input variable is \( x \), then the output is usually written with the slightly unwieldy but very descriptive \( f(x) \). Indeed it helps to visualize that the input \( x \) is processed by the function \( f \), and so when we see the symbols \( f(x) \) we are looking at the final output.

Whole research papers are written regarding how students interpret \( f(x) \). Ideally the function itself is seen as an object in its own right, and then one realizes that it can be completely defined by its actions, taken collectively. In other words, if two functions always return the same output when given the same inputs, they are considered the same functions. As is often the case, the actual mathematical definition is more abstract, and while we mention it here for further meditation, we will not use it within the rest of the text because it de-emphasizes functions as actions or processes taking inputs and deterministically returning outputs. See Footnote 41, page 124.

**Definition 2.5.1** (Abstract) A function \( f \) is a set of ordered pairs where the first in each pair comes from a given set called the domain, say \( S \) of the function, and where

\[
(\forall x_1, x_2, y_1, y_2) \left[ (x_1, x_2 \in S) \land ((x_1, y_1), (x_2, y_2) \in f) \land (x_1 = x_2) \rightarrow (y_1 = y_2) \right].
\]

To re-interpret in our earlier terms, the definition states that if the inputs processed by a particular function are the same \((x_1 = x_2)\), then the outputs will be the same \((y_1 = y_2)\). It is this deterministic nature which is key to the concept of function.
2.5.2 Functions as Processors of Inputs

We will not again consider the function itself as a set of ordered pairs throughout the rest of the text. However we will have much use for the concept of domain, so we will redefine it for our purposes.

Definition 2.5.2 For a function \( f \), we define its domain to be the set of all possible inputs which \( f \) can process.

While this will be somewhat circular, we can see that the deterministic nature of a function \( f \) with domain \( S \) can be summarized as follows:

\[
(\forall x_1, x_2 \in S) [(x_1 = x_2) \rightarrow (f(x_1) = f(x_2))].
\]

In many (but not all) cases, a function can be given by an expression, which then describes the action of the function.

Example 2.5.1 Suppose that for all \( x \in \mathbb{R} \), we define \( f(x) = x^2 + 1 \). Then

\[
\begin{align*}
    f(1) &= 1^2 + 1 = 2, \\
    f(2) &= 2^2 + 1 = 5, \\
    f\left(\frac{1}{2}\right) &= \left(\frac{1}{2}\right)^2 + 1 = \frac{5}{4} = 1.25.
\end{align*}
\]

The function in the above example can be described as a process, by which the input is first squared, and the result is added to 1. Of course it is faster to just write \( f(x) = x^2 + 1 \). However, there is something gained to writing instead

\[ f(\ ) = (\ )^2 + 1, \]

...to emphasize that what ever is in the parentheses on the left will be processed like the parentheses on the right. For instance, we note that\footnote{We will use this “empty parentheses” convention as we diagram the actions of functions later in this chapter. However, it—and the notion of diagramming functions in flow chart-like methods—is not a typical explanatory device in a mathematics course, but is more likely to be found in a computer science course. We borrow from that field when it helps illuminate our endeavors here.}

\[
\begin{align*}
    f(x^2) &= (x^2)^2 + 1 = x^4 + 1, \\
    f(x + h) &= (x + h)^2 + 1 = x^2 + 2hx + h^2 + 1,
\end{align*}
\]

and so on. What was inside the parentheses is processed in the same arithmetic manner as was \( x \) in the formula \( f(x) = x^2 + 1 \).

2.5.3 Graphs, Domain and Range

The graphical illustration of a function, when possible, is a very powerful analytical tool. When both the input and output variables are numerical, in particular real numbers, in theory a graph of the output versus the input can convey complete information about the action of the function in all circumstance. In other words, the graph can serve as a definition of the function.

Unfortunately limitations of space and resolution too often restrict our ability to plot a function with completeness and absolute precision. For that and other reasons, this textbook will...
not take a “graphing calculator first” approach to functions employed by many texts. However the graph of a function is a very powerful analytical tool.

There is much notation which is used in dealing with functions, even in the abstract. Some of it is useful here because of its descriptive nature. For instance, if we know that the outputs of \( f \) will always be contained in some set \( T \), we can write

\[
f : S \rightarrow T,
\]

read “\( f \) maps \( S \) into \( T \).” We can think of \( T \) as a target set for \( f \). The notation above implies that all possible inputs into \( f \) (i.e., all values in the domain \( S \)) will yield outputs in the set \( T \). There are usually many possible sets \( T \), because once we have a legitimate target set, any superset of it is also a target set. Target sets are usually easy to construct, but the actual set of outputs can sometimes be elusive because we have to consider all possible inputs. Still it is worth finding, so we can know exactly what kinds of outputs are possible. This set of all possible outputs is called the range of the function. If we denote this set by \( R \), for a particular function \( f \), then we can use our set notation to write

\[
R = \{ y | (\exists x \in S)[y = f(x)] \}.
\]

Another notation which is occassionaly useful is the pointwise “maps to” arrow, \( \mapsto \), where \( y = f(x) \) can also be written \( x \mapsto y \). We will have occasion to use this notation.

Note that when we write \( f : S \rightarrow T \) the arrow is not a logical implication arrow, but a visual cue that \( f \) inputs elements of \( S \) and outputs elements of \( T \).

### 2.5.4 Functions By Formulas

Very often the action of \( f \) on an arbitrary \( x \) in the domain will be given by a formula. For example, perhaps \( f \) acts on every real number by squaring the number and then adding 1. Then we could write

\[
f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = x^2 + 1. \tag{2.55}
\]

This is read, “\( f \) maps \( \mathbb{R} \) into \( \mathbb{R} \), where \( f \) of \( x \) equals \( x^2 + 1 \).” The variable \( x \) in (2.55) is a place holder, or dummy variable in which we follow through the formula to define, or even “probe” the action of \( f \) on an arbitrary input \( x \).\(^{43}\) The variable \( x \) is also often called the argument of the function \( f \) in the expression \( f(x) \). Another notation which is commonly used is (notice the difference in the arrow) is

\[
x \overset{f}{\mapsto} x^2 + 1, \tag{2.56}
\]

showing that \( x \) is “transformed” or “mapped” to the value \( x^2 + 1 \) through the function \( f \). In the notation of (2.55) we can calculate

\[
\begin{align*}
f(1) &= (1)^2 + 1 = 2, \\
f(2) &= (2)^2 + 1 = 5, \\
f(-10) &= (-10)^2 + 1 = 101,
\end{align*}
\]

\(^{43}\)An elegant term we could borrow from computer science for the formula which defines a function, as in \( f(x) = x^2 + 1 \), is function prototype. It would be appropriate, since a prototype is usually built to demonstrate a concept, such as the particular function here which is “demonstrated” on the variable \( x \) by the formula. Unfortunately this is not the way the term is used in computer science, where “function prototype” just refers to the lines of code which explain to the compiler what kinds of variables a function will input and output, without any further information on how the variable(s) will be processed.
while in the notation of (2.56) we would write

\[ f(1) \mapsto 1^2 + 2 = 2, \]
\[ f(2) \mapsto 2^2 + 1 = 5, \]
\[ f(-10) \mapsto (-10)^2 + 1 = 101. \]

This may seem simple enough with numbers, since we simply substituted \( x = 1, 2, -10 \) respectively into, say, (2.55). Our understanding of the role of \( x \) in the formula becomes more important when our inputs become more complicated or abstract. For some examples, consider

\[ f(\sqrt{x}) = (\sqrt{x})^2 + 1 = x + 1, \]
\[ f\left(\frac{1}{x}\right) = \left(\frac{1}{x}\right)^2 + 1 = \frac{1 + x^2}{x^2}, \]
\[ f(-x) = (-x)^2 + 1 = x^2 + 1, \]
\[ f(x + 2) = (x + 2)^2 + 1 = x^2 + 4x + 4 + 1 = x^2 + 4x + 5, \]
\[ f(\text{Bob}) = (\text{Bob})^2 + 1. \]

In the above, we again replaced \( x \) from (2.55) with \( \sqrt{x}, \frac{1}{x}, -x, x + 2 \) and Bob, respectively. This may seem unnatural until we again remember that \( x \) was just a placeholder in the formula which defined the action of \( f \) (always one of squaring the input, and then adding 1). We also have to be careful for which values of \( x \) the expression makes sense for. In the first example, because the first action on \( x \) is taking its square root we require \( x \geq 0 \), even though the simplified expression glosses over this requirement. We must look at the original expression for \( f(\sqrt{x}) \) to decide which \( x \)-values it is valid for. In the second example, we need \( x \neq 0 \), while the expressions for \( f(-x) \) and \( f(x + 2) \) were valid for all \( x \in \mathbb{R} \). Finally, the last expression is valid as long as Bob \( \in \mathbb{R} \).

We are given great latitude in defining functions. The only condition is that they are deterministic processes, so that if we know the input \( x \) from the domain, then the unique output \( f(x) \) is completely determined. (For each \( x \) in the domain, there is exactly one \( f(x) \) in the target set.) Of course the definition of \( f(x) \) must also make sense for each \( x \) in the domain.
2.6 Basic Functions of One Variable

To a mathematician, every function has its own story of the relationship between the input and the output. Sometimes it tells its story best descriptively; sometimes algebraically; sometimes graphically. Usually all three approaches are useful. Furthermore, sometimes a function arises for consideration from some practical problem or describes some real-world phenomenon, and other times it arises from what seem like purely mathematical considerations. In both settings, the calculus helps us ask more interesting questions about the functions, and provides for many interesting answers.\footnote{It should be pointed out that what seem like purely mathematical questions are often eventually quite useful in solving real-world problems. For instance, when Albert Einstein derived his General Relativity, it was fortunate that a relatively new field called Absolute Differential Calculus, also known as the calculus of tensors, or just tensor analysis, was already somewhat developed. That field got a tremendous boost from Einstein’s interest as well. So a field with little known application to the real world became surprisingly useful, and was paid back in kind by the sudden interest in the field that emerged from the physics community’s further development of general relativity.}

But before we look at the calculus, it is very important to understand the basic mechanisms of the simpler functions, and how they interact when we tweak or combine them to construct more complex functions. First we look at several common functions themselves, leaving the questions of how they combine to form new functions for later.

Each function considered in this section will be examined from two related perspectives:

1. The actual definition, which usually contains the motive for our interest in the function;
2. The behavior of the function, by which we mean the manner in which the outputs vary with the inputs, as often summarized by the function’s graph.

In connecting these two aspects of the functions introduced below, we will be well on our way to preparing for much more complicated functions.

2.6.1 “Linear” Functions

These are functions which, when graphed, yield a straight line.\footnote{Later we will see that “linear” in fact means the variables appear only in the first degree, but this must be clarified further. In fact “linear” is more restrictive in other contexts beyond the scope of this text.} By the nature of functions these must be nonvertical lines, or else one input would yield more than one output (infinitely many in fact). A very important property of such a graph is the slope, which we can show must be constant. This is the ratio of the “rise” of the graph to its “run.” To be precise, if we look at a point \((x, y)\) on the graph, and then perturb the horizontal \((x)\) variable by an amount \(\Delta x\) to arrive at another point \((x + \Delta x, y + \Delta y)\) on the graph, so that \(\Delta x\) is the horizontal change of position and \(\Delta y\) the vertical change, then we notice that the ratio

\[
\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}
\] (2.57)

is constant. This actually follows from similar triangles generalized slightly to allow for both positive and negative position changes.\footnote{Recall that triangles with the same interior angles are called similar. For two such triangles, the corresponding sides are proportionate. Suppose the triangles below have the same interior angles opposite \(a_1\) and \(a_2\), \(b_1\) and \(b_2\), and \(c_1\) and \(c_2\).}
used in the sciences to mean “change in.” So \( \Delta x \) is to be treated as one quantity, specifically some change in the \( x \)-variable. This can be described as
\[
\Delta x = x_{\text{final}} - x_{\text{initial}},
\]
and so \( \Delta x \) does give the change from initial to final value of the “\( x \)-variable.” When \( x \) changes from 0 to 5, the change is \( \Delta x = 5 - 0 = 5 \). A positive \( \Delta x \) means an increase in the \( x \)-variable, while a negative \( \Delta x \) means a decrease. In the discussion above, the “\( x \)-variable” is considered to change from \( x \) to \( x + \Delta x \). Of course linguistically this does not make sense, since \( x \) itself does not change, but we are to understand that the quantity which began being \( x \) is changed to \( x + \Delta x \). Similarly with \( y \) and \( y + \Delta y \).

It is not difficult to derive formulas for linear functions. Indeed, all that is necessary is for us to have knowledge of one point, say \((x_1, y_1)\), and the slope \( m \). Then any variable point \((x, y)\) on the curve, except for \((x_1, y_1)\) itself, satisfies the formula for slope \((y - y_1)/(x - x_1) = m\). If we

Then the following relationships hold:
\[
\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}
\]

From simple algebra we then get the following:
\[
\frac{a_1}{b_1} = \frac{a_2}{b_2}, \quad \frac{a_1}{c_1} = \frac{a_2}{c_2}, \quad \frac{b_1}{c_1} = \frac{b_2}{c_2}.
\]
multiply both sides by \((x - x_1)\) we get either of the following:

\[
\begin{align*}
  y - y_1 &= m(x - x_1), \\
  y &= y_1 + m(x - x_1).
\end{align*}
\]

The first equation (2.58) is called the Point-Slope formula for a line through \((x_1, y_1)\) with slope \(m\). It is commonly found in algebra and calculus textbooks. However, we will instead use the second form (where we added \(y_1\) to both sides), (2.59), calling it the Modified Point-Slope formula. It will be more similar to important equations we derive later, and so it will be useful to prepare early for such things. In terms of function notation, if \(y = f(x)\), then (2.59) becomes

\[
f(x) = y_1 + m(x - x_1).
\]

If we happen to know the \(y\)-intercept \((0, b)\), where the line intersects the \(y\)-axis, then we can use this for our point and have \(f(x) = b + m(x - 0)\), or either of the following:

\[
\begin{align*}
  y &= mx + b, \\
  f(x) &= mx + b,
\end{align*}
\]

Equation (2.61) is perhaps the earliest studied form we learn in school, but is not as natural to derive as the modified point-slope forms in calculus problems.

\[
f(x) = mx + b.
\]

### 2.6.2 Simple Powers

We begin with the function \(f(x) = x^2\). Since \(f(x) = x \cdot x\), this function can process any real number and produce an output. We are interested in analyzing how this output changes when \(x\) varies.

When \(x = 1, 2, 3, \ldots\), it becomes clear \(f(x) = x \cdot x\) will output larger and larger numbers when we input larger numbers. It is also the case when \(x = -1, -2, -3, \ldots\), and a casual observation we can make is that this particular \(f\) outputs the same for a positive value of \(x\) as it does for the additive inverse, so for instance both \(f(3)\) and \(f(-3)\) are the same, namely 9. So \(f(x)\) will take numbers with large absolute values and output large, positive numbers (larger than the inputs \(x\) if \(|x| > 1\)). What is often equally important is what occurs when \(x\) is a “smaller” number, for instance when \(-1 < x < 1\). For instance, \(f(1/2) = 1/4, f(1/10) = 1/100\), and so on. We note these two trends in the charts below:

<table>
<thead>
<tr>
<th>(x)</th>
<th>(f(x))</th>
<th>(x)</th>
<th>(f(x))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\pm 1)</td>
<td>1 (\pm 1)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(\pm \frac{1}{2})</td>
<td>(\frac{1}{4}) (\pm \frac{1}{2})</td>
<td>(\pm 2)</td>
<td>4</td>
</tr>
<tr>
<td>(\pm \frac{1}{3})</td>
<td>(\frac{1}{9}) (\pm \frac{1}{3})</td>
<td>(\pm 3)</td>
<td>9</td>
</tr>
<tr>
<td>(\pm \frac{1}{4})</td>
<td>(\frac{1}{16}) (\pm \frac{1}{4})</td>
<td>(\pm 4)</td>
<td>16</td>
</tr>
<tr>
<td>(\ldots)</td>
<td>(\ldots) (\ldots)</td>
<td>(\ldots)</td>
<td>(\ldots)</td>
</tr>
<tr>
<td>(\pm \frac{1}{10})</td>
<td>(\frac{1}{100}) (\pm \frac{1}{10})</td>
<td>(\pm 10)</td>
<td>100</td>
</tr>
<tr>
<td>(\pm \frac{1}{100})</td>
<td>(\frac{1}{10000}) (\pm \frac{1}{100})</td>
<td>(\pm 100)</td>
<td>10000</td>
</tr>
<tr>
<td>(\pm \frac{1}{1000})</td>
<td>(\frac{1}{1000000}) (\pm \frac{1}{1000})</td>
<td>(\pm 1000)</td>
<td>1000000</td>
</tr>
</tbody>
</table>

While formulas and charts are useful for understanding the action of a function, another, indespensible tool is its graph, when practical. The graph of \(y = x^2\) is given in Figure 2.7.


Theoretically, the graph illustrates all the behavior of the function, as described above, as some contemplation of the graph illustrates. Technically, the graph’s shape is that of a parabola, though we will not define the term here.\footnote{Many shapes resemble the parabola, but there is a precise geometric definition of the term (just as there are definitions for circles, ellipses, hyperbolas, and other shapes). A common mistake is to use the term parabola (or the adjective parabolic) to apply to every “U-shaped” graph, but that is incorrect. That said, in fact any graph of a polynomial of degree two, i.e., \( y = ax^2 + bx + c \) where \( a \neq 0 \), will be a parabola.

The geometric definition of a parabola is that it must be a plane figure, and consists of all points in that plane which are the same distance from a given point (focus) as from a given line (directrix) not containing the point. The reader is invited to draw what such a figure might look like. It becomes clear quickly that not all “U-shaped” curves fit the definition.}

We see already that it is not always easy to plot a function for a large range of the input variable. Figure 2.7 shows two partial graphs of this function. The first uses the same scale for \( x \) and \( y \), while the second uses very different scales for the two axes. Note that \( x \in [-10, 10] \) would require \( y \in [0, 100] \), which is not practical to graph with matched scaling in \( x \) and \( y \). Indeed, we should not always require that the input and output axes share the same scale, as they are often “incommensurable.” For instance, if \( x \) represents the length of a side of a square, then \( y = x^2 \) would represent area. If the units of \( x \) are in feet, then the units of \( y \) would need to be feet\(^2\), so indeed the units are often dissimilar. Often the horizontal axis represents a time scale and the vertical something very different, such as units of currency, power, or any other imaginable unit.

Many functions grow very rapidly, so our choices are often to plot a smaller input range, or to have a different scale for the \( y \)-axis, or perhaps to truncate the \( y \)-axis (or \( x \)-axis). The graph in Figure 2.7 actually shows the important trends in the behavior of \( f \); someone viewing the graph with the input range \([-3, 3]\) is not likely to be surprised by the behavior elsewhere, and indeed we plotted the interesting “features” of the graph. Some texts would call our partial graph “complete” because it actually presents, fairly exhaustively, all the behavior and features of the function.

We now consider other powers of \( x \). First we look at the even powers, namely \( x^2, x^4, x^6, \) etc., for comparison purposes. What we will find is that all the trends displayed by the function \( x^2 \)
are also present in \( x^4, x^6 \), and so on, but they are even more pronounced. Note that all of these output the same values at \( x = \pm 1 \) and \( x = 0 \), namely 1 and 0 respectively. But when "smaller" values are input to \( x^2 \), we saw the output was smaller still, and this will also occur with the higher, even powers of \( x \). On the other hand, larger numbers will output positive even powers of larger numbers, which will be larger still, and the higher the power the more pronounced that effect will be. The following table gives some idea of these trends.

<table>
<thead>
<tr>
<th>( x )</th>
<th>( x^2 )</th>
<th>( x^4 )</th>
<th>( x^6 )</th>
<th>( x^8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>\pm .001</td>
<td>\pm .000001</td>
<td>\pm .000000000000000001</td>
<td>\pm .000000000000000001</td>
<td>\pm .000000000000000001</td>
</tr>
<tr>
<td>\pm .01</td>
<td>\pm .0001</td>
<td>\pm .000000000000000001</td>
<td>\pm .000000000000000001</td>
<td>\pm .000000000000000001</td>
</tr>
<tr>
<td>\pm 1</td>
<td>.01</td>
<td>.0001</td>
<td>.000000000000000001</td>
<td>.000000000000000001</td>
</tr>
<tr>
<td>\pm 2</td>
<td>4</td>
<td>16</td>
<td>64</td>
<td>256</td>
</tr>
<tr>
<td>\pm 3</td>
<td>9</td>
<td>81</td>
<td>729</td>
<td>6561</td>
</tr>
<tr>
<td>\pm 4</td>
<td>16</td>
<td>256</td>
<td>4096</td>
<td>65536</td>
</tr>
<tr>
<td>\pm 10</td>
<td>100</td>
<td>10000</td>
<td>1000000</td>
<td>100000000</td>
</tr>
</tbody>
</table>

These trends are to be expected. The graphs are somewhat difficult to render because the trends occur so quickly. Figure 2.8a shows these trends, albeit incompletely.

Odd roots have similar behavior, with the exception that they are sensitive to the sign of the input. For instance, if \( f(x) = x^3 \), then \( f(2) = 8 \) while \( f(-2) = -8 \). Otherwise, again inputs \( x \in (-1,1) \) yield outputs of even smaller absolute values while inputs \( x \) with \( |x| > 1 \) yield outputs with even larger absolute values. The trends are somewhat illustrated in Figure 2.8b.

We can also compare sizes of outputs of all positive integer powers, allowing for sign difference:

\[
|x| < 1 \implies |x| > |x|^2 > |x|^3 > |x|^4 > |x|^5 > \cdots,
\]

\[
|x| > 1 \implies |x| < |x|^2 < |x|^3 < |x|^4 < |x|^5 < \cdots.
\]

For even powers we can omit the absolute values since they are redundant. Also, for \( x = 0, 1 \) all powers output the same values, namely 0 and 1 respectively, and for \( x = -1 \) each power outputs \( \pm 1 \), depending upon whether the power is even or odd.

### 2.6.3 Roots

Recall that \( \sqrt[n]{x} = x^{1/n} \), for \( n = 1, 2, 3, \cdots \). Also recall that even roots input and output only nonnegative real numbers, while odd roots can input and output any real number. Now the graph of \( y = x^{1/2} \), i.e., \( y = \sqrt{x} \) is simply the upper piece of the graph \( y^2 = x \), since

\[
y = \sqrt{x} \iff (y^2 = x) \land (y \geq 0).
\]

A less precise statement is \( y = \sqrt{x} \implies y^2 = x \), which shows the graph of \( y = \sqrt{x} \) is a subset of the graph of \( y^2 = x \). What makes this useful is that graphing \( y = \sqrt{x} \) is easily done in reference to \( y = x^2 \), in that \( x \) and \( y \) simply change roles, and we omit those points in the new graph in which \( y < 0 \). This is accomplished in Figure 2.6.3a, page 134.

That \( y = x^2 \) and \( y = \sqrt{x} \) are related by the role-reversals of \( x \) and \( y \), as long as \( x, y \geq 0 \), we can note some trends. On the graph \( y = x^2 \), when \( x > 1 \) (and thus \( y > 1 \)) a small upward
Figure 2.8: Part a shows partial graphs of $y = x^2$, $y = x^4$, $y = x^6$ and $y = x^8$. Part b shows partial graphs of $x, x^3, x^5, x^7$.

Figure 2.9: The solid graph in Figure 2.6.3a above is a partial plot of $f(x) = \sqrt{x}$. It is contained in the graph of $x = y^2$, is simply $y = x^2$ with the (horizontal) $x$ and (vertical) $y$ switching roles. The graph in Figure 2.6.3b is that of $y = \sqrt[3]{x}$, which is exactly the same as $x = y^3$, which we can draw by having $x$ and $y$ switch roles from the graph $y = x^3$ from before.
2.6. BASIC FUNCTIONS OF ONE VARIABLE

change in the value of \( x \) can cause a much faster change in the value of \( y \). This implies that on the graph \( y = \sqrt{x} \) (as part of \( x = y^2 \)), we have large increases in \( x \) do not cause such large increases in \( y \). Indeed, \( y = \sqrt{x} \) is “grows slowly” as \( x \) grows, for \( x > 1 \). When \( x \) is near to 0 (as is \( y \)), on the other hand, \( y = x^2 \) grows very slowly as \( x \) increases, and thus \( y = \sqrt{x} \) grows very quickly as \( x \) increases from 0. Both of these trends in \( y = \sqrt{x} \), i.e., quick growth followed by slow growth, are illustrated the graph in Figure 2.6.3a.

We concentrate on the square root because it is such a common function to encounter in studying calculus. The other roots are also important, but their relative sizes are left for the exercises as straightforward extensions of our analysis of the square root. However, for completeness we include the graph of \( y = \sqrt[3]{x} \), which is exactly the same graph as \( x = y^2 \), and so is readily drawn from the graph \( y = x^3 \), with \( x \) and \( y \) roles reversed. The graph is drawn in Figure 2.6.3b. Odd and even roots differ in that odd roots have domain \( x \in \mathbb{R} \) (and range \( y \in \mathbb{R} \)), where even roots can only input (and output) nonnegative numbers.

2.6.4 Reciprocal Function

The function \( f(x) = 1/x \) occurs in many contexts, alone or as a part of a more complicated function. It simply returns the reciprocal of its input. Note that this function is also an integer power, namely \( f(x) = x^{-1} \), and that it is undefined at \( x = 0 \).

The most interesting features of this function include what occurs when \( x \) is near zero, and what occurs when \( x \) is “large.” First, reciprocals of “small numbers,” such as \( \pm 0.1, \pm 0.01, \pm 0.001 \) and so on, are in fact larger numbers \( \pm 10, \pm 100, \pm 1000 \), respectively, and so on. On the
other hand, reciprocals of “large numbers,” such as $\pm 10, \pm 100, \pm 1000$ are small numbers $\pm 0.1, \pm 0.01, \pm 0.001$. The reciprocal function preserves the sign, meaning a positive input yields a positive output, while a negative input yields a negative output. Because the output’s growth near zero occurs quickly, a graph with $x$ and $y$ using the same scale is difficult to produce showing these trends with a lot of precision. In Figure 2.10a, the trends are shown but with more easily seen inputs and outputs.

The way the output of $y = 1/x$ “blows up” as $x$ nears zero is geometrically described as causing the graph to have a vertical asymptote at $x = 0$. This means that the graph geometrically becomes closer and closer to the vertical line $x = 0$. On the other hand, the way that the output nears 0 as $x$ becomes a large positive number, or a large negative number, is geometrically described as causing the graph to have a two-sided horizontal asymptote $y = 0$ as $x$ grows “large.” By two-sided we mean there is a horizontal asymptote for $x$ growing larger and positive, and the same line for an asymptote for $x$ growing larger and negative. Asymptotes and asymptotic behavior of functions is very important, and will be described in more detail later in the text.

For now, we simply mention that an asymptote is a shape which the graph grows closer and closer, without necessarily touching, for some movement in $x$. For our present example, the graph becomes more “vertical” and increasingly close in shape and behavior to the vertical line $x = 0$ for small $x$, and becomes more “horizontal” and increasingly close to the horizontal line $y = 0$ for large $x$.

### 2.6.5 Absolute Value and Piecewise Defined Functions

The absolute value is naturally the first of the piecewise defined functions one usually studies. It is defined by

$$ |x| = \begin{cases} 
  x, & \text{if } x \geq 0 \\
  -x, & \text{if } x < 0.
\end{cases} \quad (2.63) $$

Thus, $|5| = 5$ while $|-5| = -(-5) = 5$, for two examples. The spirit of (2.63) is that the absolute value function leaves nonnegative numbers unchanged, but changes the sign of negative numbers. The graph is given in Figure 2.10a. It is not as uncommon as one might think to encounter such piece-wise defined functions, which have different rules for the different ranges of inputs. The absolute value function has many interesting interpretations.

- $|x| \geq 0$, with $|x| = 0 \iff x = 0$.
- In general $|a - b|$ is the distance between $a$ and $b$ on the number line. Note that

$$ |a - b| = \begin{cases} 
  a - b, & \text{if } a - b \geq 0 \\
  -(a - b), & \text{if } a - b < 0.
\end{cases} = \begin{cases} 
  a - b, & \text{if } a \geq b \\
  b - a, & \text{if } a < b.
\end{cases} $$

Of course the distance between $a$ and $b$ is clearly $a - b$ if $a \geq b$, and is $b - a$ if $a < b$, as a quick look at the number line illustrates. (Note that the distance is always “right number minus left number.”)

- As a special case, $|x| = |x - 0|$ is the distance between $x$ and 0.
- The value of $|x|$ is sometimes considered to be a measure of the “size” of $x$; when we speak of “large $x$,” we will often write “large $|x|$,” meaning $x$-values that are “large” but may be positive or negative. Of course $|x|$ itself will never be negative.
2.6. BASIC FUNCTIONS OF ONE VARIABLE

- Since \(|x|\) does not distinguish signs, we have
  \[
  |ab| = |a| \cdot |b|, \tag{2.64}
  
  |a/b| = |a|/|b|, \tag{2.65}
  
  |a^n| = |a|^n, \text{ assuming } n \in \{\pm 1, \pm 2, \pm 3, \cdots \}. \tag{2.66}
  
  However, one must be careful. There are many cases for which
  \(|a + b| \neq |a| + |b|, |a - b| \neq |a| - |b|, \text{ and so on. The reader is invited to look at cases where replacing } \neq \text{ with } = \text{ yields true statements or false statements. In fact, what are true are the following:}
  \[
  |a + b| \leq |a| + |b|, \tag{2.67}
  
  |a - b| \geq |a| - |b|. \tag{2.68}
  
  The first, (2.67) is called the triangle inequality (for one dimension), because of how it generalizes into higher dimensions, but it should seem reasonable on its face (again, after examining the various cases where } a, b \text{ are the same sign or different signs).}

2.6.6 Exponential Functions

Here we are interested in functions of the form \(f(x) = a^x\), where \(a > 0\) but \(a \neq 1\). For these functions, the “base” is constant and the “exponent” is changing. An exponential function such as \(2^x\) will behave very differently than a polynomial function \(x^2\), for instance, so it is important to carefully distinguish the two types of functions.

Consider the function \(f(x) = 2^x\). Some of its evaluations are given below:

<table>
<thead>
<tr>
<th>(x)</th>
<th>(2^x)</th>
<th>(x)</th>
<th>(2^x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10</td>
<td>(1/1024=0.0009765625)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>-9</td>
<td>(1/512=0.001953125)</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>-8</td>
<td>(1/256=0.00390625)</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>-7</td>
<td>(1/128=0.0078125)</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>-6</td>
<td>(1/64=0.015625)</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>-5</td>
<td>(1/32=0.03125)</td>
<td>5</td>
<td>32</td>
</tr>
<tr>
<td>-4</td>
<td>(1/16=0.0625)</td>
<td>6</td>
<td>64</td>
</tr>
<tr>
<td>-3</td>
<td>(1/8=0.125)</td>
<td>7</td>
<td>128</td>
</tr>
<tr>
<td>-2</td>
<td>(1/4=0.25)</td>
<td>8</td>
<td>256</td>
</tr>
<tr>
<td>-1</td>
<td>(1/2=0.5)</td>
<td>9</td>
<td>512</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
</tr>
</tbody>
</table>

Note that for \(f(x) = 2^x\), each increase in \(x\) by 1 increases the function by a factor of 2: \(f(x) = 2^x\). The graph thus increases rapidly as \(x\) increases. Similarly, \(f(x)\) decreases by a factor of 2 when \(x\) decreases by 1: \(f(x - 1) = \frac{1}{2}f(x)\). As with polynomial functions, it is often difficult to graph exponential functions for large ranges of \(x\), but it is graphed partially in Figure 2.11a, page 138. Note that the function has a one-sided horizontal asymptote \(y = 0\) as \(x\) gets large but negative, but grows rapidly as \(x\) gets large and positive. This is also reflected in the table of values given above. There is a similar effect, more or less pronounced, for all functions \(f(x) = a^x\) when \(a > 1\).

\(^{48}\)The case \(f(x) = 1^x\) is rather uninteresting, for obvious reasons.
Later we will have a special case of the exponential functions, where the base is the irrational number
\[ e \approx 2.71828182845904523536 \] (2.69)
Thus the graph of \( e^x \) is between those of \( 2^x \) and \( 3^x \), with all of course coinciding at \( x = 0 \).

On the other hand, when we look at exponential functions \( f(x) = a^x \) with \( a \in (0, 1) \), we have the reverse effect. Consider for example \( f(x) = (1/2)^x \). For each increase in \( x \) by 1 we have a halving of the output.

<table>
<thead>
<tr>
<th>( x )</th>
<th>( (\frac{1}{2})^x )</th>
<th>( x )</th>
<th>( (\frac{1}{2})^x )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
</tr>
<tr>
<td>-10</td>
<td>1024</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>-9</td>
<td>512</td>
<td>1</td>
<td>1/2</td>
</tr>
<tr>
<td>-8</td>
<td>256</td>
<td>2</td>
<td>1/4</td>
</tr>
<tr>
<td>-7</td>
<td>128</td>
<td>3</td>
<td>1/8</td>
</tr>
<tr>
<td>-6</td>
<td>64</td>
<td>4</td>
<td>1/16</td>
</tr>
<tr>
<td>-5</td>
<td>32</td>
<td>5</td>
<td>1/32</td>
</tr>
<tr>
<td>-4</td>
<td>16</td>
<td>6</td>
<td>1/64</td>
</tr>
<tr>
<td>-3</td>
<td>8</td>
<td>7</td>
<td>1/128</td>
</tr>
<tr>
<td>-2</td>
<td>4</td>
<td>8</td>
<td>1/256</td>
</tr>
<tr>
<td>-1</td>
<td>2</td>
<td>9</td>
<td>1/512</td>
</tr>
<tr>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
</tr>
</tbody>
</table>

When \( a > 1 \), the function \( f(x) = a^x \) represents what is called exponential growth, referring to how the functions grow as \( x \) increases. For \( 0 < a < 1 \), the function \( f(x) = a^x \) represents...
2.6. BASIC FUNCTIONS OF ONE VARIABLE

Figure 2.12a. Exponential Growths

Figure 2.12b. Exponential Decays

Figure 2.12: Comparison of various exponential functions. Figure 2.12a shows $y = a^x$, where $a > 1$ are all exponential growths, which grow faster for larger $a$ as $x$ increases, and therefore shrink faster as $x$ decreases. Figure 2.12b shows $y = a^x$, where $a \in (0, 1)$, which shrinks faster for smaller $a$ as $x$ increases, and therefore grow faster as $x$ decreases.
exponential decay, referring to how the functions decrease as $x$ increases. It is not difficult to see that $3^x$ increases faster than $2^x$ as $x$ increases, but therefore also decreases faster as $x$ decreases. Similarly $(1/3)^x$ decreases faster than $(1/2)^x$ as $x$ increases, while $(1/3)^x$ increases faster as $x$ decreases. These are reflected in Figure 2.12, page 139.

Eventually we will be especially interested in $e^x$, where $e \approx 2.71828$, so the graph of $y = e^x$ will lie between the graphs of $y = 2^x$ and $y = 3^x$.

### 2.6.7 Logarithmic Functions

Logarithmic, or “log” functions also deal with powers but from the opposite direction as the exponential functions. When we write $\log_a x$, usually read, “the logarithm of $x$ in base $a$,” or simply, “log, base $a$, of $x$,” we are denoting that power of $a$ which would yield $x$. Thus,

$$
\log_2 1 = 0, \quad \log_2 2 = 1, \quad \log_2 4 = 2, \quad \log_2 8 = 3, \quad \log_2 16 = 4,
$$

and so on. To be more precise,

$$
\log_a x = y \iff a^y = x. \tag{2.70}
$$

We assume as with exponentials that $a > 0$ but $a \neq 1$. The most common logarithms encountered are the base-10 logarithm and the base-$e$ logarithm. These are so common that they have their own shorthand notations:

$$
\log_{10} x = \log x \tag{2.71}
$$

$$
\log_e x = \ln x. \tag{2.72}
$$

These are also known as the common logarithm and the natural logarithm, respectively. The common logarithm is more present in algebraic settings, while the natural logarithm is much more important in calculus.

The behaviors of these functions can be seen as we look at some input-output pairs. For $\log_2 x$, the simplest inputs are powers of 2. We see that to increase the output by 1, we must double the input (i.e., inserting another factor of 2); to decrease the output by 1 we must halve the input. With $\log x = \log_{10} x$, we must increase or decrease the number of factors of 10:

<table>
<thead>
<tr>
<th>$x$</th>
<th>$\log_2 \ x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$1/16$</td>
<td>$-4$</td>
</tr>
<tr>
<td>$1/8$</td>
<td>$-3$</td>
</tr>
<tr>
<td>$1/4$</td>
<td>$-2$</td>
</tr>
<tr>
<td>$1/2$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$1$</td>
<td>$0$</td>
</tr>
<tr>
<td>$2$</td>
<td>$1$</td>
</tr>
<tr>
<td>$4$</td>
<td>$2$</td>
</tr>
<tr>
<td>$8$</td>
<td>$3$</td>
</tr>
<tr>
<td>$16$</td>
<td>$4$</td>
</tr>
<tr>
<td>$32$</td>
<td>$5$</td>
</tr>
<tr>
<td>$64$</td>
<td>$6$</td>
</tr>
<tr>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$x$</th>
<th>$\log \ x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$1/10,000$</td>
<td>$-4$</td>
</tr>
<tr>
<td>$1/1000$</td>
<td>$-3$</td>
</tr>
<tr>
<td>$1/100$</td>
<td>$-2$</td>
</tr>
<tr>
<td>$1/10$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$1$</td>
<td>$0$</td>
</tr>
<tr>
<td>$10$</td>
<td>$1$</td>
</tr>
<tr>
<td>$100$</td>
<td>$2$</td>
</tr>
<tr>
<td>$1000$</td>
<td>$3$</td>
</tr>
<tr>
<td>$10,000$</td>
<td>$4$</td>
</tr>
<tr>
<td>$100,000$</td>
<td>$5$</td>
</tr>
<tr>
<td>$1,000,000$</td>
<td>$6$</td>
</tr>
<tr>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
</tbody>
</table>

The graphs of three logarithm functions are given in Figure 2.13, page 141. Since $y = \log_a x \iff x = a^y$, we see that $x$ and $y$ have traded roles from their roles with the exponential
functions \( y = a^x \). Thus the graph of \( y = \log_a x \) will have domain \( x > 0 \) and range \( y \in \mathbb{R} \), as well as a vertical asymptote at \( x = 0 \). (In contrast, the exponential function \( y = a^x \) has domain \( x \in \mathbb{R} \), range \( y > 0 \), and a horizontal asymptote \( y = 0 \).

Logarithms with bases \( a \in (0, 1) \) have somewhat different behavior, which will be developed in the exercises. They are rarely encountered in practice, though the reader should be aware of them.

Note that so far we have seen three different kinds of “growth”, as \( x > 0 \) increases. In order from fastest to slowest we have encountered, we have exponential functions \( a^x \) with \( a > 1 \), polynomial functions \( x^n \) with \( n \geq 1 \), and logarithm functions \( \log_a x \) with \( a > 1 \). In between the polynomial and logarithm functions are the roots \( \sqrt[n]{x} \), i.e., \( x^{1/n} \) for \( n = 2, 3, 4, \cdots \).

### 2.6.8 Sine and Cosine Functions

All trigonometric functions trace their ancestry to the sine and cosine functions, which themselves can be characterized simply as describing the variations in, respectively, vertical and horizontal positions of a moving point on the unit circle, \( x^2 + y^2 = 1 \), as the point travels with uniform speed along the circle counter-clockwise. The position of the point is determined by the angle \( \theta \), measuring the angle traveled from the starting point on the positive \( x \)-axis, i.e., from \((1, 0)\). This is shown in Figure 2.14, page 142.

While they may not seem intuitive at first, in fact radians are more “natural” than the artificial measure of degrees, where convention put a whole rotation at 360° with little explanation. “Radians” are actually measures of arc-length where the units are radius lengths. Once around a circle is \( 2\pi \) radius lengths, and so in radians we have \( 360° = 2\pi \) (rad). In order to avoid...
confusion, we sometimes write “(rad)” but in fact radians are “dimensionless,” since they are defined as a length (of arc) divided by another length (of the radius). For that reason, instead we usually simply write $360^\circ = 2\pi$.

Regardless of the units used for the angle $\theta$, when we rotate by that angle—measured against the positive $x$-axis—the horizontal ($x$) position changes with that angle, and we call that position the cosine of the angle. Similarly, the vertical ($y$) position changes with the angle, and we call that position the sine of the angle. For the general circle we have

$$\cos \theta = \frac{x}{r}, \quad (2.73)$$
$$\sin \theta = \frac{y}{r}. \quad (2.74)$$

For the unit circle, where $r = 1$, we have

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

These both change continuously, and in fact smoothly, with the angle $\theta$, and so their graphs versus $\theta$ are the familiar “sinusoidal” curves seen in Figure 2.15, page 143.

These functions are important in numerous applications of mathematics, as we will see as the text unfolds. They are also interesting in themselves for a number of reasons. First, they are periodic. A function $f$ is called periodic if there exists a number $P > 0$ so that $f(x + P) = f(x)$ for all $x$ in the domain. The smallest such number $P > 0$ is called the period, or cycle of the function. By the circular nature of the definitions of both sine and cosine functions, we see immediately that both functions have period $2\pi$ or $360^\circ$, depending upon whether we measure our angle $\theta$ in radians or degrees:

$$\cos(\theta + 2\pi) = \cos \theta, \quad (2.75)$$
$$\sin(\theta + 2\pi) = \sin \theta. \quad (2.76)$$
Figure 2.15: Shown are graphs of $\sin \theta$ and $\cos \theta$, where $\theta$ is measured in both degrees and radians. These follow quickly from the unit circle, $\cos \theta$ being the $x$-coordinate and $\sin \theta$ being the $y$-coordinate, as increasing $\theta$ causes the coordinates of the point where the terminal ray of the angle $\theta$ intersects the unit circle to oscillate.
By the nature of the unit circle these functions are also bounded, from both above and below:

\[-1 \leq \cos \theta \leq 1, \quad (2.77)\]
\[-1 \leq \sin \theta \leq 1, \quad (2.78)\]

These can be summarized as \(|\cos \theta| \leq 1\) and \(|\sin \theta| \leq 1\).

The values for sine and cosine are known exactly for several angles \(\theta\). Some are obvious from the unit circle while others come from geometry, and perhaps clearer using the general circle. Recall from geometry the proportions for sides of “30°-60°-90°” (1 : \(\sqrt{3}\) : 2) and “45°-45°-90°” (1 : 1 : \(\sqrt{2}\)). The relevant angles are drawn below in a general circle of radius 2. Note that the \(x\) and \(y\) coordinates where the terminal side intersects the circle is give in each case.

Recalling that \(\cos \theta = x/r\) and \(\sin \theta = y/r\), we then get

\[
\begin{align*}
\cos 30^\circ &= \frac{\sqrt{3}}{2}, & \cos 45^\circ &= \frac{\sqrt{2}}{2}, & \cos 60^\circ &= \frac{1}{2}; \\
\sin 30^\circ &= \frac{1}{2}, & \sin 45^\circ &= \frac{\sqrt{2}}{2}, & \sin 60^\circ &= \frac{\sqrt{3}}{2}.
\end{align*}
\]

In fact we will normally write \(\sin 45^\circ, \cos 45^\circ = \frac{1}{\sqrt{2}}\) rather than the forms above for these. Both are common.

### 2.6.9 Arcsine and Arccosine Functions

There are many times when we need to find an angle \(\theta\) whose sine or cosine is given. We can not do this definitively without having some further information, like a predetermined range we require for \(\theta\). Indeed it is not difficult to see that \(\sin 30^\circ = \sin 150^\circ\) and \(\cos 30^\circ = \cos(-30^\circ) = \cos 330^\circ\).
2.6. BASIC FUNCTIONS OF ONE VARIABLE

Exercises

1. Sketch a rough graph of \( y = \sqrt[3]{x} \) based upon the graph of \( y = x^3 \). Note that the domain is \( x \in \mathbb{R} \).

2. Using one pair of axes, sketch graphs of \( y = \sqrt{x}, \ y = \sqrt[3]{x}, \ \text{and} \ y = \sqrt[4]{x} \), displaying their shapes and relative positions for relevant ranges of \( x \).

3. Using one pair of axes, sketch graphs of \( y = \sqrt{x}, \ y = \sqrt[3]{x}, \ \text{and} \ y = \sqrt[4]{x} \), displaying their shapes and relative positions for relevant ranges of \( x \).

4. Show that any two points on the graph of a function of the form \( f(x) = mx + b \) will have slope \( m \) by computing \( \frac{f(x + \Delta x) - f(x)}{\Delta x} \).

5. Sketch a graph of \( y = \log_{1/2} x \). To do so, begin by writing a table as we did for \( \log_2 x \).
CHAPTER 2. REAL NUMBERS, ALGEBRA AND FUNCTIONS (OPTIONAL)

2.7 Functions-III: Simple Transformations of Functions

In this section we look at the effects of simple transformations of functions. In this context, by “transformations” we mean simple changes to the formula for a given function (assuming the function can be given by a formula), and the resultant change in the graphs. It should be pointed out that when we transform a function, we get a new function. Thus, we will compare the differences in formulas of these functions, and link these to the changes in their graphs. But first we will introduce a method of diagramming the actions of complex functions, similar to flow charting in computer science.

2.7.1 Diagramming Functions

Here we introduce a simple method of diagramming function actions which will be helpful later as we “parse” these functions. Part of the method involves using empty parentheses in place of the dummy variable (most often $x$). For instance, we can equally describe the action of a function $f(x) = x^2 + 1$ in either of the following ways:

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

The second is rarely found in mathematical texts, but we employ it here because it has the advantage that the action of $f$ is not mentally linked to the variable $x$. This “empty parentheses” method of describing the action of $f$ makes computations such as the following more easily understood:

$$f(x + h) = (x + h)^2 + 1 = x^2 + 2hx + h^2 + 1,$$
$$f(x^2) = (x^2)^2 + 1 = x^4 + 1,$$
$$f(\sqrt{y}) = (\sqrt{y})^2 + 1 = y + 1.$$

This also make “diagramming” $f$ simpler. Below we first diagram $f$ by showing its action on $x$ in $f(x)$. We then perform the above computations but without simplification.

The point of this is that in each case, the actual function (“the box”) we are observing is the same, namely squaring the input and then adding 1. Above we see it acting upon different inputs, but it is the same function. Each step is represented in the abstract, described as action.
2.7. FUNCTIONS-III: SIMPLE TRANSFORMATIONS OF FUNCTIONS

Upon empty parentheses. To help trace the action we may find it useful to write the result after each step, as above. Note that we could trace a number through the function as well, to find out, for instance $f(-5) = (-5)^2 + 1 = 25 + 1 = 26$. In fact this function can clearly process any number from $\mathbb{R}$ without difficulty, so we will claim $\mathbb{R}$ as its domain. (The range, $[1, \infty)$ takes a little more intuition.) Recall the the domain is the set of inputs that the function can process, and the range is the set of possible outputs.

2.7.2 Shift Transformations

Example 2.7.1 Consider the following functions:

$$f(x) = \sqrt{x},$$
$$g(x) = \sqrt{x - 2},$$
$$h(x) = \sqrt{x - 2}.$$

If we were to make tables of points $(x, y)$ to plot each of these, we would find ourselves with the three graphs in Figure 2.16. A comparison of corresponding convenient points we can plot for these functions illustrates the differences in the graphs.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
<th>$x$</th>
<th>$y$</th>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>9</td>
<td>1</td>
<td>11</td>
<td>3</td>
</tr>
</tbody>
</table>

When we diagram these functions, we notice where the transforms take place with respect to the “main” function, $\sqrt{\cdot}$, shaded gray below for emphasis.
The first function diagrammed is the basic, “main function” \( y = \sqrt{x} \).

In the next one, we see that we have decreased the output of the original function by 2. Since the \( y \)-coordinate on the graph represents the output of the graphed function, this means that the graph of \( y = \sqrt{x} - 2 \) is the same as that of \( y = \sqrt{x} \), except shifted vertically downward by 2 units.

In the third function \( y = \sqrt{x - 2} \) is a transform of the input of the original function \( y = \sqrt{x} \). Doing so means that inputs of \( y = \sqrt{x - 2} \) must be greater than the inputs of \( y = \sqrt{x} \) by 2 to guarantee the same output.

The following are true regarding transformations of functions:

- When we transform the input of a function, some horizontal aspect of the graph is changed.
- When we transform the output of a function, some vertical aspect of the graph is changed.

When the graph is moved vertically or horizontally, it is referred to as a shift or translation. These can both happen simultaneously. As a general rule, when we compare the graph of \( y = f(x) \) with that of

\[
y = f(x - h) + k,
\]

we see that the latter has the same shape, but translated by \( h \) horizontally (left if \( h < 0 \), right if \( h > 0 \)) and by \( k \) vertically (downward if \( h < 0 \), and upward if \( h > 0 \)).
**Example 2.7.2** Consider the graph of \( y = \sqrt{x + 3} - 4 \). Picking \( x \)-values so that the square root has simple outputs as before, we could construct a table and get a reasonable graph. We could also parse this function in a diagram, and notice that we decrease the values of \( x \) by 3 to have the square root output the same values as \( \sqrt{x} \), and then we decrease the output by 4 before we graph the function.

\[
\begin{array}{c|c}
 x & y \\
-3 & \sqrt{0} - 4 = 0 \\
-2 & \sqrt{1} - 4 = -3 \\
1 & \sqrt{4} - 4 = -2 \\
6 & \sqrt{9} - 4 = -1 \\
13 & \sqrt{16} - 4 = 0 \\
\end{array}
\]

When we diagram this function \( y = \sqrt{x + 3} - 4 \), the step \((x) + 3\) has the effect of shifting the function to the left 3 units because we can choose \( x \)-values which are 3 less than if that step is removed, in order to input the “main function” \( \sqrt{(x)} \) with the same values as before without that step. The last step \((x) - 4\) means that the final output (graphed on the vertical scale) is 4 less than if that step were removed. This function is of the form \( y = \sqrt{x - h} + k \), if we have \( h = -3 \) and \( k = -4 \):

\[
y = \sqrt{x - (-3)} - 4.
\]

According to the discussion regarding (2.79), page 148, the graph of this function will indeed be shifted \( h = -3 \) horizontally (3 to the left), and \( k = -4 \) vertically (4 downward).

### 2.7.3 Reflection Transformations

The next transformations of a function \( f(x) \) we consider are of the forms \(-f(x)\) and \( f(-x)\). For example, consider the functions \( y = -\sqrt{x} \) and \( y = \sqrt{-x} \).
Compared to the function \( y = \sqrt{x} \), we see that the first function \( y = -\sqrt{x} \) diagrammed and graphed above has a transformation occurring at the output of the "main" function \( \sqrt{x} \), and so we expect some vertical aspect to be changed. By contrast, the second function \( y = \sqrt{-x} \) has a transformation at the input of the \( \sqrt{\cdot} \) step, and so we expect some horizontal aspect to be different from that of \( y = \sqrt{x} \). We also see this in the graphs. (The reader should check some of the points plotted in the equations.)

As a general rule, when we multiply a function by -1, the graph is reflected vertically, through the \( x \)-axis, or "revolved" around the \( x \)-axis.\(^{49}\) If instead we replace all instances of \( x \) in a function with \( -x \), this reflects the graph horizontally, through the \( y \)-axis, i.e., "rotates" the graph around the \( y \)-axis. If we look at any two points \((x, y)\) and then consider \((-x, y)\), we see the difference is a vertical rotation; if we compare \((x, y)\) with \((-x, y)\), we see a horizontal rotation (around the \( y \)-axis) to be the change. Thus, comparing \( y = f(x) \) to

\[
\begin{align*}
y = -f(x) & : \quad \text{flipped vertically around } x\text{-axis (compared to } y = f(x)) , \\
y = f(-x) & : \quad \text{flipped horizontally around } y\text{-axis (compared to } y = f(x)) .
\end{align*}
\]

\(^{49}\)A mirror-type reflection, for our purposes, will be the same as a 180° rotation. At the expense of reading more colloquially, we will often use the term "flipped" instead of the more accepted mathematical term, "reflected."
2.7. FUNCTIONS-III: SIMPLE TRANSFORMATIONS OF FUNCTIONS

2.7.4 Rescaling: Contractions and Dilations

Now we consider comparing \( f(x) \) to \( \alpha \cdot f(x) \) and \( f(\beta x) \). Again these transformations of \( f(x) \) are, respectively, transforming the output and input of the function, which will affect the vertical and horizontal aspects of the graph, respectively. The trigonometric functions illustrate these phenomena well.

**Example 2.7.3** Consider the functions \( \sin x \), \( 3 \sin x \) and \( \sin 3x \). (This last one is read \( \sin(3x) \).)

Recall that one cycle of the sine function can be drawn by connecting the convenient points \((0,0), (\pi/2,1), (\pi,0), (3\pi/2,-1) \) and \((2\pi,0)\). With this in mind we consider the corresponding points on each of the functions above.

<table>
<thead>
<tr>
<th>( y = \sin x )</th>
<th>( y = 3 \sin x )</th>
<th>( y = \sin 3x )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x )</td>
<td>( y )</td>
<td>( x )</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \pi/2 )</td>
<td>1</td>
<td>( \pi/2 )</td>
</tr>
<tr>
<td>( \pi )</td>
<td>0</td>
<td>( \pi )</td>
</tr>
<tr>
<td>( 3\pi/2 )</td>
<td>-1</td>
<td>( 3\pi/2 )</td>
</tr>
<tr>
<td>( 2\pi )</td>
<td>0</td>
<td>( 2\pi )</td>
</tr>
</tbody>
</table>
We see that multiplying the output by 3 resulted in a vertically stretched (dilated) graph (by a factor of 3), while multiplying the input by 3 horizontally shrunk (contracted) the graph (by a factor of 3).

Note that the effect of multiplying the output of a function by some $\alpha > 1$ does indeed stretch it vertically, while multiplying the input by that same $\alpha$ will in fact “shrink” the function horizontally. The reason for this is that multiplying the input by $\alpha > 1$, in effect, makes the input change more quickly. In the above example, for the first two functions $y = \sin x$ and $y = 3\sin x$, one cycle occurs in the interval $x \in [0, 2\pi]$, while for $y = \sin 3x$ that cycle occurs for $0 \leq 3x \leq 2\pi$, i.e., $0 \leq x \leq \frac{2\pi}{3}$, or $x \in [0, 2\pi/3]$. Put another way, $y = \sin 3x$ passes through three cycles on the same interval (for $x$) in which $y = \sin x$ passes through just one.\(^{50}\)

It should also be pointed out that the function $y = \frac{1}{2}\sin x$ would look like $y = \sin x$ except contracted by a factor of $1/2$. Similarly the function $y = \sin \frac{1}{2}x$ would look like $y = \sin x$ except dilated horizontally by a factor of 2.

### 2.7.5 Multiple Simple Transformations of Basic Functions

Before considering actual examples of multiple transformations, we make the following observations:

1. Transformations which change the input of the main function will change the graph in some horizontal aspect, while transformations which change the output of the function will change some vertical aspect of the graph.

2. The transformations discussed here—namely shifts, reflections, contractions and dilations—do not radically change the actual shape of the graph. For instance, a parabola (or half-parabola) transforms to another parabola, a sinusoidal curve transforms to another sinusoidal curve, if a function’s curve is “vee-shaped,” (such as $y = |x|$) transforms to another “vee-shaped” curve.

---

\(^{50}\)Another way of comparing the two transformations of Example 2.7.3 is to rewrite the functions

\[
\begin{align*}
y = \sin x, \\
y/3 = \sin x, \\
y = \sin 3x.
\end{align*}
\]

The first being our “main” function, we see we get the same numbers on both sides of the second and third equations if we triple $y$ in the second, and divide $x$ by 3 in the third. Also, when we put the vertical transformation with the $y$-term and the horizontal transformation with the $x$-term, we see a consistency: replacing $y$ by $y/3$ stretches the graph vertically by 3, while replacing $x$ by $3x$ shrinks it horizontally by a factor of $1/3$. Because we usually solve for $y$, the transformations of the $y$-variable are algebraically inverted to put them to the other side of the equation.
3. We can trace multiple simple transformations to predict what will be the graph of the final function.

**Example 2.7.4** Consider the function \( f(x) = 2|x - 3| - 5 \). The “main” function here is \( y = |x| \), which is graphed in gray below but also in Figure 2.10b, page 135.

![Diagram showing transformations](image)

One could graph the new function in stages, as it emerges from each transformation, but this is awkward and confusing as the graphs can get quite cluttered. However, as we see the shapes of these two graphs above are not dissimilar. One could therefore plot this new function by locating the new vertex, and plotting (say) two other points (one left and one right of the vertex) to completely determine the graph of \( f(x) \) above.

It is important to notice the order of the transformations. For instance, a vertical stretching followed by a vertical shift is going to be very different from the same shift followed by the same stretching. For instance, below we have the first function representing a stretch and then a shift, while the second reverses these:

\[
\begin{align*}
y &= 2|x| + 1 \quad \text{(stretch vertically by 2, shift up by 1),} \\
y &= 2(|x| + 1) \quad \text{(shift up by 1, stretch vertically by 2).}
\end{align*}
\]

The second can also be written \( y = 2|x| + 2 \), which is a shift of one higher than the previous function \( y = 2|x| + 1 \), so they are clearly not the same.

It is not uncommon for there to be two different ways to write the same function, such as

\[
y = 2(|x| + 1) = 2|x| + 2.
\]

In such a case there is an algebraic argument that two different paths of transformation of the “main function” (in this case \(|x|\)) lead to the same final function, but that does not diminish the...
fact that for a particular function we almost always must do the written transformations in the appropriate order. Fortunately, transformations can be parsed in the same order in which the function would have been evaluated, which can be read (carefully) from the formula or from the related diagram.

Example 2.7.5 Graph \( y = -3^{x+2} + 4 \).

**Solution:** First we point out that this should be read \( y = -[3^{x+2}] + 4 \).

It is worth noting and plotting the \( y \)-intercept, \((0, -5)\). It takes a bit more then we have developed so far to find the \( x \)-intercept, namely where \( x = \log_3 4 - 2 \approx -0.7381404929 \). Later in the text, when we develop logarithms in more algebraic detail, this will be a routine computation.

2.7.6 Reflection Around Other Vertical Lines

Recall that any transform which occurs before the “main function” results in a transform of the horizontal aspect of the graph, since the horizontal variable represents the input of the function. We already established that, for instance, replacing \( x \) by \( x - 3 \) moves the function’s graph 3 units to the right. But what if we replace \( x \) by \( 3 - x \)? Note that

\[ 3 - x = -(x - 3), \]
so replacing $x - 3$ with $3 - x$ is the same as negating $x - 3$. Consider the effects of this change on the value of these quantities:

<table>
<thead>
<tr>
<th>$x - 3$</th>
<th>$x$</th>
<th>$3 - x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-3$</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>$-2$</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$-1$</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>$-1$</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>$-2$</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>$-3$</td>
</tr>
</tbody>
</table>

Note how the same $x$-value returns values for $x - 3$ and $3 - x$ which are opposites. However, for graphing purposes, this is not the simplest interpretation.

In fact it can be more useful to look at this from the opposite direction: if we want $x - 3 = k$ we need $x = 3 + k$, where if we want $-(x - 3) = k$ we need $x = 3 - k$. Both are the same distance from $x = 3$ but in opposite directions.

\[
\begin{align*}
    x - 3 &= k &\implies& & x &= 3 + k \\
    3 - x &= k &\implies& & x &= 3 - k.
\end{align*}
\]

For example, $x - 3 = 2 \implies x = 5$, while $3 - x = 2 \implies x = 1$. On a number line, the point given in the first of these equations is 2 units right of $x = 3$, while the second’s is 2 units to the left of $x = 3$. Note that this is analogous to flipping a graph around the $y$-axis, i.e., the line $x = 0$, which occurred whenever we replaced $x$ with $-x$. Here we are replacing $x - 3$ with $3 - x = -(x - 3)$. For graphing functions, the geometric upshot of this is that replacing $x - 3$ with $3 - x = -(x - 3)$ reflects the graph through the line $x = 3$.

**Example 2.7.6** Graph the function $f(x) = \log_2(3 - x)$.

**Solution:** This is a variation of the function $\log_2 x$, though it may not be entirely clear from the expression for $f(x)$ how the transform on the horizontal variable translates to a geometric transform, until we rewrite the function as

\[
f(x) = \log_2[-(x - 3)].
\]

If we parse this, we see the function $\log_2 x$ being first shifted 3 to the right, and then “flipped” horizontally around an axis which is the line $x = 3$. 

With the above example and previous examples, we can note the following about horizontal flips:

- Comparing \( y = f(x) \) to \( y = f(-x) \), the latter’s graph is flipped horizontally around the \( y \)-axis, i.e., around the line \( x = 0 \).

- Comparing \( y = f(x) \) to \( y = f(-(x-h)) = f(h-x) \), the latter is moved horizontally by \( h \), and then flipped horizontally around the line \( x = h \).

If we wish to take a graph of \( y = f(x) \) and construct a new function which is the original but reflected it through a line \( y = k \), that is a slightly more difficult analysis requiring some more sophisticated geometric argument (see exercises).
2.7. FUNCTIONS-III: SIMPLE TRANSFORMATIONS OF FUNCTIONS

**Exercises**

1.

2. Here we construct a function which is the same as \( y = f(x) \), but reflected vertically through a line \( y = k \). To do this, we first attempt it with a relatively simple function, and then argue in the abstract.

(a) Let \( f(x) = \sqrt{x} \) but defined only for \( x \in [0, 4] \). Draw this function.

(b) Draw the graph of this function as reflected through the line \( y = 3 \).

(c) Label the ordered pairs on the graph at each of the values \( x = 0, 1, 4 \). (There should be six points labeled.)

(d) Measure the distances between the corresponding points labelled above, for each case separately: \( x = 0 \), \( x = 1 \) and \( x = 4 \). (There should be three distances.)

(e) Find an equation of the new curve.

(f) Conjecture on the form of an equation for a new function which is derived from the graph of an abstract function \( y = f(x) \) but reflected through a line \( y = k \).

(g) Test your hypothesis by considering \( y = 2^x \) but reflected around the lines

(i) \( y = 2 \),

(ii) \( y = -2 \).

Graph these separately, along with their reflections and label with your conjectured form of the reflected functions. See if your conjectured formulas match their respective new functions.
### 2.8 Functions-IV: Combinations of Functions

There are infinitely many ways in which we can combine two or more functions: adding, subtracting, multiplying, dividing, and “function composition” where we take the output of one function and feed it into another, are just the most common examples. Indeed, fairly simple polynomials—even of second degree—can require a much more sophisticated analysis to determine their general behavior and therefore graphs if they consist of more than one term of positive degree.

In this section we consider a few of the various methods of combining functions in which we can have some success in predicting the behavior of the resulting function.

However, the reader should be aware that the interaction of two or more functions to produce a new function can get complicated quickly, and we may sometimes have to throw up our hands and attack the function “from the outside.” There are two general approaches for doing so that
we will develop in subsequent chapters. The first is to probe the resulting function by strategically sampling points, and we do that to some extent in Section 3.3. The other is through Calculus, with techniques beginning in Section 3.4 and building through Chapter 11. As these are more involved, it is still useful to predict the outcome of combined functions to the degree we can, and that is the subject of this section.

When we combine functions, in a sense we create a new “machine” out of the old machinery, but in sometimes complicated ways. It is therefore useful to look at our new function as a new machine crafted together from these simpler functions.

We will be particularly interested in how periodic functions interact with other functions, though there are numerous other examples of interest.

### 2.8.1 Adding Functions

The simplest case to consider is when one of the functions is a constant. We dealt with this case in Section 2.7, where we eventually found that the graph of the function \( h(x) = f(x) + k \) is exactly the same as the graph of \( f(x) \), except it has been displaced vertically by \( k \).

The next “simplest” case is adding two functions. In fact this is often not trivial, but there is some insight into realizing it as a generalization of the case of adding a constant.

**Example 2.8.1** Consider the graphs of \( y = \sin x \) and \( y = \sin x + 3 \). Upon graphing both we would see that the latter is the same as the former, but raised by 3. There are two ways of thinking about this: as the function \( y = \sin x \) “perturbed” vertically by the function \( y = 3 \); or as the function \( y = 3 \) perturbed vertically by the function \( y = \sin x \). This is illustrated in Figure 2.17, page 158.

In Figure 2.17 we introduce two new diagramming devices, namely the fixed parameter 3 signified by a circle, and the “diamond,” for binary operations—such as addition—which are those which have two inputs and one output. In part because addition is commutative, we can have those two interpretations of the combination \( y = \sin x + 3 \) of two functions: as the function \( \sin x \) “raised by three”; and as the function 3 raised at each point by the height \( \sin x \) (dropped if this is negative).

Later in the text we will emphasize the general structure of functions such as this, here meaning that it is foremost a sum, albeit a sum of two other functions. This is a type of “outside-in” approach to functions, where we look at the outward, overall structure first, and then delve into the component parts. It is an important and necessary approach to functions, particularly necessary when we develop derivative techniques in Chapter 4.

The result from adding functions easily gets more complicated, particularly when neither is constant. We can still treat either function as a “baseline” from which the other adds its own height, but by itself that analysis is not always sufficient to (easily) get a very accurate picture of the graph, though there are many worthwhile cases where we can get a general trend.

**Example 2.8.2** Consider the function \( f(x) = \frac{1}{2}x + \sin x \). Since \( \sin x \in [-1, 1] \), the unbounded function \( y = \frac{1}{2}x \) will “carry along” the sine function \( y = \sin x \). Alternatively, the sine function adds some wiggle to the function \( x \). We see how this function is constructed from the function \( x \) and the function \( \sin x \), as well as its graph, in Figure ??.

It is easy to misinterpret the graph of \( y = \frac{1}{2}x + \sin x \). At first glance it appears that we simply rotated the sine curve along the line \( y = \frac{1}{2}x \), which slants at an angle of \( \tan^{-1} \frac{1}{2} \approx 26.6^\circ \) counterclockwise from the horizontal. This is not the case. First note that for many parts of the new graph, the function seems to almost “straighten out” (such as near the origin), which is not the behavior of the sine curve at all. Also interesting is that the apparent “peaks and troughs” actually occur at different places for \( \sin x \) than for \( \frac{1}{2}x + \sin x \), and that difference is
CHAPTER 2. REAL NUMBERS, ALGEBRA AND FUNCTIONS (OPTIONAL)

Figure 2.18: Illustration of the structure of the function $f(x) = \frac{1}{2}x + \sin x$ (black), along with its graph and those of its component functions $\frac{1}{2}x$ (gray) and $\sin x$ (dashed). Note the shift of the peaks and troughs: dotted lines show the positions of the peaks and troughs of the function $f(x) = \frac{1}{2}x + \sin x$, showing that they occur at staggered positions measured along the $x$-axis, unlike those of the simple sine curve, which occur at uniform intervals. The reader can check the accuracy of the dotted lines by using horizontal straight edges to detect the peaks and troughs indicated by the dashed lines. See Example 2.8.2.

not accounted for by a rotation, which would leave the peaks and troughs differing by a fixed amount. For the sine curve $y = \sin x$, peaks occur at $x = \frac{\pi}{2} + n \cdot 2\pi$ (where $n \in \{0, \pm1, \pm2, \pm3, \cdots\}$), and are thus evenly spaced in the sense that there is always a difference of $\pi$ between a peak and its closest troughs. This even spacing does not occur with $y = \frac{1}{2}x + \sin x$, as careful scrutiny with a ruler would indicate. With the calculus of Chapter 4 we will be able to calculate the positions of peaks and troughs, and see that the peaks of the function $y = \frac{1}{2}x + \sin x$ occur at $x = \frac{2\pi}{3} + n \cdot 2\pi$, and troughs at $x = \frac{4\pi}{3} + n \cdot 2\pi$, so a peak will have one trough $4\pi/3$ to its left, and another $2\pi/3$ to its right, so the peaks and troughs are not evenly spaced between each other.

We will see how to derive the positions of these “local extrema” (peaks and troughs) once we are well into the actual calculus of Chapter 4. Even without these techniques, much of the behavior of this function $f(x) = \frac{1}{2}x + \sin x$ is predictable: that $y = \frac{1}{2}x$ will “dominate” the function $y = \sin x$, pulling it upwards as $x$ grows, while $y = \sin x$ will contribute some “wiggle” to the overall graph.

Example 2.8.3 Consider $f(x) = e^x + \cos x$, and $g(x) = x^2 + 2\cos 2x$. (Note first that $\cos x$ is very similar to the sine function, though its behavior at zero is a bit different. See Figure 2.15, page 143.)

For the first function $f(x)$, $\sin x$ dominates for $x$ much less than zero where $e^x$ is nearly constant, while $e^x$ dominates for $x$ much greater than zero. For the function $g(x)$, for “small” $x$ it is the $\sin x$ term that dominates, but the $x^2$ takes over when $x$ is large. This is illustrated in Figure 2.19.

Eventually we need calculus or other methods—sometimes graphing calculators or software suffice—in order to really acquire an accurate picture of the nature of a function. However,
even algebraic methods can shed some light on apparently complicated functions. In the next example, alternative methods of writing the function spawn alternative insights.

**Example 2.8.4** Consider the functions $f(x) = x^2$ and $g(x) = 2x + 1$, and suppose we wish to add these to get a new function $h(x) = f(x) + g(x) = x^2 + 2x + 1$. Here the graph of $f(x)$ is a parabola, while that of $g(x)$ is a line.

In this form it is not at all obvious that the graph of $h(x)$ will itself be parabolic, and indeed is just shifted from the simpler parabola which is $f(x)$. Indeed, some simple algebraic analysis explains the shape of our new function:

$$h(x) = x^2 + 2x + 1 = (x + 1)^2,$$

and so this is like the parabola $y = x^2$ except shifted horizontally to the left by one, as explained in Section 2.7.
CHAPTER 2. REAL NUMBERS, ALGEBRA AND FUNCTIONS (OPTIONAL)

Figure 2.20: The functions \( f(x) = x^2 \) (gray), and \( g(x) = 2x + 1 \) (dashed), are added and the result is \( h(x) = x^2 + 2x + 1 \) (in black). That this should be a parabola with the same shape as \( f(x) \) is not obvious from the graphs of \( f(x) \) and \( g(x) \). However, from simple algebra we see that \( h(x) = (x + 1)^2 \). The first diagram shows a construction of \( f(x) + g(x) = x^2 + (2x + 1) \), while the second diagram (far right) shows the equivalent function \( (x + 1)^2 \), which is indeed a simple horizontal shift of \( y = x^2 \). This illustrates the importance of algebraic methods for detecting equivalent and simplified versions of a function.

When we look at the “machinery” of the function \( (x + 1)^2 \), this phenomenon of equivalent forms of functions is well illustrated. See Figure 2.20, page 162.\(^{51}\)

2.8.2 Multiplying Functions

There are many interesting phenomena which occur when we multiply functions. Note that in doing so, we are altering the final output, which of course deals with the vertical aspects of its graph. Electronics provides the elegant terms “amplifying” or “amplitude modulating” a function, which occurs when we take a function and multiply it by another.

Multiplying a function by a positive constant will rescale the graph vertically, and so multiplying by a positive function will give a dynamic rescaling of the function. Multiplying a function by a negative constant will turn the graph upside-down and perhaps rescale it, and so multiplying by a negative function will also “flip” the function vertically, and rescale dynamically.

Example 2.8.5 Consider the function \( f(x) = 4 \sin x \). This will be similar to the sine function except that each output will of course be four times that of \( \sin x \). This has the effect of rescaling the graph of the function \( \sin x \) vertically by a factor of four. Note that this further has the effect of making differences in height between the corresponding outputs four times as large; if the original graph \( y = \sin x \) contained points \((x_1, y_1)\) and \((x_2, y_2)\), then the new graph will contain points \((x_1, 4y_1)\), \((x_2, 4y_2)\), and \(4y_2 - 4y_1 = 4(y_2 - y_1)\). This has the effect of making the output

\(^{51}\)We will not formally define equivalent functions, but use the term here colloquially, meaning functions which may appear different in their formulas, but are in fact the same functions. Technically, a better term would be equal functions, corresponding to the idea that, say, \( x^2 + 2x + 1 = (x + 1)^2 \), this being an equality of what appear to be two different functions because of how they are written.
Figure 2.21: The graph immediately above shows the function \( y = \sin x \) (dashed) and the function \( y = 4 \) (gray), along with the product function \( f(x) = 4\sin x \). We can consider the number 4 as an amplifier of the output of the function \( \sin x \); alternatively we can consider \( \sin x \) as a nonconstant amplifier of the function 4. The apparent symmetry in the roles of the two functions is illustrated in the second (right) function diagram (top), while the notion that the output of the sine function is altered by multiplying by 4 is illustrated in the first diagram, as we did in Section 2.7. See Example 2.8.5, page 164.
of the function change four times as rapidly as before, when we change the input. In this sense, the constant multiplier 4 serves as an amplifier of the function $\sin x$; it scales the output by a factor of 4.

However, it is useful to note that here the function $\sin x$ can be viewed as a nonconstant amplifier of the function $y = 4$, though we have to be mindful that when we multiply a function by a negative quantity we "flip" that function vertically, and when we multiply by zero the function’s height collapses entirely. See Figure 2.21, page 163.

When we multiply two functions with their own trends, each superimposes itself on the other, pointwise, as a vertical scaling and possibly “inverting” (or reflecting) factor. If we have a function $h(x) = f(x)g(x)$, it is interesting to look at $h(x)$ alternatingly as a nonconstant rescaling of $f(x)$ by the factor $g(x)$, and alternatively consider $f(x)g(x)$ as a nonconstant rescaling of $g(x)$ by $f(x)$.

**Example 2.8.6** Consider the function $h(x) = 1.2^x \sin x$. The $1.2^x$ will be an increasing amplifying factor of $\sin x$. Alternatively, the $\sin x$ will be an oscillating factor multiplying the function $1.2^x$. See Figure 2.22.

At this point it should be mentioned that, while we may be able to predict some behaviors for products of two functions, often they are too complicated to have much confidence that we can make such predictions easily. As with sums (see Example 2.8.4, page 162), for products of functions it is not uncommon to have “emergent” properties that require a more sophisticated analysis to predict. For a simple example, consider the following:
Example 2.8.7 Consider the function \( f(x) = \sin x \cos x \). Both functions are oscillating, and periodically changing signs; they are “amplifying” each other in nonconstant ways. This product will be zero anytime \( \sin x = 0 \) (\( x = 0, \pm \pi, \pm 2\pi, \pm 3\pi, \ldots \)), and anytime \( \cos x = 0 \) (\( x = \pm \pi/3, \pm 3\pi/2, \pm 5\pi/2, \ldots \)). In fact anytime one has a peak or trough, the other negates that effect by being zero.

While it is not entirely clear what the final result will be, we can make some coarse predictions. For instance,

\[
|\sin x \cos x| = |\sin x| \cdot |\cos x| \leq 1 \cdot 1 = 1,
\]

and so \( |\sin x \cos x| \leq 1 \), i.e., \( -1 \leq f(x) \leq 1 \). However, just as we used algebra to analyze \( y = x^2 + 2x + 1 \) to see it can be written in a simpler way, namely \( y = (x + 1)^2 \), there are tools from trigonometry which can help here. In particular, the sine addition identity gives us

\[
\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta,
\]

and so when \( \alpha = \beta = x \) we get the double-angle sine formula \( \sin 2x = 2 \sin x \cos x \). Dividing by 2 we see that

\[
f(x) = \sin x \cos x = \frac{1}{2} \sin 2x.
\]

This simplifies drawing this function tremendously, since we have a direct way to draw this function based upon our knowledge of the sine curve and these kinds of transformations of functions. See Figure 2.23, page 166.

Note that from this analysis we actually have the more precise bound:

\[
|f(x)| = \left| \frac{1}{2} \sin 2x \right| = \frac{1}{2} |\sin 2x| \leq \frac{1}{2} \cdot 1 = \frac{1}{2}.
\]

Indeed, the graph of the function also shows the range of the outputs is \(-\frac{1}{2} \leq f(x) \leq \frac{1}{2}\).
Figure 2.23: Graph of $f(x) = \sin x \cos x$ (in bold), which has a complicated enough structure that it is difficult to analyze easily, though some trends (oscillation, where the output is zero, sign changes, boundedness) are readily clear. When we use a trigonometric identity to rewrite the function as $f(x) = \frac{1}{2} \sin 2x$, it becomes much simpler to analyze. For instance, in its original form it is clear $|f(x)| \leq 1$, while in the rewritten form we get $|f(x)| \leq \frac{1}{2}$, which is a more precise absolute (as in absolute values) bound for $f(x)$. Also pictured are the functions $y = \sin x$ (dashed) and $y = \cos x$ (dotted). See Example 2.8.7, page 165.
2.8.3 Compositions of Functions

Here we look at functions that are constructed when we take the output of one function and immediately feed it into the input of another. The title of this subsection is not the most obvious, but it refers to the way mathematicians verbalize the notation defined below.

Definition 2.8.1 We define the composition of the function \( f \) with \( g \), notated \( f \circ g \), by

\[
(f \circ g)(x) = f(g(x)).
\]  

(2.80)

This is also denoted \( f \circ g(x) \), and it is spoken “\( f \) composed with \( g \) of \( x \),” “\( f \) circle \( g \) of \( x \),” or simply “\( f \) of \( g \) of \( x \).”

This notation on the left-hand side of Equation (2.80) is found in virtually every calculus textbook, though it is not well motivated in calculus.\(^{52}\) It is introduced here for completeness, but we will usually use the more obvious notation \( f(g(x)) \) throughout the text.

It is an important idea, that \( x \) would be processed by an composite function which has two overall steps: \( f \) and \( g \).

Example 2.8.8 Consider the functions \( y = e^{x^2} \) and \( y = (e^x)^2 \), in turn.

- \( y = e^{x^2} \): Note that the input of the exponential function, namely \( x^2 \), is never negative here, and even as \( x \) becomes more and more negative, \( x^2 \) would still be a growing, positive number, so \( e^{x^2} \) will grow large (faster than \( x \) does) as \( x \) becomes large and positive, or large and negative. Furthermore, if \( (x, y) \) is on the graph, so is \( (-x, y) \), which will give the graph a symmetry.

- \( y = (e^x)^2 \) will be greater than \( y = e^x \) when \( e^x > 1 \), i.e., for \( x > 0 \). Furthermore, \( y = (e^x)^2 \) will be positive but less than \( y = e^x \) when \( e^x \in (0, 1) \), i.e., for \( x < 0 \). This is enough to conclude that this function will “grow” faster as \( x \) grows positive, but will “shrink” faster as \( x \) becomes more and more negative.

We also should note that this function can be rewritten as \( y = e^{2x} \), which is a simple, horizontal rescaling of the function \( y = e^x \).

All of this is reflected in the diagrams and graphs in Figure 2.24.

There are enumerable, interesting ways to mix functions in “chains,” where the output of one is fed into the input of another, and possibly the output of the second fed as input to a third, and so on. While the possibilities are endless—and eventually we have to look case-by-case and possibly use some calculus in the end—a few reasonable and common examples can be pursued here, and in the exercises. Here we will consider some such functions which include absolute values within their chains, and some which contain oscillatory trigonometric functions.

Example 2.8.9 Consider the functions \( y = |x^2 - 2x + 8| \), \( y = |\log_2 x| \), and \( y = \log_2 |x| \).

- For \( y = |x^2 - 2x + 8| \), we first look at the “inner function” \( x^2 - 2x + 8 = (x-4)(x+2) \). We note that it is a quadratic function, whose graph will be parabolic, opening “upwards.” Since it is zero at \( x = 4, -2 \), its vertex must be at \( x = 1 \), i.e., the point \( (1, 1^2 - 2 \cdot 1 + 8) = (1, -9) \). Now

\[
y = |x^2 - 2x + 8| = \begin{cases} 
  x^2 - 2x + 8 & \text{if } x^2 - 2x + 8 \geq 0 \\
  -(x^2 - 2x + 8) & \text{if } x^2 - 2x + 8 < 0.
\end{cases}
\]

\(^{52}\)The notation is best motivated in Abstract Algebra and Operator Theory, which may be found in junior- or senior-level mathematics and physics courses. Most algebra and calculus texts introduce the notation with some fanfare and rarely use it subsequently.
Figure 2.24: Illustration of the functions \( y = e^{x^2} \) and \( y = (e^x)^2 = e^{2x} \). In the first case, the output of \((\ )^2\) is fed into \(e(\ )\), and in the second these are reversed. Algebraically the latter is the same as the output of \(2(\ )\) fed into \(e(\ )\), which is reflected in the graph as a horizontal contraction of the graph \( y = e^x \) by a factor of two, to arrive at \( y = e^{2x} \), giving the same result as \( y = e^{x^2} \). See Example 2.8.8.

Exercises

1. Recall Figure 2.22, page 164. Without carefully plotting many points, predict what the graph of \( y = \sin(1.2^x) \) would look like.
Chapter 3

Continuity and Limits of Functions

The concept of continuity is an important first step in the analysis leading to differential and integral calculus. It is also an important analytical tool in its own right, with significant practical applications. Fortunately the main theorems are intuitive, though their proofs can be technically challenging. Nonetheless we prove most of the continuity theorems we state, while the remaining theorems we discuss and apply without proof, since they are intuitive and useful but their proofs require background material from at least a junior level real analysis or topology course.

Related to continuity is the concept of limit, which is vital for calculus since it puts calculus on the same rigorous footing as other mathematical disciplines such as algebra and geometry. In our more modern times it has further conceptual appeal, as it is often only possible to approximate the solution to some problem, even though our method of approximation may be arbitrarily close to the actual solution if given enough computing resources. However, the real value of limits lies in its use in confronting an interesting phenomenon. This is the fact that in mathematics we at times find our analysis (algebraic, geometric or otherwise) breaking down at exactly the value of some variable where we would like to compute something; that is, we are allowed to let that variable “approach” the desired value as closely as we would like but it cannot equal that value according to our classical, pre-calculus mathematics. The relatively modern mathematical tool called “limits” can often break through the analytic barrier at that value, in turn opening us to the extensive and spectacularly useful field we call calculus.

Many examples of continuity and limits in action seem straightforward enough, but without a sufficiently deep understanding it is all too easy for students to fall victim to common errors. For this reason we introduce the rather technical definition of continuity here, and develop a method to prove continuity in cases which may seem obvious. We then employ limits for cases where continuity is “broken,” and in numerous other contexts to make calculus possible, and as an analytical tool in its own right.

Understanding limits and continuity sufficiently to avoid numerous common mistakes requires a care and depth of thought which we attempt to foster in this chapter. After continuity, we use a “forms” approach to limits, those forms themselves being ultimately intuitive but nonetheless requiring students to study them carefully and extensively to achieve satisfactory proficiency.\footnote{Unfortunately limit computations can be deceiving in that perhaps the majority of problems one first encounters do not require a deep understanding in order to “guess” their correct answers. However the interesting (and more advanced) cases tend to lie outside of those which are easily guessed, and so computing the correct answers in such cases requires a much deeper understanding. We take the approach here that it is better to heavily analyze the simpler cases, so that the later cases are more easily learned.}
3.1 Definition of Continuity at a Point

The function \( f(x) \) is continuous at \( x = a \) if and only if we can guarantee \( f(x) \) to be close to the value \( f(a) \) by restricting \( x \) to be close to \( a \). To rephrase, we say \( f(x) \) is continuous at \( x = a \) if, given any positive tolerance \( \epsilon > 0 \) we choose for \( f(x) \) as an approximation for \( f(a) \), we can then find a positive tolerance \( \delta > 0 \) for \( x \) as an approximation for \( a \) so that \( \delta \)-tolerance in \( x \) allows at most \( \epsilon \) tolerance in \( f(x) \). The definition below is very technical, but through reflection and exposure to examples, one eventually sees that this is exactly what is required.

**Definition 3.1.1** The function \( f(x) \) is continuous at the point \( x = a \) if and only if\(^2\)

\[
(\forall \epsilon > 0)(\exists \delta > 0)(\forall x)(|x - a| < \delta \implies |f(x) - f(a)| < \epsilon).
\]

This is sometimes called the *epsilon-delta* (\( \varepsilon \)-\( \delta \)) definition of continuity. (Note that all values are assumed real, i.e., \( \varepsilon, \delta, x, a, f(x), f(a) \in \mathbb{R} \), and so for instance \((\forall x)\) is short for \((\forall x \in \mathbb{R})\).)

Now let us examine the various parts of the definition.

\[
|f(x) - f(a)| < \varepsilon \tag{3.2}
\]

\[
|x - a| < \delta \tag{3.3}
\]

\[
(\forall x)(|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon) \tag{3.4}
\]

\[
(\forall \epsilon > 0)(\exists \delta > 0) \tag{3.5}
\]

(3.2): \( f(x) \) will be within \( \varepsilon \) of \( f(a) \). In other words, the function at \( x \) will be near in value to the value of the function at \( a \). How near? Less than \( \varepsilon \) distance away.

(3.3): \( x \) is within \( \delta \) of \( a \). (Otherwise the implication holds true vacuously, but that case is useless. What is important is what occurs when \( |x - a| < \delta \).

(3.4): The condition that \( x \) be within \( \delta \) of \( a \) forces \( f(x) \) to be within \( \varepsilon \) of \( f(a) \). In other words, allowing \( x \) to stray by less than \( \delta \) from \( a \) keeps \( f(x) \) within \( \varepsilon \) of \( f(a) \). By controlling \( x \) by allowing it a tolerance of less than \( \delta \), we control \( f(x) \) to have a tolerance of less than \( \varepsilon \).

(3.5): Whatever positive value of \( \varepsilon \) we choose, we can find a \( \delta \) which satisfies (3.4). In particular, *no matter how small* we choose \( \varepsilon > 0 \), we can find a positive \( \delta \) so that (3.4) is satisfied.

For a final rephrasing, we have the statement that we can control the tolerance \( \varepsilon \) in the output \( f(x) \) as much as we would like, so long as \( \varepsilon > 0 \), by controlling the tolerance \( \delta \) (which must also be positive\(^3\)) in the input variable \( x \).

---

\(^2\)Many texts abbreviate statements like (3.1) as follows:

\[(\forall \epsilon > 0)(\exists \delta > 0)(|x - a| < \delta \implies |f(x) - f(a)| < \epsilon),\]

the idea being that the \( \forall x \) is understood when we make an unquantified (in \( x \)) statement like \( |x - a| < \delta \implies |f(x) - f(a)| < \varepsilon \). For a similar example in English, consider the following two statements, usually deemed equivalent:

- All Americans have trouble speaking English.
- If \( X \) is an American, then \( X \) has trouble speaking English.

Most see both as false exactly when we can find one American (counterexample) who has no such trouble. In other words, the second statement is as much a “blanket” statement as the first, and is in fact equivalent to the first. We leave the \( \forall x \) in (3.1) for precision and to aid in negating the statement, using rules from Section 1.4.

\(^3\)Notice that \( \delta = 0 \) would be worthless for several reasons. First, the implication would be vacuously true and all functions would be continuous everywhere, since \( |x - a| < 0 \) would never be satisfied. Second, when—in reality—do we ever have a tolerance of zero in a measurement? Finally, as we explore the implications of continuity, we will see that having positive \( \delta \) is central to the spirit of what follows, particularly with regards to limits; \( \delta > 0 \) allows for some “wiggle room” for \( x \) near \( x = a \), and this wiggle room is crucial to the concept of continuity.
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

Example 3.1.1 Show (by a proof!) that the function \( f(x) = 5x - 9 \) is continuous at the point \( x = 2 \), according to the definition (3.1).

Solution: First we notice that \( f(2) = 1 \), so we are trying to show the truth of the statement

\[
(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)(|x - 2| < \delta \implies |f(x) - 1| < \varepsilon).
\]

Before completing this example, we insert here the following general strategy for all such proofs.

Strategy for Writing \( \varepsilon \)-\( \delta \) Proofs

1. Use the statement \(|f(x) - f(a)| < \varepsilon\) to see how it can be controlled by \(|x - a|\), in particular if \(|x - a|\) is a factor of \(|f(x) - f(a)|\).

2. If necessary, assume a priori\(^4\) that \( \delta \) is smaller than some fixed positive number to control the other factors contained in \(|f(x) - f(a)|\).

3. Find \( \delta \) as a function of \( \varepsilon \), i.e.,

\[
\delta = \delta(\varepsilon)
\]

with \( \varepsilon \in (0, \infty) \) (technically, the domain of this function \( \delta(\varepsilon) \)) and \( \delta \in (0, \infty) \), and such that the preliminary (or exploratory) analysis indicates that choice of \( \delta \) satisfies the definition (3.1).

4. Verify that (3.1) holds with this choice of \( \delta \), and in doing so write the actual proof.

The first three steps are analysis, or “scratch-work” to determine the form of \( \delta \). The final step is the actual proof, though elements of it are often contained in the analysis/scratch-work. Let us apply this strategy to the problem at hand.

\(^4\)Presumptive; before observations. When Step 2 is necessary, we make such suppositions not necessarily based upon observation, but to help focus our search for \( \delta \). If continuity is true, (as we will see) we will find that a legitimate \( \delta \) is still available even with the restriction. In fact, if the limit definition holds for a value of \( \delta > 0 \), it holds for any smaller positive value \( \delta \), so this is not a fatal restriction at all. Note that if \( 0 < \delta_1 < \delta_2 \), and \( |x - a| < \delta_1 \), then \( |x - a| < \delta_2 \) as well so we can always take a smaller value for \( \delta \) in our proof. The upshot is that a priori restricting the size of \( \delta > 0 \) from the start never jeopardizes our ability to prove continuity.
3.1. DEFINITION OF CONTINUITY AT A POINT

Scratch-work: We want $|f(x) - f(a)| < \varepsilon$ to follow from our choice of $\delta$. We work backwards from that statement, with $f(x) = 5x - 9$, $a = 2$, and $f(a) = f(2) = 1$.

$$|f(x) - f(a)| < \varepsilon$$  (what we need)
$$\Leftrightarrow |f(x) - 1| < \varepsilon$$
$$\Leftrightarrow |5x - 9 - 1| < \varepsilon$$
$$\Leftrightarrow |5x - 10| < \varepsilon$$
$$\Leftrightarrow 5|x - 2| < \varepsilon$$
$$\Leftrightarrow |x - 2| < \frac{\varepsilon}{5}$$  (how to get it).

We can see from this that, if we take $\delta = \frac{\varepsilon}{5}$, then $\delta > 0$ (since $\varepsilon > 0$ is assumed), and the bottom could be written $|x - 2| < \delta$. Then the implication could be read from that statement upwards to get $|f(x) - 1| < \varepsilon$. We summarize this in the proof.

Proof: For any $\varepsilon > 0$ choose $\delta = \frac{\varepsilon}{5}$. Then $\delta > 0$ exists and satisfies

$$|x - 2| < \delta \implies |f(x) - f(2)| = |(5x - 9) - 1|$$
$$= |5x - 10| = 5|x - 2| < 5\delta = 5 \cdot \frac{\varepsilon}{5} = \varepsilon,$$ q.e.d.

Note that the final line of the proof does imply that $|f(x) - f(2)| < \varepsilon$, with intermediate calculations, some of which one may wish to omit with practice.

Example 3.1.2 Show that $f(x) = 2x + 3$ is continuous at $x = -5$.

Scratch-work: Here $a = -5$ and $f(a) = f(-5) = -7$. Hence we wish to find $\delta > 0$ so that

$$|x - (-5)| < \delta \implies |f(x) - (-7)| < \varepsilon,$$ i.e.,

$$|x + 5| < \delta \implies |f(x) + 7| < \varepsilon.$$

Again we work backwards from the conclusion we wish to justify.

$$|f(x) + 7| < \varepsilon$$
$$\Leftrightarrow |2x + 3 + 7| < \varepsilon$$
$$\Leftrightarrow |2x + 10| < \varepsilon$$
$$\Leftrightarrow 2|x + 5| < \varepsilon$$
$$\Leftrightarrow |x + 5| < \frac{\varepsilon}{2}.$$

This time we take $\delta = \frac{\varepsilon}{2}$, and write the proof.

Proof: For $\varepsilon > 0$, set $\delta = \frac{\varepsilon}{2}$. Then $\delta > 0$ exists and satisfies

$$|x+5| < \delta \implies |f(x)+7| = |2x+3+7| = |2x+10| = 2|x+5| < 2\delta = 2 \cdot \frac{\varepsilon}{2} = \varepsilon,$$ q.e.d.

Proving continuity for first-degree polynomials is rather routine at any $x = a$. The strategy is the same for each, with the only complications coming from the signs of the values in question. For completeness we include one more such example.
Example 3.1.3 Show that \( f(x) = 9 - 4x \) is continuous at \( x = 2 \).

**Scratch-work:** Here \( a = 2 \), \( f(a) = f(2) = 1 \). Now we must be a little more careful, and will make use of the fact that \( |a \cdot b| = |a| \cdot |b| \).

\[
|f(x) - 1| < \varepsilon \\
\iff |9 - 4x - 1| < \varepsilon \\
\iff | - 4x + 8| < \varepsilon \\
\iff |(-4)(x - 2)| < \varepsilon \\
\iff 4|x - 2| < \varepsilon \\
\iff |x - 2| < \frac{1}{4} \varepsilon .
\]

**Proof:** For \( \varepsilon > 0 \), choose \( \delta = \frac{1}{4} \varepsilon \). Then \( \delta > 0 \) (exists) and

\[
|x - 2| < \delta \implies |f(x) - 1| = |9 - 4x - 1| = |- 4x + 8| = |(-4)(x - 2)| = 4|x - 2| < 4\delta = 4 \cdot \frac{\varepsilon}{4} = \varepsilon , \text{ q.e.d.}
\]

The function which represents a line is the easiest to confirm continuity at every point. If \( f(x) = mx + b \), where \( m \neq 0 \), it is clear from the geometric meaning of slope \( m \) that a variation (absolute value of “rise”) of less than \( \varepsilon \) in height \( f(x) \) can be achieved by allowing a variation (absolute value of “run”) of less than \( \frac{1}{|m|} \varepsilon \) in \( x \). Thus \( \delta = \frac{1}{|m|} \varepsilon \) is the largest \( \delta \) which satisfies the definition of continuity for such a function \( f(x) \). (See again our three “linear” examples above, and compare their slopes with our choices of \( \delta \).)

Example 3.1.4 Show that \( f(x) = x^2 \) is continuous at \( x = 0 \).

**Scratch-work:** Here \( a = 0 \) and \( f(a) = f(0) = 0 \). We therefore want to choose \( \delta > 0 \) such that

\[
|x - 0| < \delta \implies |f(x) - 0| < \varepsilon , \quad \text{i.e.,} \\
|x| < \delta \implies |x|^2 < \varepsilon .
\]

Again we begin with the inequality we would like to result, and see how we might get it.

\[
|x|^2 < \varepsilon \iff |x| < \sqrt{\varepsilon}.
\]

Here we do have \( \iff \) because we are dealing with only positive quantities (recall that \( \sqrt{\cdot} \) is an increasing function on \( [0, \infty) \)). Thus we have a good choice for \( \delta \), namely \( \delta = \sqrt{\varepsilon} \).

**Proof:** For \( \varepsilon > 0 \), set \( \delta = \sqrt{\varepsilon} \). Then \( \delta > 0 \) exists and satisfies

\[
|x - 0| < \delta \implies |f(x) - f(0)| = |x|^2 = |x|^2 < \delta^2 = (\sqrt{\varepsilon})^2 = \varepsilon , \text{ q.e.d.}
\]

It should be clear that we could easily modify this example to show that \( f(x) = x^n \) is continuous at \( x = 0 \) for \( n \in \mathbb{N} \). From there it is not hard to show \( f(x) = x^{m/n} \) is also continuous at \( x = 0 \), as long as \( n \) is odd and \( m, n \in \mathbb{Z} \). Once we stray from \( x = 0 \), we begin to have more difficulties, as illustrated in the next example.

---

5It is false if \( n \) is even, assuming \( m/n \) is a reduced fraction. The trouble there is that \( x^{m/n} = (\sqrt[n]{x})^m \) is undefined for \( x < 0 \), so the second part of \( (|x - 0| < \delta) \implies (|f(x) - f(0)| < \varepsilon) \) is false for such values of \( x \).
Example 3.1.5 Show that \( f(x) = x^2 \) is continuous at \( x = 4 \).

**Scratch-work:** This has a complication that the previous problem did not. To see this we first attempt to proceed as in the previous Example 3.1.4. Here \( a = 4 \) and \( f(a) = 16 \). We therefore want to choose \( \delta > 0 \) such that
\[
|x - 4| < \delta \implies |f(x) - 16| < \varepsilon.
\]

Working backwards as before we get
\[
|f(x) - 16| < \varepsilon \iff |x^2 - 16| < \varepsilon \iff |x + 4| \cdot |x - 4| < \varepsilon.
\]

We would like to be able to divide both sides by \(|x + 4|\), except that it is not constant. Here Step 2 in our strategy comes into play. We will control the \(|x + 4|\) term by assuming a priori that (in all cases) \( \delta \leq 1.5 \)
\[
(|x - 4| < \delta) \land (\delta \leq 1) \implies |x - 4| < 1 \implies -1 < x - 4 < 1 \implies 3 < x < 5.
\]

Now we add 4 to this inequality to get
\[
(|x - 4| < \delta) \land (\delta \leq 1) \implies 7 < x + 4 < 9.
\]

With \( x + 4 \) between 7 and 9, its absolute size is strictly bounded by the number with the largest absolute value, 9, i.e.,
\[
(|x - 4| < \delta) \land (\delta \leq 1) \implies 7 < x + 4 < 9 \implies |x + 4| < 9.
\]

Continuing the scratch-work, we would get
\[
(|x - 4| < \delta) \land (\delta \leq 1) \implies |f(x) - 16| = |x + 4||x - 4| \leq 9|x - 4| < 9\delta,
\]
(where we only had \( \leq \) in our inequality \(|x + 4||x - 4| \leq 9|x - 4| \) because of the case \( x = 4 \), where we have the equality \( 0 = 0 \) and now this looks similar to the earlier examples. To achieve \(|f(x) - 16| < \varepsilon\) it would be sufficient to have \( 9\delta \leq \varepsilon \), which is to say \( \delta \leq \frac{\varepsilon}{9} \).

Picking \( \delta = \frac{\varepsilon}{9} \) is not quite enough, since we assumed \( \delta \leq 1 \) (to get our estimate \(|x + 4| < 9\)), and this would be false if \( \varepsilon > 9 \). To cover both of these requirements for \( \delta \) for every given \( \varepsilon \) (as the definition requires), we choose \( \delta = \min \{1, \frac{\varepsilon}{9}\} \), i.e., the minimum of the two numbers, which will still be positive. Now we write the proof.

**Proof:** For \( \varepsilon > 0 \), choose \( \delta = \min \{1, \frac{\varepsilon}{9}\} \). Then \( \delta > 0 \) exists and satisfies
\[
|x - 4| < \delta \implies |f(x) - 16| = |x + 4| \cdot |x - 4| < 9|x - 4| < 9\delta \leq 9 \cdot \frac{\varepsilon}{9} = \varepsilon, \text{ q.e.d.}
\]

An experienced reader of mathematical proofs would be able to make sense of the proof above on its own—perhaps with minimal writing to check one inequality—but for our purposes we note that much of the explanation of the proof can be found again in the scratchwork. A key observation is that since \( \delta = \min \{1, \frac{\varepsilon}{9}\} \), we have both \( \delta \leq 1 \) and \( \delta \leq \frac{\varepsilon}{9} \). Some “steps” in our proof rely on related implications of these in turn, namely \((|x - 4| < \delta) \land (\delta \leq 1) \implies |x + 4| < 9\), which is believable on its face (as it is not hard to see that \(|x - 4| < 1 \implies x \in (3, 5) \implies
\)

\[\text{Here we chose } \delta \leq 1 \text{, but we could have chosen any positive number for the maximum we allow } \delta \text{ to be. We just need to restrict } \delta \text{ (though keeping it positive) to control the other factors of } |f(x) - 16|.,\]
$x + 4 \in (7, 9) \implies |x + 4| < 9$, and $\delta \leq \frac{1}{3} \varepsilon \implies 9\delta \leq 9 \cdot \frac{1}{3} \varepsilon = \varepsilon$. As in previous examples, it is useful that in the statement of our proof we can see the outline of the definition of continuity (3.1), page 169. Also important to note is that while we could have made a different choice for $\delta$, the obvious possibilities would still require $\delta$ to be defined as a similar type of minimum (depending upon the \textit{a priori} restriction of the form $\delta \leq M$).

**Example 3.1.6** Show that $f(x) = -x^3$ is continuous at $x = -2$.

**Scratchwork:** Here $a = -2$, and $f(a) = f(-2) = 8$, so we hope for every $\varepsilon > 0$ to find $\delta > 0$ such that $|x - (-2)| < \delta \implies |f(x) - 8| < \varepsilon$, i.e., $|x + 2| < \delta \implies |f(x) - 8| < \varepsilon$. Working backwards as before, we see (using $a^3 + b^3 = (a + b)(a^2 - ab + b^2)$) that

$$
|f(x) - f(2)| < \varepsilon \iff |(-1)(x^3 + 8)| < \varepsilon \\
\iff 1 \cdot |x^3 + 8| < \varepsilon \\
\iff |x + 2| \cdot |x^2 - 2x + 4| < \varepsilon.
$$

Here we will make a rather crude estimate\footnote{In mathematical analysis, the term \textit{estimate} often refers to a bound on the size of a quantity. For instance, $|x| < 100$ means $-100 < x < 100$, giving a lower and upper bound for $x$. In common usage, the word “estimate” often refers instead to what mathematicians and other scientists would call “approximation.”} on the size of $|x^2 - 2x + 4|$, eventually using the triangle inequality (2.35), page 98 but to do so will require as before an \textit{a priori} bound on $\delta$. So for simplicity we will again assume in all cases that $\delta \leq 1$ for what follows (and incorporate that restriction into our proof as well).

$$(|x + 2| < \delta) \land (\delta \leq 1) \implies |x + 2| \leq 1 \iff -1 < x + 2 < 1 \iff -3 < x < 1 \\
\implies |x| < 3,$$

from which we can summarize (in the first line below) and extend using the triangle inequality (in the second) to get

$$(|x + 2| < \delta) \land (\delta \leq 1) \implies |x| < 3 \\
\implies |x^2 - 2x + 4| \leq |x^2| + |-2x| + |4| = |x|^2 + 2|x| + 4 \\
< 3^2 + 2(3) + 4 \\
= 19$$

that is, $(|x + 2| < \delta) \land (\delta \leq 1) \implies |x^2 - 2x + 4| \leq 19.$

So far, under the assumptions that $|x + 2| < \delta$ and $\delta \leq 1$ we have (slightly rearranged)

$$|f(x) - 8| = |x^2 - 2x + 4| \cdot |x + 2| \leq 19|x + 2| < \frac{19\delta}{\varepsilon}$$

If we also have $\delta \leq \varepsilon/19$ we will have $|f(x) - 8| < 19 \frac{\delta}{19} \leq \varepsilon$, and our proof would be complete. However we must remember that this was possible because we already assumed $\delta \leq 1$, and so accomplish both of these equally important restrictions by setting $\delta = \text{min} \{1, \frac{\varepsilon}{19}\}$. The inequalities involved in the proof (after the “$\implies$”) include, in order, the triangle inequality, one implied by $\delta \leq 1$, $|x + 2| < \delta$ and finally $\delta \leq \frac{\varepsilon}{19}$. 


3.1. DEFINITION OF CONTINUITY AT A POINT

Proof: Let \( \varepsilon > 0 \), and set \( \delta = \min \{ 1, \frac{\varepsilon}{19} \} \). Then \( \delta > 0 \) exists and

\[
|x + 2| < \delta \implies |f(x) - f(-2)| = |x^2 - 2x + 4| \cdot |x + 2| \leq \left( |x|^2 + 2|x| + 4 \right) |x + 2| \\
\leq (3^2 + 2(3) + 4)|x + 2| \\
= 19|x + 2| < 19\delta \leq 19 \frac{\varepsilon}{19} = \varepsilon,
\]

q.e.d.

The next example requires a somewhat different bit of cleverness, and similar a priori restriction on \( \delta \).

Example 3.1.7 Show that \( f(x) = \frac{1}{x} \) is continuous at \( x = 5 \).

Scratch-work: Here \( a = 5 \) and \( f(a) = f(5) = \frac{1}{5} \). Again we work backwards:

\[
\left| f(x) - \frac{1}{5} \right| < \varepsilon \\
\iff \left| \frac{1}{x} - \frac{1}{5} \right| < \varepsilon \\
\iff \left| \frac{5 - x}{5x} \right| < \varepsilon \\
\iff \frac{1}{5} \cdot \frac{1}{|x|} \cdot |x - 5| < \varepsilon.
\]

As before, the \( |x - 5| \) will be controlled by \( \delta \), but we need to also use \( \delta \) to control the factor \( \frac{1}{|x|} \). Again we will assume a priori that \( \delta \leq 1 \).

\[
|x - 5| < \delta \implies |x - 5| < 1 \iff -1 < x - 5 < 1 \iff 4 < x < 6 \iff 4 < |x| < 6.
\]

With these bounds on \( |x| \), we also get bounds on \( \frac{1}{|x|} \) (noting that if \( z_1, z_2 > 0 \) and \( z_1 < z_2 \) then \( \frac{1}{z_1} > \frac{1}{z_2} \)):

\[
|x - 5| < 1 \implies 4 < |x| < 6 \iff \frac{1}{6} < \frac{1}{|x|} < \frac{1}{4}.
\]

With all these assumptions, then, we have (continuing from before and noting the \( x = 5 \) case)

\[
\left| f(x) - \frac{1}{5} \right| = \frac{1}{5} \cdot \frac{1}{|x|} \cdot |x - 5| \leq \frac{1}{5} \cdot \frac{1}{4} \cdot |x - 5| < \frac{1}{5} \cdot \frac{1}{4} \cdot \delta < \frac{\varepsilon}{20}.
\]

This is less than \( \varepsilon \) if \( \delta \) is no bigger than \( 20\varepsilon \). Now the analysis above also assumed \( \delta \leq 1 \), so we take

\[
\delta = \min \{ 1, 20\varepsilon \}.
\]

Now we state the proof.

Proof: For \( \varepsilon > 0 \), choose \( \delta = \min \{ 1, 20\varepsilon \} \). Then \( \delta > 0 \) exists and satisfies

\[
|x - 5| < \delta \implies |f(x) - f(5)| = \left| \frac{1}{x} - \frac{1}{5} \right| = \left| \frac{5 - x}{5x} \right| \\
= \frac{1}{5} \cdot \frac{1}{|x|} \cdot |x - 5| < \frac{1}{5} \cdot \frac{1}{4} \cdot \delta \leq \frac{1}{20} \cdot 20\varepsilon = \varepsilon,
\]

q.e.d.
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

We should note here that if we had chosen \( a = 0.5 \), then we could not use 1 as the upper bound for \( \delta \), since the function is undefined at a point within 1 of 0.5 (namely at \( x = 0 \)). For such an \( a \) we should instead assume \textit{a priori} that \( \delta < 0.25 \), or a similar number to be sure to avoid any problems with the definition of the function for any values of \( x \) in which \( |x - a| < \delta \). Indeed, we would wish to be sure that \( f(x) \) is defined within the “wiggle room” allowed by \( \delta \) in the continuity definition (3.1), page 169 at the start of this section.

\textbf{Example 3.1.8} Show that \( f(x) = x^4 \) is continuous at \( x = -2 \).

\textit{Scratch-work}: Here \( a = -2 \) and \( f(a) = 16 \). Again we will attempt to work backwards.

\[
|f(x) - f(-2)| < \varepsilon \iff |x^4 - 16| < \varepsilon \iff |x^2 + 4| \cdot |x - 2| \cdot |x + 2| < \varepsilon.
\]

Now our “\( |x - a| \)”, namely \( |x + 2| \) is controlled by \( \delta \), so we need to control the other two factors. Again let us assume that \( \delta \leq 1 \). Then

\[
|x + 2| < \delta \implies |x + 2| < 1 \iff -1 < x + 2 < 1 \iff -3 < x < -1.
\]

Note for later reference that \( |x| < 3 \).

For the \( |x - 2| \) term we can subtract 2 from the above to get \( -5 < x - 2 < -3 \), giving \( |x - 2| < 5 \). For the \( |x^2 + 4| \) term, we have \( x^2 + 4 > 0 \), so

\[
|x^2 + 4| = x^2 + 4 = |x|^2 + 4 < (3)^2 + 4 = 13.
\]

(Note that this was because \( |x| < 3 \).) So far we have

\[
\delta \leq 1 \implies |f(x) - f(-2)| = (x^2 + 4)|x - 2| \cdot |x + 2| < 13 \cdot 5 \cdot \delta.
\]

Taking \( \delta = \min \{ \frac{\varepsilon}{65}, 1 \} \) should give us a proof.

\textbf{Proof}: Let \( \varepsilon > 0 \) and choose \( \delta = \min \{ \frac{\varepsilon}{65}, 1 \} \). Then \( \delta > 0 \) and

\[
|x - (-2)| < \delta \implies |f(x) - f(-2)| = |x^4 - 16| = (x^2 + 4)|x - 2| \cdot |x + 2| \leq 13 \cdot 5 \cdot \delta \leq 26 \cdot \frac{\varepsilon}{65} = \varepsilon,
\]

q.e.d.

For our final example, we look at a case where we introduce factors in our computation to extract the \( |x - a| \) factor.

\textsuperscript{8}We could also use the triangle inequality to get \( |x - 2| \leq |x| + |2| < 3 + 2 = 5 \). This happens to give the same bound, but in more complicated cases that might not happen. Either bound would then work. For an engineering application, one would likely prefer whatever gives the larger \( \delta \), which indicates less sensitivity to tolerance in \( x \) to achieve \( \varepsilon \) tolerance in \( f(x) \).
3.1. DEFINITION OF CONTINUITY AT A POINT

Example 3.1.9  Show that \( f(x) = \sqrt{x} \) is continuous at \( x = 4 \).

Scratch-Work: As always we begin with \( |f(x) - f(a)| < \varepsilon \) and work backwards to find a case where this is implied by \( |x - a| < \delta \) for a strategically-chosen \( \delta > 0 \).

\[
|f(x) - f(4)| < \varepsilon \iff |\sqrt{x} - 2| < \varepsilon \\
\iff \frac{|\sqrt{x} - 2|}{\sqrt{x} + 2} \cdot (\sqrt{x} + 2) < \varepsilon \\
\iff \frac{1}{\sqrt{x} + 2} |x - 4| < \varepsilon.
\]

Note that the fraction \( \frac{1}{\sqrt{x} + 2} \) is (1) positive wherever it is defined, which is where \( x > 0 \), and therefore (2) maximized when the denominator is minimized, which will happen when the square root term is minimized, i.e., when \( x \) itself is minimized, but nonnegative, lest \( \sqrt{x} \) is undefined. For this case we will let \( \delta \leq 4 \), so that

\[
|x - 4| < \delta \implies |x - 4| < 4 \implies x \in (4 - 4, 4 + 4) = (0, 8)
\]

\[
\implies \frac{1}{\sqrt{x} + 2} \in \left( \frac{1}{\sqrt{0} + 2}, \frac{1}{2} \right)
\]

\[
\implies \frac{1}{\sqrt{x} + 2} < \frac{1}{2}.
\]

Using the previous computation we can say that

\[
(|x - 4| < \delta) \land (\delta \leq 4) \implies |f(x) - f(4)| = \frac{1}{\sqrt{x} + 2}|x - 4| \leq \frac{1}{2}|x - 4| < \frac{1}{2}\delta,
\]

and so we can accomplish having this be less than \( \varepsilon \) if \( \delta \leq 2\varepsilon \). We will therefore take \( \delta = \min \{4, 2\varepsilon\} \) in our proof.

\textbf{Proof:} For \( \varepsilon > 0 \), let \( \delta = \min \{4, 2\varepsilon\} \), so \( \delta > 0 \) exists and

\[
|x - 4| < \delta \implies |f(x) - f(4)| = |\sqrt{x} - 2| = \frac{1}{\sqrt{x} + 2}|x - 4| \\
\leq \frac{1}{2}|x - 4| < \frac{1}{2}\delta \leq \frac{1}{2} \cdot 2\varepsilon = \varepsilon, \text{ q.e.d.}
\]

While this section discusses the technical definition of continuity at a point \( x = a \), the reader is invited to consider where else this idea of continuity may apply. Anytime we wish to control the output of some process, the presence of continuity with respect to the input would mean that we can in principle guarantee there would be only small changes in output (of size up to \( \varepsilon \), whatever we choose that to be) by controlling the changes in the input (keeping them within \( \delta \)). For instance, if an audio amplifier’s gain changes continuously with the position of its volume control, it is much easier to make minor changes in gain by carefully moving the volume control. If (as in the case with old or dirty internal contacts) the volume control’s effect is not always continuous, it becomes much more difficult to control the gain. There are numerous other practical examples where continuity is desirable, for instance in the context of tolerances discussed previously.

However, the proofs above are technical (and the exercises somewhat difficult). Still, the reader should be encouraged that the study of subsequent sections will benefit to the extent time is spent studying this section, even if it is not mastered immediately. Furthermore, this
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

section can and should be revisited after study of future sections, so it can be seen with the benefit of a knowledge of the larger context. Of course this is true of all sections in this or any other textbook whose subject is the least bit challenging.

Exercises

1. Show that $f(x) = 9x - 11$ is continuous at $x = 2$.
2. Show that $f(x) = 9x - 11$ is continuous at $x = -2$.
3. Show that $f(x) = 3x + 1$ is continuous at $x = 5$.
4. Show that $f(x) = 6 - 2x$ is continuous at $x = -8$.
5. Show that if $m \neq 0$, then $f(x) = mx + b$ is continuous at every $x = a$. What if $m = 0$? (See Exercise 11 below.)
6. Show that $f(x) = x^3$ is continuous at $x = 0$.
7. Show that $f(x) = x^2$ is continuous at $x = 9$.
8. Show that $f(x) = x^2$ is continuous at $x = -3$.
9. Show that $f(x) = 5x^2 - 3$ is continuous at $x = 2$.
10. Show that $f(x) = \frac{1}{x^2}$ is continuous at $x = 5$.
11. Show that $f(x) = b$ (i.e., a line with slope zero) is continuous at every point $x = a$. (Hint: choosing any $\delta > 0$ will work for the continuity definition.)
12. Show that $f(x) = x^3$ is continuous at $x = 1$.
13. Show that $f(x) = x^3$ is continuous at $x = -3$.
14. Show that $f(x) = \sqrt{x}$ is continuous at $x = 9$.
15. Show that $f(x) = \sqrt{x}$ is continuous at any $a > 0$.
16. Show that $f(x) = \frac{1}{x}$ is continuous at any $a \neq 0$. (Letting $\delta \leq 1$ as in Example 3.1.7 works fine until $0 < |a| \leq 1$. For this more general case desired here, one approach is to assume a priori that $\delta \leq \frac{1}{2}|a|$, so that

$$|x - a| < \delta \implies |x - a| < \frac{1}{2}|a| \implies |x| > \frac{1}{2}|a| \implies \frac{1}{|x|} < \frac{2}{|a|}.$$  

There is some detail to showing the second implication. From this a $\delta$ can be chosen, using some minimum of two values, for each $\varepsilon > 0$.)
17. Consider $f(x) = \sqrt{x}$.
   (a) Show that $f(x)$ is continuous at $x = 0$.

---

In all fairness, it should be pointed out to the reader that some of the exercises here are likely to be quite difficult, particularly for beginning calculus students. This is because the proofs are very involved and use a large variety of methods. Furthermore, reading the proofs of the examples is very different from producing one’s own proofs from scratch.

With these exercises, students are thus advised to adopt the following general approach:
(a) attempt as many problems as possible, looking back on earlier examples for ideas;
(b) move on to the rest of the text even if few problems are completed on the first attempt; and
(c) revisit this section and its problems from time to time to attempt complete proofs for the results which were not finished previously.

With further calculus experience, the ideas and techniques should become clearer, just as the inner workings of an automobile likely make more sense—and seem more important—as one gains experience from actually driving.
3.1. Definition of Continuity at a Point

(b) Show that \( f(x) \) is continuous at \( x = 8 \).

(c) Show that \( f(x) \) is continuous at \( a \) for any \( a > 0 \).

(d) Show that \( f(x) \) is continuous at any \( a \neq 0 \).

Conclude from a–d that \( f(x) \) is continuous for all \( x \in \mathbb{R} \).

18. Recall that \( x^n - a^n = (x - a)(x^{n-1} + ax^{n-2} + a^2x^{n-3} + \cdots + a^{n-1}) \). Use this to show that \( f(x) = \sqrt[n]{x} \) is continuous for all \( x > 0 \) if \( n \in \mathbb{N} \) is even, and continuous for all \( x \in \mathbb{R} \) if \( n \in \mathbb{N} \) is odd. (For the latter you should do the case \( x = 0 \) separately.) Notice why the even case does not allow \( x = 0 \) or \( x < 0 \).

**Hint:** the triangle inequality is needed, as is the fact that \( \sqrt[n]{|x_1|} < \sqrt[n]{|x_2|} \) if \( |x_1| < |x_2| \).
3.2 Continuity Theorems

Though fundamental and technically important, using the \( \varepsilon \)-\( \delta \) definition to locate each point—or even a single point—where a function is continuous is an unwieldy approach. Fortunately there are many very general theorems which allow us to see where a function is continuous by simple inspection. The theorems below are of great use, collectively and individually, for doing just that. Here we list and prove these theorems, and later in the section we will show how to make use of them. The proofs are interesting and introduce some new techniques, but are not crucial for most of what we do. We include them here for completeness. The reader is advised to first concentrate on the theorems.

3.2.1 Basic Theorems

For each theorem below, it is useful to reflect upon some intuitions regarding what makes a function \( f(x) \) continuous at a point \( x = a \). While ultimately all continuity refers back to the technical definition found in (3.1), page 169, namely

\[
f(x) \text{ is continuous at } x = a \iff (\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)(|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon),
\]

a seemingly less precise interpretation can nonetheless be of some use, namely that if \( f(x) \) is continuous at \( x = a \), we expect that

- if \( x \) is changed very little from the value \( x = a \), then \( f(x) \) will change very little from the value \( f(a) \); sufficiently small changes away from \( x = a \) in the value of the input of \( f(x) \) will result in only small changes from \( f(a) \) in the output of \( f(x) \) very little from \( f(a) \). (This is not true if \( f(x) \) is not continuous at \( x = a \).)

It is the \( \varepsilon \)-\( \delta \) definition which makes this notion precise.

The theorems below extend this to many combinations of functions. For instance if \( f(x) \) and \( g(x) \) are both continuous at \( x = a \), then so is \( f(x) + g(x) \); if \( f(x) \) and \( g(x) \) change very little if \( x \) changes a small amount from the value \( x = a \), it is natural to expect the sum \( f(x) + g(x) \) to change little as well.

**Theorem 3.2.1** Suppose that \( f(x) \) and \( g(x) \) are continuous at \( x = a \). Then so is \( f(x) + g(x) \).

**Proof:** We are beginning under the assumption that \( f(x) \) and \( g(x) \) are continuous at \( x = a \), i.e.,

\[
(\forall \varepsilon_1 > 0)(\exists \delta_1 > 0)(\forall x)(|x - a| < \delta_1 \implies |f(x) - f(a)| < \varepsilon_1),
\]

\[
(\forall \varepsilon_2 > 0)(\exists \delta_2 > 0)(\forall x)(|x - a| < \delta_2 \implies |g(x) - g(a)| < \varepsilon_2).
\]

Define \( h(x) = f(x) + g(x) \). We want to show that \( h(x) \) is also continuous at \( x = a \).

For a given \( \varepsilon > 0 \), set both \( \varepsilon_1, \varepsilon_2 = \varepsilon/2 \). Next find corresponding \( \delta_1, \delta_2 \) satisfying the definitions above of continuity for \( f(x) \) and \( g(x) \) at \( x = a \), respectively. Finally, pick \( \delta = \min\{\delta_1, \delta_2\} \). Recalling the triangle inequality, \( |A + B| \leq |A| + |B| \), we have

\[
|x - a| < \delta \implies |h(x) - h(a)| = |f(x) + g(x) - f(a) - g(a)|
\]

\[
= |f(x) - f(a) + g(x) - g(a)| \leq |f(x) - f(a)| + |g(x) - g(a)|
\]

\[
< \varepsilon_1 + \varepsilon_2 = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad \text{q.e.d.}
\]
It is somewhat interesting to diagram \( f(x) + g(x) \) and note the flow of the tolerances which appear in the proof. We do this below, with tolerances in gray:

\[
\begin{align*}
(\vert x - a \vert < \delta &= \min\{\delta_1, \delta_2\}) \\
&
\end{align*}
\]

\[
\buildrel{Add( , )}\over{f(x) + g(x)} \xrightarrow{\vert f(x) - f(a) \vert < \frac{\varepsilon}{2}} \\
\xrightarrow{\vert g(x) - g(a) \vert < \frac{\varepsilon}{2}} \\
\vert (f(x) + g(x)) - (f(a) + g(a)) \vert < \varepsilon
\]

It is not difficult to see that this can be extended to include sums of more functions. Indeed, if \( i(x) = f(x) + g(x) + h(x) \) where \( f(x), g(x), h(x) \) are all continuous at \( x = a \), then so is the sum \((f(x) + g(x))\) by the previous theorem, and then again so will be the sum \([f(x) + g(x)] + (h(x)) = i(x)\) by that same theorem. (A proof “from scratch” can be done as well, following the same pattern but using \( \varepsilon_1, \varepsilon_2, \varepsilon_3 = \varepsilon/3 \) and \( \delta = \min\{\delta_1, \delta_2, \delta_3\} \), but it is faster to “bootstrap” later results from theorems which were already proved.)

The next theorem’s proof is less difficult, and is left as an exercise (Exercise 23, page 193):

**Theorem 3.2.2** If \( f(x) \) is continuous at \( x = a \), then so is \( Cf(x) \) for any constant \( C \).

The next theorem is the most difficult of these to prove, but will also be one of the most useful in practice.

**Theorem 3.2.3** Suppose that \( f(x) \) and \( g(x) \) are continuous at \( x = a \). Then so is \( f(x)g(x) \).

**Proof:** Let \( \varepsilon > 0 \), for the definition of continuity for the product \( h(x) = f(x)g(x) \). Now we carefully construct a \( \delta > 0 \) so that we can prove \( h(x) \) continuous at \( x = a \), i.e.,

\[
\begin{align*}
| x - a | < \delta &\implies | h(x) - h(a) | = | f(x)g(x) - f(a)g(a) | < \varepsilon \quad \text{(to prove!)}.
\end{align*}
\]

As before we are assuming that \( f(x), g(x) \) are continuous at \( x = a \), i.e.,

\[
\begin{align*}
(\forall \varepsilon_1 > 0)(\exists \delta_1 > 0)(\forall x)(| x - a | < \delta_1 \implies | f(x) - f(a) | < \varepsilon_1), \\
(\forall \varepsilon_2 > 0)(\exists \delta_2 > 0)(\forall x)(| x - a | < \delta_2 \implies | g(x) - g(a) | < \varepsilon_2).
\end{align*}
\]

First we note that what we need to control is \( | f(x)g(x) - f(a)g(a) | \), which can be
Now we prove the statement of the continuity of 
\[ f(x)g(x) - f(a)g(a) \]
expanded and then bounded using the triangle inequality as follows:
\[
= \frac{1}{2} |(f(x) - f(a))(g(x) + g(a)) + (f(x) + f(a))(g(x) - g(a))| \\
\leq \frac{1}{2} |(f(x) - f(a))(g(x) + g(a))| + \frac{1}{2} |(f(x) + f(a))(g(x) - g(a))| \\
= \frac{1}{2} |f(x) - f(a)| \cdot |g(x) + g(a)| + \frac{1}{2} |f(x) + f(a)| \cdot |g(x) - g(a)|.
\]
It is enough that the final line in the above be less than \( \varepsilon \).
We will choose \( \varepsilon_1, \varepsilon_2 \) based on the choice of \( \varepsilon \). Let us first assume \textit{a priori} that any \( \varepsilon_1, \varepsilon_2 \leq 1 \) in (3.7) and (3.8) and so with \( |x - a| < \min \{ \delta_1, \delta_2 \} \) we get by the triangle inequality
\[
|f(x) - f(a)| < 1 \implies |f(x)| = |f(a) + (f(x) - f(a))| < |f(a)| + 1, \\
|g(x) - g(a)| < 1 \implies |g(x)| = |g(a) + (g(x) - g(a))| < |g(a)| + 1.
\]
Now define \( L \) and \( M \) as follow:
\[
L = |f(a)| + 1 > 0, \\
M = |g(a)| + 1 > 0.
\]
From (3.9), (3.10) and (3.11), (3.12) we get \( |x - a| < \min \{ \delta_1, \delta_1 \} \implies \\
|f(x)| < L + 1, \quad |f(a)| < L, \\
|g(x)| < M + 1, \quad |g(a)| < M.
\]
Now we prove the statement of the continuity of \( f(x)g(x) \) at \( x = a \). For any \( \varepsilon > 0 \)
define
\[
\varepsilon_1 = \min \left\{ 1, \varepsilon, \frac{\varepsilon}{2M + 1} \right\} = \min \left\{ 1, \frac{\varepsilon}{2M + 1} \right\}, \\
\varepsilon_2 = \min \left\{ 1, \varepsilon, \frac{\varepsilon}{2L + 1} \right\} = \min \left\{ 1, \frac{\varepsilon}{2L + 1} \right\},
\]
and the respective \( \delta_1, \delta_2 \) from these \( \varepsilon_1, \varepsilon_2 \) as appear in the continuity conditions on \( f(x) \) and \( g(x) \) at \( x = a \), namely (3.7) and (3.8). Note that we used the fact that \( \varepsilon = \varepsilon/1 < \varepsilon/(2M+1) \), and similarly \( \varepsilon < \varepsilon/(2L+1) \), in (3.15) and (3.16) respectively. Finally, choose \( \delta = \min \{ \delta_1, \delta_2 \} \). Then \( \delta > 0 \) exists (because \( \delta_1, \delta_2 > 0 \)) and
\[
|x - a| < \delta \implies \\
|f(x)g(x) - f(a)g(a)| \\
\leq \frac{1}{2} |f(x) - f(a)| \cdot |g(x) + g(a)| + \frac{1}{2} |f(x) + f(a)| \cdot |g(x) - g(a)| \\
< \frac{1}{2} \varepsilon_1 \cdot (|g(x)| + |g(a)|) + \frac{1}{2} (|f(x)| + |f(a)|) \varepsilon_2 \\
\leq \frac{1}{2} \varepsilon_1 (2M + 1) + \frac{1}{2} \varepsilon_2 (2M + 1) \\
\leq \frac{1}{2} \cdot \frac{\varepsilon}{2M + 1} \cdot (2M + 1) + \frac{1}{2} \cdot \frac{\varepsilon}{2L + 1} \cdot (2L + 1) \\
= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad \text{q.e.d.}
\]
3.2. CONTINUITY THEOREMS

We could diagram \( f(x)g(x) \) as we did with \( f(x) + g(x) \), though the tolerances are more complicated. In fact the above proof is more difficult than most in this text, and requires careful examination to understand completely, or to be able to reproduce it. Contained within it lies a common method for controlling \( |f(x)g(x) - f(a)g(a)| \) by controlling \( |x - a| \) which is worth tracing. In particular we used the identity

\[
f(x)g(x) - f(a)g(a) = \frac{1}{2}(f(x) - f(a))(g(x) + g(a)) + \frac{1}{2}(f(x) + f(a))(g(x) - g(a)).
\]

All three continuity theorems so far can be described intuitively as saying, respectively (for functions continuous at \( x = a \)):

- a sum of functions which are continuous at \( a \) will also be continuous at \( a \);
- a constant multiple of a function which is continuous at \( a \) will also be continuous at \( a \);
- a product of functions which are continuous at \( a \) will also be continuous at \( a \).

Rephrased, if we can control how far functions stray from their values at \( x = a \)—by controlling how far \( x \) strays from \( a \)—then we can control how far their sums, products, and constant multiples of the functions stray from their respective values at \( x = a \) as well. For the next theorem we first need the following:

**Lemma 3.2.1** \((\forall a \in \mathbb{R})[f(x) = x \text{ is continuous at } x = a]\).

This is fairly trivial to prove, because we would just set \( \delta = \varepsilon \) in the definition of continuity. Details are left to the exercises.

**Theorem 3.2.4** *Polynomial functions are continuous at every \( a \in \mathbb{R} \).*

Hence the functions \( f(x) = x^2 + 1 \), \( g(x) = 55x^{39} + 101x - 10,000,000 \), and \( h(x) = (9 - 23x)^{15} \) are all continuous. Certainly this theorem gives a welcome relief from trying to prove these functions are continuous using \( \varepsilon \)-\( \delta \). The proof is really quite simple (and may even sound a bit flippant).

**Proof:** Any polynomial can be written

\[
f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0.
\]

Now each \( x^k \) can be written as a product of \( k \) factors of the continuous function \( g(x) = x \), so the \( x^k \) are all continuous for every \( x = a \), as are constant multiples of these, i.e., the \( a_kx^k \) terms. Of course \( h(x) = a_0 \) is a constant function and is therefore continuous (see Exercise 11, page 178), so a polynomial is just the sum of continuous functions, and is therefore continuous.

The next theorem is also very useful, and deals with compositions of functions. Compared with, say, Theorem 3.2.3 this is surprisingly simple to prove.

**Theorem 3.2.5** *Suppose that \( f(x) \) is continuous at \( x = L = g(a) \), and that \( g(x) \) is continuous at \( x = a \). Then \( f(g(x)) \) is continuous at \( x = a \).*

Put concisely, the composition of continuous functions is continuous.
**Proof:** Let $\varepsilon > 0$. We need to construct a $\delta > 0$ such that

$$|x - a| < \delta \implies |f(g(x)) - f(g(a))| < \varepsilon.$$ 

By our continuity assumptions, we know that

$$(\forall \varepsilon > 0)(\exists \delta_1 > 0)(\forall x)(|x - L| < \delta_1 \implies |f(x) - f(L)| < \varepsilon_1),$$

$$(\forall \varepsilon_2 > 0)(\exists \delta_2 > 0)(\forall x)(|x - a| < \delta_2 \implies |g(x) - g(a)| < \varepsilon_2).$$

So for this $\varepsilon$, choose $\varepsilon_1 = \varepsilon$, which gives a $\delta_1 > 0$ so that

$$|x - L| < \delta_1 \implies |f(x) - f(L)| < \varepsilon.$$ 

Next set $\varepsilon_2 = \delta_1 > 0$. This gives a $\delta_2 > 0$ so that

$$|x - a| < \delta_2 \implies |g(x) - g(a)| < \varepsilon_2 = \delta_1.$$ 

Finally, let $\delta = \delta_2$, corresponding to $\varepsilon_2$ in the continuity requirement for $g$. This gives $\delta > 0$ and the following (note the only necessary directions are $\implies$ but we note the steps for which we actually have $\iff$):

$$|x - a| < \delta \iff |x - a| < \delta_2$$

$$\implies |g(x) - g(a)| < \varepsilon_2 \quad \text{(since input } x \text{ of } g \text{ is } \delta_2\text{-close to } a)$$

$$\iff |g(x) - L| < \delta_1 \quad \text{(since } \varepsilon_2 = \delta_1)$$

$$\implies |f(g(x)) - f(L)| < \varepsilon_1 \quad \text{(since input } g(x) \text{ of } f \text{ is } \delta_1\text{-close to } L)$$

$$\iff |f(g(x)) - f(g(a))| < \varepsilon, \quad \text{q.e.d. (since } \varepsilon_1 = \varepsilon)$$

Intuitively, if $g(x)$ is changing continuously at $x = a$, and $f(\ )$ changes continuously at its input value $g(a)$, it seems reasonable that $f(g(x))$ should change continuously at $x = a$. Even less precisely, with the assumed continuity conditions of $f(\ )$ and $g(\ )$ we might expect that if $x \approx a$ then $g(x) \approx g(a)$, and then $f(g(x)) \approx f(g(a))$.

Now we have a proof, constructed as usual backwards from the function's flow, illustrated on the right. First we choose $\varepsilon > 0$ tolerance for the final output. With $\varepsilon_1 = \varepsilon > 0$ we have $\delta_1 > 0$ tolerance for the input of $f(\ )$ (near the value $g(a)$) guaranteeing $\varepsilon$ tolerance in the output of $f(\ )$. Letting $\varepsilon_2 = \delta_1 > 0$, we have $\delta_2 > 0$ tolerance in $x$ near $x = a$ allowing for at most $\varepsilon_2$ tolerance in the output of $g(x)$ to be near $g(a)$, guaranteeing $\varepsilon$-tolerance when this is fed to $f(\ )$.

We next work towards a simple theorem regarding quotients, which we argue deductively towards rather than stating and proving the steps in turn.\(^{10}\)

First recall from Exercise 16, page 178 that we have the following (stated here without proof):

\(^{10}\)When a thread contains theorems which are each difficult to prove, or contain some tangential clever technique, it is common for the discussion to proceed with a structure reading theorem-proof, theorem-proof, etc. When theorems flow from each other with minimal argument needed, the style often changes to short argument-result, short argument-result, etc.
3.2. CONTINUITY THEOREMS

Theorem 3.2.6

\[ f(x) = \frac{1}{x} \text{ is continuous for all } a \neq 0. \quad (3.17) \]

Continuing an argument which leads to a theorem on quotients, next we suppose that \( g(x) \) is continuous at \( x = a \), and that \( g(a) \neq 0 \). Then (3.17) and the previous theorem (Theorem 3.2.5 above) with \( f(x) = \frac{1}{x} \) conspire to give us that \( f(g(x)) = 1/g(x) \) is thereby continuous at \( x = a \):

Theorem 3.2.7 \( (g(x) \text{ continuous at } x = a) \land (g(a) \neq 0) \implies \frac{1}{g(x)} \text{ continuous at } x = a. \)

For arbitrary functions \( f(x) \) and \( g(x) \) which are continuous at \( x = a \), if \( g(a) \neq 0 \) we can always write \( \frac{f(x)}{g(x)} = f(x) \cdot \frac{1}{g(x)} \), a product of two functions now known to be continuous at \( x = a \), and so by our Theorem 3.2.3, page 181 that product is also continuous and we get another theorem:

Theorem 3.2.8 If \( f(x) \) and \( g(x) \) are continuous at \( x = a \), and \( g(a) \neq 0 \), then \( f(x)/g(x) \) is also continuous at \( x = a \).

This gives us a quick result on rational functions:

Theorem 3.2.9 If \( f(x) = \frac{p(x)}{q(x)} \), where \( p \) and \( q \) are polynomials, then \( f(x) \) is continuous at every \( a \in \mathbb{R} \) except where \( q(a) = 0 \), at which points \( f(x) \) is undefined and therefore discontinuous.

In other words, rational functions are continuous where defined. (The proofs of both of these are contained in the discussion above.) Next we mention the following reasonable definition:

Definition 3.2.1 If \( f(x) \) is not continuous at some point \( c \), then \( f(x) \) is called discontinuous at \( x = c \), and \( c \) is called a point of discontinuity of \( f(x) \).\footnote{The word “point” sometimes refers to (among other things) a value (point) on the real line (or an element of \( \mathbb{R} \)), and other times refers to a point in the \( xy \)-plane. For instance, one may describe \( f(x) = x^2 \) “at the point \( x = 3 \)” or “at the point \( (3,9) \).” In most cases it is an imprecision in the language which is cleared up by an understanding of the context, though occasionally two authors will have different but strong opinions on how to make the language precise.}

It is sometimes easier to list those points at which a function is discontinuous than the set of points at which it is continuous. Of course continuity at a point, and discontinuity at a point, are negations of each other.

Example 3.2.1 Find where the function \( f(x) = \frac{3x - 5}{x^2 - 1} \) is continuous.

Solution: As a rational function, \( f(x) \) is continuous except where \( x^2 - 1 = 0 \), i.e., except where \( x^2 = 1 \), i.e., except where \( x = -1, 1 \). We can effectively describe where \( f(x) \) is continuous in the following ways (all of which are equivalent):

a. \( f(x) \) is continuous except at \( x = -1, 1 \);

b. \( f(x) \) is continuous for \( x \in (-\infty, -1) \cup (-1, 1) \cup (1, \infty) \);

c. \( f(x) \) is continuous for all \( x \neq \pm 1 \).

It is important that we not simplify \( f(x) \) in any way before describing where it is continuous, lest problem points in the definition of \( f(x) \) be glossed over. For instance, consider the following:

\( f(x) = \frac{1}{x} \) is continuous for all \( a \neq 0 \).
Example 3.2.2 Find where \( f(x) = \frac{x^2 - 25}{x - 5} \) is continuous.

Solution: Clearly \( f(x) \) is not defined at \( x = 5 \), and therefore cannot possibly be continuous there. (The way in which \( f(a) \) appears in the definition of continuity requires that it exist.) Simplifying is desirable, but we must note that \( x = 5 \) is not in the domain. To be clear on this we can write

\[
f(x) = \frac{x^2 - 25}{x - 5} = \frac{(x + 5)(x - 5)}{x - 5} = x + 5, \quad x \neq 5.
\]

(Note that the cancellation step indeed required that \( x - 5 \neq 0 \), i.e., \( x \neq 5 \) because there we would essentially be dividing numerator and denominator by zero—not a valid arithmetic operation—when cancelling the \( x - 5 \) factors). Such a function is defined and continuous for \( x \in (-\infty, 5) \cup (5, \infty) \), i.e., for \( x \neq 5 \).

The proof of the next theorem was the subject of Exercise 18, page 179.

Theorem 3.2.10 Suppose that \( f(x) = \sqrt[n]{x} = x^{1/n} \), with \( n \in \mathbb{N} \).

(i) If \( n \) is odd, then \( f(x) \) is continuous at each \( x \in \mathbb{R} \).

(ii) If \( n \) is even, then \( f(x) \) is continuous at each \( x > 0 \), and discontinuous otherwise.

Thus \( f(x) = \sqrt[n]{x} \) is continuous for all \( x \in \mathbb{R} \), where \( g(x) = \sqrt[n]{x} = \sqrt[n]{x} \) is only continuous for \( x > 0 \). Though \( g(x) \) is defined at \( x = 0 \), but there is no room to the left of zero, so any interval \( (0 - \delta, 0 + \delta) \), with \( \delta > 0 \), will always have points outside the domain of \( g \) (namely \( x \in (0 - \delta, 0) \)). Thus with even roots we need to look at more than where they are defined to determine continuity. If we consider only \( x \geq 0 \), then for any \( \varepsilon > 0 \) we can find a \( \delta > 0 \) so that we can at least say \((|x - 0| < \delta) \land (x \geq 0) \implies |g(x) - g(0)| < \varepsilon \), and this is occasionally interesting, and so we have use for so-called one-sided continuity, the types of which are defined below.

### 3.2.2 One-Sided Continuity

At this point it is useful to introduce the following concepts.

**Definition 3.2.2** We call \( f(x) \) **left-continuous** at \( x = a \) if and only if

\[
(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)(x \in (a - \delta, a] \implies |f(x) - f(a)| < \varepsilon). \tag{3.18}
\]

Similarly, we call \( f(x) \) **right-continuous** at \( x = a \) if and only if

\[
(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)(x \in [a, a + \delta) \implies |f(x) - f(a)| < \varepsilon). \tag{3.19}
\]

Left-continuity at \( x = a \) only considers points at or to the left of \( a \) (\( x \leq a \)), while right-continuity considers those at or to the right of \( a \) (\( x \geq a \)). In the definition of continuity (3.1), page 169 both sides of \( x = a \) are considered. In fact that original definition of continuity can be rewritten

\[
(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)(x \in (a - \delta, a) \cup [a, a + \delta) \implies |f(x) - f(a)| < \varepsilon), \tag{3.20}
\]

from which we can see how left-continuity (3.18) and right-continuity (3.19) are related to continuity. Thus these one-sided continuity conditions are effectively two halves of the continuity requirement as written in (3.20).

One aspect of this theory which has not been mentioned explicitly as yet is that either type of continuity (the original two-sided continuity, left-continuity and right-continuity) at \( x = a \) require that \( f(a) \) must be defined, in order that \( |f(x) - f(a)| \) be true in the implications within the definitions. A function \( f(x) \) can have none of these continuity properties at \( x = a \) if \( f(a) \) is undefined. This will be an issue in identifying where some functions are continuous.
3.2. CONTINUITY THEOREMS

Example 3.2.3 Consider the function \( f(x) \) defined only on \( x \in [0, \infty) \), but on that interval given by \( f(x) = x^2 \). In other words, \( f(x) = \begin{cases} \ x^2 & \text{if } x \geq 0, \\ \text{undefined} & \text{if } x < 0. \end{cases} \)

- Then \( f(x) \) is continuous at \( x = a \) for any \( a > 0 \).
  
  - Intuitively, we can see from the graph that we have “wiggle room” in \( x \) as long as \( a > 0 \) for which \(|x - a| < a\) still has \( f(x) = x^2 \), which is clearly a continuous function; we can be guaranteed tolerances as small as we like in \( f(x) \) being close to \( f(a) \) by forcing \( x \) to be close to, but not necessarily equal to, \( a \). This is not the case if \( a = 0 \), as we would have no “wiggle room” to the left of \( x = 0 \).

- More technically this is because in the definition of continuity we can assume a priori that \( \delta \leq a \), since \( a > 0 \), and then \(|x-a| < \delta \implies f(a) \) is defined, and so we can borrow continuity results from our earlier theorem on the continuity of polynomials. To do so properly, we should consider the polynomial function \( g(x) = x^2 \) defined on all of \( \mathbb{R} \). Now \( g(x) \) is continuous for any \( x = a \), with \( a \in \mathbb{R} \). Thus for any \( a \in \mathbb{R} \), if we let \( \varepsilon > 0 \) there exists \( \delta_g > 0 \) (“the Delta for \( g \)” such that \(|x - a| < \delta_g \implies |g(x) - g(a)| < \varepsilon \). Now for our \( f(x) \) defined above, and any \( a > 0 \) we can take \( \delta = \min\{\delta_g, a\} \), so \( \delta > 0 \) exists and

\[
|x-a| < \delta \implies \left[ (|x-a| < \delta \leq \delta_g) \land (x > 0) \right] \implies |f(x) - f(a)| = |g(x) - g(a)| < \varepsilon, \quad \text{since } x > 0, \quad \text{since } \delta \leq \delta_g
\]

- \( f(x) \) is both left-continuous and right-continuous (as well as continuous in our original two-sided sense) at each \( a > 0 \). That should be clear because continuity is a stronger condition than the one-sided continuities. (See Theorem 3.2.11 below.)

- \( f(x) \) is right-continuous (but not continuous nor left-continuous) at \( x = 0 \). We only need to allow tolerance (“wiggle room”) allowing for \( x \) to vary to the right of \( x = 0 \) (see graph). In fact, \( \delta = \sqrt{\varepsilon} \) proves right-continuity: Let \( \varepsilon > 0 \), and choose \( \delta = \sqrt{\varepsilon} \), so \( \delta > 0 \) exists and

\[
x \in [0, \delta) \implies |f(x) - f(0)| = x^2 < \delta^2 = \varepsilon, \quad q.e.d.
\]

- \( f(x) \) has no continuity properties for \( x < 0 \), because the function is not defined there.

Continuity is, in fact, a “local” phenomenon; what matters for \( x = a \) is the behavior of the function “near \( x = a \).” This is because we can always restrict \( \delta > 0 \) to be as small as we like a priori, in essence ignoring what occurs outside of \((a - \delta, a + \delta)\) no matter how small we make \( \delta \), while keeping it positive. Even so restricted, we are still allowing for (and in fact requiring) a continuum containing \( a \), extending from \( a \) in one or both directions, depending upon if we are considering one-sided continuity or the original, two-sided continuity.

Of course the different types of continuities are related. We leave for an exercise the following theorem:
Theorem 3.2.11 \( f(x) \) is continuous at \( x = a \) if and only if \( f(x) \) is both left-continuous and right-continuous at \( x = a \).

There are occasions where we check continuity by checking both one-sided continuity conditions, and other settings where we require only one-sided continuity but have that easily because we have two-sided continuity, so the theorem above has its uses.

Example 3.2.4 Consider the function \( f(x) = \begin{cases} (x-2)^2 - 1 & \text{if } x \geq 2, \\ 1-x & \text{if } x < 2. \end{cases} \)

- At \( x = 0 \) this function is continuous, because it is essentially \( f(x) = 1 - x \) for \( x \in (-\infty, 2) \), and \( x = 0 \) is safely inside this interval (in and surrounded by a continuum) in which \( f(x) \) is simply the polynomial \( 1 - x \). If we wished to prove continuity at \( x = 0 \) from the definition, we can always choose \( \delta \leq 2 \) so that when we write \(|x - 0| < \delta \) we know we are within an interval in which \( f(x) \) is a fixed polynomial and therefore continuous. (In fact \( \delta = \min\{2, \varepsilon\} \) would work in the continuity definition.)

- At \( x = 3 \), this function is again continuous, because we are safely inside of \([2, \infty)\), and we could if necessary take \( \delta \leq 1 \) so that \(|x - 3| < \delta \Rightarrow x \in [2, \infty) \Rightarrow f(x) = (x - 2)^2 - 1 \), i.e., \( f(x) = x^2 - 4x + 4 - 1 = x^2 - 4x + 3 \), ultimately a polynomial, albeit one which requires a more careful choice of \( \delta > 0 \) for a given \( \varepsilon > 0 \) in the definition of continuity.

- At \( x = 2 \), we have to be more careful. Note that \( f(2) = (2 - 2)^2 - 1 = -1 \), which coincides with \( 1 - x \) where \( x = 2 \) as well. This means we could have used \( x \leq 2 \) as the condition for which \( f(x) = 1 - x \), because it did not matter which formula we applied at \( x = 2 \). Thus

\[
\begin{align*}
(i) \quad & x \in [2, \infty) \Rightarrow f(x) = (x - 2)^2 - 1, \\
(ii) \quad & x \in (-\infty, 2] \Rightarrow f(x) = 1 - x.
\end{align*}
\]

From (i) we know that \( f(x) \), being a polynomial on \([2, \infty)\), is right-continuous at \( x = 2 \); and from (ii) we know that \( f(x) \), being a polynomial on \((-\infty, 2]\), is left-continuous at \( x = 2 \). Since \( f(x) \) is both left-continuous and right-continuous at \( x = 2 \), we conclude that \( f(x) \) is continuous (in the original, two-sided sense) at \( x = 2 \).

In essence, if a function is defined piece-wise and the defining formulas agree with \( f(a) \) in their outputs at a point \( a \) on the boundary of the formulas’ respective intervals, and if the defining formulas would themselves define appropriately continuous functions (left-continuous at \( a \) in the left-piece formula, and right-continuous at \( a \) in the right-piece formula), then \( f(x) \) will be continuous at \( x = a \). In the above example, we could simply say that the two formulas to the left and right of \( x = 2 \) would both yield \( f(2) \) if we substituted \( x = 2 \) into them. Visually, the two pieces should “meet” at \((2, f(2))\), as in the graph on the right of the function in the above example.

While it is useful to consider the graphs when determining continuity, so far we have stressed the analysis without resorting to graphs. We will utilize graphs more in the next sections, but for
3.2. CONTINUITY THEOREMS

Now we will attempt to develop a “number sense” and “function sense” in a manner somewhat independent from the graphs. For instance, if we return again to the even roots \( f(x) = \sqrt[n]{x} \), (where \( n \) is even) we see that, though these functions are not continuous at \( x = 0 \), they are right-continuous at \( x = 0 \). When dealing with functions which contain radicals, for continuity we often need only see where the functions are defined, and where we have some “wiggle room” on both sides of a particular \( x = a \). Odd roots are defined and (two-sided) continuous for any real input so they do not in themselves limit where a function is continuous. Consider the following:

Example 3.2.5 For each function, find where it is continuous.

\[
\begin{align*}
  f(x) &= \frac{1}{\sqrt{x-1}} \\
  g(x) &= \sqrt{2x+1} \\
  h(x) &= \sqrt{20-4x}
\end{align*}
\]

\[
\begin{align*}
  i(x) &= \sqrt{x^2+1} \\
  j(x) &= \frac{1}{\sqrt{x^2-1}} \\
  k(x) &= 3\sqrt[3]{x^9-27x^4+x-11}.
\end{align*}
\]

**Solution:** We take each of these in turn, making use of our previous continuity theorems on sums, products, compositions and quotients of functions, as well as results on polynomials and roots.

- The function \( f(x) \) is a ratio of functions which are continuous everywhere, so we need only check where the denominator is zero. Thus \( f(x) \) is discontinuous only at \( x = 1 \). (In other words, \( f(x) \) is continuous on \((-\infty, 1) \cup (1, \infty) \), but we will be content in such a case to simply state where it is discontinuous.)

- The function \( g(x) \) is definitely continuous where \( 2x + 1 > 0 \), i.e., \( 2x > -1 \), i.e., \( x > -1/2 \). It is not continuous at \( x = -1/2 \) because it \( g(x) \) is undefined to the left of \( x = -1/2 \) as a quick check will show. \( g(x) \) is right-continuous at \( x = -1/2 \), but not continuous there.) Conclude \( g(x) \) is continuous for \( x > -1/2 \).

- The function \( h(x) \) is definitely continuous for \( 20 - 4x > 0 \), i.e., \( -4x > -20 \), i.e., \( x < 5 \). It is not continuous at \( x = 5 \) since it is undefined to the right of that point. (It is left-continuous at \( x = 5 \), but again, that was not the question.) Conclude \( h(x) \) is continuous where \( x < 5 \).

- The function \( i(x) \) is continuous everywhere, since \( x^2 + 1 \) is continuous everywhere and for all \( x \in \mathbb{R} \) we also have \( x^2 + 1 > 0 \). (In fact \( x^2 + 1 \geq 1 > 0 \).)

- The functions \( j(x) \) is definitely discontinuous at \( x = \pm 1 \) because the denominator is zero there. With that consideration, and that of the square root disallowing its input being negative, for continuous we need \( x^2 - 1 > 0 \), i.e., \( x^2 > 1 \), which occurs when \( x > 1 \) or \( x < -1 \).

- The function \( k(x) \) is continuous everywhere, being an odd root of a (continuous) polynomial.

Example 3.2.6 Find where the function \( f(x) = \frac{\sqrt{9-x^2}}{x^2-1} \) is continuous.

**Solution:** For denominator considerations, we need only that \( x \neq \pm 1 \). The numerator is definitely continuous for \( 9 - x^2 > 0 \), i.e., \( 9 > x^2 \), i.e., \(-3 < x < 3 \). It is not continuous at \( x = \pm 3 \) since both are on the edge of the domain (no “wiggle room”). Putting this together, we see \( f(x) \) is continuous for \( x \in (-3, 3) - \{-1, 1\} = (-3, -1) \cup (-1, 1) \cup (1, 3) \).
To be clear, we next look at a function which is the reciprocal of the previous function to illustrate the differences.

**Example 3.2.7** Find where the function \( f(x) = \frac{x^2 - 1}{\sqrt{9 - x^2}} \) is continuous.

**Solution:** For this function, the numerator is continuous for all \( x \in \mathbb{R} \), so there is no restriction on \( x \) implied by the form of the numerator. (Numerators can be equal to zero, while denominators cannot, for a quotient to be defined.) Clearly \( x \neq -3, 3 \) due to the denominator, lest we attempt to divide by zero. Requiring the input of the square root to be nonnegative forces \( -3 \leq x \leq 3 \), but we already disallowed \( x = \pm 3 \) due to the restriction against dividing by zero, though an alternative explanation here would be the lack of “wiggle room” either values have, \( -3 \) to the left and 3 to the right, avoiding again have a negative input for the square root function \( \sqrt{\cdot} \).

Our conclusion is that this \( f(x) \) is continuous for \( x \in (-3, 3) \).

A useful corollary to Theorem 3.2.10 is the fact that \( f(x) = |x| = \sqrt{x^2} \) is continuous for all \( x \in \mathbb{R} \). To see this, note that clearly \( |x| \) is continuous for \( x \neq 0 \), for which \( x^2 > 0 \). But there is “wiggle room” at \( x = 0 \) as well. In fact it is easier to see that \( |x| \) is both left- and right-continuous at \( x = 0 \) if we write

\[
|x| = \begin{cases} 
  x & \text{if } x \geq 0, \\
  -x & \text{if } x \leq 0.
\end{cases}
\] (3.21)

As a matter of form, we usually write definitions such as (3.21) with the “if” conditions being mutually exclusive (non-overlapping, as if we were instructing a computer), but we can see that at \( x = 0 \) both formulas give the same result. We can then see that on \( x \in [0, \infty) \) we can write \( |x| = x \), which is obviously right-continuous at \( x = 0 \). Similarly, on \( x \in (-\infty, 0] \) we can write \( |x| = -x \), which is obviously left-continuous at \( x = 0 \). Since \( |x| \) as a function is both left-continuous and right-continuous at \( x = 0 \), it is continuous there.

Furthermore, at any \( x \neq 0 \) we are “safely” within either branch: where \( x > 0 \implies |x| = x \) which is continuous, or where \( x < 0 \implies |x| = -x \) which is also continuous. From these three cases \((x = 0, x > 0, x < 0)\) we can see that \( |x| \) is continuous at each \( x \in \mathbb{R} \), as we state below:

**Theorem 3.2.12** The function \( f(x) = |x| \) is continuous for all \( x \in \mathbb{R} \).

By our theorem on function compositions (Theorem 3.2.5 on page 183), we have that the absolute value of a continuous function is also continuous:

**Theorem 3.2.13** If \( g(x) \) is continuous at \( x = a \), then \( f(x) = |g(x)| \) is also continuous at \( x = a \).

**Example 3.2.8** \( f(x) = \sqrt{x^2 - 2x + 1} \) is continuous on \( \mathbb{R} \) since \( f(x) = \sqrt{(x - 1)^2} = |x - 1| \) is the absolute value of a continuous function and is therefore continuous.

### 3.2.3 Essential versus Removable Discontinuities

Many functions which we encounter will have a discontinuity at a particular point, but where the discontinuity will be **removable**. This means that if we were to redefine the function at that point to have a well chosen output value, it would cease to be discontinuous there.
Example 3.2.9 Consider the function \( f(x) = \frac{x^2}{|x|} \). Clearly this function is undefined at \( x = 0 \), because we cannot divide by zero. Note that this function can instead be defined piece-wise:

\[
f(x) = \begin{cases} 
  x^2/x & \text{if } x > 0 \\
  x^2/(-x) & \text{if } x < 0
\end{cases}
= \begin{cases} 
  x & \text{if } x > 0 \\
  -x & \text{if } x < 0.
\end{cases}
\]

In other words, \( f(x) = |x| \) for \( x \neq 0 \). We say that \( x = 0 \) is a removable discontinuity because if we were to simply redefine \( f \) so that \( f(0) = 0 \), we would have a continuous function at \( x = 0 \).

Clearly the only value we could use to redefine the output of \( f(x) \) at \( x = 0 \) is \( f(x) = 0 \).

Example 3.2.10 Consider the function \( f(x) = \frac{x}{|x|} \). This function is also undefined at \( x = 0 \).

Defined piece-wise, we have the easily graphed function

\[
f(x) = \begin{cases} 
  x/x & \text{if } x > 0 \\
  x/(-x) & \text{if } x < 0
\end{cases}
= \begin{cases} 
  1 & \text{if } x > 0 \\
  -1 & \text{if } x < 0.
\end{cases}
\]

This is graphed on the left. In order for the function to be right-continuous at \( x = 0 \) we would require \( f(0) = 1 \), but for left-continuity we would require \( f(0) = -1 \).

Clearly we cannot have both (or we would not have a function!) and so the discontinuity at \( x = 0 \) is called essential.\(^\text{12}\)

It is usually not difficult spot removable and essential discontinuities from the graph of the function: if we can “fill in a hole” to produce a function which is continuous at the point, then the discontinuity is removable; otherwise it is essential. However, one goal of the analysis is to be able to make such determinations based upon formulas for the functions, rather than based upon a requirement that we always produce a graph first.\(^\text{13}\) At this point we are ready for our definitions:

\(^{12}\text{Another common term for essential discontinuities is nonremovable. Both terms have their respective advantages.}\)
\(^{13}\text{There are approaches for learning calculus based upon exploring most problems with graphical calculators or software first—in essence, approaching calculus first as a visual exercise—even though it does not take long for one to encounter functions complicated enough that a reliance on electronically-produced graphs is cumbersome or misleading. The calculator-based calculus instruction is especially popular in high school calculus courses, where there is an assumption that algebraic skills may be lacking, and that students will be more interested when they have a visual reference. One drawback is that the user can be fooled by limitations of a calculator display’s resolution, or the analytical shortcomings of today’s common graphing calculators.}\)
Definition 3.2.3 Given a function \( f(x) \), and a discontinuity \( x = a \) of \( f(x) \), then

- we call \( x = a \) a **removable discontinuity** if there exists \( y_0 \in \mathbb{R} \) such that, if \( f(x) \) were redefined at \( x = a \) so that \( f(a) = y_0 \), then \( f(x) \) would be continuous at \( x = a \);
- we call \( x = a \) an **essential (or nonremovable) discontinuity** if no such \( y_0 \) exists. That is, no matter how we (re)define \( f(x) \) at \( x = a \), i.e., regardless of any redefinition \( f(a) = y_0 \), the function \( f(x) \) is still discontinuous at \( x = a \).

It is assumed in the definition that \( f(x) \) is only redefined at \( x = a \), and so the output of \( f(x) \) will be the same as before for all \( x \neq a \).

Note that if \( f(x) \) has a removable discontinuity at \( x = a \), then the redefinition is unique, that is there exists exactly one \( y_0 \) such that the redefinition \( f(a) = y_0 \) removes the discontinuity of \( f(x) \) at \( x = a \). That may seem intuitive from various graphs of removable discontinuities, but we will find tools later (namely limit theorems in later sections) which make a proof of this fact easier.

Examples of essential discontinuities include **jump discontinuities** and **vertical asymptotes**:

With a jump discontinuity, the function can be redefined as either left-continuous or right-continuous, but not both. In the case above, if we filled in the “hole” we would no longer have a function, since we would have two outputs (\( y \)-values) for that particular input \( a \).

Vertical asymptotes of any kind are always essential discontinuities.

Some nonremovable discontinuities of we encounter are not terribly interesting. For instance, \( x = -1 \) is not a removable discontinuity of \( f(x) = \sqrt{x} \), because it is well outside of the domain of \( \sqrt{x} \). For that matter, \( x = 0 \) is an essential discontinuity of \( f(x) = \sqrt{x} \) by our definition, because if is on the boundary of the domain of \( \sqrt{x} \) (and no redefinition at \( x = 0 \) will change that it is a discontinuity). However \( \sqrt{x} \) is right-continuous there. We could define left-removable or right-removable discontinuities, but we will not bother to do so here. Where applicable, we will later have other language for such phenomena.
3.2. CONTINUITY THEOREMS

Exercises

For 1–16, find all $x$ for which the function is continuous.

1. $f(x) = \frac{x^2 - 4}{x^2 + 4}$.
2. $f(x) = \frac{1 + x}{1 - x^2}$.
3. $f(x) = \sqrt{3x - 7}$.
4. $f(x) = \sqrt{9 - 3x}$.
5. $f(x) = \sqrt{16 - x^2}$.
6. $f(x) = \sqrt{16 + x^2}$.
7. $f(x) = \sqrt{x^2 + 6x + 9}$.
8. $f(x) = \frac{1}{\sqrt{x^2 + 6x + 9}}$.
9. $f(x) = \sqrt[3]{25 - x^2}$.
10. $f(x) = \frac{x^2 - 1}{\sqrt[3]{25 - x^2}}$.
11. $f(x) = |x^2 - 7x + 12|$.
12. $f(x) = \sqrt{|x|}$.
13. $f(x) = \sqrt{\frac{1 - x}{1 + x^2}}$.
14. $f(x) = \frac{\sqrt{1 - x}}{1 + x^2}$.
15. $f(x) = \frac{\sqrt{1 - x}}{1 + x^2}$.
16. $f(x) = \frac{1}{\sqrt{x^2 - 1}}$.

For Exercises 17–20, determine if the functions are continuous at the indicated points, and if not, whether the discontinuity is essential or removable.

17. $f(x) = \begin{cases} -6 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 6 & \text{if } x > 0. \end{cases}$
18. $f(x) = \begin{cases} \frac{x^2 - 1}{x + 1} & \text{if } x \neq -1 \\ -2 & \text{if } x = -1, \end{cases}$ at $x = -1$.
19. $f(x) = \begin{cases} -x^2 & \text{if } x < 0 \\ 1 - \sqrt{x} & \text{if } x \geq 0, \end{cases}$ at $x = 0$.
20. $f(x) = \begin{cases} x^3 + 2 & \text{if } x < 2 \\ 8 & \text{if } x = 2 \\ 5x & \text{if } x > 2, \end{cases}$ at $x = 2$.

21. For what value of $A$ is the following function continuous on all of $\mathbb{R}$?

\[ f(x) = \begin{cases} 2x^3 & \text{if } x < 1 \\ Ax - 3 & \text{if } x \geq 1. \end{cases} \]

22. Using $\varepsilon$-$\delta$ and the definition of right-continuity, show that $f(x) = \sqrt{x}$ is right-continuous at $x = 0$.

23. Prove Theorem 3.2.2. Here is the general strategy:

(a) For $C = 0$, refer to Exercise 11, page 178.

(b) If $C \neq 0$, let $\varepsilon_1 = \frac{|C|}{2|x|}$, and find the $\delta_1$ corresponding to this $\varepsilon_1$ satisfying the definition of continuity for $f(x)$. Then take $\delta = \delta_1$.

24. Prove Lemma 3.2.1, page 183, i.e., that $f(x) = x$ is continuous at each $a \in \mathbb{R}$.


26. Prove Theorem 3.2.11. Hint: Let $\delta = \min\{\delta_1, \delta_2\}$ where $\delta_1$ and $\delta_2$ come from the one-sided continuity conditions when proving the “if,” and let $\delta_1, \delta_2 = \delta$ for the “only if.”
3.3 Continuity on Intervals

“A function φ(x) is said to be continuous between any limiting values of x, such as a and b, when to each value of x between those limits there corresponds a finite value of the function, and when an indefinitely small change in the value of x produces only an indefinitely small change in the function. In such cases the function in its passage from any one value to any other between the limits receives every intermediate value, and does not become infinite. This continuity can be readily illustrated by taking φ(x) as the ordinate of a curve, whose equation may then be written y = φ(x).”


3.3.1 Continuity on Intervals: Main Definitions and Theorems

The significance of continuity is perhaps best understood when applied to whole intervals (a, b), or [a, b], etc., rather than single points. Below we will define what it means for \( f(x) \) to be continuous on (a, b), and continuous on [a, b]. (Other cases like [a, b) would be defined in ways which would be obvious extensions once we have the (a, b) and [a, b] cases.) Continuity on open intervals is rather trivial to define, but nonetheless has interesting consequences. In practice we will focus more on continuity on finite closed intervals [a, b], properly defined, with a < b. But first we look at the open intervals.

**Definition 3.3.1** A function \( f(x) \) is said to be continuous on the open interval (a, b) if and only if \( f(x) \) is continuous at each value \( x_0 \in (a, b) \).

What is interesting about continuity on (a, b) is it implies \( f((a, b)) = \) is an interval of some kind (possibly infinite, or even a single point), and that the curve \( y = f(x) \) is a connected graph for \( a < x < b \).

**Theorem 3.3.1** If \( f(x) \) is continuous on (a, b), then \( f((a, b)) \) is an interval and the graph of \( y = f(x) \), \( a < x < b \) is a connected curve.

Unfortunately the proof is a few steps beyond the scope of this textbook and is left, for the interested reader, to a course in advanced calculus, real analysis or topology. Still, with the previous continuity sections behind us we should at least hear the ring of truth in the notion that the graph of \( f(x) \), with domain restricted to \( x \in (a, b) \), would be connected if \( f \) is continuous on (a, b). Put another way, we could draw the graph without lifting our pen from the paper. This is consistent with the elegant quote given at the top of this section, particularly the words, “...when an indefinitely small change in the value of \( x \) produces only an indefinitely small change...”

---

14Of course each \( x_0 \) may require a different \( \delta \) for a given \( \varepsilon \) in the original definition of continuity given in Section 3.1, but the above definition only requires that each individual \( x_0 \in (a, b) \) is a point at which \( f(x) \) satisfies the \( \varepsilon-\delta \). To summarize, \( (\forall x_0 \in (a, b))(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)(|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon) \).

15Notationally, if \( S \) is a subset of the domain of \( f(x) \), then \( f(S) = \{ y | (\exists x \in S)(y = f(x)) \} \). In other words, \( f(S) \) is the set of all possible outputs of \( f( ) \) if the inputs are taken from \( S \). For instance, if \( f(x) = 2x + 3 \), then \( f((0, 1)) = (3, 5) \), since \( x \in (0, 1) \implies y \in (3, 5) \), as can easily be seen by graphing \( y = 2x + 3 \) on the interval \( x \in (0, 1) \). Similarly, if \( f(x) = \sin x \) then \( f([-1, 1]) = [-1, 1] \).

16The preliminaries required to prove these are not terribly difficult, but would require a distracting amount of effort here.
in the function”; our pens would not “jump” off the page to a radically different height as we move slowly along the curve by increasing \( x \) through the interval \((a, b)\). As a consequence it is then reasonable that the range will be a (connected) interval as the theorem also states.

Figure 3.2 shows sample cases for continuity on open intervals, illustrating that the image is always an interval. Now we turn our attention to continuity on closed intervals \([a, b]\):

**Definition 3.3.2** A function \( f : [a, b] \to \mathbb{R} \), with \( a < b \) is called **continuous on the (finite) closed interval** \([a, b]\) if and only if each of the following are true:

1. \( f(x) \) is continuous on the open interval \((a, b)\), i.e., continuous at each \( x \in (a, b) \), and
2. \( f(x) \) is right-continuous at \( x = a \), and
3. \( f(x) \) is left-continuous at \( x = b \).

In other words, if we are on the interior of \([a, b]\), i.e., on \((a, b)\), then we expect our original two-sided continuity at every point therein, but then we require only right-continuity at \( a \), and left-continuity at \( b \). In this way we ignore the behavior of \( f(\ ) \) outside of inputs from strictly from \([a, b]\).

Note that we can also define continuity on \([a, b]\) as requiring our original (two-sided) continuity on \((a, b)\) and left-sided continuity at \( b \). Similarly for continuity on \([a, b]\).

Two very important results which follow from continuity on closed intervals \([a, b]\) are the Intermediate Value Theorem (IVT) and the Extreme Value Theorem (EVT). Both are wrapped up nicely in the following analog to Theorem 3.3.1, the difference being that the previous theorem required continuity on \((a, b)\), while this theorem requires continuity on \([a, b]\). That minor difference gives us a much stronger theorem (though again we omit the proof):

**Theorem 3.3.2** If \( f(x) \) is continuous on \([a, b]\), where \( a < b \), then \( f([a, b]) \) is a closed interval of the form \([c, d]\) (with the possibility that \( c = d \) in the case \( f(x) \) is constant on \([a, b]\)).

In other words, the continuous image of a closed and bounded interval \([a, b]\) is also a closed and bounded interval \([c, d]\). If we again think about graphing such a function we can see that
this is also believable. After all, we would have to “pin down” the first point \((a, f(a))\), draw continuously until we end by “pinning down” the last point \((b, f(b))\). In doing so we would somewhere draw the highest and lowest points of our curve, and these heights would be \(d\) and \(c\), respectively, from our theorem. Albeit we may hit those high and low vertical levels repeatedly, and maybe not at \(x = a\) or \(x = b\), but they will be achieved nonetheless.

The proof of Theorem 3.3.2 is also beyond the scope of this text, but the EVT and IVT follow very quickly from the theorem. We list these very important facts as corollaries.

**Corollary 3.3.1 (Extreme Value Theorem)** If \(f : [a, b] \rightarrow \mathbb{R}\) is continuous on \([a, b]\), then \(f(x)\) achieves its maximum and minimum heights at some \(x_{\text{max}}, x_{\text{min}} \in [a, b]\). In other words,

\[
\left( \exists x_{\text{min}}, x_{\text{max}} \in [a, b] \right) \left( f([a, b]) = [f(x_{\text{min}}), f(x_{\text{max}})] \right).
\]

**Corollary 3.3.2 (Intermediate Value Theorem)** If \(f(x)\) is continuous on \([a, b]\), then \(f(x)\) achieves every value between \(f(a)\) and \(f(b)\). In other words, if \(y_0\) is between \(f(a), f(b)\), then there exists \(x_0 \in [a, b]\) such that \(f(x_0) = y_0\).

(Note that the statement of IVT is contained in the second sentence of the quote from Britannica on page 194, and the EVT is also hinted at in the quote.) A sample function which demonstrates the theorem and two corollaries is given in Figure 3.3.

**Figure 3.3:** A typical graph of a function which is continuous on a finite, closed interval \([a, b]\). The image \(f([a, b])\) is also a finite, closed interval (Theorem 3.3.2), containing a maximum value \(f(x_{\text{max}})\) and a minimum value \(f(x_{\text{min}})\), which are achieved at some \(x_{\text{max}}, x_{\text{min}} \in [a, b]\) (Extreme Value Theorem, Corollary 3.3.1). Also note that all values between \(f(a)\) and \(f(b)\) (and more, for this particular graph) are in the image \(f([a, b])\), and thus achieved for some \(x\)-values between \(a\) and \(b\) (Intermediate Value Theorem, Corollary 3.3.2).
Figure 3.4: Theorems 3.3.1 and 3.3.2—and their corollaries EVT and IVT—require continuity of \( f \), without which the image can be disconnected or unbounded. In the first graph, there is a vertical asymptote at \( x = 1 \), which is a point of discontinuity. The image \((−∞,−1/2) \cup [1/2,∞)\) is disconnected (so no IVT conclusion) and unbounded (so no EVT conclusion). In the second case, there is a “jump” discontinuity at \( x = 1 \), and the image \([0,1] \cup [2,2.5]\) is also disconnected (no IVT conclusion), though bounded. It happens that the second graph does have maximum and minimum values \( f(0) = 0 \) and \( f(3.5) = 2.5 \) though this was not guaranteed because \( f(x) \) is not continuous on all of \([-1,3.5]\). These examples do not violate the corollaries IVT and EVT since both corollaries claim the truth of tautologies of the form \( P \rightarrow Q \), which is equivalent to \((\sim P) \lor Q\), and here we have \( \sim P \) in both cases, making them true to the theorems and corollaries vacuously (in form \( P \rightarrow Q \)) or trivially (using the form \((\sim P) \lor Q\)).

or minimum is actually achieved; in the third graph, the image is unbounded from above so no maximum is achieved, and no minimum is achieved either. It may happen that \( f((a,b)) \) is a closed and bounded interval, as in the second graph in that figure, but it clearly (from the other two graphs in that figure) is not guaranteed. Continuity is also required in these theorems, as we see in Figure 3.4.

Note that the first function in the figure is continuous on \([-1,1]\) because it is continuous on \((-1,1)\) and right-continuous at \(-1\). Similarly it is continuous on \((1,3]\). The second function is continuous on \([-1,1]\) and \([1,3.5]\). That is not to say it is continuous at each \( x \in [1,3.5]\), but rather that the “piece” drawn on that interval is a continuous “piece,” in the sense of Definition 3.3.2, page 195 and the discussion following that.

3.3.2 Simple Applications of the Intermediate Value Theorem

We will return to the extreme value theorem and its applications later in the text. Here we will instead look at the IVP and its usefulness in algebra. The following simple theorem can often be useful when we look at continuity considerations:

**Theorem 3.3.3** If \( I \) and \( J \) are intervals of any kind except for single points, with \( I \subseteq J \), and \( f : J \rightarrow \mathbb{R} \) is continuous on \( J \), then \( f : I \rightarrow \mathbb{R} \) is continuous on \( I \).
In other words when a function is known to be continuous on an interval, its restriction to a subinterval is also continuous on that subinterval. The proof is a matter of chasing down the definitions of continuity on the various intervals, and checking each of the cases. The following example shows a quick application.

Example 3.3.1 Show that the equation \( x^5 + 7 = x^2 \) has a solution in \( \mathbb{R} \).

Solution: First notice that
\[
x^5 + 7 = x^2 \iff x^5 - x^2 + 7 = 0.
\]
Next define \( f(x) = x^5 - x^2 + 7 \), which is a polynomial and thus continuous on all of \( \mathbb{R} \) (think of \( \mathbb{R} \) as the open interval \(( -\infty, \infty )\)). Now
\[
x^5 + 7 = x^2 \iff x^5 - x^2 + 7 = 0 \iff f(x) = 0,
\]
so solving the original equation is equivalent to finding \( x \) so that \( f(x) = 0 \). Next notice that
\[
f(-2) = (-2)^5 - (-2)^2 + 7 = -29,
\]
and
\[
f(1) = 1 - 1 + 7 = 7.
\]
Because \( f \) is continuous in \( \mathbb{R} \), it is also continuous on \([-2, 1] \subseteq \mathbb{R} \). Since \( f(-2) = -29 \) while \( f(1) = 7 \), there exists some \( x_0 \) between \(-2 \) and \( 1 \) such that \( f(x_0) = 0 \) (by IVT), and this \( x_0 \) therefore solves the original equation, q.e.d.

The technique in Example 3.3.1 is standard. It is similar to the usual algebraic trick where we try to solve an equation of the form \( g(x) = h(x) \) by instead determining where \( g(x) - h(x) = 0 \). It is often convenient to define \( f(x) = g(x) - h(x) \) and solve (or, as here, simply detect the presence of a solution of) the logically equivalent equation \( f(x) = 0 \). As we saw in the algebra sections of Chapter 2, solving \( f(x) = 0 \) is usually simpler than finding where \( g(x) = h(x) \).

We can apply the IVT more than once to a single function, as in the following example:

Example 3.3.2 Show that the equation \( x^3 - 8x^2 + 15x = -1 \) has at least three solutions.

Solution: This is equivalent to the equation \( x^3 - 8x^2 + 15x + 1 = 0 \) having at least three solutions. Defining \( f(x) = x^3 - 8x^2 + 15x + 1 \), we thus want to prove that \( f(x) = 0 \) occurs at least three times.

First we notice that \( f(x) \) is a polynomial, and therefore continuous on \( \mathbb{R} \). Next we see that
\[
\begin{align*}
f(-1) &= -23, \\
f(0) &= 1, \\
f(4) &= -3, \\
f(5) &= 1.
\end{align*}
\]
We see that \( f(x) \) must be zero for some \( x \in (-1, 0) \), another \( x \in (0, 4) \), and yet another \( x \in (4, 5) \), ultimately therefore proving that there must be at least three solutions of \( x^3 - 8x^2 + 15x = -1 \), q.e.d.

3.3.3 Polynomial Inequalities

A very important consequence of the Intermediate Value Theorem is that, for a continuous function to change sign (positive to negative or vice versa) on an interval, its values must pass through zero. We can exploit this to solve polynomial and rational inequalities. The application of the IVT to polynomial inequalities is particularly straightforward, as the next two examples illustrate.

Example 3.3.3 Solve the inequality \( x^3 - 9x^2 + 20x \leq 0 \).
3.3. CONTINUITY ON INTERVALS

Solution: Define \( f(x) = x^3 - 9x^2 + 20x \), which is a polynomial and therefore continuous on all of \( \mathbb{R} \). We wish to see where \( f(x) \leq 0 \). First we will see where \( f(x) = 0 \), to detect the possible points at which \( f(x) \) changes signs.

These points \( x = 0, 4, 5 \) are the only places where \( f(x) \) can change signs as \( x \) passes through values along the \( x \)-axis continuum, so \( f(x) \) will not change signs anywhere within the intervals \( (-\infty, 0) \), \( (0, 4) \), \( (4, 5) \) or \( (5, \infty) \). We can use this fact and a simple test, for example, to find the sign of \( f(x) \) on all of \( (-\infty, 0) \): if we know the sign of \( f(x) \) at, say, \( x = -10 \), then we know it for the whole interval \( x \in (-\infty, 0) \) because whatever is the sign at \( x = -10 \) will be the sign for that whole interval (because if it had two signs in that interval then continuity would guarantee there would be another point where \( f(x) = 0 \) in that interval, and we know that the only points where that happens, \( x = 0, 4, 5 \) are not in that interval). Furthermore, if we would like to find the sign of \( f(x) \) on the interval \( (-\infty, 0) \), it is enough to know the signs of all the factors of \( f(x) = x(x - 4)(x - 5) \), since if an odd number of the factors are negative then \( f \) is negative there, whereas if an even number of factors are negative then they “cancel” to give \( f \) a positive sign.

A nice visual device for determining where \( f(x) \) is positive, and where it is negative, is a sign chart, as illustrated below. For this particular function the signs are shown on their respective intervals, with the points \( x = 0, 4, 5 \) as boundary points. In the style of sign chart is given below, we use \( \oplus \) to represent a positive quantity, and \( \ominus \) to represent a negative quantity. We also look at the signs of the factors to determine the sign of the function:

<table>
<thead>
<tr>
<th>Function: ( f(x) = x(x - 4)(x - 5) )</th>
<th>Test ( x = )</th>
<th>-10</th>
<th>2</th>
<th>4.5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sign Factors:</td>
<td>( \ominus )</td>
<td>( \ominus \oplus )</td>
<td>( \ominus \oplus )</td>
<td>( \ominus \oplus )</td>
<td></td>
</tr>
<tr>
<td>Sign ( f(x) ):</td>
<td>( \ominus )</td>
<td>0</td>
<td>( \ominus )</td>
<td>4</td>
<td>( \ominus )</td>
</tr>
</tbody>
</table>

From the chart we see that \( f(x) < 0 \) on the first and third intervals, i.e., for \( x \in (-\infty, 0) \cup (4, 5) \). Since \( f(x) < 0 \) is equivalent to our original inequality, this is also the solution of that inequality.

The information at hand will not give us a complete picture of the graph of \( f(x) \), but it is instructive to see what the graph looks like, and how an accurate enough (for our purposes here) picture can be easily imagined from the sign chart. For this reason the graph is given in Figure 3.5.

The logic which was used in constructing the sign chart bears repeating. Since \( f \) is continuous, the only way it can change sign is to pass through zero (by IVT, see also Figure 3.5), so we chart all the \( x \)-values for which \( f(x) = 0 \). These mark boundaries of subintervals of \( \mathbb{R} \) on which \( f \) does not change sign. For each such interval, knowing the sign of \( f(x) \) at any value in the interval gives us the sign for the whole interval (since, again, it cannot change sign in the interval without there being another zero in the interval, and all such points are accounted for). If \( f(x) \) happens to be factored, we only need to check the signs of each factor to see if the negative factors “cancel” completely to leave a positive function, or if we have an odd number of negative factors to make \( f \) negative on the interval in question.
Example 3.3.3 was relatively straightforward. There can be complications, and we have to be careful to answer the given question. For instance, we do not always have strict inequalities $<, >$, but may have inclusive inequalities $\leq, \geq$.

**Example 3.3.4** Solve $x^2 \geq x + 1$.

**Solution**: First we subtract, and then define $f(x) = x^2 - x - 1$, so that

$$x^2 \geq x + 1 \iff x^2 - x - 1 \geq 0 \iff f(x) \geq 0.$$  

Now solving $f(x) = 0$ requires the quadratic formula or completing the square. We will opt for the former. Recall first that $f(x) = 0 \iff x^2 - x - 1 = 0$.

$$f(x) = 0 \iff x = \frac{-(1) \pm \sqrt{(-1)^2 - 4(1)(-1)}}{2(1)} = \frac{1}{2} \pm \frac{1}{2} \sqrt{5} \approx -0.61803, 1.61803.$$  

We will always use the exact values, but the approximate ones are also useful since we need to know where to find our test points.\(^\text{17}\)

\(^{17}\)We could factor $f(x)$ based upon the solutions to $f(x) = 0$, namely $\frac{1}{2} \pm \frac{1}{2} \sqrt{5}$:

$$f(x) = \left( x - \frac{1 + \sqrt{5}}{2} \right) \left( x - \frac{1 - \sqrt{5}}{2} \right).$$

Such an approach is perhaps more sophisticated than our method in Example 3.3.4, where we did not bother to factor $f(x)$, but is often unwieldy and requires more subtlety than necessary to solve the inequality.
3.3. CONTINUITY ON INTERVALS

Figure 3.6: Actual graph of \( f(x) = x^2 - x - 1 \). The function is zero at \( \frac{1}{2} \pm \frac{1}{2}\sqrt{5} \approx -0.61803, 1.61803 \). Compare to the sign chart in Example 3.3.4.

Function: \( f(x) = x^2 - x - 1 \)
Test \( x = \) 109 0 10
Test \( f(x) = \) -1 89

Sign \( f(x) = \) ⊕ 1.24 ⊖ ⊖ 2 ⊖ ⊕ 2 ⊕ ⊕ 2

Recall that we are searching for all points for which \( f(x) \geq 0 \). These include the cases where \( f(x) = 0 \) as well as where \( f(x) > 0 \). Therefore we include the endpoints when we report the solution: \( x \in \left(-\infty, \frac{1-\sqrt{5}}{2}\right] \cup \left[\frac{1+\sqrt{5}}{2}, \infty\right) \).

The function \( f(x) = x^2 - x - 1 \) is graphed in Figure 3.6, page 201. Compare the graph with the sign chart above.

In the first two examples, we had the function \( f(x) \) switch signs at every point where \( f(x) = 0 \). This is not always the case. It is possible for the graph to touch the axis, and retreat back to the same side. Consider the following example.

Example 3.3.5 Solve the inequality \( x^3 - 6x^2 + 9x > 0 \).

Solution: This is already a question about sign, so we will simply take \( f(x) = x^3 - 6x^2 + 9x \) and solve \( f(x) > 0 \). Now

\[
f(x) = x^3 - 6x^2 + 9x = x(x^2 - 6x + 9) = x(x - 3)^2.
\]

We see that \( f(x) = 0 \) at \( x = 0, 3 \). This gives us the following sign chart:

Function: \( f(x) = x(x-3)^2 \)
Test \( x = \) -1 2 4
Sign Factors: ⊗ ⊗ 2 ⊗ ⊗ 2 ⊗ ⊗ 2

Sign \( f(x) = \) ⊗ 0 ⊗ 3 ⊗

We see from the sign chart that \( f(x) > 0 \) for \( x \in (0, 3) \cup (3, \infty) \).
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

Figure 3.7: Graph of $f(x) = x(x - 3)^2$ as in Example 3.3.5, illustrating that a function’s height can be zero without the function changing signs there.

We can get an idea how a function might look given a sign chart as in Example 3.3.5, and indeed we see the expected behavior in the graph of the function, given in Figure 3.7.

An interesting phenomenon regarding zeroes of polynomials becomes apparent in constructing these sign charts: if $(x - a)$ appears as a factor in a polynomial $f(x)$ to an odd power, the polynomial changes signs at $a$, while if $(x - a)$ appears to an even power, the polynomial does not change signs at $x = a$. Some terminology helps to express this.

**Definition 3.3.3** If there is a polynomial $g(x)$ and a positive integer $k \in \mathbb{N}$ so that $f(x)$ can be written

$$f(x) = (x - a)^k g(x),$$

but $x - a$ is not a factor of $g(x)$—in other words $(x - a)^k$ is a factor of $f(x)$ but $(x - a)^{k+1}$ is not—then $x = a$ is called a zero, or root, of multiplicity or degree $k$ of $f(x)$.

Then we can observe the following for a polynomial $f(x)$:

1. if $x = a$ is a zero of odd multiplicity, then $f(x)$ changes sign at $x = a$;
2. if $x = a$ is a zero of even multiplicity, then $f(x)$ does not change sign at $x = a$.

In either case $f(a) = 0$ and so $x = a$ is an $x$-intercept of the function.

In Example 3.3.5 we had $x = 3$ was a zero of multiplicity 2, which is even and so $f(x)$ did not change sign (in the sense of changing from positive to negative or vice versa) as $x$ passed through the value 3, while $x = 0$ was a zero of multiplicity 1, which is odd and so the function did change sign there. If one knew the degrees of each zero, and the sign on one interval, one could construct the sign chart from that information alone, depending upon whether the function changes sign while $x$ passes from one interval to the next (left or right, all bordered by the zeroes of $f$), or does not.

It is sometimes the case that, even when $f(x)$ is factored, some of the factors might never be zero. In such a case that factor will never change signs either. One common such type of factor is of the form $x^{2k} + a$, where $k \in \mathbb{N}$ and $a > 0$. Such a factor is always positive regardless of $x \in \mathbb{R}$. Another type is a factor of the form $ax^2 + bx + c$ where $b^2 - 4ac < 0$, in which case
3.3. CONTINUITY ON INTERVALS

\( ax^2 + bx + c = 0 \) has no real solutions. By the Intermediate Value Theorem, if we determine a factor has no real zeros, we know it will not change signs. The following examples illustrates one case.

**Example 3.3.6** Solve the inequality \( x^5 \leq 25x \).

**Solution:** As before, we construct \( f(x) = x^5 - 25x \), so that

\[
 x^5 \leq 25x \iff x^5 - 25x \leq 0 \iff f(x) \leq 0.
\]

Next we factor \( f(x) = x(x^4 - 25) = x(x^2 + 5)(x^2 - 5) \). Now the first factor is zero for \( x = 0 \), the second factor is never zero, and the third is zero for \( x = \pm \sqrt{5} \approx \pm 2.23607 \). Hence we get the following sign chart:

<table>
<thead>
<tr>
<th>Function: ( f(x) = x(x^2 + 5)(x^2 - 5) )</th>
<th>Test ( x = )</th>
<th>-4</th>
<th>-1</th>
<th>1</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sign Factors: ( \Theta \Theta \Theta )</td>
<td>( \Theta \Theta \Theta \Theta \Theta )</td>
<td>( \Theta \Theta \Theta \Theta \Theta )</td>
<td>( \Theta \Theta \Theta \Theta \Theta )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sign ( f(x) ):</td>
<td>( \Theta )</td>
<td>-( \sqrt{5} )</td>
<td>( \Theta )</td>
<td>0</td>
<td>( \Theta )</td>
</tr>
</tbody>
</table>

We see that \( f(x) \leq 0 \) for \( x \in (-\infty, -\sqrt{5}] \cup [0, \sqrt{5}] \).

### 3.3.4 Rational Inequalities

When applying IVT we have to be careful that we actually do have continuity. For instance, if we define \( f(x) = 1/(x - 1) \), we see that \( f(-1) = -1/2 \), while \( f(3) = 1/2 \). However, there are no points between -1 and 3 at which \( f(x) = 0 \). The reason the IVT did not apply is that \( f(x) = 1/(x - 1) \) is not continuous in \([-1, 3]\), since it is discontinuous at \( x = 1 \). If we allow for discontinuities, then a function’s output can “jump” past a particular value, and the image be disconnected (i.e., not an interval). See the first graph in Figure 3.4, page 197 for an illustration of this phenomenon with this particular function. (The second graph also shows how discontinuity allows “jumping.”)

Now we can still use the IVT to solve rational inequalities. We simply need to analyze them and the IVT further. For instance, we can use the following corollary to that theorem:

**Corollary 3.3.3** Suppose that \( f(a) = A \) and \( f(b) = B \), where \( a < b \), and suppose further that \( C \) is between \( A \) and \( B \). Then at least one of the following must hold:

(i) there exists \( c \) between \( a \) and \( b \) so that \( f(c) = C \); or

(ii) \( f(x) \) has at least one discontinuity on \([a, b]\).

In other words, for a function to pass from the height \( A \) to the height \( B \), it must either pass through every height in between, or must be discontinuous, as proved below (see “q.e.d.”).

---

\(^{18}\)For this example we could continue factoring \( f(x) = x(x^2 + 5)(x^2 - 5) = x(x^2 + 5)(x - \sqrt{5})(x + \sqrt{5}) \). However here we will work with the partially factored form, as this will be sufficient. It is really a matter of personal taste.
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

**Proof:** We can use some symbolic logic to prove this. The idea is that

\[ f \text{ continuous on } [a, b] \implies (\text{IVT}) \implies (i) \.

Thus \( f \) continuous on \([a, b]\) \implies (i), meaning that \( f \) continuous on \([a, b]\) \implies (i) is a tautology. Finally, this is equivalent to the statement of the corollary:

\[ \sim (f \text{ continuous on } [a, b]) \lor (i), \quad \text{q.e.d.} \]

This proof is, of course, true and precise, but the corollary should also be intuitive: as we move \( x \) along the interval \([a, b]\), to pass continuously from one height to another we pass through all intermediate heights; if we are not going to pass through all intermediate heights, we have to somehow “jump” over any missed heights, requiring a discontinuity.

One implication for rational, or any other function is immediate. *For a function’s output to change signs as we vary its input, that output must pass through zero or be discontinuous.* Thus, when we produce a sign chart for a function \( f \), we include as boundary points all those points where \( f \) is zero or discontinuous. For rational functions this means we look at where the numerator is zero, and where the denominator is zero, respectively.

**Example 3.3.7** Solve \( \frac{x}{x^2 - 1} \leq 0 \).

**Solution:** Defining \( f(x) = \frac{x}{x^2 - 1} = \frac{x}{(x + 1)(x - 1)} \), we see that \( f \) is zero at \( x = 0 \), and discontinuous at \( x = \pm 1 \). We now use all three of these points to construct the sign chart.

| Function: \( f(x) = \frac{x}{(x + 1)(x - 1)} \) |
| Test \( x = \) | -2 | -0.5 | 0.5 | 2 |
| Sign Factors: | ⊗ | ⊗ | ⊗ | ⊗ |
| Sign \( f(x) \): | ⊗ | -1 | ⊗ | 0 | ⊗ | 1 | ⊗ |

Since we are interested in the points where \( f(x) \leq 0 \), we need the open intervals on which \( f(x) < 0 \), i.e., the first and third intervals, and all points where \( f(x) = 0 \), i.e., \( x = 0 \). Collecting all these we get \( x \in (-\infty, -1) \cup [0, 1) \).

We include a graph of \( f(x) = \frac{x}{(x + 1)(x - 1)} \) in Figure 3.8 to illustrate how \( f \) changes signs by passing through, or leaping over zero. There are other types of discontinuities besides vertical asymptotes, but for rational functions \( f \) in which the numerator and denominator have no common factors, vertical asymptotes are the only type of discontinuity which can occur. Notice from the graph why we include \( x = 0 \) but not \( x = \pm 1 \), when we solve \( f(x) \leq 0 \).

For another perspective justifying the technique for sign charts for rational functions, consider that a ratio of functions can only change signs if the numerator or denominator changes signs. Since both numerator and denominator are polynomials and hence continuous, they can only change signs by passing through zero. Summarizing, we conclude that a ratio of polynomials can only change signs if the numerator passes through zero or the denominator passes through zero. However, for more general functions we have to consider all possible types of discontinuities.

The method above can generalize for solving any rational inequality the same way we generalized for the polynomial case: we rewrite any inequality into an equivalent statement about signs (+/−).
3.3. CONTINUITY ON INTERVALS

Figure 3.8: Graph of \( f(x) = \frac{x}{(x+1)(x-1)} \). See Example 3.3.7. Dashed lines are vertical asymptotes at \( x = \pm 1 \), which are those values of \( x \) outside the domain of \( f(x) \). Vertical asymptotes will be properly developed later in the text.

Example 3.3.8 Solve the inequality \( \frac{x}{x^2 - 7} \leq \frac{2x}{x^2 - 9} \).

Solution: First we do as before—make this into a question about signs—by subtracting the right-hand side from the inequality, and define the difference to be \( f(x) \). Hence we have

\[
f(x) = \frac{x}{x^2 - 7} - \frac{2x}{x^2 - 9} = \frac{x(x^2 - 9) - 2x(x^2 - 7)}{(x^2 - 7)(x^2 - 9)} = \frac{-x^3 + 5x}{(x^2 - 7)(x^2 - 9)} = \frac{-x(x^2 - 5)}{(x^2 - 7)(x^2 - 9)},
\]

and we are trying to find where \( f(x) \leq 0 \). We see that \( f(x) = 0 \) for \( x = 0 \) and \( x = \pm \sqrt{5} \approx \pm 2.23607 \), and is discontinuous at \( x = \pm 3 \) and \( x = \pm \sqrt{7} \approx \pm 2.64575 \). The sign chart follows:

Function: \( f(x) = \frac{(-x)(x^2 - 5)}{(x^2 - 7)(x^2 - 9)} \)

Test \( x = \)

\[
\begin{array}{cccccccc}
-10 & -2.9 & -2.5 & -1 & 1 & 2.5 & 2.9 & 10 \\
\oplus & \oplus & \oplus & \oplus & \oplus & \oplus & \oplus & \oplus \\
\end{array}
\]

Sign Factors:

Sign \( f(x) \):

\[
\begin{array}{cccccccc}
-3 & -\sqrt{7} & -\sqrt{5} & 0 & \sqrt{5} & \sqrt{7} & 3 \\
\oplus & \oplus & \oplus & \oplus & \oplus & \oplus & \oplus \\
\end{array}
\]

It is important to note that the sign chart gives us the signs on the various open intervals. The endpoints of the interval are either where the function is zero, or undefined (the latter implying discontinuity). Since for this case we want our solution to be those points where \( f(x) \leq 0 \), we have to include those endpoints where \( f(x) = 0 \). Putting all this together, we see that

\[
x \in (-3, -\sqrt{7}) \cup [-\sqrt{5}, 0] \cup [\sqrt{5}, \sqrt{7}) \cup (3, \infty).
\]

Note where \( f \) changes signs continuously (i.e., passing through zero height) at \( 0, \pm \sqrt{5} \) and discontinuously (in fact, via vertical asymptotes) at \( x = \pm \sqrt{7}, \pm 3 \). We do not include a graph of
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

$f(x)$ here, but much of its behavior is evident by the sign chart and the way $f$ changes signs by passing through height zero at $x = 0, \pm \sqrt{5}$, and by discontinuity (in fact, via vertical asymptotes) at $x = \pm 3, \pm \sqrt{7}$.

Fortunately, most standard calculus problems require only that we know where certain functions (or more precisely, derivatives of functions) are positive ($> 0$), and where they are negative ($< 0$). For completeness here we also discussed the inclusive inequalities ($\geq, \leq$). We will round out this section with the following final example.

Example 3.3.9 Solve the inequality $\frac{x^2 + x + 1}{x^2 - 7x + 12} > 0$.

**Solution:** As usual, we first check to see where the numerator is zero (where $f(x) = 0$, if $f(x)$ is the function on the left) and where the denominator is zero (where $f(x)$ does not exist). For the numerator (NUM) we need the quadratic formula:

$$\text{NUM} = 0 \iff x = \frac{-1 \pm \sqrt{1^2 - 4(1)(1)}}{2(1)} = \frac{1 - \sqrt{-3}}{2} \notin \mathbb{R}.$$  

We see that the numerator is never zero for $x \in \mathbb{R}$, and so the numerator does not contribute any points to include in making the sign chart. For the denominator (DEN) we have

$$\text{DEN} = 0 \iff x^2 - 7x + 12 = 0 \iff (x - 3)(x - 4) = 0 \iff (x = 3) \lor (x = 4).$$

We use these two points to bound the regions of the sign chart:

| Function: $f(x) = \frac{x^2 + x + 1}{(x - 3)(x - 4)}$ |
| Test $x$ = | 0 | 3.5 | 5 |
| Sign Factors: | $\oplus / (\oplus \ominus)$ | $\oplus / (\oplus \ominus)$ | $\oplus / (\oplus \ominus)$ |
| Sign $f(x)$: | $\oplus$ | $\ominus$ | $\oplus$ |

From the sign chart we see the solution is $x \in (-\infty, 3) \cup (4, \infty)$.

**Exercises**

1. For each of the following, draw a continuous function $f(x)$ whose domain is $x \in (2, 5)$, and whose image of that set (i.e., whose range) is given.

   - (a) $(1, 4)$
   - (b) $\{3\}$
   - (c) $\mathbb{R}$
   - (d) $[1, 4]$  

   - (e) $(1, 4)$
   - (f) $(-\infty, 1)$
   - (g) $(-\infty, 1]$  

2. Repeat the previous problem except assume the domain for $f(x)$ is $x \in [2, 5]$. Some of the cases are impossible (but are interesting to attempt anyhow). See Theorem 3.3.2, page 195.

3. Draw the graph of a function $y = f(x)$, defined for $x \in [2, 5]$, where $f(x)$ is continuous on $(2, 5)$ but not on $[2, 5]$.

4. Show that $x^3 - 6x^2 + 4x + 10 = 0$ has at least three solutions, by checking values
3.3. CONTINUITY ON INTERVALS

of \( f(x) = x^3 - 6x^2 + 4x + 10 \) at various \( x \)-values.

5. Show that \( x^5 - 8x^2 + 15x = 97 \) has at least one solution in \( \mathbb{R} \).

For each of the following, solve the inequality by means of a sign chart. You may have to first rewrite the inequality.

6. \((x + 1)(x - 3) \geq 0\)
7. \(x^2 - 9 < 0\)
8. \(x^2 + 8 \leq 6x\)
9. \(x^2 + 3x \geq 2\)
10. \(x^2 - 15 > 0\)
11. \(x^2 + 15 > 0\)
12. \(x^2 + 18 > 11x\)

13. \(\frac{x}{x + 5} < 0\)
14. \(\frac{x - 7}{x + 6} \geq 0\)
15. \(\frac{x^2 - 16}{x^2 + 16} > 0\)
16. \(\frac{2x + 1}{x^2 - 16} \leq 0\)

17. \(\frac{27x - x^2}{x^2 + 11x + 30} < 0\)
18. \(\frac{(x + 2)^2}{x^2 + 4} < 1\)
19. \(\frac{x}{x + 5} > \frac{1}{x - 7}\)
20. \(x^3 + 2x^2 \leq 15x\)
21. \(\frac{x^2 - 1}{x^4 - 3x^2 - 16} \leq 0\)
22. \(\frac{2x}{x + 5} < \frac{3x}{x + 6}\)

Use a sign chart to graph the following functions, to the extent that continuity and sign of \( f \) are illustrated. You can assume a vertical asymptote will pass through any point in which the denominator is zero (after any cancellation of factors common to both numerator and denominator).

23. \(f(x) = x(x - 1)^2(x^2 - 4)\)
24. \(f(x) = x^3 - x^2 - 2x\)
25. \(f(x) = \frac{x}{x^2 - 9}\)
26. \(f(x) = \frac{x - 1}{x^2 - 9x + 8}\)
The concept of limit is fundamental to calculus. Before its development, mathematicians were able to observe many of the apparent truths of calculus, but were unable to actually prove them. The development of a theory of limits, together with the rigorous definition of $\mathbb{R}$, bridged many important theoretical gaps in calculus between what could be observed and what could be proved.

For our simplest cases, limits are just a small step away from continuity; we can compute many limits quickly based upon our knowledge of continuous functions. However, the concept of limit has been greatly expanded to have meaning in many other contexts. This section is devoted to the simplest case—closest to continuity—which is the case of a finite limit at a point.

As with continuity, we will have many theorems which should be remembered and understood, and which are intuitive on their faces though their proofs are somewhat technical. Because of this we leave the proofs until the end of the section.

---

19In fact the ancient Greek mathematician and physicist Archimedes of Syracuse, (287–212 B.C.) used many arguments which are considered calculus today for his mathematical discoveries. Without the foundations of calculus his arguments were very convincing, but fell short of proofs.

20To be sure, the proofs are always worth reading and understanding and the techniques involved are accessible and relevant to our reading here, but for now it is more important to be able to understand and apply the principles enunciated by the theorems and to understand the limit definition and examples.
3.4. FINITE LIMITS AT POINTS

$\lim_{x \to -3} g(x)$ does not exist
$\lim_{x \to -2} g(x) = 5$
$\lim_{x \to -1} g(x) = 2$
$\lim_{x \to 0} g(x) = 1$
$\lim_{x \to 3} g(x) = 2$

Figure 3.10: A function $g(x)$ to illustrate the concept of limit. (Circles and dotted lines are drawn as visual aids.) See text.

3.4.1 Definition, Theorems and Examples

A typical example of where we might use a limit is the function illustrated in Figure 3.9, on page 208, namely

$$f(x) = \frac{x^2 - 5x + 6}{x - 3}.$$

We see immediately that this function is not defined at $x = 3$. If we look at any point other than $x = 3$—though what is crucial is to look at those points near, but not at, $x = 3$—we can simplify $f(x)$ as follows:

$$f(x) = \frac{x^2 - 5x + 6}{x - 3} = \frac{(x - 3)(x - 2)}{x - 3} = x - 2, \quad x \neq 3.$$

Thus $f(x) = x - 2$ as long as $x \neq 3$. A legitimate question to ask is, as $x$ gets near the forbidden value at 3, does $f(x)$ get near a particular value? As we see from Figure 3.9, the height $f(x)$ does approach the value 1 as $x$ approaches 3. The notation we use to signify this is:

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

read, “the limit, as $x$ approaches 3, of $f(x)$, equals 1.”

Of course, one example does not define a concept, but before we give the definition we will sharpen the idea further by considering the function $g(x)$ graphed in Figure 3.10.

- $\lim_{x \to 0} g(x) = 1$. That should be clear from the graph; as $x$ nears zero, the height indeed approaches the value 1. (We will note later how continuity, as at $x = 0$, has implications for the limit.) Similarly $\lim_{x \to -1} g(x) = 2$. 

Figure 3.10: A function $g(x)$ to illustrate the concept of limit. (Circles and dotted lines are drawn as visual aids.) See text.
• \( \lim_{x \to -2} g(x) = 5 \). Such a limit is oblivious to what actually occurs at \( x = -2 \), but is instead concerned with what occurs as we approach \( x = -2 \), from both sides. In fact \( g(-2) = 1 \), but that is irrelevant. We are only interested in the value approached in the output of \( g(x) \) as its input \( x \) approaches, but does not equal, the value \(-2\).

• \( \lim_{x \to -3} g(x) \) does not exist. As we approach from the left of \( x = -3 \) the function’s value nears 4, but as we approach from the right the function’s value nears 2. For such an ambiguous case we simply decline to assign a value to the limit, instead declaring that it does not exist. Though not actually relevant, we note that \( g(-3) = 4 \).

• \( \lim_{x \to 1} g(x) = 2 \), and (coincidentally) \( \lim_{x \to 3} g(x) = 2 \). These should be clear from the graph.

Now we will give the technical definition of a finite limit at (or “about”) a point.

**Definition 3.4.1** Given a function \( f(x) \) defined on the set \( 0 < |x - a| < d \), for some \( d > 0 \) (\( f(x) \) possibly defined at the point \( x = a \) as well), and some \( L \in \mathbb{R} \), we define \( \lim_{x \to a} f(x) = L \) to mean the following:

\[
\lim_{x \to a} f(x) = L \iff (\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)(0 < |x - a| < \delta \implies |f(x) - L| < \varepsilon). \tag{3.22}
\]

If there is no such \( L \) satisfying (3.22), we say that \( \lim_{x \to a} f(x) \) does not exist.\(^{21}\)

This is sometimes called the epsilon-delta (\( \varepsilon-\delta \)) definition of limit. It differs from the definition of continuity in the implication:

\[
\text{continuity:} \quad |x - a| < \delta \implies |f(x) - f(a)| < \varepsilon; \quad \tag{3.23}
\]

\[
\text{limit:} \quad 0 < |x - a| < \delta \implies |f(x) - L| < \varepsilon. \quad \tag{3.24}
\]

A simplistic interpretation would see that the part of \( f(a) \) is played by \( L \) in the limit, so \( L \) is, in some sense, where \( f(a) \) seems it should be, at least if the function is to be continuous (which it may or may not be) at \( x = a \). Equally crucial is that in the limit definition, the implication in (3.24) is silent on the behavior at \( x = a \), i.e., when \( 0 = |x - a| \). Indeed, neither \( x = a \) nor \( f(a) \) play a role in the limit definition’s implication (3.24), while both are crucial in continuity’s implication (3.23). (See again Figure 3.10, page 209.)

This definition (3.22) of limit is nontrivial and justifies periodic revisiting. With all the cases we will study in this section it will become more and more clear that (3.22) is exactly what we need. However it would be unwieldy indeed to use this definition to prove every limit computation, particularly since such computations are ubiquitous in the rest of this text. Fortunately we know something about continuous functions, and indeed will be able to bootstrap much of our knowledge of continuity to complete nearly all of our limit calculations without working with the \( \varepsilon-\delta \) definition directly. The limit-specific theorems of this section will also assist in developing intuition and computational methods.

Our first theorem is on the uniqueness of the limit, if it exists.

**Theorem 3.4.1** If \( \lim_{x \to a} f(x) \) exists, it is unique. In other words,

\[
\left( \lim_{x \to a} f(x) = L \right) \land \left( \lim_{x \to a} f(x) = M \right) \implies L = M.
\]

\(^{21}\) Note that we will use arrows such as “\( \implies \)” and “\( \iff \)” for both implication and “approaches.” The meanings should be clear from the contexts. The valid implication “\( \implies \)” will keep its earlier meaning throughout.
Theorem 3.4.1 really says the obvious (though the proof is not so transparent): that a function cannot simultaneously be within arbitrarily small $\varepsilon$ tolerance of two values and still conform to the limit definition. Eventually the function has to “choose” between approaching $L$ or approaching $M$ (or other values, or no values, but never more than one), or the limit definition is violated. (Recall the discussion of Figure 3.10, particularly as $x \to -3$.) The proof is given at the end of the section.

The following theorem is very important for computing limits.\footnote{Many calculus texts define limits first, using $\varepsilon$-$\delta$ and then use the second statement of Theorem 3.4.2, $\lim_{x \to a} f(x) = f(a)$ as the definition of continuity of $f(x)$ at $x = a$. This is valid since the definition of limit stands alone without reference to continuity, and (as the proof of the theorem shows) their limit definition of continuity is equivalent to our earlier $\varepsilon$-$\delta$ definition.

Recall that our approach was instead to first define continuity with $\varepsilon$-$\delta$, explore continuity theorems, and then define the limit. With our approach we avoid $\varepsilon$-$\delta$ in limit calculations since those technicalities are built into the theorems (specifically in the proofs). Ours is the approach of many analysis texts, and seems to the authors less convoluted and (hopefully) more intuitive than the usual calculus textbook approach.

Eventually (Section 3.9) we do state the limit theorems which other authors build upon, but only after we exhaust the problems we can do without those methods, and after we build a strong, foundational understanding of limits in a context which is closest to continuity.}

**Theorem 3.4.2** $f(x)$ is continuous at $x = a$ if and only if $\lim_{x \to a} f(x) = f(a)$.

A glance back at Figure 3.10, page 209 (and also Figure 3.9, page 208) makes a good case for the validity of this theorem. The proof is given at the end of the section. A paraphrasing of the theorem could be: if the value which the function approaches is also where the function “ends up,” then we have continuity; if these are different then we do not. With this theorem we can compute many limits by just evaluating the functions at the limit points, if the functions are continuous there. (To help decide continuity we had several theorems in Section 3.2.)

**Example 3.4.1** Consider the following limits, where we can “plug in” the limit points because the functions are continuous at their respective limit points.

- $\lim_{x \to 3} (x^2 + 6x) = 5^2 + 6(5) = 25 + 30 = 55$,
- $\lim_{x \to -4} \sqrt{26 - x^2} = \sqrt{26 - 4^2} = \sqrt{10}$,
- $\lim_{x \to 9} \frac{1}{x-5} = \frac{1}{9-5} = \frac{1}{4}$,
- $\lim_{x \to -3} \frac{x}{x^2+1} = \frac{-3}{(-3)^2+1} = \frac{-3}{10}$.

These we could quickly calculate by simple evaluation because the functions were continuous at the points in question. However we do need to be careful not to draw the wrong conclusion from the above example; evaluating a limit by evaluating the function at the limit point is valid if and only if the function is continuous there. This is illustrated in the next example.

**Example 3.4.2** $\lim_{x \to 3} \sqrt{9 - x^2}$ does not exist because we cannot approach from the right side of 3; if $x > 3$, then $9 - x^2 < 0$ and so $\sqrt{9 - x^2}$ is undefined. This does not contradict Theorem 3.4.2, since the function is not continuous (only left-continuous) at $x = 3$.

Continuity at $x = a$ is a stronger condition than the limit existing at $x = a$: with continuity the limit must exist and equal the function value (which must also exist, i.e., be defined) there. Where limits are truly useful then is with functions which might not be continuous at a given point, but might nonetheless have a limit there. Often the function in question is equivalent to a continuous function near, but perhaps not at, the limit point. We will make repeated use of this fact through the validity of the following theorem:
Theorem 3.4.3 If \( f(x) = g(x) \) on a set \( 0 < |x - a| < d \), where \( d > 0 \), then
\[
\lim_{x \to a} f(x) = \lim_{x \to a} g(x),
\]
or both limits do not exist.

Rephrased, if \( f(x) \) and \( g(x) \) agree near, but not necessarily at, \( x = a \) then their limits (as \( x \) approaches \( a \)) will be the same. This follows quickly from the definition of limit, in which we can replace \( f \) with \( g \), assuming \( \delta \leq d \), which is the sort of thing we did in proving several continuity theorems.

Besides having several applications, this theorem also illustrates much of the nature of limits. For instance it reflects the fact that the limit has a built-in blind spot (by definition) at the point \( x = a \). Also apparent is the local nature of limits: the fact that \( f \) and \( g \) could be wildly different farther away from \( x = a \), i.e., for \( |x - a| \geq d \), is of no importance to the limit.

The most common place where we use Theorem 3.4.3 is when we calculate limits by simplifying the given function to one which is the same near \( x = a \) and has a more obvious limit there. We will consider several examples below. For our first example we revisit the limit which began this section. (See Figure 3.9, page 208.)

Example 3.4.3 Compute the limit \( \lim_{x \to 3} \frac{x^2 - 5x + 6}{x - 3} \).

Solution: Assuming \( f(x) = \frac{x^2 - 5x + 6}{x - 3} \), we see that
\[
f(x) = \frac{(x - 3)(x - 2)}{x - 3} = x - 2, \quad x \neq 3.
\]
(3.25)
In other words, for \( 0 < |x - 3| < \infty \) (thus \( 0 < |x - 3| < d \) for any positive number \( d \) as in the theorem), we have \( f(x) = g(x) \), where \( g(x) = x - 2 \), a function continuous at \( x = 3 \). Now
\[
\lim_{x \to 3} g(x) = \lim_{x \to 3} (x - 2) = 3 - 2 = 1,
\]
and since \( f \) and \( g \) agree except at \( x = 3 \), we conclude that \( \lim_{x \to 3} f(x) = 1 \) as well.

The explanation in the example above is complete and correct, but rather pedantic. Since it is one of the simpler limit problems we will come across, a less verbose explanation—but one still faithful to the spirit of the theorems used—can suffice. In this text we will write a summary version which will read as follows:
\[
\lim_{x \to 3} \frac{x^2 - 5x + 6}{x - 3} 0/0 \quad \text{ALG} \quad \lim_{x \to 3} \frac{(x - 3)(x - 2)}{x - 3} 0/0 \quad \text{ALG} \quad \lim_{x \to 3} (x - 2) = 3 - 2 = 1.
\]
The “0/0” (usually read “zero over zero”) notation over the equality symbols “=” signify that the limit is of 0/0 form, which we will discuss in the next paragraph. The “ALG” underneath the second equality symbol signifies that an algebraic rewriting was carried out which was legitimate near, but not necessarily at, the limit point (see again (3.25)). That particular step is where we used Theorem 3.4.3, with the original function playing the part of \( f \) and the function \( (x - 2) \) playing the part of \( g \). (The “ALG” under the first equality symbol just signifies we algebraically rewrote the function, so we will use “ALG” in both contexts.) We will have other comments to write in the convenient spaces above and below the equality symbols for bookkeeping purposes, so we can concisely organize and later check our work.\(^{23}\)

\(^{23}\) The comment system we use here was developed by the authors (who doubt that it is unique). It has been very useful to calculus students wishing to follow the instructor’s thinking and to clarify their own. One would usually omit such comments in professional publications, where readers are expected to have sufficient knowledge and experience to fully understand each step without explanation. Their knowledge and experience, of course, come from having practice in solving problems themselves as they learned and later applied these principles.
The previous example is a limit of a certain form, which is called the “0/0 form,” meaning that the function is a fraction in which the numerator and denominator both approach zero as \( x \) approaches the limit point. This is one of many indeterminate forms; knowing we have 0/0 form tells us nothing about the value of the limit itself, or even if it exists. The reason is that a shrinking numerator by itself tends to shrink a fraction, while a shrinking denominator alone makes a fraction grow in absolute size. Knowing these are happening simultaneously does not tell us the relative rates at which the two influences are operating. In other words, which competing influence (numerator shrinking or denominator shrinking) dominates? Or is there a compromise, as in the previous example?

There will be many other indeterminate forms, and even more determinate forms later in the text. Most of the interesting limits we will find here are of the indeterminate forms. Fortunately we can often simplify an indeterminate form to one which is not. We did so in the previous example, finding an equivalent polynomial limit.

For now it is important to note that when we have 0/0 form, we are not ready to “plug in the limit point,” but have more work to do, often algebraic, ideally finding a continuous function which is equal to the original function within the limit except possibly at the limit point. With the new, continuous function we can just “plug in” the limit point and be done. This is our strategy in the following examples. In particular note how we work towards cancelling terms in the denominator which cause the denominator to approach zero as \( x \) approaches the limit point.

Example 3.4.4 Compute \( \lim_{x \to 9} \frac{\sqrt{x} - 3}{x - 9} \).

**Solution:** Note that \( x = 9 \) is outside of the domain of the function, but the actual domain is \( x \in [0, 9) \cup (9, \infty) \) so we can certainly approach \( x = 9 \) from both directions. More casually, we can say that we can let \( x \) venture small distances to the left or right of \( x = 9 \) and the function will be defined. The usual technique for a problem such as this is to algebraically rewrite it by multiplying by \( (\sqrt{x} + 3)/(\sqrt{x} + 3) \):

\[
\lim_{x \to 9} \frac{\sqrt{x} - 3}{x - 9} = \lim_{x \to 9} \frac{\sqrt{x} - 3}{x - 9} \cdot \frac{\sqrt{x} + 3}{\sqrt{x} + 3} = \lim_{x \to 9} \frac{x - 9}{(\sqrt{x} + 3)(x - 9)} = \lim_{x \to 9} \frac{1}{\sqrt{x} + 3} = \frac{1}{6}.
\]

As before, we took a 0/0 form and algebraically manipulated it until we found a function equal to the original near, but not at, \( x = 9 \), and could then evaluate the new function at that point since it was continuous there. A quick alternative method for this limit uses some slightly clever factoring:

\[
\lim_{x \to 9} \frac{\sqrt{x} - 3}{x - 9} = \lim_{x \to 9} \frac{\sqrt{x} - 3}{(\sqrt{x} - 3)(\sqrt{x} + 3)} = \lim_{x \to 9} \frac{1}{\sqrt{x} + 3} = \frac{1}{6}.
\]

Example 3.4.5 Compute \( \lim_{x \to 4} \frac{1}{x - 4} \).

**Solution:** This is of form 0/0 as well. We need to simplify the fraction and see how things might cancel. To be rid of “fractions in the numerator” we will multiply by \( \frac{4x}{4x} \):

\[
\lim_{x \to 4} \frac{1}{x - 4} = \lim_{x \to 4} \frac{4x}{4x(x - 4)} = \lim_{x \to 4} \frac{1}{4x} = \frac{1}{16}.
\]
In the above example, the domain of the original function was \( x \neq 0, 4 \), so we could approach \( x = 4 \) locally inside the domain. Such technicalities are important, but one usually does not make mention of them while working a problem unless a quick inspection shows they need to be considered. An alternative algebraic method of simplifying the expression in this limit is to combine the fractions in the numerator:

\[
\frac{1}{x} - \frac{1}{4} = \frac{4 - x}{4x},
\]

so the function simplifies

\[
\frac{\frac{1}{x} - \frac{1}{4}}{x - 4} = \frac{4 - x}{4x} \cdot \frac{1}{x - 4} = \frac{4 - x}{4x(x - 4)}.
\]

The first method gives the same simplification a step (or two) sooner. Next consider the following rational limit.

**Example 3.4.6** Compute \( \lim_{x \to -3} \frac{x^2 - 9}{x^6 - 243} \).

**Solution:** This time we will factor both numerator and denominator, using 243 as \( 3^5 \) along the way. Recall \((a^n - b^n) = (a - b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1})\).

\[
\begin{align*}
\lim_{x \to -3} \frac{x^2 - 9}{x^6 - 243} &= \lim_{x \to -3} \frac{(x + 3)(x - 3)}{x^6 - 243} \\
&= \lim_{x \to -3} \frac{(x + 3)(x - 3)}{x(x - 3)(x^4 + 3x^3 + 9x^2 + 27x + 81)} \\
&= \lim_{x \to -3} \frac{6}{3(81 + 81 + 81 + 81 + 81)} \\
&= \frac{6}{3 \cdot 512} = \frac{2}{405}.
\end{align*}
\]

In the above example we used the factoring \( a^5 - b^5 = (a - b)(a^4 + a^3b + a^2b^2 + ab^3 + b^4) \). Recall earlier we used a version of \( a^2 - b^2 = (a-b)(a+b) \) in the form of \( x - 9 = (\sqrt{x} - 3)(\sqrt{x} + 3) \).

Similarly, we can use \( a^3 - b^3 = (a-b)(a^2+ab+b^2) \), or the related form \( a^3 + b^3 = (a+b)(a^2-ab+b^2) \). These and related techniques often work with limits containing radicals.

**Example 3.4.7** Compute \( \lim_{x \to -8} \frac{\sqrt[3]{x} + 2}{x + 8} \).

**Solution:** This can be accomplished by completing the factorization \( a^3 + b^3 = (a + b)(a^2 - ab + b^2) \) in the numerator, with \( a = \sqrt[3]{x} \) and \( b = 2 \):

\[
\begin{align*}
\lim_{x \to -8} \frac{\sqrt[3]{x} + 2}{x + 8} &= \lim_{x \to -8} \frac{x^{1/3} + 2}{x + 8} \\
&= \lim_{x \to -8} \frac{x^{1/3} + 2}{x^{2/3} - 2x^{1/3} + 4} \\
&= \lim_{x \to -8} \frac{x + 8}{x^{2/3} - 2x^{1/3} + 4} \\
&= \lim_{x \to -8} \frac{1}{(-8)^{2/3} - 2(-8)^{1/3} + 4} = \frac{1}{4 - 2(-2) + 4} = \frac{1}{12}.
\end{align*}
\]

Note that \( 1/(x^{2/3} - 2x^{1/3} + 4) \) is continuous at \( x = -8 \), since this is just \( 1/((\sqrt[3]{x})^2 - 2\sqrt[3]{x} + 4) \), the denominator’s terms being continuous and the denominator not approaching zero. An algebraic alternative is to factor the denominator using \( a^3 + b^3 = (a + b)(a^2 - ab + b^2) \), with \( a = \sqrt[3]{x} \) and \( b = 2 \):

\[
\begin{align*}
\lim_{x \to -8} \frac{\sqrt[3]{x} + 2}{x + 8} &= \lim_{x \to -8} \frac{\sqrt[3]{x} + 2}{(\sqrt[3]{x})^2 - 2x^{1/3} + 4} \\
&= \lim_{x \to -8} \frac{1}{(\sqrt[3]{x})^2 - 2x^{1/3} + 4} = \frac{1}{12}.
\end{align*}
\]

as before.
We do have to be careful that the limit exists, and there are many instances in which it will not. We will discuss a few below, and add more instances in the next section.

**Example 3.4.8** Consider the limit \( \lim_{x \to 5} \sqrt{x^2 - 25} \). This limit does not exist, because the function is not defined for \( x \in (-5, 5) \), and for the limit to exist we need to be able to approach \( x = 5 \) from the immediate left as well as the right. Indeed, the domain is \( x \in (-\infty, -5] \cup [5, \infty) \), so there is a gap to the left of \( x = 5 \) in the domain.

The previous example shows again that we cannot compute a limit by inputting the limit point just because the function is defined there. Recall we can do so if and only if the function is continuous at that point. The function \( f(x) = \sqrt{x^2 - 25} \) is only right continuous at \( x = 5 \).

**Example 3.4.9** Consider the limit \( \lim_{x \to 0} \frac{x}{\sqrt{x^2}} \). This limit is of \( 0/0 \) form, but ultimately does not exist. For the sake of argument, if it did we would have

\[
\lim_{x \to 0} \frac{x}{\sqrt{x^2}} = \lim_{x \to 0} \frac{x}{|x|}.
\]

Although \( f(x) \) is undefined at \( x = 0 \), we can use the piecewise definition of the absolute value function to rewrite the function for \( x \neq 0 \):

\[
f(x) = \frac{x}{|x|} = \begin{cases} 
  x/(x) & \text{for } x > 0 \\
  x/(-x) & \text{for } x < 0
\end{cases} = \begin{cases} 
  1 & \text{for } x > 0 \\
  -1 & \text{for } x < 0.
\end{cases}
\]

Now that function has height \(-1\) for \( x < 0 \) and height \( 1 \) for \( x > 0 \), so we get different heights when we approach from different sides. Thus the limit does not exist. This function is graphed in Figure 3.11.

**Example 3.4.10** Next consider \( \lim_{x \to -2} \frac{\sqrt{x^2}}{x} = \lim_{x \to -2} \frac{|x|}{x} \).

Here we look at two methods of computing this. First we note that the function is continuous at \( x = -2 \), so we can simply evaluate the function there:

\[
\lim_{x \to -2} \frac{\sqrt{x^2}}{x} = \lim_{x \to -2} \frac{|x|}{x} = \frac{|-2|}{-2} = \frac{2}{-2} = -1.
\]
Another method is to replace the function by another which is equal to the original near $x = -2$:
\[
\lim_{x \to -2} \frac{\sqrt{x^2}}{x} = \lim_{x \to -2} \frac{|x|}{x} = \lim_{x \to -2} \frac{-x}{x} = \lim_{x \to -1} (-1) = -1.
\]

We include the comment below the brace above for emphasis, and would normally not include it within the actual computation. The point made there is that in the limit computation above, we are not claiming that $|x|/x = -x/x$ for all $x$, but merely that we can make that substitution here because it is the case “near $x = -2$,” that is, we can replace the original function by $g(x) = (-x)/x$ because $|x| = -x$ for $x \in (-\infty, 0)$ and $-2$ is “safely” inside of $(-\infty, 0)$. Furthermore, that function could be then be replaced by the \textbf{constant} function $h(x) = -1$—which is trivially continuous—near $x = -2$.\textsuperscript{24}

\textbf{Example 3.4.11} Consider the function $f(x) = \begin{cases} 2x + 3 & \text{if } x > 4, \\ x + 4 & \text{if } -3 < x \leq 4, \\ x^2 - 8 & \text{if } x < 3. \end{cases}$

- $\lim_{x \to 5} f(x) = \lim_{x \to 5} (2x + 3) = 2(5) + 3 = 13$. This is because $5 \in (4, \infty)$, with room to the left and right, so we would claim that $x \to 5 \implies f(x) = 2x + 3 \to 13$. The important point is that $5$ is well within the interval on which $f(x)$ is defined to be the continuous function $2x + 3$.

- $\lim_{x \to -2.9} f(x) = \lim_{x \to -2.9} (x + 4) = -2.9 + 4 = 1.1$. Here $-2.9 \in (-3, 4)$, and $-2.9$ is safely within the interval on which $f(x) = x + 4$, a continuous function there. Though $-2.9$ is “only” $0.1$ units from where the function’s formula changes, that is a positive distance and we can always assume $\delta \leq 0.1$ in our limit and continuity proofs to show $f(x)$ is continuous at $x = -2.9$, and thus we can “plug in” $x = -2.9$ to our limit computation.

- $\lim_{x \to 4} f(x)$ does not exist. From the left-hand side of $x = 4$, we see $f(x) = x + 4 \to 8$, but from the right-hand side of $x = 4$ we have $f(x) = 2x + 3 \to 11$.

- $\lim_{x \to -3} f(x)$ requires us to again look at both left-hand side and right-hand side of $x = -3$. From the left we have $f(x) = x^2 - 8 \to 9 - 8 = 1$, and from the right we have $f(x) = x + 4 \to -3 + 4 = 1$. Thus $\lim_{x \to -3} f(x) = 1$ exists. (Here “$\to$” is read “approaches,” and is not to be confused with the implication operator from logic.)

The arguments made regarding the limits as $x \to 4$ and $x \to -3$ will be more systematic in the next section, and we are only showing a glimpse of them here. In the next section we will deal more with piecewise-defined functions, and how better to deal with limit points that are on the boundaries of the “pieces.” It is noteworthy that we did not require a graph to analyze any of the limits in the above example, though we may have had aspects of the graph in mind as part of our thinking.

\textsuperscript{24}The key fact that \textbf{constant} functions are continuous is often overlooked at first by students as they compute limits. Many calculus textbooks emphasize this fact in the limit context by enshrining it in a theorem, the gist of which is
\[
\lim_{x \to a} K = K.
\]

This is obvious when its meaning is understood: that if we define $f(x) = K$, where $K \in \mathbb{R}$ is a constant, then $\lim_{x \to a} f(x) = K$ as well. A quick glance at such a function—whose graph is a horizontal line at height $K$—shows that such a function is obviously continuous, so we can evaluate the limit by evaluating the (constant) function.
3.4. **FINITE LIMITS AT POINTS**

We list one more example here which avoids that issue, but shows the second method of Example 3.4.10 in action again.

**Example 3.4.12** Consider the function \( f(x) = \frac{x^2 - 9}{|x + 3|(x - 3)} \). Then

\[
\lim_{x \to 3} f(x) = \lim_{x \to 3} \frac{(x + 3)(x - 3)}{(x + 3)(x - 3)} \quad \text{ALG} \quad \lim_{x \to 3} 1 = 1;
\]

\[
\lim_{x \to -5} f(x) = \lim_{x \to -5} \frac{(x + 3)(x - 3)}{(x + 3)(x - 3)} \quad \text{ALG} \quad \lim_{x \to -5} (-1) = -1.
\]

We used the fact that \( |x + 3| = x + 3 \) for \( x \) near 3, while \( |x + 3| = -(x + 3) \) for \( x \) near \(-5\). A quick check shows this limit does not exist for \( x \to -3 \) (in a way similar to Example 3.4.9, page 215).

**Example 3.4.13** Consider the limit \( \lim_{x \to 0} \frac{1}{x} \). This limit does not exist as a finite number. The graph is given at the start of Section 3.6. It is easy to see that small inputs into this function quickly return large outputs. For instance, \( f(10^{-1}) = 10^1 \), \( f(10^{-2}) = 10^2 \), and so on, while \( f(-10^{-1}) = -10^1 \), \( f(-10^{-2}) = -10^2 \), and so on. This is sometimes verbally described as \( f(x) \) “blowing up” as \( x \) gets nearer to zero: \( 1/x \) is unbounded and negative as \( x \) approaches zero from the left, and unbounded and positive as \( x \) approaches zero from the right. The geometric result is a vertical asymptote at \( x = 0 \). When we have vertical asymptotes at our limit points, we do not have finite limits.\(^{25}\)

**3.4.2 Further Limit Notation**

We next take the opportunity to introduce some convenient notation, already alluded to previously. Unfortunately it resembles some notation from logic, but it is usually obvious which meaning is intended from the context. The notation is defined as follows:

\[
\lim_{x \to a} f(x) = L \iff [(x \to a) \implies (f(x) \to L)]. 
\]  

(3.27)

For reasons of style, it is often less awkward to write, “\( x^2 \to 9 \) as \( x \to 3 \),” than to write \( \lim_{x \to 3} x^2 = 9 \). We might also write, “as \( x \to 3 \), \( x^2 \to 9 \),” or “\( x \to 3 \implies x^2 \to 9 \).” (Arrow lengths are often chosen for convenience or aesthetics, both here and in logic notation.) It is important, but usually obvious, where the expressions would be placed in our original limit notation. The usefulness of this notation will become more apparent in the next section. One nice use we can put it to here is with the following, perhaps more elegant restatement of Theorem 3.4.2:

\[
f(x) \text{ continuous at } x = a \iff (x \to a) \implies (f(x) \to f(a)).
\]  

(3.28)

We will have much more use for this kind of notation as we look at other limit forms, and other contexts where we use limits. It is only mentioned here so that the reader will be aware of it well before it becomes ubiquitous.

\(^{25}\)In subsequent sections we will define infinite limits, and left- and right-side limits. For this section we only concern ourselves with finite limits at (or “about,” i.e., from both the left and the right) a point. Nothing we do here contradicts subsequent sections.
3.4.3 Proofs of Limit Theorems

Now we prove Theorem 3.4.2, which states that \( f(x) \) is continuous at \( x = a \) if and only if \( \lim_{x \to a} f(x) = f(a) \).

**Proof:** We prove this in two parts. First we show the “only if” part (\( \implies \)):

Suppose that \( f(x) \) is continuous at \( x = a \). Then by definition,

\[
(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x) (|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon).
\]

Now for a pair \( \varepsilon, \delta \) from continuity, we get

\[
0 < |x - a| < \delta \implies |x - a| < \delta \implies |f(x) - f(a)| < \varepsilon,
\]

and so the definition of limit works here with the \( \varepsilon, \delta \) from (assumed) continuity, and \( f(a) \) for \( L \).

For the “if” part (\( \iff \)), suppose \( \lim_{x \to a} f(x) = f(a) \). Thus

\[
(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x) (0 < |x - a| < \delta \implies |f(x) - f(a)| < \varepsilon).
\]

We only need to show that

\[
0 = |x - a| \implies |f(x) - f(a)| < \varepsilon.
\]

But that is obvious from the definition of a function, since

\[
|x - a| = 0 \iff (x = a) \implies f(x) = f(a) \implies |f(x) - f(a)| = 0 < \varepsilon.
\]

Thus the case \( 0 < |x - a| < \delta \) is taken care of by the limit assumption, and the \( 0 = |x - a| \) by the definition of function, giving us the full case \( |x - a| < \delta \), as required by the continuity definition, q.e.d.

Now we prove Theorem 3.4.1, that if \( \left( \lim_{x \to a} f(x) = L \right) \wedge \left( \lim_{x \to a} f(x) = M \right) \implies L = M \).

**Proof:** We will prove this by contradiction. Suppose that the theorem is false, i.e., that there are two numbers, \( L, M \in \mathbb{R} \), different, which satisfy the definition of the limit at \( x = a \) (if necessary see (1.21)). In other words,

\[
(\forall \varepsilon > 0)(\exists \delta_1 > 0)(\forall x) (0 < |x - a| < \delta_1 \implies |f(x) - L| < \varepsilon), \tag{3.29}
\]

\[
(\forall \varepsilon > 0)(\exists \delta_2 > 0)(\forall x) (0 < |x - a| < \delta_2 \implies |f(x) - M| < \varepsilon). \tag{3.30}
\]

Since we are assuming \( L \neq M \), we must have \( |L - M| > 0 \). Choose \( \varepsilon = \frac{1}{3}|L - M| \).

Now pick \( \delta_1 > 0 \) which satisfies the implication in (3.29) for this particular \( \varepsilon \), and \( \delta_2 > 0 \) which satisfies the implication in (3.30) for this particular \( \varepsilon \). Now define

\[
\delta = \min\{\delta_1, \delta_2\}. \tag{3.31}
\]

Now choose any \( x \) satisfying \( 0 < |x - a| < \delta \). This also gives \( 0 < |x - a| < \delta_1 \) and \( 0 < |x - a| < \delta_2 \) by (3.31). This ultimately implies

\[
|L - M| = |L - f(x) + f(x) - M| \leq |L - f(x)| + |f(x) - M| = |f(x) - L| + |f(x) - M| < \varepsilon + \varepsilon = 2\varepsilon = \frac{2}{3}|L - M| < |L - M| \implies \mathcal{F}.
\]
The reason that this is a contradiction is the we ultimately showed that if the theorem is false, then $L \neq M \implies |L - M| < |L - M|$, which is impossible. Hence the assumption that the theorem is false is itself false, proving that the theorem must be true.\footnote{We need to be careful about the last step. It is because $|L - M| > 0$ that we can say $\frac{3}{2}|L - M| < |L - M|$. Clearly this is false if $L = M$. If we do not know that $L \neq M$ (and thus $|L - M| > 0$), then we can only say $\frac{3}{2}|L - M| \leq |L - M|$, as it is always true that $0 \leq \frac{3}{2}|L - M| \leq |L - M|$ regardless of whether $L = M$ or $L \neq M$.} This completes the proof of Theorem 3.4.1, q.e.d.

From the symbolic logic point of view, it is interesting to analyze the above proof, which an advanced student would recognize immediately as a classic “proof by contradiction,” which is a minor variation of \textit{modus tollens}. We could also outline it as

$$(\sim P) \rightarrow \mathcal{F} \implies (\sim P) \iff P,$$

where

$$P : \text{Theorem 3.4.1}.$$

Of course we have to refer to our discussion of quantifiers in order to analyze what exactly would be the statement of $\sim P$, and that sets up our proof that $\sim P \implies \mathcal{F}$.

Note also that $\sim P \rightarrow \mathcal{F} \iff (\sim (\sim P)) \lor \mathcal{F} \iff P \lor \mathcal{F} \iff P$, so we actually have logical equivalence instead of $\implies$ in the above symbolic logic computation. We showed $\sim P \rightarrow \mathcal{F}$ is a tautology (by proving the truth of it), hence proving $P$. 
Consider the graph given in Figure 3.12 above.

1. For each limit, see if it exists. If not, state so. If so, evaluate it.
   
   (a) \( \lim_{x \to -2} f(x) \).
   (b) \( \lim_{x \to 0} f(x) \).
   (c) \( \lim_{x \to 3} f(x) \).
   (d) \( \lim_{x \to 6} f(x) \).

2. Find \( f(-2), f(0), f(3) \) and \( f(6) \).

3. Decide if \( f(x) \) is continuous at the given point. Explain why or why not, in light of Theorem 3.4.2.

(a) \( x = -2 \).
(b) \( x = 0 \).
(c) \( x = 3 \).
(d) \( x = 6 \).

Compute the following limits, if they exist, using Theorem 3.4.2 (page 211). If the limit does not exist, explain why.

4. \( \lim_{x \to 0} \frac{x + 2}{x - 5} \).
5. \( \lim_{x \to 5} \sqrt{x^2 - 16} \).
6. \( \lim_{x \to 4} \sqrt{x^2 - 16} \).
7. \( \lim_{x \to 3} \sqrt{x^2 - 16} \).
3.4. **FINITE LIMITS AT POINTS**

8. \( \lim_{x \to 81} \sqrt[4]{x} \).

9. \( \lim_{x \to -1} (x^{100} - x^2 + 3) \).

10. \( \lim_{x \to 0} \sqrt{x^2 + 1} \).

11. \( \lim_{x \to 0} \frac{3}{\sqrt{x}} \).

12. \( \lim_{x \to 7} ((x^2 - 9)(x + 4)) \).

13. Here we consider why the 0/0-form is indeterminate, i.e., why knowing the numerator and denominator both approach zero does not tell us immediately what the limit is. Compute each of the following limits.

   (a) \( \lim_{x \to 0} \frac{x}{2x} \).

   (b) \( \lim_{x \to 0} \frac{x}{3x} \).

   (c) \( \lim_{x \to 0} \frac{x^2}{x} \).

   (d) \( \lim_{x \to 0} \frac{x}{x^2} \).

   (e) \( \lim_{x \to 0} \frac{x}{|x|} \).

   (f) \( \lim_{x \to 0} \frac{x^2}{|x|} \).

   (g) Explain why the above computations show that the 0/0-form is indeed indeterminate.

Compute the given limits where possible. Be sure to indicate any 0/0 forms and any algebraic steps (preferably using the spaces above and below the “=” as shown in previous examples). If a particular limit does not exist, state so and explain why.

14. \( \lim_{x \to 3} \frac{x^2 + x - 12}{2x^2 - 5x - 3} \).

15. \( \lim_{x \to -25} \frac{x - 25}{\sqrt{x} - 5} \).

16. \( \lim_{x \to \sqrt[3]{2}} \frac{x^4 - 9}{x^2 - 3} \).

17. \( \lim_{x \to 0} \frac{\sqrt[4]{x} - 1}{x} \).

18. \( \lim_{x \to 0} \frac{\sqrt{x + 1} - 1}{x} \).

19. \( \lim_{x \to 1} \frac{x^8 - 1}{x^3 - 1} \).

20. \( \lim_{x \to 4} \frac{\sqrt{x} - 2}{x - 4} \).

21. \( \lim_{x \to 16} \frac{\sqrt{x} - 2}{x - 16} \).

22. \( \lim_{x \to 27} \frac{x^2 - 9}{x^3 + 27} \).

23. Let \( f(x) = \frac{(x - 2)|2x - 5|}{(x - 1)(2x - 5)} \). Compute the following where possible.

   (a) \( \lim_{x \to 2} f(x) \).

   (b) \( \lim_{x \to 0} f(x) \).

   (c) \( \lim_{x \to 3} f(x) \).

   (d) \( \lim_{x \to 5/2} f(x) \).

24. For each of the following, compute the given limit if it exists. Otherwise state and explain why it does not.

   (a) \( \lim_{x \to 3} \sqrt{x^2 - 4} \).

   (b) \( \lim_{x \to 2} \sqrt{x^2 - 4} \).

25. Compute the limit

   \[ \lim_{x \to 0} \frac{\sqrt{1 + x} + \sqrt{1 - x} - 2}{x^2} \]

   It can help to rewrite the limit

   \[ \lim_{x \to 0} \frac{1}{x} \left[ \frac{\sqrt{1 + x} - 1}{x} + \frac{\sqrt{1 - x} - 1}{x} \right] \]

26. Retrace the proof of Theorem 3.4.1 and show that the proof still works if we instead take \( \varepsilon = \frac{1}{2} |L - M| \) to arrive at a contradiction. (You should be able to then observe why the factor 1/2 is the largest factor which works in the proof of the theorem.)
3.5 One-Sided Finite Limits

Just as we found a use for one-sided continuity, so we also have use for so-called left-side (or left-hand side) and right-side (or right-hand side) limits. For instance, consider the function $f(x) = (3x^2 + x)/|x|$. Now $|x|$ can be defined piecewise to be $-x$ for $x \leq 0$, and $x$ for $x > 0$, for instance.\(^{27}\) Although $f(x)$ is undefined at $x = 0$, for $x \neq 0$ we can write

$$f(x) = \frac{3x^2 + x}{|x|} = \begin{cases} \frac{(3x^2 + x)(x)}{x} & \text{for } x > 0 \\ \frac{(3x^2 + x)(-x)}{-x} & \text{for } x < 0 \end{cases} = \begin{cases} 3x + 1 & \text{for } x > 0 \\ -3x - 1 & \text{for } x < 0. \end{cases}$$

The function is graphed in Figure 3.13. Of course this function is undefined at $x = 0$, but we might be interested in what occurs when we approach zero from one side or the other. We see that approaching zero from the left side (thus moving right towards zero) we travel along a line whose height is approaching $-1$, while from the right of zero (moving left) we travel a different line whose height approaches $1$. The notation we use to reflect this is the following:

$$\lim_{x \to 0^-} f(x) = -1, \quad \lim_{x \to 0^+} f(x) = 1.$$ 

We read the notation $x \to 0^-$ as “$x$ approaches zero from the left,” and $x \to 0^+$ as “$x$ approaches zero from the right.” To actually work such a problem, in particular without resorting to a graph, we could write the following, as the two cases reflect where $|x| = -x$ and where $|x| = x$:

$$\lim_{x \to 0^-} f(x) = \lim_{x \to 0^-} \frac{3x^2 + x}{x} = \lim_{x \to 0^-} \frac{3x^2 + x}{-x} = \lim_{x \to 0^-} \frac{0}{0} = \text{ALG, limit } (-3x - 1) = -1,$$

$$\lim_{x \to 0^+} f(x) = \lim_{x \to 0^+} \frac{3x^2 + x}{x} = \lim_{x \to 0^+} \frac{3x^2 + x}{x} = \lim_{x \to 0^+} \frac{0}{0} = \text{ALG, limit } (3x + 1) = 1.$$

Notice that $\lim_{x \to 0^-} f(x)$ does not exist, since approaching the limit point (zero) from one side gives us one value ($-1$), while approaching from the other side gives another ($1$). We next make several clarifications regarding one-sided limits.

\(^{27}\)Of course we can let $|x|$ be $x$ or $-x$ for the case $x = 0$, but that will not come to play in our discussion here.
1. When we analyze a left-side limit, we disregard behavior at and to the right of the limit point; when we analyze a right-side limit, we disregard behavior at and to the left of the limit point.

2. We have analogs (see the theorem below) of Theorem 3.2.11 (page 188), and Theorems 3.4.2 and 3.4.3 (page 212) except that here it is enough to have one sided continuity and equality of the replacement function. We also have the very important theorem in (a) below.

**Theorem 3.5.1** The following hold (where we assume that $d > 0$ where it appears):

(a) $\lim_{x \to a} f(x) = L \iff \left( \lim_{x \to a^-} f(x) = L \right) \land \left( \lim_{x \to a^+} f(x) = L \right)$.

(b) $\lim_{x \to a^-} f(x) = f(a) \iff f(x)$ is left-continuous at $x = a$.

(c) $\lim_{x \to a^+} f(x) = f(a) \iff f(x)$ is right-continuous at $x = a$.

(d) $f(x) = g(x)$ on $x \in (a - d, a) \implies \lim_{x \to a^-} f(x) = \lim_{x \to a^-} g(x)$ (or both do not exist).

(e) $f(x) = g(x)$ on $x \in (a, a + d) \implies \lim_{x \to a^+} f(x) = \lim_{x \to a^+} g(x)$ (or both do not exist).

These should become clear as we proceed, but are rather technical to prove. For now we will pursue these ideas in several examples.

**Example 3.5.1** If possible find $\lim_{x \to 5^-} \sqrt{x^2 - 25}$ and $\lim_{x \to 5^+} \sqrt{x^2 - 25}$.

**Solution:** Note that the domain of $\sqrt{x^2 - 25}$ is $(-\infty, -5] \cup [5, \infty)$. Thus we can not approach $x = 5$ from the (immediate) left, so $\lim_{x \to 5^-} \sqrt{x^2 - 25}$ does not exist.

On the other hand, the function is right-continuous at $x = 5$, so we can follow the height at $x$-values on the right of $x = 5$, and move left in $x$, and the height will change continuously to the height at $x = 5$ as we move down to that $x$-value.

$$\lim_{x \to 5^+} \sqrt{x^2 - 25} = \sqrt{5^2 - 25} = 0.$$

What gave away the fact that we could just “plug in” $x = 5$ for the right-side limit was that the values inside the square root were nonnegative for $x$ in that range. That was not the case as $x \to 5^-$; we have “wiggle room” to the right, but not to the left, of $x = 5$. For a more concise presentation, we could write

$$\lim_{x \to 5^-} \sqrt{\frac{x^2 - 25}{x^2 - 25}} \text{ does not exist},$$

$$\lim_{x \to 5^+} \sqrt{\frac{x^2 - 25}{x^2 - 25}} = \sqrt{\frac{5^2 - 25}{5^2 - 25}} = 0.$$

Note that the example $f(x) = (3x^2 + x)/|x|$ at the beginning of this section did not allow us to “plug in” $x = 0$ for either limit, because the function is neither left-continuous nor right-continuous at $x = 0$. But in both cases we replaced the given function with functions which were: $g(x) = -3x - 1$ for the left-side limit, and $g(x) = 3x + 1$ for the right-side limit.

---

28 Actually most of these can be proven quickly by modifications of the proofs of earlier theorems, but those proofs each took enough space that to modify them here would require a rather distracting effort. We leave them as exercises (but not listed in the regular exercises) for the interested reader.
Example 3.5.2 Compute if possible \( \lim_{x \to 3} \frac{(x + 2)\sqrt{x^2 - 6x + 9}}{x^2 - 7x + 12} \).

Solution: First we will do some algebraic simplification:

\[
\lim_{x \to 3} \frac{(x + 2)\sqrt{x^2 - 6x + 9}}{x^2 - 7x + 12} = \lim_{x \to 3} \frac{(x + 2)(x - 3)}{(x - 4)(x - 3)} = \lim_{x \to 3} \frac{x + 2}{x - 4} = \lim_{x \to 3} \frac{5}{-1} = -5.
\]

Because we have an absolute value which is zero at our limit point, the (continuous) function inside the absolute value may change sign there and we should check left and right limits.

\[
\lim_{x \to 3^-} \frac{(x + 2)(x - 3)}{(x - 4)(x - 3)} = \lim_{x \to 3^-} \frac{5}{-1} = -5,
\]

\[
\lim_{x \to 3^+} \frac{(x + 2)(x - 3)}{(x - 4)(x - 3)} = \lim_{x \to 3^+} \frac{5}{-1} = -5.
\]

Since the left and right limits do not agree, the original (two-sided) limit does not exist.

With absolute value problems like the above example, it is often possible to see quickly if it should be replaced by the quantity inside, or its opposite (additive inverse). One only needs to check if it is positive or negative just left or just right of the limit point (depending upon the side from which we are approaching). Since \( x - 3 \) is negative just left of \( x = 3 \), we can use \( |x - 3| = -(x - 3) \) as \( x \to 3^- \). (Recall the piecewise definition of \( |x| \).

It is not the case that the two-sided limit will not exist whenever an absolute value is involved. Consider the next two examples.

Example 3.5.3 Compute, if possible, \( \lim_{x \to 0} \frac{x^2}{|x|} \).

Solution: As before, we will see what limits we get from both sides.

\[
\lim_{x \to 0^-} \frac{x^2}{|x|} = \lim_{x \to 0^-} \frac{x^2}{-x} = 0,
\]

\[
\lim_{x \to 0^+} \frac{x^2}{|x|} = \lim_{x \to 0^+} \frac{x^2}{x} = \lim_{x \to 0^+} (x) = 0.
\]

Since both are the same, we conclude\(^\text{29}\) \( \lim_{x \to 0} \frac{x^2}{|x|} = 0 \).

Example 3.5.4 Find \( \lim_{x \to 9} \frac{(2x - 18)|2x - 23|}{x^2 - 11x + 18} \).

Solution: For this function, the expression inside the absolute value does not change signs at \( x = 9 \) and so we can use older techniques:

\[
\lim_{x \to 9} \frac{(2x - 18)|2x - 23|}{x^2 - 11x + 18} = \lim_{x \to 9} \frac{2(x - 9)(|2x - 23|)}{(x - 9)(x - 2)} = \frac{2(23 - 2x)}{x - 2} = \frac{2(23 - 18)}{9 - 2} = \frac{10}{7}.
\]

\(^\text{29}\)Actually this function simplifies to \(|x|\) for \( x \neq 0 \), since \( \frac{x^2}{|x|} = \frac{|x|^2}{|x|} = |x| \), assuming \( |x| \neq 0 \), i.e., \( x \neq 0 \).
3.5. ONE-SIDED FINITE LIMITS

**Example 3.5.5** Suppose \( f(x) = \begin{cases} x + 1 & \text{for } x > 3 \\ 5x - 11 & \text{for } x \in [-2,3] \\ x + 3 & \text{for } x < -2. \end{cases} \)

Find all left, right, and two-sided limits, where possible, at \( x = 3, x = 1 \) and \( x = -2. \)

**Solution:** First we look at \( x = 3. \) In each case the key is to see which “piece” \( x \) is on in its approach to the limit point. (See Theorem 3.5.1(d),(e), page 223.)

\[
\lim_{x \to 3^-} f(x) = \lim_{x \to 3^-} (5x - 11) = 5(3) - 11 = 4, \\
\lim_{x \to 3^+} f(x) = \lim_{x \to 3^+} (x + 1) = 4.
\]

Since these limits’ values are the same, \( \lim_{x \to 3} f(x) = 4 \) also. (See Theorem 3.5.1(a), page 223.)

Next are the limits at \( x = 1. \) In all cases the computation is the same, but we will go ahead and write them out here. (With practice, for such a case only the third would likely be computed, the other two following.)

\[
\lim_{x \to 1^-} f(x) = \lim_{x \to 1^-} (5x - 11) = 5(1) - 11 = -6, \\
\lim_{x \to 1^+} f(x) = \lim_{x \to 1^+} (5x - 11) = 5(1) - 11 = -6, \\
\lim_{x \to 1} f(x) = \lim_{x \to 1} (5x - 11) = 5(1) - 11 = -6.
\]

Those were easier because \( x = 1 \) is safely inside the interval \([-2,3]\), on which \( f \) is defined by the (continuous) function \( 5x - 12. \) Finally we turn our attention to \( x = -2. \)

\[
\lim_{x \to -2^-} f(x) = \lim_{x \to -2^-} (x + 3) = -2 + 3 = 1, \\
\lim_{x \to -2^+} f(x) = \lim_{x \to -2^+} (5x - 11) = 5(-2) - 11 = -21.
\]

Since the left and right limits do not agree, we conclude \( \lim_{x \to -2} f(x) \) does not exist.

If we collect the information in the previous example, along with the values of \( f(x) \) at the limit points, we can use Theorem 3.5.1(b),(c), and the two-sided analogs, to make some conclusions regarding various types of continuity at these points:

- At \( x = 3, \) we have \( \lim_{x \to 3} f(x) = 4 = f(3), \) so \( f(x) \) is continuous at \( x = 3. \)

- At \( x = 1, \) we have \( \lim_{x \to 1} f(x) = -6 = f(1), \) so \( f(x) \) is continuous at \( x = 1. \)

- \( \lim_{x \to -2^-} f(x) = 1 \neq f(-2) = -21, \) so \( f(x) \) is not left-continuous at \( x = -2. \)

- \( \lim_{x \to -2^+} f(x) = -21 = f(-2), \) so \( f(x) \) is right-continuous at \( x = -2. \)

- \( f(x) \) is not continuous at \( x = -2, \) since it is not both left- and right-continuous at \( x = -2. \)

Furthermore, since \( \lim_{x \to -2} f(x) \) does not exist, it cannot equal \( f(-2), \) so Theorem 3.4.2, page 211 also gives us a discontinuity at \( x = -2. \)

- Since the function is “piecewise linear,” it is continuous at all other \( x \)-values. Thus it is continuous for \( x \neq -2, \) i.e., for all \( x \in \mathbb{R} - \{-2\} = (-\infty, -2) \cup (-2, \infty). \)
Just as with continuity and limits, the graph of a function can often indicate one-sided continuity and the values of one-sided limits. If we can “ride along” the graph towards the limit point from the prescribed direction, we can visually observe if some height is approached. While the function in Example 3.5.5 is graphed in Figure 3.14, we should be able to read many of the limit and continuity properties from the graph itself.

- \( \lim_{x \to -2^-} f(x) = 1 \)
- \( \lim_{x \to -2^+} f(x) = -21 \)
- \( \lim_{x \to -2} f(x) \) DNE
- \( f(-2) = -21 \)
- \( f(x) \) is right-continuous at \( x = -2 \)
- \( \lim_{x \to -3} f(x) = 0 \)
- \( f(3) = 4 \)

Note that the continuity on \([ -2, \infty )\), by definition, means \( f(x) \) is continuous at each \( x \in (-2, \infty) \) (the open interval) and right-continuous at \( x = -2 \). (This is Definition 3.3.2, page 195.) Continuity on \((-\infty, -2)\) simply means (two-sided) continuity at each \( x \in (-\infty, -2) \) (Definition 3.3.1, page 194).
3.5. ONE-SIDED FINITE LIMITS

Exercises

1. Consider the function graphed in Figure 3.15. From looking at the graph, answer the following questions. (It is possible that a requested limit does not exist.)

(a) \( \lim_{x \to -5^+} f(x) = \)
(b) \( \lim_{x \to -5^-} f(x) = \)
(c) \( \lim_{x \to -5} f(x) = \)
(d) \( \lim_{x \to -5^+} f(x) = \)
(e) \( \lim_{x \to -5^-} f(x) = \)
(f) \( \lim_{x \to -5} f(x) = \)
(g) \( \lim_{x \to -1^-} f(x) = \)
(h) \( \lim_{x \to -1^+} f(x) = \)
(i) \( \lim_{x \to -1} f(x) = \)
(j) \( \lim_{x \to 0^-} f(x) = \)
(k) \( \lim_{x \to 0^+} f(x) = \)
(l) \( \lim_{x \to 0} f(x) = \)
(m) \( \lim_{x \to 1^-} f(x) = \)
(n) \( \lim_{x \to 1^+} f(x) = \)
(o) \( \lim_{x \to 1} f(x) = \)
(p) \( \lim_{x \to -3} f(x) = \)
(q) Is \( f(x) \) continuous at \( x = -5 \)? Is \( f(x) \) left-continuous at \( x = -5 \)? Is \( f(x) \) right-continuous at \( x = -5 \)?
(r) Repeat for \( x = -1 \).
(s) Repeat for \( x = 1 \).
(t) Repeat for \( x = 5 \).
(u) Repeat for \( x = 0 \).

2. The function in Figure 3.15 can be defined piecewise as follows:

\[
\begin{align*}
    f(x) &= \begin{cases} 
    x - 1 & \text{for } x \in (1, 5) \\
    -2(x - 1) & \text{for } x \in [-1, 1] \\
    \frac{3}{2}(x + 3) & \text{for } x \in [-5, -1].
    \end{cases}
\end{align*}
\]

Using this fact (and not referring directly to the graph), answer (a)–(t) as in the previous exercise. (For example,
\[
\lim_{x \to 2} f(x) = \lim_{x \to 2} (x - 1) = 2 - 1 = 1.
\]

For continuity and other questions here, see also Theorem 3.5.1 and the previous examples.

3. Consider the function \( f(x) = \sqrt{x^2 - 16} \).

(a) If possible compute \( \lim_{x \to -4} f(x) = \)

(b) If possible compute \( \lim_{x \to -4^+} f(x) = \)

(c) If possible compute \( \lim_{x \to -4^-} f(x) = \)

(d) Do we have continuity, left-continuity, right-continuity or neither at \( x = -4 \)?

(e) If possible compute \( \lim_{x \to -4} f(x) = \)

(f) If possible compute \( \lim_{x \to -4^+} f(x) = \)

(g) If possible compute \( \lim_{x \to -4^-} f(x) = \)

(h) Do we have continuity, left-continuity, right-continuity or neither at \( x = -4 \)?

(i) If possible compute \( \lim_{x \to 4} f(x) = \)

(j) If possible compute \( \lim_{x \to 4^+} f(x) = \)

(k) If possible compute \( \lim_{x \to 4^-} f(x) = \)

(l) Do we have continuity, left-continuity, right-continuity or neither at \( x = 5 \)?

For each of the following limits 4–8, compute its value or (if appropriate) state why it does not exist. Show details. (For some, you may need to check both left-side and right-side limits.)

4. \( \lim_{x \to 3} \frac{|x - 5|}{x + 4} \)

5. \( \lim_{x \to -1^-} \frac{|x - 1|}{x^2 + 3x - 4} \)

6. \( \lim_{x \to -4} \frac{|x + 4|}{x^2 - 16} \)

7. \( \lim_{x \to -4} \frac{|x^2 - 16|}{x - 4} \)

8. \( \lim_{x \to -4} \frac{x^2 - 8x + 16}{|x - 4|} \)

9. Consider the following function (where \( m \) and \( b \) will be determined later).

\[
f(x) = \begin{cases} 
\sqrt{x} & \text{for } x > 4 \\
mx + b & \text{for } -4 \leq x \leq 4 \\
-3 - (x + 4)^2 & \text{for } x < -4.
\end{cases}
\]

(a) Find \( \lim_{x \to 4^+} f(x) \).

(b) Find \( \lim_{x \to 4^-} f(x) \).

(c) Find \( \lim_{x \to -4^+} f(x) \).

(d) Find \( \lim_{x \to -4^-} f(x) \).

(e) Use these to find \( m \) and \( b \) so that \( f(x) \) is continuous on \( \mathbb{R} \).
3.6 Infinite Limits at Points

In this section we extend the notion of limit so it can describe quantities which are growing without bound in many circumstances. The prototype function to help define what we mean here will be the familiar \( f(x) = 1/x \), which we graph in Figure 3.16 above. As \( x \) approaches zero from the right, we have \( 1/x \) returning larger and larger positive numbers, with unbounded growth in \( 1/x \). Similarly as \( x \) approaches zero from the left we have \( 1/x \) returning larger and larger negative numbers, growing without bound. For this function we would write

\[
\lim_{x \to 0^+} \frac{1}{x} = \infty \quad \text{(or just } \infty), \quad (3.32)
\]

\[
\lim_{x \to 0^-} \frac{1}{x} = -\infty, \quad (3.33)
\]

\[
\lim_{x \to 0^0} \frac{1}{x} \text{ does not exist (} 1/0^\pm). \quad (3.34)
\]

The above limits are the most important such examples. The first two limits are particular determinate forms, \( 1/0^+ \) and \( 1/0^- \) respectively, which we will discuss later in this section. The third limit does not exist because the left-side and right-side limits do not agree, as often happened with finite limits. We will still label its form \( 1/0^\pm \). For completeness and future reference we now give definitions of what it means for a limit at a point, from the left or right, to be \( \infty \):

**Definition 3.6.1** For \( a \in \mathbb{R} \), we say

\[
\lim_{x \to a^-} f(x) = \infty \iff (\forall M \in \mathbb{R})(\exists \delta > 0)(\forall x)(0 < |x - a| < \delta \implies f(x) > M), \quad (3.35)
\]

\[
\lim_{x \to a^+} f(x) = \infty \iff (\forall M \in \mathbb{R})(\exists \delta > 0)(\forall x)(x \in (a, a + \delta) \implies f(x) > M), \quad (3.36)
\]

\[
\lim_{x \to a} f(x) = \infty \iff (\forall M \in \mathbb{R})(\exists \delta > 0)(\forall x)(x \in (a - \delta, a) \implies f(x) > M). \quad (3.37)
\]
Figure 3.17: A function $f(x)$ with $\lim_{x \to a} f(x) = \infty$. By (3.35), for any height $M$ we can find $\delta > 0$ so that $0 < |x - a| < \delta \implies f(x) > M$. In other words, we can choose any height, and then force $f(x)$ to be still higher than that height by forcing $x$ to be within some $\delta$ of $a$, but (as always with limits) never actually equal to $a$. Notice that since $f(x)$ is continuous near $x = a$ (but not at $x = a$), the limit as $x \to a$ being infinite gives a vertical asymptote there (given by the dashed vertical line at $x = a$).

In other words, to say that a limit is $\infty$ is to say that we can force $f(x)$ to be greater than any previously chosen number $M$ by forcing $x$ to be within $\delta$ of $a$ (but not equal to $a$), from one or both directions depending upon if it is a right, left, or two-sided limit. (Of course the choice of $\delta$ depends upon the choice of $M$.) A graphical example of (3.35) is given in Figure 3.17.

Proofs using these definitions are interesting, and we will include one here, but we will have numerous shortcuts based upon general observations.

**Example 3.6.1** Prove that $\lim_{x \to 0^+} \frac{1}{x} = \infty$.

**Solution:** Here we start with $M \in \mathbb{R}$ and try to work back towards some $\delta > 0$, so that $x \in (0, \delta) \implies f(x) = \frac{1}{x} > M$, as in (3.36). Upon reflection, solving this inequality for $x$ seems to be the correct strategy to find $\delta$, though we cannot immediately because we do not know if $M > 0$.

In fact this does not have to be an issue, in the sense that if we can prove that the implication is true for some pair $\delta, M > 0$, the same $\delta$ will work for lesser (in particular nonpositive) $M$ values as well; if $M_1 < M_2$, then any $\delta$ which makes the implication $x \in (0, \delta) \implies f(x) > M$ true for $M = M_2$ will also make it true for $M = M_1$, as $f(x) > M_2 \implies f(x) > M_1$.\(^{30}\) When

---

\(^{30}\)This is similar to some freedoms we have with our choices of $\delta, \varepsilon > 0$ in the continuity definition and proofs. We already took advantage of the fact that we could a priori restrict the values of $\delta$ to be bounded by a fixed positive number (we used 1 usually), because if $0 < \delta_1 < \delta_2$ and $\delta_2$ worked for $\delta$ in the implication $|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon$, then so did $\delta_1$ because $|x - a| < \delta_1 \implies |x - a| < \delta_2$. Furthermore, we could
thus streamlining the proof argument, the usual phrasing used would in this situation read, “without loss of generality, we will assume $M > 0$.”

We will eventually use the following fact, which we have used before (as in Example 3.1.7, page 175) regarding the effect of the reciprocal function on inequalities of positive numbers: $(\forall a, b > 0)[a < b \iff 1/a > 1/b]$, which itself is easily proved if one simply divides both sides of $a < b$ by $ab > 0$, giving $1/b < 1/a$, essentially $1/a > 1/b$ as claimed, written backwards.

So, assuming $M > 0$ and $x \in (0, \delta)$ for some $\delta > 0$, we now compute a value of $\delta$ which guarantees $f(x) > M$. We thus begin with $f(x) > M$ and work backwards towards $\delta$ (note $x, M > 0$):

$$f(x) > M \iff \frac{1}{x} > M \iff x < \frac{1}{M}.$$  

We will take $\delta = 1/M$ in our proof.

**Proof:** Let $M \in \mathbb{R}$. Without loss of generality assume $M > 0$. Take $\delta = 1/M$. Then

$$x \in (0, \delta) \iff 0 < x < \delta \iff f(x) = \frac{1}{x} > \frac{1}{\delta} = \frac{1}{1/M} = M, \text{ q.e.d.}$$

In this section we will not emphasize proofs, in which we would need to show that a particular limit conforms to a particular limit definition. While interesting in their own rights, from the example above we can see that such proofs will often contain interesting but perhaps distracting technicalities.

Instead we will emphasize the limit forms $1/0^+$, $1/0^-$, $1/0^\pm$ and their variations, and argue intuitively what these limits’ values will be.

However there is value in being aware of the various limit definitions as they are given precisely, and so for completeness we next list the similar definitions for the limits at a point being $-\infty$ and invite the reader to make some sense of these as well as the previous definitions.

**Definition 3.6.2** For $a \in \mathbb{R}$, we say

$$\lim_{x \to a} f(x) = -\infty \iff (\forall N \in \mathbb{R})(\exists \delta > 0)(\forall x)(0 < |x - a| < \delta \implies f(x) < N), \quad (3.38)$$

$$\lim_{x \to a^+} f(x) = -\infty \iff (\forall N \in \mathbb{R})(\exists \delta > 0)(\forall x)(x \in (a, a + \delta) \implies f(x) < N), \quad (3.39)$$

$$\lim_{x \to a^-} f(x) = -\infty \iff (\forall N \in \mathbb{R})(\exists \delta > 0)(\forall x)(x \in (a - \delta, a) \implies f(x) < N). \quad (3.40)$$

Thus, to say a limit is $-\infty$ is to say that the function’s height can be forced to be below any given (earlier) fixed level by forcing $x$ to be within some distance of $a$, ignoring the case $x = a$ as we always do with limits.

Earlier we had the *indeterminate* limit form $0/0$ (meaning that knowing the numerator and denominator both approach zero does not tell us the value of the limit). In many of these next cases we will have variations of forms $1/0^+$, $1/0^-$ and $1/0^\pm$, which are not *indeterminate*, i.e., are *determinate* and as such should become intuitive upon reflection. In the definitions below, we will write the functions NUM for the numerator, and DEN for the denominator. (Note that here “$\longrightarrow$” means “approaches.”)
Definition 3.6.3 For a function \( f(x) = \frac{\text{NUM}(x)}{\text{DEN}(x)} \), define the following forms, as \( x \) approaches the limit point in the prescribed way:

1. \( 1/0^+ \): any limit in which \( \text{NUM}(x) \to 1 \) and \( \text{DEN}(x) \to 0^+ \), the latter meaning the denominator is positive but approaching zero.

2. \( 1/0^- \): any limit in which \( \text{NUM}(x) \to 1 \) and \( \text{DEN}(x) \to 0^- \), the latter meaning the denominator is negative but approaching zero.

3. \( 1/0^\pm \): any limit in which \( \text{NUM}(x) \to 1 \) and \( \text{DEN}(x) \to 0 \), with \( \text{DEN}(x) \neq 0 \) but sometimes \( \text{DEN}(x) \) is positive, other times negative.

Theorem 3.6.1 Any limit of the form \( 1/0^+ \) will return \( +\infty \); any limit of the form \( 1/0^- \) will return \( -\infty \); and any limit of the form \( 1/0^\pm \) will not exist.

This is well illustrated in the behavior of \( f(x) = 1/x \) as zero is approached, as in Figure 3.16 at the opening of this section (page 229), and the corresponding limits (3.32), (3.33) and (3.34).

Example 3.6.2 Consider the following limits and their methods of computation:

1. \[ \lim_{x \to 5^+} \frac{1}{\sqrt{x^2 - 25}} = \frac{1}{0^+} = \infty. \]
2. \[ \lim_{x \to 9^+} \frac{10 - x}{9 - x} = \frac{1}{0^-} = -\infty. \]
3. \[ \lim_{x \to 3} \frac{1}{x - 3} \text{ does not exist } (1/0^\pm). \]
4. \[ \lim_{x \to 4} \frac{x}{(x - 4)^2} = \frac{4}{0^+} = \infty. \]

In the last limit, the denominator is positive for \( x \to 4 \) from both sides, because it is a perfect (polynomial) square. Notice also that, strictly speaking, it is of the form \( 4/0^+ \), but that is also determinate (i.e., not indeterminate). In fact, it is just \( 4 \) times a limit which is of the form \( 1/0^+ \) if we factor out the \( 4 \) (and accept that \( 4 \cdot \infty = \infty \)):

\[
\lim_{x \to 4} \frac{x}{(x - 4)^2} = 4 \lim_{x \to 4} \frac{x/4}{(x - 4)^2} = 4(1/0^+) = 4 \cdot \infty = \infty.
\] (3.41)

Notice that limits allow us to—somewhat—extend our arithmetic when it is understood that it is really a statement about limit forms. For instance, we could say that

\[
(\forall a > 0)[a \cdot \infty = \infty], \quad \text{and} \quad (\forall a > 0)[a \cdot (-\infty) = -\infty].
\] (3.42)

These mean that if part of the function approaches \( a > 0 \), and the other “approaches” \( \infty \), then so does the limit of the product approach \( \infty \). Similarly \( a \cdot (-\infty) \) gives us \( -\infty \). For (3.41), we used the fact that if a function grows positive without bound as we approach the limit point, then so will something which is roughly \( 4 \) times that function (it will just “grow” roughly four times as fast). On the other hand,

\[
(\forall a < 0)[a \cdot \infty = -\infty], \quad \text{and} \quad (\forall a < 0)[a \cdot (-\infty) = \infty].
\] (3.43)

Multiplying a function, which “blows up” as we approach the limit point, by another function which approaches a negative number, does not change the fact that the product will still blow up, but it will occur in the other direction, i.e., with the other sign.
3.6. INFINITE LIMITS AT POINTS

In terms of division, we can write\(^3\)

\[
(\forall a > 0) \left[ \left( \frac{a}{0^+} = a \cdot \frac{1}{0^+} = a \cdot \infty = \infty \right) \land \left( \frac{a}{0^-} = a \cdot \frac{1}{0^-} = a \cdot (-\infty) = -\infty \right) \right], \quad (3.44)
\]

\[
(\forall a < 0) \left[ \left( \frac{a}{0^+} = a \cdot \frac{1}{0^+} = a \cdot \infty = -\infty \right) \land \left( \frac{a}{0^-} = a \cdot \frac{1}{0^-} = a \cdot (-\infty) = \infty \right) \right]. \quad (3.45)
\]

We take this opportunity to point out that knowing where these infinite limits occur is useful in sketching a graph of the function. Consider for instance the limit from the previous example:

\[
\lim_{x \to 4} \frac{x}{(x - 4)^2} = \infty.
\]

If we were to construct a sign chart for this function, and note the vertical asymptote at \(x = 4\), we get some idea of what its graph looks like near that vertical asymptote.

In fact this function also has horizontal asymptotes, a topic which will occur in a later section. A computer-generated graph is given in Figure 3.18, page 234, and so our sign chart at least accurately displays the sign and the behavior as \(x \to 4\). (Note also the \(x\)-intercept at \(x = 0\).)

Knowing the function “blows up” at \(x = 4\), along with the sign chart, gives us the limits from both sides. For another example, we revisit the function \(f(x) = \frac{x}{(x + 1)(x - 1)}\) which we encountered in Example 3.3.7, page 204.

**Example 3.6.3** The following limits can be found by considering the sign chart for the function \(f(x) = \frac{x}{(x + 1)(x - 1)}\):

\[
\lim_{x \to -1^-} \frac{x}{(x + 1)(x - 1)} = -\infty, \quad \lim_{x \to -1^+} \frac{x}{(x + 1)(x - 1)} = \infty,
\]

\[
\lim_{x \to 1^-} \frac{x}{(x + 1)(x - 1)} = -\infty, \quad \lim_{x \to 1^+} \frac{x}{(x + 1)(x - 1)} = \infty.
\]

All of these are of forms \(1/0^+, 1/0^-, -1/0^+\) or \(-1/0^-\), since the numerators are approaching \(\pm 1\) and the denominators are approaching zero from left or right. Thus we have the function “blowing up” at \(\pm 1\), so we just need to know what are its signs as we approach \(\pm 1\) from either side. We repeat the sign chart from Example 3.3.7, page 204 for convenience, though we add features to the chart to illustrate locations of the \(x\)-intercept and vertical asymptotes.

---

\(^3\) Another way to look at these limits is to first make note that they are blowing up—for instance \(\text{NUM} \to a \neq 0\) but \(\text{DEN} \to 0\)—and then just take note of the sign of the fraction as a whole. If it is positive and blowing up, the limit must be \(\infty\); if negative and blowing up, the limit must be \(-\infty\); if both signs occur consistently as \(x\) approaches the limit point, and the fraction is blowing up in absolute size, then the limit must not exist (part of the function blows up towards \(\infty\), and another part towards \(-\infty\), as \(x \to a\)).
Figure 3.18: Partial graph of \( f(x) = \frac{x}{(x - 4)^2} \). Besides the vertical asymptote at \( x = 4 \), note the sign change at \( x = 0 \), the latter being not easily discerned due to the scale being true. (For instance, \( f(-1) = -1/25 = -0.04 \).)

Function: \[ f(x) = \frac{x}{(x + 1)(x - 1)} \]

Test \( x = \)

<table>
<thead>
<tr>
<th>(-2)</th>
<th>(-0.5)</th>
<th>(0.5)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>⊗ ⊗</td>
<td>⊗ ⊗</td>
<td>⊗ ⊗</td>
<td>⊗ ⊗</td>
</tr>
</tbody>
</table>

Sign Factors:

\[ f(x) :\]

\[ ∅ ⊗ -1 ⊗ 0 ∅ ⊗ ∅ \]

The graph of \( f(x) = \frac{x}{(x + 1)(x - 1)} \) is given in Figure 3.19, page 235.

While the sign chart can help in computing limits, its construction is often more work than required. If we are already building it for some other reason, we can go ahead and use it to compute our limits at the vertical asymptotes, but first we need to know that there is in fact “blow up” as we approach those input values, and so we need an eye to what a computation without a sign chart would look like. Thus, more examples of such computations are in order.

Example 3.6.4 Consider the following limits.

1. \[ \lim_{x \to 9^+} \frac{x}{3 - \sqrt{x}} = \frac{9/0^-}{-∞}, \]

2. \[ \lim_{x \to 6^-} \frac{3 - 2x}{x^2 + 2x - 48} = \lim_{x \to 6^-} \frac{3 - 2x}{(x + 8)(x - 6)} = \frac{-9}{(14)(0^-)} = ∞. \]
3.6. INFINITE LIMITS AT POINTS

These are fairly routine, but did use some subtlety to get the correct signs. In the first one, it was important to notice that the denominator $3 - \sqrt{x}$ is negative for $x > 9$. The second one can be thought of as a form $\frac{-1}{(x-3)^2} = \infty$. (It also helped to have the denominator factored.)

It is possible that a $0/0$ form can simplify to one of these determinate forms, as in the following example. (We will start to use the shorthand “=D.N.E.” when a limit does not exist.)

Example 3.6.5 Consider the following limits. (The algebra is the same for each so we only show details for the first limit.)

1. $\lim_{x \to 3} \frac{81 - x^4}{(x^2 - 6x + 9)^2} = \frac{0}{0}$

   $\lim_{x \to 3} \frac{(9 - x^2)(9 + x^2)}{[(x - 3)^2]^2} = \frac{0}{0}$

   $\lim_{x \to 3} \frac{(3-x)(3+x)(9 + x^2)}{(x-3)^4} = \frac{-6}{6} \cdot \frac{18}{6} = D.N.E.$

2. $\lim_{x \to 3^+} \frac{81 - x^4}{(x^2 - 6x + 9)^2} = \frac{0}{0}$

   $\lim_{x \to 3^+} \frac{-(3+x)(9 + x^2)}{(x-3)^3} = \frac{-6}{6} \cdot \frac{18}{6} = -\infty.$

3. $\lim_{x \to 3^-} \frac{81 - x^4}{(x^2 - 6x + 9)^2} = \frac{0}{0}$

   $\lim_{x \to 3^-} \frac{-(3+x)(9 + x^2)}{(x-3)^3} = \frac{-6}{6} \cdot \frac{18}{6} = \infty.$

Considerations also need to be taken for limits involving rational exponents. Recall that if $a/b$ is a simplified fraction (so we do not allow 4/6 but do allow 2/3, for example), then

$$x^{a/b} = (x^a)^{1/b} = \left(x^{1/b}\right)^a.$$  \hspace{1cm} (3.46)

Also recall that $x^{1/b} = \sqrt[b]{x}$. Finally, $(x^2)^{1/2} = \sqrt{x^2} = |x|$, for instance, while $(x^3)^{1/3} = \sqrt[3]{x^3} = x$. As usual, the odd roots are simpler to deal with than the even roots, at least for abstract computations.
Example 3.6.6 Consider the following limits:

1. \( \lim_{x \to 0^-} \frac{1}{x^{2/3}} = \lim_{x \to 0^-} \frac{1}{\sqrt[3]{x^2}} = \frac{1}{0^{+}} \infty, \)

2. \( \lim_{x \to 0^-} x^{-5/3} = \lim_{x \to 0^-} \frac{1}{\sqrt[3]{x^5}} = \frac{1}{0^{+}} \infty, \)

3. \( \lim_{x \to -4} \frac{x}{(x + 4)^{4/3}} = \lim_{x \to -4} \frac{x}{[(x + 4)^4]^{1/3}} = -4/0^{+} \infty, \)

4. \( \lim_{x \to -4^-} \frac{x}{(x + 4)^{1/3}} = -4/0^{-} \infty, \)

5. \( \lim_{x \to -4^-} \frac{x}{[(x + 4)^2]^{1/2}} = \lim_{x \to -4^-} \frac{x}{|x + 4|} = -4/0^{+} \infty. \)

There are applications where it is interesting to know what occurs in the “extreme,” or limit, case. We now consider two of these.

Example 3.6.7 According to electrostatic theory, if two protons are brought to within the distance \( d > 0 \) of each other, the magnitude of the repelling force they would exert on each other will be given by \( F = \frac{k}{d^2} \), where \( k \) is a positive constant. Then since \( k > 0 \) we can write

\[
\lim_{d \to 0^+} F = \lim_{d \to 0^+} \frac{k}{d^2} = \infty.
\]

Thus, according to electrostatic theory it would require infinite force to bring two protons together such that the distance between them is zero.\(^{32}\)

Example 3.6.8 Suppose an object lies along a line running through the center of a thin double convex lens, and further suppose that the line is perpendicular to the plane containing the outer edges of the lens. Also suppose the lens has a focal length of \( f \), and \( d_o \) is the distance from the object to the center of the lens, where \( d_o > f > 0 \). If \( d_i \) is the distance from the lens center at which the resulting image of the object would be located on the opposite side of the lens, then the thin lens equation states that

\[
\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}.
\]

Find the trend in the distance \( d_i \) as the object is placed closer and closer to the focal length. Assume \( d_o > f > 0 \) throughout.

Solution: We wish to compute \( \lim_{d_o \to f^+} d_i \), so first we solve for \( d_i \):

\[
\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o} \Rightarrow d_i = \frac{1}{\frac{1}{f} - \frac{1}{d_o}} d_o f = d_o f \frac{d_o f}{d_o - f} \Rightarrow \lim_{d_o \to f^+} d_i = \lim_{d_o \to f^+} \frac{d_o f}{d_o - f} = \frac{d_o f/0^+}{\text{ALG}} \infty.
\]

\(^{32}\)The electrostatic theory here is far from complete, because it does not take into account other forces and quantum mechanical effects, nor is it clear what it means for the distance to be zero. One should ask if \( d \) is only the distances between their “centers,” for instance, or is it truly the distances to their edges, and is the “edge” of a proton well-defined? Still it is interesting to see what does the theory say about what occurs to the force “in the limit.”
Thus, as the object approaches the focal length position, the image appears to move farther and farther away (and more rapidly away) on the opposite side of the lens. One might conclude that the image of an object placed at the focal length will never be seen on the opposite side of the lens. This effect can be observed in a laboratory to the extent we can approach the ideal of having an object lie precisely and only at the focal length, and numerically using some sample values. Suppose \( f = 20\text{cm} \). Testing several relevant values of \( d_o \), using \( d_i = d_o f / (d_o - f) \)

\[
d_o = 25\text{cm} \implies d_i = \frac{(25\text{cm})(20\text{cm})}{25\text{cm} - 20\text{cm}} = \frac{500\text{cm}^2}{5\text{cm}} = 100\text{cm},
\]

\[
d_o = 22\text{cm} \implies d_i = \frac{(22\text{cm})(20\text{cm})}{22\text{cm} - 20\text{cm}} = \frac{440\text{cm}^2}{2\text{cm}} = 220\text{cm},
\]

\[
d_o = 21\text{cm} \implies d_i = \frac{(21\text{cm})(20\text{cm})}{21\text{cm} - 20\text{cm}} = \frac{420\text{cm}^2}{1\text{cm}} = 420\text{cm},
\]

\[
d_o = 20.5\text{cm} \implies d_i = \frac{(20.5\text{cm})(20\text{cm})}{20.5\text{cm} - 20\text{cm}} = \frac{410\text{cm}^2}{0.5\text{cm}} = 820\text{cm}.
\]

Finally, we would again see (for the case \( f = 20\text{cm} \)) that (suppressing units)

\[
d_o \to 20^+ \implies d_i = \frac{d_o \cdot 20}{d_o - 20} = \frac{20}{1 - \frac{20}{d_o}} \to \frac{20}{0^+} \to \infty.
\]

**Exercises**

Compute each limit (stating which ones do not exist) without reference to a sign chart, unless otherwise instructed. In computing these limits, write the form which allows you to make the conclusion where appropriate. (See examples throughout this section.)

1. \( \lim_{x \to -5} \frac{1}{(x + 5)^2} \)
2. \( \lim_{x \to 1^+} \frac{x}{x^2 - 1} \)
3. \( \lim_{x \to 1^-} \frac{x}{x^2 - 1} \)
4. \( \lim_{x \to 1} \frac{x}{x^2 - 1} \)
5. \( \lim_{x \to -1^+} \frac{x}{x^2 - 1} \)
6. \( \lim_{x \to -1^-} \frac{x}{x^2 - 1} \)
7. \( \lim_{x \to -1} \frac{x}{x^2 - 1} \)
8. \( \lim_{x \to -2^+} \frac{x^2 - 4x + 4}{x^2 - 4} \)
9. \( \lim_{x \to -2^-} \frac{x^2 - 4x + 4}{x^2 - 4} \)
10. \( \lim_{x \to -2} \frac{x^2 - 4x + 4}{x^2 - 4} \)
11. \( \lim_{x \to -2^+} \frac{x^2 - 4x + 4}{x^2 - 4} \)
12. \( \lim_{x \to -2^-} \frac{x^2 - 4x + 4}{x^2 - 4} \)
13. \( \lim_{x \to -2} \frac{x^2 - 4x + 4}{x^2 - 4} \)
14. \( \lim_{x \to -3} \frac{x}{|x - 3|} \)
15. \( \lim_{x \to -3} \frac{x}{|x - 3|} \)
16. \( \lim_{x \to -3^+} \frac{|x - 3|}{x^2 - 9} \)
17. \( \lim_{x \to -3^-} \frac{|x - 3|}{x^2 - 9} \)
18. \( \lim_{x \to -3} \frac{|x - 3|}{x^2 - 9} \)
19. \( \lim_{x \to -3^+} \frac{|x - 3|}{x^2 - 9} \)
20. \[ \lim_{x \to -3^-} \frac{|x - 3|}{x^2 - 9} \]

21. \[ \lim_{x \to -3^+} \frac{|x - 3|}{x^2 - 9} \]

22. Suppose \( f(x) = \frac{x}{\sqrt{1 - x^2}} \).

   (a) Discuss all possible points \( a \in \mathbb{R} \) at which \( f(x) \) may have infinite limits as \( x \) approaches \( a \) from one side or both sides. List all such limits and their values.

   (b) Draw a sign chart for this function. (First, note its domain; where is it defined?)

   (c) Use all the information above to sketch a rough graph of the function.

23. Suppose \( f(x) = \frac{1}{x^4 - 9} \).

   (a) Make a sign chart for \( f(x) \).

   (b) Discuss all possible points \( a \in \mathbb{R} \) at which \( f(x) \) may have infinite limits as \( x \) approaches \( a \) from one side or both sides. List all such limits and their values, using the sign chart above.

   (c) Use all the information above to sketch a rough graph of the function.

24. According to Einstein’s Special Relativity theory, the mass of an object with resting mass \( m \) and velocity \( v \) is given by

   \[ M(v) = \frac{m}{\sqrt{1 - \frac{v^2}{c^2}}} \]

   where \( c \) is the speed of light. Assume \( m > 0 \) and \( 0 \leq v < c \). Compute the following (for (b) and (c) write your final answer as some decimal number times \( m \), using three significant digits):

   (a) \( M(0) \)

   (b) \( M(c/2) \)

   (c) \( M(0.9c) \)

   (d) \( \lim_{v \to c^-} M(v) \)

   (e) What physical insight(s) should this limit provide us?
3.7 Sandwich, Composition and Trigonometric Continuity Theorems

In this section we will state the Sandwich Theorem and use it for computing several limits, including those which prove that the trigonometric functions are continuous where defined.

3.7.1 Sandwich Theorem

Theorem 3.7.1 (Sandwich Theorem) Suppose that there exists some $d > 0$ such that for every $x \in (a - d, a) \cup (a, a + d)$, i.e., for $0 < |x - a| < d$ we have

$$f(x) \leq g(x) \leq h(x).$$

(3.48)

Then

$$\left( \lim_{x \to a} f(x) = L \right) \land \left( \lim_{x \to a} h(x) = L \right) \implies \lim_{x \to a} g(x) = L.$$

The idea is that $f$ and $h$ “sandwich” $g$ between them, and so if $f$ and $h$ both approach $L$, then $g$ has nowhere to go but $L$. This is graphed for two cases in Figure 3.20, where $L$ first is a finite real number, and then where $L = \infty$. The functions $f(x)$ and $h(x)$ can be thought of as variable lower and upper bounds for the function $g(x)$ in between by (3.48). Thus the behavior of $f(x)$ and $h(x)$ can, in some circumstances (as in the theorem) force behavior from $g(x)$. The logic of the argument for the theorem is often graphed in various ways. We will employ the style of Figure 3.21, page 240 to illustrate our arguments, except that we will not include the labels “(Hypothesis)” and “(Conclusion),” as they will become apparent in context.

There are several variations of the Sandwich Theorem, in which behavior of one or more bounding functions $f(x)$ and $h(x)$ can force behavior upon a (variably) bounded function $g(x)$.

---

33The Sandwich Theorem is also called the Squeeze Theorem and the Pinching Theorem in other texts.
(Hypothesis) As \( x \to a \): \[
\begin{align*}
f(x) &\leq g(x) \\ g(x) &\leq h(x)
\end{align*}
\]

(Conclusion): \( \therefore g(x) \to L \)

**Figure 3.21**: Figure illustrating the argument for the Sandwich Theorem.

These variations are perhaps most clearly seen by graphing their respective situations. For instance, it is easily seen that we can replace \( f(x) \leq g(x) \leq h(x) \) with \( f(x) < g(x) < h(x) \) (see again Figure 3.20 at the beginning of this section).\(^{34}\) One-sided versions of the theorem also hold, as in for instance the left-sided limit version:

\[
\left( (\exists d > 0)(\forall x) \left[ x \in (a - d, a) \implies f(x) \leq g(x) \leq h(x) \right] \right) \land \left( \lim_{x \to a^-} f(x) = L = \lim_{x \to a^-} h(x) \right)
\]

\[ \implies \lim_{x \to a^-} g(x) = L. \]

The following limit is a very traditional example for the original statement of the Sandwich Theorem. Note that it relies on the fact that \( \sin \theta \) is defined for every \( \theta \in \mathbb{R} \), and that \( -1 \leq \sin \theta \leq 1 \).

**Example 3.7.1** Compute the limit \( \lim_{x \to 0} x \sin \frac{1}{x} \).

**Solution**: Note that \( x \sin \frac{1}{x} \) is always between \( x \cdot 1 \) and \( x \cdot (-1) \), but these switch roles as top and bottom bounding functions depending upon the sign of \( x \). However, we can always write

\[
-|x| \leq x \sin \frac{1}{x} \leq |x|.
\]

By continuity of \( |x| \) and \( -|x| \), we have

\[
\lim_{x \to 0} (-|x|) = -|0| = 0, \quad \text{and} \quad \lim_{x \to 0} |x| = |0| = 0,
\]

so by the Sandwich Theorem we must conclude as well that

\[
\lim_{x \to 0} x \sin \frac{1}{x} = 0.
\]

\(^{34}\)Note that \( f(x) < g(x) < h(x) \implies f(x) \leq g(x) \leq h(x) \), so the fact that we can replace the latter with the former follows quickly from logic; if we have the strict inequalities, then we also have the non-strict inequalities and that hypotheses of the Sandwich Theorem will still hold.

\(^{35}\)We can make the argument leading to (3.49) more precise. For \( x \neq 0 \), we can say

\[
\left| x \sin \frac{1}{x} \right| = |x| \cdot \left| \sin \frac{1}{x} \right| \leq |x| \cdot 1 = |x|,
\]

\[ \implies \left| x \sin \frac{1}{x} \right| \leq |x| \]

\[ \iff -|x| \leq x \sin \frac{1}{x} \leq |x|,
\]

where the \( \iff \) comes from the general rule that \( |z| \leq K \iff -K \leq z \leq K \).
3.7. SANDWICH, COMPOSITION, TRIGONOMETRIC CONTINUITY

\[ f(x) = -|x| \]

\[ g(x) = x \sin \frac{1}{x} \]

\[ h(x) = |x| \]

**Figure 3.22:** Partial graph of \( g(x) = x \sin \frac{1}{x} \), which is bounded from above by \( h(x) = |x| \) and from below by \( f(x) = -|x| \). It oscillates wildly, running through infinitely many periods in the argument \( 1/x \) of the sine function as \( x \to 0 \), but is bounded in amplitude by functions which shrink to zero as \( x \to 0 \). \( (g(x) \text{ is undefined at } x = 0 \text{ but that fact is not apparent in the graph above.}) \)

The Sandwich Theorem argument for the limit of Example 3.7.1 above can be summarized graphically as follows:

As \( x \to 0 \):

\[ -|x| \leq x \sin \frac{1}{x} \leq |x| \]

\[ \therefore x \sin \frac{1}{x} \to 0. \]

The function \( g(x) = x \sin \frac{1}{x} \) is graphed in Figure 3.22 above, together with the bounding functions \(-|x|\) and \(|x|\). It has some interesting features which make it very valuable for later examples which clarify some limit principles. We note how the argument \( 1/x \) of the sine function here runs through infinitely many periods of sine as \( x \to 0 \), so the function oscillates with infinitely increasing rapidity as \( x \to 0 \). However the “amplitude” \(|x|\) is variable and shrinking to zero.

One use of the above function is in illustrating a rather general theorem, based upon the Sandwich Theorem, regarding limits of products where one factor approaches zero while the other factor, however else it is ill-behaved, is at least bounded and defined as we approach the limit point.
Theorem 3.7.2 Suppose that $f(x)$ is defined for $0 < |x - a| < d$ for some $d > 0$, and that for such $x$, $f(x)$ is of the form $f(x) = g(x)h(x)$ where $|g(x)| \leq M$ and $h(x) \rightarrow 0$ as $x \rightarrow a$. Then $\lim_{x \rightarrow a} f(x) = 0$.

The proof consists of noting that $-M|h(x)| \leq f(x) \leq M|h(x)|$, so $\pm M|h(x)| \rightarrow 0$ as $x \rightarrow a$ implies $f(x) \rightarrow 0$ as $x \rightarrow a$ as well:

$$
\begin{align*}
\text{As } x \rightarrow a: \quad & -M|h(x)| \leq g(x)h(x) \leq M|h(x)| \\
& \Downarrow \\
0 & \Downarrow 0 \\
\therefore f(x) = g(x)h(x) & \rightarrow 0.
\end{align*}
$$

This theorem could have been used in the previous example to compute that limit immediately. We will have more use for this theorem in the next section. For instance, there we will use “$B$” to refer to a function which is defined and bounded as we approach the limit point. Then we point out that “$B \cdot 0$” is in fact a determinate form which yields zero in the limit. For the previous example, we would note the fact that for $x \neq 0$ we have $|\sin \frac{1}{x}| \leq 1$ and so, aside from being defined, $\sin \frac{1}{x}$ is bounded as $x \rightarrow 0$, and we can write

$$
\lim_{x \rightarrow 0} x \sin \frac{1}{x} = 0.
$$

While based upon the Sandwich Theorem, the above argument is somewhat intuitive and certainly more concise.

3.7.2 “Approaches” for Independent Versus Dependent Variables

This was not such an important issue before (though a reader might have wondered about this point), so it was deferred until now, and given its own subsection here to be sure it is clarified.

The point is that when we consider the independent variable $x$ “approaching” some point, say $x \rightarrow a$, we should visualize it gradually getting closer to that point $a$—as close as we like and then even closer—but never actually achieving the value $x = a$. That is built into, for instance, the definition of limit, in the antecedant $0 < |x - a| < \delta$ of the defining implication. On the other hand, we have more flexibility in the consequent $|f(x) - L| < \varepsilon$, though we still write $f(x) \rightarrow L$. For instance, in our latest example, $x \sin \frac{1}{x} \rightarrow 0$, that function not only go closer to zero consistently, but also achieved the value zero repeatedly (infinitely many times!) as $x \rightarrow 0$. So the independent variable $x$ is forced to approach and avoid its limiting value, but the dependent variable need not actually avoid its limiting value when we write $x \rightarrow a \implies f(x) \rightarrow L$.

For another, rather trivial example, consider that $f(x) = 0 \cdot \sin x \rightarrow 0$ as $x \rightarrow \frac{\pi}{2}$. In fact, $f(x) = 0$ for all $x \in \mathbb{R}$, so the function not only approaches zero, it is never anything but zero. Still we use the notation $f(x) \rightarrow 0$.

3.7.3 “Sandwiching” From One Side

One only needs one bounding function in an infinite limit case, to have a valid Sandwich Theorem-type argument In (3.50) below, $f$ is the bounding function pushing $h$, while in (3.51) $h$ is the bounding function pushing $f$. 
Theorem 3.7.3 Suppose \( f(x) \leq h(x) \) on \( 0 < |x - a| < d \) for some \( d > 0 \). Then (separately) we have

\[
\begin{align*}
\lim_{x \to a} f(x) = \infty & \implies \lim_{x \to a} h(x) = \infty, \\
\lim_{x \to a} h(x) = -\infty & \implies \lim_{x \to a} f(x) = -\infty.
\end{align*}
\]

(3.50) (3.51)

In other words, if the lesser (lower) function blows up towards \( \infty \), then so must the greater (upper) function, while if the greater function has blowup towards \( -\infty \), then so must the lesser function. Such arguments can be verified easily by graphing the situations and seeing how the “blowup” of one function can force a similar behavior of another. It is also useful to see the following, somewhat visual style of the arguments of (3.50) and (3.51). Note that both diagrams below are, for now, hypothetical; instead of \( \therefore \), we could instead write \( \implies \).

As \( x \to a \):

- \( g(x) \leq h(x) \)
- \( f(x) \leq h(x) \)

\( \therefore h(x) \to \infty \)

\( \therefore f(x) \to -\infty \)

Of course we always have to be careful. For instance, suppose \( f(x) \leq h(x) \) and \( h(x) \to \infty \). It is not necessarily the case that \( f(x) \to \infty \) as well, since \( h(x) \) is above \( f(x) \), and thus unable to “push” \( f(x) \) up with it.

Example 3.7.2 Compute \( \lim_{x \to 2^+} \frac{x^2 + \sin x}{x - 2} \).

Solution: Since \(-1 \leq \sin x \leq 1\), and for \( x > 2 \) we have \( x^2 > 0, x - 2 > 0 \) we can write

\[
\frac{x^2 - 1}{x - 2} \leq \frac{x^2 + \sin x}{x - 2} \leq \frac{x^2 + 1}{x - 2}.
\]

(3.52)

In other words, the least that \( \frac{x^2 + \sin x}{x - 2} \) can be as \( x \to 2^+ \) is \( \frac{x^2 - 1}{x - 2} \), i.e., where \( \sin x = -1 \), and the greatest it can be is \( \frac{x^2 + 1}{x - 2} \), the case where \( \sin x = 1 \). Next we notice that

\[
\lim_{x \to 2^+} \frac{x^2 - 1}{x - 2} = \infty, \quad \lim_{x \to 2^+} \frac{x^2 + 1}{x - 2} = \infty,
\]

and so \( \lim_{x \to 2^+} \frac{x^2 + \sin x}{x - 2} = \infty \) as well.\(^{36}\)

As noted before, the first inequality in (3.52) is in fact enough to “push” the desired limit to be \( \infty \):

\(^{36}\)After the next subsection, where we prove trigonometric functions continuous where defined, the limit in Example 3.7.2 can be done directly:

\[
\lim_{x \to 2^+} \frac{x^2 + \sin x}{x - 2} = \infty,
\]

since \( x^2 + \sin x \to 4 + \sin 2 > 0 \). Though not necessary here, it is perhaps easier to see we have the correct sign if we note that \( \sin 2 \approx 0.909097426 \), and thus \( x^2 + \sin x \to 4 + \sin 2 \approx 4.909097426 \). (The sign of \( x^2 + \sin 2 \) was going to be positive as \( x \to 2 \) anyways because \( \sin 2 \in [-1, 1] \), while \( x^2 \to 4 \).)
As $x \to 2^+$:
\[
\frac{x^2 - 1}{x - 2} \leq \frac{x^2 + \sin x}{x - 2} \to \infty
\]

In fact the example above will not need this Sandwich Theorem-type argument if we note that $\sin 2 \approx 0.909297426 > 0.9$. That is because we could have written
\[
\lim_{x \to 2^+} \frac{x^2 + \sin x}{x - 2} = \lim_{x \to 2^+} \left[ \frac{x^2}{x - 2} + \frac{\sin x}{x - 2} \right] \to \infty.
\]
This used the fact that, as a form, $\infty + \infty$ yields limits which are $\infty$, as we will note in later sections and should be intuitive now. That does not mean we can avoid using such arguments altogether (recall our first example, $x \sin \frac{1}{x} \to 0$ as $x \to 0$). A small change makes the case for such a method: consider a similar limit but where the numerator of the function is now $x^2 - \sin x$. Then we still have
\[
\lim_{x \to 2^+} \frac{x^2 + \sin x}{x - 2} = \lim_{x \to 2^+} \left[ \frac{x^2}{x - 2} + \frac{\sin x}{x - 2} \right] \to \infty.
\]
that is, $\infty - \infty$ form which is indeterminate (first discussed properly in Section 3.8). However the earlier Sandwich Theorem-type argument still applies, since $-\sin x \in [-1, 1]$:

As $x \to 2^+$:
\[
\frac{x^2 - 1}{x - 2} \leq \frac{x^2 - \sin x}{x - 2} \to \infty
\]

In fact the above limit still does not require the Sandwich Theorem-type argument if it is noticed that $x^2 - \sin x \to 4 - 0.909297426 > 0$, but with practice it is a reasonably quick argument leading to the conclusion that the limit is in fact $\infty$.

### 3.7.4 Limits with Compositions of Functions

**Theorem 3.7.4** Suppose that, for some limiting behavior of $x$, we have $g(x) \to L$, and $f(x)$ continuous at $x = L$. Then for the same limiting behavior of $x$ we have $f(g(x)) \to f(L)$.

So for instance, if $f(x)$ is continuous at $L$ and $\lim_{x \to a} g(x) = L$, then
\[
\lim_{x \to a} f(g(x)) = f(L).
\]
3.7. SANDWICH, COMPOSITION, TRIGONOMETRIC CONTINUITY

i.e.,

$$\lim_{x \to a} f(g(x)) = f \left( \lim_{x \to a} g(x) \right) = f(L).$$  \hspace{1cm} (3.53)

In other words, if the limit of the “inside” function is a point of continuity of the “outside” function, then we can “move the limit notation (lim) inside.” (The reader is invited to consider this idea in light of a function diagram for $f(g(x))$.)

In fact in the interest of avoiding errors students are often discouraged from using (3.53) as it is written. One reason is that it is crucial that $f$ be continuous at the limit of $g$, or the theorem is false, as we will show in the exercises, while reading (3.53) without context may give the impression that we can always move the limit inside. Instead we will concentrate on the theorem itself, and the limit forms which arise from its proper application. Still, the theorem is very general in terms of the type of limit (basic, left, right or the “at infinity” variety we will have in Section 3.8). We will give a proof for the most basic case at the end of this section. That proof will be a simple modification of the proof of Theorem 3.2.5, page 183. The theorem can be illustrated graphically as follows:

As $x \to a$: \hspace{1cm} $f(g(x)) \hspace{1cm} \therefore$ (by continuity) \hspace{1cm} $f(L)$

Again, the diagram above is not necessarily true if $f(x)$ is not continuous at $x = L$. When $f$ is, we might not always give the justification “(by continuity)” within the diagram.

**Example 3.7.3** Suppose that $\lim_{x \to a} g(x) = 12$, while $\lim_{x \to b} g(x) = 0$. Then

- $\lim_{x \to a} \sqrt[3]{g(x)} = \sqrt[3]{12},$
- $\lim_{x \to b} \sqrt[3]{g(x)} = \sqrt[3]{0} = 0,$
- $\lim_{x \to a} \sqrt{g(x)} = \sqrt{12},$  \hspace{1cm} $\sqrt{0} = 0,$
- $\lim_{x \to b} \sqrt{g(x)}$ cannot be determined by the given information. It depends upon how $g(x)$ approaches 0 as $x \to b$:
  - If $g(x) \to 0^+$ then $\lim_{x \to b} \sqrt[3]{\sqrt{g(x)}} = \sqrt{0} = 0$: $\lim_{x \to b} \sqrt[3]{g(x)} = 0.$
  - In fact, if we just have $g(x) \geq 0$ as $x \to b$, we have the limit being $\sqrt{0} = 0.$
  - However, if $g(x) \to 0^-$, then $\lim_{x \to b} \sqrt[3]{\sqrt{g(x)}} = \sqrt[6]{0}$ does not exist.

In the first two limits above, we have $g(x) \to 12$ or 0, which are well-within the set on which $\sqrt[3]{x}$ is continuous, namely $\mathbb{R}$. For the third limit above, we recall that $\sqrt[3]{x}$ is continuous for $x > 0$, and so $g(x) \to 12 \implies \sqrt[3]{g(x)} \to \sqrt[3]{12}$. Since $\sqrt{x}$ is defined for $x \geq 0$, and is in fact right-continuous at $x = 0$, it is possible that if $g(x) \to 0$, then $\lim_{x \to b} \sqrt{g(x)}$ is zero or does not
exist. The right-continuity of $\sqrt{x}$ at $x = 0$ guarantees that $\sqrt{g(x)} \to 0$ if $g(x) \to 0^+$, but does not exist if $g(x)$ is sometimes negative as $x \to b$.

This theorem will prove to be more useful—and in fact will be crucial—in later sections, but we will make some use of it in this section as an alternative method for some limit computations for which earlier methods are not quite as efficient.

### 3.7.5 Continuity Considerations for Trigonometric Functions

Our theorem is as follows:

**Theorem 3.7.5** The six basic trigonometric functions, $\sin x$, $\cos x$, $\tan x$, $\cot x$, $\sec x$ and $\csc x$ are continuous everywhere they are defined. Thus

1. $\sin x$ and $\cos x$ are continuous for $x \in \mathbb{R}$.
2. $\tan x$ and $\sec x$ are continuous except where $\cos x = 0$, and are thus continuous for all
   \[ x \neq \pm \frac{\pi}{2}, \pm \frac{3\pi}{2}, \pm \frac{5\pi}{2}, \ldots \]
3. $\cot x$ and $\csc x$ are continuous except where $\sin x = 0$, and are thus continuous for all
   \[ x \neq 0, \pm \pi, \pm 2\pi, \pm 3\pi, \ldots \]

Considering the unit circle definitions of $\sin \theta$ and $\cos \theta$, it is reasonable that these are continuous (as functions of $\theta$). The continuity of the other trigonometric functions, which are quotients of $1$, $\sin \theta$ and $\cos \theta$ where they are defined then follows immediately. Because the results listed in Theorem 3.7.5 are intuitive, we will defer the proof until the end of the section. The proofs that $\sin \theta$ and $\cos \theta$ are continuous use the sandwich theorem and a geometric argument, and are interesting in their own rights, but for now we will concentrate our efforts in applications of the theorem.

**Example 3.7.4** The following limits follow directly from continuity of the trigonometric functions (where defined):

- $\lim_{x \to \pi} \sin x = \sin \pi = 0$.
- $\lim_{x \to \pi/4} \tan x = \tan \frac{\pi}{4} = 1$.
- $\lim_{x \to \sqrt{\pi}} \cos x^2 = \cos (\sqrt{\pi})^2 = \cos \pi = -1$.

The last limit above was computable as shown since $x^2$ is continuous on all of $\mathbb{R}$, and so is $\cos x$, so the composition $\cos x^2$ is continuous on all of $\mathbb{R}$, including $x = \sqrt{\pi}$ (see Theorem 3.2.5, page 183).

In the last example above, the cosine function was the “outer” function, but the trigonometric functions can be combined with other functions in a variety of ways.

**Example 3.7.5** Consider the following limit computations.

- $\lim_{x \to \pi} \frac{1 - \sin^2 x}{\cos x} = 0/0$.
- $\lim_{x \to \frac{\pi}{4}} \frac{\cos^2 x}{\cos x} = 0/0$.
- $\lim_{x \to \frac{\pi}{2}} \cos x = \cos \frac{\pi}{2} = 0$. 

3.7. SANDWICH, COMPOSITION, TRIGONOMETRIC CONTINUITY

\[ \lim_{x \to 0} \sqrt{1 - \cos^2 x} \quad \sqrt{1 - \cos^2 0} = \sqrt{1 - 1^2} = \sqrt{0} = 0. \]

The first was a standard \(0/0\)-form simplification using the trigonometric identity \(\cos^2 x + \sin^2 x = 1\). The second relied upon the fact that \(\cos^2 \theta < 1\) as \(x \to 0\). In fact it was enough that \(\cos x \in [-1, 1]\), so \(\cos^2 x \in [0, 1]\) and so, though we are only interested in behavior as we approach zero, in fact the expression inside the square root, \(1 - \cos^2 x\) is never negative (and is everywhere continuous): \(\mathbb{R}^\cos^2 \mathbb{R} [-1, 1] \to [0, 1] \to [0, 1]\).

**Example 3.7.6** Consider the following trigonometric limits: \(\lim x \to 0 \sin x / x = 1\). Since \(0/0\) is indeterminate, so is \(\sin \theta / \theta\). (Note that all angles are assumed to be measured in radians.)

\[ \lim_{x \to 3} \sin \left( \frac{x^2 - 9}{x^2 - 5x + 6} \right) = \lim_{x \to 3} \sin \left( \frac{(x + 3)(x - 3)}{(x - 2)(x - 3)} \right) = \lim_{x \to 3} \sin \left( \frac{x + 3}{x - 2} \right) = 3 + \frac{3}{3 - 2} = 6 \approx -0.279415498. \]

Since \(0/0\)-form is indeterminate, any “function” of it is also, so we have to deal with the argument (“inside”) of the sine function. Of course the exact answer is \(\sin 6\), but may naturally be curious about what is the approximate value of this limit so a nine-digit approximation is also included.

In Theorem 3.7.4, page 244 we had an alternative method for analyzing the previous example’s limit: we could instead exploit the everywhere-continuity of the sine function to allow manipulations such as

\[ \lim_{x \to 3} \sin \left( \frac{x^2 - 9}{x^2 - 5x + 6} \right) = \sin \left( \lim_{x \to 3} \frac{x^2 - 9}{x^2 - 5x + 6} \right) = \cdots = \sin 6. \]

We will usually opt for the first method—as in Example 3.7.6 above—where possible for reasons stated previously.\(^{37}\)

Trigonometric functions can also give rise to infinite-valued limits. In such cases it is crucial to determine from within which quadrant the argument of the function is approaching the limit point, and thus the sign of the trigonometric functions. If the argument is \(x\), then \(x \to 0^+\) means the “angle” \(x\) approaches zero from within the first quadrant (where \(x\) is the measure of an angle in standard position). For \(x \to \pi^+\), we have \(x\) approaching \(\pi\) from angles with measure slightly greater than \(\pi\), i.e., from the third quadrant. For \(x \to \frac{\pi}{2}^+, \) we are in the second quadrant, and so on. See Figure 3.23, page 248 where the part of “\(x\)” is played by the angle \(\theta\).

**Example 3.7.7** Consider the following trigonometric limits:

1. \(\lim_{x \to 0^+} \csc x = \lim_{x \to 0^+} \frac{1}{\sin x} = \infty\),
2. \(\lim_{x \to \pi^+} \csc x = \lim_{x \to \pi^-} \frac{1}{\sin x} = -\infty\),

\(^{37}\)Many textbooks prescribe exactly this second method (not preferred here) for such a problem. Instead we solved it by replacing the original function with one which was continuous at \(x = 3\), and were thus able to call upon Theorem 3.4.3, page 212. We will require this alternative kind of manipulation later, particularly with limits “at infinity,” but we will otherwise usually avoid it when possible because it requires very specific hypotheses. It is acceptable here since we know that if the limit inside exists and is finite, then the sine function is continuous there (since it is continuous everywhere!). We need more care if the outer function is not continuous on all of \(\mathbb{R}\).
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

3.7.6 Proofs

First we prove Theorem 3.7.4, page 244, which states that if \( g(x) \to L \), and \( f(x) \) is continuous at \( x = L \), then \( f(g(x)) \to f(L) \). Here we will prove the basic case where the limiting behavior is as \( x \to a \in \mathbb{R} \) (\( a \) not infinite).

**Proof:** Here we will prove the basic case where \( a \in \mathbb{R} \) (not infinite):

\[
\left( \lim_{x \to a} g(x) = L \right) \land (f(x) \text{ continuous at } x = L) \implies \left( \lim_{x \to a} f(g(x)) = f(L) \right).
\]

To show this, we have to show that for any \( \varepsilon > 0 \), we can find a \( \delta > 0 \) such that

\[
0 < |x - a| < \delta \implies |f(g(x)) - f(L)| < \varepsilon.
\]

By our continuity and limit assumptions, we know that

\[
(\forall \varepsilon_1 > 0)(\exists \delta_1 > 0)(\forall x)(|x - L| < \delta_1 \implies |f(x) - f(L)| < \varepsilon_1), \quad (3.54)
\]

\[
(\forall \varepsilon_2 > 0)(\exists \delta_2 > 0)(\forall x)(0 < |x - a| < \delta_2 \implies |g(x) - L| < \varepsilon_2). \quad (3.55)
\]
3.7. SANDWICH, COMPOSITION, TRIGONOMETRIC CONTINUITY

So for this $\varepsilon$, choose $\varepsilon_1 = \varepsilon$, which gives a $\delta_1 > 0$ so that

$$|x - L| < \delta_1 \implies |f(x) - f(L)| < \varepsilon.$$

Next set $\varepsilon_2 = \delta_1 > 0$. This gives a $\delta_2 > 0$ so that

$$0 < |x - a| < \delta_2 \implies |g(x) - g(a)| < \varepsilon_2 = \delta_1.$$

Finally, let $\delta = \delta_2$, corresponding to $\varepsilon_2$ in the limit requirement for $g(x) \to L$. This gives (with the part of “$x$” in (3.54) played by $g(x)$ in the third and fourth lines below):

$$0 < |x - a| < \delta \iff 0 < |x - a| < \delta_2 \implies |g(x) - L| < \varepsilon_2 \implies |f(g(x)) - f(L)| < \varepsilon = \varepsilon_1,$$

q.e.d.

Next we prove that the trigonometric functions $\sin x$, $\cos x$, $\tan x$, $\cot x$, $\sec x$ and $\csc x$ are continuous wherever they are defined.

**Proof:** Our “proof” will be in four parts, and will be cut somewhat shorter than a proof from “first principles” would be by using an observation about the geometry of the unit circle. In this abbreviated proof we will see the sandwich theorem in action, in particular as applied to a useful inequality, (3.56), which will be our “observation.”

The order in which we will prove our results is as follows: 1. continuity of $\sin x$ at $x = 0$, implying 2. continuity of $\cos x$ at $x = 0$, together implying 3. continuity of $\sin x$ and $\cos x$ at every $x \in \mathbb{R}$, which implies 4. continuity of the other trigonometric functions wherever they are defined.

1. $\sin x$ is continuous at $x = 0$.

Consider the unit circle graphed in Figure 3.24. Now $|\sin x|$ is the distance from the horizontal axis to a point $P$ on the terminal side of the angle. The arc is another, but non-straight path of length $|x|$ from the horizontal axis to $P$. Thus

$$|\sin x| \leq |x|,$$

which is the same as $-|x| \leq \sin x \leq |x|$. Letting $x \to 0$, we get the following:

$$-|x| \leq \sin x \leq |x| \to 0 \implies \lim_{x \to 0} \sin x = 0.$$

The Sandwich Theorem then gives us $\lim_{x \to 0} \sin x = 0$. Since $\sin 0 = 0$ as well, we have $\sin x$ is continuous at $x = 0$, q.e.d.\footnote{Recall that $f(x)$ is continuous at $x = a$ if and only if $\lim_{x \to a} f(x) = f(a)$. See Theorem 3.4.2, page 211.}
2. \( \cos x \) is continuous at \( x = 0 \). This follows immediately, since near \( x = 0 \) (so the “angle” \( x \) terminates in the first or fourth quadrants) we have \( \cos x > 0 \) and thus (again, near \( x = 0 \)) \( \cos x = \sqrt{1 - \sin^2 x} \), and so we can replace \( \cos x \) with that expression (according to Theorem 3.4.3):

\[
\lim_{x \to 0} \cos x = \lim_{x \to 0} \sqrt{1 - \sin^2 x} = \sqrt{1 - \sin^2 0} = \sqrt{1} = 1 = \cos 0, \quad \text{q.e.d.}
\]

We will take a moment here to explain why we could compute the above limit as we did. Because \( \sin x \) is continuous at \( x = 0 \), so is \( 1 - \sin^2 x \), and since that function approaches \( 1 > 0 \) as \( x \to 0 \), its square root is also continuous at \( x = 0 \).

3. \( \sin x \) and \( \cos x \) are continuous for all \( x \in \mathbb{R} \). These follow from the two results above and the trigonometric identities (??) and (??) as below:

\[
\lim_{x \to a} \sin x = \lim_{x \to a} \sin(a + (x - a)) \\
= \lim_{x \to a} (\sin a \cos(x - a) + \cos a \sin(x - a)) = \sin a \cos 0 + \cos a \sin 0 \\
= (\sin a)(1) + (\cos a)(0) = \sin a,
\]

\[
\lim_{x \to a} \cos x = \lim_{x \to a} \cos(a + (x - a)) \\
= \lim_{x \to a} (\cos a \cos(x - a) - \sin a \sin(x - a)) = \cos a \cos 0 - \sin a \sin 0 \\
= (\cos a)(1) - (\sin a)(0) = \cos a, \quad \text{q.e.d.}
\]
Here we used what we will later call a substitution argument, which will be introduced properly in Section 3.9 (though we could also invoke Theorem 3.7.4, page 244). The idea is, roughly, that $x \to a \iff x - a \to 0$ in the sense of limit (where $x$ is never actually equal to $a$, and $x - a$ is never equal to zero).

4. All six trigonometric functions are continuous where they are defined. Of course $\sin x$ and $\cos x$ were already shown continuous for all $x \in \mathbb{R}$, i.e., where defined, earlier. The other functions are defined by quotients where the numerators are either $\sin x$, $\cos x$ or 1, which are continuous everywhere, while the denominators are either $\sin x$ or $\cos x$, again continuous everywhere. Since a ratio of two functions is continuous if both numerator and denominator are continuous and the denominator is nonzero, the functions $\tan x$ and $\sec x$ are continuous except where $\cos x = 0$, and $\cot x$ and $\csc x$ are continuous except where $\sin x = 0$. Summarizing, all trigonometric functions are continuous where defined, q.e.d.
Exercises

1. Compute \( \lim_{x \to 0^+} \sqrt{x} \sin \left( \frac{1}{x} \right) \).

2. Compute \( \lim_{x \to 1^+} \frac{\sin x}{x - 1} \). (Hint: \( \sin 1 \approx 0.841470985 \).)

3. Compute using a Sandwich Theorem-type argument \( \lim_{x \to 5^+} \frac{x + \cos x}{x^2 - 25} \).

4. Compute \( \lim_{x \to 2} \cos \left( x^2 - 4x + 4 \right) \).

5. Compute the following limits.
   (a) \( \lim_{x \to \frac{\pi}{2}^+} \sec x \).
   (b) \( \lim_{x \to \frac{\pi}{2}^-} \sec x \).
   (c) \( \lim_{x \to \frac{\pi}{2}^+} \sec x \).
   (d) \( \lim_{x \to \frac{\pi}{2}^-} \sec x \).

6. Compute the following limits.
   (a) \( \lim_{x \to 0^+} \cot x \).
   (b) \( \lim_{x \to 0^-} \cot x \).
   (c) \( \lim_{x \to \pi^+} \cot x \).
   (d) \( \lim_{x \to \pi^-} \cot x \).

7. Compute \( \lim_{x \to 0^+} \sqrt{x} \sin (\csc x) \).

8. Suppose that \( f(x) \leq h(x) \) for \( 0 < |x - a| < d \), for some \( d > 0 \), and that \( \lim_{x \to a} h(x) = \infty \). By drawing several graphs, show that \( \lim_{x \to a} f(x) \) can be anything: finite, \( \infty \), \( -\infty \), or nonexistent.

9. Suppose that \( -x^3 + 2x^2 - x + 2 < f(x) < x^2 - 2x + 3 \) for all \( x \in [0, 2] \), except for \( x \neq 1 \). Find \( \lim_{x \to 1} f(x) \) if possible.

10. Suppose for all \( x \neq 2 \) we have \( 4 \leq f(x) \leq (x - 2)^2 + 4 \). Find \( \lim_{x \to 2} f(x) \) if possible.

For the following, compute each limit which exists, state which do not, and if you use the Sandwich Theorem to prove one exists, show all details.

11. \( \lim_{x \to 0} \sqrt[3]{x} \sin \frac{1}{x} \)

12. \( \lim_{x \to 0} \sqrt{x} \sin \frac{1}{x} \)

13. \( \lim_{x \to 0} \sqrt[3]{x} \sin \frac{1}{x} \)

14. \( \lim_{x \to 0} \sqrt{x^2 \sin^2 \frac{1}{x}} \)

15. \( \lim_{x \to 0^+} \sqrt[3]{x} \sin (\csc x) \)

16. \( \lim_{x \to 0^+} \sqrt[3]{x} \sin \left( \csc \left( \frac{1}{x} \right) \right) \)

17. \( \lim_{x \to 0} x^2 \cos \frac{1}{\sqrt{x}} \)

18. \( \lim_{x \to 0} \sqrt[3]{x} \sin \left( \frac{1}{x} \right) \)

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

20. \( \lim_{x \to 0} \sin x \csc x \)

21. \( \lim_{x \to 0} \cot x \csc x \)
3.8. Limits “At Infinity”

The limits we introduce here differ from previous limits in that here we are interested in the behavior of functions \( f(x) \) as \( x \) grows without bound, rather than as \( x \) approaches a finite point. There are new “forms” we will come across here, such as \( 1/\infty \), \( \infty \cdot \infty \), \( \infty/\infty \), \( 0 \cdot \infty \) and \( \infty - \infty \). (Only the first two determinate.)

The first forms we will look at are \( 1/\infty \) and \( 1/(\infty) \). For these we look again to the function \( f(x) = 1/x \). Due to the importance of this function we produce it here for the third time, in Figure 3.25. We see that as \( x \) moves to the right through values like \( x = 1, 2, 3, 10, 100, 1000, 10^6 \) and so on, the function takes on respective values \( f(x) = 1, 1/2, 1/3, 1/10, 1/100, 1/1000, 10^{-6} \) and so on. So as \( x \) grows without bound, the function’s output shrinks towards (though is never equal to, for this case) zero. A similar phenomenon occurs when we take \( x \)-values \( x = -1, -2, -3, -10, -100, -1000, -10^6 \), etc., except the values of \( f(x) \) are then \( f(x) = -1, -1/2, -1/3, -1/10, -1/100, -1/1000, -10^{-6} \) etc. So as \( x \) moves left without bound, the function values are negative numbers shrinking in absolute size. The fact that in both cases we can get as close to zero in the values of \( f(x) \) as we could like (without necessarily achieving the value zero) by choosing \( x \) large enough is reflected in the statements

\[
\lim_{x \to \infty} \frac{1}{x} = 0, \quad (3.57)
\]

\[
\lim_{x \to -\infty} \frac{1}{x} = 0. \quad (3.58)
\]

The forms \( 1/\infty \) and \( 1/(\infty) \) are determinate, both yielding zero limits. Recall that a growing denominator will produce a shrinking fraction.\(^3\) Furthermore reciprocals of large numbers give small numbers. From earlier discussions of the graph of \( y = 1/x \) we can see how, as \( x \) gets arbitrarily large, \( 1/x \) gets arbitrarily small (though never quite zero) in absolute size.

\(^3\)Unless there is another effect to counteract the growing denominator, such as a growing numerator. We will soon see that \( \infty/\infty \) is indeterminate.
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

\( L + \varepsilon \)

\( L \)

\( L - \varepsilon \)

\((\forall \varepsilon > 0)(\exists M \in \mathbb{R})(\forall x \in \mathbb{R})(x > M \rightarrow |f(x) - L| < \varepsilon)\)

Figure 3.26: Illustration of the definition of a finite limit \( L \) of a function as \( x \to \infty \).

It is common to read the left-hand side of (3.57) as, “the limit, as \( x \) approaches infinity, of \( 1/x \).” Of course \( x \) does not “get close” to \( \infty \), but the notation means that we are computing what the behavior of \( 1/x \) will be as \( x \) grows positive without bound. Similarly for \( x \to -\infty \). To make these precise, we give the following definitions.

**Definition 3.8.1** For a finite number \( L \in \mathbb{R} \), we say

\[ \lim_{x \to \infty} f(x) = L \iff (\forall \varepsilon > 0)(\exists M \in \mathbb{R})(\forall x \in \mathbb{R})(x > M \rightarrow |f(x) - L| < \varepsilon), \]

(3.59)

\[ \lim_{x \to -\infty} f(x) = L \iff (\forall \varepsilon > 0)(\exists N \in \mathbb{R})(\forall x \in \mathbb{R})(x < N \rightarrow |f(x) - L| < \varepsilon). \]

(3.60)

In (3.59), we could also write \( f((M, \infty)) \subseteq (L - \varepsilon, L + \varepsilon) \), while in (3.60), we could write \( f((-\infty, N)) \subseteq (L - \varepsilon, L + \varepsilon) \). A case of (3.59) for a particular \( \varepsilon \) is illustrated in Figure 3.26. We will leave the illustrations of (3.60) to the reader.

Next we point out that it is natural to have a notion of an infinite limit as \( x \to \infty \) or \( x \to -\infty \). For instance,

\[ \lim_{x \to \infty} x = \infty \]

(3.61)

seems quite reasonable, as does

\[ \lim_{x \to -\infty} x^2 = \infty. \]

(3.62)

There are many common functions which grow without bound as \( x \) grows without bound. Note that (3.62) can be thought of as a form \((-\infty) \cdot (-\infty)\) or \((-\infty)^2\), which reasonably yields the limit \( \infty \). On the other hand,

\[ \lim_{x \to -\infty} x^3 = -\infty, \]

(3.63)

since \( x^3 \to 0 \) and \( x^3 \) grows without bound as \( x \) grows larger, without bound but negative. We could think of the above limit as a form \((-\infty)^3\), giving the limit as \( -\infty \) as we should expect. In general, all positive powers of \( x \) will grow to \( +\infty \) as \( x \to \infty \), while even powers will grow to \( +\infty \) as \( x \to -\infty \) and odd powers will grow to \( -\infty \) as \( x \to -\infty \).\(^{41}\) Constant factors behave as

\(^{40}\)Of course \((-\infty) \cdot (-\infty)\) is a particular form representing a product of two functions which are both negative and growing without bound. The product is naturally positive and also growing without bound, the resulting limit then being \( \infty \).

\(^{41}\)Noninteger powers of \( x \) are more complicated for \( x \to -\infty \). Some approach \( +\infty \), some \( -\infty \) and some are undefined as \( x \to -\infty \). Such things will be discussed as they come up in the text.
before (see (3.44) and (3.45), page 233), as in
\[ \lim_{x \to \infty} 5x = \infty, \quad \lim_{x \to -\infty} (-3x) = -3(-\infty). \]

The definition of \( \lim_{x \to \infty} f(x) = \infty \) is given below:

**Definition 3.8.2** We make the following definition:

\[ \lim_{x \to \infty} f(x) = \infty \iff (\forall M)(\exists N)(\forall x)(x > N \implies f(x) > M). \tag{3.64} \]

In other words, for any fixed \( M \), we can force \( f(x) \) to be greater than \( M \) by taking \( x > N \), so that \( f((N, \infty)) \subseteq (M, \infty) \). The definition of \( f(x) \to -\infty \) as \( x \to \infty \), and similar definitions, are left as exercises. To see (3.64) in action, consider proving \( \lim_{x \to \infty} x^2 = \infty \). For any \( M \), we can take \( N = \sqrt{|M|} \geq 0 \) to get
\[ x > N = \sqrt{|M|} \implies f(x) = x^2 > N^2 = \left(\sqrt{|M|}\right)^2 = |M| \geq M. \]

We needed \( N \geq 0 \) so that \( x > N \implies x^2 > N^2 \).

Some relevant limit forms which occur in this and other contexts, and which are not indeterminate include the following:

1. \( \infty + a = \infty \) for any fixed \( a \in \mathbb{R} \),
2. \( \infty + \infty = \infty \),
3. \( a \cdot \infty = \infty \) if \( a > 0 \), but \( a \cdot \infty = -\infty \) if \( a < 0 \).

As before, we can perform some “arithmetic” of limit forms, though we always have to be careful (see Example 3.8.1 below).

The cases mentioned in 3. above was also mentioned in the previous sections, first on page 232. Note again that \( a \cdot \infty \) as a limit form means that we have a limit where one function is approaching \( a \), and the other \( \infty \) (positive and growing without bound in the limit), and so their product approaches \( \infty \) if \( a > 0 \), and \( -\infty \) if \( a < 0 \). If \( a = 0 \) the form is indeterminate, and we have to attempt to rewrite it algebraically to see if it can be written in a determinate form.

The above forms are relatively intuitive. The following are more subtle, and in fact indeterminate:
\[ \infty - \infty, \quad 0 \cdot \infty, \quad \infty/\infty, \quad 0/0. \]

To see the first is indeterminate, consider for instance the following limits of the form \( \infty - \infty \):\(^{42}\)

**Example 3.8.1**
\[
\lim_{x \to \infty} \left[ (x^2 + 1) - (x^2) \right] = \lim_{x \to \infty} 1 = 1,
\]
\[
\lim_{x \to \infty} \left( x^3 - x^2 \right) = \lim_{x \to \infty} x^2(x - 1) = \infty,
\]
\[
\lim_{x \to \infty} \left( x^4 - x^6 \right) = \lim_{x \to \infty} x^4(1 - x^2) = \infty.
\]

\(^{42}\) The first limit shows that it is possible to come up with a limit of the form \( \infty - \infty \) which when evaluated gives any predetermined real value we would like (just replace the number 1 with the desired value). The second and third show we can, furthermore, find limits of form \( \infty - \infty \) which return infinite limits as well.
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

The question for $\infty - \infty$ form becomes, which “infinity” is larger, i.e., which function grows faster when we have a difference $f(x) - g(x)$ of functions $f$ and $g$ which both grow without bound? Or is there ultimately a compromise? Similar examples can be found for forms $0 \cdot \infty$ (or $\infty \cdot 0$) and $\infty/\infty$. The former we look at next, with a few examples to show that it is in fact indeterminate.

Example 3.8.2 Consider the following limits:

\[
\lim_{x \to \infty} \left[ x \cdot \frac{1}{x^2} \right] = \lim_{x \to \infty} \frac{1}{x} = 0,
\]

\[
\lim_{x \to \infty} \left[ x^2 \cdot \frac{1}{x} \right] = \lim_{x \to \infty} x = \infty,
\]

\[
\lim_{x \to \infty} \left[ x \cdot \frac{5}{x} \right] = \lim_{x \to \infty} 5 = 5.
\]

Now let us turn to polynomial and rational functions. Our first theorem is the following:

Theorem 3.8.1 For a polynomial function \( p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \), where \( a_n \neq 0 \) (so the polynomial really is of degree \( n \)), we have

\[
\lim_{x \to \infty} p(x) = \lim_{x \to \infty} a_n x^n,
\]

\[
\lim_{x \to -\infty} p(x) = \lim_{x \to -\infty} a_n x^n.
\]

In other words, for \( x \to \infty \) and \( x \to -\infty \), a polynomial function’s growth is ultimately dictated by its leading (highest degree) term. Rather than prove this in general, we can see the essence of a proof in the following examples and leave the actual proof as an exercise.

Example 3.8.3 Consider the following limits. (Forms are first given above the “=,” and then simplified below.)

\[
\lim_{x \to \infty} (3x^2 - 5x + 11) = \lim_{x \to \infty} x^2 \left( 3 - \frac{5}{x} + \frac{11}{x^2} \right) \to \infty \left( 3-0+0 \right) \to \infty,
\]

\[
\lim_{x \to -\infty} (x^3 + 95x^2 - 15x + 1000) = \lim_{x \to -\infty} x^3 \left( 1 + \frac{95}{x} - \frac{15}{x^2} + \frac{1000}{x^3} \right) \to \infty \left( 1+0-0 \right) \to -\infty.
\]

When we factor out the highest power, the lower-order terms we are left with have negative powers of \( x \) which then shrink to zero, leaving only the coefficient of the highest-order term as a factor in the limit. This phenomenon is very useful when we look at rational limits as \( x \to \pm \infty \), which are often of the form $\infty/\infty$, $(-\infty)/(-\infty)$ and so on.
Example 3.8.4 Consider the following limits.

\[
\begin{align*}
\lim_{x \to \infty} \frac{3x^2 + 5x - 9}{6x + 11} &= \frac{\infty}{\infty}, \\
\lim_{x \to \infty} \frac{x^2 (3 + \frac{5}{x} - \frac{9}{x^2})}{x (6 + \frac{11}{x})} &= \frac{\infty}{\infty} \text{ ALG} \\
\lim_{x \to \infty} x \cdot \frac{3 + \frac{5}{x} - \frac{9}{x^2}}{6 + \frac{11}{x}} &= \infty, \\
\lim_{x \to \infty} \frac{9x^2 + 2x + 1}{16x^2 + 3x - 100} &= \frac{\infty}{\infty}, \\
\lim_{x \to \infty} \frac{x^2 (9 + \frac{2}{x} + \frac{1}{x^2})}{16 + \frac{3}{x} - \frac{100}{x^2}} &= \frac{\infty}{\infty} \text{ ALG} \\
\lim_{x \to \infty} \frac{9 + \frac{2}{x} + \frac{1}{x^2}}{16 + \frac{3}{x} - \frac{100}{x^2}} &= \infty, \\
\lim_{x \to -\infty} \frac{5 - 3x}{2x^2 + x + 1} &= \frac{\infty}{\infty}, \\
\lim_{x \to -\infty} \frac{x \cdot (\frac{5}{x} - 3)}{2 + \frac{1}{x} + \frac{1}{x^2}} &= \frac{\infty}{\infty} \text{ ALG} \\
\lim_{x \to -\infty} \frac{1}{x} \cdot \frac{\frac{5}{x} - 3}{2 + \frac{1}{x} + \frac{1}{x^2}} &= \infty.
\end{align*}
\]

A quick corollary—which we must be careful not to abuse—to our theorem is the following:

Theorem 3.8.2 For any rational function \( f(x) = \frac{p(x)}{q(x)} \), where \( p(x) = a_nx^n + \cdots + a_1x + a_0 \) and \( q(x) = b_mx^m + \cdots + b_1x + b_0 \), with \( a_n, b_m \neq 0 \), we have

\[
\begin{align*}
\lim_{x \to \infty} \frac{p(x)}{q(x)} &= \lim_{x \to \infty} \frac{a_nx^n}{b_mx^m}, \\
\lim_{x \to -\infty} \frac{p(x)}{q(x)} &= \lim_{x \to -\infty} \frac{a_nx^n}{b_mx^m}.
\end{align*}
\]

(3.67) (3.68)

So as we take \( x \to \infty \) or \( x \to -\infty \), the limiting behavior of a rational function is governed by the leading terms of the numerator and denominator.\(^{43}\) We will use this theorem for anticipating results, but will work the actual limits as in Example 3.8.4.\(^{44}\)

It is common for trigonometric limits, and variations of the Sandwich Theorem (originally Theorem 3.7.1, page 239) to appear with limits “at infinity.”

Example 3.8.5 Consider the limit \( \lim_{x \to -\infty} \frac{\cos x}{x} \). This yields to the Sandwich Theorem quickly:

\[\lim_{x \to -\infty} \frac{\cos x}{x} = 0.\]

\(^{43}\)Note that the “leading term” means the nonzero term of the highest degree, not necessarily the first term appearing. For instance, in the polynomial \( 6 - 5x^2 \), the leading term is \(-5x^2\).

\(^{44}\)We will continue to compute the limits longhand for three reasons. First it is good reinforcement of the underlying principles. Second, it is not entirely standard to write, for instance,

\[\lim_{x \to \infty} \frac{5x^2 + 3x - 11}{7x^2 - 9x + 1,000} = \lim_{x \to \infty} \frac{5x^2}{7x^2} = \lim_{x \to \infty} \frac{5}{7} = 5/7.\]

A reader might be confused about the whereabouts of the terms that were dropped, and generally lose confidence that the writer’s understanding is correct. Finally, the theorem requires that \( x \to \infty \) or \( x \to -\infty \), so if we reflexively drop terms we may be tempted to do so for a limit at a finite point, where the theorem does not hold. That said, it is not uncommon for a trained mathematician to simply drop all steps above and write

\[\lim_{x \to -\infty} \frac{5x^2 + 3x - 11}{7x^2 - 9x + 1,000} = \frac{5}{7}.\]
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

As \( x \to \infty \):

\[
\begin{array}{c}
-1 \\
\cos x \\
1
\end{array}
\leq \begin{array}{c}
\frac{\cos x}{x} \\
x
\end{array}
\leq \begin{array}{c}
0 \\
0
\end{array}

\therefore \frac{\cos x}{x} \to 0.

One would usually then summarize: \( \lim_{x \to \infty} \frac{\cos x}{x} = 0 \).

One could also look at the previous limit as one of a product of two functions, one which is bounded (\( \cos x \)), and the other which approaches zero (\( 1/x \)), yielding \( B \cdot 0 \) form, which is a determinate form giving zero in the limit. Furthermore we could define a form, “\( B/\infty \)” which will always yield zero since the denominator grows without bound (shrinking the fraction) while the numerator is unable to compensate (by growing the fraction) since it is bounded. We could also write \( B/\infty = B \cdot \frac{1}{\infty} = B \cdot 0 \). The “algebra” of forms is interesting and intuitive, but one needs to be careful to understand the underlying mechanisms to perform such calculations on forms.

**Example 3.8.6** Consider the limit \( \lim_{x \to \infty} (x + \sin x) \). Here we have a sum of functions, the first growing without bound and the second being bounded. Intuitively this sum should grow without bound since the function \( \sin x \) is unable to check the growth of \( x \). We can again use the Sandwich Theorem:

As \( x \to \infty \):

\[
\begin{array}{c}
x - 1 \\
x + \sin x \\
x + 1
\end{array}
\leq \begin{array}{c}
\infty \\
\infty
\end{array}
\leq \begin{array}{c}
\infty
\end{array}

\therefore (x + \sin x) \to \infty.

In fact, recall that in such a case we only need the first inequality above to form our conclusion.

We could look at the limit above as an example of a form we could define as “\( \infty + B \)” which will always give us the actual limit being \( \infty \). To see this, note that for such a case we are looking at sums \( f(x) + g(x) \) where \( g(x) \) is defined and bounded, i.e., \( |g(x)| \leq M \) for some finite fixed \( M \), and \( f(x) \to \infty \). By the boundedness of \( g(x) \), we get

\[
f(x) - M \leq f(x) + g(x) \leq f(x) + M.
\]

Since \( f(x) - M, f(x) + M \to \infty \), we would conclude \( f(x) + g(x) \to \infty \) as well.

It should be pointed out that the limits \( \lim_{x \to \infty} \sin x \) and \( \lim_{x \to \infty} \cos x \) both do not exist. This is because these functions oscillate between \(-1\) and \(1\), and do not approach any particular value to the exclusion of others (recall that a limit must be unique). However the limits above show that such functions can still be involved in limits “at infinity,” especially when their (bounded) oscillations can be checked by, or absorbed into, the influences of other functions in the limits.

The methods of of above two examples are important and should be mastered, but we can use observations about forms (proved the same ways) and have more abbreviated computations:

\[
\begin{align*}
\lim_{x \to \infty} \frac{\cos x}{x} & \quad B/\infty \quad 0, \\
\lim_{x \to \infty} (x + \sin x) & \quad \infty + B \quad \infty.
\end{align*}
\]
Here $B$ stands for any bounded function, including constants. In both cases, the “$B$” can not check the growth of the other (“$\infty$”) function, and so the other function’s influence ultimately prevails in the limit. Note that $B/\infty$ and $\infty + B$ are determinate forms. We list these and some others below. Note that the left sides are forms, and the right sides are final limit values.

\[
B/\infty = 0, \quad B + \infty = \infty, \quad B - \infty = -\infty.
\]

All these are intuitive and provable using the Sandwich Theorem and its variations. As always it is important that we are aware of technicalities. For instance, check the growth of the other ("\(\infty\)) function. Again, the last step utilized the fact that we sometimes do need to refer back to the Sandwich Theorem-type computations for these, unless the form gives us an obvious answer.

The continuity of the trigonometric functions (where they are defined) can also come into play with these limits, for instance in light of Theorem 3.7.4, page 244 on the compositions of functions, namely \((f(x)\) continuous at \(x = L) \land (g(x) \rightarrow L) \implies f(g(x)) \rightarrow f(L)\):

**Example 3.8.7** Consider \(\lim_{x \rightarrow \infty} \sec \left( \frac{x}{x^2 + 1} \right)\). From what we know of rational functions, the input of the secant function here is approaching zero. Since the secant (=1/cosine) is continuous at zero, our answer should be \(\sec 0 = 1/\cos 0 = 1/1 = 1\). For a more computational argument we might write

\[
\lim_{x \rightarrow \infty} \sec \left( \frac{x}{x^2 + 1} \right) = \sec \left( \frac{\frac{1}{x}}{1 + \frac{1}{x}} \right) = \sec \left( \frac{1}{x} \right) \cdot \frac{\sec \left( \frac{1}{x} \right)}{\sec \left( \frac{1}{x} \right)} = \sec 0 = 1.
\]

(We used the fact that \(1/x \rightarrow 0\) as \(x \rightarrow \infty\) in the denominator of the input of the secant function.) Again, the last step utilized the fact that \(\sec x\) is continuous at \(x = 0\).

---

45For example, consider \(f(x) = K/x,\ g(x) = x\) and \(x \rightarrow \infty\). This gives a limit of \(K\), i.e., \(f(x)g(x) = (K/x)\cdot x = K\rightarrow K\), and we can choose \(K\) to be anything real number we like.

46For example, this is exactly what occurs with the limit \(\lim_{x \rightarrow \infty} \left[ \left( \frac{3 + \sin x}{2} \right) \cdot x \right] = \infty\). This is because for \(x > 0\),

\[
\left( \frac{3 - 1}{2} \right) \cdot x \leq \left( \frac{3 + \sin x}{2} \right) \cdot x \leq \left( \frac{3 + 1}{2} \right) \cdot x,
\]

and so the function is sandwiched between \(x \rightarrow \infty\) and \(2x \rightarrow \infty\).
For our next example we will return to a form $\infty - \infty$. When problematic terms do not cancel from subtraction, a rewriting of the expression as a quotient will often achieve some useful cancellation or a determinate form. (This theme will return several times in the text.) In the case below, we use a conjugate multiplication step.

Example 3.8.8 Consider the limit \( \lim_{x \to \infty} \left( \sqrt{x^2 + x + 1} - x \right) \). Clearly \( x^2 + x + 1 \to \infty \), and so we are taking square roots of numbers as large as we like. In fact, since \( x > 0 \) we can write \( \sqrt{x^2 + x + 1} > \sqrt{x^2} = |x| = x \to \infty \) as \( x \to \infty \).

We solve this using the following method:

\[
\lim_{x \to \infty} \left( \sqrt{x^2 + x + 1} - x \right) \Rightarrow \lim_{x \to \infty} \left( \sqrt{x^2 + x + 1} - x \right) \cdot \frac{x + 1}{x + 1}
\]

\[
\begin{align*}
&= \lim_{x \to \infty} \frac{x^2 + x + 1 - x^2}{x^2 + x + 1 + x} \\
&= \lim_{x \to \infty} \frac{x + 1}{\sqrt{x^2 + x + 1 + x}} \\
&= \frac{\infty}{\infty} \\
&= \lim_{x \to \infty} \frac{x (1 + \frac{1}{x})}{\sqrt{1 + \frac{1}{x} + \frac{1}{x^2} + 1}} \\
&= \lim_{x \to \infty} \frac{1 + \frac{1}{x}}{\sqrt{1 + \frac{1}{x} + \frac{1}{x^2} + 1}} = \frac{1 + 0}{\sqrt{1 + 0 + 0 + 1}} = \frac{1}{2}.
\end{align*}
\]

The limit above is correct, but probably not at all obvious from the original form. Only after finding a useful fractional form could we use our earlier techniques to compute its value. Limits which are writable as ratios are often easier to solve than other forms. Here it allowed us to compare the powers of \( x \) in the numerator and denominator. In our first limit section we had many 0/0 forms which we could easily simplify to get determinate forms.

There are applications, both conceptual and practical, for limits as the input variable “blows up.” Many interesting applications involve exponential functions \( f(x) = a^x \), or their variants.

\[\text{Figure 3.27a.} \quad \text{Figure 3.27b.} \quad \text{Figure 3.27: Partial graphs of exponential functions } 2^x \text{ and } (1/2)^x, \text{ along with logarithmic functions } \log_2 x \text{ and } \log_{1/2} x, \text{ showing continuity and limiting behaviors.} \]
3.8. LIMITS “AT INFINITY”

such as \( f(x) = C \cdot a^{kx} \). These are continuous for \( x \in \mathbb{R} \), and their limits as \( x \to \infty \) or \( x \to -\infty \) are as follow:

\[
\begin{align*}
a > 1 : & \quad x \to \infty \implies a^x \to \infty, \\
& \quad x \to -\infty \implies a^x \to 0^+, \\
a \in (0, 1) : & \quad x \to \infty \implies a^x \to 0^+, \\
& \quad x \to -\infty \implies a^x \to \infty.
\end{align*}
\]

See Figure 3.27a.

This gives rise to limit forms, so for examples (recalling that \( e \approx 2.71828 > 1 \)), we can write

\[
\begin{align*}
\lim_{x \to \infty} 2x &= \infty, \\
\lim_{x \to \infty} 1.5^x &= \infty, \\
\lim_{x \to \infty} e^x &= \infty.
\end{align*}
\]

Related to the behaviors of the exponential functions are those of the logarithmic functions. Recall

\[
\log_a x = y \iff a^y = x,
\]

so when looking at \( y = \log_a x \) is the same as looking at \( x = a^y \), or \( y = a^x \) but with \( x \) and \( y \) trading roles. We can see from the graphs in Figure 3.27b that

\[
\begin{align*}
a > 1 : & \quad x \to \infty \implies \log_a x \to \infty, \\
& \quad x \to 0^+ \implies \log_a x \to -\infty, \\
a \in (0, 1) : & \quad x \to \infty \implies \log_a x \to -\infty, \\
& \quad x \to 0^+ \implies \log_a x \to \infty.
\end{align*}
\]

These are the logarithmic analogs of (3.73)–(3.76). In fact it is not immediately clear from the figure that \( \log_2 x \to \infty \) as \( x \to \infty \), but we can go back to our definition in (3.64), page 255, and so for \( M > 0 \) we can take \( N = 2^M \), and get \( x > N = 2^M \implies \log_2 x > \log_2 2^M = M \). So the logarithmic graphs do “blow up” for \( (a > 0) \land (a \neq 1) \), though they do so very slowly (for instance for \( \log_2 x > 10 \) we need \( x > 2^{10} = 1024 \)). We will thus get limit forms such as \( \log_2(\infty) \) yielding a limit of \( \infty \), \( \log_2(0^+) \) yielding \( -\infty \), and others. Recall that \( \ln x = \log_e x \), with \( e \approx 2.71828 > 1 \) and so \( \ln x \) has a similar shape and asymptotics as \( \log_2 x \), which is shown in Figure 3.27b, page 260.

We can now quickly compute some limits involving logarithms:

\[
\begin{align*}
\lim_{x \to 0^+} \sin(x) &= 0, \\
\lim_{x \to \infty} (x^2 + 5x - 9) &= \infty, \\
\lim_{x \to \infty} \frac{1}{\ln x} &= 0.
\end{align*}
\]
Note that \( \lim_{x \to \pi/2^+} \ln(\tan x) \) does not exist, because \( \tan x < 0 \) as \( x \to \pi/2^+ \), i.e., when \( x \) is in the second quadrant; recall that logarithms can only process positive numbers.

Next we consider an application of such limits. Limits as the input variable grows towards \(+\infty\) are particularly valuable in the analysis of expected long-term behaviors of different systems. It is interesting because it can describe the state of a system as it seems to mostly “settle down.” For many systems it does not take unreasonably long for the state of the system to be near its limit. Put another way, if \( x \to \infty \implies f(x) \to L \), then for large enough \( x \) we should have \( f(x) \approx L \). Thus the limit point \( L \) is interesting even though we cannot, in fact, “travel to infinity” in \( x \) (or whatever we call the input variable) to experience the limit, but may be able to experience the state of the system where \( x \) is large enough that \( f(x) \approx L \) satisfactorily.

**Example 3.8.9** In Section 4.3 we will consider electrical circuits which contain a resistor and an inductor in series, as seen below right.

With a circuit having voltage \( V \), a resistor with resistance \( R \), and an inductor with inductance \( L \), and a switch which is first “closed” at \( t = 0 \), the current flowing through the circuit will be given by the following, for \( t \geq 0 \):

\[
I(t) = \frac{V}{R} \left( 1 - e^{-tR/L} \right).
\]

(a) What is the current at time \( t = 0 \)?

(b) What are the current values at times \( t = \frac{L}{R}, \frac{2L}{R}, \frac{3L}{R} \)?

(c) As \( t \to \infty \), what value does \( I \) approach?

**Solution:**

(a) The current at \( t = 0 \) is \( I(0) = \frac{V}{R} \left( 1 - e^0 \right) = \frac{V}{R}(1 - 1) = 0 \).

(b) For the other times we get

\[
I \left( \frac{L}{R} \right) = \frac{V}{R} \left( 1 - e^{-\left(\frac{L}{R}\right) R/L} \right) = \frac{V}{R} \left( 1 - e^{-1} \right) \approx 0.63 \left( \frac{V}{R} \right),
\]

\[
I \left( \frac{2L}{R} \right) = \frac{V}{R} \left( 1 - e^{-2 \left(\frac{L}{R}\right) R/L} \right) = \frac{V}{R} \left( 1 - e^{-2} \right) \approx 0.86 \left( \frac{V}{R} \right),
\]

\[
I \left( \frac{3L}{R} \right) = \frac{V}{R} \left( 1 - e^{-3 \left(\frac{L}{R}\right) R/L} \right) = \frac{V}{R} \left( 1 - e^{-3} \right) \approx 0.95 \left( \frac{V}{R} \right),
\]

(c) Here we compute \( \lim_{t \to \infty} I(t) \):

\[
\lim_{t \to \infty} I(t) = \lim_{t \to \infty} \frac{V}{R} \left( 1 - e^{-tR/L} \right) = \lim_{t \to \infty} \frac{V}{R} \left( 1 - e^{-tR/L} \right) \frac{\frac{V}{R}(1 - e^{-\infty})}{\frac{V}{R}(1 - 0)} = \frac{V}{R}.
\]
In Chapter 4 we will introduce Ohm’s Law, which can be written \( V = IR \). Note that as \( t \to \infty \) above we have \( I \to \frac{V}{R} \), i.e., \( I = V/R \) “in the limit,” which is equivalent to Ohm’s Law. An inductor will resist any sudden voltage change, in fact countering that change with a back voltage of its own, but in the presence of a steady voltage an inductor will behave like a conductor. When \( t = 0 \) and the switch is thrown, the rate of voltage change felt by the inductor is most sudden, and for that instant no current flows as the inductor completely counters the voltage source. However its capacity to resist \( (L) \) is not unlimited, and the voltage change it experiences (and its reactance as well) fades until the inductor behaves more and more like an conductor, so that nearly all (and in fact all, in the limit for the ideal case) of the resistance in the circuit comes from the resistor.

Sometimes computing the limit as the input approaches infinity is also useful just to indicate what could theoretically occur if the variable were allowed to get large enough. For instance, if some process’s output is logarithmic, with a base greater than 1, even though growth may be very slow it is theoretically possible for it to be as large as we like. Just that piece of information is at times valuable.

In the meantime, the limit computations in the exercises help us to gain further “number sense” and “function sense,” as we explore more aspects of the behaviors of functions so we can better analyze these things both in theory, and in the practice of analyzing real-world problems.

### Exercises

For problems 1–15, compute the limits where they exist (and if not, state so), showing all steps. You may wish to use Theorem 3.8.1 (page 256) or Theorem 3.8.2 (page 257) to anticipate an answer, but perform all the computations as in Examples 3.8.3 and 3.8.4, starting on page 257.

1. \( \lim_{x \to \infty} x^5 \)
2. \( \lim_{x \to -\infty} x^5 \)
3. \( \lim_{x \to -\infty} x^4 \)
4. \( \lim_{x \to -\infty} (x^4 - 5x^5) \).
5. \( \lim_{x \to -\infty} (x^4 - 5x^5) \).
6. \( \lim_{x \to -\infty} (x^4 - 5x^6) \).
7. \( \lim_{x \to -\infty} (x^4 - 5x^6) \).
8. \( \lim_{x \to -\infty} \frac{1}{x^3} \).
9. \( \lim_{x \to -\infty} \frac{1}{x^3} \).
10. \( \lim_{x \to -\infty} \frac{x^2}{x^4 + 1} \).
11. \( \lim_{x \to -\infty} \frac{1 - 2x^2}{x^3 + x^2 + x + 9} \).
12. \( \lim_{x \to -\infty} \frac{3x^2 + 5x - 11}{2x^2 - 27x + 100} \).
13. \( \lim_{x \to -\infty} \frac{3x^2 + 5x - 11}{2x^2 - 27x + 100} \).
14. \( \lim_{x \to -\infty} \frac{x^2 + 3x - 7}{x + 5} \).
15. \( \lim_{x \to -\infty} \frac{x^2 + 3x - 7}{x + 5} \).

16. Compute the limits, showing logical steps to justify answers.
   (a) \( \lim_{x \to -\infty} (x + \cos x) \).
   (b) \( \lim_{x \to -\infty} (x - \cos x) \).
   (c) \( \lim_{x \to -\infty} (x + \cos x) \).
   (d) \( \lim_{x \to -\infty} (x^2 + \cos x) \)

17. Compute the limits, showing logical steps to justify answers.
(a) \( \lim_{x \to \infty} \frac{\sin x}{x^2} \).

(b) \( \lim_{x \to \infty} \frac{x + \sin x}{x^2 + 1} \).

(c) \( \lim_{x \to \infty} \frac{x^2}{x - \sin x} \).

(d) \( \lim_{x \to \infty} \frac{x^2 + 2x + 1 - \sin x}{3x^2 + 2x - 1} \).

(e) \( \lim_{x \to \infty} x \sin x \).

18. Compute the limit
   \( \lim_{x \to \infty} \left( \sqrt{x^2 + 3x + 9} - x \right) \).

19. Compute the limit
   \( \lim_{x \to \infty} \left( \sqrt{x^2 - 5x + 9} - x \right) \).
   It is actually simpler than the previous limit (one line!).

20. Compute the following.
    (a) \( \lim_{x \to \infty} \frac{\sin x}{x} \).
    (b) \( \lim_{x \to \infty} \cos \frac{1}{x} \).
    (c) \( \lim_{x \to \infty} \left[ \frac{x^2 + 2x + 9}{6x^2 - 11x + 45} \right] \).

21. Write definitions for the following (see (3.64), page 255). It may help to graph situations where these are true.
    (a) \( \lim_{x \to \infty} f(x) = \infty \).
    (b) \( \lim_{x \to -\infty} f(x) = \infty \).
    (c) \( \lim_{x \to \infty} f(x) = -\infty \).
    (d) \( \lim_{x \to -\infty} f(x) = -\infty \).

22. Prove that \( \lim_{x \to \infty} \sqrt{x} = \infty \) using the the definition found in (3.64), page 255. (For a hint, see the example in the paragraph immediately following that definition.)

23. Prove Theorem 3.8.1, page 256.


25. A 10Ω resistor and a variable resistor \( R \) are placed in parallel in a circuit. The equivalent resistance \( R_p \) is related by the equation
   \[ \frac{1}{R_p} = \frac{1}{10} + \frac{1}{R}, \]
   where all resistances are in ohms (Ω):

   \[ R_p \quad \frac{10\Omega}{\text{}} \quad R \]

   (a) Solve for \( R_p \).
   (b) Compute \( \lim_{R \to \infty} R_p \), the limiting value of \( R_p \) as \( R \to \infty \).

26. An employee can produce approximately
   \( N(x) = \frac{50x + 7}{2x + 5} \) items per day on the production line after \( x \) days on the job. In Chapter 4 we will be able to show that this is an increasing function. Find the maximum number of items that can be produced per day by computing \( \lim_{x \to \infty} N(x) \).

27. A series circuit consisting of a voltage source, a resistor, a capacitor (initially discharged) and a switch is diagrammed below.

   \[ V^+ \quad R \quad C \]

   If the switch is first closed (“on”) at \( t = 0 \) the charge on the capacitor is given by
   \[ q(t) = CV \left( 1 - e^{-t/RC} \right). \]
   (a) What is the charge at \( t = 0 \)?
(b) Expressed as a percentage of $CV$, what is the charge at $t = RC$, $t = 2RC$, $t = 3RC$?
(c) What is the trend in the charge as $t \to \infty$?

29. Compute the following:

(a) $\lim_{x \to 0^+} \ln \csc x$
(b) $\lim_{x \to 0^+} \frac{1}{\ln x}$
(c) $\lim_{x \to 0^-} \ln \left( \frac{1}{5 - x} \right)$
(d) $\lim_{x \to \infty} \ln \left( \ln x \right)$
(e) $\lim_{x \to \infty} \ln \left( \ln \left( \ln x \right) \right)$
(f) $\lim_{x \to \infty} x \ln \left( \frac{1}{x} \right)$
3.9 Further Limit Theorems and Trigonometric Limits

In this section we wrap up our discussion of fundamental methods for computing limits of functions. We also look at some very general theorems on limits, ranging from very intuitive to rather sophisticated. Into the mix we introduce and prove an interesting trigonometric limit, (3.81) which will be the basis for another fundamental trigonometric limit, (3.86) and many consequent trigonometric limits we could not have proved with previous methods. Finally we will develop our most sophisticated (so far) limit method, which is substitution.\footnote{Actually, we will add other methods after we develop derivatives. Still, what we finish in this section are the \textit{foundational} methods for limits of functions. Even with the methods we will introduce after derivatives, we can not avoid the requirements of the methods of this section or this chapter. Indeed, the later methods are quite powerful where applicable, but have very limited scopes and therefore cannot replace what we develop here.}

3.9.1 Simple Limit Theorems

Many of the limit computations that we have already performed in this textbook were based upon what we knew from continuity arguments. We took that approach because it is more intuitive than the usual treatment found in most calculus textbooks. (See Footnote 22, page 211.) In fact, the more common treatment is to instead rely upon theorems about limits independent of possible underlying continuity (or continuity of replacement functions). These basic facts we combine below into one theorem, the different parts of which can be proved in ways very similar to the proofs of corresponding continuity theorems (see especially Section 3.2).

**Theorem 3.9.1** Suppose that, as $x$ approaches some value, we have $f(x) \to L \in \mathbb{R}$ and $g(x) \to M \in \mathbb{R}$. Then

(i) $f(x) + g(x) \to L + M$,

(ii) $f(x) - g(x) \to L - M$,

(iii) $Cf(x) \to CL$,

(iv) $f(x)g(x) \to LM$,

(v) and if $M \neq 0$, then $f(x)/g(x) \to L/M$.

(vi) $\lim_{x \to a} C = C$, where $C$ is any fixed constant.

The theorem above assumes $L, M \in \mathbb{R}$, so $L$ and $M$ exist and are finite. The theorem can be extended to include several (but not all!) cases where $L$ or $M$ do not exist or are infinite. Furthermore the theorem extends in the obvious ways to numerous forms, so for instance if $f(x) \to \infty$ and $g(x) \to \infty$, then $f(x) + g(x) \to +\infty$. We still have to be careful: for such $f$ and $g$ we have $f(x) - g(x)$ is of form $\infty - \infty$, which is indeterminate, as demonstrated in Example 3.8.1, page 255.

Textbooks include part (vi) for theoretical reasons we will consider momentarily, but also give it special emphasis because it is so simple it sometimes confuses. For an example illustrating how to interpret (vi) properly, consider the statement that $\lim_{x \to -5} 10 = 10$. What this means is that a function $h(x)$, where $h(x) = 10$ for all $x \in \mathbb{R}$, will give $\lim_{x \to a} h(x) = 10$ regardless of $a$ (finite or infinite). This $h(x)$ is a “constant” function, which always returns the value 10 regardless of the input; its graph is the horizontal line $y = 10$. When graphed it is clear that...
3.9. FURTHER LIMIT THEOREMS AND TRIGONOMETRIC LIMITS

\[ h(x) \longrightarrow 10 \text{ as } x \to a \text{ (regardless of } a, \text{ chosen in advance).} \] Rather than explicitly defining such \( h \), it is customary to simply write, for example, \( \lim_{x \to a} 10 = 10 \).

The usual method—found in most of today’s calculus textbooks—for obtaining a preliminary theory of finite limits is in fact based upon Theorem 3.9.1 above using the following scheme:

1. First prove (with \( \varepsilon-\delta \) or by graphical demonstration) that \( \lim_{x \to a} x = a \).

2. Then use (i), (iii), (iv) and (vi) from the theorem above repeatedly to show that polynomials \( p(x) \) have the property that \( \lim_{x \to a} p(x) = p(a) \), using the limit version of the argument given in the proof of Theorem 3.2.4, page 183 but without reference to continuity (defined later in that approach).

3. From there (v) gives that rational functions \( p(x)/q(x) \), i.e., where \( p \) and \( q \) are polynomials are continuous where defined, that being where \( q(x) \neq 0 \).

4. Then one looks at rational cases with \( 0/0 \) form, mentioning Theorem 3.4.3, page 212 on replacing functions with other functions that agree near the limit point and are continuous there.

5. One then progresses through the more sophisticated cases (radicals, infinite limits, limits at infinity, Sandwich Theorem, etc.).

6. Define continuity at \( x = a \) by the criterion that \( \lim_{x \to a} f(x) = f(a) \). (Our definition of continuity is logically equivalent, according to Theorem 3.4.2, page 211.)

This is mentioned here so the reader will be aware of this common alternative treatment. Though we did not follow that logical scheme (instead opting for the advanced calculus and real analysis style of continuity before limits), we will occasionally have use for Theorem 3.9.1 above in the rest of the text. Now we will look at an (admittedly) abstract example.

**Example 3.9.1** Suppose that \( \lim_{x \to 3} f(x) = 5 \) and \( \lim_{x \to 3} g(x) = 7 \).

Then

\[
\begin{align*}
\lim_{x \to 3} [f(x) + g(x)] &= \frac{5 + 7}{5 + 7} = 12, \\
\lim_{x \to 3} [f(x) - g(x)] &= \frac{5 - 7}{5 - 7} = -2, \\
\lim_{x \to 3} [4f(x)] &= \frac{4 \cdot 5}{4 \cdot 5} = 20,
\end{align*}
\]

\[
\lim_{x \to 3} [f(x)g(x)] = \frac{5 \cdot 7}{5 \cdot 7} = 35,
\]

\[
\lim_{x \to 3} \frac{f(x)}{g(x)} = \frac{5/7}{5/7} = \frac{5}{7},
\]

\[
\lim_{x \to 3} [191] = 191.
\]

The last equation above, of course, has nothing to do with the functions \( f \) or \( g \). For a more concrete example, consider the following:

**Example 3.9.2** Compute \( \lim_{x \to \infty} \left[ 17 + \frac{x}{x^2 + 1} + \frac{3x^2 + 5x + 9}{2 - x^2} + \frac{2x - 3}{6x - 5} \right] \).

**Solution:** There are several subexpressions here whose limits exist. In this particular example we are lucky that we can partition the whole expression into such well-behaved subexpressions:
Chapter 3. Continuity and Limits of Functions

Example 3.9.3 Compute \( \lim_{x \to \infty} \left[ \frac{x - 2}{x^3 - 9x + 5} \cdot \frac{x^4 + 10x^3 - 9x^2 + 11x + 5}{2x^2 + x - 9} \right] \).

Solution: As it stands, the form of this limit is \( 0 \cdot \infty \) (by Theorem 3.8.2, page 257 and the thinking surrounding that result) which is indeterminate. One brute-force method of computing this limit is to combine the two fractions into one, but this requires some lengthy multiplication calculations. Instead we offer the two methods below, which work well because it is a limit at infinity.

1. We can factor the largest power of \( x \) which appears in each term and cancel:

\[
\lim_{x \to \infty} \left[ \frac{x - 2}{x^3 - 9x + 5} \cdot \frac{x^4 + 10x^3 - 9x^2 + 11x + 5}{2x^2 + x - 9} \right] = \lim_{x \to \infty} \left[ \frac{x (1 - \frac{2}{x})}{x^3 (1 - \frac{9}{x^2} + \frac{5}{x})} \cdot \frac{x^4 (1 + \frac{10}{x} - \frac{9}{x^2} + \frac{11}{x^3} + \frac{5}{x^4})}{x^2 (1 + \frac{1}{x} - \frac{2}{x^2})} \right] = \lim_{x \to \infty} \left[ \frac{x^5 (1 - \frac{2}{x}) (1 + \frac{10}{x} - \frac{9}{x^2} + \frac{11}{x^3} + \frac{5}{x^4})}{x^5 (1 - \frac{9}{x^2} + \frac{5}{x^4}) (2 + \frac{1}{x} - \frac{2}{x^2})} \right] = \frac{1 \cdot 1}{1 \cdot 2} = \frac{1}{2}.
\]

2. Another method is to observe what the leading terms of the numerator and denominator polynomials would be if we were to multiply and simplify them. For a limit at infinity, as we know we need only look at the highest-order terms in the numerator and denominator:

\[
\lim_{x \to \infty} \left[ \frac{x - 2}{x^3 - 9x + 5} \cdot \frac{x^4 + 10x^3 - 9x^2 + 11x + 5}{2x^2 + x - 9} \right] = \lim_{x \to \infty} \frac{x^5 + \cdots - 10}{2x^5 + \cdots - 45} = \frac{1}{2}.
\]

The highest- and lowest-order terms are the easiest to compute for a polynomial product (whereas—recall—the intermediate-order terms may be complicated sums). In fact it is only the highest-order terms which are relevant, again, because the limit is at infinity.

Next we consider what to do if our partition of the limit’s function yields subexpressions whose limits do not necessarily exist. It is true that knowing that one of the component limits
3.9. FURTHER LIMIT THEOREMS AND TRIGONOMETRIC LIMITS

269

does not exist can sometimes allow us to conclude that the entire limit does not. However this
is not always the case. In the next example we give an argument where the nonexistence of
a component limit can, in that context, imply nonexistence of the full limit. We follow that
example with one in which nonexistence of a component limit does not wreck the full limit.
Both types should become intuitive, but the perennial lesson that we must be careful not to be
too cavalier with our limit arguments should be apparent below.

Example 3.9.4 Suppose that \( \lim_{x \to a} f(x) = 5 \) and \( \lim_{x \to a} g(x) \) D.N.E. Then

\[
\lim_{x \to a} [f(x) + g(x)] \equiv D.N.E. \quad \text{if} \quad \lim_{x \to a} f(x) = 5 \text{ and } \lim_{x \to a} g(x) \text{ D.N.E.}
\]

To see this, we can argue that since \( g(x) \) is not approaching a well-defined limit value for
whatever reason (perhaps being undefined near \( x = a \), or oscillating, or having different left
and right limits), then adding \( f(x) \), which is approaching a number, will not compensate for the
behavior of \( g(x) \), and the final limit can not exist. A more rigorous argument is given next.

**Proof:** Suppose again \( \lim_{x \to a} f(x) = 5 \) and \( \lim_{x \to a} g(x) \) D.N.E. We will prove by
contradiction that \( \lim_{x \to a} [f(x) + g(x)] \) can not exist, for suppose that it does. Then
according to our general Theorem 3.9.1, page 266 since \( \lim_{x \to a} [f(x) + g(x)] \) exists
(by our assumption to be contradicted), we have

\[
f(x) + g(x) \to L \implies -f(x) + (f(x) + g(x)) \to -5 + L \quad \text{(by Theorem 3.9.1(i))}
\]

\[
\implies (-f(x) + f(x)) + g(x) \to -5 + L \quad \text{(algebra)}
\]

\[
\implies (0 + g(x)) \to -5 + L \quad \text{(algebra near limit point)}
\]

\[
\iff g(x) \to -5 + L \text{ exists (algebra).}
\]

But that contradicts our original information that \( \lim_{x \to a} g(x) \) must not exist. Thus
we have to conclude that our assumption \( \lim_{x \to a} [f(x) + g(x)] \) exists—which leads to
a contradiction—must be false, and so \( \lim_{x \to a} [f(x) + g(x)] \) does not exist, q.e.d.

In the above example, the fact that the limit of \( f(x) \) was finite meant that it could not—
through simple addition—compensate for “bad behavior” of \( g(x) \) in the limit. It is possible,
however, for \( f \) to still compensate, in the sense that \( f(x) \) can also simply dominate \( g(x) \) in
the expression \( f(x) + g(x) \) if \( g(x) \) is defined but bounded and \( f(x) \) blows up. (The reader may wish
to revisit (3.69)–(3.72), page 259 as we work the next example.)

Example 3.9.5 Suppose that \( f(x) \to \infty \) and \( |g(x)| \leq M \in \mathbb{R} \) for some \( M > 0 \) (but real
and therefore finite) as \( x \to a \). Then we employ a Sandwich Theorem argument, with \(-M + f(x) \leq
f(x) + g(x) \leq M + f(x) \), with the first and third terms approaching \( \infty \) (though the first of the
three is enough) carrying \( f(x) + g(x) \) to an infinite limit as well. We could write

\[
\lim_{x \to a} [f(x) + g(x)] \equiv \infty + B \quad \infty.
\]

For a specific example, consider: \( \lim_{x \to \infty} (x + \sin x) \equiv \infty + B \quad \infty. \)

---

49 At \( x = a \) we may have that \( f(x) \) is undefined, so there it may be incorrect to say \(-f(x) + f(x) = 0 \). However,
near the limit point we know \( f(x) \) is defined, from the assumption \( f(x) \to 5 \) as \( x \to a \). Implied in our definition
of limits ((3.22), page 210) is that \( f(x) \) exists for \( 0 < |x - a| < \delta \), for some \( \delta > 0 \).

50 Note how we used \( P \equiv (\sim P) \implies \sim P. \)
Actually we worked exactly this example during the discussion of the form \( B + \infty \), introduced on page 259.

The lesson to be gleaned from the above examples is that if one part of the function “inside the limit” has nonexistent limit (for the prescribed approach in the independent variable), sometimes we can conclude the same about the whole limit and sometimes we can not. We usually need to dig deeper into the behavior of the other parts of the function, and take into account how their influences combine. Sometimes the form is enough to determine the actual limit (or its nonexistence). With experience, the various cases become intuitive. (Recall also Example 3.7.1, page 241 and its associated Figure 3.22.)

### 3.9.2 A Trigonometric Limit

An interesting limit, which is surprisingly useful for future results, is the following:

**Theorem 3.9.2** With \( \theta \) given in radians, we have the following limit:

\[
\lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1.
\] (3.81)

**Proof:**

The proof relies upon the Sandwich Theorem and a geometric observation which is given in Figure 3.28. In that figure, \( \theta \) is the radian measure of the angle which terminates in the first quadrant. The observation involves three areas defined by this angle \( \theta \): a right triangle, contained within a circular wedge, which is in turn contained in another right triangle. The smaller triangle has “base” \( \cos \theta \) and “height” \( \sin \theta \), and thus has area \( \frac{1}{2} \cos \theta \sin \theta \). For the circular wedge, recall that a wedge with radius \( r \) and radian-measure angle \( \theta \) has area \( \frac{1}{2} r^2 \theta \). The other triangle has base 1 and height \( \tan \theta \). To see this, note that it is similar to the smaller triangle, and so we have the
3.9. FURTHER LIMIT THEOREMS AND TRIGONOMETRIC LIMITS

proportion of sides: \( \sin \theta / \cos \theta = h/1 \), where \( h \) is the height of the larger triangle. It follows that \( h = \tan \theta \). Thus the larger triangle has area \( \frac{1}{2} \cdot 1 \cdot \tan \theta \). Now for \( \theta \in [0, \pi/2) \), we have

\[
\frac{1}{2} \cos \theta \sin \theta \leq h \leq \frac{1}{2} \tan \theta.
\]

As \( \theta \to 0^+ \), we have \( \sin \theta > 0 \) so we can not only multiply by 2, but also then divide by \( \sin \theta \), giving us

\[
\cos \theta \leq \frac{\theta}{\sin \theta} \leq \frac{1}{\cos \theta}.
\]  

(3.82)

This gives us a Sandwich Theorem type of argument as \( \theta \to 0^+ \):

\[
\begin{align*}
\text{As } \theta \to 0^+ : & \quad \cos \theta \leq \frac{\theta}{\sin \theta} \leq \frac{1}{\cos \theta} \\
& \quad \downarrow \quad \downarrow \\
& \quad 1 \quad 1
\end{align*}
\]

giving us \( \frac{\theta}{\sin \theta} \to 1 \) as \( \theta \to 0^+ \), i.e.,

\[
\lim_{\theta \to 0^+} \frac{\theta}{\sin \theta} = 1.
\]  

(3.83)

Next we use dispatch with the left-side limit. In fact the same inequality (3.82) holds as \( \theta \to 0^- \), because all three expressions are the same if we replace \( \theta \) with \( -\theta \). This follows because all three functions are “even,” i.e., \( \cos(-\theta) = \cos \theta \), \( (-\theta)/\sin(-\theta) = \theta/\sin \theta \), and \( 1/\cos(-\theta) = 1/\cos \theta \), and \( \theta \to 0^+ \iff (-\theta) \to 0^- \). With this one can perform the above computations with \( (-\theta) \to 0^+ \), and thus \( \theta \to 0^- \).

Perhaps a less convoluted approach is to more explicitly borrow the substitution method from upcoming Subsection 3.9.3. Here we let \( \phi = -\theta \) so that \( \theta \to 0^- \iff \phi \to 0^+ \). Thus

\[
\lim_{\phi \to 0^+} \frac{-\phi}{\sin(-\phi)} = \lim_{\phi \to 0^+} \frac{-\phi}{-\sin \phi} = \lim_{\phi \to 0^+} \frac{\phi}{\sin \phi} = 1,
\]  

(3.84)

the last limit being, of course, (3.83) with the variable renamed. Putting (3.83) together with (3.84), of course, gives us

\[
\lim_{\theta \to 0} \frac{\theta}{\sin \theta} = 1.
\]  

(3.85)

Finally, we can then get our result based upon the above limit, looking at the reciprocal function:

\[
\lim_{\theta \to 0} \frac{\sin \theta}{\theta} = \lim_{\theta \to 0} \left( \frac{1}{\frac{\theta}{\sin \theta}} \right)^{1/1} = 1, \quad \text{q.e.d.}
\]

Notice that the limit we proved, (3.81), says something about how fast \( \sin \theta \) approaches zero as \( \theta \) approaches zero: that \( \sin \theta \) and \( \theta \) approach zero at approximately the same rate. In fact, many physics problems use the approximation \( \sin \theta \approx \theta \) (following from \( (\sin \theta)/\theta \approx 1 \)) for \( |\theta| \) small and in radians. This approximation is graphed in Figure 3.29, page 272 but using \( x \) instead
of \( \theta \) as the independent (domain) variable. We will see how this very important approximation, and the underlying limit, arise from other calculus techniques in later chapters.\footnote{While (3.85) is the immediate result of our analysis, and is interesting in its own right, most texts (as here) go ahead and present the reciprocal limit (3.81). The reason is that \( \lim_{\theta \to 0} (\sin \theta / \theta) \) gives a (perhaps) more intuitive comparison of the behavior of \( \sin \theta \) versus that of the independent variable \( \theta \). We have already had many limits of ratios of functions \( f(x) / y(x) \) which compare the two functions’ behaviors in the sense of limits. A very useful way to analyze a function is to compare it to its input variable by ratios \( f(x) / x \), \( f(x) / (x - a) \), or \( (f(x) - f(a)) / (x - a) \), for examples. Though this can often be accomplished instead by looking the reciprocals of these, it is less intuitive to do so.}

Many other interesting limits follow from this limit (3.81), as we can see in the following example. Note that the second limit computation below utilizes the trigonometric identities

\[
(1 - \cos x)(1 + \cos x) = 1 - \cos^2 x = \sin^2 x.
\]

**Example 3.9.6** Consider the following limits, which require both our basic trigonometric limit (3.81) and our general limit theorem, Theorem 3.9.1 with which we began the section, on page 266.

\[
\begin{align*}
\lim_{x \to 0} \frac{\tan x}{x} &= \lim_{x \to 0} \frac{\sin x}{x \cos x} = \lim_{x \to 0} \left[ \frac{\sin x}{x} \cdot \frac{1}{\cos x} \right]^{1/2} = 1. \\
\lim_{x \to 0} \frac{1 - \cos x}{x} &= \lim_{x \to 0} \left[ \frac{1 - \cos x}{x} \cdot \frac{1 + \cos x}{x + \cos x} \right] = \lim_{x \to 0} \frac{1 - \cos^2 x}{x(1 + \cos x)} = \lim_{x \to 0} \frac{\sin^2 x}{x(1 + \cos x)} \\
&= \lim_{x \to 0} \left[ \frac{\sin x}{x} \cdot \frac{\sin x}{1 + \cos x} \right]^{1/2} = 1 \cdot 0 \cdot 2 = 0.
\end{align*}
\]

The first limit is also used in physics in the form \( \tan \theta \approx \theta \) for \( |\theta| \) small. The second limit occurs enough to warrant being set aside as its own theorem, though as we see above it is easily derived from the more basic \( \lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1 \).

**Theorem 3.9.3** The following limit (proved above) holds:

\[
\lim_{\theta \to 0} \frac{1 - \cos \theta}{\theta} = 0.
\] (3.86)
3.9. FURTHER LIMIT THEOREMS AND TRIGONOMETRIC LIMITS

Figure 3.30: Partial graphs of \( y = x \) and \( y = 1 - \cos x \), showing how \( 1 - \cos x \) approaches zero faster than \( x \) as \( x \to 0 \)—so much so that \( (1 - \cos x)/x \to 0 \) (see (3.86)). The graph of \( y = \frac{1}{2}x^2 \) is also given (thinner curve), illustrating how \( (1 - \cos x)/x^2 \approx \frac{1}{2} \), i.e., \( \cos x \approx 1 - \frac{1}{2}x^2 \) for \( x \) small, as derived in Example 3.9.7.

In other words, \( 1 - \cos \theta \) shrinks to zero faster than \( \theta \) does, as \( \theta \to 0 \). This is reasonable when we see the graphs of \( y = 1 - \cos x \) and \( y = x \), given (as darker curves) in Figure 3.30.\(^{52}\)

For the rest of this section, we will assume (3.81) and (3.86) and compute other trigonometric limits based upon these, the work from previous sections, and Theorem 3.9.1. As with all limits, it is important to be careful and not jump to incorrect conclusions; our basic trigonometric limits (3.81) and (3.86)—which the reader should commit to memory—given here again as

\[
\lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1, \quad \lim_{\theta \to 0} \frac{1 - \cos \theta}{\theta} = 0,
\]

are very specific in their scopes. Moreover, with trigonometric limits it is sometimes still necessary to exploit the algebraic identities among those functions. Consider the following (perhaps surprising) trigonometric limit calculation:

**Example 3.9.7** Compute \( \lim_{x \to 0} \frac{1 - \cos x}{x^2} \).

**Solution:** A first, perhaps more obvious attempt is quickly seen to be a dead-end:

\[
\lim_{x \to 0} \frac{1 - \cos x}{x^2} = \lim_{x \to 0} \left[ \frac{1 - \cos x}{x} \cdot \frac{1}{x} \right] = 0 \cdot (\pm \infty)
\]

of the indeterminate form \( 0 \cdot (\pm \infty) \). As happens so frequently with limits, we look for some other way of rewriting the function. The usual method of computing this limit is to again exploit the fact that

\[
(1 - \cos x)(1 + \cos x) = 1 - \cos^2 x = \sin^2 x.
\]

By multiplying the function inside the limit by \( (1 + \cos x)/(1 + \cos x) \), we can compute the limit as follows:

\[
\lim_{x \to 0} \frac{1 - \cos x}{x^2} = \lim_{x \to 0} \left[ \frac{1 - \cos x}{x^2} \cdot \frac{1 + \cos x}{1 + \cos x} \right] = \lim_{x \to 0} \frac{1 - \cos^2 x}{x^2(1 + \cos x)} = \lim_{x \to 0} \frac{\sin^2 x}{x^2(1 + \cos x)}
\]

\[
= \lim_{x \to 0} \left[ \frac{\sin x}{x} \cdot \frac{\sin x}{x} \cdot \frac{1}{1 + \cos x} \right] = 1 \cdot 1 \cdot \frac{1}{1 + 1} = \frac{1}{2}.
\]

\(^{52}\)One could also say the \( \cos \theta \to 1 \) very rapidly as \( \theta \to 0 \), but of course neither description is as precise as (3.86), i.e., that \( (1 - \cos \theta)/\theta \to 0 \) as \( \theta \to 0 \).
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

In the last step of the above example, we used Theorem 3.9.1 (the limit of a product being the
product of the limits, when they exist), and the fact that the factor \(1/(1 + \cos x)\) is continuous
at \(x = 0\).

One could argue using this limit that, for \(|x| \text{ small, } 1 - \cos x \approx \frac{1}{2} x^2\), which is also illustrated
in Figure 3.30, page 273. This fact we will derive later in the form \(\cos x \approx 1 - \frac{1}{2} x^2\), which is
sometimes used in applications. It is much more accurate than the earlier approximation that
\(\cos x \approx 1 \text{ for } x \text{ small (though they coincide at } x = 0)\).

We will compute many more trigonometric limits in this section, but first we need a new
limit technique which we introduce next.

3.9.3 Limits by Substitution

To bring our analytical methods of computing limits to the next level, we now develop some
substitution techniques. These techniques require delicacy, but with care they are quite powerful
and sophisticated. To motivate the discussion, we first give two examples below:

Example 3.9.8 Compute \(\lim_{x \to 0} \frac{\sin 5x}{x}\).

Solution: The usual method is to multiply by \(\frac{5}{5}\) as below.

\[
\lim_{x \to 0} \frac{\sin 5x}{x} = \lim_{x \to 0} \frac{5 \sin 5x}{5x} = \lim_{x \to 0} \left[ 5 \cdot \frac{\sin 5x}{5x} \right] = 5 \cdot 1 = 5.
\]

Notice that the original limit was of the form \((\sin 0)/0\) (i.e., \(0/0\)) which is indeterminate. It
is important in our basic trigonometric limit (3.81), i.e., \(x \to 0 \implies (\sin x)/x \to 1\), that the
two “zeros” are terms approaching zero at the same rate (though even that is not always quite
enough, as we will eventually see). By multiplying the fraction by \(5/5\), we were able to get a form
“\(5 \cdot [(\sin 0)/0]\),” but where the rates of the “zeros” were exactly the same. Most presentations of
the computation of the above limit are exactly as given above, but it is to be understood that
we are transforming this by way of a substitution. One might instead write

\[
\lim_{x \to 0} \frac{\sin 5x}{x} = \lim_{x \to 0} \left[ 5 \cdot \frac{\sin 5x}{5x} \right] = \lim_{\theta \to 0} \left[ 5 \cdot \frac{\sin \theta}{\theta} \right] = 5 \cdot 1 = 5,
\]

where \(\theta = 5x\), and \(x \to 0 \iff \theta \to 0\). The nature of the mechanism whereby \(\theta \to 0 \iff x \to 0\)
is crucial, but we will explain this in due course. The next example displays a slightly more
sophisticated argument.

Example 3.9.9 Compute \(\lim_{x \to 0} \frac{\cos x^2 - 1}{x^4}\). (Compare to Example 3.9.7, page 273.)

Solution: Here we will make a substitution \(\theta = x^2\), so that \(x \to 0 \implies \theta \to 0^+\). Then
we can write

\[
\lim_{x \to 0} \frac{\cos x^2 - 1}{x^4} = \lim_{\theta \to 0^+} \frac{\cos \theta - 1}{\theta^2} = \lim_{\theta \to 0^+} \left[ \frac{\cos \theta - 1}{\theta^2} \cdot \frac{\cos \theta + 1}{\cos \theta + 1} \right] = \lim_{\theta \to 0^+} \frac{\cos^2 \theta - 1}{\theta^2 (\cos \theta + 1)}
\]

\[
= \lim_{\theta \to 0^+} \frac{-\sin^2 \theta}{\theta^2 (\cos \theta + 1)} = \lim_{\theta \to 0^+} \left[ \frac{-\sin \theta \cdot \sin \theta}{\theta \cdot \theta} \cdot \frac{1}{\cos \theta + 1} \right] = -1 \cdot 1 \cdot \frac{1}{1 + 1} = -\frac{1}{2}.
\]
3.9. FURTHER LIMIT THEOREMS AND TRIGONOMETRIC LIMITS

Notice that in the above example we used the fact that full limits imply one-sided limits. Specifically, above we used that
\[
\lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1 \implies \lim_{\theta \to 0^+} \frac{\sin \theta}{\theta} = 1.
\]

Now we consider circumstances where we can make a substitution to compute a limit. First we take another look at what it means for \( x \to a \). (Subsection 3.7.2, page 242 began this discussion.) When we write \( x \to a \) under “\( \lim \),” we understand this to mean that \( x \) gets arbitrarily close to but not equal to arbitrarily close to a continuum of values, so that \( x \) does not “skip over” any values as we approach \( a \) either, as we see from the definition of \( \lim_{x \to a} f(x) \), as for example in the case this limit is a finite number \( L \):

\[
(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)[0 < |x - a| < \delta \implies |f(x) - L| < \varepsilon].
\]

In order to substitute algebraically to produce another limit, say \( \lim_{u \to \beta} F(u) \) and claim it is the same as \( \lim_{x \to a} f(x) \), we would like to be sure that \( x \to a \implies u \to \beta \), and that the latter “approach” is still somehow a proper approach. Usually what is obvious is only that \( x \to a \implies u \to \beta \) in the sense that \( \lim_{x \to a} u = \beta \), which allows for the approach of \( u \) to \( \beta \) to be quite sloppy. It may be that the approach \( u \) takes to \( \beta \) is one-sided, or has gaps as \( x \to a \), or actually achieves \( u = \beta \) while we require \( x \neq a \). Indeed when we developed the convention that \( \lim_{x \to a} u = \beta \) would also be written \( x \to a \implies u \to \beta \), we interpreted the approach of \( u \) to \( \beta \) quite loosely, while that of \( x \) to \( a \) was more strict. To force \( u \to \beta \) to occur in the same way that \( x \to a \) requires much more structure in the relationship between \( x \) and \( u \) than we usually wish to have to accommodate. One notable example where this does happen is when \( u = mx + b \), where \( m \neq 0 \), giving that \( x \to a \iff u \to (ma + b) \), and both approaches are proper in every way. Another example is \( x \to 0 \implies x^2 \to 0^+ \) (note the absence of the converse \( \iff \)). Other cases can be quite complicated.

What we can do to produce a useful theorem is to not make a strong statement about the relationship between \( x \) and \( u \) but rather qualify the result in a different way (which also makes a proof easier to formulate, were we to include it), still yielding a useful result. What we settle on here is the following, the proof of which is similar to that of Theorem 3.7.4, page 244, and is left as an exercise for the interested reader.

**Theorem 3.9.4 (Limit Substitution Theorem)** Suppose that the variables \( x \) and \( u \) are related in such a way that

(a) \( \lim_{x \to a} u = \beta \),

(b) \( (\exists \delta > 0)(\forall x)[0 < |x - a| < \delta \implies (u \neq \beta)] \),

(c) \( (\exists \delta > 0)(\forall x)[0 < |x - a| < \delta \implies (f(x) = F(u))] \).

Then
\[
\lim_{x \to a} f(x) = \lim_{u \to \beta} F(u) \quad \text{if this second limit exists.} \quad (3.87)
\]

There are many theorems we will come across where we have an equality like (3.87) which is qualified by the criterion that the second quantity exists, or otherwise makes sense in the context. To remove that criterion would require a much more complicated set of conditions than (a)–(c) above. That the second limit exists is key, but also that \( u \to \beta \) to be an “approach” of a similar
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

kind in the sense that not only should \( u \to \beta \), but \( u \neq \beta \) as \( x \to a \) according to (b).\(^{53}\) In fact, the validity of (a) and (c) usually are clear from the actual algebraic substitution; it is (b) that requires a bit more scrutiny but is also usually not difficult to see.

**Example 3.9.10** Compute \( \lim_{x \to \pi/2} \frac{\sin(x - \pi/2)}{x - \pi/2} \).

Here we make the substitution \( u = x - \pi/2 \), so that \( x \to \pi/2 = \Rightarrow u \to 0 \) in the sense of the hypotheses (a)–(c) of the theorem. Thus we can write (as long as the final limit exists!)

\[
\lim_{x \to \pi/2} \frac{\sin(x - \pi/2)}{x - \pi/2} = \lim_{u \to 0} \frac{\sin u}{u} = 1.
\]

In the earlier Examples 3.9.8 and 3.9.9 above we used \( \theta \) to play the role of \( u \) in the theorem, but that is mostly a matter of taste. We will usually use \( u \) because it is the traditional variable of substitution later in the calculus.\(^{54}\)

The theorem can be extended in obvious ways to cases such as

\[
x \to a = \Rightarrow u \to \beta^+ \quad x \to a^+ = \Rightarrow u \to \beta^+
\]

and so on.

It should be pointed out that if \( x \to a \) implies a “proper” one-sided approach, for instance \( u \to \beta^+ \), then we can perform a substitution based upon that, except then we would compute the one-sided limit \( \lim_{u \to \beta^+} F(u) \). This was the case in Example 3.9.9, page 274. Furthermore, a one-sided approach in \( x \) to \( a \) may yield a proper approach in a variable \( u \) of some kind. Infinite “approaches” can also arise from substitutions, or give rise to substitutions, as we see below.

**Example 3.9.11** Consider \( \lim_{x \to \pi/2^+} \frac{2\tan^2 x + 3\tan x + 7}{\tan^2 x - 6\tan x + 30} \).

Here we let \( u = \tan x \). Then \( x \to \pi/2^+ = \Rightarrow u \to -\infty \). (Because we will never have \( u = -\infty \), we do not need an analog to the criterion (b) in the Limit Substitution Theorem, Theorem 3.9.4, page 275.) Thus

\[
\lim_{x \to \pi/2^+} \frac{2\tan^2 x + 3\tan x + 7}{\tan^2 x - 6\tan x + 30} = \lim_{u \to -\infty} \frac{2u^2 + 3u + 7}{u^2 - 6u + 30} = \lim_{u \to -\infty} \frac{u^2 (2 + \frac{3}{u} + \frac{7}{u^2})}{u^2 (1 - \frac{6}{u} + \frac{30}{u^2})} = \frac{2 + 0 + 0}{1 - 0 + 0} = 2.
\]

In fact the above example can also be computed by multiplying numerator and denominator by \( \cos^2 x \). A similar computation is left to the exercises.

The next example illustrates that if one type of indeterminate form is not easily dealt with, an algebraic manipulation can likely give one which is more easily computed.

**Example 3.9.12** Compute \( \lim_{x \to \infty} x \sin \frac{1}{x} \).

**Solution:** Note that \( \frac{1}{x} \to 0^+ \) as \( x \to \infty \), so the form here is essentially \( \infty \cdot \sin 0 = \infty \cdot 0 \) (more precisely, \( \infty \cdot \sin 0^+ = \infty \cdot 0^+ \)), which is indeterminate (see Example 3.8.2, page 256).\(^{53}\)

\(^{53}\)Later in the subsection we will have an example to show why (b) is necessary. In that example, \( x \to a = \Rightarrow u \to \beta \), but \( u \) oscillates, passing through the value \( \beta \) infinitely many times as \( x \to a \). In that example the naive substitution is invalid: the new limit exists but the original does not, so obviously they are not equal.

\(^{54}\)Actually \( \theta \) is becoming increasingly common as a variable of substitution, and we will have occasion to use it as we did in our first substitution examples.
However we can rewrite the limit with a power of $x$ in the denominator, instead of having $x$ as a multiplicative factor. Then we will perform a substitution and use (3.81) for the final computation.

$$\lim_{x \to \infty} x \sin \frac{1}{x^\infty} = \lim_{x \to \infty} \frac{\sin \frac{1}{x}}{x} \quad (\text{form } 0^+ / 0^+).$$

Now we let $u = 1/x$, so that $x \to \infty \implies u \to 0^+$ properly, so that

$$\lim_{x \to \infty} x \sin \frac{1}{x} = \lim_{u \to 0^+} \frac{\sin u}{u} = 1.$$

Substitution is a very powerful method, but sometimes the mechanics of it become needlessly complicated and a certain amount of “hand-waving” becomes appropriate. For instance, depending upon the author and the audience the $u$-limit above might be omitted, but the substitution principle should be understood. When we think of the limits

$$\lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1, \quad \lim_{\theta \to 0} \frac{1 - \cos \theta}{\theta} = 0,$$

what is important is that the “$\theta$” inside the sine and cosine functions approaches (but never achieves!) zero at the same rate as the “$\theta$” in the denominator. Thus one may just write

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

since the substitution $\theta = \sin x$ gives us $x \to 0 \implies \theta \to 0$ properly, and $\lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1$ is known. To be cautious one can explicitly write this substitution step:

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

where $\theta = \sin x \to 0$ is a manner consistent with our theorem (as $x \to 0$), as an examination of the graph (or just the nature) of the sine curve indicates.

The next example shows how we can save some effort by just noting the approaches to zero are at the same rate. Still we should have the limit substitution theorem, and its criteria (a)–(c) in mind. We will show the careful substitution method, and then the more terse argument.

**Example 3.9.13** Consider the limit $\lim_{x \to 0} \frac{1 - \cos x^2}{x}$. Again we know what this limit would be if we had $x^2$ in the denominator, so we will rewrite the function in a form where that is the case, and compensate in the numerator:

$$\lim_{x \to 0} \frac{1 - \cos x^2}{x} = \lim_{x \to 0} \frac{x(1 - \cos x^2)}{x^2} = \lim_{x \to 0} \left[ x \cdot \frac{1 - \cos x^2}{x^2} \right].$$

At this point we could use a substitution, say $u = x^2$, and so $u \to 0 \implies u \to 0^+$, but we have different expressions for our function as $x \to 0^+$ and $x \to 0^-$:

$$\lim_{x \to 0^+} \left[ x \cdot \frac{1 - \cos x^2}{x^2} \right] = \lim_{u \to 0^+} \left[ u \cdot \frac{1 - \cos u}{u} \right] = 0^+, \quad \lim_{x \to 0^-} \left[ x \cdot \frac{1 - \cos x^2}{x^2} \right] = \lim_{u \to 0^-} \left[ -u \cdot \frac{1 - \cos u}{u} \right] = 0^-.$$
This is all correct, but instead we will simply rewrite the function so that we get the same rate of approach to zero inside the cosine and in the denominator:

\[ \lim_{x \to 0} \frac{1 - \cos x^2}{x} = \lim_{x \to 0} \left[ x \cdot \frac{1 - \cos x^2}{x^2} \right] = 0. \]

Rather than making the substitution, we rewrote the function and noticed we have matching rates of approach to zero in the \((1 - \cos x^2)/x^2\) term: we see that \(x^2 \to 0^+\) as in our limit substitution theorem so we can just cite the result \((1 - \cos \theta)/\theta \to 0\) as \(\theta \to 0\) (which includes left and right limits).

As we saw in the above example, sometimes substitution requires us to consider cases, where instead we could do some hand-waving (based upon anticipating what would happen if we did perform the substitution). The next example perhaps makes the case for selective, “informed handwaving” more strongly:

**Example 3.9.14** Consider \( \lim_{x \to 0} \frac{\sin 2x}{\sin 5x} \).

**Solution:** We need a \(2x\) to oppose the \(\sin 2x\), and \(5x\) to oppose the \(\sin 5x\), all the while not actually changing the value of the function. We do this by multiplying the numerator by \(2x/2x\), and the denominator by \(5x/5x\), with factors arranged strategically:

\[ \lim_{x \to 0} \frac{\sin 2x}{\sin 5x} = \lim_{x \to 0} \frac{2x}{5x} \cdot \frac{\sin 2x}{\sin 5x} = \lim_{x \to 0} \frac{2}{5} \cdot \frac{\sin 2x}{5x} \cdot \frac{2x}{5x} \cdot \frac{2}{5} = \frac{2}{5}. \]

To actually use a verbose substitution method, we would have to factor the function first into three factors, \(\frac{2}{5} \cdot \frac{\sin 2x}{2x} \cdot \frac{5x}{\sin 5x}\) and invoke two separate substitutions, one each for the second and third factors, and use our general limit Theorem 3.9.1, page 266. Such an approach would be correct, but quite cumbersome.

When we have \(x \to a \implies u \to \beta\), it is the exceptional cases when the \(u\)-variable approach to \(\beta\) is problematic. We always need to be aware that we require \(u \neq \beta\) as \(x \to a\), but again it is rare that there is a problem when the substitutions are routine.\(^{55}\) The next two examples give an idea of what kind of substitutions to avoid.

**Example 3.9.15** We claim that the limit \( \lim_{x \to 0} \frac{\sin \left[ x \sin \frac{1}{x} \right]}{x \sin \frac{1}{x}} \) does not exist.

Now the function \( x \sin \frac{1}{x} \) is plotted in Figure 3.22, page 241 and has two relevant features for our discussion here:

1. \(x \sin \frac{1}{x} \to 0\) as \(x \to 0\), which was proved by the Sandwich Theorem, however

2. \(x \sin \frac{1}{x}\) has infinitely many zeroes (that is, points \(x\) where the function is zero) as \(x \to 0\).

From the first point, we see that if we let \(\theta = x \sin \frac{1}{x}\), then \(x \to 0 \implies \theta \to 0\). Unfortunately, \(\theta = 0\) infinitely many times as \(x \to 0\). Thus, even though \(\lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1\), we cannot say that the original limit is the same value. Indeed, the original limit does not exist, because as we have \(x \to 0\) there are infinitely many points where the original function is not even defined, forcing us to conclude

\[ \lim_{x \to 0} \frac{\sin \left[ x \sin \frac{1}{x} \right]}{x \sin \frac{1}{x}} \text{ does not exist.} \]

\(^{55}\)In fact it is not necessary for \(u \neq \beta\) if \(f\) is continuous at \(x = a\), but then the value of the limit is just \(f(a)\). These more advanced limit techniques such as substitution are for dealing with the cases that continuity is “broken,” or the quantities are not all finite.
The reason the above limit does not exist compares to the reason \( \lim_{x \to 5} \sqrt{x^2 - 25} \) does not exist, though the latter is more obvious: that \( \sqrt{x^2 - 25} \) is undefined for part of the path of approach, in particular for \( x \in (-5, 5) \) so that \( \sqrt{x^2 - 25} \) is undefined as \( x \to 5^- \). With our example’s limit function above, the function does not exist everywhere that the denominator, \( x \sin \frac{1}{x} \) is zero, i.e., wherever \( 1/x = n\pi \), or \( x = \pi/n \) where \( n \in \mathbb{Z} \setminus \{0\} \), and this happens infinitely many times inside of any interval \((\pi, 0)\) or \((0, \pi)\) if \( \pi > 0 \). As with \( \lim_{x \to 5} \sqrt{x^2 - 25} \), we cannot approach the limit point on a continuum (from both sides in the example \( \sqrt{x^2 - 25} \)) and be sure to stay in the domain of the function, and thus we are forced to conclude the limit does not exist.

For a less subtle example, consider the following.

**Example 3.9.16** Suppose \( g(x) = \begin{cases} 
  x - 1 & \text{if } x \in (1, \infty) \\
  0 & \text{if } x \in [-1, 1] \\
  x + 1 & \text{if } x \in (-\infty, -1). 
\end{cases} \)

This is graphed in Figure 3.31. Next consider the limit

\[
\lim_{x \to 0} \frac{\sin(g(x))}{g(x)}.
\]

If we let \( u = g(x) \), then one could say \( x \to 0 \implies u \to 0 \), but the latter will not be a proper approach, because \( u = 0 \) for all \( x \in [-1, 1] \) (not just at the limit point \( x = 0 \)). In other words, \( u \) is not simply “approaching” zero for \( x \) near zero, but rather \( u \) is the constant zero for such \( x \) close enough to zero. Here \( (\sin(g(x))/g(x)) \) is undefined for \( x \in [-1, 1] \), which contains the final path of \( x \) as \( x \to 0 \). Thus we cannot say \( \lim_{x \to 0}(\sin(g(x))/g(x)) = \lim_{u \to 0}(\sin(u))/u \), as the former limit DNE while the latter is just 1.

### 3.9.4 Epilogue

The point of this section is that there are limits which we can rightfully summarize how variables and functions approach values, though we have to be very clear to do so in the spirit of limits: we see what values are approached without necessarily having those values actually occur. We also have to consider how the values are approached (including for our approaches “to infinity”): along a continuum of values, or in fits and spurts. At times one kind of “approach” is necessary, and at other times in a computation either way suffices. This depends mainly upon what is used as an input variable and what is used as an output variable, and if there are variables “in between” (as in the case of substitution), how are they behaving. With training and depth of
understanding many complicated limits can be dispatched quickly, but attempts do so without the requisite understanding can (and likely will) cause some serious errors in limit computations.

Our development of limits of functions ends here temporarily, and resumes in Chapter 9. There these techniques will be revisited and expanded, and new techniques based upon the calculus developed in between will also be introduced.

We will use the limit techniques developed here in the meantime, and the concept of limit will be embedded in much of what we do throughout the text.

Concerning the expansion of the techniques developed here, the interested student might enjoy perusing Section 9.1 at this point, as it does not depend on any calculus developed in between.

We also offer two sections regarding limits of sequences next, though they will not be necessary until much later in the text.

**Exercises**

1. Compute \( \lim_{x \to 0} \frac{\sin 9x}{x} \).
2. Compute \( \lim_{x \to 0} \frac{x}{\sin 9x} \).
3. Compute \( \lim_{x \to 0} \frac{\sin 9x}{\sin 7x} \).
4. Compute \( \lim_{x \to 0} \frac{\sin x^2}{x} \).
5. Compute \( \lim_{x \to 0} \frac{\sin x}{x^2} \).
6. Compute \( \lim_{x \to 0} \frac{\sin 5x}{x^3} \).
7. Compute \( \lim_{x \to 0} \sqrt{\frac{\sin x}{x}} \).
8. Compute \( \lim_{x \to 0} \sqrt{\frac{\sin x}{x^2}} \).
9. Compute \( \lim_{x \to 0} \frac{\tan x}{x} \).
10. Compute \( \lim_{x \to 0} \frac{\tan 2x}{x} \).
11. Compute \( \lim_{x \to 0} \frac{1 - \cos 2x}{x} \).
12. Compute \( \lim_{x \to 0} \frac{1 - \cos 2x}{x^2} \).
13. Compute \( \lim_{x \to \frac{\pi}{2}} \frac{1 - \cos 2x}{x^2} \).
14. Compute \( \lim_{x \to 0} \frac{1 - \cos x}{x^4} \).
15. Compute \( \lim_{x \to 0} \frac{\sin (\tan x)}{x} \).
16. Compute \( \lim_{x \to \frac{\pi}{4}} \frac{\sin (\tan x)}{\tan x} \).
17. Compute \( \lim_{x \to 0} \frac{\sin |x|}{x} \).
18. Show that \( \lim_{x \to 0} \frac{\sin |x|}{x} \) does not exist.
19. Compute \( \lim_{x \to 0} \frac{x}{\sin x} \).
20. Compute \( \lim_{x \to 0} \frac{x^2}{\sin x} \).
21. Compute \( \lim_{x \to -\infty} \frac{\sin e^x}{e^x} \).
22. Compute \( \lim_{x \to -\infty} \frac{\sin e^x}{e^x} \).
23. Compute \( \lim_{x \to -\infty} \frac{1 - \cos e^x}{e^x} \).
24. Compute \( \lim_{x \to 0} \frac{1 - \cos (x \sin \frac{1}{x})}{x \sin \frac{1}{x}} \).
25. Recompute the limit of Example 3.9.11, page 276, this time by multiplying the numerator and denominator of the function by \( \cos^2 x \).
26. Compute \( \lim_{x \to 0} \frac{2 \csc^2 x + 3 \csc x + 11}{5 \csc^2 x + 4 \csc x + 7} \) using a substitution argument.
27. Compute \( \lim_{x \to 0} \frac{2 \csc^2 x + 3 \csc x + 11}{5 \csc^2 x + 4 \csc x + 7} \) without a substitution argument.
28. Compute \( \lim_{x \to \infty} \left[ \frac{2x - 9}{3x + 5} \cdot \frac{6x^2 - 9x + 10}{x^2 - 7x + 8} + \frac{7x^2 + 6}{3x^3 + 4x - 3} \cdot \cos \left( \frac{1}{x} \right) \right]. \)

29. Compute \( \lim_{x \to -\infty} \left[ \frac{x^2 + 8x - 9}{5x^2 - 6x + 42} + \sin e^x \right]. \)

30. Compute \( \lim_{x \to \infty} \left[ \frac{x}{x^2 + 1} + \cos x \right]. \)

31. Compute \( \lim_{x \to \infty} \left[ \frac{x^2 + 1}{x - 1} + \cos x \right]. \)
3.10 Limits of Sequences: A First Look

Sequences offer a different set of challenges than functions on intervals, and so this section and the next are devoted to the nuances particular to sequences. In fact, a whole course could be devoted to sequences, but the amount required for this text is more modest. What we need for Riemann Sums in Chapter 6 is very similar to some earlier concepts (in particular, limits as $x \to \infty$) and can be quickly dispatched here. Some of our study of sequences here will be required to prepare for the development of series, introduced in Chapter 10 and continued in Chapter 11. Recall for what follow that $\mathbb{N} = \{1, 2, 3, \ldots\}$ and $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$.

3.10.1 Definitions and First Examples

Here we give a formal definition of an infinite sequence.

**Definition 3.10.1** An infinite sequence is a function whose domain is $\mathbb{N}$, or a similar (bounded from below) infinite subset of $\mathbb{Z}$, and whose range is a subset of $\mathbb{R}$. For instance, any function $f: \mathbb{N} \to \mathbb{R}$ will define a sequence. In such a case, for $n \in \mathbb{N}$ we usually write, for instance, $f(n) = a_n$; the subscript $n$ is used in place of the argument of the function, and the letter $a$ is used to "name" the sequence. Another notation used to denote such a sequence is as follows:

$$\{a_n\}_{n=1}^\infty = a_1, a_2, a_3, \ldots.$$  \hspace{1cm} (3.88)

The notation on the left is read, "the sequence (of numbers) $a_n$ as $n$ ranges from 1 to $\infty$." The notation on the right is useful only to show whatever pattern may be contained in the sequence, and though common, is not ideal (since it is open to misinterpretation). A simple example where a sequence is "named," a formula is given for each term, and the pattern is established as in the right hand side of (3.88) is

$$\{a_n\}_{n=1}^\infty = \left\{\frac{1}{n}\right\}_{n=1}^\infty = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots.$$  \hspace{1cm} (3.89)

So here $a_n = 1/n$ defines each term of the sequence.

For a sequence as in (3.89), we can see immediately that the terms are shrinking in size, and we would naturally enough want to describe this by stating something to the effect that

$$\lim_{n \to \infty} a_n = \lim_{n \to \infty} \frac{1}{n} \overset{1/\infty}{\longrightarrow} 0.$$  

First we need definitions for such a statement to make sense. Consider the following definition for a finite limit of a sequence:

**Definition 3.10.2** For a sequence $\{a_n\}_{n=1}^\infty$ and a finite number $L \in \mathbb{R}$, we define $\lim_{n \to \infty} a_n = L$ by the following:

$$\lim_{n \to \infty} a_n = L \iff (\forall \varepsilon > 0)(\exists N \in \mathbb{N})(\forall n) \left[ n > N \implies |a_n - L| < \varepsilon \right].$$  \hspace{1cm} (3.90)

Furthermore, in such a case we say the sequence $\{a_n\}_{n=1}^\infty$ converges to $L$. 


3.10. LIMITS OF SEQUENCES: A FIRST LOOK

Figure 3.32: Illustration of the definition of a finite limit of a sequence, (3.90). In this type of graph, \( a_n \) versus \( n \) is plotted. If the sequence converges to \( L \), then once a tolerance \( \varepsilon > 0 \) is chosen, we can find \( N \) (depending upon \( \varepsilon \)) so that \( n > N \) \( \implies |a_n - L| < \varepsilon \), i.e., \( a_n \in (L - \varepsilon, L + \varepsilon) \). Here the points are labeled by their heights, i.e., the values of the respective sequence elements \( a_n \).

Compare this definition to (3.59). In fact, this is nearly identical to the definition for \( \lim_{x \to \infty} f(x) = L \), except that the part of the continuous variable \( x \) is now played by the discrete variable \( n \). Another common illustration of this is given in Figure 3.32.

Other definitions from Section 3.8 have analogs for infinite sequences. For instance the case where \( f(x) \to \infty \) as \( x \to \infty \), (3.64), becomes the following definition:

**Definition 3.10.3** For an infinite sequence \( \{a_n\}_{n=1}^{\infty} \), we say the sequence diverges to infinity, and write \( \lim_{n \to \infty} a_n = \infty \), according to the following:

\[
\lim_{n \to \infty} a_n = \infty \iff (\forall M)(\exists N)(\forall n)(n > N \to a_n > M).
\]  

(3.91)

Thus we can find a “tail end” of the sequence to be within \( \varepsilon \) of \( L \) if we travel far enough down the sequence. Another common illustration of this is given in Figure 3.32.

We can employ all the relevant theorems for limits in \( x \) on intervals through their analogs in \( n \). The theorem which shows earlier function limit theorems imply their sequence counterparts is the following:

\[56\] It would seem strange to say a sequence converges to \( \infty \), since the verb to converge indicates getting close (or approaching). Of course we can not really “get close” to infinity.
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

Figure 3.33: Illustration of a sequence \( \{a_n\}_{n=1}^{\infty} \) (the dots) where \( a_n = 1 + \frac{\cos n\pi}{n} \), along with the function (the curve) \( f(x) = 1 + \frac{\cos \pi x}{x} \). Since \( f(x) \to 1 \) as \( x \to \infty \), that behavior carries the sequence, i.e., implying \( a_n \to 1 \) as \( n \to \infty \).

Theorem 3.10.1 Suppose that \( \lim_{x \to \infty} f(x) = L \), where \( L \) is either a real number, \( \infty \) or \( -\infty \), and \( a_n = f(n) \) for \( n \in \mathbb{N} \). Then \( \lim_{n \to \infty} a_n = L \) as well. In symbolic logic, we can write

\[
\left( \lim_{x \to \infty} f(x) = L \right) \land \left( \forall n \in \mathbb{N} \right) a_n = f(n) \implies \left( \lim_{n \to \infty} a_n = L \right).
\] (3.92)

We will not prove this since first, it is intuitive on its face upon reflection, and second, because it is not difficult to see that it should be the case based upon the similarities of the limit definitions with functions \( f : [1, \infty) \to \mathbb{R} \) and sequences \( \{a_n\}_{n=1}^{\infty} \). A nice way to summarize Theorem 3.92 is the phrase “the function carries the sequence.” Figure 3.33 shows this mechanism in action for a particular example (Example 3.10.1 below). In fact there is much more information in the statement \( f(x) \to L \) than in \( a_n \to L \), since the function outputs values for a continuum of \( x \)-values, including between the positive integer values. However, as we will eventually see, the sequence can not “carry” the function. But for now let us look at a simple example of the theorem.

Example 3.10.1 Suppose \( a_n = 1 + \frac{\cos n\pi}{n} \). Then, just as

\[
\lim_{x \to \infty} \left[ 1 + \frac{\cos \pi x}{x} \right] \xrightarrow{1+\frac{B}{\infty}} 1 + 0 = 1,
\]

we have the sequence \( \{a_n\}_{n=1}^{\infty} \) converges to 1 also, i.e.,

\[
\lim_{n \to \infty} a_n = \lim_{n \to \infty} \left[ 1 + \frac{\cos \pi n}{n} \right] \xrightarrow{1+\frac{B}{\infty}} 1 + 0 = 1.
\]

This is illustrated in Figure 3.33. (Recall that the “\( B \)” refers to a term which is “bounded,” in this case referring to the fact that \( |\cos \pi x| \leq 1 \), and so \( B/\infty \) is a determinate form which yields zero in the limit.)

In the above example, we assumed that \( B/\infty \) is a determinate form for sequences, which is the case with functions of a continuous variable (such as our usual \( x \)). Theorem 3.10.1 guarantees that this does, indeed, work with sequences as with functions defined on \([1, \infty)\). We will not
3.10. LIMITS OF SEQUENCES: A FIRST LOOK

bother to chase through the details here, but with what we have developed earlier in the chapter, it should ring true.\(^{57}\) The limits of other sequences can also be found using earlier methods (though later we will see the limitations of Theorem 3.10.1).

Example 3.10.2 Consider the sequence \(\{a_n\}_{n=1}^{\infty} = \left\{ \frac{n^3 - n}{n^2 + 1} \right\}_{n=1}^{\infty} \). Now

\[
\lim_{n \to \infty} a_n = \lim_{n \to \infty} \frac{n^3 - n}{n^2 + 1} = \lim_{n \to \infty} \frac{n^2 \left( n - \frac{1}{n} \right)}{n^2 \left( 1 + \frac{1}{n} \right)} = \lim_{n \to \infty} \frac{n - \frac{1}{n}}{1 + \frac{1}{n}} = \frac{\infty - 0}{1 + 0} = \infty.
\]

Thus the sequence \(\{a_n\}\) diverges to \(\infty\).

Anytime we can meaningfully replace \(n \in \mathbb{N}\) with \(x \in [1, \infty)\) (or a similar interval, unbounded from above), and the limit as \(x \to \infty\) exists, we can use the methods of earlier sections. However, there are examples where the sequence is better behaved than the function.

Example 3.10.3 Define \(\{a_n\}_{n=1}^{\infty}\) so that \(a_n = n \sin n \pi\). Now

\[
\lim_{x \to \infty} x \sin(x \pi) \text{ does not exist,}
\]

since the function \(f(x) = x \sin(x \pi)\) oscillates in an unbounded way as \(x \to \infty\). This behavior is illustrated in Figure 3.34. However, a closer look at \(\{a_n\}_{n=1}^{\infty}\) shows that

\[
\{a_n\}_{n=1}^{\infty} = 0, 0, 0, 0, 0, \ldots \implies \lim_{n \to \infty} a_n = 0.
\]

The example above does not contradict Theorem 3.10.1; the theorem is of the form \(P \rightarrow Q\), while here we have \((\sim P)\), a case the theorem does not address.

Because we can often replace \(a_n = f(n)\) with a function \(f : [1, \infty) \rightarrow \mathbb{R}\) which has an existing limit as \(x \to \infty\), many of the theorems which were available to us for functions on the continuum have relevance here as well. For example, the Sandwich Theorem (page 239) and its variants apply to sequences as well (just replacing \(x\) with \(n\)). Consider for example the following:

Example 3.10.4 Consider \(\left\{ \frac{n^2 \sin(n^2 + 1)}{n^3 + 1} \right\} \). Using algebraic methods we can write

\[
\lim_{n \to \infty} \frac{n^2 \sin(n^2 + 1)}{n^3 + 1} = \lim_{n \to \infty} \frac{n^2 \sin(n^2 + 1)}{n^3 \left( 1 + \frac{1}{n^3} \right)} = \lim_{n \to \infty} \left[ \sin(n^2 + 1) \cdot \frac{1}{n} \left( 1 + \frac{1}{n^3} \right) \right] = 0.
\]

The form \(B \cdot 0 = 0\), recall, represents a bounded function multiplying one which approaches zero, implying the product approaches zero, but the argument was essentially a Sandwich Theorem argument; here one could write \(-\frac{n^2}{n^3 + 1} \leq a_n \leq \frac{n^2}{n^3 + 1}\), and since \(\pm \frac{n^2}{n^3 + 1} \to 0\) as \(n \to \infty\), we conclude \(a_n \to 0\). Other more explicitly sandwich theorem applications will occur as we proceed.

Example 3.10.5 Consider the sequence \(\{a_n\}_{n=1}^{\infty} = \left\{ \sqrt[6]{\frac{6n - 1}{n}} \right\}_{n=1}^{\infty}\).

\[
\lim_{n \to \infty} a_n = \lim_{n \to \infty} \sqrt[6]{\frac{6 - \frac{1}{n}}{1}} = \lim_{n \to \infty} \sqrt[6]{6 - 1} = \sqrt[6]{6} - 0 = \sqrt[6]{6}.
\]

\(^{57}\) It is interesting to modify earlier proofs for functions of \(x\) to the language of sequences and thus actually prove the same theorems for sequences, but we will not do so here. In fact, earlier in the chapter we skipped many of the proofs for the cases \(x \to \infty\) because they produced forms we analyzed for limits with \(x \to a\) where \(a\) was a finite number.
CHAPTER 3. CONTINUITY AND LIMITS OF FUNCTIONS

Example 3.10.6 Consider the sequence \( \{a_n\}_{n=1}^\infty = \{n \sin n\pi\}_{n=1}^\infty \) is just a sequence of zeros (given by the dots above), while the function \( f(x) = x \sin(x\pi) \) oscillates with growing distance between peaks and adjacent troughs. As \( x \to \infty \), there is no limit for \( f(x) \), though \( a_n \to 0 \) as \( n \to \infty \). This does not contradict Theorem 3.10.1, since that states that the function can carry the sequence, not that the sequence can carry the function. (Note \(-x \leq x \sin(\pi x) \leq x\), so \( f(x) \) lies between the dashed lines \( y = \pm x \).)

Example 3.10.7 Consider \( a_n = \cos\left(n \sin \frac{1}{n}\right) \).

Since \( \cos x \) is continuous for all \( x \in \mathbb{R} \), we can first work “inside” that function to compute (as we have done before):

\[
\lim_{n \to \infty} n \sin \frac{1}{n} = \lim_{n \to \infty} \frac{\sin \frac{1}{n}}{\frac{1}{n}} = \lim_{\theta \to 0^+} \frac{\sin \theta}{\theta} = 1,
\]

where we made the substitution \( \theta = \frac{1}{n} \to 0^+ \) as \( n \to \infty \). Thus \( \cos\left(n \sin \frac{1}{n}\right) \to \cos 1 \approx 0.540302306 \) as \( n \to \infty \).

We usually do not mention such a theorem as we make these computations, but its statement provides a useful fact to call upon in abstract arguments.\(^{58}\) A professional mathematician might
phrase the theorem something like, “continuous functions preserve convergence of a sequence,” or “continuous functions take (or map) convergent sequences to convergent sequences.” Still, an example like the above may be summarized without direct reference to the theorem:

\[
\lim_{n \to \infty} \cos \left( n \sin \frac{1}{n} \right) = \cos(\infty \cdot 0) = \lim_{n \to \infty} \cos \left( \frac{1}{n} \right) = \frac{1}{2} \cos(1).
\]

In doing so recall that we have to be sure that the argument of the sine function approached zero at the same rate as the denominator, as occurs here.

### 3.10.2 Subsequences

As Example 3.10.3 (page 285) showed, there is more to studying sequences than just a rephrasing of our earlier study of limits “at infinity” that we began in Section 3.8. In fact the ability to easily assign a function on \([1, \infty)\) with the same limiting behavior as \(a_n\) is a rather special (though very common) case. Consider the following example.

**Example 3.10.8** Define \(a_n = (-1)^n + 1\), for \(n = 1, 2, 3, \cdots\).

First note that it is impossible to have \(n\) replaced by \(x \in [1, \infty)\) in the formula, since \((-1)^x\) is undefined for \(x\) of the form \(p/q\), where \(p, q \in \mathbb{N}\), \(p/q\) is simplified and \(q\) is even. It \((-1)^x\) is also undefined for \(x \in \mathbb{R} - \mathbb{Q}\), i.e., irrational values of \(x\). For this sequence we really do have to analyze it on its face. One method is to list the terms and see the pattern:

\[
\{a_n\}_{n=1}^{\infty} = (-1) + 1, 1 + 1, -1 + 1, 1 + 1, -1 + 1, 1 + 1, \cdots
\]

\[= \quad 0, \quad 2, \quad 0, \quad 2, \quad 0, \quad 2, \cdots.
\]

We see that the sequence never converges to a unique number, and therefore the limit does not exist, and so the sequence is divergent.

The example above shows more than just the fact that there are sequences which are best analyzed on their faces. It also shows that sequences diverge in ways other than towards \(\infty\) or \(-\infty\). Finally we see that the sequence is really a union of two sequences, one which is all zeros and the other is all twos. These can be listed as \(a_{2n}\) and \(a_{2n-1}\):

\[
\{a_{2n}\}_{n=1}^{\infty} = a_2, a_4, a_6, \cdots = 2, 2, 2, \cdots,
\]

\[
\{a_{2n-1}\}_{n=1}^{\infty} = a_1, a_3, a_5, \cdots = 0, 0, 0, \cdots.
\]

What we have above are two examples of subsequences of a given sequence \(\{a_n\}_{n=1}^{\infty}\). Now we give the formal definition.

**Definition 3.10.4** Given an infinite sequence \(\{a_n\}_{n=1}^{\infty}\) and a set of natural numbers \(n_1 < n_2 < n_3 < n_4 < \cdots\), another sequence \(\{b_k\}_{k=1}^{\infty} = \{a_{n_k}\}_{k=1}^{\infty}\) is called a subsequence of the original sequence \(\{a_n\}_{n=1}^{\infty}\).

Thus the sequence \(b_k = a_{n_k}\) is gotten by moving along the sequence \(a_n\) and picking out the terms \(a_{n_1}, a_{n_2}, a_{n_3}, \cdots\).

The following theorem contains three results which partially explain the convergence relationships among a sequence and its subsequences. Though all three results are very closely related, their emphases are somewhat different and so we list them separately. The final result, Theorem 3.10.3c, will be the most useful.
Theorem 3.10.3 Given a sequence $\{a_n\}_{n=1}^{\infty}$.

(a) If $a_n \rightarrow L$, then for any subsequence $\{a_{nk}\}_{k=1}^{\infty}$, we also have $a_{nk} \rightarrow L$ as well (as $k \rightarrow \infty$).

(b) Moreover, $a_n \rightarrow L$ if and only if for every subsequence $\{a_{nk}\}_{k=1}^{\infty}$, we have $a_{nk} \rightarrow L$.

(c) Finally, suppose that $n_{k_i} < n_{k_2} < n_{k_3} \cdots$ and $m_{j_1} < m_{j_2} < m_{j_3} < \cdots$, where $\{n_k\} \cup \{m_j\} = \{1, 2, 3, 4, \cdots\}$, and that $b_k = a_{n_k}$, $c_j = a_{m_j}$. In other words, $\{b_k\}$ and $\{c_j\}$ are subsequences of $\{a_n\}$ which exhaust that whole sequence. Then (as $k, j, n \rightarrow \infty$)

$$b_k \rightarrow L \land (c_j \rightarrow L) \iff (a_n \rightarrow L).$$

Rephrased, if we have two subsequences who collectively contain the whole original sequence, then the convergence of the original sequence to $L$ is equivalent to the convergence of the subsequences also to $L$.

The first (a) can be rephrased to state that “the sequence carries all subsequences.” The second statement (b) needs no rephrasing, and is mostly useful for junior or senior real analysis courses. The last one (c) is the most useful here, stating that if we want to check for a limit of the full sequence, it is enough to check two subsequences whose members exhaust the full sequence. For completeness we include a proof below, but the truths of these should be apparent on their faces upon reflection.\(^{59}\)

**Proof:** We will prove all of these for the case that the limit in question is some finite $L \in \mathbb{R}$. Simple modifications of the proof given here will cover the cases $L = \pm \infty$.

(a) First assume $a_n \rightarrow L$ and $b_k = a_{n_k}$ is any subsequence. Now we make the observation that $n_k \geq k$, since the $k$th choice as we move down the original sequence cannot happen before we come to the $k$th term of that sequence. Put another way, if $n_1, n_2, n_3, \cdots \in \mathbb{N}$, and $n_1 < n_2 < n_3 < \cdots$, then $n_k \geq k$. By the convergence of the original sequence to $L$ we have

$$(\forall \varepsilon > 0)(\exists N)(\forall n > N)[|a_n - L| < \varepsilon]. \quad (3.93)$$

But then once $\varepsilon$ and $N$ are chosen so that (3.93) holds true, we have

$$k > N \implies n_k \geq k > N \implies |a_{n_k} - L| < \varepsilon \iff |b_k - L| < \varepsilon.$$

In other words, for any $\varepsilon > 0$, the $N$ from (3.93) gives us $k > N \implies |b_k - L| < \varepsilon$. Thus the definition for $a_{n_k} = b_k \rightarrow L$ holds. This shows that every subsequence of $\{a_n\}_{n=1}^{\infty}$ also converges to $L$, q.e.d.

(b) By (a), we know that $a_n \rightarrow L \implies a_{n_k} \rightarrow L$. We need to show the converse ($\iff$), i.e., that if every every subsequence converges to $L$, this forces the sequence to also converge to $L$. But this is trivial, since, if every subsequence of $\{a_n\}_{n=1}^{\infty}$ converges to $L$, then the subsequence $a_{n_k}$ defined by $n_1 = 1, n_2 = 2, n_3 = 3, \cdots$ must converge to $L$. But that is just the statement $\{a_n\}_{n=1}^{\infty}$ converges to $L$, q.e.d. (In other words, the sequence is itself a subsequence, so if every subsequence converges to $L$, then so must the sequence itself, q.e.d.)

\(^{59}\)Indeed the reader should not let the proof distract, but can in good conscience bypass the proof for the moment, and return later when more familiar with sequences in general.
3.10. LIMITS OF SEQUENCES: A FIRST LOOK

(c) By (a) (and, for that matter, (b)), we have (\(\epsilon\)-). To show (\(\implies\)), suppose \(b_k, c_j \to L\) and that the union of these subsequences is the full sequence \(\{a_n\}_{n=1}^{\infty}\). For \(\epsilon > 0\), choose \(n_b\) and \(n_c\) so that

\[
\begin{align*}
n > n_b & \implies |b_n - L| < \epsilon, \\
n > n_c & \implies |c_n - L| < \epsilon.
\end{align*}
\]

Now choose \(N_b, N_c\) so that \(a_{N_b}\) is the term chosen from the original sequence to represent \(b_{n_b}\) in the first subsequence, and \(a_{N_c}\) from the original is the term \(c_{n_c}\) in the second subsequence. In other words, the \(n_b\)th term in \(\{b_k\}_{k=1}^{\infty}\) is the \(N_b\)th term in \(\{a_n\}_{n=1}^{\infty}\), and the \(n_c\)th term in \(\{c_j\}_{j=1}^{\infty}\) is the chosen to be the \(N_c\)th term in \(\{a_n\}_{n=1}^{\infty}\). Now let \(N = \max\{N_b, N_c\}\). Then

\[
\begin{align*}
n > N & \implies (a_n = b_k, \text{ some } k > n_b) \lor (a_n = c_j, \text{ some } k > n_c) \\
& \implies (|a_n - L| = |b_k - L| < \epsilon) \lor (|a_n - L| = |c_j - L| < \epsilon) \\
& \implies |a_n - L| < \epsilon, \ q.e.d.
\end{align*}
\]

Note where the proof of (c) required that the subsequences \(\{b_k\}_{k=1}^{\infty}, \{c_j\}_{j=1}^{\infty}\) exhaust the original sequence \(\{a_n\}_{n=1}^{\infty}\) (two lines up from “q.e.d.”). A quick look at the proof indicates the following corollary can be proved the same way:

**Corollary 3.10.1** Given a sequence and any finite set of subsequences whose entries exhaust the original sequence (in the sense of Theorem 3.10.3c), we then have the limit of the original sequence is \(L\) if and only if each subsequence also has limit \(L\).

It sometimes occurs that we analyze a sequence \(\{a_n\}_{n=1}^{\infty}\) by looking at the odd- and even-indexed subsequences separately, since together they exhaust the original sequence. This is especially useful when the original sequence can be simplified differently for odd and even terms. Note how

\[
\{a_n\}_{n=1}^{\infty} = \{a_{2n-1}\}_{n=1}^{\infty} \cup \{a_{2n}\}_{n=1}^{\infty}.
\]

If instead we wish to look at three subsequences, each of every third term, we could write

\[
\{a_n\}_{n=1}^{\infty} = \{a_{3n-2}\}_{n=1}^{\infty} \cup \{a_{3n-1}\}_{n=1}^{\infty} \cup \{a_{3n}\}_{n=1}^{\infty},
\]

and so on.

**Example 3.10.9** Discuss the limiting behavior of the sequence \(\left\{ \frac{(-1)^n n}{n+1} \right\} \).

**Solution:** Because this sequence alternates signs, we will look at the even and odd terms separately.

\[
a_{2n-1} = \frac{(-1)^{2n-1}(2n-1)}{(2n-1)+1} = \frac{(-1)(2n-1)}{2n} = \frac{(-1)n \left(2 - \frac{1}{n}\right)}{n(2)} \\
= \frac{-1 \cdot \frac{2}{2}}{} = -1;
\]

\[
a_{2n} = \frac{(-1)^{2n}(2n)}{(2n)+1} = \frac{2n}{2n+1} = \frac{n(2)}{n \left(2 + \frac{1}{n}\right)} \\
= \frac{2}{2 + \frac{1}{n}} \to \frac{2}{2} = 1.
\]
Thus the subsequence \( \{a_{2n-1}\}_{n=1}^{\infty} \) of odd terms approaches \(-1\) while the subsequence \( \{a_{2n}\}_{n=1}^{\infty} \) of even terms approaches \(1\). Since two subsequences have different limits, we conclude the original sequence diverges.\(^6\)

In the example above we used the fact that odd powers of \((-1)\) (e.g., \((-1)^{2n-1}\)) yield \((-1)\), and even powers (e.g., \((-1)^{2n}\)) yield \(1\). Some other ways to achieve alternation of signs use trigonometric functions, which have the conceptual advantage that they are continuous on all of \(\mathbb{R}\):

\[
\cos n\pi = \begin{cases} 
-1, & n \text{ odd} \\
1, & n \text{ even}, 
\end{cases}
\]

\[
\sin \frac{(2n-1)\pi}{2} = \begin{cases} 
1, & n \text{ odd} \\
-1, & n \text{ even}. 
\end{cases}
\]

**Exercises**

1. Show that an alternating sequence \(\{a_n\}\) converges if and only if \(\{a_n\}\) converges if and only if \(|a_n| \to 0\), and thus \(a_n \to 0\).

\(^6\)This is a case one can say the sequence diverges “by oscillation,” meaning one subsequence goes to \(L\), another to \(M \neq L\), and these subsequences exhaust the original sequence. See also Example 3.10.8, page 287.
3.11 Sequences II

Here we examine some of the more sophisticated arguments regarding sequences. In particular, we will revisit the least upper bound property of $\mathbb{R}$ (page 83), and its implications for bounded and so-called monotonic sequences. These will be of crucial importance theoretically for many of the convergence theorems for series in Chapter 10 (and thus Chapter 11). First we need some definitions.

**Definition 3.11.1** We call a sequence $\{a_n\}_{n=1}^{\infty}$

- **nondecreasing** if and only if $a_1 \leq a_2 \leq a_3 \leq \cdots$, i.e.,
  \[(\forall n \in \mathbb{N}) [a_n \leq a_{n+1}] ;\]

- **nonincreasing** if and only if $a_1 \geq a_2 \geq a_3 \geq \cdots$, i.e.,
  \[(\forall n \in \mathbb{N}) [a_n \geq a_{n+1}] ;\]

- **increasing** if and only if $a_1 < a_2 < a_3 < \cdots$, i.e.,
  \[(\forall n \in \mathbb{N}) [a_n < a_{n+1}] ;\]

- **decreasing** if and only if $a_1 > a_2 > a_3 > \cdots$, i.e.,
  \[(\forall n \in \mathbb{N}) [a_n > a_{n+1}] .\]

Note that

- $\{a_n\}$ increasing $\implies \{a_n\}$ nondecreasing,
- $\{a_n\}$ decreasing $\implies \{a_n\}$ nonincreasing.

We will often have increasing or decreasing sequences, but it turns out that our important theorems only require that the sequences are either nondecreasing or nonincreasing, and so these two weaker categories (and thus all four categories) above are collected into one concept:

**Definition 3.11.2** Any sequence $\{a_n\}_{n=1}^{\infty}$ which is either nondecreasing or nonincreasing is called monotonic.

Again we point out that any of the four categories of sequences from the first definition are therefore monotonic.\(^{61}\)

We also need to recall some definitions regarding boundedness of a set of numbers, except this time the set will be a sequence.

---

\(^{61}\)It should be pointed out that there are two camps of writers when it comes to classifying monotonic sequences. Some writers—from the other camp—use “increasing” more loosely to mean what we call “nondecreasing” (meaning “never decreasing”) here, and similarly use “decreasing” for our “nonincreasing.” On its face this seems inaccurate, but one can argue from negations that these uses make perfect sense. For example, if one thinks of an increasing sequence as one which “never decreases,” we need the definition

\[\sim [\exists n \in \mathbb{N}] (a_n > a_{n+1}) \equiv [\forall n \in \mathbb{N}] (a_n \leq a_{n+1}) ;\]

the right-hand side of which is exactly this other camp’s definition of “increasing.”

Using these different definitions of “increasing” and “decreasing,” this other camp (which will not coincide with this text’s terminology) can say a monotonic sequence is thus one which is increasing (throughout the entire sequence), or decreasing (throughout the entire sequence).

In order to distinguish the cases where $a_n < a_{n+1}$ and $a_n \leq a_{n+1}$, this other camp calls the former case “strictly increasing,” and the latter (again) simply “increasing.”
Definition 3.11.3 A sequence \( \{a_n\}_{n=1}^{\infty} \) is called

- **bounded from above** if and only if \( (\exists M \in \mathbb{R}) (\forall n \in \mathbb{N}) [a_n \leq M] \);
- **bounded from below** if and only if \( (\exists m \in \mathbb{R}) (\forall n \in \mathbb{N}) [m \leq a_n] \);
- **bounded** if and only if it is bounded from above and below, i.e., if and only if \( (\exists m, M \in \mathbb{R}) (\forall n \in \mathbb{N}) [m \leq a_n \leq M] \).

Our crucial theorem—and main result of this section—is the following:

**Theorem 3.11.1** A bounded, monotonic sequence converges.

The proof we give here contains many smaller results which are interesting in their own rights. For instance, we have

1. A nondecreasing function which is bounded from above necessarily converges.
2. A nonincreasing function which is bounded from below necessarily converges.

For this particular theorem we will defer the proof until the end of the section, to avoid distraction from the usual intuition, which is quite visual.
Chapter 4

The Derivative

The earlier chapters are the analytical preludes to calculus. This chapter begins the study calculus proper, starting with the study of differential calculus, also known as the calculus of derivatives. We will develop all of the fundamental derivative computations in this chapter.

Once we complete our development of derivatives and see many of their most immediate applications in this chapter and in Chapter 5, we will then look towards integral calculus, where roughly speaking we see how to reverse what we do here.\footnote{Integral calculus begins with Chapter 6, with the advanced computational techniques being introduced in Chapter 7. Subsequent chapters develop several topics which are either offshoots of differential and integral calculus, or are greatly extended by these. As we will see, differential calculus addresses rates at which quantities change with respect to each other, while integral calculus addresses how quantities (or the changes in quantities) accumulate. Once we lay the foundations of both differential and integral calculus, we will further develop and apply both in very diverse circumstances for the remainder of the text.}

As we will see, differential calculus addresses rates at which quantities change with respect to each other, while integral calculus addresses how quantities (or the changes in quantities) accumulate. Once we lay the foundations of both differential and integral calculus, we will further develop and apply both in very diverse circumstances for the remainder of the text.

4.1 The Derivative: Rates of Change, Velocity and Slope

Suppose that we are passengers in a car driving west to east along a highway. Further suppose that we cannot see the speedometer (measuring speed) but the highway is marked at regular intervals so we can measure our position accurately.\footnote{Here we take a function and find its derivative. Later we take the derivative and determine what was the underlying function. This reverse process is often a more formidable task, as we will see in later chapters. However a thorough understanding of the material in this chapter greatly simplifies the learning of integral calculus.}\footnote{Alternatively, we have a very accurate odometer or global positioning device in plain view.}

Using a stopwatch, we see that we traveled a total of 130 miles in 2 hours for the whole trip. Then we would say our average velocity (with positive measured in the eastward direction) was

\[
\frac{130 \text{ mi}}{2 \text{ hr}} = 65 \text{ mi/hr}.
\]

Now suppose that during the trip we would like to know our actual velocity at a particular time \(t_1\). The average for the whole trip does not usually reflect the velocity at any particular time \(t_1\) with acceptable accuracy, since we could have been stopped for a break at that particular time, or speeding up to pass a truck, or even driving in reverse (for a negative velocity). One way to attempt to approximate the velocity at time \(t_1\) is to begin our stopwatch at \(t_1\), measure how far we traveled in the next minute, and calculate the average velocity for the time interval \(t \in [t_1, t_1 + 1 \text{ minute}]\).
At this point some notation will be useful. We will take the position at time \( t \) to be \( s(t) \). It is a function, which we will call the position function; its input is time \( t \) and its output is our position at time \( t \). We will denote a change in \( t \) by \( \Delta t \), read “delta \( t \).” With \( t_1 \) as the initial time in our experiment to approximate velocity, and \( t_2 = t_1 + \Delta t \) as the final time, we see the change is indeed \( t_2 - t_1 = \Delta t \). The average velocity over any time interval \([t_1, t_2]\) is thus

\[
\frac{s(t_2) - s(t_1)}{t_2 - t_1} = \frac{s(t_1 + \Delta t) - s(t_1)}{\Delta t}.
\]

If we take \( \Delta s = s(t_1 + \Delta t) - s(t_1) \) to be the change in \( s(t) \) which results from the change in \( t \) from \( t_1 \) to \( t_1 + \Delta t = t_2 \), then we have the average velocity also equal to \((\Delta s)/(\Delta t)\), i.e.,

\[
\frac{s(t_2) - s(t_1)}{t_2 - t_1} = \frac{s(t_1 + \Delta t) - s(t_1)}{\Delta t} = \frac{\Delta s}{\Delta t}.
\]

This is akin to the old-fashioned “rate equals distance divided by time” that is taught in grade school.\(^3\) With this we can get back to the problem of attempting to find the velocity at time \( t_1 \). If we let \( \Delta t \) be equal to one minute, then we look to see how far we traveled in that one minute, and find the average velocity for that minute. If the velocity did not change very much in that time interval, then the average velocity will closely approximate the actual velocity, which we will denote \( v(t_1) \), where \( v(t) \) is the actual velocity at time \( t \) (what we would have read on the speedometer—except for a possible sign difference—were it available):

\[
\frac{\Delta s}{\Delta t} = \frac{s(t_1 + 1 \text{ minute}) - s(t_1)}{1 \text{ minute}} \approx v(t_1).
\]

On the other hand, many things can happen in a minute which can cause the velocity to change significantly. Perhaps we have a true velocity of 65 miles/hour at \( t_1 \), but then slow to a stop at a toll booth during that minute, and thus unacceptably underestimate \( v(t_1) \) as approximated by the average velocity for the time interval \([t_1, t_2]\). If possible, it would likely be much better to measure how far we traveled in the first second after \( t_1 \), since most cars cannot change velocity as significantly in such a time interval except in catastrophic circumstances (e.g., collisions). Thus\(^5\)

\[
v(t_1) \approx \frac{s(t_1 + 1 \text{ second}) - s(t_1)}{1 \text{ second}}.
\]

Following the same line of thinking, it seems reasonable that we can better approximate the actual value of \( v(t_1) \) by taking the average velocity over an interval \([t_1, t_1 + \Delta t]\) with smaller and smaller values of \( \Delta t \) (such as one minute, one second, 0.001 seconds, etc.). For this reason we actually define the velocity at time \( t_1 \) by the following, 0/0-form limit:

\[
v(t_1) = \lim_{\Delta t \to 0} \frac{s(t_1 + \Delta t) - s(t_1)}{\Delta t} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t}.
\]

Recall that we have to consider \( \Delta t \to 0^- \) as well as \( \Delta t \to 0^+ \) in this calculation. This is not unreasonable, as we could also approximate \( v(t_1) \) by considering how far we went in the minute, second, 0.001 second, etc., ending with \( t_1 \). Now we state the formal definition of velocity.

\(^{3}\)Note that we take \( \Delta t \) as one quantity. It is not “\( \Delta \) times \( t \)” One can read \( \Delta t \) to be synonymous with the change in \( t \). Occasionally we will write \( \Delta t = (\Delta t) \) to remove ambiguity and reinforce that it is one quantity. (Here \( \Delta \) is the capital Greek letter delta.)

\(^{4}\)The grade school formula is lacking in that it always assumes velocity is constant, and does not distinguish between “distance” and “displacement,” or “distance” and “position.” (Distance only carries a nonnegative sign.) It is only mentioned here because of its familiarity.

\(^{5}\)Of course we need to convert units to be consistent, e.g., 1 second = (1/3600) hour, and so on.
4.1. DERIVATIVE: RATE OF CHANGE, VELOCITY, SLOPE

Figure 4.1: Here we trace the one-dimensional motion \( s(t) = t^2 + 1 \) as a position on a number line for times \( t = -3, -2, -1, 0, 1, 2, 3 \) (note the progression of those positions in the figure). The velocities \( v = 2t \) are also given. The graph reflects how the particle comes in from the right for negative \( t \), stops at \( t = 0 \) (\( s = 1, v = 0 \)), and moves back out towards the right for positive \( t \) (faster and faster as \( t > 0 \) increases). Note that no “-axis” appears explicitly.

**Definition 4.1.1** Given a position function \( s(t) \), define the **velocity** (or instantaneous velocity, to distinguish it from average velocity) at a time \( t \) to be the function given by the limit

\[
v(t) = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = \lim_{\Delta t \to 0} \frac{s(t + \Delta t) - s(t)}{\Delta t},
\]

(4.1)

for each \( t \) for which the limit (4.1) exists, and where we also define

\[
\Delta s = s(t + \Delta t) - s(t).
\]

(4.2)

For now it is the second part of (4.1) that will be most useful. If we are lucky enough to know an algebraic formula for \( s(t) \) as a function, then we can use the limit to calculate \( v(t) \). This describes a situation known to physicists as **one-dimensional motion**. Note how (4.1) allows for \( \Delta t \to 0^- \) as well as \( \Delta t \to 0^+ \). Note also that for continuous \( s(t) \)—a reasonable assumption in classical physics—limits of form (4.1) will be of \( 0/0 \) form.

**Example 4.1.1** Suppose that position is given by \( s(t) = t^2 + 1 \). We can use (4.1) to calculate the velocity function for any fixed \( t \) as follows. As this limit will be a \( 0/0 \) form, we perform algebra to attempt to cancel the \( \Delta t \) factor in the denominator.\(^6\)

\[
v(t) = \lim_{\Delta t \to 0} \frac{s(t + \Delta t) - s(t)}{\Delta t}
\]

\[
= \lim_{\Delta t \to 0} \frac{((t + \Delta t)^2 + 1) - (t^2 + 1)}{\Delta t}
\]

\[
= \lim_{\Delta t \to 0} \frac{(t^2 + 2t\Delta t + (\Delta t)^2 + 1) - (t^2 + 1)}{\Delta t}
\]

\[
= \lim_{\Delta t \to 0} \frac{2t\Delta t + (\Delta t)^2}{\Delta t}
\]

\[
= \lim_{\Delta t \to 0} (2t + \Delta t)
\]

\[
= 2t.
\]

We showed that \( s(t) = t^2 + 1 \implies v(t) = 2t \).

Some position and velocity data are given for various times \( t \) in Figure 4.1. Note that when \( s > 0 \) the position is to the right of \( s = 0 \) (as is always the case here), and when \( s < 0 \) position is to the left. Also, when \( v > 0 \) the motion is to the right, and when \( v < 0 \) the motion is to the left. For example, at time \( t = 2 \) we have the position \( s(2) = 2^2 + 1 = 5 \), and velocity \( v(2) = 2(2) = 4 \). If \( s \) is measured in meters, and \( t \) in seconds, then the

\(^6\)Here we treat \( t \) as a constant in the calculation of \( v(t) \); \( t \) is fixed while \( \Delta t \to 0 \).
The ability to find a nonconstant velocity function is a tremendous leap from the grade school notion of “rate = distance/time.” Having limits at our disposal made it possible. Note that \( v(t) \) is really a limit of an expression of the form “position change/time change,” i.e., \((\Delta s)/(\Delta t)\) so it has some of the spirit of the grade school notion.

Such limits are useful in more than just “position \( \implies \) velocity” problems; we will have use for them throughout the text in numerous contexts. Because they are ubiquitous we generalize the notation and call the functions which arise from these limits derivatives. (The following definition should be committed to memory.)

**Definition 4.1.2** Given any quantity \( Q \) which is a function of the variable \( x \), i.e., \( Q = Q(x) \).

- The derivative of \( Q \) with respect to \( x \) is the function \( Q'(x) \), read “\( Q \)-prime of \( x \),” defined by
  \[
  Q'(x) = \lim_{\Delta x \to 0} \frac{Q(x + \Delta x) - Q(x)}{\Delta x},
  \]
  wherever that limit exists and is finite.
- If this limit does not exist or is infinite at a given \( x_0 \), we say \( Q'(x_0) \) does not exist. If the limit does exist as a finite number at \( x = x_0 \), we say \( Q(x) \) is differentiable at \( x_0 \).
- \( Q'(x) \) is also called the instantaneous rate of change of \( Q(x) \) with respect to \( x \).\(^7\)

To define the derivative \( Q'(x) \) at a given value \( x \), we require not only that the limit (4.3) exists, but also that it is finite (i.e., exists as a real number). We will make more use of the term differentiable in later sections where its justification is clearer.

Note that in most cases we expect the limit (4.3) which defines the derivative to be of \( 0/0 \) form, requiring the usual techniques of algebraic simplification to compute.

**Definition 4.1.3** We also define the average rate of change over an interval as before: if the initial value of \( x \) is \( x_0 \) (pronounced “\( x \)-naught” or “\( x \) sub(script) zero”), and the final value is \( x_f \), then the average rate of change of \( Q(x) \) with respect to \( x \) for \( x \in [x_0, x_f] \) or \( x \in [x_f, x_0] \) (depending upon whether \( x_0 < x_f \) or \( x_0 > x_f \)) is given by the difference quotient

\[
\frac{Q(x_f) - Q(x_0)}{x_f - x_0} = \frac{Q(x_0 + \Delta x) - Q(x_0)}{\Delta x} = \frac{\Delta Q}{\Delta x},
\]

where

\[
\Delta x = x_f - x_0, \quad \Delta Q = Q(x_f) - Q(x_0) = Q(x_0 + \Delta x) - Q(x_0).
\]

So we see that the derivative (4.3) is just the limit of the average rate of change in \( Q \) given in (4.4) on an interval with endpoints \( x \) and \( x + \Delta x \), assuming that limit as \( \Delta x \to 0 \) is finite.

With this notation, we can rewrite the (instantaneous) velocity function for a given \( s(t) \) as:

\[
v(t) = s'(t).
\]

\(^7\)“Instantaneous” rate of change means the rate of change “at that instant,” as opposed to an average rate of change of the output variable which occurs over an entire interval’s length of values of the input variable.
4.1. DERIVATIVE: RATE OF CHANGE, VELOCITY, SLOPE

Because there are so many contexts, there are many different notations. They each have their places and all are worth knowing.8

Example 4.1.2 Suppose our car has a very accurate fuel gauge and a very accurate odometer. Let \( V(s) \) be the volume of fuel in the tank at a particular position \( s \), as read from the odometer.9 Then

\[
V'(s) = \lim_{\Delta s \to 0} \frac{\Delta V}{\Delta s} = \lim_{\Delta s \to 0} \frac{V(s + \Delta s) - V(s)}{\Delta s}
\]

represents the instantaneous rate of fuel flow in terms of volume per unit length. If \( s \) is in units of miles and \( V \) is in units of gallons, this would be a flow rate in gallons per mile. (If we prefer miles/gallon, we can take the reciprocal.) So the derivative can also represent flow of a fluid.

Notice that the fuel should be leaving the tank whenever the engine is running, so \( V \) should be decreasing as we drive. This gives \( \Delta V < 0 \) when \( \Delta s > 0 \), giving \( \Delta V/\Delta s < 0 \), and thus \( V'(s) \leq 0 \). However the fuel running through the engine is exactly the fuel leaving the tank, and so the actual flow rate we would report would be \(-V'(s)\), a positive quantity, for any particular position \( s \).

There are countless other applications of the derivative. All we need is a quantity \( Q \) as a function of another quantity \( x \), to measure the rate that \( Q \) changes as \( x \) changes. The average rate of change of \( Q \) with respect to \( x \) is again \( (\Delta Q)/(\Delta x) \), and the instantaneous rate is the number we get when we take \( (\Delta Q)/(\Delta x) \) and let \( \Delta x \to 0 \), giving the rate “at that instant.”

4.1.1 Slope of a Curve

All of these applications have their own interpretations. Interestingly enough, the analytic geometric interpretation of the derivative of a function unifies them all in one graphical setting. For the bulk of the remainder of this section, we will concentrate on the significance of the derivative \( f'(x) \) to the graph of \( y = f(x) \).

First we will consider a very simple case. Suppose

\[
f(x) = mx + b,
\]

where \( m, b \in \mathbb{R} \) are fixed constants. Then

\[
f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{[m(x + \Delta x) + b] - [mx + b]}{\Delta x} = \lim_{\Delta x \to 0} \frac{m\Delta x}{\Delta x} = \lim_{\Delta x \to 0} m = m.
\]

Thus, when \( y = f(x) \) is the line \( y = mx + b \) we get \( f'(x) = m \); if the function is a line then the derivative is its slope.

Recall that the slope of a line measures how rapidly that line rises or falls as we move along the line and to the right. In other words, slope measures the rate of change in \( y \) with respect to \( x \). That rate is constant on a line, but changes on most curves. Still, if we look closely at a point \((a,f(a))\) on the graph of \( y = f(x) \), we can often associate a slope with the curve there.10 To

---

8In a later section we will introduce the very powerful Leibniz notation for the derivative \( Q'(x) \), which we will then write \( dQ/dx \) (notice the resemblance to \( \Delta Q/\Delta x \)).

9Technically odometers display total distance traveled and not position, but here we can use the reading as a position under the assumption all travel is in the same, positive direction.

10Just as a naive, and nearsighted, observation of the Earth’s surface can lead us to believe the Earth is flat, if we were standing on a curve at \((a,f(a))\), and very focused on the curve at and around that point, we might believe we are looking at a constant slope. The actual slope is what we approach when we focus more and more on that point, by letting \( \Delta x \to 0 \).
measure this slope, again we would if effect measure the way \( y = f(x) \) changes (instantaneously) with respect to \( x \) at \( x = a \). With this motivation, we make the following definition:

**Definition 4.1.4** Given a function \( f(x) \), the slope of the graph of \( y = f(x) \) at any point \((a, f(a))\) on the graph is given by \( f'(a) \), assuming this derivative exists there.\(^{11}\)

It is important to note that a function and its derivative give two types of information about the graph of the function:

- \( f(x) \) gives the height of the graph for a particular \( x \)-value;

- \( f'(x) \) gives the slope of the graph at that \( x \)-value.

For instance, we saw in Example 4.1.1, page 295 (using different variables) that \( f(x) = x^2 + 1 \implies f'(x) = 2x \). When we graph \( y = x^2 + 1 \), i.e., when we graph the function \( f(x) = x^2 + 1 \), the function \( f(x) \) gives the height at each \( x \), and the derivative \( f'(x) = 2x \) gives the slope there. This is illustrated in Figure 4.2.

Also illustrated in that figure, and of geometric interest, is the tangent line to the graph of \( y = f(x) \) at a given point \((a, f(a))\) on the curve. This is just the line through \((a, f(a))\) with the same slope as the curve, i.e., with slope \( f'(a) \).

**Definition 4.1.5** The line through \((a, f(a))\) with slope \( f'(a) \), i.e., the same slope as the function at \( x = a \), is the tangent line to the graph of \( y = f(x) \) through \((a, f(a))\).

Three separate tangent lines are drawn in Figure 4.2. A formula for the tangent line through \((a, f(a))\) presents itself immediately, since we have a point \((a, f(a))\), and a slope \( f'(a) \), the modified point-slope form (2.59), page 131 gives us:

\[
y = f(a) + f'(a)(x - a). \tag{4.8}
\]

For the function \( f(x) = x^2 + 1 \) in Figure 4.2, (4.8) gives us

- At \( x = 1 \): \((a, f(a)) = (1, 2)\), \( f'(a) = f'(1) = 2 \), so the tangent line is \( y = 2 + 2(x - 1) \).

- At \( x = -2 \): \((a, f(a)) = (-2, 5)\), \( f'(a) = f'(-2) = -4 \), so the tangent line is \( y = 5 - 4(x + 2) \).

- At \( x = 0 \): \((a, f(a)) = (0, 1)\), \( f'(a) = f'(0) = 0 \), so the tangent line is \( y = 1 + 0(x - 0) \), or simply \( y = 1 \).

This tangent line form (4.8) appears repeatedly in this text; much of Section 5.4 is devoted to it.

**Example 4.1.3** Consider the function \( f(x) = \sqrt{2x + 1} \). Then, using a conjugate multiplication

---

\(^{11}\)Recall \( f'(a) \) exists means exactly that \( \lim_{\Delta x \to 0} \frac{f(a + \Delta x) - f(a)}{\Delta x} \) exists as a limit and is finite.
4.1. DERIVATIVE: RATE OF CHANGE, VELOCITY, SLOPE

Figure 4.2: The graph of \( f(x) = x^2 + 1 \), along with the tangent lines to the graph at \( x = -2, 0, 1 \). A tangent line is a line through a point \((a, f(a))\) on the curve which has the same slope as the curve’s at that point. The height at each \( x \) is given by \( f(x) = x^2 + 1 \), while the slope is given by \( f'(x) = 2x \).

Figure 4.3: The graph of \( f(x) = \sqrt{2x + 1} \), along with a few slopes. Notice the behavior of the slope, in particularly how \( f'(x) = \frac{1}{\sqrt{2x + 1}} \to \infty \) as \( x \to -1/2^+ \), and how \( f'(x) \to 0^+ \) as \( x \to \infty \).
(third line below) we can compute:

\[
\begin{align*}
  f'(x) &= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \\
  &= \lim_{\Delta x \to 0} \frac{\sqrt{2(x + \Delta x) + 1} - \sqrt{2x + 1}}{\Delta x} \\
  &= \lim_{\Delta x \to 0} \frac{\sqrt{2(x + \Delta x) + 1} - \sqrt{2x + 1}}{\Delta x} \cdot \frac{\sqrt{2(x + \Delta x) + 1} + \sqrt{2x + 1}}{\sqrt{2(x + \Delta x) + 1} + \sqrt{2x + 1}} \\
  &= \lim_{\Delta x \to 0} \frac{(2(x + \Delta x) + 1) - (2x + 1)}{\Delta x} \left( \frac{1}{\sqrt{2(x + \Delta x) + 1} + \sqrt{2x + 1}} \right) \\
  &= \lim_{\Delta x \to 0} \frac{2x + 2\Delta x + 1 - 2x - 1}{\Delta x} \left( \frac{1}{\sqrt{2(x + \Delta x) + 1} + \sqrt{2x + 1}} \right) \\
  &= \lim_{\Delta x \to 0} \frac{2\Delta x}{\Delta x} \left( \frac{1}{\sqrt{2(x + \Delta x) + 1} + \sqrt{2x + 1}} \right) \\
  &= \lim_{\Delta x \to 0} \frac{2}{\sqrt{2(x + \Delta x) + 1} + \sqrt{2x + 1}} = \frac{2}{2\sqrt{2x + 1}} = \frac{1}{\sqrt{2x + 1}}.
\end{align*}
\]

To summarize, \( f(x) = \sqrt{2x + 1} \implies f'(x) = \frac{1}{\sqrt{2x + 1}}. \)

We can make several observations about the form of this derivative. First, note that \( f(-1/2) = 0 \) exists, but \( f'(-1/2) \) does not. Of course for \( x = -1/2 \) we cannot take \( \Delta x \to 0^- \) or we would be taking square roots of negative numbers. Still, it is interesting to notice that \( f'(x) \to \infty \) as \( x \to -1/2^+ \). Furthermore, as \( x \to \infty \), we have \( f'(x) \to 0^+ \), so the function becomes less sloped as we move \( x \) farther to the right. This is all reflected in the graph, as illustrated in Figure 4.3.
4.1. DERIVATIVE: RATE OF CHANGE, VELOCITY, SLOPE

301

1 2 3

−1 2

1

−1

2

1

f'(1/2) = −4

f'(1) = −1

f'(2) = −1/4

f'(-2) = −1/4

f'(-1) = −1

f'(-1/2) = −4

Figure 4.4: Illustration of the graph of \( f(x) = \frac{1}{x} \), from Example 4.1.4 showing also data from its derivative \( f'(x) = −\frac{1}{x^2} \). From the graph or from the formula for \( f'(x) \), we can notice that \( f'(x) < 0 \) for all \( x \neq 0 \), and observe the behavior of \( f'(x) \) as \( x \to \infty \), as \( x \to −\infty \), and as \( x \to 0 \).

Example 4.1.4 Consider the function \( f(x) = \frac{1}{x} \). We will find its slope everywhere that it is defined. The method below involves multiplying the numerator and denominator by a factor which will remove the fractions in the numerator.

\[
f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\frac{1}{x + \Delta x} - \frac{1}{x}}{\Delta x} = \lim_{\Delta x \to 0} \frac{1}{\Delta x} \cdot \frac{x(x + \Delta x)}{x(x + \Delta x)} = \lim_{\Delta x \to 0} \frac{-\Delta x}{(\Delta x)(x + \Delta x)} = \lim_{\Delta x \to 0} \frac{-1}{x(x + \Delta x)} = -\frac{1}{x^2}.
\]

Summarizing, \( f(x) = \frac{1}{x} \implies f'(x) = -\frac{1}{x^2} \).

We see some interesting features of this derivative as well. For instance, it is always negative, so the graph is always sloping downwards, where the derivative exists (and the function is defined). Furthermore, it is the same at \( x = a \) as \( x = −a \), which is a result of the symmetry. Finally, we note that \( f'(x) \to 0 \) as \( x \to \pm\infty \), and \( f'(x) \to −\infty \) as \( x \to 0 \). This is indeed reflected in the graph in Figure 4.4.

4.1.2 Marginal Cost

As we will see, the derivative has many applications. While geometrically it represents the slope of the curve of a function, we have already seen it can represent any instantaneous rate of change of a quantity \( Q(x) \) with respect to its input variable \( x \). A very good illustration is
an object’s velocity \( v(t) = s'(t) \) when the function \( s(t) \) is its position and \( t \) is time, but another very different example was \( V'(s) \) measuring the change of volume \( V \) in a vehicle’s fuel tank as its position \( s \) varies. Nearly any field which deals with numerical quantities has some use for the derivative, which measures how those quantities change with respect to each other. One such field is business economics.

For example, in economics the precise definition of marginal cost is the cost of the next item, or \((x + 1)st\) item, after \( x \) items are already produced and their costs for production already paid. In function notation, if \( C(x) \) is the total cost of producing \( x \) items, then (upon reflection) it is clear that the marginal cost at that level—that is, the cost of the \((x + 1)st\) item—would be \( C(x + 1) - C(x) \). One might assume that with the infrastructure required to produce the first \( x \) items already in place, this next item would be relatively inexpensive—though probably not free—to produce. However there are exceptions to this, where it could be much more expensive than the previous items if more infrastructure is suddenly needed to produce that next item. For instance, if one jetliner can carry a maximum of 400 passengers, the 400th may be nearly free for the airline to seat and fly (after the previous 399 are provided for), but the 401st would likely be very expensive since another plane is required (even if ultimately the 401 passengers would be divided more evenly between the planes). That would be a case where \( C(401) - C(400) \) is quite large.

In practice, rather than computing the actual marginal cost \( C(x + 1) - C(x) \), a proxy (substitute) for this is often used, usually the instantaneous rate of change of total cost \( C(x) \) with respect to the number of units \( x \), that is, \( C'(x) \), where \( C(x) \) is a function whose formula seems to make computational sense for \( x \) on actual intervals (and not just at nonnegative integer values). Thus many textbooks instead make the following definition (without the parenthetical):\(^{12}\)

**Definition 4.1.6** If \( C(x) \) is the total cost function for producing \( x \) items, then the (proxy) marginal cost function is given by \( C'(x) \).

This assumes \( C'(x) \) makes sense as a function, usually requiring some formula for \( C(x) \). Then \( C''(x) \) will have dimensions which are units of money per units of items produced; if \( C \) is

\(^{12}\)In many fields there will be a precise definition and a different, “working definition” which itself may change from context to context. For example, in statistics a numerical datum said (as a working definition) to be in the 75th percentile is often understood to be one for which 75% of the other data is below it. Elsewhere being in the 75th percentile is described as meaning 25% of the other data is above it. However these cannot both be exact because if 75% is below and 25% is above, when we add in the datum in question, we would have over 100% of our data, which is impossible. If the data set is very large, so that one datum is much less than 1% of the total observations, the discrepancy is minor enough to ignore. For a much smaller sample the discrepancy is large. (The third highest of four data could be described as being in the 50th percentile by one definition, and 75th in the other.)

Often one measurement is used as a proxy for the measurement which matches the exact definition, and if that proxy is used more and more, it is sometimes given as the definition. This is indeed the case with marginal cost; there are textbooks which give only the proxy, derivative definition. The two definitions are connected by the approximation below, in which the precise definition of marginal cost at level \( x \) is on the left-hand side, and its usual proxy which is often given as the definition is on the right in the following:

\[
\frac{C(x + 1) - C(x)}{1} \approx C'(x).
\]

While this may be somewhat confusing (or simply unsatisfying), it is similar to the idea that, if \( s'(t_1) = 5 \text{ ft/sec} \), we expect to have gone approximately 5 ft in the next second after \( t_1 \), though the more exact answer would be \( s(t_1 + 1 \text{ sec}) - s(t_1) \). In a factory producing 10,000 cars it is reasonable to assume the approximation above using \( C'(x) \) is quite good, assuming a reasonable formula for \( C(x) \) which makes sense on intervals, while an aerospace firm producing a small number of stealth fighter jets would likely use the exact definition \( C(x + 1) - C(x) \).

With the derivative rules we will derive as this chapter unfolds, we will see that it is often easier and more efficient to compute \( s'(t_2) \) than to compute the difference between \( s(t) \)-values at the end and start of the interval \( t \in [t_1, t_1 + 1 \text{ sec}] \). That is not the case here, where we find derivatives from the limit definition.
in dollars and $x$ is in items produced, then $C'(x)$ will be in dollars per item. This is because the limit inherits the units from the quotient in the definition

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

In contrast, the units of the actual (non-proxy) marginal cost $C(x + 1) - C(x)$ will be in dollars only, though since it is for one item, it can be still considered a dollar amount “per item.”

**Example 4.1.5** Consider the cost function $C(x) = -3x^2 + 90x + 500$, representing the cost (in hundreds of dollars) of manufacturing $x$ cases of a particular product. Assume this function is valid for $0 \leq x \leq 25$. Find the average rate of change of cost per case as $x$ ranges from 10 to 15 cases. Also find the instantaneous rate of change in cost when 10 cases are manufactured.

**Solution:** The average rate of change of cost per case on $[10, 15]$ will be given by the difference quotient $(C(15) - C(10))/(15 - 10)$. (See Equation 4.4 on page 296.) Thus we compute:

$$C(15) = -3(15)^2 + 90(15) + 500 = 1175,$$

$$C(10) = -3(10)^2 + 90(10) + 500 = 1100,$$

and so the average rate of change of cost per case for $10 \leq x \leq 15$ is

$$\frac{C(15) - C(10)}{15 - 10} = \frac{1175 - 1100}{5} = \frac{75}{5} = 15.$$

Since $C$ is in hundreds of dollars, and the quantities in the denominator above are in cases of product, it follows that the units of this final quantity will be in hundreds of dollars per case. We could report our answer as 15 hundred dollars per case, or $1500/case.

To compute the instantaneous rate at $x = 10$, we will compute $C'(x)$ in the abstract and input $x = 10$ (instead of computing the special case $C'(10)$ directly, for which $C(x + \Delta x)$ would be $C(10 + \Delta x)$, and $C(x)$ would be $C(10)$).

$$C'(x) = \lim_{\Delta x \to 0} \frac{C(x + \Delta x) - C(x)}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{[-3(x + \Delta x)^2 + 90(x + \Delta x) + 500] - [-3x^2 + 90x + 500]}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{[-3(x^2 + 2x\Delta x + (\Delta x)^2) + 90x + 90\Delta x + 500] + 3x^2 - 90x - 500}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{-3x^2 - 6x\Delta x - 3(\Delta x)^2 + 90x + 90\Delta x + 500 + 3x^2 - 90x - 500}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{-6x\Delta x - 3(\Delta x)^2 + 90\Delta x}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{(\Delta x)[-6x - 3\Delta x + 90]}{\Delta x}$$

$$= \lim_{\Delta x \to 0} [-6x - 3\Delta x + 90]$$

$$\implies C'(x) = -6x + 90.$$  

Thus $C'(10) = -6(10) + 90 = 30$. Because of the form of the limit, since $C(x + \Delta x)$ and $C(x)$ are in hundreds of dollars and $\Delta x$ is in cases (of product), the quotients above are all in units of hundreds of dollars/case, and so therefore is the limit $C'(x)$. Thus $C'(10) = 30$ hundreds of dollars/case, and so we can instead report that the instantaneous rate of change at $x = 10$ cases is $3000/case.
The derivative computation above may seem tedious, but fortunately techniques of later sections will make the derivation

\[ C(x) = -3x^2 + 90x + 500 \implies C'(x) = -6x + 90 \]

as simple as writing just that. In Exercise 18 of this section it is noted that there is a general formula for the derivative of a function of the form \( f(x) = ax^2 + bx + c \), depending only upon the coefficients. That and other rules for derivatives will make computations such as the one above almost trivial.

We should note that with the exact definition, at \( x = 10 \) we would compute the marginal cost to be (in hundreds of dollars per unit)

\[
C(11) - C(10) = [\{-3(11)^2 + 90(11) + 500\} - \{-3(10)^2 + 90(10) + 500\}] = 1127 - 1100 = 27.
\]

This is well approximated by our computation of \( C'(10) = 30 \) above. Economics being rather far from an exact science anyhow, using \( C'(10) = -6(10) + 90 = 30 \) seems like a reasonable proxy for the computation of the exact marginal cost of \( C(11) - C(10) = 27 \) given above. From time to time we will make some use out of the idea that for small \( \Delta x \),

\[
\frac{f(x + \Delta x) - f(x)}{\Delta x} \approx f'(x),
\]

depending upon how “small” is \( \Delta x \), and upon how quickly \( (f(x + \Delta x) - f(x))/\Delta x \to f'(x) \) as \( \Delta x \to 0 \). While the notation will change, this idea is present in many contexts in this text. A special case of this, then, is with the marginal cost but with \( \Delta x = 1 \) item:

\[
\frac{C(x + 1) - C(x)}{1} \approx C'(x).
\]

The computation above would automatically contain the units of money/item, which is understood when the question is asked, “how much would the next product cost?” but flows better from the approximation above (by the presence of the denominator, even if it is simply 1).

It should also be pointed out that there are similar, related quantities in economics, such as total revenue and marginal revenue, and total profit and marginal profit. Whether total or marginal, it is assumed that profit is the difference between revenue and cost. Hence one often sees the equations \( P = R - C \) and \( P' = R' - C' \), relating the total versions of profit, revenue and cost, and also the derivative proxies for the marginal versions of profit, revenue and cost.

4.1.3 Some Further Applications

In Definition 4.1.2 page 296 we saw that we can consider \( Q'(x) \) as the derivative of \( Q(x) \) with respect to \( x \), whatever \( Q \) and \( x \) represent, and \( Q'(x) \) is given by (4.3):

\[
Q'(x) = \lim_{\Delta x \to 0} \frac{Q(x + \Delta x) - Q(x)}{\Delta x}.
\]

The real power of the definition lies in the interpretation and intuition in the meaning of the derivative \( Q'(x) \), as the instantaneous rate of change of \( Q(x) \) with respect to \( x \). So far we have had four applications:

- velocity \( v(t) = s'(t) \), the instantaneous rate of change in position \( s(t) \) with respect to the time \( t \),
4.1. DERIVATIVE: RATE OF CHANGE, VELOCITY, SLOPE

- the rate $V'(s)$ of fuel volume change with respect to position (or distance traveled) $s$,
- the slope $f'(x)$ of the graph of $y = f(x)$ at a given value of $x$, which also represents the
  change in the height of the function $f(x)$ with respect to the horizontal position $x$, and
- the (proxy) marginal cost $C'(x)$, measuring the rate of change in cost with respect to the
  number of items produced.

Throughout this chapter and the next we will consider further contexts for the derivative. Indeed, the derivative arguably offers a unique insight into each setting, though all of these
insights are ultimately unified mathematically in the limit formula for $Q'(x)$, as the instantaneous
rate of change of $Q(x)$ with respect to $x$.

As with any application problem, some care must be taken to be sure that the natural quanti-
ties considered are the same quantities which are usually considered in the common conversation.
For instance, recall that if the volume of fuel in a tank is $V(s)$, the fuel flow is actually $-V'(s)$.
(See Example 4.1.2, page 297.)

Example 4.1.6 Suppose a liquid is stored in an inverted conical container where the height of
the cone is twice its radius. Find a formula for the instantaneous rate of change of the volume $V$
with respect to the height $h$ of the liquid in the cone.

Solution: What we seek here is $V'(h)$, and so we first need to find a formula for $V(h)$. In
general, the volume of a cone (such as that represented by the liquid in the tank) is given
by $V = \frac{1}{3} \pi r^2 h$, where $r$ is the radius of the “base” of the cone and $h$ is the “height” of
the cone. For the whole tank, we would have $h = 2r$, and so by similar triangles this will be the case regardless of the height of the fluid in the cone.

We can use this to substitute $r = \frac{1}{2} h$. Putting these together we get $V$ as a function of $h$,
namely $V = \frac{1}{3} \pi \left(\frac{1}{2} h\right)^2 h = \frac{1}{12} \pi h^3$, or

$$ V(h) = \frac{1}{12} \pi h^3. $$

Now we compute $V'(h)$ in the usual way:

$$ V'(h) = \lim_{\Delta h \to 0} \frac{V(h + \Delta h) - V(h)}{\Delta h} = \lim_{\Delta h \to 0} \frac{\frac{1}{12} \pi (h + \Delta h)^3 - \frac{1}{12} \pi h^3}{\Delta h} $$

$$ = \lim_{\Delta h \to 0} \left[ \frac{\pi}{12} \left( h^3 + 3h^2 \Delta h + 3h(\Delta h)^2 + (\Delta h)^3 - h^3 \right) \right] = \lim_{\Delta h \to 0} \left[ \frac{\pi}{12} \frac{(\Delta h)(3h^2 + 3h(\Delta h) + (\Delta h)^2)}{\Delta h} \right] $$

$$ = \lim_{\Delta h \to 0} \left[ \frac{\pi}{12} (3h^2 + 3h \Delta h + (\Delta h)^2) \right] = \frac{\pi}{12} (3h^2) = \frac{\pi}{4} h^2. $$

Thus $V'(h) = \frac{\pi}{4} h^2$.

As an example, the instantaneous rate of change of the volume $V$ when $h = 5cm$ is

$$ V'(5 \text{cm}) = \frac{\pi}{4} (5 \text{cm})^2 = \frac{25\pi}{4} \frac{\text{cm}^3}{\text{cm}} \approx 19.6 \frac{\text{cm}^3}{\text{cm}}. $$

The units first appear as cm$^2$, but we write the units as cm$^3$/cm because the rate should be in
units of volume change per unit of height length change.
Exercises

For problems 1–16, use the definition of the derivative,
\[ f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}, \]
to find \( f'(x) \) for the given function.
1. \( f(x) = 5 - 2x \).
2. \( f(x) = 10 \).
3. \( f(x) = 2x^2 + 3 \).
4. \( f(x) = 3x^2 - 5x + 9 \).
5. \( f(x) = \sqrt{x} \).
6. \( f(x) = 3x + 2 \).
7. \( f(x) = \sqrt{9 - 5x} \).
8. \( f(x) = \frac{1}{x} \).
9. \( f(x) = \frac{2}{\sqrt{x}} \).
10. \( f(x) = x^{3/2} \). (Hint: Rewrite as \( \sqrt{x^3} \).)
11. \( f(x) = 2x^3 \).
12. \( f(x) = \sqrt{x} + 1 \). (Hint: We made use of the difference of two squares (see Section 2.2) in Example 4.1.3. Here you will need to use the difference of two cubes in a similar manner.)
13. \( f(x) = x^4 \).
14. \( f(x) = \frac{x}{x + 1} \). (Hint: easier if \( f(x) \) is rewritten using long division.)
15. \( f(x) = \frac{x + 1}{x - 1} \).
16. \( f(x) = x^{2/3} \).
17. Suppose \( s(t) = -16t^2 + 15t + 20 \) describes the height of a projectile in free fall.
   a. Find the velocity function \( v(t) \), using (4.1), page 295.
   b. What is the projectile’s velocity when \( t = 0? \ t = 10? \)
   c. Find \( t \) so that the projectile is stationary (i.e., \( v = 0 \)).
   d. How high is the projectile when it is stationary?
18. Consider the general quadratic function \( f(x) = ax^2 + bx + c \).
   a. Use the definition of derivative to find a formula for the derivative of the general quadratic function \( f(x) = ax^2 + bx + c \).
   b. Assuming \( a \neq 0 \), this represents a parabola. Assuming also that the slope is zero at the vertex, find a general formula for the \( x \)-coordinate of the vertex.
   c. Find a general formula for the \( y \)-coordinate of the vertex, and thus a formula for the point \((x,y)\) at the vertex.
19. The cost (in hundreds of dollars) from manufacturing \( x \) cases of a product is \( C(x) = 8x - 0.2x^2 \).
   a. Find the average rate of change of cost between 10 and 20 cases are manufactured.
   b. Find the marginal cost when 15 cases are manufactured.
   c. Do the same for 18 cases.

For each of the following, find the tangent line to the graph at the given point \( x = a \) for the given function. (See (4.8), page 298 and the examples following.)
20. \( f(x) = x^2 - 9 \), \( a = 4 \).
21. \( f(x) = x^3 \), \( a = -1 \).
22. \( f(x) = \sqrt{x} \), \( a = 9 \).
23. \( f(x) = \frac{1}{x} \), \( a = \frac{1}{10} \).
4.1. DERIVATIVE: RATE OF CHANGE, VELOCITY, SLOPE

Inserted so that page numbers for Calculus 1 students’ notes would be correct.
4.2 First Differentiation Rules; Leibniz Notation

In this section we derive rules which let us quickly compute the derivative function \( f'(x) \) for any polynomial function \( f(x) \), and for \( \sin x \) and \( \cos x \). Along the way we will derive a few general (though not comprehensive) rules for derivatives. We will also introduce the very powerful Leibniz notation for derivatives, and show how knowing the derivative helps us to further analyze a function. One consequence is that we can more accurately graph a function’s behavior by hand.

4.2.1 Positive Integer Power Rule

We will often be interested in finding derivatives of functions \( f(x) = x^n \). Fortunately there is a simple rule which covers all such functions. It is usually called the power rule, as stated below. (Recall \( \mathbb{N} = \{1, 2, 3, 4, 5, 6, \ldots \} \).

**Theorem 4.2.1** \((f(x) = x^n) \land (n \in \mathbb{N}) \implies f'(x) = n \cdot x^{n-1} \).

Note that implicit in this theorem is that the derivative of \( x^n \) exists for every \( x \in \mathbb{R} \)—i.e., “exists everywhere”—since it is equal to \( nx^{n-1} \), defined everywhere.

**Proof:** The proof we give here depends upon the binomial expansion (2.36), page 99. It is important to remember that \( x \) is a fixed number in the limit, and \( \Delta x \) is the variable approaching zero as far as the limit is concerned. With that in mind,

\[
f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}
= \lim_{\Delta x \to 0} \frac{(x + \Delta x)^n - x^n}{\Delta x}
= \lim_{\Delta x \to 0} \frac{x^n + nx^{n-1}\Delta x + \frac{n(n-1)x^{n-2}(\Delta x)^2}{2} + \cdots + (\Delta x)^n}{\Delta x} - x^n
= \lim_{\Delta x \to 0} \frac{nx^{n-1}\Delta x + \frac{n(n-1)x^{n-2}(\Delta x)^2}{2} + \cdots + (\Delta x)^n}{\Delta x}
= \lim_{\Delta x \to 0} \left(nx^{n-1} + \frac{n(n-1)}{1 \cdot 2}x^{n-2}(\Delta x)^2 + \cdots + (\Delta x)^{n-1}\right)
= nx^{n-1} + 0 + \cdots + 0
= nx^{n-1}, \quad \text{q.e.d.}
\]

The only term which survives in the limit in the fifth line is the \( nx^{n-1} \) term because the others have positive integer powers of \( \Delta x \), which is approaching zero. We will see later that this power rule is actually much more general. In fact, it can be used for \( n \in \mathbb{R} \) but we need some more advanced methods to prove such generality. For now we will apply it only to \( n \in \mathbb{N} \).

**Example 4.2.1** Here we list the derivatives of some of the positive integer powers of \( x \). The first case listed below (\( n = 1 \)) does follow from the proof, though we would be reading the statement of the theorem for that case \( f(x) = x \), \( \implies f'(x) = 1x^0 = 1 \). Again we do not really wish to say \( x^0 = 1 \) regardless of \( x \), for several technical reasons (though it is fine as long as \( x > 0 \)), but we see how the formula naively gives us what we want for \( n = 1 \). The rest of the table is more straightforward:

<table>
<thead>
<tr>
<th>( f(x) )</th>
<th>( x )</th>
<th>( x^2 )</th>
<th>( x^3 )</th>
<th>( x^4 )</th>
<th>( x^{100} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f'(x) )</td>
<td>1</td>
<td>2x</td>
<td>3x^2</td>
<td>4x^3</td>
<td>100x^{99}</td>
</tr>
</tbody>
</table>
4.2. FIRST DIFFERENTIATION RULES; LEIBNIZ NOTATION

4.2.2 Leibniz Notation

We will find that other derivative rules will be unwieldy to write with our present notation. Thus we will introduce the very powerful Leibniz notation and use it except in a few settings where our present (prime) notation is simpler.

**Definition 4.2.1** \( \frac{d}{dx} f(x) = f'(x) \).

This is also written \( \frac{df(x)}{dx} \), and sometimes shortened to \( df/dx \) when it is clearly understood that \( f \) is a function of \( x \). The symbol \( \frac{d}{dx} \) is a differential operator which takes a function of \( x \) and returns the derivative with respect to \( x \). (Thus \( \frac{d}{dx} \) takes a function as its input, and returns a function as its output.) When we are interested in position and velocity, we can write \( s'(t) = \frac{ds(t)}{dt} \), or when it is understood that \( s = s(t) \), we might write

\[
v = \frac{ds}{dt}.
\]

(4.9)

Notice that the notation resembles difference quotients, because, with our definition of derivatives, we have

\[
\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{\Delta f(x)}{\Delta x},
\]

\[
\frac{ds}{dt} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t}.
\]

Similarly for any such related quantities. With Leibniz notation our power rule becomes:

\[
\frac{d}{dx} (x^n) = nx^{n-1}.
\]

(4.10)

If we would like to compute \( f'(a) \), i.e., the derivative at a particular point, in the Leibniz notation we use a vertical line which is read, “evaluated at,” as in

\[
f'(a) = \frac{d}{dx} f(x) \bigg|_{x=a}.
\]

So, for example, \( \frac{dx^5}{dx} = 5x^4 \), and the slope at \( x = 1 \) of the function \( f(x) = x^5 \) is given by

\[
f'(1) = 5 \cdot 1^4 = 5, \quad \text{or} \quad \frac{dx^5}{dx} \bigg|_{x=1} = 5x^4 \bigg|_{x=1} = 5 \cdot 1^4 = 5.
\]

Note how the Leibniz notation is often assumed to act like a fraction: \( \frac{d}{dx} (x^5) = \frac{dx^5}{dx} \). However the \( d \) in the numerator, and (separately) the \( dx \) in the denominator are treated as inviolable; we do not ever break those terms up further.

---

\(^{13}\)Named for Gottfried Wilhelm Leibniz (July 1, 1646–November 14, 1716), a German mathematician and philosopher. Most credit him and English philosopher, mathematician, physicist and theologian Sir Isaac Newton, (December 25, 1642–March 20, 1727) with independently discovering calculus. Much is written about rivalries between the “Newton camp” and the “Leibniz camp,” regarding who discovered what first. Newton’s notation for derivative used a dot above the function, as in \( \frac{ds}{dt} = s' \), a notation still used in some physics textbooks.

\(^{14}\)Often the “\( x = \)” is omitted when the variable is obvious, as in \( \frac{dx^5}{dx} \bigg|_{x=1} = 5x^4 \bigg|_{x=1} = 5 \cdot 1^4 = 5.\)
Note the flexibility of the Leibniz notation in the following:

\[
\frac{dx^3}{dx} = 3x^2, \quad \frac{du^3}{du} = 3u^2, \quad \frac{dt^3}{dt} = 3t^2.
\]

These are actually the same rule (with different variables): that the cube of a quantity changes with respect to that quantity at the (instantaneous) rate of 3 times the square of the quantity, be it \(x\), \(u\) or \(t\). Put another way, if the horizontal axis is given by \(t\), and we graph the height \(t^3\) on the vertical axis, then the slope is always 3\(t^2\).

The Leibniz notation also keeps us from making the mistake of trying to use the derivative rules (such as the power rule) to compute, for example, \(\frac{du^3}{dx}\). Since the variables (\(u\) and \(x\)) do not match, the power rule cannot be used directly.\(^{15}\)

With the power rule (4.10) and a few other results we can quickly calculate the derivatives of polynomials. Much of this chapter will be devoted to calculating derivatives using known rules, which save an enormous amount of time when compared to calculating derivatives using limits of difference quotients as in the previous section.

### 4.2.3 Sum and Constant Derivative Rules

**Theorem 4.2.2 (Sum Rule)** Suppose \(\frac{d}{dx} f(x)\) and \(\frac{d}{dx} g(x)\) exist. Then

\[
\frac{d}{dx} (f(x) + g(x)) = \frac{d}{dx} f(x) + \frac{d}{dx} g(x). \tag{4.11}
\]

In other words the derivative of a sum is the sum of the respective derivatives. Before we give the proof, it is worth mentioning that some texts write this using the prime notation:

\[(f + g)' = f' + g'.\]

**Proof:** Assume that \(\frac{d}{dx} f(x)\) and \(\frac{d}{dx} g(x)\) exist at a particular \(x\). Then

\[
\frac{d}{dx} (f(x) + g(x)) = \lim_{\Delta x \to 0} \frac{(f(x + \Delta x) + g(x + \Delta x)) - (f(x) + g(x))}{\Delta x}
\]

\[
= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x) + g(x + \Delta x) - g(x)}{\Delta x}
\]

\[
= \lim_{\Delta x \to 0} \left( \frac{f(x + \Delta x) - f(x)}{\Delta x} + \frac{g(x + \Delta x) - g(x)}{\Delta x} \right)
\]

\[
= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} + \lim_{\Delta x \to 0} \frac{g(x + \Delta x) - g(x)}{\Delta x}
\]

\[
= \frac{df(x)}{dx} + \frac{dg(x)}{dx}, \quad \text{q.e.d.}
\]

The reason that we could break this into two limits legitimately is because the two limits both existed and were finite by assumption. (See Theorem 3.9.1, page 266.)

\(^{15}\)Later in the text we will have the chain rule, which helps us get around the problem of computing \(\frac{du}{dx}\), and similar derivatives where the variable in the numerator does not match the variable of the denominator, i.e., the variable of the differential operator (here \(d/dx\)). There we will see some of the true power of the Leibniz notation, as we compute for instance

\[
\frac{du^3}{dx} = \frac{du^3}{du} \cdot \frac{du}{dx} = 3u^2 \cdot \frac{du}{dx}
\]

Notice how we apparently multiplied and divided by \(du\) to achieve the second expression.
Example 4.2.2 \[ \frac{d}{dx}(x^3 + x^2 + x) = \frac{dx^3}{dx} + \frac{dx^2}{dx} + \frac{dx}{dx} = 3x^2 + 2x + 1. \]

With practice, one learns to skip the first step in the example above. Note how \( \frac{dx}{dx} = 1 \), as one might hope. This reflects the fact that \( x \) and \( x \) change at the same rate (i.e., the ratio of their rates of change is always 1). Put another way, the slope of the line \( y = x \) is always 1.

The next theorem is usually given separately for emphasis.

Theorem 4.2.3 \textit{The derivative of a constant is zero; if a function is defined by \( f(x) = C \) for all \( x \in \mathbb{R} \), where \( C \) is some fixed constant, then \( f'(x) = 0 \) for all \( x \in \mathbb{R} \).} Written two different ways, we thus have:

\[ f(x) = C \implies f'(x) = 0; \quad (4.12) \]
\[ \frac{d}{dx}C = 0. \quad (4.13) \]

There are several ways to see this. From the difference quotient limit definition (4.3) (page 296), regardless of \( \Delta x \neq 0 \) the difference quotient \( \frac{f(x + \Delta x) - f(x)}{\Delta x} = \frac{C - C}{\Delta x} = 0/\Delta x = 0 \), so it remains zero in the limit. From another perspective regarding what we know about lines we have that \( f(x) = C \) is a line of slope \( m = 0 \). From a qualitative standpoint this theorem is reasonable since constants have rate of change zero (hence the term \textit{constant}) with respect to \( x \). Some texts write\(^{16} \) \( (C)' = 0 \). With this theorem we can write, for example,

\[ \frac{d}{dx}(x^3 + 28) = \frac{d}{dx}(x^3) + \frac{d}{dx}(28) = 3x^2 + 0 = 3x^2. \]

With little or no practice one learns to write simply \( \frac{d}{dx}(x^3 + 28) = 3x^2 \).

We need just one more result before we can find derivatives of arbitrary polynomials. This answers the question of what to do with the coefficients of a polynomial, and multiplicative constants in general.

Theorem 4.2.4 \textit{Multiplicative constants are preserved in the derivative. In other words,}

\[ \frac{d}{dx}(C \cdot f(x)) = C \cdot \frac{d}{dx}f(x). \quad (4.14) \]

The proof is left as an exercise. It follows from the fact that multiplicative constants “go along for the ride” in limits as well. (See again Theorem 3.9.1, page 266.) For a simple example, we have

\[ \frac{d}{dx}(5x^7) = 5 \cdot \frac{d}{dx}(x^7) = 5 \cdot 7x^6 = 35x^6. \]

Again, with very little practice one learns to compute such a derivative in one step. Note how the derivative operator \( \frac{d}{dx} \) treats additive constants (which do not survive) differently from

\(^{16}\)One weakness of Taylor’s “prime” notation is that we do not know what variable we are taking the derivative with respect to. For instance, in an earlier example we have fuel volume \( V \) as a function of position \( s \), and so \( dV/ds \) measured the flow rate of fuel per mile. However, since \( s = s(t) \), we have \( V = V(s(t)) \), so ultimately \( V = V(t) \), i.e., \( V \) can be written as a (algebraically different) function of \( t \) instead, in which case we can calculate \( dV/dt \), measuring the flow rate with respect to time. So when asked to calculate \( V' \), or even \( V'(5) \), there is this ambiguity which is not present in the Leibniz notation.

If one wrote \( V'(s) \), it probably would be understood to be \( dV/ds \) and not \( dV/dt \). Similarly, \( V'(5 \text{ seconds}) \) would be understood to mean \( dV/dt \) evaluated at \( t = 5 \) seconds.
multiplicative constants (which do survive). We can combine the power rule (4.10), (4.13), and (4.14) to quickly compute the derivative of any given polynomial (where \( n \in \mathbb{N} \)):

\[
\frac{d}{dx} \left( a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 \right) = a_n \cdot nx^{n-1} + a_{n-1} \cdot (n-1)x^{n-2} + \cdots + a_2 \cdot 2x + a_1.
\]

To be clear on the logic, note that we first use the sum rule to break this into a sum of derivatives of the \( a_k x^k \), \( k = 1, \cdots, n \) and \( a_0 \), calculating the derivatives of the \( a_k x^k \) terms in turn, each time using the fact that the multiplicative constants \( a_k \) are along for the ride in what are otherwise simple power rules: \( \frac{d}{dx} (a_k x^k) = a_k \cdot k x^{k-1} \). The final term \( a_0 \) is an additive constant with derivative zero and thus does not appear on the right hand side of (4.15).

**Example 4.2.3** To see how (4.15) can be carried out quickly, we list a couple of brief examples:

\[
\begin{align*}
\frac{d}{dx} (5x^4 + 9x^2 + 13x + 47) &= 5 \cdot 4x^3 + 9 \cdot 2x^1 + 13 \cdot 1 + 0 = 20x^3 + 18x + 13. \\
\frac{d}{dx} (9 - 6x + 5x^{11}) &= 0 + (-6) \cdot 1 + 5 \cdot 11x^{10} = -6 + 55x^{10}.
\end{align*}
\]

One learns quickly to “think” but not necessarily write the first computational step in such problems. Note that the negative sign also “goes along for the ride,” since it is just a factor of \(-1\). In fact we could list a difference rule,

\[
\frac{d}{dx} (f(x) - g(x)) = \frac{d}{dx} f(x) - \frac{d}{dx} g(x),
\]

but that would be redundant given the sum rule, and how a multiplicative constant, even if negative, is preserved in the derivative.

We need to also point out that to use (4.15), we need to have the function written in the form of the left hand side of that equation, i.e., expanded and not left factored. Consider for instance the following example.

**Example 4.2.4** \( \frac{d}{dx} [(x^2 + 1)^2] = \frac{d}{dx} [x^4 + 2x^2 + 1] = 4x^3 + 4x. \)

In the above we needed to multiply out the polynomial. Thus \( \frac{d}{dx} [(x^2 + 1)^2] \neq 2(x^2 + 1)^1 \), since we are taking the derivative with respect to \( x \) and not \( (x^2 + 1) \).

It is also important to note that derivatives do not allow variable quantities to “go along for the ride.” Thus \( \frac{d}{dx} [x \cdot x^3] \neq x \cdot \frac{d}{dx} [x^3] \). Indeed, \( x \cdot \frac{d}{dx} x^3 = x \cdot 3x^2 = 3x^3 \), while \( \frac{d}{dx} [x \cdot x^3] = \frac{d}{dx} x^4 = 4x^3 \).

Finally we point out again that it should be clear from the previous examples that using (4.15) for such computations is much simpler than using the original definition of the derivative (letting \( \Delta x \rightarrow 0 \) in a limit of difference quotients as in (4.3), page 296) to calculate derivatives of polynomials.

### 4.2.4 Increasing and Decreasing Functions; Graphing Polynomials

Recall that while \( f(x) \) gives the height of the graph \( y = f(x) \) at a particular value of \( x \), the derivative \( f'(x) \) gives the slope there. If the slope is positive the graph is “sloping upwards”; if negative the graph is “sloping downwards.” Another way to speak of such things is to discuss functions which are increasing or decreasing on an interval, say \((a, b)\).
Definition 4.2.2 Consider a function \( f(x) \) with an interval \((a, b)\) contained within its domain.

1. We say \( f(x) \) is increasing on \( (a, b) \) if and only if \((\forall x, y \in (a, b))(x < y \iff f(x) < f(y))\).

2. We say \( f(x) \) is decreasing on \( (a, b) \) if and only if \((\forall x, y \in (a, b))(x < y \iff f(x) > f(y))\).

(Note that it is possible that a function is not consistently increasing or consistently decreasing on a given interval.)

Clearly, for an increasing function on \( (a, b) \), the height increases as \( x \) increases through the interval. Similarly, for a decreasing function on \( (a, b) \), the height decreases as \( x \) increases through the interval. If we know exactly where a function is increasing, and where it is decreasing, that information can be of great help in plotting or analyzing the function. To see what this has to do with derivatives we state the following theorem. Its proof relies on the Mean Value Theorem which will be introduced in a later section. However, it should already have the ring of truth given what we know of derivatives and slopes.

Theorem 4.2.5 Suppose \( f(x) \) is defined for \( x \in (a, b) \), and \( f'(x) \) exists for \( x \in (a, b) \). Then

1. \((\forall x \in (a, b))(f'(x) > 0) \implies f(x) \text{ is increasing on } (a, b)\);

2. \((\forall x \in (a, b))(f'(x) < 0) \implies f(x) \text{ is decreasing on } (a, b)\).

(Again, if \( f'(x) \) changes sign on \( (a, b) \), then neither of these hold.)

Example 4.2.5 To see how we might use this to graph polynomials, consider the graph of the function \( f(x) = x^3 - 3x \). This function is continuous on all of \( \mathbb{R} = (-\infty, \infty) \). Also notice that

\[
\begin{align*}
\lim_{x \to -\infty} f(x) &= \lim_{x \to -\infty} x^3 \left(1 - \frac{3}{x^2}\right) \xrightarrow{x \to -\infty} -\infty, \\
\lim_{x \to \infty} f(x) &= \lim_{x \to \infty} x^3 \left(1 - \frac{3}{x^2}\right) \xrightarrow{x \to \infty} \infty.
\end{align*}
\]

If we draw a sign chart for \( f(x) \), showing where the function is positive and where it is negative, we can get some idea of what the graph looks like. To construct a sign chart for any function we look at all the possible points where the function can change signs. Recall that the Intermediate Value Theorem (Corollary 3.3.2, page 196) implies a function \( f(x) \) can only change signs, as we increase \( x \), by either passing through zero height or having a discontinuity. Since our particular \( f(x) \) here is continuous on all of \( \mathbb{R} \), we look to where \( f(x) = 0 \) to divide \( \mathbb{R} \) into intervals of constant sign. Now \( f(x) = x^3 - 3x = x(x^2 - 3) \) is zero for \( x = 0, \pm\sqrt{3} \). This gives us four intervals on which \( f(x) \) does not change signs. We can test for the sign of \( f(x) \) at a single point in each interval to get the sign of \( f(x) \) on that interval. Doing so as we did in Section 3.3, we construct the sign chart for \( f(x) \):

<table>
<thead>
<tr>
<th>Function: ( f(x) = x(x^2 - 3) )</th>
<th>Test ( x )</th>
<th>-10</th>
<th>-1</th>
<th>1</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sign Factors: ( \ominus \oplus )</td>
<td>( \ominus \oplus )</td>
<td>( \ominus \oplus )</td>
<td>( \oplus \ominus )</td>
<td>( \oplus \ominus )</td>
<td></td>
</tr>
<tr>
<td>Sign ( f(x) ): ( \ominus )</td>
<td>-( \sqrt{3} )</td>
<td>( \oplus )</td>
<td>0</td>
<td>( \ominus )</td>
<td>( \sqrt{3} )</td>
</tr>
</tbody>
</table>
Figure 4.5: Rough graph of \( f(x) = x^3 - 3x \) based upon its sign chart and behavior as \( x \to \pm \infty \). In particular we do not know the exact locations of the local maximum(s) or minimum(s) without investigating the derivative of \( f(x) \).

From the sign chart and the behavior as \( x \to \pm \infty \) we can get some idea of what the graph of \( f(x) \) looks like. That information is reflected, however imprecisely, in Figure 4.5. A serious drawback to such a graph is that we know from the Extreme Value Theorem (Corollary 3.3.1, page 196) that there will be a value in \([-\sqrt{3}, 0]\) which is a local maximum, and another in \([0, \sqrt{3}]\) which is a local minimum, but we do not know exactly where these are from the sign chart of the function (we will formally define the boldface terms shortly). However, a sign chart for the derivative of \( f(x) \) can possibly give us this information.

Since \( f(x) = x^3 - 3x \), it follows quickly that \( f'(x) = 3x^2 - 3 \). Recall that on intervals where \( f' > 0 \), the function \( f \) is increasing, while on those intervals on which \( f' < 0 \), the function is decreasing. Since \( f'(x) \) is also an easily factored polynomial, constructing its sign chart is easy. Note \( f'(x) = 3x^2 - 3 = 3(x - 1)(x + 1) \) is zero exactly where \( x = \pm 1 \).

<table>
<thead>
<tr>
<th>Test</th>
<th>( x = )</th>
<th>( -2 )</th>
<th>( 0 )</th>
<th>( 10 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sign ( f'(x) = )</td>
<td>( \oplus \oplus \oplus )</td>
<td>( \oplus \oplus \oplus )</td>
<td>( \oplus \oplus \oplus )</td>
<td></td>
</tr>
<tr>
<td>Sign ( f'(x) ):</td>
<td>INC</td>
<td>DEC</td>
<td>INC</td>
<td></td>
</tr>
<tr>
<td>Behavior of ( f(x) ):</td>
<td>( \nearrow )</td>
<td>( \searrow )</td>
<td>( \nearrow )</td>
<td></td>
</tr>
</tbody>
</table>

Here we used “INC” to abbreviate increasing, which we also signified by the arrow pointing upwards (\( \nearrow \)), and we used “DEC” and (\( \searrow \)) to signify decreasing. From this we deduce that we get a local maximum at \((-1, f(-1)) = (-1, 2)\), and a local minimum at \((1, f(1)) = (1, -2)\). These two bits of information allow us to draw a more accurate sketch of the graph of \( f(x) = x^3 - 3x \), as illustrated in Figure 4.6. That graph is computer-generated, but we can get a very accurate picture of the function’s general behavior by plotting the information we have gathered: the sign
4.2. FIRST DIFFERENTIATION RULES; LEIBNIZ NOTATION

Figure 4.6: Partial graph of $f(x) = x^3 - 3x$ showing the sign of $f(x)$, the limiting behavior as $x \to \pm\infty$, and the sign of $f'(x)$ (which indicates also the locations of local extrema). The $x$-intercepts (where $f(x) = 0$), the local maximum and local minimum points are also illustrated, as are the facts that $x \to \infty \implies f(x) \to \infty$, and $x \to -\infty \implies f(x) \to -\infty$.

It is important to distinguish the meanings of a sign chart for $f(x)$, and one for $f'(x)$. The former just tells us where the function is below or above the $x$-axis; the latter tells us where the function is increasing and where the function is decreasing.

In the above we used the following terms, which we now define:

**Definition 4.2.3** Given a function $f(x)$.

1. We call a point $x_0$ a **local maximum** of $f(x)$ if and only if

   $$\exists (a,b) \ni x_0 \forall x \in (a,b) \quad (f(x) \leq f(x_0)).$$

   (4.16)

2. We call a point $x_0$ a **local minimum** of $f(x)$

   $$\exists (a,b) \ni x_0 \forall x \in (a,b) \quad (f(x) \geq f(x_0)).$$

   (4.17)

In other words, $x_0$ is a local maximum of $f(x)$ if there is an open interval containing $x_0$ in which the function is never greater than $x_0$ on that interval. Local minimum is defined analogously. If $f(x)$ is continuous in an open interval around $x_0$, and $f'$ exists in that interval, then a change of signs of $f'$ at $x_0$ indicates one of these local extrema. If, for instance, $f' > 0$ to the left of $x_0$ and $f' < 0$ to the right, then $f$ increases before and decreases after $x_0$, making $x_0$ a local maximum. This can be seen in the derivative sign chart and graph above for our example function $f(x) = x^3 - 3x^2$.

**Example 4.2.6** Use the derivative to determine where the graph of $f(x) = x^4 - 6x^2 + 8x$ is increasing, and where it is decreasing. Use this information to sketch a graph of $y = f(x)$.

---

17It is also worth noticing that $f'(x) = 3(x + 1)(x - 1) \to \infty$ for both $x \to \infty$ and $x \to -\infty$, and so the slope of $f(x)$ grows larger as $x \to \pm\infty$. This is not the case with all graphs (see Figure 4.3, page 299 for example), but it is a nice feature to notice when plotting a graph such as Figure 4.6 above.
Solution: We wish to know where \( f'(x) > 0 \) and \( f'(x) < 0 \), so we compute this derivative and construct its sign chart.

\[
f'(x) = \frac{d}{dx} [x^4 - 6x^2 + 8x]
\]

\[
\Rightarrow f'(x) = 4x^3 - 12x + 8
\]

\[
\Rightarrow f'(x) = 4(x^3 - 3x + 2).
\]

To construct the sign chart we need to solve \( x^3 - 3x + 2 = 0 \). While solving a third-degree polynomial equation can be quite difficult, in this case we are somewhat fortunate that \( x = 1 \) is one solution:

\[
(x - 1) x^2 + x - 2
\]

\[
x^3 - 3x + 2
- x^3 + x^2
\]

\[
x^2 - 3x
- x^2 + x
\]

\[
-2x + 2
2x - 2
0
\]

Hence, we have

\[
f'(x) = 4(x - 1)(x^2 + x - 2)
\]

\[
\Leftrightarrow f'(x) = 4(x - 1)(x + 2)(x - 1)
\]

\[
\Leftrightarrow f'(x) = 4(x - 1)^2(x + 2).
\]

From this we get \( f'(x) = 0 \Leftrightarrow x \in \{-2, 1\} \). We use this for our sign chart:

<table>
<thead>
<tr>
<th>Function: ( f'(x) = 4(x - 1)^2(x + 2) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test ( x )</td>
</tr>
<tr>
<td>Sign ( f'(x) )</td>
</tr>
<tr>
<td>Sign ( f'(x) ):</td>
</tr>
<tr>
<td>Behavior of ( f(x) ):</td>
</tr>
</tbody>
</table>

The graph of \( f(x) \) will clearly have a local minimum at \( x = -2 \). At \( x = 1 \) we have a curious situation where the graph is increasing on the intervals \((-2, 1)\) and \((1, \infty)\), but has slope zero at \( x = 1 \). This is illustrated in Figure 4.7, showing that the curve briefly “levels off” at \( x = 1 \).

In fact, \( f(x) \) here can be said to be increasing on \([-2, \infty)\) and decreasing on \((\infty, -2]\). Whether a function is increasing or decreasing (or neither) is technically a description of its behavior on an interval, not at a particular point. (See Definition 4.2.2, page 313.)

Note also that

\[
\lim_{x \to \pm\infty} f(x) = \lim_{x \to \pm\infty} [x^4 - 6x^2 + 8x] = \lim_{x \to \pm\infty} \left[ x^4 \left( 1 - \frac{6}{x^2} + \frac{8}{x^3} \right) \right] = \lim_{x \to \pm\infty} \left[ x^4 \left( 1 - \frac{6}{x^2} + \frac{8}{x^3} \right) \right] = \infty.
\]
4.2. FIRST DIFFERENTIATION RULES; LEIBNIZ NOTATION

Figure 4.7: Partial graph of \( f(x) = x^4 - 6x^2 + 8x \). From the derivative’s sign chart and form \( f'(x) = 4x^3 - 12x + 8 = 4(x - 1)^2(x + 2) \), it is clear that \( f'(x) = 0 \) at \( x = -2, 1 \), but while the function has a local (indeed “global”) minimum at \( x = -2 \), the graph only levels out at \( x = 1 \), and continues to increase on all of \([-2, \infty)\).

We can easily see one \( x \)-intercept occurs at \( x = 0 \) (since \( f(0) = 0 \) clearly), and from our computer-generated graph or from experimentation and the Intermediate Value Theorem, another occurs just to the right of \( x = -3 \).\(^\text{18}\) Computing also

\[
\begin{align*}
    f(-2) &= (-2)^4 - 6(-2)^2 + 8(-2) = 16 - 24 - 16 = -24, \\
    f(1) &= (1)^4 - 6(1)^2 + 8(1) = 1 - 6 + 8 = 3,
\end{align*}
\]

we get the graph in Figure 4.7.

It is important to note that having \( f'(x) = 0 \) does not imply that there is a local maximum or minimum point there. We should always consider the full sign chart of \( f'(x) \). Local extrema do exist at points of continuity \( x = a \) on the graph if \( f'(x) \) changes sign at \( x = a \). In the above example, \( x = 1 \) is a zero of multiplicity 2 for \( f'(x) \), and so \( f' \) does not change sign there, and there is no local extremum at \( x = 1 \). At \( x = -2 \) we have a zero of multiplicity 1 for \( f'(x) \), and so \( f' \) does change sign there, giving us a local extremum (a minimum in that case).

4.2.5 Derivatives of Sine and Cosine

In this subsection we show how sin \( x \) and cos \( x \) are both differentiable, compute their derivatives, and apply them to functions involving the chain rule. We will prove the following theorem.

**Theorem 4.2.6** The functions sin \( x \) and cos \( x \) are differentiable for all \( x \in \mathbb{R} \), and—when \( x \) is

\(^{18}\)In Section 5.5 we will see how to find \( x \)-intercepts of most functions to as much accuracy as we desire.
measured in radians—their derivatives are given by:

\[
\frac{d}{dx} \sin x = \cos x, \quad \text{and} \quad \frac{d}{dx} \cos x = -\sin x.
\]  

These should seem reasonable given the respective graphs of Figure 4.8. For instance, for the sine curve we have the following data:

\[
x = 0, \quad \frac{\pi}{2}, \quad \pi, \quad \frac{3\pi}{2}, \quad 2\pi
\]

\[
y = \sin x = 0, \quad 1, \quad 0, \quad -1, \quad 0
\]

\[
\frac{dy}{dx} = \cos x = 1, \quad 0, \quad -1, \quad 0, \quad 1
\]

Looking at the graph of \( \sin x \) as drawn in Figure 4.8, the slopes at these points and the values for \( \cos x \) seem at least compatible. Similarly for the cosine curve:

\[
x = 0, \quad \frac{\pi}{2}, \quad \pi, \quad \frac{3\pi}{2}, \quad 2\pi
\]

\[
y = \cos x = 1, \quad 0, \quad -1, \quad 0, \quad 1
\]

\[
\frac{dy}{dx} = -\sin x = 0, \quad -1, \quad 0, \quad 1, \quad 0
\]

We will prove the derivative formula for \( \sin x \), and leave the derivative of \( \cos x \) as an exercise. (The two computations are very similar.)

**Proof** (4.18): The proof is based upon the following:

\[
\sin(\alpha + \beta) = \sin\alpha \cos\beta + \cos\alpha \sin\beta, \quad \lim_{\theta \to 0} \frac{\sin\theta}{\theta} = 1, \quad \lim_{\theta \to 0} \frac{1 - \cos\theta}{\theta} = 0,
\]

which are, respectively, a trigonometric identity, (3.81) from page 270, and (3.86) from page 272. We will use the limit-definition of the derivative, expand using the
4.2. FIRST DIFFERENTIATION RULES; LEIBNIZ NOTATION

formula for \(\sin(\alpha + \beta)\), and rearrange the terms so we can use the trigonometric limits above.

\[
f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\sin(x + \Delta x) - \sin x}{\Delta x}
= \lim_{\Delta x \to 0} \frac{\sin x \cos \Delta x + \cos x \sin \Delta x - \sin x}{\Delta x}
= \lim_{\Delta x \to 0} \frac{\sin x(\cos \Delta x - 1) + \cos x \sin \Delta x}{\Delta x}
= \lim_{\Delta x \to 0} \left[ \sin x \cdot \frac{\cos \Delta x - 1}{\Delta x} + \cos x \cdot \frac{\sin \Delta x}{\Delta x} \right]
= \sin x \cdot 0 + \cos x \cdot 1
= \cos x, \quad \text{q.e.d.}
\]

Now we combine what we know into other examples.

**Example 4.2.7** Find \(f'(x)\) if \(f(x) = x^2 + \sin x - 3 \cos x\).

**Solution:**

\[
f'(x) = \frac{d}{dx} [x^2 + \sin x - 3 \cos x] = 2x + \cos x - 3(-\sin x) = 2x + \cos x + 3 \sin x.
\]

We can also use these derivatives to find where functions involving \(\sin x\) and \(\cos x\) are increasing/decreasing, and thus find any local extrema (that is, local maxima and minima).

**Example 4.2.8** Consider the function \(f(x) = \sin x - \cos x\). Find where \(f(x)\) is increasing and where \(f(x)\) is decreasing, and use this information to plot \(f(x)\).

**Solution:** Here \(f'(x) = \cos x - (-\sin x) = \cos x + \sin x\). Since this is defined and continuous everywhere, we will check where it is zero to detect where it (\(f'(x)\) here) possibly changes signs. The technique below works anytime we are interested in solving \(a \sin x + b \cos x = 0\), where \(a, b \neq 0\):

\[
\cos x + \sin x = 0 \iff \sin x = -\cos x
\iff \frac{\sin x}{\cos x} = -1
\iff \tan x = -1.
\]

The reason we can divide by \(\cos x\) is because there are no solutions where \(\cos x = 0\), because such solutions would require also \(\sin x = 0\), and these cannot be zero simultaneously because (recall) \(\sin^2 x + \cos^2 x = 1\). So we are looking for \(x \in \mathbb{R}\) such that \(\tan x = -1\). This occurs in the second quadrant (if \(x\) represents an angle in standard position) and in the fourth quadrant, with reference angles \(\pi/4\):
Thus we are looking for angles \( x = \frac{3\pi}{4} + n\pi \), where \( n = 0, \pm 1, \pm 2, \pm 3, \cdots \). Now \( f(x) = \cos x + \sin x \) is \( 2\pi \)-periodic, so we can analyze one period to see what the graph should look like. We will use the points \( x = -\pi/4, 3\pi/4, 7\pi/4 \) for our sign chart, and declare the pattern from there:

\[
f'(x) = \cos x + \sin x
\]

<table>
<thead>
<tr>
<th>Test ( x = )</th>
<th>0</th>
<th>( \pi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f'(x) = )</td>
<td>1 + 0</td>
<td>-1 + 0</td>
</tr>
</tbody>
</table>

Sign \( f' \): \( \oplus \) \( \ominus \)
Behavior of \( f \): \( \nearrow \) \( \searrow \)

Because this behavior continues, we see a local maximum at \( (\frac{3\pi}{4}, f(\frac{3\pi}{4})) = (\frac{3\pi}{4}, \sqrt{2}) \), since

\[
f(3\pi/4) = \sin \frac{3\pi}{4} - \cos \frac{3\pi}{4} = \frac{\sqrt{2}}{2} - \frac{-\sqrt{2}}{2} = \frac{\sqrt{2} + \sqrt{2}}{2} = \frac{2\sqrt{2}}{2} = \sqrt{2}.
\]

This local maximum height then repeats every \( 2\pi \) in both (left and right) directions. Similarly, because of the sign chart and the fact that this function (and its derivative and its derivative’s sign chart) repeats every \( 2\pi \), we have a local minimum at, for instance, \( (\frac{7\pi}{4}, f(\frac{7\pi}{4})) = (\frac{7\pi}{4}, -\sqrt{2}) \), which also repeats every \( 2\pi \) in both directions. This function is graphed in Figure 4.9.

**Example 4.2.9** Let \( f(x) = x + \sin x \). Find where \( f(x) \) is increasing and where \( f(x) \) is decreasing.

**Solution:** Here \( f'(x) = 1 + \cos x = 0 \) when \( \cos x = -1 \), which is at \( x = \pm \pi, \pm 3\pi, \pm 5\pi, \cdots \). A partial sign chart is given below:
4.2. FIRST DIFFERENTIATION RULES; LEIBNIZ NOTATION

\[
\sqrt{2} - \sqrt{2} - \frac{\pi}{4} - \frac{\pi}{4} - \frac{5\pi}{4} + \frac{7\pi}{4}
\]

Figure 4.9: Partial graph of \( f(x) = \sin x - \cos x \), showing for instance the local minima at \( x = 7\pi/4 \) and \( x = -\pi/4 \), and the local maxima at \( x = 3\pi/4 \) and \( x = -5\pi/4 \). Each local extremum is repeated every \( 2\pi \). See Example 4.2.8.

\[
f'(x) = 1 + \cos x
\]

Test \( x = -2\pi \) \( 0 \) \( 2\pi \) \( 4\pi \)

\[
\text{sign } f'(x): \quad \oplus \quad -\pi \quad \oplus \quad \pi \quad \oplus \quad 3\pi \quad \oplus
\]

So this function is actually always increasing, only briefly having zero slope at the odd multiples of \( \pi \). Note that these points occur at \((\pi, \pi), (3\pi, 3\pi), (7\pi, 7\pi)\), etc., and \((-\pi, -\pi), (-3\pi, -3\pi), (-7\pi, -7\pi)\), etc. This function is graphed in Figure 4.10, showing this behavior.

4.2.6 Absolute Values and Piece-wise Defined Functions

Here we begin by exploring how one would compute derivatives for piece-wise defined functions such as \( \frac{d}{dx}|x| \). This is indeed a piece-wise defined function, as we have

\[
|x| = \begin{cases} 
  x & \text{if } x \geq 0, \text{ i.e., } x \in [0, \infty) \\
  -x & \text{if } x < 0, \text{ i.e., } x \in (-\infty, 0).
\end{cases}
\Rightarrow \frac{d|x|}{dx} = \begin{cases} 
  1 & \text{if } x > 0, \text{ i.e., } x \in (0, \infty) \\
  -1 & \text{if } x < 0, \text{ i.e., } x \in (-\infty, 0).
\end{cases}
\]

If we are safely “within” one of the “pieces,” then the derivatives are easily computed using other rules. For instance, if \( f(x) = |x| \), then

\[
f'(5) = \left. \frac{d}{dx}(x) \right|_{x=5} = 1, \\
f'(-3) = \left. \frac{d}{dx}(-x) \right|_{x=-3} = -1.
\]

We can do this, because when we look at the definition of the derivative, with \( |\Delta x| \) small enough, \( x + \Delta x \) will still be within the interior of the interval \([0, \infty)\) or \((-\infty, 0)\). However, \( f'(0) \) will not
CHAPTER 4. THE DERIVATIVE

Figure 4.10: Partial graph of \( f(x) = x + \sin x \). The derivative being \( f'(x) = 1 + \cos x \), which is positive except at \( x = \pm \pi, \pm 3\pi, \pm 5\pi, \ldots \), the function is always increasing, momentarily “leveling off” at these points where \( f'(x) = 0 \).

For \( f'(0) \) to exist, these two limits would have to be equal (and thus equal to \( f'(0) \)), but these limits are different and so \( f'(0) \) does not exist.

However, the above process is somewhat more involved than necessary. We can simply note that the left-side limit (\( \Delta x \to 0^- \)) would yield the same as \( \frac{d}{dx}(-x) = -1 \) and the right-side limit would yield the same as \( \frac{d}{dx}(1) = 1 \), which are not equal and so the

**Example 4.2.10** Find \( f'(x) \) if \( f(x) = \begin{cases} 1, & \text{if } x \geq \pi/2, \\ \sin x, & \text{if } -\pi/2 < x < \pi/2, \\ -x, & \text{if } x \leq -\pi/2. \end{cases} \)

**Solution:** Note first that \( f(x) \) is continuous at \( x = \pi/2 \), so there is a possibility of a derivative there. It is not continuous at \( x = -\pi/2 \), as a quick check of left and right limits there would show they do not match. We can therefore say that \( f'(\pi/2) \) does not exist. That said, we do at least have

\[
\begin{align*}
x \in (\pi/2, \infty) & \implies f'(x) = \frac{d}{dx}(1) = 0, \\
x \in (-\pi/2, \pi/2) & \implies f'(x) = \frac{d}{dx}(\sin x) = \cos x, \\
x \in (-\infty, -\pi/2) & \implies f'(x) = \frac{d}{dx}(-x) = -1.
\end{align*}
\]
What is left to check is $f'(\pi/2)$. This is simpler than one might think at first, because when we look at the definition of $f'(\pi/2)$ using limits, when computing this with $\Delta x \to 0^+$ it would be exactly the same computation as if the function in question were the (constant) function 1, and so that limit would be 0. When taking $\Delta x \to 0^-$, that limit computation would be exactly the same as if the function were $\sin x$, and we know its derivative is $\cos x$, and $\cos(\pi/2) = 0$. Thus the left and right limits can be found using the functions in the branches which define $f(x)$, assuming that the function is continuous there. (Otherwise some difficulty would be found when trying to take the whole limit as $\Delta x \to 0^\pm$.)

Summarizing,

$$f'(x) = \begin{cases} 
0, & \text{if } x \geq \pi/2, \\
\cos x, & \text{if } -\pi/2 < x < \pi/2, \\
-1, & \text{if } x < -\pi/2, 
\end{cases}$$

and does not exist otherwise (i.e., at $x = -\pi/2$).

Functions such as these indicate why it is important to recall the original definition of $f'(x)$, namely

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

To make the argument above more formally, one could write,

$$\lim_{\Delta x \to 0^+} \frac{f(\pi/2 + \Delta x) - f(\pi/2)}{\Delta x} = \lim_{\Delta x \to 0^+} \frac{1 - 1}{\Delta x} = 0,$$

$$\lim_{\Delta x \to 0^-} \frac{f(\pi/2 + \Delta x) - f(\pi/2)}{\Delta x} = \lim_{\Delta x \to 0^-} \frac{\sin(\pi/2 + \Delta x) - \sin(\pi/2)}{\Delta x} = \cos \frac{\pi}{2} = 0.$$

The final computation came from what we know of the derivative of $\sin x$, which is a two-sided limit and so definitely suffices for the one-sided limit needed there.

We were allowed to use $\sin(\pi/2)$ where we did since this equals 1, and so does $f(\pi/2)$. We were allowed to use $\sin(\pi/2 + \Delta x)$ because for $\Delta x < 0$ (but small), we are on that middle branch of the function’s domain where the function returns the sine of the input. These two limits above being the same, we can declare the limit as $\Delta x \to 0$ to exist and to be that value, i.e., $f'(\pi/2) = 0$.

Similar but simpler computations, not requiring us to do the two one-sided limits separately, yield the derivatives on the interiors of the branches on which $f(x)$ is defined by the different formulas.

A similar but two-sided attempt at $x = -\pi/2$ would have complications which cannot be overcome (impossible already just due to the lack of continuity there—recall the derivative cannot exist without continuity), and so ultimately that limit will not exist and so $f'(-\pi/2)$ does not exist.
Exercises

1. Find the following derivatives. Expand polynomials first where necessary.
   
   \( \frac{d}{dx} [x^2 - 199x + 27] \).
   
   \( \frac{d}{dx} \left[ \frac{1}{2} x^2 + 2x \right] \).
   
   \( \frac{d}{dt} [t^7 - 19t + 10^6] \).
   
   \( \frac{d}{dx} [(x + 9)(x - 3)] \).
   
   \( \frac{d}{dy} [10 - 9y^8] \).
   
   \( \frac{d}{dx} (2x + 5)^2 \).
   
2. Show that
   
   \( \frac{d}{dx} (x^2 \cdot x^3) \neq \left( \frac{d}{dx} (x^2) \right) \left( \frac{d}{dx} (x^3) \right) \).
   
   Why does this not violate Theorem 4.2.4 (i.e., (4.14))?  

3. Give an alternate proof of the integer power rule, Theorem 4.10 by using
   
   \( a^n - b^n = (a - b) \left( \sum_{k=0}^{n-1} a^{n-k} b^k \right) \).
   
   (Hint: \( a^n \) will be your \( f(x + \Delta x) \) term, and \( b^n \) will be \( f(x) \).)

4. Suppose \( s(t) = 3t^2 - 2t + 19 \). Find \( v(t) \). Also find when the particle is moving to the right (\( v > 0 \)), and when it is moving left (\( v < 0 \)).

5. Graph the function \( f(x) = x^4 - 4x^2 \), showing all \( x \)-intercepts, all local maxima and minima. (See Example 4.2.5, page 313.)

6. Use \( \cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta \) to prove (4.19), page 318:
   
   \( \frac{d \cos x}{dx} = -\sin x \).

7. Graph \( f(x) = \sin x + \cos x \) for \( x \in [-2\pi, 2\pi] \), showing where this function is increasing and where it is decreasing.

8. Graph \( f(x) = x + 2 \cos x \) over a reasonable interval, showing where this function is increasing and where it is decreasing. Also show its behavior as \( x \to \pm \infty \).

9. \( 1(f) \) above must be expanded (“multiplied out”) before using the power rule. The answer is \( 8x + 20 \). Compute \( 1(f) \) above using instead the technique in Footnote 15, page 310. (There \( u = 2x + 5 \)).

10. The voltage \( V \) across a resistor in an electrical circuit is the product of the current \( I \) and the resistance \( R \). (This is Ohm’s Law, discussed at length in Section 4.3.) If both the current and resistance vary with time \( t \) and are given by
   
   \( I = 3 + 2t + 0.1t^2, \)
   
   \( R = 20 - 0.2t, \)
   
   find the time rate of change of \( V \) (i.e., \( \frac{dV}{dt} \)) when \( t = 1.5 \). Here \( V \) is in Volts and \( t \) in seconds.

11. Find the general rate of change \( \frac{dV}{ds} \) of the volume of a cube with respect to the length \( s \) of one side, given \( V = s^3 \). Then find the specific rate when \( s = 10 \) cm. What are its units?

12. Find the rate of change \( \frac{dA}{dr} \) of the area of a circle with respect to its radius, recalling that \( A = \pi r^2 \). Does the formula for \( \frac{dA}{dr} \) look familiar? What is its value when \( r = 15 \) cm?
4.3 First Applications of Derivatives

Geometrically, the derivative $f'(x)$ represents the slope of the curve $y = f(x)$. That alone has many uses, but only gives the beginning of the story. The simplest physics example of derivatives is in one-dimensional motion, where we have $s(t)$ giving the position of an object along some line which we can treat as an axis, and its derivative which we denote $s'(t)$, $v(t)$, or $ds/dt$ giving the object’s velocity at time $t$.

The obvious visualizations of these two applications—slope and velocity—of the derivative are different. (See for instance Figure 4.2, page 299 for the slope application (with tangent lines as well), and Figure 4.1, page 295 for illustrations of these.)

It is useful to consider possible applications of the derivative in all the varied contexts, which are legion, and which will have somewhat different intuitions and visualizations that arise naturally. All are in fact unified by the theme of instantaneous change of one quantity with respect to another upon which it depends. That was the spirit of Definition 4.3, page 296 from the first section of this chapter.

In this section we explore several contexts for derivatives. The themes here, and others, will be revisited as we develop rules for computing derivatives of more complicated functions.

4.3.1 Velocity: Vertical and Horizontal

We can revisit our earlier discussion of velocity and other applications in light of the recently introduced differentiation (here meaning derivative computing) rules. For instance, if we have a particle with a position function $s(t) = 6t^2 - 9t + 15$, we immediately get the velocity function:

$$s(t) = 6t^2 - 9t + 15 \implies v = \frac{ds}{dt} = \frac{d}{dt}[6t^2 - 9t + 15] = 12t - 9.$$  

If $s$ is in meters and $t$ in seconds, then $v = \frac{ds}{dt}$ is in meters/second.

If instead, $s(t) = \sin t$, then $v(t) = \frac{ds(t)}{dt} = \frac{d}{dt}\sin t = \cos t$, so a sinusoidal motion gives a similar periodic velocity.\(^{19}\) If $s$ is in meters and $t$ in seconds, then a more explicit formula for $s$ would be $s(t) = (\sin(t/\sec))$ meters, and then $\frac{ds}{dt}(\cos(t/\sec))$ meters/second, though many texts would simply write $s(t) = \sin t$, explain $s$ is in meters and $t$ in seconds, and assume the reader understands where the units would appear in the more explicit formula, and how they carry over in the derivative. In short, anytime we know position $s(t)$, we can compute velocity $v(t) = \frac{ds(t)}{dt}$.

So far all velocities we have looked at were one-dimensional. If we care to consider two-dimensional velocity—velocity in the $xy$-plane for instance—we would look at cases where position would be given by $(x(t), y(t))$, and then we can look separately at how horizontal position changes with time, and how vertical position (height) changes with time to have the horizontal and vertical velocities:

$$\begin{align*}
x(t) &= \text{horizontal position at time } t, \\
y(t) &= \text{vertical position at time } t.
\end{align*} \implies \begin{align*}
x'(t) &= \text{horizontal velocity at time } t, \\
y'(t) &= \text{vertical velocity at time } t.
\end{align*}$$

These can indeed be analyzed separately, though the total motion is of course important. Where a motion is clearly horizontal we may opt to use $x(t)$ instead of $s(t)$, and when clearly vertical we may opt for $y(t)$. If motion is a combination of both, we will usually use $x(t)$ and $y(t)$, respectively.

\(^{19}\)Any curve which has the same shape as a sine curve is usually referred to as sinusoidal. A cosine curve can be considered to be a sine curve that has been shifted horizontally and so “sinusoidal” can describe either, though if an author or speaker used the term “cosinusoidal” the meaning would be clear enough.
Example 4.3.1  From ground level, an object thrown straight up at 96 ft/sec at time $t = 0$ (in seconds) will have vertical position if feet given by $y(t) = 96t - 16t^2$, until it hits the ground.

(a) Find how long it will be in the air.

(b) Starting at $t = 0$, find its velocity at each subsequent second until it hits the ground.

Solution: The motion of this object is purely vertical, since $x(t)$ is constant. Part (a) is an algebra problem, where we solve for $t$ when $y(t) = 0$:

$$0 = 96t - 16t^2 \iff 0 = 16t(6 - t) \iff t \in \{0, 6\}.$$

For Part (b) we compute $\frac{dy}{dt} = \frac{d}{dt} (96t - 16t^2) = 96 - 32t$. From this we can easily make a chart:

<table>
<thead>
<tr>
<th>$t$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{dy}{dt}$</td>
<td>96</td>
<td>64</td>
<td>32</td>
<td>0</td>
<td>-32</td>
<td>-64</td>
<td>-96</td>
</tr>
</tbody>
</table>

Note that when $\frac{dy}{dt} > 0$ the object is rising ($y$ is increasing with $t$), and when $\frac{dy}{dt} < 0$ the object is falling. Similarly when $\frac{dx}{dt} > 0$ motion is to the right, and when $\frac{dx}{dt} < 0$ motion is to the left.

Also note that we can deduce the maximum height of the object from the velocity, if we observe that it will reach its maximum height at the moment it is neither rising nor falling, i.e., that moment when $\frac{dy}{dt} = 0$, which is at $t = 3$, at which time the height is $y(3) = 96(3) - 16(3)^2 = 278 - 144 = 134$, i.e., $y = 134$ ft.

Example 4.3.2  An object is thrown horizontally at time $t = 0$ off of a 200 ft building at 50 ft/sec. How far from the base of the building will it fly, and at what vertical velocity will it hit the ground? The equations for its position before it strikes the ground will be

$$x(t) = 50t,$$

$$y(t) = 200 - 16t^2.$$

Here $x, y$ are in feet and $t$ is in seconds.\footnote{Note that $y(t) = 200 - 16t^2$, if we include all the units, would actually read $y(t) = 200$ ft $- \frac{16t^2}{sec^2}$. This would imply also that $\frac{dy}{dt}$ will be in ft/sec. Units are usually omitted from the equation when it is well understood what they should be.}

Solution: The object will fly until $y(t) = 0$, for which we solve for $t$:

$$200 - 16t^2 = 0 \iff 16t^2 = 200 \iff t = \pm \sqrt{200/16} = \pm \sqrt{2} \cdot 10/16 = \pm \frac{10}{4} \sqrt{2} = \pm \frac{5}{2} \sqrt{2}.$$

Since we do not know anything about the period $t < 0$, and in any event are not interested in such times for this problem, we will conclude that the only relevant algebraic answer here is $t = \frac{5}{2} \sqrt{2} \approx 3.54$. At that time,

$$x \left( \frac{5}{2} \sqrt{2} \right) = 50 \cdot \frac{5}{2} \sqrt{2} = 75 \sqrt{2} \approx 106.$$

We conclude that the object will land approximately 106 ft from the base of the building. We also compute $\frac{dy}{dt} = -32t$, so

$$y' \left( \frac{5}{2} \sqrt{2} \right) = -32 \cdot \frac{5}{2} \sqrt{2} = -80 \sqrt{2} \approx -113,$$

that is, the vertical speed at which the object hits the ground will be approximately 113 ft/sec, though $\frac{dy}{dt} < 0$ when the object hits the ground, so it is (clearly) falling.

It should be noted that the above two examples use formulas for motion which will be developed later, and which assume that air resistance is negligible, and that the ground is flat.
4.3. FIRST APPLICATIONS OF DERIVATIVES

4.3.2 Fluid Flow

The relationship between fluid (in this case liquid) flow and total volume is another which can be described using derivatives. If we have the volume $V$ of a liquid in a tank given as a function of time $t$, then the net fluid flow rate into or out of the tank would be the same (up to a sign $+/-$) as the rate of change of the volume of liquid in the tank with respect to time. If we desire the instantaneous flow rate, that would be the instantaneous rate of change of volume with respect to time, i.e., the derivative of volume with respect to time.

For instance, if the fluid is coming into the tank at 5 ft$^3$/min and none is leaving the tank, then $\frac{dV}{dt} = 5$ ft$^3$/min. If fluid is flowing only out of the tank at 2 ft$^3$/min, then $\frac{dV}{dt} = -2$ ft$^3$/min. Note that a positive rate of change of $V$ with time $t$, as measured by $\frac{dV}{dt}$, means that the fluid volume is increasing within the tank (fluid is, on net, flowing into the tank). On the other hand, a negative $\frac{dV}{dt}$ means fluid is (on net) flowing out of the tank.

Example 4.3.3 Suppose the volume of water in a tank is given by $V(t) = t(10 - t)$, $0 \leq t \leq 10$. Then the volume of liquid in the tank is changing at a rate (for $0 < t < 10$) of

$$\frac{dV}{dt} = \frac{d}{dt} [10t - t^2] = 10 - 2t = 2(5 - t).$$

If $V$ is in gallons and $t$ is in minutes, then $\frac{dV}{dt}$ is in gallons/minute. The maximum volume of the tank occurs at $t = 5$, since before then we have $\frac{dV}{dt} > 0$, while after $t = 5$ we have $\frac{dV}{dt} < 0$. The actual maximum volume is then $V(5) = 5(10 - 5) = 25$. If we were to graph $V(t)$ for this situation, we would produce a picture as in Figure 4.11.

Note from the graph how the volume increases when $\frac{dV}{dt} > 0$, i.e., when $t \in (0, 5)$, and decreases when $\frac{dV}{dt} < 0$, on $t \in (5, 10)$. In fact, one often looks where a derivative is zero to find where a function might be maximized, or for that matter minimized. This will be discussed at length as we continue to discuss derivatives.

There are countless other real-world examples of instantaneous rates of change we can investigate using just the rules of this section. However, we will be much better equipped to pursue
applications after we develop the other differentiation rules in the next sections. Eventually we will tackle many and varied such problems to illustrate how the calculus is applied to real-world questions.

4.3.3 Energy, Work and Power

Work is defined as a force acting over a distance. In the metric system of units, force is defined in Newtons N, and distance in meters m. The work accomplished in pushing an object with 25N over a distance of 4m would then be $25N \cdot 4m = 100N \cdot m$.

While the $N \cdot m$ is an acceptable unit of work, for convenience this was renamed as the joule, written J:\[^{21}\]

$$J = N \cdot m.$$  \hspace{1cm} (4.20)

Thus we could rework our original example and claim that the work performed by using 25N over 4m would be $25N \cdot 4m = 100J$.

Energy is defined as the ability to do work, and quantitatively is considered equal to the work that can be performed by the energy when it is applied. Similarly, work done is the same as the energy spent performing the work. Thus we could also say that 100J of energy was expended in moving the object discussed above.

It is well known that under everyday circumstance—not involving, say, nuclear energy where energy is produced by converting mass to energy—we have energy “conserved,” meaning that it is neither created nor destroyed, but converted to different forms or passed from one object to another. Energy forms include heat, light, electrical, magnetic, kinetic (energy due to motion), sound (acoustic), chemical and nuclear.

For conceptual reasons, mechanical energy is often broadly categorized into two types:

**Potential Energy:** energy an object is capable of making available to do work. This is sometimes called stored energy.

**Kinetic Energy:** the energy an object possesses due to its motion.

We will show later in the textbook that kinetic energy is equal to $\frac{1}{2}mv^2$, where $m$ is the mass of the object and $v$ is its velocity. Therefore in Figure 4.12, page 329 we can actually solve for $v$ when the object hits the ground because the energy put into the object will be $wh$, where $h$ is the height, and the energy it has at the bottom of its path will be $\frac{1}{2}mv^2$, where $w = m \cdot 9.8m/sec^2$. With mass $m$ in kilograms (kg), $w$ will be in kg $\cdot m/sec^2$, or newtons N.

*Power* is a measure of how quickly energy is available (or spent). One could describe power as a rate of “flow” of energy. The unit of power is the joule/second, also known as the watt, written W.\[^{22}\] If, in the example in the above paragraphs, the 100J of work was performed in 2 seconds, then the average power used during those seconds was $(100J)/(2sec) = 50W$.

---

\[^{21}\]Named for James Prescott Joule (1818–1889), and English physicist who discovered the relationship between heat and work.

\[^{22}\]Named for James Watt (1736–1819), a Scottish engineer and inventor, known especially for the unit which bears his name, and for his work in greatly improving steam engines which considerably hastened the Industrial Revolution, as his worked unlocked the potential power of steam-powered engines. While studying power, he developed the concept of horsepower as a unit which, depending upon the underlying definition, varies from between 735W to 750W.
4.3. FIRST APPLICATIONS OF DERIVATIVES

Figure 4.12: Potential and Kinetic Energy Examples. When the weight $w$, in Newtons N, is at height 0 and not moving, it has zero potential and kinetic energy. If work is performed to raise it up 3m where it stops, that work is $3m \cdot w$, with units of joules (J), and it would then have potential energy $3w$ in joules and zero kinetic energy. If it is then allowed to drop, it will release this potential energy in the form of kinetic energy gravity will perform work on it and it will have kinetic energy due to its motion. At the bottom of its path, this object will have given up all of its potential energy to kinetic energy $3w$.

4.3.4 Electrical Current

The relationship between electrical charge and electrical current is analogous to the relationship between volume and fluid flow, with one difference: there are two types of charges, namely negative and positive. This adds some complications because we are not accustomed to discussing negative volumes. Furthermore, in most practical cases the total charge of a system is zero, which means also that we need to be able to quantify charges. The standard unit of charge is the coulomb C, and the standard unit of current is the ampere, or amp A, which is the same as a coulomb/second. (These will be revisited and further explained in subsequent paragraphs.) Electrical phenomena being in many ways more complicated—and often more misunderstood—than simple liquid flow phenomena, some further background is offered here.

What we usually call electricity is the movement of electric charges, usually through a conductor. Putting the charges into motion requires energy, which can then be transmitted through conductors and then harnessed in various ways. The elementary particle which usually carries this electrical energy is the electron, the particle which actually moves through a conductor by jumping from atom to atom or joining one atom and causing it to expel another electron to join another atom farther along the conductor.

When not jumping from one atom to another, an electron is usually found in shells surrounding a core, or nucleus of an atom. Inside a nucleus itself will be one or more protons (several for atoms in a conductor), each of which carries a charge equally large, but opposite to the charge of one electron. The number of protons within the nucleus of an atom will decide which element the atom represents, be it Hydrogen (one proton), Helium (two protons), or so on within the Periodic Table of Elements. In a solid conductor, the nuclei are somewhat fixed to their locations, while electrons in the atoms’ outermost shells are somewhat free to move among the neighboring atoms.

As a general rule, an object with more electrons than protons will have a net negative charge,

---

23The reader will likely note the repeated occurrence of words like “usually” in this subsection. This is because in high-energy cases, charges can be forced—or freed—to behave very differently from how they behave within a conductor, or for that matter within an insulator (the opposite of a conductor). In fact, “conductor” and “insulator” are relative terms: some materials conduct better than others, and some insulate better than others. With enough energy, materials that would normally act as insulators can be forced to conduct electricity, as happens in some lightning phenomena.
where one with more protons than electrons will have a net positive charge. The naturally preferred (stable) state is neutrality, where the number of positive charges equals the number of negative charges and we say there is zero net charge. This is preferred because opposite charges attract while similar charges repel, and so an object with an excess of electrons—and therefore a net negative charge—will tend to repel its own electrons until enough leave that it becomes neutral. Similarly, an object with a deficit of electrons (that is, less electrons than protons) will have a net positive charge and tend to attract electrons in order to achieve neutrality. However, as we will discuss momentarily, there are other tendencies which can overcome this tendency towards neutrality.

One formula which describes the strength of the attraction and repulsion is Coulomb’s Law: \[ F = k \frac{q_1 q_2}{r^2}, \]  
(4.21)
where \( q_1, q_2 \) are two “point charges” (meaning \( q_1 \) is in effect concentrated at one point, and \( q_2 \) likewise at another point), \( r \) is the distance between them, and \( F \) is the force each exerts on the other. Here \( k > 0 \) is some constant which depends upon the units involved and the medium in which the charges reside. Note that if \( q_1, q_2 \) have the same sign, then this force is positive, which signifies that the charges are pushing against each other, where if \( q_1, q_2 \) have opposite signs, the force is negative and they pull towards each other. (The forces do not necessarily cause movement, because there may be some physical barrier or constraint that keeps the charges in place.)

As mentioned previously, the standard unit of charge is the coulomb, C. One coulomb of charge is the charge of approximately \( 6.24 \times 10^{18} \) protons:\[ 1 \text{ C} \approx 6.24 \times 10^{18} \text{ proton charges}, \]  
(4.22)

\(-1 \text{ C} \approx 6.24 \times 10^{18} \text{ electron charges}. \]

The standard unit of current flow in a conductor is coulombs/second, also known as amperes, or amps, denoted A: \[ 1 \text{ A} = 1 \text{ C/sec}. \]  
(4.23)

Electrical current is almost always due to the movement of electrons within a conductor, with the positive nuclei of the atoms within the conductor relatively fixed in their positions. However, due to the convention that these nuclei are assigned a positive charge, and the naïve intuition that the notation suggests charge should move from positive (adding) to negative (subtracting), a concept of conventional current emerged, which follows the progression of positive charge rather than electron progression. In many electronic settings, particularly involving transistors, this...
can more intuitively explain the behaviors of electronic circuits. However, the net movement of positive charge is counter to the movement of negative charges, i.e., counter to the movement of the free electrons. This is better explained with some more background.

First we recall that in most electrically stable states, an atom will have as many protons as electrons, and so the net charge will be zero. If it were to have more electrons than protons, it would have a net negative charge and tend to repel its own electrons until it has a neutral charge; if it had less electrons than protons, it would have a positive charge and tend to attract electrons from elsewhere. However, there are other factors which influence how many electrons an atom will tend to favor in its electron shells. In particular, due to quantum mechanical considerations atoms also have a preference for “filled” electron shells, whereby each shell contains a maximum number of electrons, depending upon the “level” of the shell. One common example of this is ordinary table salt, which is composed of sodium and chlorine: sodium has an extra electron in an outer shell, where chlorine would need one electron to fill its outer shell. The tendency for chlorine to fill this outer shell is quite strong, and can overcome the coulomb forces having an extra electron would create. Sodium’s partially empty, outer electron shell being relatively unstable means that it cannot overcome the tendancy of chlorine to fill its outer shell, and so sodium gives up an electron to chlorine. The result is a positive ion (electrically non-neutral atom) of sodium, and a negative ion of chlorine, which are then attracted together by the coulomb forces. Lattices of sodium and chlorine then create salt chrystals. (It should be pointed out that pure water has very stable electron shells and does not generally conduct electricity, whereas salt water conducts electricity readily, due to the presence of the ions.)

A more thorough discussion of electron behavior at the atomic and molecular level is beyond the scope of this text (the interested reader should research “valence electrons,” or “electron shells”), but for now it should suffice to note that the electron shells of an atom are most stable when the net charge is zero (implying the same number of protons and electrons), and all shells are filled. If an outer shell has only one electron, a relatively small amount of energy can pry it from the atom to roam to other atoms. In doing so it causes the first atom to become positively charged and likely to attract another electron, perhaps from a second atom which had one only weakly bound in its outer shell. In this way charges can move from atom to atom, giving us an electrical current.

For example, a popular conducting material is copper, which has 29 protons, and thus 29 electrons when it is electrically neutral, which is its preferred state absent an outside influence. In such an atom, 28 electrons are bound tightly in full electron shells and one electron is alone in its own shell. Similarly, silver (47 protons) has 46 electrons in filled shells and one lone electron in its own shell to roam, and gold (79 protons) has 78 electrons in filled shells and one in its own shell. Electron shell principles being complicated, silver is actually the best elemental conductor, followed closely by copper, and then gold. For reasons of economy copper is the most popular, though gold is often used for contacts between conductors due to its excellent resistance to corrosion and its ability to be produced in very thin coatings for that purpose. All three materials (and many others, in particular compounds of more than one element) are able to “give up” an electron so it can carry electrical energy to other areas of the conductor.

One upshot of electrons carrying current from atom to atom is that their movement produces a counter movement (conventional current) of positive charge from atom to atom, in the opposite direction. To see this, consider for instance if we have three collinear atoms lying along the x-axis, atoms which we dub, from left to right, Atom 1, Atom 2 and Atom 3. Suppose that there is an object with a negative charge to the left of Atom 1, and one with a positive charge to the right of Atom 3. If the positive charge of the object on the right is strong enough, it could cause Atom 3 to lose one electron, leaving it with a positive charge. Atom 2 would “see” this charge, and perhaps give up an electron to fill the “hole” in Atom 3. At this point, Atom 1 would “see” the positive charge of Atom 2 and could give up an electron to fill the “hole” in Atom 2, and
CHAPTER 4. THE DERIVATIVE

Figure 4.13: Illustration of the motivation for two types of current paradigms: electronic current which follows the movement of electrons, and conventional current which follows the movement of positive charges. As electrons move right, positive charges (holes) left by the electrons move left. Electronic current is movement of electrons, where conventional current is movement of the positive charges.

so Atom 1 would now have a positive charge. Atom 1 might then attract an electron from the negative object to its left to fill its “hole.” In the meantime, Atom 3 might have given its newly acquired electron to the positive object to its right (assuming it still has a positive charge after receiving the first electron from Atom 3), and started the whole process over again. What we observe then, is that

the positive charges actually move to the left, while the electrons (negative charges) move to the right.

As quickly as electrical current in a conductor moves, it matters little if the process was started by “holes” moving from the right, or electrons moving from the left. Nonetheless, the reader should be aware that there are two types of “current” referenced in the literature: the electronic current, i.e., the movement of the electrons, and the conventional current, which is the movement of the positive charge (or positive “influence,” or “electron hole”). This is illustrated in Figure 4.13, which shows the placements of charges along a conductor (drawn as a line connecting atom “circles”) at different times.

A battery is a device for storing electrical energy, and consists of a cathode which is capable of holding a positive charge, an anode which is capable of holding a negative charge, and a separator.27 Battery theory is somewhat complicated by the fact that the chemical composition of the anode and cathode can sometimes change depending upon whether they are charged or neutral, and indeed many battery manufacturers exploit this chemical change to produce batteries which can better hold stored electrical energy (in the form of charges). The interested reader is encouraged to research such things elsewhere. The simplistic explanation is that, when

27The names “cathode” and “anode” correspond to names given to the different charges ions can carry: positive ions are also called cations, and negative ions called anions. Incidentally, the term ion was coined by Michael Faraday, whose name will appear in a later footnote.
fully charged, if the cathode has a positive charge of \( q \), then the anode will have a negative charge of \(-q\). If the battery is placed in a circuit allowing electrons to flow from the anode to the cathode, these charges may be able to perform useful work, and in the process will somewhat neutralize the charges of the anode and cathode, thus partially discharging the battery. Depending upon the chemical change which occurs, the battery may be rechargeable, reversing the process and thus letting the battery again store energy.

Charging occurs when an external supply of electrical energy is applied to the battery in such a way that a positive wire from the charger (or similar source) is attached to the cathode’s terminal, and the negative wire of the charger is attached to the anode’s terminal. Once the battery is charged, a device requiring electrical power can be connected to the cathode and anode, and electrons will tend to flow from the negative anode, through the device, to the positive cathode, so that the anode and cathode can become more neutral, reversing the internal chemical changes that charging caused.

The electrical source charging the battery must be rated at an appropriate voltage, so at this point it is necessary to define voltage.\(^28\) It is a measure of energy per coulomb of electrons, so higher voltage means more energetic electrons, regardless of their number. Total energy then depends upon voltage and the number of electrons. If an electrical supply is rated at 5 volts then each coulomb of electrons carries 5 joules of energy: \( 5V = 5J/C \). Since power, given in watts \( W \), is the measure of how quickly the energy is available or spent, and one watt is equal to one joule per second, we can also write:

\[
1W = 1J/sec = 1J/C \cdot 1C/sec = 1\text{volt} \cdot 1\text{A}.
\]

Thus one watt is also the electrical power available in one coulomb of current operating at one volt. This will be revisited later.

Batteries are fairly efficient storage devices of energy. While coaxing the cathode and anode to take on electrical charge requires some energy, nearly all of that energy can then be returned from the battery through the electricity which it returns when powering devices.\(^29\)

The symbols for charge and current are, respectively, \( Q \) and \( I \), with standard units being coulombs \( C \) for charge \( Q \) and Amperes, or Amps \( A \), for the current \( I \). In the end, the question of current is analogous to our earlier question of fluid flow; the computations acknowledge that current in the wire—that is, out of the battery—reflects the opposite of what is occurring in the battery. In other words, a positive measurement of current from the battery would reflect a negative change in charge:

\[
I = -\frac{dQ}{dt}.
\]  

(4.24)

However, with a battery the electrons move from the anode, through the circuit, and to the cathode so the net charge of the battery does not change. Only the segregation of charges diminishes (or increases, in the case of a charging battery). Also, the flow of electrons (electronic current) is, again, opposite of the flow of positive charges (conventional current). To simplify things, we usually concentrate on one of the battery’s electrodes, and for that matter usually the cathode because the charge there is given as positive.

\(^{28}\)Named after Alessandro Giuseppe Antonio Anastasio Gerolamo Umberto Volta (1745–1827), an Italian physicist generally considered to be the inventor of the first electrochemical cell battery in 1800.

\(^{29}\)A relatively small amount of energy is lost through heat and destructive chemical reactions. Some more is lost because of incomplete discharging of the battery under normal usage; once a battery’s voltage drops significantly due to discharge, it is usually not capable of continuing to power a device which it was designed to power. The relative cost of this inability to completely discharge is mitigated if the battery can survive a high number of recharges.
Example 4.3.4 A battery is connected to a charger in such a way that the total charge of the battery’s cathode is given by \( Q = (t + \sin t) \), where \( Q \) is in coulombs and \( t \) is in seconds. What is the current as measured coming from the battery’s cathode during this charging?

Solution: Current in coulombs/second, i.e., amps, will be given by

\[
I = -\frac{dQ}{dt} = -\frac{d}{dt}(t + \sin t) = -(1 + \cos t).
\]

For instance, the following graph shows some values for \( Q \), \( \frac{dQ}{dt} \) and \( I = -\frac{dQ}{dt} \):

In the above example, note that the battery is, over time, accumulating (positive) charge at the cathode and anode, as the charger is sending electrons into the anode and pulling them from the cathode of the battery. The accumulation is not occurring at a constant rate, so for instance at \( t = \pi \) the charger is sending no charges, and the total charge in the battery’s anode levels off briefly. At \( t = 2\pi \) the charger is pulling electrons at the cathode at a rate of 1A (so they come out of the battery with a negative charge, hence \( I = -1A \)), and we can see the charge \( Q \) building more rapidly at that time, as the slope of \( Q \) is steeper there.

4.3.5 Ohm’s Law

One of the most basic laws in electrical theory is Ohm’s Law.\(^{30}\) It describes how much current \( I \) a voltage of \( V \) can pass through a material which resists the passage of current. There are many reasons to employ standard resistors in a circuit, one reason being to regulate voltages and currents passing through various components. Incandescent light bulbs and heating elements in electrical heaters are basically resistors which rob the flowing charges of some of their energy, and in the process give off energy as light heat.

Resistance is measured in standard units of ohms which, in a bit of a play on words with a certain Greek letter, are given the symbol Ω. The usual form given for Ohm’s law is

\[
V = IR. \tag{4.25}
\]

This can describe a simple circuit such as that drawn in Figure 4.14, page 335. Note that a resistor will be shown as a zig-zagging line, signifying the difficulty with which charges can pass, compared to the straight-line conductors which are assumed to have negligible resistance. Also note the symbol for a battery on the left of the circuit.

\(^{30}\)Named for Georg Simon Ohm (1789–1854), a German physicist and teacher, mostly self-taught, with initial home-schooling from his also self-taught father. The law which bears his name was first found in his 1827 treatise on his complete theory of electricity.
4.3. FIRST APPLICATIONS OF DERIVATIVES

Figure 4.14: Figure showing a circuit with a voltage $V$, in this case coming from a battery, sending current $I$ through the circuit which has resistance $R$. The current which can pass through the resistor (and hence through the whole circuit) is given by $I = V/R$. Ohm’s Law is usually given in the equivalent form $V = IR$. In the figure above, current is drawn as “conventional current,” which proceeds clockwise, while actual electrons are moving counterclockwise.

From the equation, we can see that if a voltage $V$ placed across a circuit with resistance $R$, the circuit will allow $I = V/R$ current to flow. For a fixed voltage, doubling the resistance would halve the current allowed to flow. On the other hand, for a fixed resistance, doubling the voltage would double the current flow.

Note that (4.25) clearly holds when $V, I, R = 1$, that is $V = 1$ volt, $I = 1$ amp and $R = 1$ ohm. Since we can also write $R = V/I$, we can use this example of a circuit with 1 ohm of resistance and 1 volt of voltage causing 1 amp of current to flow, and note therefore that

$$1\text{Ω} = \frac{1\text{ volt}}{1\text{A}}.$$  \hfill (4.26)

i.e., to overcome 1 ohm of resistance requires one volt per amp of current flowing.

Example 4.3.5 If we have a circuit which always has 5 ohms of resistance, but $I = (2 \cos t)\text{A}$, where $t$ is in seconds, then find what must be

1. the voltage $V$ in the circuit, as a function of time;
2. the rate of change of the current with respect to time;
3. the rate of change of the voltage with respect to time.

Solution: When we know we are using standard units, so $V = IR$ here, we have the option of carrying units through our computation, or suppressing them and interpreting the results in the end, with or without units. Here we will include units, though to do so can clutter the computations somewhat.

1. $V = IR = (2 \cos t)\text{A} \cdot 5\Omega = 10(\cos t)\text{AΩ}$, which by (4.26) can be rewritten $V = 10(\cos t)\text{volts}$.
2. Since $I = (2 \cos t)\text{A}$, we can write $\frac{dI}{dt} = \frac{d}{dt}[(2 \cos t)\text{A}] = -2(\sin t)\text{A/sec}$.
3. Similarly, $\frac{dV}{dt} = \frac{d}{dt}[10(\cos t)\text{volts}] = -10(\sin t)\text{volts/seg}$.

It is also perfectly acceptable to perform the above computations without carrying units throughout, just noting that the final answers are in volts, A/sec and volts/sec, respectively. If standard units are used for the variables within the computation, the result will be in appropriate, standard units. However, it is sometimes useful for error detection to use units throughout, though one has to also be aware that the Leibniz notation implies introduction of units as well. For instance, $dV/dt$ will be in units of volts/second.
CHAPTER 4. THE DERIVATIVE

Figure 4.15: Symbol for an AC voltage source represented by a cycle of a sine wave inside of a circle, and a typical graph of the voltage across the terminals. At the beginning of the labeled cycle, there is no voltage, but then the top terminal has an increasing positive voltage, relative to the bottom terminal which therefore becomes increasingly negative. This condition is maximized at one-fourth of the way through the cycle and then decreases until the voltage is again zero half-way through the cycle. Then the top terminal becomes increasingly negative relative to the bottom terminal (now positive), until this negative voltage is maximized three-fourths through the cycle, at which time the voltage begins decreasing (in absolute value) until it is again zero at the end of the cycle.

Example 4.3.6 Find the time rate of change of voltage $V$ across a resistor (as in Figure 4.14) at time $t = 1$ sec if $R = 20 - 0.1t^2$ and $I = 3 + 0.2t^3$.

Solution: Here we again use $V = IR$, and compute

$$V = IR = (3 + 0.2t^3) (20 - 0.1t^2) = 60 + 4t^3 - 0.3t^2 - 0.02t^5$$

$$\Rightarrow \frac{dV}{dt} = 12t^2 - 0.6t - 0.1t^4$$

$$\Rightarrow \frac{dV}{dt} \bigg|_{t=1} = 12 - 0.6 - 0.1 = 11.3$$

with the units being in volts/second.

4.3.6 Electromagnetic Inductance

Two very important discoveries, which revolutionized our technology, are electromagnetism and inductance. First it was discovered that current through a wire causes a magnetic field around the wire (see Footnote 26, page 330). Coiling wire makes the effect more pronounced, and in particular wire coiled around an iron core (rather than around an “air core”) allows the iron atoms to align their magnetic axes to make the effect even further pronounced.

Later it was discovered that if a magnet moves in relationship to a wire (or coil), a voltage is “induced” in the wire.\(^{31}\) This led to many applications, including generators and transformers.

In a generator, coils of wire and a magnetic source are moved relative to each other (by the work done by some external source of energy), thus generating a voltage in the wires. There are

\(^{31}\)This was discovered in 1831–1832 by Michael Faraday (1791–1867), an English physicist and chemist who gave us Faraday’s Law, namely, “The induced electromotive force [voltage] in any closed circuit is equal to the time rate of change [derivative] of the magnetic flux [field] through the circuit.” We will not pursue the most theoretical aspects of magnetic fields, but the upshot is that a moving magnetic field can indeed “induce” a voltage in a wire passing through it. This can be accomplished by having a fixed magnet and a moving wire, a fixed wire and a moving magnet, a fixed coil and a variable electromagnet, or any combination of these.
many designs for these, but for ease of efficiency most give “alternating current,” or AC, output by design. Put simply, the coils or magnets spin in a circular fashion, and since magnets are polarized (one side of each is an “N” pole while the other is a “S” pole), voltage is induced in one direction when the coil moves from near the N to near the S poles, and the other direction when it moves from near the S to near the N poles. Due to the sinusoidal nature of the path of the spinning magnets (or spinning coils), we get a “sine wave” when we graph the induced voltage $V$ against time $t$. Each complete revolutions $N \rightarrow S \rightarrow N$ which occurs is called a cycle, and the standard unit for the rate of these occurring in cycles/second is the hertz, or Hz.$^{32}$

For future reference we digress here briefly to point out that voltage which does not change direction is referred to as “direct current” voltage, or DC voltage, the “current” part referring to the single direction which this voltage can force a current into traveling. Batteries produce DC voltage and currents, as do some generators, though AC generators tend to be more efficient and reliable because they can operate without energy wasting “brushes” and “commutators,” needed to switch the direction of the induced current in the coils as they pass by the different poles of the generator’s magnets. Most electronic circuits ultimately require DC to operate, but AC to DC converters are simple and efficient enough that devices which plug into electrical outlets can easily contain their own converters. The extra efficiencies of AC power generation (and brushless AC motors as well) have resulted in the near extinction of large-scale DC generation.$^{33}$

Another device for exploiting inductance is the transformer, the theory of which is somewhat complicated but we will discuss some of the better known phenomena.

A transformer is basically two or more coils of wire wrapped around the same core, typically iron in power or audio applications, or air in higher-frequency applications. While there are many physical configurations, the basic idea is the same: one coil of the transformer, called the primary coil, produces the moving magnetic field because it is connected to a varying (usually AC) source of voltage. The other coils are called secondary coils, and a voltage is induced in these coils due to the moving magnetic field produced by the changing voltage in the primary.

If for instance the primary coil of a transformer is connected to an AC source, the resulting magnetic field in and around the core is constantly changing. When the voltage rises in one direction in the primary coil, there is a growing magnetic field with a definite N-S orientation, which then collapses and grows again with an opposite (S-N) orientation as the voltage in the coil changes direction a half-cycle later. It is this constant changing of the magnetic field which simulates a moving magnet, or (equivalently) simulates the secondary coil moving through a fixed field, and thus induces a voltage in any secondary coil.

What makes transformers especially useful is that the ratio of windings between the primary and a secondary coil will be the same as the ratio of the voltages present across the terminals of the coils.$^{34}$ For instance, if the primary coil has 100 turns and is connected to a 90-volt AC

---

$^{32}$Named for Heinrich Rudolf Hertz (1857–1894), a German physicist who was the first to demonstrate the existence of electromagnetic waves. His work in radio wave detection and experiments showing how they pass through some objects and bounce off of others, and his work with antennas, he considered to be useful only for verifying equations already known from the very important Treatise on Electricity & Magnetism of James Clerk Maxwell (Scottish physicist and mathematician, 1831–1879), whose equations also inspired Albert Einstein (German-born, later American physicist 1879–1955) to pursue his relativity theories.

In particular Hertz did not foresee implications of his work towards radio or radar.

$^{33}$It was Serbian-American inventor and mechanical and electrical engineer Nikola Tesla (1856–1943) who is credited with developing the first modern AC electric power generating system. Initially working for American inventor Thomas Alva Edison (1847–1931) to improve Edison’s DC generating facilities, Tesla resigned apparently over a broken promise of much higher wages on Edison’s part, working independently on his AC projects (including his famous “Tesla coil”) at first while surviving as a laborer, to eventually work with Edison’s rival George Westinghouse, Jr. (1846–1914) to develop AC power generation systems, which ultimately became standard for most countries.

$^{34}$The usual formula is $N_1/N_2 = V_1/V_2$, where $N_i$ is the number of windings in the $i$th coil, and $V_i$ is the voltage in the $i$th coil.
Primary
Voltage: 90V
(Resistance: 45Ω)
(Current: 2A)
(Wattage: 180W)

Secondary
Voltage: 180V
(Resistance: 180Ω)
(Current: 1A)
(Wattage: 180W)

Figure 4.16: An example of the theory of transformers. If we assume there are twice as many windings in the secondary coil (depicted as the coil on the right of the transformer core) as the primary coil, a 90-volt AC supply (abbreviated 90 V AC) connected to the primary coil induces a 180 V AC supply in the secondary coil. With a resistance of 180Ω, the secondary will flow 180 volts/(180Ω) = 1A of current, and thus 180 volts × 1A = 180W, of power is consumed (or equivalently 180 J/sec of energy). This requires that the same 180W of power flow in the primary (since energy is conserved), and so the primary winding will act as a 45Ω resistor, and the primary circuit will flow 2A of current at 90VAC to achieve its 180W of power. The values in the primary which are dependent upon the action in the secondary’s circuit are shown in parentheses.

source, and the secondary coil has only 50 turns, then the voltage coming from the secondary coil will be 90 volts · (50turns)/(100turns) = 45 volts. On the other hand, a transformer with a 100-turn primary coil connected to a 90-volt AC supply, and a 200-turn secondary coil will induce 90 volts · (200turns)/(100turns) = 180 volts.

One might naively think a transformer can produce new energy by raising the voltage of an AC supply, but this is not the case. If the secondary coil is not connected to any device which can make use of the energy, the electrical energy flowing through the primary is doing nothing useful. If the secondary is experiencing some voltage $V_2$, based upon the primary’s voltage and the ratio of the number of windings, and if the secondary is connected to a device which puts the moving charges to some use, then in the theoretical case of a 100% efficient transformer which loses no energy through heat, vibrations or other losses, then the actual power flowing through the primary will be the same as that flowing out of the secondary (or the sum of all power flows out of all secondaries).

For instance, if the secondary has twice as many windings as the primary, and the primary is a 90-volt AC source, then the secondary will have a 180-volt induced voltage. If the secondary is connected to an object which uses 1A of current, then the secondary will be outputting a quantity of power which we can compute:

$$180 \text{ volts} \cdot 1\text{A} = 180\text{W}.$$  

At the same time, the source supplying voltage to the primary coil will need to send 2A through the primary at 90 volts, to achieve the same power, that is, the same flow of energy:

$$90 \text{ V} \cdot 2\text{A} = 180\text{W}.$$  

So while the secondary has electrons which are twice as energetic, it has only half as many of them passing through in a given time. See Figure 4.16 for an illustration of this scenario.

Electric energy companies use high ratio “step-up” transformers to send their energy over power wires (also known as “lines”) at very high voltages, because there is less “line loss” of energy that way (ultimately due to having fewer electrons but very energetic electrons), and then these companies use “step-down” transformers closer to their clients to bring voltages down to useful levels. Higher voltages require the electrical lines to be farther apart, so that charges will not “arc,” meaning jump across the air gaps or insulators between lines, so there are practical limits to how high these voltages can be raised in the long distance power lines using transformers.
There are many interesting and important phenomena which occur in the context of transformers, besides the voltage multiplication. For instance, if two coils are wrapped around a core using the same orientation (windings in the same direction), then when current is flowing in one direction on the primary side, the induced current actually flows in the opposite direction. This ultimately explains the feedback between the secondary and primary coils, which is why the actual flow through the primary depends upon the flow through the secondary, so that the wattage is the same in both. A complete explanation is beyond the scope of this text.

One other effect is that of “eddy currents,” which occur in iron-core transformers because the iron itself conducts electricity, and so the alternating magnetic field will induce currents within the core, heating the core and thus leaching some of the energy sent to the primary coil. Clever construction using layering techniques to make the electrically conducting iron pieces small, insulated from each other, and strategically shaped helps to reduce these currents and their effects. Another effect is hysteresis, meaning the resistance in the core of having its magnetic orientation changed. These effects, plus some loss from resistance in the windings themselves, and some heat is generated, which demonstrates that transformers are not 100% efficiency in transferring power from the primary’s circuit to the secondary’s. However, commercial power application transformer efficiencies can approach 98%, which is quite high compared to other energy transforming methods (such as chemical to mechanical, heat to mechanical, or solar to electric).

Finally, we will discuss inductors, which are in some ways simply one-sided transformers, i.e., transformers without a secondary coil.

As we can imagine from the discussion of eddy currents, no conductor is immune from induced voltages when magnetic fields are changing. So consider a coil of wire which is wrapped around an iron core. If an AC voltage forces current through the coil, then each winding in some ways also acts as an albeit small secondary coil. Therefore a reactionary voltage is induced in the other direction from that which is passed through the coil originally. In this way, an inductor will act like a resistor to AC current, though it will act more or less like a conductor with a DC current. However, when a DC circuit with an inductor wired in “series” is first “turned on,” the inductor will resist relatively strongly because of the DC source’s attempted immediate build-up of the magnetic field within the inductor. The inductor will resist this just as it would resist an AC voltage due to the counter current attempts by its various coil windings. Similarly, if the DC source is “turned off,” there will be a sudden drop in the magnetic field in the inductor and this too will be resisted, resulting in a brief output of voltage from the inductor itself when the DC source is first removed.

The final effect of all of this is summarized in the equation $V = L \frac{dI}{dt}$, where $L$ is the inductance.

It should also be pointed out that anytime power is taken from an electrical outlet, this requires more current to pass through any transformers between the outlet and the generators, and therefore more power output from the generators at the power company. This in turn implies there must be more current flowing from, and therefore through, the coils of the generators. Now the coils within the generators will themselves therefore become stronger magnets and will more strongly resist the movement of the generators due to the generators’ own magnets more strongly repelling the coils carrying the extra current.

Restated, the more power that clients use from the electrical utility, the harder it is for that utility’s generators to turn, and therefore the more power that is needed from the original energy sources, be they heat engines, hydroelectric dams turbines, wind turbines, or whichever forms they may take.

To better understand hysteresis, consider the following. If a direct current passes through the primary coils of a transformer, the atoms in the core will align their magnetic poles, and the core will somewhat stay “magnetized” even if the current is turned off. If a DC current is then passed into the primary but in the other direction, the core will somewhat lag in how quickly it is remagnetized into the new orientation, in effect “fighting” the remagnetization effort briefly.

If an AC voltage is applied to the primary coil of an iron-core transformer, the core’s continually lagging magnetic orientation change will somewhat decrease the effectiveness of the primary voltage in changing the magnetic field as the voltage changes direction, thus decreasing the efficiency of the inductance phenomenon.
given in units of henries, H.\textsuperscript{37} Note that we can write instead

\[ L = \frac{V}{(\frac{dI}{dt})}, \]  

(4.27)

and so a 1H inductor will require 1 volt to pass a change of current of 1A/sec. On the other hand, a 5H inductor would require 5 volts to force a change of 1A/sec through an inductor.

\textsuperscript{37}Named for Joseph Henry (1797–1878), an American scientist who discovered “self inductance” by inductors, and (independently of Michael Faraday) the “mutual inductance” seen in transformers.
4.3. FIRST APPLICATIONS OF DERIVATIVES

Exercises

1. In a circuit such as in Figure 4.14, page 335, assuming that $V$ is constant, say, $V = 5$ volts, use Ohm’s Law (4.25) to find

   (a) $\lim_{R \to 0^+} I$
   (b) $\lim_{R \to \infty} I$

2. In a particular electrical circuit, the charge on the cathode is given by $Q = 20 - 3t$. What is the rate of current flowing out of the anode? (Assume current is conventional current.)

3. Consider the voltage $V$ across a variable resistor with $R = 20 - 0.2t$ and $I = 3 + 2t + 0.1t^2$. Here $V$ is in volts, $R$ is in ohms, $I$ is in amps and $t$ is in seconds.

   (a) Find $V$ as a function of $t$.
   (b) Find $\frac{dV}{dt}$ at $t = 1.5$ sec.

4. A cube is a solid object bounded by six identical squares (faces of the cube) of uniform side length $s$. The point where three faces meet is a vertex of the cube, consisting of three right angles. The volume of a cube of side length $s$ is therefore $V = s^3$. Find the rate of change of the volume of a cube with respect to side length $s$ when $s = 10$ cm.

5. A circle is the set of all points in a plane which are a fixed distance $r > 0$, called the circle’s radius, from a fixed point in the plane, called the circle’s center. The area bounded by the circle, usually called the “area of the circle,” is given by $A = \pi r^2$. Find the rate of change in $A$ when $r = 15$ cm.

6. A sphere is the set of all points in space which are the same distance $r > 0$, called the sphere’s radius, from a fixed point called the center of the sphere. A sphere will form the boundary of a volume $V = \frac{4}{3} \pi r^3$, this $V$ usually referred to as the “volume of the sphere.”

   (a) Find the rate of change of the volume of a sphere with respect to its radius.
   (b) At what radius will $\frac{dV}{dr} = 20\pi$ cm$^3$/cm?
The chain rule is perhaps the most important of the differentiation rules. It is immensely rich in application, and very elegantly stated when notation is chosen wisely. In this section we will look closely at the rule itself, and the underlying intuition of the rule.

The mechanics of applying the rule are of utmost importance, but as with other calculus principles, understanding the intuition aids in determining when and how to apply the chain rule. Because of the importance and theoretical richness of the chain rule, the reader is encouraged to revisit this section from time to time, to reinforce this most important topic.

In its simplest form, the chain rule dictates how we must calculate derivatives of compositions of functions, i.e., functions of the form

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

especially when we know how to calculate $f'(a)$ and $g'(a)$. With it we will be able to calculate derivatives for a much wider class of functions, find slopes on implicit curves, and to find so-called related rates relationships among variables.

We will first state the chain rule using our “prime” notation—first introduced by English mathematician Brook Taylor (1685–1731)—and then re-write it in terms of the more powerful and explanatory Leibniz notation. (Taylor’s name appears prominently in Chapter 11, which bears his name.)

### 4.4.1 Chain Rule in Prime Notation

**Theorem 4.4.1 (Chain Rule)** Suppose that $h(x) = f(g(x))$, where $g'(a)$ exists and $f'(g(a))$ exists. Then

$$h'(a) = [f'(g(a))]g'(a).$$

Another way of writing this in Taylor’s prime notation is

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

Note that there is an “outer” function, namely $f$, and an “inner function” $g$. So the chain rule is sometimes stated that we compute the derivative of the outer function with respect to the inner function, that is $f'(g(x))$, and then multiply by the derivative of the inner function, i.e., by $g'(x)$. That is a common, colloquial way of expressing the chain rule. It should be mentioned that “multiplying by the derivative of the inner function” is the step that is most commonly forgotten in such derivative problems.

**Example 4.4.1** Compute $h'(x)$ if $h(x) = (x^2 + 3x)^2$.

**Solution:** Besides the unwieldy “limit definition,” thus far there are two possible methods for computing $f'(x)$ here:

- Expand the function first, and then compute the derivative, as we would need to do if we had little more than the power rule to rely upon:

$$h(x) = x^4 + 6x^3 + 9x^2 \implies h'(x) = 4x^3 + 18x^2 + 18x.$$

- Use the chain rule. Here the “outer function” is $f(x) = x^2$ (squaring the input), and the “inner function” is $g(x) = x^2 + 3x$. Note that $f'(x) = 2x$ while $g'(x) = 2x + 3$. Using the chain rule we would have

$$h'(x) = [f'(g(x))]g'(x) = [2(g(x))']\ g'(x) = [2(x^2 + 3x)]\ (2x + 3).$$

Note that this gives $h'(x) = 2(2x^3 + 9x^2 + 9x) = 4x^3 + 18x^2 + 18x$ as before.
In Chapter 2 we made some use of a visual device that utilized “empty parentheses” to illustrate the actions of functions. For the above derivative, we can look at the following:

<table>
<thead>
<tr>
<th>outer function</th>
<th>is</th>
<th>((x)^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>its derivative</td>
<td>is</td>
<td>(2x)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>inner function</th>
<th>is</th>
<th>((x)^2 + 3(x))</th>
</tr>
</thead>
<tbody>
<tr>
<td>its derivative</td>
<td>is</td>
<td>(2(x) + 3)</td>
</tr>
</tbody>
</table>

The chain rule for this problem could then read:

\[
[(x^2 + 3x)^2]' = 2((x^2 + 3x)) \cdot [2(x) + 3] = 2(x^3 + 3x) \cdot (2x + 3).
\]

As we will see, there are simpler ways of looking at the chain rule than labeling an “outer function” \(f(x)\) and an “inner function” \(g(x)\), calculating \(f'\) and \(g'\), and evaluating at \(g(x)\) and \(x\), respectively. Still, there are advantages over expanding the function first. For instance, expansion might not be so easy. Moreover the final answer using the chain rule is somewhat factored which helps in, say, constructing its sign chart. The next example will demonstrate some of these advantages more dramatically.

**Example 4.4.2** Find \(h'(x)\) if \(h(x) = (x^3 + 27x + 9)^55\).

**Solution:** Certainly we do not want to multiply this out to use earlier rules. Instead we just notice that this is a composition of two functions, with the “outer” function being \(f(x) = x\) and the “inner” function being \(g(x) = x^3 + 27x + 9\). Now \(f'(x) = 55x^{54}\), while \(g'(x) = 3x^2 + 27\).

Thus

\[
h'(x) = [f'(g(x))g'(x)] = 55(g(x))^{54} \cdot g'(x) = 55(x^3 + 27x + 9)^{54} \cdot (3x^2 + 27).
\]

Even if we had somehow expanded the original, 165-degree polynomial first, and then calculated the derivative, it is unlikely we would have noticed that our resulting 164-degree polynomial answer factors so nicely.

The chain rule has much to say about derivatives of functions which contain trigonometric functions in their structures. Below are two examples where the “outer function” and “inner function” are squaring and sine functions, respectively and then vice-versa.

**Example 4.4.3** Find \(h'(x)\) if \(h(x) = \sin^2 x\).

**Solution:** Note that \(h(x) = (\sin x)^2\), so the “outer function” is \(f(x) = x^2\), while the “inner function” is \(\sin x\). Next note that \(f'(x) = 2x\) and \(g'(x) = \cos x\). Thus

\[
h'(x) = [f'(g(x))g'(x)] = 2(\sin x) \cdot \cos x, \quad \text{or alternatively,}\n\]

\[
\frac{d}{dx}(\sin x)^2 = 2(\sin x) \cdot \cos x.
\]

**Example 4.4.4** Suppose instead \(h(x) = \sin x^2\). Here \(f(x) = \sin x\), \(g(x) = x^2\), \(f'(x) = \cos x\), \(g'(x) = 2x\). Hence

\[
h'(x) = [f'(g(x))g'(x)] = \cos g(x) \cdot g'(x) = \cos x^2 \cdot 2x = 2x \cos x^2, \quad \text{or alternatively,}\n\]

\[
\frac{d}{dx}\sin(x^2) = \cos(x^2) \cdot 2x = 2x \cos x^2.
\]

(It is customary to write the polynomial factor before the trigonometric function factor in the final answer, so it is clearer what terms are inside the trigonometric function, and which are multiplying the trigonometric function.)
Note how it is crucial to identify the outer function and the inner function. It is also important that the inner function \( g(x) \) is inputted into the derivative \( f' \) of the outer function.

While we identified the outer and inner functions by name, in fact naming an “\( f(x) \)” and “\( g(x) \)” is unwieldy and unnecessary. Before long (though usually not immediately) the pattern of these computations does become natural enough. (Observe in particular the second methods in each of the two examples given above.)

To show that the chain rule makes sense from the limit-definition standpoint as a derivative rule, we next offer a partial proof of the chain rule. Note how the way the limit is re-written reflects the ultimate statement of the chain rule. It also reflects much of the intuition of the rule.

### 4.4.2 Partial Proof of Chain Rule

We will not completely prove the chain rule in this context because of some technical difficulties which arise when proving the rule in its most general form. However, we will look at a proof in perhaps the most common case, which is the case that \( g(x + \Delta x) - g(x) \to 0 \) “properly” (so that we also have \((3 \delta > 0)[0 < |\Delta x| < \delta \to g(x + \Delta x) - g(x) \neq 0]\). In such a case we can safely divide and multiply by \( g(x + \Delta x) - g(x) \) in the limit definition of the derivative to get

\[
\frac{d}{dx} [f(g(x))] = \lim_{\Delta x \to 0} \frac{f(g(x + \Delta x)) - f(g(x))}{\Delta x} = \lim_{\Delta x \to 0} \left( f(g(x + \Delta x)) - f(g(x)) \right) \frac{g(x + \Delta x) - g(x)}{\Delta x}.
\]

Now we claim that this limit is \( f'(g(x))g'(x) \). The second term \((II)\) clearly has limit \( g'(x) \), by definition. Under the assumption that \( g(x + \Delta x) - g(x) \to 0 \) properly as \( \Delta x \to 0 \), we can substitute \( \Delta g(x) = g(x + \Delta x) - g(x) \to 0 \), and rewrite the limit \((I)\)

\[
\lim_{\Delta g(x) \to 0} \frac{f(g(x) + \Delta g(x)) - f(g(x))}{\Delta g(x)} = f'(g(x)).
\]

Note that the computation above is also correct even if \( \Delta g(x) \to 0^+ \) or \( \Delta g(x) \to 0^- \) properly, because the existence of the two-sided limit represented by \( f'(g(x)) \) is assumed to exist in our statement of the chain rule theorem. Thus \( \frac{d}{dx} f(g(x)) = (I) \cdot (II) = f'(g(x)) \cdot g'(x) \), q.e.d.

### 4.4.3 Leibniz Notation and the Chain Rule

Before we see how the chain rule is stated with Leibniz notation, first we will make some observations about that notation. For example, the following three formulas say the same thing—how the square of a quantity changes with respect to the quantity—albeit with different variables:

\[
\frac{d x^2}{dx} = 2x, \quad \frac{d t^2}{dt} = 2t, \quad \frac{d u^2}{du} = 2u.
\]

See Figure 4.17 for a graphical interpretation of this fact. It is important that in each equation the variables matched. (Note, for instance, that \( du^2/dx \neq 2u \), as we shall soon see, the problem

\[\text{When we rewrite } f(g(x + \Delta x)) = f(g(x) + \Delta g(x)), \text{ we were justified because} \]

\[
g(x + \Delta x) = (g(x + \Delta x) - g(x)) + g(x) = \Delta g(x) + g(x) = g(x) + \Delta g(x).
\]
being that the variables involved do not match, so the original power rule, namely \( \frac{d}{dx} x^n = nx^{n-1} \) or equivalently \( \frac{d}{du} u^n = nu^{n-1} \), does not apply.

Now suppose we have a differentiable function \( u = u(x) \) and want to take the derivative of \( (u(x))^2 \). The “outer” function is \( f(x) = x^2 \) (i.e., squaring what is inside), while the “inner” function is \( u = u(x) \). Consider how we find the derivative of \( (u(x))^2 \) (a chain rule problem) with Taylor’s “prime” notation (4.29), and with Leibniz’s notation:

Taylor: \[
\frac{d}{dx} (u(x))^2 = 2(u(x)) \cdot u'(x)
\]

Leibniz: \[
\frac{du^2}{dx} = \frac{du^2}{du} \cdot \frac{du}{dx} = 2u \cdot \frac{du}{dx}.
\]

The two notations say the same thing, but the Leibniz notation has several advantages, two of which we point out here:

- Resemblance to algebraic manipulations: it appears that we simply decomposed \( \frac{du^2}{dx} \) by dividing and multiplying by \( du \), yielding derivatives that made sense and could be calculated by known rules (because the variables matched): the first by the power rule, and the second by whatever method could give us \( du/dx \).

- Variable of differentiation appears explicitly: we know when we are taking the derivative with respect to \( x \), and when it is with respect to \( u \).

Compare (4.31) to our partial proof from the last subsection. Before giving more computational examples, we will look at another argument for the validity of the chain rule. We begin with a very simple example.

**Example 4.4.5** Suppose we have a vehicle which always achieves a fuel efficiency rating of 35 mile/gallon, and each gallon costs $1.40. Then we can ask what is the cost per mile. We can think of this situation as total cost \( C \) being a function of total gallons consumed \( g \), i.e., \( C = C(g) \) and total gallons as a function of total miles, i.e., \( g = g(m) \). Ultimately cost is then a (composite) function of miles, i.e., \( C = C(g(m)) \). Now cost per mile will be cost per gallon
times gallons per mile. In other words, the rate of change in $C$ with respect to total miles $m$ is

\[
\frac{dC}{dg} \cdot \frac{dg}{dm} = \frac{dC}{dm}
\]

This example is simple because these rates do not change. Still, it is reasonable that even if these rates $dC/dg$ and $dg/dm$ only hold for an instant in time, during that instant $dC/dm$ will still be the product $dC/dg \cdot dg/dm$ as above. (This is not a proof, but an argument for reasonableness.)

For that reason, we can similarly do the following (but where the “instant” is a particular value of $x$):

\[
\frac{d}{dx} (x^3 + 1)^2 = \frac{d}{dx} (x^3 + 1)^2 \cdot \frac{d(x^3 + 1)}{dx} = 2(x^3 + 1)^1 \cdot 3x^2.
\]  

That $d(x^3 + 1)^2/d(x^3 + 1) = 2(x^3 + 1)$ is not much different from the argument in Figure 4.17, page 354, except the “horizontal axis” would be $(x^3 + 1)$ while the vertical would be $(x^3 + 1)^2$.

The derivative with respect to $x$—which is a “hidden” variable upon which the others depend—is found by compensation, that is, the rate of change of $(x^3 + 1)^2$ with respect to $x$ is found by first finding its rate of change with respect to $(x^3 + 1)$, and then multiplying by a compensating factor which is the rate of change of $(x^3 + 1)$ with respect to $x$. The idea is very similar to (4.32).

To review (4.33), there we want to know how $(x^3 + 1)^2$ changes with respect to $x$, so we first ask how does $(x^3 + 1)^2$ change with respect to $(x^3 + 1)$, i.e., how does the square of a quantity change with respect to that quantity (power rule)—and multiply by how $(x^3 + 1)$ changes as $x$ changes. With this kind of argument we can extend the power rule to have a chain rule version,

\[
\frac{du^n}{dx} = \frac{du^n}{du} \cdot \frac{du}{dx}, \quad \text{i.e.,}
\]

\[
\frac{du^n}{dx} = nu^{n-1} \frac{du}{dx}.
\]

**Example 4.4.6** With (4.34) we can quickly calculate the following derivatives, which would be more difficult without the chain rule:

\[
\begin{align*}
\frac{d}{dx} (29x - x^2)^4 &= 4(29x - x^2)^3 \cdot \frac{d}{dx} (29x - x^2) = 4(29x - x^2)^3(29 - 2x), \\
\frac{d}{dx} (x + 21)^9 &= 9(x + 21)^8 \cdot \frac{d}{dx} (x + 21) = 9(x + 21)^8 \cdot 1 = 9(x + 21)^8. \quad \text{Note that occasionally the derivative } \frac{du}{dx} \text{ of the “inner function” is just 1.}
\end{align*}
\]

This is somewhat similar to compensating for unmatched units in physics or chemistry problems. If we travel 60 miles in 75 minutes, and we want our average speed in miles/hour, we can first find miles/minute, and then multiply by a compensating factor which relates minutes to hours:

\[
\text{speed} = \frac{60 \text{ mile}}{75 \text{ minute}} = 0.8 \frac{\text{mile}}{\text{minute}} \cdot \frac{60 \text{ minute}}{1 \text{ hour}} = 48 \text{ mile/hour}
\]

Since the question was how the distance relates to hours, we first did the easy computation relating distance to minutes, and then compensated by multiplying how minutes relate to hours.
\[ \frac{d}{dx}(5x - 9)^8 = 8(5x - 9)^7 \cdot \frac{d}{dx}(5x - 9) = 8(5x - 9)^7 \cdot 5 = 40(5x - 9)^7. \]

When we computed \( \frac{d}{dx}(5x - 9)^8 \), we could have written a Leibniz-style decomposition/expansion,

\[ \frac{d}{dx}(5x - 9)^8 = \frac{d}{d(5x - 9)}(5x - 9)^8 \cdot \frac{d}{dx}(5x - 9) = 8(5x - 9)^7 \cdot 5, \]

(4.35)

and this is quite correct. However it is not standard practice to write the middle step. Indeed most authors prefer to avoid having complicated expressions in the denominator of a differential operator, preferring denominators as in \( \frac{d}{dx} \), \( \frac{d}{dt} \), etc. In this text we will still occasionally write as in (4.35) for clarity (which is akin to using a truth table to show a style of argument is valid), but more often we will just state the chain rule version (4.34) of the power rule with the correct terms in place of the general \( u \).

The reader is encouraged in the strongest possible terms to always write the first step of “power rule with chain rule” problems, as shown in the next examples below. This will avoid common errors, especially as derivative computations become more and more complex, and such practice will reinforce the proper use of the chain rule version of the power rule, (4.34). We now list several further examples of this rule, showing in the first step what the rule says literally for each problem, before the “inner function’s” derivative is computed.

- \( \frac{d}{dx}(2x + 9)^3 = 3(2x + 9)^2 \cdot \frac{d}{dx}(2x + 9) = 3(2x + 9)^2 \cdot 2 = 6(2x + 9)^2. \)

- \( \frac{d}{dx} \sin^2 x = \frac{d}{dx}(\sin x)^2 = 2(\sin x) \frac{d}{dx} \sin x = 2 \sin x \cos x. \) Note that the final answer is understood to mean \( 2(\sin x)(\cos x) \).

- \( \frac{d}{dx} (x^2 + \cos x)^4 = 4(x^2 + \cos x)^3 \frac{d}{dx} (x^2 + \cos x) = 4(x^2 + \cos x)^3(2x - \sin x). \)

- \( \frac{d}{dx} (3x^2 + 6x + 7)^7 = 7 (3x^2 + 6x + 7)^6 \cdot \frac{d}{dx} (3x^2 + 6x + 7)
\]
\[ = 7 (3x^2 + 6x + 7)^6 (6x + 6) = 7(3x^2 + 6x + 7)^6 \cdot 6(x + 1) = 42(3x^2 + 6x + 7)(x + 1). \]

Note that the chain rule version of the power rule, \( \frac{d}{dx} u^n = nu^{n-1} \cdot \frac{du}{dx} \), does not contradict the earlier power rule that \( \frac{d}{dx} x^n = nx^{n-1} \). For instance, when “\( u \)” is equal to \( x \), we can write

\[ \frac{d}{dx} x^9 = 9x^8 \cdot \frac{dx}{dx} = 9x^8 \cdot 1 = 9x^8, \]

which agrees with our original power rule, which here would give us \( \frac{d}{dx} x^9 = 9x^8 \). Thus the chain rule version of the power rule in fact generalizes the original power rule.

### 4.4.4 Chain Rule Derivatives of Sine and Cosine

Now we look at derivatives of \( \sin u \) and \( \cos u \) with respect to \( x \), assuming \( u = u(x) \), i.e., that \( u \) is actually a function of \( x \). Recall (4.18) and (4.19) from page 318: \( \frac{d}{dx} \sin x = \cos x \) and \( \frac{d}{dx} \cos x = -\sin x \). The chain rule versions of the derivatives of sine and cosine follow easily from the Leibniz notation:

\[ \frac{d}{dx} \sin u = \frac{d}{du} \sin u \cdot \frac{du}{dx} = \cos u \cdot \frac{du}{dx}, \]
\[ \frac{d}{dx} \cos u = \frac{d}{du} \cos u \cdot \frac{du}{dx} = -\sin u \cdot \frac{du}{dx}, \]
4.4. CHAIN RULE I

As interesting as the above equations are to behold, in practice one usually uses summary versions:

\[
\frac{d \sin u}{dx} = \cos u \cdot \frac{du}{dx}, \quad \frac{d \cos u}{dx} = -\sin u \cdot \frac{du}{dx},
\]

(4.36) (4.37)

Thus derived, we are free to use (4.36) and (4.37) where applicable. In the next example we show two methods of computing a particular derivative: using a Leibniz-style decomposition, and applying (4.36) directly.

Example 4.4.7 If \( f(x) = \sin x^2 \), that is \( f(x) = \sin(x^2) \), then we can compute \( f'(x) \) the following two ways:

\[
\frac{d \sin x^2}{dx} = \frac{d \sin x^2}{dx^2} \cdot \frac{dx^2}{dx} = \cos x^2 \cdot 2x = 2x \cos x^2, \quad (4.38)
\]

\[
\frac{d \sin x^2}{dx} = \cos x^2 \cdot \frac{dx^2}{dx} = \cos x^2 \cdot 2x = 2x \cos x^2. \quad (4.39)
\]

In the second method (4.39) for computing the derivative above, we did just insert \( u = x^2 \) into (4.36), but the justification for that can also be seen in the first method (4.38) with the Leibniz-style decomposition. Note also that we computed this same derivative in Example 4.4.4, page 352 using the prime notation.

Example 4.4.8 Compute \( \frac{df(z)}{dz} \) if \( f(z) = \cos(z^3 + \sin z) \).

Solution: Here the names of the variables have changed, but the principle of the chain rule is the same. Again we will compute this two ways, the second using (4.37), except with \( z \) in place of \( x \):

\[
f'(z) = \frac{d}{dz} \cos(z^3 + \sin z) = \frac{d \cos(z^3 + \sin z)}{dz} \cdot \frac{d(z^3 + \sin z)}{dz} = -\sin(z^3 + \sin z) \cdot (3z^2 + \cos z),
\]

\[
f'(z) = \frac{d}{dz} \cos(z^3 + \sin z) = -\sin(z^3 + \sin z) \cdot \frac{d(z^3 + \sin z)}{dz} = -\sin(z^3 + \sin z) \cdot (3z^2 + \cos z).
\]

Accepted practice is to compute the above derivative using the latter method, and so that is the method students should eventually strive to reproduce. As in an earlier discussion involving the power rule, one can think of this example as using the thought pattern that says the derivative of cosine is minus sine... multiplied by the derivative of what is inside the cosine. That is perhaps an over-simplification, and should be informed by awareness of what we get from the Leibniz-style expansion.

To be clear on what (4.36) and (4.37) say, and why these should hold, consider the following abstract equations, which are in fact restatements of (4.36):\(^{40}\)

\[
\frac{d \sin u}{dw} = \frac{d \sin u}{du} \cdot \frac{du}{dw}, \quad \frac{d \sin \theta}{d\xi} = \frac{d \sin \theta}{d\theta} \cdot \frac{d\theta}{d\xi},
\]

\[
\frac{d \sin x}{dt} = \frac{d \sin x}{dx} \cdot \frac{dx}{dt} = \cos x \cdot \frac{dx}{dt}.
\]
Note that in all the cases, the decomposition’s first factor let us use the known derivative formula for sine—because the variables matched—and then we compensated for introducing the new variable’s derivative (as a fraction of sorts) with the second factor.

We now point out that it is quite common for the chain rule to apply more than once in a particular problem. Our next example below shows a case of a function within a function within a function, and the example following will be a sum of two functions, each requiring a chain rule.

**Example 4.4.9** Compute \( \frac{d}{dx} \left[ \sin^3(4x) \right] \).

**Solution:** Note that the function can be rewritten \([\sin 4x]^3\). Now we compute the derivative, first applying the power rule version of the chain rule (4.34), page 355, and then (4.36), page 357. Then we show the same computation using the Leibniz-style decomposition.

\[
\frac{d[\sin 4x]^3}{dx} = 3[\sin 4x]^2 \cdot \frac{d[\sin 4x]}{dx} \\
= 3 \sin^2 4x \cdot \cos 4x \cdot \frac{d(4x)}{dx} \\
= 3 \sin^2 4x \cos 4x \cdot 4 = 12 \sin^2 4x \cos 4x.
\]

\[
\frac{d[\sin 4x]^3}{dx} = \frac{d[\sin 4x]^3}{d[\sin 4x]} \cdot \frac{d[\sin 4x]}{dx} \\
= 3[\sin 4x]^2 \cdot \cos 4x \cdot 4 = 12 \sin^2 4x \cos 4x.
\]

So the Leibniz-style decomposition will work for longer “chains” of functions within functions. But so will the abbreviated chain rules which say \( \frac{du}{dx} = 3u^2 \cdot \frac{du}{dx} \), and \( \frac{d[\sin u]}{dx} = \cos u \cdot \frac{du}{dx} \), which was the first approach in the computations above: after applying the power rule, the “inner” derivative called another chain rule.\(^{41}\) Again, it is best to strive for the abbreviated approach in practice, though both approaches are worth studying (and of course the decomposition explains the abbreviated approach).

**Example 4.4.10** Find \( f'(x) \) if \( f(x) = \sin^2 x + \cos^2 x \).

**Solution:** The larger structure of this function is that of a sum of two functions, so first we use the sum rule, which tells us to add (and thus first compute) the derivatives of \( \sin^2 x \) and \( \cos^2 x \). These are then both chain rules.

\[
f'(x) = \frac{d}{dx} \left[ \sin^2 x + \cos^2 x \right] \\
= \frac{d}{dx} \left[ \sin^2 x \right] + \frac{d}{dx} \left[ \cos^2 x \right] \\
= \frac{d}{dx} \left[ (\sin x)^2 \right] + \frac{d}{dx} \left[ (\cos x)^2 \right] \\
= 2(\sin x) \frac{d}{dx} \sin x + 2(\cos x) \frac{d}{dx} \cos x \\
= 2 \sin x \cos x + 2 \cos x (- \sin x) \\
= 2 \sin x \cos x - 2 \cos x \sin x \\
= 0.
\]

\(^{40}\)Just to be sure, it should be pointed out that when we write for instance \( \cos x \cdot \frac{dx}{dt} \), we mean that the \( \frac{dx}{dt} \) is outside of the cosine function, i.e., we mean \( (\cos x) \cdot \frac{dx}{dt} \). Note that many texts assume this meaning without making it explicit with the dot “\( \cdot \)" and simply write \( \cos x \frac{dx}{dt} \). As a matter of style, it is assumed the derivative \( \frac{dx}{dt} \) is not part of the argument of the cosine function in such a case.

\(^{41}\)This phenomenon of “rules calling other rules” occurs repeatedly throughout the rest of the textbook. We will see it occasionally in this section, and it will become the norm in later sections.
4.4. CHAIN RULE I

Actually this is what we should hope would be the answer, for the original function we are taking the derivative of is actually constant:

\[
\frac{d}{dx} [\sin^2 x + \cos^2 x] = \frac{d}{dx} [1] = 0.
\]

It happens frequently in calculus that it is advantageous to algebraically rewrite a function before taking its derivative. In fact we did that each time we took a derivative of \( \sin^2 x = (\sin x)^2 \), the latter notation being more obvious in illustrating the composition (function inside of a function) structure of the original function. For calculus to be consistent (which it is, so no need to fear!), we should be able to rewrite the function and get the same derivative, as long as we rewrite the function correctly. The differentiation (derivative computing) rules are eventually sufficient to compute the derivative no matter how the function is rewritten, but some forms of a given function are easier to deal with than others.

4.4.5 Power Rule for Rational Powers

With the chain rule we have enough theoretical development to show that the power rule actually holds for any constant power which is a rational number \( \frac{p}{q} \) (where \( p, q \in \mathbb{Z} \), and of course \( q \neq 0 \)). Recall that the set of all rational numbers was denoted \( \mathbb{Q} \), for “quotients,” i.e., fractions, of integers. We already proved the power rule for powers \( n \in \{0, 1, 2, 3, \ldots \} \), and that result is used in the proof for rational powers which we leave the proof until the end of this section, so we can expeditiously come to examples. But first we state the theorem.

**Theorem 4.4.2 (Power Rule for Rational Powers)** For any \( r \in \mathbb{Q} - \{0\} \) (i.e., nonzero rational numbers),

\[
\frac{d}{dx} x^r = r x^{r-1}, \quad (4.40)
\]

\[
\frac{d}{dx} u^r = r u^{r-1} \cdot \frac{du}{dx}. \quad (4.41)
\]

**Example 4.4.11** For example, the following (which was an exercise with difference quotient limits in Section 4.1) yields quickly to the power rule:

\[
\frac{d}{dx} \sqrt{x} = \frac{d}{dx} x^{1/2} = \frac{1}{2} x^{1/2 - 1} = \frac{1}{2} x^{-1/2} = \frac{1}{2 \sqrt{x}}.
\]

In fact, this particular derivative occurs often enough that it, along with its chain rule version, deserves special attention (and should be committed to memory):

\[
\frac{d}{dx} \sqrt{x} = \frac{1}{2 \sqrt{x}}. \quad (4.42)
\]

\[
\frac{d}{dx} \sqrt{u} = \frac{1}{2 \sqrt{u}} \cdot \frac{du}{dx}. \quad (4.43)
\]

**Example 4.4.12** Find \( f'(x) \) for \( f(x) = \sqrt{x^2 + 1} \).

\[
f'(x) = \frac{d}{dx} \sqrt{x^2 + 1} = \frac{1}{2 \sqrt{x^2 + 1}} \cdot \frac{d}{dx} (x^2 + 1) = \frac{1}{2 \sqrt{x^2 + 1}} \cdot (2x) = \frac{x}{\sqrt{x^2 + 1}}.
\]
Note that in the above example the “outer” function was the square root, while the “inner” function is \( x^2 + 1 \). One could write (though again, it is not standard practice):

\[
\frac{d}{dx} \sqrt{x^2 + 1} = \frac{d\sqrt{x^2 + 1}}{dx(x^2 + 1)} \cdot \frac{d(x^2 + 1)}{dx} = \frac{1}{2\sqrt{x^2 + 1}} \cdot 2x = \frac{x}{\sqrt{x^2 + 1}}.
\]

**Example 4.4.13** Suppose \( f(x) = \sqrt{x + \sqrt{x}} \). Then the Leibniz decomposition would look like

\[
\frac{df(x)}{dx} = \frac{d\sqrt{x + \sqrt{x}}}{dx} \cdot \frac{d(x + \sqrt{x})}{dx} = \frac{1}{2\sqrt{x + \sqrt{x}}} \left(1 + \frac{1}{2\sqrt{x}}\right).
\]

Again—and especially with practice—one would usually not write the Leibniz decomposition in the first step, but should instead write

\[
\frac{d}{dx} \sqrt{x + \sqrt{x}} = \frac{1}{2\sqrt{x + \sqrt{x}}} \cdot \frac{d(x + \sqrt{x})}{dx} = \frac{1}{2\sqrt{x + \sqrt{x}}} \left(1 + \frac{1}{2\sqrt{x}}\right).
\]

A common mistake in the above example is to think of \( \sqrt{x} \) as the inner function, since geometrically it somehow appears to be innermost. In fact the inner function is actually the whole of \( x + \sqrt{x} \).

**Example 4.4.14** Suppose \( f(x) = \frac{2}{(x^3 - 9x + 7)^2} \). Then

\[
f'(x) = \frac{d}{dx} \left[2(x^3 - 9x + 7)^{-2}\right]
= 2 \cdot (-7)(x^3 - 9x + 7)^{-8} \frac{d}{dx}(x^3 - 9x + 7)
= -14(x^3 - 9x + 7)^{-8}(3x^2 - 9)
= \frac{-14(3x^2 - 9)}{(x^3 - 9x + 7)^5}.
\]

In the previous example we were able to use the chain rule version (4.41) of the power rule once we wrote the function as a power of a polynomial, albeit negative and with a multiplicative constant along for the ride. The next example calls for a rewriting (for simplicity), and then calls the chain rule twice.

**Example 4.4.15** \( f(x) = \sqrt[3]{\frac{1}{x + \sqrt{x^3 + 9}}} \)

\[
f'(x) = \frac{d}{dx} \left[3\sqrt[3]{\frac{1}{x + \sqrt{x^3 + 9}}}\right]
= \frac{d}{dx} (x + \sqrt{x^3 + 9})^{-1/3}
= \frac{1}{3} (x + \sqrt{x^3 + 9})^{-4/3} \cdot \frac{d}{dx} (x + \sqrt{x^3 + 9})
= \frac{1}{3} (x + \sqrt{x^3 + 9})^{-4/3} \cdot \left[1 + \frac{1}{2\sqrt{x^3 + 9}} \cdot \frac{d}{dx}(x^3 + 9)\right]
= \frac{1}{3} (x + \sqrt{x^3 + 9})^{-4/3} \cdot \left[1 + \frac{3x^2}{2\sqrt{x^3 + 9}}\right]
\]
In the calculation above, we first rewrote the expression as a $\frac{1}{3}$ power, then used the chain rule version of the power rule, and used the chain rule again in calculating the derivative of that “inner” function.

**4.4.6 Proof Of Power Rule for Rational Powers**

We now divert temporarily to prove the power rule for rational numbers $r \in \mathbb{Q} = \mathbb{R} \setminus \{0\}$, from which the chain rule version also follows. The technique used will recur several times throughout the text. Fortunately it is somewhat self-explanatory, assuming knowledge of the general chain rule.

The proof is in two steps, the first being a proof in the case of negative integer powers, from which we can eventually recover all rational power cases.

**Proof:** Now we will use the chain rule to show that the power rule, $\frac{d}{dx} x^r = rx^{r-1}$, holds also for any rational power $r = p/q$, with $p, q$ nonzero integers. (The case $p = 0$ is trivial and the case $q = 0$ is meaningless.) The proofs below are included for completeness, and also because they foreshadow a method we will use extensively later in the text, that method being *implicit differentiation*.

First we will show that the power rule holds for $y = x^n$ for any negative integer exponents $n$. In such cases we can write $y = x^{-m}$ for a positive integer exponent $m$ (namely $-n$). But then $y^{-1} = x^m$. Furthermore we already showed in Section 4.1 (Example 4.1.4, page 301) that the derivative definition gives us $\frac{dy^{-1}}{dy} = -1/y^2$ (though the variable used in the proof there was $x$). Using this and the chain rule, we get

\[
  y^{-1} = x^m \\
  \implies \frac{d}{dx} [y^{-1}] = \frac{d}{dx} [x^m] \\
  \implies \frac{1}{y^2} \frac{dy}{dx} = mx^{m-1} \\
  \implies \frac{dy}{dx} = -y^2mx^{m-1} \\
  = -(x^n)^2(-n)x^{-n-1} \\
  = nx^{2n-n-1} = nx^{n-1}, \quad \text{q.e.d.}
\]

It is important to that we interpret the first implication correctly. Recall that $y = x^n$, so $y$ is a function of $x$. But then so is $y^{-1}$ and, in fact, $y^{-1}$ and $x^m$ are the same functions of $x$. Hence, if $y^{-1}$ and $x^m$ were graphed versus $x$, the graphs would be the same, so the slopes at each $x$-value would be the same. Therefore $y^{-1}$ and $x^m$ have the same derivative with respect to $x$.

Now we will use the chain rule in a similar way to compute the derivatives of rational powers of $x$. Suppose $y = x^{p/q}$, where $p, q \in \mathbb{Z} - \{0\}$ are nonzero integers, and that $r = p/q$ is in simplified form. Then we can raise both sides of $y = x^{p/q}$ to the power...
CHAPTER 4. THE DERIVATIVE

$q$ to get

\[ y^q = x^p \]

\[ \Rightarrow \quad \frac{dy^q}{dx} = \frac{dx^p}{dx} \]

\[ \Rightarrow \quad qy^{q-1} \frac{dy}{dx} = px^{p-1} \]

\[ \Rightarrow \quad \frac{dy}{dx} = \frac{p}{q} \frac{x^{p-1}}{x^{q-1}} = \frac{p}{q} \frac{x^{q-p}}{x^{p-q}} = \frac{p}{q} x^{q-p} \]

which can be rewritten \( \frac{dy}{dx} = rx^{r-1} \). Thus \( y = x^r \) implies the form that we sought to prove for the derivative.

In fact, once we have logarithms we can define \( y = x^r \) for all \( r \in \mathbb{R} \), and find again that the derivative is given by the same formula as in our power rules here.

### 4.4.7 Further Examples

**Example 4.4.16** Suppose charge \( q_1 \) is fixed and \( q_2 \) is moving away from charge \( q_1 \), and \( r \) is the distance between them. When \( r = 0.1 \text{m} \), we have \( F = 0.6 \text{N} \). Find \( \frac{dF}{dr} \) when \( r = 0.1 \text{m} \) and when \( r = 0.3 \text{m} \).

**Solution:** Recall Coulomb’s Law: \( F = k \frac{q_1 q_2}{r^2} \). Since \( F = k \frac{q_1 q_2}{r^2} = (kq_1q_2) r^{-2} \), and

\[ 0.6 = \frac{kq_1 q_2}{(0.1)^2} \]

\[ \Leftrightarrow \quad 0.6 = (kq_1q_2) \cdot 100 \]

\[ \Leftrightarrow kq_1q_2 = 0.006 \]

\[ \Rightarrow F = \frac{kq_1 q_2}{r^2} = \frac{0.006}{r^2}. \]

(Recall that \( k, q_1 \) and \( q_2 \) are all constants.) From this we can deduce

\[ \frac{dF}{dr} = \frac{d}{dr} \left[ \frac{0.006}{r^2} \right] = \frac{d}{dr} \left[ 0.006r^{-2} \right] = -2(0.006)r^{-3} = -0.012r^{-3} \]

\[ \Rightarrow \quad \frac{dF}{dr} = -0.012 \frac{1}{r^3}. \]

With that, we have

\[ \left. \frac{dF}{dr} \right|_{r=0.1} = -0.012 \frac{1}{0.1^3} = -12, \]

\[ \left. \frac{dF}{dr} \right|_{r=0.3} = -0.012 \frac{1}{0.3^3} \approx -0.44. \]

Note that for \( r > 0 \) we have \( \frac{dF}{dr} < 0 \), so \( F \) decreases as \( r \) increases. Furthermore, the rate of decrease in \( F \) itself slows.

\[ \text{Note that p or q (but not both, since p/q is simplified) could be negative, but what we are about to do is justified by the previous result that the power rule also works for negative integer exponents.} \]
4.4. CHAIN RULE I

Example 4.4.17 Supply and demand for an item are usually related to its price: demand will be high if the price is low; producers will supply a large number of an item if the price is high. The price elasticity of demand is given by the formula

$$ \varepsilon = \left( \frac{P}{q} \right) \left| \frac{dq}{dP} \right|, $$

(4.44)

where $P$ is the price and $q$ is the quantity for a given demand equation which relates $P$ and $q$. Demand is said to be elastic if $\varepsilon > 1$, inelastic if $\varepsilon < 1$, and unitary if $\varepsilon = 1$. Describe the elasticity of the following demand curve equations:

(a) $q = 100 - \frac{1}{3}P$

(b) $q = \sqrt{1200 - P}$

Solution:

(a) For $q = 100 - \frac{1}{3}P$, we have $\frac{dq}{dP} = -\frac{1}{3}$, and so

$$ \varepsilon = \left( \frac{P}{q} \right) \left| \frac{dq}{dP} \right| = \left( \frac{P}{100 - \frac{1}{3}P} \right) \left| -\frac{1}{3} \right| = \frac{P}{100 - \frac{1}{3}P} \cdot \frac{1}{3} = \frac{P}{300 - P}.$$

We can either solve directly for the three cases $P/(300 - P) < 1, > 1, = 1$ or we can introduce $F(P) = \frac{P}{300 - P} - 1$ and see where this is negative, positive and zero—in short, construct a sign chart for $F(P)$:

$$ F(P) = \frac{P}{300 - P} - 1 = \frac{P - (300 - P)}{300 - P} = \frac{2P - 300}{300 - P}. $$

we can see this is zero when $2P - 300 = 0 \iff 2P = 300 \iff P = 150$. This corresponds to $\varepsilon = 1$, so demand is unitary ($\varepsilon = 1$) when $P = 150$.

Before constructing the sign chart, we notice that $q = 100 - \frac{1}{3}P$ requires that $P < 300$, lest $q$ be negative. Furthermore, we can assume $P > 0$ as well. With these constraints in mind, we produce our sign chart:

| Function: $F(P) = \frac{2(P - 150)}{300 - P}$ |
| Test $P = 50$ $200$ |
| Sign Factors: $\Theta/\Theta$ $\Theta/\Theta$ |
| Sign $F(P)$: $0$ $\Theta$ $150$ $\Theta$ $300$ |

From the chart we get $F(P) < 0$ on $0 < P < 150$, $F(P) = 0$ at $P = 150$ and $F(P) > 0$ on $150 < P < 300$. Equivalently, we get

- (inelastic) $\varepsilon < 0 \iff 0 < P < 150$,
- (unitary) $\varepsilon = 0 \iff P = 150$,
- (elastic) $\varepsilon > 0 \iff 150 < P < 300$.

This formula is the derivative version of an alternative definition, in which $\varepsilon$ is defined as

$$ \frac{(\Delta q)/q}{(\Delta P)/P} \cdot 100\% \cdot \frac{(\Delta q)/q}{(\Delta P)/P} = \text{percent change in } q \text{ percent change in } P. $$

Using the first expression in the line above, and replacing $\Delta$ with $d$, we see this is approximately $(P/q) \cdot (dq/dP)$. However, this quantity being almost always negative, economists tend to verbalize its absolute value instead. Since $P$ and $q$ are nonnegative (or at least it stretches the imagination to find scenarios where they are negative), using absolute values with the derivative's output generally leaves a nonnegative value for $\varepsilon$. 

---

43This formula is the derivative version of an alternative definition, in which $\varepsilon$ is defined as

$$ \frac{(\Delta q)/q}{(\Delta P)/P} = \frac{100\% \cdot (\Delta q)/q}{100\% \cdot (\Delta P)/P}. $$

Using the first expression in the line above, and replacing $\Delta$ with $d$, we see this is approximately $(P/q) \cdot (dq/dP)$. However, this quantity being almost always negative, economists tend to verbalize its absolute value instead. Since $P$ and $q$ are nonnegative (or at least it stretches the imagination to find scenarios where they are negative), using absolute values with the derivative's output generally leaves a nonnegative value for $\varepsilon$. 

---
(b) For the case \( q = \sqrt{1200 - P} \), we have

\[
\varepsilon = \left( \frac{P}{q} \right) \left| \frac{dq}{dP} \right| = \left( \frac{P}{\sqrt{1200 - P}} \right) \left| \frac{1}{2\sqrt{1200 - P}} \cdot \frac{d}{dP}(1200 - P) \right|
\]

\[
= \frac{P}{\sqrt{1200 - P}} \left| \frac{-1}{2\sqrt{1200 - P}} \right| = \frac{P}{2(1200 - P)}.
\]

We can use the same technique as we did in (a) to find where \( \varepsilon < 1, > 1, = 1 \), though this time we will opt for solving these directly. Note that we must have \( 0 < P < 1200 \) due to the definition of \( q \). Under that assumption, we have:

\[
\varepsilon = 1 \iff \frac{P}{2400 - 2P} = 1 \iff P = 2400 - 2P \iff 3P = 2400 \iff P = 800,
\]

\[
\varepsilon < 1 \iff \frac{P}{2400 - 2P} < 1 \iff P < 2400 - 2P \iff 3P < 2400 \iff P < 800,
\]

\[
\varepsilon > 1 \iff \frac{P}{2400 - 2P} > 1 \iff P > 2400 - 2P \iff 3P > 2400 \iff P > 800,
\]

Remembering our constraint \( 0 < P < 1200 \), we can write in fact that

- (inelastic) \( \varepsilon < 0 \iff 0 < P < 800 \),
- (unitary) \( \varepsilon = 0 \iff P = 800 \),
- (elastic) \( \varepsilon > 0 \iff 800 < P < 1200 \).

It should be noted that we could multiply both sides of the inequalities by \( 2400 - 2P \) because we were operating under the assumption that \( 0 < P < 1200 \), and this implies \( 2400 - 2P > 0 \). If we are unsure of the sign of an expression we wish to multiply by to manipulate an inequality, it is safer to use a sign chart method, which avoids multiplying both sides of an inequality by a variable quantity (of unknown sign). Also, in a case like (b) we see nearly the same computation completed three times, and the sign chart somewhat decreases this redundancy. However, the observant student might combine both ideas, finding where \( \varepsilon = 1 \) and simply testing the other regions in simpler cases such as the above.

Example 4.4.18 Suppose an object slides on a straight, frictionless track. Suppose further that one end of a spring is attached to a fixed point at the end of the track, and that the other end of the spring is attached to the sliding object. If the object is otherwise free to slide, its motion will be periodic, of a type known as simple harmonic motion.

Suppose that the equilibrium point for object attached to the spring is at position \( x = 0 \), \( A \) is the maximum positive displacement of the object, \( x(t) \) is the position of the object at any given time \( t \), and \( x(0) = A \), we can give the position by
4.4. CHAIN RULE I

\[ x = A \cos \left( \frac{2\pi}{T} t \right), \]

where \( T \) is the period of the motion, that is, the time required for one complete cycle before the motion repeats. This equation would hold, for instance, if the object were pulled to position \( A \) and then released at time \( t = 0 \).

Find the velocity and acceleration for any time \( t > 0 \).

**Solution:** As usual we will define velocity \( v(t) = \frac{dx}{dt} \), and acceleration \( a(t) = \frac{dv}{dt} \). From the chain rule we get the following computations (note \( \frac{2\pi}{T} \) is a constant):

\[
\begin{align*}
v(t) &= \frac{d}{dt} \left[ A \cos \left( \frac{2\pi}{T} t \right) \right] \\
&= -A \sin \left( \frac{2\pi}{T} t \right) \cdot \frac{d}{dt} \left[ \frac{2\pi}{T} t \right] \\
&= -A \sin \left( \frac{2\pi}{T} t \right) \cdot \frac{2\pi}{T}
\end{align*}
\]

\[ \implies v(t) = -\frac{2\pi A}{T} \sin \left( \frac{2\pi}{T} t \right), \]

\[
\begin{align*}
a(t) &= \frac{d}{dt} v(t) \\
&= \frac{d}{dt} \left[ -\frac{2\pi A}{T} \sin \left( \frac{2\pi}{T} t \right) \right] \\
&= -\frac{2\pi A}{T} \cos \left( \frac{2\pi}{T} t \right) \cdot \frac{d}{dt} \left[ \frac{2\pi}{T} t \right] \\
&= -\frac{2\pi A}{T} \cos \left( \frac{2\pi}{T} t \right) \cdot \frac{2\pi}{T}
\end{align*}
\]

\[ \implies a(t) = -A \left( \frac{2\pi}{T} \right)^2 \sin \left( \frac{2\pi}{T} t \right). \]

It is common in the study of oscillating functions to define \( \omega = \frac{2\pi}{T} \) so that the above computations can be summarized

\[ x = A \cos(\omega t) \]

\[ \implies v = -A \omega \sin(\omega t) \]

\[ \implies a = -A \omega^2 \cos(\omega t). \]

In such a case, the period is given by \( T = \frac{2\pi}{\omega} \). Note also that \( a(t) = -\omega^2 x(t) \). This is a defining feature of *simple harmonic oscillators*, namely that the force is proportional to but in the opposite direction of the position. (Recall force = mass \( \times \) acceleration.) This is the case with most springs, which conform to Hooke’s Law, \( F = -kx \), where \( k > 0 \) is a constant, \( x \) is the position relative to the equilibrium (rest) position of the spring and \( F \) is the force it exerts.\(^44\)

\(^{44}\)Named for Robert Hooke(1635–1703), an English inventor, professor, natural philosopher and architect. His 1678 statement of his law in words read, *Ut tensio, sic vis* (Latin), meaning, "As the extension, so the force". Hooke was a contemporary of Newton’s, and was very much involved in many of the significant discussions of his time. For instance, Newton gave him some credit for reviving Newton’s interest in astronomical mechanics, leading to Newton’s development of the inverse square law of gravitational attraction, which mirrors Coulomb’s electrostatic law (4.21), page 330, though there are no cases of gravitational repulsion in Newton’s law (only attraction).
CHAPTER 4. THE DERIVATIVE

Exercises

For 1–6, use the power rule to compute the derivative, after re-writing the problem. In particular, you can use \( \frac{d}{dx}(ax^n) = a \cdot nx^{n-1} \).

1. \( \frac{d}{dx} \left[ \frac{1}{x^{11}} \right] \)

2. \( \frac{d}{dx} \left[ \frac{1}{\sqrt{x}} \right] \)

3. \( \frac{d}{dx} \left[ \sqrt[3]{x^4} \right] \)

4. \( \frac{d}{dx} \left[ \frac{6}{x} \right] \)

5. \( \frac{d}{dt} \left[ \frac{1}{2t^2} \right] \)

6. \( \frac{d}{dy} \left[ \sqrt[3]{9y} \right] \) Recall \((ab)^n = a^n b^n \) anytime \( a, b \geq 0 \).

For 7–10, compute the given derivative.

7. \( \frac{d}{dx} \left[ (1 - 9x)^{11} \right] \)

8. \( \frac{d}{dx} \left[ 27(3x^2 - 10x + 55)^2 \right] \).

9. \( \frac{d}{dx} \left[ \sqrt[3]{2x^5 - 1} \right] \)

10. \( \frac{d}{dx} \left( 3x + 1 \right)^2 \) (Do this two ways: using the chain rule, and by first expanding the square. Compare your answers.)

For 11–14, compute \( f'(x) \).

11. \( f(x) = (x + 5)^{100} \).

12. \( f(x) = (2x + 5)^{100} \).

13. \( f(x) = \frac{1}{(x^4 - x + 1)^3} \).

14. \( f(x) = \sqrt{x + \sqrt{x + \sqrt{x}}} \).

Compute the following derivatives.

15. \( \frac{d}{dx} \left[ \sin z \right] \)

16. \( \frac{d}{dt} \left[ \cos \theta \right] \)

17. \( \frac{d}{dt} \left[ x^7 \right] \)

18. \( \frac{d}{dx} \left[ \sin \left( \cos x \right) \right] \frac{d}{dx} \left[ \cos x \right] \)

For 19–23, compute the given derivative.

19. \( \frac{d}{dx} \left[ \sin \sqrt{x} \right] \)

20. \( \frac{d}{dx} \left[ \sin x \right] \)

21. \( \frac{d}{dx} \left[ \sin \left( \cos x \right) \right] \)

22. \( \frac{d}{dx} \left[ \cos^3 x \right] \)

23. \( \frac{d}{dx} \left[ \cos(x + \cos x) \right] \)

24. Find \( h'(9) \) if \( h(x) = f(g(x)), g(9) = 5, g'(9) = 2, \) and \( f'(5) = 7 \).

25. On the unit circle, \( y^2 = 1 - x^2 \). If we take either the upper semicircle or the lower semicircle, then \( y \) is also a function of \( x \). Find the tangent line to the graph at the point \( (3/5, 4/5) \) by finding \( \frac{dy}{dx} \) two ways:

(a) Using \( y = \sqrt{1 - x^2} \) for the upper semicircle, and the chain rule.

(b) By applying \( \frac{d}{dx} \) to both sides of \( y^2 = 1 - x^2 \), as we did in the proof of Theorem 4.4.2, and then solving for \( \frac{dy}{dx} \), and plugging into that expression \( (x, y) = (3/5, 4/5) \).
26. Using $\sec x = (\cos x)^{-1}$,
   
   (a) derive $\frac{d}{dx} \sec x = \sec x \tan x$.

   (b) Use (a) and the chain rule to compute $\frac{d}{dx} \sqrt{x^2 + 1}$.

27. Show that if $k$, $q_1$, $q_2$ are constant in Coulomb’s law (4.21), page 330, then $\frac{dE}{dr} = -2 \frac{E}{r}$.

28. As a charge $q_2$ moves away from a stationary charge $q_1$, the instantaneous rate of change of the Coulomb force $F$ with respect to $r$ is measured to be 5N/m.

   (a) Find the magnitude of the force between the charges when they are 0.3m apart. (See Example 4.4.16, page 362.)

   (b) If we are given that $k = 9 \times 10^9$ N·m²/C, and $q_2 = 3 \times 10^{-6}$C, find the equation for the force between the two particles when separated by a distance $r$.

   (c) Find $\frac{dF}{dr}$ from (b) when $r = 0.25$M.

29. The gravitational force of attraction between any two masses is directly proportional to the product of the masses and inversely proportional to the square of the distance between them. In particular,

   $$F = G \frac{m_1 m_2}{r^2}, \quad (4.45)$$

   where $G \approx 6.67 \times 10^{-11}$N(m²/kg)² is the universal gravitational constant, $m_1$, $m_2$ are in kg, and $r$ is in m. Find the rate of change of $F$ with respect to distance $r$ (i.e., find $dF/dr$) when $r = 0.5$m.

30. An object moves along the $x$-axis according to the equation $s(t) = t + 2 \sin t$. Find those values of $t$ when the object is moving to the left, i.e., find where $s'(t) < 0$, within the interval $t \in [0, 2\pi]$.

31. Suppose that the cost in thousands of dollars to manufacture $x$ thousands of items is given by

   $$C(x) = 0.003x^3 + 0.2x^2 + 0.3x + 70,$$

   where $x \in [0, 10]$.

   (a) Find the (proxy) marginal cost function, $C'(x)$.

   (b) Determine the marginal cost at $x = 4$ (in units of thousands of dollars per thousands of items, which simplifies to dollars per item).\(^{45}\)

   (c) Find the average rate of change of cost in going from $x = 3.9$ to $x = 4.1$.

   (d) Compare the answers to parts (b) and (c).

32. Consider the following alternating current circuit:

   $$\begin{align*}
   V & \quad \text{R} \\
   & \quad \text{L}
   \end{align*}$$

   The current $I$ where the resistor $R$ and an inductor $L$ are present is given by

   $$I = \frac{V}{\sqrt{R^2 + (2\pi f L)^2}},$$

   where $V$ is the voltage and $f$ is the current’s frequency of alternation (in cycles/second, or Hertz, also written Hz). Assuming that $V$, $R$ and $f$ are constants, find an equation for the instantaneous rate of change of current $I$ with respect to inductance $L$.

---

\(^{45}\)Here marginal cost would represent the cost of the next thousand items.
4.5 Product, Quotient and Other Trigonometric Rules

In this section we first introduce the rule for the differentiation of a product of two functions. From that and the chain rule, we will derive a rule for differentiating a quotient of two functions. With a quotient rule we will be able to use the rules for $\sin x$ and $\cos x$ to derive rules for $\tan x$ and $\cot x$. For completeness we will also compute the rules for $\sec x$ and $\csc x$ and thus finish our rules for the six basic trigonometric functions.

The rules for calculating the derivative of a product or a quotient are not as simple as for a sum or difference. However they are straightforward when applied correctly.

4.5.1 Product Rule Stated and First Applied

We begin with the statement and some discussion of the product rule, followed by several examples demonstrating its mechanics. The actual proof we leave until the next subsection.

**Theorem 4.5.1 (Product Rule)** At each $x$ for which $\frac{d}{dx}f(x)$ and $\frac{d}{dx}g(x)$ exist, so does the derivative $\frac{d}{dx}(f(x) \cdot g(x))$ exist, and it is given by

$$\frac{d}{dx}[f(x) \cdot g(x)] = f(x)\frac{d}{dx}g(x) + g(x)\frac{d}{dx}f(x). \tag{4.46}$$

Though we defer the proof, we can make a couple of observations.

- This is not simply the product of the two derivatives. (See for example Exercise 2 in Section 4.2, page 324.)

- Recall that multiplicative constants are preserved in the derivative: $\frac{d}{dx}(Cf(x)) = C\frac{d}{dx}f(x)$. One could say that the constant “amplifies” the function $f(x)$ by the factor $C$, causing the same amplification factor the rate of change, or derivative, of the new function $Cf(x)$. (For example, $C = 2$ doubles the function, and thus doubles the rate of change.)

Next notice that the first term of (4.46) treats $f(x)$ as though it were a constant amplifying the change in (i.e., derivative of) $g(x)$, while the second term treats $g(x)$ as though it were a constant amplifying the change in $f(x)$. In this way the product rule accounts for the changes in each function, as amplified by the other. A close scrutiny of the proof shows how this emerges.

Our first example shows how the product rule gives us what we expect for a very simple case.

**Example 4.5.1** Let $f(x) = x^5$. Then $f'(x) = 5x^4$ from the power rule. But we can also write $f(x) = x^3 \cdot x^2$, from which the product rule gives

$$\frac{d}{dx}[x^3 \cdot x^2] = x^3 \cdot \frac{d}{dx}(x^2) + x^2 \cdot \frac{d}{dx}(x^3) = x^3 \cdot 2x + x^2 \cdot 3x^2 = 2x^4 + 3x^4 = 5x^4.$$

Of course the product rule will be of much more use than proving things we already knew. The next example requires the product rule (or some very clever tricks with difference quotients!):

**Example 4.5.2** Suppose $f(x) = x^2 \sin x$. This is a product of two differentiable functions. Its derivative is given by

$$f'(x) = \frac{d}{dx}(x^2 \sin x) = x^2 \cdot \frac{d}{dx}(\sin x) + \sin x \cdot \frac{d}{dx}(x^2) = x^2 \cos x + \sin x \cdot 2x = x(x \cos x + 2 \sin x)$$

The last step was just an algebraic one, factoring the final answer as much as possible.

---

46If the functions are not differentiable, this fact appears as we take the derivatives on the right-hand side of the product rule statement (4.40). Thus we usually just apply the rule—instead of checking differentiability first.
Now we list several simple examples to illustrate the basic mechanics of the product rule.

- \( \frac{d}{dx} \left[ (3x^2 + 5x - 9)(5x^3 + 7x^2 + 27x - 4) \right] \)
  
  \[
  = (3x^2 + 5x - 9) \frac{d}{dx} (5x^3 + 7x^2 + 27x - 4) + (5x^3 + 7x^2 + 27x - 4) \frac{d}{dx} (3x^2 + 5x - 9)
  \]

- \( \frac{d}{dx} (x \cos x) = x \frac{d}{dx} \cos x + \cos x \cdot \frac{d}{dx} x = x(-\sin x) + \cos x \cdot 1 = -x \sin x + \cos x. \)

- \( \frac{d}{dt}[PV] = P \cdot \frac{dV}{dt} + V \cdot \frac{dP}{dt}. \)

One of the interesting aspects of the calculus is the various ways that the consistency of differentiation (derivative-taking) rules can be seen by using different strategies for particular derivatives. Earlier we showed how to use the product rule to compute \( \frac{d}{dx}(x^2 \cdot x^3) = \cdots = 5x^4, \) which we can also compute with the power rule \( \frac{d}{dx}(x^5) = 5x^4. \) We can also see how the product rule gives us the behavior of multiplicative constants in derivative computations. For example, instead of writing \( \frac{d}{dx}(2 \sin x) = 2 \frac{d}{dx} \sin x = 2 \cos x, \) we can consider \( 2 \sin x \) to be the product of two functions, and compute the derivative using the product rule if we care to:

\[
\frac{d}{dx}(2 \sin x) = 2 \frac{d}{dx} \sin x + \sin x \cdot \frac{d}{dx} (2) = 2 \cos x + \sin x \cdot 0 = 2 \cos x + 0 = 2 \cos x,
\]

as before. Of course the rule on multiplicative constants (page 311) is faster.

Because the product rule calls for the computation of derivatives of the factors, it often “calls” upon other rules to compute these component derivatives. Conversely, other rules may call upon the product rule. As we saw with the chain rule, it is crucial that we look at the overall structure of a function to see which rule to apply first, and then work our way in towards the inner structures as the differentiation rules require in their turns; we compute derivatives “from the outside to the inside.” The next two examples are product rules first, which then call the chain rule.

**Example 4.5.3** Suppose \( f(x) = \sin x^2 \cos x^3. \) This is foremost a product of two functions, so we need the product rule first.\(^{47}\)

\[
f'(x) = \frac{d}{dx} \left[ \sin x^2 \cos x^3 \right]
= \sin x^2 \cdot \frac{d}{dx} \cos x^3 + \cos x^3 \cdot \frac{d}{dx} \sin x^2 \quad \text{(Product Rule)}
= \sin x^2 \left( -\sin x^3 \cdot \frac{d}{dx} x^3 \right) + \cos x^3 \cdot \left( \cos x^2 \cdot \frac{d}{dx} x^2 \right) \quad \text{(Chain Rule, twice)}
= (\sin x^2)(-\sin x^3)(3x^2) + \cos x^3 \cos x^2 \cdot 2x
= -3x^2 \sin x^2 \sin x^3 + 2x \cos x^3 \cos x^2. \quad \text{(Rearrangement)}
\]

Thus, when we took the derivatives called for by the product rule, these required the chain rule.

(We could have factored the final computation but it is not necessary.)

\(^{47}\)Note that \( \sin x^2 \cos x^3 \) is taken to be a product. Indeed, it is understood that the sine and cosine functions here are separate factors. Also note that \( x^2 \) is the input of the sine, and \( x^3 \) the input of the cosine. The convention is to understand this function, as written, in the following way:

\[
\sin x^2 \cos x^3 = (\sin x^2)(\cos x^3) = \left[ \sin(x^2) \right] \left[ \cos(x^3) \right].
\]
For a polynomial example, consider the following:

**Example 4.5.4** \( f(x) = (x^2 + 2x + 3)^2(x^2 + 1)^3 \). Without the product rule we would be forced to carry out the multiplications, but since this is written as a product of two functions, we can instead use the product rule. (The calculus is finished in four lines; the rest is algebra, which is optional.)

\[
\begin{align*}
    f'(x) &= \frac{d}{dx} [(x^2 + 2x + 3)^2(x^2 + 1)^3] \\
    &= (x^2 + 2x + 3)^2 \cdot \frac{d}{dx}(x^2 + 1)^3 + (x^2 + 1)^3 \cdot \frac{d}{dx}(x^2 + 2x + 3) \\
    &= (x^2 + 2x + 3)^2 \cdot 3(x^2 + 1)^2 \cdot \frac{d}{dx}(x^2 + 1) + (x^2 + 1)^3 \cdot 2(x^2 + 2x + 3) \frac{d}{dx}(x^2 + 2x + 3) \\
    &= 6x(x^2 + 2x + 3)^2(x^2 + 1)^2 + (4x + 4)(x^2 + 1)^3(x^2 + 2x + 3) \\
    &= (x^2 + 2x + 3)(x^2 + 1)^2 \left[ 6x^3 + 12x^2 + 18x + 4x^3 + 4x + 4x^2 + 4 \right] \\
    &= (x^2 + 2x + 3)(x^2 + 1)^2 \left[ 10x^3 + 16x^2 + 22x + 4 \right].
\end{align*}
\]

Again, the statement of the product rule here called for derivatives of the factors, and each of those required a chain rule. Note how one factor of \((x^2 + 2x + 3)\) and two factors of \((x^2 + 1)\) could be factored from each term. Such algebraic manipulations are useful, for instance, if a sign chart for the derivative is desired.

It is also possible that a product rule can occur within a chain rule, as in the following.

**Example 4.5.5** Suppose \( f(x) = \sqrt{\sin x \cos x} \). Then

\[
\begin{align*}
    f'(x) &= \frac{d}{dx} \sqrt{\sin x \cos x} \\
    &= \frac{1}{2\sqrt{\sin x \cos x}} \cdot \frac{d}{dx}(\sin x \cos x) \\
    &= \frac{1}{2\sqrt{\sin x \cos x}} \cdot \left[ \sin x \frac{d}{dx} \cos x + \cos x \frac{d}{dx} \sin x \right] \\
    &= \frac{1}{2\sqrt{\sin x \cos x}} \cdot [\sin x (-\sin x) + \cos x \cos x] \\
    &= \frac{-\sin^2 x + \cos^2 x}{2\sqrt{\sin x \cos x}}.
\end{align*}
\]

This is not the only method for solving this problem, but it is perhaps the most straightforward.\(^\text{48}\)

We can also use these product rule-derived derivatives to help graph functions.

\(^\text{48}\)Actually, with trigonometry we can rewrite the problem and the answer, using \( \sin 2\theta = 2\sin \theta \cos \theta \) and \( \cos 2\theta = \cos^2 \theta - \sin^2 \theta \). Below, “\( \Rightarrow \)” represents another chain rule problem (calling yet another chain rule).

\[
\begin{align*}
    f(x) &= \sqrt{\frac{1}{2} \sin 2x} \quad \Rightarrow \quad f'(x) = \cdots = \frac{\cos 2x}{2\sqrt{\frac{1}{2} \sin 2x}}.
\end{align*}
\]
Example 4.5.6 \( f(x) = x\sqrt{1-x^2} \). This function we will differentiate and then graph.

\[
f'(x) = x \cdot \frac{d}{dx} \sqrt{1-x^2} + \sqrt{1-x^2} \cdot \frac{d}{dx} (x)
= x \cdot \frac{1}{2\sqrt{1-x^2}} \cdot d\sqrt{1-x^2} + \sqrt{1-x^2} \cdot 1
= x \cdot \frac{1}{2\sqrt{1-x^2}} \cdot (-2x) + \sqrt{1-x^2}
= \frac{-x^2}{\sqrt{1-x^2}} + \sqrt{1-x^2}.
\]

To graph this function we would like to know where it is increasing and where it is decreasing and thus locate local extrema. But even before delving into the derivative, we can first notice the domain of \( f(x) \) is \(-1 \leq x \leq 1\), and the function (height) itself is zero at \( x = 0, \pm 1 \) (x-intercepts).

Next we proceed to see where \( f' > 0 \) and \( f' < 0 \). For such a task, it is best if the derivative is written as a single fraction:

\[
f'(x) = \frac{-x^2}{\sqrt{1-x^2}} + \sqrt{1-x^2} \cdot \frac{\sqrt{1-x^2}}{\sqrt{1-x^2}} = \frac{-x^2 + 1 - x^2}{\sqrt{1-x^2}} = \frac{1 - 2x^2}{\sqrt{1-x^2}} \text{ Optional } \frac{(1-\sqrt{2}x)(1+\sqrt{2}x)}{\sqrt{1-x^2}}.
\]

Now we see that this is undefined (\( f' \text{ DNE} \)) except for \(-1 < x < 1\). The fraction which is \( f' \) is zero exactly where the numerator is zero and the denominator is not. Thus

\[
f'(x) = 0 \iff 1 - 2x^2 = 0 \iff 1 = 2x^2 \iff \frac{1}{2} = x^2 \iff x = \pm \frac{1}{\sqrt{2}} \approx \pm 0.7071.
\]

From this we can make a sign chart for \( f' \) to see where \( f \) is increasing/decreasing.

\[
\begin{array}{c|c|c|c}
Test & x & -0.9 & 0 & 0.9 \\
Factors f'(x): & \Theta/\oplus & 0 & \Theta/\oplus & \Theta/\oplus \\
Sign f'(x): & -1 & \Theta & -\frac{1}{\sqrt{2}} & \oplus & \frac{1}{\sqrt{2}} & \ominus & 1 \\
Behavior of f(x): & \text{DEC} & \text{INC} & \text{DEC} & \\
\end{array}
\]

We see a local minimum at \( x = -1/\sqrt{2} \), and a local maximum at \( x = 1/\sqrt{2} \). The actual points are

\[
\left(-\frac{1}{\sqrt{2}}, f\left(-\frac{1}{\sqrt{2}}\right)\right) = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \cdot \sqrt{1/2}\right) \approx \left(-0.7071, -\frac{1}{2}\right),
\]

\[
\left(\frac{1}{\sqrt{2}}, f\left(\frac{1}{\sqrt{2}}\right)\right) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \cdot \sqrt{1/2}\right) \approx \left(0.7071, \frac{1}{2}\right).
\]

All this behavior leads us to the graph, which is given in Figure 4.18. Notice the (computer generated) graph there also reflects that \( f'(x) = \frac{1 - 2x^2}{\sqrt{1-x^2}} \rightarrow -\infty \text{ as } x \rightarrow -1^+ \text{ or } x \rightarrow 1^- \).
4.5.2 Product Rule Proof

For completeness we include here a proof of the product rule.

The proof of the product rule is not accomplished so much by “brute force,” but instead utilizes some clever rewriting.

It is normal for questions of why one thinks of the “trick” used to make it work, but these should not immediately distract from the fact that it does. Many of the proofs used today have been condensed over the decades, or even centuries since the first proofs, and are therefore quite short because revisits to earlier proofs naturally lead us to shortcuts. As a result, proofs often look less like the natural paths of discovery and more like terse explanations. Nonetheless there is knowledge to be gained from even these short proofs—for instance, the “trick” may be useful in another context—and so they are worth reading and understanding, though again we will almost always just quote the results—without reference to their proofs—when solving problems.

The proof of the product rule depends upon another theorem which is intuitive, is important in its own right, and has its own short, somewhat clever proof. In sum, the theorem says that to have a well-defined slope at $x = a$, a function must also be continuous there.

**Theorem 4.5.2** $(f'(a)$ exists$) \implies (f(x)$ is continuous at $x = a)$.  

**Proof:** Recall that

$$f'(a) \text{ exists} \iff f'(a) = \lim_{\Delta x \to 0} \frac{f(a + \Delta x) - f(a)}{\Delta x} \in \mathbb{R},$$

i.e., the limit exists as a (finite) real number. We need to show that this implies $f(x)$ is continuous at $x = a$, which is equivalent to $\lim_{x \to a} f(x) = f(a)$ (Theorem 3.4.2, page 211). First we re-write this limit using the substitution $x = a + \Delta x$, which gives $x \to a \iff \Delta x \to 0$ properly. Then we perform an algebraic expansion of the argument of the limit by subtracting and adding $f(a)$ (which exists and is real or the
above limit could not exist and be finite), and divide and multiply by $\Delta x$, to get

$$
\lim_{x \to a} f(x) = \lim_{\Delta x \to 0} f(a + \Delta x) = \lim_{\Delta x \to 0} \left[ f(a + \Delta x) - f(a) + f(a) \right] = \lim_{\Delta x \to 0} \left[ \frac{f(a + \Delta x) - f(a)}{\Delta x} \cdot \Delta x + f(a) \right] = f'(a) \cdot 0 + f(a) = f(a), \quad \text{q.e.d.}
$$

Note that the last line of the proof used the fact that the difference quotient approached the finite number $f'(a)$, and so the limit form was “$f'(a) \cdot 0 + f(a)$,” yielding $f(a)$.

This theorem is sometimes described as “differentiability implies continuity.” In fact differentiability is a stronger criterion than continuity. This result is also interesting in its contrapositive form (recall $P \implies (\sim Q) \iff (\sim P)$):

$$
f(x) \text{ discontinuous at } x = a \implies f'(x) \text{ DNE at } x = a.
$$

To paraphrase, at the point $x = a$, to have a tangent line the function must be continuous, and equivalently, a function which is discontinuous can not have a tangent line. Now we use Theorem 4.5.2 and some algebraic tricks to prove the product rule.

**Proof: (Product Rule)** Suppose $f(x)$ and $g(x)$ are both differentiable at a given $x$, i.e., $f'(x)$ and $g'(x)$ both exist. Then $f$ and $g$ are both continuous at $x$, and

$$
\frac{d}{dx} \left[ f(x)g(x) \right] = \lim_{\Delta x \to 0} \frac{f(x + \Delta x)g(x + \Delta x) - f(x)g(x)}{\Delta x} = \lim_{\Delta x \to 0} \left[ f(x + \Delta x) \cdot \frac{g(x + \Delta x) - g(x)}{\Delta x} + g(x) \cdot \frac{f(x + \Delta x) - f(x)}{\Delta x} \right] = f(x)g'(x) + g(x)f'(x), \quad \text{q.e.d.}
$$

The last line of the proof follows because as $\Delta x \to 0$, by continuity (which the previous theorem gives us from the differentiability) we have $f(x + \Delta x) \to f(x)$, and the two difference quotients approach $f'(x)$ and $g'(x)$ respectively, while $g(x)$ is constant in the limit (which is in $\Delta x$, not $x$). The middle two lines were simply algebra, with the “clever trick” in the second line using the fact that $AB - CD = A(B - D) + D(A - C)$, except here it was with $f(x + \Delta x)g(x + \Delta x) - f(x)g(x) = \underbrace{f(x + \Delta x)g(x + \Delta x)}_A - \underbrace{f(x)g(x)}_B = \underbrace{f(x)g'(x)}_A + \underbrace{g(x)f'(x)}_B$.

### 4.5.3 Quotient Rule

We often need to find derivatives of functions of the form $h(x) = f(x)/g(x)$. We can rewrite these as $h(x) = f(x)(g(x))^{-1}$, and use the product rule, which will then call the chain rule, to
get

\[ h'(x) = \frac{d}{dx} \left[ f(x)(g(x))^{-1} \right] \]

\[ = f(x) \frac{d}{dx} [(g(x))^{-1}] + (g(x))^{-1} \frac{d}{dx} f(x) \]

\[ = f(x) \left[ (-1)(g(x))^{-2} \frac{d}{dx} g(x) \right] + (g(x))^{-1} \frac{d}{dx} f(x) \]

\[ = -\frac{f(x) \frac{d}{dx} g(x)}{(g(x))^2} + \frac{\frac{d}{dx} f(x)}{g(x)} \]

\[ = -\frac{f(x) \frac{d}{dx} g(x)}{(g(x))^2} + \frac{g(x) \frac{d}{dx} f(x)}{(g(x))^2}. \]

Combining the two fractions and putting the term with the negative sign (−) second, we can write:

**Theorem 4.5.3** If \( f \) and \( g \) are differentiable at \( x \), and \( g(x) \neq 0 \), then

\[ \frac{d}{dx} \left[ \frac{f(x)}{g(x)} \right] = \frac{g(x) \frac{d}{dx} f(x) - f(x) \frac{d}{dx} g(x)}{(g(x))^2}. \] (4.47)

As noted in the derivation, this rule is actually redundant given the availability of the product and quotient rules. However, it is useful especially because the resulting derivative emerges as an already combined fraction.

**Example 4.5.7** Find \( f'(x) \) if \( f(x) = \frac{\sin x}{x} \).

**Solution**: Using the quotient rule we have

\[ f'(x) = x \frac{d}{dx} \frac{\sin x}{x} - \frac{\sin x \frac{d}{dx} x}{(x)^2} = \frac{x \cos x - \sin x}{x^2}. \]

The quotient rule is especially useful for rational functions, i.e., ratios of polynomials.

**Example 4.5.8** Suppose \( f(x) = \frac{x^3 - 8}{x^2 - 9} \). Then

\[ f'(x) = \frac{(x^2 - 9) \frac{d}{dx} (x^3 - 8) - (x^3 - 8) \frac{d}{dx} (x^2 - 9)}{(x^2 - 9)^2} \]

\[ = \frac{(x^2 - 9)(3x^2) - (x^3 - 8)(2x)}{(x^2 - 9)^2} \]

\[ = \frac{3x^4 - 27x^2 - 2x^4 + 16x}{(x^2 - 9)^2} \]

\[ = \frac{x^4 - 27x^2 + 16x}{(x^2 - 9)^2}. \]

As with the product rule, the quotient rule can be embedded within a chain or product rule, or vice-versa. The following uses a rule derived in the Exercise 26a, page 367, namely that \( \frac{d}{dx} \sec x = \sec x \tan x \).
4.5. PRODUCT, QUOTIENT AND OTHER TRIGONOMETRIC RULES

Example 4.5.9 Suppose \( f(x) = \sec \left( \frac{x}{x-1} \right) \). Then

\[
f'(x) = \frac{d}{dx} \sec \left( \frac{x}{x-1} \right) = \sec \left( \frac{x}{x-1} \right) \tan \left( \frac{x}{x-1} \right) \cdot \frac{d}{dx} \left( \frac{x}{x-1} \right)
= \sec \left( \frac{x}{x-1} \right) \tan \left( \frac{x}{x-1} \right) \cdot \frac{(x-1)(1)-(x)(1)}{(x-1)^2}
= \sec \left( \frac{x}{x-1} \right) \tan \left( \frac{x}{x-1} \right) \cdot \frac{x-1-x}{(x-1)^2}
= -\frac{1}{(x-1)^2} \sec \left( \frac{x}{x-1} \right) \tan \left( \frac{x}{x-1} \right).
\]

In fact the quotient rule computation for this particular problem could have been avoided through long division, giving \( \frac{d}{dx} \left( \frac{x}{x-1} \right) = \frac{d}{dx} \left( 1 + \frac{1}{x-1} \right) \), making for a simple power/chain rule, but the quotient rule was straightforward, and left that factor as a single fraction. For an example of chain rules inside a quotient rule, consider the next example. Recall \( \frac{d}{dx} \sqrt{x} = \frac{d}{dx} x^{1/2} = \frac{1}{2} x^{-1/2} = 1/(2\sqrt{x}) \).

Example 4.5.10 Suppose \( f(x) = \frac{\cos 2x}{\sqrt{x^2-1}} \). Then

\[
f'(x) = \frac{\sqrt{x^2-1} \frac{d}{dx} \cos 2x - \cos 2x \cdot \frac{d}{dx} \sqrt{x^2-1}}{(\sqrt{x^2-1})^2}
= \frac{\sqrt{x^2-1} \cdot (-\sin 2x \cdot \frac{d}{dx} (2x)) - \cos 2x \cdot \frac{1}{2\sqrt{x^2-1}} \cdot \frac{d}{dx} (x^2-1)}{x^2-1}
= \frac{x^2-1}{x^2-1} \cdot (\sin 2x \cdot (-2)) - \frac{\cos 2x \cdot 2x}{2\sqrt{x^2-1}} \cdot \frac{\sqrt{x^2-1}}{x^2-1}
= -2(x^2-1) \sin 2x - x \cos 2x
= \frac{(x^2-1)^{3/2}}{(x^2-1)^{3/2}}.
\]

It is an interesting exercise in both calculus and algebraic simplification to derive the same conclusion using \( f(x) = \cos 2x \cdot (x^2-1)^{-1/2} \) and the product rule (which would also call the chain rule twice).
4.5.4 Tangent, Cotangent, Secant and Cosecant Rules

The following are derivative rules for the remaining trigonometric functions. These rules are given in both simple (“matching variable”) and chain rule versions.

\[
\begin{align*}
\frac{d}{dx} \tan x &= \sec^2 x, \\
\frac{d}{dx} \cot x &= -\csc^2 x, \\
\frac{d}{dx} \sec x &= \sec x \tan x, \\
\frac{d}{dx} \csc x &= -\csc x \cot x,
\end{align*}
\]

These should all be memorized. It may help to notice patterns when comparing the derivatives of tangent and cotangent, secant and cosecant, and how these are similar to the comparison of sine and cosine derivatives. In short, these formulas come in function/cofunction pairs.

We will prove the derivative of \( \tan x \) is \( \sec^2 x \) and leave the rest as exercises. The chain rule versions then follow. To see the formula for the tangent, we rewrite it as the quotient of sine and cosine, and use the quotient rule.

\[
\frac{d}{dx} \tan x = \frac{d}{dx} \left( \frac{\sin x}{\cos x} \right) = \frac{\cos x \cdot \frac{d}{dx} \sin x - \sin x \cdot \frac{d}{dx} \cos x}{(\cos x)^2} = \frac{\cos x \cos x - (\sin x)(-\sin x)}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x.
\]

Chain Rule Version: \( \frac{d}{dx} \tan u = \frac{d}{du} \tan u \cdot \frac{du}{dx} = \sec^2 u \cdot \frac{du}{dx} \).

The fourth line above used some trigonometric identities \( \sin^2 \theta + \cos^2 \theta = 1, 1/\cos \theta = \sec \theta \). The last line was our usual chain rule argument, given \( \frac{d}{dx} \tan x = \sec^2 x \) was already proved. With (4.48)–(4.51), and derivatives of sine and cosine from earlier((4.18), (4.19), page 318), we finally have derivatives of all six trigonometric functions.

Every student of calculus should memorize the derivatives of the trigonometric functions, and be able to derive these new ones from knowing the derivatives of \( \sin x \) and \( \cos x \).

Now we can apply these. First we look at some of the simpler examples.

**Example 4.5.11**

- \( \frac{d}{dx} \sec x^9 = \sec x^9 \tan x^9 \cdot \frac{d}{dx} (x^9) = \sec x^9 \tan x^9 \cdot 9x^8 = 9x^8 \sec x^9 \tan x^9 \).

- \( \frac{d}{dx} x^2 \tan x = x^2 \cdot \frac{d}{dx} \tan x + \tan x \cdot \frac{d}{dx} (x^2) = x^2 \sec^2 x + 2x \tan x \).

- \( \frac{d}{dx} \left( \frac{x}{\tan x} \right) = \frac{d}{dx} (x \cot x) = x \cdot \frac{d}{dx} (\cot x) + \cot x \cdot \frac{d}{dx} (x) = (x)(-\csc^2 x) + \cot x \)
  \( = -x \csc^2 x + \cot x \).
4.5. PRODUCT, QUOTIENT AND OTHER TRIGONOMETRIC RULES

Figure 4.19: Partial graph of \( y = \tan x \). Since \( y = \frac{\sin x}{\cos x} \), there are vertical asymptotes at each \( x \)-value where \( \cos x = 0 \) (and \( \sin \theta = \pm 1 \)). Recall that \( \tan x \) is positive if \( x \) represents an angle in the first or third quadrants, and negative in the second and fourth quadrants, so the quadrants represented by the \( x \)-values are labeled QI–QIV. Note that \( \frac{d}{dx} \tan x = \sec^2 x \) which is positive where defined (coinciding with where \( \tan x \) is defined), and thus \( \tan x \) is an increasing function where defined.

Note that we turned a quotient rule into a product rule for the last derivative problem.

Before continuing with more complicated examples, we should briefly consider why it makes sense graphically that \( \frac{d}{dx} \tan x = \sec^2 x \). The graph of \( f(x) = \tan x \) is given in Figure 4.19, and so we can consider its derivative formula in light of that graph. Now that derivative is always positive where defined; \( \sec^2 x > 0 \), and in fact \( \sec^2 x = \frac{1}{\cos^2 x} \geq 1 \). Thus \( \tan x \) is always increasing in any interval on which it is defined. Furthermore \( \sec^2 x \to \infty \) as \( x \to \pm \frac{\pi}{2}, \pm \frac{3\pi}{2}, \pm \frac{5\pi}{2} \), etc., and that has implications for the graph. Of course \( \tan x = \sin x / \cos x \) has vertical asymptotes at each of those \( x \)-values (where \( \cos x = 0 \) and \( \sin x = \pm 1 \)). Finally, note that \( \frac{d\tan x}{dx} \bigg|_{x=0} = \sec^2 x \bigg|_{x=0} = \frac{1}{\cos^2 0} = 1 \), for instance, so the slope through the \((0, \tan 0) = (0, 0)\) is 1. That slope repeats every \( \pi \) in both directions, due to the \( \pi \)-periodic nature of the tangent.

Example 4.5.12 Here is a typical chain rule problem involving the tangent.

\[
\frac{d}{dx} \tan \sqrt{x^2 - 1} = \sec^2 \sqrt{x^2 - 1} \cdot \frac{d}{dx} \sqrt{x^2 - 1} = \sec^2 \sqrt{x^2 - 1} \cdot \frac{1}{2\sqrt{x^2 - 1}} \cdot \frac{d(x^2 - 1)}{dx} \\
= \frac{\sec^2 \sqrt{x^2 - 1}}{2\sqrt{x^2 - 1}} \cdot 2x = \frac{x \sec^2 \sqrt{x^2 - 1}}{\sqrt{x^2 - 1}}.
\]

Note that, tempting as it may be, the radicals above cannot be combined or canceled in any of the steps; one is outside the secant-squared function, and the other is safely quarantined inside. Note also that squaring the secant function does not alter its argument \( \sqrt{x^2 - 1} \).

Example 4.5.13 Another chain rule problem is the following. Note \( \cot^3 x = (\cot x)^3 \).

\[
\frac{d}{dx} \cot^3 x = 3(\cot x)^2 \frac{d}{dx} \cot x = 3 \cot^2 x(-\csc^2 x) = -3 \cot^2 x \csc^2 x.
\]
Example 4.5.14 One product rule problem is the following:

\[
\frac{d}{dx}(\sec x \tan x) = \sec x \cdot \frac{d}{dx} \tan x + \tan x \frac{d}{dx} \sec x \\
= \sec x \sec^2 x + \tan x \sec x \tan x \\
= \sec^3 x + \sec x \tan^2 x.
\]

There is so much algebraic structure built into the trigonometric functions that such an answer can be rewritten many different ways. For instance, since \(\tan^2 x + 1 = \sec^2 x\) our final answer can be written

\[
\sec x(\sec^2 x + \tan^2 x) = \sec x(\sec^2 x + \sec^2 x - 1) = \sec x(2 \sec^2 x - 1),
\]

for instance. Another alternative is \(\sec x(\tan^2 x + 1 + \tan^2 x) = \sec x(2 \tan^2 x + 1)\). When we study integration particularly, it is important to consider such options.

4.5.5 Putting Rules Together—Carefully

This subsection is just a reminder that, when computing the derivative of a complicated function it is quite possible to use several of the previous differentiation rules. In such cases we need to recognize which rules apply, and then exactly how to invoke them.

It is easy to lapse into intellectual laziness by skipping steps, but this is an error-prone habit which does not save any time in the long run. Some steps can be combined into other steps with little risk, especially with practice (the sum and additive and multiplicative constant rules come to mind). However, the quotient, product, power, trigonometric, and all forms of the chain rule should be written out in their own steps before we compute the derivatives internal to these rules. For instance, it is tempting to compute all at once:

\[
\frac{d}{dx}[(x^3 + 9x^2 + \sin 2x)(\tan x^5)]
\]

\[
= (x^3 + 9x^2 + \sin 2x)(\sec^2 x^5 \cdot 5x^4) + \tan x^5 \cdot (3x^2 + 18x + \cos 2x \cdot 2).
\]

However, this approach has a couple of disadvantages which become more important as functions become increasingly complicated. First, we have to keep track of what rule applies where, without the benefit of breaking it into steps. Second, if we would like to check our work we can try to re-read what we wrote, but we find ourselves again performing the same mental gymnastics we did the first time, and likely repeating any mistakes we made that first time. We can go a long way towards avoiding these difficulties by writing out all the steps. Since each step invokes a single differentiation rule (though we may apply different rules to different terms in the same “step”), much of our work is recopying the line above, which takes very little time. Care in “bookkeeping” will translate into clearer thinking and less error (and easier error correction!).

Consider the following approach to the problem above:

\[
\frac{d}{dx}[(x^3 + 9x^2 + \sin x)(\tan x^5)]
\]

\[
= (x^3 + 9x^2 + \sin 2x) \frac{d}{dx} (\tan x^5) + (\tan x^5) \cdot \frac{d}{dx} (x^3 + 9x^2 + \sin 2x) \\
= (x^3 + 9x^2 + \sin 2x) \sec^2 x^5 \cdot \frac{d}{dx} x^5 + \tan x^5 \cdot (3x^2 + 18x + \cos 2x \cdot \frac{d}{dx} 2x) \\
= (x^3 + 9x^2 + \sin 2x)(\sec^2 x^5 \cdot 5x^4) + \tan x^5 \cdot (3x^2 + 18x + \cos 2x \cdot 2) \\
= 5x^4(3x^2 + 9x^2 + \sin 2x) \sec^2 x^5 + (3x^2 + 18x + 2 \cos 2x) \tan x^5.
Ultimately this given function is a product, so the first step was exactly the statement of the product rule. In the next line, we begin to take the derivatives demanded within the product rule, and find some power rules (with the multiplicative constants along for the ride—i.e., with the rule that multiplicative constants are preserved), and a couple of chain rules which we write out exactly. Next we compute the derivatives of the “inside” functions demanded by the chain rule, and finally we do some algebra to make the result more presentable. Note how we can re-read this with great assurance that it is correct, since each step follows differentiation rules in obvious (though the terms themselves may be complicated) ways.

In the examples which follow we will continue to write out all the steps, while being careful to invoke them in the proper order. In effect, we work from the outside (large-structure) inwards.

**Example 4.5.15** Find $f'(x)$ if $f(x) = 2 \sin^3 5x + \csc \sqrt{x^2 + 1}$.

This is first a sum, and then there are several chain rules which come into play.

\[
f'(x) = \frac{d}{dx} (2 \sin^3 5x) + \frac{d}{dx} \left( \csc \sqrt{x^2 + 1} \right)
\]

\[
= 2 \cdot 3 \sin^2 5x \cdot \frac{d}{dx} \sin 5x - \csc \sqrt{x^2 + 1} \cot \sqrt{x^2 + 1} \cdot \frac{1}{2\sqrt{x^2 + 1}} \cdot \frac{d}{dx} (x^2 + 1)
\]

\[
= 6 \sin^2 5x \cos 5x \cdot 5 - \frac{x \csc \sqrt{x^2 + 1} \cot \sqrt{x^2 + 1}}{\sqrt{x^2 + 1}}
\]

\[
= 30 \sin^2 5x \cos 5x - \frac{x \csc \sqrt{x^2 + 1} \cot \sqrt{x^2 + 1}}{\sqrt{x^2 + 1}}.
\]

**Example 4.5.16** Find $f'(x)$ if $f(x) = \sin^3 (x^2 \tan x)$.

This is first a power and chain rule problem, since $f(x) = \left[ \sin (x^2 \tan x) \right]^3$. After a second, trigonometric chain rule we will have a product rule. It is not necessary to notice all this structure from the beginning; it all becomes apparent as we dissect the function, applying the appropriate differentiation rules as we go:

\[
f'(x) = 3 \left[ \sin (x^2 \tan x) \right]^2 \cdot \frac{d}{dx} \left[ \sin (x^2 \tan x) \right]
\]

\[
= 3 \sin^2 (x^2 \tan x) \cdot \cos(x^2 \tan x) \cdot \frac{d}{dx} (x^2 \tan x)
\]

\[
= 3 \sin^2 (x^2 \tan x) \cdot \cos(x^2 \tan x) \cdot \left( x^2 \frac{d}{dx} \tan x + \tan x \cdot \frac{d}{dx} x^2 \right)
\]

\[
= 3 \sin^2 (x^2 \tan x) \cos (x^2 \tan x) \cdot (x^2 \sec^2 x + \tan x \cdot 2x) \quad \text{(Now Rearrange)}
\]

\[
= 3 (x^2 \sec^2 x + 2x \tan x) \sin^2 (x^2 \tan x) \cos (x^2 \tan x).
\]

In some cases it is best to simplify algebraically before applying differentiation rules. For an obvious illustration of this, consider the following.
Example 4.5.17 Compute \( \frac{d}{dx}(\cos x \sec x) \).

**Solution:** We have two methods that come to mind immediately:

\( \frac{d}{dx}(\cos x \sec x) = \cos x \cdot \frac{d}{dx}(\sec x) + \sec x \cdot \frac{d}{dx}(\cos x) = \cos x \sec x \tan x + \sec x(-\sin x) \)

\( = 1 \cdot \tan x - \tan x = 0. \)

\( \frac{d}{dx}(\cos x \sec x) \bigg|_{x=1} = 0. \)

As mentioned before, there is sometimes more than one “natural” method of computing a derivative. The rules are consistent, and if correctly applied different methods will yield the same result. Of course in the above example, the second method is much faster. The next example is perhaps not so obvious.

Example 4.5.18 Here we compute \( \frac{d}{dz} \left( \frac{1 + \frac{z+1}{z-1}}{1 - \frac{z+1}{z-1}} \right) \). A method of brute force would be to perform the quotient rule and continue from there:

\[ \left(1 - \frac{z+1}{z-1}\right) \frac{d}{dz} \left(1 + \frac{z+1}{z-1}\right) - \left(1 + \frac{z+1}{z-1}\right) \frac{d}{dz} \left(1 - \frac{z+1}{z-1}\right) \]

\[ \left(1 - \frac{z+1}{z-1}\right)^2, \]

from which we need to compute two more quotient rules. However, if we instead simplify the function from the beginning, our work is greatly simplified too:

\[ \frac{d}{dz} \left( \frac{1 + \frac{z+1}{z-1}}{1 - \frac{z+1}{z-1}} \right) = \frac{d}{dz} \left( \frac{z-1 + (z+1)}{z-1 - (z+1)} \right) = \frac{d}{dz} \left( \frac{2z}{-2} \right) = \frac{d}{dz}(-z) = -1. \]

In applications especially, we are often led to complicated expressions for functions, for which we then need to compute the derivatives. It is always better to look out for such cases in which algebraic simplification from the beginning will simplify our calculus tasks, as well as give us a look at a simpler form of the original function. As complicated as the above function first appeared, it was simply the function \(-z\) (where both original and simplified are both defined, which for this problem means \(z \neq 1\)). Note that even if we computed this derivative the longer way, we might not recognize our answer to be simply \(-1\).
For 1–8, compute the derivatives two different ways:

(a) by using the rule called upon by the way the function is originally written; and

(b) by first simplifying the function, and then computing the derivative of the simplified function in the obvious way (using already established derivative formulas).

(c) Show that the answers are the same.

For instance, in 1, first use the product rule, and then compute $\frac{d}{dx} \left[ x^2 \cdot x^2 \right]$ with the power rule, and finally show that the answers are the same (e.g., both $4x^3$ for that case).

1. $\frac{d}{dx} \left[ x^2 \cdot x^2 \right]$
2. $\frac{d}{dx} \left[ x^9 \right] \over x^3$
3. $\frac{d}{dx} \left[ \cos x \sec x \right]$
4. $\frac{d}{dx} \left[ \cos x \tan x \right]$
5. $\frac{d}{dx} \left[ \frac{1}{\cos x} \right]$
6. $\frac{d}{dx} \left[ \tan x \cot x \right]$
7. $\frac{d}{dx} \left[ \sin x \sec x \right]$
8. $\frac{d}{dx} \left[ \frac{1}{\sin x} \right]$

For 9–24, compute the given derivatives.

9. $\frac{d}{dx} \left[ \tan^2 x \right]$
10. $\frac{d}{dx} \left[ \frac{x^2 + 1}{\sin x - 1} \right]$
11. $\frac{d}{dx} \left[ \sqrt{1 - \csc 2x} \right]$
12. $\frac{d}{dx} \left[ x \sin x \cos x \right]$
13. $\frac{d}{dx} \left[ \frac{x^3 - 7x + 5}{x^2 - 3} \right]$
14. $\frac{d}{dx} \left[ \frac{1 + \frac{1}{x+1}}{1 - \frac{1}{x+1}} \right]$
15. $\frac{d}{dx} \left[ \sin^2 x \cos^3 x \right]$
16. $\frac{d}{dx} \left[ x^2 \cos x^2 \right]$
17. $\frac{d}{dx} \left[ \frac{x^4 + 3}{x - 1} \right]$
18. $\frac{d}{dx} \left[ \sec^4 (x^3 + 2x) \right]$
19. $\frac{d}{dx} \left[ \frac{\sin x + 3}{\cos x + 2} \right]$
20. $\frac{d}{dx} \left[ \sec 3x \cot 5x \right]$
21. $\frac{d}{dx} \left[ \frac{x}{\sqrt{x^2 + 1}} \right]$
22. $\frac{d}{dx} \left[ \frac{(x + 5)^3}{(x - 4)^6} \right]$. Factor the numerator in your answer to simplify.
23. $\frac{d}{dx} \left[ 3 \cot^2 9x - \sqrt{\cos 6x + 1} \right]$
24. $\frac{d}{dx} \left[ \tan(x + \tan(x + \tan x)) \right]$
25. Show that $\frac{d}{dx}[f(x)g(x)h(x)] = f(x)g(x)h'(x) + f(x)g'(x)h(x) + f'(x)g(x)h(x)$. What do you think will be the derivative of $f(x)g(x)h(x)i(x)$? (See notes on the intuition of the product rule, immediately following its introduction as Theorem 4.5.1 at the start of this section.)
26. A simple, Earth-bound pendulum of length $L$ in meters has a period of oscillation of $T$ seconds. That is, every $T$ (seconds) the pendulum completes one cycle. The period varies with the length by the equation $T = 2\pi \sqrt{L/g}$, where $g = 9.8\text{m/sec}^2$. Find the rate of change of $T$ with respect to $L$ when $T = 3\text{sec}$.

27. If the cost to manufacture $x$ items is $C(x)$, then we define the average cost per item to be

$$\overline{C}(x) = \frac{C(x)}{x}, \quad (4.52)$$

and from this define the marginal average cost by $\overline{C}'(x)$ Find the average cost per item function, and the marginal average cost function, if

$$C(x) = \frac{x^2 + 3x}{x + 4} + 100.$$ 

How do you interpret marginal average cost?

28. The intensity $I$ of a 100 watt (100W) light bulb is given by

$$I(x) = \frac{7.92W}{x^2},$$

where $x$ is the distance in meters from the bulb. (The W in the equation above can be suppressed in actual computations.)

(a) What are the units of $I$?

(b) Find $\frac{dI}{dx}$. What are its units?

(c) At what distance (accurate to the nearest cm=0.01m) will the rate of change of intensity be equal to $-3\text{W/m}$?

29. A variable resistance $R$ is given in a circuit and the voltage is found by the formula

$$V = \frac{5R + 10}{R + 3}.$$ 

Find the instantaneous rate of change of $V$ with respect to $R$ when $R = 6\Omega$. ($V$ will be measured in volts.)

30. A double convex converging lens will focus an object $p$ distance in front of the lens to an image a distance $q$ from the lens on the opposite side of the lens. The lens has a focal length of $f$. (Here all distances are in cm.) A well-known formula in optics gives

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}.$$ 

Find an equation which represents the rate of change of $q$ with respect to $p$, assuming $f$ is constant. (Hint: Solve the given equation for $q$ and then differentiate—that is take its derivative—while assuming that $f$ is constant. From this you can determine $dq/dp$.)
4.6 Chain Rule II: Implicit Differentiation

In this section we will apply the chain rule in a setting of so-called implicit functions, which are more general than the explicit functions we have dealt with so far. In later sections we will use the chain rule in still a third context, related rates.

In this section we first consider a brief summary and re-examination of all differentiation rules, including the chain rule. We then perform some simple differentiation problems which are of the type typically encountered in a first study of implicit functions. We then proceed to the topic of the chain rule in still a third context, related rates.

4.6.1 Review of Differentiation Rules

At this point we have many differentiation rules. If we look at all the rules so far, they can be summed up as follows. First we had the very general rules, regardless of $f(x)$ and $g(x)$, and fixed constants $C \in \mathbb{R}$ as long as the expressions on the right existed:

\[
\begin{align*}
\frac{d}{dx} [Cf(x)] &= C \cdot \frac{df(x)}{dx} \\
\frac{d}{dx} [f(x) + g(x)] &= \frac{df(x)}{dx} + \frac{dg(x)}{dx} \\
\frac{d}{dx} [f(x) \cdot g(x)] &= f(x) \cdot \frac{dg(x)}{dx} + g(x) \cdot \frac{df(x)}{dx} \\
\frac{d}{dx} [f(x)] &= \frac{df(x)}{dx} \\
\frac{d}{dx} [g(x)] &= \frac{dg(x)}{dx} \\
\frac{d}{dx} \left[\frac{f(x)}{g(x)}\right] &= \frac{g(x) \cdot \frac{df(x)}{dx} - f(x) \cdot \frac{dg(x)}{dx}}{[g(x)]^2} \\
\frac{d}{dx} \left[Cf(x) + g(x)\right] &= C \cdot \frac{df(x)}{dx} + \frac{dg(x)}{dx} \\
\frac{d}{dx} \left[f(x) \cdot g(x)\right] &= f(x) \cdot \frac{dg(x)}{dx} + g(x) \cdot \frac{df(x)}{dx} \\
\frac{d}{dx} \left[\frac{f(x)}{g(x)}\right] &= \frac{g(x) \cdot \frac{df(x)}{dx} - f(x) \cdot \frac{dg(x)}{dx}}{[g(x)]^2}
\end{align*}
\]

Then we had rules for specific functions, and their chain rule versions:

\[
\begin{align*}
\frac{d}{dx} [x^n] &= n \cdot x^{n-1} \\
\frac{d}{dx} \sin x &= \cos x \\
\frac{d}{dx} \cos x &= -\sin x \\
\frac{d}{dx} \tan x &= \sec^2 x \\
\frac{d}{dx} \cot x &= -\csc^2 x \\
\frac{d}{dx} \sec x &= \sec x \tan x \\
\frac{d}{dx} \csc x &= -\csc x \cot x \\
\frac{d}{dx} \sqrt{x} &= \frac{1}{2\sqrt{x}} \\
\frac{d}{dx} \left[\frac{1}{x}\right] &= -\frac{1}{x^2}
\end{align*}
\]

(Note that the bottom four rules were just special cases of the power rules.)

While we will not venture here to reinvent the discussion of the previous section, we will at least notice that when the variable of differentiation matches the input variable of the function,
we can use the simple derivative rules for that particular function, and when they do not, we multiply by the derivative of that variable with respect to the variable of differentiation:

\[
\frac{df(x)}{dx} = f'(x),
\]

\[
\frac{df(u)}{dx} = f'(u) \cdot \frac{du}{dx},
\]

(4.53)

(4.54)

For one example of (4.54), justified with Leibniz notation, we had \( f(x) = \sin x \) and \( u = u(x) \) yielding the following:

\[
\frac{d}{dx} \sin u = \frac{d}{dx} \sin u \cdot \frac{du}{dx} = \cos u \cdot \frac{du}{dx}.
\]

The decomposition in the middle step allowed us to use the sine derivative formula because the variable in the sine—\( u \)—matched the differential operator \( \frac{du}{dx} \) that appeared in the first term of the Leibniz-style decomposition. Writing that step will become more burdensome if the input of sine is more complicated, so we will skip that step in most future computations.

For a few slightly more complicated examples, consider the following, noting that the previous rules still apply if we assume that \( y \) is a function of \( x \) (else Footnote 50, page 383 applies):

**Example 4.6.1** Consider the following three derivative computations.

- \( \frac{d}{dx} (y^2) = 2y \cdot \frac{dy}{dx} \).
- \( \frac{d}{dx} \sqrt{y} = \frac{1}{2\sqrt{y}} \cdot \frac{dy}{dx} \).
- \( \frac{d}{dx} (y \sin y) = y \cdot \frac{d}{dx} (\sin y) + \sin y \cdot \frac{d}{dx} (y) = y \cdot \cos y \cdot \frac{dy}{dx} + \sin y \cdot \frac{dy}{dx} \).

Notice that the last example required first the product rule, and so the first step was to write—exactly—the statement of the product rule, where the product is \( y \cdot \sin y \). This is ultimately a product of two functions of \( x \), since we are assuming \( y \) is a function of \( x \), and thus \( y \sin y \) is also ultimately a function of \( x \). As with earlier derivatives, it is crucial to write out precisely what the general rules dictate for the given function. This becomes even more critical in the following.

**Example 4.6.2** Consider the following derivative computations:

- \( \frac{d}{dx} (xy^2) = x \cdot \frac{dy^2}{dx} + y^2 \cdot \frac{dx}{dx} = x \cdot 2y \cdot \frac{dy}{dx} + y^2 (1) = 2xy \cdot \frac{dy}{dx} + y^2, \)

- \( \frac{d}{dx} (x^2 + y^2)^2 = 2(x^2 + y^2) \cdot \frac{d}{dx} (x^2 + y^2) = 2 \left( x^2 + y^2 \right) \left( 2x + 2y \cdot \frac{dy}{dx} \right) = 4 \left( x^2 + y^2 \right) \left( x + y \cdot \frac{dy}{dx} \right) \).

Notice that we skipped an explicit writing of the “sum rule” step in the second computation above:

\[
\frac{d}{dx} (x^2 + y^2) = \frac{d}{dx} (x^2) + \frac{d}{dx} (y^2) = 2x + 2y \cdot \frac{dy}{dx}.
\]

The first term, \( \frac{d}{dx} (x^2) \) was a simple power rule because the variables matched.\(^{51}\) For the second term they do not match, so we need the chain rule to compute \( \frac{d}{dx} (y^2) = 2y \cdot \frac{dy}{dx} \).\(^{52}\) Notice also

\(^{51}\)We can write \( \frac{d}{dx} (x^2) = 2x \cdot \frac{dx}{dx} \) but then, of course, “inner function’s” derivative is just \( \frac{dx}{dx} = 1 \).

\(^{52}\)One usually does not continue to write, for example,

\[
\frac{d}{dx} \left( y^2 \right) = \frac{d}{dy} \left( y^2 \right) \cdot \frac{dy}{dx} = 2y \cdot \frac{dy}{dx},
\]
4.6. CHAIN RULE II: IMPLICIT DIFFERENTIATION

that in the first computation we used the fact that \( \frac{dx}{dx} = 1 \), which is reasonable on its face, but technically describes the fact that the function \( f(x) = x \) is a line with slope 1, i.e., \( f'(x) = 1 \), regardless of \( x \).\(^{53}\)

**Example 4.6.3** Consider the following derivative computation:

\[
\frac{d}{dx} \sin xy = \cos xy \cdot \frac{d(xy)}{dx} \\
= \cos xy \cdot \left( x \frac{dy}{dx} + y \frac{dx}{dx} \right) \\
= \cos xy \cdot \left( x \frac{dy}{dx} + y \right) \\
= x \cos xy \frac{dy}{dx} + y \cos xy.
\]

The above example used the chain rule first, and then the product rule “called” by the chain rule. Note that the first step could have been written

\[
\frac{d}{dx} \sin xy = \frac{d}{dx} \sin xy \cdot \frac{d(xy)}{dx}.
\]

Of course it is important to note that, in the Leibniz notation, \( d(xy) \) is taken to be one, encapsulated quantity. It is not a multiplication of three quantities, so for instance \( \frac{d(xy)}{dx} \neq y \).

Indeed, \( \frac{d(xy)}{dx} = x \cdot \frac{dy}{dx} + y \cdot \frac{dx}{dx} = x \cdot \frac{dy}{dx} + y \) by the product rule. However the chain rule does allow for the apparent division and multiplication by \( d(xy) \) in the derivative notation in (4.55) above.

Two more points should be emphasized regarding the product rule computation above, which we repeat here:

\[
\frac{d(xy)}{dx} = x \frac{dy}{dx} + y \frac{dx}{dx}.
\]

First, note that (4.56) is exactly what the product rule says for this product \( xy \). Thus it should be established true from the explicit result of the product rule. Next, when computing \( \frac{d}{dx}(y) \), we get exactly what we would from the chain rule, or from the power rule. Thinking of \( y \) as a function “raised to the power 1,” we might write:

\[
\frac{d}{dx}(y) = \frac{d}{dy}(y) \cdot \frac{dy}{dx} = 1 \cdot \frac{dy}{dx}.
\]

Thus once again the derivative rules are self-consistent, as is Leibniz notation, properly interpreted.

4.6.2 Implicit Functions and Their Derivatives

In this subsection we discuss the context and concepts of implicit differentiation, following one example through the discussion as an illustration. In subsequent subsections we will discuss more numerous examples.

\(^{53}\)Of course \( \frac{dx}{dx} = 1 \) also follows from the power rule, loosely interpreted: \( \frac{d}{dx}(x^1) = 1x^0 \), which seems to give 1, though technically \( x^0 = 1 \) only for \( x > 0 \) (for reasons we will see later in the text). Still, in this case the “formal answer” \( x^0 = 1 \), that is, 1, is in fact the correct answer for \( \frac{d(x^1)}{dx} \).
CHAPTER 4. THE DERIVATIVE

Figure 4.20: The graph of \( x^2 + y^2 = 25 \) gives \( y \) as a function of \( x \) locally except at \((-5, 0)\) and \((5, 0)\). Some possible open rectangles centered at various \((x_0, y_0)\) on the graph are drawn, within each of which (separately) \( y \) is a function of \( x \).

In previous sections we were interested in functions \( y = f(x) \), i.e., where \( y \) was given explicitly as a function of \( x \). However, there are relationships and graphs of interest where \( y \) is related to \( x \) implicitly, for instance, by an equation which contains both \( y \) and \( x \) as variables, except that \( y \) is not solved for as an explicit function of \( x \). Indeed, the graph of the equation (i.e., the graph of all \((x, y)\) which satisfy the equation) might not represent a function at all. However, it is quite possible for \( y \) to be a function of \( x \) locally near a point \((x_0, y_0)\) on the graph, and to speak of “slope” there, found through so-called implicit differentiation with the equation describing the curve.

**Definition 4.6.1** We will say \( y \) is locally (or implicitly) a function of \( x \) near \((x_0, y_0)\) if and only if there exist \( \delta, \varepsilon > 0 \) and an open rectangle

\[
\{ (x, y) \mid (|x - x_0| < \delta) \land (|y - y_0| < \varepsilon) \}
\]

such that, within that open rectangle, the graph represents \( y \) as a function of \( x \).

A simple example is a circle such as \( x^2 + y^2 = 25 \). Except at \((5, 0)\) and \((-5, 0)\), we can find a small enough, open rectangle around each point \((x_0, y_0)\) on the graph so that \( y \) is a function of \( x \) inside that rectangle. Note that it is not necessary that there exists a vertical line touching the graph inside the rectangle, but if one does touch the graph inside the rectangle it can only do so at one point.

Now suppose we would like to find the slope of the graph of \( x^2 + y^2 = 25 \), without solving for \( y \) first.\footnote{For this one we can solve for \( y \), almost: \( y = \pm \sqrt{25 - x^2} \), with the “plus” case being the upper semicircle and the “minus” case being the lower semicircle. For many curves we cannot isolate \( y \), or cannot do so easily. Even when we can, it is sometimes easier to use the equation as it stands than in its “solved” form.} If we are at a point \((x_0, y_0)\) satisfying the equation \( x^2 + y^2 = 25 \), and for which we can (in principle) find an open rectangle centered at \((x_0, y_0)\) in which \( y = y(x) \), i.e., in which \( y \) is a function of \( x \), then inside such a rectangle, we have that \( y^2 = (y(x))^2 \) is a (composite) function of \( x \), and \( x^2 + y^2 \) is therefore also a function of \( x \). Furthermore, 25 is a (rather trivial)
function of $x$ and, inside the rectangles, $x^2 + y^2$ and 25 are in fact the same function of $x$:

\[
\begin{array}{ll}
\text{function of } x & 25 \\
\text{in each rectangle} & \text{same function of } x
\end{array}
\]

Because these are the same functions of $x$ in an open rectangle, they have the same derivatives with respect to $x$ there. Thus we can state

\[
\frac{d}{dx} (x^2 + y^2) = \frac{d}{dx} (25) .
\]  (4.58)

Now suppose that we restrict our analysis to just such a rectangle in which $y$ is a function of $x$. In the above we are taking the derivative with respect to $x$. The first term is a simple power rule, but the second is a chain rule version of the power rule since the variables do not match, while the other side of the equation is a constant. Thus, (4.58) becomes

\[2x + 2y \frac{dy}{dx} = 0.\]

The process above is called implicit differentiation, since we are differentiating both sides of an equation which represents a curve in the $xy$-plane, on which $y$ is locally a function of $x$, given “implicitly” (not explicitly) by the equation. The differentiation process is simply a manifestation of the chain rule, though the name “implicit differentiation” does give the process context.\(^{55}\) It is $\frac{dy}{dx}$ that we want, i.e., the slope on the original curve. Fortunately we can now solve for it:

\[2x + 2y \frac{dy}{dx} = 0 \iff 2y \frac{dy}{dx} = -2x \iff \frac{dy}{dx} = \frac{-2x}{2y} = -\frac{x}{y}.
\]

Some more notation is now needed for when we wish to evaluate this at particular points on the curve. The value of $\frac{dy}{dx}$ at such $(x_0, y_0)$ on the curve is written

\[\left. \frac{dy}{dx} \right|_{(x_0, y_0)}.
\]

Out loud, this is said, “$\frac{dy}{dx}$ evaluated at the point $(x_0, y_0)$.” Looking back at the circle, we can

\(^{55}\)The term implicit differentiation unfortunately is sometimes used to refer to any computation of a derivative of an expression with respect to one variable, where the expression contains other variables. For instance sometimes the term is used when the variable of differentiation is “hidden” in the original equation, as is $t$ in $PV = C \implies \frac{d}{dt}(PV) = \frac{d}{dt}(C) \implies P \frac{dV}{dt} + V \frac{dP}{dt} = 0$ (assuming $C$ is constant). In all cases it is implied that the other variables are functions of the variable of differentiation (or the differential operator with respect to that variable cannot be meaningfully applied), though technically “implicit differentiation” should only refer to differentiation on curves on which one variable is “implicitly” (i.e., not explicitly) given as a function (at least “locally”) of the variable of differentiation.

In fact in all of these cases, we are simply applying the chain rule in different contexts.
CHAPTER 4. THE DERIVATIVE

find the slope at any \((x_0, y_0)\) which is on the graph:

\[
\left. \frac{dy}{dx} \right|_{(3,4)} = \frac{-x}{y} \bigg|_{(3,4)} = \frac{-3}{4},
\]

\[
\left. \frac{dy}{dx} \right|_{(0,5)} = \frac{-x}{y} \bigg|_{(0,5)} = \frac{-0}{5} = 0,
\]

\[
\left. \frac{dy}{dx} \right|_{(-4,3)} = \frac{-x}{y} \bigg|_{(-4,3)} = \frac{-(-4)}{3} = \frac{4}{3}, \quad \text{etc.}
\]

A quick check of the graph in Figure 4.20, page 386 shows that these slope calculations ring true. In fact, it is interesting to notice what happens if we attempt to evaluate \(\frac{dy}{dx}\) at \((5,0)\). If we naively insert \((5,0)\) into \(\frac{dy}{dx} = \frac{-x}{y}\), we would get \(-\frac{5}{0}\), which is undefined. Thus the derivative there does not exist, which is born out by the graph in the sense that there is no defined slope there. (Note that in the Euclidean geometry sense the tangent to the circle is a vertical line there.) Similarly for \((-5,0)\).

We might also be interested in equations of tangent lines to this graph. For instance, at \((x, y) = (-4,3)\), we have the slope

\[
\left. \frac{dy}{dx} \right|_{(-4,3)} = \frac{-x}{y} \bigg|_{(-4,3)} = \frac{4}{3},
\]

and so with the point \((-4,3)\) we have the line shown at right, with equation

\[
y = 3 + \frac{4}{3}(x + 4).
\]

Finding the slope at points on the circle above does not require the implicit differentiation technique, that is finding \(dy/dx\) from an implicit curve where \(y\) is not an explicit function of \(x\). This is not an absolute requirement here because we can write \(y\) as an explicit function of \(x\) and compute the derivative, near any \((x_0, y_0)\) on the curve except at \((\pm 5,0)\). Indeed, on the upper semi-circle we have \(y = \sqrt{25-x^2}\), and so

\[
\frac{dy}{dx} = \frac{d}{dx} \sqrt{25-x^2} = \frac{1}{2\sqrt{25-x^2}} \frac{d}{dx} (25-x^2) = \frac{-2x}{2\sqrt{25-x^2}} = \frac{-x}{\sqrt{25-x^2}}.
\]

Notice that (on the upper semicircle) this is the same as the derivative acquired through implicit differentiation:

\[
\frac{dy}{dx} = \frac{-x}{\sqrt{25-x^2}} = \frac{-x}{y}.
\]

A similar analysis shows that on the lower semicircle we also get two equivalent forms of the derivative:

\[
\frac{dy}{dx} = \frac{d}{dx} \left[ -\sqrt{25-x^2} \right] = \cdots = \frac{x}{\sqrt{25-x^2}} = \frac{x}{-y} = \frac{-x}{y}.
\]
However it is arguably easier to compute the derivative with the implicit form of the curve as we did originally.

Next we consider other examples where it is not clearly possible to solve for \( y \) as an explicit function of \( x \).

**Example 4.6.4** Consider the curve \((x^2 + y^2)^{3/2} = 2xy\), which is graphed in Figure 4.21, page 390. (Much later in the text we will see how this graph came about.) First we find \( \frac{dy}{dx} \) by applying the differential operator \( \frac{d}{dx} \) to both sides:

\[
(x^2 + y^2)^{3/2} = 2xy
\]

\[
\Rightarrow
\frac{d}{dx} \left[ (x^2 + y^2)^{3/2} \right] = \frac{d}{dx} [2xy]
\]

\[
\Rightarrow \frac{3}{2} (x^2 + y^2)^{1/2} \cdot \frac{d}{dx} (x^2 + y^2) = 2 \left( \frac{dy}{dx} + \frac{dy}{dx} \right)
\]

\[
\Rightarrow \frac{3}{2} (x^2 + y^2)^{1/2} \cdot \left( 2x + 2y \frac{dy}{dx} \right) = 2 \left( \frac{dy}{dx} + y \cdot 1 \right)
\]

\[
\Rightarrow \frac{3}{2} (x^2 + y^2)^{1/2} (2x) + \frac{3}{2} (x^2 + y^2)^{1/2} \left( 2y \frac{dy}{dx} \right) = 2x \frac{dy}{dx} + 2y
\]

\[
\Rightarrow 3x \sqrt{x^2 + y^2} + 3y \sqrt{x^2 + y^2} \frac{dy}{dx} = 2x \frac{dy}{dx} + 2y
\]

\[
\Rightarrow 3y \sqrt{x^2 + y^2} \cdot \frac{dy}{dx} - 2x \frac{dy}{dx} = 2y - 3x \sqrt{x^2 + y^2}
\]

\[
\Rightarrow \left[ 3y \sqrt{x^2 + y^2} - 2x \right] \frac{dy}{dx} = 2y - 3x \sqrt{x^2 + y^2}
\]

\[
\Rightarrow \frac{dy}{dx} = \frac{2y - 3x \sqrt{x^2 + y^2}}{3y \sqrt{x^2 + y^2} - 2x}.
\]

Now we will compute the slope for two points on the curve, \((\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})\) and \((\frac{3}{4}, \frac{\sqrt{3}}{4})\), which are labeled in the aforementioned figure. (The reader should verify that these points are actually on the original curve.)

\[
\begin{align*}
\text{at } & \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \quad \frac{dy}{dx} = \frac{2y - 3x \sqrt{x^2 + y^2}}{3y \sqrt{x^2 + y^2} - 2x} \\
& = \frac{\sqrt{2} - \frac{3}{\sqrt{2}} \cdot 1}{\frac{3}{\sqrt{2}} - \sqrt{2}} = \frac{\frac{3\sqrt{2} - \sqrt{2}}{\sqrt{2}}}{\frac{3}{\sqrt{2}} - \sqrt{2}} = 1.
\end{align*}
\]

This should seem reasonable given the position of this point on the graph. For the computation of \( \frac{dy}{dx} \) at \((\frac{3}{4}, \frac{\sqrt{3}}{4})\) we will be more brief.

\[
\begin{align*}
\text{at } & \left( \frac{3}{4}, \frac{\sqrt{3}}{4} \right) \quad \frac{dy}{dx} = \frac{2 \cdot \frac{\sqrt{3}}{4} - 3 \cdot \frac{3}{4} \sqrt{\frac{12}{16}}}{3 \cdot \frac{\sqrt{3}}{4} \sqrt{\frac{12}{16}} - 2 \cdot \frac{3}{4}} \\
& = \frac{\sqrt{3} - \frac{9}{4} \cdot \sqrt{3}}{\frac{3}{4} \cdot \sqrt{\frac{12}{16}} - \frac{3}{2}} = \frac{\sqrt{3} - \frac{9}{4} \cdot \sqrt{3}}{\frac{3}{4} \cdot \sqrt{\frac{12}{16}} - \frac{3}{2}} = 8 = \frac{4\sqrt{3} - 9\sqrt{3}}{9 - 12} \\
& = \frac{-5\sqrt{3}}{-3} = \frac{5\sqrt{3}}{3} \approx 2.88675135.
\end{align*}
\]

This should also seem reasonable from the graph.
Figure 4.21: Graph of \((x^2 + y^2)^{3/2} = 2xy\). The slope \(dy/dx\) at each point \((x, y)\) on the curve is calculated in Example 4.6.4, page 390. The points \((x, y) = (\sqrt{2}, \sqrt{2})\) and \((x, y) = (4, \sqrt{7})\) are plotted as well, for which slopes were computed in that example.

The argument in the proof of the power rule for rational powers of \(x\), Theorem 4.4.2 on page 359, used this implicit differentiation technique. As in that proof, we can compute \(dy/dx\) without worrying about actually finding the open rectangles which give \(y\) locally as a function of \(x\). The rectangles are important in justifying the technique, but if something goes wrong, it will usually show up in the final form of the computed derivative.

**Example 4.6.5** Find \(dy/dx\) for the graph \(5x + x^2 + y^2 + xy = \tan y\). This one will require two chain rule calls, for the \(y^2\) and \(\tan y\) terms, and a product rule for the \(xy\) term. As before, if we are careful in writing out the product rule we are less likely to have errors.

\[
\frac{d}{dx} \left[5x + x^2 + y^2 + xy\right] = \frac{d}{dx} \tan y
\]

\[
\Rightarrow 5 + 2x + 2y \cdot \frac{dy}{dx} + \left[x \frac{dy}{dx} + y \frac{dx}{dx}\right] = \sec^2 y \cdot \frac{dy}{dx}
\]

\[
\Rightarrow 5 + 2x + 2y \frac{dy}{dx} + x \frac{dy}{dx} + y = \sec^2 y \frac{dy}{dx}
\]

\[
\Rightarrow 5 + 2x + y = \sec^2 y \frac{dy}{dx} - 2y \frac{dy}{dx} - x \frac{dy}{dx}
\]

\[
\Rightarrow 5 + 2x + y = \left(\sec^2 y - 2y - x\right) \frac{dy}{dx}
\]

\[
\Rightarrow \frac{5 + 2x + y}{\sec^2 y - 2y - x} = \frac{dy}{dx}
\]

To find the tangent line through \((0, 0)\), which is on the graph, we then compute

\[
\left.\frac{dy}{dx}\right|_{(0,0)} = \left.\frac{5 + 2x + y}{\sec^2 y - 2y - x}\right|_{(0,0)} = \frac{5 + 0 + 0}{1 - 0 - 0} = 5,
\]

and so the tangent line through \((0, 0)\) has equation \(y = 5x\).
Though we do not have the tools to prove it here, it is true that we will always be able to solve for $\frac{dy}{dx}$ in these problems. The basic idea of a proof is that $\frac{dy}{dx}$ will always be a factor in the terms in which it appears, and will only appear to the first degree, so we are basically solving a *linear* equation in the “variable” $\frac{dy}{dx}$, albeit with nonconstant coefficients. So in fact it is no different fundamentally than solving $Ax + By + C = Dx + Ey + F$ for $y$ (or for $x$, for that matter). One moves all terms containing that variable to one side, the other terms to the other side, factors the variable from the side which then contains it, and divides by the other factor.

Of course our implicit differentiation technique begins with calculus steps and ends with algebra steps. The entire process can be summarized by five steps:

1. **Apply $\frac{d}{dx}$ to both sides of the equation describing the curve.**

2. **Complete all differentiation steps, flushing out all terms with a factor $\frac{dy}{dx}$.**

3. **Put all terms with $\frac{dy}{dx}$ factors on one side, other terms on the other side of the equation.**

4. **Factor the $\frac{dy}{dx}$ from the side which contains it, and finally,**

5. **Divide by the remaining factor, leaving $\frac{dy}{dx}$ on one side by itself.**

**Example 4.6.6** Consider the equation $x = \sin y$, pictured below and to the right.

The slope $\frac{dy}{dx}$ is then computed as follows:

\[
\begin{align*}
  x &= \sin y \\
  \implies \quad \frac{d}{dx}[x] &= \frac{d}{dx}[\sin y] \\
  \implies \quad 1 &= \cos y \frac{dy}{dx} \\
  \implies \quad \frac{1}{\cos y} &= \frac{dy}{dx}.
\end{align*}
\]

So for instance the slope at $\left(\frac{\pi}{2}, \frac{\pi}{2}\right)$ is given by $\frac{dy}{dx} = 1/\cos \frac{\pi}{6} = 1/\sqrt{3/2} = 2/\sqrt{3}$, and the equation of the tangent line there is

\[
y = \frac{\pi}{6} + \frac{2}{\sqrt{3}} \left( x - \frac{1}{2} \right).
\]

It is also worth noting what happens when $\cos y = 0$, $(y = \pm \frac{\pi}{2}, \pm \frac{3\pi}{2}, \pm \frac{5\pi}{2}, \ldots, x = \pm 1)$ and its implications for $\frac{dy}{dx}$, in light of the graph, given on the right.

The next example is quite long, but with moderate persistence is also quite do-able.
Example 4.6.7 Find $\frac{dy}{dx}$ on the graph of $y^3 \sec \sqrt{x^2 + y^2} = \cos 2x$.

\[
\frac{d}{dx} \left[ y^3 \sec \sqrt{x^2 + y^2} \right] = \frac{d}{dx} \left[ \cos 2x \right]
\]

\[\Rightarrow \quad y^3 \frac{d}{dx} \left[ \sec \sqrt{x^2 + y^2} \right] + \sec \sqrt{x^2 + y^2} \cdot \frac{dy}{dx} = -2 \cos 2x \cdot \frac{d}{dx} \left[ 2x \right]
\]

\[\Rightarrow \quad y^3 \sec \sqrt{x^2 + y^2} \tan \sqrt{x^2 + y^2} \cdot \frac{d}{dx} \left[ \sqrt{x^2 + y^2} \right]
\]

\[+ \sec \sqrt{x^2 + y^2} \cdot 3y^2 \cdot \frac{dy}{dx} = -2 \cos 2x \cdot 2
\]

\[\Rightarrow \quad y^3 \sec \sqrt{x^2 + y^2} \tan \sqrt{x^2 + y^2} \cdot \frac{1}{2\sqrt{x^2 + y^2}} \cdot \frac{d}{dx} \left[ x^2 + y^2 \right]
\]

\[+ 3y^2 \sec \sqrt{x^2 + y^2} \cdot \frac{dy}{dx} = -2 \cos 2x
\]

Next we factor the $\frac{dy}{dx}$:

\[\left( y^3 \sec \sqrt{x^2 + y^2} \tan \sqrt{x^2 + y^2} \cdot \frac{dy}{dx} + 3y^2 \sec \sqrt{x^2 + y^2} \cdot \frac{dy}{dx} \right) \frac{dy}{dx}
\]

\[= -2 \cos 2x - \frac{xy^3 \sec \sqrt{x^2 + y^2} \tan \sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2}}
\]

Next we factor the $\frac{dy}{dx}$:

\[\left( y^4 \sec \frac{x^2 + y^2}{\sqrt{x^2 + y^2}} \tan \frac{x^2 + y^2}{\sqrt{x^2 + y^2}} \cdot \frac{dy}{dx} + 3y^2 \sec \frac{x^2 + y^2}{\sqrt{x^2 + y^2}} \cdot \frac{dy}{dx} \right) \frac{dy}{dx}
\]

\[= -2 \cos 2x - \frac{xy^3 \sec \sqrt{x^2 + y^2} \tan \sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2}}
\]

Finally, we divide to solve for $\frac{dy}{dx}$:

\[\frac{dy}{dx} = \frac{-2 \cos 2x - \frac{xy^3 \sec \sqrt{x^2 + y^2} \tan \sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2}}}{y^4 \sec \frac{x^2 + y^2}{\sqrt{x^2 + y^2}} \tan \frac{x^2 + y^2}{\sqrt{x^2 + y^2}} + 3y^2 \sec \frac{x^2 + y^2}{\sqrt{x^2 + y^2}}}
\]
If we would like, we can multiply the numerator and denominator by $\sqrt{x^2 + y^2}$ to get:

$$\frac{dy}{dx} = \frac{-2\sqrt{x^2 + y^2}\sin 2x - xy^3 \sec \sqrt{x^2 + y^2} \tan \sqrt{x^2 + y^2}}{y^4 \sec \sqrt{x^2 + y^2} \tan \sqrt{x^2 + y^2} \tan \sqrt{x^2 + y^2} + 3y^2 \sqrt{x^2 + y^2} \sec \sqrt{x^2 + y^2}}.$$

Although this latest example may seem tedious, no particular step is conceptually difficult. Success in such a project is nearly as much dependent upon our bookkeeping skills as upon understanding of derivative rules.

The next example requires some interpretation of the derivative computation.

**Example 4.6.8** Show that there are no horizontal tangent lines for the curve $xy^5 + x^5 y = 2$.

**Solution:** As before, we apply $\frac{d}{dx}$ to both sides:

$$xy^5 + x^5 y = 2$$

$$\Rightarrow \quad \frac{d}{dx} [xy^5 + x^5 y] = \frac{d}{dx} (2)$$

$$\Rightarrow \quad x \frac{dy^5}{dx} + y^5 \frac{dx}{dx} + x^5 \frac{dy}{dx} + y \frac{dy^5}{dx} = 0$$

$$\Rightarrow \quad x \cdot 5 y^4 \frac{dy}{dx} + y^5 + x^5 \frac{dy}{dx} + y \cdot 5 x^4 = 0$$

$$\Rightarrow \quad \left[5xy^4 + x^5\right] \frac{dy}{dx} = -y^5 - 5x^4 y$$

$$\Rightarrow \quad \frac{dy}{dx} = -\frac{y^5 + 5x^4 y}{5xy^4 + x^5}$$

$$\Rightarrow \quad \frac{dy}{dx} = -\frac{y(y^4 + 5x^4)}{x(5y^4 + x^4)}.$$

To have a horizontal tangent, we would need the numerator above to be zero, i.e., $y(y^4 + 5x^4) = 0$. This would imply $y = 0$ or $y^4 + 5x^4 = 0$.

1. **Case $y = 0$:** This never happens on the curve $x^5 y + xy^5 = 2$, because it would imply $0 = 2$, a contradiction.

2. **Case $y^4 + 5x^4 = 0$:** This would require $x, y = 0$, since $(y^4 \geq 0) \cap (5x^4 \geq 0) \implies y^4 + 5x^4 \geq 0$, and we get equality only when both (nonnegative) terms in the sum are zero, thus requiring $x, y = 0$, which is impossible again due to the fact that we require $(x, y)$ to be on the curve $xy^5 + x^5 y = 2$.

The above example demonstrates the importance of applying the derived formula for $dy/dx$ to only points $(x, y)$ which are on the original curve, though the formula otherwise “makes sense” for nearly any pair $(x, y)$. The actual curve is the setting in which the formula for $dy/dx$ is meaningful. The curve is given in Figure 4.22, page 394.

**4.6.3 A Mistake to Avoid**

Before finishing this section a remark is in order. It is tempting to try to simplify an algebraic equation by taking derivatives of both sides, especially in the case of polynomial equations. However, in such cases we are looking for points where one side’s expression equals the other, which is very different from saying the two sides are the same functions of, say, $x$. Rephrased, their outputs coinciding at a particular $x$-value is indeed different from saying the functions on each side are the same functions (therefore having the same derivatives).
Figure 4.22: Graph of the curve $xy^5 + x^5y = 2$ from Example 4.6.8, page 393. The curve can have no horizontal tangents, though we can see that $\frac{dy}{dx} \to 0$ as $x \to \pm\infty$. (We also see $\frac{dy}{dx} \to \infty$ as $x \to 0$. We did not prove this fact, though we can see that $y$ must get large quickly as $x \to 0$ if the equation $xy^5 + x^5y = 2$ is to hold, since $x \to 0$ alone would shrink $xy^5$ and $x^5y$, and so $y$ must grow to compensate if the sum is to be fixed at 2.)

Figure 4.23: The graphs of $y = f(x)$ and $y = g(x)$ intersect at $x = \pm 2$, which is the solution to $f(x) = g(x)$. However, it is clear from the picture that, while $f(x) = g(x)$ at $x = \pm 2$, the derivatives (slopes) $f'(x)$ and $g'(x)$ at those two points are not the same. In fact, $f'(x) = g'(x)$ at $x = 0$ only. A look at the graphs above indicates that the slopes of $y = f(x)$ and $y = g(x)$ at $x = 0$ do appear to agree, but those at $x = 2$ and $x = -2$ do not. Thus $f(a) = g(a) \nRightarrow f'(a) = g'(a)$. 

\[ f(x) = x^2 - 2x + 1 \]
\[ g(x) = 5 - 2x \]
Example 4.6.9 For a very simple case, consider the equation
\[ x^2 - 2x + 1 = 5 - 2x. \] (4.59)
This succumbs easily to the earlier methods:
\[
\begin{align*}
  x^2 - 2x + 1 &= 5 - 2x \\
  \iff x^2 &= 4 \\
  \iff x &= \pm 2.
\end{align*}
\]
so the solution is simply \( x = \pm 2 \). Now suppose instead we tried to take derivatives of both sides of (4.59):
\[
\begin{align*}
  2x - 2 &= -2 \\
  \iff 2x &= 0 \\
  \iff x &= 0.
\end{align*}
\]
We see that we get the incorrect answer, if our goal was to solve the original equation \( x^2 - 2x + 1 = 5 - 2x \). If we let \( x = 2 \) this equation becomes \( 2^2 - 2 \cdot 2 + 1 = 5 - 2 \cdot 1 \), i.e., \( 1 = 3 \), which is clearly false.

Thus if we are solving \( f(x) = g(x) \), it does not follow that \( f'(x) = g'(x) \). It is true if they are the same functions, i.e., same heights everywhere (and we use this fact in our implicit differentiation process), then they have the same derivatives but when we solve algebraic equations we are only interested in those points where the graphs of the two (usually different) functions intersect. It is unlikely that they would share the same slopes there as well as the heights. Hence it is important to use algebraic arguments where appropriate, and calculus arguments where appropriate. See Figure 4.23, page 394.

4.6.4 An Application

Example 4.6.10 When two resistors with values \( R_1 \) and \( R_2 \) are connected in parallel, the resulting resistance \( R \) is given by the equation
\[
\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}. \] (4.60)
Suppose \( R_2 = R_1 + 4 \Omega \).

1. Find an equation for \( R \) in terms of \( r = R_1 \).
2. Show that \( r^2 = 2rR + 4R - 4r \).
3. Find \( dR/dr \) in terms of \( r \) using the equation you found in part 1 directly.
4. Find \( dR/dr \) in terms of \( r \) by differentiating the equation in part 2. Use part 1 to show this is the same as the answer in part 3.

Solution:

1. After making the substitution \( R_2 = R_1 + 4 \) we have
\[
\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_1 + 4}.
\]
Using \( r = R_1 \) this then becomes
\[
\frac{1}{R} = \frac{1}{r} + \frac{1}{r + 4}.
\]

Taking reciprocals of both sides gives us
\[
R = \frac{1}{\frac{1}{r} + \frac{1}{r + 4}} = \frac{r(r + 4)}{(r + 4) + r}
\]
\[
\Rightarrow R = \frac{r^2 + 4r}{2r + 4}.
\]

2. Multiplying the above equation by \((2r + 4)\) gives us
\[
2rR + 4R = r^2 + 4r
\]
\[
\Leftrightarrow \quad 2rR + 4R - 4r = r^2
\]
\[
\Leftrightarrow r^2 = 2rR + 4R - 4r, \; q.e.d.
\]

3. We find \( dR/dr \) using the quotient rule:
\[
\frac{dR}{dr} = \frac{(2r + 4)\frac{d}{dr}(r^2 + 4r) - (r^2 + 4r)\frac{d}{dr}(2r + 4)}{(2r + 4)^2}
\]
\[
= \frac{(2r + 4)(2r + 4) - (r^2 + 4r)(2)}{[2(r + 2)]^2}
\]
\[
= \frac{4r^2 + 16r + 16 - 2r^2 - 8r}{4(r + 2)^2}
\]
\[
= \frac{2r^2 + 8r + 16}{4(r + 2)^2}
\]
\[
= \frac{2(r^2 + 4r + 8)}{4(r + 2)^2}
\]
\[
\Rightarrow \frac{dR}{dr} = \frac{r^2 + 4r + 8}{2(r + 2)^2}.
\]

4. Using part 2, we apply \( \frac{d}{dr} \) to both sides:
\[
\frac{d}{dr} \left[ r^2 \right] = \frac{d}{dr} \left[ 2rR + 4R - 4r \right]
\]
\[
\Rightarrow \quad 2r = 2r \frac{d}{dr} (R) + R \frac{d}{dr} (2r) + 4 \frac{dR}{dr} - 4
\]
\[
\Rightarrow \quad 2r = 2r \frac{dR}{dr} + 2R + 4 \frac{dR}{dr} - 4
\]
\[
\Rightarrow \quad 2r + 4 - 2R = (2r + 4) \frac{dR}{dr}
\]
\[
\Rightarrow \quad \frac{dR}{dr} = \frac{2r + 4 - 2R}{2r + 4}.
\]

Now, using \( R = \frac{r^2 + 4r}{2r + 4} \) we get
\[
\frac{dR}{dr} = \frac{2r + 4 - 2 \cdot \frac{r^2 + 4r}{2r + 4}}{2r + 4} \cdot \frac{2r + 4}{2r + 4} = \frac{(2r + 4)^2 - 2(r^2 + 4r)}{(2r + 4)^2}.
\]
4.6. CHAIN RULE II: IMPLICIT DIFFERENTIATION

which was one of our intermediate results in the computations in part 3 (specifically, the second line there), and so continued simplification would give us the same as the final simplification in part 3 as well, q.e.d.

Exercises

For Exercises 1–10 compute \( \frac{dy}{dx} \) by implicit differentiation.

1. \( Ax + By = C \), where \( A, B \) and \( C \) are consants.
2. \( \sin x + \cos y = \frac{1}{2} \).
3. \( x^2 - y^2 = 9 \). Find also the tangent line at (5, -4).
4. \( x^2 + 3x + y^2 = 16 + 7y \).
5. \( x + \tan x = y^3 + \sec y \).
6. \( x^2 + 2xy + y^2 = \cos x \).
7. \( \sin x^2 = \cos y^3 \).
8. \( x \sin y + y \sin x = 2 \).
9. \( \tan \frac{x}{y} = 9 \).
10. \( \sin(xy) + y = \csc x - 11 \).

11. Consider the curve graphed below given by \( (x^2 + y^2)^3 = (x^2 - y^2)^2 \).
   (a) Find \( \frac{dy}{dx} \).
   (b) Find the equation of the tangent line at \( \left( \frac{1}{4}, \frac{\sqrt{3}}{4} \right) \).
   (c) What can you say about the slope \( \frac{dy}{dx} \) as \( y \to -1 \)?
   (d) What can you say about the slope \( \frac{dy}{dx} \) as \( x \to -1^- \) or \( x \to -1^+ \)?

12. Consider the curve \( xy = 1 \).
   (a) Use implicit differentiation to find \( \frac{dy}{dx} \).
   (b) Use \( y = \frac{1}{x} \) to find \( \frac{dy}{dx} \).
   (c) Show algebraically how these are the same answers.

13. Consider the algebraic equation \( x^2 = 9 \).
   (a) Define functions \( f(x) = x^2 \) and \( g(x) = 9 \) and graph them together on one grid.
   (b) What is its solution of the original equation, i.e., of \( f(x) = g(x) \)?
   (c) What is the graphical significance of the solution of \( f(x) = g(x) \)?
   (d) Now consider the equation \( f'(x) = g'(x) \) for these two functions \( f \) and \( g \). What is the solution of this new equation \( f'(x) = g'(x) \)?
   (e) Explain the significance of the solution of \( \frac{d}{dx} f(x) = \frac{d}{dx} g(x) \) for this particular example.
   (f) Explain why, if a particular \( x \) satisfies \( f(x) = g(x) \), we cannot expect that it also satisfies \( f'(x) = g'(x) \).
4.7 Arctrigonometric Functions and their Derivatives

In this and the next three sections, we will explore derivatives of the last of our standard classes of functions. Presently we will look at the arctrigonometric functions, while in the next three sections we will look respectively at exponential functions, logarithmic functions, and related techniques. While we will need to be mindful of the exact natures of all these functions, with their derivative formulas we will usually be able to simply apply the new formulas together with the previous rules, and in so doing nearly finish our study of computing derivatives. However, their algebraic properties and exact natures cannot be ignored. Indeed, they are often exploited to simplify our calculus computations.

The arctrigonometric functions are also called the \textit{inverse trigonometric functions}.

### 4.7.1 One-to-one Functions and Inverses, Briefly

Recall what it means for a function $y = f(x)$ to be invertible. It is also described as \textit{one-to-one}, meaning that for $f : S \to \mathbb{R}$, i.e., where $S$ is the domain of $f$, we have

$$\forall x_1, x_2 \in S \left[ (x_1 = x_2) \iff (f(x_1) = f(x_2)) \right]$$

(4.61)

Note that for $f$ to be a function we already have $\forall x_1, x_2 \in S \left[ (x_1 = x_2) \to (f(x_1) = f(x_2)) \right]$, i.e., if we know the input we (in principle) know the output.

For the function to be one-to-one means we also have $\leftarrow$, i.e., if we know the output we can deduce the input. When we have such an $f$, we call its \textit{inverse} $f^{-1}$, the function defined by the property

$$\forall x \in S \forall y \in f(S) \left[ f(x) = y \iff f^{-1}(y) = x \right].$$

(4.62)

Here the superscript “$-1$” is not an exponent (power), but an indication that we are looking at the “reverse” of the process that takes $x$ to $y$.\footnote{Anytime we have an “exponent” which is not $-1$, we assume it is a true exponent. However, the “exponent” $-1$, when immediately modifying a function given by name such as $f$, $g$, etc., or any of the named trigonometric functions, is reserved to denote instead (what is taken to be) the inverse function.}

Note that just as $f^{-1}$ reverses $f$, so does $f$ reverse $f^{-1}$, as shown in the diagrams below left:

Thus, if $f$ is one-to-one, we have

$$\forall x \in S \ [f^{-1}(f(x)) = x],$$

(4.63)

$$\forall y \in f(S) \ [f(f^{-1}(y)) = y].$$

(4.64)

In other words, the functions $f$ and $f^{-1}$ “undo” each other; for a one-to-one function $f$ with inverse $f^{-1}$, we have that their mappings are reversible:

$$x \overset{f}{\to} y \overset{f^{-1}}{\to} x,$$

$$y \overset{f^{-1}}{\to} x \overset{f}{\to} y,$$

which as mappings of sets would look like

$$S \overset{f}{\to} f(S) \overset{f^{-1}}{\to} S.$$
4.7. ARCTRIGONOMETRIC FUNCTIONS AND THEIR DERIVATIVES

The usual algebraic way to attempt to invert a function \( f \), if possible, is to solve the equation \( y = f(x) \) for \( x \), producing an equation of the form \( x = f^{-1}(y) \). Often in the process of attempting to calculate \( f^{-1}(y) \) in this way we will discover if it in fact exists, i.e., if \( f \) is one-to-one. There can be many technicalities, depending upon the original function, but also many computations are straightforward.

Note also that there is often confusion regarding the naming of input and output variables. If \( x \) is the input of \( f \) and \( y \) the output, then it is useful to let \( y \) be the input of \( f^{-1} \) and \( x \) the output, hence \( y = f(x) \iff x = f^{-1}(y) \) when \( f \) is one-to-one. However the inverse is a function in its own right so the usual convention is to discuss \( f^{-1}(x) \) just as we discuss \( f(x) \). See the following example.

Example 4.7.1 Consider the function \( f(x) = 6x - 9 \). Compute \( f^{-1}(x) \), and from that also compute \( f^{-1}(f(x)) \) and \( f(f^{-1}(x)) \).

Solution: The usual method is essentially to write \( y = f(x) \) and try to solve for \( x \). If it can be done uniquely, then our equation will be of the form \( x = f^{-1}(y) \).

\[
y = f(x) \iff y = 6x - 9 \iff \frac{y + 9}{6} = x.
\]

Since we want to use \( x \) for the input and \( y \) for the output, regardless of the function, here we would write

\[
f(x) = 6x - 9, \\
f^{-1}(x) = \frac{1}{6}(x + 9).
\]

Note that another way to demonstrate the functions’ actions is with empty parentheses:

\[
f(\ ) = 6(\ ) - 9, \\
f^{-1}(\ ) = \frac{1}{6}(\ + 9).
\]

Using \( x \) as the input variable again for both functions, and working “outside-in,” we have

\[
f^{-1}(f(x)) = \frac{1}{6}f(x) + 9 = \frac{1}{6}(6x - 9 + 9) = \frac{1}{6}(6x) = x, \\
f(f^{-1}(x)) = \left[\left(\frac{1}{6}x + 9\right) - 9\right] = 6\left(\frac{1}{6}(x + 9)\right) - 9 = (x + 9) - 9 = x.
\]

Note how applying \( f^{-1} \) indeed “undoes” the process of applying \( f \) to an element in the domain of \( f \), while applying \( f \) “undoes” the process of applying \( f^{-1} \) to an element of \( f(S) \). That is the essence of (4.63) and (4.64) from before.

4.7.2 The Arctrigonometric Functions and Their Derivatives

In Table 4.1, page 402 we summarize the most important conclusions of what a careful development of the arctrigonometric functions would yield, including their domains, their ranges, and their derivatives. We save the actual derivations for later subsections, since at this stage we are most interested in rules for computing derivatives involving these functions. However, the detailed derivations are ultimately quite important and should themselves be studied for a couple

\[57\] Alternatively, one writes \( x = f(y) \) and solves for \( y = f^{-1}(x) \).
of reasons. First, the technicalities involved there will reappear many times in later chapters. Second, the derivations are also quite interesting—and useful as exercises—because, as occurs in many applied problems, the techniques used form an interesting mix of algebra, trigonometry, and our calculus techniques, particularly implicit differentiation.\footnote{Indeed, the farther along one gets in mathematical or applied studies, the more one has to borrow from a growing diversity of subjects. It is very often the “technicalities”—themselves often first found in derivations, footnotes or otherwise parenthetically—which play key roles in solving interesting problems, in both pure and applied mathematics.}

However in this subsection we concentrate on their derivatives, though we keep an eye towards their actual definitions, including their domains and ranges. Armed with the derivative formulas, we will be able to greatly expand the class of functions we can differentiate.

Now the trigonometric functions repeat periodically, so there is generally no guarantee for instance that \( \sin x_1 = \sin x_2 \) implies \( x_1 = x_2 \) (not true!). The way we nonetheless endeavor to “invert” the trigonometric functions is to temporarily restrict their domains so that they are forced to be one-to-one. For instance, \( \sin x \) is one-to-one on \( x \in [-\pi/2, \pi/2] \), i.e.,

\[
\left( \forall x_1, x_2 \in \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right] \right) \left[ (\sin x_1 = \sin x_2) \iff (x_1 = x_2) \right].
\]

This is kind of domain restriction is explained in subsequent subsections, where for each of the six standard trigonometric functions we look for a subset of its domain on which (1) the function is one-to-one, and (2) the set of all outputs covers the whole range of the original function.\footnote{It is akin to the problem trying to invert the function \( f(x) = x^2 \), which is not one-to-one on all of \( \mathbb{R} = (-\infty, \infty) \), since for instance \( f(-5) = f(5) \), while \(-5 \neq 5 \). Since it is impossible to truly invert \( f(x) = x^2 \), instead we define the “principal square root” \( g(y) = \sqrt{y} \), which does give us an inverse to \( f(x) \) on the set \( x \geq 0 \), i.e., \( x \in [0, \infty) \):

\[
\left( \forall x_1, x_2 \in [0, \infty) \right) \left[ (x_1 = x_2) \iff (x_1^2 = x_2^2) \right].
\]

While the above is not true if we replace \([0, \infty)\) with the whole domain \( \mathbb{R} \) of \( f \), it is still useful to define such a \( g \). For instance, knowing \( f(x) = K \) gives us \( x = \pm \sqrt{K} = \pm g(K) \), so the value of \( g(K) \) is still useful in finding a particular \( x \), regardless of whether we ultimately need \( x = \sqrt{K} \) or \( x = -\sqrt{K} \). Which one is chosen usually follows from further information contained within the problem.}

There are some other issues which turn out to be easier to accommodate, such as consistency and compatibility with the other six trigonometric functions, and which will be explained as we develop the theory.

Of course the simplest uses for the arctrigonometric functions arise from solving trigonometric equations. For instance, it often happens in applications that we need to solve an equation such as \( \tan \theta = x \) for the variable \( \theta \). For this and other reasons, arctrigonometric functions become of interest. These are functions which take a number, such as \( x \), and return an angle such as \( \theta \) (also represented by a number, but the distinction is important) for which \( x \) is some given trigonometric function of that angle. So if we are interested in knowing an angle \( \theta \) such that \( \tan \theta = x \), there is an arctrigonometric function \( \arctan x \) which will give us such an angle \( \theta \) so that indeed \( \tan \theta = x \). Three arctrigonometric functions, namely arcsine, arccosine and arctangent are built into scientific calculators, along with the original trigonometric functions sine, cosine and tangent. These functions allow us to move from angle to trigonometric function of the angle, and back, almost.

The trouble is that there is important ambiguity about the angle \( \theta \) in such a problem, namely that if some angle \( \theta \) solves \( \tan \theta = x \), so do all angles of the form \( \pi n + \theta \), where \( n \in \mathbb{Z} \) (i.e., \( \theta \pm \pi, \theta \pm 2\pi, \theta \pm 3\pi, \cdots \)). Presently calculators only output a single angle for such a problem. Still, knowing one solution is quite useful, as it indirectly gives us such information as the reference angle of the other solutions. For instance, if we use a calculator for a solution to \( \tan \theta = 5 \), one solution will be given by the calculator to be \( \theta = \tan^{-1} 5 \approx 1.373400767 \) (or approximately \( 78.69000254^\circ \)). If we need an angle in the third quadrant (not practical in most right-triangle trigonometry, but useful in many applications nonetheless) we can instead use, for one example,
4.7. ARCTRIGONOMETRIC FUNCTIONS AND THEIR DERIVATIVES

\[ \tan^{-1} 5 + \pi \approx 4.514993421 \] (or \( \tan^{-1} 5 + 180^\circ \approx 258.6900675^\circ \) if we want to work in degrees). The notations \( \arctan x \) and \( \tan^{-1} x \) are used interchangeably.\(^{60}\)

Leaving the derivation for later, for now we list the derivatives in basic and chain rule forms.

\[
\begin{align*}
\frac{d}{dx} \sin^{-1} x &= \frac{1}{\sqrt{1 - x^2}}, & \frac{d}{dx} \sin^{-1} u &= \frac{1}{\sqrt{1 - u^2}} \cdot \frac{du}{dx}, \\
\frac{d}{dx} \cos^{-1} x &= \frac{-1}{\sqrt{1 - x^2}}, & \frac{d}{dx} \cos^{-1} u &= \frac{-1}{\sqrt{1 - u^2}} \cdot \frac{du}{dx}, \\
\frac{d}{dx} \tan^{-1} x &= \frac{1}{x^2 + 1}, & \frac{d}{dx} \tan^{-1} u &= \frac{1}{u^2 + 1} \cdot \frac{du}{dx}, \\
\frac{d}{dx} \cot^{-1} x &= \frac{-1}{x^2 + 1}, & \frac{d}{dx} \cot^{-1} u &= \frac{-1}{u^2 + 1} \cdot \frac{du}{dx}, \\
\frac{d}{dx} \sec^{-1} x &= \frac{1}{|x|\sqrt{x^2 - 1}}, & \frac{d}{dx} \sec^{-1} u &= \frac{1}{|u|\sqrt{u^2 - 1}} \cdot \frac{du}{dx}, \\
\frac{d}{dx} \csc^{-1} x &= \frac{-1}{|x|\sqrt{x^2 - 1}}, & \frac{d}{dx} \csc^{-1} u &= \frac{-1}{|u|\sqrt{u^2 - 1}} \cdot \frac{du}{dx}.
\end{align*}
\]

Remarkably all these derivatives are “algebraic” in nature, meaning that they involve only multiplication, division, polynomials and radicals.\(^{61}\) These emerging derivative forms, developed next and summarized in Table 4.1, page 402 will prove crucial in the later development of antiderivatives.

For our first examples we observe the following simple, similar chain rule computations. Note how \( \sin^{-1} x^2 \) is interpreted to be \( \sin^{-1} (x^2) \), etc.

**Example 4.7.2** Here we compute examples of each form where the “inner function” is \( x^2 \).

\[
\begin{align*}
\frac{d}{dx} \sin^{-1} x^2 &= \frac{1}{\sqrt{1 - (x^2)^2}} \cdot \frac{d}{dx} x^2 = \frac{1}{\sqrt{1 - x^4}} \cdot 2x = \frac{2x}{\sqrt{1 - x^4}}, \\
\frac{d}{dx} \cos^{-1} x^2 &= \frac{-1}{\sqrt{1 - (x^2)^2}} \cdot \frac{d}{dx} x^2 = \frac{-1}{\sqrt{1 - x^4}} \cdot 2x = \frac{-2x}{\sqrt{1 - x^4}}, \\
\frac{d}{dx} \tan^{-1} x^2 &= \frac{1}{(x^2)^2 + 1} \cdot \frac{d}{dx} x^2 = \frac{1}{x^4 + 1} \cdot 2x = \frac{2x}{x^4 + 1}, \\
\frac{d}{dx} \cot^{-1} x^2 &= \frac{-1}{(x^2)^2 + 1} \cdot \frac{d}{dx} x^2 = \frac{-1}{x^4 + 1} \cdot 2x = \frac{-2x}{x^4 + 1}, \\
\frac{d}{dx} \sec^{-1} x^2 &= \frac{1}{|x^2|\sqrt{(x^2)^2 - 1}} \cdot \frac{d}{dx} x^2 = \frac{1}{x^2\sqrt{x^4 - 1}} \cdot 2x = \frac{2x}{x^2\sqrt{x^4 - 1}} = \frac{2}{x\sqrt{x^4 - 1}}, \\
\frac{d}{dx} \csc^{-1} x^2 &= \frac{-1}{|x^2|\sqrt{(x^2)^2 - 1}} \cdot \frac{d}{dx} x^2 = \frac{-1}{x^2\sqrt{x^4 - 1}} \cdot 2x = \frac{-2x}{x^2\sqrt{x^4 - 1}} = \frac{-2}{x\sqrt{x^4 - 1}}.
\end{align*}
\]

\(^{60}\)Again note that the \(-1\) “exponent” in \( \tan^{-1} x \) is not actually an exponent in the sense of “power,” but is rather a notation borrowed from the study of inverse functions. Indeed, we have a name for \( (\tan x)^{-1} \), specifically \( \cot x \).

\(^{61}\)“Algebraic” is in contrast to “transcendental,” the latter referring to trigonometric, arctrigonometric, logarithmic and exponential functions, for instance. Note that the absolute value can be considered algebraic, since \( |x| = \sqrt{x^2} \).
<table>
<thead>
<tr>
<th>Function</th>
<th>Inputs (Domain)</th>
<th>Outputs (Range)</th>
<th>Outputs Graphed</th>
<th>Derivative</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sin^{-1} x$</td>
<td>$x \in [-1, 1]$</td>
<td>$\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$</td>
<td>$d \sin^{-1} x \over dx = \frac{1}{\sqrt{1-x^2}}$</td>
<td></td>
</tr>
<tr>
<td>$\cos^{-1} x$</td>
<td>$x \in [-1, 1]$</td>
<td>$\theta \in [0, \pi]$</td>
<td>$d \cos^{-1} x \over dx = -\frac{1}{\sqrt{1-x^2}}$</td>
<td></td>
</tr>
<tr>
<td>$\tan^{-1} x$</td>
<td>$x \in \mathbb{R}$</td>
<td>$\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$</td>
<td>$d \tan^{-1} x \over dx = \frac{1}{x^2 + 1}$</td>
<td></td>
</tr>
<tr>
<td>$\cot^{-1} x$</td>
<td>$x \in \mathbb{R} - {0}$</td>
<td>$\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) - {0}$</td>
<td>$d \cot^{-1} x \over dx = -\frac{1}{x^2 + 1}$</td>
<td></td>
</tr>
<tr>
<td>$\sec^{-1} x$</td>
<td>$x \in (-\infty, -1] \cup [1, \infty)$</td>
<td>$\theta \in [0, \pi] - {\frac{\pi}{2}}$</td>
<td>$d \sec^{-1} x \over dx = \frac{1}{</td>
<td>x</td>
</tr>
<tr>
<td>$\csc^{-1} x$</td>
<td>$x \in (-\infty, -1] \cup [1, \infty)$</td>
<td>$\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] - {0}$</td>
<td>$d \csc^{-1} x \over dx = -\frac{1}{</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 4.1: Summary of arctigonometric functions, their domains and ranges (given in both interval notation and graphed as angles through the unit circle), and their derivatives. Note that all angles displayed in the “Outputs Graphed” column are assumed to be between $-\frac{\pi}{2}$ and $\pi$. 
4.7. ARCTRIGONOMETRIC FUNCTIONS AND THEIR DERIVATIVES

Some patterns emerge quickly. First, these obviously come in pairs, with the difference between members of each pair being a factor of $-1$. The other patterns are more complicated and involve the forms themselves, which are summarized in the equations (4.65)–(4.70).

In later chapters it will be especially important to recognize chain rule derived derivatives. The forms where the “outer function” is an arctrigonometric function are important to recognize, though it can happen that the form is obscured by the simplification. In the next example, the outer and inner functions from above are switched.

Example 4.7.3 Here we compute examples where the “outer function” is “squaring” in each.

\[
\frac{d}{dx} (\sin^{-1} x)^2 = 2 (\sin^{-1} x) \frac{d}{dx} \sin^{-1} x = 2 \sin^{-1} x \cdot \frac{1}{\sqrt{1-x^2}} = \frac{2 \sin^{-1} x}{\sqrt{1-x^2}},
\]
\[
\frac{d}{dx} (\cos^{-1} x)^2 = 2 (\cos^{-1} x) \frac{d}{dx} \cos^{-1} x = 2 \cos^{-1} x \cdot \frac{-1}{\sqrt{1-x^2}} = \frac{-2 \cos^{-1} x}{\sqrt{1-x^2}},
\]
\[
\frac{d}{dx} (\tan^{-1} x)^2 = 2 (\tan^{-1} x) \frac{d}{dx} \tan^{-1} x = 2 \tan^{-1} x \cdot \frac{1}{x^2+1} = \frac{2 \tan^{-1} x}{x^2+1},
\]
\[
\frac{d}{dx} (\cot^{-1} x)^2 = 2 (\cot^{-1} x) \frac{d}{dx} \cot^{-1} x = 2 \cot^{-1} x \cdot \frac{-1}{x^2+1} = \frac{-2 \cot^{-1} x}{x^2+1},
\]
\[
\frac{d}{dx} (\sec^{-1} x)^2 = 2 (\sec^{-1} x) \frac{d}{dx} \sec^{-1} x = 2 \sec^{-1} x \cdot \frac{1}{|x| \sqrt{x^2-1}} = \frac{2 \sec^{-1} x}{|x| \sqrt{x^2-1}},
\]
\[
\frac{d}{dx} (\csc^{-1} x)^2 = 2 (\csc^{-1} x) \frac{d}{dx} \csc^{-1} x = 2 \csc^{-1} x \cdot \frac{-1}{|x| \sqrt{x^2-1}} = \frac{-2 \csc^{-1} x}{|x| \sqrt{x^2-1}}.
\]

Some miscellaneous examples of derivative computations using these follow:

- \[
\frac{d}{dx} \left[ \frac{\cos^{-1} x}{x} \right] = \frac{x \cdot \frac{d}{dx} \cos^{-1} x - \cos^{-1} x \cdot \frac{dx}{dx}}{(x)^2} = \frac{-\frac{x}{\sqrt{1-x^2}} - \frac{1}{\sqrt{1-x^2}}}{x^2} = \frac{-x}{x^2} - \cos^{-1} x \cdot \frac{x^2}{x^2} = \frac{-x}{x^2} - \cos^{-1} x,\]
- \[
\frac{d}{dx} \left[ x \sec^{-1} x \right] = x \frac{d}{dx} \sec^{-1} x + \sec^{-1} x \cdot \frac{dx}{dx} = x \cdot \frac{1}{|x| \sqrt{x^2-1}} + \sec^{-1} x.
\]

If we happen to know $x > 0$, this simplifies to \[
\frac{x}{x \sqrt{x^2-1}} + \sec^{-1} x = \frac{1}{\sqrt{x^2-1}} + \sec^{-1} x.
\]

If $x < 0$ we instead get \[
\frac{x}{-x \sqrt{x^2-1}} + \sec^{-1} x = \frac{1}{\sqrt{x^2-1}} + \sec^{-1} x.
\]

- \[
\frac{d}{dx} \left[ \tan^{-1}(\tan x) \right] = \frac{1}{(\tan x)^2+1} \cdot \frac{d}{dx} \tan x = \frac{1}{\tan^2 x + 1} \cdot \sec^2 x = \frac{1}{\sec^2 x} \cdot \sec^2 x = 1.
\]

A careful look at this particular function, especially in light of our later development, would reveal that for all for every interval on which $\tan x$ is defined, there exists and integer $n$ so that, $\tan^{-1}(\tan x) = x + n\pi$, so it is reasonable that this derivative above should be 1 (since $n\pi$ will be constant on such an interval). This sort of thing occurs on occasion when dealing with functions and their derivatives. Indeed sometimes it is the calculus considerations which first lead us to such simplifications of the functions.
\[ \frac{d}{dx} \sqrt{x + \tan^{-1} 2x} = \frac{1}{2\sqrt{x + \tan^{-1} 2x}} \cdot \frac{d}{dx} [x + \tan^{-1} 2x] \]
\[ = \frac{1}{2\sqrt{x + \tan^{-1} 2x}} \left[ 1 + \frac{1}{(2x)^2 + 1} \cdot \frac{d}{dx}(2x) \right] = \frac{1}{2\sqrt{x + \tan^{-1} 2x}} \cdot \left[ \frac{2}{4x^2 + 1} \right]. \]

\[ \frac{d}{dt} \sec^{-1} \left( \frac{1}{t} \right) = -\frac{1}{|t|\sqrt{(\frac{1}{t})^2 - 1}} \cdot \frac{d}{dt} \left( \frac{1}{t} \right) = -\frac{1}{|t|\sqrt{\frac{1}{t^2} - 1}} = -\frac{1}{\sqrt{1 - t^2}}. \]

Now we will note that \( t^2 = |t|^2 \), and \( |1/t| = 1/|t| \), and so we can simplify this somewhat, continuing:
\[ = \frac{1}{|t|\sqrt{(\frac{1}{t})^2 - 1}} = -\frac{1}{|t|\sqrt{\frac{1}{t^2} - 1}} = -\frac{1}{\sqrt{1 - t^2}}. \]

That derivative may seem familiar, because it is the same as the derivative of \( \cos^{-1} t \) (with respect to \( t \)). This is no accident, because (as we will see) we define \( \sec^{-1} x \) to be that angle \( \theta \in [0,\pi] \) (excluding \( \pi/2 \) but that turns out not to be a problem here) such that \( x = \sec \theta \). Then \( 1/x = \cos \theta \). We also define (later) \( \cos^{-1} t \) to be that angle \( \theta \in [0,\pi] \) such that \( \cos \theta = z \), so these two arctrigonometric functions, arcsecant and arccosine, both return similar angles. Moreover, if \( \theta \in [0,\pi] \) and \( \sec \theta = \frac{1}{t} \), then \( \cos \theta = t \); that is, if \( \theta = \sec^{-1} \frac{1}{t} \), then \( \theta = \cos^{-1} t \). If we had noticed that previously, we could have made shorter work of this derivative:
\[ \frac{d}{dt} \sec^{-1} \left( \frac{1}{t} \right) = \frac{d}{dt} \cos^{-1} t = -\frac{1}{\sqrt{1 - t^2}}. \]

\[ \frac{d}{dz} \left[ \frac{1}{\cot^{-1} z} \right] = -\frac{1}{(\cot^{-1} z)^2} \cdot \frac{d}{dz} \cot^{-1} z = -\frac{1}{(\cot^{-1} z)^2} \cdot \frac{-1}{z^2 + 1} = \frac{1}{(z^2 + 1)(\cot^{-1} z)^2}. \]

Compare the last two derivatives examples above, where in the first we could take advantage of a reciprocal relationship, but not in the second. This is because with the arctrigonometric, i.e., inverse trigonometric, functions the reciprocal relationships occur with the inputs; with trigonometric functions the reciprocal relationships occur in the outputs. For instance, \( \sec x = 1/\cos x \), while \( \sec^{-1} x = \cos^{-1}(1/x) \). That is because the arcsecant inputs a secant of an angle, which is the reciprocal of what the arccosine inputs, namely the cosine of an angle (both arctrigonometric functions returning that same angle).

We could use these facts to simplify \( \frac{d}{dx} \left[ \frac{1}{\sec x} \right] = \frac{d}{dx} \cos x = -\sin x \) instead of using a power or chain rule, while we used \( \frac{d}{dt} \sec^{-1} \frac{1}{t} = \frac{d}{dt} \cos^{-1} t \) in one of the above examples here.

The previous general derivative rules mix easily with the rules for the arctrigonometric functions, and so one who knows the previous rules need only have a list of these new derivative rules to compute derivatives involving both sets of rules. Moreover, one can then begin to notice some peculiarities of these functions which emerge from the derivative computations. Many of these peculiarities are then predictable, at least in retrospect (tautologically!), when one completely understands the geometric and algebraic natures of the arctrigonometric functions. We now delve into those natures in turn, and note some of their peculiarities. With greater understanding of the functions come more opportunities to take advantage of them preemptively, just as one might notice immediately (before invoking four chain rule computations) that \( \frac{d}{dx} (\cos^2 9x + \sin^2 9x) = 0 \) because, after all, \( \cos^2 9x + \sin^2 9x = 1. \)
4.7. ARCTRIGONOMETRIC FUNCTIONS AND THEIR DERIVATIVES

405

\(\theta\) 

\(\pi/2\) 

\(-\pi/2\) 

4.7. ARCTRIGONOMETRIC FUNCTIONS AND THEIR DERIVATIVES

\[\cos \theta, \sin \theta\]

\(-\pi/2 \leq \theta \leq \pi/2\)

\(-1 \leq \sin \theta \leq 1\)

Figure 4.24: Diagram showing that part of the unit circle used to construct the arcsine function. For all \(x \in [-1, 1]\), there exists a unique \(\theta \in [-\pi/2, \pi/2]\) so that \(\sin \theta = x\).

4.7.3 The Arcsine Function and Its Derivative

We begin our development with the inverse trigonometric functions which can be found on scientific calculators: arcsine, arccosine and arctangent (a.k.a. \(\sin^{-1}\), \(\cos^{-1}\) and \(\tan^{-1}\) respectively). Though these are the most intuitive in their derivations, in fact calculus applications find the most use for the arcsine, arctangent and arcsecant, so the derivation of the arcsecant will also be presented in full. For completeness we will include results for arccosecant and arccotangent.

We begin with the arcsine function. We are thus interested in finding a function which we will call \(\sin^{-1} x\), or \(\arcsin x\), so that \(\sin^{-1} x\) returns an angle \(\theta\) so that \(\sin \theta = x\). Now the range (of outputs) of \(\sin \theta\) is \([-1, 1]\), so our function \(\sin^{-1} x\) should be able to input any such \(x \in [-1, 1]\) and return an angle \(\theta\) whose sine is \(x\). As shown in Figure 4.24, all such outputs \(\sin \theta \in [-1, 1]\) are achieved if we restrict the input of the sine function to \(\theta \in [-\pi/2, \pi/2]\). Moreover, for each \(x \in [-1, 1]\) there exists a unique \(\theta \in [-\pi/2, \pi/2]\) such that \(\sin \theta = x\). Thus we make the following definition:

Definition 4.7.1 For every \(x \in [-1, 1]\), define

\[\sin^{-1} x = \text{“that } \theta \in [-\pi/2, \pi/2] \text{ such that } \sin \theta = x.”\] (4.71)

Note that \(\sin(\sin^{-1} x) = x\), because in that computation we are taking the sine of an angle—albeit inside of \([-\pi/2, \pi/2]\)—whose sine is \(x\), and so its sine is, naturally, \(x\).\(^62\) Using this fact, we can note that \(y = \sin^{-1} x \implies \sin y = \sin(\sin^{-1} x) \iff \sin y = x\), i.e.,

\[y = \sin^{-1} x \implies \sin y = x.\] (4.72)

The graph of \(y = \sin^{-1} x\) is thus a subset of the graph of \(\sin y = x\). This is shown in Figure 4.25.

Using (4.72), we can now derive the derivative of the arcsine function. Eventually we will

\(^{62}\)However \(\sin^{-1}(\sin x)\) is \(x\) if and only if \(x \in [-\pi/2, \pi/2]\), though \(\sin^{-1}(\sin x)\) will at least share the same reference angle as \(x\).
need to refer to a variation of Figure 4.24, but the initial computations are chain rules in nature:

\[
\begin{align*}
  y &= \sin^{-1} x \\
  \implies \sin y &= \sin(\sin^{-1} x) \\
  \implies \sin y &= x \\
  \implies \frac{d}{dx}(\sin y) &= \frac{d}{dx}(x) \\
  \implies \cos y \cdot \frac{dy}{dx} &= 1 \\
  \implies \frac{dy}{dx} &= \frac{1}{\cos y}.
\end{align*}
\]

At this point we need to rewrite this derivative in terms of \( x \) using \( y = \sin^{-1} x \):

\[
y = \sin^{-1} x \implies \frac{dy}{dx} = \frac{1}{\cos(\sin^{-1} x)}.
\] (4.73)

Now recall that \( \sin^{-1} x \in \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right] \), and in fact \( \sin^{-1} x \) is that angle \( \theta \in \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right] \) so that \( \sin \theta = x \).

There are two basic cases for this \( \theta \): that \( \theta \) is in Quadrant I or Quadrant IV. (If \( \theta \) is axial then the analysis of either case will still work.) These cases are given in Figure 4.26.

It is important to construct angles \( \theta \) with the proper representative triangles: the sine of \( \theta \) must is labeled \( x \), here representing the vertical displacement, either positive or negative. The hypotenuse is positive (since it is always a distance, not a displacement), and the quadrants are correct. The third side is constructed to be consistent with both the Pythagorean Theorem and the quadrant, the latter required to get the correct sign for the displacement represented by that.
4.7. ARCTRIGONOMETRIC FUNCTIONS AND THEIR DERIVATIVES

With two sides of such a triangle, and the quadrant known, the third side can be drawn from the Pythagorean Theorem. However note that the hypoteneuse must be positive since it is a length, while the other two sides are displacements and therefore can be positive or negative, as determined by the quadrant in which the terminal side of the angle falls.

In both cases, $x > 0$ and $x < 0$, we see that the third side of the representative triangle is $\sqrt{1-x^2}$ since it is a positive—that is, rightward—displacement in the horizontal direction. From this we can read off $\cos(\sin^{-1} x) = \cos \theta = \sqrt{1-x^2}$.

Inserting this information into our derivative computation (4.73) gives us $y = \sin^{-1} x \implies \frac{dy}{dx} = 1/\cos(\sin^{-1} x) = 1/\sqrt{1-x^2}$. We give this result in summary form, and then give the chain rule version:

\[
\frac{d}{dx} \sin^{-1} x = \frac{1}{\sqrt{1-x^2}}. \tag{4.74}
\]

\[
\frac{d}{dx} \sin^{-1} u = \frac{1}{\sqrt{1-u^2}} \cdot \frac{du}{dx}. \tag{4.75}
\]

As usual, the latter can be decomposed into

\[
\frac{d}{dx} \sin^{-1} u = \frac{d}{du} \sin^{-1} u \cdot \frac{du}{dx} = \frac{1}{\sqrt{1-u^2}} \cdot \frac{du}{dx}.
\]

Note that (4.74) only makes sense for $x \in (-1, 1)$, and that $\frac{d}{dx} \sin^{-1} x = \frac{1}{\sqrt{1-x^2}} \to \infty$ as $x \to 1^-$ and as $x \to -1^+$. This is borne out by the graph of $y = \sin^{-1} x$ given in Figure 4.25, page 406. That graph also reflects how $\frac{d}{dx} \sin^{-1} x = \frac{1}{\sqrt{1-x^2}} > 0$ for all $x \in (-1, 1)$, that is, how $\sin^{-1} x$ is increasing in that interval (and in fact, in the whole domain $x \in [-1, 1]$).

### 4.7.4 The Arccosine and Its Derivative

The development for the arccosine function mirrors that of the arcsine, except that we will take the range of the arccosine function to contain angles in Quadrants I and II, specifically $\theta \in [0, \pi]$. That is because such angles form exactly the kind of set we need so that the cosine is a one-to-one function with outputs covering the whole range $[-1, 1]$. 

---

**Figure 4.26**: Illustration of the two cases for representative triangles for angles $\theta = \sin^{-1} x$.

With two sides of such a triangle, and the quadrant known, the third side can be drawn from the Pythagorean Theorem. However note that the hypoteneuse must be positive since it is a length, while the other two sides are displacements and therefore can be positive or negative, as determined by the quadrant in which the terminal side of the angle falls.
Definition 4.7.2 For every $x \in [-1, 1]$, define

$$
\cos^{-1} x = \text{“that } \theta \in [0, \pi] \text{ such that } \cos \theta = x."
$$  \hspace{1cm} (4.76)

The arccosine function is given (in bold) in Figure 4.27. The computation of the derivative of $\cos^{-1} x$ is similar to that for the arcsine, eventually referring to an illustration of two cases, namely $\theta = \cos^{-1} x$ terminating in Quadrant I and $\theta = \cos^{-1} x$ terminating in Quadrant II.

From Figure 4.28 we see that the sine of the angle $\theta = \cos^{-1} x$ is $\sqrt{1-x^2}$ regardless of in which of the two quadrants $\theta$ terminates. Continuing our earlier computation, we have

$$
y = \cos^{-1} x \implies \frac{dy}{dx} = \frac{-1}{\sin y} = \frac{-1}{\sin(\cos^{-1} x)}.
$$
4.7. ARCTRIGONOMETRIC FUNCTIONS AND THEIR DERIVATIVES

\[ \theta = \cos^{-1} x \in [0, \pi]. \]

In both cases, the vertical side represents a positive displacement of \( \sqrt{1 - x^2} \).

Collecting this with its chain rule version, we have

\[ \frac{d}{dx} \cos^{-1} x = \frac{-1}{\sqrt{1 - x^2}}, \quad (4.77) \]

\[ \frac{d}{dx} \cos^{-1} u = \frac{-1}{\sqrt{1 - u^2}} \cdot \frac{du}{dx}. \quad (4.78) \]

The derivative of the arccosine function is negative for \( x \in (-1, 1) \), and so the function itself is decreasing. Also, the derivative approaches \( -\infty \) as \( x \to 1^- \) and as \( x \to -1^+ \).

Another derivation for the arccosine function relies upon the fact that, with these definitions of \( \sin^{-1} x \) and \( \cos^{-1} x \), we have an identity

\[ \sin^{-1} x + \cos^{-1} x = \frac{\pi}{2}. \quad (4.79) \]

This is verified easily when \( 0 < x < 1 \) because it reflects that the acute angles of a right triangle sum to \( \pi/2 \). Some checking (which we omit here) shows that (4.79) also holds for other cases in which \( x \in [-1, 1] \). With (4.79) we can take derivatives of both sides and easily get that \( \frac{d}{dx} \cos^{-1} x = -\frac{d}{dx} \sin^{-1} x \), which reflects why the derivatives of \( \cos^{-1} x \) and \( \sin^{-1} x \) are the same except for the sign.

### 4.7.5 The Arctangent Function and Its Derivative

**Definition 4.7.3** For every \( x \in \mathbb{R} \), define

\[ y = \tan^{-1} x = \text{“that } \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \text{ such that } \tan \theta = x.” \quad (4.80) \]

Thus the range we will use for the arctangent function is \( \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \). The graph of \( y = \tan^{-1} x \) is a subset of the graph of \( x = \tan y \), as illustrated in Figure 4.29.

This function \( y = \tan^{-1} x \) has some interesting features. For instance, the domain of \( \tan^{-1} x \) is all of \( \mathbb{R} \), so it makes sense to consider its behavior “at infinity.” From the graph we can see
that

\[ x \to \infty \implies \tan^{-1} x \to \frac{\pi}{2}, \quad (4.81) \]
\[ x \to -\infty \implies \tan^{-1} x \to \left(\frac{-\pi}{2}\right)^+. \quad (4.82) \]

These become important in later sections, as we consider more limits and discuss improper integrals.

In finding the derivative of the arctangent function, we proceed as in the earlier derivations, with Figure 4.30 giving the relevant triangles. Note how we construct the triangles, where \( \theta = \tan^{-1} x \). In doing so we must ensure that the quadrants and signs of the (vertical and horizontal) displacements are consistent with both the Pythagorean Theorem, and the quadrant of the terminal side of \( \theta = \tan^{-1} x \). Now we proceed with the derivative computation:
Figure 4.30: Illustration of the two cases for representative triangles for angles \( \theta = \tan^{-1} x \).

In constructing the triangles, note that we want \( \tan \theta = x \), the horizontal displacement to be positive (we picked 1), and the hypotenuse to be positive (as is always the case, since it is a distance, unlike the other two sides which represent horizontal or vertical displacements). Note also that these triangles reside within circles of radius \( \sqrt{x^2 + 1} \), which are therefore not generally unit circles.

\[
\begin{align*}
y &= \tan^{-1} x \\
\implies \tan y &= x \\
\implies \frac{d}{dx}(\tan y) &= d(x) \\
\implies \sec^2 y \cdot \frac{dy}{dx} &= 1 \\
\implies \frac{dy}{dx} &= \cos^2 y \\
\implies \frac{dy}{dx} &= (\cos y)^2 = \left( \frac{1}{\sqrt{x^2 + 1}} \right)^2 = \frac{1}{x^2 + 1}.
\end{align*}
\]

Summarizing this, and the chain rule version, we have

\[
\begin{align*}
\frac{d}{dx} \tan^{-1} x &= \frac{1}{x^2 + 1}, \quad (4.83) \\
\frac{d}{dx} \tan^{-1} u &= \frac{1}{u^2 + 1} \cdot \frac{du}{dx}. \quad (4.84)
\end{align*}
\]

Note that \( \frac{d}{dx} \tan^{-1} x = \frac{1}{x^2 + 1} > 0 \) for all \( x \in \mathbb{R} \), and so the arctangent function is increasing everywhere (see Figure 4.29). Moreover, the slope of the graph becomes gentler as \( |x| \) grows:

\[
\begin{align*}
\lim_{x \to \infty} \left[ \frac{d}{dx} \tan^{-1} x \right] &= \lim_{x \to \infty} \frac{1}{x^2 + 1} = 0, \\
\lim_{x \to -\infty} \left[ \frac{d}{dx} \tan^{-1} x \right] &= \lim_{x \to -\infty} \frac{1}{x^2 + 1} = 0.
\end{align*}
\]

This reflects the behavior of the slope of \( y = \tan^{-1} x \) as the graph approaches its horizontal asymptotes.
4.7.6 The Arcsecant Function and Its Derivative

We define the arcsecant to be consistent with the arccosine. We begin with the fact that

$$\sec \theta = x \iff \cos \theta = \frac{1}{x}.$$  

Since the range of \( \cos \theta \) is \( |\cos \theta| \leq 1 \), it follows that the range of secant is \( |\sec \theta| = \left| \frac{1}{\cos \theta} \right| \geq 1 \). We will define the arcsecant so that its domain (input) is the same as the output of \( \sec \theta \), and that its range is the same as \( \cos^{-1} \frac{1}{x} \) (almost that of arccosine, except that \( \frac{1}{2} \neq 0 \)).

**Definition 4.7.4** For \( x \in (-\infty, -1] \cup [1, \infty) \), define

$$\sec^{-1} x = \text{“that } \theta \in \left[0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right] \text{ such that } \sec \theta = x.”$$  \hspace{1cm} (4.85)

There are some complications which arise in the derivation of the arcsecant function’s derivative. That the derivation is not as straightforward as the others’ is not surprising when we consider that the arcsecant function is not even continuous on its domain. Indeed, when we pick enough of \( x = \sec y \) to cover all possible values for \( x \) but not so much that \( y \) is no longer a function of \( x \), we are forced to take two separate branches of the graph \( x = \sec y \). In Figure 4.31, that part of the graph of \( x = \sec y \) which defines \( y = \sec^{-1} x \) is highlighted.

**Figure 4.31:** Partial graph of \( x = \sec y \) in gray, with that part of the graph which represents \( y = \sec^{-1} x \), i.e., that of \( \{(x, y) \mid x = \sec y, \quad y \in [0, \frac{\pi}{2}) \cup (\frac{\pi}{2}, \pi]\} \) in black.
4.7. ARCTRIGONOMETRIC FUNCTIONS AND THEIR DERIVATIVES

\[ \theta = \left\{ \begin{array}{ll}
|1 - \sqrt{1 - x^2}| & \text{if } x \geq 1 \\
-\left| \frac{1}{x} \right| - \sqrt{1 - x^2} & \text{if } x \leq -1
\end{array} \right. \]

Case \( x \geq 1 > 0 \)

Case \( x \leq -1 < 0 \)

Figure 4.32: Illustration of the two cases for representative triangles for angles \( \theta = \sec^{-1} x \in [0, \pi] - \{ \frac{\pi}{2} \} \). We draw the triangles so that \( \sec \theta = x \), i.e., \( \cos \theta = \frac{1}{x} \). In doing so, however, we must be sure that the hypotenuse is always positive, and that the other two sides have appropriate signs. In particular, we need the hypotenuse to be \( x \) when \( x \geq 1 \) and \(-x\) when \( x \leq -1 \). In both cases the hypotenuse can be written \(|x|\). Also, in both cases the vertical side represents a positive displacement of \( \sqrt{x^2 - 1} \).

Now we derive \( \frac{dy}{dx} \sec^{-1} x \).

\[
\sec y = x \quad \Rightarrow \quad \frac{d}{dx}(\sec y) = \frac{d}{dx}(x) \quad \Rightarrow \quad \sec y \tan y \frac{dy}{dx} = 1 \quad \Rightarrow \quad \frac{dy}{dx} = \cos y \cot y = \cos(\sec^{-1} x) \cot(\sec^{-1} x).
\]

Referring to Figure 4.32, we see that the hypotenuse—which always must be positive—is \( x \) when \( x \) is positive, and \(-x\) when \( x \) is negative. In both cases, we can summarize the hypotenuse as represented by the quantity \(|x|\). Thus we can continue the implications above to get

\[
y = \sec^{-1} x \quad \Rightarrow \quad \frac{dy}{dx} = \cos(\sec^{-1} x) \cot(\sec^{-1} x) = \begin{cases} 
\frac{1}{x} & \text{if } x \geq 1, \\
-\frac{1}{x} & \text{if } x \leq -1.
\end{cases}
\]

Now we summarize these, using the fact that \( x = |x| \) in the expression for \( x \geq 1 \), while \(-x = |x|\) as well in the expression for \( x \leq -1 \). We also include the chain rule version:

\[
\frac{d}{dx} \sec^{-1} x = \frac{1}{|x|\sqrt{x^2 - 1}}, \quad (4.86)
\]

\[
\frac{d}{dx} \sec^{-1} u = \frac{1}{|u|\sqrt{u^2 - 1}} \frac{du}{dx}. \quad (4.87)
\]

Notice that (4.86) implies that \( y = \sec^{-1} x \) is increasing wherever it is differentiable, since its derivative is positive wherever defined. Notice also the limiting behavior of \( \sec^{-1} x \) as \( x \to \infty \).
and as $x \to -\infty$ (see Figure 4.31, page 412):

$$
x \to \infty \implies \sec^{-1} x \to \frac{\pi}{2}^-,
\frac{d}{dx} \sec^{-1} x \to 0; \tag{4.88}
$$

$$
x \to -\infty \implies \sec^{-1} x \to \frac{\pi}{2}^+,
\frac{d}{dx} \sec^{-1} x \to 0. \tag{4.89}
$$

A closely related derivative, which is left as an exercise, is the following (with the chain rule form included):

$$
\frac{d}{dx} \sec^{-1} |x| = \frac{1}{x \sqrt{x^2 - 1}}, \tag{4.90}
$$

$$
\frac{d}{dx} \sec^{-1} |u| = \frac{1}{u \sqrt{u^2 - 1}} \cdot \frac{du}{dx}. \tag{4.91}
$$

Equation (4.90) can be proved using the chain rule and the fact that $|x|$ is the same as $x$ for $x > 0$, and $-x$ for $x < 0$. Both cases, $x > 0$ and $x < 0$ (actually $x \geq 1$, $x \leq 1$ to be precise) should be proved separately. Equation (4.91) then follows from the chain rule. These forms are preferred to (4.86) and (4.87) when we compute antiderivatives in later sections.

### 4.7.7 Geometric and Algebraic Arguments

As explained in the context of the derivative computation $\frac{d}{dt} \sec^{-1}(1/t) = \frac{d}{dt} \cos^{-1} t$, on page 404, and in light of the angles the various arctrigonometric functions output, the following should be clear:

$$
\cos^{-1} \left(\frac{1}{x}\right) = \sec^{-1} x, \quad \sec^{-1} \left(\frac{1}{x}\right) = \cos^{-1} x, \tag{4.92}
$$

$$
\sin^{-1} \left(\frac{1}{x}\right) = \csc^{-1} x, \quad \csc^{-1} \left(\frac{1}{x}\right) = \sin^{-1} x, \tag{4.93}
$$

$$
\tan^{-1} \left(\frac{1}{x}\right) = \cot^{-1} x, \quad \cot^{-1} \left(\frac{1}{x}\right) = \tan^{-1} x. \tag{4.94}
$$

This can be useful if the argument of the arctrigonometric function is computationally more difficult to deal with than its reciprocal.

**Example 4.7.4** First we use a reciprocal relationship to compute $\frac{d}{dx} \tan^{-1} \frac{1}{x^2}$.

$$
\frac{d}{dx} \tan^{-1} \frac{1}{x^2} = \frac{-1}{(x^2)^2 + 1} \cdot \frac{d}{dx} x^2 = \frac{-2x}{x^4 + 1}. \tag{4.95}
$$

Alternatively, we can draw in the plane the angle $\theta = \tan^{-1} \frac{1}{x^2}$ and observe that one of the other trigonometric functions of $\theta$ is simpler:

$$
\theta = \tan^{-1} \frac{1}{x^2}, \quad \tan \theta = \frac{1}{x^2} > 0 \quad (\text{Quadrant I})
$$

$$
\cot \theta = x^2, \quad \theta = \cot^{-1} x^2.
$$
From the picture we see that \( \tan^{-1} \frac{1}{x} = \theta = \cot^{-1} x^2 \), and so we can use this reasoning, rather than the reciprocal relationships (4.92)–(4.94) to perform the “ALG” step in the derivative computation above. Note how we filled in the length of the hypotenuse based upon the Pythagorean Theorem, though it was not necessary for this particular example (see remarks below). Note also that since \( x \neq 0 \), we have \( 1/x^2 > 0 \), so the angle \( \theta \) must be in the first quadrant.

Because the reciprocal relationships (4.92)–(4.94) are fairly straightforward, for the above example the graphical method of deriving \( \tan^{-1} \frac{1}{x} = \cot^{-1} x^2 \) is unnecessarily involved. However, we present it here for this simpler problem to introduce the method in the context of actual derivative computations.\(^{\text{63}}\)

When using a graphical method to approach such a problem, the basic approach is, more or less, the following:

(a) to draw representative angles \( \theta \) (representing the arctrigonometric function in question),

(b) to draw triangles in the appropriate quadrants with the given trigonometric function value for the angle,

(c) to fill in the missing side length (if the hypotenuse) or displacement (if one of the other “legs”),

(d) to read off some other, computationally simpler trigonometric function of the angle, and finally,

(e) to rewrite the angle as either the arctrigonometric function of the simpler expression in (4), or some related angle.

Next we consider this approach for an example in which the reciprocal relations are insufficient.

**Example 4.7.5** Compute \( \frac{d}{dx} \tan^{-1} \left[ \frac{x}{\sqrt{9 - x^2}} \right] \).

**Solution:** Here we draw triangles for the two possible cases, \( x > 0 \) and \( x < 0 \), i.e., where \( \theta \) has positive tangent and where \( \theta \) has negative tangent, respectively. The zero tangent case can be absorbed into either. (To be more precise, the cases can be \( x \in [0, 3) \) and \( x \in (-3, 0), \) for example.) In both diagrams below we have

\[
\theta = \tan^{-1} \left[ \frac{x}{\sqrt{9 - x^2}} \right]
\]

Note how he signs associated with the three sides of the triangle are consistent with the definition of \( \theta \), the sign of \( \tan \theta \) and the geometry of the plane. Also in both cases we note the

\(^{\text{63}}\)Essentially this technique of drawing representative triangles and angles was already used in our derivations of the derivative formulas for the arctrigonometric functions. See Figure 4.26, page 407, and Figure 4.32, page 413. We re-introduce the geometric perspective repeatedly because of its versatility in application.
\[ \sin \theta = \frac{x}{3}, \text{ and in fact these angles } \theta \text{ are in the correct intervals for the range (output) of } \arcsin e. \]  
Thus we can write \( \theta = \sin^{-1} \frac{x}{3}. \) The differentiation problem then becomes

\[
\frac{d}{dx} \tan^{-1} \left[ \frac{x}{\sqrt{9-x^2}} \right] = \frac{d}{dx} \sin^{-1} \left( \frac{x}{3} \right) = \frac{1}{\sqrt{1 - \left( \frac{x}{3} \right)^2}} \cdot \frac{d}{dx} \left( \frac{x}{3} \right) = \frac{1}{\sqrt{1 - \frac{x^2}{9}}} \cdot \frac{1}{3}
\]

Notice that the calculus was finished by the end of the first line above, but often the algebraic simplifications are worthwhile.

It should be pointed out that this geometric method is not always useful. For instance, it requires us to find the third side of the “triangle,” and on inspection it is very possible that none of the trigonometric functions of the angle \( \theta \) will be substantially simpler than the original. Indeed it is usually the case the others will be more difficult. However, perhaps because these function often arise from geometric arguments regarding some physical or abstract analysis, the technique is useful enough to be worth developing. Furthermore, the simpler technique of considering the angle from the reciprocal trigonometric function as in our first solution in Example 4.7.4, page 415.

Even when the technique does yield a simpler expression for a trigonometric function for \( \theta \), there can still be complications from the quadrants. Most complications, in fact, turn out not to matter, as for instance two \(-1\) factors cancel. But care must be taken to ensure proper representation of the angle and the sides of the triangles.

**Example 4.7.6** Compute \( \frac{d}{dx} \sin^{-1} \left( \frac{\sqrt{x^2 - 25}}{x} \right) \).

**Solution:** Here we draw angles \( \theta \in [-\pi/2, \pi/2] \) such that \( \sin \theta = \frac{\sqrt{x^2 - 25}}{x} \). Recall that the hypotenuse must be positive. For that reason, the second triangle will require us to add factors of \(-1\) to two of the sides.

In both cases, the simplest trigonometric functions are in fact the secants of \( \theta \). Unfortunately, \( \text{arcsecant} \) has its range in the first and second quadrants. Note that in the second diagram above, the angle \( \phi = \sec^{-1} \frac{x}{5} = \sec^{-1} \frac{\sqrt{x^2 - 25}}{5} \), and so \( \theta = \phi - \pi \). This will be needed for the second
4.7. ARCTRIGONOMETRIC FUNCTIONS AND THEIR DERIVATIVES

We see both computations gave the same expression for the derivative, but it was not obvious from the outset. While there were geometric complications, and some algebraic cleverness needed, it is arguably easier (particularly with practice) than a direct approach:

\[
\frac{d}{dx} \sin^{-1} \left[ \frac{\sqrt{x^2 - 25}}{x} \right] = \frac{5}{|x| \sqrt{x^2 - 25}} \quad (all \ cases).
\]

Note the repeated use of the chain rule and the common algebraic techniques, in particular using \( \sqrt{x^2} = |x| \), \( x^2 = |x|^2 \), and the technique for simplifying the second fraction which arose from the chain rule.

It should also be pointed out that, if we had made early substitutions for \( |x| \), i.e., \( x \) when \( x > 0 \) and \( -x \) when \( x < 0 \), we would have different expressions for the final answer in the above example, but they can be reconciled by performing the reverse substitutions strategically.
closely related to some techniques required in Chapter 7. There the focus is with the sine, tangent, secant and their respective arctrigonometric functions, and in fact a representative triangle in the first quadrant is enough, except for the important case in which the secant is used for the substitution. In that case, both first and second quadrant cases must be considered separately.

The geometric method can also give an efficient method in many applied problems in which an expression in one trigonometric function can be written in terms of another.

**Exercises**

1. Compute the following derivatives:
   
   (a) \( \frac{d}{dx} \sqrt{\sin^{-1} x} \)
   
   (b) \( \frac{d}{dx} \sqrt{\cos^{-1} x} \)
   
   (c) \( \frac{d}{dx} \sqrt{\tan^{-1} x} \)
   
   (d) \( \frac{d}{dx} \sqrt{\cot^{-1} x} \)
   
   (e) \( \frac{d}{dx} \sqrt{\sec^{-1} x} \)
   
   (f) \( \frac{d}{dx} \sqrt{\csc^{-1} x} \)

2. Compute and simplify the following derivatives. Note that necessarily \( x, \sqrt{x} \geq 0 \) for each of these.

   (a) \( \frac{d}{dx} \sin^{-1} \sqrt{x} \)
   
   (b) \( \frac{d}{dx} \cos^{-1} \sqrt{x} \)
   
   (c) \( \frac{d}{dx} \tan^{-1} \sqrt{x} \)
   
   (d) \( \frac{d}{dx} \cot^{-1} \sqrt{x} \)
   
   (e) \( \frac{d}{dx} \sec^{-1} \sqrt{x} \)
   
   (f) \( \frac{d}{dx} \csc^{-1} \sqrt{x} \)

3. Compute the following derivatives:

   (a) \( \frac{d}{dx} \left[ x \csc^{-1} x \right] \) (assume \( x > 0 \))
   
   (b) \( \frac{d}{dx} \left[ \tan^{-1} \frac{x}{x} \right] \)

4. Compute \( \frac{d}{dx} \sec^{-1} \left( \frac{1}{x} \right) \) two ways:

   (a) Directly, using the chain rule, and

   (b) Rewriting the function as an arccosine function.

   (c) Show that the answers are in fact the same. (You may need to consider two cases, \( x \) positive and \( x \) negative.)

5. Compute \( \frac{d}{dx} \left[ \sin^{-1} \left( \frac{x}{\sqrt{x^2 + 1}} \right) \right] \) by drawing a representative triangle and re-writing the function. Be sure to consider both cases: \( x \geq 0 \) and \( x < 0 \).

6. Use the technique of the previous problem to compute each of the following:

   (a) \( \frac{d}{dx} \tan^{-1} \frac{1}{x} \)
   
   (b) \( \frac{d}{dx} \sin^{-1} \frac{1}{x} \)
   
   (c) \( \frac{d}{dx} \sec^{-1} \frac{5}{x} \)
4.7. ARCTRIGONOMETRIC FUNCTIONS AND THEIR DERIVATIVES

(d) \( \frac{d}{dx} \cos^{-1} \left( \frac{x}{\sqrt{25 + x^2}} \right) \). Assume \( x > 0 \).

(e) \( \frac{d}{dx} \sin^{-1} \left( \frac{x-1}{x} \right) \). Assume \( x > 0 \) first. Then modify for \( x < 0 \).

7. Prove (4.90), that

\[ \frac{d}{dx} \sec^{-1} x = \frac{1}{x\sqrt{x^2 - 1}}. \]

To do so, use the previous formulas for the derivative of the arcsecant function and consider the cases \( x > 0 \) and \( x < 0 \) (or more precisely, \( x > 1 \) and \( x < -1 \)) separately and show (4.90) holds for both cases.

8. Compute \( \frac{d}{dx} [\sin^{-1} x + \cos^{-1} x] \). Explain why we should have known the answer would be simple given the algebraic relationship between the arcsine and arccosine functions. (See (4.79), page 409.)

9. Consider the general problem of \( f(x) \) being one-to-one with inverse function \( f^{-1}(x) \). Next consider that

\[ y = f(x) \implies \frac{dy}{dx} = f'(x) \]

\[ x = f^{-1}(y) \implies \frac{dx}{dy} = (f^{-1})'(y). \]

Next, using \( x = f^{-1}(y) \), apply \( \frac{d}{dx} \) to both sides to show that

\[ \frac{dx}{dy} = \left( \frac{dy}{dx} \right)^{-1}. \]

\( ^{64} \) While there is general agreement about how to define the arccosine, arcsecant and arccotangent functions, there is some disagreement regarding how best to define the arcsecant, arccosecant and arccotangent. Fortunately these are rarely used, with the arcsecant being somewhat the exception. If we define the arcsecant as in the body of this Section, we have sec\(^{-1} x = \cos^{-1} \left( \frac{1}{x} \right) \), which means we do not have to change the angles when switching between the two functions to describe the same angle. However, with our definition the derivative has the absolute value in the expression, a complication absent from this alternative definition. The trade-off is that the other definition given in Problem 10 has the less natural range of outputs but a simpler derivative. At present, our approach seems to be the more popular.

The approach of Problem 10 is then extended to the arcsecant and arccotangent by using

\[ \csc^{-1} x = \frac{x}{\pi} - \sec^{-1} x, \]

\[ \cot^{-1} x = \frac{x}{\pi} - \tan^{-1} x. \]

This makes for derivatives consistent with ours, except for the lack of an absolute value in the expression in the arccosecant as well as the arcsecant.

The best justification for our approach, in fact, is in the ability to use (4.92)–(4.94), page 414.

10. Some texts define sec\(^{-1} x \) to be that \( \theta \in \left( 0, \frac{\pi}{2} \right) \cup \left( \pi, \frac{3\pi}{2} \right) \) so that sec\( \theta = x \). In other words, \( \theta \) is chosen from Quadrants I and III.

(a) Produce a graph as in Figure 4.31, page 412, with the appropriate part of the curve highlighted, to show that the highlighted curve does indeed represent a one-one function.

(b) Modify Figure 4.32, page 413 showing the same cases in which sec\( \theta = x \).

(c) Use this modified graph to show that, with this definition, in fact we have

\[ \frac{d}{dx} \sec^{-1} x = \frac{1}{x\sqrt{x^2 - 1}}. \]

(d) Explain why, with this definition, we will no longer always have sec\(^{-1} x = \cos^{-1} \left( \frac{1}{x} \right) \).
4.8 Exponential Functions

In this section we look at a function $f(x) = e^x$, where $e$ is a very important, irrational number, approximated by $e \approx 2.7182818$. What makes this function interesting, among other reasons, is that $f'(e^x) = e^x$. In fact, only functions which are constant multiples of $e^x$ are their own derivatives. Before arguing that such a function exists, we will look briefly at exponential functions $a^x$ in general, after which we will concentrate on $e^x$.

When we are finished with this section, we will then look at the inverse functions of the exponential functions. These inverses are better known as logarithms, and their derivatives fill an important gap in the theory. Furthermore, these functions have many useful algebraic properties which we can exploit before we compute the derivatives. In fact, in many problems which do not explicitly include logarithms, we can introduce them to exploit their properties to make some derivative computations proceed much faster.

Much of what we do in this section relies upon the fact that $a^x$ is an everywhere continuous, differentiable function for any $a > 0$. To actually prove this requires integral calculus (the second part of this text), but this fact is believable through observations. For our purposes here we will work from this assumption ($a^x$ continuous and differentiable for $a > 0, x \in \mathbb{R}$), see how it is reasonable, and defer the proof.

4.8.1 Exponential Functions

Of course we will have use for the algebraic rules of exponential functions, which we quickly re-list here. Assuming $a, b > 0$, $r, s \in \mathbb{R}$, $m \in \{1, 2, 3, \ldots\}$ we have

\[
\begin{align*}
a^r a^s &= a^{r+s}, \\
a^r a^{-s} &= a^{r-s}, \\
a^0 &= 1, \\
(ab)^r &= a^r b^r, \\
\left(\frac{a}{b}\right)^r &= \left(\frac{a}{b}\right)^r.
\end{align*}
\]

In this subsection we look at functions $f(x) = a^x$. In order for this function to have domain $x \in \mathbb{R}$, we restrict ourselves to $a > 0$. (Think of what $(-2)^x$ would be for $x = \frac{1}{2}, \frac{1}{4}, \frac{3}{2}, \pi$, etc.) We will generally avoid the case $a = 1$ as well, as the function $1^x$ is rather trivial.

We will look at two cases separately, namely $a > 1$ and $a \in (0, 1)$. For our prototypes, we will look at $a = 2$ and $a = \frac{1}{2}$ specifically, and argue that the same trends hold for similar $a$'s in their respective cases.

**Example 4.8.1** Consider the function $f(x) = 2^x$. There are two trends—which are in fact reflections of each other—that we will observe with this function: what happens as $x \to \infty$ and

---

65 In fact, $e$ is arguably at least equal in importance to $\pi$, though its importance is not as easily accessible.

66 The proof that $a^x$ is continuous and differentiable on all of $\mathbb{R}$ is interesting and worthwhile, but it will be offered only as a last subsection in a later chapter and section. It is included there mostly for completeness, as it could be a distraction from the main thrust of the text. The proof is long, and follows a path which is essentially backwards from how we most easily learn these functions. It relies upon an alternative definition of logarithms, proves all their properties still hold with that definition, and then considers the exponentials as inverses of the logarithms. It makes for more mathematically cohesive theory, but is counterintuitive in its path of discovery.
4.8. EXPONENTIAL FUNCTIONS

what happens as \( x \to -\infty \). We will do so by incrementing by 1 in each direction to observe the trends.

\[
\begin{align*}
2^0 &= 1 & 2^{-1} &= \frac{1}{2} & = 0.5 \\
2^1 &= 2 & 2^{-2} &= \frac{1}{4} & = 0.25 \\
2^2 &= 4 & 2^{-3} &= \frac{1}{8} & = 0.125 \\
2^3 &= 8 & 2^{-4} &= \frac{1}{16} & = 0.0625 \\
2^4 &= 16 & 2^{-5} &= \frac{1}{32} & = 0.03125 \\
2^5 &= 32 & 2^{-6} &= \frac{1}{64} & = 0.015625 \\
2^6 &= 64 & 2^{-7} &= \frac{1}{128} & = 0.0078125 \\
2^7 &= 128 & 2^{-8} &= \frac{1}{256} & = 0.00390625 \\
2^8 &= 256 & 2^{-9} &= \frac{1}{512} & = 0.001953125 \\
2^9 &= 512 & 2^{-10} &= \frac{1}{1024} & = 0.0009465625 \\
2^{10} &= 1024 & \text{etc.}
\end{align*}
\]

These trends continue as \( x \to \infty \) and \( x \to -\infty \). They are also predictable since

\[
2^{x+1} = 2^x \cdot 2, \quad 2^{x-1} = 2^x \cdot \frac{1}{2}.
\]

In other words, for every increment of one to the right, the height of the function is multiplied by a factor of 2; a movement of one to the left lowers the function’s height by half.

To Compute \( 2^r \) for rational numbers \( r = \frac{p}{q} \) where \( q \in \mathbb{N} \) is to compute \( 2^{p/q} = \sqrt[q]{2^p} \). For an irrational number \( s \in \mathbb{R} - \mathbb{Q} \), we simply take any sequence of rational numbers \( r_1, r_2, \ldots \) so that \( r_n \to s \) and define \( 2^s = \lim_{n \to \infty} 2^{r_n} \). In doing so we can eventually define \( 2^x \) for any \( x \in \mathbb{R} \), thus achieving the first graph in Figure 4.33, page 422.

To be sure, the argument above is not rigorous. With later techniques we can eventually have a rigorous proof, but the graph should be somewhat convincing, at least in its behavior at the integers. In fact, we can eventually prove that

- \( f(x) = 2^x \) is continuous for all \( x \in \mathbb{R} \),
- \( f(x) = 2^x \) is one-to-one as a function \( f : \mathbb{R} \to (0, \infty) \).

Taking these facts and the graph for granted, we also notice the following limiting behavior:

\[
\begin{align*}
x &\to \infty \quad \Rightarrow \quad 2^x \to \infty, \quad (4.95) \\
x &\to -\infty \quad \Rightarrow \quad 2^x \to 0^+. \quad (4.96)
\end{align*}
\]

We now contrast this behavior with that of the related function \( g(x) = \left(\frac{1}{2}\right)^x \).
Figure 4.33: Partial graphs of $y = 2^x$ and $y = \left(\frac{1}{2}\right)^x = 2^{-x}$. In both, a move to the right or left by one unit causes a change in the height of the graph, by a factor of 2. Such functions whose values change by a (positive) constant factor with each increment are called exponential. Increasing exponential functions are said to represent exponential growth, while decreasing exponential functions represent exponential decay.

Example 4.8.2 Consider $g(x) = \left(\frac{1}{2}\right)^x$. Some points on the graph of this function are indicated below:

$2^0 = 1$ \quad $2^{-1} = \left(\frac{1}{2}\right)^{-1} = 2$

$2^1 = \frac{1}{2} = 0.5$ \quad $2^{-2} = \left(\frac{1}{2}\right)^{-2} = 4$

$2^2 = \frac{1}{4} = 0.25$ \quad $2^{-3} = \left(\frac{1}{2}\right)^{-3} = 8$

$2^3 = \frac{1}{8} = 0.125$ \quad $2^{-4} = \left(\frac{1}{2}\right)^{-4} = 16$.

Such a function shrinks in height by a factor of $1/2$ with each increment of one unit to the right in $x$, and increases by a factor 2 with each increment of one unit to the left. Thus the behavior of $g(x) = \left(\frac{1}{2}\right)^x$ is thus just the opposite of that of $f(x) = 2^x$. This is not surprising, when we realize one is just the reflection of the other in the sense that $g(x) = f(-x)$:

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

This function $g(x)$ is illustrated by the second graph in Figure 4.96.

Any function of the form $f(x) = a^x$ where $a > 1$ represents a function which increases by the factor $a > 1$ with every increment to the right. Such behavior will ultimately imply $a^x \to \infty$ as $x \to \infty$, and $a^x \to 0^+$ as $x \to -\infty$. On the other hand, we get the opposite if $a \in (0, 1)$,
4.8. EXPONENTIAL FUNCTIONS

for we can then write \( a^x = \left( \frac{1}{a} \right)^{-x} \), which is of the form \( b^{-x} \) where \( b = \frac{1}{a} > 1 \). This limiting behavior is summarized below:

\[
\begin{align*}
\text{if } a > 1 & \implies \\
& \begin{cases} 
  a^x \to \infty & \text{as } x \to \infty, \\\n  a^x \to 0^+ & \text{as } x \to -\infty,
\end{cases} \\
\text{if } a \in (0, 1) & \implies \\
& \begin{cases} 
  a^x \to 0^+ & \text{as } x \to \infty, \\\n  a^x \to \infty & \text{as } x \to -\infty.
\end{cases}
\end{align*}
\]

The rate at which this limiting behavior occurs depends upon \( a \). For instance, \( 3^x \) increases faster than \( 2^x \) as \( x \) increases, and therefore decreases faster as \( x \) decreases. Since \( 2^x \) and \( 3^x \) agree at \( x = 0 \), and are positive everywhere, we thus have

\[
x > 0 \implies 3^x > 2^x > 0, \\
x < 0 \implies 0 < 3^x < 2^x.
\]

Similarly \( \left( \frac{1}{3} \right)^x \) shrinks faster than \( \left( \frac{1}{2} \right)^x \) as \( x \) increases, with the opposite occurring as \( x \) decreases.

Next we see how dramatic this difference in growth, of \( 2^x \) versus \( 3^x \), in two ways:

**Example 4.8.3** Show that \( 3^x \) grows faster than \( 2^x \) as \( x \to -\infty \).

**Solution:** Note first that both \( 3^x \) and \( 2^x \) are both increasing as \( x \) increases, and have the same value at \( x = 0 \). That is, \( y = 3^x \) and \( y = 2^x \) both contain the point \((0, 1)\). Furthermore,

\[
\lim_{x \to -\infty} (3^x - 2^x) = \lim_{x \to -\infty} \left[ 2^x \left( \left( \frac{3}{2} \right)^x - 1 \right) \right] = \infty. \\
\]

Thus \( 3^x \) is an increasing distance above \( 2^x \), and in fact that distance increases without bound. Alternatively, we can show \( 3^x \) grows significantly faster than \( 2^x \) by noting that

\[
\lim_{x \to -\infty} \frac{3^x}{2^x} = \lim_{x \to -\infty} \left( \frac{3}{2} \right)^x = \infty,
\]

so \( 3^x \) is a nonconstant multiple of \( 2^x \), and that multiple (namely \( \left( \frac{3}{2} \right)^x \)) is blowing up as \( x \to -\infty \).

### 4.8.2 Derivative of a Special Exponential Function

If we look back at the (computer generated) graphs in Figure 4.33, page 422 it is not unreasonable to expect that slope can be defined along these curves. Indeed, that is the case, though again we must wait to have more tools with which to prove it. Still, if we take for granted that \( a > 0 \implies \frac{d}{dx}(a^x) \) exists, we can perform the following computation. Note that \( a^x \) is constant in the limit (which varies \( \Delta x \), and not \( x \) itself).

\[
f(x) = a^x \implies f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = a^x \cdot \lim_{\Delta x \to 0} \frac{f(0 + \Delta x) - f(0)}{\Delta x} = a^x \cdot f'(0),
\]
CHAPTER 4. THE DERIVATIVE

Figure 4.34: Graphs of $2^x$, $e^x$ and $3^x$. Only $y = e^x$ has slope 1 at $x = 0$, which implies ultimately that $\frac{d}{dx}(e^x) = e^x$. See the discussion leading to (4.97)—(4.99).

To get the last line from the one immediately prior is just to recognize the definition of $f'(0)$. In all cases $a > 0$ we see that

$$f(x) = a^x \implies f'(x) = a^x \cdot f'(0). \quad (4.97)$$

So if $f(x) = a^x$, then $f'(x)$ is a constant multiple of the original function $a^x$, that constant—namely $f'(0)$—depending upon $a$.

Now, perhaps with the aid of a computer to approximate $f'(0)$ for various functions $f(x) = a^x$, it can be determined (for now without proof) that

$$f(x) = 2^x \implies f'(0) \approx 0.69314718,$$
$$f(x) = 3^x \implies f'(0) \approx 1.09861229.$$

Due to the nature of the functions, it is not hard to see that the larger the base $a$, the greater the slope of the graph of $a^x$. So for $b \in (2, 3)$, we get the slope of $b^x$ at $x = 0$ should be between that of $2^x$ and $3^x$ at $x = 0$. It is then reasonable to believe that for some number $e \in (2, 3)$, we will get $f'(0) = 1$, so that (4.97) becomes

$$f(x) = e^x \implies f'(x) = e^x \cdot 1 = e^x. \quad (4.98)$$

In fact, that number $e$ does exist (and we will find other ways to derive it much later), and $e^x$ is graphed along with $2^x$ and $3^x$ in Figure 4.34. We will refer to the function $e^x$ as the **natural exponential function**. Thus we have a derivative formula, with the chain rule version following as always:

$$\frac{d}{dx} e^x = e^x, \quad \frac{d}{dx} e^u = e^u \cdot \frac{du}{dx}. \quad (4.99)$$
The number $e$ is irrational, but can be given approximately by\footnote{We list 50 places after the decimal point because many students get the wrong impression when seeing the standard $e \approx 2.718281828$, leading them to leap to the conclusion that there is a pattern in the decimal representation. Of course, the number $2.71828$ is a “repeating decimal” and therefore rational, unlike $e$ which is irrational. (An interesting algebra exercise is to show that $2.71828 = 271,801/99,990$, an obviously rational number.)}

$$e \approx 2.71828 \; 18288 \; 50459 \; 23536 \; 02874 \; 71352 \; 66249 \; 77572 \; 47093 \; 69996.$$

The function $e^x$, together with $2^x$ and $3^x$, is given in Figure 4.34. Later in the text we will show that only constant multiples of $e^x$ are their own derivative functions. For now we can include such functions in derivative problems. In the next two examples we show some simple chain rule problems where the exponential is, alternatively, the “outer” and “inner” function.

**Example 4.8.4** Here we compute derivatives of functions where the natural exponential function is the “outer” function.

$$\frac{d}{dx} [e^{\sin^{-1} x}] = e^{\sin^{-1} x} \cdot \frac{d}{dx} \sin^{-1} x = e^{\sin^{-1} x} \cdot \frac{1}{\sqrt{1-x^2}} = \frac{e^{\sin^{-1} x}}{\sqrt{1-x^2}},$$

$$\frac{d}{dx} [e^{\cos^{-1} x}] = e^{\cos^{-1} x} \cdot \frac{d}{dx} \cos^{-1} x = e^{\cos^{-1} x} \cdot \frac{-1}{\sqrt{1-x^2}} = -\frac{e^{\cos^{-1} x}}{\sqrt{1-x^2}},$$

$$\frac{d}{dx} [e^{\tan^{-1} x}] = e^{\tan^{-1} x} \cdot \frac{d}{dx} \tan^{-1} x = e^{\tan^{-1} x} \cdot \frac{1}{x^2 + 1} = \frac{e^{\tan^{-1} x}}{x^2 + 1},$$

$$\frac{d}{dx} [e^{\cot^{-1} x}] = e^{\cot^{-1} x} \cdot \frac{d}{dx} \cot^{-1} x = e^{\cot^{-1} x} \cdot \frac{-1}{x^2 + 1} = -\frac{e^{\cot^{-1} x}}{x^2 + 1},$$

$$\frac{d}{dx} [e^{\sec^{-1} x}] = e^{\sec^{-1} x} \cdot \frac{d}{dx} \sec^{-1} x = e^{\sec^{-1} x} \cdot \frac{1}{|x| \sqrt{x^2 - 1}} = \frac{e^{\sec^{-1} x}}{|x| \sqrt{x^2 - 1}},$$

$$\frac{d}{dx} [e^{\csc^{-1} x}] = e^{\csc^{-1} x} \cdot \frac{d}{dx} \csc^{-1} x = e^{\csc^{-1} x} \cdot \frac{-1}{|x| \sqrt{x^2 - 1}} = -\frac{e^{\csc^{-1} x}}{|x| \sqrt{x^2 - 1}}.$$

**Example 4.8.5** Here we compute derivatives of functions where the natural exponential function is the “inner” function. Note that $e^x > 0$ and so $|e^x| = e^x$.

$$\frac{d}{dx} \sin^{-1} e^x = \frac{1}{\sqrt{1-(e^x)^2}} \cdot \frac{d}{dx} e^x = \frac{1}{\sqrt{1-e^{2x}}} \cdot e^x = \frac{e^x}{\sqrt{1-e^{2x}}},$$

$$\frac{d}{dx} \cos^{-1} e^x = \frac{-1}{\sqrt{1-(e^x)^2}} \cdot \frac{d}{dx} e^x = \frac{-1}{\sqrt{1-e^{2x}}} \cdot e^x = -\frac{e^x}{\sqrt{1-e^{2x}}},$$

$$\frac{d}{dx} \tan^{-1} e^x = \frac{1}{(e^x)^2 + 1} \cdot \frac{d}{dx} e^x = \frac{1}{e^{2x} + 1} \cdot e^x = \frac{e^x}{e^{2x} + 1},$$

$$\frac{d}{dx} \cot^{-1} e^x = \frac{-1}{(e^x)^2 + 1} \cdot \frac{d}{dx} e^x = \frac{-1}{e^{2x} + 1} \cdot e^x = -\frac{e^x}{e^{2x} + 1},$$

$$\frac{d}{dx} \sec^{-1} e^x = \frac{1}{|e^x| \sqrt{(e^x)^2 - 1}} \cdot \frac{d}{dx} e^x = \frac{1}{e^x \sqrt{e^{2x} - 1}} \cdot e^x = \frac{1}{\sqrt{e^{2x} - 1}},$$

$$\frac{d}{dx} \csc^{-1} e^x = \frac{-1}{|e^x| \sqrt{(e^x)^2 - 1}} \cdot \frac{d}{dx} e^x = \frac{-1}{e^x \sqrt{e^{2x} - 1}} \cdot e^x = -\frac{1}{\sqrt{e^{2x} - 1}}.$$
Example 4.8.6 Compute \( \frac{d}{dx}(e^x)^2 \) two different ways: first by the chain rule in the obvious way, and then by instead rewriting the function using properties of exponents.

Solution:

(a) \( \frac{d}{dx}(e^x)^2 = 2(e^x)\frac{d}{dx}e^x = 2e^x \cdot e^x = 2e^{2x} \).

(b) \( \frac{d}{dx}(e^x)^2 = \frac{d}{dx}e^{2x} = e^{2x} \cdot \frac{d}{dx}2x = e^{2x} \cdot 2 = 2e^{2x} \).

Of course we expect to be able to rewrite a function algebraically, when convenient, before computing a derivative, and achieve the same result, though perhaps the forms look dissimilar.

Example 4.8.7 Compute \( \frac{d}{dx}\left[ \frac{e^x}{e^{8x}} \right] \) two different ways: first by the chain rule in the obvious way, and then by instead rewriting the function using properties of exponents.

Solution: \( \frac{d}{dx}\left[ \frac{e^x}{e^{8x}} \right] = \frac{d}{dx}e^{-7x} = -7 \frac{d}{dx}e^{-7x} = -7e^{-7x} \).

The above example could have been computed using the quotient rule (an interesting exercise), but the algebraic simplification made this easier. Now we list several further examples, all of which should be self-explanatory given previous rules and the derivative formula for the natural exponential function.

- \( \frac{d}{dx}(5e^x) = 5 \cdot \frac{d}{dx}e^x = 5e^x \).
- \( \frac{d}{dx}(e^x)^2 = e^x \cdot \frac{d}{dx}e^x = e^x \cdot 2e^x = 2e^{2x} \).
- \( \frac{d}{dx}\sin e^x = \cos e^x \cdot \frac{d}{dx}e^x = e^x \cos e^x \).
- \( \frac{d}{dx}\sin^{-1} e^x = \frac{1}{\sqrt{1-\left(e^x\right)^2}} \cdot \frac{d}{dx}(e^x) = \frac{1}{\sqrt{1-e^{2x}}} \cdot e^x = \frac{e^x}{\sqrt{1-e^{2x}}} \).
- \( \frac{d}{dx}\left[ \frac{e^x}{x^2} \right] = \frac{\frac{d}{dx}(e^x) - e^x \cdot \frac{d}{dx}x^2}{(x^2)^2} = \frac{x^2e^x - e^x \cdot 2x}{x^4} = \frac{xe^x(x-2)}{x^3} \).
- \( \frac{d}{dx}(e^{\csc x}) = e^{\csc x} \cdot \frac{d}{dx}(\csc x) = e^{\csc x}(-\csc x \cot x) = -e^x \csc x \cot x \).
- \( \frac{d}{dx}(e^{2x} \sin 3x) = e^{2x} \frac{d}{dx} \sin 3x + \sin 3x \frac{d}{dx} e^{2x} = e^{2x} \cos 3x \frac{d}{dx} 3x + \sin 3x \cdot e^{2x} \frac{d}{dx} 2x = 3e^{2x} \cos 3x + 2e^{2x} \sin 3x = e^{2x}(3 \cos 3x + 2 \sin 3x) \).
- \( \frac{d}{dx}\sec^{-1} e^x = \frac{1}{|e^x| \sqrt{(e^x)^2 - 1}} \cdot \frac{d}{dx} (e^x) = \frac{1}{e^x \sqrt{e^{2x} - 1}} \cdot e^x = \frac{1}{\sqrt{e^{2x} - 1}} \).

Here we used the fact that \( e^x > 0 \) for all \( x \), so that \(|e^x| = e^x\).

- \( \frac{d}{dx}\sqrt{e^{5x}} = \frac{d}{dx} \left[ \left(e^{5x}\right)^{1/3} \right] = \frac{d}{dx} (e^{(5/3)x}) = e^{(5/3)x} \cdot \frac{d}{dx} \left[ (5/3)x \right] = \frac{5}{3} e^{(5x/3)} \).

These are all exercises involving the old differentiation rules, combined with our newest derivative formula (4.99). In the last problem above, we simplified first to avoid calling an extra chain rule.
4.8.3  A Note on Differences Between Polynomials, Exponentials

It should be well noted that these functions $a^x$ in general, and $e^x$ in particular, are very different from any of the other functions we had previously. Even though they involve powers, in the past the $x$-variable was part of the base, and not the exponent. Compare the behavior of $x^2$ and $2^x$, for instance, as well as the natures of their derivatives. For a more dramatic example, consider

$$\frac{d}{dx} x^{20} = 20x^{19},$$

$$\frac{d}{dx} 2^{x} = 2^x \cdot k,$$

$k = \frac{d 2^x}{dx} \bigg|_{x=0}.$

Not only are the power rule and exponential rule formulas not the same, but the power rule in the polynomial example decreases the power for the derivative, which does not occur in the exponential problem. The two functions share at most a vague resemblance in their behaviors. (Both increase without bound, but that is almost where the similarities end.) It is very different to raise the (variable) $x$ to a constant power, than to take a constant raised to the (variable) $x$th power. In fact we will be able to show later that any exponential growth will trump any polynomial growth:

$$\lim_{x \to \infty} \frac{a^x}{x^n} = \infty \quad \text{for } a > 1, \text{ and any fixed } n.$$  (4.101)

Thus, though the trend would not show itself until $x$ is almost unimaginably large, we have for example

$$\lim_{x \to \infty} \frac{(1.0000000000000000001)^x}{x^{1000}} = \infty.$$

It would require special programming, with a very large number of significant digits allowed, to see this trend with the help of a computer, though in a later chapter we will be able to prove the limit above with ease.

---

68 A common mistake among novice calculus students is treat exponential functions as if they are similar to polynomials when, for instance, computing derivatives. It is important to notice, for example, that

$$\frac{d}{dx} 2^x \neq x \cdot 2^{x-1}.$$  

The above derivative is not a power rule, but an exponential rule which we will derive later (essentially finding the $k$ in our already-derived formula that $\frac{d}{dx} a^x = k \cdot a^x$, where $a > 0$). Power rules assume the reverse of the exponential rules: that the variable is in the base, not the exponent (and that the exponent is a constant).
CHAPTER 4. THE DERIVATIVE

Exercises

1. Compute the following derivatives (in pairs):
   (a) \( \frac{d}{dx} e^{\sin x}, \frac{d}{dx} \sin e^x \)
   (b) \( \frac{d}{dx} e^{\cos x}, \frac{d}{dx} \cos e^x \)
   (c) \( \frac{d}{dx} e^{\tan x}, \frac{d}{dx} \tan e^x \)
   (d) \( \frac{d}{dx} e^{\cot x}, \frac{d}{dx} \cot e^x \)
   (e) \( \frac{d}{dx} e^{\sec x}, \frac{d}{dx} \sec e^x \)
   (f) \( \frac{d}{dx} e^{\csc x}, \frac{d}{dx} \csc e^x \)

2. Compute the following derivatives by first rewriting the original function using properties of exponents. (These will still require chain rules.)
   (a) \( \frac{d}{dx} \sqrt{e^x} \)
   (b) \( \frac{d}{dx} \left[ \frac{1}{e^x} \right] \)
   (c) \( \frac{d}{dx} (e^{2x})^9 \)
   (d) \( \frac{d}{dx} (e^{x^3}) \)
   (e) \( \frac{d}{dx} \left[ \frac{e^{5x}}{e^{3x}} \right] \)

3. Compute the derivatives.
   3. \( \frac{d}{dx} (e^{-x}) \)
   4. \( \frac{d}{dx} \left( e^{\sqrt{x}} \right) \)

5. \( \frac{d}{dx} \left( e^{-1/x^2} \right) \)
6. \( \frac{d}{dx} (2e^x + 9)^4 \)
7. \( \frac{d}{dx} \left( \sin^2 e^{x^3} \right) \)
8. \( \frac{d}{dx} \left( x^3 e^x \right) \) (factor your answer)
9. \( \frac{d}{dx} \left( \frac{e^{2x}}{x^2 + 1} \right) \)
10. \( \frac{d}{dx} \left( \frac{e^x}{5} \right) \) (rewrite as a multiplication first)
11. \( \frac{d}{dx} \left( \frac{e^x - e^{-x}}{2} \right) \)
12. \( \frac{d}{dx} e^x \)
13. \( \frac{d}{dx} (xe^x - e^x) \) (simplify your answer)

14. Consider \( f(x) = e^{-x^2} \). This curve is one example of a normal (also called bell) curve. Make a sign chart for \( f'(x) \) (using the fact that \( e^r > 0 \) for all \( r \in \mathbb{R} \)). Use this to show that there is a maximum at \( x = 0 \). Draw this function, showing the limiting behavior of \( f(x) \) as \( x \to \pm \infty \).

15. For general \( a > 1 \), use the facts that \( a^x > 0 \) for all \( x \), and that \( x > 0 \implies a^x > 1 \) to prove that \( a^x \) is increasing, i.e., \( x_1 < x_2 \implies a^{x_1} < a^{x_2} \). (Hint: consider \( a^{x_1} - a^{x_2} \), factor, and show this is positive.)
4.9 The Natural Logarithm I

In this section we introduce the function which is the inverse to $e^x$, namely the *natural logarithm* of $x$, denoted $\ln x$. Its derivative, as in the case with the arctrigonometric functions, is surprisingly simple and algebraic, and has nothing algebraically to do with logarithms, trigonometric, arctrigonometric or exponential functions.

We will derive the derivative of the natural logarithm, and look at several examples of derivatives involving this function. Before exploring the calculus of the natural logarithm, we will review the algebra of general logarithm functions. In doing so, we will see examples where what would be difficult derivatives can be found more quickly when we rewrite the function to a more convenient form for calculus, using the algebraic properties of logarithms.

In the next section, we will use algebraic techniques to extend the results here to compute derivatives of more general logarithm and exponential functions, such as $\log_a x$ and $a^x$. We will also develop a technique known as logarithmic differentiation which will allow faster differentiation in many cases, as well as differentiation of functions of the form $f(x)^{g(x)}$ (where the base and the exponent are both allowed to vary).

4.9.1 Algebra of Logarithms

For $a \in (0, 1) \cup (1, \infty)$ we define below the logarithm—with base $a$—of $x$, written $\log_a x$ (usually read, “log base $a$ of $x$”), as described below:

**Definition 4.9.1** For $x > 0$, $\log_a x$ is that number $y$ so that $a^y = x$. In other words, $\log_a x$ is that power of $a$ which yields $x$.

**Example 4.9.1** Consider the following logarithm computations:

- $\log_2 8 = 3$ since $2^3 = 8$,
- $\log_3 9 = 2$ since $3^2 = 9$,
- $\log_{10} \frac{1}{100} = -2$ since $10^{-2} = \frac{1}{100}$,
- $\log_{16} 2 = \frac{1}{4}$ since $16^{1/4} = 2$,
- $\log_{27} 9 = \frac{2}{3}$ since $27^{2/3} = 9$,
- $\log_4 \frac{1}{8} = -\frac{3}{2}$ since $4^{-3/2} = \frac{1}{8}$.

An observation which follows very quickly from the definition of the logarithms is the following:

$$\log_a a^x = x. \quad (4.102)$$

We will make repeated use of that observation. For now we note that with $(4.102)$ we can perform computations as above by instead rewriting the argument of the logarithm as a power of $a$: 
Example 4.9.2 We compute the following examples using (4.102).

\[
\log_2 8 = \log_2 2^3 = 3, \\
\log_{10} 1000 = \log_{10} 10^3 = 3, \\
\log_8 4 = \log_8 8^{2/3} = 2/3, \\
\log_3 \frac{1}{81} = \log_3 3^{-4} = -4, \\
\log_a a = \log_a a^1 = 1.
\]

Now we list some properties of logarithms based upon the definition. We also show how they mirror the related properties of exponents. In the table below, assume \(M = a^m\) and \(N = a^n\).

<table>
<thead>
<tr>
<th>Logarithmic Property</th>
<th>Exponential Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (\log_a (MN) = \log_a M + \log_a N)</td>
<td>1. (a^m a^n = a^{m+n})</td>
</tr>
<tr>
<td>2. (\log_a \frac{M}{N} = \log_a M - \log_a N)</td>
<td>2. (\frac{a^m}{a^n} = a^{m-n})</td>
</tr>
<tr>
<td>3. (\log_a M^p = p \cdot \log_a M)</td>
<td>3. ((a^m)^p = a^{mp})</td>
</tr>
<tr>
<td>4. (\log_a 1 = 0)</td>
<td>4. (a^0 = 1)</td>
</tr>
<tr>
<td>5. (\log_a \frac{1}{M} = -\log_a M)</td>
<td>5. (a^{-m} = \frac{1}{a^m})</td>
</tr>
</tbody>
</table>

The first logarithmic property reflects the first exponential property in the following way. When we say \(M = a^m\) and \(N = a^n\), we can think of these as stating that while \(M\) represents \(m\) factors of \(a\), and \(N\) represents \(n\) factors of \(a\), it follows that \(MN\) represents \(m+n\) factors of \(a\):

\[
\log_a \left( \frac{M}{N} \right) = \log_a \frac{M}{a^m} + \log_a \frac{N}{a^n}, \text{ i.e.,} \\
\log_a a^{m+n} = \log_a a^m + \log_a a^n, \text{ i.e.,} \\
m + n = m + n.
\]

This makes perfect sense if \(m, n \in \{0, 1, 2, 3, 4, \cdots\}\), but we can also make sense of having a half-factor of \(a\), which is then \(a^{1/2}\), i.e., the square root of \(a\). We can also talk about having \(-3\) factors of \(a\), which is like removing 3 factors, or dividing by \(a^3\), which is the same as having a factor of \(a^{-3}\). Extending this to all real powers of \(a\), we can say that by a number representing \(m\) factors of \(a\) means that number being \(a^m\), as we would have computed in the previous section. (For that discussion, see page 421.)

The second property says, roughly, that if we have \(m\) factors of \(a\), and we divide by \(n\) factors of \(a\), then we are left with \(m - n\) factors of \(a\).

The third says that if \(M\) represents \(m\) factors of \(a\), then \(M^p\) represents \(p \cdot m\) factors of \(a\).

The fourth can be interpreted as meaning, in the context of multiplication\(^{69}\) (which is arguably the context of exponents and therefore ultimately logarithms), having no factors of \(a\) is...
4.9. THE NATURAL LOGARITHM I

the same as being left with the factor 1 only. Of course it also follows from our definition of logarithms, since 1 is the zeroth power of \(a\).

The fifth property can be achieved by the second (with help from the fourth) or from the third:

\[ \log_a \frac{1}{M} = \log_a 1 - \log_a M = 0 - \log_a M = -\log_a M, \]

\[ \log_a (M^{-1}) = -1 \log_a M = -\log_a M. \]

Note that the algebraic properties of logarithms follow because logarithms are about counting factors of the base \(a\), and so products, quotients and powers are more relevant to logarithms. Thus there are no simple, general rules for expanding \(\log_a (M + N)\), for instance, because logs are not about addition of numbers per se.

However there is one last, crucial property which requires mention, since we may be interested in computing approximations to numbers such as \(\log_2 3\), though our calculating devices generally do not come equipped with the function \(\log_2( )\). Thus we need a “change of base” formula, which follows. In what is below, we assume \(a, b \in (0, 1) \cup (1, \infty)\), which (as explored in the exercises) is what we require of “proper bases” for a logarithm. In the next section we will prove the change of base formula:

\[ \log_a x = \frac{\log_b x}{\log_b a}. \] (4.103)

If we let \(b = 10\) we can compute on a standard scientific calculator

\[ \log_2 3 = (\log_{10} 3)/(\log_{10} 2) \approx 1.584962501. \]

This is reasonable, since \(2^1 = 1\) and \(2^2 = 4\), and \(f(x) = 2^x\) is continuous on all of \(\mathbb{R}\), so for some \(x \in (1, 2)\) we have \(2^x = 3\). (See Figure 4.33, page 422.)

It should be pointed out that most scientific calculators have two logarithmic keys: \(\log_{10} x\), labeled “\(\log x\)” and \(\log_e x\) given by “\(\ln x\)” Though many sciences use the logarithm with base 10, for calculus it turns out that \(\ln x\) is much more useful, as we will see.\(^70\)

One interesting aspect of (4.103) is that every function \(\log_a x\) is a constant multiple of every other such function. For instance, with \(a \in (0, 1) \cup (1, \infty)\) and \(b = e\) in (4.103), we have

\[ \log_a x = \frac{\log_e x}{\log_e a} = \frac{\ln x}{\ln a}. \] (4.104)

In this section we will develop the derivative of the natural logarithm, and with (4.104) we will therefore have derivatives of all the other logarithms.

We will see that algebraically the logarithm function in any base \(a \in (0, 1) \cup (1, \infty)\) is the inverse function of the exponential function with the same base. Thus the range of the exponential becomes the domain of the logarithm, and the domain of the exponential becomes the range of the logarithm. The range and domain of the exponential functions being \((0, \infty)\) and \(\mathbb{R}\), respectively, we have the following for general \(a \in (0, 1) \cup (1, \infty)\):

\[ \log_a a^x = x, \quad x \in \mathbb{R}, \] (4.105)

\[ a^{\log_a x} = x, \quad x > 0, \] (4.106)

\[ y = \log_a x \iff x = a^y. \] (4.107)

\(^70\)One has to be careful when reading formulas which contain “\(\log x\),” because while most texts mean by this \(\log_{10} x\), there are some which will mean \(\log_e x\), i.e., \(\ln x\). The problem is akin to knowing if \(\sin x\) is a function of \(x\) in radians, or in degrees. We will always write \(\ln x\) for \(\log_e x\) and \(\log x\) for \(\log_{10} x\).
4.9.2 Graph of the Natural Logarithm Function

We will use (4.107), with the case \( a = e \), to both graph, and eventually differentiate \( y = \ln x \).

The case \( a = e \) is important enough that it bears emphasis:

\[
y = \ln x \iff x = e^y. \tag{4.108}
\]
\[
\ln e^x = x, \quad x \in \mathbb{R}, \tag{4.109}
\]
\[
e^{\ln x} = x, \quad x > 0, \tag{4.110}
\]

According to (4.109), we see that graphing \( y = \ln x \) is the same as graphing \( x = e^y \). Of course, graphing \( x = e^y \) is the same as graphing \( y = e^x \), except that \( x \) and \( y \) have changed roles, so \((0, 1)\) being in the graph of \( y = e^x \) means that \((1, 0)\) is in the graph of \( x = e^y \), i.e., \( y = \ln x \). In Figure 4.35 we have both \( y = e^x \) (in gray), and \( x = e^y \), i.e., \( y = \ln x \), in thick black.

Referring to the graph in Figure 4.35, we see also the following limiting behavior:

\[
x \to 0^+ \iff \ln x \to -\infty, \tag{4.111}
\]
\[
x \to \infty \iff \ln x \to \infty. \tag{4.112}
\]

These follow from \( x \to -\infty \iff e^x \to 0^+ \), and \( x \to \infty \iff e^x \to \infty \), respectively. The growth in \( \ln x \) shown in the graph is indeed unbounded, but it is a very slow type of growth. In fact, such slow growth is dubbed *logarithmic growth*. We will show in a later chapter that this growth is slower than any positive power of \( x \), so for example

\[
\lim_{x \to \infty} \frac{\ln x}{x^{0.0000000001}} = 0.
\]

We can replace 0.0000000001 with any other positive number and have the same limit. More generally,

\[
\lim_{x \to \infty} \frac{\log_a x}{x^s} = 0, \quad \text{for } a > 1, \text{ and } s > 0. \tag{4.113}
\]
4.9. THE NATURAL LOGARITHM I

This is the logarithmic version of (4.101), page 427.

4.9.3 Derivative of the Natural Logarithm

We use (4.109), that is
\[ y = \ln x \iff x = e^y, \]
so we use implicit differentiation as follows:
\[
\begin{align*}
\frac{d}{dx} \ln x &= \frac{d}{dx} x \\
\frac{d}{dx} (e^y) &= \frac{d}{dx} (x) \\
e^y \frac{dy}{dx} &= 1 \\
\frac{dy}{dx} &= \frac{1}{e^y} = \frac{1}{x}.
\end{align*}
\]

We summarize this and the chain rule version:
\[
\begin{align*}
\frac{d}{dx} (\ln x) &= \frac{1}{x}, \\
\frac{d}{dx} (\ln u) &= \frac{1}{u} \cdot \frac{du}{dx}.
\end{align*}
\]

Thus the derivative of the natural logarithm is in fact a simple power, albeit the \(-1\) power.

As nice as (4.114) and (4.115) appear to be, they are incomplete. The first clue is that the derivative formulas seem to be defined as long as \(x\) (or \(u\)) is nonzero, while the logarithm required positive \(x\) (or \(u\)). Thus we know what kind of function gives derivative \(\frac{1}{x}\) as long as \(x > 0\), but \(\frac{1}{x}\) exists for \(x \neq 0\), so we should like to know what function gives rise to derivative \(\frac{1}{x}\) for \(x < 0\) as well. (For example, what kind of position gives velocity \(\frac{1}{t}\) even when \(t < 0\)?) The solution to this is to consider the functions \(\ln |x|\), and more generally \(\ln |u|\), which are defined as long as \(|x|\) and \(|u|\) are simply nonzero. In fact, our more general derivative formulas are the following:
\[
\begin{align*}
\frac{d}{dx} \ln |x| &= \frac{1}{x}, \\
\frac{d}{dx} \ln |u| &= \frac{1}{u} \cdot \frac{du}{dx}.
\end{align*}
\]

If we know \(x\) or \(u\), respectively, is positive then the absolute values above are redundant. In considering (4.116), note that the graph of \(y = \ln |x|\), given in Figure 4.36, page 434, shows how the derivative \(\frac{1}{x}\) gives reasonable slopes at several points.

We now prove (4.116), assuming (4.114) and (4.115), as follows. For convenience we first assume
\[
f(x) = \ln |x|.
\]

1. Case \(x \in (0, \infty)\): Here \(f(x) = \ln |x| = \ln x\), so \(f'(x) = \frac{1}{x}\) as before (see (4.114) above).

---

\(^{71}\)It may not be so obvious that \(x \to \infty \implies \ln x \to \infty\) from the graph. Recall
\[
\lim_{x \to \infty} f(x) = \infty \iff (\forall N)(\exists M)[x > M \implies f(x) > N].
\]

Thus we need to show that we, for any \(N\), can force \(\ln x > N\) by taking \(x > M\) for some \(M\). Later we will show that \(\ln x\) is an increasing function on its domain \((0, \infty)\), so we can take \(M = e^N\), so \(x > M \implies \ln x > \ln M = \ln e^N = N\). So for example if we want to show that eventually, as we move rightward on the \(x\)-axis, we have \(\ln x > 10,000,000,000\), we just take \(x > e^{10,000,000,000} (= M)\), a very large number but certainly finite.
Fig. 4.36: Partial graph of $f(x) = \ln |x|$ with slopes at some sample points.

2. Case $x \in (-\infty, 0)$: Here $f(x) = \ln |x| = \ln(-x)$, so we can let $u = -x > 0$, and use (4.115) as follows:

$$x < 0 \implies f(x) = \ln(-x) \implies f'(x) = \frac{1}{-x} \frac{d(-x)}{dx} = \frac{1}{-x} (-1) = \frac{1}{x}.$$

3. Thus in both cases we have $f'(x) = \frac{1}{x}$, q.e.d.

Now we can put these derivative formulas to use.

**Example 4.9.3** We compute the following derivatives:

- $\frac{d}{dx}(x \ln x) = x \cdot \frac{d}{dx} \ln x + (\ln x) \frac{d}{dx}(x) = x \cdot \frac{1}{x} + \ln x \cdot 1 = 1 + \ln x.$

- $\frac{d}{dx} \ln \cos x = \frac{1}{\cos x} \cdot \frac{d}{dx} \cos x = \frac{1}{\cos x} \cdot (-\sin x) = -\tan x.$

- $\frac{d}{dx} \ln |\cos x| = \frac{1}{\cos x} \cdot \frac{d}{dx} \cos x = \frac{1}{\cos x} \cdot (-\sin x) = -\tan x.$

Compare to the previous computation. Note that the domain for this second function is $\{x \mid \cos x > 0\}$, while here it is $\{x \mid \cos x \neq 0\}$. Where $\cos x > 0$ these are the same function.

- $\frac{d}{dx} \sin(\ln x) = \cos(\ln x) \cdot \frac{d}{dx}(\ln x) = \cos(\ln x) \cdot \frac{1}{x} = \frac{\cos(\ln x)}{x}.$

- $\frac{d}{dx} \sin(\ln |x|) = \cos(\ln |x|) \cdot \frac{d}{dx}(\ln |x|) = \cos(\ln |x|) \cdot \frac{1}{x} = \frac{\cos(\ln |x|)}{x}.$

Compare to the previous computation. Note that the domain here is $x \neq 0$, while there it was $x > 0$. For $x > 0$ they are the same function.
4.9. THE NATURAL LOGARITHM I

Example 4.9.4 Consider the following computation:

\[
\frac{d}{dx} \ln \sqrt{x} = \frac{1}{\sqrt{x}} \cdot \frac{d}{dx} \sqrt{x} = \frac{1}{\sqrt{x}} \cdot \frac{1}{2\sqrt{x}} = \frac{1}{2x}.
\]

Notice that we get one-half the answer we would have had if we had taken the derivative of \(\ln x\). In fact, this should be the case can be seen from one of the properties of logarithms:

\[
\frac{d}{dx} \ln \sqrt{x} = \frac{d}{dx} \ln(x^{1/2}) = \frac{1}{2} \frac{d}{dx} \ln x = \frac{1}{2} \cdot \frac{1}{x} = \frac{1}{2x}.
\]

Of course we had the same result using either method of computation. The lesson from this latest example is that we can at times use the properties of logarithms to algebraically rewrite the function in such a way that the derivative computation is simpler. For review, and emphasis on the natural logarithm, we revisit the properties of logarithms as applied to the special case \(a = e\). In what is below, \(M, N > 0\) and \(p \in \mathbb{R}\).

\[
\ln(MN) = \ln M + \ln N, \quad (4.118)
\]
\[
\ln \frac{M}{N} = \ln M - \ln N, \quad (4.119)
\]
\[
\ln(M^p) = p \cdot \ln M, \quad (4.120)
\]
\[
\ln 1 = 0, \quad (4.121)
\]
\[
\ln \frac{1}{M} = -\ln M, \quad (4.122)
\]
\[
\log_a M = \frac{\ln M}{\ln a}. \quad (4.123)
\]

The absolute value can be introduced easily into the arguments of the natural logarithm. Below we need only assume \(M, N \neq 0\).

\[
\ln |MN| = \ln(|M| \cdot |N|) = \ln |M| + \ln |N|, \quad (4.118)
\]
\[
\ln |M/N| = \ln(|M|/|N|) = \ln |M| - \ln |N|, \quad (4.121)
\]
\[
\ln |M^p| = \ln |M|^p = p \cdot \ln |M| \quad (4.120)
\]

Example 4.9.5 Find \(f'(x)\) if \(f(x) = \ln |x \sin x|\).

Solution: We will compute \(f'(x)\) using two methods:

1. \(f'(x) = \frac{1}{x \sin x} \cdot \frac{d}{dx} (x \sin x) = \frac{1}{x \sin x} \left[ x \cdot \frac{d}{dx} \sin x + \sin x \cdot \frac{dx}{dx} \right] = \frac{x \cos x + \sin x}{x \sin x} = \frac{x \cos x}{x \sin x} + \frac{\sin x}{x \sin x} = \cot x + \frac{1}{x}.
\]

2. If instead we first re-write \(f(x) = \ln |x| + \ln |\sin x|\), we get

\[
f'(x) = \frac{1}{x} + \frac{1}{\sin x} \cdot \frac{d}{dx} \sin x = \frac{x}{x} + \frac{1}{\sin x} \cdot \cos x = \frac{1}{x} + \cot x.
\]

\(^{72}\)Note that \(|M^p| = |M|^{|p|}\) if \(p < 0\), as for example

\[
|2^{-3}| = \left| \frac{1}{8} \right| = \frac{1}{8}, \quad |2^{-3}| = 2^3 = 8.
\]
The first method required a chain rule calling a product rule. The second required logarithm identity and a simple chain rule. As often occurs, algebraic re-writing of the function made the calculus easier.

**Example 4.9.6** Find \( \frac{d}{dx} \ln |\sin^4 x \cos^6(x^2)| \).

**Solution:** If we did not wish to use the algebraic properties of logarithms first, we would need a chain rule, calling a product rule, calling two chain rules, one of those calling yet another chain rule. It is certainly do-able, but not desirable if it can be avoided. Instead we will use the algebraic properties to expand the function to make for a simpler differentiation process. Consider the following:

\[
\frac{d}{dx} \ln |\sin^4 x \cos^6(x^2)| = \frac{d}{dx} \left[ \ln |(\sin x)^4 + \ln |(\cos(x^2))|^6 \right]
\]

\[
= \frac{d}{dx} \left[ 4 \ln |\sin x| + 6 \ln |\cos(x^2)| \right]
\]

\[
= 4 \cdot \frac{1}{\sin x} \cdot \frac{d}{dx} \sin x + 6 \cdot \frac{1}{\cos(x^2)} \cdot \frac{d}{dx} \cos(x^2)
\]

\[
= 4 \cos x \sin x + 6 \cos(x^2) \cdot (-\sin(x^2) \cdot 2x)
\]

\[
= 4 \cot x - 12x \tan(x^2).
\]

Note that the first two lines are algebraic in nature. (Note also that \( \cos(x^2) \) is not a product.)

The natural logarithm transforms products to sums, quotients to differences, and powers to multiplying factors. Each of these transformations leaves us with simpler derivative rules. With some practice, the expansion of the logarithm becomes natural and quick (we will strive for a single step!), after which the differentiation steps are relatively easy.

**Example 4.9.7** Compute \( \frac{d}{dx} \ln \left| \frac{x^3 \cos x}{\sqrt{1 + x^2}} \right| \).

**Solution:** We will write all the steps in expanding the function, but it should become clearer with practice that the final rewriting of the function can be anticipated from the original form (see previous paragraph). The first three steps below show the algebraic expansion, from which the calculus carries us to our final answer.

\[
\frac{d}{dx} \ln \left| \frac{x^3 \cos x}{\sqrt{1 + x^2}} \right| = \frac{d}{dx} \left[ \ln |x^3 \cos x| - \ln \sqrt{1 + x^2} \right]
\]

\[
= \frac{d}{dx} \left[ \ln |x^3| + \ln |\cos x| - \ln(1 + x^2)^{1/2} \right]
\]

\[
= \frac{d}{dx} \left[ 3 \ln |x| + \ln |\cos x| - \frac{1}{2} \ln(1 + x^2) \right]
\]

\[
= 3 \cdot \frac{1}{x} + \frac{1}{\cos x} \cdot \frac{d}{dx} \cos x - \frac{1}{2} \cdot \frac{1}{1 + x^2} \cdot \frac{d}{dx} (x^2 + 1)
\]

\[
= 3 \cdot \frac{\sin x}{\cos x} - \frac{1}{2(1 + x^2)} \cdot (2x)
\]

\[
= \frac{3}{x} - \tan x - \frac{x}{1 + x^2}.
\]

Note how we did not use absolute values in the \( \ln \sqrt{x^2 + 1} \) term (first line), since \( x^2 + 1 \geq 1 > 0 \).
It is interesting to note how the absolute values inside of the logarithms “disappear” in the derivative step. For that reason some texts will skip them altogether, instead assuming that the quantity inside a logarithm is positive. But to be careful, even if the argument of the ln in the derivative step. For that reason some texts will skip them altogether, instead assuming that the x in the above example, since it is possible that \( x^2 \cos x > 0 \) while the two factors are negative, making the subsequent logarithms undefined. Note how the following are all correct uses of the absolute values:

- \( \frac{d}{dx} \ln x^2 = \frac{d}{dx} |2 \ln |x|| = \frac{2}{x} \). Note \( x^2 \geq 0 \), and \( \ln x^2 \) is defined for all \( x \neq 0 \). Also, \( x^2 = |x|^2 \). We could have been more verbose:

\[
\frac{d}{dx} \ln x^2 = \frac{d}{dx} \ln |x|^2 = \frac{d}{dx} (2 \ln |x|) = 2 \cdot \frac{1}{x} = \frac{2}{x}.
\]

- \( \frac{d}{dx} [\ln x^3] = \frac{d}{dx} [3 \ln x] = \frac{3}{x} \), but is undefined if \( x < 0 \).

- \( \frac{d}{dx} [\ln |x|^3] = \frac{d}{dx} [3 \ln |x|] = \frac{3}{x} \), defined for all \( x \neq 0 \).

While the algebraic properties are what often distinguish the natural and other logarithms, in a calculus setting it is also important to note the chain rule properties, particularly when the natural logarithm is the “outside” function, but it is on occasion the “inside” function as well, so we include the following examples to illustrate both patterns.\(^73\)

**Example 4.9.8** Note that \(-1\) is not an exponent in \( \sin^{-1} x \), \( \cos^{-1} x \) and so on.

\[
\frac{d}{dx} \ln \sin^{-1} x = \frac{1}{\sin^{-1} x} \cdot \frac{d}{dx} \sin^{-1} x = \frac{1}{\sin^{-1} x} \cdot \frac{1}{\sqrt{1-x^2}} = \frac{1}{(\sin^{-1} x) \sqrt{1-x^2}}.
\]

\[
\frac{d}{dx} \ln \tan^{-1} x = \frac{1}{\tan^{-1} x} \cdot \frac{d}{dx} \tan^{-1} x = \frac{1}{\tan^{-1} x} \cdot \frac{1}{x^2 + 1} = \frac{1}{(\tan^{-1} x) (x^2 + 1)}.
\]

\[
\frac{d}{dx} \ln \sec^{-1} x = \frac{1}{\sec^{-1} x} \cdot \frac{d}{dx} \sec^{-1} x = \frac{1}{\sec^{-1} x} \cdot \frac{1}{|x| \sqrt{x^2 - 1}} = \frac{1}{(\sec^{-1} x) |x| \sqrt{x^2 - 1}}.
\]

\[
\frac{d}{dx} \sin^{-1} \ln x = \frac{1}{\sqrt{1-(\ln x)^2}} \cdot \frac{d}{dx} \ln x = \frac{1}{\sqrt{1-(\ln x)^2}} \cdot \frac{1}{x} = \frac{1}{x \sqrt{1-(\ln x)^2}}.
\]

\[
\frac{d}{dx} \tan^{-1} \ln x = \frac{1}{(\ln x)^2 + 1} \cdot \frac{d}{dx} \ln x = \frac{1}{(\ln x)^2 + 1} \cdot \frac{1}{x} = \frac{1}{x [(\ln x)^2 + 1]}.
\]

\[
\frac{d}{dx} \sec^{-1} \ln x = \frac{1}{|\ln x| \sqrt{(\ln x)^2 - 1}} \cdot \frac{d}{dx} \ln x = \frac{1}{|\ln x| \sqrt{(\ln x)^2 - 1}} \cdot \frac{1}{x} = \frac{1}{x |\ln x| \sqrt{(\ln x)^2 - 1}}.
\]

It bears repeating that the “−1” superscript on the trigonometric functions is not an exponent, so for instance

\[\ln (\sin^{-1} x) \neq -1 \cdot \ln (\sin^{-1} x).\]

Also, to be clear, when we write for instance \( \ln \sin^{-1} x \) this is taken to mean \( \ln(\sin^{-1} x) \), and \( \sin^{-1} \ln x \) means \( \sin^{-1}(\ln x) \).\(^74\)

\(^73\)However, the most interesting and important patterns are left for the exercises. It is very important for the student to see the mechanics of differentiation with logarithmic functions.

\(^74\)This is mentioned because students sometimes make the mistake of interpreting an expression like \( \ln \sin^{-1} x \) to be a product. A moment’s reflection shows that it cannot be a product, since in this expression \( \ln \) requires an input.
Exercises

1. For each of the following, compute the derivative two ways:
   (i) by using the derivative rules called by the original form. (See Example 4.9.4 on page 435);
   (ii) by using properties of the natural logarithm to rewrite the function and then compute the derivative.
   (a) \( \frac{d}{dx} \ln x^4 \)
   (b) \( \frac{d}{dx} \ln x^5 \)
   (c) \( \frac{d}{dx} \ln |x^5| = \frac{d}{dx} \ln |x| \)
   (d) \( \frac{d}{dx} \ln \left| \frac{1}{x} \right| \)
   (e) \( \frac{d}{dx} \ln \sqrt{|x|} = \frac{d}{dx} \ln \sqrt{x} \)
   (f) \( \frac{d}{dx} \ln e^x \)
   (g) \( \frac{d}{dx} \ln e^{x^2} \)

2. Consider \( f(x) = e^{\ln x} \). Note that \( f(x) = x \) (though its domain is restricted to \( x > 0 \)). Find \( f'(x) \) using the original (unsimplified) form, and show that the form of the derivative can be simplified to \( f'(x) = 1 \) (as we should expect).

3. Compute and simplify the following derivatives (using (4.117), page 433):
   (a) \( \frac{d}{dx} \ln |\sin x| \)
   (b) \( \frac{d}{dx} \ln |\cos x| \)
   (c) \( \frac{d}{dx} \ln |\tan x| \) (do not re-write first)
   (d) \( \frac{d}{dx} \ln |\cot x| \) (do not re-write first)
   (e) \( \frac{d}{dx} \ln |\sec x| \) (it is interesting to both use this form, and alternatively to rewrite the function first)
   (f) \( \frac{d}{dx} \ln |\csc x| \) (see previous comment)

Compute the desired derivative for 4–10.

4. \( \frac{d}{dx} (\ln x)^2 \)
5. \( \frac{d}{dx} \sqrt{\ln x} \)
6. \( \frac{d}{dx} \sin(\ln x) \)
7. \( \frac{d}{dx} \tan(\ln x) \)
8. \( \frac{d}{dx} \sec(\ln |x|) \)
9. \( \frac{d}{dx} \ln |\ln x| \)
10. \( \frac{d}{dx} \ln(\ln(\ln x)) \)

For Exercises 11–15, compute \( f'(x) \) for each of the following by first rewriting \( f(x) \) by expanding the logarithms.

11. \( f(x) = \ln |x^3(x^2 + 3x)^{20}| \)
12. \( f(x) = \ln \left| (2x + 9)^3(3x^2 + 5x)^9(2 - 7x)^{10} \right| \)
13. \( f(x) = \ln \left| \frac{9x - 1}{2x + 4} \right| \)
14. \( f(x) = \ln \left| \frac{(3x + 5)^7}{(7x^2 + 2)^8} \right| \)
15. \( f(x) = \ln \left| \frac{x^2 \sin^3 x}{\cos^4 2x \sqrt{x^2 - 9}} \right| \)

16. Compute and simplify the following derivatives (which cannot take advantage of the properties of logarithms).
   (a) \( \frac{d}{dx} (x \ln x - x) \)
   (b) \( \frac{d}{dx} \sin(\ln |\cos x|) \)
4.9. THE NATURAL LOGARITHM I

(c) \( \frac{d}{dx} \ln(x^2 + 1) \). Why is \( \ln |x^2 + 1| \) the same as \( \ln(x^2 + 1) \)?

17. Show that \( \frac{d}{dx} \ln |\sec x + \tan x| = \sec x \).
(Use the simplest derivative strategy. The key is in the simplification of the derivative.)

18. Compute the following two derivatives directly (using the chain rule formula for the derivative of the natural logarithm), show that they are the same, and explain why we should have expected they were the same:
\[
\frac{d}{dx} \ln |\sec x|, \quad \frac{d}{dx} \ln |\cos x|.
\]

19. Compute the following derivatives and show that they simplify as described.

(a) \( \frac{d}{dx} \ln |\sec x + \tan x| = \sec x \).
(b) \( \frac{d}{dx} \ln |\csc x + \cot x| = -\csc x \).

20. For \( f(x) = (g(x))^h(x) \), where \( g(x) > 0 \) derive the following formula for \( f'(x) \):
\[
f'(x) = h(x)(g(x))^{h(x)-1}g'(x) + (g(x))^{h(x)}h'(x) \ln(g(x)). \tag{4.124}
\]
To do so, follow the following steps:

(a) Use the idea that \( a = e^{\ln a} \) (so what is \( a^2 \)) to show
\[
f(x) = e^{[\ln(g(x))-h(x)]}.
\]
(b) Find \( f'(x) \) using this form.
(c) Simplify \( f'(x) \) from the step above to achieve (4.124).

21. Assume \( h(x) = n \) is constant. Then rewrite \( f(x) = (g(x))^h(x) \) and use (4.124) to compute \( f'(x) \) for this case.

22. Repeat the previous problem supposing instead that \( g(x) = a \) is a constant, and \( h(x) \) is allowed to vary.

23. Now we consider the rationale for not allowing \( a = 1 \) to be the base of a logarithm. To do so, we consider how we would attempt to develop a function \( \log_a x \).

(a) First, recall how we found the graph of \( y = \ln x = \log_e x \) based upon the graph of \( y = e^x \). Show what would happen if we attempted to do this for 1 instead of \( e \) as the base.
(b) Separately, consider the definition of the logarithm function, and decide what would be the domain of \( \log_1 x \).
(c) Separately still, explain why (4.123) would preclude \( a = 1 \).
4.10 The Natural Logarithm II: Further Results

In this section we use the properties of the natural logarithm, and its relationship to exponential functions and other logarithms, to pursue further differentiation problems. The first technique we will develop is called logarithmic differentiation, in which we actually introduce the natural logarithm into problems which can benefit from its presence. In later subsections, we use change of base-type techniques to rewrite problems into forms for which we have formulas. In doing so, some new and more general differentiation rules will emerge.

4.10.1 Logarithmic Differentiation

Because the natural logarithm takes products to sums, quotients to differences, and powers to multiplying factors, we have already seen that many derivative problems involving natural logarithms can be much reduced in complexity through algebraic expansion of a relevant logarithm. By properly introducing the natural logarithm into differentiation problems involving products, quotients and powers, we can take advantage of the logarithm’s algebraic expansion properties. Introducing the logarithm into the problem inserts its own complications, but these are relatively minor compared to previous, brute-force methods for many of these computations. We begin with an example to illustrate the method.

Example 4.10.1 Find \( f'(x) \) if \( f(x) = \frac{x \sin^3 x}{\sqrt{x^2 + 1}} \).

Solution: The technique below is similar to our implicit differentiation, except that first we apply the function \( \ln |\cdot| \) to both sides in the following sense:

\[
\ln |f(x)| = \ln \left| \frac{x \sin^3 x}{\sqrt{x^2 + 1}} \right|
\]

The whole point of doing so is to be able to take advantage of the algebraic properties of logarithms (so that the differentiation steps will be easier).

\[
\Rightarrow \quad \ln |f(x)| = \ln |x| + \ln |\sin^3 x| - \ln \left| (x^2 + 1)^{1/2} \right|
\]

Now we differentiate, i.e., apply \( \frac{d}{dx} \) to both sides. Note where we use \( \frac{d}{dx} \ln |u| = \frac{1}{u} \cdot \frac{du}{dx} \).

\[
\Rightarrow \quad \frac{d}{dx} \ln |f(x)| = \frac{d}{dx} \left[ \ln |x| + \ln |\sin^3 x| - \ln \left| (x^2 + 1)^{1/2} \right| \right]
\]

\[
\Rightarrow \quad \frac{1}{f(x)} \frac{df(x)}{dx} = \frac{1}{x} + 3 \cdot \frac{1}{\sin x} \cdot \frac{d}{dx} \sin x - \frac{1}{2} \cdot \frac{1}{x^2 + 1} \cdot \frac{d}{dx} (x^2 + 1)
\]

\[
\Rightarrow \quad \frac{f'(x)}{f(x)} = \frac{1}{x} + 3 \cos x \cdot \frac{1}{\sin x} - \frac{2x}{2(x^2 + 1)}
\]

\[
\Rightarrow \quad \frac{f'(x)}{f(x)} = \frac{1}{x} + 3 \cot x - \frac{x}{x^2 + 1}
\]
4.10. THE NATURAL LOGARITHM II: FURTHER RESULTS

We wanted \( f'(x) \), so we next multiply both sides by \( f(x) \), which gives \( f'(x) \) in terms of both \( x \) and \( f(x) \). We then finish by substituting the original form for \( f(x) \).

\[
f'(x) = f(x) \left[ \frac{1}{x} + 3 \cot x - \frac{x}{x^2 + 1} \right] = \frac{x \sin^3 x}{\sqrt{x^2 + 1}} \left[ \frac{1}{x} + 3 \cot x - \frac{x}{x^2 + 1} \right].
\]

While the process above did require several steps, those steps were arguably simpler than those in our earlier methods, which would have called for a quotient rule calling one product rule and two chain rules. Furthermore with practice the process of logarithmic differentiation can be streamlined to be much faster. For instance, if we can anticipate the final logarithmic expansion, and also use the shortcut

\[
\frac{d \ln |u(x)|}{dx} = \frac{u'(x)}{u(x)}, \tag{4.125}
\]

particularly when \( u'(x) \) is simple to compute, then we can consolidate steps from the previous example. (For clarity, in fact the last step from before becomes two steps below.)

\[
f(x) = \frac{x \sin^3 x}{\sqrt{x^2 + 1}}
\]

\[\Rightarrow\] \( \ln |f(x)| = \ln |x| + 3 \ln |\sin x| - \frac{1}{2} \ln |x^2 + 1| \)

\[\Rightarrow\] \( \frac{f'(x)}{f(x)} = \frac{1}{x} + 3 \cdot \frac{\cos x}{\sin x} - \frac{1}{2} \cdot \frac{2x}{x^2 + 1} \)

\[\Rightarrow\] \( f'(x) = f(x) \left[ \frac{1}{x} + 3 \cot x - \frac{x}{x^2 + 1} \right] \)

\[\Leftarrow\] \( f'(x) = \frac{x \sin^3 x}{\sqrt{x^2 + 1}} \left[ \frac{1}{x} + 3 \cot x - \frac{x}{x^2 + 1} \right]. \)

To be sure, more steps can and arguably should be written when the technique is first learned, but it should appear true that the abbreviated process above is much preferable to our earlier techniques, and that such efficiency is an achievable goal.

The process of logarithmic differentiation does not replace our earlier methods entirely. Indeed, it is only immediately useful for finding \( f'(x) \) if \( f(x) \) is the type of function whose natural log (of its absolute value to be more precise) can be expanded in a useful way. For instance, the function above yielded nicely to the process because it consisted of powers of functions, combined through multiplication and division. However, it would not be advantageous, for instance, to attempt logarithmic differentiation on a function such as \( f(x) = \sec x + \tan x + 9x^2 - \frac{1}{x} \), since this is a sum, and there is no algebraic expansion for the natural log of a sum.\(^{75}\)

The logarithmic differentiation process for finding \( f'(x) \) for a given \( f(x) \) is as follows:

1. Beginning with the equation which defines \( f(x) \), apply \( \ln |\cdot| \) to both sides.

2. Expand the logarithm on the right-hand side (may be consolidated into Step 1).

3. Apply \( \frac{d}{dx} \) to both sides. the left-hand side will be \( \frac{f'(x)}{f(x)} \).

4. Multiply the resulting equation by \( f(x) \), thus solving for \( f'(x) \) in terms of \( x \) and \( f(x) \).

\(^{75}\)One could rewrite \( f(x) = y_1 + y_2 + y_3 + y_4 \), for instance, and perform logarithmic differentiation on each \( y_k \) separately to find each \( y'_k \), and thus have \( f'(x) = y'_1 + y'_2 + y'_3 + y'_4 \), and indeed sometimes this is necessary. For the example this note refers to, that would certainly not decrease the work required to find \( f'(x) \).
5. Substitute the original formula for \( f(x) \) on the right-hand side (may be consolidated into Step 4).

We will see that there are other times where applying \( \ln \left| \cdot \right| \) to both sides before differentiation is advantageous besides just for finding certain derivatives \( f'(x) \), but for now another example of this first type is called for.

**Example 4.10.2** Find \( f'(x) \) if \( f(x) = \frac{5 \sin 2x \cos^3 4x}{\sqrt{1 + \tan 6x}} \).

**Solution:** We proceed as before, this time being more verbose in our chain rule computations.

\[
f(x) = \frac{5 \sin 2x \cos^3 4x}{\sqrt{1 + \tan 6x}}
\]

\[
\Rightarrow \quad \ln |f(x)| = \ln \left| \frac{5 \sin 2x \cos^3 4x}{(1 + \tan 6x)^{1/2}} \right|
\]

\[
\Rightarrow \quad \ln |f(x)| = \ln |5| + \ln |\sin 2x| + 3 \ln |\cos 4x| - \frac{1}{3} \ln |1 + \tan 6x|
\]

\[
\Rightarrow \quad \frac{d}{dx} \ln |f(x)| = \frac{\cos 4x}{\sin 2x} \cdot \frac{d}{dx} \left[ \ln |5| + \ln |\sin 2x| + 3 \ln |\cos 4x| - \frac{1}{3} \ln |1 + \tan 6x| \right]
\]

\[
\Rightarrow \quad \frac{f'(x)}{f(x)} = 0 + \frac{1}{\sin 2x} \cdot \frac{d}{dx} \left[ \ln |5| + \ln |\sin 2x| + 3 \ln |\cos 4x| - \frac{1}{3} \ln |1 + \tan 6x| \right]
\]

\[
\Rightarrow \quad \frac{f'(x)}{f(x)} = \frac{1}{\sin 2x} \cdot \frac{d}{dx} \left[ \ln |5| + \ln |\sin 2x| + 3 \ln |\cos 4x| - \frac{1}{3} \ln |1 + \tan 6x| \right]
\]

\[
\Rightarrow \quad f'(x) = 2 \cot 2x + 3 \tan 4x \cdot (-4) - \frac{6 \sec^2 6x}{3(1 + \tan 6x)}
\]

\[
\Rightarrow \quad f'(x) = f(x) \left[ 2 \cot 2x - 12 \tan 4x - \frac{2 \sec^2 6x}{1 + \tan 6x} \right]
\]

\[
\Rightarrow \quad f'(x) = \frac{5 \sin 2x \cos^3 4x}{\sqrt{1 + \tan 6x}} \left[ 2 \cot 2x - 12 \tan 4x - \frac{2 \sec^2 6x}{1 + \tan 6x} \right].
\]

A few more notes about the process are in order.

(i) The absolute values introduced with the natural logarithm vanish in the derivative step. For this reason some textbooks do not include them, but technically they should be included. One nice feature of the process is that the correct answer can be found even when absolute values are (naively?) omitted.

(ii) The answer this process delivers is of a different form than our earlier methods, but they are algebraically the same, except that the answer here may need to be expanded and simplified to be, technically, completely correct. For instance, if \( \cos 4x = 0 \) this answer appears undefined because of the \( \tan 4x \) term, though \( \cos 4x = 0 \) does not necessarily break the differentiability. When we distribute the \( f(x) \) factor across the brackets in our final answer, a factor of \( \cos 4x \) will cancel the denominator in the \( \tan 4x \) term. Algebraically that is not correct if \( \cos 4x = 0 \), but in fact the naively simplified form—with the \( \cos 4x \) term (as well as the \( \sin 2x \) in the \( f(x) \) and \( \cot 2x \) terms) canceled—ultimately gives the correct derivative \( f'(x) \).
4.10. THE NATURAL LOGARITHM II: FURTHER RESULTS

(iii) Related to the previous item, note that we cannot technically compute \( \ln |f(x)| \) wherever \( f(x) = 0 \) (such as when \( \sin 2x = 0 \) or \( \cos 4x = 0 \)), but this too gets glossed over in the differentiation process, especially in the final answer if \( f(x) \) is distributed across the brackets, and the offending terms canceled.

Thus in some ways logarithmic differentiation is better than expected, in that even when certain things technically should be going wrong in the process (such as being outside the domain of the natural logarithm), in the end—at least in the simplified answer—we get the correct result.

Logarithmic differentiation gives a nice proof of the generalized product rule.\(^{76}\)

**Theorem 4.10.1** For \( f(x) = g_1(x)g_2(x)g_3(x) \cdots g_n(x) \), we have

\[
f'(x) = (g_1'(x)g_2(x)g_3(x) \cdots g_n(x)) + (g_1(x)g_2'(x)g_3(x) \cdots g_n(x)) + \cdots + (g_1(x)g_2(x)g_3'(x) \cdots g_n(x)). \tag{4.126}
\]

A proof in the case \( f(x) = g_1(x)g_2(x)g_3(x) \) shows the pattern of argument for the general case.

\[
\frac{d}{dx} |\ln f(x)| = \frac{d}{dx} |\ln g_1(x)| + \ln |g_2(x)| + \ln |g_3(x)|
\]

\[
f'(x) = g_1(x)g_2(x)g_3(x)
\]

\[
\ln |f(x)| = \ln |g_1(x)||g_2(x)||g_3(x)|
\]

\[
\frac{d}{dx} |\ln f(x)| = \frac{d}{dx} |\ln g_1(x)| + \ln |g_2(x)| + \ln |g_3(x)|
\]

\[
f'(x) = g_1'(x) + g_2'(x) + g_3'(x)
\]

\[
n f(x) = g_1(x)g_2(x)g_3(x)
\]

\[
\frac{d}{dx} |\ln f(x)| = \frac{d}{dx} |\ln g_1(x)| + \ln |g_2(x)| + \ln |g_3(x)|
\]

\[
\frac{f'(x)}{f(x)} = \frac{g_1'(x)}{g_1(x)} + \frac{g_2'(x)}{g_2(x)} + \frac{g_3'(x)}{g_3(x)}
\]

\[
f'(x) = f(x) \left[ \frac{g_1'(x)}{g_1(x)} + \frac{g_2'(x)}{g_2(x)} + \frac{g_3'(x)}{g_3(x)} \right]
\]

\[
f'(x) = g_1(x)g_2(x)g_3(x) \left[ \frac{g_1'(x)}{g_1(x)} + \frac{g_2'(x)}{g_2(x)} + \frac{g_3'(x)}{g_3(x)} \right], \text{ q.e.d.}
\]

**Example 4.10.3** Thus we can perform the following quickly, this time in primed notation.

\[
\frac{d}{dx} (x^3e^x \sin x) = (x^3)'e^x \sin x + x^3(e^x)' \sin x + x^3e^x(\sin x)'
\]

\[
= 3x^2e^x \sin x + x^3e^x \sin x + x^3e^x \cos x
\]

\[
= x^2e^x(3 \sin x + x \sin x + x \cos x)
\]

If such a product is more complicated, we should revert to Leibniz notation:

\[
\frac{d}{dx} (f(x)g(x)h(x)) = f(x)g(x) \cdot \frac{d h(x)}{dx} + f(x)h(x) \cdot \frac{d g(x)}{dx} + g(x)h(x) \cdot \frac{d f(x)}{dx}
\]

With more complicated \( f, g \) or \( h \), this has the advantage that the derivative computations are the rightmost factors in each term, so keeping them separate from the other terms, and expanding as we call up the various rules, is more convenient. To take full advantage of the Leibniz style, we therefore have to write the terms of a generalized product rule in a different order than given in the theorem above.

\(^{76}\) See notes on the roles of the various factors in the product rule, page 368. Comments there generalize to more general products.
Example 4.10.4 Later in the text we will often be computing derivatives with respect to time \( t \), though that variable might not explicitly appear in the problem. Still it will make sense to apply \( \frac{d}{dt} \) to quantities which do, in fact, depend upon time \( t \). So for instance there is the formula from chemistry that \( PV = k \cdot T \), where \( k \) is a constant, \( P, V \) and \( T \) are pressure, volume and absolute temperature, respectively. This is the equation for a so-called ideal gas with some fixed number of particles. In a mathematically sophisticated advanced chemistry text or article, it is not unusual to see a computation like

\[
PV = k \cdot T \implies \frac{1}{P} \cdot \frac{dP}{dt} + \frac{1}{V} \cdot \frac{dV}{dt} = \frac{1}{T} \cdot \frac{dT}{dt}.
\]

Without logarithmic differentiation this might seem rather mysterious, and difficult to prove easily. However, one well versed in the technique would likely see the truth of this implication quickly, being practiced enough in the middle steps to anticipate the outcome:

\[
PV = k \cdot T \implies \ln(PV) = \ln(kT) \\
\implies \ln P + \ln V = \ln k + \ln T \\
\implies \frac{d}{dt} [\ln P + \ln V] = \frac{d}{dt} [\ln k + \ln T] \\
\implies \frac{1}{P} \cdot \frac{dP}{dt} + \frac{1}{V} \cdot \frac{dV}{dt} = 0 + \frac{1}{T} \cdot \frac{dT}{dt}, \quad \text{q.e.d.}
\]

We would have a different equation involving these functions (of \( t \)) \( P, V, T \) if we simply applied \( \frac{d}{dt} \) to both sides rather than first applying the natural logarithm function. If we chose that strategy we would require the product rule on the left side. In fact the equations we get with either method are equivalent under the original assumption, that \( PV = k \cdot T \). To show the two equations involving the derivatives in fact say the same thing is an exercise in algebra.

Note also that we normally apply \( \ln | \cdot | \) to both sides when intending to perform logarithmic differentiation, but here we could just apply \( \ln(\cdot) \) because the quantities involved are never negative.

Recalling that powers become multiplying factors, another quick example, this time from basic electricity, would be

\[
P = \frac{E^2}{R} \implies \ln P = 2 \ln E - \ln R \\
\implies \frac{1}{P} \cdot \frac{dP}{dt} = \frac{2}{E} \cdot \frac{dE}{dt} - \frac{1}{R} \cdot \frac{dR}{dt}.
\]

With practice one is quite likely to be confident enough to skip the middle step.

4.10.2 Bases Other Than \( e \): Logarithms

In this subsection we look at derivatives of functions \( \log_a x \) for more general \( a \). This is a simple application of our change of base formula (4.103), page 431 (note that \( \frac{1}{\ln a} \) is a constant):

\[
\frac{d}{dx} \log_a x = \frac{d}{dx} \left[ \frac{\ln x}{\ln a} \right] = \frac{d}{dx} \left[ \frac{1}{\ln a} \cdot \ln x \right] = \frac{1}{\ln a} \cdot \frac{1}{x} = \frac{1}{x \ln a}.
\]

We can also perform the computation above with \( |x| \) replacing \( x \), and the same argument which worked with \( \frac{d}{dx} \ln |x| = \frac{1}{x} \) will work here. In fact, since this function \( \log_a x \) is just a constant
multiple of \( \ln x \), all the rules we had for \( \ln x \) work here, with the constant carrying through. Thus the absolute value and chain rule versions are just

\[
\frac{d}{dx} \log_a |x| = \frac{d}{dx} \left[ \frac{1}{\ln a} \ln |x| \right] = \frac{1}{x \ln a},
\]  
(4.127)

\[
\frac{d}{dx} \log_a |u| = \frac{d}{dx} \left[ \frac{1}{\ln a} \ln |u| \right] = \frac{1}{u \ln a} \cdot \frac{du}{dx}.
\]  
(4.128)

**Example 4.10.5** Consider the following derivative computations:

- \[
\frac{d}{dx} \log_{10} x = \frac{1}{x \ln 10}.
\]

- \[
\frac{d}{dx} \log_2 \left| \sqrt{x^2 + 1} \right| = \frac{d}{dx} \left[ \log_2 |x| + \frac{1}{2} \log_2 (x^2 + 1) \right] = \frac{1}{x \ln 2 + \frac{1}{(x^2 + 1) \ln 2} \cdot \frac{d}{dx} (x^2 + 1)}
\]

- \[
\frac{d}{dx} \log_3 \left| \frac{\sin x}{2x + 5} \right| = \frac{1}{\sin x \ln 3} \cdot \frac{d}{dx} \left[ \log_3 |\sin x| - \log_3 |2x + 5| \right]
\]

- \[
\frac{d}{dx} \log_3 (\log_5 x) = \frac{1}{\log_5 x \ln 3} \cdot \frac{d}{dx} \left[ \log_5 (\ln x) - \log_5 (\ln 5) \right] = \frac{1}{\ln x \ln 3} \cdot \frac{d}{dx} (\ln x) - 0 = \frac{1}{\ln x \ln 3 \cdot x}.
\]

For this last example, we can cut short the calculus steps using algebraic properties of logarithms first:

\[
\frac{d}{dx} \log_3 (\log_5 x) = \frac{1}{\ln x \ln 3} \cdot \frac{d}{dx} \left[ \log_5 x \ln 3 = \frac{1}{\ln x} \cdot \frac{1}{\ln x} \cdot \frac{1}{x} = \frac{1}{\ln x \ln 3 \cdot x} \right].
\]

**4.10.3 Bases other than \( e \): Exponential Functions**

Here we look at derivatives of exponential functions \( a^x \) (\( a > 0 \)) and functions involving these. Two computations of the derivative of \( a^x \) are offered. Both illustrate useful techniques which are worth remembering. The first technique is logarithmic differentiation. We find \( \frac{dy}{dx} \) under the assumption \( y = a^x \). Note that \( y > 0 \) so no absolute values are needed.

\[
y = a^x
\]

\[
\Rightarrow \quad \ln y = \ln a^x
\]

\[
\Leftrightarrow \quad \ln y = x \ln a
\]

\[
\Rightarrow \quad \frac{d}{dx} \ln y = \frac{d}{dx} \left[ (\ln a)x \right]
\]

\[
\Rightarrow \quad \frac{1}{y} \frac{dy}{dx} = \ln a
\]

\[
\Rightarrow \quad \frac{dy}{dx} = y \ln a
\]

\[
\Leftrightarrow \quad \frac{dy}{dx} = a^x \ln a.
\]
Alternatives, we can instead use the fact that \( a = e^{\ln a} \), and compute \( \frac{d}{dx} [a^x] \) using the chain rule. Note that \( a^x = [e^{\ln a}]^x = e^{(\ln a)x} \).

\[
\frac{d}{dx} a^x = \frac{d}{dx} [e^{\ln a}]^x = \frac{d}{dx} e^{(\ln a)x} = e^{(\ln a)x} \cdot \frac{d}{dx} (\ln a) = e^{(\ln a)x} \cdot \ln a = a^x \ln a.
\]

From either method, we get the derivative of \( a^x \), and its chain rule version:

\[
\frac{d}{dx} a^x = a^x \ln a, \quad (4.129)
\]
\[
\frac{d}{dx} a^u = a^u \ln a \cdot \frac{du}{dx}, \quad (4.130)
\]

Note that if \( a = e \), we have \( \ln a = \ln e = 1 \), giving us \( \frac{d}{dx} e^x = e^x \ln e = e^x \cdot 1 = e^x \) as before.

We proved (4.129), from which (4.130) follows. While the first technique was the (now) familiar logarithmic differentiation, the second was to change the base, rewriting \( a^x = (e^{\ln a})^x = e^{(\ln a)x} \). This can be exploited in other venues, but in particular it allows one to use the simpler and ubiquitous derivative formula for \( e^x \) by an algebraic rewriting, rather than relying upon the more obscure (but not unimportant) formula (4.129).

**Example 4.10.6** We can now compute the following derivatives using (4.129) and (4.130):

- \( \frac{d}{dx} 2^x = 2^x \ln 2 \).
- \( \frac{d}{dx} 3^5x = 3^{5x} \ln 3 \cdot \frac{d}{dx} [5x] = 3^{5x} \ln 3 \cdot 5 = 5(\ln 3)3^{5x} \).
- \( \frac{d}{dx} \sin^{10} x = \cos^{10} x \cdot \frac{d}{dx} [10^x] = \cos 10^x \cdot 10^x \ln 10 = 10^x \ln 10 \cos 10^x \).
- \( \frac{d}{dx} x^2 10^x = x^2 \cdot \frac{d}{dx} [10^x] + 10^x \cdot \frac{d}{dx} [x^2] = x^2 \cdot 10^x \ln 10 + 10^x \cdot 2x. \)
  
  \[= x \cdot 10^x \left[ x \ln 10 + 2 \right]. \]
- \( \frac{d}{dx} [2^x] = 2^x \ln 2 \cdot \frac{d}{dx} [3^x] = 2^x \ln 2 \cdot 3^x \ln 3 \cdot 3^x = 2^x 3^x \ln 2 \ln 3. \)
- \( \frac{d}{dx} [\tan^{-1} 2x] = \frac{1}{(2x)^2 + 1} \cdot \frac{d}{dx} [2^x] = \frac{1}{4x^2 + 1} \cdot 2^x \ln 2 = \frac{2^x \ln 2}{4x^2 + 1}. \)
- \( \frac{d}{dx} [2^\tan^{-1} x] = 2^\tan^{-1} x \ln 2 \cdot \frac{d}{dx} [\tan^{-1} x] = 2^\tan^{-1} x \ln 2 \cdot \frac{1}{x^2 + 1} = \frac{2^\tan^{-1} x \ln 2}{x^2 + 1}. \)
- \( \frac{d}{dx} [\ln 10^x] = \frac{d}{dx} [x \ln 10] = \ln 10, \) following from the usual power rule with a constant multiple \( \ln 10 \) carrying through the computation.

Note that an alternative, arguably superior strategy would be to first re-write the function:

\[
\frac{d}{dx} [\ln 10^x] = \frac{d}{dx} [x \ln 10] = \ln 10.
\]

While the last derivative above allowed us to first use the rules of logarithms and exponents, it should be pointed out that there are not useful rewritings for every possible case. In fact that last derivative computation was the only one above for which there was a useful way to algebraically rewrite the problem. Now it happens that most of those above would be fine
candidates for logarithmic differentiation if we wanted to avoid the formulas for derivatives of exponential functions in bases other than $e$, namely (4.129) and (4.130). Indeed only the arctangent example above, namely $\frac{d}{dx} \tan^{-1} 2x$, could not be computed directly with logarithmic differentiation. However, using our formulas for $\frac{d}{dx} a^x$ and $\frac{d}{dx} -\ln a$ will get us our results much more expeditiously.

Before continuing, it should be noted that the log and exponential functions in bases $a \in (0, 1) \cup (1, \infty)$ all have similar derivatives to those in base $e$, except for the extra factor of either $\ln a$ or $\frac{1}{\ln a}$, and so it is simply a matter of dividing or multiplying by $\ln a$:

\[
\frac{d}{dx} \ln |x| = \frac{1}{x}, \quad \frac{d}{dx} e^x = e^x, \quad \frac{d}{dx} x^a = a^x \ln a.
\]

For the case $a = e$, we have $\ln a = \ln e = 1$, giving simple cases of the more general formulas.

### 4.10.4 Derivative of $(f(x))^g(x)$

So far we have two derivative rules for functions which can loosely be defined as powers:

\[
\frac{d}{dx} [x^n] = n \cdot x^{n-1}, \quad n \in \mathbb{R} \quad \text{ (variable base, fixed exponent)},
\]

\[
\frac{d}{dx} [a^x] = a^x \ln a, \quad a \geq 0 \quad \text{ (fixed base, variable exponent)}.
\]

Crucially, in each of those cases either the base or the exponent is fixed, i.e., constant. Furthermore, we should expect very different derivative formulas for these since the functions behave very differently. For instance, if $n \in \{1, 2, 3, \cdots\}$, that function $x^n$ has polynomial growth, while if $a > 1$ the function $a^x$ has exponential growth, which is eventually much faster. There were other cases, but for the moment let us consider these. So for instance $x^2$ grows without bound, but $2^x$ grows even faster, though we will have to wait for another chapter to actually prove this fact. Consider then a function like $x^x$ in which the base grows without bound, and so does the exponent. Here these two growth conspire in such a way that this new function will grow much faster than either $x^2$ or $2^x$. This is not difficult to believe, for suppose $x = 100$.

\[
\frac{100^2}{10,000} = 0.01 \ll \frac{2^{100}}{1.26 \times 10^{30}} \ll \frac{100^{100}}{10^{1000}}.
\]

Here the notation "$\ll$" means "much less than," and as such is usually used subjectively, in much the same way that "$\approx$" is also subjective. It is used here for emphasis. The numerical results above give just a small glimpse of the relative growth rates of these three functions, $x^2$, $2^x$ and $x^x$.

Now we are interested in computing derivatives of functions in which there is a base and an exponent, but both are allowed to vary. The usual method is logarithmic differentiation, but a

---

Note that $x^x$ is only continuous for $x > 0$. That is because, while it is defined for many negative numbers, it is undefined for many more. For instance, $(-3)^{-3}$ and $(-1)^{-1}$ make sense, but $(-1/2)^{-1/2}$ and $(-\pi)^{-\pi}$ do not. Nor does $0^0$, unless we care to define it to be some number. In fact many algebra books do define it to be 1, but we will see in the next chapter that there are other choices which make equal sense, so we will decline to define $0^0$.

One way to define $x^x$ and see that it is continuous for $x > 9$ is to rewrite this function as $x^x = (e^{\ln x})^x = e^{x \ln x}$.

In this last form we see that the only thing that can “break” the continuity is for $x$ to be nonpositive. We will use this kind of technique for rewriting such a function on occasion in what follows, and indeed used it before in one computation of $\frac{d}{dx}$. 

---

77Note that $x^x$ is only continuous for $x > 0$. That is because, while it is defined for many negative numbers, it is undefined for many more. For instance, $(-3)^{-3}$ and $(-1)^{-1}$ make sense, but $(-1/2)^{-1/2}$ and $(-\pi)^{-\pi}$ do not. Nor does $0^0$, unless we care to define it to be some number. In fact many algebra books do define it to be 1, but we will see in the next chapter that there are other choices which make equal sense, so we will decline to define $0^0$.

One way to define $x^x$ and see that it is continuous for $x > 9$ is to rewrite this function as $x^x = (e^{\ln x})^x = e^{x \ln x}$.

In this last form we see that the only thing that can “break” the continuity is for $x$ to be nonpositive. We will use this kind of technique for rewriting such a function on occasion in what follows, and indeed used it before in one computation of $\frac{d}{dx}$. 

simply use the product rule.

could problem. Even for that term, while we used the formula for the derivative of an exponential function

This would be a somewhat long but fairly simple problem using the older rules, but logarithmic differentiation

it appears that the change we measure by applying $d$ to $\ln y$ is $\frac{dy}{dx}$. This would be a somewhat long but fairly simple problem using the older rules, but logarithmic differentiation

448

CHAPTER 4. THE DERIVATIVE

formula is possible, and it carries some interesting intuition. In fact types of problems are often
taught only as logarithmic differentiation problems, so the reader should be both aware of that

and able to compute these through logarithmic differentiation, but the formula we will derive is

also worth knowing.

Example 4.10.7 Find $\frac{d}{dx}[x^{\sin x}]$.  

Solution: For convenience we (equivalently) find $\frac{d}{dx}$ where $y = x^{\sin x}$.

$$y = x^{\sin x} \iff \ln y = \ln [x^{\sin x}]$$

$$\iff \ln y = \sin x \cdot \ln x$$

$$\iff \frac{dy}{dx} = \frac{d}{dx}[\sin x \cdot \ln x]$$

$$\iff \frac{1}{y} \cdot \frac{dy}{dx} = \sin x \cdot \frac{d}{dx}[\ln x] + \ln x \cdot \frac{d}{dx}[\sin x]$$

$$\iff \frac{dy}{dx} = \sin x + \ln x \cos x$$

$$\iff \frac{dy}{dx} = y \left[ \frac{\sin x}{x} + \ln x \cos x \right] = x^{\sin x} \left[ \frac{\sin x}{x} + \ln x \cos x \right].$$

It is very interesting to note here that the derivative simplifies to

$$\frac{dy}{dx} = \frac{\sin x \cdot x^{\sin x}}{x} + x^{\sin x} \ln x \cdot \cos x$$

$$= (\sin x) x^{\sin x - 1} + x^{\sin x} \ln x \cdot \frac{d}{dx}[\sin x].$$

In the sum above, the first term is what we would have if we assumed naively that the $\sin x$
exponent were a constant and we used the power rule $\frac{d}{dx}[x^u] = u \cdot x^{u-1}$ with $u = \sin x$. The

second term is what we would have if we instead assumed naively that the base $x$ were a constant,

and we used the formula for the derivative of an exponential function $\frac{d}{dx}[e^a] = e^a \ln a \cdot \frac{d}{dx}[a]$. So it

appears that the change we measure by applying $\frac{d}{dx}$ has two components, the first assuming

that the exponent is constant and measuring that part of the change we get from the base

changing, and the second assuming the base is constant and measuring the change we get from

the exponent changing. This is very much akin to our earlier interpretation of the product rule.

(See again the notes on the roles of the various factors in the product rule, page 368.) This gives us a general formula, which is slightly more complicated than one might anticipate from the

above example because of a chain rule involved in computing the component from the change in the base:

$$\frac{d}{dx} \left[ (f(x))^{g(x)} \right] = g(x) \cdot (f(x))^{g(x)-1} \cdot \frac{d}{dx}[f(x)] + (f(x))^{g(x)} \ln f(x) \cdot \frac{d}{dx}[g(x)]. \quad (4.131)$$

78One could in fact use logarithmic differentiation to derive the power rule, product rule, or quotient rule, as

we will see in the exercises. However, we do not abandon these rules since they are convenient, efficient, and

can be easily implemented when called by other rules, while logarithmic differentiation requires whole sides of

equations to be products, quotients, or powers, not for instance sums, differences, or other combinations which
can not have their logarithms expanded. Consider attempting, for instance, logarithmic differentiation on the

problem of finding $\frac{d}{dx}$ if

$$y = \sin (x^2 + 1) + \tan^{-1} x - \cos e^x + \sec x \tan x.$$
Again, the first term is computed as though the exponent were constant and the power rule employed, while the second term is computed as though the base were constant and the exponential rule used. In both cases, chain rule versions were necessary to be most general.

Two possible proofs come to mind. One is to rewrite the original function with a constant base, so that

\[(f(x))^{g(x)} = \left[e^{\ln f(x)}\right]^{g(x)} = e^{(\ln f(x))(g(x))},\]

and use the chain rule, with a product rule inside, finally rewriting the result with the original base.

More in the spirit of how these problems are usually presented is a proof using logarithmic differentiation, which is what we used in the example above. Below we consolidate some of the steps. Note that we assume \(f(x) > 0\) for continuity’s sake.

\[
y = (f(x))^{g(x)} \implies \ln y = g(x) \ln f(x)\]
\[
\implies \frac{1}{y} \cdot \frac{dy}{dx} = g(x) \cdot \frac{d}{dx} \ln f(x) + \ln f(x) \cdot \frac{dg(x)}{dx}\]
\[
\implies \frac{1}{y} \cdot \frac{dy}{dx} = g(x) \cdot \frac{1}{f(x)} \cdot \frac{d f(x)}{dx} + \ln f(x) \cdot \frac{dg(x)}{dx}\]
\[
\implies \frac{dy}{dx} = y \left[g(x) \cdot \frac{1}{f(x)} \cdot \frac{d f(x)}{dx} + \ln f(x) \cdot \frac{dg(x)}{dx}\right]\]
\[
\implies \frac{dy}{dx} = (f(x))^{g(x)} \left[g(x) \cdot \frac{1}{f(x)} \cdot \frac{d f(x)}{dx} + \ln f(x) \cdot \frac{dg(x)}{dx}\right]\]
\[
\implies \frac{dy}{dx} = g(x) \cdot (f(x))^{g(x)-1} \cdot \frac{d f(x)}{dx} + (f(x))^{g(x)} \ln f(x) \cdot \frac{d}{dx}[g(x)],\]

q.e.d. While the formula above is rather long, it is easy enough to memorize if it is remembered in spirit: again, the first term treats \(g(x)\) as a constant, and the second term \(f(x)\) as a constant. Still, it is predictable that many students would feel more comfortable either using logarithmic differentiation, or the change of base (to \(e\)) strategy.

It is interesting to note that this new derivative formula in fact generalizes the power and exponential rules.

1. If \(g(x)\) is constant, so \(\frac{d}{dx}[g(x)] = 0\), then we get the power rule, in its chain rule version.

2. If \(f(x)\) is constant, so \(\frac{d}{dx}[f(x)] = 0\), then we get the exponential rule, in its chain rule version.

But for now we need to consider more examples.

**Example 4.10.8** Find \(\frac{d}{dx}[(\ln x)^x]\).

**Solution:** While we can use logarithmic differentiation here, we will use our general formula, first treating the exponent \(x\) as constant, and then the base \(\ln x\) as constant.

\[
\frac{d}{dx}[(\ln x)^x] = x(\ln x)^{x-1} \frac{d}{dx}[\ln x] + (\ln x)^x \ln(\ln x) \cdot \frac{d}{dx}[\ln x]\]
\[
= x(\ln x)^{x-1} \cdot \frac{1}{x} + (\ln x)^x \ln(\ln x) \cdot 1\]
\[
= (\ln x)^{x-1} + (\ln x)^x \ln(\ln x).^{79}\]
Example 4.10.9 Compute $\frac{d}{dx}[xe^x]$.

Solution:

$$\frac{d}{dx}[xe^x] = e^x [xe^x - 1] + xe^x \ln x \cdot e^x$$

Admittedly it is rare to find the formula (4.131) for $\frac{d}{dx}(f(x)g(x))$ in a standard calculus textbook. One could argue that the formula is sufficiently complicated that it is better to always use the logarithmic differentiation technique, since it is more flexible and, after all, has the means to prove the formula. But (4.131) is included here because it contains some nice intuition and gives a shorter route the desired derivatives. Indeed, while these types of derivatives are relatively rare in most practices, for those with the need to compute them routinely the formula could have good utility.

\footnote{In most mathematical literature, the short-hand for $\ln(\ln x)$ is simply $\ln \ln x$. It is not uncommon to see such things as $\ln \ln \ln x \ln \ln x$, meaning $[\ln(\ln(\ln x))][\ln(\ln x)]$, for instance.}
4.10. THE NATURAL LOGARITHM II: FURTHER RESULTS

Exercises

For Exercises 1–8, compute the given derivative.

1. \( \frac{d}{dx} [4\sin x] \)
2. \( \frac{d}{dx} [\sin 4^x] \)
3. \( \frac{d}{dx} \log_2 |\tan x| \)
4. \( \frac{d}{dx} \tan (\log_2 x) \)
5. \( \frac{d}{dx} \log_{4^3} x \)
6. \( \frac{d}{dx} \log_{3^x} x \)
7. \( \frac{d}{dx} [\sin^2 x \cos^4 x] \)
8. \( \frac{d}{dx} [3^{2x}] \)

For Exercises 9–14, use logarithmic differentiation to prove the given formula. For example, to show \( \frac{dx^n}{dx} = nx^{n-1} \), first set \( y = x^n \) and then use logarithmic differentiation to find \( \frac{dy}{dx} \).

9. \( \frac{d}{dx} [a^x] = a^x \ln a \), assuming \( a > 0 \).
10. \( \frac{d}{dx} [uv] = u \cdot \frac{dv}{dx} + v \cdot \frac{du}{dx} \).
11. \( \frac{d}{dx} \left[ \frac{u}{v} \right] = \frac{v \cdot \frac{du}{dx} - u \cdot \frac{dv}{dx}}{v^2} \).
12. \( \frac{d}{dx} [x^n] = n \cdot x^{n-1} \).
13. \( \frac{d}{dx} \sec x = \sec x \tan x \), using the derivative formula for \( \cos x \) and the fact that \( \sec x = 1/\cos x \).
14. \( \frac{d}{dx} e^x = e^x \)

For Exercises 15–20, use logarithmic differentiation to compute the following.

15. \( \frac{d}{dx} \sqrt{\frac{1 + x}{1 - x}} \)
16. \( \frac{d}{dx} [(10x^2 + 9)] \)
17. \( \frac{d}{dx} \left[ \left( x^2 + 4 \right)^3 \left( x^3 + 5 \right)^6 \right] \)
18. \( \frac{d}{dx} \left[ \frac{x^5}{(x^4 + 8x - 3)^3} \right] \)
19. \( \frac{d}{dx} [e^{2x} \sin 5x \cos 9x] \)
20. \( \frac{d}{dx} \left[ \frac{\sin^5 x \cdot \sqrt{x^2 - 2}}{2(6x - 7)^4(2x + 5)^3} \right] \)

For Exercises 21–24, compute \( \frac{dy}{dx} \) for each. You may use either logarithmic differentiation, or (4.131), page 448. (It would be useful to use both and examine how the results are in fact the same.)

21. \( y = x^x \)
22. \( y = x^{1/x} \)
23. \( y = x^{x^x} = (x)(x^x) \)
24. \( y = \left( \frac{\sin x}{x} \right)^x \)
4.11 Further Interpretations of the Derivative

If one asks a calculus student the meaning of the derivative, the most common answer will likely be “the slope of the tangent line,” or perhaps “the slope of the curve,” which expresses the geometric significance of the derivative in the context of the graph of a function such as \( y = f(x) \). On the other hand, if one asks an engineer, scientist or economist, one is more likely to hear “the rate of change.” There are many quantities which are discussed in even casual conversation which are rates of change, even if they are not called that explicitly. Included in those are speed (or velocity) and acceleration, population (or any other) growth or decay, power, current, miles per gallon, and for that matter just about any concept which includes the word “per.” Recall that a rate of change has meaning only if it includes the answer to “with respect to what?” We already discussed how velocity \( v \) is a rate of change of position \( s \) with respect to time \( t \), so \( v = ds/dt \). This is an instantaneous rate of change of position with respect to time, and is thus the result of computing the limit of \((\Delta s)/(\Delta t)\) as \( \Delta t \to 0 \).

In general, a derivative \( dQ/dx \) can always be viewed as this instantaneous rate of change of \( Q \) with respect to \( x \), assuming \( Q \) is ultimately a function of \( x \). Now \( dQ/dx \) is again a limit of difference quotients \((\Delta Q)/(\Delta x)\) (see the discussion beginning with (4.4), page 296), so it carries with it the units of \( Q \) divided by the units of \( x \) and the intuition of being a limit of difference quotients.

In this short section we consider several kinds of such rates of change, in order to construct an intuition connecting these rates of change with the relevant derivatives.

4.11.1 Velocity and Acceleration

Recall that if \( s(t) \) is the position along a numbered axis at time \( t \), then velocity is given by

\[
v(t) = \lim_{\Delta t \to 0} \frac{s(t + \Delta t) - s(t)}{\Delta t} = \frac{ds(t)}{dt},
\]

or \( dv/dt \) for short, and by the nature of the difference quotient this will carry units of length/time: if \( s \) is in meters (m) and \( t \) is in seconds (sec), then \( dv/dt \) will be in m/sec.

Furthermore, with acceleration meaning a change in velocity, we have

\[
a(t) = \lim_{\Delta t \to 0} \frac{v(t + \Delta t) - v(t)}{\Delta t} = \frac{dv(t)}{dt},
\]

or simply \( dv/dt \), which will have units of \( \text{m/sec}/\text{sec} \), or \( \text{m/sec}^2 \). While those units may see a bit odd, mathematically they are correct and can perhaps make more sense when stated as “meters per second per second.”

4.11.2 Power

Many science textbooks define energy as the ability to do work, and work as a force (push or pull) acting on an object through a distance travelled. Numerically, the energy required to do 2 units of work will be the same 2 units of energy, though there is usually not a 100% transfer of energy required to do that work.

4.11.3 Fluid Flow and Volume

4.11.4 Population Growth and Decay

This is one of many applications of calculus which often uses a continuum model for a phenomenon which is actually discrete. While there are differing opinions on what constitutes a
4.11. FURTHER INTERPRETATIONS OF THE DERIVATIVE

discrete variable, it can be contrasted to a continuous one, which can, for instance, take on any values within an interval of positive length. For instance, length, (time) duration, and temperature are such that if they can take on values $a$ and $b$, then they can take on any value within $[a, b]$. Discrete variables, by contrast, can not. They will have a finite number of values they can take on, or perhaps they are allowed to take on only integer values, or some subset thereof. Examples include the possible answers on a true-or-false question, hotel ratings (zero to five “stars”), a government identity number, and so on.

Another example is the number of individuals in a population. If one has all the data available to analyze an actual population, there is little need for calculus, except perhaps for predictive analysis. However, if one wished to model a hypothetical population, the calculus can be useful.

**Example 4.11.1** Suppose a population of 100 individuals is expected to double every 25 years. What is the growth rate, in persons per year, at the start, ten years later, 25 years later and 50 years later?

**Solution:** A reasonable model for the size of the population at year $t$ after the 100 individuals form the population would be

$$P = 100 \cdot 2^{t/25}.$$ 

Note how $P(0) = 100$, $P(25) = 200$, $P(50) = 400$, and so on. This is a continuous approximation of a discrete problem. We can then compute

$$P'(t) = \frac{dP}{dt} = 100 \cdot 2^{t/25} \ln 2 \cdot \frac{1}{25} \left[ \frac{1}{25} t \right]$$

$$= 4 \ln 2 \cdot 2^{t/25}.$$

From this we can compute

$$P'(0) = 4 \ln 2 \approx 2.772$$

$$P'(10) = 4 \ln 2 \cdot 2^{10/25} \approx 3.658$$

$$P'(25) = 4 \ln 2 \cdot 2 \approx 5.545$$

$$P'(50) = 4 \ln 2 \cdot 2^{2} \approx 11.090$$

Note that $P'(50) = \left. \frac{dP}{dt} \right|_{t=50}$, and so it will carry units of persons/year.

The above example is theoretically correct but conceptually awkward. First, it is inconceivable to produce 11.090 persons in a year. Second, this is actually an instantaneous rate of change, so it seems more intuitive in some ways to instead discuss the actual number of persons added during the year, so instead of $P'(50)$ it might be more interesting to compute

$$P(50) - P(49) = 100 \cdot 2^2 - 100 \cdot 2^{49/25} \approx 10.938,$$

though that is still not quite consistent with reality because persons only come in whole numbers. At this point one might remember that this is likely to be an approximation already, and so a researcher might instead report from either observation that there will be an annual rate of approximately 11 persons/year at time $t = 50$ years. However, note that the computation which uses the calculus is a shorter computation than the one which computes the actual value $(\Delta P)/(\Delta t)$.

---

80The interval to use to compute $(\Delta P)/(\Delta t)$ for the above example is someone open to opinion. We used $(P(50) - P(49))/1$ but it is reasonable to opt for $(P(51) - P(50))/1$, or perhaps $(P(50.5) - P(49.5))/1$, this last one centering at $t = 50$. Either way the calculus computation $P'(50)$ is a reasonable approximation.
A more reasonable continuous approximation with what is technically a discrete variable is radioactive decay. It is impractical to count the actual number of radioactive atoms in a visible sample, so once we choose a scale of unit where the actual numbers we deal with are not large, for practical purposes we can assume the possible values of the number of atoms forms a continuum.

4.11.5 Consumption

4.11.6 Marginal Cost and Profit
In this chapter we develop methods for analyzing functions extensively by exploiting the information contained in their derivatives. We also investigate higher-order derivatives, where the second derivative is the derivative of the derivative function, the third derivative is the derivative of the second derivative, and so on. We will pay particular attention to the theorems and intuitions involving a function’s first and second derivatives.

Of course a very natural and general approach for analyzing functions is to consider their graphs in the Cartesian Plane, and much of this chapter is devoted to describing their graphs by analyzing their derivatives. While the plane is an abstract setting, the analysis will have useful and intuitive import to applied problems as well.

Fortunately we will be able to prove all of the results presented in this chapter, referring only to previously proved results and already mentioned facts we borrow from more advanced studies. However, we have already used some of the intuition without proof, for example when we noted that it seemed reasonable to believe

\[ f'(x) > 0 \text{ on } (a,b) \implies f(x) \text{ increasing on } (a,b), \]
\[ f'(x) < 0 \text{ on } (a,b) \implies f(x) \text{ decreasing on } (a,b) \]

and stated this as Theorem 4.2.5, page 313. In this chapter we will finally prove this theorem, using the so-called Mean Value Theorem (MVT). In fact we will have a slightly modified version in which \((a,b)\) is replaced by a closed interval \([a,b]\).

The Mean Value Theorem may at first seem to be a mere curiosity. In fact it is astonishingly useful for proving many first derivative implications which are perhaps more intuitive but otherwise surprisingly difficult to prove. Furthermore, the MVT lends itself to many practical problems, particularly those involving average rates of change and inequalities.

The graphical significance of a function’s first derivative as slope was discussed in the previous chapter, and should seem straightforward. The significance of the second derivative is nearly as straightforward, though it is often necessary to break that analysis into more cases which depend upon the value (or sign) of the first derivative. Yet higher-order derivatives’ effects on graphs are much less easily intuited, as their graphical significances rely upon the values (or signs) of all previous derivatives, so we will not pursue these very far in this chapter. By contrast, the second derivative is somewhat intuitive, and is important enough for its own analysis, albeit in the context of being the derivative of the more easily understood first derivative.
5.1 Higher-Order Derivatives and Graphing

In this section, we introduce notation for higher-order derivatives, develop some intuition regarding the so-called second derivative of a function, and use that to aid in graphing functions. We also use the opportunity to combine several analytical aspects of a function in order to sketch graphs which illustrate more and more aspects of the function’s behavior.

### 5.1.1 Higher-Order Derivatives (Derivatives of Derivatives)

From physics we learn that it is useful to take the derivative of velocity to find acceleration. Velocity being a derivative itself—of position—acceleration is thus the derivative of the derivative of position. There are a few ways of writing this:

\[ a(t) = v'(t) = (s'(t))' = s''(t), \]

\[ a = \frac{dv}{dt} = \frac{d}{dt} \left( \frac{ds}{dt} \right) = \frac{d^2s}{(dt)^2} = \frac{d^2s}{dt^2}. \]

The final form shown above is a bit of an abuse of the notation, in that our previous understanding of \( dt^2 = 2t \frac{dt}{dt} \) is not the same as \( (dt)^2 \), but as often happens it has become acceptable to shorten the notation in this context.\(^1\) Note that \( s''(t) \) is usually read, “s double-prime of t.”

If we wished for a third derivative of a function \( f(x) \), we can write it as \((f'''(x)) = f'''(x)\), or as \( \frac{d}{dx} \left[ \frac{d^2f(x)}{dx^2} \right] = \frac{d^3f(x)}{dx^3} \). Beyond the third derivative, we have further notations, chosen for convenience. For instance, the fourth derivative can be written with lower-case Roman numeral-type notation as \( f^vi(x) \), or using a regular numeral but in parentheses, as in \( f^{(iv)}(x) \), to distinguish the fourth-order derivative of \( f(x) \) from a simple power such as \((f(x))^4\). For another example,

\[ f^{vi}(x) = f^{(10)}(x) = \frac{d^6f(x)}{dx^6}. \]

To be clear, the last expression above is short for \( \frac{d}{dx} \left[ \frac{d}{dx} \left[ \frac{d}{dx} \left[ \frac{d}{dx} \left[ \frac{d}{dx} \left[ \frac{d}{dx} \left[ \frac{d}{dx} \left[ f(x) \right] \right] \right] \right] \right] \right] \right] = \frac{d^6f(x)}{(dx)^6} \).

For polynomials, each time we differentiate we get a lower-degree polynomial; for exponential functions \( e^x \) and the like, and the simplest trigonometric functions \( \sin x \) and \( \cos x \), each differentiation yields a function roughly as complicated as the original function; for functions which require the chain rule, product rule or quotient rule, each differentiation step can yield a much more complicated function than before that step. A simple case of this phenomenon follows.

#### Example 5.1.1 Find the first two derivatives of \( f(x) = \sin x^2 \).

**Solution:**

\[ f(x) = \sin x^2 \implies f'(x) = \cos x^2 \cdot 2x = 2x \cos x^2 \]

\[ \implies f''(x) = 2x \frac{d^2 \cos x^2}{dx} + \cos x^2 \cdot \frac{d^2 x}{dx} \]

\[ = 2x(-\sin x^2)(2x) + 2 \cos x^2 \]

\[ = -4x^2 \sin x^2 + 2 \cos x^2. \]

---

\(^1\)It should be pointed out that one of the Leibniz notation’s strengths for first derivatives—the way its fractional form allows chain rule derivatives to be consistent with multiplicative rules (particularly cancellation) of arithmetic—does not extend to higher-order derivatives. It is usually not a problem but the reader should be aware of this limitation. So for instance \( \frac{d^2 \sin u}{dx^2} \neq \frac{d^2 \sin u}{dx^2} \cdot \frac{d^2 x}{dx} \). Instead one would have to write

\[ \frac{d^2 \sin u}{dx^2} = \frac{d}{dx} \left[ \frac{d \sin u}{dx} \right] = \frac{d}{dx} \left[ \cos u \cdot \frac{du}{dx} \right] = \cos u \cdot \frac{d^2 u}{dx^2} + \left( \frac{du}{dx} \right) \left( -\sin u \cdot \frac{du}{dx} \right) = \cos u \cdot \frac{d^2 u}{dx^2} - \sin u \left( \frac{du}{dx} \right)^2, \]

and so a product rule is necessary to complete the computation.
5.1. HIGHER-ORDER DERIVATIVES AND GRAPHING

Still higher-order derivatives of \( f(x) = \sin x^2 \) would be even more complicated. However for polynomials, when written in expanded (not factored) form it is clear that each differentiation step lowers the degree by one, until all subsequent derivatives are zero:

**Example 5.1.2** Compute all orders of derivatives for \( f(x) = x^5 + 3x^4 - 8x^3 + 9x - 12 \).

**Solution:** We take these in turn.

\[
\begin{align*}
f(x) &= x^5 + 3x^4 - 8x^3 + 9x - 12 \\
\Rightarrow f'(x) &= 5x^4 + 12x^3 - 24x + 9 \\
\Rightarrow f''(x) &= 20x^3 + 36x^2 - 24 \\
\Rightarrow f'''(x) &= 60x^2 + 72x \\
\Rightarrow f^{(4)}(x) &= 120x + 72 \\
\Rightarrow f^{(5)}(x) &= 120 \\
\Rightarrow f^{(n)}(x) &= 0 \text{ for all } n \geq 6.
\end{align*}
\]

Clearly, for an \( n \)th degree polynomial \( f(x) \), we will have \( m > n \Rightarrow f^{(m)}(x) = 0 \), which is a useful observation in later chapters.

We will have further opportunities to compute higher-order derivatives throughout the text. We will pay particularly close attention to \( f''(x) \) and its significance in this section.

5.1.2 Graphical Significance of the Second Derivative

For an open interval \((a,b)\), and a function \( f(x) \) defined for all \( x \in (a,b) \), we have\(^2\)

\[
\begin{align*}
f' > 0 \text{ on } (a,b) &\implies f \text{ is strictly increasing on } (a,b), \\
f' < 0 \text{ on } (a,b) &\implies f \text{ is strictly decreasing on } (a,b).
\end{align*}
\]

Similarly, if the derivative of \( f' \) is positive on \((a,b)\) then \( f' \) must be strictly increasing on \((a,b)\), while if the derivative of \( f' \) is negative on \((a,b)\) then \( f' \) must be strictly decreasing on \((a,b)\). More formally,

\[
\begin{align*}
f'' > 0 \text{ on } (a,b) &\implies f' \text{ is strictly increasing on } (a,b), \\
f'' < 0 \text{ on } (a,b) &\implies f' \text{ is strictly decreasing on } (a,b).
\end{align*}
\]

What this means graphically is illustrated in Figure 5.1, page 458. The shape of the curve is called **concave up** (or “u-shaped”) when \( f'' > 0 \), and **concave down** (or “n-shaped”) when \( f'' < 0 \).

When \( f', f'' \) exist and are nonzero on an interval, there are four possible combinations of signs for \( f' \) and \( f'' \), as illustrated below:

\[
\begin{array}{cccc}
\text{\( f' > 0 \)} & \text{\( f' > 0 \)} & \text{\( f' < 0 \)} & \text{\( f' < 0 \)} \\
\text{\( f'' > 0 \)} & \text{\( f'' < 0 \)} & \text{\( f'' > 0 \)} & \text{\( f'' < 0 \)}
\end{array}
\]

\(^2\)To be more precise, by \( f' > 0 \text{ on } (a,b) \implies f \text{ strictly increasing on } (a,b) \) we mean that

\[
(\forall x \in (a,b))[f'(x) > 0] \implies (\forall x_1, x_2 \in (a,b))[x_1 < x_2 \implies f(x_1) < f(x_2)].
\]

A similar implication is valid in the case we might abbreviate as \( f' < 0 \implies f \text{ strictly decreasing}. \)
Theorem 5.1.2 (Second Derivative Test)
Suppose the second derivative, we can now also consider the following (see again Figure 5.1):

\[ f'' > 0 \implies f' \text{ strictly increasing} \]
\[ f'' < 0 \implies f' \text{ strictly decreasing} \]

curve is “concave up”

Figure 5.1: Illustrations of the effects of the sign of \( f'' \) on the shape of the graph of a function.

Figure 5.1 also illustrates how the signs of the first two derivatives of a function can often be used to find local maximum and minimum points. The tests are intuitive and easily remembered in substance, but since they are also well known by name we will mention them here.

**Theorem 5.1.1 (First Derivative Test):** If \( f(x) \) is continuous on \((a,b)\), and \( x_0 \in (a,b) \) is such that

- \( f' < 0 \) on \((a,x_0)\) and \( f' > 0 \) on \((x_0,b)\), then \((x_0, f(x_0))\) is a local minimum;
- \( f' > 0 \) on \((a,x_0)\) and \( f' < 0 \) on \((x_0,b)\), then \((x_0, f(x_0))\) is a local maximum.

We have used this principle for graphs in the previous chapter. With our consideration of the second derivative, we can now also consider the following (see again Figure 5.1):

**Theorem 5.1.2 (Second Derivative Test)** Suppose \( f'(x_0) = 0 \). Then the following hold.

- If \( f''(x_0) > 0 \), then \((x_0, f(x_0))\) is a local minimum.
- If \( f''(x_0) < 0 \), then \((x_0, f(x_0))\) is a local maximum.

Note that if \( f''(x_0) = 0 \), the theorem is silent. With both \( f'(x_0) = 0 \) and \( f''(x_0) = 0 \), a graph tends to be more “flat” at \( x_0 \) than if we have only \( f'(x_0) = 0 \). Consider the following examples where this occurs.

1. Consider \( f(x) = x^4 \implies f'(x) = 4x^3 \implies f''(x) = 12x^2 \). Then \( f'(0) = 0 \), and \( f''(0) = 0 \).

   Since \( f(x) > 0 \) for \( x \neq 0 \), we have that \((0, f(0)) = (0, 0)\) is a local (actually global) minimum (also following from \( f' < 0 \) on \((−\infty, 0)\), but \( f' > 0 \) on \((0, \infty)\)).

2. On the other hand, for the case \( f(x) = -x^4 \), we have \( f'(x) = -4x^3 \), \( f''(x) = -12x^2 \), so \( f'(0) = 0 \) and \( f''(0) = 0 \), but \( f(x) < 0 \) for all \( x \neq 0 \), so \((0, f(0)) = (0, 0)\) is a local (again, actually global) maximum point.

3. \( f(x) = x^3 \implies f'(x) = 3x^2 \implies f''(x) = 6x \). In this case \( f'(0) = 0 \) and \( f''(0) = 0 \), but that point \((0, f(0)) = (0, 0)\) is neither a local maximum nor a local minimum, since \( x < 0 \implies f(x) < 0 \), while \( x > 0 \implies f(x) > 0 \). “From the left,” \((0,0)\) appears to be a maximum point, where from the right \((0,0)\) appears to be a minimum point, so it is neither. (See the graph of \( y = x^3 \) on page 460.)
It is interesting to look at points on the graph of a function \( f(x) \) where concavity changes, i.e., the curve’s bend (or “concavity”) changes from “upward” to “downward” or vice versa. Such a point is called an inflection point, assuming it does actually lie on the graph:

**Definition 5.1.1** If a point \((x_0, f(x_0))\) on the graph of \( y = f(x) \) is such that

1. \( f(x) \) is continuous at \( x_0 \), and
2. \( f''(x) \) has one sign (positive or negative) in some interval \((x_0 - \delta, x_0)\) and the opposite sign in \((x_0, x_0 + \delta)\), for some \( \delta > 0 \),

then \((x_0, f(x_0))\) is called an **inflection point** of the graph of \( y = f(x) \).

Some of our common functions illustrate these phenomena relating to the sign of the second derivative. For instance, for \( f(x) = \sin x \), we have \( f'(x) = \cos x \) and \( f''(x) = -\sin x \). Thus for this case \( f''(x) = -f(x) \), so where \( f(x) \) is positive, \( f''(x) \) is negative and vice versa. As \( \sin x \) is continuous for all \( x \in \mathbb{R} \), the points where \( f''(x) \) changes signs are inflection points, namely at \( x = \pm \pi, \pm 2\pi, \pm 3\pi, \ldots \), as seen below:

![Graph of \( \sin x \)](image)

By contrast, the function \( f(x) = \ln x \) is concave down wherever defined, since \( f(x) = \ln x \implies f'(x) = \frac{1}{x} \implies f''(x) = -x^{-2} < 0 \) where \( \ln x \) is defined. For that matter, this is also true of \( f(x) = \ln |x| \), as graphed below:

![Graph of \( \ln |x| \)](image)

Even the functions \( x^2 \) and \( x^3 \) below are also interesting for their first and second derivative properties. Note that \( y = x^2 \) is always concave up, while \( y = x^3 \) is always increasing, but has an inflection point at \((0, 0)\), where it changes from concave down to concave up.
Functions which are more complicated can require a more systematic search for features such as intervals on which they increase/decrease, where they are concave up/down, and other interesting behavior such as x-intercepts or y-intercepts, and asymptotic behaviors. Where possible, it is useful to have sign charts for

- $f(x)$, to find where the graph is above/below the x-axis,

- $f'(x)$, to find where $f(x)$ is increasing/decreasing, and

- $f''(x)$, to find where $f(x)$ is concave up/down.

If some of these features are extraordinarily difficult to derive, one can sometimes sketch the graph of $y = f(x)$ illustrating the features that were more easily found and leaving the remaining features unaddressed (or having the sketch be vague regarding those). The order in which one derives the various features is not necessarily important either, as one is often reminded to consider previously overlooked features as the graph is drawn. For instance, the y-intercept, i.e., the point $(0, f(0))$ where applicable, is often noted as the graph is drawn, because it is fairly trivial to find in practice and a distraction to note too early in the process of analyzing a function’s graph’s characteristics. Nonetheless we will be somewhat systematic in analyzing functions’ graphs as we proceed, at least with the derivative computations and their ramifications.

**Example 5.1.3** Graph $f(x) = 2x^3 - 3x^2$, illustrating the signs of $f(x)$, $f'(x)$, $f''(x)$, all x- and y-intercepts, and the behavior as $x \to \pm\infty$.

**Solution:** We begin with the intercepts, noting that $f(x) = x^2(2x - 3) = 0$ at $x = 0, 3/2$. While we could construct a sign chart for $f(x)$, in fact once we note the locations of these intercepts, the addition of the sign chart for $f'(x)$ will be enough to determine the signs of $f(x)$ elsewhere. For example, if $f(x_0) = 0$ but $f' > 0$ in an interval containing $x_0$, then within that interval we must have $f < 0$ to the left of $x_0$, and $f > 0$ to the right of $x_0$.

Now $f'(x) = \frac{d}{dx} [2x^3 - 3x^2] = 6x^2 - 6x = 6x(x - 1) = 0$ at $x = 0, 1$. From this we construct a sign chart for $f'(x)$ (test points omitted below):

![Sign Chart](image-url)
5.1. HIGHER-ORDER DERIVATIVES AND GRAPHING

\[ f'(x) = 6x(x-1) \]

Factors: \( \bigoplus \) \( \bigoplus \) \( \bigoplus \)

Sign \( f' \): \( \bigoplus \) 0 \( \bigoplus \) 1 \( \bigoplus \)

Graph behavior: / \( \bigcirc \) \( \bigcirc \) \( \bigcirc \)

From the sign chart, clearly we have \((0, f(0)) = (0, 0)\) is a local maximum of \(f(x)\), while \((1, f(1)) = (1, -1)\) is a local minimum. There are no further local extrema.

(As an aside, recalling \(f(x) = 0 \iff x \in \{0, 3/2\}\), and since \(f(x)\) is continuous for all \(x \in \mathbb{R}\), from the arrows above we can see that \(f(x) < 0\) on \((-\infty, 0) \cup (0, 3/2)\), and \(f(x) > 0\) on \((1, \infty)\). This will also become apparent—and indeed more obvious—when the graph is actually drawn to reflect the signs of \(f', f''\) and the \(x\)-intercepts.)

Next we consider the sign of \(f''(x)\), by noting that \(f''(x) = \frac{d^2}{dx^2}(f'(x)) = \frac{d}{dx} [6x^2 - 6x] = 12x - 6 = 0\) for \(x = 1/2\), and is otherwise continuous, so we can construct the following sign chart for \(f''(x)\):

\[ f''(x) = 12x - 6 \]

Sign \(f''\): \( \bigoplus \) \( \bigoplus \)

Concavity of graph: Down \(1/2\) Up

From the sign chart, we have an inflection point at \((1/2, f(1/2)) = (1/2, -3/4)\).

It is sometimes useful to construct a combined sign chart—or function behavior chart—using all of the points used on the sign charts for \(f'(x)\) and \(f''(x)\). (It is often easier to defer the inclusion of signs of \(f(x)\) until we are ready to construct the final graph.) There are various ways to organize this, such as the chart below:

\[
\begin{array}{cccccc}
& f' > 0 & f' < 0 & f' < 0 & f' > 0 & \\
& f'' < 0 & f'' < 0 & f'' > 0 & f'' > 0 & \\
\end{array}
\]

Shape of graph:

\[ \begin{array}{c}
\bigcirc \ 1/2 \ 1 \end{array} \]

From the shapes given above, and the fact that \(f(x)\) is a polynomial and therefore defined and continuous on all of \(\mathbb{R}\), we can deduce its behavior \(x \to \pm \infty\), so it is not so necessary to compute those limits, but we do so here to verify these facts (note dominance of the \(x^3\) terms):

\[
\lim_{x \to \infty} f(x) = \lim_{x \to \infty} (2x^3 - 3x^2) = \lim_{x \to \infty} \left[ x^3 \left( 2 - \frac{3}{x} \right) \right] = \infty - \infty,
\]

\[
\lim_{x \to -\infty} f(x) = \lim_{x \to -\infty} (2x^3 - 3x^2) = \lim_{x \to -\infty} \left[ x^3 \left( 2 - \frac{3}{x} \right) \right] = -\infty - \infty.
\]

Again, these limits are indicated by the sign chart above as well. Putting all of these facts together we get the actual graph, plotted below, with the local maximum \((0, 0)\), local minimum \((1, -1)\) and inflection point \((1/2, -1/2)\) all highlighted. Also recall the \(x\)-intercepts at \(x = 0, 3/2\).
The graph above is, of course, computer-generated but a reasonable “sketch” can be made to illustrate those features relevant to the calculus: where the graph is increasing or decreasing, and where it is concave up or down. One often makes a point to graph the \( y \)-intercept—where the graph intersects the \( y \)-axis (i.e., where \( x = 0 \)—as well, but in this case that point is also an \( x \)-intercept.

Recall that a function can change signs as \( x \) increases by either (1) passing through the value 0 in its output, or (2) having \( x \) pass through a value where \( f(x) \) is discontinuous. It is similar with derivatives, though we are interested in where the derivative is (1) passing through the value 0, or (2) is discontinuous, which for our cases will be those places where \( f'(x) \) does not exist (as a real number). These are the points where it can change signs. (See the graph of \( y = \ln |x| \) on page 459, and consider what occurs there at \( x = 0 \).)

**Example 5.1.4** Sketch the graph of \( f(x) = x - 3x^{2/3} \), indicating where the graph is increasing/decreasing, concave up/down, and any asymptotic behavior.

**Solution:** Rather than making a sign chart for \( f(x) \), we will instead note all intercepts and let other aspects imply where \( f(x) > 0 \) and where \( f(x) < 0 \).

\[
\frac{x - 3x^{2/3}}{f(x)} = 0 \iff x^{2/3}(x^{1/3} - 3) = 0 \iff (x^{2/3} = 0) \lor (x^{1/3} = 3) \iff x \in \{0, 27\}.
\]

Thus \( x \)-intercepts are \((0, 0)\) and \((27, 0)\), the former also being the \( y \)-intercept. Next we create a sign chart for \( f'(x) \).

\[
f'(x) = \frac{d}{dx} \left( x - 3x^{2/3} \right) = 1 - 2x^{-1/3} = \frac{x^{1/3} - 2}{x^{1/3}}.
\]

Thus \( f'(x) = 0 \iff x^{1/3} = 2 \iff x = 8 \), while \( f'(x) \) does not exist \( \iff x = 0 \). We need both points for our sign chart. (Note \( f(x) \) is continuous on all of \( \mathbb{R} \).)

<table>
<thead>
<tr>
<th>factors:</th>
<th>( \Theta/\Theta )</th>
<th>( \Theta/\Theta )</th>
<th>( \Theta/\Theta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>sign ( f' ):</td>
<td>( \oplus )</td>
<td>0</td>
<td>( \oplus )</td>
</tr>
<tr>
<td>graph behavior:</td>
<td>/</td>
<td>( \downarrow )</td>
<td>/</td>
</tr>
</tbody>
</table>

Since \( f(x) \) is continuous (so, for instance it has no vertical asymptotes or jump discontinuities), we must have for \( f(x) = x - 3x^{2/3} \):
• a local maximum at \((0, f(0)) = (0, 0)\);
• a local minimum at \((8, f(8)) = (8, 8 - 3 \cdot 8^{2/3}) = (8, 8 - 3 \cdot 4) = (8, -4)\).

We might wish to keep in mind when we graph this that, while \(f'(8) = 0\) so the graph’s slope is horizontal there, \(f'(0)\) does not exist so the graph is not smooth at \(x = 0\). In fact, the graph “approaches vertical” as \(x \to 0\), as we can observe from the following:

\[
\lim_{x \to 0^-} f'(x) = \lim_{x \to 0^-} \frac{x^{1/3} - 2}{x^{1/3}} = -2/0^- \to \infty, \\
\lim_{x \to 0^+} f'(x) = \lim_{x \to 0^+} \frac{x^{1/3} - 2}{x^{1/3}} = -2/0^+ \to -\infty.
\]

Next we consider concavity by computing \(f''(x)\) and constructing its sign chart.

\[
f''(x) = \frac{d}{dx} \left(1 - 2x^{-1/3}\right) = \frac{2}{3}x^{-4/3} = \frac{2}{3(\sqrt[3]{x})^4}.
\]

This does not exist at \(x = 0\) (which we actually should have known because \(f'(0)\) does not exist), but \(x \neq 0 \implies f''(x) > 0\), so this curve is concave up on any interval which does not contain \(x = 0\). We can make a combined sign chart for \(f'\) and \(f''\) as before (though with practice this is not necessary except for more complicated cases):

\[
\begin{array}{ccc}
 f' > 0 & f' < 0 & f' > 0 \\
 f'' > 0 & f'' > 0 & f'' > 0 \\
\end{array}
\]

shape of graph:

Recalling again that the graph is of a continuous function, we can finally sketch it, though we might also wish to include the other \(x\)-intercept at \(x = 27\), which makes for a somewhat compacted scale for a reasonable graph. Note the minimum point \((8, -4)\), and the \(x\)-intercepts at \(x = 0, 27\) are highlighted.
This function has no vertical or horizontal asymptotes, but we can see by the concavity and slopes that \( x \to \infty \implies f(x) \to \infty \), while \( x \to -\infty \implies f(x) \to -\infty \). We can also use limits to compute those behaviors:

\[
\lim_{x \to \infty} f(x) = \lim_{x \to \infty} \left( \frac{x^{2/3}}{3} \left( x^{1/3} - 3 \right) \right) = \infty, \\
\lim_{x \to -\infty} f(x) = \lim_{x \to -\infty} \left( \frac{x^{2/3}}{3} \left( x^{1/3} - 3 \right) \right) = -\infty.
\]

These are reflected in the graph as well.

The graph above does illustrate that we have to consider points where \( f' \) does not exist when we construct a sign chart for \( f'' \). There was a change of signs at \( x = 0 \) where \( f' \) does not exist (and another at \( x = 8 \), where \( f' = 0 \)) and indeed there was a local maximum at \( x = 0 \) (and minimum at \( x = 8 \)). Similarly we have to check for concavity changes where \( f'' = 0 \), and also where \( f'' \) does not exist. Now visually the concavity for \( x < 0 \) is very subtle, and barely visible on the graph above, which connected 2000 computer-generated points. The graph is more obviously concave up for that part of it drawn above where \( x > 0 \). (Interestingly, though it is concave up, it must never intersect the straight line \( y = x \) for \( x \neq 0 \), since the height difference at each \( x \neq 0 \) is \( 3x^{2/3} > 0 \), which is in fact getting larger as \( x \to \pm \infty \).)

**Example 5.1.5** Graph \( f(x) = e^{-x^2} \), illustrating signs of \( f' \) and \( f'' \), as well as asymptotic behavior.

**Solution:** At some point we might wish to note \( f(x) > 0 \) for all \( x \in \mathbb{R} \), due to the output range of the exponential function, and that

\[
\lim_{x \to \pm \infty} f(x) = 0.
\]

Thus the function has a two-sided horizontal asymptote \( y = 0 \), as \( x \to \pm \infty \) (or, more precisely, \( x \to \pm \infty \implies y = f(x) \to 0^+ \)). Next we compute its derivative:

\[
f'(x) = \frac{d}{dx} e^{-x^2} = e^{-x^2} \cdot \frac{d}{dx} -2xe^{-x^2}.
\]

Since \( f'(x) \) exists for all \( x \in \mathbb{R} \), we need only check where \( f'(x) = 0 \), which occurs at \( x = 0 \). From this and the fact that the exponential part of \( f'(x) \), namely \( e^{-x^2} \) will be positive for all \( x \in \mathbb{R} \), we can easily make a sign chart for \( f'(x) = -2xe^{-x^2} \):

\[
\begin{array}{c|cc}
\text{sign of factors:} & \oplus & \ominus \\
\text{sign } f': & \oplus & 0 & \ominus \\
\text{graph behavior:} & / & / & /
\end{array}
\]

We can see clearly that there is a local (in fact global) maximum at \( (0, f(0)) = (0, e^0) = (0, 1) \). This also happens to be the y-intercept.
Next we interest ourselves in the concavity, so we compute
\[
 f''(x) = \frac{d}{dx} \left[ -2xe^{-x^2} \right] = -2x \frac{d}{dx} \left( e^{-x^2} \right) + e^{-x^2} \frac{d}{dx} (-2x) \\
 = -2xe^{-x^2} \cdot (-2x) - 2e^{-x^2} = 4x^2 e^{-x^2} - 2e^{-x^2} \\
 = 2e^{-x^2} (2x^2 - 1).
\]

From this we get \( f''(x) = 0 \iff 2x^2 = 1 \iff x^2 = \frac{1}{2} \iff x = \pm \frac{1}{\sqrt{2}} \). Note that \( \pm \frac{1}{\sqrt{2}} \approx 0.7071 \).

\[
 f''(x) = 2e^{-x^2}(2x^2 - 1)
\]

We detected two inflection points, namely

- \((-1/\sqrt{2}, f(-1/\sqrt{2})) = (-1/\sqrt{2}, e^{-1/2}) \approx (-0.7071, 0.6065), \text{ and}\)
- \((1/\sqrt{2}, f(1/\sqrt{2})) = (1/\sqrt{2}, e^{-1/2}) \approx (0.7071, 0.6065)\)

From the two sign charts we can make a combined sign chart (or synthesize all this on our final graph)

We detected two inflection points, namely

- \((-1/\sqrt{2}, f(-1/\sqrt{2})) = (-1/\sqrt{2}, e^{-1/2}) \approx (-0.7071, 0.6065), \text{ and}\)
- \((1/\sqrt{2}, f(1/\sqrt{2})) = (1/\sqrt{2}, e^{-1/2}) \approx (0.7071, 0.6065)\)

From the two sign charts we can make a combined sign chart (or synthesize all this on our final graph)

Putting all this together, with the limiting behavior, but also noticing the symmetry—namely that if \((x, y)\) is on the graph, then so is \((-x, y)\) because \(f(x) = e^{-x^2}\) will output the same value for inputs of \(\pm x_0\) regardless of \(x_0 \in \mathbb{R}\)—and the \(y\)-intercept \((0, e^0) = (0, 1)\), we can construct a sketch of the graph:
Highlighted are the maximum point at \((0, f(0)) = (0, 1)\), and the two inflection points, namely \((\pm1/\sqrt{2}, e^{-1/2}) \approx (\pm0.7071, 0.6065)\).

The curve above is related to the bell-shaped normal probability distributions one encounters in the fields of probability and statistics.

Rational functions also exhibit some interesting asymptotic behaviors, and can have complicated second derivatives. The following has no vertical asymptotes, but does exhibit horizontal asymptotes and some interesting slope and concavity behavior.

**Example 5.1.6** Follow the directions for the previous example for the function \(f(x) = \frac{x}{x^2 + 1}\).

**Solution:** First we notice that \(x^2 + 1 > 0\) for all \(x \in \mathbb{R}\), so this function is defined, and for that matter continuous, on all of \(\mathbb{R}\). Furthermore, \(f(x) < 0\) for \(x < 0\), while \(f(x) > 0\) for \(x > 0\). We might also notice that

\[
\lim_{x \to \pm\infty} f(x) = \lim_{x \to \pm\infty} \frac{x}{x^2 (1 + \frac{1}{x^2})} = \lim_{x \to \pm\infty} \frac{1}{x (1 + \frac{1}{x^2})} 0,
\]

that is, \(f(x)\) has a two-sided horizontal asymptote \(y = 0\). Next we compute

\[
f'(x) = \frac{(x^2 + 1)\frac{dx}{dx} - x \cdot 2x(x^2 + 1)}{(x^2 + 1)^2} = \frac{x^2 + 1 - x \cdot 2x}{(x^2 + 1)^2} = \frac{1 - x^2}{(x^2 + 1)^2}.
\]

Rewriting this somewhat we can more easily construct a sign chart for \(f'(x)\):

<table>
<thead>
<tr>
<th>signs factors:</th>
<th>(\otimes/\oplus)</th>
<th>(\oplus/\otimes)</th>
<th>(\oplus/\oplus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sign (f'):</td>
<td>(\otimes) -1</td>
<td>(\oplus) 1</td>
<td>(\otimes)</td>
</tr>
<tr>
<td>graph behavior</td>
<td>(\backslash)</td>
<td>(\backslash)</td>
<td>(\backslash)</td>
</tr>
</tbody>
</table>

From this we can clearly detect

- a local minimum at \((-1, f(-1)) = (-1, -1/2)\), and
- a local maximum at \((1, f(1)) = (1, 1/2)\).

At this point we could have a reasonably good sketch, but from the above \(f'\) sign chart and the limiting behavior of \(f(x)\) as \(x \to \pm\infty\), we should expect inflection points as well. To find those we must compute \(f''(x)\), which will be more involved. For that computation we will use one of the intermediate forms for \(f'(x)\):
5.1. HIGHER-ORDER DERIVATIVES AND GRAPHING

\[ f''(x) = \frac{d}{dx} \left( \frac{1 - x^2}{(x^2 + 1)^2} \right) = \frac{(x^2 + 1)^2 \cdot \frac{d}{dx}(1 - x^2) - (1 - x^2) \cdot \frac{d}{dx}(1 + x^2)^2}{[(1 + x^2)^2]^2} \]
\[ = \frac{(x^2 + 1)^2(-2x) - (1 - x^2) \cdot 2(1 + x^2)(2x)}{(1 + x^2)^4} \]
\[ = \frac{(-2x)(x^2 + 1)[x^2 + 1 + 2(1 - x^2)]}{(x^2 + 1)^4} \]
\[ = \frac{(-2x)(3 - x^2)}{(x^2 + 1)^3} \]
\[ \Rightarrow f''(x) = \frac{2x(x^2 - 3)}{(x^2 + 1)^3}. \]

The sign chart for \( f''(x) \) should reflect possible sign changes of \( f''(x) \) at \( x = 0, \pm \sqrt{3} \):

\[ f''(x) = \frac{2x(x^2 - 3)}{(x^2 + 1)^3} \]

signs factors:  \( \ominus \ominus / \oplus \)  \( \ominus \ominus / \oplus \)  \( \ominus \ominus / \oplus \)  \( \ominus \ominus / \oplus \)

sign \( f'' \):  \( \ominus \)  \( -\sqrt{3} \)  \( \oplus \)  \( 0 \)  \( \ominus \)  \( \sqrt{3} \)  \( \oplus \)
concavity:  down  \( -\sqrt{3} \)  up  \( 0 \)  down  \( \sqrt{3} \)  up

Clearly we have inflection points at \( (0, f(0)) = (0, 0) \), at \( (\sqrt{3}, f(\sqrt{3})) = (\sqrt{3}, \frac{1}{4}\sqrt{3}) \), and at \( (-\sqrt{3}, f(-\sqrt{3})) = (-\sqrt{3}, -\frac{1}{4}\sqrt{3}) \), i.e., inflection at approximately \( (0, 0), (1.732, 0.433) \) and \( (-1.732, -0.433) \). Putting these two sign charts for \( f' \) and \( f'' \) together, we get the following chart showing general shapes of pieces of the graph:

\[ f' < 0 \quad f' < 0 \quad f' > 0 \quad f' > 0 \quad f' < 0 \quad f' < 0 \]
\[ f'' < 0 \quad f'' > 0 \quad f'' > 0 \quad f'' < 0 \quad f'' < 0 \quad f'' > 0 \]

From the above charts, the asymptotic behavior and the locations of local extrema and inflection points we can sketch the graph:
The above graph also shows so-called “symmetry with respect to the origin,” the idea being that an inverting lens at the origin—which would take the image in Quadrants I, IV, and reflect and vertically flip that image and place the result in Quadrants III, II respectively—would give us the complete graph in all four quadrants. Thus if \((x, y)\) is on the graph, so is \((-x, -y)\). Put another way, \(f(-x) = -f(x)\) for all \(x\) in the domain, which here means for all \(x \in \mathbb{R}\).

The computations do show that if computing \(f'\) involves a quotient rule, then computing \(f''\) will likely involve a more complicated quotient rule. Also note that the form of \(f(x)\) or \(f'(x)\) which is most convenient for creating a sign chart might not be the simplest form from which to compute the next higher-order derivative. The next example illustrates this also, and involves both a vertical asymptote, and another linear asymptote which is neither vertical nor horizontal. Such linear but nonvertical and nonhorizontal asymptotes are called by various names such as oblique asymptotes or slant asymptotes.

**Example 5.1.7** Graph \(f(x) = \frac{x^2}{x - 4}\) illustrating the signs of \(f(x)\), \(f'(x)\) and \(f''(x)\), together with all asymptotic behavior.

**Solution:** Perhaps the most apparent features of this function’s graph will be its vertical asymptote at \(x = 4\), and its lack of any horizontal asymptotes since the degree of the numerator is greater than that of the denominator.

A sign chart for \(f(x) = \frac{x^2}{x - 4}\) is easily enough constructed, since \(f(x) = 0\) only at \(x = 0\), and \(f(x)\) is discontinuous only at \(x = 4\) (the vertical asymptote):

\[
f(x) = \frac{x^2}{x - 4}
\]

<table>
<thead>
<tr>
<th>signs factors:</th>
<th>(\oplus/\oplus)</th>
<th>(\oplus/\oplus)</th>
<th>(\oplus/\oplus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sign (f):</td>
<td>(\ominus) 0</td>
<td>(\oplus) 4</td>
<td>(\ominus)</td>
</tr>
</tbody>
</table>

In fact, because the numerator has degree at least (actually greater than) the degree of the denominator, we can perform polynomial long division, and we will see that our derivative
computations will be made easier:

\[
\begin{array}{c|c|c|c|c|c|c|c|c|c}
& & & & & & & & & \\
\hline
x - 4 & x^2 & 0 & x & 0 & -x^2 & 0 & 4x & 0 & 16 \\
\hline
& & & & & & & & & \\
\end{array}
\]

For differentiation and some other purposes (explained next), it is clearly easier to use the form

\[f(x) = x + 4 + \frac{16}{x - 4},\]

than to use the original form \(f(x) = \frac{x^2}{x - 4}\). Before we compute derivatives, note that

\[|\lambda| \text{ large} \implies f(\lambda) = x + 4 + \frac{16}{x - 4} \approx x + 4,
\]

and so the vertical distance between the graphs of \(y = f(x)\) and \(y = x + 4\) shrinks towards zero as \(x \to \pm \infty\). Thus we have an oblique asymptote, which is the line \(y = x + 4\), as an asymptote as \(x \to \pm \infty\). This will be an important feature to include in our graph. Now we compute

\[f'(x) = \frac{d}{dx} \left[ x + 4 + \frac{16}{x - 4}\right] = 1 - \frac{16}{(x - 4)^2} \cdot \frac{d(x - 4)}{dx}
\]

\[\implies f'(x) = 1 - \frac{16}{(x - 4)^2}.
\]

For the sake of building a sign chart we can either find where \(f' = 0\) and where \(f'\) does not exist from the expression above, or we can recombine terms:

\[f'(x) = \frac{(x - 4)^2 - 16}{(x - 4)^2} = \frac{x^2 - 8x + 16 - 16}{(x - 4)^2} = \frac{x(x - 8)}{(x - 4)^2}.
\]

Thus we need \(x = 0, 4, 8\) as our dividing points on our sign chart for \(f'(x)\):

\[f'(x) = \frac{x(x - 8)}{(x - 4)^2}
\]

\begin{tabular}{c|c|c|c|c|c|c|c}
signs factors: & \(\ominus\ominus\)/\(\ominus\) & \(\ominus\ominus\)/\(\ominus\) & \(\ominus\ominus\)/\(\ominus\) & \(\ominus\ominus\)/\(\ominus\) & \(\ominus\ominus\)/\(\ominus\) & \(\ominus\ominus\)/\(\ominus\) & \(\ominus\ominus\)/\(\ominus\) \\

sign \(f'\): & \(\ominus\) & 0 & \(\ominus\) & \(\ominus\) & 8 & \(\ominus\) & \(\ominus\)

graph behavior: & \(\nearrow\) & \(\searrow\) & \(\nearrow\) & \(\searrow\) & \(\nearrow\)
\end{tabular}

Noting that \(x = 4\) is a vertical asymptote of \(f(x)\) and the only discontinuity of \(f(x)\), we can see clearly that we have two local extrema (since \(x = 0, 8\) are not discontinuities):

- local maximum at \((0, f(0)) = (0, 0)\), and a
- local minimum at \((8, f(8)) = (8, 64/(8 - 4)) = (8, 16, 16).\)
We can use a simpler form of $f'$ to compute $f''$:

$$f''(x) = \frac{d}{dx} \left[ 1 - 16(x - 4)^{-2} \right] = 32(x - 4)^{-3}$$

$$\implies f''(x) = \frac{3}{2(x - 4)^3}.$$  

The sign chart for $f''$ follows readily, with $x = 4$ being the only dividing point.

$$f''(x) = \frac{3}{2(x - 4)^3}$$

<table>
<thead>
<tr>
<th>sign $f''$:</th>
<th>⊖</th>
<th>⊕</th>
</tr>
</thead>
<tbody>
<tr>
<td>concavity:</td>
<td>down</td>
<td></td>
</tr>
</tbody>
</table>

Combining the two charts above would be fairly straightforward. Graphing the function to illustrate signs of $f'$ and $f''$, along with the asymptotes, would yield the following:

To review, we were given $f(x) = x^2/(x - 4)$, which as it stands would require a quotient rule to find $f'(x)$, but polynomial long division yielded $f(x) = x + 4 + 16/(x - 4)$, which is not only easier to differentiate, but also makes more obvious the oblique asymptote since $|x|$ large implies $f(x) \approx x + 4$. However, the original form of $f(x)$ is better for constructing a sign chart of $f(x)$.

When presented with a rational function $f(x) = p(x)/q(x)$, where $p$ and $q$ are polynomials, if the degree of $p$ is at least as large as that of $q$, so polynomial long division is possible, it is usually worth considering the form of $f(x)$ which arises from that long division, for purposes of differentiation and possible asymptotic behavior. The same can be said about $f'(x)$, though the effects of its asymptotics usually can be found in observations for $f(x)$ and $f''(x)$.

Nonlinear asymptotes are also a possibility. Consider the following example.
Example 5.1.8 Graph \( f(x) = \frac{x^3}{x + 2}, \) illustrating signs of \( f(x), f'(x) \) and \( f''(x) \) as well as all asymptotic behavior.

Solution. First we note the vertical asymptote at \( x = -2, \) and then we construct a sign chart for \( f(x): \)

\[
\begin{array}{c|cccc}
\text{signs factors} & \Theta/\Theta & \Theta/\Theta & \Theta/\Theta & \Theta/\Theta \\
\text{sign } f: & \Theta & -2 & 0 & \Theta \\
\end{array}
\]

Anticipating the derivative computation, we again see that a viable alternative to the quotient rule is to perform division first. To do so here we take a minor shortcut, using \( a^3 + b^3 = (a + b)(a^2 - ab + b^2) \) by adding and subtracting what would be the \( "a^2b" \) term in the numerator:

\[
f(x) = \frac{x^3 + 8 - 8}{x + 2} = \frac{(x + 2)(x^2 - 2x + 4) - 8}{x + 2} = x^2 - 2x + 4 - \frac{8}{x + 2}.
\]

From this we note the very important asymptotic feature, namely that for large \( |x| \) we have

\[
f(x) \approx x^2 - 2x + 4,
\]

which is a vertically-opening parabola. We can find the vertex using algebra techniques, or from realizing that at that vertex the slope is zero, so \( 0 = \frac{d}{dx}(x^2 - 2x + 4) = 2x - 2 = 2(x - 1) \Rightarrow x = 1. \) The vertex of the parabola \( y = x^2 - 2x + 4 \) is thus \( (1, 1^2 - 2(1) + 4) = (1, 3). \) (If extra accuracy is desired for the graph of \( f(x), \) several points on this parabolic asymptote can be drawn.) To compute \( f'(x) \) we could use the form from the division, i.e.,

\[
f'(x) = \frac{d}{dx} \left[ x^2 - 2x + 4 - \frac{8}{x + 2} \right] = 2x - 2 + \frac{8}{(x + 2)^2},
\]

or we could use the original form of the function to arrive at an already-combined fraction:

\[
f'(x) = \frac{d}{dx} \left[ \frac{x^3}{x + 2} \right] = \frac{(x + 2)(3x^2) - x^3(1)}{(x + 2)^2} = \frac{3x^3 + 6x^2 - x^3}{(x + 2)^2} = \frac{2x^3 + 6x^2}{(x + 2)^2} = \frac{2x^2(x + 3)}{(x + 2)^2}.
\]

From this we can see our sign chart will have possible sign changing borders at \( x = -3, -2, 0. \)

\[
\begin{array}{c|cccc}
\text{signs factors} & \Theta/\Theta & \Theta/\Theta & \Theta/\Theta & \Theta/\Theta \\
\text{sign } f': & \Theta & -3 & 2 & 0 \\
\text{graph behavior} & \downarrow & / & / & / \\
\end{array}
\]

Since \( x = -3 \) is a point of continuity of \( f(x), \) from the sign chart above this must be a local minimum, specifically at \( (-3, f(-3)) = (-3, -27/(-1)) = (-3, 27). \) Next we compute \( f''(x): \)

\[
f''(x) = \frac{d}{dx} \left[ 2x^2 + 8(x + 2)^{-2} \right] = 2 - 16(x + 2)^{-3} \cdot 1 = \frac{2(x + 2)^3 - 16}{(x + 2)^3}
\]

\[
\Rightarrow f''(x) = \frac{2[(x + 2)^3 - 8]}{(x + 2)^3}.
\]
While we could further factor the numerator, it is reasonably clear that the numerator’s value is zero when $x + 2 = 2$, i.e., when $x = 0$. The denominator is zero when $x = -2$.

$$f''(x) = \frac{2[(x + 2)^3 - 8]}{(x + 2)^3}$$

<table>
<thead>
<tr>
<th>signs factors:</th>
<th>$\ominus$/\ $\ominus$</th>
<th>$\ominus$/\ $\oplus$</th>
<th>$\oplus$/\ $\oplus$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{sign } f''$:</td>
<td>$\ominus$</td>
<td>$\ominus$</td>
<td>$\oplus$</td>
</tr>
<tr>
<td>concavity:</td>
<td>up</td>
<td>down</td>
<td>up</td>
</tr>
</tbody>
</table>

Clearly the concavity changes at the vertical asymptote $x = -2$, and then at the point $(0, f(0)) = (0, 0)$ on the graph. Combining these two charts into one, we get the following:

| $f' < 0$ | $f' > 0$ | $f' > 0$ | $f' > 0$ |
| $f'' > 0$ | $f'' > 0$ | $f'' < 0$ | $f'' > 0$ |

<table>
<thead>
<tr>
<th>shape of graph:</th>
<th>$\ominus$</th>
<th>$\ominus$</th>
<th>$\oplus$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-3$</td>
<td>$2$</td>
<td>$0$</td>
<td></td>
</tr>
</tbody>
</table>

Recalling the vertical asymptote at $x = -2$, the nonlinear asymptote $y = x^2 - 2x + 4$, the local minimum at $(-3, 27)$, and the inflection point (also the $x$-intercept and $y$-intercept) at $(0, 0)$, we can reasonably sketch the graph, which is given below (computer-generated):

Note that we did not compute the limits as $x \to \pm\infty$, which are simple enough to compute, but which are also consequences of the other analyses, such as sign charts of $f'$ and $f''$.  

\footnote{Technically $(0, 0)$ is the only actual “inflection point,” since it is on the curve while there is no point on the curve where $x = -2$. Clearly it is important to note that the concavity changes at the discontinuity at $x = -2$ as well.}
When given a piecewise-defined function, a general approach is to consider each piece separately. The absolute value function is defined piecewise (one way for nonnegative inputs, and another way for negative inputs) but its properties can be at times seen in a simpler light. Consider the following example.

**Example 5.1.9** Graph \( f(x) = |x^3 - 3x| \), illustrating where the function is increasing, decreasing, concavity and asymptotic behaviors.

**Solution:** We will do so by first analyzing and graphing the function inside the absolute values, temporarily calling it \( g(x) = x^3 - 3x \), and the graph of \( f(x) = |g(x)| \) will follow quickly from that of \( g(x) \). Note that \( g(x) = x(x^2 - 3) \), and so \( g(x) = 0 \iff x \in \{0, -\sqrt{3}, \sqrt{3}\} \).

\[
g(x) = x(x^2 - 3)
\]

**signs factors:** ⊕⊕ ⊕⊕ ⊕⊕ ⊕⊕

\[\text{sign } g: \quad \oplus -\sqrt{3} \ominus 0 \ominus \sqrt{3} \oplus\]

Next we consider \( \frac{dg(x)}{dx} = \frac{d}{dx}(x^3 - 3x) = 3x^2 - 3 = 3(x^2 - 1) \), which is clearly zero at \( x = \pm 1 \). The sign chart for \( \frac{dg(x)}{dx} \) then follows:

\[
\frac{dg(x)}{dx} = 3(x + 1)(x - 1)
\]

**signs factors:** ⊕⊕⊕ ⊕⊕⊕ ⊕⊕⊕

\[\text{sign } \frac{dg(x)}{dx}: \quad \oplus -1 \ominus 1 \oplus\]

\[\text{graph of } g: \quad / \backslash \backslash /\]

This inner function \( g(x) \) clearly has a local maximum at \((-1, g(-1)) = (-1, (-1)^3 - 3(-1)) = (-1, 2)\), and a local minimum at \((1, g(1)) = (1, 1^3 - 3(1)) = (1, -2)\).

Next we note \( \frac{d^2g(x)}{dx^2} = \frac{d}{dx}(3x^2) = 6x \), which is clearly negative for \( x < 0 \) and positive for \( x > 0 \), hence we have an inflection point at \((0, g(0)) = (0, 0)\). Combining our first and second derivative information we make a combined chart (not drawn to scale):

\[
g' > 0 \quad g' < 0 \quad g' < 0 \quad g' > 0
\]

\[
g'' < 0 \quad g'' < 0 \quad g'' > 0 \quad g'' > 0
\]

We use the chart and locations of extrema and the inflection point to graph the function \( g(x) \) below in gray, and then use the definition of the absolute value function which will give us to graph

\[
f(x) = |g(x)| = \begin{cases} 
g(x) & \text{if } g(x) \geq 0, \\
-g(x) & \text{if } g(x) < 0.
\end{cases}
\]
Points of interest on the graph of $f(x) = |x^3 - 3x|$ include the two maxima, namely $(±1, 2)$, the three minima at $(0, 0)$ and $(±\sqrt{3}, 0)$, those also being $x$-intercepts, and inflection points $(±\sqrt{3}, 0)$. (Note that the inflection point of $g(x)$ is not an inflection point of $f(x) = |g(x)|$.)

In the above example we strayed somewhat from the spirit of the previous examples, where we found directly the possible points where $f$ is increasing, decreasing, concave up and concave down. The previous methods can be employed directly with this function $f(x) = |x^3 - 3x|$, though it requires some work to find a piecewise-defined version of the function, in this case

$$f(x) = \begin{cases} 
-(x^3 - 3x) & \text{if } x < -\sqrt{3}, \\
 x^3 - 3x & \text{if } -\sqrt{3} \leq x \leq 0, \\
-(x^3 - 3x) & \text{if } 0 < x < \sqrt{3}, \\
 x^3 - 3x & \text{if } x \geq \sqrt{3}.
\end{cases}$$

When we look for possible points where $f'$ may change sign, we find ourselves computing, depending upon the interval, $\frac{d}{dx} \left[ \pm (x^3 - 3x) \right] = \pm (3x^2 - 3) = 0 \iff x = \pm 1$. One can then make a sign chart for each interval based upon these things, but the boundaries of the intervals are also possible places for $f'$ to change signs, so we must also consider the three points $0, \pm \sqrt{3}$. Similarly with $f''$ signs. This is unavoidable with some piecewise-defined functions, but those in which the final function stage is the absolute value function can be more easily graphed by considering the “inside” function’s graph first, and reflecting any parts with negative heights across the $x$-axis and into the upper quadrants. In the original analysis used in the example (analyzing the “inner” function first), some of the maxima, minima and inflection points often emerge from the graph rather than from the the derivative analysis of the “inner” function.

Since there are infinitely many functions we can consider, there are many different functional behaviors we can encounter. Even relatively simple combinations of well understood functions can yield new functions with “emergent properties” which are not always clear at first glance from the new function’s formula. Our derivative tests and simple observations regarding asymptotics can be employed to detect many of these emergent properties of both simple and complicated functions.

While we have mainly dealt with algebraic functions (simple combinations of powers of $x$, including fractional powers), we have also seen a simple observation regarding a trigonometric function, a simple logarithmic graph and one case which involved an exponential function. Everything we did here can also apply to trigonometric functions, inverse trigonometric functions, piecewise-defined functions and so on.

Some combinations of functions will, unfortunately, yield results where it is algebraically impossible to find exactly all solutions of $f = 0, f' = 0, f'' = 0$ and so on. In Section 5.5, we will have other techniques to approximate solutions to many such otherwise algebraically unwieldy equations, accurate (when the technique succeeds) to as many decimal places for which we care.
to carry out necessary computations, thus extending our ability to use the techniques of this section.
Before and after we develop those techniques we will have several other uses for the derivative.

**Exercises**

1. Show that \((fg)'' = f''g + 2f'g' + fg''\). 
   In other words, derive the formula
   \[
   \frac{d^2}{dx^2}(f(x)g(x)) = f''(x)g(x) + 2f'(x)g'(x) + f(x)g''(x).
   \]
   What do you suppose is the formula for \((fg)'''\)?

For 2–19, graph the function

2. \(f(x) = x^4 - x^2\)
3. \(f(x) = \frac{x^2 - 1}{x^2 + 1}\)
4. \(f(x) = \frac{x^2 + 1}{x^2 - 1}\)
5. \(f(x) = x(\sqrt{x} - 3)\)
6. \(f(x) = x \ln x\). You can assume \(\lim_{x \to 0^+} f(x) = 0\).
7. \(f(x) = xe^{-x^2}\). You can assume \(\lim_{x \to \pm \infty} f(x) = 0\).
8. \(f(x) = x + \ln |x|\)
9. \(f(x) = \frac{x^3}{x^2 - 1}\)
10. \(f(x) = \frac{x^4}{x^2 - 1}\)
11. \(f(x) = \left(\frac{x^{1/3}}{1}ight) \left(\frac{x^{1/3}}{2}\right)\)
12. \(f(x) = \frac{x^2 - x + 2}{x - 3}\)
13. \(f(x) = (x^2 - 9)^{2/3}\)
14. \(f(x) = \frac{x^2 + 1}{x^2}\)
15. \(f(x) = \frac{x^4 - 1}{x^2}\)
16. \(f(x) = |x^2 - 2x - 3|\)
17. \(f(x) = x\sqrt{9 - x^2}\)
18. \(f(x) = \frac{x^2 + 1}{x^2 - 9}\)
19. \(f(x) = \frac{x^2 - 4}{x^2 - 9}\)

For Exercises 20–22 ketch the graph of \(f(x)\) on the given interval \([a, b]\) illustrating the signs of \(f(x), f'(x), f''(x)\).

20. \(f(x) = x\sqrt{16 - x^2}\), on \([-3, 3]\).
21. \(f(x) = |x^2 + 4x + 3|\), on \([-2, 2]\).
22. \(f(x) = x - \sin x\), on \([0, 2\pi]\).
23. The position of a moving object at time \(t\) is given by \(s(t) = 2t^3 + 3t^2 - 6t\), where \(t \in [0, 10]\).
   (a) Find the time interval(s) over which the object’s velocity is positive.
   (b) Find the time interval(s) over which the object’s acceleration is positive.
24. The position of a moving object at time \(t\) is given by \(s(t) = t^5 - 5t^4\), where \(t \in [0, 10]\).
   (a) Find the time interval(s) over which the object’s velocity is positive.
   (b) Find the time interval(s) over which the object’s acceleration is positive.
25. Determine the \(x\)-coordinate of the only inflection point of a cubic polynomial
   \(f(x) = ax^3 + bx^2 + cx + d\), where \(a \neq 0\).
26. Consider a function of the form \( f(x) = (x + 3)^n \). Determine any inflection point(s) of the graph of \( f(x) \) when

(a) \( n = 1 \)
(b) \( n = 2 \)
(c) \( n = 3 \)
(d) \( n = 4 \)
(e) \( n = 5 \)

(f) Make a conjecture as to when \( f(x) = (x + 3)^n \) has an inflection point given that \( n \) is a positive integer.

27. What condition(s) must be placed on \( a, b, c, d, e \in \mathbb{R} \) so that the graph of \( f(x) = ax^4 + bx^3 + cx^2 + dx + e \) has 2 inflection points?

28. Make a sketch of a function \( f(x) \) which satisfies all of the following conditions:

- \( f(0) = f(4) = 0 \),
- \( f'(x) > 0 \) if \( x < 2 \),
- \( f'(2) = 0 \),
- \( f'(x) < 0 \) if \( x > 2 \),
- \( f''(x) < 0 \) for all \( x \in \mathbb{R} \).

29. Make a sketch of a function \( f(x) \) which satisfies all of the following conditions:

- \( f(0) = f(4) = 0 \),
- \( f'(x) > 0 \) if \( x < 2 \),
- \( f'(2) \) does not exist,
- \( f'(x) < 0 \) if \( x > 2 \),
- \( f''(x) > 0 \) if \( x \neq 2 \).

30. Consider a function \( f(x) \) which has a two-sided vertical asymptote at \( x = 2 \). Answer each of the following. (Hint: consider the graphs.)

(a) If \( f'(x) > 0 \) for \( x \in (0, 2) \cup (2, 4) \), compute \( \lim_{x \to 2^-} f(x) \) and \( \lim_{x \to 2^+} f(x) \).

(b) If \( f''(x) > 0 \) for \( x \in (0, 2) \cup (2, 4) \), compute those same limits, \( \lim_{x \to 2^-} f(x) \) and \( \lim_{x \to 2^+} f(x) \).

31. Sketch the graph for \( f(x) = \frac{|x|}{1 + |x|} \), illustrating the signs of \( f(x) \), \( f'(x) \) and \( f''(x) \). Be sure to illustrate any asymptotes. (Hint: find a piecewise definition of \( f(x) \).)
5.2 Extrema on Closed Intervals

In this section we see how to locate the extrema (maxima and minima) of a continuous function \( f : [a, b] \rightarrow \mathbb{R} \). The main theorem is Theorem 5.2.2, page 480, stating that \( f \) will achieve maximum and minimum values somewhere in \([a, b]\), and that any such point where an extremum is achieved will either be an endpoint \( a \) or \( b \), or an interior point \( x \in (a, b) \) where \( f'(x) \) is zero or does not exist. The theorem is ultimately quite intuitive. However, to prove the theorem requires a somewhat lengthy argument, so the actual argument is given its own subsection directly below, and the formal statement of the theorem opens Subsection 5.2.2, page 480. That subsection also contains theoretical examples, and then applied examples follow in Subsection 5.2.3.

5.2.1 Argument for Main Theorem

We wish to find those points in \([a, b]\) at which a continuous function \( f : [a, b] \rightarrow \mathbb{R} \) achieves its maximum and minimum values possible for those \( x \in [a, b] \). Recall from Section 3.3 that \( f([a, b]) \), the set of all possible outputs of \( f(\cdot) \) resulting from inputs from \([a, b]\), will itself be a finite, closed interval, and thus will have a maximum and minimum value, i.e., there will indeed exist some \( x_{\min}, x_{\max} \) in the interval \([a, b]\) so that \( f([a, b]) = [f(x_{\min}), f(x_{\max})] \), i.e.,

\[
(\exists x_{\min}, x_{\max} \in [a, b]) (\forall x \in [a, b]) [f(x_{\min}) \leq f(x) \leq f(x_{\max})]. \quad (5.1)
\]

This was the essence of the Extreme Value Theorem (EVT), which was Corollary 3.3.1, page 196. The next theorem is stated in a manner reflective of its proof. However, its equivalent forms given in the corollaries are more useful in applications.

**Theorem 5.2.1** Suppose \( f : [a, b] \rightarrow \mathbb{R} \) is continuous, and that \( f(x) \) achieves its maximum or minimum value for the interval \([a, b]\) at an interior point \( x_0 \in (a, b) \). Under these assumptions, if \( f'(x_0) \) exists then \( f'(x_0) = 0 \), i.e.,

\[
f'(x_0) \text{ exists } \implies f'(x_0) = 0.
\]

**Proof:** First we look at the case \( f : [a, b] \rightarrow \mathbb{R} \) continuous, \( x_0 \in (a, b) \), and \( f(x_0) \) is the maximum value of \( f(x) \) on \([a, b]\). We need to show that if \( f'(x_0) \) exists then \( f'(x_0) = 0 \). So we (further) suppose \( f'(x_0) \) exists. This means

\[
\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = f'(x_0) \text{ exists}.
\]

Thus the left and right side limits exist and must agree with this two-sided limit:

\[
\lim_{\Delta x \to 0^-} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = f'(x), \quad (5.2)
\]

\[
\lim_{\Delta x \to 0^+} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = f'(x). \quad (5.3)
\]

Now we carefully consider each of these two limits, (5.2) and (5.3). In particular we look at the signs of these. In both limits above, since \( f(x_0) \) is a maximum, we have \( f(x_0 + \Delta x) \leq f(x_0) \), implying \( f(x_0 + \Delta x) - f(x_0) \leq 0 \). Thus the numerators in the limits are both nonpositive.

---

\[\text{It is very strongly suggested that the reader at least briefly review Section 3.3 at some point during the reading of this current section. It would likely be particularly helpful to reconsider the figures in that section to recall why it is necessary to have a continuous function and a closed interval for the conclusion of the Extreme Value Theorem to be guaranteed.}\]
Next we look specifically at (5.2). Since $\Delta x \to 0^-$, we are looking at limits of fractions with denominators $\Delta x < 0$. Thus we have

$$f'(x_0) = \lim_{\Delta x \to 0^-} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \leq 0. \quad (5.4)$$

On the other hand, when we instead look closely at the form of the limit in (5.3), we see

$$f'(x_0) = \lim_{\Delta x \to 0^+} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \geq 0. \quad (5.5)$$

So from (5.4) and (5.5) we have $(f'(x_0) \geq 0)$ and $(f'(x_0) \leq 0)$, i.e., $0 \leq f'(x_0) \leq 0$, so we have to conclude $f'(x_0) = 0$, q.e.d. for the case of an interior maximum.

It is useful to consider graphically why inequalities (5.4) and (5.5) should hold, as well as the conclusion of Theorem 5.2.1:

$$[f'(x_0) \text{ exists}] \implies [f'(x_0) = 0] \iff [f'(x_0) \text{ does not exist}] \lor [f'(x_0) = 0].$$

We thus rewrite Theorem 5.2.1, in the form of the very important corollary below:

**Corollary 5.2.1** Suppose $f : [a, b] \to \mathbb{R}$ is continuous, and achieves its maximum or minimum output value for $[a, b]$ at an interior point $x_0 \in (a, b)$. Then

$$[f'(x_0) = 0] \lor [f'(x_0) \text{ does not exist}].$$

So any interior (non-endpoint) extremum of a continuous function on a closed interval $[a, b]$ must occur at points where $f'$ is zero or does not exist. This is illustrated in Figure 5.3, page 479.

Because of the two possibilities for the derivative at an interior extremum, it is convenient to define terminology for those interior points where $f'(x) = 0$ or $f'(x)$ does not exist. We collectively call such points **critical points**.

**Definition 5.2.1** For $f(x)$ defined and continuous on $(a - \delta, a + \delta)$, for some $\delta > 0$, the point $x = a$ is called a **critical point** of $f(x)$ if and only if $f'(a) = 0$ or $f'(a)$ does not exist.

It helps later to establish now that, when analyzing a function $f(x)$ where we consider only those $x \in [a, b]$, by **critical points** we mean only those **interior** points $x_0 \in (a, b)$, where $f'(x_0)$ is zero or nonexistent, thus allowing $x_0$ to have some “wiggle room” both to the left and to the right within the interval in question (which allows for derivative computations). Using this definition, we can again rewrite Theorem 5.2.1, and Corollary 5.2.1: 

---

5In the proof above it was necessary that $x_0 \in (a, b)$, so that there is room to the left and right of $x_0$ in the interval $(a, b)$, where $f(x)$ is defined, so we can have $\Delta x \to 0^-$ and $\Delta x \to 0^+$ and eventually, for small enough $|\Delta x|$, have $x + \Delta x \in (a, b)$—where $f(x)$ is defined—as well. In contrast, if $x_0 \in \{a, b\}$, i.e., $x_0$ is an endpoint, we cannot say anything of its derivative without further information.
5.2. EXTREMA ON CLOSED INTERVALS

Figure 5.2: Illustration of behavior of difference quotients (defined on page 296) $(\Delta y)/(\Delta x)$, representing slopes of “secant” line segments (gray) joining points on the curve, where one endpoint is at a maximum or minimum. Near a maximum, these line segments have non-negative slope immediately to the left, and nonpositive slope immediately to the right, of the maximum. The reverse is true at a minimum: nonpositive slope to the left, and nonnegative slope to the right. These observations are crucial in the proof that $f' = 0$ or does not exist at such interior maxima or minima.

Figure 5.3: Illustration that interior extrema can only occur at points where $f'$ is zero or does not exist. In fact, the same is true of local extrema (not all labeled above).
For \( f(x) \) continuous on \([a, b] \), any interior extremum must occur at a critical point.

This is still not quite the final and most useful version of this theorem. To arrive at the most complete statement, let us first recall (from Section 3.3) that if \( f(x) \) is continuous on \([a, b] \), it will necessarily achieve its maximum and minimum values for that interval on that interval. Now we can import our theorem, which states that if one of these is achieved in the interior of \([a, b] \), i.e., on \((a, b)\), then that point must be a critical point:

Assuming that \( x_0 \) is any point in \([a, b]\) at which \( f \) achieves either its maximum or minimum value for that interval \([a, b]\), then

\[
x_0 \in [a, b] \iff (x_0 \in (a, b)) \lor (x_0 \in \{a, b\})
\]

\[
\implies (x_0 \text{ a critical point of } f) \lor (x_0 \in \{a, b\}).
\]

Corollary 5.2.1, page 478

Thus \( x_0 \) must be a critical point of \( f \) in \((a, b)\), or an endpoint of the interval \([a, b]\).

So when we look for a maximum or minimum value of a continuous \( f \) on \([a, b] \), we need only look at the set of critical points in \((a, b)\), and at the endpoints \( a \) and \( b \), for our candidates for \( x_{\text{min}} \) and \( x_{\text{max}} \). See below.

### 5.2.2 Main Theorem and Examples

We tie all this together in the following theorem, which includes for context the Extreme Value Theorem (EVT), first given as Corollary 3.3.1, page 196.

**Theorem 5.2.2** For any function \( f(x) \), continuous on a closed interval \([a, b]\),

1. (EVT) there exist \( x_{\text{min}}, x_{\text{max}} \in [a, b] \) such that

\[
(\forall x \in [a, b]) [f(x_{\text{min}}) \leq f(x) \leq f(x_{\text{max}})]; \tag{5.6}
\]

2. furthermore, any such \( x_{\text{min}}, x_{\text{max}} \in [a, b] \) satisfying (5.6) must be a critical point of \( f \) in \((a, b)\), or an endpoint \( a, b \) of the interval \([a, b]\).

In other words, as long as the function is continuous on \([a, b]\), we need only scrutinize the critical points—where \( f' \) is zero or does not exist—in \((a, b)\), and the points \( a, b \) themselves. Checking the function values at these points exhausts all possibilities for the maximum and minimum values of \( f(x) \). The analysis also yields the locations of the extrema, as \( x \)-values.

**Example 5.2.1** Find the maximum and minimum values of \( f(x) = x^3 - x \) for \( x \in [0, 10] \), and their locations (as \( x \)-values).

**Solution:** Clearly \( f(x) \) is continuous on \([0, 10]\) (it is a polynomial), and \( f'(x) = 3x^2 - 1 \) exists throughout \((0, 10)\). We first find the critical points, and then check the values of \( f \) at any critical points in \((0, 10)\), and then the endpoints \( x = 0, 10 \).

\[
f'(x) = 0 \iff 3x^2 - 1 = 0
\]

\[
\iff 3x^2 = 1
\]

\[
\iff x^2 = 1/3 \iff x = \pm 1/\sqrt{3}.
\]
5.2. EXTREMA ON CLOSED INTERVALS

Now \(-1/\sqrt{3} \not\in (0, 10)\), so the only relevant critical point is \(x = 1/\sqrt{3} \approx 0.577350269 \approx 0.577\).

Now we check the functional (output) values at the one critical point, and the two endpoints.

**critical point:** \(f\left(\frac{1}{\sqrt{3}}\right) = \left(\frac{1}{\sqrt{3}}\right)^3 - \frac{1}{\sqrt{3}} = \frac{1}{3\sqrt{3}} - \frac{1}{\sqrt{3}} = \frac{1 - 3}{3\sqrt{3}} = \frac{-2}{3\sqrt{3}} \approx -0.3849.\)

**endpoints:** \(f(0) = 0^3 - 0 = 0\), and \(f(10) = 10^3 - 10 = 1000 - 10 = 990.\)

From this we get that the minimum value of \(f\) on \([0, 10]\) is \(-2/3\sqrt{3} \approx -0.3849\), occurring at \(x = 1/\sqrt{3} \approx 0.577\), and the maximum value is 990 occurring at \(x = 10\).

To summarize geometrically, the highest and lowest points on the graph of the function \(f(x) = x^3 - x\) for \(x \in [0, 10]\) occur at the points:

- **maximum:** \((10, 990)\)
- **minimum:** \((1/\sqrt{3}, -2/3\sqrt{3}) \approx (0.577, -0.3849)\)

We could have graphed \(y = f(x)\), but we would have discovered possible locations of the extrema from the sign chart, but it is not necessary since the method here gave us only three points to check the function’s values.

**Example 5.2.2** Find the maximum and minimum values of \(f(x) = \sin x + \cos x\) over the interval \(x \in [0, 2\pi]\).

**Solution:** We note that this function is continuous, and \(f'(x) = \cos x - \sin x\) exists for all real \(x\), so the main theorem’s requirement that \(f\) is continuous on \([a, b]\) and differentiable on \((0, 2\pi)\) is satisfied. Now we find those critical points within \((0, 2\pi)\), and examine \(f\) evaluated there and at the endpoints of our interval.

\[f'(x) = 0 \iff \cos x - \sin x = 0 \iff \cos x = \sin x \iff 1 = \tan x \iff x \in \{\pi/4, 5\pi/4\}.\]

**critical points:**

\[f(\pi/4) = \sin \frac{\pi}{4} + \cos \frac{\pi}{4} = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} = \sqrt{2} \approx 1.41421356,\]

\[f(5\pi/4) = \sin \frac{5\pi}{4} + \cos \frac{5\pi}{4} = -\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} = -\sqrt{2} \approx -1.41421356.\]

**endpoints:**

\[f(0) = \sin 0 + \cos 0 = 0 + 1 = 1,\]

\[f(2\pi) = \sin 2\pi + \cos 2\pi = 0 + 1 = 1.\]

- **maximum:** \((\pi/4, \sqrt{2})\).
- **minimum:** \((-\pi/4, -\sqrt{2})\).
Example 5.2.3 \( f(x) = x^{2/3}, \ x \in [-8, 8] \).

Solution: This function is defined and continuous on all of \( \mathbb{R} \), so it is continuous on \([-8, 8]\). Now
\[
\frac{df}{dx}(x) = \frac{2}{3} x^{-1/3} = \frac{2}{3 \sqrt[3]{x}}.
\]

While this is never zero, we note that \( \frac{df}{dx}(x) \) does not exist at \( x = 0 \), which is in our interval. Now we compute \( f \) at the critical point and the endpoints.

Critical point: \( f(0) = 0^{2/3} = 0 \).

Endpoints:
\[
\begin{align*}
\quad f(-8) &= (-8)^{2/3} = \left[ (-8)^{1/3} \right]^2 = (-2)^2 = 4,
\quad f(-8) &= (8)^{2/3} = \left[ (8)^{1/3} \right]^2 = (2)^2 = 4.
\end{align*}
\]

Thus the minimum value of \( f(x) \) on \([-8, 8]\) is \( f(0) = 0 \), and the maximum value is \( f(\pm 8) = 4 \).

Example 5.2.4 Find the maximum and minimum values of \( f(x) = x - \sqrt{x} \) for the interval \( x \in [0, 4] \).

Solution: Note that \( f \) is continuous on \([0, 4]\) and differentiable on \((0, 4)\) (see derivative computation below), so we can proceed as before. First we compute
\[
f'(x) = 1 - \frac{1}{2\sqrt{x}} = \frac{2\sqrt{x} - 1}{2\sqrt{x}},
\]
and see \( f'(x) = 0 \) when \( \sqrt{x} = \frac{1}{2} \), i.e., when \( x = \frac{1}{4} \in [0, 4] \). (We might note \( f'(0) \) does not exist, but \( x = 0 \) being an endpoint in this example, we don’t also classify it as a critical point, a term we reserve for interior points where \( f' = 0 \) or does not exist.) Thus we check the relevant points and note the maximum and minimum values and their locations.

\[\text{The approximation } x = 1/\sqrt{3} \approx 0.577350269 \text{ is given so that some idea of the actual value is visible, and also to illustrate that, indeed, } 1/\sqrt{3} \in [0, 10]. \text{ For the sake of illustration, the approximate values of both inputs and outputs of the functions will be given routinely, though the analysis should use the exact values when possible before stating approximations. Indeed, we will usually refrain from using any approximations in actual calculations throughout this text, excepting for where it is similarly illustrative.} \]
critical point: $f(1/4) = 1/4 - \sqrt{1/4} = 1/4 - 1/2 = -1/4$.

endpoints: $f(0) = 0 - \sqrt{0} = 0$, $f(4) = 4 - \sqrt{4} = 4 - 2 = 2$.

maximum: $(4, f(4)) = (4, 2)$.

minimum: $(1/4, f(1/4)) = (1/4, -1/4)$.

Example 5.2.5 Find the maximum and minimum values of $f(x) = \frac{x}{x^2 + 1}$ for the interval $x \in [0, 2]$.

Solution: Note that the function is continuous over all of $\mathbb{R}$, since $x^2 + 1 \geq 1 > 0$, so that $(\forall x \in \mathbb{R})[x^2 + 1 \neq 0]$. Using the quotient rule to compute the derivative we get

$$f'(x) = \frac{(x^2 + 1)\frac{d}{dx}(x) - x \frac{d}{dx}(x^2 + 1)}{(x^2 + 1)^2} = \frac{x^2 + 1 - x(2x)}{(x^2 + 1)^2} = \frac{1 - x^2}{(x^2 + 1)^2},$$

which is defined everywhere, and $f'(x) = 0$ for $x = \pm 1$. The only critical point in $[0, 2]$ is thus at $x = 1$.

critical point: $f(1) = \frac{1}{1^2 + 1} = \frac{1}{2} = 0.5$.

endpoints: $f(0) = 0/(0^2 + 1) = 0$, $f(2) = 2/(2^2 + 1) = 2/5 = 0.4$.

maximum: $(1, f(1)) = (1, 1/2)$.

minimum: $(0, f(0)) = (0, 0)$.

Quite often we are only interested in either a maximum or minimum value of a function, and at times we do not need to have a closed, bounded interval. Consider the following example:

Example 5.2.6 Find the minimum value of $f(x) = 1 + x \ln x$.

Solution: Note that the domain is not specified, so we will look at the natural domain of $f(x)$, for which the natural logarithm present within the function requires $x \in (0, \infty)$.

Next we note that

$$f'(x) = x \cdot \frac{d}{dx} \ln x + (\ln x) \frac{dx}{dx} = x \cdot \frac{1}{x} + \ln x = 1 + \ln x.$$

Thus $f'(x) = 0 \iff \ln x = -1 \iff x = e^{-1}$.

We can then make a sign chart for $f'(x)$ if we like, though for this particular example we might instead note that

$$f''(x) = \frac{d}{dx} (1 + \ln x) = \frac{1}{x} > 0 \quad \text{for all } x > 0,$$

and so this function is concave up on its entire domain. From this we know that the one critical point, namely $(e^{-1}, f(e^{-1}))$, is a global minimum. The global minimum value of $f(x)$ is then

$$f(e^{-1}) = 1 + e \ln \left(\frac{1}{e}\right) = 1 - \frac{1}{e} = \frac{e - 1}{e}.$$

If we would like an approximation of the actual point, we can use $\left(\frac{1}{e}, \frac{e - 1}{e}\right) \approx (0.367879, 0.632121)$. 
The graph of the function \( f(x) = 1 + x \ln x \) is given at right. In a later chapter we will be able to show \( x \to 0^+ \implies 1 + x \ln x \to 1 \), but that observation is not necessary for the analysis above. Indeed it was enough to know \( f(x) \) is defined exactly for \( x > 0 \), is concave up everywhere it is defined, and has zero derivative at \( x = 1/e \) is enough to conclude that point is a global minimum.

5.2.3 Applications (Max/Min Problems)

There are many applications where we wish to maximize or minimize some quantity which is dependent upon another quantity in some functional relationship. In other words, there is some “output” \( A(x) \) depending upon an input \( x \), where we wish to choose \( x \) to maximize or minimize \( A(x) \). There is often some natural, or practical domain of values \( x \) allowed by the context. Once we translate the application’s parameters into the language of functions, we can find the desired extrema (or conclude one does not exist) based upon the techniques already employed in this section. Some “applications” do come from a mathematical setting, as in our first example below, while others are first set in a more “real world” context.

Example 5.2.7 Find the dimensions of the rectangle with the largest area with its base along the \( x \)-axis and which fits within the upper semicircle of \( x^2 + y^2 = 9 \).

Solution: It is useful to begin with an illustration:

We will take \( x \) to be the \( x \)-coordinate of the upper rightmost point of the rectangle, so \( x \) also represents half of the horizontal length of the rectangle. The total area is then given by \( A = 2xy \). We need \( A \) to be a function of one variable only, so that we can take a derivative and use the previous techniques. From the equation of the circle, \( y = \sqrt{9 - x^2} \), and so

\[
A = 2xy = 2x \sqrt{9 - x^2}.
\]

Note that for this scenario, we must assume \( x \in [0, 3] \). Clearly the endpoints \( x \in \{0, 3\} \) both yield zero area, and so neither will yield the global maximum. We thus look for critical points.

\[
A = 2x \sqrt{9 - x^2} \implies \frac{dA}{dx} = 2x \frac{d}{dx} \sqrt{9 - x^2} + \sqrt{9 - x^2} \frac{d}{dx} 2x = 2x \frac{1}{2 \sqrt{9 - x^2}} (-2x) + 2 \sqrt{9 - x^2} = \frac{-2x^2 + 2(9 - x^2)}{\sqrt{9 - x^2}} = \frac{2(9 - 2x^2)}{\sqrt{9 - x^2}}.
\]

From this we get \( dA/dx = 0 \iff 9 - 2x^2 = 0 \iff 9 = 2x^2 \iff x^2 = 9/2 \). Solutions of this equation are \( x = \pm 3/\sqrt{2} \), but in our model we have \( x \in [0, 3] \) and so \( x = 3/\sqrt{2} \approx 2.12 \) is
our relevant solution of $dA/dx = 0$. It is not difficult to see that $x = 3/\sqrt{2}$ must therefore yield the maximum value for $A$, since clearly $A(3/\sqrt{2}) > 0$ while the endpoints $x = 0, 3$ both yield $A = 0$. Since all we need to check are critical points and endpoints, $x = 3/\sqrt{2}$ must yield the maximum.

However that does not quite solve our original problem, which asked for the dimensions of the rectangle with maximum area. The dimensions could be of the form $2x \times y$, where $x = 2/\sqrt{3}$ and

$$y = \sqrt{9 - x^2} = \sqrt{9 - (3/\sqrt{2})^2} = \sqrt{9 - 9/2} = \sqrt{9/2} = 3/\sqrt{2}.$$  

So for this problem we get the maximum area occurs when $x = y = 3/\sqrt{2}$, but then the dimensions are $2 \cdot 3/\sqrt{2} \times 3/\sqrt{2}$, or $3\sqrt{2} \times 3/\sqrt{2} \approx 4.24 \times 2.12$.\footnote{Note that the rectangle we found was twice as wide (horizontally) as it is tall. This proportion will be the case regardless of the radius of the semicircle (see Exercise 1, page 486).}
Exercises

1. Show that when we replace the radius 3 in Example 5.2.7 (page 484) with a general radius \( R \), then the rectangle with the largest area will still be twice as wide as it is tall.
5.3. THE MEAN VALUE THEOREM

5.3.1 Linear Interpolation and Average Rate of Change

Suppose we are given a function $f : [a, b] \rightarrow \mathbb{R}$ which is continuous on $[a, b]$. We can connect the two points $(a, f(a))$ and $(b, f(b))$ by a line, which is itself a function we call the linear interpolation of $f$ between the two points. Graphically it is also called a secant line of the curve $y = f(x)$, meaning literally a line connecting two points on a circle, but generalized to mean a line connecting any two points on any given graph.

The equation of the linear interpolation is readily calculated. It has slope $(f(b) - f(a))/(b-a)$, and so when we note that $(a, f(a))$ is one point on the graph, we can quickly write $y = l(x)$,
where
\[ l(x) = f(a) + \left[ \frac{f(b) - f(a)}{b - a} \right] (x - a). \] (5.7)

Compare for instance to (4.8), The linear interpolation is useful if we wish to approximate \( f(x) \) for some \( x \) between or near \( a \) and \( b \), but only have knowledge of \( f \) evaluated at \( a \) and \( b \). In fact, we can also approximate a value for \( x \) given the desired output value for \( f(x) \) by solving for \( x \) in \( y = l(x) \). While the accuracy of the interpolation varies too much to be reliable without further information on the behavior of \( f \) on \((a, b)\), often there are tables of pairs \((x, y)\) for functions which behave somewhat linearly between such pairs, and linear interpolation (using \( y = l(x) \)) is a fairly standard procedure.

At this point we might again define the \textit{average rate of change of} \( f(x) \) \textit{over the interval} \([a, b]\) as we did briefly in Chapter 4, specifically (4.4), page 296. It is the net change in \( f \) over \([a, b]\), divided by the length of the interval, to give the average change in \( f(x) \) per unit length traversed in \( x \), namely
\[
\text{“average change in } f(x) \text{ over } [a, b]\text{”} = \frac{f(b) - f(a)}{b - a}. \] (5.8)

It is also the slope \( l' \) of the linear interpolation of \( f \) on the same interval, i.e., the slope of the line through \((a, f(a))\) and \((b, f(b))\). (The reader should verify \( l'(x) = (f(b) - f(a))/(b - a) \) by mentally differentiating (5.7).)

\section{5.3.2 Mean Value Theorem}

Here we give the statement of the Mean Value Theorem, and then look at an important Lemma known as Rolle’s Theorem, which we then use in the proof of the Mean Value Theorem. Recall that \( f \) being differentiable means that \( f' \) exists.

\begin{theorem}[Mean Value Theorem] Suppose \( f : [a, b] \rightarrow \mathbb{R} \) is continuous on \([a, b]\) and differentiable on \((a, b)\). Then there exists \( \xi \in (a, b) \) such that
\[ f'(\xi) = \frac{f(b) - f(a)}{b - a}. \] (5.9)
\end{theorem}

This is illustrated in Figure 5.5, page 487, in which there are two such values of \( \xi \), namely \( \xi_1 \) and \( \xi_2 \).

Before we begin to prove this, note that what this says is that there is a point between \( a \) and \( b \) on which the slope of the curve matches that of the linear interpolation from \((a, f(a))\) to \((b, f(b))\), i.e., matches the average rate of change of \( f(x) \) over \([a, b]\).

See Figure 5.5 for an illustration of this theorem. Note the line \( y = l(x) \), called the interpolation line connecting \((a, f(a))\) and \((b, f(b))\). It is also called a \textit{secant line} to the graph because it connects two points on the graph.

\footnote{Compare for instance to (4.8), page 298. Indeed, anytime we have two points \((x_1, y_1)\) and \((x_2, y_2)\), we have the slope being \((y_2 - y_1)/(x_2 - x_1)\), and if we take this and one of the points, say \((x_1, y_1)\), we get the equation of the line in the form
\[ \frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1} \implies y - y_1 = \left[ \frac{y_2 - y_1}{x_2 - x_1} \right] (x - x_1) \implies y = y_1 + \left[ \frac{y_2 - y_1}{x_2 - x_1} \right] (x - x_1). \] In (5.7) we have this, but with the two points being \((a, f(a))\) and \((b, f(b))\), instead of \((x_1, y_1)\) and \((x_2, y_2)\).}

\footnote{This idea of \textit{average rate of change} over an interval is perhaps most intuitively illustrated when we look at average velocities over intervals of time. It is not hard to see how \([s(t_f) - s(t_0)]/(t_f - t_0)\) represents the average velocity over the interval \([t_0, t_f]\) (where the particle is, minus where it was, all divided by how much time it took). With the variable and function names changed, this is just}
5.3. **THE MEAN VALUE THEOREM**

Now we turn to the proof, but in two stages. First we state and prove what is known as Rolle’s Theorem, which is perhaps a bit more intuitive than the Mean Value Theorem.

**Theorem 5.3.2 (Rolle’s Theorem)** *(Used to prove the Mean Value Theorem.)* Suppose that \( f : [a, b] \rightarrow \mathbb{R} \) is continuous on \([a, b]\) and differentiable on \((a, b)\). If \( f(a) = f(b) \), then there exists \( \xi \in (a, b) \) such that \( f'(\xi) = 0 \).

**Proof:** It is useful to consider two cases. First, that \( f(x) \) is constant, with the common value \( f(a) = f(b) \) on all of \([a, b]\), in which case \( f'(\xi) = 0 \) for all \( \xi \in (a, b) \), and the conclusion of the theorem holds true.

For the other case, we suppose \( f(x) \) is not a constant function on \([a, b]\), so it must achieve a maximum or minimum value for \( x \in (a, b) \) (other than \( f(a) = f(b) \)), at some point \( \xi \in (a, b) \). Thus \( \xi \) is a critical point of \( f(x) \), and since \( f'(\xi) \) exists, we must conclude \( f'(\xi) = 0 \), q.e.d.

We now finish the proof of the Mean Value Theorem, by applying Rolle’s Theorem to the difference between the function \( f(x) \) and its linear interpolation \( l(x) \) between \((a, f(a))\) and \((b, f(b))\).

**Proof:** (MVT) Suppose \( f(x) \) is continuous on \([a, b]\) and differentiable on \((a, b)\). Define \( l(x) \) to be the linear interpolation between \((a, f(a))\) and \((b, f(b))\) as in (5.7), page 488. Now consider the function

\[
g(x) = f(x) - l(x).
\]

We note that \( g(x) \) is continuous on \([a, b]\), differentiable on \((a, b)\). Furthermore, we note that \( g(a), g(b) = 0 \), so there exists some point in \( \xi \in (a, b) \) such that \( g'(\xi) = 0 \), implying \( f'(\xi) - l'(\xi) = 0 \), or \( f'(\xi) = l'(\xi) \). Now \( l'(\xi) = (f(b) - f(a))/(b - a) \), so

\[
f'(\xi) = l'(\xi) = \frac{f(b) - f(a)}{b - a}, \quad \text{q.e.d.}
\]

5.3.3 **Applications to Graphing Theorems**

Recall that we “observed” in Chapter 4 that \( f' > 0 \) on an interval implied \( f \) was increasing on that interval. Similarly \( f' < 0 \) on the interval showed \( f \) was decreasing on that interval. With the Mean Value Theorem at our disposal, we are now in a position to prove these.

**Theorem 5.3.3** Suppose \( a < b \), \( f(x) \) is continuous on \([a, b]\) and \( f'(x) \) exists on \((a, b)\).

1. If \( f'(x) > 0 \) on \([a, b]\), then \( f(x) \) is increasing on \([a, b]\).
2. If \( f'(x) < 0 \) on \([a, b]\), then \( f(x) \) is decreasing on \([a, b]\).

**Proof:** First we prove (1). Let \( a \leq x_1 < x_2 \leq b \). Then there exists \( \xi \in (x_1, x_2) \) such that

\[
\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\xi) > 0 \implies f(x_2) - f(x_1) = f'(\xi) \cdot (x_2 - x_1) > 0
\]

\[
\implies f(x_2) > f(x_1).
\]
This being true for all \( x_1, x_2 \in [a, b] \) where \( x_1 < x_2 \) shows \( f(x) \) is increasing on \([x_1, x_2]\). Case (2) is an easy modification of the argument to prove case (1):

\[
\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\xi) < 0 \implies f(x_2) - f(x_1) = f'(\xi) \cdot (x_2 - x_1) < 0
\]

\[
\implies f(x_2) < f(x_1).
\]

Another useful and intuitive theorem is the following.

**Theorem 5.3.4** Suppose \( f(x) \) is continuous on \([a, b]\) and differentiable on \((a, b)\). Moreover, suppose \( f'(x) = 0 \) for all \( x \in (a, b) \). Then \( f(x) \) is constant on \([a, b]\).

**Proof:** The proof follows a similar strategy as above. Let \( a \leq x_1 < x_2 \leq b \). Then there exists \( \xi \in (x_1, x_2) \) such that

\[
\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\xi) = 0 \implies f(x_2) - f(x_1) = 0 \implies f(x_2) = f(x_1).
\]

While the theorem itself is interesting, it seems obvious enough to be almost worthless. But in fact it has a very nice corollary, which is that if two functions have the same derivative on an interval, they must differ by a constant.

**Corollary 5.3.1** Suppose \( f(x) \) and \( g(x) \) are continuous on \([a, b]\), differentiable on \((a, b)\), and \( f'(x) = g'(x) \) on \((a, b)\). Then there exists \( C \in \mathbb{R} \) such that \( f(x) = g(x) + C \).

**Proof:** Consider the function \( h(x) = f(x) - g(x) \), which is also continuous on \([a, b]\), differentiable on \((a, b)\), and \( h'(x) = f'(x) - g'(x) = 0 \) on \((a, b)\), so there exists \( C \in \mathbb{R} \) such that, on \([a, b]\) we have

\[
h(x) = C \iff f(x) - g(x) = C \iff f(x) = g(x) + C.
\]

**Example 5.3.1** Note how, on an interval such as \((-\pi/2, \pi/2)\) we have

\[
\frac{d}{dx} \tan^2 x = 2 \tan x \cdot \frac{d}{dx} \tan x = 2 \tan x \sec^2 x = 2 \sec^2 x \tan^2 x,
\]

\[
\frac{d}{dx} \sec^2 x = 2 \sec x \cdot \frac{d}{dx} \sec x = 2 \sec x \sec x \tan x = 2 \sec^2 x \tan x.
\]

Thus \( \frac{d}{dx} \tan^2 x = \frac{d}{dx} \sec^2 x \) on \((-\pi/2, \pi/2)\), leading us to conclude \( \tan^2 x = \sec^2 x + C \). Indeed, with \( C = -1 \) we have one of our basic trigonometric identities, namely \( \tan^2 x = \sec^2 x - 1 \). (Of course, this identity actually holds true anywhere \( \cos x \neq 0 \).)

Graphically, two functions with the same derivatives on an interval will be the same shape. One will just be a vertical translation of the other.
5.3. Numerical Applications

The existence of the value $\xi$ with the prescribed derivative (equal to the average rate of change) can imply many things. Knowing bounds for the derivative over an interval means we know bounds for $f'(\xi)$, and thus the average rate of change. Consider the following example.

**Example 5.3.2** Suppose $f(x)$ is continuous on $[2, 10]$, differentiable on $(2, 10)$, and $f'(x) > 3$ on $(2, 10)$. If $f(2) = 5$, find a lower bound for $f(10)$.

**Solution:** There exists $\xi \in (2, 10)$ such that $f'(\xi) = (f(10) - f(2))(10 - 2)$. Now
\[
\frac{f(10) - f(2)}{10 - 2} = f'(\xi) > 3 \implies f(10) - f(2) = f'(\xi)(10 - 2) > 3(8) = 24
\implies f(10) > f(2) + 24 = 5 + 24 = 29.

We conclude that $f(10) > 29$.

Note that it was important that we multiplied the first equation by $10 - 2 > 0$.

The example is akin to asking, if $s = 5$ ft when $t = 2$ sec, and $v > 3$ ft/sec at all times $t \in (2\text{sec}, 10\text{sec})$, then where must $s(10\text{sec})$ be? If our starting position is at 5 ft, and our velocity is greater than 3 ft/sec for a time interval of length 8 sec, then we must have traveled more than 24 ft in that time, leaving us to the right of the position marked 29 ft.

In fact this method works for more general problems.

**Example 5.3.3** Supposing $f(x)$ is continuous on $[3, 15]$, and differentiable on $(3, 15)$ in such a way that $-2 \leq f'(x) < 4$ on that interval, and $f(3) = 6$, find an interval containing $f(15)$.

**Solution:** According to the Mean Value Theorem, there exists $\xi \in (3, 15)$ such that $f(\xi) = (f(15) - f(3))/(15 - 3)$. We can use this to find bounds for $f(15)$ as follows:
\[
\frac{f(15) - f(3)}{15 - 3} = f(\xi) \in [-2, 4] \implies -2 \leq \frac{f(15) - f(3)}{15 - 3} < 4
\implies -2 \leq \frac{f(15) - 6}{12} < 4
\implies -12 \leq f(15) - 6 < 48
\implies -6 \leq f(15) < 54.

In fact, one can use a geometric interpretation of $f'$ to construct lines of the slopes given by bounds on $f'(x)$ for $3 \leq x \leq 15$, and the values of $f(x)$ will fall between these lines. This is illustrated in Figure 5.6.
Figure 5.6: Illustration of the situation in Example 5.3.3. There, \( f'(x) \in [-2, 4], \) \( f(3) = 6, \) and bounds for \( f(15) \) are desired. Geometrically, the function is bounded by the two lines shown above, throughout all of \( 3 < x \leq 15. \) The final bounds of \( f(x) \) at \( x = 15 \) are labeled above. The upper line represents the “extreme” case \( f'(x) = 4 \) throughout \([3, 15]\), and the lower line the other “extreme” case where \( f'(x) = -2 \) on throughout all of \([3, 15]\).
5.4 Differentials and the Linear Approximation Method

The main idea of this section is to show how the tangent at \((a, f(a))\) can be used to approximate the function \(f(x)\) near \(x = a\). But because differential calculus (calculus of derivatives) is not just about tangent lines, we will first show intuitively how the same idea is reasonable in other contexts. Finally, this will give yet another chance to interpret the Leibniz notation, this time giving \(dx\) and \(df(x)\) numerical significance, consistent with but beyond the meaning we gave these before, when we only considered these symbols when put together formally in a fraction \(\frac{df(x)}{dx} = f'(x)\).

We will begin this section with some intuitive examples to motivate the technique. We will then show how these are specific cases of linear approximations. At that point we will introduce differentials to show their geometric meaning.

5.4.1 Some Approximation Problems

The following examples should be intuitive, and eventually we will see that the underlying idea in each is the same. We will discuss this principle after we introduce the examples.

Example 5.4.1 Suppose a car travels along a highway, and a passenger notices the speedometer reads 72 miles per hour. About how far will the car travel in the next second?

Solution: We have very little information here, but a reasonable assumption to make is that the car will not change speed very much in that one second. To the extent that that is true, we can say the distance traveled is given approximately by the following:

\[
\text{Distance} \approx \frac{72 \text{ mile}}{\text{hour}} \cdot 1 \text{ second} = \frac{72 \text{ mile}}{1 \text{ hour}} \cdot \frac{5280 \text{ foot}}{1 \text{ mile}} \cdot \frac{1 \text{ hour}}{3600 \text{ second}} \cdot 1 \text{ second} = 105.6 \text{ foot}.
\]

The idea in the above example is akin to the grade school formula \((\text{distance}) = (\text{rate}) \cdot (\text{time})\), but we made an assumption here that the velocity would be approximately constant over that one second. The first form of the answer is technically consistent, but unit conversions were included to make the answer more intuitive. Still, the gist of the method was to assume that the rate of position change (72 mile/hour), at the time that rate was sampled, would be approximately the rate for the entire one-second time interval in question.\(^{10}\)

Example 5.4.2 Suppose a manufacturer’s research shows that the profit from making \(x\) of a particular item should be

\[
P(x) = -0.004x^3 + 10x^2 - 1000.
\]

Suppose further that the manufacturer is initially planning on a production run of 100 items. How much more profit would he make if he produced 101 items instead?

\(^{10}\)The difference between interpolation and approximation is subtle but important. Interpolation refers to taking some data and using it to approximate where other points may lie, particularly but not limited to those in between the given data.

Interpolation is thus a kind of approximation, but there are others. For instance, in this Section 5.4 we instead use a function’s output and derivative at a given input, and assume the derivative is approximately constant to approximate function values elsewhere. So we are not “connecting” data points—as we do when we interpolate—to predict others, but look at the function’s properties at a single point to predict its values at others. There is an entire literature on the topic of approximation theory, and here we have two basic examples: linear interpolation from two data, and linear approximation from one datum, but including information about the derivative there.
Solution 1: If the model is correct, the actual extra profit from making that 101st item would be the difference in profit from making 101 items and the profit from making 100 items:

\[ P(101) - P(100) = [-0.004(101)^3 + 10(101)^2 - 1000] - [-0.004(100)^3 + 10(100)^2 - 1000] \]
\[ = 96888.796 - 95000 \]
\[ = 1888.796. \]

Of course this should be rounded to hundredths of a dollar (that is, to the nearest cent), so according to this model the manufacturer would make an extra \$1888.80 from that 101st item.

Of course, the model itself is likely an approximation based upon research. The apparent precision of the expected extra profit is open to further scrutiny (as with any economic model). In any event, another approach, this time definitely an approximation, is given below:

Solution 2: Note that

\[ P'(x) = \frac{d}{dx} [-0.004x^3 + 10x^2 - 1000] \]
\[ = -0.012x^2 + 20x. \]

Furthermore, \( P'(100) = -0.012(100)^2 + 20(100) = 1880 \) (dollars/item). In other words, the profit is changing at \$1880/item when the number of items is 100. We can use this to approximate that the next (101st) item will cause a growth in the profit of approximately \$1880.

In the above example, note that the second method offered a very good approximation for the difference \( P(101) - P(100) \), by considering how quickly \( P(x) \) was changing with \( x \) (quantified by \( P'(x) \)) when \( x = 100 \), and—assuming that \( P(x) \) continued to change at approximately that rate for \( x \) near 100—used that to approximate the actual change in the value of \( P(x) \) as \( x \) changes from 100 to 101. This is in the same spirit as in the previous example (Example 5.4.1). Below are two reasons why one may wish to use the approximating method of Solution 2 instead of the more exact, Solution 1 method\(^1\):

- It was easier (and faster) to compute \( P'(100) \) item \( \cdot \) (1 item) than to compute the difference \( P(101 \) item \) \( - \) \( P(100 \) item \). The former (approximation) was a second-degree polynomial with two terms, while the latter (and actual) was a third-degree polynomial with three terms, evaluated twice.

- The original model was only an approximation, so this approximation to an approximation might not have lost too much confidence in accuracy to be useful.

Note that since we have an approximation for the difference \( P(101) - P(100) \approx P'(100) \cdot 1 \), we can write this as an approximation for \( P(101) \), rather than the difference:

\[ P(101) \approx P(100) + P'(100)(1), \]

where the trailing 1 represents how many units away from \( x = 100 \) we traveled to get to \( x = 101 \). If we strayed too far from \( x = 100 \), the profit change may stray far from \( P'(100) \), and our approximation will be less accurate.

In the next example we will use a similar strategy which again lets us use a simple function to approximate a computationally more complicated one.

\(^{11}\)Of course the reasons given here are usually balanced by fact that, if the exact answer is easily available, it may be far preferable to an approximation. We will discuss this further as we progress through this section.
Example 5.4.3 Suppose a laser at ground level points to the base of a building 300 feet away. If the laser beam is then turned so that it still points to the building, but with an angle of elevation of $5^\circ$, then approximately how high on the building is the point illuminated by the beam?

Solution: The height $h$ on the building is a function of $\theta$, the angle of elevation of the beam, that is, the angle formed by the beam and the horizontal. Since $h/(300 \text{ foot}) = \tan \theta$, we can write

$$h(\theta) = 300 \text{ foot} \cdot \tan \theta. \tag{5.10}$$

If we use radian measure for $\theta$, then

$$h'(\theta) = 300 \text{ foot} \sec^2 \theta.$$

For $\theta = 0$, we have $h'(0) = 300 \text{ foot}$. It is instructive to note that the units of $dh(t)/dt$ are formally foot/radian (though we often omit the ultimately dimensionless unit of radian). With this in mind, we can note that $5^\circ = 5^\circ \cdot \frac{\pi (\text{rad})}{180^\circ} = \frac{\pi}{36} \text{ (radian)}$, and so

$$h \left( \frac{\pi}{36} \right) \approx h(0) + h'(0) \frac{\pi}{36} \text{ (radian)}$$

$$= 0 + 300 \text{ foot} \cdot \frac{\pi}{36} \text{ (radian)}$$

$$= \frac{300\pi}{36} \text{ foot}$$

$$\approx 26.18 \text{ foot}.$$

The actual height of the laser beam is $300 \text{ foot} \cdot \tan 5^\circ \approx 26.25 \text{ foot}$.

If we look at the above example closely, we see that replacing $5^\circ$ with any angle $\theta$ in radian measure, then we can claim the approximation

$$h(\theta) \approx 300 \text{ foot} \cdot \theta. \tag{5.11}$$

From the example we see that this approximation is very good for $\theta = 5^\circ = \pi/36 \text{ (radian)}$. In fact it compares well even for larger $\theta$, but certainly not for all $\theta$. Clearly from (5.10), $h(\theta) = 300 \text{ ft} \cdot \tan \theta \rightarrow \infty$ as $\theta \rightarrow \frac{\pi}{2}^-$, which does not happen with our approximation (5.11). What happened is that the rate of change $dh/d\theta = 300 \text{ sec}^2 \theta$ did not stay at all constant, and in fact blew up also as $\theta \rightarrow \frac{\pi}{2}^-$.  

5.4.2 Linear Approximation

In all these examples, we approximated a “future” measurement of a function based upon its presently known value at a particular point, and how fast it was changing at that point. For a function $f(x)$, the known value was at some $x = a$, where we had data on $f(a)$ and on how fast $f(x)$ changes with respect to $x$ at $x = a$; that is, we knew $f(a)$ and $f'(a)$. Based upon these, we could find an approximation of $f(a + \Delta x)$ by thinking of $\Delta x$ as a “run,” with $f(a + \Delta x) - f(a)$ being the resulting “rise.” Hence $(f(a + \Delta x) - f(a))/\Delta x$ is “rise/run,” which is assumed to be approximately $f'(a)$, which—upon multiplying by $\Delta x$ gives us:

$$f(a + \Delta x) - f(a) \approx f'(a) \Delta x.$$  

Solving for $f(a + \Delta x)$, we would have

$$f(a + \Delta x) \approx f(a) + f'(a) \Delta x. \tag{5.12}$$

For a more useful formula, we now let $x = a + \Delta x$, so that $\Delta x = x - a$, and then (5.12) gives us the following:
Definition 5.4.1 For a function \( f(x) \), and a real number \( a \) for which \( f(a) \), \( f'(a) \) exist, the linear approximation of \( f(x) \) at \( a \) is given by

\[
f(x) \approx f(a) + f'(a)(x - a).
\]

Note that the right-hand side of (5.13) is exactly the expression for the tangent line to \( y = f(x) \) at the point \( x = a \), given by (4.8), that is, \( y = f(a) + f'(a)(x - a) \). We can write (5.13) in a colloquial way as follows: Where is \( f \) at \( x \)? The approximate answer is, where it was at \( a \), plus how fast it was changing at \( a \) multiplied by how far we traveled (+/-) from \( a \).

Example 5.4.4 Use the linear approximation of \( f(x) = \sqrt[3]{x} \) to approximate \( \sqrt[3]{8.5} \).

**Solution:** Here \( f(x) = x^{1/3} \), and \( f'(x) = \frac{1}{3}x^{-2/3} \). Using \( a = 8 \) we get

\[
f(8) = 8^{1/3} = 2,
\]

\[
f'(8) = \frac{1}{3}(8)^{-2/3} = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12}.
\]

Thus, for \( x \) near 8, we have

\[
f(x) \approx f(8) + f'(8)(x - 8), \text{ i.e.,}
\]

\[
f(x) \approx 2 + \frac{1}{12}(x - 8).
\]

Using this we get

\[
\sqrt[3]{8.5} = f(8.5)
\]

\[
\approx 2 + \frac{1}{12}(8.5 - 8)
\]

\[
= 2 + \frac{1}{12} \cdot \frac{1}{2}
\]

\[
= 2 + \frac{1}{24} = 49/24 = 2.041666666 \ldots.
\]

Thus \( \sqrt[3]{8.5} \approx 2.0417 \).

The actual value of \( \sqrt[3]{8.5} \) is approximately 2.04082755, so our linear approximation is accurate to three significant digits when \( x = 8.5 \). Any linear approximation is a statement regarding \( x \) close to the point \( x = a \) (at which the linear approximation is just the tangent line to the graph). Such an approximation is likely to worsen in accuracy as we leave the immediate vicinity of \( x = a \), although the degree to which this happens depends upon the function—in particular, how closely the graph follows the trend of the tangent line at \( x = a \), as we move away from \( x = a \).

\( ^{12} \)That the approximation technique should give rise to the tangent line should not be surprising, since \( f'(a) \) measures the (instantaneous) rate of change of \( y \) with respect to \( x \), as well as the slope of the tangent line at \( x = a \). In our examples, we used respectively how position changed with time, how profit changed with the number of items produced, and how one leg of a right triangle changed with its opposite angle. The connection to the derivative was apparent in each, and the derivative is geometrically the slope of the curve, which also defines the slope of the tangent line.
Below is a list of values of $\sqrt[3]{x}$ as approximated by this method, and than as computed directly (first 8 digits shown).

\[
\begin{align*}
  f(8) & \approx 2 & f(8) &= 2 \\
  f(8.5) & \approx 2.0416667 & f(8.5) &= 2.0408275 \ldots \\
  f(9) & \approx 2.0833333 & f(9) &= 2.0800838 \ldots \\
  f(10) & \approx 2.1666667 & f(10) &= 2.1544346 \ldots \\
  f(11) & \approx 2.2500000 & f(11) &= 2.2239800 \ldots \\
  f(12) & \approx 2.3333333 & f(12) &= 2.2894284 \ldots \\
  f(13) & \approx 2.4166667 & f(13) &= 2.3513346 \ldots \\
  & \vdots & & \vdots \\
  f(20) & \approx 3.0000000 & f(20) &= 2.7144176 \ldots \\
  f(30) & \approx 3.8333333 & f(30) &= 3.1072325 \ldots \\
  & \vdots & & \vdots \\
  f(64) & \approx 6.6666667 & f(64) &= 4 \cdots \\
  f(1000) & \approx 84.6666667 & f(1000) &= 10.
\end{align*}
\]

We see that the approximation based upon the behavior at $x = 2$ (i.e., the linear approximation at $x = 2$) stays reasonably close to the actual values of $f(x)$ until we stray far from $x = 8$.\footnote{Of course, “close” and “far” are subjective measures of proximity, and acceptable tolerances differ from context to context.} The actual graph of $f(x) = \sqrt[3]{x}$, together with the tangent line emanating from $(8, f(8))$ are graphed in Figure 5.7, page 498.

Again looking at Figure 5.7, we see that the tangent line at $x = 2$ does stay somewhat close to the curve for a while as $x$ increases past $x = 2$. However, we see a different behavior as $x$ moves left of $x = 2$. Consider the following comparisons of the linear approximation and actual value of $f(x)$ for $x < 2$:

\[
\begin{align*}
  f(7) & \approx 1.9166667 & f(7) &= 1.9129311 \cdots \\
  f(6) & \approx 1.8333333 & f(6) &= 1.8171205 \cdots \\
  f(5) & \approx 1.75 & f(5) &= 1.7099759 \cdots \\
  f(4) & \approx 1.6666667 & f(4) &= 1.5874010 \cdots \\
  f(3) & \approx 1.5833333 & f(3) &= 1.4422495 \cdots \\
  f(2) & \approx 1.5 & f(2) &= 1.2599210 \cdots \\
  f(1) & \approx 1.4166667 & f(1) &= 1 \\
  f(0) & \approx 1.3333333 & f(0) &= 0 \\
  & \vdots & & \vdots \\
  f(-8) & \approx 0.6666667 & f(-8) &= -2.
\end{align*}
\]

As we can see from the graph in Figure 5.7, the function and the tangent line are quite close when $|x - 8|$ is small, but is unreliable as we stray from $x = 8$.

One of the most useful linear approximations in physics is used to approximate $\sin x$ for small $|x|$, i.e., when $|x - 0|$ is small, i.e., when $x$ is near zero:
Figure 5.7: Partial graph of \( f(x) = \sqrt{x} \), along with the linear approximation (tangent line) at \( x = 8 \). The two graphs are very close to each other near \( x = 8 \) (and coincide at \( x = 8 \)), but part company as we stray farther from \( x = 8 \). They may (and in fact do) eventually come together again, but that is coincidence, while the approximation is known to be useful near \( x = 8 \) by the nature of the tangent line.
5.4. DIFFERENTIALS AND THE LINEAR APPROXIMATION METHOD

14Note that $|x| < 1$ in radians corresponds to, approximately $|x| < 57^\circ$, but to use this to approximate $\sin 48^\circ$, for instance, we need to convert back to radians:

$$
\sin 48^\circ = \sin \frac{48^\circ \cdot \pi}{180^\circ} = \sin \frac{48\pi}{180} \approx \frac{48\pi}{180} \approx 0.837758041.
$$

The actual value of $\sin 48^\circ$ is 0.7431448, when rounded to seven places.

Example 5.4.6 Consider the curve $y^2 - x^2 = 9$. Approximate $y$ as a function of $x$ near $(4, -5)$ using a linear approximation.

Figure 5.8: Partial graph of $f(x) = \sin x$, and the linear approximation at $x = 0$, which is $y = x$. Though not clear from the printed resolution here, the functions only coincide at $x = 0$. The proof of that fact is left as an exercise.

Example 5.4.5 Find the linear approximation for $f(x) = \sin x$ at $x = 0$.

Solution: We will use the formula $f(x) \approx f(a) + f'(a)(x - a)$ with $a = 0$ and $f(x) = \sin x$. Now

$$f(0) = \sin 0 = 0,$$
$$f'(0) = \cos 0 = 1.$$

With this data, (5.13) becomes

$$f(x) \approx f(0) + f'(0)(x - 0) = 0 + 1(x - 0), \text{ i.e.,}$$
$$f(x) \approx x.$$

The graphs of $y = \sin x$ and its linear approximation at $x = 0$, namely $y = x$, are given in Figure 5.8. The approximation is very good for $|x| < 1$.14

5.4.3 Linear Approximations and Implicit Functions

Because we can find $\frac{dy}{dx}$ on implicit curves written as equations, we can find tangent lines and therefore linear approximations. In such a case it is not for $y$ as a function of $x$, but rather for $y$ as a local function on $x$. Still the method is valid. We offer two examples here.

Example 5.4.6 Consider the curve $y^2 - x^2 = 9$. Approximate $y$ as a function of $x$ near $(4, -5)$ using a linear approximation.
CHAPTER 5. USING DERIVATIVES TO ANALYZE FUNCTIONS; APPLICATIONS

Figure 5.9: Graph for Example 5.4.6, showing a partial graph of the curve $y^2 - x^2 = 9$ and the tangent line at $(4, -5)$, which is also the linear approximation to the local (implicit) function defined near there.

Solution: We use the usual implicit differentiation technique as in Section 4.6:

$$y^2 - x^2 = 9$$
$$\Rightarrow \quad \frac{d}{dx}[y^2 - x^2] = \frac{d}{dx}[9]$$
$$\Rightarrow \quad 2y \cdot \frac{dy}{dx} - 2x = 0$$
$$\Rightarrow \quad 2y \frac{dy}{dx} = 2x$$
$$\Rightarrow \quad \frac{dy}{dx} = \frac{2x}{2y} = \frac{x}{y}$$

For the point $(4, -5)$—which a quick check shows is on the curve—we have slope

$$\left. \frac{dy}{dx} \right|_{(4,-5)} = \frac{x}{y} \bigg|_{(4,-5)} = \frac{4}{-5} = -\frac{4}{5}.$$  

The tangent line is given by

$$y = -5 - \frac{4}{5}(x - 4).$$

Thus, on the curve near $(4, -5)$, we have $y \approx -5 - \frac{4}{5}(x - 4)$.

This graph and the linear approximation at $(4, -5)$ is given in Figure 5.9.

Example 5.4.7 Recall that in Example 4.6.5, on page 390 we had the implicit curve given by $5x + x^2 + y^2 + xy = \tan y$, and we found

$$\frac{dy}{dx} = \frac{5 + 2x + y}{\sec^2 y - 2y - x},$$

and the tangent line to the curve at $(0, 0)$ had slope 5. Thus near $(0, 0)$, we can say that the local function is given by $y \approx 0 + 5(x - 0)$, or $y \approx 5x$, which is a huge advantage over trying to find an actual $y$ for a given $x$ near $x = 0$. 
5.4. DIFFERENTIALS AND THE LINEAR APPROXIMATION METHOD

5.4.4 Differentials

Below we define differentials. Eventually we will give numerical and geometric meaning to all of the terms in the definition, but first we define them only formally:

**Definition 5.4.2** Given a function \( f(x) \), the differential of \( f(x) \) is given by

\[
 df(x) = f'(x) \, dx, \tag{5.14}
\]

\( dx \) being the differential of \( x \), and where the prime, \( ' \) represents that the derivative is taken with respect to the underlying variable, which here is \( x \).

This is consistent with our previous use of Leibniz notation:

\[
 \frac{dx}{dx} = 1 \quad \text{and} \quad \frac{df(x)}{dx} = f'(x) \quad \iff \quad df(x) = f'(x) \, dx, \tag{5.15}
\]

Now we look at some quick computations which follow from this definition:

- \( dx = d(x) = (x)' \, dx = 1 \cdot dx = dx \), as we would hope, and
- \( \frac{df(x)}{dx} = f'(x) \quad \iff \quad df(x) = f'(x) \, dx \), at least formally, where we (again formally) multiplied both sides by \( dx \).

All of these become old-fashioned derivative problems if we divide these equations by \( dx \). We can fashion differential versions of all of our rules by taking the derivative rules and multiplying both sides by \( dx \). In particular, there are product, quotient and chain rules:

\[
 d(uv) = u \, dv + v \, du, \tag{5.15}
\]

\[
 d \left[ \frac{u}{v} \right] = \frac{v \, du - u \, dv}{v^2}, \tag{5.16}
\]

\[
 df(u(x)) = f'(u(x))u'(x) \, dx. \tag{5.17}
\]

With these we can compute many more differentials directly:

- \( d(x \tan x) = x \, d \tan x + \tan x \cdot dx = x \sec^2 x \, dx + \tan x \, dx \), which is exactly the result of our efforts just above to compute this differential.
- \( d \sin x^2 = \cos x^2 \cdot 2x \, dx = 2x \cos x^2 \, dx \).

In fact notice that (5.17) is completely consistent with the idea that \( df(u) = f'(u) \, du \):
By definition: $df(u) = f'(u) \, du$.

If it happens that $u$ is a function of $x$, i.e., $u = u(x)$, then

$$df(u) = df(u(x)) = \frac{df(u(x))}{dx} \, dx = \frac{df(u(x))}{du} \cdot \frac{du(x)}{dx} \, dx = f'(u(x)) \, u'(x) \, dx = f'(u) \, du.$$

We now have two methods for computing quantities such as $\frac{d \sin x^2}{dx^2}$:

1. considering $x^2$ as a variable in its own right:

$$\frac{d \sin x^2}{x^2} = \cos x^2;$$

2. using the definition of differentials (Definition 5.4.2, page 501):

$$\frac{d \sin x^2}{x^2} = \frac{(\sin x^2)' \, dx}{(x^2)' \, dx} = \frac{\cos x^2 \cdot 2x \, dx}{2x \, dx} = \cos x^2.$$

This again shows the robustness of the Leibniz notation, with differentials as well as derivatives. Computations of differentials become ubiquitous as we develop integration techniques in later sections.

So far we have only looked at these differentials formally. Now we will emphasize that these differentials can in fact be interpreted numerically.

Recall that $\frac{df(x)}{dx}$ measures how $f(x)$ changes as $x$ changes. More precisely, the fraction $\frac{df(x)}{dx}$ gives us the instantaneous rate of change in $f(x)$ as $x$ changes, at a particular value of $x$. (This is akin to $\frac{ds(t)}{dt}$ measuring velocity—that is, how position $s(t)$ is changing as $t$ is changing—at a particular value of $t$.) Note also that $\frac{df(x)}{dx}$ can be interpreted as the slope of the graph of $f(x)$ at the particular value $x$. Hence $\frac{df(x)}{dx}$ represents an instantaneous “rise/run.”

Now we will let $dx$ represent a “run,” i.e., a change in $x$ from a fixed value of $x$, while $df(x)$ will be the resultant “rise,” at the rate of $\frac{df(x)}{dx}$, i.e., $df(x)$ will represent the “rise” along the tangent line which ran through $(x, f(x))$. This is illustrated in Figure 5.10.

The same justification for using linear approximations for functions allows us to use $df(x)$ to approximate an actual change in the function $f(x)$, as we perturb the $x$-value by a small quantity $dx$. In fact, the actual change, the linear approximation, and the differentials are all related. If we call the perturbation in $x$ by both names $\Delta x$ and $dx$, we get:

$$f(x + \Delta x) = f(x) + \frac{f(x + \Delta x) - f(x)}{\Delta x} \cdot \Delta x$$

$$\approx f(x) + f'(x) \Delta x$$

$$= f(x) + \frac{df(x)}{dx} \cdot dx$$

$$= f(x) + df(x).$$

This reflects exactly what occurs with the linear approximation:

$$f(x) \approx f(a) + f'(a)(x - a),$$

$$\left. \frac{df(x)}{dx} \right|_{x=a} \cdot dx.$$
Figure 5.10: Illustration of the geometric meaning of $\frac{df(x)}{dx}$, giving further, geometric meaning to both $dx$ and $df(x)$. For any value $x$ in the domain, if $f'(x)$ is defined, we have a tangent line of slope $f'(x) = \frac{df(x)}{dx}$. Algebraically we could then write $df(x) = f'(x)dx$, where $dx$ can represent a “run” and $df(x)$ the resultant “rise” along the tangent line. Note that $df(x)$ is an approximation of the actual “rise” $\Delta f(x)$ (not necessarily positive) of the function, as the input changes from $x$ to $x + dx = x + \Delta x$. Moreover that tangent line’s rise $df(x)$ is the same as what is given by the linear approximation, as in (5.19).

where the part of $x + \Delta x$ in (5.18) is played by $x$ in (5.19), the part of $x$ is played by $a$, and $\Delta x$ is played by $(x - a)$. Furthermore, $f'(a)(x - a)$ represents $df(x)$ evaluated at $x = a$ with $dx = (x - a)$, so the right-hand side of (5.19) $f(a)$ plus the approximate perturbation of $f(x)$ from $f(a)$ as $x$ strays from $a$. 
5.5 Newton’s Approximation Method

While there are many types of equations we can solve algebraically—and students typically spend much time in their young years learning to solve them—there are numerous equations we cannot solve algebraically. We can often demonstrate that there exist solutions, and perhaps even discover the number of solutions, without actually knowing the exact solutions. Within industry approximate solutions often suffice, though the level of precision needed varies from problem to problem. When the problem is very complex, or an approximate solution is needed quickly, efficiency becomes an important consideration. Efficiency, accuracy, and prospects for success are all factors when choosing a method, and what works well for one problem might not work well—or even at all—for another.

In this section we will look at an approximating scheme attributed to Newton, though the modern interpretation has been ascribed to Thomas Simpson (of Simpson’s Rule, Section 8.3). We will also examine other methods for comparison.

5.5.1 Approximating Methods Applied to an Example

Consider the equation

\[ x = \cos x. \]  \hspace{1cm} (5.20)

We can tell from the first figure of Figure 5.11 that there is a solution of this equation somewhere in \( x \in [0, 2] \). In fact, by the Intermediate Value Theorem (page 196) we can look at

\[ f(x) = x - \cos x \]  \hspace{1cm} (5.21)

and note that \( f(x) \) is continuous everywhere, \( f(0) = 0 - 1 = -1 < 0 \), while \( f(\pi/2) = \pi/2 - 0 = \pi/2 > 0 \), so there is a solution of \( f(x) = 0 \) in \( x \in [0, \pi/2] \), meaning there is a solution of

\[ x = \cos x \quad \iff \quad x - \cos x = 0 \]

Figure 5.11: Figure showing the existence of a solution of \( x = \cos x \) which occurs where the graphs of \( y = x \) and \( y = \cos x \) intersect, or equivalently a solution of \( f(x) = 0 \) where \( f(x) = x - \cos x \).
5.5. **NEWTON’S APPROXIMATION METHOD**

$x = \cos x$ there as well. Rather than solving the original equation (5.20), we will instead solve the equivalent equation (5.21). To do so, we will consider three methods in turn.

**Solve by graphing:** If we have access to a device such as a graphing calculator, we can make successive magnifications of the window near the apparent solution, and thereby “zoom in” to the solution repeatedly, in the meantime reading off better and better approximations from the provided axes or curve tracing features of the device. While visually appealing, it is not an easy process to train a primitive device to implement.

**Method of Bisection:** This method is based on the Intermediate Value Theorem, where we check the sign of $f(x)$ at various points, and once we detect a sign change on an interval, we check the sign of $f(x)$ at the midpoint of the interval to see which subinterval (left or right) contained the sign change, and then we continue, each time halving the length of the interval with the sign change. In this way we can ultimately get as close to the actual value as we wish.

For instance, if we wish to attempt this method with our current problem, we would compute

\[
\begin{align*}
  f(0) & = -1 < 0 \\
  f(2) & = 2 - \cos 2 \approx 1.090702573 > 0 \\
\end{align*}
\]

(sign change in $[0, 2]$).

Next we consider the value of $f(x)$ at the midpoint of $[0, 2]$, and include the previous values for clarity:

\[
\begin{align*}
  f(0) & = -1 < 0 \\
  f(1) & = 1 - \cos 1 \approx 0.4596796941 > 0 \\
  f(2) & = 2 - \cos 2 \approx 1.090702573 > 0 \\
\end{align*}
\]

(sign change in $[0, 1]$).

Next we compute $f(x)$ at the midpoint of $[0, 1]$, again including the two endpoints for clarity:

\[
\begin{align*}
  f(0) & = -1 < 0 \\
  f(0.5) & = 0.5 - \cos 0.5 \approx -0.3775825619 < 0 \\
  f(1) & = 1 - \cos 1 \approx 0.4596796941 > 0 \\
\end{align*}
\]

(sign change in $[0.5, 1]$).

Next we compute $f(x)$ and the midpoint of $[0.5, 1]$ and compare it to the endpoints:

\[
\begin{align*}
  f(0.5) & = 0.5 - \cos 0.5 \approx -0.3775825619 < 0 \\
  f(0.75) & = 0.75 - \cos 0.75 \approx 0.0183111311 > 0, \\
  f(1) & = 1 - \cos 1 \approx 0.4596796941 > 0 \\
\end{align*}
\]

(sign change in $[0.5, 0.75]$).

Next we compute $f(x)$ and the midpoint of $[0.5, 0.75]$, namely $\frac{1}{2}(0.5 + 0.75) = 0.625$:

\[
\begin{align*}
  f(0.5) & = 0.5 - \cos 0.5 \approx -0.3775825619 < 0 \\
  f(0.625) & = 0.625 - \cos 0.625 \approx -0.1859631195 < 0 \\
  f(0.75) & = 0.75 - \cos 0.75 \approx 0.0183111311 > 0 \\
\end{align*}
\]

(sign change in $[0.625, 0.75]$).

This tells us the solution is somewhere in $[0.625, 0.75]$, whose midpoint is $\frac{1}{2}(0.625 + 0.75) = 0.6875$, and so on.

If we take $x_0 = 0$, $x_1 = 2$, and then consider the midpoints of the relevant intervals generated by the sign changes to be $x_2 = 1$, $x_3 = 0.5$, $x_4 = 0.75$, $x_5 = 0.625$, and so on, we see a sequence of $x$-values which get progressively closer to the actual solution. A sign chart for the $f(x_n)$ values helps to illustrate this phenomenon:
While this method can be easily programmed into a computer or similar device, it is not usually the most efficient computationally, meaning it takes more steps to get the same accuracy as the next method, which takes some advantage of the actual slope of the function, i.e., the function’s derivative.

**Newton’s Method:** With Newton’s method, we begin with one guess, which we will call \(x_1\).

We then look at the tangent line to \(f(x)\) at \((x_1, f(x_1))\), the slope of this line being \(f'(x_1)\).

We then take \(x_2\) to be the \(x\)-intercept of this tangent line (which would be a solution if the curve’s slope were constant!), in essence “following” the tangent line to the \(x\)-axis to find \(x_2\). We then take the tangent line at \((x_2, f(x_2))\) and follow it to the \(x\)-axis to find \(x_3\), and follow the tangent line at \((x_3, f(x_3))\) to the \(x\)-axis to find \(x_4\), and so on, to produce a sequence \(x_1, x_2, x_3, \ldots\), which will often converge very quickly towards a point \(x\) which solves \(f(x) = 0\). This is illustrated for our example above, namely solving \(x = \cos x\), i.e., \(f(x) = x - \cos x = 0\) using \(x_1 = 0\), and the \(x_n\) following for \(n = 2, 3, 4, \ldots\). This is illustrated in Figure 5.12.

Algebraically, for Newton’s Method we let \(x_n\) be our “\(n\)th guess,” and then its tangent line to \(y = f(x)\) at \(x = x_n\) is given by

\[
y = f(x_n) + f'(x_n)(x - x_n),
\]

which when we set equal to zero—to find the \(x\)-intercept—giving us

\[
f(x_n) + f'(x_n)(x - x_n) = 0 \iff f'(x_n)(x - x_n) = -f(x_n)
\]

\[
\iff x - x_n = -\frac{f(x_n)}{f'(x_n)}
\]

\[
\iff x = x_n - \frac{f(x_n)}{f'(x_n)}
\]

Taking \(x_{n+1}\) to be this \(x\)-value at which the tangent line through \((x_n, f(x_n))\) intersects the \(x\)-axis, we get the following, recursive formula:

\[
x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.
\]

(5.22)

For our present example in which \(f(x) = x - \cos x\) and \(f'(x) = 1 + \sin x\), our recursion relationship becomes

\[
x_{n+1} = x_n - \frac{x - \cos x}{1 + \sin x}.
\]

So far we have considered \(x_1 = 0\) to be our "first guess." Below we list values of \(x_n\) where we begin first with \(x_1 = 0\), and then also consider \(x_1 = 1, x_1 = 2\) and \(x_1 = 10\). Note that approximations are given to ten significant digits.
5.5. **NEWTON’S APPROXIMATION METHOD**

Figure 5.12: Figure showing two iterations of Newton’s Method for finding approximate solutions of \( f(x) = x - \cos x = 0 \), i.e., solutions of \( x = \cos x \), with an initial guess of \( x_1 = 0 \).

<table>
<thead>
<tr>
<th>( n )</th>
<th>( x_n )</th>
<th>( n )</th>
<th>( x_n )</th>
<th>( n )</th>
<th>( x_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>-1.35124412</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>-0.4965581783</td>
<td>2</td>
<td>-15.13524412</td>
</tr>
<tr>
<td>3</td>
<td>0.7503638678</td>
<td>3</td>
<td>2.131003844</td>
<td>3</td>
<td>16.07064125</td>
</tr>
<tr>
<td>4</td>
<td>0.7391128909</td>
<td>4</td>
<td>0.6896627208</td>
<td>4</td>
<td>-10.28559751</td>
</tr>
<tr>
<td>5</td>
<td>0.7390851334</td>
<td>5</td>
<td>0.7396529975</td>
<td>5</td>
<td>-4.80680582</td>
</tr>
<tr>
<td>6</td>
<td>0.7390851332</td>
<td>6</td>
<td>0.7390852044</td>
<td>6</td>
<td>-0.35795113</td>
</tr>
<tr>
<td>7</td>
<td>0.7390851332</td>
<td>7</td>
<td>0.7390851332</td>
<td>7</td>
<td>3.34826027</td>
</tr>
<tr>
<td>8</td>
<td>0.7390851332</td>
<td>8</td>
<td>0.7390851332</td>
<td>8</td>
<td>-2.095887668</td>
</tr>
<tr>
<td>9</td>
<td>0.7390851332</td>
<td>9</td>
<td>0.7390851332</td>
<td>9</td>
<td>9.740319348</td>
</tr>
</tbody>
</table>

For this particular example, it is clear that a good “first guess” \( x_1 \) can cause the algorithm to converge quickly to a stable value. However, using \( x_1 = 4 \) here leads to unstable oscillation in the sequence \( \{x_n\} \), and no apparent convergence. The behavior is not always predictable at first glance (the reader is invited to note that \( x_1 = 2 \) actually converges faster to the stable value of \( x_n \) than does letting \( x_1 = 0 \), though the solution is closer to 0 than 2). To be sure we have an approximate solution, one can check \( \cos 0.7390851332 \approx 0.7390851332 \), as desired.

Note that when the method does converge, it often does so quite quickly: we managed to get answers accurate to 10 significant digits after 5–6 iterations. To use our method of bisection and have accuracy of \( \pm 10^{-10} \) we would need to bisect the interval \([0, 2]\) some \( n \) times, giving a subinterval of length \( 2/2^n \) where \( 2/2^n \leq 10^{-10} \), i.e., \( 2^{1-n} \leq 10^{-10} \), which means \((1 - n) \log 2 \leq -10\), or \((n - 1) \log 2 \geq 10\), or \( n \geq 1 + 10/\log 2 \approx 34.2 \), so we would need 35 iterations of the Method of Bisection to achieve the same accuracy for this problem. That is certainly not a problem for a computer program to accomplish quickly, but the 5–6 steps required
here show Newton’s Method to be more efficient for this particular problem.\footnote{Indeed such problems were—and still are—typically solved by students using pencil-and-paper and very simple “scientific” calculators in homework and exam settings for decades.}

In the typical application of Newton’s Method, one attempts to solve an equation that has been put into the form \( f(x) = 0 \), makes a “guess” for \( x_1 \), runs the algorithm\footnote{Three common ways to run the algorithm are through the use of a calculator, a computer programming language, or a spreadsheet. With most graphing calculators today, we can use the \( \text{ANS} \) and \( \text{ENTER} \) keys to run the recursive steps easily. For instance, if we type our \( x_1 \) value \( x_1 \text{[ENTER]} \) the display shows 0 as our “answer.” We then can type \( \text{ANS}-(\text{ANS}-(\cos(\text{ANS})))/(1+\sin(\text{ANS})) \text{[ENTER]} \) and the calculator will return the output from entering our first “answer” \( (x_1 = 0) \) after it is run through the right-hand side of (5.22), namely \( x_2 \). At that point, \( x_2 \) is our new “answer,” so if we type \( \text{ENTER} \) again it will repeat the line above with \( \text{ANS}=x_2 \), giving our new “answer,” namely \( x_3 \) and so on. Repeatedly pressing \( \text{ENTER} \) then outputs \( x_4, x_5 \) and so on.} and inspects for convergence. If there will be convergence, it is usually readily apparent. If not, that is usually clear as well and another attempt to choose an \( x_1 \) that will cause convergence. There are some functions which will not allow for convergence, and a method such as the Method of Bisection can be attempted. (We will consider such situations later.) The electronic ability to graph the function in question can be very helpful as well, even if only to determine where to begin whichever method is chosen.

It is also the case that there can be several solutions of \( f(x) = 0 \), and perhaps some graphical analysis needs to be done at first so we know how many solutions exist, and approximately where they are located.

Example 5.5.1 Find all real solutions of \( x^3 = 1 - 3x^2 \).

\textbf{Solution:} We begin as before by making this a question about “zeros” of a function \( f(x) = x^3 + 3x^2 - 1 \). Assuming we had no graph to work with, we could next seek out intervals on which the function changes signs:

\[
\begin{array}{c|ccccccc}
\multicolumn{2}{c}{x} & -3 & -2 & -1 & 0 & 1 & 2 & 3 \\
\hline
f(x) & -1 & 3 & 1 & -1 & 3 & 19 & 53 \\
\end{array}
\]

We see sign changes in \([-3,-2], [-1,0], \) and \([0,1]\). We know from Algebra that there can be at most three solutions of \( f(x) = 0 \) for the case that \( f(x) \) is a third-degree polynomial, so we need only look for approximations of the solutions in these three intervals. We will choose \( x_1 \) to be \( 3 \), and so on. This will in fact make it easier to try new first guesses for \( A1 \) (think “\( x_1 \)”\), since editing \( A1 \) to be a different number will automatically “update” \( A2, A3 \) and so on.

Using either the programmable features of a graphing calculator, or an actual computer programming language are alternative strategies, as is simply using calculator memory (or re-typing) and processing the numbers by brute force.

Of course, regardless of the chosen method, we have to be sure that the calculating device is reading angles in radians.
5.5. **NEWTON’S APPROXIMATION METHOD**

The midpoint in each case. We will use these in our recursive formula, which for this problem becomes:

\[ x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \]

\[ \Leftrightarrow x_{n+1} = x_n - \frac{x^3 + 3x^2 - 1}{3x^2 + 6x}. \]

Using \( x_1 = 2.5, -0.5, 0.5 \) respectively we generate the following tables:

<table>
<thead>
<tr>
<th>( n )</th>
<th>( x_n )</th>
<th>( n )</th>
<th>( x_n )</th>
<th>( n )</th>
<th>( x_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-2.5</td>
<td>1</td>
<td>-0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>-3.066666667</td>
<td>2</td>
<td>-0.666666667</td>
<td>2</td>
<td>0.533333333</td>
</tr>
<tr>
<td>3</td>
<td>-2.900875604</td>
<td>3</td>
<td>-0.652777778</td>
<td>3</td>
<td>0.532090643</td>
</tr>
<tr>
<td>4</td>
<td>-2.879719904</td>
<td>4</td>
<td>-0.6527036468</td>
<td>4</td>
<td>0.5320888862</td>
</tr>
<tr>
<td>5</td>
<td>-2.879385325</td>
<td>5</td>
<td>-0.6527036447</td>
<td>5</td>
<td>0.5320888862</td>
</tr>
<tr>
<td>6</td>
<td>-2.879385242</td>
<td>6</td>
<td>-0.6527036447</td>
<td>6</td>
<td>0.5320888862</td>
</tr>
<tr>
<td>7</td>
<td>-2.879385242</td>
<td>7</td>
<td>-0.6527036447</td>
<td>7</td>
<td>0.5320888862</td>
</tr>
</tbody>
</table>

We conclude that approximate solutions of \( x^3 - 1 - 3x^2 \) are \( x \approx -2.879385242, -0.6527036447 \) and 0.5320888862. The graph of \( f(x) = x^3 + 3x^2 - 1 \) is given in Figure 5.13.

From both the geometric interpretation (of following the tangent lines back to the \( x \)-axis) and the recursion formula that a flatter curve, i.e., smaller \( f'(x_n) \) will cause \( x_{n+1} \) to be more distant from \( x_n \). In the above example, at \( x = 0 \) we would have \( f'(x) = 0 \) and therefore the tangent line would never intercept the \( x \)-axis. We would also find ourselves with a zero denominator in the recursion formula.

A common example showing graphically how the \( x_n \) can diverge is given next.

**Example 5.5.2** Suppose we wish to use Newton’s Method to solve \( \sqrt{x} = 0 \). While we can see the solution is \( x = 0 \), instead we will suppose we do not know this, and instead make our first
Figure 5.14: Figure showing the tangent lines producing the succession of the \(x_1, x_2, x_3, \ldots\) generated by the Newton’s Method recursion formula for \(f(x) = \sqrt{x} = 0\) if \(x_1 = 1\), as in Example 5.5.2. As the tangent lines become more horizontal, the \(x_n\) values diverge from each other.

guess of \(x_1 = 1\). Our recursion relation becomes

\[
x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^{1/3}}{3x_{n-2/3}} = x_n - 3x_n.
\]

Running our algorithm, we get

<table>
<thead>
<tr>
<th>(n)</th>
<th>(x_n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>-2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>-8</td>
</tr>
<tr>
<td>5</td>
<td>16</td>
</tr>
</tbody>
</table>

and so on. Indeed we can see a simple pattern where \(x_{n+1} = -2x_n\), and the \(x_n\) diverge. This coincides with the slopes of the tangent lines shrinking in size, the effect of which is shown graphically in Figure 5.14.

As has been pointed out, when the slope of the curve decreases in size, the method is likely to produce points \(x_n\) which are more spread apart, and therefore less likely convergent. However, the method can be quite useful when algebraic methods are not up to the task. It is helpful to have a general idea of the behavior of the function’s graph, such as the number of solutions.
Exercises

1. Newton’s Method produces a method to compute square roots which was used for many years by students who had no access to calculators. For instance, to compute $\sqrt{3}$, one would look at the function $f(x) = x^2 - 3$ and use the method to find where $f(x) = 0$. In this way, use Newton’s Method to compute $\sqrt{3}$, $\sqrt{24}$ and $\sqrt{36}$ accurate to $\pm 10^{-9}$ or better.

2. As above, use Newton’s Method to compute $\sqrt{9}$ accurate to $\pm 10^{-9}$ or better.

3. Consider the equation $\tan x = x$.
   (a) Sketch rough graphs of $y = \tan x$ and $y = x$. Note that $x = 0$ is clearly one solution.
   (b) Use Newton’s Method to find two other positive solutions.

4. Using only a scientific (non-graphing) calculator or similar device, graph $f(x) = x^3 + 2x^2 - 5x - 1$. To do so:
   (a) Find the critical points (using the Quadratic Formula) and produce a sign chart for $f'(x)$. Use this to identify local extrema.
   (b) Find any inflection points and produce a sign chart for $f''(x)$.
   (c) Use Newton’s Method to find all (three) $x$-intercepts.
   (d) Sketch the graph, reflecting all of this information, along with the $y$-intercept.
5.6 Chain Rule III: Related Rates

For many students, related rates is the topic for which calculus “comes alive,” because it involves dynamic changes (changes with time) of the variables involved. Basically, those variables are considered ultimately to be functions of time \( t \) (often a “hidden variable”), and their time derivatives describe their rate of dynamic change. The calculus that is applied is based mostly upon the chain rule, using the principles explored previously. For a complicated situation, there is usually some geometric or algebraic work to set the stage for the calculus, but the calculus itself is straightforward.

5.6.1 The Basic Idea; An Example

When variables which are related by some equations, or constraints, if those variables are dynamic (changing with time) then the rates at which they change are related, and we can use our usual calculus to discover those relations. The strategy is to apply \( \frac{d}{dt} \) to both sides of the relationship, and the usual differentiation rules (particularly the chain rule) will give us a relationship among the rates of change of the variables, though the variables themselves (without derivatives) may also be present.

For example, suppose a rocket’s height at time \( t \) seconds after the rocket lifts off is given by

\[ y = \frac{1}{320} t^2. \]  

(5.23)

In order to have consistent units, if \( y \) is the height in miles of the rocket, then we would in fact have

\[ y = \frac{1}{320 \text{ sec}^2/\text{mi}} \cdot t^2. \]  

(5.24)

This way, for \( t \) given in seconds we see that \( y \) will indeed be given in miles. From the above equation we can compute the rocket’s vertical velocity (by applying \( \frac{d}{dt} \) to both sides) to be

\[
\frac{dy}{dt} = \frac{1}{320 \text{ sec}^2/\text{mi}} \cdot 2t \Rightarrow \frac{dy}{dt} = \frac{1}{160} \text{ mi/sec} \cdot t.
\]  

(5.25)

Note that if \( t \) is in units of seconds, then \( \frac{dy}{dt} \) will be in miles/sec. This makes for simple enough computations of the rocket’s velocity at any variety of values of \( t > 0 \), as long as the rocket’s height is still given by (5.24). A short table of such values of \( y \) and \( \frac{dy}{dt} \) is given below:

<table>
<thead>
<tr>
<th>( t )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>\cdots</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y )</td>
<td>0</td>
<td>1/320</td>
<td>4/320</td>
<td>9/320</td>
<td>16/320</td>
<td>25/320</td>
<td>\cdots</td>
<td>0.3125</td>
<td>1.25</td>
<td>2.8125</td>
<td>5</td>
<td>7.8</td>
</tr>
<tr>
<td>( \frac{dy}{dt} )</td>
<td>0</td>
<td>1/160</td>
<td>2/160</td>
<td>3/160</td>
<td>4/160</td>
<td>5/160</td>
<td>\cdots</td>
<td>0.0625</td>
<td>0.125</td>
<td>0.1875</td>
<td>0.025</td>
<td>0.3125</td>
</tr>
</tbody>
</table>

Note that the final column’s entries including units are, respectively, 50 seconds, 7.8 miles, and 0.3125 miles/second. If we prefer to read that last \( (t = 50 \text{ sec}) \) entry in miles/hour, we simply convert:

\[
0.3125 \frac{\text{mi}}{\text{sec}} \cdot 3600 \frac{\text{sec}}{\text{hour}} = 1125 \text{ mile/hour}.
\]
5.6. CHAIN RULE III: RELATED RATES

In practice, for brevity one usually does not write units at each step, resting assured that the calculus is completely consistent with units. If the units are introduced correctly from the beginning, as we did when we rewrote (5.24) more carefully to include units, we can follow through with them as when we then derived the resulting equation (5.25). The student interested in studying how units follow through in the computations is encouraged to do so, on occasion. In doing so it is important to notice how they carry over into derivatives from the difference quotients. For instance,

\[
\frac{dy}{dt} \bigg|_{t=t_0} = \lim_{\Delta t \to 0} \frac{y(t_0 + \Delta t) - y(t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t},
\]

and so \(dy/dt\) will have units which are those of \(y\) divided by those of \(t\). Most of the time we will not include units in our computations, but there will be times where we either may verify they are valid, or include them if they are on balance instructive.

5.6.2 Choosing the Form of the Relationship

For a slightly more complicated example, now suppose we have a camera at ground level 2 miles from the site of the rocket’s liftoff, as seen in Figure 5.15. If we would like to find the speed in degrees/sec that the camera’s angle of elevation has to change to keep the rocket centered in the viewing area, we have to relate the angle of elevation \(\theta\) to the variable \(y\), which is simple enough:

\[
\tan \theta = \frac{y}{2 \text{ mi}}.
\]

So that units work easily with the calculus, we will temporarily measure \(\theta\) in radians, and we can differentiate this equation with respect to \(t\) as follows:

\[
\frac{d}{dt} \tan \theta = \frac{d}{dt} \left[ \frac{y}{2 \text{ mi}} \right] \Rightarrow \sec^2 \theta \cdot \frac{d\theta}{dt} = \frac{1}{2 \text{ mi}} \cdot \frac{dy}{dt}
\]
\[
\Rightarrow \frac{d\theta}{dt} = \frac{1}{2 \text{ mi}} \cos^2 \theta \cdot \frac{dy}{dt} \tag{5.26}
\]

If we were asked for the rate of change in the camera’s angle after the rocket had flown for 1 second, we would be looking for \(d\theta/dt\) for that instant, but note that we would need to find both \(y\) and \(\theta\) for that instant to use them in (5.26). Using a little bit of cleverness, we could instead note the hypotenuse in Figure 5.15 is \(\sqrt{y^2 + 4}\), making \(\cos \theta = \frac{2}{\sqrt{y^2 + 4}}\), and so (5.26)
can also be written
\[
\frac{d\theta}{dt} = \frac{1}{2} \cdot \frac{4}{y^2 + 4} \cdot \frac{dy}{dt}
\]
\[
\Rightarrow \frac{d\theta}{dt} = \frac{2}{y^2 + 4} \cdot \frac{dy}{dt}.
\]
(5.27)

This gives us a relationship among \( \frac{d\theta}{dt}, \frac{dy}{dt} \) and \( y \).

As an alternative, we could have used arctrigonometric functions from the beginning:
\[
\theta = \tan^{-1} \frac{y}{2}
\]
\[
\Rightarrow \frac{d\theta}{dt} = \frac{1}{y^2 + 1} \cdot \frac{1}{2} \cdot \frac{dy}{dt} = \frac{2}{y^2 + 4} \cdot \frac{dy}{dt}.
\]
Note that this is the same as (5.27). Ignoring units for now, if we recall from (5.23) and (5.25) that \( y = \frac{1}{320} t^2 \) and \( \frac{dy}{dt} = \frac{1}{160} t \), we would get:
\[
\frac{d\theta}{dt} = \left( \frac{1}{320} t^2 \right)^2 + 4 \cdot \frac{1}{160} \cdot t = \frac{t}{\left( \frac{1}{320} t^2 \right)^2 + 4} \cdot \frac{1}{80} \cdot \frac{1}{320^2} = \frac{1280t}{t^4 + 409,600}.
\]
(5.28)

Now we could have instead found \( \theta \) as a function of \( t \) using (5.23):
\[
\theta = \tan^{-1} \frac{1}{320} t^2
\]
\[
= \tan^{-1} \frac{t^2}{640}
\]
\[
\Rightarrow \frac{d\theta}{dt} = \frac{t^2}{640^2 + 1} \cdot \frac{2t}{640} \cdot \frac{640 \cdot 2t}{t^4 + 640^2} = \frac{1280t}{t^4 + 409,600}.
\]
(5.29)

We see that each of our methods yield equivalent results for \( \frac{d\theta}{dt} \), but there can be some variation in the complexity of the computation. Usually, when the relationship among variables is manipulated to be less complicated in one respect, the work to actually use the relationship among rates (i.e., “related rates”) is made more complicated; the hard work can be done early or late.

For instance, suppose we wish to know how fast the camera’s angle is rising at \( t = 1 \) second. We can use (5.29) with \( t = 1 \) immediately to get
\[
\left. \frac{d\theta}{dt} \right|_{t=1} = \frac{1280(1)}{(1)^4 + 409,600} = \frac{1280}{409,601},
\]
with the units in radians/second. To convert this to degrees/second, we multiply
\[
\frac{1280}{409,601} \text{ rad/sec} \cdot \frac{180^\circ}{\pi \text{ rad}} \approx 0.179^\circ/\text{sec}.
\]

If we were instead going to use (5.27), we would compute
\[
\left. \frac{d\theta}{dt} \right|_{t=1} = \frac{2}{y^2 + 4} \cdot \frac{dy}{dt},
\]
but we would need to find $y$ and $\frac{dy}{dt}$ at $t = 1$. This is not terribly difficult because we have (5.23) and (5.25), which give us

\[
y|_{t=1} = \frac{1}{320}(1)^2 = \frac{1}{320},
\]

\[
\frac{dy}{dt}|_{t=1} = \frac{1}{160}(1) = \frac{1}{360},
\]

and so we have

\[
\frac{d\theta}{dt}|_{t=1} = \frac{2}{y^2 + 4} \cdot \frac{1}{160} = \frac{1}{102400} + 4 \cdot \frac{1}{80} = \frac{1280}{409600} = \frac{1}{409600},
\]

as before in radians/second, i.e., approximately $0.179^\circ/\text{sec}$. This indeed does illustrate that simpler forms for the relationships among rates often require extensive ad hoc computations to use for actual numerical purposes. If one were to produce a table of values of $\frac{d\theta}{dt}$ for various values of $t$, it is best to use the explicit formula (5.29). We do so below:

<table>
<thead>
<tr>
<th>$t$ (in sec)</th>
<th>0</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{d\theta}{dt}$ (in rad/sec)</td>
<td>0.016</td>
<td>0.031</td>
<td>0.042</td>
<td>0.045</td>
<td>0.040</td>
<td>0.031</td>
<td>0.023</td>
<td>0.017</td>
<td>0.010</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>$\frac{d\theta}{dt}$ (in $^\circ$/sec)</td>
<td>0.894</td>
<td>1.748</td>
<td>2.390</td>
<td>2.575</td>
<td>2.291</td>
<td>1.804</td>
<td>1.344</td>
<td>0.988</td>
<td>0.551</td>
<td>0.073</td>
<td></td>
</tr>
</tbody>
</table>

Note how the camera at first has zero angular velocity, how that angular velocity increases for a time, and how it then decreases until it almost stops, as the rocket is flying very fast but the camera will be almost pointing vertically. This should seem somewhat intuitive.

While the explicit formula (5.29) for $\frac{d\theta}{dt}$ as a function of $t$ is best used to produce such a table (and a spreadsheet program is very well suited for the purpose), there is intuition to be gained from the other forms as well, as for instance $y$ and $\frac{dy}{dt}$ appear in (5.27). In fact that formula is still correct even if we change the formula for $y$ as a function of $t$. Thus (5.29) is specific to this particular rocket’s motion, while (5.27) holds correct for any such vertically launching rocket, assuming the camera is two miles from the launch pad.

**Example 5.6.1** Figure 5.16 on page 516 shows another example where the form of the relation among variables determines the resultant relationship involving rates. Consider the following relations—equivalent in for the situation in the figure—and the results of applying $\frac{d}{dt}$ to both sides.

\[
x^2 + y^2 = r^2 \quad \Rightarrow \quad \frac{d}{dt} \left( x^2 + y^2 \right) = 2x \frac{dx}{dt} + 2y \frac{dy}{dt} = 2r \frac{dr}{dt}.
\]

\[
r = \sqrt{x^2 + y^2} \quad \Rightarrow \quad \frac{dr}{dt} = \frac{1}{2\sqrt{x^2 + y^2}} \left( 2x \frac{dx}{dt} + 2y \frac{dy}{dt} \right),
\]

\[
y = \sqrt{r^2 - x^2} \quad \Rightarrow \quad \frac{dy}{dt} = \frac{1}{2\sqrt{r^2 - x^2}} \left( 2r \frac{dr}{dt} - 2x \frac{dx}{dt} \right),
\]

\[
x = \sqrt{r^2 - y^2} \quad \Rightarrow \quad \frac{dx}{dt} = \frac{1}{2\sqrt{r^2 - y^2}} \left( 2r \frac{dr}{dt} - 2y \frac{dy}{dt} \right).
\]

The first of these four above is the simplest to compute derivatives with respect to time. The next has the advantage that $r$ does not appear on the right, or the left for that matter except in
Figure 5.16: If \( x, y \) and \( r \) are the legs and hypotenuse of a right triangle, then \( x^2 + y^2 = r^2 \) according to the Pythagorean Theorem. This implies the \( 2x \frac{dx}{dt} + 2y \frac{dy}{dt} = 2r \frac{dr}{dt} \), if \( x, y \) and \( r \) are functions of \( t \). In fact, if we allow \( x \) and \( y \) to be respectively horizontal and vertical displacements from the origin \((0, 0)\), with \( r \) the distance from \((0, 0)\), the Pythagorean Theorem still holds and so does the equation relating the rates.

The above example gives a reasonably stark illustration that choosing the form of the relationship to differentiate can have implications for how difficult it is to compute related rates. However, the previous example where we found \( \frac{d\theta}{dt} \) also illustrates how the price for ease of differentiation is sometimes paid in difficulty of actually using the consequent rates relation. Fortunately, they are consistent and thus it is often the case that any correct relation can be differentiated into a useful relation among rates.

To be clear, in Example 5.6.1 one of the differentiations can be broken into smaller steps. The second one could read

\[
\begin{align*}
\frac{dr}{dt} &= \sqrt{x^2 + y^2} \\
\Rightarrow \quad \frac{d}{dt}[r] &= \frac{d}{dt}\sqrt{x^2 + y^2} \\
\Rightarrow \quad \frac{dr}{dt} &= \frac{1}{2\sqrt{x^2 + y^2}} \left( \frac{d}{dt}[x^2 + y^2] \right) \\
\Rightarrow \quad \frac{dr}{dt} &= \frac{1}{2\sqrt{x^2 + y^2}} \left[ \frac{d}{dt}[x^2] + \frac{d}{dt}[y^2] \right] \\
\Rightarrow \quad \frac{dr}{dt} &= \frac{1}{2\sqrt{x^2 + y^2}} \left[ 2x \frac{dx}{dt} + 2y \frac{dy}{dt} \right] \\
\Rightarrow \quad \frac{dr}{dt} &= \frac{1}{2\sqrt{x^2 + y^2}} \left[ 2x \frac{dx}{dt} + 2y \frac{dy}{dt} \right],
\end{align*}
\]

as before. If desired this can be somewhat simplified:

\[
\frac{dr}{dt} = \frac{1}{2\sqrt{x^2 + y^2}} \cdot 2 \left[ x \frac{dx}{dt} + y \frac{dy}{dt} \right] = \frac{x \frac{dx}{dt} + y \frac{dy}{dt}}{\sqrt{x^2 + y^2}}.
\]

Many important steps followed from the chain rule, such as when we computed \( \frac{dr^2}{dt} = 2x \frac{dx}{dt} \).\(^{17}\)

\(^{17}\)Some texts also call this process *implicit differentiation*, mainly because it resembles the work we did in Section 4.6. Related Rates and implicit differentiation are both really just specific applications of the chain rule.
Thus an equation relating variables holding true forces another equation relating the rates of change of these variables to hold true, hence the title of this section.

Example 5.6.2 Using the illustration in Figure 5.16, suppose at some particular time \(t_0\) we have \(x = 30\) m and \(y = 40\) m, but \(x\) is decreasing at \(20\) m/sec, and \(y\) is increasing at \(10\) m/sec. How is \(r\) changing at that time?

Solution: First we collect the numerical data we are given:

\[
\begin{align*}
x &= 30 \text{ m}, & \frac{dx}{dt} &= -20 \text{ m/sec} & \text{(since \(x\) is decreasing)}, \\
y &= 40 \text{ m}, & \frac{dy}{dt} &= 10 \text{ m/sec} & \text{(since \(y\) is increasing)}. \\
\end{align*}
\]

We do not have \(r\) immediately (though we can find it algebraically from \(x^2 + y^2 = r^2\)), so we will use the above computation for \(\frac{dr}{dt}\):

\[
\frac{dr}{dt} = \frac{x \frac{dx}{dt} + y \frac{dy}{dt}}{\sqrt{x^2 + y^2}} = \frac{30 \cdot (-20) + 40 \cdot (10)}{\sqrt{(30)^2 + (40)^2}} = \frac{-200}{50} = -4,
\]

so the distance \(r\) is decreasing at a rate of 4 m/sec at that instant. (The reader should verify that in the computation above, if units are included, the answer will be in units of m/sec.)

5.6.3 Examples: Finding a Relationship Among Variables

Before one answers the question of which form of the relationship among variables to differentiate for a particular application, one must first find at least one such relationship. The clues can often come from the geometry or other given constraints. Some of the more classical examples are given here.

Example 5.6.3 Given an inverted (downward-pointing), right circular conical tank with base diameter 10 meters and height 15 meters, and suppose it is being filled with water at a rate of 2 meter\(^3\) per minute. Find how quickly the water level is rising when the depth of the water is (a) 1 meter, (b) 2 meters, (c) 5 meters and (d) approaching 10 meters, and (e) approaching 0 meters.

Solution: We will let \(h\) be the height of the water in the tank, and note that we are seeking \(\frac{dh}{dt}\) for certain instances of \(h\). We are also given a rate of volume change, which we can write as \(\frac{dV}{dt} = 2\) (volume increases at 2 m\(^3\)/min), so a natural thing to do here is to relate \(V\) to \(h\) and differentiate. In this and many other cases, it is best to begin with a picture:\(^{18}\)

---

\(^{18}\)In practice, the picture and the equations are usually constructed in parallel, as the equations demand more information from the picture and vice versa.
The next step is to find a relationship between \( V \) and \( h \). This can perhaps most easily be done by introducing a third variable \( r \), which will be the radius of the water cone when the depth is \( h \). We can then write \( V \) in terms of both \( r \) and \( h \) from a formula in geometry:

\[
V = \frac{1}{3} \pi r^2 h.
\]

In fact we can differentiate this with respect to \( t \), but we would then have a \( \frac{dr}{dt} \) term, which would be somewhat problematic (since we only wish to have terms involving \( V \), \( h \) or their time derivatives). From the geometry of similar triangles we can “eliminate” \( r \):

\[
\frac{r}{5} = \frac{h}{15}.
\]

Putting this into our volume formula we now have

\[
V = \frac{1}{3} \pi \left( \frac{h}{15} \right)^2 h
\]

\[
\Rightarrow V = \frac{\pi}{675} h^3.
\]

Now we can take time derivatives:

\[
\frac{d}{dt} V = \frac{d}{dt} \left[ \frac{\pi}{675} h^3 \right] = \frac{\pi}{675} \cdot 3h^2 \frac{dh}{dt} = \frac{\pi}{225} h^2 \frac{dh}{dt}.
\]

We are given \( \frac{dV}{dt} \) and wish to solve for \( \frac{dh}{dt} \) for several different values of \( h \). In the abstract we have

\[
\frac{dh}{dt} = \frac{225}{\pi h^2} \cdot \frac{dV}{dt},
\]

which, after using our given information that \( \frac{dV}{dt} = 2 \), we have

\[
\frac{dh}{dt} = \frac{225}{\pi h^2} \cdot 2 = \frac{450}{\pi h^2}.
\]

From this we can compute

\[
\frac{dh}{dt} \bigg|_{h=1} = \frac{450}{\pi} \approx 143 \text{ m/min},
\]

\[
\frac{dh}{dt} \bigg|_{h=5} = \frac{450}{25\pi} \approx 5.73 \text{ m/min},
\]

\[
\frac{dh}{dt} \bigg|_{h=2} = \frac{450}{4\pi} \approx 35.8 \text{ m/min},
\]

\[
\frac{dh}{dt} \bigg|_{h=9} = \frac{450}{81\pi} \approx 1.77 \text{ m/min}.
\]

Furthermore, it is clear that

\[
x \to 10^- \implies \frac{dh}{dt} = \frac{450}{\pi h^2} \to \frac{450}{100\pi} \approx 1.43 \text{ m/min},
\]

\[
x \to 0^+ \implies \frac{dh}{dt} = \frac{450}{\pi h^2} \to \infty.
\]
5.6. CHAIN RULE III: RELATED RATES

We see a rapid rising of the height \( h \), indeed theoretically “infinitely” fast at the beginning, and then a slowing of the rising in \( h \) as it takes more water to fill higher levels of the tank.

In the above example, we could have “left” \( r \) in the equation for \( V \) and later used \( \frac{1}{15} r = \frac{1}{15} h \) when solving for \( \frac{dh}{dt} \), but it was arguably easier to have only \( V \) and \( h \) being related algebraically, so (they and) their derivatives could be related after the differentiation step. As mentioned before, there can be many correct paths to arrive at an answer to a related rate question, but it is often the case that one path can be much more efficient, or at least more straightforward, than another.

Students solving a typical related rate problem will often find a significant part of the difficulty lying in the steps taken before derivatives are computed. In other words, finding an actual equation relating the variable in question can be the most difficult part. In Subsection 5.6.1 we used the Pythagorean Theorem, in Subsection 5.6.2 we used right triangle trigonometry, and in Example 5.6.3 we used the formula for the volume of a cube and some facts regarding similar triangles.

5.6.4 Further Examples

There are occasions where we know some of the rates at a particular instant but do not necessarily know the entire scope of the relationship among the variables. The Leibniz notation can be particularly helpful here, though one has to be sure that, say, when writing \( \frac{dQ}{dx} \) that \( Q \) can be in theory ultimately a function of \( x \) (or whatever other variables are involved).

Example 5.6.4 Suppose we have an ideal gas, which is governed by the equation

\[
PV = nRT,
\]

where \( P \) is pressure, \( V \) is volume, \( n \) is the number of particles and \( T \) is the absolute temperature (for instance, in Kelvins), and \( R \) is a physical constant which also includes within it the necessary units to make the equation valid from a dimensional standpoint. Note that each of these quantities must be nonnegative. If the units of \( P \) are atmospheres, of \( V \) are liters, of \( n \) are moles and of \( T \) are Kelvins, then the research shows that we can take

\[
R \approx 0.08206 \frac{\text{L \cdot atm}}{\text{mol \cdot K}}.
\]

While we could take time derivatives \( \frac{d}{dt} \) of both sides, i.e., \( \frac{d}{dt}(PV) = \frac{d}{dt}(nRT) \), using the product rule and noting that \( R \) is a constant, it is a bit faster to use logarithmic differentiation and by taking the natural logarithm of both sides first:

\[
\ln(PV) = \ln(nRT) \Rightarrow \ln P + \ln V = \ln n + \ln k + \ln T
\]

\[
\frac{d}{dt} \frac{1}{P} \frac{dP}{dt} + \frac{1}{V} \frac{dV}{dt} = \frac{1}{n} \frac{dn}{dt} + \frac{1}{T} \frac{dT}{dt}.
\]

Technically, \( n \) is a “discrete” variable since it can only take on integer values, and so its derivative does not make sense. However, in practice we use units such as mols (1 mol \( \approx 6.02214179 \times 10^{23} \)) of particles, and so in practice \( n \) takes on decimal values with theoretically tens of significant digits possible, and as a result it is quite acceptable to behave as if \( n \) were a “continuous” variable.\(^{19}\) There are eight quantities in this derivative formula, so if we know any seven we can

\(^{19}\) In theory, the time derivative of an integer-valued, or discrete-valued function will be either zero (any open time interval on which the function is constant), or will not exist (at any point where the function has a discontinuity). Consider for instance \( f(x) = \lceil x \rceil \), whose derivative is zero except at the integer values of \( x \), at which it is undefined.
use algebra to find the eighth. In fact, since we also have the original relationship $PV = nRT$, we can eliminate any of the variables $P, V, n, T$ as well. Similarly, if we know one of these is held constant, then its derivative is zero and so we would have one less to be concerned about.

For instance, suppose we keep constant temperature and number of particles. Then we have

$$\frac{1}{P} \frac{dP}{dt} = -\frac{1}{V} \frac{dV}{dt}.$$ 

Using then $V = nRT/P$ we would have

$$\frac{1}{P} \frac{dP}{dt} = -\frac{P}{nRT} \frac{dV}{dt} \implies \frac{dP}{dt} = -\frac{P^2}{nRT} \frac{dV}{dt}.$$ 

One interesting aspect of this is that if volume increases ($dV/dt > 0$), then pressure decreases ($dP/dt < 0$). While intuitive, it not only falls out of the equations, but we can also compute the actual (instantaneous) rate at which that occurs.

If at a particular instant we have $P = 2.0\text{atm}$ and is decreasing at $0.52\text{ atm/min}$, and $V=5.0\text{ L}$, we can find out the change in $V$:

$$\frac{dV}{dt} = -\frac{V}{P} \frac{dP}{dt} = -\frac{5.0\text{ L}}{2.0\text{ atm}} \cdot -0.52\text{ atm/min} \approx 1.3\text{ L/min}.$$ 

We see that the volume at that instance is increasing at $1.3\text{ L/min}$.

The above example illustrates how differentiation with respect to time can elicit relationships among variables and their rates of change with respect to time even though we do not have explicit knowledge of the exact nature of their dependence on time: we do not have any formulas for how $P, V, n$ or $T$ actually depend on time but we can still say something about the relationship among their time derivatives. We can do that because we know how the variables themselves relate to each other, and because it is clear that these variables are, ultimately, functions of time: $P = P(t), V = V(t), n = n(t),$ and $T = T(t)$. Thus their derivatives $dP/dt, dV/dt, dn/dt$ and $dT/dt$ all make sense, even if we do not know formulas for them.

### 5.6.5 A Modeling Example

While some of the beauty of calculus is that it gives exact answers for idealized problems, a reality of engineering and science is that real-world problems are not ideal, and measurements are not exact. While one one hand this can be somewhat unsatisfying, on the other hand it sometimes allows us to relax the connection between the exact circumstances to find a simpler—even if somewhat more approximate—mathematical model. We take our model for what it is worth, but it may be worth a great deal in information and intuition regarding the actual, real-world problem.

**Example 5.6.5** Suppose we have a coil of sheet metal of thickness 1 millimeter, rolled on a 10 centimeter cylinder, mounted on a horizontal spindle. Assuming that the coil begins with 2 meter diameter, and is unrolled at 1.5 meters/second, find an equation relating the speed with which the coil spins as it unwinds as a function of the length that has been removed.
5.6. CHAIN RULE III: RELATED RATES

Solution: We are looking for a formula for $d\theta/dt$ in terms of the length $s$ of metal that has been pulled from the coil. There are many approaches to this problem, though each approach is likely to consider the same relations among the variables.

One problem with this situation is that in the ideal case, there is no “nice” formula for the relation between $\theta$ and $s$. If we consider the problem of how the metal was coiled onto the cylinder, we realize that while the inner most winding was mounted to a perfect circular cylinder, the next winding starts on the “bump” where the first coil ends, and the metal runs into itself. Theoretically the metal must bend around this bump, requiring a small extra amount of metal, but experience indicates that the “bump” will be cushioned and diffused by the slightly bending metal until each layer is rolled upon what is very nearly a perfect cylinder. We will ignore this.

From Cramer’s Rule we have

$$s = r\theta \text{ when } r \text{ is approximately constant},$$

but $r$ is increasing with $\theta$ as well, making $s$ increase with $\theta \cdot \theta$. While not a precise argument, we will see that this is reasonable as we look at the first few cases of $\theta = 2\pi n$.

For the first layer, $s = 10 \text{ cm} \cdot \theta$, $0 \leq \theta < 2\pi$. After that first layer, the new radius is 10.1 cm, and so we can take $s = 10 \text{ cm} \cdot \theta + 10.1 \text{ cm} \cdot (\theta - 2\pi), 2\pi \leq \theta < 4\pi$. This continues, and so we see a pattern (in cm):

$$0 \leq \theta < 2\pi : s = 10\theta,$$

$$2\pi \leq \theta < 4\pi : s = 10 \cdot 2\pi + 10.2(\theta - 2\pi),$$

$$4\pi \leq \theta < 6\pi : s = 10 \cdot 2\pi + 10.2 \cdot 2\pi + 10.4(\theta - 4\pi),$$

$$6\pi \leq \theta < 8\pi : s = 10 \cdot 2\pi + 10.2 \cdot 2\pi + 10.4 \cdot 2\pi + 10.6(\theta - 6\pi).$$

As mentioned before, this is already not exact, since it does not take into account the small length of metal needed to bend over the “bump” that should occur at each $\theta = 2\pi n$. However, if we consider some data points $(\theta, s)$, we see $(0, 0), (2\pi, 20\pi), (4\pi, 40.4\pi), (6\pi, 61.2\pi)$.

We now attempt to connect these points with a quadratic function

$$s = a\theta^2 + b\theta + c,$$

and use these data to find the coefficients $a, b, c$, and further test them on subsequent data. Using the point $(0, 0)$ we see that $0 = c$. Next we use $(2\pi, 20\pi)$ and $(4\pi, 40.4\pi)$. These give a system

$$
\begin{align*}
4\pi^2 a + 2\pi b &= 20\pi, \\
16\pi^2 a + 4\pi b &= 40.4\pi.
\end{align*}
$$

From Cramer’s Rule we have

$$a = \frac{20\pi}{40.4\pi} - \frac{2\pi}{4\pi} = \frac{80\pi^2 - 80.8\pi^2}{16\pi^3 - 32\pi^3} = -\frac{0.8\pi^2}{16\pi^3} = \frac{1}{20\pi},$$

$$b = \frac{4\pi^2}{40.4\pi} - \frac{20\pi}{4\pi} = \frac{161.6\pi^3 - 320\pi^3}{16\pi^3 - 32\pi^3} = -\frac{158.4\pi^3}{16\pi^3} = 9.9.$$

This gives us the quadratic model

$$s = \frac{1}{20\pi} \theta^2 + 9.9\theta. \quad (5.30)$$
Note that for \( \theta = 6\pi \), we would have
\[
s = \frac{36\pi^2}{20\pi} + 9.9 \cdot 6\pi = 1.8\pi + 59.4\pi = 61.2\pi \quad \text{as we required,}
\]
i.e., \((6\pi, 61.2\pi)\) is one such point in the graph of the function.

Equation (5.30) “smoothes out” our model for the length of metal coming off of the roll as \( \theta \) increases despite the presence of “bumps” where the depth of the metal on the roll jumps when given in units of layers of metal. It also lets us relate the time derivatives:

\[
\frac{ds}{dt} = \left( \frac{\theta}{10\pi} + 9.9 \right) \frac{d\theta}{dt},
\]

(5.31)

From this we get

\[
\frac{d\theta}{dt} = \frac{\frac{ds}{dt}}{\frac{\theta}{10\pi} + 9.9}.
\]

(5.32)

Note that here \( s \) is given in cm, \( \theta \) in radians and time \( t \) in seconds.
Chapter 6

Basic Integration

In this chapter we will consider the problem of recovering a function from knowledge of its derivative, or, equivalently, for a given function we will try to find another function whose derivative is the given function's. The general process is called antidifferentiation, or integration. The meaning of the first term is obvious: we are working backwards from the derivative to the function. The meaning of the other name for the process will become clearer in Section 6.3.

The main purpose of this chapter is to develop the first, basic techniques for computing antiderivatives \( F(x) \) for a given function \( f(x) \), i.e., given \( f(x) \) we look for \( F(x) \) so that

\[
F'(x) = f(x).
\]  

(6.1)

As we will see early in this chapter and throughout the next, antidifferentiation (finding some such \( F(x) \)), also known as integration, is less straightforward than differentiation (finding \( f'(x) \)). However, there are easily as many applications of antidifferentiation as there are of differentiation so it is a worthwhile process. In the first section we will limit ourselves to two applications:

1. Given the slope \( f'(x) \), find the function \( f(x) \) by antidifferentiation;\(^1\)

   Moreover, given \( f''(x) \), find \( f'(x) \), and then \( f(x) \).

2. Given velocity \( v \), find position \( s \).

   Moreover, given acceleration \( a \), find \( v \) and then \( s \).

Later in the chapter we will look at the geometric significance of antiderivatives \( F(x) \) of a function \( f(x) \). Just as the geometric meaning of slope gave us a useful perspective for arriving at derivative theorems (mean value theorem, first derivative test, etc.), so too will the antiderivatives benefit from geometric analysis. To make that analysis will require us to consider another major theoretical device, namely Riemann sums, which—together with the Fundamental Theorem of Calculus—will open the topic of integration to innumerable applications. To illustrate the reasonableness of the Fundamental Theorem of Calculus, we will again look closely at the velocity-position connection as well, in Chapter 8.

In Chapter 7 we will develop more advanced techniques of antidifferentiation, so that we can use all of our integration techniques in Chapter 8, which is devoted to applications.

For now we will concentrate on the actual computation of antiderivatives of the more basic types. In the first section we will limit ourselves to those which arise from our known derivative formulas. In subsequent sections we will then explore the substitution technique, which is the antidifferentiation analog to the chain rule in differentiation, and is thus arguably the most important of the integration techniques. It will be developed at length.

\(^1\)In this application, the part of \( f(x) \) in (6.1) is played by \( f'(x) \), while the part of \( F(x) \) is played by \( f(x) \).
6.1 First Indefinite Integrals (Antiderivatives)

In this section we introduce antiderivatives, which are exactly what the name implies. These are also called indefinite integrals for reasons which will eventually become clear.

6.1.1 Indefinite Integrals and Constants of Integration

Definition 6.1.1 Consider a function \( f(x) \) which is defined on an open interval \((a, b)\). Another function \( F(x) \), also defined on \((a, b)\) is called an antiderivative of \( f(x) \) on the same interval if and only if \( F'(x) = f(x) \) on \((a, b)\).

If instead \( f(x) \) is defined on a closed interval \([a, b]\), we still call \( F : [a, b] \rightarrow \mathbb{R} \) an antiderivative of \( f(x) \) on \([a, b]\) if and only if

\[
\begin{align*}
F'(x) &= f(x), & \text{for } x \in (a, b), \\
\lim_{\Delta x \to 0^+} \frac{F(a + \Delta x) - F(a)}{\Delta x} &= f(a), & \text{and} \\
\lim_{\Delta x \to 0^-} \frac{F(b + \Delta x) - F(b)}{\Delta x} &= f(b).
\end{align*}
\]

In other words, on an open interval we require \( F'(x) = f(x) \), while on the closed interval we also require the right derivative of \( F(x) \) to be \( f(a) \) at \( x = a \), and the left derivative of \( F(x) \) to be \( f(b) \) at \( x = b \).

Notice that (by Theorem 4.5.2, page 372) the definition implies that \( F(x) \) is continuous on the interval in question (since where \( F' = f \) exists, \( F \) must be continuous). For a simple example, consider \( f(x) = 2x + 3 \) on any open (or nontrivial closed) interval. An antiderivative of \( f(x) \) can be \( F(x) = x^2 + 3x \), since then \( F'(x) = 2x + 3 = f(x) \). However, another perfectly good antiderivative can be \( F(x) = x^2 + 3x + 5 \), or \( F(x) = x^2 + 3x - 100,000 \), since the derivative of the trailing constant term will always be zero. In logical terms we can write

\[
\begin{align*}
F(x) &= x^2 + 3x \quad \Rightarrow \quad F'(x) = 2x + 3, \\
F(x) &= x^2 + 3x + 5 \quad \Rightarrow \quad F'(x) = 2x + 3, \\
F(x) &= x^2 + 3x - 100,000 \quad \Rightarrow \quad F'(x) = 2x + 3, \\
F(x) &= x^2 + 3x + C, \text{ some } C \in \mathbb{R} \quad \iff \quad F'(x) = 2x + 3.
\end{align*}
\]

That (6.2) is an equivalence we will prove shortly. To signify that equivalence, we write

\[ \int (2x + 3) \, dx = x^2 + 3x + C, \quad C \in \mathbb{R}. \quad (6.3) \]

We call the right hand side of (6.3) the most general antiderivative, or just the antiderivative, of \( 2x + 3 \) (with respect to \( x \)). It is also called the indefinite integral of \( 2x + 3 \) (again with respect to \( x \)), and we will eventually migrate to using that term as our default.\(^2\) The constant \( C \) is called the constant of integration, since it must be included to achieve all solutions to the question of what is an antiderivative of \( f(x) = 2x + 3 \).

It is useful to note that there is an analogy to the operation of differentiation contained in the symbols:

\(^2\)The indefinite integral has a strong connection to the very important definite integral, which is a measure of accumulated change as computed from the instantaneous rate of change. This computation is our eventual goal and—to restate the introduction to this chapter—the connection between antiderivatives (indefinite integrals) and accumulated change (definite integrals) is precisely the subject of the Fundamental Theorem of Calculus.
6.1. FIRST INDEFINITE INTEGRALS (ANTIDERIVATIVES) 

- \( \frac{d}{dx} \left( \right) \) symbolizes computing the derivative of “( )” with respect to \( x \);
- \( \int \left( \right) \; dx \) symbolizes computing the antiderivative of “( )” with respect to \( x \).

Just as \( \frac{d}{dx} \) was considered a “differential operator,” \( \int \left( \right) \; dx \) is considered an “integral operator,” inputting a function of \( x \) and outputting its general antiderivative. When we write (6.3), that is, \( \int (2x + 3) \; dx = x^2 + 3x + C \), the expression on the left can be broken into:

1. “\( \int \),” the integral symbol, introduced by Leibniz from an old style, German “long s” (so-called for pronunciation, not length) for reasons we will discuss later;
2. “\( 2x + 3 \),” the integrand, whose antiderivatives we seek; and
3. “\( dx \),” the differential of \( x \), signifying we are computing the antiderivative respect to \( x \).

The integral symbol \( \int \) and the differential \( dx \) together form the integral operator \( \int \left( \right) \; dx \). When there is no ambiguity, the parentheses are omitted. Also it is common to treat the differential \( dx \) as a multiplier of the integrand, for reasons which will become more clear after Section 6.3, and so it is common to see notation such as

\[
\int f(x) \; dx = \frac{1}{2} x^2 + C, \quad \int \frac{dx}{x^2} = -\frac{1}{x} + C.
\]

These are shorthand for \( \int f(x) \; dx = \frac{1}{2} x^2 + C \) and \( \int \left( \frac{1}{x^2} \right) \; dx = -\frac{1}{x} + C \).

We now note that, indeed, all antiderivatives of \( 2x + 3 \) are necessarily of the form \( x^2 + 3x + C \). To prove this, on some interval define \( F(x) = x^2 + 3x \), which we can easily see is an antiderivative of \( 2x + 3 \) on that interval. Next suppose \( G(x) \) is another such antiderivative, i.e., that \( G'(x) = 2x + 3 \), on that same interval. Then on that interval, \( F \) and \( G \) must differ by a constant:

\[
\frac{d}{dx} \left[ (G(x) - F(x)) \right] = G'(x) - F'(x) = (2x + 3) - (2x + 3) = 0 \implies G(x) - F(x) = C,
\]

for some \( C \in \mathbb{R} \). Thus any antiderivative \( G(x) \) must be of the form \( G(x) = F(x) + C \), q.e.d.\(^3\)

To be clear on the notation, we now insert the following definition.

**Definition 6.1.2** If \( F(x) \) is an antiderivative of \( f(x) \), with respect to \( x \), on the interval \( I \), then on that interval we write

\[
\int f(x) \; dx = F(x) + C,
\]

where \( C \) is an arbitrary constant of integration. The process of computing an antiderivative is called integration (while the process of computing a derivative is called differentiation).

**Example 6.1.1** Consider \( f(x) = 2 \sin x \cos x \). One antiderivative is \( F(x) = \sin^2 x \), since

\[
F'(x) = \frac{d}{dx} \sin^2 x = \frac{d}{dx} (\sin x)^2 = 2 \sin x \cdot \frac{d}{dx} \sin x = 2 \sin x \cos x.
\]

\(^3\) Recall that a function with the zero function for its derivative on an interval must be constant on that interval, i.e., \( h' = 0 \) in \((a, b)\) implies \( h(x) \) is constant in \((a, b)\). Applying this to \( h = F - G \) we can argue:

\[
(\forall x \in I)[(F(x) - G(x))'] = 0 \implies (\exists C \in \mathbb{R})[\forall x \in I][F(x) - G(x) = C].
\]
However, another antiderivative is $G(x) = -\cos^2 x$, since
\[ G'(x) = \frac{d(-\cos^2 x)}{dx} = - \frac{d(\cos x)^2}{dx} = - \left[ 2 \cos x \cdot \frac{d\cos x}{dx} \right] = - [2 \cos x(-\sin x)] = 2 \sin x \cos x. \]

Note that
\[ F(x) - G(x) = \sin^2 x - (-\cos^2 x) = \sin^2 x + \cos^2 x = 1, \]
so we see that $F$ and $G$ do actually differ by a constant. To report the most general antiderivative of $f(x) = 2 \sin x \cos x$, either of the following are valid (but understood to have different “C’s”):
\[ \int 2 \sin x \cos x \, dx = \sin^2 x + C, \]
\[ \int 2 \sin x \cos x \, dx = -\cos^2 x + C. \]

Especially when dealing with trigonometric functions—with all their interconnectedness through various identities—it is common to find very different-looking forms of the general antiderivative, all of which differ by constants from each other. It is occasionally important to be alert for apparent discrepancies which are explained by this nature of the general antiderivative.

**Example 6.1.2** Suppose $f(x) = x + 1$. Then both forms below are general antiderivatives:
\[ \int (x + 1) \, dx = \frac{1}{2} x^2 + x + C, \]
\[ \int (x + 1) \, dx = \frac{1}{2} (x + 1)^2 + C. \]

We can see this by taking derivatives of each. We can also see this if we label $F(x) = \frac{1}{2} x^2 + x$ and $G(x) = \frac{1}{2} (x + 1)^2$, so the antiderivatives above are just $F(x)$ and $G(x)$ plus constants, respectively, and then compute
\[ F(x) - G(x) = \left[ \frac{1}{2} x^2 + x \right] - \left[ \frac{1}{2} (x^2 + 2x + 1) \right] = -\frac{1}{2}, \]
so these do differ by a constant, as expected.

### 6.1.2 Power Rule for Integrals

Where the rules for computing derivatives were straightforward (which is not to say immediately “easy”), those for computing antiderivatives are not so algorithmic. Indeed the methods are varied. Nonetheless, they are necessary to learn for a reasonably complete understanding of standard calculus, and we begin with the **power rule for integrals**:
\[ \int x^n \, dx = \frac{1}{n + 1} x^{n+1} + C, \quad n \neq -1, \quad \text{(6.5)} \]
\[ \int \frac{1}{x} \, dx = \ln |x| + C. \quad \text{(6.6)} \]

Here the intervals in question are those upon which $x^n$ is defined. To check (6.6), we simply notice $\frac{d}{dx} \ln |x| = \frac{1}{x}$. For (6.5) we compute:
\[ \frac{d}{dx} \left[ \frac{1}{n + 1} x^{n+1} \right] = \frac{1}{n + 1} \cdot (n + 1) x^{(n+1)-1} = x^n, \quad \text{q.e.d.} \]
In performing the above computation, we must notice that \( n + 1 \) and \( \frac{1}{n + 1} \) are constants, and so are preserved throughout the computation (and do not, for instance, require product/quotient/chain rules since they do not vary). We also note that the formula to the right of (6.5) is meaningless for the case \( n = -1 \), so we need (6.6) for that case.

When checking any antiderivative by differentiation, it is customary to not include the arbitrary additive constant, since its derivative is zero. However, it is certainly correct to include it, as in \( \frac{d}{dx} \ln |x| + C = \frac{1}{x} + 0 = \frac{1}{x} \).

We can apply the power rule immediately as in the following:

\[
\begin{align*}
\int x^2 \, dx &= \frac{1}{3}x^3 + C, \\
\int x^3 \, dx &= \frac{1}{4}x^4 + C, \\
\int \frac{1}{x^2} \, dx &= \int x^{-2} \, dx = \frac{1}{-1}x^{-1} + C = -x^{-1} + C = \frac{-1}{x} + C, \\
\int \frac{x}{x^2} \, dx &= \int \frac{1}{x} \, dx = \ln |x| + C.
\end{align*}
\]

We can check all of these by taking derivatives of our answers, with respect to \( x \) (i.e., by applying \( d/dx \)). As with derivatives, many functions which are powers of the variable are not explicitly written as such. Furthermore as with derivatives, the variable name in antiderivative formulas does not matter as long as it is matched in the differential:

\[
\begin{align*}
\int t^9 \, dt &= \frac{1}{10}t^{10} + C, \\
\int \sqrt{u} \, du &= \int u^{1/2} \, du = \frac{1}{\frac{3}{2} + 1}u^{\frac{3}{2} + 1} + C = \frac{2}{3}u^{3/2} + C, \\
\int \frac{1}{z^{5/4}} \, dz &= \int z^{-5/4} \, dz = \frac{1}{\frac{3}{5} + 1}z^{\frac{3}{5} + 1} + C = \frac{z^{-1/4}}{-\frac{4}{5}} + C = \frac{-4}{\sqrt[4]{z}} + C.
\end{align*}
\]

To check these, we can apply respectively \( d/dt \), \( d/du \) and \( d/dz \). Before we go on we note that (6.5), interepreted formally, applies to the case \( n = 0 \) as well: \( \int x^0 \, dx = \frac{1}{0+1}x^{0+1} + C \), i.e.,

\[
\int 1 \, dx = x + C. \tag{6.7}
\]

Taking the derivative of the right-hand side of (6.7) quickly shows it is in fact true. This integral is often, perhaps at first confusingly, abbreviated without the factor “1” included:

\[
\int dx = x + C. \tag{6.8}
\]

As with the chain rule computations, the differential \( dx \) is formally treated as a factor as in the following:

\[
\int \frac{dx}{x^3} = \int x^{-3} \, dx = \frac{1}{-2}x^{-2} + C = \frac{-1}{2x^2} + C.
\]

In fact, it is a very useful exercise to check these antiderivatives by quick mental calculations. Since the derivative formulas have been used extensively to this point, the processes of computing antiderivatives can be well-informed by their connections to known derivative techniques. In particular, mistakes in computing antiderivatives can often be immediately corrected. Perhaps even more importantly, the approximate form of an antiderivative can often be anticipated. For instance, we know that the derivative of a fourth-degree polynomial is necessarily a third-degree polynomial. It should be clear (though some argument is necessary to prove) that the antiderivative of a third-degree polynomial is necessarily a fourth-degree polynomial. Anticipating the final form is also very useful in substitution problems, which are introduced in Section 6.4 and ubiquitous thereafter.
Later we will see this treatment of \( dx \) justified and exploited in several contexts.

Now we state two very general results which may seem obvious, but are worth exploring with some care because of a technical consideration regarding the constant of integration.\(^5\) These are also useful to us immediately because they allow us to use the power rule multiple times to compute the derivatives of polynomials.

**Theorem 6.1.1** Suppose that \( F'(x) = f(x) \) and \( G'(x) = g(x) \) on some interval in consideration, and that \( k \in \mathbb{R} \) is a fixed constant. Then on that same interval we have

\[
\int [f(x) + g(x)] \, dx = F(x) + G(x) + C, \tag{6.9}
\]

\[
\int [k \cdot f(x)] \, dx = kF(x) + C. \tag{6.10}
\]

where \( C \in \mathbb{R} \) is a constant of integration.

These can be proved by taking derivatives. For instance, \( \frac{d}{dx}[kF(x)] = k \cdot \frac{d}{dx}F(x) = kF'(x) = kf(x) \). Note that this theorem can be rewritten

\[
\int [f(x) + g(x)] \, dx = \int f(x) \, dx + \int g(x) \, dx, \tag{6.11}
\]

\[
\int kf(x) \, dx = k \int f(x) \, dx, \quad k \neq 0, \tag{6.12}
\]

\[
\int 0 \, dx = C. \tag{6.13}
\]

A moment’s reflection shows that we do need (6.12) and (6.13) to catch both cases summarized in (6.10), else we lose the constant of integration if we let \( k = 0 \) in (6.12). (Note that, indeed, \( \frac{d}{dx}C = 0 \) which verifies (6.13).) More importantly, these new forms (6.11)–(6.13) are not inconsistent with those of the theorem when we consider the arbitrary constants. For instance, if we assume \( F'(x) = f(x) \) and \( G'(x) = g(x) \) as in the theorem, then we can write (6.11) as follows:

\[
\int f(x) \, dx + \int g(x) \, dx = (F(x) + C_1) + (G(x) + C_2)
\]

\[
= F(x) + G(x) + (C_1 + C_2) = F(x) + G(x) + C,
\]

where \( C_1 \) and \( C_2 \) are arbitrary constants, and their sum will also be an arbitrary constant which we can name \( C \).

With the above theorem and the power rule, we can now compute the indefinite integrals of polynomials and other linear combinations of powers (that is, sums of constant multiples of powers of \( x \)).

**Example 6.1.3** Consider the following integrals:

(a) \[
\int (4x^2 - 9x + 7) \, dx = 4 \cdot \frac{x^3}{3} - \frac{9x^2}{2} + 7 \cdot x + C = \frac{4}{3}x^3 - \frac{9}{2}x^2 + 7x + C,
\]

\(^5\)Many calculus students become careless about this constant of integration and just “tack it on the end” when computing antiderivatives. However, its correct placement is crucial in several contexts, so it is useful to be vigilant from the beginning of integration study. Carelessness in its placement can cause trouble already in this section, but will be particularly troublesome in subsequent third-semester calculus and in differential equations studies.
6.1. FIRST INDEFINITE INTEGRALS (ANTIDERIVATIVES) 529

(b) \[ \int \left( \frac{3x - 9}{x^3} \right) \, dx = \int (3x^{-2} - 9x^{-3}) \, dx = 3 \cdot x^{-1} \cdot \frac{1}{-1} - 9 \cdot x^{-2} \cdot \frac{1}{-2} + C = -\frac{3}{x} + \frac{9}{2x^2} + C, \]

(c) \[ \int (5x^2)^3 \, dx = \int 125x^6 \, dx = 125 \cdot \frac{x^7}{7} + C = \frac{125}{7} x^7 + C, \]

(d) \[ \int (x^2 + 7)^2 \, dx = \int (x^4 + 14x^2 + 49) \, dx = \frac{x^5}{5} + 14 \cdot \frac{x^3}{3} + 49x + C, \]

(e) \[ \int \sqrt{2x} \, dx = \int \sqrt{2} \sqrt{x} \, dx = \int 2^{1/2} x^{1/2} \, dx = 2^{1/2} \cdot \frac{x^{3/2}}{3/2} + C = \frac{2\sqrt{2}}{3} x^{3/2} + C. \]

When we use an integration rule such as the power rule for integrals, as with derivatives it is important that the variable which the antiderivative is with respect to matches the term in the function. For instance,

\[ \int x^3 \, dx = \frac{1}{4} x^4 + C, \quad \text{(6.14)} \]
\[ \int w^3 \, dw = \frac{1}{4} w^4 + C, \quad \text{(6.15)} \]
\[ \int (5x - 11)^3 \, dx \neq \frac{1}{4} (5x - 11)^4 + C. \quad \text{(6.16)} \]

What goes wrong in (6.16) is the integral analog to what goes wrong below (which is that we need the chain rule to make the variables of differentiation match):

\[ \frac{d}{dx} [(5x - 11)^4] \neq 4(5x - 11)^3; \]
\[ \frac{d}{dx} [(5x - 11)^4] = 4(5x - 11)^3 \cdot \frac{d(5x - 11)}{dx} = 4(5x - 11)^3 \cdot 5 \neq 4(5x - 11)^3, \]

The problem is that the differential, \( dx \), is that of \( x \) and not \( 5x - 11 \). In the next section we will address a kind of integral version of the chain rule (commonly known as integration by substitution for reasons which will be clear later), which would make short work of this integral. Without it, we may need to expand \( (5x - 11)^3 \) as a polynomial, or guess the solution and check that it works, and possibly make adjustments. Either way, it should suffice to point out that we need to take care in using integral formulas such as the integral power rule, page 526.\(^6\)

6Without giving away the substitution technique, we will note here that we can rewrite the integral in (6.16) so the differential matches the term \( (5x - 11) \). The argument would be the analog of our early chain rule expansions, such as (4.33), page 355. The idea is that \( d(5x - 11) = 5dx \) (recall the meaning of \( d(5x - 11)/dx \)), and so \( dx = d(5x - 11)/5 \), which allows us to rewrite the integral as follows:

\[ \int (5x - 11)^3 \, dx = \int (5x - 11)^3 \cdot \frac{d(5x - 11)}{5} = \frac{1}{5} \cdot \frac{1}{4} \cdot (5x - 11)^4 + C = \frac{1}{20} (5x - 11)^4 + C. \]

Except for the factor of \( \frac{1}{5} \) in the second interval, that rewriting had an integral power rule form.

Our integration by substitution method will be more systematic than the above computation. Consequently it will read better, and be less error-prone. That method will then be called upon extensively from then on.

6.1.3 Finding \( C \) (Where Possible)

Many times we are interested in a particular antiderivative. This is then a question of finding the particular "\( C \)" we need. Recall that all antiderivatives of a function (on a particular interval) differ by a constant, so here we use some other information (where available) to "fix," i.e., determine, the constant.
Example 6.1.4 Find \( f(x) \) so that \( f'(x) = 2x \) and \( f(3) = 7 \).

Solution: For a problem such as this, it is common to write

\[
f(x) = \int f'(x) \, dx,
\]

where it is understood that we will eventually find the exact antiderivative so that the function is well-defined. For our particular problem, one might continue to write

\[
f(x) = \int 2x \, dx = 2 \cdot \frac{x^2}{2} + C = x^2 + C.
\]

Now we find the particular \( C \) for this case, and we do this by inputting the “datum” (sometimes called “data point”) \( f(3) = 7 \). Graphically this means that the point \((3, 7)\) is on the curve. Since \( f(x) = x^2 + C \), we can find \( C \) using this datum:

\[
f(3) = 7 \iff 3^2 + C = 7 \iff 9 + C = 7 \iff C = -2.
\]

Thus \( f(x) = x^2 - 2 \).

A graphical way of interpreting the example above is to realize that all the curves \( y = x^2 + C \) are parabolas, and in fact are just vertical shifts of the curve \( y = x^2 \). Our task in Example 6.1.4 was then to find which shift satisfies both \( f'(x) = 2x \) and \( f(3) = 7 \). In Figure 6.1, \( y = x^2 + C \) is graphed for various values of \( C \). Once we require the graph to pass through a particular point—in this case the point \((3, 7)\), we “pin down” a particular curve, i.e., we determine exactly one curve from the family of curves, as graphed in Figure 6.1, page 531.

Finding a particular antiderivative is also very useful in kinematics. For instance, if we know the velocity function \( s'(t) = v(t) \), we can find the position function \( s \) if we are also given one position datum to “fix” the constant. With the understanding that the constant is to be determined, it is often written:

\[
s(t) = \int s'(t) \, dt = \int v(t) \, dt.
\] (6.17)

A common datum to prescribe is that \( s(0) = s_0 \) (where \( s_0 \) is some fixed number), but any data which “pins down” the function will suffice.

Example 6.1.5 Suppose \( v = t^2 + 11t - 25 \), and \( s(1) = 4 \). Find \( s(t) \).

Solution:

\[
s(t) = \int v(t) \, dt = \int (t^2 + 11t - 25) \, dt = \frac{t^3}{3} + \frac{11t^2}{2} - 25t + C.
\]

Using \( s(1) = 4 \) we get

\[
\frac{1^3}{3} + \frac{11 \cdot 1^2}{2} - 25(1) + C = 4 \iff \frac{1}{3} + \frac{11}{2} - 25 + C = 4,
\]

and so \( C = 4 + 25 - \frac{11}{2} - \frac{1}{3} = 29 - \frac{35}{6} = \frac{174}{6} - \frac{35}{6} = \frac{139}{6} \). Finally, this gives us

\[
s(t) = \frac{t^3}{3} + \frac{11t^2}{2} - 25t + \frac{139}{6}.
\]
6.1. FIRST INDEFINITE INTEGRALS (ANTIDERIVATIVES)

Figure 6.1: Partial view of the family of curves \( y = x^2 + C \) satisfying \( dy/dx = 2x \). For Example 6.1.4, page 530, we needed to find the value of \( C \) satisfying \( f'(x) = 2x \), i.e., \( f(x) = x^2 + C \), so that \( (3, 7) \) was on the curve. Note that the slopes of all curves given above are the same for a given \( x \)-value, but only one passes through \( (3, 7) \), namely \( f(x) = x^2 - 2 \).

Now we will derive a well-known formula of physics, namely one-dimensional motion assuming constant acceleration.

Example 6.1.6 Suppose that acceleration is given by a constant, say \( s''(t) = a \) (where \( a \) is fixed, i.e., \( a(t) = a \) is constant). Suppose further that \( s(0) = s_0 \) and \( v(0) = v_0 \). Now we work “backwards” from the acceleration towards the position function (via the velocity function) as follows:

\[
v(t) = s'(t) = \int s''(t) \, dt = \int a \, dt = at + C_1.
\]

(Note that the last computation required that acceleration, \( a \), be constant.) Using \( v(0) = v_0 \), we then have

\[
a \cdot 0 + C_1 = v_0 \iff C_1 = v_0.
\]

This gives us the following equation, which itself is well known to physics students:

\[
v(t) = at + v_0. \tag{6.18}
\]

Now we integrate (6.18), again taking care to treat constants and variables correctly:

\[
s(t) = \int s'(t) \, dt = \int v(t) \, dt = \int (at + v_0) \, dt = a \cdot \frac{t^2}{2} + v_0 t + C_2.
\]

Finally, using \( s(0) = s_0 \), we get

\[
a \cdot \frac{0^2}{2} + v_0(0) + C_2 = s_0 \iff C_2 = s_0.
\]
Thus
\[ s = \frac{1}{2} at^2 + v_0 t + s_0. \] (6.19)

It is important to note that (6.19) followed under the special condition that acceleration is constant (such as occurs when an object is in freefall in a constant gravitational field, with no other resistance). Nonconstant acceleration will not give (6.18) or (6.19). However, the method for computing \( v \) and \( s \), given \( a \), is the same when \( a \) is not constant:

1. find \( v(t) = \int a(t) \, dt \), using one datum regarding velocity at a particular time, to fix the constant of integration;
2. find \( s(t) = \int v(t) \, dt \), using another datum regarding position at a particular time, to fix the second constant of integration.

Actually, two position data can fix the constants as well, since we can just carry the first constant into the second calculation, and then we will have two equations with two unknowns (the constants of integration), and then solve for both constants.

**Example 6.1.7** Suppose \( a(t) = 3t^2 \), \( s(0) = 3 \) and \( s(1) = 5 \). Find \( v(t) \) and \( s(t) \).

**Solution**: First we will find \( v(t) \), to the extent that we can:

\[
v(t) = \int a(t) \, dt = \int 3t^2 \, dt = 3 \cdot \frac{t^3}{3} + C_1 = t^3 + C_1.\]

Next we find the form of \( s(t) \):

\[
s(t) = \int v(t) \, dt = \int (t^3 + C_1) \, dt = \frac{t^4}{4} + C_1 t + C_2.\]

So we know that \( s(t) = \frac{1}{4} t^4 + C_1 t + C_2 \), for some \( C_1, C_2 \). Using the facts that \( s(0) = 3 \) and \( s(1) = 5 \), we get the following system of two equations in two unknowns:

\[
\begin{cases}
\frac{9}{4} + C_1(0) + C_2 = 3 \\
\frac{1}{4} + C_1(1) + C_2 = 5
\end{cases} \iff \begin{cases}
C_1 + C_2 = 3 \\
C_1 + C_2 = \frac{19}{4}
\end{cases}
\]

From the second form of the system, we see \( C_2 = 3 \), and so \( C_1 = \frac{19}{4} - 3 = \frac{7}{4} \). Putting all this together we first get

\[ s(t) = \frac{t^4}{4} + \frac{7t}{4} + 3, \]

from which we can calculate \( v(t) = s'(t) \) (or just read \( v(t) \) off of our first integral calculation, inserting \( C_1 = \frac{7}{4} \)) to get

\[ v(t) = t^3 + \frac{7}{4}. \]

### 6.1.4 First Trigonometric Rules

With every derivative formula for functions comes an analogous antiderivative formula, which is more or less the derivative formula in reverse. Sometimes the reverse is more obvious than other times. For instance, the power rule formula for derivatives is sometimes seen algorithmically as "multiply by the exponent (‘bring the power down’) and decrease the exponent by one,” as in

\[ \frac{d}{dx} [x^n] = n \cdot x^{n-1}. \]
If we are careful to reverse the process, we need to do the inverse steps in reverse order: _increase_ the exponent by one, and then _divide_ by the exponent. That is the essence of (6.5):

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

(Note that this requires $n \neq -1$, which we will handle later. As the reader may suspect, it will involve logarithms.) By other sophisticated arguments, in the next section we will see a kind of reverse chain rule. A bit later in the text we will also come across what can be loosely called a reverse product rule called integration by parts, although there really is no good analog of the product rule with integrals _per se_.

The formulas presented in this subsection are immediate consequences of our trigonometric derivative formulas. For instance, we have the following pair of formulas:

$$\frac{d}{dx} \sin x = \cos x \iff \int \cos x \, dx = \sin x + C,$$

$$\frac{d}{dx} \cos x = -\sin x \iff \int (-\sin x) \, dx = \cos x + C.$$  

This second integration formula is more awkward than necessary, since it is more likely we would like an antiderivative for $\sin x$ directly. We could multiply both sides by $-1$, and rename the new constant $C$, or just notice that $\frac{d}{dx}(-\cos x) = \sin x$, to come to the formula

$$\int \sin x \, dx = -\cos x + C.$$  

As before, we can always check these by taking the derivative of the right-hand side. Recalling our six basic trigonometric derivative formulas, and making adjustments for negative sign placements

---

7Integration by parts is really an integration technique which takes advantage of the product rule for derivatives—or more precisely a permutation of the product rule for derivatives—but is not itself a product rule for integrals; it does not by itself give a formula for $\int f(x)g(x) \, dx$. Instead it gives a formula which can be summarized by

$$\int f(x)g'(x) \, dx = f(x)g(x) - \int g(x)f'(x) \, dx,$$

which follows from integrating—i.e., applying $\int (\cdots) \, dx$ to both sides of—the following rearrangement of the product rule,

$$f(x)g'(x) = [f(x)g(x)]' - g(x)f''(x).$$

Because the product rule for derivatives is what makes the technique of integration by parts valid, many authors describe it as a kind of analog of the product rule, though again, it is not a direct formula for the integral of a product like we had for the derivative of a product.

Still, it is a very useful technique which we will spend some time developing in a later chapter, when we have other methods to draw upon for the inevitable intermediate computations.

8There will be still several other techniques which are not at all simple reverses of derivative rules, and for which checking by differentiating (computing the derivative of) the answer is as difficult as, or more difficult than, the integration technique itself. Those sophisticated techniques are for later chapters.
as above, we have the following pairs of derivative/integral formulas:

\[
\begin{align*}
\frac{d}{dx} \sin x &= \cos x & \iff & \int \cos x \, dx = \sin x + C, \\
\frac{d}{dx} \cos x &= -\sin x & \iff & \int \sin x \, dx = -\cos x + C, \\
\frac{d}{dx} \tan x &= \sec^2 x & \iff & \int \sec^2 x \, dx = \tan x + C, \\
\frac{d}{dx} \cot x &= -\csc^2 x & \iff & \int \csc^2 x \, dx = -\cot x + C, \\
\frac{d}{dx} \sec x &= \sec x \tan x & \iff & \int \sec x \tan x \, dx = \sec x + C, \\
\frac{d}{dx} \csc x &= -\csc x \cot x & \iff & \int \csc x \cot x \, dx = -\csc x + C.
\end{align*}
\]


With these and our previous rules, we have some limited ability to compute integrals involving trigonometric functions.

**Example 6.1.8** Consider the following integrals:

1. \[
\int \left[ x^2 + \sin x - \frac{1}{x} \right] \, dx = \frac{x^3}{3} - \cos x - \ln |x| + C
\]
2. \[
\int \cos w \, dw = \sin w + C
\]
3. \[
\int \frac{\sin x}{\cos^2 x} \, dx = \int \left( \frac{1}{\cos x} \cdot \frac{\sin x}{\cos x} \right) \, dx = \int \sec x \tan x \, dx = \sec x + C
\]
4. \[
\int \tan^2 x \, dx = \int (\sec^2 x - 1) \, dx = \tan x - x + C.
\]

In fact we are fortunate if a trigonometric integral has a form which is just the derivative of one of the six basic trigonometric functions. When it is the case, it often requires some rewriting, as in the latter pair of integrals above.

### 6.1.5 Integrals Yielding Inverse Trigonometric Functions

These follow from derivative formulas, though the third requires some eventual explanation:

\[
\begin{align*}
\int \frac{1}{\sqrt{1 - x^2}} \, dx &= \sin^{-1} x + C, \\
\int \frac{1}{x^2 + 1} \, dx &= \tan^{-1} x + C, \\
\int \frac{1}{x\sqrt{x^2 - 1}} \, dx &= \sec^{-1} |x| + C.
\end{align*}
\]

(6.26) (6.27) (6.28)

Note how we employ only three of the six arctrigonometric functions in (6.26)–(6.28). In fact these are sufficient. Recall for instance that the arccosine and arcsine have derivatives which differ by the factor \(-1\). For simplicity, it is much more commonly written \(\int \frac{1}{\sqrt{1 - x^2}} \, dx = -\sin^{-1} x + C\), rather than using the arccosine function as the antiderivative, i.e., \(\int \frac{1}{\sqrt{1 - x^2}} \, dx = \cos^{-1} x + C\).
though the latter is certainly legitimate. Indeed, since $\sin^{-1} x + \cos^{-1} x = \frac{\pi}{2}$, we see $\cos^{-1} x$ and $-\sin^{-1} x$ differ by a constant. In fact one could rewrite (6.26) as $\int \frac{dx}{\sqrt{1-x^2}} = -\cos^{-1} x + C$.

The choice can sometimes depend upon which range of angles we wish the antiderivative function to output, though for this case we can adjust that with the constant $C$.

Similarly one usually writes $\int \frac{dx}{x^2+1} = -\tan^{-1} x + C$, though $\cot^{-1} x + C$ (for a “different” $C$) is also legitimate. Analogously for arsecant and arccosecant; we usually avoid the arcsecant function as an antiderivative.

But for these last two there is another small complication. Note how Equation (6.28) has the absolute value on the antiderivative rather than the $x$-term of the denominator in the integrand, so it does not appear to be just a restatement of the derivative rule for the arsecant: $\frac{d}{dx} \sec^{-1} x = \frac{1}{x\sqrt{x^2-1}}$. To see that (6.28) is still correct, note that $|x| = x$ if $x > 0$, and $|x| = -x$ if $x < 0$.

Taking the derivative of $\sec^{-1} |x|$ for those two cases, as we did in the computation of $\frac{d}{dx} \ln |x|$ (see page 433) we can see that we do get $\frac{1}{x\sqrt{x^2-1}}$ both times. But it should also be noted that, while not often seen, it would be legitimate to have the absolute value inside, rather than outside, the integral, as in $\int \frac{dx}{x\sqrt{x^2-1}} = \sec^{-1} x + C$.\(^9\)

Before listing some examples, we last make note of the convention mentioned earlier (see page 527), of using $dx$ as a factor. So our new integration formulas are often written:

\[
\int \frac{dx}{\sqrt{1-x^2}} = \sin^{-1} x + C,
\]
\[
\int \frac{dx}{x^2+1} = \tan^{-1} x + C,
\]
\[
\int \frac{dx}{x\sqrt{x^2-1}} = \sec^{-1} |x| + C.
\]

The above formulas will become much more important in future sections. For now it is important to realize that these particular function forms do have (relatively) simple antiderivatives. At this point in the development we are not prepared to make full use of these forms, but we need to be aware of them. Sometimes a simple manipulation produces a function containing one of these forms.

**Example 6.1.9** Compute $\int \frac{x^2}{x^2+1} \, dx$.

**Solution:** With the aid of long division, we can see that\(^{10}\)

\[
\frac{x^2}{x^2+1} = 1 - \frac{1}{x^2+1},
\]

\(^9\)To further complicate things, we could notice that $\sec^{-1} x = \cos^{-1} \frac{1}{x}$, so it can occur that computational software will output an arccosine function, as in $\int \frac{dx}{x\sqrt{x^2-1}} = \cos^{-1} \frac{1}{|x|} + C$, and then this can be rewritten (with a “different $C$”) $\int \frac{1}{x\sqrt{x^2-1}} \, dx = -\sin^{-1} \frac{1}{|x|} + C$. The software might also omit the absolute values, theoretically assuming $x > 0$. Still, the standard written computation would output the expected $\sec^{-1} |x| + C$.

\(^{10}\)A popular alternative technique for a fraction like that in our integrand is to strategically add and subtract a term in the numerator, which produces a term in the numerator identical to (or a multiple of) the denominator, and the extra term, from which we can make two fractions:

\[
\frac{x^2}{x^2+1} = \frac{x^2+1-1}{x^2+1} = \frac{x^2+1}{x^2+1} - \frac{1}{x^2+1} = 1 - \frac{1}{x^2+1}.
\]

While this is a very useful technique for such a simple case, it does not easily extend easily to more complicated cases. Long division—when the degree of the numerator is at least that of the denominator—can always be employed. However, we will eventually have need of a technique similar to that given above in this footnote, though that setting will be much more complicated and we will need all other advanced integration techniques to make full use of it.
and so

\[\int \frac{x^2}{x^2 + 1} \, dx = \int \left[ 1 - \frac{1}{x^2 + 1} \right] \, dx = x - \tan^{-1} x + C.\]

### 6.1.6 Integrals Yielding Exponential Functions

To finish our list of integrals which arise from differentiation formulas, we list those yielding exponential functions. Below \(a \in (0, 1) \cup (1, \infty)\).

\[
\int e^x \, dx = e^x + C, \quad (6.29)
\]
\[
\int a^x \, dx = \frac{a^x}{\ln a} + C. \quad (6.30)
\]

The first of these, (6.29) is the more obvious. Both can be verified through differentiation, as we often do with these simpler antiderivative computations. Recalling that \(\frac{d}{dx} [a^x] = a^x \ln a\), and that \(\ln a\) is a constant, we compute:

\[
\frac{d}{dx} \left[ \frac{a^x}{\ln a} \right] = \frac{1}{\ln a} \cdot \frac{da^x}{dx} = \frac{1}{\ln a} \cdot a^x \ln a = a^x, \quad \text{q.e.d.}
\]

**Example 6.1.10** We compute some antiderivatives involving these and other rules. (Some “simplifications” are matters of preference.)

- \(\int [1 + x + e^x] \, dx = x + \frac{x^2}{2} + e^x + C,\)
- \(\int 2^x \, dx = \frac{2^x}{\ln 2} + C,\)
- \(\int (3 \cdot 2^x + x^2 + e^x + e^x) \, dx = 3 \cdot \frac{2^x}{\ln 2} + \frac{x^3}{3} + e^x + \frac{1}{e + 1} e^x + C,\)
- \(\int \frac{2^x - 3^x}{5^x} \, dx = \int \left[ \frac{2^x}{5^x} - \frac{3^x}{5^x} \right] \, dx = \int \left[ \left( \frac{2}{5} \right)^x - \left( \frac{3}{5} \right)^x \right] \, dx = \frac{\left( \frac{2}{5} \right)^x}{\ln \frac{2}{5}} - \frac{\left( \frac{3}{5} \right)^x}{\ln \frac{3}{5}} + C,\)
- \(\int 5^{2x+1} \, dx = \int 5 \cdot (5^2)^x \, dx = 5 \int 25^x \, dx = \frac{5 \cdot 25^x}{\ln 25} + C = \frac{5^{2x+1}}{2 \ln 5} + C,\)
- \(\int (1 + 3^x)^2 \, dx = \int [1 + 2 \cdot 3^x + 3^{2x}] \, dx = \int [1 + 2 \cdot 3^x + 9^x] \, dx = x + \frac{2 \cdot 3^x}{\ln 3} + \frac{9^x}{\ln 9} + C.\)

In the last integral, we used the fact that \((3^x)^2 = 3^{2x} = 3^{2x} = (3^2)^x = 9^x.\)

In later sections we develop the initial techniques which allow us to find antiderivatives of functions which are not known to be derivatives of common functions, but are related to them. In particular we look for reverse chain rules but the technique to do so has many more applications, and is used extensively in the rest of the text.
### Exercises

Verify the following antiderivative formulas by differentiation (applying \( \frac{d}{dx} \) to the right-hand sides).

1. \[ \int e^{2t} \, dt = \frac{1}{2} e^{2t} + C \] 
2. \[ \int \frac{e^{\sqrt{x}}}{\sqrt{x}} \, dx = 2e^{\sqrt{x}} + C \] 
3. \[ \int \cos 2x \, dx = \frac{1}{2} \sin 2x + C \] 
4. \[ \int \ln x \, dx = x \ln x - x + C \] 
5. \[ \int \sec x \, dx = \ln |\sec x + \tan x| + C \] 
6. \[ \int \tan^{-1} x \, dx = x \tan^{-1} x - \frac{\ln(x^2 + 1)}{2} + C \]
7. \[ \int \sin^2 x \, dx = \frac{1}{2} (x - \sin x \cos x) + C \] 
8. \[ \int \sec 7x \tan 7x \, dx = \frac{1}{7} \sec 7x + C \] 
9. \[ \int \tan x \, dx = \ln |\sec x| + C \] 
10. \[ \int xe^x \, dx = xe^x - e^x + C \] 
11. \[ \int \frac{1}{x (1 + (\ln x)^2)} \, dx = \tan^{-1}(\ln x) + C \] 
12. \[ \int \frac{2}{x^2 - 1} \, dx = \ln \left|\frac{x - 1}{x + 1}\right| + C \] (Hint: expand the logarithms first.)

Calculate the following indefinite integrals, in most cases by rewriting the integrand to achieve a form found earlier in the section.

13. \[ \int \frac{1}{x^{2/3}} \, dx \] 
14. \[ \int \frac{1}{\cos^2 x} \, dx \] 
15. \[ \int (x^2 + 3x + 9) \, dx \] 
16. \[ \int (x^2 + 1)^3 \, dx \] 
17. \[ \int \sqrt{9w} \, dw \] 
18. \[ \int \frac{(x + 1)^2}{x^2} \, dx \] 
19. \[ \int (-5 \sin x) \, dx \] 
20. \[ \int 10^x \, dx \] 
21. \[ \int \cot^2 t \, dt \]

22. Suppose \( a(t) = -2 \cos t, \ v(0) = 3 \) and \( s(0) = 7 \). Find \( s(t) \).

For 23–28, find the function satisfying the given criteria.

23. \( f'(x) = 3x^2 + 2x + 5, \ f(1) = 2 \) 
24. \( g'(z) = \sqrt{z}, \ g(4) = \frac{16}{5} \) 
25. \( s'(t) = \frac{1}{t^2 + 1}, \ s(1) = \pi/2 \) 
26. \( f'(x) = 1 + \cos x, \ f(\pi/2) = 6 \) 
27. \( f'(x) = \sec x \tan x, \ f(0) = 4 \) 
28. \( f'(x) = 5e^x, \ f(\ln 3) = 11 \)
6.2 Summation (Sigma) Notation

This short section looks at the Σ-notation common for signifying summations. For example,

\[ \sum_{i=1}^{n} a_i = a_1 + a_2 + a_3 + \cdots + a_{n-1} + a_n. \]  \hspace{1cm} (6.31)

It is no accident that the “Greek S” is used to signify a sum. In the above, \( i \) is the index of summation, ranging along the integers from 1 to \( n \), i.e., \( i = 1, 2, 3, \ldots, n \). If \( a_1 = 3 \), \( a_2 = 5 \),
\( a_3 = 1 \), \( a_4 = 7 \), and \( a_5 = -2 \), we can write

\[ \sum_{i=1}^{5} a_i = a_1 + a_2 + a_3 + a_4 + a_5 = 3 + 5 + 1 + 7 - 2 = 14. \]

Often the summation desired is of the form \( \sum_{i=1}^{n} f(i) \), where \( f: \{1, 2, \ldots, n\} \rightarrow \mathbb{R} \), though the domain of \( f \) need not begin at \( i = 1 \). For instance,

\[ \sum_{i=1}^{5} i^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 1 + 4 + 9 + 16 + 25 = 55, \]
\[ \sum_{i=0}^{3} (2i + 1) = \sum_{i=0}^{3} 2i + \sum_{i=0}^{3} 1 = [2(0) + 1] + [2(1) + 1] + [2(2) + 1] + [2(3) + 1] = 1 + 3 + 5 + 7 = 16, \]
\[ \sum_{i=0}^{4} 2i = \sum_{i=0}^{4} 2i = 2 + 4 + 6 + 8 + 10 = 5(2) = 10. \]

Some care must be taken regarding the range of values for the index, which in the above is \( i \) for each example. Other indices are also common, usually lower-case Latin or Greek letters. It should also be pointed out that these summations are, in many ways, operators, notationally similar to the derivative and integral operators \( \frac{d}{dx} [ \ ] \) and \( \int \) \( dx \), so some of the same notational conventions apply. Consider for instance the following:

\[ \sum_{i=1}^{5} i + 9 = \left( \sum_{i=1}^{5} i \right) + 9 = (1 + 2 + 3 + 4 + 5) + 9 = 24, \]
\[ \sum_{i=1}^{5} (i + 9) = (1 + 9) + (2 + 9) + (3 + 9) + (4 + 9) + (5 + 9) = 10 + 11 + 12 + 13 + 14 = 60, \]
\[ \sum_{i=1}^{3} 2i = \sum_{i=1}^{3} 2i = 2 + 4 + 6, \]
\[ \sum_{i=1}^{5} (i + 1)^2 = 2^2 + 3^2 + 4^2 + 5^2 + 6^2 = 4 + 9 + 16 + 25 + 36 = 90, \]
\[ \left( \sum_{i=1}^{5} (i + 1) \right)^2 = (2 + 3 + 4 + 5 + 6)^2 = 20^2 = 400. \]

The notation has many uses. It can be found nearly anywhere a large number of similar quantities are routinely added, such as in accounting or on spreadsheets. In statistics, if we have
6.2. **SUMMATION (SIGMA) NOTATION**

Data $x_1, x_2, \ldots, x_n$, then the sample mean (or average) is given by

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i,$$

while the sample standard deviation is then given by

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}}.$$

Some arithmetic properties of summations can be seen with little difficulty. In particular, we have

$$\sum_{i=1}^{n} (a_i + b_i) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i,$$

$$\sum_{i=1}^{n} ka_i = k \cdot \sum_{i=1}^{n} a_i,$$

$$\sum_{i=1}^{n} k = nk.$$

The first is simply a regrouping, the second the distributive law, while the third is a simple counting of terms added. These are more clearly demonstrated below, with the aid of some extra parentheses:

$$\sum_{i=1}^{n} (a_i + b_i) = (a_1 + b_1) + (a_2 + b_2) + \cdots + (a_n + b_n)$$

$$= (a_1 + a_2 + \cdots + a_n) + (b_1 + b_2 + \cdots + b_n) = \left( \sum_{i=1}^{n} a_i \right) + \left( \sum_{i=1}^{n} b_i \right),$$

$$\sum_{i=1}^{n} (ka_i) = ka_1 + ka_2 + \cdots + ka_n$$

$$= k (a_1 + a_2 + \cdots + a_n) = k \cdot \left( \sum_{i=1}^{n} a_i \right),$$

$$\sum_{i=1}^{n} k = k + k + \cdots + k = n \cdot k, \quad \text{q.e.d.}$$

Note that in (6.34), the number of terms is important, so that $\sum_{i=0}^{n} k = (n + 1)k$, while $\sum_{i=2}^{5} k = 4k$. In all cases the summation is a sum of several copies of the same constant.

It is interesting to note that there are formulas which one can derive for sums of positive
powers of the index. The first few are as follow:

\[ \sum_{i=1}^{n} i = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}, \]  
(6.35)

\[ \sum_{i=1}^{n} i^2 = 1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}, \]  
(6.36)

\[ \sum_{i=1}^{n} i^3 = 1^3 + 2^3 + 3^3 + \cdots + n^3 = \frac{n^2(n+1)^2}{4}. \]  
(6.37)

We will prove (6.37) after some examples, and leave the computations of the others to the exercises. With these, we can compute for example

\[ \sum_{i=1}^{5} i^2 = \frac{5(5+1)(2(5)+1)}{6} = \frac{5 \cdot 6 \cdot 11}{6} = 55, \]

which we can also compute directly: \(1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 55\). With these formulas (6.35)–(6.37) immediately above, together with the more obvious formulas (6.32)–(6.34), we can more readily compute some commonly occurring summations without resorting to adding each individual term.

**Example 6.2.1** Consider the following computations (see (6.35)–(6.37) and (6.32)–(6.34)):

- \( \sum_{i=1}^{100} i = \frac{100(100+1)}{2} = 5050 \)
- \( \sum_{i=1}^{10} (3 + i^2) = \sum_{i=1}^{10} 3 + \sum_{i=1}^{10} i^2 = 10 \cdot 3 + \frac{10(10+1)(2(10)+1)}{6} \]
  \[ = 30 + \frac{10 \cdot 11 \cdot 21}{6} = 30 + 385 = 415, \]
- \( \sum_{i=1}^{40} [(2i + 3)^2] = \sum_{i=1}^{40} [4i^2 + 12i + 9] = 4 \sum_{i=1}^{40} i^2 + 12 \sum_{i=1}^{40} i + \sum_{i=1}^{40} 9 \]
  \[ = 4 \cdot \frac{(40)(41)(81)}{6} + 12 \cdot \frac{(40)(41)}{2} + 40 \cdot 9 = 88,560 + 9,840 + 360 = 98,760, \]
- \( \sum_{i=1}^{100} i^2 = \sum_{i=1}^{100} i^2 - \sum_{i=1}^{49} i^2 = \frac{(100)(101)(201)}{6} - \frac{(49)(50)(99)}{6} \]
  \[ = 338,350 - 40,425 = 29,7925. \]

*This last summation we were able to rewrite as a difference, the terms we do not need having cancelled.*

Sometimes there is value in “adjusting the indices.” While this is not well motivated in this section, we consider it here for purposes of early introduction. Note that the following
sums are the same:

\[ \sum_{i=1}^{5} a_i = a_1 + a_2 + a_3 + a_4 + a_5, \]

\[ \sum_{i=0}^{4} a_{i+1} = a_1 + a_2 + a_3 + a_4 + a_5, \]

\[ \sum_{i=0}^{4} a_{i-2} = a_3 + a_4 + a_5 + a_6 + a_7. \]

This can simplify computations such as the following:

\[ \sum_{i=1}^{10} (i + 3)^2 = \sum_{i=4}^{13} i^2 \]

(check the actual numbers being summed)

\[ = \sum_{i=1}^{13} i^2 - \sum_{i=1}^{3} i^2 \]

\[ = \frac{(13)(14)(27)}{6} - \frac{(3)(4)(7)}{6} = \frac{1}{6} (4914 - 84) = \frac{4830}{6} = 805. \]

While the above example may not be all that compelling, with larger numbers it is likely to be more convincing. Furthermore, this technique of re-indexing the numbers in the sum has its usefulness in many other contexts, and will be refined and expanded as needed later.

Finally, for completeness we now look at the typical method of proving (6.37), page 540. This type of formula is usually proven using mathematical induction. What this entails is (1) directly proving one or more of the “first” cases, and (2) showing that anytime we have, say, the \( n \)th case we automatically get the \( (n + 1) \)st case. So we define statements \( P_1, P_2, P_3, \ldots \) so that the statement \( P_n \) is defined as follows:

\[ P_n \rightarrow \sum_{i=1}^{n} i^3 = \frac{n^2(n + 1)^2}{4}. \]

Now on its face, \( P_n \) could be true or false for a particular \( n \), but our strategy will be to show that

1. \( P_1 \) is true, and that
2. \( P_n \implies P_{n+1} \).

This second part is called the induction step.

Before continuing, an explanation of mathematical induction is in order. As a general rule, if we prove \( P_1 \) is true, and that for all \( n \in \mathbb{N} \) we have \( P_n \implies P_{n+1} \) (i.e., \( P_n \) being true would force \( P_{n+1} \) to be true also), then by this implication (2) since \( P_1 \) is true we must have that \( P_2 \) is true, and by the implication (2) we further have \( P_3 \) is true, and from that we next have \( P_4 \) is true, and so on. This eventually proves, say, \( P_{1,000,000} \) is true because there is this “chain of truth” from \( P_1 \) to \( P_2 \) to \( P_3 \) and so on, so after 999,999 such steps we would reach the conclusion that \( P_{1,000,000} \) is also true. We are forced to conclude then that \( P_n \) is true for each \( n \in \mathbb{N} \), because the truth of \( P_n \) would be reached after \( n - 1 \) invocations of the induction step (2). That is the essence of the strategy known as “proof by induction” in mathematics, though there are variations of it.
Getting back to this particular proof, the statement $P_1$ would be that
\[ \sum_{i=1}^{1} i^3 = \frac{1^2(2)^2}{4}, \]
which is clearly true because it is equivalent to $1^3 = \frac{1^2(2)^2}{4}$, i.e., $1 = 1$, which is true (obviously).

The induction step (2) has a simple, yet sophisticated little proof. We want to show $P_n \implies P_{n+1}$, so to do that we suppose (hypothetically) $P_n$, i.e., that the “$n$th case” is true, and then show that this would imply the “$(n+1)$st” case follows.

So if $P_n$ is true, i.e. \[ \sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4}, \]
then
\[ \sum_{i=1}^{n+1} i^3 = \left( \sum_{i=1}^{n} i^3 \right) + (n+1)^3 = \frac{n^2(n+1)^2}{4} + (n+1)^3 = \left( n^2 + 4(n+1) \right) \frac{1}{4}, \]
which is the statement $P_{n+1}$ when examined closely:
\[ P_{n+1} : \quad \sum_{i=1}^{n+1} i^3 = \frac{(n+1)^2((n+1)+1)^2}{4}. \]

Thus we showed $P_n \implies P_{n+1}$, proving the induction step (2), and so with $P_1$ being true we get $P_n$ is true for all $n \in \mathbb{N}$, q.e.d.

Since this proof served the purpose to introduce induction as a proof style, as well as to use the method for a particular example (proving (6.37)), it is more verbose than it would be if the author assumes the readers know the technique. It is a deep enough method that most students are only expected to become comfortable with it after repeated exposure.\(^{11}\)\(^{12}\)

\(^{11}\)A more streamlined proof by induction would look more like the following.

**Theorem:** For any $n \in \mathbb{N}$, \[ \sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4} \]

**Proof:** First we note that it is true for $n = 1$: $1^3 = \frac{1^2(2)^2}{4}$. Next we assume it is true for the $n$th case. Then
\[ \sum_{i=1}^{n+1} i^3 = \sum_{i=1}^{n} i^3 + (n+1)^3 = \frac{n^2(n+1)^2}{4} + (n+1)^3 = \left( n^2 + 4(n+1) \right) \frac{1}{4}, \]
Thus by induction the formula holds for all $n \in \mathbb{N}$, q.e.d.

\(^{12}\)Another variation of induction is where one begins by proving the first few cases, say $P_1, P_2, P_3$, and then uses a “weaker” induction step, where the implication might use more than the information from the statement immediately prior, i.e., where instead of $(\forall n \in \mathbb{N})[P_n \implies P_{n+1}]$, one uses
\[ [(\forall k \in \mathbb{N}) (k \leq M \implies P_k)] \implies P_{M+1}. \]
In other words, the induction step is to show that the next statement $P_{M+1}$ is true under the assumption that all previous statements in the list $P_1, P_2, \ldots, P_M$ are true, rather than just $P_M$.

The basic spirit is that the truth of the later statements is “bootstrapped” off of the truths of the previous statements, and that for any statement $P_n$, no matter how large is $n \in \mathbb{N}$, its truth will be established in finitely many valid implication steps from the truths of the previous statements. Thus there are many possible variations of structure for induction proofs. (Some even prove even and odd cases separately, for instance.)
6.2. **SUMMATION (SIGMA) NOTATION**

**Exercises**

For Exercises 1–8, compute the sum by writing out every term and simplifying the result.

1. \[\sum_{i=1}^{6} i\]

2. \[\sum_{i=0}^{5} 2^i\]

3. \[\sum_{i=0}^{5} 2^{-i}\]

4. \[\sum_{i=4}^{10} i^2\]

5. \[\sum_{i=1}^{11} (-1)^i\]

6. \[\sum_{i=1}^{12} \cos \frac{n\pi + 1}{2}\]

7. \[\sum_{i=1}^{6} \left[\frac{1}{i+1} - \frac{1}{i-1}\right]\] (notice what cancels when you write out the terms in long-hand)

8. \[\sum_{i=1}^{10} 2 + 3\]

For Exercises 9–13, use techniques similar to those of Example 6.2.1, page 540 to compute the sums:

9. \[\sum_{i=1}^{200} i\]

10. \[\sum_{i=100}^{200} i\]

11. \[\sum_{i=1}^{20} (2i + 1)\]

12. \[\sum_{i=1}^{10} [i(i - 1)]\]

13. \[\sum_{i=1}^{20} [i(2i - 3)^2]\]

14. By writing out all of the terms, prove the following and compute the sum (using the second form of the sum):

\[\sum_{i=0}^{9} (i + 1)^3 = \sum_{i=1}^{10} i^3.\]

15. By writing out several terms, show that the following computation is valid and compute the sum:

\[\sum_{i=14}^{30} (i - 1)^3 = \sum_{i=13}^{29} i^3\]

\[= \sum_{i=1}^{29} i^3 - \sum_{i=1}^{12} i^3.\]

16. Using a strategy similar to the problems above, compute the following:

(a) \[\sum_{i=3}^{20} (i - 2)^2\]

(b) \[\sum_{i=5}^{20} (i + 2)^2\] (for this one, you may wish to write a summation as a difference of two other summations)

17. Prove (6.35) by induction.

18. Prove (6.36) by induction.
6.3 Riemann Sums and the Fundamental Theorem of Calculus

Anytime a theorem is called “fundamental” in its field, we expect it to be somewhat deep, ultimately intuitive, very important, and not trivial to prove. These all apply to the Fundamental Theorem of Calculus (FTC), as discussed here. An actual proof of the theorem is beyond the scope of this text, and will not be found here.\footnote{Most if not all science and engineering calculus textbook authors attempt an argument for why the FTC is true. Some give partial proofs which are as intuitive as possible, while others give proofs that are more technical, but closer to an actual proof. So far, none have offered a complete proof without having one large gap which requires junior or senior level Real Analysis to fill. This textbook is no different. Here we opt for intuitive arguments, and later outline a proof which is closer to an actual proof, but neither the intuitive, nor the more technical, of the arguments given here constitute a rigorous proof. That is left for junior or senior level classes.}

In fact, the theorem presented here is technically known as the Second Fundamental Theorem of Calculus, because its proof usually comes after the proof of the First Fundamental Theorem of Calculus, discussed later. However, this “Second” theorem is used more than the first, and is arguably more intuitive, and will therefore be what we mean in most cases when discussing “The Fundamental Theorem of Calculus.”

Instead of attempting a proof, we present a case where it is, more or less, obvious (or at least very believable), and then generalize somewhat to less obvious cases. Along the way there are several concepts to define and explore, and the explanation bears careful study and several revisits.

We begin with the twin concepts of relative and percent error, show how they stay controlled within summations, show how the study of approximate displacements leads us to Riemann Sums and one case of the FTC, and then generalize for the full conclusion of one part of the FTC. There is another part, which we leave for a later section.
6.3. RIEMANN SUMS AND THE FUNDAMENTAL THEOREM OF CALCULUS

6.3.1 Absolute, Relative and Percent Errors

For a simple example of these three types of errors, consider a man who weighs 200 lbs, weighed on a scale which indicates his weight to be 210 lbs. In such a case we say the absolute error is 10 lbs. For a 200 lbs man this seems “relatively” small, but if we are weighing a newborn child, 10 lbs is clearly an unacceptable error. Thus it is also important to note what fraction of his weight the error represents, so we compute the relative error, namely \( \frac{10 \text{ lbs}}{200 \text{ lbs}} \), or 0.05. Now as a percentage (or “parts per hundred”), we multiply by 100\%\) (which is just another expression for 1) and find the percent error to be 5\%. Note that the relative and percent error are unitless, in the sense that the “lbs” cancel.

Put colloquially, these errors are defined as follow:

\[
\text{Absolute Error} = (\text{Actual Quantity}) - (\text{Measured Quantity}) \tag{6.38}
\]

\[
\text{Relative Error} = \frac{\text{Absolute Error}}{(\text{Actual Quantity})} \tag{6.39}
\]

\[
\text{Percent Error} = \left( \frac{\text{Relative Error}}{1} \right) \cdot 100\% \tag{6.40}
\]

Some texts will define the absolute error to include absolute values of the quantity on the right-hand side of (6.38), hence the name. One can then also define relative absolute error, and percent absolute error, in the obvious ways. However in all cases it is informative to have a sign \((+/−)\) associated to the error, and so we will still use the term error rather than absolute (in the sense of absolute values) error, when there is no confusion.

Relative and percent errors are easily visualized, and in fact judging when a relative or percentage error is “small” or “large” is fairly easy given an accurate illustration of the quantities involved. See Figure 6.2, page 544.

Now we consider what would be the cumulative effect on a summation if the measured amount were consistently a given percentage higher than the actual.

Example 6.3.1 Suppose the actual quantity we desire to know is \( \sum_{i=1}^{n} a_i \), where we attempt to measure each of the \( a_i \), and each is measured to be \( b_i \), where \( b_i \) is a 5\% overestimation of \( a_i \) in each case. Then the measured summation will be

\[
\sum_{i=1}^{n} b_i = \sum_{i=1}^{n} (1.05a_i) = 1.05 \sum_{i=1}^{n} a_i.
\]

In other words, if the \( b_i \) all overestimate the respective \( a_i \) by exactly 5\%, then the summation of the \( b_i \) overestimates the summation of the \( a_i \) by exactly 5\%.

The example above simply illustrates the distributive property of multiplication over summations. We can conclude similarly that a consistent underestimation of \( a_i \) by 5\% would result in exactly a 5\% underestimation of the summation. More generally, since these would be the “extreme” cases, we can further state that if, in the sense of absolute values, the percent absolute error in \( a_i \) is less than 5\% \((+/−)\), then the percent absolute error in the sum must also be less than 5\% \((+/−)\). This leads us to the more general conclusion:

Theorem 6.3.1 If each \( a_i \) is estimated by a respective \( b_i \) within \( p\% \) error, then it follows that \( \sum a_i \) is also estimated within \( p\% \) error by \( \sum b_i \).

That fact will be important in the next step of our argument for the validity of the FTC.
### 6.3.2 A Physics Example

Here we consider an abstract motion problem. We wish to find the net displacement of an object in one-dimensional motion, over the time interval \([t_0, t_f]\). For a classical problem, the velocity \(v(t)\) over this time interval should be continuous. For technical reasons explained later, we also assume that it is positive, i.e., \(v(t) > 0\) on \([t_0, t_f]\).

Now the actual net displacement over the time interval is given by \(s(t_f) - s(t_0)\). Intuitively it will also be positive, since the velocity is assumed to be positive.

Now we consider a scheme for approximating this net displacement, based upon the velocity function. We do this by partitioning the time interval into subintervals with endpoints

\[ t_0 < t_1 < t_2 < t_3 < \cdots < t_{n-1} < t_n = t_f. \]

The width of the \(i\)th subinterval \([t_{i-1}, t_i]\) will then be

\[ \Delta t_i = t_i - t_{i-1}. \]  \(\text{(6.41)}\)

If that interval is short enough, then the velocity change over the interval will be small, and in fact we expect the percent change in the velocity to be small, giving a small percent error in assuming velocity is approximately constant. A small percentage error resulting from assuming \(v \approx v(t^*_i)\) for some \(t^*_i \in [t_{i-1}, t_i]\) will allow us to assume that to the same level of percentage error, the net displacement over that interval can be approximated by

\[ s(t_i) - s(t_{i-1}) \approx v(t^*_i)\Delta t_i, \]

where again \(t^*_i \in [t_{i-1}, t_i]\) is a point in the interval at which we sample the velocity.

Our scheme is thus to approximate the net displacement on each subinterval \([t_{i-1}, t_i]\), and sum these.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Sample</th>
<th>Width</th>
<th>Approximate Displacement</th>
<th>Actual Displacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>([t_0, t_1])</td>
<td>(t^*_1)</td>
<td>(\Delta t_1)</td>
<td>(v(t^*_1)\Delta t_1)</td>
<td>(s(t_1) - s(t_0))</td>
</tr>
<tr>
<td>([t_1, t_2])</td>
<td>(t^*_2)</td>
<td>(\Delta t_2)</td>
<td>(v(t^*_2)\Delta t_2)</td>
<td>(s(t_2) - s(t_1))</td>
</tr>
<tr>
<td>([t_2, t_3])</td>
<td>(t^*_3)</td>
<td>(\Delta t_3)</td>
<td>(v(t^*_3)\Delta t_3)</td>
<td>(s(t_3) - s(t_2))</td>
</tr>
<tr>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
</tr>
<tr>
<td>([t_{n-2}, t_{n-1}])</td>
<td>(t^*_{n-1})</td>
<td>(\Delta t_{n-1})</td>
<td>(v(t^*<em>{n-1})\Delta t</em>{n-1})</td>
<td>(s(t_{n-1}) - s(t_{n-2}))</td>
</tr>
<tr>
<td>([t_{n-1}, t_n])</td>
<td>(t^*_n)</td>
<td>(\Delta t_n)</td>
<td>(v(t^*_n)\Delta t_n)</td>
<td>(s(t_n) - s(t_{n-1}))</td>
</tr>
</tbody>
</table>

When we now sum the last two columns, respectively, we get much cancellation in the last column, resulting in the approximation:

\[
\sum_{i=1}^{n} v(t^*_i)\Delta t_i \approx (s(t_1) - s(t_0)) + (s(t_2) - s(t_1)) + (s(t_3) - s(t_2)) \\
+ \cdots + (s(t_{n-1}) - s(t_{n-2})) + (s(t_n) - s(t_{n-1})),
\]

which, after the mostly “middle” terms cancel, simplifies to

\[
\sum_{i=1}^{n} v(t^*_i)\Delta t_i \approx s(t_f) - s(t_0). \]  \(\text{(6.42)}\)

However it is not clear how good the above approximation actually is. For that we turn to our earlier note, that it is reasonable we can choose intervals small enough that the velocity changes
6.3. Riemann Sums and the Fundamental Theorem of Calculus

no more than $p \% \,(+/\-)$, for any $p > 0$. In doing so, we are assured that the net displacement
over each interval is within $p \%$ of the actual for that interval, and so by our Theorem 6.3.1,
page 545, the sum on the left of (6.42) is within $p \%$. Now we reason that the percent error
will shrink to zero as the interval lengths shrink to zero, and we will argue that\footnote{There is one caveat to this reasoning, which is that on an interval where the velocity may be momentarily zero, our choice of $v(t^*_i)$ could be off by 100\%, and if the actual displacement were zero and we chose some $t^*_i$ such that $v(t^*_i) \neq 0$, then our error is an infinite percent. Thus we have to rely on the fact that we can then choose the interval small enough that the absolute error is as small as we like, and keep the percentage error small in the other intervals. This is partially alleviated by our assumption that $v(t) > 0$ on $[t_0, t_f]$, but we would like our analysis to work under less restrictive conditions. Later illustrations will help show that this ability is reasonable.}

$$\lim_{\max(\Delta x_i \to 0^+)} \sum_{i=1}^{n} v(t^*_i) \Delta t_i = s(t_f) - s(t_0).$$

At this point we note that $s(t)$ is an antiderivative of $v(t)$, which is no accident. More
generally, we will have (6.46) below for any continuous function $f : [a, b] \to \mathbb{R}$.

6.3.3 General Riemann Sums, FTC

The sum on the right-hand side of (6.42) is one example of what is known as a Riemann Sum.\footnote{Named for Georg Friedrich Bernhard Riemann, 1826–1866, a German mathematician with very important contributions to calculus and differential geometry, the latter of which laid important groundwork for later physicists, such as Albert Einstein in his derivation of the equations of general relativity. Riemann’s work is therefore one example of how the work of curious mathematicians can produce mathematical results which long predate many real-world physical problems which give the mathematics its deeper relevance.}

More generally, for $f(x)$ defined on $[a, b]$, we define Riemann sums to be any sum of the form

$$\sum_{i=1}^{n} f(x_i^*) \Delta x_i,$$

where we partition $[a, b]$ into subintervals with endpoints

$$a = x_0 < x_1 < x_2 < x_3 < \cdots < x_{n-1} < x_n = b,$$

and $\Delta x_i = x_i - x_{i-1}$ is the width of the $i$th subinterval. What we have argued in the context of velocities is actually one part of the Fundamental Theorem of Calculus (or FTC):

**Theorem 6.3.2** (Fundamental Theorem of Calculus, Part 1): For $f(x)$ continuous on $[a, b]$, and $F(x)$ being an antiderivative of $f(x)$ on $[a, b]$, and Riemann sums as above, we have

$$\lim_{\max(\Delta x_i \to 0^+)} \left( \sum_{i=1}^{n} f(x_i^*) \Delta x_i \right) = F(b) - F(a).$$

As written above in (6.45), for the moment we will assume each $\Delta x_i > 0$, or we would more carefully write our limit to be as $\max \{|\Delta x_i|\} \to 0^+$. Also note that $\max(\Delta x_i) \to 0^+ \implies n \to \infty$, so we shrink all the subintervals’ lengths, and therefore increase their number. At this point we introduce very important some notation (which the reader should memorize eventually):

$$\int_{a}^{b} f(x) \, dx \overset{\text{definition}}{=} \lim_{\max(\Delta x_i \to 0^+)} \left( \sum_{i=1}^{n} f(x_i^*) \Delta x_i \right),$$

$$F(x) \bigg|_{a}^{b} \overset{\text{definition}}{=} F(b) - F(a).$$
With definitions (6.47) and (6.48), we can rewrite the Fundamental Theorem of Calculus (6.46) as

\[
\int_a^b f(x) \, dx = F(b) - F(a).
\]

(6.49)

To distinguish the integral symbol \( \int \) in this context from its use in Section 6.1, the quantity on the left-hand sides of (6.47) and (6.49) is called the definite integral of \( f(x) \) with respect to \( x \), from \( x = a \) to \( x = b \).

**Definition 6.3.1** For a function \( f(x) \), continuous on \([a, b]\), define the definite integral of \( f \) over the interval \([a, b]\) by the following notation and its numerical definition given by the equation

\[
\int_a^b f(x) \, dx \overset{\text{definition}}{=} \lim_{\max(\Delta x_i) \to 0^+} \left( \sum_{i=1}^n f(x_i^*) \Delta x_i \right).
\]

By the Fundamental Theorem of Calculus, this can be computed using (6.49).

Of course the FTC gives a very strong connection between the two uses of the symbol \( \int \), used for antiderivatives when no endpoints are given, and for the limit above which is the same as the difference of any antiderivative as evaluated at the endpoints \( a, b \).

Note that (6.49) requires that \( f \) be continuous on \([a, b]\), and \( F \) be an antiderivative there, in the sense of Definition 6.1.1, page 524.

The geometric interpretation of this limit is that it describes the “signed area” between a function \( f(x) \) and the \( x \)-axis, along the interval \( x \in [a, b] \). Recall that a function \( f(x) \) gives the height of the curve \( y = f(x) \) at a specific value of \( x \). This “height” can be positive, negative or zero at a given value of \( x \). For the moment we only consider nonnegative functions, with therefore nonnegative heights, which yield nonnegative areas bounded on one side by the graph of the given function, and on the other side by the \( x \)-axis over the given interval \([a, b]\).

Since the heights of a function tend to vary, we cannot simply use a “base times height” formula for computing one such area in question. However we can approximate the area using rectangles whose heights are derived from the function, and whose bases lie along the \( x \)-axis. As before, we break the interval \([a, b]\) in question into a partition of \( n \) subintervals with \( n + 1 \) endpoints \( x_0, x_1, \ldots, x_n \) so that

\[
a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b,
\]

and sample the height of the function on each interval, by choosing \( n \) values \( x_i^* \in [x_{i-1}, x_i] \), whose height is \( f(x_i^*) \), to represent the height of an approximating rectangle for the area between the function’s graph and the \( i \)th interval \([x_{i-1}, x_i]\). The area of this \( i \)th approximating rectangle will be \( f(x_i^*) \Delta x_i \), where

\[
\Delta x_i = x_i - x_{i-1}.
\]

(6.50)

Adding the areas of all such approximating rectangles gives us a Riemann Sum approximation of the total area between the curve and the interval \([a, b]\) on the \( x \)-axis:

\[
\text{Shaded Area} \approx \sum_{i=0}^n f(x_i^*) \Delta x_i.
\]

One such approximation scheme is illustrated in Figure 6.3, page 549. That scheme uses \( x_i^* \) to be the midpoint of the \( i \)th interval \([x_{i-1}, x_i]\). It also uses a constant width \( \Delta x_i \) for each interval \([x_{i-1}, x_i]\).

---

16 The endpoints \( a \) and \( b \) in the definite integral are often referred to as the lower and upper limits of integration, perhaps an unfortunate term since “limit” usually refers to very different concepts. Perhaps better words in this context would be boundary points, or endpoints or terms of similar spirit.
6.3. RIEMANN SUMS AND THE FUNDAMENTAL THEOREM OF CALCULUS

Area \approx f(x_1^*)(x_1 - x_0) + f(x_2^*)(x_2 - x_1) + f(x_3^*)(x_3 - x_2) + f(x_4^*)(x_4 - x_3) \\
= f(x_1^*)\Delta x_1 + f(x_2^*)\Delta x_2 + f(x_3^*)\Delta x_3 + f(x_4^*)\Delta x_4 \\
= \sum_{i=1}^{4} f(x_i^*)\Delta x_i, \quad \text{where } \Delta x_i = x_i - x_{i-1}.

Figure 6.3: Figure for general Riemann Sum, in the case of a positive function \( f \). The actual area between the curve and the \( x \)-axis on some interval \([a, b]\) is approximated by a sum of areas of rectangles, where for each subinterval interval \([x_{i-1}, x_i]\) is approximated by sampling one height \( f(x_i^*) \) of the function in the interval, with \( x_i^* \in [x_{i-1}, x_i] \) (the \( i \)th subinterval). The area of the \( i \)th rectangle will be \( f(x_i^*)(x_i - x_{i-1}) = f(x_i^*)\Delta x_i \). When we add these together we get a Riemann Sum, approximating the total area.

The next figure, namely Figure 6.4 shows two schemes for approximating the same area. In both, a right-endpoint approximation is used, where \( x_i^* = x_i \), which has the advantage of simplicity and is therefore the most common, but has the disadvantage that it is often unlikely that the right endpoint of an interval is likely be where we expect to find the “average” height to be found for that interval.

Nonetheless, it is not difficult to see that whatever rule we use for choosing the \( x_i^* \) values, as the width of rectangles decreases and consequently the number of rectangles increases, so does the accuracy of the Riemann sums increase in approximating the actual area.\(^{17}\) Indeed, Figure 6.4 shows how much error can be reduced when the number of rectangles increases. In that case, since the function is increasing, using the right-endpoint method whereby \( x_i^* = x_i \) we get that the Riemann Sums overestimate the actual areas. However, we decrease the percent error when we increase the number of rectangles. According to the Fundamental Theorem of Calculus, when we let \( \max\{\Delta x_i\} \to 0^+ \), and therefore \( n \to \infty \) we will get a value which is equal to \( F(b) - F(a) \), where the original interval is \([a, b]\) and \( F \) is an antiderivative of the function \( f \) on that interval. Intuitively (looking graphically at our approximation schemes), it seems also

\(^{17}\)A similar phenomenon occurs with Riemann Sums used to approximate displacement

\[ s(t_f) - s(t_0) \approx \sum_{i=1}^{n} v(t_i^*)\Delta t_i. \]

Our argument here will be that we can make the percent error small in each time interval by shrinking the maximum allowable size of all intervals. If the percent error is small on each interval, so will be the percent error of the sum, and our approximation above will be within that percent error.
true that as $\max\{\Delta x_i\} \to 0^+$ and $n \to \infty$ we also get

$$
\sum f(x_i^*) \Delta x_i = \sum f(x_i) \Delta x_i \to \text{Shaded Area}.
$$

The FTC applies for any choices of $x_i^* \in [x_{i-1}, x_i]$, and so applies to the case $x_i^* = x_i$, and so by FTC the shaded area should be equal to $F(b) - F(a)$:

\[
\text{Shaded Area} = \lim_{\max\{\Delta x_i\} \to 0^+} \left( \sum_{i=1}^{n} f(x_i^*) \Delta x_i \right) \\
= \lim_{n \to \infty} \left( \sum_{i=1}^{n} f(x_i) \Delta x_i \right) \\
= \text{Definition} \int_{a}^{b} f(x) \, dx \quad \text{FTC} \quad F(x) \bigg|_{a}^{b} \quad \text{Definition} \quad F(b) - F(a).
\]

In our case the star above can be removed and $x_i$ inserted for $x_i^*$.

### 6.3.4 Computing Areas: Geometric Interpretation of FTC

The above is a very long argument, which the reader is advised to revisit frequently. The upshot is that the geometric interpretation of $\int_{a}^{b} f(x) \, dx$ is that this represents the signed area between the curve $y = f(x)$, $a \leq x \leq b$, and that the $x$-axis; the function $f(x)$ gives the height at each $x \in [a, b]$, with the “base” (of the region whose area we are computing) being the interval $[a, b]$ as it is contained within the $x$-axis. See again Figure 6.4.

However, that area is “signed” because if $f(x) < 0$ on all of $[a, b]$, then $\int_{a}^{b} f(x) \, dx$ will be negative as well, as we can see because each $f(x_i^*) \Delta x$ will be negative but its absolute value will be approximately the area between $f(x)$ and the $i$th interval $[x_{i-1}, x_i]$, and this approximation will improve as $n \to \infty$ and $\Delta x \to 0^+$. And so when the curve is below the $x$-axis (thus having
negative height), the “area” will be represented by a negative number. If part of the curve is above, and another part below, the \(x\)-axis, there will be some area “cancellation.”

It will

In this subsection we will compute signed areas bounded by curves and the \(x\)-axis, and also look into some physics problems involving displacements, by which we mean changes in position.

**Example 6.3.2** Find the area bounded by the parabola \(y = x^2\) and the \(x\)-axis along the interval \(0 \leq x \leq 2\).

**Solution:** While it helps to draw this to visualize the situation, it is not actually necessary. The function \(f(x) = x^2\) is nonnegative, so any Riemann Sum approximation of the area will not contain negative “heights” of the rectangles. Once we are sufficiently convinced that shrinking widths and growing numbers of such rectangles will, in the limit, approach the actual area, we can invoke the FTC to compute the area, as is illustrated below:

\[
\text{Area} = \int_0^2 x^2 \, dx = \frac{1}{3} x^3 \bigg|_0^2 = \left[ \frac{1}{3} (2)^3 \right] - \left[ \frac{1}{3} (0)^3 \right] = \frac{8}{3}.
\]

If instead we want to compute limits of Riemann Sums directly, we would divide the interval \([0, 2]\) into \(n\) subintervals with endpoints \(0 = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = 2\), and let \(n \to \infty\). The width of each subinterval would be \(\Delta x = \frac{2 - 0}{n} = \frac{2}{n}\). Furthermore, we can take any \(x_i \in [x_{i-1}, x_i]\) so we will take \(x_i^* = x_i\) (the right endpoint) for each interval, which we further compute to be \(x_i = 0 + i\Delta x = \frac{2i}{n}\), for \(i = 1, 2, \cdots, n\). Thus

\[
\text{Area} = \lim_{n \to \infty} \sum_{i=1}^{n} f(x_i) \Delta x = \lim_{n \to \infty} \sum_{i=1}^{n} \left[ \left( \frac{2i}{n} \right)^2 \frac{2}{n} \right]
\]

\[
= \lim_{n \to \infty} \sum_{i=1}^{n} \frac{8i^2}{n^3} = \lim_{n \to \infty} \frac{8}{n^3} \sum_{i=1}^{n} i^2
\]

\[
= \lim_{n \to \infty} \left[ \frac{8}{n^3} \cdot \frac{n(n+1)(2n+1)}{6} \right] = \lim_{n \to \infty} \frac{8(2n^3 + 3n^2 + n)}{6n^3} = \frac{16}{6} = \frac{8}{3}.
\]

For the right-endpoint Riemann Sums approximating an area or a net displacement, where we wish to have a partition of the interval \([a, b]\) into \(n\) pieces of equal length, with endpoints labeled \(a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b\), we will always have

\[
\Delta x = \frac{b-a}{n},
\]

\[
(6.51)
\]

\[
x_i = a + i \cdot \Delta x.
\]

\[
(6.52)
\]

We also used (6.36), page 540, namely \(\sum_{i=1}^{n} i^2 = n(n+1)(2n+1)/6\). Note that when we write \(\sum_{i=1}^{n} f(x_i) \Delta x\), in that expression \(n\) is a constant, and so if it appears as a factor (multiplier) inside of the summation then it can be brought out (factored). However, no term involving \(i\) can be factored outside of the summation, because \(i\) is not constant within the summation, but changes values in the range \(i = 1, 2, \cdots, n\).
Example 6.3.3  Find the total signed area bounded by the curve \( y = x^3 \) and the \( x \)-axis for the interval \(-2 \leq x \leq 2\).

Solution: If we follow the precedent from the previous example, we get

\[
\text{Area} = \int_{-2}^{2} x^3 \, dx = \left[ \frac{1}{4} x^4 \right]_{-2}^{2} = \left[ \frac{1}{4} (2)^4 \right] - \left[ \frac{1}{4} (-2)^4 \right] = \frac{16}{4} - \frac{16}{4} = 0.
\]

This seems odd until we note that there should be a cancellation of two “areas” which are identical, except that their signs are opposites. We can calculate the individual areas separately:

\[
\int_{-2}^{0} x^3 \, dx = \left[ \frac{1}{4} x^4 \right]_{-2}^{0} = \frac{0^4}{4} - \frac{(-2)^4}{4} = -4,
\]

\[
\int_{0}^{2} x^3 \, dx = \left[ \frac{1}{4} x^4 \right]_{0}^{2} = \frac{2^4}{4} - \frac{0^4}{4} = 4.
\]

Total Area = \(-4 + 4 = 0\).

If we are to believe that we can extend the general geometric notion that the area of a region should be the same as the sum of non-overlapping subregions whose union is the original (whole) region, we should accept the first equality given below, and therefore the final computation based upon those above:

\[
\int_{-2}^{2} x^3 \, dx = \int_{-2}^{0} x^3 \, dx + \int_{0}^{2} x^3 \, dx = -4 + 4 = 0.
\]

This example above illustrates how areas of different sign can “cancel” each other, and that we can if we wish break up a particular area computation into sub-area computations. When we have an antiderivative formula for the entire interval (such as \([-2, 2]\) in the above example) there is no need. However, sometimes we have antiderivative formulas for individual subintervals (Example 6.3.4 below) and other times there are geometric considerations which make a computation simpler. For instance, in the above example we could have noted the symmetry (with respect to the origin) of the odd function \( f(x) = x^3 \), and the symmetry of the interval, and noted that there was exactly as much “positive area” as there was “negative area,” and therefore the total area would be zero.\(^{18}\)

We used the following intuitive theorem, which we state without proof:

**Theorem 6.3.3** If \( f(x) \) is continuous on \([a, c]\), and \( b \in (a, c) \), then

\[
\int_{a}^{c} f(x) \, dx = \int_{a}^{b} f(x) \, dx + \int_{b}^{c} f(x) \, dx. \quad (6.53)
\]

\(^{18}\)In later sections it is important to not use the argument about “cancelling areas” if there is a chance one of the areas is infinite, as can happen near vertical asymptotes, for instance. We want to be careful not to be tempted to compute \( \infty - \infty \) as being zero, for instance. (See for instance Example 3.8.1, page 255 and the relevant discussions.)
Example 6.3.4 Compute the area under the curve of the function

\[ f(x) = \begin{cases} 
  x^2 & \text{if } x \leq 1, \\
  \sqrt{x} & \text{if } x \geq 1
\end{cases} \]

over the interval \([0, 2]\).

Solution: Here we have a function which is given by one formula for one interval of \(x\)-values, and another formula for another interval, and the area we wish to compute lies along an interval which overlaps both of these. In a case such as this, we break the area into two pieces, where each has a valid simple formula for the bounding function. Here we will use the following:

\[
\int_0^2 f(x) \, dx = \int_0^1 f(x) \, dx + \int_1^2 f(x) \, dx
\]

\[
= \int_0^1 x^2 \, dx + \int_1^2 x^{1/2} \, dx
\]

\[
= \left[ \frac{1}{3} x^3 \right]_0^1 + \left[ \frac{2}{3} x^{3/2} \right]_1^{2}
\]

\[
= \left( \frac{(1)^3}{3} - \frac{0^3}{3} \right) + \left( \frac{2}{3} (2)^{3/2} - \frac{2}{3} (1)^{3/2} \right)
\]

\[
= \frac{1}{3} + \frac{4\sqrt{2} - 2}{3}
\]

\[
= \frac{4\sqrt{2} - 1}{3} \approx 1.5522875.
\]

Note that in the above example, either formula was valid for computing \(f(1)\), in the sense that \(f(1) = 1 = (1)^2 = \sqrt{1}\). Indeed the function \(f(x)\) was continuous (so the FTC applies), as are both functions \(x \mapsto x^2\) and \(x \mapsto \sqrt{x}\) at \(x = 1\). Thus there was no difficulty in using the formula \(f(x) = x^2\) for \([0, 1]\) and \(f(x) = \sqrt{x}\) for \([1, 2]\), even though \(x = 1\) is shared by them.

In fact, for a single point such as \(x = 1\), the “area” under the curve will be zero, so we are allowed some flexibility in using whatever formula for \(f(x)\) matches everywhere in the interval, except perhaps at a finite number of points (themselves determining zero area between the curve and the \(x\)-axis). It is especially useful if we use a formula for \(f(x)\) which represents a continuous function on the interval, so we can employ the FTC and go searching for an antiderivative.

Example 6.3.5 Suppose \(f(x) = \begin{cases} 
  x^2, & \text{if } x \neq 1, \\
  5, & \text{if } x = 1
\end{cases} \). Find \(\int_0^3 x^2 \, dx\).

Solution: Here we have a single point at which the function is discontinuous, namely \(x = 1\). However, we should be able to convince ourselves that the area under that single point is zero, and so it can be ignored:

\[
\text{Area} = \int_0^3 x^2 \, dx = \left[ \frac{x^3}{3} \right]_0^3 = \frac{27}{3} - 0 = 9.
\]
In fact if we go back to our Riemann Sum definition of \( \int_0^3 f(x) \, dx \), we would see that even if we chose \( x_i^* = 1 \) for one of our intervals, the term \( f(x_i^*) \Delta x \) would have its influence shrink to zero in the limit as \( n \to \infty \), i.e., as \( \Delta x \to 0^+ \). We will use that same idea in the next example.

**Example 6.3.6** Suppose \( f(x) = \frac{|x|}{x} \), and we wish to find \( \int_{-1}^1 f(x) \, dx \). The function is undefined at \( x = 0 \), but intuitively the “area under the curve at \( x = 0 \)” is itself zero, because the width of that one point is zero.\(^{19}\) So we can let \( f(0) \) be redefined to be any finite value, and compute the integral as in the previous example, ignoring the possible presence of \( f(0) \Delta x \) in the Riemann sums whose limits we are ultimately computing.

However, we will have different expressions for \( f(x) \) for the cases \( x < 0 \) and \( x > 0 \), at least if we want expression forms for which we can use our antiderivative formulas. So for this example we look at \( \int_{-1}^0 f(x) \, dx \) and \( \int_0^1 f(x) \, dx \) separately. Except at \( x = 0 \) the expressions for \( f(x) \) have well-known antiderivatives, and so we “fill in” \( f(0) \) for each one separately, with the values that would make \( f(x) \) continuous at \( x = 0 \) on the respective intervals:

\[
\int_{-1}^1 f(x) \, dx = \int_{-1}^0 f(x) \, dx + \int_0^1 f(x) \, dx \\
= \int_{-1}^1 (-1) \, dx + \int_0^1 1 \, dx \\
= (-x)^1_{-1} + (x)^1_0 \\
= 0 - [-(-1)] + [1 - 0] \\
= -1 + 1 = 0.
\]

That the areas would “cancel,” and indeed what their values are such that they would cancel, is clear when this function is graphed.

### 6.3.5 Physics Application: Net Displacement versus Distance Traveled

Recall that, over an interval \( t_0 \leq t \leq t_f \), the net displacement of an object with position function \( s(t) \) is given by \( s(t_f) - s(t_0) \). Note that the fundamental theorem of calculus, and the definition of velocity given by \( v(t) = s'(t) \), gives us

\[
s(t_f) - s(t_0) = \int_{t_0}^{t_f} s'(t) \, dt = \int_{t_0}^{t_f} v(t) \, dt.
\]  

(6.54)

If we wished to measure instead the total distance traveled by the object, we have to be sure that a negative-only velocity would still result in a positive distance traveled (even though the displacement would be negative).

\[
\int_{t_0}^{t_f} |v(t)| \, dt
\]

### 6.3.6 Infinitesimals

There is an elegant viewpoint often used in interpreting definite integrals \( \int_a^b f(x) \, dx \), which calls upon a once incompletely understood notion from the early days of calculus, that viewpoint\(^{19}\)This is a subtle point which can easily be over-generalized, i.e., one can draw too many conclusions from this observation that \( \int_0^0 f(x) \, dx = 0 \) regardless of \( f(x) \). In fact the integral makes no sense if \( f(0) \) is undefined, but we expect the area to be zero if \( f(0) \) is any real number, so it seems not unreasonable to disregard the behavior of \( f(x) \) at a single point.
6.3. RIEMANN SUMS AND THE FUNDAMENTAL THEOREM OF CALCULUS

being namely that of the infinitesimals. For such an interpretation to be correct in a particular context, it is best to refer back to the viewpoint of Riemann Sums and their limits.

The idea of considering a quantity to be “infinitesimally small” is in some sense absurd, but worth considering if in a way to rescue that mindset and put it on firm footing. Consider the notation which gives us $s(t_f) - s(t_0) = \int_{t_0}^{t_f} v(t) \, dt$, which when properly understood ($s$ is an antiderivative of $v$, the integral is a limit of Riemann Sums) is actually intuitive, and some would say obvious, albeit after much reflection. (See the discussion leading up to (6.43), page 547.)

Now let us somewhat dissect this notation as it stands. First note that

$$ds(t) = v(t) \, dt$$

from our previous derivative and differential notations. One looking at this in terms of infinitesimals would say that “$ds(t)$ is an infinitesimal change of position at time $t$ caused by an infinitesimal change in $v$ in time, when the velocity was $v(t)$.” Note that there is an assumption that velocity is, for these purposes, constant (or close enough to constant) as time changes by this infinitesimal amount $dt$, and so the change of position would be $v(t) \, dt$.

This idea that the resulting infinitesimal change in $s$, namely $ds(t)$, would be the same as $v(t) \, dt$ in fact does become more accurate as $dt \to 0$, in the sense that if $ds(t)/dt$ exists, then it must be $v(t)$, and moreover, the actual change in $s$ will be approximated better and better—in terms of percent error—by $v(t) \, dt$ when $dt$ shrinks. Indeed, when $ds/dt$ exists there is a shrinkage to zero in the percent error in using $ds(t)$ to approximate the actual change (namely $\Delta s$) in $s(t)$ resulting from the change in $t$ to $dt$ (also known as $\Delta t$), and so writing $\Delta s(t) \approx ds(t) = v(t) \, dt$ becomes closer to 100% accurate as $dt \to 0$.20 (Of course if we added all the $\Delta s$ terms for as $t$ ranges from $t_0$ to $t_f$, they would sum to $s(t_f) - s(t_0)$.)

This thinking allows one to (naively) look at $\int_{t_0}^{t_f} v(t) \, dt$ as an infinite sum of infinitesimal quantities $ds(t)$, one such infinitesimal for each $t \in [t_0, t_f]$, and these somehow accumulating to represent the actual quantity $s(t_f) - s(t_0)$:

$$\int_{t_0}^{t_f} v(t) \, dt = s(t_f) - s(t_0).$$

Again, this makes sense if we also keep in mind that this integral represents a limit of Riemann Sums of the form $\sum_{i=1}^{n} s(t^*_i) \Delta t_i$, as $\max \{ \Delta t_i \} \to 0^+$ and $n \to \infty$.

When looking at $\int_{a}^{b} f(x) \, dx$, one considers “infinitesimal rectangles of infinitesimal widths $dx$, these rectangles having signed areas $f(x) \, dx$, at each value of $x \in [a, b]$.”

As we will see eventually, this kind of analysis is quite powerful for discovery purposes in a multitude of circumstances beyond displacement and area problems, though to be sure of its validity for other cases a Riemann Sum analysis should be included, where one sees if a percentage error argument is convincing.

A simple example is using infinitesimals to find the area of a circle of radius $R$. One could consider breaking such a circle up into concentric circles of radii $r \in [0, R]$, each such circle having circumference $2\pi r$, but given also an infinitesimal “thickness” of $dr$ in the perimeter. The area of the actual curve of such a circle (not its interior) would arguably be approximately $dA = 2\pi r \, dr$, that is the perimeter (circumference) multiplied by the thickness of that perimeter. This will not be exact, because if we “unrolled” a circle’s perimeter which was given some thickness, we would not have a rectangle, but it would be likely a trapezoid which would be very nearly rectangular.

---

20This is arguably false if $\Delta s = 0$, but we have argued before that that technicality can be resolved because of the $\Delta t \to 0^+$ in the limit of the Riemann Sums, so while percent error may be undefined, absolute error from those seemingly problematic terms will shrink to zero, since those terms are of the form $v(t^*_i) \Delta t_i.$
The area would be very near to that of a rectangle with length $2\pi r$ and height $dr$ (the thickness). “Adding” all these up, we would get

\[
\text{Area of Circle} = \int_{r=0}^{r=R} dA(r) = \int_{0}^{R} 2\pi r \, dr = \pi r^2 \bigg|_{r=0}^{r=R} = \pi R^2 - \pi (0)^2 = \pi R^2,
\]

as we should expect. Countless other examples can be found, where we don’t need the exact formula for a “piece” of the accumulated quantity we need, but if we have an approximation which has percentage error that shrinks to zero when we break our quantity (such as displacement or area) into pieces whose number approaches infinity but whose individual contributions shrink to zero, then our integral formula for that desired cumulative quantity is correct. This is more obvious when the definite integral in question is viewed as a limit of Riemann sums, but the use of infinitesimals has its appeal.
6.4 Substitution With Power Rule

Substitution in general is the most important of the integrating techniques, finding its way into the other techniques as well. While we introduce it here, for now we limit the scope to power rules.

Before looking at this method formally, consider the following antiderivative statements, each of which refer to the same power rule (perhaps most familiar in the first case):

\[
\int x^2 \, dx = \frac{x^3}{3} + C, \\
\int u^2 \, du = \frac{u^3}{3} + C, \\
\int (\sin x)^2 \, d(\sin x) = \frac{(\sin x)^3}{3} + C.
\]

The last integral is simply asking for an antiderivative of \((\sin x)^2\) with respect to \(\sin x\). Indeed, we can check the answer as before:

\[
\frac{d}{d \sin x} \left[ \frac{1}{3} (\sin x)^3 \right] = \frac{1}{3} \cdot 3(\sin x)^2 = (\sin x)^2,
\]

as we expect. Of course we usually take derivatives and antiderivatives with respect to a variable, and not a function. However the integral in (6.55) is not so unlikely to occur as one might think. Recall that \(df(x) = f'(x) \, dx\) is the definition of the differential (see (5.14), page 501). Thus \(d \sin x = \cos x \, dx\), and the integral in (6.55) can be written instead

\[
\int (\sin x)^2 \cos x \, dx = \int (\sin x)^2 \frac{d \sin x}{dx} \, dx = \int (\sin x)^2 d \sin x = \frac{1}{3} (\sin x)^3 + C.
\]

Indeed, it is not hard to see that the chain rule gives us \(\frac{d}{dx} \left[ \frac{1}{3} (\sin x)^3 \right] = \frac{1}{3} \cdot 3(\sin x)^2 \cos x = \sin^2 x \cos x\).

In this section we will concentrate on integrals of the form

\[
\int u^n \, du = \begin{cases} 
\frac{1}{n+1} \cdot u^{n+1} + C & \text{if } n \neq -1, \\
\ln |u| + C & \text{if } n = -1.
\end{cases}
\]

As anticipated in the discussion above, the content of the differential \(du\) may be more expansive than what we may expect from a single variable. The point of this section is to recognize when we have the form (6.56), and how to go about rewriting the integral into the proper form.

The reader should be forewarned: this method requires a fair amount of practice. It is not a simple algorithm. For each problem, the reader has to decide which substitution will produce an integral which can be computed with known rules. Here we will limit ourselves to the power rule (6.56), but in subsequent sections we will delve into many other rules, and it is not always obvious which rule should be used for a given integral. With practice one learns to look for clues, and anticipate what will occur several steps ahead, to see if there is indeed an integration rule which can apply.\(^{21}\)

\(^{21}\)In fact, often there is no rule which will produce an antiderivative, and then some approximation scheme will be necessary. Still, it is most desirable to have an exact antiderivative, and we can find one often enough that it is well worth studying these techniques.
6.4.1 The Technique

Here we will look at some of the simpler problems of integration by substitution. As we proceed, several observations will be made regarding the method.

Example 6.4.1 Compute the indefinite integral \( \int (x^2 + 1)^7 \cdot 2x \, dx \).

Solution: The technique is to introduce a new variable, \( u \), with which we can write the original integral in a simpler form. We also have to take into account what will be the new differential, namely \( du \):

\[
\begin{align*}
  u &= x^2 + 1 \\
  du &= 2x \, dx.
\end{align*}
\]

(Recall that if \( u \) is a function of \( x \), then \( du = u'(x) \, dx \), consistent with \( \frac{du}{dx} = u'(x) \). Using this information, we can replace all the terms in the original integral: the \( (x^2 + 1)^7 \) becomes \( u^7 \), and the terms \( 2x \, dx \) collectively become \( du \) (see the above implication arrow). Thus

\[
\int (x^2 + 1)^7 \cdot 2x \, dx = \int u^7 \, du = \frac{1}{8} u^8 + C.
\]

This is all true, but we introduced \( u \), while the original question asked for an antiderivative with respect to \( x \). We only need to replace \( u \) in the final answer, using again \( u = x^2 + 1 \). Summarizing,

\[
\int (x^2 + 1)^7 \cdot 2x \, dx = \int u^7 \, du = \frac{1}{8} u^8 + C = \frac{1}{8} (x^2 + 1)^8 + C.
\]

Note that we can check our answer in the above example by computing the derivative of the answer (using the chain rule), yielding \( \frac{d}{dx} \left[ \frac{1}{8} (x^2 + 1)^8 \right] = \frac{1}{8} \cdot 8(x^2 + 1)^7 \cdot 2x = (x^2 + 1)^7 \cdot 2x \) as hoped. In fact, integration by substitution, at least in its simplest forms, is often called a type of reverse chain rule. Indeed, we can rewrite (6.56) as follows:

\[
\int u^n \, du = \int u^n \cdot \left( \frac{du}{dx} \right) \, dx = \int \left[ u^n \cdot \frac{du}{dx} \right] \, dx = \begin{cases} 
  \frac{1}{n+1} \cdot u^{n+1} + C & \text{if } n \neq 1, \\
  \ln|u| + C & \text{if } n = -1. 
\end{cases} \tag{6.57}
\]

To see that this is correct, if we take the derivative of the answers with respect to \( x \), we see that we do indeed get \( u^n \cdot \frac{du}{dx} \) from the chain rule.

In short, with integration by substitution we try to pick some function we call \( u \), so that

1. a main part of the integrand can be written as a simple function of \( u \)—one for which we know the antiderivative with respect to \( u \)—and, equally crucial, so that

2. the remaining variable terms of the integral can be safely absorbed in \( du \) (except for multiplicative constants, which we will see add only a slight complication).

If these are both satisfied, our substitution of \( u \) and \( du \) terms gives us a new, simple integral (entirely in terms of \( u \) and \( du \)).

When working such a problem (as opposed to, say, publishing a problem and solution for professional consumption), a useful format is to (1) write the original integral, (2) write the substitution function \( u \), with its differential \( du \) both on different lines than the original integral, (3) write the new form of the original integral, i.e., in \( u \) and \( du \), (4) compute the antiderivative
of this new integral in $u$ as a continuation of the first step, and (5) resubstitute to arrive at the antiderivative in $x$. Hence a typical homework-style presentation of Example 6.4.1 might look like the following (with the choice of $u$ and resulting $du$ offset and below the original integral):

$$\int (x^2 + 1)^7 \cdot 2x \, dx = \int u^7 \, du = \frac{1}{8} u^8 + C = \frac{1}{8} (x^2 + 1)^8 + C$$

$u = x^2 + 1$

$\Rightarrow \, du = 2x \, dx$

That part of the integral which we hope to absorb into $du$ is underlined in the original integral, the computation of $du$, and the corresponding term in the new integral. The rest of the integral was just $(x^2 + 1)^7 = u^7$. In fact, when we choose $u$ so that a major portion of the integral can be written $u^n$, then any other factors which are variable, along with the differential $dx$, must be absorbed in $du$ or the substitution will fail (because the resulting integral will contain both $x$ and $u$ and no antiderivative rules will apply). We will continue to use this kind of spatial organization when we integrate by substitution in the examples below.

**Example 6.4.2** Compute the indefinite integral $\int \sin^2 x \, \cos x \, dx$.

**Solution:** Note that this integral can be written $\int (\sin x)^2 \, \cos x \, dx$. Now we proceed:

$$\int \sin^2 x \, \cos x \, dx = \int u^2 \, du = \frac{1}{3} u^3 + C = \frac{1}{3} \sin^3 x + C.$$  

$u = \sin x$

$\Rightarrow \, du = \cos x \, dx$

Before continuing we will make a very minor change to the integral in the first example (Example 6.4.1, page 558), and show a simple way to extend our method to handle this.

**Example 6.4.3** Compute $\int x(x^2 + 1)^7 \, dx$.

**Solution:** Here we will make the same substitution as before, but the $du$ will have an extra factor of 2. Since constant factors are relatively easy to handle in derivative and antiderivative problems in general, we should not expect this extra factor of 2 to cause much difficulty. It will simply mean one extra step in the substitution computations.\(^{22}\)

\[^{22}\text{There is another method used by some texts to handle a problem such as this, which is to simply introduce the needed factor of 2 in the integral to complete the differential } \int 2x \, dx, \text{ and compensate for the insertion of the new factor by simultaneously inserting a factor of } \frac{1}{2}, \text{ which is simply carried through the rest of the calculation:}\]

$$\int x(x^2 + 1)^7 \, dx = \int \frac{1}{2} (x^2 + 1)^7 \cdot 2x \, dx = \frac{1}{2} \int u^7 \, du = \frac{1}{2} \cdot \frac{1}{8} u^8 + C = \frac{1}{16} (x^2 + 1)^8 + C,$$

where again $u = x^2 + 1$, $du = 2x \, dx$.

This method is appealing because one rewrites the integrand into a form where it is, more or less, clearly a derivative of a chain rule function (perhaps multiplied by a constant, as with $\frac{1}{2}$ here).

We will avoid this method because, though it is not so challenging for simpler problems, it quickly becomes unreasonably difficult if an integral is complicated. Furthermore, the method presented in this text—in the author’s opinion—makes for much better preparation for more advanced methods, such as trigonometric substitution and integration by parts.
\[
\int x(x^2 + 1)^7 \, dx = \int u^7 \cdot \frac{1}{2} \, du = \frac{1}{2} \cdot \frac{1}{8} u^8 + C = \frac{1}{16} (x^2 + 1)^8 + C
\]

\[
\begin{align*}
& \quad u = x^2 + 1 \\
\Rightarrow & \quad du = 2x \, dx \\
\Rightarrow & \quad \frac{1}{2} \, du = x \, dx
\end{align*}
\]

This time the extra nonconstant and differential terms of the original integral were, collectively, \(x \, dx\). Though that product is not exactly \(du\), it is a constant times \(du\). In our substitution we took an extra step and solved, again collectively, for \(x \, dx = \frac{1}{2} \, du\).

The preceding example shows that we need to be flexible when looking for a possible power rule application. Not every integral where we can use the power rule will be of the strict form (6.57), page 558. Indeed, we need to be especially vigilant to notice that an integral may be of the form

\[
\int k \cdot u^n \, du = \int k \cdot u^n \cdot \left(\frac{du}{dx}\right) \, dx.
\]

(6.58)

So when we make a substitution, we try not to be distracted by extra or missing multiplicative constants, as they will work themselves out in the substitution and final integration steps.

**Example 6.4.4** Compute \(\int x^3 \cos^5 x^4 \sin x^4 \, dx\).

**Solution:** It is perhaps more obvious how to proceed if we rewrite the integral in the form \(\int x^3 (\cos x^4)^5 \sin x^4 \, dx\). Then we see that the \(u^n\) term will be \((\cos x^4)^5 = u^5\), where \(u = \cos x^4\).

Next we need to see if \(du\) can absorb the other nonconstant terms:

\[
\begin{align*}
& \quad \int x^3 \cos^5 x^4 \sin x^4 \, dx = \int u^5 \left(\frac{-1}{4}\right) \, du = \frac{-1}{4} \cdot \frac{1}{6} u^6 + C = -\frac{1}{24} (\cos x^4)^6 + C.
\end{align*}
\]

\[
\begin{align*}
& \quad \quad u = \cos x^4 \\
& \quad \quad \quad \quad du = -\sin x^4 \cdot 4 x^3 \, dx \\
& \quad \quad \quad \quad \frac{1}{4} \, du = x^3 \sin x^4 \, dx
\end{align*}
\]

We should point out that the method above would not have worked without both the \(x^3\) and the \(\sin x^4\) terms in the integral, for the \(du\) would have variable terms not in the original integral.

Also, it is possible to compute the integral in the previous by using two substitution steps instead of one. For instance, a student recognizing that \(x^4\), and a multiple of its derivative in the form of \(x^3\), both appear, might first make a substitution of the form \(u = x^4\):

\[
\begin{align*}
& \quad \int x^3 \cos^5 x^4 \sin x^4 \, dx = \int \cos^5 u \, \sin u \, \frac{1}{4} \, du = \frac{1}{4} \int \cos^5 u \sin u \, du.
\end{align*}
\]

\[
\begin{align*}
& \quad \quad u = x^4 \\
& \quad \quad \quad \quad du = 4 x^3 \, dx \\
& \quad \quad \quad \quad \frac{1}{4} \, du = x^3 \, dx
\end{align*}
\]
At this point, we have a simpler integral which itself requires a substitution:

\[
\frac{1}{4} \int \cos^5 u \sin u \, du = \frac{1}{4} \int w^5 (-dw) = -\frac{1}{4} \cdot \frac{1}{6} w^6 + C = -\frac{1}{24} (\cos u)^6 + C.
\]

\[
w = \cos u \\
\Rightarrow \quad dw = -\sin u \, du \\
\Rightarrow \quad -dw = \sin u \, du
\]

Of course this gives the answer in terms of \(u\), so we substitute back again, in terms of \(x\).

Summarizing,

\[
\int x^3 \cos^5 x^4 \sin x^4 \, dx = \cdots = -\frac{1}{24} u^6 + C = -\frac{1}{24} \cos^6 u + C = -\frac{1}{24} \cos^6 x^4 + C.
\]

The second approach is longer, but it has the advantage that we are not trying to rewrite the integral in one, all-encompassing (and thus more complicated) substitution step. Indeed it is sometimes desirable to simplify an integral with substitution even if the resulting integral cannot be evaluated immediately. With most examples we will use the first method, but the student working problems should be aware that the option of successive substitutions is perfectly valid.

Next we look at a few very common types of examples where the power of \(n\) is \(1/2\), \(-1\) and \(-2\). These appear often enough that it is worth some effort to remember them specifically.

**Example 6.4.5** Compute \(\int \frac{x}{\sqrt{x^2 - 9}} \, dx\).

**Solution** Here we will take \(u = x^2 - 9\), since the \(du = 2x \, dx\) can absorb both the \(dx\) and the extra factor of \(x\):

\[
\int \frac{x}{\sqrt{x^2 - 9}} \, dx = \int u^{-1/2} \cdot \frac{1}{2} \, du = \frac{1}{2} \cdot 2u^{1/2} + C = \sqrt{x^2 - 9} + C.
\]

\[
u = x^2 - 9 \\
\Rightarrow \quad du = 2x \, dx \\
\Rightarrow \quad \frac{1}{2} \, du = x \, dx
\]

**Example 6.4.6** Compute \(\int \frac{\sin x}{\cos x} \, dx\).

**Solution** Here we will take \(u = \cos x\), since \(du = -\sin x \, dx\) will absorb the other terms.

\[
\int \frac{\sin x \, dx}{\cos x} = \int \frac{1}{u} (-du) = -\ln |u| + C = -\ln |\cos x| + C.
\]

\[
u = \cos x \\
\Rightarrow \quad du = -\sin x \, dx \\
\Rightarrow \quad -du = \sin x \, dx
\]

Note that if we instead took \(u = \sin x\), then \(du = \cos x \, dx\), but \(\cos x\) is not a *multiplicative* factor in the original integral; the desired factor is \(\frac{1}{\cos x}\), which is not contained in the \(du\) term if \(u = \sin x\).
CHAPTER 6. BASIC INTEGRATION

It should be remembered that checking these antiderivatives is as simple as computing the derivative of the answer. Here
\[
\frac{d}{dx} [-\ln |\cos x|] = -\frac{1}{\cos x} \cdot \frac{d}{dx} \cos x = -\frac{1}{\cos x} (-\sin x) = \frac{\sin x}{\cos x},
\]
as we hope. Of course our original integrand, and the derivative above, can both be written \(\tan x\).

Note that we can write
\[
-\ln |\cos x| = \ln |\sec x|,
\]
so many calculus books contain the integration formula
\[
\int \tan x \, dx = \ln |\sec x| + C. \tag{6.59}
\]
(It is also interesting to “fill in the dots” for the computation \(\frac{d}{dx} \ln |\sec x| = \cdots = \tan x\), verifying (6.59). See also Exercises 15 and 16, page 567.)

**Example 6.4.7** Compute \(\int \frac{e^{3x}}{(e^{3x} + 4)^2} \, dx\).

**Solution:** Here we note that the numerator of the integrand, namely \(e^{3x}\), is the derivative of \(e^{3x} + 4\), except for a multiplicative constant. Thus we will let \(u = e^{3x} + 4\):
\[
\int \frac{e^{3x}}{(e^{3x} + 4)^2} \, dx = \int \frac{1}{u^2} \cdot \frac{1}{3} \, du = \frac{1}{3} \int u^{-2} \, du = \frac{1}{3} \left(-1\right)u^{-1} + C
\]
\[
\Rightarrow \quad du = e^{3x} \cdot 3 \, dx
\]
\[
\Rightarrow \quad \frac{1}{3} \, du = e^{3x} \, dx.
\]
At this point we notice three common forms of integration by substitution:
\[
\int \frac{u'(x)}{\sqrt{u(x)}} \, dx = 2\sqrt{u(x)} + C, \tag{6.60}
\]
\[
\int \frac{u'(x)}{u(x)} \, dx = \ln |u(x)| + C, \tag{6.61}
\]
\[
\int \frac{u'(x)}{|u(x)|^2} \, dx = -\frac{1}{u(x)} + C. \tag{6.62}
\]
In all three cases, \(u'(x) \, dx = du\), and we have simple power rules. In the first and third cases there are multiplicative constants which occur. There is no real need to memorize these, but they occur often enough that their “mechanics” should become familiar. For that reason these three results can become, if not memorized, then at least easily cited.

The method also works for cases where the \(du\) term is just a constant multiple of \(dx\):

**Example 6.4.8** Compute \(\int \frac{1}{(6 - 2x)^3} \, dx\).

**Solution:**
\[
\int \frac{1}{(6 - 2x)^3} \, dx = \int u^{-5} \cdot \frac{-1}{2} \, du = -\frac{1}{2} \cdot \frac{1}{4} u^{-4} + C = \frac{1}{8}(6 - 2x)^{-4} + C
\]
\[
\Rightarrow \quad du = -2 \, dx
\]
\[
\Rightarrow \quad -\frac{1}{2} \, du = dx
\]
\[
= \frac{1}{8(6 - 2x)^4} + C.
\]
In the case that \( du = dx \), this can often be anticipated and the experienced calculus student might omit the middle steps:

Example 6.4.9 Compute \( \int (x + 9)^4 \, dx \).

Solution:

\[
\int (x + 9)^4 = \int u^4 \, du = \frac{1}{5}u^5 + C = \frac{1}{5}(x + 9)^5 + C.
\]

\[
\begin{align*}
\begin{cases}
u = x + 9 \\
\implies du = dx
\end{cases}
\end{align*}
\]

Another way to look at the example above is to realize that \( d(x + 9) = dx \), so we can write

\[
\int (x + 9)^4 \, dx = \int (x + 9)^4 \cdot d(x + 9) = \frac{1}{5}(x + 9)^5 + C.
\]

In other words, \( dx \) is the same as \( d(x + 9) \), so we get the same if we interpret the original integral as an antiderivative with respect to \( (x + 9) \).

Indeed this is a shortcut one learns with practice—thinking but perhaps not writing the second step—but at first it is still best to write out the full substitution, as in the example above, at least until one is proficient in the method as presented here. Of course this is the analog to a chain rule where the “inner” derivative is 1:

\[
\frac{d}{dx} \left[ \frac{1}{5}(x + 9)^5 \right] = \frac{1}{5} \cdot 5(x + 9)^4 \cdot \frac{d(x + 9)}{dx} = \frac{1}{5} \cdot 5(x + 9)^4 \cdot 1 = (x + 9)^4,
\]

q.e.d.

6.4.2 A Slight Twist on the Method

Recall our second example, namely Example 6.4.3 on page 560: \( \int x(x^2 + 1)^7 \, dx \). We used a substitution \( u = x^2 + 1 \) because \( du = 2x \, dx \) contained the extra factor of \( x \) in the integrand. The substitution eventually gave us \( \int u^7 \cdot \frac{1}{2} \, du \), which was a simple power rule. Of course we could have “simply” expanded the original function

\[
x(x^2 + 1)^7 = x(x^2 + 7x^4 + 21x^6 + 35x^8 + 35x^{10} + 21x^{12} + 7x^{14} + 1)
\]

\[
= x^3 + 7x^5 + 21x^7 + 35x^9 + 35x^{11} + 21x^{13} + 7x^{15} + x,
\]

and integrated “term by term.” However the substitution method was arguably easier, and the answer’s simple form, \( \frac{1}{10}(x^2 + 1)^8 + C \) would probably not be recognizable from a strategy which expands the integrand first.

Now consider the integral \( \int x(x - 1)^{3/2} \, dx \). Here we can not simply “expand” the integrand (even by brute force, as above), because of the fractional power term \( (x - 1)^{3/2} \), which is algebraically more difficult to deal with than positive integer powers. Furthermore, if we let \( u = x - 1 \), then \( du = dx \), but this differential term cannot absorb the extra factor \( x \). The key is to then notice that the original substitution offers a way out: that extra factor \( x \) can be rewritten \( u + 1 \) (since \( u = x - 1 \iff u + 1 = x \)). Below we show how this can be utilized. Indeed we will expand the new integrand, but what is interesting is how the algebraic difficulties of the \( (x - 1)^{3/2} \) term (namely that this is of the form \( (a + b)^r, \ r \notin \mathbb{N} \)) is transferred to the \( x \) term which, being a positive integer power, is then easier to handle. Below we write this out in the standard example format:

\[\text{Note that the change in } x + 9 \text{ is the same as the change in } x.\]
Example 6.4.10 Compute \( \int x(x-1)^{3/2} \, dx \).

Solution: Here we substitute for \( dx \) and \( x \). Both substitutions are calculated below, but separately. (This time we underline the substitution for \( x \), instead of the differential part.) Once the substitutions are completed, we can perform the multiplication to get two simple power rules:

\[
\int x(x-1)^{3/2} \, dx = \int (u+1)u^{3/2} \, du = \int \left( u^{5/2} + u^{3/2} \right) \, du = \frac{2}{7}u^{7/2} + \frac{2}{5}u^{5/2} + C
\]

\[
\left\{
\begin{align*}
\quad u &= x-1 \\
\quad du &= dx
\end{align*}\right\}
\quad \Rightarrow \quad u+1 = x
\]

\[
Also, \quad \left\{
\begin{align*}
\quad u &= x-1 \\
\quad \frac{1}{2} \, du &= \frac{1}{2} \, dx
\end{align*}\right\}
\quad \Rightarrow \quad x = \frac{1}{2}(u-1)
\]

Though the answer above is correct, one often factors the final answer:

\[
= \frac{2}{35}(x-1)^{5/2}[5(x-1)-7] + C = \frac{2}{35}(x-1)^{5/2}(5x-12) + C.
\]

Example 6.4.11 Compute \( \int \frac{x}{\sqrt{2x+1}} \, dx \).

Solution: We will work this problem twice using two different substitutions. The first is perhaps the most obvious, but the second has some appeal as well.

\[
\int \frac{x}{\sqrt{2x+1}} \, dx = \int \frac{\frac{1}{2}(u-1)}{u^{1/2}} \cdot \frac{1}{2} \, du = \frac{1}{4} \int \left( u^{1/2} - u^{-1/2} \right) \, du
\]

\[
\left\{
\begin{align*}
\quad u &= 2x+1 \\
\quad du &= 2 \, dx
\end{align*}\right\}
\quad \Rightarrow \quad \frac{1}{2} \, du = dx
\]

\[
Also, \quad \left\{
\begin{align*}
\quad u &= 2x+1 \\
\quad x &= \frac{1}{2}(u-1)
\end{align*}\right\}
\]

Again one might factor, simplify and rearrange the variable parts of the answer to arrive at

\[
= \frac{1}{6}(2x+1)^{1/2}[(2x+1) - 3] + C = \frac{1}{6}\sqrt{2x+1}(2x-2) + C = \frac{1}{3}(x-1)\sqrt{2x+1} + C.
\]

For the alternative substitution we let \( u = \sqrt{2x+1} \). Note how much of the integrand is then
absorbed into $du$ (due to the relationship between the square root and its derivative).

\[
\int \frac{x}{\sqrt{2x + 1}} \, dx = \int \frac{1}{2} \left( u^2 - 1 \right) \, du = \frac{1}{2} \left[ \frac{1}{3} u^3 - u \right] + C = \frac{1}{6} u^3 - \frac{1}{2} u + C
\]

\[
\begin{align*}
\text{Let } u &= \sqrt{2x + 1} \\
\Rightarrow \quad du &= \frac{1}{2 \sqrt{2x + 1}} \cdot 2 \, dx \\
\Rightarrow \quad du &= \frac{1}{\sqrt{2x + 1}} \, dx \\
\text{with } u &= \sqrt{2x + 1} \\
\Rightarrow \quad u^2 &= 2x + 1 \\
\Rightarrow \quad \frac{1}{2} (u^2 - 1) &= x
\end{align*}
\]

It is important to notice that we used the same equation for $u$ to calculate $du$, within a given strategy (though $u$, $du$ were different for the different strategies). Also the reader should begin to see that we can make some rather interesting substitutions, so long as we are consistent when replacing every term inside the integral. In doing so, it will become apparent if (1) it is even possible to use a given substitution to rewrite the integral, and (2) even if so, is the new integral one which we can actually compute.

### 6.4.3 Other Miscellaneous Power Rule Substitutions

So far we have concentrated on algebraic (polynomial and rational-power), exponential and trigonometric functions in our substitution problems. It is also worth examining how power rules can arise from integrals involving logarithmic and arctrigonometric functions, which we do in this subsection.

**Example 6.4.12** Compute $\int \frac{(\ln x)^5}{x} \, dx$.

**Solution:** He we see a factor $\frac{1}{x}$, which is the derivative of $\ln x$, so the latter will be $u$:

\[
\begin{align*}
\int \frac{(\ln x)^5}{x} \, dx &= \int u^5 \, du = \frac{1}{6} u^6 + C = \frac{(\ln x)^6}{6} + C. \\
\text{Let } u &= \ln x \\
\Rightarrow \quad du &= \frac{1}{x} \, dx
\end{align*}
\]

Note that, as a general rule, if we have a function $f(x)$ with antiderivative $F(x)$, then we have\footnote{In fact we can replace $\ln x$ with $\ln |x|$ in throughout the above.}

\[
\int \frac{f(\ln x)}{x} \, dx = \int f(u) \, du = F(u) + C = F(\ln x) + C. \quad (6.63)
\]

\[
\begin{align*}
\text{Let } u &= \ln x \\
\Rightarrow \quad du &= \frac{1}{x} \, dx
\end{align*}
\]

Similar formulas apply to the arctrigonometric functions. Rather than list and commit to memorize them, it is better to look at the general idea that if, say, $\sin^{-1} x$ occurs in an integral, we would look immediately to see if its derivative, $\frac{1}{\sqrt{1-x^2}}$, also appears. Similarly for all functions.
Example 6.4.13 Compute \( \int \frac{1}{\sqrt{1 - x^2 \sin^{-1} x}} \, dx \).

Solution: Note here that if \( u = \sin^{-1} x \) then our \( du \) below will account for \( \frac{1}{\sqrt{1 - x^2}} \, dx \):

\[
\int \frac{1}{\sqrt{1 - x^2 \sin^{-1} x}} \, dx = \int \frac{1}{u} \, du = \ln |u| + C = \ln |\sin^{-1} x| + C.
\]

\[
\Rightarrow du = \frac{1}{\sqrt{1 - x^2}} \, dx
\]

Example 6.4.14 Compute \( \int \frac{\sec^{-1} x}{x \sqrt{x^2 - 1}} \, dx \). Assume \( x > 0 \) (or more precisely, \( x \geq 1 \)).

Solution:

\[
\int \frac{\sec^{-1} x}{x \sqrt{x^2 - 1}} \, dx = \int u \, du = \frac{1}{2} u^2 + C = \frac{\left(\sec^{-1} x\right)^2}{2} + C.
\]

\[
\Rightarrow du = \frac{1}{x \sqrt{x^2 - 1}} \, dx
\]

In the example above, if instead \( x < 0 \) (actually \( x < -1 \)), we would replace \( x \) by \(-|x|\) in the denominator of the integrand, giving eventually \(-\frac{1}{2}(\sec^{-1} x) + C\) for the antiderivative.
6.4. SUBSTITUTION WITH POWER RULE

Exercises

Compute the given indefinite integrals.

1. \( \int \cos^3 x \sin x \, dx \)

2. \( \int \cos^3 2x \sin 2x \, dx \)

3. \( \int x^3(x^4 + 10) \, dx \)

4. \( \int [(2x + 1)(x^2 + x + 9)] \, dx \)

5. \( \int \frac{\sec^2 x \, dx}{\sqrt{\tan x}} \)

6. \( \int \frac{\ln x}{x} \, dx \)

7. \( \int \frac{1}{x \ln x} \, dx \)

8. \( \int \frac{x}{x^2 + 4} \, dx \)

9. \( \int \frac{(\tan^{-1} x)^2}{1 + x^2} \, dx \)

10. \( \int \cot x \, dx \)

11. \( \int \sec^2 x \tan x \, dx \)

12. \( \int \frac{1}{\sqrt{x}} \cdot (5 + \sqrt{x})^{10} \, dx \)

13. \( \int \frac{\cos^{-1} x}{\sqrt{1 - x^2}} \, dx \)

14. Perform the following computation without rearrangement for both cases.
   (a) \( \int \sec x \cdot \sec x \tan x \, dx \)

   (b) \( \int \tan x \sec^2 x \, dx \)

   (c) Explain why, though the answers “look” different, in fact are the same.

15. Without rearrangement, compute \( \int \frac{\sec x \tan x}{\sec x} \, dx \).

16. Perform the previous computation after first simplifying the fraction (perhaps into sines and cosines).

17. Using the methods of Subsection 6.4.2, compute the following integral:

   \( \int x\sqrt{1-x} \, dx \).

18. Compute \( \int x(x + 5)^5 \, dx \)

19. Compute \( \int \frac{x}{\sqrt{2x + 1}} \, dx \)
6.5 Second Trigonometric Rules

We first looked at the simplest trigonometric integration rules—those arising from the derivatives of the trigonometric functions—in Section 6.1 (Subsection 6.1.4, page 534 to be more precise). Here we will complete the trigonometric rules in which one of the six basic trigonometric functions is the “outer” function. In fact we have four of the six antiderivatives we need: sine and cosine come quickly from the derivative formulas, and tangent and cotangent come from substitution arguments. As it turns out, secant and cosecant require a little more cleverness, and while we will not derive these from first principles, we will show that checking them is a quick and interesting derivative computation. Unfortunately (or fortunately, whatever your perspective) there are variations of the antiderivatives of tangent, cotangent, secant and cosecant. We will choose one form for each, but the well-informed student must be aware of the others to be prepared to discuss calculus topics among students with different backgrounds.25

6.5.1 Antiderivatives of the Six Trigonometric Functions

The antiderivatives of the six basic trigonometric functions are as follow:

\[
\int \sin x \, dx = -\cos x + C, \tag{6.64}
\]
\[
\int \cos x \, dx = \sin x + C, \tag{6.65}
\]
\[
\int \tan x \, dx = \ln |\sec x| + C, \tag{6.66}
\]
\[
\int \cot x \, dx = -\ln |\csc x| + C, \tag{6.67}
\]
\[
\int \sec x \, dx = \ln |\sec x + \tan x| + C, \tag{6.68}
\]
\[
\int \csc x \, dx = -\ln |\csc x + \cot x| + C. \tag{6.69}
\]

The first four of these can be verified mentally through quick derivative computations if the student is well enough versed in differentiation. The last two require some more care, but are

\[
\int \tan x \, dx = -\ln |\cos x| + C,
\]
\[
\int \cot x \, dx = \ln |\sin x| + C,
\]

which after all have slightly simpler verifications by differentiation than (6.66) and (6.67). Here we have opted to use the latter, slightly more difficult formulas for a few reasons. First, they are themselves quite popular. Second, the reader used to (6.66) and (6.67) will be less likely to be confused when presented the simpler alternatives by a colleague (or future professor) with a different background, while the reader used to those simpler alternatives may have some initial difficulty if similarly presented our forms here. Finally, there is so much added structure, both calculus and algebraic, found in the context of the secant and cosecant functions so it is important to be familiar and comfortable with them.

Admittedly, however, if (6.66) and (6.67) were not so common we would likely opt for the simpler forms.
somewhat interesting to check. For instance, we can verify (6.68) as follows:

\[
\frac{d \ln |\sec x + \tan x|}{dx} = \frac{1}{\sec x + \tan x} \cdot \frac{d (\sec x + \tan x)}{dx} = \frac{1}{\sec x + \tan x} \cdot (\sec x \tan x + \sec^2 x) = \frac{\sec x(\sec x + \tan x)}{\sec x + \tan x} = \sec x,
\]

\[\text{q.e.d.}\]

It is not entirely obvious how one would derive antiderivatives of the secant and cosecant functions, and so it is important to memorize those especially. Indeed it is likely these were discovered through experimentation, and such results are often very time consuming to reproduce from first principles if one has to re-invent “the trick,” one of which will be explored in the exercises. In fact we will later show a popular alternative antiderivative for the cosecant, and a not-so-popular alternative for the secant. The alternatives for the tangent and cotangent are similar in popularity to those we will use for our standards.

There is little we can do with just (6.64)–(6.69) as they stand, but we nonetheless explore a few examples quickly.

**Example 6.5.1** Below are two quick antiderivative computations involving our basic trigonometric integral formulas.

- \[\int \frac{\sin^2 x + \cos x}{\sin x} \, dx = \int (\sin x + \cot x) \, dx = -\cos x - \ln |\csc x| + C.\]
- \[\int (x + \sec x) \, dx = \frac{x^2}{2} + \ln |\sec x + \tan x| + C.\]

**Example 6.5.2** Suppose \(v(t) = 1 + \tan t,\) and \(s(\pi/3) = 7\). Find \(s(t),\) and the range of \(t\) for which the solution is valid.

**Solution:** We know that \(s(t)\) is an antiderivative of \(v(t),\) so we write the following, realizing that we will use our one datum \((s(\pi/3) = 7)\) to find the additive constant later.

\[s(t) = \int v(t) \, dt = \int (1 + \tan t) \, dt = t + \ln |\sec t| + C.\]

So far \(s(t) = t + \ln |\sec t| + C,\) and \(s(\pi/3) = 7,\) so

\[7 = \frac{\pi}{3} + \ln |\sec \frac{\pi}{3}| + C \iff 7 = \frac{\pi}{3} + \ln 2 + C \iff 7 = \frac{\pi}{3} + \ln 2 + C,
\]

and so \(C = 7 - \frac{\pi}{3} - \ln 2.\) Hence

\[v(t) = t + \ln |\sec t| + 7 - \frac{\pi}{3} - \ln 2.\]
It is interesting to note that most modern tables of integrals do not use the common variable \( x \) in the formulas, but instead use \( u \), which is the most common variable for substitution type problems. This is because substitution is so ubiquitous that it is assumed the reader might not need a form exactly as it is in the table, but rather needs one which becomes one of the forms (or a constant multiple of one of the forms) found in the table only after a substitution. In that spirit, the standard method of listing the antiderivatives of the basic six trigonometric functions is as follows:

\[
\int \sin u \, du = -\cos u + C, \quad (6.70) \\
\int \cos u \, du = \sin u + C, \quad (6.71) \\
\int \tan u \, du = \ln |\sec u| + C, \quad (6.72) \\
\int \cot u \, du = -\ln |\csc u| + C, \quad (6.73) \\
\int \sec u \, du = \ln |\sec u + \tan u| + C, \quad (6.74) \\
\int \csc u \, du = -\ln |\csc u + \cot u| + C. \quad (6.75)
\]

Of course these are just our previous formulas (6.64)–(6.69), from page 568, but with the entire integral written in the variable \( u \) instead of \( x \). However, each of these properly interpreted contains a reverse chain rule, also known as a substitution-type form. So for instance, if \( u = u(x) \), then \( du = u'(x) \, dx \) and so we can read (6.72) as

\[
\int_{\tan u}^{u'(x)} \tan u(x) \, du = \ln |\sec u(x)| + C,
\]

verified by differentiation:

\[
\frac{d}{dx} \ln |\sec u(x)| = \frac{1}{\sec u(x)} \cdot \frac{d}{dx} \sec u(x) = \frac{1}{\sec u(x)} \cdot \sec u(x) \tan u(x) \cdot \frac{du(x)}{dx} = \tan u(x) \cdot u'(x),
\]

q.e.d. So forms (6.70)–(6.75) are all forms in which a basic trigonometric function of some function \( u(x) \), and the derivative \( u'(x) \), and the differential \( dx \) are the nonconstant factors of the integral. We now look at several examples.

**Example 6.5.3** Compute \( \int x \sin x^2 \, dx \).
Solution: As often occurs, the form is not exact but a constant multiple of one of our forms, this time (6.70), and furthermore the order of the factors is changed. Here we see that the factor \( x \) is a constant multiple of \( u'(x) \) if \( u(x) = x^2 \), so the extra factor of \( x \) can be “absorbed” in the differential \( du \) after the substitution. This ultimately leaves us with the problem of finding the antiderivative of a sine function.

\[
\int \frac{1}{x} \sin x^2 \, dx = \int \sin u \cdot \frac{1}{2} \, du = -\frac{1}{2} \cos u + C = -\frac{1}{2} \cos x^2 + C.
\]

\[
u = x^2
\]

\[
\Rightarrow \quad du = 2x \, dx
\]

\[
\Leftrightarrow \quad \frac{1}{2} \, du = x \, dx
\]

The above example can be quickly checked by differentiation.

**Example 6.5.4** Compute \( \int e^x \cot e^x \, dx \).

Solution: Here we see the derivative of the argument \( e^x \) of the cotangent function is also present as a multiplicative factor.

\[
\int e^x \cot e^x \, dx = \int \cot u \, du = -\ln |\csc u| + C = -\ln |\csc e^x| + C.
\]

\[
u = e^x
\]

\[
\Rightarrow \quad du = e^x \, dx
\]

**Example 6.5.5** Compute \( \int \frac{\sec \sqrt{x}}{\sqrt{x}} \, dx \).

Solution: Here the factor \( \frac{1}{\sqrt{x}} \) is in fact a constant multiple of the derivative of \( \sqrt{x} \), the argument of the trigonometric function. Thus we take \( u = \sqrt{x} \), and then the resulting \( du \) will absorb the \( \frac{1}{\sqrt{x}} \) term:

\[
\int \frac{\sec \sqrt{x}}{\sqrt{x}} \, dx = \int \sec \sqrt{x} \cdot \frac{1}{\sqrt{x}} \, dx = \int \sec u \cdot \frac{1}{2} \, du = 2 \ln |\sec u + \tan u| + C
\]

\[
u = \sqrt{x}
\]

\[
\Rightarrow \quad du = \frac{1}{2 \sqrt{x}} \, dx
\]

\[
\Leftrightarrow \quad 2 \, du = \frac{1}{\sqrt{x}} \, dx
\]

**Example 6.5.6** Compute \( \int \frac{\cos(1 + 4 \ln x)}{x} \, dx \).

Solution: Here we see the derivative of \( (1 + 4 \ln x) \) appearing as a factor as well, except for
CHAPTER 6. BASIC INTEGRATION

a constant factor.

\[
\int \frac{\cos(1 + 4 \ln x)}{x} \, dx = \int \cos(1 + 4 \ln x) \cdot \frac{1}{x} \, dx = \int \cos u \cdot \frac{1}{4} \, du = \frac{1}{4} \sin u + C
\]

\[
u = 1 + 4 \ln x
\]

\[
\Rightarrow \quad du = 4 \cdot \frac{1}{x} \, dx
\]

\[
\Leftrightarrow \quad \frac{1}{4} \cdot du = \frac{1}{x} \, dx
\]

\[
\int \cos(1 + 4 \ln x) \cdot \frac{1}{x} \, dx = \frac{1}{4} \sin(1 + 4 \ln x) + C.
\]

**Example 6.5.7** Compute \( \int x^2 \csc (\cos x^3) \sin x^3 \, dx \).

**Solution:** To be clear, first we note that the integrand is the product of three factors:

\[
x^2 \cdot \csc (\cos x^3) \cdot \sin x^3,
\]

so the argument of the cosecant is \( \cos x^3 \). Now we will compute this two different ways. The first method requires two substitutions, which is an option that students must be aware is legitimate, assuming all computations are made carefully and consistently.

**Method 1.** Here we will first make a substitution \( u = x^3 \) to yield a simpler integral without any polynomial factors, though our new integral will still require some work.

\[
\int x^2 \csc (\cos x^3) \sin x^3 \, dx = \int \csc(\cos u) \sin u \cdot \frac{1}{3} \, du
\]

\[
\Rightarrow \quad du = 3 \cdot \frac{1}{x^2} \, dx
\]

\[
\Leftrightarrow \quad \frac{1}{3} \, du = \frac{1}{x^2} \, dx
\]

So at this point our problem reduces to computing \( \int \csc(\cos u) \sin u \cdot \frac{1}{3} \, du \). To do so we use another substitution, noting that \( \sin u \) is the derivative—up to a multiplicative constant—of \( \cos u \) (with respect to \( u \) this time). To remain consistent this second substitution must use a new variable (lest we give one letter two different meanings within the same problem, which would be contradictory!). So we call our new variable something other than \( u \) or \( x \). A commonly used variable at this stage is \( w \):

\[
\frac{1}{3} \int \csc(\cos u) \sin u \, du = \frac{1}{3} \int \csc w \cdot (-1) \, dw = -\frac{1}{3} \left[ \ln |\csc w + \cot w| \right] + C
\]

\[
\Rightarrow \quad dw = -\sin u \, du
\]

\[
\Leftrightarrow \quad (-1) \, dw = \sin u \, du
\]

\[
= \frac{1}{3} \ln |\csc (\cos u) + \cot (\cos u)| + C
\]

\[
= \frac{1}{3} \ln |\csc (\cos x^3) + \cot (\cos x^3)| + C.
\]

Note how we computed the antiderivative in \( w \), which we then replaced by its expression in \( u \), and finally by the definition of \( u \) in terms of \( x \).
Method 2. If we can see far enough ahead, we can combine both substitutions into one. For clarity we will use a different variable—namely $z$—here (though by convention one would usually use $u$). We choose $z = \cos x^3$, noting that its derivative, requiring the chain rule, will have a $\sin x^3$ and a $x^2$ term (ignoring multiplicative constants), which leaves us with a constant multiple of $\int \csc z \, dz$, for which we have a formula.

$$
\int x^2 \csc (\cos x^3) \sin x^3 \, dx = \int \csc z \cdot \frac{-1}{3} \, dz = -\frac{1}{3} \cdot [-\ln |\csc z + \cot z|] + C
$$

$$
z = \cos x^3
\implies \quad dz = -\sin x^3 \cdot 3x^2 \, dx
\iff \quad -\frac{1}{3} \, dz = x^2 \sin x^3 \, dx
$$

$$
= \frac{1}{3} \ln |\csc (\cos x^3) + \cot (\cos x^3)| + C.
$$

In the previous example, the second method in fact just combines the two substitutions from the first method into one. Indeed, the formula for $w$ in terms of $x$ is the same as that of $z$:\footnote{Similarly, though perhaps not so obviously, when we recall these variables are all functions of $x$ we also have $dw = dz$:

$$
dw = \frac{dw}{du} \cdot \frac{du}{dx} \cdot dx = (-\sin u) \cdot 3x^2 \, dx = -\sin x^3 \cdot 3x^2 \, dx = dz.
$$

Again we see the power of the Leibniz notation in what is essentially a chain rule. Of course we should expect that $w = z \implies dw = dz$. But also we see that while there are obvious algebraic consistencies in our substitution method, there are also consequent calculus consistencies which, while more subtle, are still correct when we perform all the computations correctly.}

$$
w = \csc(u) = \csc (\cos x^3) = z.
$$

When one is well practiced in substitution the second method will likely be chosen. However, it is important also for the student to realize that even if a substitution does not achieve an integral that can be immediately computed, that does not mean that the particular substitution need be abandoned. If the new integral is simpler, then the first substitution can be worthwhile. In fact in later sections we will on occasion require multiple substitutions. Of course it is important that all steps be carried out carefully, accurately and consistently.

In this section we concentrated on those integrals which reduce to integrals of a single trigonometric function, perhaps with the aid of a substitution. In Chapter 7 (and more specifically Section 7.3) we will look at the many techniques for computing those integrals which contain several factors of trigonometric functions, and no other factors. The techniques of our present section will be called upon often, but these are only a small part of the needed knowledge for computing the “trigometric integrals” of the later sections. But in fact we have some other techniques already. For instance, there were the first trigonometric integral formulas we had in Section 6.1, Subsection 6.1.4 which arose from the derivative rules for the six basic trigonometric functions. In fact we had one other technique for dealing with some trigonometric integrals, which was substitution in the case we could rewrite the trigonometric integral as a power-rule type integral.\footnote{In fact there will be several other substitution type arguments we will make for trigonometric integrals besides those which yield power rules.}
Example 6.5.8 Compute \( \int \frac{\sin 3x}{\cos^3 3x} \, dx \).

Solution: Here we see that we have the derivative of the cosine function is present as a factor, and we are left with a power of the cosine:

\[
\int \frac{\sin 3x}{\cos^3 3x} \, dx = \int (\cos 3x)^{-5} \sin 3x \, dx = \int u^{-5} \cdot \frac{-1}{3} \, du
\]

\[
\begin{align*}
\Rightarrow \quad & u = \cos 3x \\
\Rightarrow \quad & du = -\sin 3x \cdot 3 \, dx \\
\Leftrightarrow \quad & -\frac{1}{3} \, du = \sin 3x \, dx
\end{align*}
\]

\[
= -\frac{1}{3} \cdot \frac{1}{4} u^{-4} + C
= \frac{1}{12} \cos^{-4} 3x + C
= \frac{1}{12} \sec^4 3x + C.
\]

The integration techniques we encounter throughout the book are many and varied. We will see later how a slight change in a problem can substantially change the result, its difficulty, or the technique used to achieve it. We have seen this phenomenon before. Consider for instance

\[
\begin{align*}
\int \frac{x}{x^2 + 1} \, dx &= \frac{1}{2} \ln (x^2 + 1) + C, \\
\int \frac{1}{x^2 + 1} \, dx &= \tan^{-1} x + C, \\
\int \frac{x}{\sqrt{1 - x^2}} \, dx &= -\sqrt{1 - x^2} + C, \\
\int \frac{1}{\sqrt{1 - x^2}} \, dx &= \sin^{-1} x + C, \\
\int \sec x \, dx &= \ln |\sec x + \tan x| + C, \\
\int \sec^2 x \, dx &= \tan x + C, \\
\int \sec^3 x \, dx &= \frac{1}{2} (\sec x \tan x + \ln |\sec x + \tan x|) + C.
\end{align*}
\]

In fact this last problem will have to wait until Chapter 7, and is quite long and technical. Even the verification by differentiation is nontrivial, and requires one to employ some trigonometric identity along the way. Suffice for now to simply note that the techniques, and results, for even these first three powers of the secant are all very different.

Computing antiderivatives in good time requires the ability to recognize which technique will work for a particular problem. That in turn requires a fairly complete knowledge of the techniques, even to the extent that one can anticipate the outcomes of several later steps. Of course practice is one key to gaining this understanding. For that reason this chapter will contain one section in which the exercises’ required techniques are purposely randomized, by method as well as difficulty.
Exercises

1. By differentiation, verify each of our basic six trigonometric integrals in the forms we use, (6.64)–(6.69). For reference see the proof for the secant, page 569.
   
   (a) \[ \int \sin x \, dx = -\cos x + C \]
   
   (b) \[ \int \cos x \, dx = \sin x + C \]
   
   (c) \[ \int \tan x \, dx = \ln |\sec x| + C \]
   
   (d) \[ \int \cot x \, dx = -\ln |\csc x| + C \]
   
   (e) \[ \int \sec x \, dx = \ln |\sec x + \tan x| + C \]
   
   (f) \[ \int \csc x \, dx = -\ln |\csc x + \cot x| + C \]

2. Compute the following integrals.
   
   (a) \[ \int x \sec (x^2 + 1) \, dx \]
   
   (b) \[ \int \frac{\tan(ln x)}{x} \, dx \]
   
   (c) \[ \int \frac{\sin \left(\frac{1}{x^2}\right)}{x^2} \, dx \]
   
   (d) \[ \int \sqrt{x} \csc (x\sqrt{x}) \, dx \]
   
   (e) \[ \int x^3 e^{5x^4} \cot \left(\frac{6e^{5x^4}}{7}\right) \, dx \]
   
   (f) \[ \int \frac{x}{\sin 3x^2} \, dx \]
   
   (g) \[ \int 2^x \cos(3 \cdot 2^x) \, dx \]

   (h) \[ \int (1 - \sec x)^2 \, dx \]
   
   (i) \[ \int (\tan 7x - 2)^2 \, dx \]
   
   (j) \[ \int \frac{1 - \sin^2 x}{\cos x} \, dx \]
   
   (k) \[ \int \frac{\sin^2 x}{\cos x} \, dx \] (for fun)
   
   (l) \[ \int \frac{(1 - \cos x)^2}{\sin x} \, dx \] (for fun)

3. Derive our formula (above) for the integral of \( \sec x \) by the following algebraic device, namely multiplying and dividing by \( \sec x + \tan x \) within the integral, i.e.,

\[
\int \sec x \, dx = \int \sec x \cdot \frac{\sec x + \tan x}{\sec x + \tan x} \, dx,
\]

and then using an appropriate substitution argument.

4. By differentiation, verify each of the following alternative integration formulas.
   
   (a) \[ \int \tan x \, dx = -\ln |\cos x| + C. \]
   
   (b) \[ \int \cot x \, dx = \ln |\sin x| + C. \]
   
   (c) \[ \int \sec x \, dx = -\ln |\sec x - \tan x| + C. \]
   
   (d) \[ \int \csc x \, dx = \ln |\csc x - \cot x| + C. \]


6.6 Substitution with All Basic Forms

In this section we will add to our forms for substitution and recall some rather general guidelines for substitution. Except for our four new trigonometric forms from Section 6.5, all forms in this chapter derive directly from derivative rules. These comprise what we call here the basic integration rules. Each is based upon a single function specific to the rule. So for instance, we will have in our list the following:

\[
\int \frac{1}{u^2 + 1} \, du = \tan^{-1} u + C.
\]

As before, the usual variable of integration in the given problem will likely be \(x\), but the form may be ultimately as above, except for multiplicative constants, where \(u = u(x)\) and then \(du = u'(x) \, dx\) contains another factor from the original integral. So for instance we might see

\[
\int \frac{x}{x^4 + 1} \, dx = \int \frac{1}{(x^2)^2 + 1} \cdot x \, dx = \int \frac{1}{u^2 + 1} \cdot \frac{1}{2} \, du = \frac{1}{2} \tan^{-1} u + C
\]

One clue that we might try \(u = x^2\) was that its derivative was a factor in the integrand in the form of the factor \(x\), again excepting multiplicative constants, and so we wrote that factor separately next to the differential \(dx\). As it turned out, the rest of the integrand could indeed be written as a function of \(u = x^2\).

Reading the problem above backwards, the arctangent is the “outer function” of a chain rule differentiation problem, and \(x^2\) was the “inner function.” Put in terms of integration, the form was \(\int \frac{1}{u^2 + 1} \, du\), excepting multiplicative constants, with the “inner function” \(u = x^2\). The arctangent appeared because of the ultimate form of the integral, in terms of \(u = x^2\).

But note that the arctangent can also appear as the “inner function,” which we may wish to set equal to \(u\). So for instance,

\[
\int \frac{(\tan^{-1} x)^3}{x^2 + 1} \, dx = \int (\tan^{-1} x)^2 \cdot \frac{1}{x^2 + 1} \, dx = \int u^2 \, du = \frac{u^3}{3} + C
\]

In all these cases, we are looking for some \(u = u(x)\) so that

- one major (nonconstant) factor of the integral can be simply written \(f(u)\), i.e., \(u\) is an “inner function” of some composite function \(f(u(x))\) which appears in the integrand,
- the remaining factors of the integrand will collectively be a constant multiple of \(du = u'(x) \, dx\),
- and finally, so that \(\int f(u) \, du\) is an integral we can compute, i.e., we know the antiderivative of \(f\).
So of course identifying \( u \) is the key, and in doing so we have to be sure its derivative is also present, and finally that we are left with an integral—albeit in \( u \)—which we can handle.  \(^{29}\)

### 6.6.1 List of Basic Forms

\[
\begin{align*}
\int u^n \, du &= \frac{u^{n+1}}{n+1} + C, \quad n \neq -1, \\
\int \frac{1}{u} \, du &= \ln |u| + C, \\
\int \sin u \, du &= -\cos u + C, \\
\int \cos u \, du &= \sin u + C, \\
\int \tan u \, du &= \ln |\sec u| + C, \\
\int \cot u \, du &= -\ln |\csc u| + C, \\
\int \sec u \, du &= \ln |\sec u + \tan u| + C, \\
\int \csc u \, du &= -\ln |\csc u + \cot u| + C, \\
\int \frac{1}{\sqrt{1-u^2}} \, du &= \sin^{-1} u + C, \\
\int \frac{1}{u^2+1} \, du &= \tan^{-1} u + C, \\
\int \frac{1}{u\sqrt{u^2-1}} \, du &= \sec^{-1} |u| + C, \\
\int e^u \, du &= e^u + C, \\
\int a^u \, du &= \frac{a^u}{\ln a} + C, \\
\int \sec^2 u \, du &= \tan u + C, \\
\int \csc^2 u \, du &= -\cot u + C, \\
\int \sec u \tan u \, du &= \sec u + C, \\
\int \csc u \cot u \, du &= -\csc u + C.
\end{align*}
\]

\(^{29}\)This all assumes that there is a substitution which will make the integral into one of these simple forms. It is not always the case. One which occurs in probability and other subjects is \( \int e^{x^2} \, dx \), which can not be changed by substitution into a useful form for simple integration. In fact it will not succumb to any of the methods in this or the next chapter. We will eventually find a way to deal with this integral, in Chapter 11. In the meantime it is actually a good exercise to see why this integral can not be forced into any of our methods. Indeed, seeing what goes wrong in such a case very well complements seeing what goes right in the cases where substitution, and later methods, do achieve an answer.
It can not be stressed too much that each form given assumes that a substitution may be required. So again, the following formulas say the same:

\[
\int e^u \, du = e^u + C, \quad \int e^{u(x)} u'(x) \, dx = \int e^{u(x)} \cdot \frac{du(x)}{dx} \, dx = e^{u(x)} + C.
\]

Recognizing when we have such a form is again key to using these formulas.

**Example 6.6.1** Compute \( \int xe^{x^2} \, dx \).

**Solution:** Here we see the derivative of \( x^2 \) appearing as the factor \( x \), except for a constant multiple. Hence we let \( u = x^2 \), and the \( du \) will contain the other factor \( x \), leaving an integral in one of our standard forms, namely (6.87), and nothing else except a multiplicative constant.

\[
\int xe^{x^2} \, dx = \int e^u \cdot \frac{1}{2} \, du = \frac{1}{2} e^u + C
\]

\( u = x^2 \)

\[\Rightarrow\]

\[du = 2x \, dx\]

\[\Leftrightarrow\]

\[\frac{1}{2} du = x \, dx\]

**Example 6.6.2** Compute \( \int \frac{x^3}{\sqrt{1-x^8}} \, dx \).

**Solution:** At first it is tempting to let \( u = 1-x^8 \) and hope this will become a power rule, except that such \( u \) implies \( du = -8x^7 \), which is very different from a constant multiple of the other factor here, namely \( x^3 \).

In fact, the other factor can be a good source of information about what to set equal to \( u \). Indeed the factor \( x^3 \) will be part of the differential of \( u = x^4 \), and then we can recognize a form which will yield an arcsine ultimately, i.e., form (6.84).

\[
\int \frac{x^3}{\sqrt{1-x^8}} \, dx = \int \frac{1}{\sqrt{1-(x^4)^2}} \cdot x^3 \, dx = \int \frac{1}{\sqrt{1-u^2}} \cdot \frac{1}{4} \, du = \frac{1}{4} \sin^{-1} u + C
\]

\( u = x^4 \)

\[\Rightarrow\]

\[du = 4x^3 \, dx\]

\[\Leftrightarrow\]

\[\frac{1}{4} \, du = x^3 \, dx\]

As with the power rule, there are occasions where the derivative of our \( u \) is a nonzero constant, and thus a constant multiple of every other nonzero constant. While these integrals are arguably easier than the others we encounter here, their relative simplicity can be a source of confusion.

**Example 6.6.3** Compute \( \int \csc 5x \, dx \).

**Solution:** Here we simply let \( u = 5x \).

\[
\int \csc 5x \, dx = \int \csc u \cdot \frac{1}{5} \, du = \frac{1}{5} \ln | \csc u + \cot u | + C
\]

\( u = 5x \)

\[\Rightarrow\]

\[du = d \, dx\]

\[\Leftrightarrow\]

\[\frac{1}{5} du = d \, dx\]
Example 6.6.4 Compute $\int 2^x \cdot 3^x \, dx$.

**Solution:** Here we will need (6.88) eventually, but first we simply notice that the factor $3^x$ is a constant multiple of the exponent in the first factor, so we let $u = 3^x$.

$$\int 2^x \cdot 3^x \, dx = \int 2^u \cdot \frac{1}{\ln 3} \, du = \frac{1}{\ln 3} \cdot \frac{1}{\ln 2} \cdot 2^u + C$$

$$\begin{align*}
\Rightarrow \quad & du = 3^x \ln 3 \, dx \\
\Leftrightarrow \quad & \frac{1}{\ln 3} \, du = 3^x \, dx
\end{align*}$$

The example above shows the importance of following the various constant factors through the integration. Students who rely upon guessing the answers, without performing the formal substitution steps, are much more likely to misplace one or more constant factors.

For further practice we consider more basic examples.

Example 6.6.5 Compute $\int \sec \sqrt{x} \sqrt{x} \, dx$.

**Solution:** The key here is that the derivative of the argument of the secant is also present as a factor. Recall $\frac{d}{dx}(\sqrt{x}) = \frac{1}{2\sqrt{x}}$, which is obvious when the radicals are written as $1/2$-powers.

$$\begin{align*}
\int \sec \sqrt{x} \sqrt{x} \, dx &= \int \sec \sqrt{x} \cdot \frac{1}{\sqrt{x}} \, dx = \int \sec u \cdot 2 \, du = 2 \ln |\sec u + \tan u| + C \\
\Rightarrow \quad & du = \frac{1}{2} \cdot x^{-1/2} \, dx \\
\Leftrightarrow \quad & = 2 \ln |\sec \sqrt{x} + \tan \sqrt{x}| + C.
\end{align*}$$

In the next section we will further explore the arctrigonometric antiderivatives by considering further complications. For now we only look at two such complications, first involving the arctangent (though the arcsine has a similar potential complication), and then the arcsecant, which has the same complications as the others and then one more.

Example 6.6.6 Compute $\int \frac{x^2}{1 + 25x^6} \, dx$.

**Solution:** There are two clues directing our choice of $u$. First, we see the factor $x^2$, which is a multiple of the derivative of $x^3$. Then we see that the denominator can be written as $1 + (5x^3)^2$, which we can put into the form yielding the arctangent, namely (6.85).

$$\begin{align*}
\int \frac{x^2}{1 + 25x^6} \, dx &= \int \frac{1}{1 + (5x^3)^2} \cdot x^2 \, dx = \int \frac{1}{1 + u^2} \cdot \frac{1}{15} \, du = \frac{1}{15} \tan^{-1} u + C \\
\Rightarrow \quad & du = 15x^2 \, dx \\
\Leftrightarrow \quad & = \frac{1}{15} \tan^{-1} (5x^3) + C.
\end{align*}$$

The complication in this example is fairly benign: that the $u$ term contains a multiplicative constant. Here we wanted $25x^6$ to be $u^2$ for the form (6.85), so we took $u = 5x^3$. Fortunately

\[\text{Note that we could also have used } u = -5x^3, \text{ but then } du = -15x^2 \, dx, \text{ and so our answer would ultimately be (as the reader should verify)} - \frac{1}{15} \tan^{-1} (-5x^3) + C. \text{ In fact that is the same as the answer we got, since the arctangent is an “odd” function, that is, } \tan^{-1}(-z) = -\tan^{-1} z.\]
this was consistent with the $du$ containing the $x^2$ term of the integrand, and that form (6.85) could actually be used. In the next section we will see how to deal with cases where the additive constant in the denominator of the integrand is not 1. For now we look at another complication which is somewhat specific to the arcsecant form.

**Example 6.6.7** Compute $\int \frac{1}{x \sqrt{9x^2 - 1}} \, dx$.

**Solution:** Because a new complication needs to be explained while the problem is solved, the organization will be slightly different than previous exercises, but every technique used below has appeared previously. Note that we are trying to fit this integral into form (6.86).

Here we want $u^2 = 9x^2$ so we have $\sqrt{u^2 - 1}$ as one factor in the denominator of our integrand. Thus will let $u = 3x \Rightarrow du = 3 \, dx \iff \frac{1}{3} \, du = dx$. Our integral so far is then

$$\int \frac{1}{x \sqrt{u^2 - 1}} \cdot \frac{1}{3} \, du.$$ 

Now, all of our integral formulas require just one variable, but in fact the integral above makes sense because of the relationship between $x$ and $u$. But to use a formula we have to put it all into the new variable, namely $u$. So there is one term left, which is the factor $x$ on the bottom, which has to be put into $u$-terms. For that we go back to our original substitution and note that $u = 3x \iff x = \frac{1}{3} \, u$. Now we continue:

$$\int \frac{1}{x \sqrt{9x^2 - 1}} \, dx = \int \frac{1}{x \sqrt{u^2 - 1}} \cdot \frac{1}{3} \, du$$ 

$$= \int \frac{1}{\frac{u}{3} \sqrt{u^2 - 1}} \cdot \frac{1}{3} \, du$$ 

$$= \int \frac{1}{u \sqrt{u^2 - 1}} \, du$$ 

$$= \sec^{-1} |u| + C$$ 

$$= \sec^{-1} |3x| + C.$$ 

With practice, our process of solving the above problem could more resemble the following:

$$\int \frac{1}{x \sqrt{9x^2 - 1}} \, dx = \int \frac{1}{x \sqrt{(3x)^2 - 1}} \, dx = \int \frac{1}{u \sqrt{u^2 - 1}} \cdot \frac{1}{3} \, du$$ 

$$\Longrightarrow u = 3x \quad \Rightarrow \quad du = 3 \, dx$$ 

$$\Rightarrow \frac{1}{3} \, du = dx, \quad u = 3x$$ 

Also, 

$$\frac{1}{3} u = x$$ 

The next example illustrates why it is important to recall that $(a^n)^m = a^{mn}$. In particular for the example below, $e^{2x} = (e^x)^2$. 

Example 6.6.8 Compute $\int \frac{e^x}{\sqrt{1 - e^{2x}}} \, dx$.

**Solution:** Here we note a form $\int \frac{1}{\sqrt{1 - u^2}} \cdot \frac{du}{dx} \cdot dx$ if we let $u = e^x$.

\[
\int \frac{e^x}{\sqrt{1 - e^{2x}}} \, dx = \int \frac{1}{\sqrt{1 - e^{2x}}} \cdot e^x \, dx
\]

Let $u = e^x \Rightarrow du = e^x \, dx$.

\[
\int \frac{1}{\sqrt{1 - u^2}} \, du = \sin^{-1} u + C = \sin^{-1} e^x + C.
\]

On the other hand, we have to be careful that we do not try to “read” arc-trigonometric or other forms into what may be simple power rules.

Example 6.6.9 Compute $\int \frac{e^{2x}}{\sqrt{1 - e^{2x}}} \, dx$.

**Solution:** Here we note that the derivative of the function under the radical is also a factor, leaving only constant multiple factors:

\[
\int \frac{e^{2x}}{\sqrt{1 - e^{2x}}} \, dx = \int (1 - e^{2x})^{-1/2} \cdot e^{2x} \, dx = \int u^{-1/2} \cdot \frac{1}{2} \, du
\]

Let $u = 1 - e^{2x} \Rightarrow du = -e^{2x} \cdot 2 \, dx$.

\[
\Rightarrow -\frac{1}{2} \, du = e^{2x} \, dx
\]

\[
\Rightarrow \frac{1}{2} \, du = e^{2x} \, dx
\]

\[
\int u^{-1/2} \, du = \sqrt{1 - e^{2x}} + C = -\sqrt{1 - e^{2x}} + C.
\]
Exercises

1. Compute \( \int \frac{1}{(x^2 + 1) \tan^{-1} x} \, dx \)
2. Compute \( \int \frac{1 + e^x}{x + e^x} \, dx \)
3. Compute \( \int \frac{\sec \ln x}{x} \, dx \)
4. Compute \( \int \frac{\sqrt{\ln 9x}}{x} \, dx \)
5. Compute \( \int \frac{x^2}{x^3 + 1} \, dx \)
6. Compute \( \int \frac{x^3 + 1}{x^2} \, dx \)
7. Compute \( \int \frac{x^2}{x^6 + 1} \, dx \)
8. Compute \( \int \frac{\sec x \tan x}{\sqrt{1 - \sec x}} \, dx \)
9. Compute \( \int \frac{e^x}{(1 + e^x)^2} \, dx \)
10. Compute \( \int \csc^2 3x \cot^6 3x \, dx \)
11. Compute \( \int \frac{1}{\sqrt{x(x + 1)}} \, dx \)
12. Compute \( \int 2x^3 3x \, dx \)
13. Compute \( \int \frac{2x}{1 + 4x} \, dx \)
14. Compute \( \int 5x \sec 5x \, dx \)
15. Compute \( \int \frac{1}{\sqrt{1 - 5x^2}} \, dx \)
16. Compute \( \int \frac{x - 1}{\sqrt{1 - 4x^2}} \, dx \)
17. Compute \( \int xe^{x^2} \, dx \) two ways.
18. In the spirit of the previous exercise, compute \( \int e^{1/x} x^2 \, dx \) two ways, i.e., using two different substitutions.
19. Compute \( \int \sec^2 x \tan x \, dx \) two different ways.
    (a) Let \( u = x^2 \) (as in Example 6.6.1).
    (b) Instead let \( u = e^{x^2} \).
20. Though it can not be rewritten into a basic form, it is interesting to attempt some substitutions for \( \int e^{x^2} \, dx \).
    (a) Let \( u = x^2 \) and attempt to rewrite the integral with this substitution.
    (b) Let \( u = x \) and attempt to rewrite the integral with this substitution.
    Explain why this is never useful for a substitution attempt for any integral of type \( \int f(x) \, dx \).
6.7 Further Arctrigonometric Forms

Here we will still use the same arctrigonometric forms we had before, namely

\[ \int \frac{1}{\sqrt{1 - u^2}} du = \sin^{-1} u + C, \]  
(6.93)
\[ \int \frac{1}{u^2 + 1} du = \tan^{-1} u + C, \]  
(6.94)
\[ \int \frac{1}{u\sqrt{u^2 - 1}} du = \sec^{-1} |u| + C. \]  
(6.95)

What is different in this section is that our integrals will need to be algebraically rewritten into these forms, and this will require more than the previous substitution.

Each of the integrals in (6.93), (6.94) and (6.95) need to be exactly as they are stated. For instance, replacing \(1 - u^2\) with \(1 + u^2\) or \(u^2 - 1\) in (6.93) will give completely different antiderivatives. In fact, even the domain of the integrand would be completely different with any such changes! Similar changes would substantially alter the results in (6.94) and (6.95).

In this section we will have integrands which we can algebraically rewrite so they conform to one of the forms (6.93), (6.94) or (6.95). In fact there are only a couple of algebraic “tricks” which we introduce here. The first of these is to force the denominators to have the additive constant 1, where originally there may be another constant. This is accomplished through simple factoring techniques. The second technique is “completing the square,” where appropriate, and then using the first technique to finish rewriting the integrand. With substitution there will often be further multiplicative constants to accommodate as well.

6.7.1 Factoring to Achieve “1”

Each of our integrals (6.93), (6.94) and (6.95) have the number 1 conspicuously appearing in the denominator, near the \(u^2\) term. Any other nonzero constant there will have an effect on the vertical and horizontal scaling of the function in ways we can not ignore in the formula. To compensate is fairly straightforward: factor the constant, and see what should be called “\(u^2\).”

**Example 6.7.1** Compute \( \int \frac{1}{9 + x^2} \, dx \).

**Solution:** Our first priority is to rewrite this so we have a form \(1 + u^2\) in the denominator.

\[ \int \frac{1}{9 + x^2} \, dx = \int \frac{1}{9(1 + \frac{x^2}{9})} \, dx = \int \frac{1}{9} \cdot \frac{1}{1 + \frac{x^2}{9}} \, dx. \]

The factor \(\frac{1}{9}\) can simply be carried along for the rest of the computation. The denominator of the other factor can be written \(1 + u^2\) (same as \(u^2 + 1\) in our formula) if we take \(u^2 = \frac{x^2}{9}\), which can be accomplished letting \(u = \frac{x}{3}\).

\[
\begin{align*}
\int \frac{1}{9 + x^2} \, dx &= \frac{1}{9} \cdot \int \frac{1}{1 + \frac{x^2}{9}} \, dx = \frac{1}{9} \int \frac{1}{1 + u^2} \cdot 3 \, du \\
\Rightarrow \quad du &= \frac{1}{3} \, dx \\
\Leftrightarrow \quad 3 \, du &= \, dx \\
&= \frac{3}{9} \tan^{-1} u + C = \frac{1}{3} \tan^{-1} \left( \frac{x}{3} \right) + C.
\end{align*}
\]
The above example illustrated much of the process: algebraically manipulate by factoring to achieve “1” in the appropriate place, and then pick $u$ so the other term is $u^2$. It is slightly more complicated with the forms yielding arcsine and arcsecant, due to the presence of the radical. In the next example we will show more detail than one might normally write.

**Example 6.7.2** Compute \( \int \frac{1}{\sqrt{25 - 4x^2}} \, dx \).

**Solution:** Here we must find a way to replace the constant 25 with 1 instead. We factor as before, but respect the operation of the radical as well.

\[
\int \frac{1}{\sqrt{25}} \, dx = \int \frac{1}{\sqrt{25}} \, dx = \int \frac{1}{\sqrt{25}} \, dx = \int \frac{1}{5} \cdot \frac{1}{\sqrt{1 - \frac{4x^2}{25}}} \, dx.
\]

So the factor 25 under the radical becomes the factor 5 outside the radical. Otherwise it is the same process as the previous example. Now we continue, using a substitution which will result in $u^2 = \frac{4x^2}{25}$. For simplicity we take $u = \frac{2x}{5}$.

\[
\int \frac{1}{\sqrt{25 - 4x^2}} \, dx = \frac{1}{5} \int \frac{1}{\sqrt{1 - \frac{4x^2}{25}}} \, dx = \frac{1}{5} \int \frac{1}{\sqrt{1 - u^2}} \cdot \frac{5}{2} \, du
\]

This example above again illustrates the role of the number 1 in the denominator, but also suggests a couple of new points that we make here. First, it is not obvious where the factors 2 and 5—being the square roots of the 4 and 25 appearing in the original—will be present in the final answer. There is a pattern for the arctangent form, and a different one for the arcsine form, but patterns can be forgotten if not used often enough, where the logic of manipulating the integral algebraically to get one of the three basic arctrigonometric forms should still be reproducible after the patterns—which we will explore at the end of this section—are forgotten. Second, we are approaching the boundary between integrals which are easily checked with differentiation, and those where the differentiation has at least as many algebraic difficulties as the integration. In such cases, it is usually better to have carefully written each integration step so it can be audited for accuracy, rather than risk algebraic error in testing our answer. Consider a verification of the answer in this latest example (readers’ steps may vary):

\[
\frac{d}{dx} \left[ \frac{1}{2} \sin^{-1} \left( \frac{2x}{5} \right) \right] = \frac{1}{2} \cdot \frac{1}{\sqrt{1 - \left( \frac{2x}{5} \right)^2}} \cdot \frac{d}{dx} \left[ \frac{2x}{5} \right] = \frac{1}{2} \cdot \frac{1}{\sqrt{1 - \frac{4x^2}{25}}} \cdot \frac{2}{5}
\]

\[
= \frac{1}{5} \cdot \frac{1}{\sqrt{25 - 4x^2}} = \frac{1}{\sqrt{25}} \cdot \frac{1}{\sqrt{1 - \frac{4x^2}{25}}} = \frac{1}{\sqrt{25 - 4x^2}}, \quad \text{q.e.d.}
\]

While such a verification is certainly possible, it is not likely one to be performed “mentally” with much confidence, as we may have been able to do with many previous computations. Indeed there are enough constants to be accommodated that this verification should be done in careful writing. In most of Chapter 7 we will see much more complicated rewritings of integrals,
and verification will usually be much better accomplished by checking our individual steps in integration rather than by differentiating of our answers.

Our next example just takes this theme one step further. Recall that substitution in the arcsecant form had a slight complication, which was that the \( u \)-variable appeared both inside and outside the radical. This caused a minor complication in Example 6.6.7, page 580 for instance. A similar problem will occur in this next example.

Example 6.7.3 Compute \( \int \frac{1}{x \sqrt{81x^2 - 16}} \, dx \).

**Solution:** As with the previous two examples, it is necessary to have a 1 in the place presently occupied by 16, so we will factor the 16 from the radical. The other algebraic difficulties will be taken care of by the substitution. Indeed, the remaining term under the radical must be \( u^2 \), and the rest of the form will follow, with residual multiplicative constants.

\[
\int \frac{1}{x \sqrt{81x^2 - 16}} \, dx = \int \frac{1}{4x \sqrt{\frac{81x^2}{16} - 1}} \, dx = \int \frac{1}{4} \cdot \frac{1}{\frac{9}{4}} \sqrt{u^2 - 1} \cdot \frac{4}{9} \, du
\]

\[
\begin{align*}
u &= \frac{9x}{4} \\
\Rightarrow \quad du &= \frac{9}{4} \cdot dx \\
\text{also,} \quad u &= \frac{9x}{4} \\
\text{so,} \quad x &= \frac{4u}{9}
\end{align*}
\]

Note how the term \( \frac{81x^2}{16} \) under the radical became simply \( u^2 \), and then the term \( x \) outside the radical became \( \frac{4u}{9} \), both consistent with \( u = \frac{9x}{4} \). Also note that a factor of \( \frac{4}{9} \) in the denominator canceled with the same factor multiplying the differential \( du \).

This latest example again illustrates the points made before: that having the 1-term in the denominator is the key to the whole process, that the rest is taken care of by the substitution which follows and finally, that checking by differentiation is nontrivial.

Another minor complication is that the numbers we must factor might not be perfect squares. The process is exactly the same, though perhaps some more care is required.

Example 6.7.4 Compute \( \int \frac{1}{\sqrt{5 - 2x^2}} \, dx \).

**Solution:** The process is exactly the same as before. The key is to factor the denominator to have a 1 in the place of the 5:

\[
\int \frac{1}{\sqrt{5 - 2x^2}} \, dx = \int \frac{1}{\sqrt{\frac{5}{2} - \frac{2}{5} x^2}} \, dx = \frac{1}{\sqrt{\frac{5}{2}}} \int \frac{1}{\sqrt{1 - \left( \frac{\sqrt{2}}{5} \cdot x \right)^2}} \, \cdot \frac{\sqrt{2}}{5} \cdot du
\]

\[
\begin{align*}
u &= \frac{\sqrt{5}}{2} \cdot x \\
\Rightarrow \quad du &= \frac{\sqrt{2}}{5} \cdot dx \\
\text{so,} \quad \frac{\sqrt{5}}{2} \cdot du &= dx
\end{align*}
\]

\[
\frac{1}{\sqrt{\frac{5}{2}}} \sin^{-1} u + C = \frac{1}{\sqrt{2}} \sin^{-1} \left( \frac{\sqrt{2}}{5} \cdot x \right) + C.
\]
6.7.2 Completing the Square

In the previous subsection our first concern after identifying our target form was to rewrite the integrand to have the number 1 in the appropriate place in the denominator. In this subsection our first concern is identifying, except for a multiplicative constant, what will be \( u^2 \). We do this by completing the square first, and then fixing the form to have the number 1 where we need it, and working from there as before. As there are differing levels of difficulty in such problems, we will begin with one of the simplest and continue from there. It should be noted that the completing the square step is sometimes needed before determining that one of our three forms, (6.93), (6.94) or (6.95), can even be achieved. If not, and we are fortunate, another earlier method may work, though usually we should notice that before attempting the method here. If no earlier method will work, there may be a method available in Chapter 7 that will solve the problem.

Recall that when completing the square, one adds and subtracts \((b/2)^2\), where the original polynomial is \( x^2 + bx \), or more generally \( x^2 + bx + c \):

\[
x^2 + bx + c = x^2 + bx + \left( \frac{b}{2} \right)^2 - \left( \frac{b}{2} \right)^2 + c
\]

\[
= \left( x + \frac{b}{2} \right)^2 - \left( \frac{b}{2} \right)^2 + c.
\]

As we will see, the fact that the coefficient of \( x^2 \) was 1 was key to the computation above. If not, the leading coefficient will be factored from the \( x^2 \) and \( x \) terms. Our first examples will not require that initial factoring.

**Example 6.7.5** Compute \( \int \frac{1}{x^2 + 2x + 2} \, dx \).

**Solution:** The hope is that we can somehow write the denominator as \( 1 + u^2 \), perhaps multiplied by some nonzero constant, without introducing any more variable factors. For this one we are unusually fortunate. Note that here “\( b \)” is 2.

\[
\int \frac{1}{x^2 + 2x + 2} \, dx = \int \frac{1}{x^2 + 2x + \left( \frac{2}{2} \right)^2 - \left( \frac{2}{2} \right)^2 + 2} \, dx = \int \frac{1}{x^2 + 2x + 1 - 1 + 2} \, dx
\]

\[
u = x + 1 \quad \Rightarrow \quad du = dx
\]

\[
\int \frac{1}{(x+1)^2 + 1} \, dx = \int \frac{1}{u^2 + 1} \, du = \tan^{-1} u + C = \tan^{-1}(x + 1) + C.
\]

What made this last example particularly simple was that the additive constant outside of the perfect square was already 1, which is of course key to our arctigonometric antiderivative forms. If not, we have to perform some division.

**Example 6.7.6** Compute \( \int \frac{1}{x^2 + 6x + 17} \, dx \).

**Solution:** Here \( b = 3 \), so we add and subtract \((\frac{b}{3})^2 = 9\).

\[
\int \frac{1}{x^2 + 6x + 17} \, dx = \int \frac{1}{x^2 + 6x + 9 - 9 + 17} \, dx = \int \frac{1}{(x + 3)^2 + 8} \, dx = \frac{1}{8} \int \frac{1}{\left( \frac{x+3}{\sqrt{8}} \right)^2 + 1} \, dx
\]

\[
u = \frac{x + 3}{\sqrt{8}} \quad \Rightarrow \quad du = \frac{1}{\sqrt{8}} \, dx
\]

\[
\Leftrightarrow \sqrt{8} \, du = dx
\]

\[
= \frac{1}{8} \int \frac{1}{u^2 + 1} \cdot \sqrt{8} 
\]

\[
= \frac{1}{\sqrt{8}} \tan^{-1} u + C = \frac{1}{\sqrt{8}} \tan^{-1} \left( \frac{x + 3}{\sqrt{8}} \right) + C.
\]
As this last example illustrates, the final form of the antiderivative can be more complicated when completing the square is required. While it would be an interesting exercise to verify the answer by differentiation, perhaps verifying each individual step in the solution process would be a more efficient means of verifying the answer we derived.

For simplicity we will continue with arctangent forms for the moment, as we look at the next complication, which is that the coefficient of $x^2$ is not equal to 1. In such a case we factor that leading coefficient out of the entire polynomial, or at least out of the $x^2$ and $x$ terms. It is then important to perform the addition and subtraction steps of completing the square within the factor with the $x^2$ and $x$ terms; the addition and subtraction of the $(b/2)^2$ in the process must occur simultaneously and beside each other. Note that such a term has a different effect inside parentheses (or brackets) compared to outside, so we must have the addition and subtraction steps together in order that numerically they have no net effect.

**Example 6.7.7** Compute $\int \frac{1}{5x^2 - 4x + 9} \, dx$.

**Solution:** Our first priority is to have the coefficient of $x^2$ to be 1, after which we complete the square and finish the problem.

\[
\int \frac{1}{5x^2 - 4x + 9} \, dx = \int \frac{1}{5 \left[ x^2 - \frac{4}{5}x + \frac{9}{5} \right]} \, dx = \int \frac{1}{5 \left[ x^2 - \frac{4}{5}x + \left( \frac{2}{5} \right)^2 - \left( \frac{2}{5} \right)^2 + \frac{9}{5} \right]} \, dx
\]

\[
= \int \frac{1}{5 \left( \left( x - \frac{2}{5} \right)^2 - \frac{4}{25} + \frac{9}{25} \right)} \, dx = \frac{1}{5} \int \frac{1}{\left( \left( x - \frac{2}{5} \right)^2 + \frac{41}{25} \right)} \, dx
\]

(Note how we added and subtracted $(2/5)^2$ together, both within the brackets.) Now we need to manipulate this integral so there is a 1 in place of the fraction $\frac{41}{25}$, which we do as before, by factoring. Continuing,

\[
\int \frac{1}{5x^2 - 4x + 9} \, dx = \cdots
\]

\[
= \frac{1}{5} \int \left[ \left( x - \frac{2}{5} \right)^2 + \frac{41}{25} \right] \, dx
\]

\[
= \frac{1}{5} \int \frac{1}{\frac{41}{25} \left( x - \frac{2}{5} \right)^2 + 1} \, dx
\]

\[
u = \frac{5}{\sqrt{41}} \left( x - \frac{2}{5} \right)
\]

\[
\Rightarrow \quad du = \frac{5}{\sqrt{41}} \, dx
\]

\[
\Leftrightarrow \quad \frac{\sqrt{41}}{5} \, du = dx
\]

\[
= \frac{1}{\sqrt{41}} \tan^{-1} u + C = \frac{1}{\sqrt{41}} \tan^{-1} \left[ \frac{5}{\sqrt{41}} \left( x - \frac{2}{5} \right) \right] + C.
\]
6.8 Hyperbolic Functions

The algebraic and differential structure embedded in the trigonometric functions made for some surprising, but useful derivative and integral formulas involving the arctrigonometric functions. As it happens, there is another genre of functions with a similar, yet distinct structure, that genre being the so-called hyperbolic functions.

In fact the hyperbolic functions are somewhat redundant, in that they are based upon exponential functions and the more interesting integration formulas can be arrived at through other methods. However, exploiting these functions can greatly simplify certain types of integration problems, as we will see.

We will begin with definitions, derivatives and some identities involving the hyperbolic functions. We will then look at their graphs, and consider what their “inverses” would look like, and how they play out in derivative and integral formulas. Along the way we will compare them to their trigonometric counterparts and see how a sign ($\pm$) here or there can make a crucial difference in an integration problem.

6.8.1 Hyperbolic Functions and Their Basic Intrinsic Structures

We begin with the hyperbolic sine and hyperbolic cosine functions. These can be defined geometrically, but unlike their trigonometric counterparts, these also have straightforward definitions in terms of our earlier, familiar functions (though it is not obvious from the geometry!):

\[
\sinh x = \frac{1}{2} (e^x - e^{-x}) = \frac{e^x - e^{-x}}{2}, \quad (6.96)
\]

\[
\cosh x = \frac{1}{2} (e^x + e^{-x}) = \frac{e^x + e^{-x}}{2}. \quad (6.97)
\]

The others are defined in terms of these, just as with the trigonometric functions. Before we define the others, we will notice a couple of relationships which are similar, though distinct, from what occurs with the trigonometric functions. The first result is algebraic:

**Theorem 6.8.1** For all $x \in \mathbb{R}$, we have

\[
cosh^2 x - \sinh^2 x = 1. \quad (6.98)
\]

For the proof, we just expand the left-hand side:

\[
cosh^2 x - \sinh^2 x = \left(\frac{e^x + e^{-x}}{2}\right)^2 - \left(\frac{e^x - e^{-x}}{2}\right)^2
\]

\[
= \frac{e^{2x} + 2e^x e^{-x} + e^{-2x}}{4} - \frac{e^{2x} - 2e^x e^{-x} + e^{-2x}}{4}
\]

\[
= \frac{e^{2x} + 2e^0 + e^{-2x} - e^{2x} + 2e^0 - e^{-2x}}{4}
\]

\[
= \frac{2 + 2}{4}
\]

\[
= 1, \text{ q.e.d.}
\]

Of course this is the hyperbolic analog of the basic trigonometric identity $\cos^2 \theta + \sin^2 \theta = 1$. As we will see, the difference in the signs between the trigonometric identity and (6.98) makes for analogous, but significantly distinct results throughout the hyperbolic development. We next look at the derivatives of these:
Theorem 6.8.2 For $x \in \mathbb{R}$, we have
\[
\frac{d}{dx} \sinh x = \cosh x, \quad (6.99)
\]
\[
\frac{d}{dx} \cosh x = \sinh x. \quad (6.100)
\]
These are simple chain rule computations. For the first case, we have
\[
\frac{d}{dx} \left[ \frac{1}{2} \left( e^x - e^{-x} \right) \right] = \frac{1}{2} \left( \frac{d}{dx} \left[ e^x - e^{-x} \right] \right) = \frac{1}{2} \left( e^x - e^{-x} (-1) \right) = \frac{1}{2} (e^x + e^{-x})
\]
i.e., $\frac{d}{dx} \sinh x = \cosh x$. That $\frac{d}{dx} \cosh x = \sinh x$ is similar. Note how these compare to derivative formulas for $\sin x$ and $\cos x$. 

Chapter 7

Advanced Integration Techniques

Before introducing the more advanced techniques, we will look at a shortcut for the easier of the substitution-type integrals. Advanced integration techniques then follow: integration by parts, trigonometric integrals, trigonometric substitution, and partial fraction decompositions.

7.1 Substitution-Type Integration by Inspection

In this section we will consider integrals which we would have done earlier by substitution, but which are simple enough that we can guess the approximate form of the antiderivatives, and then insert any factors needed to correct for discrepancies detected by (mentally) computing the derivative of the approximate form and comparing it to the original integrand. Some general forms will be mentioned as formulas, but the idea is to be able to compute many such integrals without resorting to writing the usual $u$-substitution steps.

Example 7.1.1 Compute $\int \cos 5x \, dx$.

**Solution:** We can anticipate that the approximate form$^1$ of the answer is $\sin 5x$, but then

$$\frac{d}{dx} \sin 5x = \cos 5x \cdot \frac{d}{dx} (5x) = \cos 5x \cdot 5 = 5 \cos 5x.$$  

Since we are looking for a function whose derivative is $\cos 5x$, and we found one whose derivative is $5 \cos 5x$, we see that our candidate antiderivative $\sin 5x$ gives a derivative with an extra factor of 5, compared with the desired outcome. Our candidate antiderivative’s derivative is 5 times too large, so this candidate antiderivative, $\sin 5x$ must be 5 times too large. To compensate and arrive at a function with the proper derivative, we multiply our candidate $\sin 5x$ by $\frac{1}{5}$. This gives us a new candidate antiderivative $\frac{1}{5} \sin 5x$, whose derivative is of course $\frac{1}{5} \cos 5x \cdot 5 = \cos 5x$, as desired. Thus we have

$$\int \cos 5x \, dx = \frac{1}{5} \sin 5x + C.$$  

It may seem that we wrote more in the example above than with the usual $u$-substitution method, but what we wrote could be performed mentally without resorting to writing the details.

In future sections, an integral such as the above may occur as a relatively small step in the execution of a more advanced and more complicated method (perhaps for computing a much more difficult integral). This section’s purpose is to point out how such an integral can be quickly dispatched, to avoid it becoming a needless distraction in the more advanced methods.

$^1$In this section, by approximate form we mean a form which is correct except for multiplicative constants.
7.1. **SUBSTITUTION-TYPE INTEGRATION BY INSPECTION**

### 7.1.1 The Method

The method used in all the examples here can be summarized as follows:

1. Anticipate the form of the antiderivative by an *approximate form* (correct up to a multiplicative constant).
2. Differentiate this approximate form and compare to the original integrand function;
3. If Step 1 is correct, i.e., the approximate form’s derivative differs from the original integrand function by a multiplicative constant, insert a compensating, reciprocal multiplicative constant into the approximate form to arrive at the actual antiderivative;
4. For verification, differentiate the answer to see if the original integrand function emerges.

For instance, some general formulas which should be quickly verifiable by inspection (that is, by reading and mental computation rather than with paper and pencil, for instance) follow:

\[
\int e^{kx} \, dx = \frac{1}{k} e^{kx} + C, \tag{7.1}
\]
\[
\int \cos kx \, dx = \frac{1}{k} \sin kx + C, \tag{7.2}
\]
\[
\int \sin kx \, dx = -\frac{1}{k} \cos kx + C, \tag{7.3}
\]
\[
\int \sec^2 kx \, dx = \frac{1}{k} \tan kx + C, \tag{7.4}
\]
\[
\int \csc^2 kx \, dx = -\frac{1}{k} \cot kx + C, \tag{7.5}
\]
\[
\int \sec kx \tan kx \, dx = \frac{1}{k} \sec kx + C, \tag{7.6}
\]
\[
\int \csc kx \cot kx \, dx = \frac{1}{k} \csc kx + C, \tag{7.7}
\]
\[
\int \frac{1}{ax + b} \, dx = \frac{1}{a} \ln |ax + b| + C. \tag{7.8}
\]

**Example 7.1.2** The following integrals can be computed with *u*-substitution, but also are computable by inspection:

- \[ \int e^{7x} \, dx = \frac{1}{7} e^{7x} + C, \]
- \[ \int \frac{1}{5x - 9} \, dx = \frac{1}{5} \ln |5x - 9| + C, \]
- \[ \int \sin 5x \, dx = -\frac{1}{5} \cos 5x + C; \]
- \[ \int \cos \frac{x}{2} \, dx = 2 \sin \frac{x}{2} + C, \]
- \[ \int \sec^2 \pi x \, dx = \frac{1}{\pi} \tan \pi x + C, \]
- \[ \int \sec 6x \cot 6x \, dx = -\frac{1}{6} \csc 6x + C. \]

While it is true that we can call upon the formulas (7.1)–(7.8), the more flexible strategy is to anticipate the form of the antiderivative and adjust accordingly. For instance, we have the following antiderivative form, written two ways:

\[
\int \frac{1}{u} \, du = \ln |u| + C,
\]
\[
\int \frac{f'(x)}{f(x)} \, dx = \ln |f(x)| + C.
\]
(As usual, the second form is the same as the first where \( u = f(x) \).) So when we see an integrand which is a fraction with the numerator being the derivative of the denominator except for multiplicative constants, we know the antiderivative will be, approximately, the natural log of the absolute value of that denominator.

**Example 7.1.3** Consider \( \int \frac{x}{x^2 + 1} \, dx \)

Here we see that the derivative of the denominator of the integrand is present—up to a multiplicative constant—in the numerator. Our candidate approximate form can then be given by \( \ln |x^2 + 1| = \ln(x^2 + 1) \). Now we differentiate to see what constant factor we need to insert to get the correct derivative:

\[
\frac{d}{dx} \ln(x^2 + 1) = \frac{1}{x^2 + 1} \cdot 2x = \frac{2x}{x^2 + 1}.
\]

To correct for the extra factor of 2 and thus get the correct derivative, we insert the factor \( \frac{1}{2} \):

\[
\frac{d}{dx} \left[ \frac{1}{2} \ln(x^2 + 1) \right] = \frac{1}{2} \cdot \frac{1}{x^2 + 1} \cdot 2x = \frac{x}{x^2 + 1},
\]

as desired. Thus

\[
\int \frac{x}{x^2 + 1} \, dx = \frac{1}{2} \ln(x^2 + 1) + C.
\]

To be sure, a quick mental check by differentiation verifies the answer.

Of course there are many other forms. Recall we had many other integration formulas, as in Subsection 6.6.1, page 577. For instance, it is not difficult to see, or check, that

\[
\int \frac{1}{u^2 + 1} \, du = \tan^{-1} u + C \quad \Rightarrow \quad \int \frac{1}{a^2x^2 + 1} \, dx = \frac{1}{a} \tan^{-1}(ax) + C.
\]

**Example 7.1.4** For instance, we have the following integral computations, which can be seen by relatively easily taking derivatives of the presented antiderivatives.

- \( \int \frac{1}{9x^2 + 1} \, dx = \frac{1}{3} \tan^{-1} 3x + C \),
- \( \int \sec 3x \, dx = \frac{1}{3} \ln |\sec 3x + \tan 3x| + C \),
- \( \int \frac{1}{\sqrt{1 - \sqrt{x^2}} \, dx = \frac{1}{\sqrt{3}} \sin^{-1} \left( \sqrt{\frac{x}{3}} \right) + C \),
- \( \int \tan 2x \, dx = \frac{1}{2} \ln |\sec 2x| + C \).

The method can be used for somewhat more complicated integrals as well, though there does come a point where it seems more natural to simply execute the full substitution method, which is more “constructive” than our method here. However, our “approximate and correct” (read as verbs) method here can be reasonably employed on still more complicated integrals.

**Example 7.1.5** Consider \( \int \frac{1}{\sqrt{5x - 9}} \, dx \).

Of course this can be rewritten \( \int (5x - 9)^{-1/2} \, dx \). Now it is crucial that a complete substitution, \( u = 5x - 9 \Rightarrow du = 5 \, dx \), etc., would show that \( du \) and \( dx \) agree except for a multiplicative constant, so we know that the integral—up to a multiplicative constant—is of approximate form \( \int u^{-1/2} \, du \), which calls for the power rule.

The approximate form of the antiderivative is thus \( u^{1/2} = (5x - 9)^{1/2} \), which we write in \( x \) and then differentiate:

\[
\frac{d}{dx} (5x - 9)^{1/2} = \frac{1}{2} (5x - 9)^{-1/2} \cdot 5,
\]
which has extra factors (compared to our original integrand) of collectively \( \frac{2}{5} \). To cancel their effects we include a factor \( \frac{2}{5} \) in our actual, reported antiderivative. Thus

\[
\int \frac{1}{\sqrt{5x} - 9} \, dx = \frac{2}{5} (5x - 9)^{1/2} + C = \frac{2}{5}\sqrt{5x - 9} + C.
\]

Note that a quick derivative computation, albeit involving a (simple) chain rule, gives us the correct function \( \frac{1}{\sqrt{5x} - 9} \).

**Example 7.1.6** Consider

\[
\int 7x \sin^5 x^2 \cos x^2 \, dx.
\]

For such an antiderivative, our ability to guess the form depends upon our expertise with the original substitution method. Each of these were of a form \( \int f(u) \, K \, du \), where we could anticipate both \( u \) and \( f \), with \( du \) accounting for remaining terms, and \( K \in \mathbb{R} \) which we can initially ignore by taking our shortcut path described in this section. Looking ahead, the student well-versed in substitution will expect \( u = \sin x^2 \), and the integral being of the approximate form \( \int u^5 \, du \) (times a constant). Thus we will have an approximate antiderivative of \( u^6 \) (times a constant), i.e., the approximate form should be \( \sin^6 x^2 \). Now we differentiate this and see what compensating factors must be included to reconcile with the original integrand:

\[
\frac{d}{dx} (\sin x^2)^6 = 6(\sin x^2)^5 \cdot \cos x^2 \cdot 2x = 12x \sin^5 x^2 \cos x^2.
\]

Of course we want 7 in the place of the 12 (or separately, 2 \cdot 6), so we multiply by \( \frac{7}{12} \) (or again, \( 7 \cdot \frac{1}{6} \)). With this we have

\[
\int 7x \sin^5 x^2 \cos x^2 \, dx = \frac{7}{12} \sin^6 x^2 + C.
\]

It would be perfectly natural to forego this method of “guess and adjust” in favor of the old-fashioned substitution method for this problem. Indeed the full substitution method has some advantages (see the next subsection). For instance, it is more “constructive,” and thus less error-prone; one is less tempted to skip steps while employing substitution, while one might attempt a purely mental derivative computation of the answer here and thus easily be off by a factor. It is important that each student find the comfortable level of brevity for himself or herself.

But the the method of this section is still often worthwhile.

**Example 7.1.7** Compute \( \int x^3 \sin x^4 \, dx \).

_Solution:_ This is of the approximate form \( \int \sin u \, du \), with \( u = x^4 \). The approximate form of the solution is thus \( \cos x^4 + C \) (or \( -\cos x^4 + C \), but these differ by a multiplicative constant \(-1\)), which has derivative \( -\sin x^4 \cdot 4x^3 \). We introduce a factor of \( -\frac{1}{4} \) to compensate for the extra factor of \(-1\):

\[
\int x^3 \sin x^4 \, dx = -\frac{1}{4} \cos x^4 + C,
\]

which can be quickly verified by differentiation.

---

2Notice that we are assuming fluency in the chain rule as we compute the derivative of \( \sin^6 x^2 \), rather than writing out every step as we did in Chapter 4. Each student must gage personal ability to omit steps.

3It is the author’s experience that students in engineering and physics programs are often more interested in arriving at the answer quickly, while mathematics and other science students usually prefer the presentation of the full substitution method. The latter are somewhat less likely to be wrong by a multiplicative constant, though the former tend to progress through the topics faster. There are, of course, spectacular exceptions, and each group benefits from camaraderie with the other.
Example 7.1.8 Compute $\int x\sqrt{9-x^2} \, dx$.

**Solution:** It is advantageous to read this integral as $\int \left(9-x^2\right)^{1/2} \, dx$, which is of approximate form $\int u^{1/2} \, du$ (where $u = 9 - x^2$). These observations, and the approximate form $(9-x^2)^{3/2}$ of the integral, can be gotten by this mental observation we are developing in this section. The approximate antiderivative’s derivative is $\frac{3}{2}(9-x^2)^{1/2} \cdot (-2x)$, which has an extra factor of $-3$ (after cancellation). Thus

$$\int x\sqrt{9-x^2} \, dx = -\frac{1}{3} \left(9-x^2\right)^{3/2} + C.$$

### 7.1.2 Limitations of the Method

There are two very important points to be made about the limitations of the method. The first point is illustrated by an example, and the second by making several related points.

1. **It is imperative that the derivative of the approximate form differs from the original function to be integrated by at most a multiplicative constant.**

   In particular, an extra variable function cannot be compensated for. To illustrate this point, and simultaneously warn against a common mistake, consider

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

   The mistake to avoid here is to take erroneously the approximate solution to be $\ln(x^2 + 1)$, which we then notice has derivative

   $$\frac{d}{dx} \ln(x^2 + 1) = \frac{1}{x^2 + 1} \cdot 2x.$$ 

   Unfortunately we cannot compensate by dividing by the extra factor $2x$, because

   $$\frac{d}{dx} \left[\frac{\ln(x^2 + 1)}{2x}\right] = \frac{2x \cdot \frac{d\ln(x^2 + 1)}{dx} - \ln(x^2 + 1) \cdot \frac{d(2x)}{dx}}{(2x)^2} = \frac{2x \cdot \frac{2x}{x^2 + 1} - 2\ln(x^2 + 1)}{4x^2},$$

   which is guaranteed (by the presence of the non-cancelling logarithm in the result) to be something other than our original function $\frac{1}{x^2 + 1}$. The method does not work because multiplicative functions do not “go along for the ride” in derivative (or antiderivative) problems the way multiplicative constants do. Of course we knew from before that

   $$\int \frac{1}{x^2 + 1} \, dx = \tan^{-1} x + C,$$

   so this integral is not really suitable for a substitution argument, but is rather a special form in its own right.

---

4 Alternatively, a product rule computation can be used:

$$\frac{d}{dx} \left[\frac{1}{2x} \ln(x^2 + 1)\right] = \frac{1}{2x} \frac{d\ln(x^2 + 1)}{dx} + \ln(x^2 + 1) \cdot \frac{d}{dx} \left[\frac{1}{2x}\right] = \frac{1}{2x} \cdot \frac{2x}{x^2 + 1} - \frac{1}{2x^2} \ln(x^2 + 1),$$

which eventually gives the original function for the first product, but the second part of the product rule is a complication we cannot rid ourselves of easily, though a partial solution to this problem of extra, function-type factors in the integrand is given in the next section.
(II) This method can not totally replace the earlier substitution method.

(a) The skills used in the substitution method will be needed for later methods. In particular, the idea which is crucial and recurring is that the entire integral in \( x \) is often replaced by one in \( u \) (for instance)—including the \( dx \)-term and possibly others replaced by \( du \)—and, if a definite integral, the interval of integration also represents \( u \)-values.

(b) If an integral is difficult enough, the more “constructive” substitution method is less error-prone than is this shortcut style developed here.

(c) Anyhow, the idea of the substitution method is embedded in this method; anticipating what to set equal to \( u \) is equivalent to guessing the approximate form of the integral in \( u \), and thus the approximate form of the antiderivative.

(d) When using numerical and other methods with definite integrals, a substitution can sometimes make for a much simpler integral to be approximated or otherwise analyzed, even if the antiderivative is never computed. For instance, with \( u = x^2 \), giving then \( du = 2x \, dx \), we can write

\[
\int_{-1}^{2} xe^{x^4} \, dx = \frac{1}{2} \int_{1}^{4} e^{u^2} \, du.
\]

None of our usual techniques will yield antiderivatives for either integrand (as the reader is invited to try), but numerical methods such Riemann sums, Trapezoidal and Simpson’s Rules can find approximations for the definite integrals. The latter form of the integral (in \( u \)) will yield accurate numerical results more easily than the former (in \( x \)).

To summarize, the method here has us making an educated guess about the form of the antiderivative, perhaps writing down our guess as our tentative (or “candidate”) answer, taking its derivative, and, assuming it is the same as the integrand except for some multiplicative constant(s), inserting other multiplicative constants into our answer to adjust for discrepancies. It will only work if the tentative antiderivative has derivative equal to some constant times the original integrand.

The method is not sophisticated, but will be useful for streamlining later, much longer integration techniques introduced in the rest of this chapter.
Exercises

For each of the following, first attempt to compute the antiderivative by finding an approximate form of the antiderivative, differentiating it, and inserting a constant factor to compensate for any extra or missing constants. If that method is too unwieldy, compute the integral by the substitution method, showing all details.

1. \[ \int (x^2 + 1)^7 \cdot 2x \, dx \]
2. \[ \int \cos x^4 \cdot 4x^3 \, dx \]
3. \[ \int 15x^2 \sec^2 5x^3 \, dx \]
4. \[ \int \sec \sqrt{x} \tan \sqrt{x} \cdot 2x \, dx \]
5. \[ \int \csc^2 \left( \frac{1}{x^2} \right) \cdot x^2 \, dx \]
6. \[ \int \tan^7 x \sec^2 x \, dx \]
7. \[ \int \frac{x}{(x^2 + 1)^3} \, dx \]
8. \[ \int (2x - 11)^9 \, dx \]
9. \[ \int \cos 5x \, dx \]
10. \[ \int \csc 9x \cot 9x \, dx \]
11. \[ \int \cos x \sin x \, dx \] (See #13)
12. \[ \int \tan^3 5x \sec^2 5x \, dx \]
13. \[ \int \sin x \cos x \, dx \] (See #11)
14. \[ \int \sin^3 x \cos x \, dx \]
15. \[ \int \tan^9 x \sec^2 x \, dx \]
16. \[ \int x \sin x^2 \, dx \]
17. \[ \int x^3 \cdot \sqrt{x^4 - 2} \, dx \]
18. \[ \int (x^3 + x^2)^4 (3x^2 + 2x) \, dx \]
19. \[ \int \sec^5 3x \cdot \sec 3x \tan 3x \, dx \]
20. \[ \int \frac{\sin \sqrt{x}}{\sqrt{x}} \, dx \]
21. \[ \int x^2 \sin x^3 \cos^5 x^3 \, dx \]
22. \[ \int \frac{\sin^3 \left( \frac{1}{x^2} \right) \cos \left( \frac{1}{x^2} \right)}{x^2} \, dx \]
23. \[ \int e^x \cos e^x \, dx \]
24. \[ \int xe^{x^2} \, dx \]
25. \[ \int e^{2x} \sin e^{2x} \, dx \]
26. \[ \int e^{-x} \sec^2 e^{-x} \, dx \]
27. \[ \int e^{5x} \, dx \]
28. \[ \int \frac{e^x}{e^{2x} + 1} \, dx \]
29. \[ \int \frac{e^{3x}}{\sqrt{1 - e^{6x}}} \, dx \]
30. \[ \int \frac{dx}{\sqrt{e^{2x} - 1}} \] (Hint: multiply the integrand by \( e^x / e^x \).)
31. \[ \int e^{4x} (9 + e^{4x})^{10} \, dx \]
32. \[ \int xe^{-2x^2} \, dx \]
33. \[ \int \frac{e^{1/x}}{x^2} \, dx \]
7.1. SUBSTITUTION-TYPE INTEGRATION BY INSPECTION

34. \( \int \frac{e^{\sqrt{x}}}{\sqrt{x}} \, dx \)

35. \( \int e^{3\cos 2x} \sin 2x \, dx \)

36. \( \int \frac{\cos x}{\sin x + 1} \, dx \)

37. \( \int \frac{\cos x}{\sin x} \, dx \)

38. \( \int \frac{\sin x}{\cos x} \, dx \)

39. \( \int \frac{2x + 1}{x^2 + x} \, dx \)

40. \( \int \frac{x}{x^2 + 1} \, dx \)

41. \( \int \frac{1}{x^2 + 1} \, dx \)

42. \( \int \frac{1}{x \ln x} \, dx \)

43. \( \int \frac{e^{2x}}{1 + e^{2x}} \, dx \)

44. \( \int \frac{e^{2x}}{1 + e^{4x}} \, dx \)

45. \( \int \frac{\sec^2 x}{1 + \tan x} \, dx \)

46. \( \int \frac{\sin(\ln x)}{x} \, dx \)

47. \( \int \frac{\ln x}{x} \, dx \)

48. \( \int \frac{1}{x \sqrt{1 - (\ln x)^2}} \, dx \)

49. \( \int \frac{1}{x(1 + (\ln x)^2)} \, dx \)

50. \( \int \frac{1}{x |\ln x| \sqrt{(|\ln x|^2 - 1)}} \, dx \)

51. \( \int \frac{\sec^2(\ln x)}{x} \, dx \)

52. \( \int \frac{(9 + \ln x)^6}{x} \, dx \)

53. \( \int \frac{1}{x(\ln x)^2} \, dx \)

54. \( \int \cot 2x \, dx \)

55. \( \int \frac{\tan(\ln x)}{x} \, dx \)

56. \( \int \csc \frac{x}{9} \, dx \)

57. \( \int x^2 \sec 5x^3 \, dx \)
7.2 Integration By Parts

While integration by substitution in its elementary form takes advantage of the chain rule, by contrast integration by parts exploits the product rule. In applications it is a bit more complicated than substitution, and there are perhaps more variations on the theme than with substitution, at least at the college calculus level. For these reasons, being fluent in this method usually requires seeing more steps ahead than substitution required. However it can be similarly mastered with practice.

7.2.1 The Idea by Examples

Suppose that we need to find an antiderivative of the function \( f(x) = x \sec^2 x \). It is not hard to see that normal substitution in not going to easily yield a formula for our desired antiderivative \( F(x) \):

\[
\int x \sec^2 x \, dx = F(x) + C.
\]

However, a clever student might notice that \( x \sec^2 x \) contains terms that could have arisen from a product rule derivative such as:

\[
\frac{d}{dx} [x \tan x] = x \cdot \frac{d}{dx} \tan x + \tan x \cdot \frac{dx}{dx} = x \sec^2 x + \tan x.
\]

If we rearrange the terms above, we can rewrite this product rule derivative computation as follows:

\[
x \sec^2 x = \frac{d}{dx} [x \tan x] - \tan x. \tag{7.9}
\]

In fact (7.9) above is perhaps where the spirit of the method is most on display: that the given function to be integrated is indeed one part of a product rule derivative. If we are fortunate, the other part of the product rule formula is easier to integrate, because the derivative term, namely \( \frac{d}{dx} [x \tan x] \) is trivial to integrate (see below). Indeed, if we take antiderivatives of both sides of (7.9), we then get the desired formula for the antiderivative \( \int x \sec^2 x \, dx \):

\[
\int x \sec^2 x \, dx = \int \left[ \left( \frac{d}{dx} [x \tan x] \right) - \tan x \right] \, dx
\]

\[
= x \tan x - \int \tan x \, dx
\]

\[
= x \tan x - \ln |\sec x| + C.
\]

From such as the above emerges a method whereby we identify our given function (here \( x \sec^2 x \)) as a part of a product rule computation (\( \frac{d}{dx} [x \tan x] \)), and integrate our original function by instead (trivially) integrating the product rule derivative term (again \( \frac{d}{dx} [x \tan x] \)), and then integrating the other part \( (\tan x) \) of the product rule output. Often the other, hidden part of the underlying product rule is easier to integrate than the original function, and therein lies much of the usefulness of the method.

While we will have a formal procedure to implement the method, one more example from first principles can further illustrate its spirit.
Example 7.2.1 Compute $\int x \cos x \, dx$.

Solution: The integrand is one part of a product rule computation for $\frac{d}{dx}[x \sin x]$, namely

$$\int x \cos x \, dx = \int \left[ \frac{d}{dx}[x \sin x] \right] \, dx = x \sin x + \cos x + C.$$

It is interesting to compute the derivative of our answer, and see how the term we desire $(x \cos x)$ emerges, and how other terms which naturally emerge cancel each other. This is left to the reader.

7.2.2 The Technique in Its Simpler Applications

Recall that when we completely developed the substitution method, the underlying principle—the chain rule—was not written out in complete derivative form, but rather in the more compact differential form. Having supposed that $F$ was an antiderivative of $f$, i.e., $F' = f$, we eventually settled on writing the argument below without the first two integrals:

$$\int f(u(x))u'(x) \, dx = \int f(u(x)) \cdot \frac{du(x)}{dx} \, dx = \int f(u) \, du = F(u) + C = F(u(x)) + C.$$  

At first we did write the first steps because the proof was in the chain rule: $\frac{d}{dx}F(u(x)) = F'(u(x))u'(x) = f(u(x))u'(x)$. However, we eventually opted for the differential form, though for most it takes some practice for it to seem natural.

We will adopt differential notation in integration by parts as well. For instance, recall that the product rule for derivatives,

$$\frac{d}{dx}(uv) = u \cdot \frac{dv}{dx} + v \cdot \frac{du}{dx}, \quad (7.10)$$

can be rewritten in differential form

$$d(uv) = u \, dv + v \, du. \quad (7.11)$$

This came from multiplying both sides of (7.10) by $dx$, all along assuming $u$ and $v$ are in fact functions of $x$. Now we rearrange (7.11) as follows:

$$u \, dv = d(uv) - v \, du. \quad (7.12)$$

Equation (7.12) is perhaps the best equation to visualize the principle behind the eventual integration formula, because it is an easy step from the product rule. The actual formula quoted in most textbooks is still two steps away. First we integrate both sides:

$$\int u \, dv = \int d(uv) - \int v \, du. \quad (7.13)$$

Next we notice that the first integral on the right hand side of (7.13) is simply $\int d(uv) = uv + C$. Since there is also a constant present in the second integral on the right hand side of (7.13), we can omit mentioning the first constant and arrive at our final working formula for our integration by parts technique:

$$\int u \, dv = uv - \int v \, du. \quad (7.14)$$
Most textbooks and instructors use the formula above in exactly this form (7.14). It should be memorized, though its derivation—particularly from (7.12)—should also not be forgotten. Furthermore, anytime it is used it is good practice to write the (boxed) formula (7.14) within the problem at the point at which it is used.

Next we look at an example of the actual application of (7.14). In the example below, the arrangement of terms is as one would work the problem with pencil and paper, except for the implication arrows which we will omit in subsequent problems, and possibly the under-braces which become less and less necessary with practice.

**Example 7.2.2** Compute \( \int x e^x \, dx \).

**Solution:** First recall our boxed formula (7.14). For this problem, we write

\[
\begin{align*}
  u &= x & dv &= e^x \, dx \\
  du &= dx & v &= e^x \\
\end{align*}
\]

\[
\int x \, e^x \, dx = uv - \int v \, du \\
= (x)(e^x) - \int (e^x) \, dx \\
= xe^x - e^x + C.
\]

It is interesting to note that we choose \( u \) and \( dv \), and then compute \( du \) and \( v \), with one qualification. That is that \( v \) is not unique; the computation from \( dv \) to \( v \) is of an antidifferentiation nature and so we really only know \( v \) up to an additive constant. In fact any \( v \) so that \( dv = e^x \, dx \) (we took \( v = e^x \implies dv = e^x \, dx \)) will work in (7.14). Any additive constant, while legitimate, will eventually cancel in the final computation, and so we usually omit it. For instance, if we had chosen \( v = e^x + 100 \), we would have had

\[
\begin{align*}
  uv - \int v \, du &= x(e^x + 100) - \int (e^x + 100) \, dx \\
  &= xe^x + 100x - e^x - 100x + C \\
  &= xe^x - e^x + C,
\end{align*}
\]

with the same final answer as before. For most cases, we will just assume that the additive constant is zero and we will use the simplest antiderivative for \( v \). We will also not continue to write the implication arrows as they are technical and perhaps confusing.

Now we will revisit the first example we gave in Subsection 7.2.1, using what will be our basic style for this method.

**Example 7.2.3** Compute \( \int x \sec^2 x \, dx \).

**Solution:** First, as is standard for these, we complete a chart like in the previous example.

\[
\begin{align*}
  u &= x & dv &= \sec^2 x \, dx \\
  du &= dx & v &= \tan x
\end{align*}
\]
\[
\int x \sec^2 x \, dx = \int u \, dv = uv - \int v \, du = x \tan x - \int \tan x \, dx = x \tan x - \ln |\sec x| + C.
\]

Since this method is more complicated than substitution, there are more complicated considerations in how to apply it. First of course, one should attempt an earlier, simpler method. But if those fail, and integration by parts is to be attempted, the following guidelines for choosing \(u\) and \(dv\) should be considered for our formula \(\int u \, dv = uv - \int v \, du\):

1. \(u\) and \(dv\) must account for all factors of the original integral, and no more.
   
   1.5. Of course, \(dv\) must contain the differential term (for example, \(dx\)) as a factor, but can contain more terms.

2. \(v = \int dv\) should be computable with relative ease.

3. \(du = u' (x) \, dx\) (assuming the original integral was in \(x\)) should not be overly complicated.

4. The integral \(\int v \, du\) should be simpler than the original integral \(\int u \, dv\).

The next example illustrates the importance of consideration 2 above.

**Example 7.2.4** Compute \(\int x^3 \sin x^2 \, dx\).

**Solution:** We do not want to make \(u = \sin x^2\), because then \(dv = x^3 \, dx\), giving \(du = 2x \cos x^2\) and \(v = \frac{1}{2} x^4\), and our \(\int v \, du\) will be \(\int \frac{1}{2} x^4 \cos x^2 \, dx\), which is worse than our original integral.

We will instead take \(u\) to be some power of \(x\), but not all of \(x^3\), else the terms remaining for \(dv\) would be \(dv = \sin x^2 \, dx\), which we cannot integrate with ordinary methods.

What we will settle on is \(dv = x \sin x^2 \, dx\), because its integral is an easy substitution we can short-cut as in Section 7.1. We leave the remaining terms, collectively \(x^2\), for \(u\):

\[
\begin{align*}
\ u & = x^2 \\
\ dv & = x \sin x^2 \, dx \\
\ du & = 2x \, dx \\
\ v & = -\frac{1}{2} \cos x^2
\end{align*}
\]

---

5Of course with practice one can see ahead whether or not integration by parts is likely to achieve an answer for a particular integral.

6Later, in a twist on the method, we will see that the we do not require \(\int v \, du\) be easier than the original, \(\int u \, dv\), in all cases, but it is desirable in most cases.

7In fact, we cannot compute \(\int \sin x^2 \, dx\) using any kind of substitution or parts, or any other method of this text for that matter, and arrive at an antiderivative in simple terms of the functions we know so far such as powers, exponentials, logarithms, trigonometric or hyperbolic functions or their inverses. However, when we study series we will find other expressions with which we can fashion an antiderivative of \(\sin x^2\).
\[
\int x^3 \sin x^2 \, dx = \int \frac{x^2 \cdot x \sin x^2 \, dx}{dv}
\]
\[
= uv - \int v \, du
\]
\[
= (x^2) \left(-\frac{1}{2} \cos x^2\right) - \int \left(-\frac{1}{2} \cos x^2\right) 2x \, dx
\]
\[
= -\frac{x^2}{2} \cos x^2 + \int \cos x^2 \cdot x \, dx
\]
\[
= -\frac{x^2}{2} \cos x^2 + \frac{1}{2} \sin x^2 + C.
\]

We omitted the details of computing \(v\) from \(dv\), and computing the last integral, in the spirit of Section 7.1.

This last example shows how the requirement that \(v = \int dv\) (up to an additive constant) be computable helps to guide us to the proper choice of \(u\) and \(dv\). It was lucky that the second integral was easily computable (which would not have been the case if the original integral were, say, \(\int x^2 \sin x^2 \, dx\) or \(\int x^4 \sin x^2 \, dx\)), but anyhow we can not even get to the second integral if we can not compute \(v\).

**Example 7.2.5** Compute \(\int \frac{x^9}{\sqrt{1 - x^5}} \, dx\).

**Solution:** This is similar to the previous example, in the sense that computability of \(v = \int dv\) dictates our choices of \(u\) and \(dv\).

\[
u = x^5 \quad dv = x^4 (1 - x^5)^{-1/2} \, dx
\]
\[
du = 5x^4 \, dx \quad v = \frac{2}{5} (1 - x^5)^{1/2}
\]

\[
\int \frac{x^9}{\sqrt{1 - x^5}} \, dx = \int uv - \int v \, du
\]
\[
= -\frac{2}{5} x^5 (1 - x^5)^{1/2} + 2 \int x^4 (1 - x^5)^{1/2} \, dx
\]
\[
= -\frac{2}{5} x^5 \sqrt{1 - x^5} + 2 \cdot \frac{-1}{5} \cdot \frac{2}{3} (1 - x^5)^{3/2} + C
\]
\[
= -\frac{2}{5} x^5 \sqrt{1 - x^5} - \frac{4}{15} (1 - x^5)^{3/2} + C
\]
\[
= -\frac{2}{15} \sqrt{1 - x^5} [3x^5 + 2(1 - x^5)] + C
\]
\[
= -\frac{2}{15} (x^5 + 2) \sqrt{1 - x^5} + C.
\]

As will often be the case in this section and subsequent sections, to check our answers it will usually be much easier to carefully check each step in our work of computing the integrals, than to compute the derivatives of our answers. (If we do wish to check answers by differentiation, it is often simpler to do so with a nonsimplified expression of the solution.)
7.2.3 Repeated Use of Integration by Parts

The next examples show a different lesson: that it is sometimes appropriate to integrate by parts more than once in a given problem. At each step, the hope is that the integral \( \int v \, du \) is simpler than the integral it came from (\( \int u \, dv \)) in the integration by parts formula. Sometimes the new integral \( \int v \, du \) is indeed simpler, but not so much so that it can be integrated readily. Indeed, at times that second integral also needs integration by parts, and so on, until we arrive at an integral that we can compute easily.

Example 7.2.6 Compute \( \int x^2 \cos 3x \, dx \).

Solution: The \( x^2 \) term is complicating our integral, and so we reduce its effect somewhat by an integration by parts step.

\[
\begin{align*}
    u &= x^2 & dv &= \cos 3x \, dx \\
    du &= 2x \, dx & v &= \frac{1}{3} \sin 3x \\
\end{align*}
\]

\[
\int x^2 \cos 3x \, dx = uv - \int v \, du = \frac{1}{3} x^2 \sin 3x - \int \frac{2}{3} x \sin 3x \, dx. \tag{7.15}
\]

While we still cannot compute this last integral directly with old methods, it is better than the original in the sense that our trigonometric function is multiplied by a first-degree polynomial, where in the original the polynomial was second-degree. One more application of integration by parts and there will be no polynomial factor at all in the new integral.

A strict use of the language would force us to introduce two new variables other than \( u \) and \( v \), but since they have “disappeared” in the present form of our answer, namely \( \frac{1}{3} x^2 \sin 3x - \frac{2}{3} \int x \sin 3x \, dx \), it is not considered such bad form to “reset” (or “recycle”) \( u \) and \( v \) for another integration by parts step, this time involving the integral \( \int x \sin 3x \, dx \):

\[
\begin{align*}
    u &= x & dv &= \sin 3x \, dx \\
    du &= dx & v &= -\frac{1}{3} \cos 3x \\
\end{align*}
\]

\[
\int x \sin 3x \, dx = uv - \int v \, du = -\frac{x}{3} \cos 3x + \frac{1}{3} \int \cos 3x \, dx = -\frac{x}{3} \cos 3x + \frac{1}{3} \cdot \frac{1}{3} \sin 3x + C_1.
\]

Now we insert this last result into our original computation (7.15):

\[
\int x^2 \cos 3x \, dx = \frac{x^2}{3} \sin 3x - \frac{2}{3} \left[ -\frac{x}{3} \cos 3x + \frac{1}{9} \sin 3x + C_1 \right] = \frac{x^2}{3} \sin 3x + \frac{2}{9} x \cos 3x - \frac{2}{27} \sin 3x + C.
\]
Several lessons can be gleaned from the example above. (1) It is very important for proper
"bookkeeping," as these problems can beget several "subproblems," and the proper placement
of resulting terms is crucial for getting the correct final answer; (2) It is not unknown to use
integration by parts more than once in a problem; (3) If we can take as many antiderivatives of
a function \( f(x) \) as we like (i.e., the antiderivative, the antiderivative of the antiderivative, etc.),
then for an integral \( \int x^n f(x) \, dx \) we can let \( u = x^n \), and integration by parts will yield a second
integral with a reduction in the power of \( x \), namely
\[
\int x^n f(x) \, dx = x^n F(x) - \int nx^{n-1} F(x) \, dx
\]  
where \( F' = f \). If this process is repeated often enough, the complicating polynomial factors will
have their degrees diminished until there is no polynomial factor, and we can finally finish the
integration.

7.2.4 Integrals of Certain Other Functions

Here we look at cases where functions appear whose derivatives are known, but whose antideriva-
tives might not be standard knowledge of the average calculus student. In these cases we choose
that function to be \( u \), and the other terms to constitute \( dv \), assuming of course we can compute
\( v = \int dv \).

**Example 7.2.7** Compute \( \int (x^2 + 1) \ln x \, dx \).

*Solution:* We cannot let \( dv = \ln x \, dx \), since as yet we do not know the antiderivative of \( \ln x \).
(Even if we did, such a choice for \( dv \) would not be advantageous, as Example 7.2.8 will help to
show later.) So we have little choice but to let \( \ln x \) be \( u \).

\[
\begin{align*}
  u &= \ln x \\
  dv &= (x^2 + 1) \, dx \\
  du &= \frac{1}{x} \, dx \\
  v &= \frac{x^3}{3} + x
\end{align*}
\]

\[
\int (x^2 + 1) \ln x \, dx = uv - \int v \, du
\]

\[
= (\ln x) \left( \frac{x^3}{3} + x \right) - \int \left( \frac{x^3}{3} + x \right) \frac{1}{x} \, dx
\]

\[
= \frac{1}{3} (\ln x)(x^3 + 3x) - \int \left( \frac{x^2}{3} + 1 \right) \, dx
\]

\[
= \frac{1}{3} (x^3 + 3x) \ln x - \frac{1}{9} x^3 - x + C.
\]

Next we have an interesting case where we are forced to take \( dv = dx \).

**Example 7.2.8** Compute \( \int \ln x \, dx \).

*Solution:* Here we can not let \( dv = \ln x \, dx \), for computing \( v = \int dv \) would be the same as
computing the whole, original integral. As in Example 7.2.7, we also note that placing \( \ln x \) in
the \( u \)-term makes for a simple enough \( du \) term. Thus we write

\[
\begin{align*}
  u &= \ln x \\
  dv &= dx \\
  du &= \frac{1}{x} \, dx \\
  v &= x
\end{align*}
\]
so that

\[
\int \ln x \, dx = uv - \int v \, du = (\ln x)(x) - \int (x) \left( \frac{1}{x} \, dx \right) = x \ln x - \int 1 \, dx = x \ln x - x + C.
\]

In fact the same type of computation will be used for finding antiderivatives of arctrigonometric functions (though arcsecant and arccosecant need later techniques for finding the resulting second integral \( \int v \, du \)).

**Example 7.2.9** Compute \( \int \sin^{-1} x \, dx \).

**Solution:** Again we have no choice but to let \( u = \sin^{-1} x \), and \( dv = dx \).

\[
u = \sin^{-1} x \quad dv = dx \quad du = \frac{1}{\sqrt{1 - x^2}} dx \quad v = x
\]

For brevity, we label the desired integral \( (I) \), so here \( (I) = \int \sin^{-1} x \, dx \). (The second integral below is computed “by inspection.”)

\[
(I) = uv - \int v \, du = x \sin^{-1} x - \int \frac{x}{\sqrt{1 - x^2}} \, dx = x \sin^{-1} x + (1 - x^2)^{1/2} + C.
\]

### 7.2.5 An Indirect Method

The following method, summarized at the end, is useful in surprisingly many settings.

**Example 7.2.10** Compute \( \int e^{2x} \cos 3x \, dx = (I) \).

**Solution:** Here we again name the desired integral \( (I) \) for brevity in later steps. It should be obvious (especially after a few attempts) that simple substitution methods will not work. So we attempt an integration by parts.

**Step 1.** We will let the trigonometric function be part of \( dv \):

\[
u = e^{2x} \quad dv = \cos 3x \, dx \quad du = 2e^{2x} \, dx \quad v = \frac{1}{3} \sin 3x
\]

So far, after some rearrangement and simplifying, we have

\[
(I) = uv - \int v \, du = \frac{1}{3} e^{2x} \sin 3x - \frac{2}{3} \int e^{2x} \sin 3x \, dx = (II) \quad (7.17)
\]

[II]
This does not seem any easier than the first integral, so perhaps we might continue, but this time let the trigonometric function be $u$ and the exponential (along with $dx$) be contained in $dv$.

**Step 2—First Attempt.** Compute $(\text{II}) = \int e^{2x} \sin 3x \, dx$ in light of the comments at the end of the first step.

\[
u = \sin 3x, \quad dv = e^{2x} \, dx, \quad du = 3 \cos 3x \, dx, \quad v = \frac{1}{2} e^{2x}.
\]

\[
(\text{II}) = uv - \int v \, du = \frac{1}{2} e^{2x} \sin 3x - \frac{3}{2} \int e^{2x} \cos 3x \, dx.
\]

Combining this with the conclusion (7.17) of Step 1 gives us:

\[
(I) = \frac{1}{3} e^{2x} \sin 3x - \frac{2}{3} \left[ \frac{1}{2} e^{2x} \sin 3x - \frac{3}{2} \int e^{2x} \cos 3x \, dx \right]

= \frac{1}{3} e^{2x} \sin 3x - \frac{2}{3} e^{2x} \sin 3x + \int e^{2x} \cos 3x \, dx

= \int e^{2x} \cos 3x \, dx.
\]

Unfortunately that puts us right back where we started. However, a minor change in our effort above will eventually lead us to the solution. While keeping Step 1, our next step towards a solution is to replace Step 2 by the same strategy as used in Step 1, namely that we use the exponential function for $u$ and the trigonometric function in the $dv$ term.

**Step 2—Second Attempt.** Again we attempt to compute $(\text{II}) = \int e^{2x} \sin 3x \, dx$, though with different choices of $u$ and $dv$.

\[
u = e^{2x}, \quad dv = \sin 3x \, dx, \quad du = 2e^{2x} \, dx, \quad v = -\frac{1}{3} \cos 3x
\]

\[
(\text{II}) = uv - \int v \, du

= -\frac{1}{3} e^{2x} \cos 3x + \frac{2}{3} \int e^{2x} \cos 3x \, dx. \quad (7.18)
\]

It may seem that (7.18) is also a dead end, since it contains the original integral. But this attempt is different. In fact, when we combine (7.18) with (7.17) we get

\[
(I) = \frac{1}{3} e^{2x} \sin 3x - \frac{2}{3} (\text{II})

= \frac{1}{3} e^{2x} \sin 3x - \frac{2}{3} \left[ -\frac{1}{3} e^{2x} \cos 3x + \frac{4}{9} \int e^{2x} \cos 3x \, dx \right]

= \frac{1}{3} e^{2x} \sin 3x + \frac{2}{9} e^{2x} \cos 3x - \frac{4}{9} \int e^{2x} \cos 3x \, dx,
\]

\[
(\text{I})
\]
which we can summarize by the following equation:

\[
(I) = \frac{1}{3} e^{2x} \sin 3x + \frac{2}{9} e^{2x} \cos 3x - \frac{4}{9} (I).
\]  

Now we are ready to derive \((I)\), not by another calculus computation, but in fact by simple algebra: we solve for it.

**Step 3.** Solve (7.19) for \((I)\). First we add \(\frac{4}{9} (I)\) to both sides of (7.19):

\[
\frac{13}{9} (I) = \frac{1}{3} e^{2x} \sin 3x + \frac{2}{9} e^{2x} \cos 3x + C_1.
\]  

Here we include \(C_1\) because in fact each \((I)\) in (7.19) represents all antiderivatives, which differ from each other by additive constants. Now (7.19) made sense because of the fact that there are (hidden) additive constants on both sides of that equation (though on the right side they are multiplied by \(-\frac{4}{9}\), but that still yields additive constants). Solving (7.20) for \((I)\) we now have

\[
(I) = \frac{9}{13} \left[ \frac{1}{3} e^{2x} \sin 3x + \frac{2}{9} e^{2x} \cos 3x + C_1 \right] = \frac{3}{13} e^{2x} \sin 3x + \frac{2}{13} e^{2x} \cos 3x + C,
\]  

where \(C = \frac{9}{13} C_1\).

Of course with this our original problem is solved:

\[
\int e^{2x} \cos 3x \, dx = \frac{3}{13} e^{2x} \sin 3x + \frac{2}{13} e^{2x} \cos 3x + C.
\]

What is important to understand about the example above is that sometimes, though we cannot perhaps directly compute a particular integral, it may happen that an indirect method gives us the answer. Here we found an equation, namely (7.19), which our desired integral satisfies, and for which \((I)\) could be solved algebraically. We must be open to the possibility—indeed, the opportunity—of finding a desired quantity by such indirect methods, as well as direct computations.

It should be pointed out that we could have computed the integral in Example 7.2.10 by instead letting \(u\) be the trigonometric function, and \(dv = e^{2x} \, dx\) in both Steps 1 and 2. In fact it is usually best to pick similar choices for \(u\) and \(dv\) when an integration by parts will take more than one step. (Recall the discussion for \(\int x^n f(x) \, dx\).)

The method of Example 7.2.10, namely solving for \((I)\) after an integration by parts step, is available perhaps more often than one would think, though it is not a method of first resort.

The next example is also one in which we will eventually “solve” for the integral algebraically.

**Example 7.2.11** Compute \(\int \sin^2 x \, dx = (I)\).

\[^{8}\text{In fact, two simultaneous appearances of } (I) \text{ do not have to have the same additive constants, so } (I) - (I) = C_2, \text{ not zero.}\]
Solution: The only reasonable choice here seems to be to let \( u = \sin x \) and \( dv = \sin x \, dx \), if we are to integrate this by parts.\(^9\)

\[
\begin{align*}
  u &= \sin x & dv &= \sin x \, dx \\
  du &= \cos x \, dx & v &= -\cos x
\end{align*}
\]

\[
(\mathcal{I}) = \int uv - \int v \, du = -\sin x \cos x + \int \cos^2 x \, dx.
\]

We could perform the same integration by parts with the second integral, which might or might not yield an equation we can solve for \((\mathcal{I})\) (as the reader is invited to explore), but instead we will use the fact that \(\cos^2 x = 1 - \sin^2 x\):

\[
(\mathcal{I}) = -\sin x \cos x + \int (1 - \sin^2 x) \, dx \\
= -\sin x \cos x + x - \int \sin^2 x \, dx \\
= x - \sin x \cos x - (\mathcal{I}).
\]

Adding \((\mathcal{I})\) to both sides we get\(^10\)

\[
2(\mathcal{I}) = -\sin x \cos x + x + C_1
\]

\[
\implies (\mathcal{I}) = \frac{1}{2} (x - \sin x \cos x) + C.
\]

7.2.6 Miscellaneous Considerations

First we look at a definite integral arising from integration by parts. It should be pointed out that the general formula will look like the following:\(^11\)

\[
\int_{x=a}^{x=b} u \, dv = uv \bigg|_{x=a}^{x=b} - \int_{x=a}^{x=b} v \, du.
\]  

\(^9\)The integral in Example 7.2.11 can also be computed directly if we first use the trigonometric identity \(\sin^2 x = \frac{1}{2} (1 - \cos 2x)\), and then the identity \(\sin 2x = 2 \sin x \cos x\):

\[
\int \sin^2 x \, dx = \int \frac{1}{2} (1 - \cos 2x) \, dx = \frac{1}{2} x - \frac{1}{4} \sin 2x + C \\
= \frac{1}{2} x - \frac{1}{4} \cdot 2 \sin x \cos x + C = \frac{1}{2} x - \frac{1}{2} \sin x \cos x + C,
\]

which is the same as the answer in the text of Example 7.2.11. In Section 7.3 we will opt for this alternative method, and indeed will make quite an effort to exploit the algebraic properties of the trigonometric functions wherever possible, but some integrals there will still require integration by parts.

\(^10\)In fact many textbooks do not bother writing the \(C_1\) term, preferring to remind the student at the end that an indefinite integral problem necessitates a “+C.”

\(^11\)Some texts leave out the “\(x=\)” parts, assuming they are understood, but we will continue to use the convention that, unless otherwise stated, the “limits of integration” should correspond to values of the differential’s variable. Another popular way to write (7.22) avoids the issue:

\[
\int_a^b u(x) v'(x) \, dx = u(x)v(x) \bigg|_a^b - \int_a^b v(x)u'(x) \, dx.
\]
Example 7.2.12 Compute \( \int_{-\pi}^{\pi} x \sin x \, dx \).

**Solution:** The antiderivative is an easier case than many of our previous examples, but care has to be taken to keep track of all the signs (+/−) in computing the definite integral:

\[
\begin{align*}
    u &= x & dv &= \sin x \, dx \\
    du &= dx & v &= -\cos x
\end{align*}
\]

\[
\int_{-\pi}^{\pi} x \sin x \, dx = \left. (-x \cos x) \right|_{-\pi}^{\pi} + \int_{-\pi}^{\pi} \cos x \, dx
\]

\[
= [-\pi \cos \pi] - [-(\pi) \cos(-\pi)] + \sin x \bigg|_{-\pi}^{\pi}
\]

\[
= (-\pi)(-1) - (\pi)(-1) + \sin \pi - \sin(-\pi)
\]

\[
= \pi + \pi + 0 - 0
\]

\[
= 2\pi.
\]

In the example above, we could also have noticed that \( \int_{-\pi}^{\pi} \cos x \, dx \) is zero because we are integrating over a whole period \([ -\pi, \pi ]\) of \( \cos x \), and both \( \sin x \) and \( \cos x \) have definite integral zero over any full period \([ a, a + 2\pi ]\). (Think of their graphs, or their definite integrals over any such period.)

It is typical to compute that part \( u(x)v(x) \big|_{a}^{b} \) separately, but one could instead separately compute the entire antiderivative, and then evaluate at the two limits and take the difference:

\[
\int_{-\pi}^{\pi} x \sin x \, dx = \left. (-x \cos x + \sin x) \right|_{-\pi}^{\pi}
\]

\[
= (\pi + 0) - (-\pi + 0)
\]

\[
= 2\pi.
\]

The choice of method is a matter of bookkeeping preferences, and perhaps whether or not part of the right-hand side of (7.22) is particularly simple. If not, it is reasonable to solve the indefinite integral \( \int x \sin x \, dx \) as a separate matter, and then write the definite integral with the formula for the antiderivative inserted, as in \( \int f(x) \, dx = F(x) \big|_{a}^{b} \) and so on as above.

The next example gives us several options for computing the new integral \( \int v \, du \) along the way, though in each the original choices of \( u \) and \( dv \) are the same.

Example 7.2.13 Compute \( \int x \tan^{-1} x \, dx = (I) \).

**Solution:** Again we have little choice on our selection of \( u \) and \( dv \).

\[
\begin{align*}
    u &= \tan^{-1} x & dv &= x \, dx \\
    du &= \frac{1}{x^2 + 1} \, dx & v &= \frac{1}{2} x^2
\end{align*}
\]

\[
(I) = uv - \int v \, du
\]

\[
= \frac{1}{2} x^2 \tan^{-1} x - \frac{1}{2} \int \frac{x^2}{x^2 + 1} \, dx.
\]
Now this last integral can be found by first rewriting the integrand using either polynomial long
division, or by using a little cleverness:

\[
\frac{x^2}{x^2 + 1} = \frac{x^2 + 1 - 1}{x^2 + 1} = \frac{x^2 + 1}{x^2 + 1} - \frac{1}{x^2 + 1} = 1 - \frac{1}{x^2 + 1}.
\]

Polynomial long division would have yielded the same result, and is a bit more straightforward,
but the technique we used here is good to have available. Either way we then have

\[
(I) = \frac{1}{2} x^2 \tan^{-1} x - \frac{1}{2} \int \left(1 - \frac{1}{x^2 + 1}\right) \, dx
\]

\[
= \frac{1}{2} x^2 \tan^{-1} x - \frac{1}{2} x + \frac{1}{2} \tan^{-1} x + C.
\]

Though our choice of \(u\) and \(dv\) was limited, our choice of \(v\) was actually not as limited. Recall
that we could have chosen any \(v = \frac{1}{2} x^2 + C_1\). While previously we chose \(C_1 = 0\) for its apparent
simplicity, for this particular integral we could have actually saved ourselves some effort if we
had chosen \(v\) more strategically, with a different \(C_1\):

\[
u = \tan^{-1} x \quad dv = x \, dx
\]

\[
du = \frac{1}{x^2 + 1} \, dx \quad v = \frac{1}{2} (x^2 + 1)
\]

In effect we chose \(C_1 = \frac{1}{2}\). This gives us

\[
(I) = uv - \int v \, du
\]

\[
= \frac{1}{2} (x^2 + 1) \tan^{-1} x - \int \frac{1}{2} \frac{(x^2 + 1)}{x^2 + 1} \, dx
\]

\[
= \frac{1}{2} (x^2 + 1) \tan^{-1} x - \int \frac{1}{2} \, dx
\]

\[
= \frac{1}{2} (x^2 + 1) \tan^{-1} x - \frac{1}{2} x + C,
\]

which is the same as before, though slightly rearranged.

Though rare, and not crucial, strategically adding a particular constant to the natural choice
for \(v\) can on occasion make for easier computations.

Integration by parts is a very important technique to have at one’s disposal when tackling
integration problems. As has been demonstrated here, such problems can range from fairly simple
computations which are perhaps just slightly more involved than more routine substitution-type
integration problems, to some which utilize or indeed require much more clever applications of
the technique. At the center of all these is a rearranged product rule,

\[
u \cdot \frac{dv}{dx} = \frac{d(uv)}{dx} - v \cdot \frac{du}{dx},
\]

which we can integrate to get

\[
\int u \cdot \frac{dv}{dx} \, dx = \int \frac{d(uv)}{dx} \, dx - \int v \cdot \frac{du}{dx} \, dx,
\]

or our working formula which then follows:

\[
\int u \, dv = uv - \int v \, du.
\]
Of course the applications can be quite clever and detailed, but they all are based upon this relatively simple formula.

**Example 7.2.14** An amount of money $P$ invested at an annual interest rate $r$ (written as a decimal), compounded continuously for $t$ years will then be worth $A(t) = Pe^{rt}$. Solving for $P$ we would get the present value of $A$ dollars $t$ years from now to be

$$P = Ae^{-rt}.$$ 

The idea is that a smaller amount of money $P$ today (presently) could be invested to return the amount $A$ in $t$ years. Naturally so we expect $P < A$.

If instead we have revenue flowing into the account at a constant yearly rate of $R$ dollars per year (but flowing at that yearly rate constantly) for $T$ years, then the present value $P$ is instead given by

$$P = \int_0^T Re^{-rt} \, dt.$$ 

This can also be seen by examining the general case, where the revenue flow rate is dependent upon time, so $R = R(t)$. Then then the present value is given by

$$P = \int_0^T R(t)e^{-rt} \, dt. \quad (7.23)$$

This is reasonable because it represents the accretion of the present values of each injected revenue $R(t)$ for $t \in [0, T]$, because:

- $R(t) \, dt$ represents the revenue injected at time $t$ at a rate of $R(t)$ for an infinitesimal time interval of length $dt$;
- $R(t)e^{-0.05t} \, dt$ therefore represents the present value of that revenue, each of these to be accumulated in the integral in (7.23).

Use (7.23) to calculate the present value $P$ of the first three years of revenue when $R(t) = 2000 + 100t$ (in dollars per year) and the compounded interest rate is 5% (compounded continuously).

**Solution:** Given $R(t) = 2000 + 100t$ and $r = 0.05$, then

$$P = \int_0^3 (2000 + 100t)e^{-0.05t} \, dt.$$ 

We can break this into two integrals—one of which requires integration by parts—or we can wrap both integrals into one integration by parts process. Here we choose the latter approach.

$$u = 2000 + 100t \quad dv = e^{-0.05t} \, dt$$

$$du = 100 \, dt \quad v = \frac{1}{-0.05}e^{-0.05t} = -20e^{-0.05t}.$$ 

---

12To be clear, note that we can have a velocity of 60 miles/hour for any period of time, including significantly less than an hour, or even as an instantaneous rate. Here we have a constant (and instantaneous) flow rate of $R$ dollars per year. It does not mean that we invest $R$ dollars in a lump sum at the end of each year, but that we have a constant flow which would amount to $R$ dollars if it were allowed to proceed for an entire year, just as 60 miles/hour would accumulate to 60 miles after one hour. Next we let $R$ vary, as in $R = R(t)$. The rate will still be in dollars per year, but it will be an instantaneous rate.
From this we get

\[
P = \left(2000 + 100t\right)\left(-20e^{-0.05t}\right)^3 + \left(\int_0^3 20e^{-0.05t} \cdot 100 dt\right)\left(−2000(20 + t)e^{-0.05t} + 2000e^{-0.05t}\right)^3
\]

\[
= \left(-80,000 - 2000t\right)e^{-0.05t}\left(-80,000e^{-0.15} + 80,000e^0 \approx -74,021 + 80,000\right)
\]

\[
\Rightarrow P \approx 5,979.
\]

Ultimately, the computation in this example means that we would have to invest a lump sum of approximately $5979 at the same 5% (compounded continuously) rate today to accumulate the total value of our more complicated investment strategy with a nonconstant revenue stream \(R(t)\) in the same three years (at the same interest rate), with that total being approximately (in part because we have rounded to the nearest dollar for \(P \approx 5979\)) \(A = 5979e^{0.05(3)} \approx 6947\).

Another way to compute this total (future) value $6947 of our investment scheme after three years is to use the following computation (shown in general and for our specific case):

\[
A(T) = \int_0^T R(t)e^{r(T-t)} dt, \quad A(3) = \int_0^3 (2000 + 100t)e^{0.05(3-t)} dt. \tag{7.24}
\]

This is because we can look at the quantities inside the integral in the following way:

- \(R(t) = (2000 + 100t)\) is the rate of revenue (investment) flowing into the account per unit time at time \(t\).
- \(R(t) dt = (2000 + 100t) dt\) represents the amount of new revenue invested during an infinitesimal time interval of length \(dt\) at time \(t\).
- \(T - t = 3 - t\) is the total length in years that the investment made at time \(t\) earns interest.

Thus the integral (7.24) represents the accretion of values of each injected revenue \(R(t)\) for all times \(t \in [0, T]\). For our case of \(R(t) = 2000 + 100t, r = 0.05\) and \([0, T] = [0, 3]\) we integrate this as before:

\[
\begin{align*}
  u &= 2000 + 100t, & dv &= e^{0.05(3-t)} dt, \\
  du &= 100 dt, & v &= -20e^{0.05(3-t)},
\end{align*}
\]

\[
A(3) = \int_0^3 (2000 + 100t)e^{0.05(3-t)} dt = -20(2000 + 100t)e^{0.05(3-t)}\left|_0^3\right. + 2000 \int_0^3 e^{0.05(3-t)} dt
\]

\[
= -20(2000 + 100t)e^{0.05(3-t)}\left|_0^3\right. - 40,000e^{0.05(3-t)}\left|_0^3\right.
\]

\[
= -80,000 \left( e^0 - e^{-15} \right) - 6000
\]

\[
\approx 12,947 - 6000 = 6947,
\]

so indeed we see that the value $6947 of the revenue stream in 3 years (calculated by (7.24) above) is the same as what our calculated present value \(P \approx 5979\) would be worth in 3 years if invested all at once now at the same 5% interest rate, compounded continuously.
7.2. INTEGRATION BY PARTS

While in this model revenue $R(t)$ is a continuous function approximating what in reality is a discrete depositing and interest phenomenon, we can nonetheless approximate (very well in fact) with an integral what would be the present and future values of a somewhat complicated revenue injection and investment scheme, running for several years. This is one of many ways in which calculus is applied in business and economic settings.

**Exercises**

Compute the following integrals, all of which can be computed “by parts.”

1. $\int x \sin x \, dx$
2. $\int x^2 \sin x \, dx$
3. $\int x \cos 3x \, dx$
4. $\int x \sec x \tan x \, dx$
5. $\int x \sec^2 x \, dx$
6. $\int x \ln x \, dx$
7. $\int x \tan^{-1} x \, dx$
8. $\int x \sec^{-1} x \, dx$, $x > 1$
9. $\int x \sec^{-1} x \, dx$, $x < 1$
10. $\int x \sqrt{1-x} \, dx$. (Parts optional)
11. $\int \frac{x}{\sqrt{1-x}} \, dx$ (Parts optional)
12. $\int xe^x \, dx$
13. $\int xe^{x/2} \, dx$
14. $\int x^3 e^{-x^2} \, dx$
15. $\int x^5 \sin x^3 \, dx$
16. $\int x^2 e^{3x} \, dx$
17. $\int \ln x \, dx$
18. $\int \tan^{-1} x \, dx$
19. $\int \sin^{-1} x \, dx$
20. $\int x \sqrt{1-x^2} \sin^{-1} x \, dx$
21. $\int x^3 \sin 2x \, dx$
22. $\int (\ln x)^2 \, dx$
23. $\int \sin^2 5x \, dx$ (Parts optional)
24. $\int \cos^2 x \, dx$ (Parts optional)
25. $\int e^{5x} \cos 2x \, dx$
26. $\int \sec^3 x \, dx$
27. The current flowing in a particular circuit as a function of time $t$ is given by $i = e^{-3t} \sin t$. Determine the charge $q$ which has passed through the circuit in $[0, t]$. Recall that $i = \frac{dq}{dt}$.
28. The slope of a curve is given by $\frac{dy}{dx} = x^3 \sqrt{1 + x^2}$. Find the equation of the curve if it passes through the point $(0, 1)$. 
29. The root-mean-square value of a function $f(x)$ over an interval $[0, T]$ is given by

$$f_{\text{rms}}(x) = \sqrt{\frac{1}{T} \int_{0}^{T} [f(x)]^2 \, dx}. \quad (7.25)$$

Find the root-mean-square value of the function $f(x) = \sqrt{x} \sin^{-1} x$ over $[0, 1]$.

30. Find the root-mean-square value of the function $f(x) = \sqrt{e^x \cos x}$ over $[0, \pi/2]$.

31. Find the present value $P$ of the first 2 years of revenue when $R(t) = (800 + 10 \sin t)$ dollars per year and the compounded interest rate (continuously compounded) is 4.5%. (See Example ?????)
7.3 Trigonometric Integrals

We have already looked at two basic types of trigonometric integrals: those arising from the derivatives of the trigonometric functions (Subsection 6.1.4, page 532), and those of the elementary substitution types in Section 6.5. In this section we are mainly interested in computing integrals where the integrands are combinations of powers of trigonometric functions. In such cases, the angles of each trigonometric function appearing are all the same. Another important topic considered here is how to deal with trigonometric combinations where the angles differ, and we will examine how to deal with several of those cases.

In the first examples where the angles agree, we rearrange the terms in the integrand and use the three basic trigonometric identities to write the entire integral as function of one trigonometric function, and its differential as the final factor. A substitution step then leads to one or more power rules. Unfortunately this only leads to a solution if the combinations of powers are of a few simple forms. Still, these combinations occur often enough to warrant study.

After we look at those simplest forms, we look at other combinations of powers where the angles agree. Techniques include other algebraic manipulations, as well as integration by parts.

In the final forms, where the angles do not agree, we look at several trigonometric identities which help us to rewrite the integrals into simpler forms.

7.3.1 Sample Problems

These first three examples illustrate an approach we develop in Subsections 7.3.2, 7.3.3 and 7.3.4.

Example 7.3.1 Compute $\int \tan^2 x \, dx$.

Solution: $\int \tan^2 x \, dx = \int (\sec^2 x - 1) \, dx = \tan x - x + C$.

The example above used the facts that $\tan^2 x = \sec^2 x - 1$, and that we know the antiderivative of $\sec^2 x$ (where we might not have known the antiderivative of $\tan^2 x$ immediately). The integral above does not in itself contain a general method. Indeed there is no general method, but there are ways to rewrite many trigonometric integrals to make their computations more elementary.

Example 7.3.2 Compute $\int \sin^6 x \cos^5 x \, dx$.

Solution: Here we will use the fact that $\cos^2 x = 1 - \sin^2 x$, and so $\cos^{2k} x = (1 - \sin^2 x)^k$. Eventually we will take $u = \sin x$, implying $du = \cos x \, dx$:

$$\int \sin^6 x \cos^5 x \, dx = \int \sin^6 x \cos^4 x \cos x \, dx$$

$$\begin{align*}
\text{Let } u &= \sin x \\
du &= \cos x \, dx
\end{align*}$$

$$= \int \sin^6 x (1 - \sin^2 x)^2 \cos x \, dx$$

$$= \int u^6 (1 - u^2)^2 \, du = \int u^6 (1 - 2u^2 + u^4) \, du$$

$$= \int [u^6 - 2u^8 + u^{10}] \, du$$

$$= \frac{u^7}{7} - \frac{2u^9}{9} + \frac{u^{11}}{11} + C$$

$$= \frac{1}{7} \sin^7 x - \frac{2}{9} \sin^9 x + \frac{1}{11} \sin^{11} x + C.$$
Example 7.3.3 Compute $\int \sec^5 x \tan^3 x \, dx$.

Solution: Here we will borrow a factor of secant, and another of tangent, to form the functional part of $du$, where $u = \sec x$:

$$\int \sec^5 x \tan^3 x \, dx = \int \sec^4 x \tan x \, dx$$
$$= \int \sec^4 x (\sec^2 x - 1) \tan x \, dx$$
$$= \int u^4 (u^2 - 1) \, du$$
$$= \int [u^6 - u^4] \, du$$
$$= \frac{u^7}{7} - \frac{u^5}{5} + C$$
$$= \frac{1}{7} \sec^7 x - \frac{1}{5} \sec^5 x + C.$$

Now we look at these three specific techniques more closely and generalize them.

7.3.2 Odd Powers of Sine or Cosine

Here we are interested in the cases of integrals

$$\int \sin^m \theta \cos^n \theta \, d\theta. \quad (7.26)$$

where either $m$ or $n$ is odd. Suppose, for example, that $m$ is odd, so that we can write $m = 2k + 1$ for some integer $k$. Then we rewrite the form $(7.26)$ as

$$\int \sin^m \theta \cos^{2k+1} \theta \, d\theta = \int \sin^m \theta \cos^{2k} \theta \cos \theta \, d\theta.$$

The cosine term which we “peeled away” becomes the functional part of the $du$, where $u = \sin \theta$ (so $du = \cos \theta \, d\theta$). We then write the rest of the integral in terms of $u = \sin \theta$. To do so we use

$$\sin^2 \theta + \cos^2 \theta = 0$$
$$\iff \cos^2 \theta = 1 - \sin^2 \theta$$
$$\iff \cos^{2k} \theta = (1 - \sin^2 \theta)^k.$$

Using this fact in the integral above, and setting $u = \sin \theta$, we get

$$\int \sin^m \theta \cos^{2k+1} \theta \, d\theta = \int \sin^m \theta \cos^{2k} \theta \cos \theta \, d\theta$$
$$= \int \sin^m \theta (1 - \sin^2 \theta)^k \cos \theta \, d\theta$$
$$= \int u^m (1 - u^2)^k \, du.$$

This yields a polynomial integrand, which we may then wish to expand before computing (with a sequence of power rules).
Similarly, if there is an odd power of the sine, we can use the fact that \( \sin^2 \theta = 1 - \cos^2 \theta \), and eventually using \( u = \cos \theta \), to rewrite such an integral

\[
\int \sin^{2k+1} \theta \cos \theta \, d\theta = \int \sin^{2k} \theta \cos \theta \sin \theta \, d\theta \\
= \int (1 - \cos^2 \theta)^k \cos^n \theta \sin \theta \, d\theta \\
= \int (1 - u^2)^k u^n (-du) \\
= -\int (1 - u^2)^k u^n \, du.
\]

In both of these it was crucial that we had an odd number of factors of either the sine or cosine, since “peeling off” one factor then leaves an even number, which can be easily written in terms of the other trigonometric function. The peeled off factor is then the functional part of the differential after substitution.

Note that while any even power of a sine or cosine function can be written entirely in terms of the other, this is not the case with odd powers.\(^\text{13}\) This technique works because removing a factor from an odd power of sine or cosine, both provides the functional part of \( du \) and leaves an even power, which we write in terms of the other function which is then \( u \) in the substitution.

**Example 7.3.4** Compute \( \int \sin^5 x \cos^4 x \, dx \).

**Solution:** Here we see an odd number of sine factors, as so we peel one away to be part of the differential term, and write the entire integral in terms of the cosine:

\[
\int \sin^5 x \cos^4 x \, dx = \int \sin^4 x \cos^4 x \sin x \, dx \\
= \int \sin^2 x \, \cos^4 x \sin x \, dx \\
= \int (1 - \cos^2 x)^2 \cos^4 x \sin x \, dx.
\]

(In most future computations we will skip the second line above.) Now we take

\[
u = \cos x \\
\implies du = -\sin x \, dx \\
\iff -du = \sin x \, dx.
\]

With the substitution we will have a polynomial to integrate. To summarize and finish the

---

\(\text{13}\) Consider the trigonometric identity \( \sin^2 \theta + \cos^2 \theta = 1 \). When solved for either the sine or cosine function, we get one of the following:

\[
\sin \theta = \pm \sqrt{1 - \cos^2 \theta}, \\
\cos \theta = \pm \sqrt{1 - \sin^2 \theta}.
\]

We see the ambiguity in the \( \pm \), and the introduction of a radical which itself can very much complicate an integral. However, when we raise these to even powers the radicals and the \( \pm \) both disappear, and we are left with sums of nonnegative, integer powers.
problem, we have:

\[
\int \sin^5 x \cos^4 x \, dx = \int (1 - \cos^2 x)^2 \cos^4 x \sin x \, dx \\
= \int (1 - u^2)^2 u^4 (-du) \\
= - \int (1 - 2u^2 + u^4) u^4 \, du \\
= - \int (u^4 - 2u^6 + u^8) \, du \\
= - \frac{1}{5} u^5 + \frac{2}{7} u^7 - \frac{1}{9} u^9 + C \\
= - \frac{1}{5} \cos^5 x + \frac{2}{7} \cos^7 x - \frac{1}{9} \cos^9 x + C.
\]

It should be clear that one can not easily differentiate the final answer and immediately recognize the original integrand. This is because some trigonometric identities were used to get an integrand form which was computable using these methods. Indeed, it is best to check the validity of the steps from the beginning, rather than to differentiate a tentative answer. However, it is an interesting exercise—left to the interested reader—in trigonometric identities to perform the differentiation, and then validate that the answer there is the original integrand.

It is not necessary that the angle is always \(x\). However, for this technique we do require the angles inside the trigonometric functions to always match, and for the approximate differential of the variable of substitution to be present.

**Example 7.3.5** Compute \( \int \sin^4 5x \cos^3 5x \, dx \).

**Solution:** Here there is an odd number of cosine terms, and we act accordingly.

\[
\int \sin^4 5x \cos^3 5x \, dx = \int \sin^4 5x \cos^2 5x \cos 5x \, dx \\
= \int \sin^4 5x (1 - \sin^2 5x) \cos 5x \, dx.
\]

Here we have

\[
u = \sin 5x \\
\implies du = 5 \cos 5x \, dx \\
\iff \frac{1}{5} du = \cos 5x \, dx.
\]

Now we begin again, incorporating this new information into our computation:

\[
\int \sin^4 5x \cos^3 5x \, dx = \int \sin^4 5x (1 - \sin^2 5x) \cos 5x \, dx \\
= \int u^4 (1 - u^2) \cdot \frac{1}{5} \, du \\
= \frac{1}{5} \int (u^4 - u^6) \, du \\
= \frac{1}{5} \cdot \frac{1}{5} u^5 - \frac{1}{5} \cdot \frac{1}{7} u^7 + C \\
= \frac{1}{25} \sin^5 5x - \frac{1}{35} \sin^7 5x + C.
\]
The technique works even if only one of the trigonometric functions sine or cosine appears, as long as it is to an odd power.

**Example 7.3.6** Compute \( \int \sin^3 7x \, dx \).

**Solution:** Here we can still peel off a sine factor to be the functional part of our differential, and then write the remaining factors in terms of the cosine.

\[
\int \sin^3 7x \, dx = \int \sin^2 7x \sin 7x \, dx
\]

\[
= \int (1 - \cos^2 7x) \sin 7x \, dx.
\]

Using the substitution \( u = \cos 7x \), so \( du = -7 \sin 7x \, dx \), implying \( -\frac{1}{7} \, du = \sin 7x \, dx \), we get

\[
\int \sin^3 7x \, dx = \int (1 - \cos^2 7x) \sin 7x \, dx
\]

\[
= \int (1 - u^2) \cdot -\frac{1}{7} \, du
\]

\[
= -\frac{1}{7} \left[ u - \frac{1}{3} u^3 \right] + C
\]

\[
= -\frac{1}{7} \cos 7x + \frac{1}{21} \cos^3 7x + C.
\]

Furthermore, not all the powers need to be positive integer powers, as long as one is odd.

**Example 7.3.7** Compute \( \int \frac{\cos^7 x}{\sqrt{\sin x}} \, dx \).

**Solution:** Here we have an odd number of cosine terms, so we will peel one off to be the functional part of our differential. That is, we will have \( u = \sin x \), so \( du = \cos x \, dx \). Thus

\[
\int \frac{\cos^7 x}{\sqrt{\sin x}} \, dx = \int \frac{\cos^6 x}{\sqrt{\sin x}} \cos x \, dx
\]

\[
= \int \frac{(1 - \sin^2 x)^3}{\sqrt{\sin x}} \cos x \, dx
\]

\[
= \int \frac{(1 - u^2)^3}{\sqrt{u}} \, du
\]

\[
= \int \frac{1 - 3u^2 + 3u^4 - u^6}{u^{1/2}} \, du
\]

\[
= \int \left[ u^{-1/2} - 3u^{3/2} + 3u^{7/2} - u^{11/2} \right] \, du
\]

\[
= 2u^{1/2} - 3 \cdot \frac{2}{5} u^{5/2} + 3 \cdot \frac{2}{9} u^{9/2} - \frac{2}{13} u^{13/2} + C
\]

\[
= 2u^{1/2} \left[ 1 - \frac{3}{5} u^2 + \frac{1}{3} u^4 - \frac{1}{13} u^5 \right] + C
\]

\[
= 2\sqrt{\sin x} \left[ 1 - \frac{3}{5} \sin^2 x + \frac{1}{3} \sin^4 x - \frac{1}{13} \sin^6 x \right] + C.
\]
It is possible that both powers are odd, and either function can be peeled off, and the integral written in terms of the other. However, if one of these odd powers is greater than the other, it is more efficient to peel off a factor from the lower power, as the next example demonstrates.

**Example 7.3.8** Compute \( \int \sin^3 x \cos^7 x \, dx \).

**Solution:** We will consider both methods for computing this antiderivative. First we peel off a sine to be part of the differential, and let \( u = \cos x \) (so \( du = -\sin x 
\),
\begin{align*}
\int \sin^3 x \cos^7 x \, dx &= \int \sin^2 x \cos^7 x \sin x \, dx \\
&= \int (1 - \cos^2 x) \cos^7 x \sin x \, dx \\
&= \int (1 - u^2) u^7 (\sin x \, du) \\
&= -\int (u^7 - u^9) \, du \\
&= -\frac{1}{8} u^8 + \frac{1}{10} u^{10} + C \\
&= -\frac{1}{8} \cos^8 x + \frac{1}{10} \cos^{10} x + C.
\end{align*}

Next we instead peel off a cosine factor, and let \( w = \sin x \) (so that \( dw = \cos x \, dx \)).
\begin{align*}
\int \sin^3 x \cos^7 x \, dx &= \int \sin^3 x \cos^6 x \cos x \, dx \\
&= \int \sin^3 x (1 - \sin^2 x)^3 \cos x \, dx \\
&= \int w^3 (1 - w^2)^3 \, dw \\
&= \int w^3 (1 - 3w^2 + 3w^4 - w^6) \, dw \\
&= \int (w^3 - 3w^5 + 3w^7 - w^9) \, dw \\
&= \frac{1}{4} w^4 - \frac{3}{6} w^6 + \frac{3}{8} w^8 - \frac{1}{10} w^{10} + C \\
&= \frac{1}{4} \sin^4 x - \frac{1}{2} \sin^6 x + \frac{3}{8} \sin^8 x - \frac{1}{10} \sin^{10} x + C.
\end{align*}

As we see in the above example, there can be different valid choices for some integrals. The answers may look very different, but that is a reflection of the wealth of trigonometric identities available. In fact the antiderivatives, excluding the arbitrary constants, need not be equal but the difference should be accounted for in the constants.\(^{14}\)

\(^{14}\)Recall the integral \( \int 2 \sin x \cos x \, dx \), for which one can let either \( u = \sin x \) or \( u = \cos x \), yielding
\begin{align*}
\int 2 \sin x \cos x \, dx &= \sin^2 x + C_1, \quad \text{or} \\
\int 2 \sin x \cos x \, dx &= -\cos^2 x + C_2.
\end{align*}

Since these differ by a constant, specifically \( \sin^2 x = -\cos^2 x + 1 \), both are valid. But clearly \( \sin^2 x \neq -\cos^2 x \).
7.3.3 Even Powers of Secant or Odd Powers of Tangent

This technique of peeling off some factors of a trigonometric function to be part of the \(du\) (after substitution) has two workable versions for integrals of the type

\[
\int \sec^n \theta \tan^m \theta \, d\theta. \tag{7.27}
\]

These rely upon the following facts from trigonometry and calculus:

\[
\tan^2 \theta + 1 = \sec^2 \theta, \quad \frac{d}{d\theta} \tan \theta = \sec^2 \theta, \\
\sec^2 \theta - 1 = \tan^2 \theta, \quad \frac{d}{d\theta} \sec \theta = \sec \theta \tan \theta.
\]

The techniques we will employ are as follow:

1. If an integral of the form (7.27) contains an odd power of tangent, we peel off a factor \(\sec \theta \tan \theta\) to be the functional part of the differential. This leaves an even power of tangent, which can be written as a power of \((\sec^2 \theta - 1)\).

2. If an integral of the form (7.27) contains an even power of secant, we peel off a factor \(\sec^2 \theta\) to be the functional part of the differential. The remaining, still even power of secant is then written as a power of \((\tan^2 \theta + 1)\).

Example 7.3.9 Compute \(\int \sec^6 x \tan^8 x \, dx\).

\textbf{Solution:} Here we have an even number of secant factors, and so we can peel off two. Eventually we will let \(u = \tan x\), implying \(du = \sec^2 x \, dx\).

\[
\int \sec^6 x \tan^8 x \, dx = \int \sec^4 x \tan^8 x \sec^2 x \, dx \\
= \int (\tan^2 x + 1)^2 \tan^8 x \sec^2 x \, dx \\
= \int (u^2 + 1)^2 u^8 \, du \\
= \int (u^4 + 2u^2 + 1) u^8 \, du \\
= \int (u^{12} + 2u^{10} + u^8) \, du \\
= \frac{1}{13} u^{13} + \frac{2}{11} u^{11} + \frac{1}{9} u^9 + C \\
= \frac{1}{13} \tan^{13} x + \frac{2}{11} \tan^{11} x + \frac{1}{9} \tan^9 x + C.
\]

Example 7.3.10 Compute \(\int \sec^7 2x \tan^5 2x \, dx\).

\textbf{Solution:} Here we have an odd number of tangent factors, so we peel off a \(\sec 2x \tan 2x\) factor to be the functional part of the differential. Eventually we then have \(u = \sec 2x\), giving
du = 2 sec 2x tan 2x dx and thus \( \frac{1}{2} du = \sec 2x \tan 2x \, dx \).

\[
\int \sec^7 2x \tan^5 2x \, dx = \int \sec^6 2x \tan^4 2x \sec 2x \tan 2x \, dx
\]
\[
= \int \sec^6 2x (\sec^2 2x - 1)^2 \sec 2x \tan 2x \, dx
\]
\[
= \int u^6 (u^2 - 1)^2 \cdot \frac{1}{2} \, du
\]
\[
= \frac{1}{2} \int u^6 (u^4 - 2u^2 + 1) \, du
\]
\[
= \frac{1}{2} \int (u^{10} - 2u^8 + u^6) \, du
\]
\[
= \frac{1}{2} \left[ \frac{1}{11} u^{11} - \frac{2}{9} u^9 + \frac{1}{7} u^7 \right] + C
\]
\[
= \frac{1}{22} \sec^{11} 2x - \frac{1}{9} \sec^9 2x + \frac{1}{14} \sec^7 2x + C.
\]

In fact this last example could be computed by first rewriting the integral in terms of cosines and sines:

\[
\int \frac{\sin^5 2x}{\cos^{12} 2x} \, dx = \int \frac{\sin^4 2x}{\cos^{12} 2x} \sin 2x \, dx = \int \frac{(1 - \cos^2 2x)^2}{\cos^{12} 2x} \sin 2x \, dx
\]
\[
= \int \frac{(1 - u^2)^2}{u^{12}} \cdot \frac{-1}{2} \, du = \frac{-1}{2} \int u^{-12} (1 - 2u^2 + u^4) \, du, \text{ etc.}
\]

Thus, the relationships involving the secant and tangent are not required in this last example. However, rewriting the integral in the previous problem, Example 7.3.9, in terms of sines and cosines would not yield an odd power of either. Thus Example 7.3.9 illustrates an integral which does benefit from the extra structure (algebraic and calculus) of the secant-tangent relationship.

**Example 7.3.11** Compute \( \int \tan^4 x \, dx \).

**Solution:** Here we look at two solutions. In the first, instead of exploiting the fact that there are an even number of factors of secant (namely zero) present here, we will repeatedly use the fact that \( \tan^2 \theta + 1 = \sec^2 \theta \). (In the second line, we let \( u = \tan x \).)

\[
\int \tan^4 x \, dx = \int \tan^2 x (\sec^2 x - 1) \, dx
\]
\[
= \int \tan^2 x \sec^2 x \, dx - \int \tan^2 x \, dx
\]
\[
= \frac{1}{3} \tan^3 x - \int \tan^2 x \, dx
\]
\[
= \frac{1}{3} \tan^3 x - \int (\sec^2 x - 1) \, dx
\]
\[
= \frac{1}{3} \tan^3 x - \tan x + x + C.
\]

Of course the other method is to “peel off” a factor of \( \sec^2 x \), which we do even though it does not really appear. To have it appear, we will multiply and divide the integrand by \( \sec^2 x \). Then
we will let \( u = \tan x \). A long division will give us the sum of powers in our final integral below. \(^{15}\)

\[
\int \tan^4 x \, dx = \int \frac{\tan^4 x}{\sec^2 x} \, dx = \int \frac{\tan^4 x}{\tan^2 x + 1} \sec^2 x \, dx = \int \frac{u^4}{u^2 + 1} \, du = \int \left( u^2 - 1 + \frac{1}{u^2 + 1} \right) \, du = \frac{1}{3} u^3 - u + \tan^{-1} u + C_1 = \frac{1}{3} \tan^3 x - \tan x + \tan^{-1}(\tan x) + C_1 = \frac{1}{3} \tan^3 x - \tan x + x + C.
\]

Here we did have the extra complication of long division. Furthermore, to see that the two answers were the same we had to notice \( \tan^{-1}(\tan x) = x + n\pi \), where \( n \in \mathbb{Z} \), i.e., \( n \) is an integer. Thus the final constant \( C \) takes into account \( C_1 - n\pi \), still a constant.

### 7.3.4 Even Powers of Cosecant or Odd Powers of Cotangent

Here we just point out that a similar relationship exists between the cosecant and cotangent, as exists between the secant and tangent. We briefly look at two examples to illustrate this. The integral type is

\[
\int \csc^m \theta \cot^n \theta \, d\theta. \tag{7.28}
\]

We begin with the following facts from trigonometry and calculus:

\[
\cot^2 \theta + 1 = \csc^2 \theta, \quad \frac{d}{d\theta} \cot \theta = -\csc^2 \theta, \\
\csc^2 \theta - 1 = \cot^2 \theta, \quad \frac{d}{d\theta} \csc \theta = -\csc \theta \cot \theta.
\]

The techniques we will employ mirror those used for the secant-tangent integrals:

1. If an integral of the form \((7.28)\) contains an odd power of cotangent, we peel off a factor \( \csc \theta \cot \theta \) to be the functional part of the differential. This leaves an even power of cotangent, which can be written as a power of \( \csc^2 \theta - 1 \).

2. If an integral of the form \((7.28)\) contains an even power of cosecant, we peel off a factor \( \csc^2 \theta \) to be the functional part of the differential. The remaining even power of cosecant is then written as a power of \( \cot^2 \theta + 1 \).

\(^{15}\)A clever student might also note that

\[
\frac{u^4}{u^2 + 1} = \frac{u^4 - 1 + 1}{u^2 + 1} = \frac{u^4 - 1}{u^2 + 1} + \frac{1}{u^2 + 1} = \frac{(u^2 - 1)(u^2 + 1)}{u^2 + 1} + \frac{1}{u^2 + 1} = u^2 - 1 + \frac{1}{u^2 + 1}.
\]

However, such “clever” methods are difficult to generalize, and polynomial long division is more straightforward.
Example 7.3.12 Compute \( \int \csc^8 x \cot^2 x \, dx \).

Solution: We see an even number of cosecants, so we peel off two to be part of the differential.

\[
\int \csc^8 x \cot^2 x \, dx = \int \csc^6 x \cot^2 x \csc^2 x \, dx = \int (\csc^2 x)^3 \cot^2 x \csc^2 x \, dx = \int (\cot^2 x + 1)^3 \cot^2 x \csc^2 x \, dx.
\]

Taking \( u = \cot x \), giving \( du = -\csc^2 x \, dx \), so \(-du = \csc^2 x \, dx\), we get

\[
\int \csc^8 x \cot^2 x \, dx \int (\cot^2 x + 1)^3 \cot^2 x \csc^2 x \, dx
\]

\[
= \int (u^2 + 1)^3 u^2 (-du)
\]

\[
= -\int (u^6 + 3u^4 + 3u^2 + 1) u^2 \, du
\]

\[
= -\int (u^8 + 3u^6 + 3u^4 + u^2) \, du
\]

\[
= -\frac{1}{9} u^9 - \frac{3}{7} u^7 - \frac{3}{5} u^5 - \frac{1}{3} u^3 + C
\]

\[
= \frac{1}{9} \cot^9 x - \frac{3}{7} \cot^7 x - \frac{3}{5} \cot^5 x - \frac{1}{3} \cot^3 x + C.
\]

Example 7.3.13 Compute \( \int \csc^3 \frac{x}{2} \cot^3 \frac{x}{2} \, dx \).

Solution: Here the cotangent appears to an odd power, so we will peel of one cosecant and one cotangent.

\[
\int \csc^3 \frac{x}{2} \cot^3 \frac{x}{2} \, dx = \int \csc^2 \frac{x}{2} \cot^2 \frac{x}{2} \csc \frac{x}{2} \cot \frac{x}{2} \, dx = \int \csc^2 \frac{x}{2} \left( \csc^2 \frac{x}{2} - 1 \right) \csc \frac{x}{2} \cot \frac{x}{2} \, dx.
\]

Now we let \( u = \csc \frac{x}{2} \), implying \( du = -\csc \frac{x}{2} \cot \frac{x}{2} \cdot \frac{1}{2} \, dx \), whence \(-2 \, du = \csc \frac{x}{2} \cot \frac{x}{2} \, dx\). Our integral then becomes

\[
\int \csc^3 \frac{x}{2} \cot^3 \frac{x}{2} \, dx = \int \csc^2 \frac{x}{2} \left( \csc^2 \frac{x}{2} - 1 \right) \csc \frac{x}{2} \cot \frac{x}{2} \, dx
\]

\[
= \int u^2 (u^2 - 1) (-2) \, du
\]

\[
= -2 \int (u^4 - u^2) \, du
\]

\[
= -\frac{2}{5} u^5 + \frac{2}{3} u^3 + C
\]

\[
= -\frac{2}{5} \csc^5 \frac{x}{2} + \frac{2}{3} \csc^3 \frac{x}{2} + C.
\]
7.3.5 Even Powers of Sine and Cosine

Now we turn our attention to the question of integration when both powers of sine and cosine are even. There are two standard methods for handling this: integration by parts, and “half-angle formulas.” The former is more useful when the powers are small than when they are large, and the latter is perhaps more general.\(^{16}\)

**Example 7.3.14** Compute \(\int \sin^2 x \, dx\) using integration by parts.

**Solution**: This exact computation was performed in Example 7.2.11, page 608. So that it is in front of us here, we summarize that computation:

\[
(\mathcal{I}) = \int \sin x \sin x \, dx = (\sin x)(-\cos x) - \int (-\cos x) \cos x \, dx = -\sin x \cos x + \int \cos^2 x \, dx
\]

\[
= -\sin x \cos x + \int (1 - \sin^2 x) \, dx = -\sin x \cos x + x - \int \sin^2 x \, dx
\]

\[
= x - \sin x \cos x - (\mathcal{I}).
\]

At this point we add \((\mathcal{I}) = \int \sin^2 x \, dx\) to both sides to get

\[
2 \int \sin^2 x \, dx = x - \sin x \cos x + C_1
\]

\[
\implies \int \sin^2 x \, dx = \frac{1}{2} \left(x - \sin x \cos x\right) + C.
\]

The method above works well for integrating \(\sin^2 x\) or \(\cos^2 x\), but higher, even powers become more cumbersome. For this reason it is common to opt for alternatives involving slightly more sophisticated trigonometric identities. There is some redundancy in the list below, as (7.29), (7.30), (7.31) and (7.33) together imply the others.

\[
\sin(-\theta) = -\sin \theta, \quad \cos(-\theta) = \cos \theta, \quad \sin(A + B) = \sin A \cos B + \sin B \cos A \tag{7.29}
\]

\[
\sin(A - B) = \sin A \cos B - \sin B \cos A \tag{7.30}
\]

\[
\cos(A + B) = \cos A \cos B - \sin A \sin B \tag{7.31}
\]

\[
\cos(A - B) = \cos A \cos B + \sin A \sin B \tag{7.32}
\]

\[
\sin 2\theta = 2 \sin \theta \cos \theta \tag{7.33}
\]

\[
\cos 2\theta = \cos^2 \theta - \sin^2 \theta \tag{7.34}
\]

\[
\cos 2\theta = 2 \cos^2 \theta - 1 \tag{7.35}
\]

\[
\cos 2\theta = 1 - 2 \sin^2 \theta. \tag{7.36}
\]

It is left for the exercises to show that

1. Equation (7.32) follows from replacing \(B\) with \(-B\) in (7.31),

---

\(^{16}\)In today’s calculus texts, integration by parts is less prominently presented for such integrals, while half-angle methods are more popular among authors. We present both here for the lower powers, as some of the phenomena found in the integration by parts for such integrals are found later in this section.
2. Similarly, (7.34) follows from (7.33).

3. Equation (7.35) follows from (7.31), if we let \( A, B = \theta \).

4. Similarly (7.36) follows from (7.33).

5. (7.37) and (7.38) follow from (7.36) and the identity \( \sin^2 \theta + \cos^2 \theta = 1 \).

Now (7.37) and (7.38) can be rewritten as follow:

\[
\cos 2\theta + 1 = 2 \cos^2 \theta,
\]
\[
2 \sin^2 \theta = 1 - \cos 2\theta.
\]

Dividing each of these by 2 gives us the so-called half-angle formulas:\(^{17}\)

\[
\cos^2 \theta = \frac{1}{2} (1 + \cos 2\theta),
\]
\[
\sin^2 \theta = \frac{1}{2} (1 - \cos 2\theta).
\]

Using (7.40), we see

\[
\int \sin^2 x \, dx = \int \frac{1}{2} (1 - \cos 2x) \, dx = \frac{1}{2} \left( x - \frac{1}{2} \sin 2x \right) + C.
\]

For reasons which will be clear in the next section, it is often desirable that the angle in the final answer agree with the original angle, in this case \( x \). For that we use the double-angle formula (7.35), to get

\[
\int \sin^2 x \, dx = \frac{1}{2} \left( x - \frac{1}{2} \sin 2x \right) + C = \frac{1}{2} \left( x - \frac{1}{2} \cdot 2 \sin x \cos x \right) + C
\]

\[
= \frac{1}{2} x - \frac{1}{2} \sin x \cos x + C.
\]

This agrees with the answer we obtained through integration by parts, in Example 7.3.14, page 625.

**Example 7.3.15** Compute \( \int \sin^4 x \, dx \).

\(^{17}\)Equations (7.39) and (7.40) are called half-angle formulas because the angle \( \theta \) on the left is half of the angle \( 2\theta \) on the right. In fact, knowing the location of the terminal side of an angle does not tell us where its half is located. Indeed, 90° and 450° are coterminal, but their half angles, 45° and 225° are not. This is reflected in what are given as “half-angle” formulas in most trigonometry texts (compare to (7.39) and (7.40)):

\[
\cos \frac{\alpha}{2} = \pm \sqrt{\frac{1 + \cos \alpha}{2}},
\]
\[
\sin \frac{\alpha}{2} = \pm \sqrt{\frac{1 - \cos \alpha}{2}}.
\]

However, knowing where an angle terminates does determine where twice the angle terminates, as is reflected in (7.35)–(7.38).
Solution: Here we use the half-angle formulas repeatedly, until our integral has no positive, even powers of sine or cosine:

\[
\int \sin^4 x \, dx = \int (\sin^2 x)^2 \, dx = \int \left( \frac{1}{2} (1 - \cos 2x) \right)^2 \, dx \\
= \frac{1}{4} \int (1 - 2\cos 2x + \cos^2 2x) \, dx \\
= \frac{1}{4} \int \left[ 1 - 2\cos 2x + \frac{1}{2}(1 + \cos 4x) \right] \, dx \\
= \frac{1}{4} \int \left[ \frac{3}{2} - 2\cos 2x + \frac{1}{2}\cos 4x \right] \, dx \\
= \frac{1}{4} \left[ \frac{3}{2} x - \sin 2x + \frac{1}{2} \cdot \frac{1}{4}\sin 4x \right] + C \\
= \frac{3}{8} x - \frac{1}{4}\sin 2x + \frac{1}{32}\sin 4x + C.
\]

This answer is correct, but if we want to match the angles to the original \(x\), we can use some double-angle formulas (7.35) and (7.36):

\[
\int \sin^4 x \, dx = \frac{3}{8} x - \frac{1}{4}\sin 2x + \frac{1}{32}\sin 4x + C \\
= \frac{3}{8} x - \frac{1}{4} \cdot 2\sin x \cos x + \frac{1}{32} \cdot 2\sin 2x \cos 2x + C \\
= \frac{3}{8} x - \frac{1}{2}\sin x \cos x + \frac{1}{16} (2\sin x \cos x)(\cos^2 x - \sin^2 x) + C,
\]

which can again be simplified and rewritten in several ways.

Example 7.3.16 Compute \( \int \sin^2 3x \cos^2 3x \, dx \).

Solution:

\[
\int \sin^2 3x \cos^2 3x \, dx = \int \frac{1}{2} (1 - \cos 6x) \cdot \frac{1}{2} (1 + \cos 6x) \, dx = \frac{1}{4} \int (1 - \cos^2 6x) \, dx \\
= \frac{1}{4} \int \left[ 1 - \frac{1}{2}(1 + \cos 12x) \right] \, dx = \frac{1}{4} \int \left[ \frac{1}{2} - \frac{1}{2}\cos 12x \right] \, dx \\
= \frac{1}{4} \left[ \frac{3}{2} x - \frac{1}{2} \cdot \frac{1}{12}\sin 12x \right] + C = \frac{x}{8} - \frac{1}{96} \sin 12x + C.
\]

(We could have used \(1 - \cos^2 6x = \sin^2 6x \) after the first line.) As before, if we would like to have our answer in terms of the original angle, we need to utilize the double angle formulas (7.35) and (7.36):

\[
\int \sin^2 3x \cos^2 3x \, dx = \frac{x}{8} - \frac{1}{96} \sin 12x + C \\
= \frac{x}{8} - \frac{1}{96} \cdot 2\sin 6x \cos 6x + C \\
= \frac{x}{8} - \frac{1}{48} (2\sin 3x \cos 3x)(\cos^2 3x - \sin^2 3x) + C \\
= \frac{x}{8} - \frac{1}{24} \sin 3x \cos 3x(\cos^2 3x - \sin^2 3x) + C.
\]
Another method would be to have used the identity
\[
\sin \theta \cos \theta = \frac{1}{2} \sin 2\theta, \tag{7.41}
\]
which follows quickly from the double-angle formula \(\sin 2\theta = 2 \sin \theta \cos \theta\), i.e., (7.35). From this we can compute
\[
\int \sin^2 3x \cos^2 3xdx = \int (\sin 3x \cos 3x)^2 dx
\]
\[
= \int \frac{1}{4} \sin^2 6x dx
\]
\[
= \frac{1}{4} \int \frac{1}{2} (1 - \cos 12x) dx
\]
\[
= \frac{1}{8} \left( x - \frac{1}{12} \sin 12x \right) + C.
\]
While the trigonometric parts of these antiderivatives look very different, the polynomial parts are all \(\frac{1}{8}x\). The trigonometric parts often take interesting manipulations via identities to verify that they are the same, or off by an additive constant. The conclusion is that two students seeking antiderivatives of trigonometric functions may well have very different answers, though the non-trigonometric parts are usually the same.

### 7.3.6 Miscellaneous Problems and Methods-I

There are many trigonometric integrals that either do not fit one of the above categories, or for which those methods are unwieldy. We will look at several such here. The reader should realize, however, that we cannot exhaust all possibilities here, and a particular problem may have a particularly clever solution which does not generalize well to other problems.

The methods of the earlier subsections are all standard and any successful calculus student is expected to know them. The first few examples here are of this class also, in that the better students should be able to handle these without resorting to references. We will, however, eventually have methods in this subsection which such a student should be aware of, but is understandably less likely to be able to recite from memory. All are derivable, but again, the latter are somewhat more obscure and even an excellent students might prefer to use a reference. It is important, however, that all students be aware of these latter classes of problems, and the available methods of solution, regardless of whether a reference is used ultimately.

**Example 7.3.17** Compute \(\int \sec^3 x \, dx\).

**Solution:** Here we have an odd number of secants, and an even (zero) number of tangents. Unfortunately our earlier methods called for an even number of secants or an odd number of tangents. We could notice that the integrand represents an odd number \((-3)\) of cosines, and then with \(u = \sin x\), we could write
\[
\int \sec^3 x \, dx = \int \frac{1}{\cos^4 x} dx = \int \frac{1}{\cos^4 x} \cos x \, dx = \int \frac{1}{(1 - \sin^2 x)^2} \cos x \, dx = \int \frac{1}{(1 - u^2)^2} du,
\]
but in fact we have yet to discuss how to integrate that final form. (We will in Section 7.5, and while it will be somewhat long, it will be a straightforward computation there). Instead we will
next try integration by parts. Since the integrand contains an easily integrated \( \sec^2 x \, dx \) factor, we will let that be \( dv \):

\[
\begin{align*}
  u &= \sec x & dv &= \sec^2 x \, dx \\
  du &= \sec x \tan x \, dx & v &= \tan x
\end{align*}
\]

\[
\int \sec^3 x \, dx = uv - \int v \, du = \sec x \tan x - \int \sec x \tan^2 x \, dx.
\]

Now it is tempting to do another parts step, with \( dv = \sec x \tan x \, dx \), but—as happened in some previous examples—we would then have our original integral on the left, and the same on the right (with all other functions cancelling). What works here is to instead use one of the basic trigonometric identities at this step:

\[
\int \sec^3 x \, dx = \sec x \tan x - \int \sec x \tan^2 x \, dx
\]

Adding \( \int \sec^3 x \, dx \) to both sides gives

\[
2 \int \sec^3 x \, dx = \sec x \tan x + \ln |\sec x + \tan x| + C_1
\]

\[
\Rightarrow \int \sec^3 x \, dx = \frac{1}{2}(\sec x \tan x + \ln |\sec x + \tan x|) + C.
\]

When integrating arbitrary integer powers of the trigonometric functions, a common technique is to make use of so-called reduction formulas. These are derived using integration by parts, often incorporating the kind of computation above. For instance, let us consider the general problem of integrating \( \sec^n x \), where \( n \geq 3 \). Such an integral contains within its integrand the factor \( \sec^2 x \), which we use in the \( dv \) term. Integration by parts can proceed as follows:

\[
\begin{align*}
  u &= \sec^{n-2} x & dv &= \sec^2 x \, dx \\
  du &= (n-2) \sec^{n-3} x \tan x \, dx & v &= \tan x
\end{align*}
\]

i.e., \( du = (n-2) \sec^{n-2} x \tan x \, dx \)

giving us

\[
\int \sec^n x \, dx = \int \frac{\sec^{n-2} x \sec^2 x \, dx}{u} \frac{dv}{dv}
\]

\[
= \sec^{n-2} x \tan x - \int (n-2) \sec^{n-2} x \tan^2 x \, dx
\]

\[
= \sec^{n-2} x \tan x - \int (n-2) \sec^{n-2} x (\sec^2 x - 1) \, dx
\]

\[
= \sec^{n-2} x \tan x - (n-2) \int \sec^n x \, dx + (n-2) \int \sec^{n-2} x \, dx
\]

\[
\Rightarrow (n-1) \int \sec^n x \, dx = \sec^{n-2} x \tan x + (n-2) \int \sec^{n-2} x \, dx.
\]
Now we can divide by \((n - 1)\) to get a general reduction formula
\[
\int \sec^n x \, dx = \frac{1}{n - 1} \sec^{n-2} x \tan x + \frac{n - 2}{n - 1} \int \sec^{n-2} x \, dx. \tag{7.42}
\]
This is called a reduction formula because the resulting integral is of a lower power of secant. A quick inspection reveals that this formula is valid for \(n = 2\) as well, so it is in fact valid for \(n \geq 2\).

**Example 7.3.18** Compute \(\int \sec^5 x \, dx\).

**Solution:** Here we will invoke the formula twice: once for \(n = 5\), and then again for \(n = 3\) to deal with the resulting integral. That will give an integral of secant to the first power, which is one which should be already memorized.

\[
\int \sec^5 x \, dx = \frac{1}{4} \sec^3 x \tan x + \frac{3}{4} \int \sec^3 x \, dx \quad (n=5 \text{ in (7.42)})
\]

\[
= \frac{1}{4} \sec^3 x \tan x + \frac{3}{4} \left\{ \frac{1}{2} \sec x \tan x + \frac{1}{2} \int \sec x \, dx \right\} \quad (n=3 \text{ in (7.42)})
\]

\[
= \frac{1}{4} \sec^3 x \tan x + \frac{3}{8} \sec x \tan x + \frac{3}{8} \ln |\sec x + \tan x| + C.
\]

Other reduction formulas which can be arrived at similarly include
\[
\int \sin^n x \, dx = -\frac{\sin^{n-1} x \cos x}{n} + \frac{n - 1}{n} \int \sin^{n-2} x \, dx, \tag{7.43}
\]
\[
\int \cos^n x \, dx = \frac{\cos^{n-1} x \sin x}{n} + \frac{n - 1}{n} \int \cos^{n-2} x \, dx, \tag{7.44}
\]
\[
\int \tan^n x \, dx = \frac{1}{n - 1} \tan^{n-1} x - \int \tan^{n-2} x \, dx. \tag{7.45}
\]

In fact (7.45) is simpler, not requiring one to “solve” for the integral, but just using that \(\tan^n x = \tan^{n-2} x (\sec^2 x - 1)\). It is left as an exercise.

**Example 7.3.19** Compute \(\int \cos^6 5x \, dx\).

**Solution:** Here we cannot use the formula (7.44) directly, because our angle does not match our differential. To compensate, we will perform a substitution step first. Specifically, we will let \(u = 5x\), so \(du = 5dx\) and thus \(\frac{1}{5} \, du = dx\), giving

\[
\int \cos^6 5x \, dx = \int \cos^6 u \cdot \frac{1}{5} \, du
\]

\[
= \frac{1}{5} \int \cos^6 u \, du
\]

\[
= \frac{1}{5} \left\{ \frac{\cos^5 u \sin u}{6} + \frac{5}{6} \int \cos^4 u \, du \right\} \quad (n = 6 \text{ in (7.44)})
\]

\[
= \frac{\cos^5 u \sin u}{30} + \frac{5}{6} \left\{ \frac{\cos^3 u \sin u}{4} + \frac{3}{4} \int \cos^2 u \, du \right\} \quad (n = 4 \text{ in (7.44)})
\]

\[
= \frac{\cos^5 u \sin u}{30} + \frac{5}{6} \left\{ \frac{\cos^3 u \sin u}{24} + \frac{1}{8} \left[ \frac{\cos u \sin u}{2} + \frac{1}{2} \int 1 \, du \right] \right\} \quad (n = 2 \text{ in (7.44)})
\]

\[
= \frac{\cos^5 u \sin u}{30} + \frac{5}{6} \left\{ \frac{\cos^3 u \sin u}{24} + \frac{1}{8} \left[ \cos u \sin u \frac{1}{2} + \frac{1}{16} u + C \right] \right\}
\]

\[
= \frac{\cos^5 5x \sin 5x}{30} + \frac{5}{6} \left\{ \frac{\cos^3 5x \sin 5x}{24} + \frac{1}{8} \left[ \cos 5x \sin 5x \frac{1}{2} + \frac{5x}{16} + C \right] \right\}.
\]
Clearly reduction formulas can be very useful. Indeed they provide an iterative method for reducing certain integral computations, step by step, until—hopefully—easily manageable integrals appear. In fact in both examples above, we used the reduction formula for one more step than necessary, because it was easier than recomputing \( \int \sec^3 x \, dx \) or \( \int \cos^2 u \, du \) as before. Furthermore, many of the technical details for finding these integrals are built into the reduction formulas.

As useful as the reduction formulas are, they have a couple of minor drawbacks. First, when the angle does not match the differential, some substitution needs to be performed to compensate. Second—and more serious—is that any attempt to memorize these is likely to result in error. Thus the student of calculus needs to learn the earlier methods and be able to perform such calculations unaided, and also know that these reduction formulas (and others) are available and know how to use them.\(^\text{18}\)

Now we consider integrals of following three forms, where \( m \neq n \):

\[
\int \sin mx \sin nx \, dx, \quad \int \sin mx \cos nx \, dx, \quad \int \cos mx \cos nx \, dx.
\]

What distinguishes these is that the angles of the trigonometric functions do not agree. There are two methods for computing these: integration by parts, and utilizing the following trigonometric identities:

\[
\begin{align*}
\sin A \cos B &= \frac{1}{2} \left[ \sin(A - B) + \sin(A + B) \right], \quad (7.46) \\
\sin A \sin B &= \frac{1}{2} \left[ \cos(A - B) - \cos(A + B) \right], \quad (7.47) \\
\cos A \cos B &= \frac{1}{2} \left[ \cos(A - B) + \cos(A + B) \right]. \quad (7.48)
\end{align*}
\]

For obvious reasons, these are called product-sum formulas. These follow from adding or subtracting (7.31), (7.32), (7.33), and (7.34), which we repeat here for reference:

- \( \sin(A + B) = \sin A \cos B + \sin B \cos A \)
- \( \sin(A - B) = \sin A \cos B - \sin B \cos A \)
- \( \cos(A + B) = \cos A \cos B - \sin A \sin B \)
- \( \cos(A - B) = \cos A \cos B + \sin A \sin B \)

For instance, (7.46) follows from adding the first two of these and solving for \( \sin A \cos B \). All three, (7.46)–(7.48), are left to the exercises.

**Example 7.3.20** Compute \( \int \sin 2x \cos 5x \, dx \).

**Solution**: Here we use (7.46), with \( A = 2x \) and \( B = 5x \):

\[
\int \sin 2x \cos 5x \, dx = \int \frac{1}{2} \left[ \sin(-3x) + \sin 7x \right] \, dx \\
= \int \frac{1}{2} \left[ - \sin 3x + \sin 7x \right] \, dx \\
= \frac{1}{3} \cos 3x - \frac{1}{7} \cos 7x + C.
\]

---

\(\text{18}\)Such formulas can be found in most engineering/science calculus texts, as well as books containing tables of integration formulas such as the CRC Standard Mathematical Tables and Formulae.
The method above has the drawback that the solution does not contain the same angles as the integrand. One can get back to the original angles using the formulas

\[
\cos 3x = \cos(5x - 2x) = \cos 5x \cos 2x + \sin 5x \sin 2x, \\
\cos 7x = \cos(5x + 2x) = \cos 5x \cos 2x - \sin 5x \sin 2x.
\]

Alternatively, an integration by parts argument leaves intact the angles. It requires two integration by parts steps, and we need to solve for the integral. Furthermore, we have to make the analogous substitution for \( u \) both times, and for \( dv \) both times. By analogous, here we mean using the same angle, \( 2x \) or \( 5x \), as the argument of the trigonometric function both times. If we always let the \( u \)-term have angle \( 2x \), and the \( dv \)-term have angle \( 5x \), eventually the solution there will be

\[
\int \sin 2x \cos 5x \, dx = \frac{5}{21} \sin 2x \sin 5x + \frac{2}{21} \cos 2x \cos 5x + C.
\]

While it is far from obvious that these results are the same, it is an interesting exercise to check this last computation by computing the derivative of our answer here, so see how it simplifies to the integrand.

7.3.7 Miscellaneous Problems and Methods-II

It is often the case that trigonometric integrals arise in the process of using another technique, and so the form of the integral may be more awkward than we have illustrated here so far. (This will be the case particularly in Section 7.4.) In such cases it is often necessary to experiment with rewriting the integral using whatever trigonometric identities apply. One also has to be aware that a substitution or integration-by-parts argument may be required eventually. Even then, if it is a multi-step application of integration by parts, it helps to recall when it helps to continue the “parts” step, and when it is better to use a trigonometric identity and solve for the integral algebraically, as when we integrated \( \sec^3 x \) or \( \csc^3 x \) (see Example 7.3.17, completed on page 629). In fact, once part of the process yields such an integral, it helps to know that such a complication was probably inevitable and we might as well deal with it from there, rather than attempt to bypass that problem.

With experience one learns to look ahead a few steps and anticipate which algebraic manipulation or identity will yield positive progress towards a form we can integrate (even if it requires some cleverness such as needed to integrate \( \sec^3 x \)), but it is not uncommon to require multiple attempts to integrate a trigonometric integral before finding a strategy that will ultimately succeed.

**Example 7.3.21** Compute \( \int \tan^2 \theta \sec \theta \, d\theta \).

**Solution:** We will attempt to compute this integral two different methods. First we will integrate by parts, noting that

\[
\int \tan^2 \theta \sec \theta \, d\theta = \int \tan \theta \cdot \sec \theta \tan \theta \, d\theta.
\]

\[
u = \tan \theta \quad dv = \sec \theta \tan \theta \, d\theta \\
du = \sec^2 \theta \, d\theta \quad v = \sec \theta
\]
\[ \int \tan^2 \theta \sec \theta \, d\theta = \int \underbrace{\tan \theta \sec \theta \tan \theta \, d\theta}_{uv} \]
\[ = \tan \theta \sec \theta - \int \underbrace{\sec \theta \sec^2 \theta \, d\theta}_{dv} \]
\[ = \sec \theta \tan \theta - \int \sec^3 \theta \, d\theta \]
\[ = \sec \theta \tan \theta - \int \tan^2 \theta \sec \theta \, d\theta - \int \sec \theta \, d\theta \]
\[ = \sec \theta \tan \theta - \int \tan^2 \theta \sec \theta \, d\theta - \ln |\sec \theta + \tan \theta| + C_1 \]
\[ \Rightarrow 2 \int \tan^2 \theta \sec \theta \, d\theta = \sec \theta \tan \theta - \ln |\sec \theta + \tan \theta| + C_1 \]
\[ \Rightarrow \int \tan^2 \theta \sec \theta \, d\theta = \frac{1}{2} \sec \theta \tan \theta - \frac{1}{2} \ln |\sec \theta + \tan \theta| + C. \]

Because an intermediate step included the integral of \( \sec^3 \theta \), which by itself would require an integration by parts technique that includes solving for the integral, it was likely that part of the solution would include our original integral on the left. Rather than beginning an integration by parts, the simple experiment of rewriting \( \sec^3 \theta \) as \( \sec^2 \theta \sec \theta \), and ultimately \( (\tan^2 \theta + 1) \sec \theta \), was successful in producing a form in which we could solve for our original integral.

The other method of attack is to write the entire integral in terms of \( \sec x \) in the first place:

\[ \int \tan^2 x \sec x \, dx = \int (\sec^2 x - 1) \sec x \, dx \]
\[ = \int (\sec^3 x - \sec x) \, dx. \]

While the second function in the integral has a standard and known antiderivative, as mentioned earlier there seems to be no way to avoid integrating \( \sec^3 x \) or similarly complicated terms, so for it we then either use a reduction formula or integrate by parts, using the indirect method where we solve for the integral. The latter approach was built into our first (successful) approach in this example.

Alternatively we could have computed \( \int \sec^3 \theta \, d\theta \) elsewhere and inserted the computation here. That computation could have been carried out as before (Example 7.3.17), or through the appropriate reduction formula (7.42), page 630.

The next example can be computed in two ways, but the first that we show here in fact requires a technique from a later section, namely Section 7.5. We show it here anyways in anticipation of the time when the student can call upon the ideas of more than one of these sections for a single problem, as we did previously (in using integration by parts to solve a trigonometric integral, for instance).

**Example 7.3.22** Compute \( \int \frac{\cos^4 x}{\sin x} \, dx \).
CHAPTER 7. ADVANCED INTEGRATION TECHNIQUES

Solution 1: Note that this integral does in fact contain an odd number of sine factors, even though that number is \(-1\). We can therefore make the integral one in terms of cosine only as the integrand, with the differential containing the sine:

\[
\int \frac{\cos^4 x}{\sin x} \, dx = \int \frac{\cos^4 x}{\sin^2 x} \cdot \sin x \, dx
= \int \frac{\cos^4 x}{1 - \cos^2 x} \cdot \sin x \, dx.
\]

At this point, we have \(u = \cos x \implies du = -\sin x \, dx \implies -du = \sin x \, dx\), and then we eventually use some long division of polynomials to get

\[
\int \frac{\cos^4 x}{\sin x} \, dx = \int \frac{\cos^4 x}{1 - \cos^2 x} \cdot \sin x \, dx = -\int \frac{u^4}{1 - u^2} \, du = \int \frac{u^4}{u^2 - 1} \, du
= \int \left[ u^2 + 1 + \frac{1}{u^2 - 1} \right] \, du.
\]

At this point we need a technique from the Partial Fractions section (Section 7.5) to expand the fraction still inside the integral to get

\[
\frac{1}{u^2 - 1} = -\frac{1}{2} \cdot \frac{1}{u + 1} + \frac{1}{2} \cdot \frac{1}{u - 1}.
\]

While we mention only the final result of that computation, it is easily enough verified by combining the fractions on the right-hand side. To finish our integral computation, we would then write

\[
\int \frac{\cos^4 x}{\sin x} \, dx = \int \left[ u^2 + 1 - \frac{1}{2} \cdot \frac{1}{u + 1} + \frac{1}{2} \cdot \frac{1}{u - 1} \right] \, du
= \frac{1}{3}u^3 + u - \frac{1}{2}\ln |u + 1| + \frac{1}{2}|u - 1| + C
= \frac{1}{3}\cos^3 x + \cos x - \frac{1}{2}\ln |\cos x + 1| + \frac{1}{2}\ln |\cos x - 1| + C.
\]

Solution 2: An alternative approach is to rewrite the entire integral in terms of the sine function:

\[
\int \frac{\cos^4 x}{\sin x} \, dx = \int \left( \frac{\cos^2 x}{\sin x} \right)^2 \, dx = \int \left( \frac{1 - \sin^2 x}{\sin x} \right)^2 \, dx
= \int \frac{1 - 2\sin^2 x + \sin^4 x}{\sin x} \, dx = \int \left( \csc x - 2\sin x + \sin^3 x \right) \, dx
= -\ln |\csc x + \cot x| + 2\cos x + \int \sin^3 x \, dx.
\]

This last integral is then one with an odd power of sine, so we use our previous techniques, with \(u = \cos x \in [-1, 1]\):

\[
\int \frac{1}{1 - u^2} \, du = \tanh^{-1} u + C.
\]
7.3. TRIGONOMETRIC INTEGRALS

\[ u = \cos x \implies du = -\sin x \, dx: \]

\[
\int \sin^3 x \, dx = \int \sin^2 x \cdot \sin x \, dx = \int \frac{(1 - \cos^2 x) \sin x \, dx}{1 - u^2} = -du
\]

\[ = -\cos x + \frac{1}{3} \cos^3 x + C. \]

Thus

\[
\int \frac{\cos^4 x}{\sin x} \, dx = -\ln |\csc x + \cot x| + 2 \cos x - \cos x + \frac{1}{3} \cos^3 x + C
\]

\[ = -\ln |\csc x + \cot x| + \cos x + \frac{1}{3} \cos^3 x + C. \]

So far the major theme of this section is that we often benefit from rewriting trigonometric integrals using trigonometric identities, but with an eye towards the forms that make the calculus more agreeable. We developed some guidelines, but we should also be aware that there are simple cases which might not fit neatly into the guidelines, or might have an easier manipulation than the guidelines would direct.

**Example 7.3.23** Consider \( \int \frac{\cos^2 x}{\sin x} \, dx. \)

This integral does have an odd number of sines, so we could multiply the numerator and denominator by \( \sin x \), with the numerator’s to be used in the \( du \) term, but (as the reader is invited to verify) the resulting integral is not as simple as we get from the following.

\[
\int \frac{\cos^2 x}{\sin x} \, dx = \int \frac{1 - \sin^2 x}{\sin x} \, dx
\]

\[ = \int (\csc x - \sin x) \, dx
\]

\[ = \ln |\csc x - \cot x| + \cos x + C. \]

Trigonometry students know to consider replacing \( \cos^2 x \) with \( 1 - \sin^2 x \) at any such opportunity, to see if it is advantageous. Calculus students should do the same.

**Example 7.3.24** Compute \( \int \frac{1}{1 + \sin x} \, dx. \)

**Solution:** Here we consider multiplying the integrand by \( (1 - \sin x)/(1 - \sin x) \) to see if it is advantageous. (It is.)

\[
\int \frac{1}{1 + \sin x} \, dx = \int \frac{1}{1 + \sin x} \cdot \frac{1 - \sin x}{1 - \sin x} \, dx
\]

\[ = \int \frac{1 - \sin x}{\cos^2 x} \, dx
\]

\[ = \int (\sec^2 x - \sec x \tan x) \, dx
\]

\[ = \tan x - \sec x + C. \]

The second integral could have been done fairly easily using sines and cosines, but would have taken more steps.
Example 7.3.25 Compute \( \int \frac{\tan x}{\sec x - 1} \, dx \).

Solution: Here we do a similar computation but using secants and tangents.

\[
\int \frac{\tan x}{\sec x - 1} \, dx = \int \frac{\tan x \cdot \sec x + \tan x}{\sec^2 x - 1} \, dx
= \int \frac{\tan x \cdot \sec x + \tan x}{\tan^2 x} \, dx
= \int (\csc x + \cot x) \, dx
= \ln |\csc x - \cot x| - \ln |\csc x| + C.
\]

That there are so many techniques here is not surprising, since even if we keep the angles the same, there are six trigonometric functions and a great many identities relating them. The calculus facts guide us somewhat in where to look for identities which can aid in finding a particular antiderivative, but even with these twenty-five examples there are still unexplored possibilities. However these give a reasonable sample of the types of trigonometric integrals we are likely to encounter in the rest of the text.
**Exercises**

Evaluate the following integrals.

1. \( \int \sin x \cos x \, dx \)
2. \( \int \sin^2 x \cos x \, dx \)
3. \( \int \sin x \cos^2 x \, dx \)
4. \( \int \sin^3 x \cos^2 x \, dx \)
5. \( \int \sin^4 x \cos^3 x \, dx \)
6. \( \int \frac{\sin^3 x}{\cos^2 x} \, dx \)
7. \( \int \frac{\sin^3 x}{\cos^2 x + 1} \, dx \)
8. \( \int \sin^4 x \cos^5 x \, dx \)
9. \( \int \sin^3 2x \cos^{15} 2x \, dx \)
10. \( \int \frac{\sin^2 x}{\cos x} \, dx \)
11. \( \int \sin x \ln |\sin x| \, dx \)
12. \( \int \cos x \ln |\sin x| \, dx \)
13. Prove (7.45), page 630.
7.4 Trigonometric Substitution

In this section we explore how integrals can sometimes be solved by making some clever substitutions involving trigonometric functions, even though the original integrals themselves do not involve such functions.

7.4.1 Introduction

Before developing the general mechanics, we look at a few examples below for motivation.

Example 7.4.1 Compute \( \int \frac{1}{\sqrt{1-x^2}} \, dx \), using a nontrivial substitution method.

Solution: We already know the answer to this integral, because we know \( \frac{d}{dx} \sin^{-1} x = \frac{1}{\sqrt{1-x^2}} \), so a “clever” but unnecessary substitution \( u = \sin^{-1} x \) would yield the antiderivative quickly.

But let us imagine for a moment that we do not have this antiderivative at our disposal, and need to tackle this integral from first principles. Can we accomplish this without first developing a theory of derivatives of arc-trigonometric functions?

The answer is yes, if we are clever in a different way, which offers a more general method we can apply to a class of more complicated integrals, as we will see. This substitution method is to notice that \(-1 < x < 1\) is required in the integral, and that is nearly the same as the range of \( \sin \theta \), namely \(-1 \leq \sin \theta \leq 1\), and is in fact contained in it. So we make a substitution, albeit somewhat “backwards” from what we are used to, where \( x \) will be a function of \( \theta \) rather than a function of \( x \) being explicitly some new variable:

\[
\int \frac{1}{\sqrt{1-x^2}} \, dx = \int \frac{1}{\sqrt{1-\sin^2 \theta}} \cos \theta \, d\theta = \int \frac{\cos \theta}{\cos \theta} \, d\theta = \int d\theta = \theta + C = \sin^{-1} x + C
\]

\( x = \sin \theta \quad \Rightarrow \quad dx = \cos \theta \, d\theta \)

The above computation is completely correct, but there are a few technicalities to check.

1. Why are we allowed to take the nonnegative case above, when we know \( \sqrt{1-\sin^2 \theta} = \sqrt{\cos^2 \theta} = |\cos \theta| \)? In other words, usually \( \cos \theta = \pm \sqrt{1-\sin^2 \theta} \), but we took the “+” case.

2. Why can we automatically say \( x = \sin \theta \iff \theta = \sin^{-1} x \)? After all, there are infinitely many angles with the same sine, and they need not necessarily even be coterminal when we graph them in standard position. (Example: \( \sin \frac{\pi}{4} = \sin \frac{3\pi}{4} = \sin \frac{9\pi}{4} = \sin 11\pi 4 \), etc.)

The answer to both questions lies in the values for \( \theta \) that we can choose when we make the substitution. In fact we were negligent by not fixing the range of \( \theta \) from the outset, but we...
will see there is a standard practice in which we will take exactly those angles $\theta$ (or a subset of them) which are the same as the range of the relevant arc-trigonometric function. For the problem above, it is assumed $\theta \in \left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$, since $\sin \theta$ is one-to-one for these values, and $\sin \theta$ ranges over the whole range of the sine function for these values, i.e., for the same reason the arcsine is defined to output these values.

Thus we have mappings from $x$ to $\theta$ and back. (When we actually compute the integral, we concentrate on the second mapping more than the first, until we compute the antiderivative, in which case we may need the first mapping to “get back” to $x$.)

When we look at the values of $\theta$ and their terminal points on the unit circle, all doubt about our casual computation $\sqrt{1 - \sin^2 \theta} = \sqrt{\cos^2 \theta} = |\cos \theta| = \cos \theta$ is laid to rest, because cosine is nonnegative in our range of angles $\theta$:

Now “$x$” here does not signify the coordinate on the horizontal axis. Indeed, for this particular case “$x$” is the same as $\sin \theta$, which is the vertical (“$y$”) coordinate where our angle in standard position (initial side being the positive horizontal axis) pierces the unit circle.

Note that having the angles $\theta$ ranging from $-\pi/2$ to $\pi/2$ means that $-1 \leq \sin \theta \leq 1$ is equivalent to $\theta = \sin^{-1} x$, and also implies $\cos \theta \geq 0$ and so $\cos \theta = \sqrt{1 - \sin^2 \theta}$, where normally we only have $\cos \theta = \pm \sqrt{1 - \sin^2 \theta}$:

1. $(\forall \theta \in \left[\frac{-\pi}{2}, \frac{\pi}{2}\right]) \left[ x = \sin \theta \iff \theta = \sin^{-1} x \right]$, and
2. $(\forall \theta \in \left[\frac{-\pi}{2}, \frac{\pi}{2}\right]) \left[ \cos \theta = \sqrt{1 - \sin^2 \theta} \right]$.

The example above illustrates another useful approach for some integrals:

1. Make a substitution to rewrite the integral in terms of an angle $\theta$ (with appropriate range of $\theta$ matching the range of values for $x$).
2. Compute the resultant trigonometric integral.
3. Rewrite the antiderivative in terms of $x$. 
Drawing diagrams on an appropriate circle as above will be quite useful in subsequent problems. The process can not only clarify somewhat our substitution process, but it can also allow us to check that we have correct signs for all the various cases for $\theta$. Moreover, there are times when we need to read actual values of other trigonometric functions of $\theta$, but in terms of $x$.

The process above may seem unnecessarily complicated, especially for an integral for which we knew the answer from the beginning. However, this advanced technique generalizes to integrals which do not succumb to previous methods (though they should always be considered first!). For instance we would be hard-pressed to compute the antiderivative in the next example using earlier techniques.

Example 7.4.2 Compute $\int \frac{\sqrt{1-x^2}}{x} \, dx$.

Solution: Note first that this integral will not simply yield to earlier techniques. (The reader is welcome to try, to see where those methods eventually fall short.)

Note also that, due to the square root, we require $-1 \leq x \leq 1$. In fact we also cannot have $x = 0$, but that constraint will be consistent with our new integral upon substitution. Ultimately it is the radical which is giving us the most difficulty here.

The key to solving this problem is to again realize that $[-1, 1]$ is exactly the range of the function $\sin \theta$, so we will again use the substitution $x = \sin \theta$ in the integral above, with the understanding that $\theta \in [-\pi/2, \pi/2]$ (excluding zero due to the denominator). This time we will again draw the diagram, but will label the various parts of the resulting triangles for future reference. Note that in the second drawing, $\sin \theta = x < 0$ while $\cos \theta = \sqrt{1-x^2}$ is still a positive quantity, and again, where the angle’s terminal ray pierces the unit circle is the point $(\cos \theta, \sin \theta)$.

Note that we can now read all trigonometric functions of $\theta$ from the diagrams. For instance, $\tan \theta = x/\sqrt{1-x^2}$, regardless of which of the two quadrants contains the terminal ray of $\theta$.

Now we perform the substitution, noting as usual that the differential must also be accounted for:
7.4. TRIGONOMETRIC SUBSTITUTION

Now we have a trigonometric form of the antiderivative, but of course the original integral was in $x$ and not $\theta$. Our diagram allows us to read the other trigonometric functions of $\theta$ in terms of $x$. We note that it does not matter, for this example, which of the two quadrants contains $\theta$:

$$\int \frac{\sqrt{1-x^2}}{x} \, dx = \ln |\csc \theta - \cot \theta| + \cos \theta + C = \ln \left| \frac{1}{x} - \frac{\sqrt{1-x^2}}{x} \right| + \sqrt{1-x^2} + C$$

Unlike the first example, we are not likely to be anxious to check our work by taking the derivative of our solution (though it would be an interesting exercise, particularly to see how terms cancel), so instead we strive to be careful and clear about our derivation, so we can minimize errors and easily re-read our computations to verify our result.

It should be noted at the outset that the trigonometric integrals which arise here may require some re-writing before they succumb to our trigonometric integral methods of Section 7.3; a problem which naturally gives rise to trigonometric substitution (as in the previous example) may or may not yield a simple trigonometric integral. However, all trigonometric integrals we

\[21\] Another student might instead use $\int \csc = -\ln |\csc \theta + \cot \theta + C_1$, but the final answer would be the same:

$$\int \frac{\sqrt{1-x^2}}{x} \, dx = -\ln |\csc \theta + \cot \theta + \cos \theta + C = -\ln \left| \frac{1}{x} + \frac{\sqrt{1-x^2}}{x} \right| + \sqrt{1-x^2} + C$$

which is the same as the answer written in the example.
will encounter here are of classes we considered in Section 7.3, so ultimately those techniques equipped us for our work here.

Before we delve into other trigonometric substitutions, we will perform one more involving the sine.

**Example 7.4.3** Compute \( \int \sqrt{9 - 25x^2} \, dx \).

**Solution:** As always, we should look to see if previous methods apply. They do not, without extraordinary cleverness, though it is interesting to note that simple substitution would work if we had an extra factor of \( x \) in the integrand. (We do not.)

Now in previous examples, we wanted to exploit the trigonometric identity \( 1 - \sin^2 \theta = \cos^2 \theta \). Here we will do the same, except we will multiply this equation by 9, giving us \( 9 - 9\sin^2 \theta = 9\cos^2 \theta \).

So we wish to have \( 25x^2 = 9\sin^2 \theta \), i.e., \( x^2 = \frac{9}{25}\sin^2 \theta \). We get this if we let \( x = \frac{3}{5}\sin \theta \). Below is what a student (or professor) might hand-write to compute this integral:

\[
\int \sqrt{9 - 25x^2} \, dx = \int \sqrt{9 - 25 \cdot \frac{9}{25}\sin^2 \theta} \cdot \frac{3}{5}\cos \theta \, d\theta
\]

\[
25x^2 = 9\sin^2 \theta \quad \Rightarrow \quad x^2 = \frac{9}{25}\sin^2 \theta
\]

\[
\Rightarrow \quad dx = \frac{3}{5}\cos \theta \, d\theta
\]

\[
\therefore \quad \int \sqrt{9 - 25x^2} \, dx = \int \sqrt{9 - 9\sin^2 \theta} \cdot \frac{3}{5}\cos \theta \, d\theta
\]

\[
= \int \sqrt{9(1 - \sin^2 \theta)} \cdot \frac{3}{5}\cos \theta \, d\theta
\]

\[
= \int \sqrt{\frac{9}{5}\cos^2 \theta} \, d\theta
\]

\[
= \int \frac{3}{5}\cos \theta \, d\theta
\]

\[
= \frac{9}{10}\theta + \frac{9}{10} \cdot \frac{1}{2} \cdot 2\sin \theta \cos \theta + C
\]

\[
= \frac{9}{10} \theta + \frac{9}{10} \sin \theta \cos \theta + C.
\]

We require the final trigonometric antiderivative to be in terms of the angle \( \theta \) (and not \( 2\theta \) for example), so that we can read the trigonometric functions of \( \theta \) from the diagram. In constructing the diagram, we need to solve our substitution for \( \sin \theta \):

\[
x = \frac{3}{5}\sin \theta \quad \Rightarrow \quad \sin \theta = \frac{5x}{3}.
\]

Note that our diagram below is not constructed on the unit circle, but on a more general circle of positive radius (in this case 3). As before, we have to consider both quadrants in which \( \theta \) may fall.
From these and our expression for \( \sin \theta \) above we can complete our integral computation:

\[
\int \sqrt{9 - 25x^2} \, dx = \frac{9}{10} \theta + \frac{9}{10} \sin \theta \cos \theta + C
\]

\[
= \frac{9}{10} \sin^{-1} \left( \frac{5x}{3} \right) + \frac{9}{10} \cdot \frac{5x}{3} \sqrt{9 - 25x^2} + C
\]

\[
= \frac{9}{10} \sin^{-1} \left( \frac{5x}{3} \right) + \frac{x}{2} \sqrt{9 - 25x^2} + C.
\]

### 7.4.2 The General Approach

There are cases where a different trigonometric substitution is appropriate and useful. In fact the choices are mutually exclusive, and the form to be used can be deduced from the range of \( x \)-values in the domain of the original integrand, though one instead usually sees how problematic terms would simplify. In the chart below, we use \( x \), though "\( x \)" can be \( x \) or \( 5x \), or similar, with the necessary algebra and calculus to compensate. Also for simplicity, we assume \( a > 0 \):\(^{22}\)

<table>
<thead>
<tr>
<th>Integral contains:</th>
<th>Substitute:</th>
<th>Motivation:</th>
<th>Range of ( x ):</th>
<th>Range of ( \theta ):</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sqrt{a^2 - x^2} )</td>
<td>( x = a \sin \theta )</td>
<td>( \sqrt{a^2 - x^2} = a \cos \theta )</td>
<td>( -a \leq x \leq a )</td>
<td>( \theta \in [-\pi/2, \pi/2] )</td>
</tr>
<tr>
<td>( \sqrt{a^2 + x^2} )</td>
<td>( x = a \tan \theta )</td>
<td>( \sqrt{a^2 + x^2} = a \sec \theta )</td>
<td>( x \in \mathbb{R} )</td>
<td>( \theta \in (-\pi/2, \pi/2) )</td>
</tr>
<tr>
<td>( \sqrt{x^2 - a^2} )</td>
<td>( x = a \sec \theta )</td>
<td>( \sqrt{x^2 - a^2} = a</td>
<td>\tan \theta</td>
<td>)</td>
</tr>
</tbody>
</table>

The graphs of the positions of \( \theta \) are important for verification purposes, as well as for filling in with \( x \)-expressions the trigonometric functions of \( \theta \) that usually arise from the trigonometric form of the antiderivative.

\(^{22}\)In fact, sometimes these substitutions are useful even when the radicals are not present, particularly for the tangent case, such as in computing \( \int \frac{1}{\sqrt{a^2 + x^2}} \, dx \). So we can perhaps make a general statement that these substitutions should be considered if we have powers of \( (a^2 - x^2) \), \( (a^2 + x^2) \) or \( (a^2 - x^2) \), respectively, when previous methods will not work, and the range of values for \( x \) and \( \theta \) are compatible.
Example 7.4.4 Compute \( \int \frac{1}{\sqrt{36 + x^2}} \, dx \).

**Solution:** Here we will let \( x = 6 \tan \theta \). Note that as part of the method (but usually left unstated), \(-\pi/2 < \theta < \pi/2 \iff -\infty < x < \infty \).

\[
\int \frac{1}{\sqrt{36 + x^2}} \, dx = \int \frac{1}{\sqrt{36 + 36 \tan^2 \theta}} \cdot 6 \sec^2 \theta \, d\theta = \int \frac{6 \sec^2 \theta}{\sqrt{36 \sec^2 \theta}} \, d\theta = \int \frac{6 \sec^2 \theta}{6 \sec \theta} \, d\theta
\]

\[
x = 6 \tan \theta \quad \implies dx = 6 \sec^2 \theta \, d\theta
\]

Thus \( \int \frac{1}{\sqrt{36 + x^2}} \, dx = \ln |\sec \theta + \tan \theta| + C_1 \)

\[
= \ln \left| \frac{\sqrt{36 + x^2}}{6} + \frac{x}{6} \right| + C_1
\]

\[
= \ln \left| \frac{\sqrt{36 + x^2} + x}{6} \right| - \ln 6 + C_1
\]

\[
= \ln \left| \sqrt{36 + x^2} + x \right| + C.
\]

It is customary to have the solution in simplest possible form, which is why we expanded the logarithm and absorbed the \(-\ln 6\) term into the constant. (One usually only puts the subscript on the \(C_1\) term after further lines reveal it is useful.)

As we see in the drawings of the relevant angles, when we use a tangent-type substitution we get the same, simple expressions for the derived sides regardless of which of the two quadrants holds the terminal ray of \( \theta \). It is not always the case with the secant substitutions.

Example 7.4.5 Compute \( \int \sqrt{x^2 - 9} \, dx \).

**Solution:** Here we let \( x = 3 \sec \theta \).

\[
\int \sqrt{x^2 - 9} \, dx = \int \sqrt{9 \sec^2 \theta - 9} \cdot 3 \sec \theta \tan \theta \, d\theta
\]

\[
x = 3 \sec \theta \quad \implies dx = 3 \sec \theta \tan \theta \, d\theta
\]

\[
= \int 3|\tan \theta| \cdot 3 \sec \theta \tan \theta \, d\theta
\]
Recall that our choices for \( \theta \) terminate in either the first or second quadrant when we use a secant-type substitution, but while \( \tan \theta \geq 0 \) in Quadrant I, we have \( \tan \theta \leq 0 \) in Quadrant II. These coincide with the cases \( x \) positive and \( x \) negative (or more precisely \( x \geq 3 \) and \( x \leq -3 \)), respectively. Below we graph the two cases, noting that \( \sec \theta = x/3 \), i.e., \( \cos \theta = 3/x \), but also that the “hypotenuse” must be positive in each case, as dictated by trigonometric theory.

The signs of the two “legs” of the representative triangles are also useful in checking the expression for the hypotenuse. Note \( \sec \theta = x/3 = (-x)/(-3) \), and also that \( -x > 0 \) for the case \( \theta \) in QII. (The hypotenuse is a radius, and therefore always positive.)

For this particular example, the antiderivatives for the two cases differ by a factor of \(-1\), so we do most of the work by finding the antiderivative for one case, and changing sign for the other. For simplicity we will compute the antiderivative for the case \( \theta \) in Quadrant I first.

1. Case \( x \geq 3 \): \[ \int \sqrt{x^2 - 9} \, dx = 9 \int \tan^2 \theta \sec \theta \, d\theta. \]

This requires integration by parts, of the type where we “solve” for the integral. We will use

\[
\begin{align*}
    u &= \tan \theta & \quad dv &= \sec \theta \tan \theta \, d\theta \\
    du &= \sec^2 \theta \, d\theta & \quad v &= \sec \theta.
\end{align*}
\]

\[
(I) = uv - \int v \, du = \sec \theta \tan \theta - \int \sec^3 \theta \, d\theta
\]

\[
= \sec \theta \tan \theta - \int (\tan^2 \theta + 1) \sec \theta \, d\theta = \sec \theta \tan \theta - \int \sec \theta \, d\theta - \int \tan^2 \theta \sec \theta \, d\theta
\]

\[
= \sec \theta \tan \theta - \ln |\sec \theta + \tan \theta| - (I).
\]

Solving for \((I)\), we have

\[
2(I) = \sec \theta \tan \theta - \ln |\sec \theta + \tan \theta| + C_1
\]

\[
\Rightarrow (I) = \frac{1}{2} \sec \theta \tan \theta - \frac{1}{2} \ln |\sec \theta + \tan \theta| + C.
\]

Using our definition of \( \sec \theta \) and the previous diagram (for QI), we have

\[
(I) = \frac{1}{2} \cdot \frac{x}{3} \sqrt{x^2 - 9} \cdot \frac{3}{3} - \frac{1}{2} \ln \left| \frac{x}{3} + \sqrt{\frac{x^2 - 9}{3}} \right| + C_2
\]

\[
= \frac{x}{18} \sqrt{x^2 - 9} - \frac{1}{2} \ln \left| x + \sqrt{x^2 - 9} \right| - \ln 3 + C_1
\]

\[
= \frac{x}{18} \sqrt{x^2 - 9} - \frac{1}{2} \ln \left| x + \sqrt{x^2 - 9} \right| + C.
\]
2. Case \( x \leq -3 \): That is, \( \theta \) in QII, where we will have the same antiderivative in \( \theta \) except for a sign. Here \( \tan \theta \leq 0 \), so \( |\tan \theta| = -\tan \theta \), so the trigonometric form of the antiderivative is the same as above except for an extra factor of \((-1)\):

\[
(I) = \int 3(-\tan \theta) \cdot 3 \sec \theta \tan \theta \, d\theta
\]

\[
= -\frac{1}{2} \sec \theta \tan \theta + \frac{1}{2} \ln |\sec \theta + \tan \theta| + C
\]

\[
= -\frac{1}{2} \cdot \frac{x}{3} \sqrt{x^2 - 9} + \frac{1}{2} \ln \left| \frac{x}{3} + \frac{\sqrt{x^2 - 9}}{-3} \right| + C_3
\]

\[
= \frac{x\sqrt{x^2 - 9}}{18} + \frac{1}{2} \ln \left| x - \sqrt{x^2 - 9} \right| - \ln 3 + C_3
\]

\[
= \frac{x\sqrt{x^2 - 9}}{18} + \frac{1}{2} \ln \left| x - \sqrt{x^2 - 9} \right| + C.
\]

Summarizing,

\[
\int \sqrt{x^2 - 9} \, dx = \begin{cases} 
\frac{1}{18} x\sqrt{x^2 - 9} - \frac{1}{2} \ln \left| x + \sqrt{x^2 - 9} \right| + C, & x \geq 3, \\
\frac{1}{18} x\sqrt{x^2 - 9} + \frac{1}{2} \ln \left| x - \sqrt{x^2 - 9} \right| + C, & x \leq -3.
\end{cases}
\]

Often, in a problem like this latest example we will know from the start what will be the range of interest of values of \( x \). For instance, if we know \( x \geq 0 \) (more precisely, \( x \geq 3 \)) we can finish the problem by drawing one diagram only, namely that in the first quadrant. If this were a definite integral, for instance, we would know which range of \( x \) we need by looking at the endpoints of our integral.

It should be emphasized that only the secant-type trigonometric substitutions require us to check both quadrants, because \( \sqrt{\sec^2 \theta - 1} = \pm \tan \theta \) for the range of \( \theta \) we use.\(^{23}\)

Sometimes, particularly for tangent substitutions, no radical is present in the original integral yet the trigonometric substitution is the method of choice.

**Example 7.4.6** Compute \( \int \frac{1}{(x^2 + 9)^2} \, dx \).

**Solution:** Note that this is very different from the case where we have \((x^2 + 9)^1\) in the denominator, which would eventually yield a arctangent, namely \( \frac{1}{9} \tan^{-1} \frac{x}{3} + C \). But it gives us

\(^{23}\)There is an approach where one assumes \( \theta \in [0, \pi/2) \cup [\pi, 3\pi/2) \) or similar QI and QIII angles, and then \( \sqrt{\sec^2 \theta - 1} = \tan \theta \) since \( \tan \theta \geq 0 \) in those quadrants. However we avoid this because sometimes the antiderivative includes the angle \( \theta \) itself, which for our case would be \( \theta = \sec^{-1} \frac{x}{3} = \cos^{-1} \frac{3}{x} \). This other approach simply redefines the arccosecant as well, so \( \sec^{-1} z \in [0, \pi/2] \cup [\pi, 3\pi/2) \). The approach has many advantages (for instance the derivative of the arccosecant does not contain an absolute value), but at a cost of such conveniences as \( \sec^{-1} z = \cos^{-1} \frac{1}{z} \). One also redefines the arccosecant in that approach to arc-trigonometric functions.
a hint of what to do, namely let \( x = 3 \tan \theta \).

\[
\int \frac{1}{(x^2 + 9)^2} \, dx = \int \frac{1}{(9 \tan^2 \theta + 9)} \cdot 3 \sec^2 \theta \, d\theta
\]

\[
x = 3 \tan \theta \quad \Rightarrow \quad \frac{dx}{d\theta} = 3 \sec^2 \theta
\]

\[
dx = 3 \sec^2 \theta \, d\theta
\]

\[
\int \frac{3 \sec^2 \theta \, d\theta}{9 \sec^2 \theta} = \int \frac{3 \sec^2 \theta \, d\theta}{81 \sec^4 \theta}
\]

\[
= \frac{1}{27} \int \cos^2 \theta \, d\theta
\]

\[
= \frac{1}{27} \cdot \frac{1}{2} \int (1 + \cos 2\theta) \, d\theta
\]

\[
= \frac{1}{27} \left[ \frac{1}{2} \theta + \frac{1}{2} \sin 2\theta \right] + C
\]

\[
= \frac{1}{54} \theta + \frac{1}{27} \cdot \frac{1}{2} \cdot 2 \sin \theta \cos \theta + C.
\]

Now the integral is solved in terms of \( \theta \), so we need to get back to \( x \), which we again do by looking at diagrams of the relation between \( x \) and \( \theta \), namely that \( \tan \theta = \frac{x}{3} \), and \( \theta \in (-\pi/2, \pi/2) \).

![Diagram](image)

Finishing off our integral is now fairly easy:

\[
\int \frac{1}{(x^2 + 9)^2} \, dx = \frac{1}{54} \theta + \frac{1}{27} \cdot \sin \theta \cos \theta + C
\]

\[
= \frac{1}{54} \tan^{-1} x + \frac{1}{27} \cdot \frac{x}{\sqrt{x^2 + 9}} \cdot \frac{3}{\sqrt{x^2 + 9}} + C
\]

\[
= \frac{1}{54} \tan^{-1} x + \frac{1}{9} \cdot \frac{x}{x^2 + 9} + C.
\]

While this is a powerful approach, it is not always the technique of choice. For example,

\[
\int \frac{x}{\sqrt{x^2 - 9}} \, dx = \sqrt{x^2 - 9} + C,
\]

from a simple substitution, or perhaps even an anticipation of the general form followed by a check of multiplicative constants. Similarly, if we replaced \( x \) with \( x^2 \) in the numerator we could integrate by parts. If we found ourselves integrating by parts twice, for instance (perhaps it is \( x^3 \) in the numerator) it may well be more efficient to use trigonometric substitution, or an integration by parts step may ultimately require it! But trigonometric substitution is usually not the method of choice if previous methods apply.
Example 7.4.7 Compute \( \int \frac{\sqrt{x^2 - 5}}{x} \, dx \) assuming that \( x \geq \sqrt{5} \).

**Solution:** Here we let \( x^2 = 5 \sec^2 \theta \), i.e., \( x = \sqrt{5} \sec \theta \), giving \( dx = \sqrt{5} \sec \theta \tan \theta \, d\theta \). Note that we are assuming positive \( x \), so \( \theta \in [0, \pi/2) \) so \( \theta \) terminates in QI, where \( \tan \theta \geq 0 \). Thus,

\[
\int \frac{\sqrt{x^2 - 5}}{x} \, dx = \int \frac{\sqrt{5 \sec^2 \theta - 5}}{\sqrt{5} \sec \theta} \cdot \sqrt{5} \sec \theta \tan \theta \, d\theta \\
= \int \frac{\sqrt{5} \tan \theta}{\sqrt{5} \sec \theta} \cdot \sqrt{5} \sec \theta \tan \theta \, d\theta \\
= \sqrt{5} \int \tan^2 \theta \, d\theta \\
= \sqrt{5} \int (\sec^2 \theta - 1) \, d\theta \\
= \sqrt{5} \tan \theta - \sqrt{5} \theta + C \\
= \sqrt{5} \cdot \frac{\sqrt{x^2 - 5}}{\sqrt{5}} - \sqrt{5} \sec^{-1} \frac{x}{\sqrt{5}} + C \\
= \sqrt{x^2 - 5} + \sec^{-1} \frac{x}{\sqrt{5}} + C.
\]

**Exercises**

1. Use trigonometric substitution to derive the general formula \((a > 0)\)

\[\int \frac{dx}{\sqrt{a^2 - x^2}} = \sin^{-1} \frac{x}{a} + C.\]

2. Use trigonometric substitution to derive the general formula \((a > 0)\)

\[\int \frac{dx}{a^2 + x^2} = \frac{x}{a} \tan^{-1} \frac{x}{a} + C.\]

3. Use trigonometric substitution to derive the general formula for \(x > a > 0\):

\[\int \frac{dx}{x \sqrt{x^2 - a^2}} = \frac{x}{a} \sec^{-1} \frac{x}{a} + C.\]

4. Compute \( \int \frac{\sqrt{1 - x^2}}{x^2} \, dx.\)

5. Compute \( \int (9 - x^2)^{3/2} \, dx.\)

6. Compute \( \int \frac{1}{(25 + 9x^2)^{5/2}} \, dx.\)

7. Compute \( \int \sqrt{x^2 + 2x} \, dx. \) (Complete the square.)
7.5 Partial Fractions and Integration

In this section we are interested in techniques for computing integrals of the form

$$\int \frac{P(x)}{Q(x)} \, dx,$$  \hfill (7.49)

where $P(x)$ and $Q(x)$ are polynomials. This is not in general a simple problem, unless the integral in (7.49) is from very particular classes. However, with the techniques we explore here, we can break $\frac{P(x)}{Q(x)}$ into simpler fractions whose integrals are relatively easy. To see an advantage of such an approach, consider the following example.

**Example 7.5.1** Compute $\int \frac{5x - 1}{x^2 - x - 2} \, dx$.

**Solution:** Note that the numerator is not a simple (constant number) multiple of the derivative of the denominator, so substitution will not give a simple $\int \frac{1}{u} du$-form.

However, it happens that

$$\frac{5x - 1}{x^2 - x - 2} = \frac{5x - 1}{(x - 2)(x + 1)} = \frac{2}{x + 1} + \frac{3}{x - 2}. \hfill (7.50)$$

We well see later a technique for deriving the right-hand side of (7.50). This gives us

$$\int \frac{5x - 1}{x^2 - x - 2} \, dx = \int \left( \frac{2}{x + 1} + \frac{3}{x - 2} \right) \, dx$$

$$= 2 \ln |x + 1| + 3 \ln |x - 2| + C$$

$$= \ln |(x + 1)^2(x - 2)^3| + C.$$  

In this section we will develop the methods needed to find expansions of fractions such as in (7.50). The idea is to reverse the high school algebra exercises, which would have us combine sums or differences of fractions into a single fraction. For purposes of integral calculus, it is almost always better to instead deal with several, simpler fractions than their combination into a single, more complicated fraction.

A significant amount of our work in this section will be algebraic, specifically, developing the method of decomposing a fraction $P(x)/Q(x)$ into simpler, “partial fractions.” In order for the method to work, we will require $P(x)$ to have lower degree than $Q(x)$ (If the degree of $P$ is at least that of $Q$, we can use long division to write the function as a polynomial plus $r(x)/Q(x)$, where the degree of $r$ is less than that of $Q$.)

Though not crucial for the calculus, we will spend the next subsection looking roughly at the theory behind the general form of these partial fraction decompositions (PFD’s), in hopes it will help reinforce the rules themselves. In Subsection 7.5.2 we will definitively write the rules for PFD’s without reference to integrals. Finally, we will see how to solve for the coefficients of a particular PFD, in the context of computing antiderivatives of these.

### 7.5.1 Theory Behind the Forms of PFD’s (Optional)

The argument here is usually omitted from calculus texts, and instead left to linear algebra courses (if discussed at all!). However, the basic intuition is not difficult so we include it here, though the real work is in later subsections. In all of these we are looking at functions

$$\frac{P(x)}{Q(x)}, \quad P \text{ and } Q \text{ polynomials, degree } P < \text{degree } Q. \hfill (7.51)$$
Before stating the general rules for PFD’s, we look at several examples illustrating the underlying theory. (Note that a linear combination of two functions \( f(x) \) and \( g(x) \) is a function of the form \( a \cdot f(x) + b \cdot g(x) \), where \( a, b \in \mathbb{R} \) are constants.)

**Example 7.5.2** For this example, we will argue in steps.

1. Consider all functions of the form

\[
\frac{ax + b}{(x + 1)(x - 2)} .
\]

(7.52)

2. Now there are two linearly independent\(^{24}\) functions, specifically \( \frac{1}{(x + 1)(x - 2)} \) and \( \frac{1}{(x + 1)(x - 2)} \) which—with linear combinations—can give us any such function of form (7.52). Indeed,

\[
\frac{ax + b}{(x + 1)(x - 2)} = a \cdot \left[ \frac{1}{(x + 1)(x - 2)} \right] + b \cdot \left[ \frac{x}{(x + 1)(x - 2)} \right].
\]

In a linear-algebraic sense, we would say the functions of the form (7.52) form a 2-dimensional space (or 2-dimensional vector space), because to specify such a function requires two constants, \( a \) and \( b \).

3. Now instead consider another 2-dimensional space of functions given by linear combinations of the form

\[
\frac{A}{x + 1} + \frac{B}{x - 2} = A \cdot \left[ \frac{1}{x + 1} \right] + B \cdot \left[ \frac{1}{x - 2} \right].
\]

(7.53)

4. The functions \( \frac{1}{x + 1} \) and \( \frac{1}{x - 2} \) are indeed also linearly independent, so the set of all functions of the form (7.53) also forms a 2-dimensional vector space; to specify any such function requires specifying two constants, \( A \) and \( B \).

5. Now notice that

\[
\frac{A}{x + 1} + \frac{B}{x - 2} = \frac{A(x - 2) + B(x + 1)}{(x + 1)(x - 2)} = \frac{(A + B)x + (-2A + B)}{(x + 1)(x - 2)},
\]

which is of form (7.52) with \( a = A + B \) and \( b = -2A + B \). In other words, any function of the form (7.53) can also be written in the form (7.52).

6. This tells us that the two-dimensional space of functions \( \frac{A}{x + 1} + \frac{B}{x - 2} \) is contained in the two-dimensional space of functions \( \frac{ax + b}{(x + 1)(x - 2)} \). It is a fact of linear algebra that the only way for a two-dimensional space to be contained in another two-dimensional space is for them to be the same spaces. (Think about a plane being contained in another plane, and realize that they must then be the same plane.)

7. Finally, since (by 6 above) the space of all functions of the form \( \frac{ax + b}{(x + 1)(x - 2)} \) is the same as the space of all functions of the form \( \frac{A}{x + 1} + \frac{B}{x - 2} \), it follows that any function of the form \( \frac{ax + b}{(x + 1)(x - 2)} \) can also be written in the form \( \frac{A}{x + 1} + \frac{B}{x - 2} \).

---

\(^{24}\)We call functions \( f_1, f_2, \ldots, f_n \) linearly independent if and only if it is impossible to write any of these as linear combinations of the others. In other words, we exclude cases where there exist constants \( a_1, \ldots, a_k, a_{k+1}, \ldots, a_n \in \mathbb{R} \) such that

\[
f_k = a_1f_1 + a_2f_2 + \cdots + a_{k-1}f_{k-1} + a_{k+1}f_{k+1} + \cdots + a_nf_n, \quad \text{i.e.,}
\]

\[
(\forall x)[f_k(x) = a_1f_1(x) + a_2f_2(x) + \cdots + a_{k-1}f_{k-1}(x) + a_{k+1}f_{k+1}(x) + \cdots + a_nf_n(x)].
\]
Notice that functions of the form (7.52) are indeed also of the form \(P(x)/Q(x)\) where \(P\) is of degree less than \(Q\), since the degree of \(P\) is at most 1 (zero if \(a = 0\)) and the degree of \(Q\) is 2.

The argument above guarantees that a PFD such as (7.50) exists. It is more desirable for integration purposes to have form (7.53) than (7.52).

**Example 7.5.3** An argument similar to that of the previous example shows that the following forms give exactly the same set of functions:

\[
\frac{ax^2 + bx + c}{(x + 1)(x + 2)(x + 3)} = \frac{A}{x + 1} + \frac{B}{x + 2} + \frac{C}{x + 3}. \tag{7.54}
\]

Of course \(a, b, c\) are likely to differ from \(A, B, C\). Here the underlying sets of linearly independent functions are, respectively,

\[
U = \left\{ \frac{x^2}{(x + 1)(x + 2)(x + 3)}, \frac{x}{(x + 1)(x + 2)(x + 3)}, \frac{1}{(x + 1)(x + 2)(x + 3)} \right\},
\]

\[
V = \left\{ \frac{1}{x + 1}, \frac{1}{x + 2}, \frac{1}{x + 3} \right\}.
\]

Both sets of vectors span\(^{25}\) 3-dimensional spaces. It is not hard to see that functions on the right-hand side of (7.54) can also be in the form on the left. Indeed, if we combine the fractions on the right, we get

\[
\frac{A}{x + 1} + \frac{B}{x + 2} + \frac{C}{x + 3} = \frac{A(x + 2)(x + 3) + B(x + 1)(x + 3) + C(x + 1)(x + 2)}{(x + 1)(x + 2)(x + 3)},
\]

which gives us a polynomial in the numerator with degree at most 2, as on the left-hand side of (7.54). Here we have the span of \(V\) contained in the span of \(U\), though they are both 3-dimensional spaces. Thus they must be the same spaces (we have one 3-dimensional space inside of another, so they must be the same!), so in fact, anything written like the right-hand side of (7.54) can be written like the right-hand side. (In a later subsection we will show how to find \(A, B, C\) if we are given \(a, b, c\).)

It should be clear that integrating a function written like the right-hand side of (7.54) is likely much simpler than integrating one in the form on the left.

**Example 7.5.4** Next we argue that the following forms describe the same (space of) functions:

\[
\frac{ax^2 + bx + c}{(x + 1)^3} = \frac{A}{x + 1} + \frac{B}{(x + 1)^2} + \frac{C}{(x + 1)^3}. \tag{7.55}
\]

The underlying sets of linearly independent functions are, respectively, \(\left\{ \frac{x^2}{(x + 1)^3}, \frac{x}{(x + 1)^3}, \frac{1}{(x + 1)^3} \right\}\) and \(\left\{ \frac{1}{x + 1}, \frac{1}{x + 2}, \frac{1}{x + 3} \right\}\), both spanning 3-dimensional spaces. To show they are the same spaces, we note that

\[
\frac{A}{x + 1} + \frac{B}{(x + 1)^2} + \frac{C}{(x + 1)^3} = \frac{A(x + 1)^2 + B(x + 1) + C}{(x + 1)^3} = \frac{Ax^2 + (14A + B)x + (49A + C)}{(x + 1)^3},
\]

\(^{25}\)The noun form of span has a precise technical meaning. The span of “vectors” \(v_1, v_2, \ldots, v_n\) is the set of all possible linear combinations of those vectors. Thus for example

\[
\text{Span} \left\{ \frac{1}{x + 1}, \frac{1}{x + 2}, \frac{1}{x + 3} \right\} = \left\{ a \cdot \frac{1}{x + 1} + b \cdot \frac{1}{x + 2} + c \cdot \frac{1}{x + 3} \mid a, b, c \in \mathbb{R} \right\}.
\]

We would then say that the functions (vectors, in the linear algebra sense) \(\frac{1}{x + 1}, \frac{1}{x + 2}, \frac{1}{x + 3}\), taken together, span the set described above.
thus simplifying to the form on the left-hand side of (7.55), with \( a = A, \ b = 14A + B, \) and \( c = 49A + C, \) all constants. Arguing as before, the underlying sets of linearly independent functions must span the same 3-dimensional spaces, so anything of the form on the left-hand side of (7.55) can be written also in the form on the right-hand side.

**Example 7.5.5** For our last example, we claim the following forms give the same functions:

\[
\frac{ax^4 + bx^3 + cx^2 + dx + e}{x^3(x^2 + 1)} = \frac{Ax}{x^2} + \frac{B}{x^3} + \frac{C}{x^3} + \frac{Dx + E}{x^2 + 1}.
\]  

(7.56)

Here the spanning sets of linearly independent functions are respectively

\[
U = \left\{ \frac{x^4}{x^3(x^2 + 1)}, \frac{x^3}{x^3(x^2 + 1)}, \frac{x^2}{x^3(x^2 + 1)}, \frac{x}{x^3(x^2 + 1)}, \frac{1}{x^3(x^2 + 1)} \right\},
\]

\[
V = \left\{ \frac{1}{x}, \frac{1}{x^2}, \frac{1}{x^3}, \frac{x}{x^2 + 1}, \frac{1}{x^2 + 1} \right\}.
\]

Again, the form on the right of (7.56) can be rewritten as below, simplifying into the form on the left-hand side of (7.56) as

\[
\frac{Ax^2(x^2 + 1) + Bx(x^2 + 1) + C(x^2 + 1) + (Dx + E)(x^2 + 1)}{x^3(x^2 + 1)}
\]

\[
= \frac{Ax^4 + (B + D)x^3 + (A + C + E)x^2 + (B + D)x + (C + E)}{x^3(x^2 + 1)}.
\]  

(7.57)

(Note that this has numerator of degree at most 4.) So anything that is a linear combination of functions of \( V \) can also be written as a linear combination of functions of \( U \). Because both sides of (7.56) must therefore describe exactly the same functions in a 5-dimensional vector space, it follows that anything written in the form on the left-hand side of (7.56) can also be written in the form on the right-hand side.

Note that it is a little tricky to find the new coefficients of \( 1, x, x^2, x^3, \) etc., in the new numerator of (7.57). They all follow from gathering what we would get if we multiplied out the original numerator, but it would be easy to miss a term. (The reader should verify the computation (7.57).)

In the next subsection we generalize the logic of the examples above to write exact rules for the form of a PFD based upon the original fraction’s denominator. Then in Subsection 7.5.3 we look at three methods of finding the coefficients, \( A, \ B, \ C, \) etc., of the PFD expansion, and immediately apply the methods to problems of computing integrals of such functions.

---

26Of course we can also simplify the functions in \( U \) as the five linearly independent functions

\[
U = \left\{ \frac{x}{x^2 + 1}, \frac{1}{x^2 + 1}, \frac{1}{x(x^2 + 1)}, \frac{1}{x^2(x^2 + 1)}, \frac{1}{x^3(x^2 + 1)} \right\}.
\]
7.5.2 Partial Fraction Decompositions: The Rules

It is a fact of algebra (corollary to the Fundamental Theorem of Algebra) that any polynomial with real coefficients can be factored uniquely—up to rearrangement of multiplicative constants—into powers of linear terms \((ax + b)^n\) and powers of irreducible quadratic\(^{27}\) terms \((ax^2 + bx + c)^m\) with real coefficients. So for instance,

\[
x^3 - x^2 + x - 1 = (x - 1)(x^2 + 1),
\]

and there is no other way to factor it, except for instance \(2(x - 1)(\frac{1}{2}x^2 + \frac{1}{2})\), etc. With that in mind, the rules for partial fraction decompositions follow.

First, we are given a rational function \(\frac{P(x)}{Q(x)}\), where \(P\) and \(Q\) are polynomials.

0. In 1 and 2 below we assume \(\deg P < \deg Q\). If the \(\deg P \geq \deg Q\), then first we apply polynomial long division to achieve

\[
\frac{P(x)}{Q(x)} = p(x) + \frac{r(x)}{Q(x)},
\]

where \(p, r\) are polynomials and \(\deg r < \deg Q\). Then the following rules apply to \(\frac{r(x)}{Q(x)}\).

1. If \((ax+b)^n\), where \(a \neq 0\) occurs as a factor in \(Q(x)\), then the partial fraction decomposition (PFD) of \(\frac{P(x)}{Q(x)}\) will contain terms

\[
\frac{A_1}{ax + b} + \frac{A_2}{(ax + b)^2} + \cdots + \frac{A_n}{(ax + b)^n},
\]

2. If \(ax^2 + bx + c\) is an irreducible quadratic, and \((ax^2 + bx + c)^m\) occurs as a factor in \(Q(x)\), then the PFD of \(\frac{P(x)}{Q(x)}\) will contain terms

\[
\frac{A_1x + B_1}{ax^2 + bx + c} + \frac{A_2x + B_2}{(ax^2 + bx + c)^2} + \cdots + \frac{A_mx + B_m}{(ax^2 + bx + c)^m}.
\]

The application of these rules can be somewhat confusing at first, so we will look at several examples before proceeding to solve for the coefficients \(A_1, B_1\), etc. For the second case, we will mostly be interested in irreducible quadratics of the form \(x^2 + k^2\), where \(k > 0\). Note that we will usually use letters without subscripts, such as \(A, B, C\), and so on for our PFD coefficients (to be found later).

**Example 7.5.6** Write the partial fraction decompositions for the given rational functions:

\[
\frac{3x^5 - 11x^3 + 15x - 2}{(x + 1)^2(x - 3)^4} = \frac{A}{x + 1} + \frac{B}{(x + 1)^2} + \frac{C}{x - 3} + \frac{D}{(x - 3)^2} + \frac{E}{(x - 3)^3} + \frac{F}{(x - 3)^4}.
\]

\[
\frac{1}{(x + 1)^2(x - 3)^4} = \frac{A}{x + 1} + \frac{B}{(x + 1)^2} + \frac{C}{x - 3} + \frac{D}{(x - 3)^2} + \frac{E}{(x - 3)^3} + \frac{F}{(x - 3)^4}.
\]

\(^{27}\)It is easy to see when a quadratic term is “irreducible over the real numbers,” meaning we cannot write it as \((ex + f)(gx + h)\), where \(e, f, g, h \in \mathbb{R}\), the latter being equivalent to there being real numbers \(\alpha, \beta\) such that the polynomial is zero there (i.e., \(ae - f\beta = -h\)). Using the quadratic formula, it is plain that no such real solutions to the quadratic being zero occur if and only if \(b^2 - 4ac < 0\) (i.e., when the term under the radical in the quadratic formula is negative).
In both cases, we had a polynomial of degree less than 6 divided by a polynomial of degree 6, so the exception to “Rule 0” is not invoked. We also had two factors of \(x + 1\) in the denominator, so we needed a constant over the first power, plus another constant over the second power of \(x + 1\). With \((x - 3)\) appearing to the third power in the denominator, we needed a constant over each of the first, second, and third powers of \(x - 3\). (Of course the choice of constants \(A, B, C, D, E\) will be different for these two functions above, but the abstract form of their PFD’s is the same.)

We do not want to be redundant in our PFD’s, so if \(Q(x)\) contains the factor \((x - 3)^4\) but does not contain \((x - 3)^5\), for instance, we require constants divided by \((x - 3), (x - 3)^2, (x - 3)^3\) and \((x - 3)^4\) (but not \((x - 3)^5\)). Now one could say that such a \(Q(x)\) also contains \((x - 3)^2\), but we do not then require in our PFD constants divided by \((x - 3)\) and \((x - 3)^2\) \(\text{again}\), since these are already taken care of by those required by the factor \((x - 3)^4\) in \(Q(x)\).

To rephrase the rules in light of the last paragraph, if exactly \(n\) factors of \((ax + b)\) appear in \(Q(x)\), then the PFD contains terms \(\frac{A}{x + a} + \cdots + \frac{A}{(ax + b)^n}\). If exactly \(m\) factors of \((ax^2 + bx + c)\) appear, with \(b^2 - 4ac < 0\), then the PFD contains terms \(\frac{A}{ax^2 + bx + c} + \cdots + \frac{A}{(ax^2 + bx + c)^m}\).

**Example 7.5.7** Here are more PFD expansion forms. (We do not solve for the coefficients yet.)

\[
\begin{align*}
\frac{x^4 + x + 1}{x^3(x^2 + 9)^2} &= \frac{A}{x} + \frac{B}{x^2} + \frac{C}{x^3} + \frac{Dx + E}{x^2 + 9} + \frac{Fx + G}{(x^2 + 9)^2}. \\
\frac{1}{(x^2 + 1)(x^2 + 4)} &= \frac{Ax + B}{x^2 + 1} + \frac{Cx + D}{x^2 + 4}. \\
\frac{x^5 - 8}{x(2x + 1)^2(9 - x)^3} &= \frac{A}{x} + \frac{B}{2x + 1} + \frac{C}{(2x + 1)^2} + \frac{D}{9 - x} + \frac{E}{(9 - x)^2} + \frac{F}{(9 - x)^3}. \\
\frac{2}{x^2 - 5} &= \frac{2}{(x - \sqrt{5})(x + \sqrt{5})} = \frac{A}{x - \sqrt{5}} + \frac{B}{x + \sqrt{5}}. \\
\frac{1}{x^4 - 1} &= \frac{1}{(x^2 - 1)(x^2 + 1)} = \frac{1}{(x + 1)(x - 1)(x^2 + 1)} = \frac{A}{x + 1} + \frac{B}{x - 1} + \frac{Cx + D}{x^2 + 1}. \\
\frac{1}{x(x - 3)(3x - 9)} &= \frac{1}{3x(x - 3)^2} = \frac{A}{x} + \frac{B}{x - 3} + \frac{C}{(x - 3)^2}.
\end{align*}
\]

The last two PFD’s above required us to factor the denominators before we started to implement the rules. Note that we must be careful to identify factors which are truly distinct. Consider the following:

\[
\frac{1}{x(x - 3)(3x - 9)} = \frac{1}{3x(x - 3)^2} = \frac{A}{x} + \frac{B}{x - 3} + \frac{C}{(x - 3)^2}.
\]

The factors \((x - 3)\) and \((3x - 9)\) were not really distinct factors, but were constant multiples of each other. If we do not notice this we will find ourselves attempting a PFD with \(\frac{A}{x} + \frac{B}{x - 3} + \frac{C}{(x - 3)^2}\), but the “B” and “C” terms are not independent, so we will miss one dimension of possibilities for our PFD. Note also that the factor \(\frac{1}{x}\) can be included in the first PFD term (i.e., we could replace \(\frac{1}{x}\) with \(\frac{1}{x^2}\), of course giving a different “A”), or its influence absorbed into the \(A, B\) and \(C\) terms. We will usually opt for the latter approach (as we did above).

**7.5.3 Finding the Coefficients for PFD’s**

There are two main methods, and one auxiliary method, for finding the coefficients \(A, B, \text{etc.}\), for a PFD. The most efficient method for a particular PFD is usually a mixture of the two main
methods; perhaps the first method can be used to find \( A \) and \( C \), and the second to find \( B \), for instance. Efficiently computing the coefficients is thus somewhat of an art.\(^{28}\)

The methods are based upon some properties of polynomials. Consider two polynomials

\[
\begin{align*}
  f(x) &= a_nx^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0, \\
  g(x) &= b_mx^m + b_{m-1}x^{m-1} + \cdots + b_2x^2 + b_1x + b_0.
\end{align*}
\]

The statement that \( f(x) = g(x) \) is an “equality of polynomials,” i.e., that \( f(x) \) and \( g(x) \) are the same polynomial is equivalent to each of the following two conditions (separately):\(^{29}\)

1. \((\forall x \in \mathbb{R})[f(x) = g(x)]\). In other words, \( f \) and \( g \) are the same functions.

2. \((\forall i \in \{1, 2, \ldots, \max\{m, n\}\})[a_i = b_i]\), that is, all the coefficients are the same. (Note that it is possible, for instance, that \( m < n \), in which case we just take \( b_{m+1}, \ldots, b_n = 0 \).

Furthermore, if \( f(x) \) and \( g(x) \) are the same polynomials, then \( f'(x) = g'(x) \), \( f''(x) = g''(x) \), \( f'''(x) = g'''(x) \cdots \), in the sense of being the same polynomials, and so 1 and 2 from above apply to these derivatives as well.

Our first method will exploit 1, the second 2, and the auxiliary method will make use of final observation about derivatives. The first method essentially “probes” the two polynomials at different points, usually chosen strategically, to get some quick information out of a polynomial equality, and is often called an “evaluation method.” The second method is often referred to as “comparing coefficients,” and can also be useful for finding quick information. The auxiliary method exploits the fact that the first methods can be applied to the derivatives (of any order) of \( f \) and \( g \) to get further information quickly.

**Example 7.5.8** Compute the integral \( \int \frac{1}{x^2 - 5x + 6} \, dx \).

**Solution:** Here we have a degree-0 polynomial divided by a degree-2 polynomial, so the PFD rules apply. Now one usually writes the PFD form of the integrand, complete with the unknown coefficients, before proceeding to the methods of computing the coefficients. In other words, our first step would be to write:

\[
\int \frac{1}{x^2 - 5x + 6} \, dx = \int \frac{1}{(x-2)(x-3)} \, dx = \int \left[ \frac{A}{x-2} + \frac{B}{x-3} \right] \, dx.
\]

The next two lines can be skipped with practice, though the first time one works this section they are worth writing so the mechanics of the method can be understood and reinforced. First we write the algebraic step (PFD) which was contained in the rewriting of the integrands above:

\[
\frac{1}{(x-2)(x-3)} = \frac{A}{x-2} + \frac{B}{x-3}.
\]

This came from the fact that we have \( x - 2 \) as a factor in the denominator, but only once, and the same for \( x - 3 \). Next we multiply both sides by the denominator on the left:

\[
(x-2)(x-3) \left[ \frac{1}{(x-2)(x-3)} \right] = (x-2)(x-3) \left[ \frac{A}{x-2} + \frac{B}{x-3} \right].
\]

---

\(^{28}\)In fact either method is—strictly speaking—sufficient, and indeed there are textbooks which teach only one or the other method. However, trying to fit a particularly complicated PFD into any single method will make for much more difficult computations than are necessary. That said, for computer programming one would likely choose one method and let the computer calculate the coefficients by “brute force.”

\(^{29}\)Some texts use the notation \( f(x) \equiv g(x) \), read “\( f(x) \) is identically equal to \( g(x) \).” In other words, \( f(x) \) and \( g(x) \) are the same functions. The word *identically* is in the spirit of, for instance, trigonometric identities, so one could write for example \( \sin^2 \theta + \cos^2 \theta \equiv 1 \).
On the left, the whole denominator cancels and we have the numerator of the original fraction. (This always happens.) On the right we have to distribute the \((x-2)(x-3)\) across the sum in the brackets. For the “\(A\)” term the \((x-2)\) cancels, while for the “\(B\)” term the \((x-3)\) cancels, giving us an equality of polynomials:

\[
1 = A(x-3) + B(x-2). \tag{7.58}
\]

Because this is an equality of polynomials \((f(x) = g(x)\) where \(f(x) = 1\) and \(g(x) = A(x-3) + B(x-2)\)), it must be true for any \(x \in \mathbb{R}\). Now we choose two values of \(x\) strategically.

\[
\begin{align*}
\text{when } x = 3 & : 1 = A(3-3) + B(3-2) \implies 1 = B \\
\text{when } x = 2 & : 1 = A(2-3) + B(2-2) \\
& \implies 1 = -A \implies A = -1
\end{align*}
\]

Now we summarize what we have so far, and compute the desired integral:

\[
\int \frac{1}{x^2 - 5x + 6} \, dx = \int \left[ \frac{-1}{x-2} + \frac{1}{x-3} \right] \, dx \\
= -\ln|x-2| + \ln|x-3| + C \\
= \ln \left| \frac{x-3}{x-2} \right| + C.
\]

(The last step is not necessary, but for reasons of style many textbooks combine logarithmic terms into a single logarithm.)

Because (7.58) was an equality of polynomials (meaning the polynomial on the left is the same polynomial as that on the right\(^{30}\)), we could substitute any number for \(x\) in (7.58) and still have a true statement. Fortunately, there were choices which could eliminate an unknown, leaving an equation in the other unknown which is easily solved.

The second method for finding \(A\) and \(B\) (not preferred here but not terribly difficult here either) is to look at the coefficients of the polynomials on the left-hand side and right-hand side of (7.58), and realize that the coefficients of the various powers of \(x\) must agree for these to be the same polynomials. Though perhaps not necessary for this simple case, one sometimes expands the right-hand side and collects like terms

\[
1 = (A + B)x + (-3A - 2B).
\]

From this or just reading from (7.58), we can in turn set equal the coefficients of the \(x^1\) terms and the constant (some say \(x^0\)) terms to get the following system of two equations in two unknowns:

\[
\begin{cases}
0 = A + B \\
1 = -3A - 2B.
\end{cases} \tag{7.59}
\]

The first equation came from the fact that there is no \(x^1\)-term on the left-hand side of (7.59), or alternatively, the \(x^1\)-term is \(0x^1\) on the left. To solve such a system one might add three

\(^{30}\)It should be pointed out that when we write a PFD, for instance

\[
\frac{1}{(x-2)(x-3)} = \frac{A}{x-2} + \frac{B}{x-3}
\]

we mean that these are the same functions as well, so once we find \(A\) and \(B\), the right-hand side would simplify to become the left-hand side. To find \(A\) and \(B\) we actually solve the polynomial equality (7.58) for \(A\) and \(B\).

Note also the distinction between “equations” such as \(2x - 1 = 5\), which is true only for \(x = 3\), and “equalities” such as \((x+1)^2 = x^2 + 2x + 1\), true for all \(x\), meaning the function \((x+1)^2\) is the same as the function \(x^2 + 2x + 1\).
times the first equation to the second, to get \( B = 1 \), and use that information in the first to get \( A = -1 \), as before.

Whenever the denominator of our function \( P(x)/Q(x) \) has a linear factor \((ax + b)\), evaluating the associated polynomial equality—such as (7.58)—at that \( x \)-value which makes this linear factor zero (namely \( x = -b/a \)) will quickly yield one of the coefficients, since all but one term in the polynomial equation will have \((ax + b)\) as a factor, and therefore vanish at \( x = -b/a \). Thus this first method should always be employed to find that coefficient if the denominator \( Q \) has a linear factor. If the denominator has all linear terms to the first power, then this “evaluation” method will quickly yield all coefficients.

**Example 7.5.9** Compute \( \int \frac{2x^2 - 3x + 2}{x(x+5)(2x+1)} \, dx \).

**Solution:** It is important to notice that the numerator is degree 2, and the denominator degree 3, so the PFD rules do apply.

\[
\int \frac{2x^2 - 3x + 2}{x(x+5)(2x+1)} \, dx = \int \left[ \frac{A}{x} + \frac{B}{x+5} + \frac{C}{2x+1} \right] \, dx.
\]

Eventually we will cease writing the next two lines, but to be sure we will include them here so that the logic is clear:

\[
\frac{2x^2 - 3x + 2}{x(x+5)(2x+1)} = \frac{A}{x} + \frac{B}{x+5} + \frac{C}{2x+1}
\]

\[
\Rightarrow x(x+5)(2x+1) \left[ \frac{2x^2 - 3x + 2}{x(x+5)(2x+1)} \right] = x(x+5)(2x+1) \left[ \frac{A}{x} + \frac{B}{x+5} + \frac{C}{2x+1} \right]
\]

\[
\Rightarrow 2x^2 - 3x + 2 = A(x+5)(2x+1) + B(x+5) + Cx(x+5).
\]

Into this last line we can now enter values for \( x \) which will quickly yield the coefficients.

\[
x = 0 : \quad 2 = A(5)(1) \quad \Rightarrow \quad A = \frac{2}{5}
\]

\[
x = -5 : \quad 50 + 15 + 2 = B(-5)(-9) \quad \Rightarrow \quad 67 = 45B \quad \Rightarrow \quad B = \frac{67}{45}
\]

\[
x = -\frac{1}{2} : \quad 2 \cdot \frac{1}{4} - 3 \cdot \left( -\frac{1}{2} \right) + 2 = C \left( -\frac{1}{2} \right) \left( -\frac{1}{2} + 5 \right)
\]

\[
\Rightarrow \quad \frac{1}{2} + \frac{3}{2} + 2 = C \left( \frac{9}{2} \right)
\]

\[
\Rightarrow \quad 4 = \frac{9}{4} C \quad \Rightarrow \quad C = \frac{16}{9}
\]

Putting this together with our original integral, we get

\[
\int \frac{2x^2 - 3x + 2}{x(x+5)(2x+1)} \, dx = \int \left[ \frac{2/5}{x} + \frac{67/45}{x+5} + \frac{-16/9}{2x+1} \right] \, dx
\]

\[
= \frac{2}{5} \ln |x| + \frac{67}{45} \ln |x+5| - \frac{16}{9} \cdot \frac{1}{2} \ln |2x+1| + C
\]

\[
= \frac{2}{5} \ln |x| + \frac{67}{45} \ln |x+5| - \frac{8}{9} \ln |2x+1| + C.
\]
Next we look at an example where all factors of $Q(x)$ are linear, but one of these linear factors appears to the second power.

**Example 7.5.10** Compute \[ \int \frac{x + 1}{x^2(x - 5)(x + 4)} \, dx. \]

**Solution:** This time we will describe but omit the explicit multiplication step in the PFD. \[ \int \frac{x + 1}{x^2(x - 5)(x + 4)} \, dx = \int \left[ \frac{A}{x} + \frac{B}{x^2} + \frac{C}{x - 5} + \frac{D}{x + 4} \right] \, dx, \]

where \[ \frac{x + 1}{x^2(x - 5)(x + 4)} = \frac{A}{x} + \frac{B}{x^2} + \frac{C}{x - 5} + \frac{D}{x + 4}. \]

Multiplying by $x^2(x - 5)(x + 4)$ then gives us
\[ x + 1 = Ax(x - 5)(x + 4) + B(x - 5)(x + 4) + Cx^2(x + 4) + Dx^2(x - 5). \] (7.60)

With this equation, choosing $x = 0, 5, -4$ will yield three of the four coefficients.

- $x = 0$: \[ 1 = B(-5)(4) \implies B = -\frac{1}{20} \]
- $x = 5$: \[ 6 = C(5^2)(9) \implies 6 \cdot \frac{3}{5} = C \implies C = \frac{2}{75} \]
- $x = -4$: \[ -3 = D((-4)^2)(-9) \implies -3 = D[-16 \cdot 9] \implies 3 = D \cdot 16 \cdot 3 \implies D = \frac{1}{48} \]

This exhausts the evaluations which give equations in one coefficient. Next we have several methods for finding $A$.

**Method 1.** Compare coefficients. In particular, we look at the highest-order $x$-terms which appear—at least initially—in the polynomial equality, which for (7.60) means the $x^3$-terms. Here we have no $x^3$-terms on the left, and on the right, even without a complete expansion, we can see that the $x^3$-terms will be $A + C + D$. (The middle-order terms are more difficult to read from (7.60).) Fortunately we already know the values of $C$ and $D$, so we have enough information to find $A$: \[ x^3 \text{-term:} \quad 0 = A + C + D \]
\[ \implies 0 = A + \frac{2}{75} + \frac{1}{48} = A + \frac{2 - 48 + 75}{24} = A + \frac{29}{24} \]
\[ \implies -\frac{57}{1200} = A \implies A = -\frac{19}{400} \]

\[31\] The pattern of cancellation, when we multiply the PFD by the original denominator $Q(x)$, should become second nature with a small amount of practice. That said, it is important to remember what we are doing (multiplying by $Q(x)$) to get from the PFD to the polynomial equality, and how the various factors cancel (or do not cancel) in that multiplication.
From this we can complete the integration:

\[
\int \frac{x + 1}{x^2(x - 5)(x + 4)} \, dx = \int \left[ -\frac{19}{400} \ln |x| - \frac{1}{20} \ln |x - 5| + \frac{1}{48} \ln |x + 4| + C \right] dx
\]

Method 2. One can instead evaluate the polynomial equality (7.60) at still another \( x \)-value, though no such value will produce \( A \) alone:

\[ x = 1 : \quad 2 = A(1)(-4)(5) + B(-4)(5) + C(1)^2(5) + D(1)^2(-4). \]

Since we already know \( B, C \) and \( D \), we can insert that information and solve for \( A \).

Method 3. This will be more useful later, but this method (referred to earlier as the auxiliary method) certainly applies. The idea is that we apply \( \frac{d}{dx} \) to both sides of (7.60), which is valid because the left-hand side and right-hand side of (7.60) are the same functions.

In order to use this method, it is useful to recall the generalized product rule. For three functions \( u(x), v(x) \) and \( w(x) \), for instance, we have

\[(uvw)' = u'vw + uv'w + uw'.\]

For reference we recall (7.60), from which we then compute the derivatives. Equation (7.60) reads:

\[ x + 1 = Ax(x - 5)(x + 4) + B(x - 5)(x + 4) + Cx^2(x + 4) + Dx^2(x - 5). \]

\[
\frac{d}{dx} : \quad 1 = A[(1)(x - 5)(x + 4) + x(1)(x + 4) + x(x - 5)(1)]
\]
\[ + B[(1)(x + 4) + (x - 5)(1)]
\[ + C[(2x)(x + 4) + x^2(1)]
\[ + D[(2x)(x - 5) + x^2(1)]. \]

Now when we evaluate this at \( x = 0 \), we see that we get

\[ 1 = A[(1)(-5)(4)] + B[(1)(4) + (-5)(1)] + 0 + 0. \]

This gives us \( 1 = -20A - B \), so then \( A = (1 + B)/(-20) = (19/20)/(-20) = -19/400 \), as before.

A simple principle buried in the third method is the following:

**Theorem 7.5.1** If \((x - a)^m\), where \( m > 1 \) is a factor of a polynomial \( f(x) \), then \((x - a)^{m-1}\) is a factor of \( f'(x) \).

For a proof, we note that \((x - a)^m\) being a factor of \( f(x) \) is equivalent to \( f(x) = (x - a)^mg(x) \), where \( g(x) \) is another polynomial. Thus

\[ f'(x) = (x - a)^mg'(x) + m(x - a)^{m-1}(1)g(x) = (x - a)^{m-1} \left[ (x - a)g'(x) + mg(x) \right]. \]
so indeed \((x - a)^{m-1}\) is a factor of \(f'(x)\).

Now evaluating both sides of (7.60) at \(x = 0\) caused those terms with \(x\) and \(x^2\) factors to vanish, leaving an equation with \(B\) only. When we differentiate (7.60), those terms with \(x^2\) factors still vanish at \(x = 0\)—because one power of \(x\) remains—leaving only the \(B\)-term (as before) and the \(A\)-term (which had a factor \(x\) but not \(x^2\)). Already knowing \(B\), we could solve for \(A\).

A few guidelines for efficiently finding the PFD coefficients should be made at this point.

1. When linear factors are present in \(Q(x)\), it is best to exhaust this method for finding some of the coefficients easily. This means evaluating the relevant polynomial equality at each value for which \(Q(x) = 0\).

2. When those values are exhausted, we should next compare coefficients of the powers of \(x\), particularly the highest power which occurs on the right-hand side.

3. If \((ax + b)^m\) is a factor of \(Q(x)\), where \(m > 1\), and the first two methods fail to get all coefficients, then differentiation of the polynomial equality may yield more coefficients.

4. If there are still coefficients to be found, then further evaluations, differentiations, or coefficient comparisons should be implemented.

Example 7.5.11  Compute

\[
\int \frac{5x^3 - 17x^2 + 19x - 13}{(x + 1)(x - 2)^3} \, dx.
\]

Solution: As usual we start with the PFD, since the denominator has higher degree than the numerator.

\[
\frac{5x^3 - 17x^2 + 19x - 13}{(x + 1)(x - 2)^3} = \frac{A}{x + 1} + \frac{B}{x - 2} + \frac{C}{(x - 2)^2} + \frac{D}{(x - 2)^3}
\]

\[
5x^3 - 17x^2 + 19x - 13 = A(x - 2)^3 + B(x + 1)(x - 2)^2 + C(x + 1)(x - 2) + D(x + 1) \quad (7.61)
\]

\[
x = -1: \quad -5 - 17 - 19 - 13 = A(-27) \quad \Rightarrow -54 = -27A \quad \Rightarrow \boxed{A = 2}
\]

\[
x = 2: \quad 5(8) - 17(4) + 19(2) - 13 = D(3) \quad \Rightarrow -3 = 3D \quad \Rightarrow \boxed{D = -1}
\]

\[
\text{\(x^3\)-term:} \quad 5 = A + B \quad \Rightarrow 5 = 2 + B \quad \Rightarrow \boxed{B = 3}
\]

While we could perform another evaluation \((x = 0\) comes to mind), or look at another coefficient (prone to error), instead we will differentiate (7.61):

\[
15x^2 - 34x + 19 = A[3(x - 2)^2(1)] + B[(1)(x - 2)^2 + (x + 1) \cdot 2(x - 2)(1)] + C[(1)(x - 2) + (x + 1)(1)] + D(1).
\]

Now we evaluated at \(x = 2\):

\[
15(4) - 34(2) + 19 = C(3) + D \quad \Rightarrow 11 = 3C - 1 \quad \Rightarrow \boxed{C = 4}
\]
7.5. PARTIAL FRACTIONS AND INTEGRATION

Thus

\[
\int \frac{5x^3 - 17x^2 + 19x - 13}{(x + 1)(x - 2)^3} \, dx = \int \left( \frac{2}{x + 1} + \frac{3}{x - 2} + \frac{4}{(x - 2)^2} - \frac{1}{(x - 2)^3} \right) \, dx
\]

\[
= 2 \ln |x + 1| + 3 \ln |x - 2| - \frac{4}{x - 2} - \frac{1}{2} \cdot \frac{1}{(x - 2)^2} + C
\]

\[
= \ln |(x + 1)^2(x - 2)^3| - \frac{4}{x - 2} + \frac{1}{2(x - 2)^2} + C.
\]

A quick corollary to our Theorem 7.5.1 is that if \((x - a)^m\) is a factor of a polynomial \(f(x)\), then for \(k < m\) we have \((x - a)^{m-k}\) is a factor of \(f^{(k)}(x) = \frac{d^k}{dx^k} f(x)\). This follows from repeated applications of the theorem, which can be paraphrased as saying that we lose at most one factor of \((x - a)\) for each derivative we take, until we run out of factors of \((x - a)\). If our latest example had \((x - 2)^4\) in the denominator, we could have taken a second derivative of the corresponding polynomial equation, and then those terms with \((x - 2)^3\) or \((x - 2)^4\) will still be zero at \(x = 2\), but the other terms would likely be nonzero.\(^{32}\)

Now we turn our attention to PFD’s where the denominators contain irreducible quadratic factors.\(^{33}\) One problem with such factors is that they are nonzero for any \(x \in \mathbb{R},^{34}\) so the evaluation method’s usefulness is limited in these cases. For such PFD’s, we will need to rely more upon the coefficient comparison method to find our coefficients.

Example 7.5.12 Compute \(\int \frac{4x^3 - 7x^2 + 31x - 38}{x^4 + 13x^2 + 36} \, dx\).

Solution: PFD rules apply since the degree of the numerator is less than that of the denominator. We need to begin by factoring the denominator of the integrand, after which we can write the general form of the PFD.

\[
\int \frac{4x^3 - 7x^2 + 31x - 38}{x^4 + 13x^2 + 36} \, dx = \int \frac{4x^3 - 7x^2 + 31x - 38}{(x^2 + 4)(x^2 + 9)} \, dx = \int \left( \frac{Ax + B}{x^2 + 4} + \frac{Cx + D}{x^2 + 9} \right) \, dx.
\]

Now taking the second equation, we underlying PFD becomes the polynomial equality

\[
4x^3 - 7x^2 + 31x - 38 = (Ax + B)(x^2 + 9) + (Cx + D)(x^2 + 4).
\] (7.62)

Now we look at the coefficients.\(^{35}\)

\[
\begin{align*}
\text{x^3-term:} & \quad 4 = A + C \\
\text{x^2-term:} & \quad -7 = B + D \\
\text{x^1-term:} & \quad 31 = 9A + 4C \\
\text{x^0-term:} & \quad -38 = 9B + 4D.
\end{align*}
\]

\(^{32}\)Note that it is quite possible that \(x - a\) is not a factor of a polynomial \(f(x)\), but is a factor of \(f'(x)\). That is the case when \(x = a\) is a critical point of \(f(x)\). For example, \(x - 1\) is not a factor of \(f(x) = x^2 - 2x + 21\), but is a factor of \(f'(x) = 2x - 2 = 2(x - 1)\).

Note also that the theorem applies to any linear factor \(ax + b\), where \(a \neq 0\), since \(ax + b = a \left( x + \frac{b}{a} \right) \). Thus if \(x^m\) is a factor of a polynomial \(f(x)\), the \(x^{m-1}\) is a factor of \(f'(x)\), etc., as is the case if we replace \(x^m\) with \((ax + b)^m = a^m (x + \frac{b}{a})\).

\(^{34}\)Until the next section, we will not be able to integrate the general case where we have an irreducible quadratic factor to a power greater than 1, with some exceptional cases.

\(^{33}\)Recall that if \(f(x)\) is a polynomial of degree \(\geq 1\), then \(f(a) = 0 \iff (x - a)\) is a factor of \(f(x)\).

\(^{35}\)Note that the constant (“x^0”) term equation is what we would get if we evaluated (7.62) at \(x = 0\). It is easy to see that this is always the case.
Though this looks like (and is) four equations in four unknowns, in fact it “decouples” into two systems, each with two unknowns, since the first and third equations have only $A$ and $C$, and the second and fourth have $B$ and $D$ only. We solve these in turn.

\[
\begin{align*}
4 &= A + C \\
31 &= 9A + 4C
\end{align*}
\]

For the first system, we multiply the first equation by $-9$ and add to the second, to get $-5 = 0A - 5C \implies C = 1$. From that we have the original first equation giving $A = 4 - C = 4 - 1 = 3$.

For the second system, we do the same, that is, multiply the first equation by $-9$ and add to the second, giving $63 - 38 = -5D \implies 25 = -5D \implies -5 = D$. From the original first equation in that system, we then get $B = -7 - D = -7 + 5 = -2$.

Now we compute the integral, noting that it is easier if we break the PFD into four distinct terms:

\[
\int \frac{4x^3 - 7x^2 + 31x - 38}{(x^2 + 4)(x^2 + 9)} \, dx = \int \left[ \frac{3x}{x^2 + 4} - \frac{2}{x^2 + 4} + \frac{x}{x^2 + 9} - \frac{5}{x^2 + 9} \right] \, dx
\]

\[
= \frac{3}{2} \ln(x^2 + 4) - \frac{2}{2} \tan^{-1} \frac{x}{2} + \frac{1}{2} \ln(x^2 + 9) - \frac{5}{3} \tan^{-1} \frac{x}{3} + C
\]

\[
= \ln \sqrt{(x^2 + 4)(x^2 + 9)} - \tan^{-1} \frac{x}{2} - \frac{5}{3} \tan^{-1} \frac{x}{3} + C.
\]

In the example above, we used the following common integration formula, which is particularly useful in problems encountered in this section. It is derivable with the usual substitution methods, and not too difficult to verify by differentiation. The formula is the following:

\[
\int \frac{1}{x^2 + a^2} \, dx = \frac{1}{a} \tan^{-1} \frac{x}{a} + C.
\]  

(7.63)

We also used

\[
\int \frac{x}{x^2 + k^2} \, dx = \frac{1}{2} \ln(x^2 + k^2) + C,
\]

assuming $k \neq 0$. Note that we do not need absolute values inside the logarithm since $x^2 + k^2 \geq k^2 > 0$.

When we have irreducible quadratic factors in the denominator $Q(x)$, it is likely that we will need to compare coefficients. After all, there are no real numbers which will make all but one of those coefficients vanish. (We can make two vanish with $x = 0$, but that still leaves two.) If linear terms are also present, however, the evaluation method will yield one or more of the coefficients quickly.

**Example 7.5.13** Compute \( \int \frac{12x^4 + 190x^2 + 13x - 6}{(2x - 1)(x^2 + 16)} \, dx \).

**Solution:** First we note that the numerator has degree which is not less than the denominator, so we must use long division. To do so we need to expand the denominator: \((2x - 1)(x^2 + 16) = 2x^3 - x^2 + 32x - 16\).

\[
12x^2 - 14x + 31 = (A)(x^2 + 9) + (Ax + B)(2x) + (C)(x^2 + 4) + (Cx + D)(2x),
\]

which, when we consider the datum $x = 0$ gives $31 = 9A + 4C$. Both of these we had before. More derivatives, evaluated at $x = 0$, give multiples of the other two equations in our system (four equations in four unknowns).

\[
36\text{Or something equivalent to comparing coefficients. For instance, } x = 0 \text{ gives } -38 = 9B + 4D, \text{ and one derivative of (7.62) gives}
\]

\[
12x^2 - 14x + 31 = (A)(x^2 + 9) + (Ax + B)(2x) + (C)(x^2 + 4) + (Cx + D)(2x),
\]

which, when we consider the datum $x = 0$ gives $31 = 9A + 4C$. Both of these we had before. More derivatives, evaluated at $x = 0$, give multiples of the other two equations in our system (four equations in four unknowns).
7.5. PARTIAL FRACTIONS AND INTEGRATION

Now through polynomial long division we get
\[
\frac{12x^4 + 190x^3 + 13x - 6}{(2x - 1)(x^2 + 16)} = \frac{12x^4 + 190x^3 + 13x - 6}{2x^3 - x^2 + 32x - 16} = 6x + 3 + \frac{x^2 + 13x + 42}{2x^3 - x^2 + 32x - 16}. \tag{7.64}
\]

Refactoring our denominator, our integral now becomes
\[
\int \left[ 6x + 3 + \frac{x^2 + 13x + 42}{(2x - 1)(x^2 + 16)} \right] dx = \int \left[ 6x + 3 + \frac{A}{2x - 1} + \frac{Bx + C}{x^2 + 16} \right] dx.
\]

The first two terms are easy enough. For our PFD, we need only concern ourselves with the remaining fraction:
\[
\frac{x^2 + 13x + 42}{(2x - 1)(x^2 + 16)} = \frac{A}{2x - 1} + \frac{Bx + C}{x^2 + 16}.
\]

The corresponding polynomial equation is then
\[
x^2 + 13x + 42 = A(x^2 + 16) + (Bx + C)(2x - 1). \tag{7.65}
\]

We begin with an evaluation, followed by a coefficient comparison.

\[
x = \frac{1}{2} : \quad \frac{1}{4} + \frac{13}{2} + 42 = A \left( \frac{1}{4} + 16 \right)
\]
\[
\quad \Rightarrow \quad 1 + 26 + 16 = 65 \frac{A}{4}
\]
\[
\quad \Rightarrow \quad 195 = 65A \quad \Rightarrow \quad A = 3
\]

\[
x^2\text{-term :} \quad 1 = A + 2B
\]
\[
\quad \Rightarrow \quad 1 = 3 + 2B \quad \Rightarrow \quad B = -1.
\]

Perhaps the simplest next step is to find \(C\) by evaluation of (7.65) at, say, \(x = 0\):
\[
x = 0 : \quad 42 = 16A - C
\]
\[
\quad \Rightarrow \quad 42 = 16(3) - C
\]
\[
\quad \Rightarrow \quad C = 16(3) - 33 = 48 - 16 \quad \Rightarrow \quad C = 6.
\]

Thus our original integral, including the polynomial terms, becomes
\[
\int \frac{12x^4 + 190x^3 + 13x - 6}{(2x - 1)(x^2 + 16)} \, dx = \int \left[ 6x + 3 + \frac{3}{2x - 1} - \frac{x}{x^2 + 16} + \frac{6}{x^2 + 16} \right] dx
\]
\[
= 3x^2 + 3x + \frac{3}{2} \ln |2x - 1| - \frac{1}{2} \ln (x^2 + 16) + \frac{6}{4} \tan^{-1} \frac{x}{4} + C
\]
\[
= 3x(x + 1) + 3 \ln \sqrt{|2x - 1|} - \ln \sqrt{x^2 + 16} + \frac{3}{2} \tan^{-1} \frac{x}{4} + C.
\]

The second from the last line was complete; the last line just gives some alternative styles for the particular terms.

Of course with any new technique, we have to be sure that we do not neglect the earlier methods.
7.6 Miscellaneous Methods

In this section we will use completing the square, and other methods to rewrite several types of integrals into forms where we can more easily use either partial fractions or trigonometric substitution. We will also look at examples where a substitution will bring us to such forms. Finally, we will consider the use of integration tables, which can be found in numerous publications, but which require some sophistication to be used properly.
Chapter 8

Applications of Definite Integrals
I: General Arguments

Here we look at many further quantities which give rise to antidifferentiation.

Most modern calculus textbooks contain numerous excellent examples of applications of definite integrals. Indeed, most examples which are likely to be seen in further studies of the physical sciences can trace back to calculus textbook-type problems.

Here we will make an attempt to accomplish the presentations of the usual topics, and others. We will also do the following:

1. Re-introduce the notion of infinitesimals into the physical analysis of these problems. This was a traditional approach which has fallen out of favor. While we will refer back to the Riemann sums for each case, that will be more for a “spot-check” of the reasoning behind the infinitesimals, and perhaps some proofs. However, the first introduction to most topics will be through infinitesimals. Besides, they make for prettier pictures!

Students are unlikely to be inspired by the Riemann sum proofs, which are often used to show that the quantity represented by the integral has the integrand as derivative. The proofs are technical, and leave a student to believe he would never come up with it himself. The differentials “cut to the chase,” and can be proved later after guessed, rather than derived from Riemann sums.

2. Finite Riemann sums used for numerical approximations of the quantities involved. This contains all the intuition (short of the proofs) contained in the most textbook developments of these integrals.

3. Explanations of why some guesses for the infinitesimals will not work.

4. Explanation of how to tell—by sight—if a particular differential is correct, and whether the Riemann sum form will actually converge to the desired quantity as the partition is refined.

8.1 Riemann Sums and Approximations of Cumulative Quantities

Suppose we had some data on the velocity of an object, and we wish to approximate its net displacement over a time interval. Suppose the data we have is the following:
8.2 Other Complications

Consider cases where the function is only piecewise continuous, and try to recover a continuous antiderivative. Consider, perhaps, some cases from electricity and magnetism regarding fields or potentials across interfaces.

Consider two “antiderivatives” of

\[ f(x) = \begin{cases} 
1, & x > 0, \\
-1, & x < 0.
\end{cases} \]

One is continuous, the other not.

Can always recover a continuous antiderivative from a piecewise continuous function, so long as we don’t have vertical asymptotes, for instance.

8.3 Approximation Methods for Definite Integrals

8.4 Physical Applications of Definite Integrals

Recall that anytime we have a continuous function \( f(x) \) on a closed interval \([a,b]\), the definite integral

\[
\int_a^b f(x) \, dx = \lim_{\max\{\Delta x_i\} \to 0} \left( \sum_{i=1}^{n} f(x_i^*) \, \Delta x_i \right) = F(b) - F(a), \quad (8.1)
\]

where \( F(x) \) is continuous on \([a,b]\), and \( F'(x) = f(x) \) on \((a,b)\). The first of the two equations above is the definition of the antiderivative, and the other is the Fundamental Theorem of Calculus.

While the above is eventually intuitive when one considers the context of the signed area problem—with the percentage error of each nonzero area’s approximating rectangle shrinking to zero as \( \max\{\Delta x_i\} \to 0 \) (and therefore \( n \to \infty \)—there are other problems which become equally intuitive but for different contexts.

Indeed anytime a quantity can be better and better approximated by Riemann sums, in such a way that the percentage error of each step decreases to zero as the size of that step does as well, we can argue that the quantity in question should be the limit of the Riemann sums—as the maximum interval length approaches zero—and therefore can be represented by the definite integral, which we can compute with the Fundamental Theorem of Calculus. In other words, if we must accept the first equality below, we must accept the second:

\[
\text{desired quantity} = \lim_{\max\{\Delta x_i\} \to 0} \left( \sum_{i=1}^{\infty} q(x_i^*) \, \Delta x_i \right) = \int_a^b q(x) \, dx. \quad (8.2)
\]

This in turn will then equal \( Q(b) - Q(a) \), where \( Q' = q \). There are countless examples where the integral calculus lets us compute such “desired” quantities in this way. The signed area between a curve and the \( x \)-axis is one example, as are other geometric examples, but those are only the beginning.

8.4.1 Geometric Applications

Some of the simplest examples come from geometry.
Example 8.4.1 It was known in ancient times that one could take any circle, divide the circumference by the diameter, and always get the same number which, of course, we know today as $\pi \approx 3.14159265$. We usually write the circumference formula using the radius instead of the diameter, writing for instance $s = 2\pi r$. From this we can use an integration argument to find a formula for the area of a circle. We do this by setting up an approximating Riemann sum.

In this computation, we will let $r$ be the radius of the circle, but we will break up the total area into areas of concentric rings with outer and inner radii, such that the $i$th ring will have inner radius $\rho_{i-1}$ and outer radius $\rho_i$, with $i = 1, 2, \cdots, n$.\footnote{It is customary to substitute the Greek version of a letter for the Latin version in cases such as this. We cannot use $r$ for a variable radius when we are simultaneously using $r$ for the radius of the circle, so we instead use $\rho$ for the variable radius. Using the Greek version respects the similarities in the quantities while noting that they are different. The Greek version is often used as well to signify a “temporary” variable introduced into a computation but not present in the final answer.}
Chapter 9

Improper Integrals and Advanced Limit Techniques

In this chapter we will develop some more advanced techniques for computing limits. In the first section, we will look at many determinate forms. In the next section we compute many limits of indeterminate forms 0/0 or \( \infty/\infty \) using a very useful—but not universally applicable—technique known as L'Hôpital's Rule. Next we look at other indeterminate forms, revisiting \( \infty - \infty \) and 0 \( \cdot \) \( \infty \), and introducing for the first time \( (0^+)^0, \infty^0 \), and \( (1)\infty \) using a simple strategy to find related (but not equivalent) fractional forms on which to apply L'Hôpital's Rule, and how to use the conclusions there to determine the original limit. Finally we apply these and previous methods to so-called improper integrals, meaning those for which we relax the rules that the integrand must be a continuous function \( f(x) \) on a closed and bounded interval \([a, b]\). Examples of useful integrals \( \int_a^b f(x) \, dx \) which break these rules include

\[
\int_{-2}^{2} \frac{1}{x^{2/3}} \, dx, \quad \int_{0}^{1} \ln x \, dx, \quad \int_{1}^{\infty} \frac{1}{x^2} \, dx, \quad \int_{-\infty}^{\infty} \frac{1}{x^2 + 1} \, dx,
\]

among many others. In all of the above, either the function \( f(x) \) is not continuous in the entire range of integration \([a, b]\) (particularly because of the presence of vertical asymptotes there), or those “limits of integration” \( a \) and \( b \) are not finite. We will eventually make sense of such integrals by intuitive approaches which will require all of our previous limit methods, as well as those developed here in the earlier sections of this Chapter.

9.1 Some Asymptotics of Functions

It will be important to be able to spot the “forms” of the limits we will encounter in this chapter and beyond. This was also the case in Chapter 3, but we have encountered many more functions since then.

In this section we first take note of the behaviors of functions near vertical asymptotes, and as \( |x| \to \infty \). These behaviors we will collectively call the asymptotics of the functions. In our first subsection we will look at particular functions. Then we will look at compositions and combinations of these functions, in cases where we can somewhat quickly read the behavior of these more complicated functions as following from the behaviors of the underlying functions.

The functions in Figure 9.1, page 670 have all been encountered earlier in the text. Taking these functions in turn, we can list their asymptotic and other limiting behaviors, as read off of their respective graphs.
Figure 9.1.a: Graph of $y = e^x$

Figure 9.1.b: Graph of $y = \ln x$

Figure 9.1.c: Graph of $y = \sin^{-1} x$

Figure 9.1.d: Graph of $y = \cos^{-1} x$

Figure 9.1.e: Graph of $y = \tan^{-1} x$

Figure 9.1.f: Graph of $y = \sec^{-1} x$

Figure 9.1: Graphs of some important functions, illustrating asymptotic behaviors.
9.1. SOME ASYMPTOTICS OF FUNCTIONS

1. $e^x$:
   (a) $x \to \infty \implies e^x \to \infty$
   (b) $x \to -\infty \implies e^x \to 0^+$

2. $\ln x$ (note the relationship between this and $e^x$, which is the inverse of $\ln x$):
   (a) $x \to \infty \implies \ln x \to \infty$
   (b) $x \to 0^+ \implies \ln x \to -\infty$

3. $\sin^{-1} x$:
   (a) $x \to -1^+ \implies \sin^{-1} x \to \left(-\frac{\pi}{2}\right)^+$
   (b) $x \to 1^- \implies \sin^{-1} x \to \left(\frac{\pi}{2}\right)^-$

4. $\cos^{-1} x$:
   (a) $x \to -1^+ \implies \cos^{-1} x \to \pi^-$
   (b) $x \to 1^- \implies \cos^{-1} x \to 0^+$

5. $\tan^{-1} x$:
   (a) $x \to -\infty \implies \tan^{-1} x \to \left(-\frac{\pi}{2}\right)^+$
   (b) $x \to \infty \implies \tan^{-1} x \to \left(\frac{\pi}{2}\right)^-$

6. $\sec^{-1} x$:
   (a) $x \to -\infty \implies \sec^{-1} x \to \left(\frac{\pi}{2}\right)^+$
   (b) $x \to \infty \implies \sec^{-1} x \to \left(\frac{\pi}{2}\right)^-$

These give rise to limit forms, so "$e^\infty = \infty$," "$e^{-\infty} = 0^+,$" and so on. It is also important to remember where these, and all basic functions, are continuous (i.e., for which values $a$ we have $x \to a \implies f(x) \to f(a)$), and what their limiting behavior (from any direction) is at any points within their domains, or approachable from within their domains.

The above limiting behaviors are all clear from the graphs. We can now apply these to more complicated limits where relevant.

**Example 9.1.1** Consider the following limits:

\[
\lim_{x \to \infty} \tan^{-1}(\ln x) = \frac{\pi}{2}, \\
\lim_{x \to 0^+} \tan^{-1}(\ln x) = -\frac{\pi}{2}.
\]

When looking at these combinations of functions, it is important to look “inside-out” to see how the “inner” and “component” functions behave in the limit. To emphasize this, we will sometimes illustrate computations such as the above in ways such as the following:

\[
\lim_{x \to \infty} \tan^{-1}(\ln x) = \frac{\pi}{2}, \\
\lim_{x \to 0^+} \tan^{-1}(\ln x) = -\frac{\pi}{2}.
\]

Yet another way to compute these limits is through substitution. For instance, in this last limit we can write set $u = \ln x$, so that $x \to 0^+ \implies u \to -\infty$, (similar to Theorem 3.9.4, page 275) and so

\[
\lim_{x \to 0^+} \tan^{-1}(\ln x) = \lim_{u \to -\infty} \tan^{-1} u = -\frac{\pi}{2}.
\]

It does help to have the graphs of both of these functions ($\ln$ and $\tan^{-1}$) in mind when performing the computations. However it is more important to know the limiting behaviors, as they may not appear immediately.

**Example 9.1.2** Consider the following limits:

- $\lim_{x \to \infty} \ln x = \infty$.
- $\lim_{x \to \infty} \ln(\ln x) = \ln(\infty) \to \infty$. 

CHAPTER 9. IMPROPER INTEGRALS AND ADVANCED LIMIT TECHNIQUES

\[ \lim_{x \to \infty} \ln(\ln(\ln x)) \quad \text{i.e., } \ln(\infty) \]

With this last one, we can also argue “pictorially.” As \( x \to \infty \), we have

\[
\begin{array}{c}
\infty \\
\downarrow \\
\ln(\ln x) \\
\uparrow \\
\infty
\end{array}
\]

The function \( \ln(\ln(\ln x)) \) does indeed grow to infinity as \( x \) does so, but the growth is very, very slow. If we wish this function to be greater than 10, we would need \( x > e^{e^{10}} \), which is a real number to be sure, but well beyond most readily available computational devices. (Note that \( e^{10} \approx 22026 \), and \( e^{e^{10}} \approx 10^{9565} \), so the exponential of that number is astronomically huge.)

Note that neither \( \lim_{x \to -\infty} \tan^{-1}(\ln x) \) nor \( \lim_{x \to 0} \tan^{-1}(\ln x) \) exist, since the natural logarithm is not defined as \( x \) “travels the path” prescribed by the limit.

The above visual representation lends itself better to handwritten mathematics, for instance in notebooks or on a chalkboard, so we will make somewhat limited use of it here.

Many of the limits which occur naturally in subsequent sections are compositions of functions as above. We will consider a few more of these before moving on to limiting behaviors of other combinations (products, quotients, etc.) of functions.

**Example 9.1.3** Consider the following limit computations. In many, it is important to keep in mind the graphs of the functions involved.

- \[ \lim_{x \to \frac{\pi}{2}} \ln(x) \quad \text{Does Not Exist.} \]
- \[ \lim_{x \to 0^+} e^{e^x} \quad \text{Does Not Exist.} \]
- \[ \lim_{x \to \infty} e^{x} \quad \text{Does Not Exist.} \]
- \[ \lim_{x \to 0^+} \ln(x) \quad \text{Does Not Exist.} \]
- \[ \lim_{x \to 0} \ln(x) \quad \text{Does Not Exist.} \]

Note: cosine is continuous everywhere, including at \( \frac{1}{3} \), so if its argument approaches \( \frac{1}{3} \), then the cosine approaches \( \cos \frac{1}{3} \approx 0.9449 \) (with the angle \( 1/3 \) measured in radians).

In fact, numbers that large are routinely found in statistical mechanics, also known as modern thermodynamics, where it is not so important how large is a particular number per se, but what is often crucial is how it compares to another (large) number.
9.1. SOME ASYMPTOTICS OF FUNCTIONS

\[ \lim_{x \to \frac{\pi}{2}^+} e^{\tan x} = e^{-\infty} 0. \]
\[ \lim_{x \to 0^+} \sec^{-1} \ln x = \frac{\sec^{-1}(-\infty)}{2}. \]
\[ \lim_{x \to 1^+} \ln (\ln x) = \ln(0^+) = -\infty. \]
\[ \lim_{x \to \infty} \sin x = \sin \infty \text{ Does Not Exist.} \]
\[ \lim_{x \to \infty} \tan^{-1}(\frac{1}{e^x}) = \frac{\pi}{2}. \]
\[ \lim_{x \to \infty} (x + \sin x) = \infty + B \infty. \]

The last two limits above require some explanation. That \( \lim_{x \to \infty} \sin x \) does not exist is because the sine function continues to oscillate between \(-1\) and \(1\) as \( x \to \infty \), and thus never approaches settling on a particular value. The last limit is infinite because the sine, while oscillating, nevertheless is bounded in its output within \([-1, 1]\), and can thus never overcome the effects of the term \( x \), and we know \( x \to \infty \) in the limit.

Learning how to compute these types of limits is a matter of remembering how the individual functions are behaving for the limiting value of the variable \( x \) in all the above, and how they are “put together” to form the complete function. The arguments are not difficult, but do require practice.

Exercises

Compute the following limits.

1. \( \lim_{x \to \infty} 2^x \)
2. \( \lim_{x \to \infty} 2^{-x} \)
3. \( \lim_{x \to \infty} \left( \frac{2}{3} \right)^x \)
4. \( \lim_{x \to \infty} \left( \frac{2}{3} \right)^{-x} \)
5. \( \lim_{x \to \infty} 1.001^x \)
6. \( \lim_{x \to \infty} .99^x \)
7. \( \lim_{x \to \infty} .99^{1-x} \)
8. \( \lim_{x \to \infty} \tan^{-1} \sqrt{x} \)
9. \( \lim_{x \to \infty} \tan^{-1}(1 - e^x) \)
10. \( \lim_{x \to 1^+} \ln(\sec^{-1} x) \)
11. \( \lim_{x \to \infty} \frac{\cos x}{x} \)
12. \( \lim_{x \to \infty} \tan^{-1}(e^x) \)
13. \( \lim_{x \to \infty} \tan^{-1}(e^{-x}) \)
14. \( \lim_{x \to \infty} e^{-x^2} \)
15. \( \lim_{x \to \infty} e^{-x^2} \)
16. \( \lim_{x \to \infty} \ln \left( \frac{x^2 + 9}{x^2 - 1} \right) \)
17. \( \lim_{x \to \infty} \frac{1}{x} \)
9.2 L’Hôpital’s Rule

In this section we introduce a very powerful technique for computing limits of the forms 0/0 and ∞/∞, and their variants. These are very important limits in calculus for several reasons, but overall they measure the relative shrinkage or growth of two quantities in the limit: the numerator and the denominator of the function. In particular, the derivative (where it exists) is a 0/0 form limit:

\[ f'(a) = \lim_{\Delta x \to 0} \frac{f(a + \Delta x) - f(a)}{\Delta x}. \]

It measures the instantaneous change in \( f(x) \) (as \( x \to a \)) per unit change in \( x \). Clearly knowing that the limit is of form 0/0 is not enough to tell us the actual value of the limit, or nearly all of our derivatives would be the same (which they are not). In the earlier development of derivatives, we used algebraic and, in the case of the sine function, sandwich theorem and geometric arguments to compute such limits. With L’Hôpital’s Rule we will have another tool available.

The forms 0/0 were, in fact, cases where the outcome depended upon which function approached zero faster: if the numerator was much faster in approaching zero, then its effect was stronger and the fraction shrunk in size to zero; if the denominator was much faster in approaching zero, then its effect was stronger and the fraction blew up to produce a limit of \( \infty \), \(-\infty \), or perhaps nonexistent if both “blowups” were present; and if the effects were proportional, some finite limit could be the result. So for instance we had limits like

\[ \lim_{x \to 0} \frac{\sin x}{x} = 1, \]

\[ \Rightarrow \lim_{x \to 0} \frac{1 - \cos x}{x} = \lim_{x \to 0} \frac{1 - \cos^2 x}{x(1 + \cos x)} = \lim_{x \to 0} \frac{\sin^2 x}{x(1 + \cos x)} = \lim_{x \to 0} \frac{\sin x}{1 + \cos x} = 1 \cdot \frac{0}{2} = 0, \]

\[ \Rightarrow \lim_{x \to 0} \frac{x^2}{(1 - \cos x)^2} = \lim_{x \to 0} \frac{1}{1 - 2 \cos x + 1} = \frac{1}{1 - 0 + 1} = \frac{1}{2}. \]

The first was proven using a geometric argument, the second followed from the first with some algebra (namely multiplying both the numerator and the denominator of the second limit’s function by \( (1 + \cos x) \)), and the third from the second by noting its form is 1/0+. So in the first, the tending towards zero in the numerator and denominator is at very nearly the same rate (for small \( x \)), in the second the numerator approaches zero much faster than the denominator, and in the third it is the denominator which approaches zero faster.

A similar ratio-style comparison of rates occurs with limits of the form ∞/∞. The following are fairly routine:

\[ \lim_{x \to \infty} \frac{x}{x^2 + 1} = 0, \]

\[ \lim_{x \to \infty} \frac{x^2}{x + 1} = \infty, \]

\[ \lim_{x \to \infty} \frac{3x^2 + 9}{5x^2 - x + 1} = \frac{3}{5}. \]

The trick to seeing these was to factor powers of \( x \) from the numerator and denominator and see which effects can cancel each other in the tug-of-war between the numerator and denominator, and analyze the behavior of the terms that are left. So for example one can write

\[ \lim_{x \to \infty} \frac{3x^2 + 9}{5x^2 - x + 1} = \lim_{x \to \infty} \frac{x^2 \left(3 + \frac{9}{x^2}\right)}{x^2 \left(5 - \frac{1}{x} + \frac{1}{x^2}\right)} = \lim_{x \to \infty} \frac{3 + \frac{9}{x^2}}{5 - \frac{1}{x} + \frac{1}{x^2}} = \frac{3 + 0}{5 - 0 + 0} = \frac{3}{5}. \]

\[^2\text{By variants we mean that, for instance, } 0^+/0^-, \infty/(-\infty), \text{ etc.}\]
At times, however, these algebraic analyses are either unwieldy or at least appear to fail, and our main tool to break through such an impass is l’Hôpital’s Rule. The beauty of l’Hôpital’s Rule is that it measures the relative rates of convergence of the numerator and denominator towards their limiting values by comparing their derivatives, which after all do measure their rates of change. So the rule is intuitive, though not trivial to verify. The rule is stated as follows:

**Theorem 9.2.1** Suppose that for some limiting value of \( x \) we have \( f(x), g(x) \to 0 \) or \( f(x), g(x) \to \infty \). Then, if there exists \( L \in \mathbb{R} \cup \{-\infty, \infty\} \) such that \( \frac{f'(x)}{g'(x)} \to L \) exists, it follows that \( \frac{f(x)}{g(x)} \to L \) as well.

So for instance, if we have a \( 0/0 \) form limit we can look instead at the limits of the derivatives of the numerator and denominator of our original function. Similarly with \( \infty/\infty \). Rephrased:

1. If \( \lim_{x \to a} \frac{f(x)}{g(x)} \) is of \( 0/0 \) form, then (see notes below)
   \[
   \lim_{x \to a} \frac{f(x)}{g(x)} \quad \frac{0/0}{0/0} \quad \lim_{x \to a} \frac{f'(x)}{g'(x)} \quad \text{if this second limit exists.}
   \]

2. If \( \lim_{x \to a} \frac{f(x)}{g(x)} \) is of \( \infty/\infty \) form, then
   \[
   \lim_{x \to a} \frac{f(x)}{g(x)} \quad \frac{\infty/\infty}{\infty/\infty} \quad \lim_{x \to a} \frac{f'(x)}{g'(x)} \quad \text{if this second limit exists.}
   \]

Of course the hope is that the second limit will not only exist (or the theorem does not apply!), but will be easier to compute. Several observations are in order before demonstrating the theorem.

1. It is important that the second limit exists. If not, some other method has to be found to compute the original limit, which might or might not exist. In other words, if the second limit does not exist, that is not enough to say anything about the original limit.

2. **This rule requires \( 0/0 \) or \( \infty/\infty \) form, or the rule has no input on the limit.** (Attempting to use the rule for other limit forms is a very common mistake.)

3. If a previous method seems to show potential for success, it should also be considered.

4. The above also works for one-sided limits, limits at infinity, and when the second limit is infinite.

This last point shows how robust l’Hôpital’s Rule is, in that it can work for many cases. However, there are cases for which it does not, and it is certainly not a rule to use on every limit.

Now we look at our first example of how one usually employs the rule.

---

\(^{3}\)Guillaume François Antoine, Marquis de l’Hôpital (originally l’Hospital, 1661–February 2, 1704), French mathematician, who apparently never claimed actual credit for the rule he published anonymously, which many believe was in fact discovered by his teacher, the Swiss mathematician Johann Bernoulli (July 27, 1667–January 1, 1748), who was also in his employ. Apparently the two agreed that Bernoulli would produce mathematical results and l’Hôpital would publish them after paying Bernoulli. L’Hôpital gave some nonspecific credit to Bernoulli for contributing to the work, in particular for the 1696 textbook *Analysis of the Infinitely Small to Understand Curves (l’Analyse des Infiniment Petits pour l’Intelligence des Lignes Courbes)*, so it is not entirely clear who discovered the rule. However, l’Hôpital was apparently a respectable mathematician in his own right, solving the famous “Brachistochrone Problem,” for finding the “curve of fastest descent,” independently (and fairly simultaneously) from Sir Isaac Newton and others.
Example 9.2.1 \( \lim_{x \to 0} \frac{\sin x}{x} \) o/o LHR \( \lim_{x \to 0} \frac{\frac{d}{dx} \sin x}{\frac{d}{dx} x} \) = \( \lim_{x \to 0} \frac{\cos x}{1} \) = \frac{\cos 0}{1} = 1. (See footnote.⁴)

Our application of the rule does not read in the same logical order that the logic of the rule is stated. In fact when we invoke the rule, as in our first “=,” we are doing so provisionally; we write “=" until we are proven wrong (if ever) by the subsequent limit not existing. Because the new limit did exist our first “=” is vindicated, and we declare the computation finished (see the original statement of l’Hôpital’s Rule). We will see how to deal with cases where the second limit does not exist, or is no easier to compute, later in this section.

Other examples where we had previous techniques also illustrate the validity of l’Hôpital’s Rule, such as the next two examples.

Example 9.2.2 \( \lim_{x \to 2} \frac{x^2 - 4}{x^2 + x - 6} \) o/o LHR \( \lim_{x \to 2} \frac{2x}{2x + 1} \) = \frac{4}{5}.

Alternatively, \( \lim_{x \to 2} \frac{x^2 - 4}{x^2 + x - 6} \) o/o ALG \( \lim_{x \to 2} \frac{(x - 2)(x + 2)}{(x - 2)(x + 3)} \) o/o ALG \( \lim_{x \to 2} \frac{x + 2}{x + 3} \) = \frac{4}{5}.

Example 9.2.3 \( \lim_{x \to \infty} \frac{2x + 5}{3x + 4} \) \( \lim_{x \to \infty} \frac{2}{3} \) = \frac{2}{3}.

Alternatively, \( \lim_{x \to \infty} \frac{2x + 5}{3x + 4} \) \( \lim_{x \to \infty} \frac{x}{3x + 4} \) \( \lim_{x \to \infty} \left( \frac{2}{3} + \frac{5}{x} \right) \) \( \lim_{x \to \infty} \frac{2}{3} + \frac{5}{x} \) = \frac{2}{3} + 0 = \frac{2}{3}.

In fact any of the limits at infinity for rational functions can be computed either way, though (as we discuss for other examples later), we may need to employ l’Hôpital’s Rule several times for a rational function with higher-degree polynomials in the numerator and denominator if we wish to use that rule. In fact neither method is as efficient as simply learning the asymptotic rules for rational functions, namely comparing degrees of the numerator and denominator, but those rules were based upon the algebraic approach, and l’Hôpital’s Rule is usually unnecessarily lengthy.

The more obvious utility of l’Hôpital’s Rule is in cases where our algebraic attempts fail to cancel the factors which cause the common limiting behavior of the numerator and denominator. We compute two such limits in the next example. Note that \( \ln x \) cannot algebraically “cancel,” in full or in part, with a power of \( x \).

Example 9.2.4 Consider the following limits. (See Figure 9.2.)

- \( \lim_{x \to \infty} \frac{\ln x}{x} \) \( \lim_{x \to \infty} \frac{1}{x} \) \( \lim_{x \to \infty} \frac{1}{x} = 0 \).

- \( \lim_{x \to \infty} \frac{\ln x}{\sqrt{x}} \) \( \lim_{x \to \infty} \frac{1}{2\sqrt{x}} \) \( \lim_{x \to \infty} \frac{1}{2\sqrt{x}} \) \( \lim_{x \to \infty} \left( \frac{1}{x} \cdot 2\sqrt{x} \right) \) = \( \lim_{x \to \infty} \frac{2}{\sqrt{x}} \) \( \lim_{x \to \infty} \frac{2}{ \sqrt{x}} \) = 0.

The above example illustrates that \( \ln x \) grows towards \( \infty \) much more slowly than \( x \) (as \( x \to \infty \)), and in fact much more slowly than \( \sqrt{x} \). While the first limit seems very reasonable in light of Figure 9.2, because \( y = \ln x \) and \( y = x \) seem to diverge so quickly and irreversibly,

⁴In fact there is a circular argument here, because we used geometry to prove that \( \lim_{x \to 0} \frac{\sin x}{x} = 1 \) (Theorem 3.9.2, page 270), and used that limit to prove that \( \frac{d}{dx} \sin x = \cos x \) (Theorem 4.2.6, page 318), so we can hardly use that derivative to prove the original limit! So Example 9.2.1 is not a proof of the limit, yet we consider this example merely to demonstrate l’Hôpital’s Rule and to verify that it is consistent with our original result for this limit.
by contrast the shapes of \( y = \ln x \) and \( y = \sqrt{x} \) appear too similar to draw the same conclusion from their graphs, especially with the scale used in the figure. However, when we compare their slopes, respectively \( \frac{1}{x} \) and \( \frac{1}{2 \sqrt{x}} \), we see the latter shrinks much more slowly than the former, as our limit computations verify. l'Hôpital's Rule gives us a more sensitive tool to compare the numerator and denominator, at least for the second problem in Example 9.2.4 above.

In fact \( \ln x \) grows more slowly than any positive power of \( x \), i.e., more slowly than \( x^s \) for any \( s > 0 \). The proof is left to the exercises. Note that we can also “flip” the function in that previous limit and still use l'Hôpital's Rule:

\[
\lim_{x \to \infty} \frac{\sqrt{x}}{\ln x} = \lim_{x \to \infty} \frac{1}{2 \sqrt{x}} = 0
\]

Since the final limit actually does exist—albeit as \( \infty \)—we conclude the original limit is also \( \infty \). Again this shows \( \sqrt{x} \) grows much faster than \( \ln x \), as \( x \to \infty \).

How quickly \( \ln u \to 0 \) as \( u \to 1 \) is also interesting, though can be more complicated.

Example 9.2.5 (See Figure 9.3.)

\[
\lim_{x \to 0} \frac{\ln(1+7x)}{x} = \lim_{x \to 0} \frac{1+7x}{x} = \frac{1+7}{1} = 8
\]

Though not vital for computing such limits, some intuition regarding l'Hôpital's Rule for \( 0/0 \) forms can be seen from this and similar examples. Indeed, for any \( 0/0 \)-form example where both functions have legitimate linear approximations at the limit point, such as the functions \( \ln(1+7x) \) and \( x \) as \( x \to 0 \), these approximations help illustrate the nature of l'Hôpital's Rule for such \( 0/0 \) cases. For our two relevant functions above, we have \( \ln(1+7x) \approx 0 + 7x = 7x \), i.e., \( l(x) = 0 + 7x \) is the linear approximation of \( \ln(1+7x) \) near \( x = 0 \), and \( y = x \) is its own linear approximation (everywhere). From these it seems we can state that \( \ln(1+7x)/x \approx 7x/x = 7 \), this being arguably

\footnote{Recall \( f(x) \approx f(x_0) + f'(x_0)(x-x_0) \) is the linear approximation of \( f(x) \) centered at \( x_0 \), assuming \( f'(x_0) \) exists. If \( f(x_0) = 0 \), then the linear approximation is just \( f(x) \approx f'(x_0)(x-x_0) \).}
a better and better approximation as \( x \to 0 \), giving some intuition why l’Hôpital’s Rule makes sense. However we also need the derivatives to approach \( \frac{d}{dx} \ln(1 + 7x) \big|_{x=0} \) and \( \frac{d}{dx} \big|_{x=0} \) as \( x \to 0 \), for this approximation to equal the new limit (of the ratio of derivatives) at the limit point, but for this example both derivatives are continuous at \( x = 0 \) (the interested reader should verify) so the limit of derivatives computed in the example and described in l’Hôpital’s Rule is consistent with this observation, at least for this case. See Figure 9.3.

If the limit point is not \( x = 0 \), then the same kind of ratio of derivatives occurs again, which a glance at the linear approximations will show (because the “\( f(x_0) \)” term is zero, see Footnote 5, page 677). If the limit point is not a point of continuity of the functions (and their derivatives), or the limit point is not finite, then this intuition needs to be modified. We will not look at all possible cases of l’Hôpital’s Rule here, as it would be a rather long distraction.

There are cases in which l’Hôpital’s Rule, in whatever form, is invoked multiple times to compute a single limit. If the forms continue to be \( 0/0 \) or \( \infty/\infty \) until the final limit, and it exists, then it will equal the original. It is important to verify that we have one of these forms at each step we apply l’Hôpital’s Rule.
9.2. L'HÔPITAL'S RULE

Example 9.2.6 Consider the following limits.

\[ \lim_{x \to 0} \frac{e^x - x - 1}{x^2} \quad \lim_{x \to 0} \frac{e^x - 1}{x} \quad \lim_{x \to 0} \frac{e^x}{x} \quad \lim_{x \to 0} \frac{e^{x^2}}{2x} \quad \lim_{x \to 0} \frac{e^{x^2/2}}{L^{HR}} = \frac{1}{2}. \]

\[ \lim_{x \to 0} \frac{\sin x - xe^{x^2}}{x^2} \quad \lim_{x \to 0} \frac{\cos x - (x \cdot 2xe^{x^2} + e^{x^2})}{2x} \quad \lim_{x \to 0} \frac{\cos x - 2xe^{x^2} - e^{x^2}}{2x} = 0 = 0. \]

There are times when l'Hôpital's Rule seems useless because either the limit of derivatives does not exist, or it is no simpler than the original limit. In such cases, usually previous methods need to be employed.

Example 9.2.7 If we try to compute \( \lim_{x \to \infty} \frac{2^x}{3^x} \), it is quite simple using earlier methods:

\[ \lim_{x \to \infty} \frac{2^x}{3^x} = \lim_{x \to \infty} \left(\frac{2}{3}\right)^x = 0, \]

since we know that \( y = a^x \to 0^+ \) as \( x \to \infty \) if \( a \in (0, 1) \).

But if we attempt to use l'Hôpital's Rule to compute this limit, we have

\[ \lim_{x \to \infty} \frac{2^x}{3^x} = \lim_{x \to \infty} \frac{2^x \ln 2}{3^x \ln 3} = \lim_{x \to \infty} \frac{2^x (\ln 2)^2}{3^x (\ln 3)^2} = \cdots, \]

which, while true, does not get us any closer to the correct answer.

Example 9.2.8 Attempting to compute \( \lim_{x \to \infty} \frac{\sqrt{3x^2 + 4}}{x} \) using l'Hôpital's Rule finds us no better off than we were at the start.

\[ \lim_{x \to \infty} \frac{\sqrt{3x^2 + 4}}{x} = \lim_{x \to \infty} \frac{2\sqrt{3x^2 + 4}}{1} = \lim_{x \to \infty} \frac{3x}{\sqrt{3x^2 + 4}}. \]

Clearly this is no simpler than the original. We could attempt a strategy where we try to “solve for the limit” algebraically, as we sometimes did for the antiderivative with integration by parts—here we would have \( L = \frac{3}{2} \)—except we are not sure (without previous methods) that we are dealing with a finite, or even existent, limit \( L \) for which to solve. So instead we look to previous methods (note \( x > 0 \)):

\[ \lim_{x \to \infty} \frac{\sqrt{3x^2 + 4}}{x} = \lim_{x \to \infty} \sqrt{\frac{3 + \frac{4}{x^2}}{x}} = \lim_{x \to \infty} \sqrt{\frac{3}{x^2} + \frac{4}{x^2}} = \sqrt{3} + 0 = \sqrt{3}. \]

In fact, that is consistent with our algebraic equation \( L = 3/L \implies L^2 = 3 \). Note that there is another strategy, where we use the continuity of the square root to move the “limit” inside:

\[ \lim_{x \to \infty} \frac{\sqrt{3x^2 + 4}}{x} = \lim_{x \to \infty} \sqrt{\frac{3x^2 + 4}{x^2}} = \sqrt{\lim_{x \to \infty} \frac{3x^2 + 4}{x^2}} = \sqrt{3}. \]

As with l'Hôpital's Rule, bringing the limit inside the radical required that the final limit existed, which it did.

Next we look at an example where the new limit (of derivatives) in l'Hôpital's Rule does not exist, though the original limit does.

Example 9.2.9 Consider \( \lim_{x \to \infty} \frac{\sin x + x}{\cos x + x} \). First note that \(-1 \leq \sin x \leq 1, \) and \(-1 \leq \cos x \leq 1, \) so as \( x \to \infty \) we have \( \sin x + x \to \infty \) and \( \cos x + x \to \infty \). This can be seen by using the sandwich theorem, but we really only need one side of it (the other being \( 1 + x \geq \sin x + x \)) because we have

\[ \sin x + x \geq -1 + x \to \infty \quad \text{as} \quad x \to \infty, \]

\[ \lim_{x \to \infty} \frac{1}{\sqrt{e^x - x - 1}} = \lim_{x \to \infty} \frac{1}{x^2} = 0. \]
implying the greater quantity $\sin x + x \to \infty$ as well. Similarly for $\cos x + x$.

Using our old methods, where we let “$B$” stand for any bounded term when useful (such as for $\sin x$ since $|\sin x| \leq 1$), we might write

$$
\lim_{x \to \infty} \frac{\sin x + x}{\cos x + x} = \lim_{x \to \infty} \frac{x \left( \frac{\sin x}{x} + 1 \right)}{x \left( \frac{\cos x}{x} + 1 \right)} = \lim_{x \to \infty} \frac{\sin x + 1}{\cos x + 1} \frac{(\infty/\infty) + 1}{(0/0) + 1} = 0 + 1 = 1.
$$

However, we should note that if we attempted l’Hôpital’s Rule, we would be next analyzing $\lim_{x \to \infty} \frac{\cos x + 1}{-\sin x + 1}$, which does not exist because both numerator and denominator are oscillating, but not in a way in which these effects would cancel. This is readily seen by the periodic nature of the new quotient, but we have another reason to conclude this second limit does not exist: the fact that $-\sin x + 1 \to 0^{\pm}$ for many instances where $\cos x + 1 \to 1$ (namely, when $x \to \frac{\pi}{2} + 2n\pi$, $n \in \{0, \pm 1, \pm 2, \cdots \}$), making the function periodically undefined as $x$ increases, again leading us to conclude the new limit does not exist (though the original does).

The above example demonstrates that we can not use l’Hôpital’s Rule directly if the second limit does not exist. However, there are times we can still use l’Hôpital’s Rule for some cases, if we are willing to look at both left and right limits.

**Example 9.2.10** Consider $\lim_{x \to 0} \frac{\sin x}{x^2}$. If we first attempt to use l’Hôpital’s Rule, we would write

$$
\lim_{x \to 0} \frac{\sin x}{x^2} = \frac{0/0}{LHR ?} \lim_{x \to 0} \frac{\cos x}{2x} \text{ does not exist (1/0), so l’Hôpital’s Rule does not, technically, apply. However, we note that if we had taken left and right limits separately, l’Hôpital’s Rule would apply to each:}
$$

$$
\lim_{x \to 0^-} \frac{\sin x}{x^2} = \frac{0/0}{LHR} \lim_{x \to 0^-} \frac{\cos x}{2x} = \frac{1/0^-}{2x} = -\infty,
$$

$$
\lim_{x \to 0^+} \frac{\sin x}{x^2} = \frac{0/0}{LHR} \lim_{x \to 0^+} \frac{\cos x}{2x} = \frac{1/0^+}{2x} = \infty.
$$

From these computations we can see that left and right limits for the original function differ (one is $-\infty$ and the other $\infty$), so we conclude that $\lim_{x \to 0} \frac{\sin x}{x^2}$ does not exist.

L’Hôpital’s Rule is a very useful method to add to one’s arsenal, but it does require discretion for when to use it, and when not to use it. Sometimes it is useful, or just necessary, for only part or parts of the limit, but not the whole limit collectively.

**Example 9.2.11** Compute $\lim_{x \to 0} \frac{(x^2 + 7x - 9) \sin x}{4x^2 + 6x^2 + 7x}$.

**Solution:** Here we factor the function in question into the product of two functions:

$$
\lim_{x \to 0} \frac{(x^2 + 7x - 9) \sin x}{4x^2 + 6x^2 + 7x} = \lim_{x \to 0} \left[ \frac{x^2 + 7x - 9}{4x^2 + 6x + 7} \cdot \frac{\sin x}{x} \right] = \left( \lim_{x \to 0} \frac{x^2 + 7x - 9}{4x^2 + 6x + 7} \right) \left( \lim_{x \to 0} \frac{\sin x}{x} \right)
$$

$$
= \frac{-9}{7} \cdot \lim_{x \to 0} \frac{\cos x}{1} = -\frac{9}{7} \cdot 1 = -\frac{9}{7}.
$$

Note that this was only allowed because both limits existed, though you can always factor out a any part that has a finite, nonzero limit and find the limit of what remains, existing or not.
In fact we could have used l'Hôpital’s Rule on the entire function above, but we would have endured needless complications, in this case the product rule repeatedly.

The following limit is already broken into three summed parts, and while we could combine them into one fraction with which to use l'Hôpital’s Rule, it is much easier to analyze the terms separately, and if each limit exists as a finite number, then the limit of the sum will be the sum of the limits.

**Example 9.2.12** Compute \( \lim_{x \to \infty} \left( \tan^{-1} x + \frac{(\ln x)^2}{x} - e^{1/x} \right) \).

**Solution:** Here again we break the function into pieces that have known limits, and a piece that needs to be analyzed further.

\[
\lim_{x \to \infty} \left( \tan^{-1} x + \frac{(\ln x)^2}{x} - e^{1/x} \right) = \frac{\pi}{2} + \left( \lim_{x \to \infty} \frac{(\ln x)^2}{x} \right) - e^0 = \frac{\pi}{2} + \left( \lim_{x \to \infty} \frac{2 \ln x \cdot \frac{1}{x}}{1} \right) - 1
\]

\[
= \frac{\pi}{2} - 1 + \lim_{x \to \infty} \frac{2 \ln x}{x} = \frac{\pi}{2} - 1 + \lim_{x \to \infty} \frac{2}{x} = \frac{\pi}{2} - 1 + 0
\]

\[
= \frac{\pi}{2} - 1.
\]

We do have to be a little careful in breaking a limit’s function into various parts. For instance, if doing so gives us another indeterminate form we have to adopt another strategy, often recombining the function in some expedient way (as in the next section). For instance, if the “pieces” of the function have limits that would collectively read \( \infty - \infty \), or \( 0 \cdot \infty \), etc., then we have to find some way to combine them to either cancel competing influences (as much as possible), or employ other techniques, including but not limited to l'Hôpital’s Rule.

It also happens sometimes that a \( 0/0 \) or \( \infty/\infty \) form occurs within a function, and we are wise to attempt to use LHR “inside” the function to eventually lead us to the final conclusion regarding a limit’s value.

**Example 9.2.13** Compute \( \lim_{x \to \infty} \cos \left( \frac{\ln x}{x} \right) \).

**Solution:** Here we do not have a function which is a quotient per se, but we can compute the limit inside the function, and if it approaches a point for which we can conclude to have a determinate “form,” we can deduce the actual limit. Here,

\[
\lim_{x \to \infty} \frac{\ln x}{x} = \lim_{x \to \infty} \frac{1}{x} = 0/0 \quad \text{LHR}
\]

\[
\lim_{x \to \infty} \frac{1}{x} = \lim_{x \to \infty} \frac{(1/\infty)/1}{0}
\]

\[
\implies \lim_{x \to \infty} \cos \left( \frac{\ln x}{x} \right) = \cos 0 = 1.
\]

At this point a warning regarding l'Hôpital’s Rule is in order. While the rule involves derivatives and quotients, it does not invoke the quotient rule; we are not taking the derivative of the function itself, but of its numerator and denominator separately. Misreading l'Hôpital’s Rule and applying the quotient rule is another common mistake for novice calculus students. However the more common mistake is to try to use LHR when we do not have a \( 0/0 \) or \( \infty/\infty \) form.
Chapter 9. Improper Integrals and Advanced Limit Techniques

Note: Historical information was obtained from the internet site “Wikipedia,” located at the URL http://www.wikipedia.com.

Exercises

Compute the following limits.

1. Compute \( \lim_{x \to 0} \frac{\tan^{-1} x}{x} \).
2. Compute \( \lim_{x \to \pi/4} \frac{\tan x - 1}{x - \frac{\pi}{4}} \).
3. Compute \( \lim_{x \to 1} \frac{\tan^{-1} x - \pi}{x - 1} \).
4. Compute \( \lim_{x \to 1} \ln x \cdot \frac{1}{x - 1} \).
5. Compute \( \lim_{x \to 0} \frac{\ln x}{x} \).
6. Compute \( \lim_{x \to 0} \frac{\sin 5x}{\sin 3x} \).
7. Compute \( \lim_{x \to 0} \frac{e^x - \frac{1}{2}x^2}{x^2} \) two ways:
   (a) By using l'Hôpital’s Rule directly.
   (b) By simplifying the fraction and then using l'Hôpital’s Rule.
8. Compute \( \lim_{x \to 0} \frac{\sin x}{x} \).
9. Show that \( \lim_{x \to \infty} \frac{\ln x}{x^s} = 0 \) for any \( s > 0 \).
10. Show that \( \lim_{x \to \infty} \frac{e^x}{x^n} = \infty \) for any \( n \in \{1, 2, 3, \ldots\} \). (You will need to “employ” l'Hôpital’s Rule \( n \) times, or at least declare some pattern eventually.)
11. Compute \( \lim_{x \to 0^+} \frac{\ln(1 + x)}{x} \).
12. Compute \( \lim_{x \to \infty} \frac{\ln(1 + \frac{1}{x})}{x} \), two ways.
   (a) With l'Hôpital’s Rule directly.
   (b) Using the previous limit. (Let \( u = \frac{1}{x} \).)
13. Compute \( \lim_{x \to \infty} \sqrt{4x^2 + 9x + 7} \).
14. Compute \( \lim_{x \to \infty} \frac{2^x}{x^2} \).
15. Compute \( \lim_{x \to 0} \frac{2^x}{x^2} \).
16. Compute \( \lim_{x \to 0} \frac{x^2}{2^x} \).
17. Compute \( \lim_{x \to \infty} xe^{-x} \).
18. Compute \( \lim_{x \to 0} \tan^{-1} \left( \frac{\sin x}{x} \right) \).
19. Compute \( \lim_{x \to \infty} \sin \left( \frac{\pi x^2 - 4x + 6}{1 - 2x^2} \right) \).
9.3 Other Indeterminate Forms

Besides the forms $0/0$ and $\infty/\infty$, for which we can use old methods or l'Hôpital’s Rule, there are other forms which are also indeterminate. Here we will be mostly interested in the following, which the reader should confirm seem indeterminate (meaning just knowing that the limit is of such a form is not enough to determine the limit’s value).

1. $\infty - \infty$.
2. $0 \cdot \infty$.
3. $(0^+)^0$.
4. $\infty^0$.
5. $1^\infty$.

While the methods of Chapter 3 will sometimes work here, we will usually need something that attacks these new problems with slightly different strategies. So far our only new tool has been l’Hôpital’s Rule, and we will use it again here extensively, but it requires a quotient. So to summarize our method of attack with these problems, we attempt to rewrite the problems so that we have a quotient which can be analyzed by l’Hôpital’s Rule or previous strategies. In fact, having a quotient is often useful even if we do not need l’Hôpital’s Rule.

9.3.1 $\infty - \infty$, $0 \cdot \infty$

We also include variants such as $0 \cdot (-\infty)$, $0^+ \cdot (-\infty)$, etc., and when the order of the terms is switched (e.g., $\infty \cdot 0$). For obvious reasons, do not include definite forms like $\infty + \infty = \infty$, or $\infty \cdot \infty = \infty$, or for that matter $0 - 0 = 0$, $0 \cdot 0 = 0$.

Example 9.3.1 Compute \( \lim_{x \to \infty} \sqrt{x^2 - 5x + 100} - x \).

Solution: Here we have $\infty - \infty$, since $x^2 - 5x + 100 \to \infty$ and therefore $\sqrt{x^2 - 5x + 100} \to \infty$. In fact this type of limit was already analyzed in Chapter 3 (specifically Example 3.8.8, page 260). Note that we do solve this by re-writing the limit as a quotient.

\[
\lim_{x \to \infty} \left[ \sqrt{x^2 - 5x + 100} - x \right] = \lim_{x \to \infty} \left[ \left( \sqrt{x^2 - 5x + 100} - x \right) \cdot \frac{\sqrt{x^2 - 5x + 100} + x}{\sqrt{x^2 - 5x + 100} + x} \right] \\
= \lim_{x \to \infty} \frac{x^2 - 5x + 100 - x^2}{\sqrt{x^2 - 5x + 100} + x} = \lim_{x \to \infty} \frac{-5x + 100}{\sqrt{x^2 - 5x + 100}} \\
= \lim_{x \to \infty} x \cdot \left( \frac{-5 - \frac{100}{x}}{\sqrt{1 - \frac{5}{x} + \frac{100}{x^2}} + 1} \right) = \lim_{x \to \infty} \frac{-5 - \frac{100}{x}}{\sqrt{1 - \frac{5}{x} + \frac{100}{x^2}} + 1} \\
= \frac{-5 - 0}{\sqrt{1 - 0 + 0} + 1} = -\frac{5}{2}.
\]

The example above did not require l’Hôpital’s Rule, and in fact l’Hôpital’s Rule is not helpful even after we have a quotient. (See Example 9.2.8, page 679.)

The next example is of the form $-\infty \cdot 0^+$ (sometimes just written $-\infty \cdot 0$, which is similar to $\infty \cdot 0$ except for the sign).
Example 9.3.2 Compute \( \lim_{x \to 0^+} x \ln x \).

**Solution:** Since this is of form \(0 \cdot (-\infty)\), which is indeterminate, our strategy is to first algebraically rewrite the function as a quotient. This will give us a form on which we can apply l'Hôpital’s Rule.

\[
\lim_{x \to 0^+} x \ln x = \lim_{x \to 0^+} \frac{\ln x}{\frac{1}{x}} = \lim_{x \to 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x \to 0^+} \left( -x \right) = 0.
\]

In the above example, the influence of \(x \to 0\) is stronger than that of \(\ln x \to -\infty\). To reach this conclusion, we needed a quotient to compare properly the growth of these two functions, because a quotient allows tools such as l'Hôpital’s Rule.

It should be pointed out that we could have placed the natural logarithm function in the denominator, but that would not give us a satisfactory application of l'Hôpital’s Rule:

\[
\lim_{x \to 0^+} x \ln x = \lim_{x \to 0^+} \frac{\ln x}{\frac{1}{x}} = \lim_{x \to 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x \to 0^+} \left( -x \right) = 0.
\]

We see that our new limit is no simpler, and is arguably harder than the the original.

Example 9.3.3 Compute \( \lim_{x \to \infty} x^2 e^{-x} \).

**Solution:** This is actually simpler than the previous example.

\[
\lim_{x \to \infty} x^2 e^{-x} = \lim_{x \to \infty} \frac{0}{x} = 0.
\]

Example 9.3.4 Consider the following \(\infty \cdot 0\) limits. Note that sometimes we have no choice but to put a logarithmic term in the new denominator:

\[
\begin{align*}
\bullet \quad & \lim_{x \to \infty} x \ln(\cos(1/x)) = \lim_{x \to \infty} \frac{\ln(\cos(1/x))}{1/x} = \lim_{x \to \infty} \frac{\frac{1}{x}}{\cos(1/x)} \cdot (-\sin(1/x)) \cdot \frac{1}{x^2} \\
& = \lim_{x \to \infty} \tan(1/x) = \tan 0 = 0.
\end{align*}
\]

\[
\begin{align*}
\bullet \quad & \lim_{x \to \infty} (\ln x) \ln(\cos(1/x)) = \lim_{x \to \infty} \frac{\ln(\cos(1/x))}{1/x} = \lim_{x \to \infty} \frac{\frac{1}{x}}{\cos(1/x)} \cdot (-\sin(1/x)) \cdot \frac{1}{x^2} \\
& = \lim_{x \to \infty} \frac{-2 \ln x \cdot \frac{1}{x^2}}{\cos(1/x)} \cdot \frac{1}{x^2} + \cot(1/x) \cdot \frac{1}{x^2} = \lim_{x \to \infty} \frac{-2 \ln x}{\csc(1/x) + x \cot(1/x)} \\
& = \lim_{x \to \infty} \frac{-2 \ln x}{\csc(1/x) \cot(1/x)} \cdot \frac{1}{x^2} + \cot(1/x) + x \left( -\csc^2(1/x) \cdot \frac{1}{x^2} \right) \cdot \frac{1}{x^2} \\
& = \lim_{x \to \infty} \frac{-2x}{\csc(1/x) \cot(1/x) + x^2 \cot(1/x) + x \csc^2(1/x)} \\
& = \lim_{x \to \infty} \frac{-2/ \left( -\csc(1/x) \cot(1/x) \cdot \frac{1}{x^2} \csc(1/x) + x^2 \left( -\csc^2(1/x) \cdot \frac{1}{x^2} \right) + \csc^2(1/x) \right)}{-2/\infty} = 0.
\end{align*}
\]
(It was important that all terms in the denominator have been added, that the trigonometric functions are all approaching \( \infty \), and so is \( \frac{1}{x^2} \).) Note how we do what we can to get \( \ln x \) by itself, without other factors, so that it does not reappear when we invoke l’Hôpital’s Rule.

Note also that since \( x \to \infty \) faster than \( \ln x \to \infty \), we could have used the first limit computation above to anticipate this one. In other words, if \( x \ln(\cos(1/x)) \to 0 \), and for large enough \( x \) we have \( 0 \leq \ln x < x \), we expect \( (\ln x)(\ln(\cos(1/x))) \to 0 \) as well.

Finally note that we cannot use l’Hôpital’s Rule for the \( 0 \cdot \infty / \infty \), because we need to know we have \( 0/0 \) or \( \infty / \infty \). After a bit of algebra, we got \( \infty / \infty \), used l’Hôpital’s Rule, then used more algebra until we had a determinate form \( 0/\infty = 0 \). Most problems are not that involved.

\[
\begin{align*}
\lim_{x \to \infty} x^2 \ln(\cos(1/x)) &= \lim_{x \to \infty} \frac{x^2 \ln(x)}{\cos(1/x)} \\
&= \lim_{x \to \infty} \frac{\tan(1/x)}{\cos(1/x)} \\
&= \lim_{x \to \infty} \frac{\sec^2(1/x) \cdot \frac{1}{x^2}}{1} \\
&= \lim_{x \to \infty} \frac{1}{\sec^2(1/x)} = \frac{1}{2}.
\end{align*}
\]

\[
\begin{align*}
\lim_{x \to \infty} x^3 \ln(\cos(1/x)) &= \lim_{x \to \infty} x \cdot x^2 \ln(\cos(1/x)) \\
&= \lim_{x \to \infty} \frac{x \cdot x^2 \ln(x)}{\cos(1/x)} \\
&= \lim_{x \to \infty} \frac{\sec^2(1/x) \cdot \frac{1}{x^2}}{1} \\
&= \lim_{x \to \infty} \frac{-x}{6 \cos^2(1/x)} = -\infty.
\end{align*}
\]

If this last limit seems surprising, note that \( \cos(1/x) < 1 \) for large \( x \), and so \( \ln(\cos(1/x)) < \ln 1 = 0 \), so the form is more precisely stated as \( \infty \cdot 0^- \).

The above limits from Example 9.3.2 to this last one illustrate again that \( 0 \cdot \infty \) form is indeterminate (more information than just the form is required), and that the general strategy is to rewrite the function as a ratio (fraction), though there is some cleverness as to the form of the ratio that will help us to find the actual limit. The algebraic rewriting may be necessary in intermediate steps as well. And of course we can only use l’Hôpital’s Rule if we have a variation of \( 0/0 \) or \( \infty/\infty \) forms.

### 9.3.2 \( 1^\infty, \ 0^0, \ \infty^0 \) and similar forms.

The technique for all of these is to apply the natural logarithm to the function, thus bringing the exponent down as a factor. Some adjustment must be made, because we are then calculating the limit of the natural logarithm of the original function, but that is relatively easy (but crucial) to deal with.

**Example 9.3.5** Compute \( \lim_{x \to \infty} \left( 1 + \frac{7}{x} \right)^x \).

**Solution:** This is of the form \( 1^\infty \). If we define \( y \) to be the function inside the limit, then

\[
\lim_{x \to \infty} \left( 1 + \frac{7}{x} \right)^x = \lim_{y \to \infty} y = \lim_{x \to \infty} e^{\ln y}.
\]
Once we know that \( \ln y \to L \), we would have \( y \to e^L \), assuming \( L \) is finite. (Even if \( L \) is infinite we will have ways of dealing with this.)

Set \( y = \left( 1 + \frac{7}{x} \right)^x \)

\[ \Rightarrow \ln y = x \ln \left( 1 + \frac{7}{x} \right) \]

\[ \Rightarrow \ln y = \lim_{x \to \infty} x \ln \left( 1 + \frac{7}{x} \right) \]

\[ \lim_{x \to \infty} \ln y = \lim_{x \to \infty} x \ln \left( 1 + \frac{7}{x} \right) \]

\[ = \lim_{x \to \infty} \frac{\ln \left( 1 + \frac{7}{x} \right)}{\frac{1}{x}} \cdot \frac{7}{x^2} = \lim_{x \to \infty} \frac{7}{1 + \frac{7}{x}} = \frac{7}{1 + 0} = 7. \]

This is not our final answer. So far we know where the natural log of our function approaches: \( \ln y \to 7 \). The actual function’s limit is computed next:

\[ \lim_{x \to \infty} \left( 1 + \frac{7}{x} \right)^x = \lim_{x \to \infty} y = \lim_{x \to \infty} e^{\ln y} e^7. \]

The general strategy is to bring the exponent down as a multiplier, which can be put (in the form of its reciprocal) into a denominator, at which point we can use l’Hôpital’s Rule and other techniques. However we must realize that the process has us computing the limit of \( \ln y \), if \( y \) is our original function, so we must report instead the limit of \( y = e^{\ln y} \).

It is worth noting that the “7” in Example 9.3.5 can be replaced by any number. In fact we have the following:

\[ \lim_{x \to \infty} \left( 1 + \frac{\xi}{x} \right)^x = e^\xi, \quad (9.1) \]

\[ \lim_{x \to \infty} \left( 1 + \frac{1}{x} \right)^x = e^1 = e. \quad (9.2) \]

These give alternative definitions of \( e^\xi \) and \( e \), respectively.

**Example 9.3.6** Compute \( \lim_{x \to \infty} x^{1/\sqrt{x}} \).

**Solution:** This is of the form \( \infty^0 \) (or more precisely \( \infty^0^+ \)). We compute this limit as in the previous example.

Set \( y = x^{1/\sqrt{x}} \)

\[ \Rightarrow \ln y = \ln x^{1/\sqrt{x}} = \frac{\ln x}{\sqrt{x}} \]

\[ \Rightarrow \lim_{x \to \infty} \ln y = \lim_{x \to \infty} \frac{\ln x}{\sqrt{x}} \]

\[ = \lim_{x \to \infty} \frac{\ln x}{\sqrt{x}} \]

\[ = \lim_{x \to \infty} \frac{1}{\sqrt{x}} \cdot \frac{2x}{2x} = \lim_{x \to \infty} \frac{1}{\sqrt{x}} \cdot \frac{2x}{2x} = \lim_{x \to \infty} \frac{2}{\sqrt{x}} = 0. \]

Thus \( \lim_{x \to \infty} x^{1/\sqrt{x}} = \lim_{x \to \infty} y = \lim_{x \to \infty} e^{\ln y} = e^0 = 1. \)
Example 9.3.7 Compute $\lim_{x \to 0^+} (\cos x)^{1/x}$, $\lim_{x \to 0^+} (\cos x)^{1/x^2}$ and $\lim_{x \to 0^+} (\cos x)^{1/x^3}$.

Solution. These three are very similar, but we write all the computations for the sake of comparison. All are of the form $1^\infty$.

- $\lim_{x \to 0^+} (\cos x)^{1/x}$:
  
  Set $y = (\cos x)^{1/x}$

  \[
  \Rightarrow \ln y = \frac{1}{x} \ln(\cos x)
  \]

  \[
  \Rightarrow \lim_{x \to 0^+} \ln y = \lim_{x \to 0^+} \frac{\ln(\cos x)}{x} \xrightarrow{LHR} 0/0 \implies \lim_{x \to 0^+} \frac{1}{\cos x} (-\sin x) = \lim_{x \to 0^+} \frac{-\sin x}{\cos x} = 0/1 = 0.
  \]

  Thus $\lim_{x \to 0^+} (\cos x)^{1/x} = \lim_{x \to 0^+} y = \lim_{x \to 0^+} e^{\ln y} = e^0 = 1$.

- $\lim_{x \to 0^+} (\cos x)^{1/x^2}$:
  
  Set $y = (\cos x)^{1/x^2}$

  \[
  \Rightarrow \ln y = \frac{1}{x^2} \ln(\cos x)
  \]

  \[
  \Rightarrow \lim_{x \to 0^+} \ln y = \lim_{x \to 0^+} \frac{\ln(\cos x)}{x^2} \xrightarrow{LHR} 0/0 \implies \lim_{x \to 0^+} \frac{1}{\cos x} (-\sin x) \frac{2x}{2x} = \lim_{x \to 0^+} \frac{-\tan x}{2x} = -1/2.
  \]

  Thus $\lim_{x \to 0^+} (\cos x)^{1/x^2} = \lim_{x \to 0^+} y = \lim_{x \to 0^+} e^{\ln y} = e^{-1/2} = \frac{1}{\sqrt{e}}$.

- $\lim_{x \to 0^+} (\cos x)^{1/x^3}$:
  
  Set $y = (\cos x)^{1/x^3}$

  \[
  \Rightarrow \ln y = \frac{1}{x^3} \ln(\cos x)
  \]

  \[
  \Rightarrow \lim_{x \to 0^+} \ln y = \lim_{x \to 0^+} \frac{\ln(\cos x)}{x^3} \xrightarrow{LHR} 0/0 \implies \lim_{x \to 0^+} \frac{1}{\cos x} (-\sin x) \frac{3x^2}{3x^2} = \lim_{x \to 0^+} \frac{-\tan x}{3x^2} = -\infty.
  \]

  Thus $\lim_{x \to 0^+} (\cos x)^{1/x^3} = \lim_{x \to 0^+} y = \lim_{x \to 0^+} e^{\ln y} = e^{-\infty} = 0$.

Summarizing, though these were all $1^\infty$, we have a case where the influence of the 1 is stronger, one where they are relatively equal, and one where the $\infty$ is stronger (coupled with the fact that “1” is in each of these cases precisely stated as $1^-$):

\[
\lim_{x \to 0^+} (\cos x)^{1/x} = 1, \\
\lim_{x \to 0^+} (\cos x)^{1/x^2} = 1/\sqrt{e}, \\
\lim_{x \to 0^+} (\cos x)^{1/x^3} = 0.
\]
9.4 Improper Integrals

In this section we look at definite integrals where one or more boundaries of integration—here sometimes called endpoints of integration—is improper, meaning not a point of continuity of the integrand. We also consider cases where there is some discontinuity, for instance a vertical asymptote, inside the interval of integration. For any such internal point or endpoint, the basic idea is to “sneak up” on that point, seeing what value an approximating integral approaches if we replace that point by a variable endpoint, which then approaches the offending point using a limit process.

9.4.1 First Examples

Example 9.4.1 Compute \( \int_1^\infty \frac{1}{x^2} \, dx \).

Solution: This integral is “improper” at \( x = \infty \), because we can not claim \( 1/x^2 \) is continuous on \([1, \infty)\), since \( \infty \) is not a proper included endpoint for such an interval. Instead we consider \( \int_1^\beta \frac{1}{x^2} \, dx \), for \( \beta \in (1, \infty) \), so that \( 1/x^2 \) is continuous on \([1, \beta]\) and the Fundamental Theorem of Calculus applies to \( \int_1^\beta \frac{1}{x^2} \, dx \), and finally we see the limiting behavior of this definite integral as \( \beta \to \infty \):

\[
\lim_{\beta \to \infty} \int_1^\beta \frac{1}{x^2} \, dx = \lim_{\beta \to \infty} \left[ -\frac{1}{x} \right]_1^\beta = \lim_{\beta \to \infty} \left[ -\frac{1}{\beta} - \left( -\frac{1}{1} \right) \right] = 0 + 1 = 1.
\]

From this we conclude that

\[
\int_1^\infty \frac{1}{x^2} \, dx = 1.
\]

The strategy is illustrated in Figure 9.4. Because the limit of “proper” integrals exists, we express its meaning the following equivalent ways:

---

6The usual term is “limit” of integration, which is somewhat unfortunate because it can erroneously conjure an association with limits in the sense of Chapter 3 (“Limits and Continuity”). In fact in this section we will use a lot of limit techniques from that chapter and the previous sections of this chapter, but not always on the original boundaries of our integration, so the phrase “limits of integration” is avoided here.
9.4. IMPROPER INTEGRALS

Figure 9.5: Illustration for Example 9.4.2, analyzing \( \int_1^{\beta} \frac{1}{x} \, dx \) by computing the limit of \( \int_1^{\beta} \frac{1}{x} \, dx \) as \( \beta \to \infty \). In this case the limit is infinite, so the integral diverges. To be more precise, the integral diverges to infinity.

- \( \int_1^{\infty} \frac{1}{x^2} \, dx \) converges to 1; or, simply
- \( \int_1^{\infty} \frac{1}{x} \, dx = 1. \)

The term convergence appears in a few different contexts in this textbook, namely with improper integrals, sequences and series. To converge in our context here means that the relevant limit exists as a finite number. To not converge is to diverge, meaning either the limit does not exist, or is not finite. More concisely, that an improper integral diverges means that the limit does not exist as a finite number.

Note that it is quite possible for the total area to be finite, even though the length of the interval is infinite. With an integral such as in Example 9.4.1 above, what matters is whether or not the function shrinks fast enough as \( x \to \infty \) (so the limit for \( \beta \to \infty \) is finite).

Example 9.4.2 Determine if \( \int_1^{\infty} \frac{1}{x} \, dx \) converges or diverges, and if it converges compute its value.

Solution: The technique is the same as the previous example’s: we look at closed intervals \([1, \beta]\), on which the integrand \( 1/x \) is continuous so the Fundamental Theorem of Calculus applies, and we compute the limit of these integrals as \( \beta \to \infty \):

\[
\lim_{\beta \to \infty} \int_1^{\beta} \frac{1}{x} \, dx = \lim_{\beta \to \infty} \ln x \bigg|_1^{\beta} = \lim_{\beta \to \infty} \ln \beta \xrightarrow{\infty} \infty.
\]

Therefore \( \int_1^{\infty} \frac{1}{x} \, dx \) diverges.

To be more precise, we would say the above integral diverges to infinity. For practical reasons we are mainly interested in integrals that converge (to finite numbers). Moreover, there are those that diverge, but not to \( \infty \) or \(-\infty\). All nonconvergent cases like these can be put into one class and labeled as “divergent.”

9.4.2 The Background and Method

Recall that the Fundamental Theorem of Calculus tells us, among other things, that if \( f(x) \) is continuous on the (finite, closed) interval \([a, b]\) (\(a\) and \(b\) being real numbers), then
CHAPTER 9. IMPROPER INTEGRALS AND ADVANCED LIMIT TECHNIQUES

Figure 9.6: Illustration for Example 9.4.3, analyzing \( \int_0^1 \frac{1}{\sqrt{x}} \, dx \) by computing the limit of \( \int_0^\alpha \frac{1}{\sqrt{x}} \, dx \) as \( \alpha \to 1^+ \). In this case the limit is finite, namely 2, so the integral converges to 2.

- there exists a continuous function \( F(x) \) defined on \([a, b]\) (and beyond the endpoints if so desired), such that for all \( x \in (a, b) \) we have \( F'(x) = f(x) \), and
  \[
  \lim_{\Delta x \to 0^+} \frac{F(a + \Delta x) - F(a)}{\Delta x} = f(a),
  \]
  \[
  \lim_{\Delta x \to 0^-} \frac{F(b + \Delta x) - F(b)}{\Delta x} = f(b).
  \]
- Moreover, \( \int_a^b f(x) \, dx = F(b) - F(a) \).

(Recall that, technically \( \int_a^b f(x) \, dx \) denotes a limit of Riemann Sums over \([a, b]\) with \( \max\{\Delta x_i\} \to 0 \).) So the Fundamental Theorem of Calculus requires a closed interval of the form \([a, b]\), and a function which is continuous on that closed interval. Two types of problems with this are dealt with here:

1. one or both of the endpoints is not finite, as in our first two examples above, or
2. the function may have a discontinuity in \([a, b]\), at an endpoint or internally (in \((a, b)\)).

In fact, an integral may have several such features to deal with in assigning a value to it. In all cases, the strategy is to break the integral into subintegrals in which at most one endpoint is “improper” (infinite or a discontinuity, the latter usually a vertical asymptote), and then for each such subintegral we look at definite integrals with one variable endpoint approaching the “improper” endpoint of the subintegral, from the correct side, and such that the Fundamental Theorem of Calculus can be applied to the approximating subintegrals.

We now illustrate this with several more such improper integrals. Our next integral is improper because one of the endpoints is improper. See Figure 9.6.

**Example 9.4.3** Compute, if it converges, \( \int_0^1 \frac{1}{\sqrt{x}} \, dx \).

**Solution:** Here the integral is improper at \( x = 0 \), where the function has a vertical asymptote. We thus look at definite integrals on \([\alpha, 1]\) where \( \alpha \in (0, 1) \), and so we can use the Fundamental Theorem of Calculus with this function on these intervals, and let \( \alpha \to 0^+ \).

\[
\lim_{\alpha \to 0^+} \int_0^1 \frac{1}{\sqrt{x}} \, dx = \lim_{\alpha \to 0^+} 2 \sqrt{x}\bigg|_0^1 = \lim_{\alpha \to 0^+} (2 - 2 \sqrt{\alpha}) = 2 - 0 = 2
\]

\[
\Rightarrow \int_0^1 \frac{1}{\sqrt{x}} \, dx = 2.
\]
This last example again shows that, while one dimension of the area may extend without bounds (upwards in this case as $x \to 0^+$), it is possible for the total area to be finite, in this case because of the speed at which the function’s graph approaches the vertical asymptote as the function’s height increases.

Example 9.4.4 Compute, if possible, $\int_{-\infty}^{\infty} \frac{1}{x^2 + 1} \, dx$.

Solution: The integrand is continuous on all of $\mathbb{R}$, since $x^2 + 1 > 0$, but both endpoints are improper. To deal with this we break the range of integration into subintervals in which at most one endpoint is improper. So we look at the following, but taken provisionally (see Figure 9.7):

$$\int_{-\infty}^{\infty} \frac{1}{x^2 + 1} \, dx = \int_{-\infty}^{0} \frac{1}{x^2 + 1} \, dx + \int_{0}^{\infty} \frac{1}{x^2 + 1} \, dx.$$ 

By provisionally we mean that, if both integrals on the right converge, then the equation makes sense and is true. If either of the integrals on the right diverges, then we declare that the original integral diverges. Now we look at these two integrals separately.

$$\int_{-\infty}^{0} \frac{1}{x^2 + 1} \, dx : \quad \lim_{\alpha \to -\infty} \int_{\alpha}^{0} \frac{1}{x^2 + 1} \, dx = \lim_{\alpha \to -\infty} \tan^{-1} x \bigg|_{\alpha}^{0} = \lim_{\alpha \to -\infty} \left( \tan^{-1} 0 - \tan^{-1} \alpha \right) = 0 - \frac{-\pi}{2} = \frac{\pi}{2},$$

$$\int_{0}^{\infty} \frac{1}{x^2 + 1} \, dx : \quad \lim_{\beta \to 0} \int_{0}^{\beta} \frac{1}{x^2 + 1} \, dx = \lim_{\beta \to 0} \tan^{-1} x \bigg|_{0}^{\beta} = \lim_{\beta \to 0} \left( \tan^{-1} \beta - \tan^{-1} 0 \right) = \frac{\pi}{2} - 0 = \frac{\pi}{2}.$$

Since both subintegrals converge to finite numbers, the original integral will be their sum:

$$\int_{-\infty}^{\infty} \frac{1}{x^2 + 1} \, dx = \int_{-\infty}^{0} \frac{1}{x^2 + 1} \, dx + \int_{0}^{\infty} \frac{1}{x^2 + 1} \, dx = \frac{\pi}{2} + \frac{\pi}{2} = \pi.$$
In this last example we could also make use of the symmetry to get that the two subintegrals are the same, so
\[ \int_{-\infty}^{\infty} \frac{1}{x^2 + 1} \, dx = 2 \int_{0}^{\infty} \frac{1}{x^2 + 1} \, dx = 2 \cdot \frac{\pi}{2} = \pi. \]
However, the first “=” in the line above is also provisional, valid only if the second integral converges.

It should also be pointed out that the endpoint 0 common to both subintegrals was, theoretically, arbitrary. We could have just as easily used \( \int_{1}^{\infty} + \int_{-\infty}^{1} \), or any other real number; the goal is to first break the original integral into two on which there is only one improper endpoint.

**Example 9.4.5** Next we list several examples of how we would have to partition some improper integrals into subintegrals, each having at most one improper endpoint.

Note that each equation below is provisional, i.e., true if each of the subintegrals converges (to a finite number). In the interests of space and clarity, we will not write the integrands or differentials beyond the initial statement of the integral in question.

- \( \int_{-\infty}^{\infty} \frac{1}{x^2 + 1} \, dx = \int_{-\infty}^{0} \, + \int_{0}^{\infty}, \) Improper at \(-\infty, \infty\).
- \( \int_{-1}^{1} \frac{1}{x^{1/3}} \, dx = \int_{-1}^{0} \, + \int_{0}^{1}, \) Improper at 0.
- \( \int_{-\infty}^{\infty} \frac{1}{x^3} \, dx = \int_{-\infty}^{-1} \, + \int_{-1}^{0} \, + \int_{0}^{1} \, + \int_{1}^{\infty}, \) Improper at \(-\infty, 0, \infty\).
- \( \int_{0}^{\infty} \frac{1}{x \ln x} \, dx = \int_{0}^{1/2} \, + \int_{1/2}^{1} \, + \int_{1}^{\infty}, \) Improper at 0,1,\( \infty\).
- \( \int_{0}^{4} \frac{1}{(x-1)(x-2)} \, dx = \int_{0}^{1} \, + \int_{1}^{3/2} \, + \int_{3/2}^{2} \, + \int_{2}^{4}, \) Improper at 1,2.

**9.4.3 Infinite Areas do not Cancel**

One has to be careful in making symmetry arguments, especially if claiming that two parts of an integral “cancel.” It is not too difficult to see that
\[ \int_{-\alpha}^{\alpha} \frac{x}{x^2 + 1} \, dx = 0 \]
for any \( \alpha \in \mathbb{R} \). While we actually compute this below, it is also easy to visualize for \( \alpha > 0 \), because the integrand is an odd function. See Figure 9.8.\(^7\) However, we can not simply let \( \alpha \to \infty \) and conclude that \( \int_{-\infty}^{\infty} \frac{x}{x^2 + 1} \, dx \) will also be zero. The trouble is that the integral on \((-\infty, 0]\) diverges to \(-\infty\), while the integral on \([0, \infty)\) diverges to \(\infty\), so we can not add them and claim that they cancel. To see why, consider the following computations:

\(^7\)Recall that \( f \) is an odd function if and only if \( f(-x) = -f(x) \), for all \( x \) in the domain of \( f \).
Figure 9.8: Illustration for \( \int_{-\infty}^\infty \frac{x}{x^2 + 1} \, dx \). This is improper at both \(-\infty\) and \(\infty\), so we must consider separately \( \int_0^\alpha \frac{x}{x^2 + 1} \, dx \) and \( \int_\beta^0 \frac{x}{x^2 + 1} \, dx \), respectively letting \( \alpha \to -\infty \) and \( \beta \to \infty \). It turns out the first diverges to \( -\infty \) and the second to \( \infty \), so the original integral also diverges. It would be incorrect to only look at \( \int_\alpha^- \frac{x}{x^2 + 1} \, dx \), and let one limit collectively compute the sum of each, because it would depend heavily on the fact that we would be counting “areas,” both positive and negative, precisely the way that they would cancel. However, the areas could be counted other ways collectively to achieve other results—for instance \( \int_\alpha^{2\alpha} \frac{x}{x^2 + 1} \, dx \to \ln 2 \neq 0 \) (see text). Though both appear to the naïve observer to “cover” the entire interval \((-\infty, \infty)\) in the limit, they do not return the same “total area.” For these and other reasons we decline to assign a value to this improper integral, and instead declare it to be divergent.

\[
\begin{align*}
\lim_{\alpha \to -\infty} \int_{-\alpha}^\alpha \frac{x}{x^2 + 1} \, dx &= \lim_{\alpha \to -\infty} \frac{1}{2} \ln(x^2 + 1) \bigg|_{-\alpha}^\alpha = \lim_{\alpha \to -\infty} \left[ \frac{1}{2} \ln(\alpha^2 + 1) - \frac{1}{2} \ln(\alpha^2 + 1) \right] = 0. \\
\lim_{\alpha \to -\infty} \int_{-\alpha}^{2\alpha} \frac{x}{x^2 + 1} \, dx &= \lim_{\alpha \to -\infty} \frac{1}{2} \ln(4\alpha^2 + 1) \bigg|_{-\alpha}^{2\alpha} = \lim_{\alpha \to -\infty} \frac{1}{2} \left[ \ln(4\alpha^2 + 1) - \ln(\alpha^2 + 1) \right] \\
&= \lim_{\alpha \to -\infty} \frac{1}{2} \ln \frac{4\alpha^2 + 1}{\alpha^2 + 1} = \lim_{\alpha \to -\infty} \ln \sqrt{\frac{4\alpha^2 + 1}{\alpha^2 + 1}} = \ln \sqrt{4} = \ln 2.
\end{align*}
\]

Note that both integrals cover all of \((-\infty, \infty)\) in the limits, yet they return different values. So the “total area” is different if you count faster on one side of zero than the other. For this and other reasons, we must look at two separate subintegrals for this problem.

First we analyze \( \int_0^\infty \frac{x}{x^2 + 1} \, dx \).

\[
\lim_{\beta \to -\infty} \int_0^\beta \frac{x}{x^2 + 1} \, dx = \lim_{\beta \to -\infty} \frac{1}{2} \ln(x^2 + 1) \bigg|_0^\beta = \lim_{\beta \to -\infty} \frac{1}{2} \left[ \ln(\beta^2 + 1) - \ln 1 \right] = \infty.
\]

Thus \( \int_0^\infty \frac{x}{x^2 + 1} \, dx \) diverges to \( \infty \). By a similar computation or symmetry, \( \int_0^\infty \frac{\alpha}{x^2 + 1} \, dx \) diverges to \(-\infty\), but that is not necessary to notice this to conclude that \( \int_{-\infty}^\infty \frac{x}{x^2 + 1} \, dx \) diverges. Indeed, it was enough that the subintegral \( \int_0^\infty \frac{x}{x^2 + 1} \, dx \) diverges, to conclude the full integral \( \int_{-\infty}^\infty \frac{x}{x^2 + 1} \, dx \) does so as well. (The difference is that the full integral does not diverge towards either \(-\infty\) or \(\infty\).)
9.4.4 More Complicated Integrals

In this subsection we examine ways to organize our solution of an improper integral problem if the antidifferentiation or limit steps are complicated. There are basically five steps to any improper integral computation (some with substeps):

1. Identify improper points within the range of integration, and if necessary break the integral into subintegrals for which only one endpoint is improper.

2. For each of these subintegrals, write a definite integral on a closed interval in which one endpoint is fixed, and the other is allowed to vary, approaching the improper endpoint from within the interval.

Perform the following two steps collectively, one subintegral at a time, and if any subintegral diverges, we can stop and declare the original integral diverges.

3. Compute the relevant antiderivative for the subintegral.

4. Compute the limit as the variable endpoint approaches the improper endpoint’s value from within the interval where the function is continuous (using the Fundamental Theorem of Calculus and limit techniques).

5. If all subintegrals converge, sum them to get the original integral’s value.

The antidifferentiation step may be complicated enough that it is advantageous to compute the antiderivative without the endpoints present. Similarly a limit problem may arise which is better not done inline with the general flow of the logic.

Example 9.4.6 Compute \( \int_{0}^{1} \ln x \, dx \). (See Figure 9.9.)

\textbf{Solution:} This is improper at \( x = 0 \), so we will eventually look at \( \lim_{\alpha \to 0^+} \int_{\alpha}^{1} \ln x \, dx \), but first we need to compute the relevant antiderivative.
9.4. IMPROPER INTEGRALS

\[
\int \ln x \, dx = x \ln x - \int x \cdot \frac{1}{x} \, dx = x \ln x - \int dx = x \ln x - x + C.
\]

\[
u = \ln x \quad dv = dx
\]
\[
\begin{array}{c}
u = \ln x \\
dv = \frac{1}{x} \, dx
\end{array}
\]

Now we write, provisionally,

\[
\int_0^1 \ln x \, dx = \lim_{\alpha \to 0^+} \int_\alpha^1 \ln x \, dx = \lim_{\alpha \to 0^+} \left( (x \ln x - x) \right)_{\alpha}^1 = \lim_{\alpha \to 0^+} (0 - 1 - \alpha \ln \alpha - \alpha).
\]

The only term in that limit that needs detailed analysis is the \( \alpha \ln \alpha \) term, so rather than attempting to wrap it into the other terms (in one unnecessarily complicated fraction), we compute it separately.

\[
\lim_{\alpha \to 0^+} \alpha \ln \alpha = \lim_{\alpha \to 0^+} \frac{\ln \alpha}{\frac{1}{\alpha}}.
\]

\[
\begin{align*}
\text{ALG} & \quad \lim_{\alpha \to 0^+} \alpha \ln \alpha \\
\text{LHR} & \quad \lim_{\alpha \to 0^+} \frac{\ln \alpha}{\frac{1}{\alpha}} = \lim_{\alpha \to 0^+} (-\alpha) = 0.
\end{align*}
\]

Filling this result into the limit above it, we get

\[
\int_0^1 \ln x \, dx = \lim_{\alpha \to 0^+} \int_\alpha^1 \ln x \, dx = \lim_{\alpha \to 0^+} (0 - 1 - \alpha \ln \alpha - \alpha) = 0 - 1 - 0 - 0 = -1.
\]

When we bring techniques from several different topics in calculus to bear on a single problem like that above, it is often useful to work the various dimensions of the problem somewhat separately. So the problem above required that we (1) identify improper points in the range of integration, (2) set up the appropriate limits of integrals, (3) compute the relevant antiderivative(s), (4) compute the limits that emerge from that process, and (5) draw a conclusion from all of that work. Since the antiderivative can be a relatively extensive computation itself, it is better to not attempt it on a definite integral (with endpoints). Computing the desired limits in the same line as the antiderivative computation can also be difficult, especially since there may be several terms in the limits, some of which may need techniques such as l’Hôpital’s Rule to compute. One is led to conclude that it is better to break these tasks into sequestered subtasks until their results are needed.

**Example 9.4.7** Compute \( \int_0^\infty e^{-x} \sin 4x \, dx \). (See Figure 9.10.)

**Solution:** As usual we consider \( \int_0^\beta e^{-x} \sin 4x \, dx \), and let \( \beta \to \infty \). However it is nontrivial to compute the needed antiderivative by hand, so we make that a “problem within the problem.” Recall that this is the type of integral which we integrate by parts twice, and then have an
equation involving the integral which we can solve for the integral.

\[
(I) = \int e^{-x} \sin 4x \, dx = u v - \int v \, du = -e^{-x} \sin 4x + 4 \int e^{-x} \cos 4x \, dx
\]

\[
\begin{align*}
\frac{du}{dx} &= 4 \cos 4x \\
\frac{dv}{dx} &= e^{-x} \\
\frac{du}{dx} &= -4 \sin 4x \\
\frac{dv}{dx} &= -e^{-x}
\end{align*}
\]

\[
I = -e^{-x} \sin 4x + 4 \left[ uv - \int v \, du \right]
\]

\[
\begin{align*}
&= -e^{-x} \sin 4x + 4 \left[ -e^{-x} \sin 4x + 4 \left( -e^{-x} \cos 4x - 4 \int e^{-x} \sin 4x \, dx \right) \right] \\
&= -e^{-x} \sin 4x - 4e^{-x} \cos 4x - 16(I)
\end{align*}
\]

\[
\begin{align*}
\Rightarrow 17(I) &= -e^{-x} \sin 4x + 4 \cos 4x + C_1 \\
\Rightarrow (I) &= \frac{-e^{-x}}{17} \left( \sin 4x + 4 \cos 4x \right) + C.
\end{align*}
\]

With the relevant antiderivative known, we now can now compute

\[
\lim_{\beta \to \infty} \int_0^\beta e^{-x} \sin 4x \, dx = \lim_{\beta \to \infty} \left. \frac{-e^{-x}}{17} \left( \sin 4x + 4 \cos 4x \right) \right|_0^\beta
\]

\[
= \lim_{\beta \to \infty} \left[ \frac{-e^{-\beta}}{17} \left( \sin 4\beta + 4 \cos 4\beta + \frac{1}{17}(0 + 4) \right) \right]
\]

\[
= \lim_{\beta \to \infty} \frac{\sin 4\beta + 4 \cos 4\beta}{17e^{\beta}} + \frac{4}{17}
\]

\[
\overset{(B/\infty) + 1/17}{=} 0 + \frac{4}{17} = \frac{4}{17}.
\]

From this we get \[
\int_0^\infty e^{-x} \sin 4x \, dx = \frac{4}{17}.
\]

This last example required more difficult antidifferentiation steps. We could also consider integrals with more points where they are improper, but they are more rare in practice. Usually we only need to worry about one or two points where an integral is improper, and once we discover them, we simply break the original into subintegrals with one improper endpoint each, and use a limiting process as developed here.
Examples of applications include certain total work (energy) computations. If we have a force such as gravity, between two stationary objects, varying as the distance squared between them, we have force given by \( \frac{k}{x^2} \), where \( x \) is that distance and \( k \) is a constant. Then the (infinitesimal) work to move one of the objects the distance \( dx \) at the position \( x \) is given by \( \frac{k}{x^2} dx \) (or, putting it roughly, force times distance). Next we note that the work required to pull the second object infinitely far from the first would be given by

\[
\int_{s_0}^{\infty} \frac{k}{x^2} dx = \lim_{\beta \to \infty} \int_{s_0}^{\beta} \frac{k}{x^2} dx = \lim_{\beta \to \infty} \left[ -\frac{k}{\beta} + \frac{k}{s_0} \right] = \frac{k}{s_0},
\]

where \( s_0 > 0 \) is the original distance between them. Just as in Example 9.4.1 at the beginning of this section (page 688), we have that the integral is finite. This total work can be equated to the kinetic energy it would need to “escape” the stationary object, and from that kinetic energy \( \frac{1}{2}mv^2 \) (where \( m = \text{mass} \) we could find the necessary velocity of the second object for it to have that energy (ability to do the work) in kinetic form. The required velocity \( v \) is then called the escape velocity.

In the definition of derivative,

\[
f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}
\]

we needed a limit to “break through” the fact that we would really like \( \Delta x \) to be zero (but it can not). So too do we need limits to “break through” improper points in a definite integral, where otherwise the Fundamental Theorem of Calculus does not hold (because of a lack of continuity of the function, a key requirement of the theorem). For escape velocity, we have to break through the fact that you need to be able to compute how much energy is needed for the second object to move away from the first “forever,” and not be recaptured. Other physical applications can also be found for improper integrals, but we will not pursue any more in this chapter.

### 9.4.5 A Complication With Oscillatory and Other Integrals

This subsection makes a point which is usually left to graduate-level real analysis courses (and so the reader should not be too anxious if it seems overly technical). This point has been hinted at previously, and is only included here for full disclosure. The point is that any integral is only said to converge if we can rearrange the order in which we count the signed area between the curve and the \( x \)-axis any way we like, and still come out with the same expression for total signed area.

So for instance, if an improper integral is to converge we should be able to add up all the positive area first, and then add up the negative area separately, and when we combine them we get a finite expression which is the same as if we found another way to sum the areas.

Consider again Example 9.4.7, page 696. There are infinitely many subintervals on which that function \( e^{-x} \sin 4x \) is positive, and infinitely many on which it is negative. Our limit, however, has us sliding \( \beta \) towards infinity (\( \beta \to \infty \), along the way allowing the areas counted to alternate between positive and negative, partially cancelling each other in turn as \( \beta \) increases. It is not clear at this point that, in fact, the positive areas alone sum to a finite number, as do the negative areas. This is equivalent\(^8\) to the statement that it does not matter in which order the areas are summed, as long as any chosen area in the integral is eventually part of the sum; we will always get the same answer for \( \int_{0}^{\infty} e^{-x} \sin 4x \, dx \).

\(^8\)That these two ideas are equivalent is a fact usually pondered for some time by the average mathematics graduate student.
A general theory can be proposed, but first a little notation is necessary. Rather than giving endpoints (a.k.a. limits) of integration, one notation has the integral written as “over the set.” So for instance, the definite integral of \( f(x) \) with respect to \( x \) over the interval \([a, b]\) is written

\[
\int_a^b f(x) \, dx = \int_{[a,b]} f(x) \, dx.
\]

If there is a function whose area is defined by disjoint intervals, this second notation is better. For instance, if we would like to find the combined area under \( y = x^2 \) for \(-2 \leq x \leq -1\) and \(1 \leq x \leq 2\), we write one of the following:

\[
\int_{-2}^{-1} x^2 \, dx + \int_1^2 x^2 \, dx = \int_S x^2 \, dx,
\]

where \( S = [-2, -1] \cup [1, 2] \).

With this notation we can mention that the necessary and sufficient condition for an improper integral of a function \( f(x) \) over a set \( S \) to converge is

\[
\int_S |f(x)| \, dx < \infty. \tag{9.3}
\]

Before explaining why this is reasonable, first we make two subintegrals:

\[
\int_S |f(x)| \, dx = \int_{S_+} |f(x)| \, dx + \int_{S_-} |f(x)| \, dx \tag{9.4}
\]

where \( S_+ = \{x \mid f(x) \geq 0\} \) and \( S_- = \{x \mid f(x) < 0\} \).

This integral \( \int_S |f(x)| \, dx \) does not allow the positive and negative parts of the original integral to cancel: if the positive area is infinite, the total integral is greater than or equal to the integral over \( S_+ \), and is therefore infinite as well. Similarly with the negative part (though its sign is changed in (9.3) and (9.4)).

A function that satisfies (9.3) is called \( L^1(S) \) (pronounced, “ell-one of \( S \)”), and will also satisfy that \( \int_S f(x) \, dx \) converges, and we can use the limiting process for improper integrals if it is indeed improper. Now

\[
\int_0^\infty e^{-x} \sin 4x \, dx \leq \int_0^\infty e^{-x} \, dx = \lim_{\beta \to \infty} \int_0^\beta e^{-x} \, dx = \lim_{\beta \to \infty} [-e^{-\beta} + e^0] = 1 < \infty,
\]

so \( e^{-x} \sin 4x \in L^1([0, \infty)) \), and we can compute the integral as we did in the example. However,

\[
\int_1^\infty \frac{\sin x}{x^2} \, dx \quad \text{converges}, \quad \text{while} \quad \int_1^\infty \frac{\sin x}{x} \, dx \quad \text{diverges},
\]

though for both cases \( \lim_{\beta \to \infty} \left( \int_1^\beta f(x) \, dx \right) \) exists, as we will see when we study “alternating series” later. The reason for the different outcomes is that the first integral’s convergence does not rely on the sequential cancellation from the alternating \((+/-)\) areas (but on the shrinking in absolute value of the integrand), where the second does.

It is beyond the scope of the text to pursue \( L^1\)-function theory further, but the same themes will come up for a time when we develop alternating series and absolute convergence.\(^9\)

\(^9\)For reference, \( f \in L^p(S) \) means \( \int_S |f(x)|^p \, dx < \infty \). When \( p = 1 \) we have our previous definition of \( L^1(S) \).
Chapter 10

Series of Constants

In this chapter we consider “infinite sums,” which we call series, (in both the singular and the plural) such as
\[ \sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \cdots. \] (10.1)
The “sum” above begins with \(a_1\), but we will often begin with a term \(a_0\), or \(a_2\), etc. It is not the beginning terms which determine if we can in fact compute such a sum, but rather it is the infinite “tail” of the series. This is reasonable because we can always, in principle, add as many terms together as we like, so long as there are finitely many of them. As with other calculus concepts, the tool which breaks the finite/infinite barrier is limit. Indeed, to make sense of a sum such as (10.1), we consider the \(N\)th partial sum,
\[ S_N = \sum_{n=1}^{N} a_n = a_1 + a_2 + \cdots + a_{N-1} + a_N, \] (10.2)
and then look at the sequence of these partial sums \(S_1, S_2, S_3, \cdots\), i.e.,
\[ S_1 = a_1, \]
\[ S_2 = a_1 + a_2, \]
\[ S_3 = a_1 + a_2 + a_3, \]
\[ S_4 = a_1 + a_2 + a_3 + a_4, \]
and so on. To determine if (10.1) makes sense is then considered (by definition) to be equivalent to determining the behavior of the sequence \(\{S_N\}_{N=1}^{\infty}\). We say the series (10.1) converges to \(S \in \mathbb{R}\) if and only if \(S_N \to S\) as \(N \to \infty\).

In a few cases we will actually be able to compute a simple formula for \(S_N\), and thus be able to compute the series by taking \(N \to \infty\). However, in many cases we cannot find a compact formula for \(S_N\). In those cases we have to develop other methods for determining if the series converges at all, and if so, how to approximate the value of the series with as much precision as we require (short of exactness), by determining how large we require \(N\) to be so that we can approximate the full series by \(S_N\).

In this text we will take more steps than most other texts in developing the theory of series, since this topic is the source of much confusion for students. Indeed we devote this entire chapter to the topic of series of constant terms, leaving nonconstant terms for their own chapter. The concepts are intuitive—at times perhaps deceptively so—but require practice so that,
for example, one can recognize when and where to apply a particular test of convergence or divergence.

In the next chapter we will look at functions defined by series

\[
f(x) = \sum_{n=0}^{\infty} a_n (x - a)^n = a_0 + a_1 (x - a) + a_2 (x - a)^2 + a_3 (x - a)^3 + \cdots.
\]  

(10.3)

In fact most functions we have dealt with in this text can be written in the form above, at least on some open intervals, so such functions are very important theoretically. However there are very important functions which can be written in the form (10.3) but not by using the functions from our usual library (powers, logs, exponential, trigonometric and arctrigonometric functions). We deal extensively with functions of the form (10.3), also known as power series, in the next chapter.

One such function which can be represented by a series of form (10.3), as we will see, is that antiderivative \( F(x) \) of \( e^{x^2} \) whose graph passes through the origin, i.e., so that \( F(0) = 0 \). We will see how this can be given by the following, with \( a = 0 \) in (10.3):

\[
F(x) = \int_0^x e^{t^2} \, dt = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)n!}
\]

\[
= \frac{x}{1 \cdot 0!} + \frac{x^3}{3 \cdot 1!} + \frac{x^5}{5 \cdot 2!} + \frac{x^7}{7 \cdot 3!} + \cdots
\]

\[
= x + \frac{x^3}{3} + \frac{x^5}{10} + \frac{x^7}{42} + \cdots.
\]

While this might appear complicated or intimidating, in fact it is a rather simple computation, though we must build some theory first. Of course, before dealing with series with variable terms, we must first develop a theory of series of constants, to which we devote the rest of this chapter.
10.1 Series and Partial Sums

As mentioned in the introduction to this chapter, the convergence of a series is defined as equivalent to the convergence of its partial sums. For convenience, we will define the Nth partial sum to be the sum of all terms of the underlying sequence up to the term whose subscript is N. Thus if the sequence is series \( \sum_{n=k}^{\infty} a_n \), with underlying sequence \( \{a_n\}_{n=k}^{\infty} \), then

\[
S_N = \sum_{n=k}^{N} a_n = a_k + a_{k+1} + \cdots + a_N.
\]  

(10.4)

So for a series \( a_0 + a_1 + a_2 + \cdots \), the partial sum \( S_N = a_0 + a_1 + \cdots + a_N \) would actually have \( N + 1 \) terms, though we will still call it the \( N \)th partial sum. (Of course if \( N < k \) we do not define an \( N \)th partial sum.)

Example 10.1.1 Consider the series \( \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2 + 1} \), and find the first five partial sums.

Solution: We do this directly:

\[
S_1 = \sum_{n=1}^{1} \frac{(-1)^n}{n^2 + 1} = \frac{(-1)^1}{1^2 + 1} = \frac{-1}{2}
\]

\[
S_2 = \sum_{n=1}^{2} \frac{(-1)^n}{n^2 + 1} = \frac{(-1)^1}{1^2 + 1} + \frac{(-1)^2}{2^2 + 1} = \frac{-1}{2} + \frac{1}{5} = \frac{-3}{10} = 0.3
\]

\[
S_3 = \sum_{n=1}^{3} \frac{(-1)^n}{n^2 + 1} = \frac{(-1)^1}{1^2 + 1} + \frac{(-1)^2}{2^2 + 1} + \frac{(-1)^3}{3^2 + 1} = \frac{-3}{10} + \frac{-1}{10} = \frac{2}{10} = 0.2
\]

\[
S_4 = \sum_{n=1}^{4} \frac{(-1)^n}{n^2 + 1} = \frac{(-1)^1}{1^2 + 1} + \frac{(-1)^2}{2^2 + 1} + \frac{(-1)^3}{3^2 + 1} + \frac{(-1)^4}{4^2 + 1} = S_3 + \frac{1}{17} = \frac{44}{170} \approx 0.2588235294
\]

\[
S_5 = \sum_{n=1}^{5} \frac{(-1)^n}{n^2 + 1} = S_4 + \frac{(-1)^5}{5^2 + 1} = \frac{44}{170} + \frac{-1}{26} \approx 0.220361991.
\]

Note we used the simple recursion relationship for partial sums of a series: given a series \( \sum_{n=k}^{\infty} a_n \), and \( N \geq k \) we have

\[
S_{N+1} = S_N + a_{N+1},
\]

(10.5)

that is,

\[
S_{N+1} = \sum_{n=k}^{N+1} a_n = a_k + a_{k+1} + \cdots + a_N + a_{N+1} = \sum_{n=k}^{N} a_n + a_{N+1} = S_N + a_{N+1}, \text{ q.e.d.}
\]

In a later section we will see that the series in the above example does in fact converge, though we can only approximate its exact value here by computing \( S_N \) for large values of \( N \).
10.1.1 Telescoping Series

Telescoping series do occur on occasion, but the main reason they are included in most calculus textbooks is that their partial sums simplify in nice ways, leaving us able to compute their limits and thus the whole series. Indeed, the behavior of telescoping series is unusually “nice”—rivaled only by that of the much more important geometric series we will see later in this section—and therefore well-suited for early examples of the general notion of series.

The simplest type of telescoping series is one in which the terms added are themselves sums of two terms, constructed in such a way that there is cancellation such as the following:

\[
\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} \left[ b_n - b_{n-1} \right] = (b_1 - b_0) + (b_2 - b_1) + (b_3 - b_2) + \cdots \quad (10.6)
\]

After a careful examination of the terms which appear in (10.6), it seems that all cancel except for \( -b_0 \). However we must be even more careful since there are infinitely many terms we are claiming we can cancel. The correct approach is to carefully examine the partial sums:

\[
S_1 = b_1 - b_0,
\]
\[
S_2 = b_1 - b_0 + b_2 - b_1 = b_2 - b_0,
\]
\[
S_3 = b_1 - b_0 + b_2 - b_1 + b_3 - b_2 = b_3 - b_0,
\]
\[
S_4 = b_1 - b_0 + b_2 - b_1 + b_3 - b_2 + b_4 - b_3 = b_4 - b_0,
\]
and so on, whereby we can conclude that, for this simplest type of example (10.6), we have

\[
S_n = b_n - b_0. \quad (10.7)
\]

Now such a series will therefore converge if and only if \( \left\{ b_n \right\}_{n=1}^{\infty} \) converges. If \( b_n \to B \in \mathbb{R} \) as \( n \to \infty \), then by (10.7) we have \( S_n \to B - b_0 \), whence \( \sum_{n=1}^{\infty} [b_n - b_{n-1}] = B - b_0 \).

More complicated telescoping series also occur, though the basic idea is that the partial sums can be written in such a way that all but a few terms found in the partial sums eventually cancel, and where we can compute the limits of those terms which do not.\(^1\) Rather than memorizing the sample telescoping forms (10.6) and (10.7), it is better to consider each example separately, writing out the terms of \( S_N \) for enough values of \( N \) that the pattern emerges.

**Example 10.1.2** Consider the series \( \sum_{n=1}^{\infty} \left[ \frac{1}{n+1} - \frac{1}{n} \right] \). Compute the form of each partial sum \( S_N \) (as a function of \( N \)), and the value of the series if it converges.

**Solution:** We will write out a few partial sums longhand, from which the pattern will emerge.

---

\(^1\)It is interesting to visualize why the term *telescoping* is used to describe such a series. One of the *Webster’s* dictionaries defines the intransitive verb form of telescope as follows:

*to slide together, or into something else, in the manner of the tubes of a jointed telescope.*

For another example, a “telescoping antenna” comes to mind. Both can “collapse” to be much shorter than when fully extended. The reader should keep such images in mind as we consider so-called telescoping series.
Indeed, all but two terms will cancel in each of the following.

\[
S_1 = \left[ \frac{1}{2} - 1 \right] = \frac{1}{2} - 1,
\]

\[
S_2 = \left[ \frac{1}{2} - 1 \right] + \left[ \frac{1}{3} - \frac{1}{2} \right] = \frac{1}{3} - 1,
\]

\[
S_3 = \left[ \frac{1}{2} - 1 \right] + \left[ \frac{1}{3} - \frac{1}{2} \right] + \left[ \frac{1}{4} - \frac{1}{3} \right] = \frac{1}{4} - 1,
\]

\[
S_4 = \left[ \frac{1}{2} - 1 \right] + \left[ \frac{1}{3} - \frac{1}{2} \right] + \left[ \frac{1}{4} - \frac{1}{3} \right] + \left[ \frac{1}{5} - \frac{1}{4} \right] = \frac{1}{5} - 1.
\]

From this we do indeed see a pattern in which

\[
S_N = \frac{1}{N+1} - 1.
\]

Taking \( N \to \infty \), we see \( S_N = \frac{1}{N+1} - 1 \to 0 - 1 = -1 \), and so we conclude that the series converges to \(-1\), i.e.,

\[
\sum_{n=1}^{\infty} \left[ \frac{1}{n+1} - \frac{1}{n} \right] = -1.
\]

Sometimes we need to do a little more work to detect a telescoping series, and its formula for \( S_N \). Note that the general term of the added sequence terms, namely \( \frac{1}{n+1} - \frac{1}{n} \), in our series above looks like a partial fraction decomposition if the variable is \( n \). For that reason, when the general term can be written in a PFD, the series may in fact be telescoping. This is the case with the following example.

**Example 10.1.3** Consider the series \( \sum_{n=2}^{\infty} \frac{1}{n^2 - 1} \). Compute a general formula for the \( N \)th partial sum \( S_N \), and compute its limit, if \( S_N \) converges, thereby computing the series.

**Solution:** Note first that there is no \( S_1 \) here. That said, the technique which we will use for this is to first look at the partial fraction decomposition (PFD) for \( \frac{1}{n^2 - 1} \). Of course we need the denominator factored, giving us the form

\[
\frac{1}{n^2 - 1} = \frac{1}{(n+1)(n-1)} = \frac{A}{n+1} + \frac{B}{n-1}.
\]

Multiplying by \((n+1)(n-1)\) in the second equation then gives us

\[
1 = A(n-1) + B(n+1).
\]

Now we use the usual methods for computing the coefficients \( A \) and \( B \):

\[
\begin{align*}
&n = 1: \quad 1 = B(2) \quad \Rightarrow \quad B = \frac{1}{2}; \\
&n = -1: \quad 1 = A(-2) \quad \Rightarrow \quad A = -\frac{1}{2}.
\end{align*}
\]

From this we can rewrite our series

\[
\sum_{n=2}^{\infty} \left[ \frac{-1/2}{n+1} + \frac{1/2}{n-1} \right] = \sum_{n=2}^{\infty} \left[ \frac{1}{2} \left( \frac{-1}{n+1} + \frac{1}{n-1} \right) \right].
\]
CHAPTER 10. SERIES OF CONSTANTS

There is no $S_1$, so we begin with $S_2$. (For space considerations we do not write out all terms at each line.)

\[ S_2 = \frac{1}{2} \left( \frac{-1}{3} + 1 \right) = \frac{1}{2} \left( \frac{-1}{3} + 1 \right), \]

\[ S_3 = \frac{1}{2} \left( \frac{-1}{3} + 1 \right) + \frac{1}{2} \left( \frac{-1}{4} + \frac{1}{2} \right) = \frac{1}{2} \left( \frac{-1}{3} + 1 + \frac{-1}{4} + \frac{1}{2} \right), \]

\[ S_4 = \frac{1}{2} \left( \frac{-1}{3} + 1 \right) + \frac{1}{2} \left( \frac{-1}{4} + \frac{1}{2} \right) + \frac{1}{2} \left( \frac{-1}{5} + \frac{1}{3} \right) = \frac{1}{2} \left( 1 + \frac{1}{2} - \frac{1}{4} - \frac{1}{5} \right), \]

\[ S_5 = S_4 + \frac{1}{2} \left( \frac{-1}{6} + \frac{1}{4} \right) = \frac{1}{2} \left( 1 + \frac{1}{2} - \frac{1}{5} - \frac{1}{6} \right), \]

\[ S_6 = S_5 + \frac{1}{2} \left( \frac{-1}{7} + \frac{1}{5} \right) = \frac{1}{2} \left( 1 + \frac{1}{2} - \frac{1}{6} - \frac{1}{7} \right), \]

\[ S_7 = S_6 + \frac{1}{2} \left( \frac{-1}{8} + \frac{1}{6} \right) = \frac{1}{2} \left( 1 + \frac{1}{2} - \frac{1}{7} - \frac{1}{8} \right), \]

By this point a pattern has clearly emerged, and it can be written

\[ S_N = \frac{1}{2} \left( 1 + \frac{1}{2} - \frac{1}{N} - \frac{1}{N+1} \right), \]

and so $S_N \to \frac{1}{2} \left[ 1 + \frac{1}{2} - 0 - 0 \right] = \frac{1}{2} \cdot \frac{3}{2} = \frac{3}{4}$ as $N \to \infty$. We can thus conclude that

\[ \sum_{n=2}^{\infty} \frac{1}{n^2-1} = \sum_{n=2}^{\infty} \left[ \frac{1}{2} \left( \frac{-1}{n+1} + \frac{1}{n-1} \right) \right] = \frac{3}{4}. \]

Example 10.1.4 Find $S_N$ and discuss the convergence (or divergence) of the series

\[ \sum_{n=0}^{\infty} \sqrt{n+1} - \sqrt{n}. \]

Solution:

\[ S_0 = \sqrt{1} - \sqrt{0} = \sqrt{1} - \sqrt{0}, \]

\[ S_1 = \sqrt{1} - \sqrt{0} + \sqrt{2} - \sqrt{1} = \sqrt{2} - \sqrt{0}, \]

\[ S_2 = \sqrt{1} - \sqrt{0} + \sqrt{2} - \sqrt{1} + \sqrt{3} - \sqrt{2} = \sqrt{3} - \sqrt{0}, \]

\[ S_3 = \sqrt{1} - \sqrt{0} + \sqrt{2} - \sqrt{1} + \sqrt{3} - \sqrt{2} + \sqrt{4} - \sqrt{3} = \sqrt{4} - \sqrt{0}, \]

and so on, so that

\[ S_N = \sqrt{N+1} - \sqrt{0} = \sqrt{N+1} \to \infty \text{ as } N \to \infty. \]

Thus the series diverges (to infinity, to be more descriptive).

Note that we could simplify our earlier expressions for $S_N$, since for instance $\sqrt{0} = 0$, $\sqrt{1} = 1$ and $\sqrt{4} = 2$, but to do so would more likely obscure the pattern of cancellation.
10.1. SERIES AND PARTIAL SUMS

10.1.2 Geometric Series

The class of series considered here is arguably the most important we will encounter. Many important series analyses depend upon how a particular series compares to, or mimics the behavior of, an appropriately chosen geometric series. As with the telescoping series, the geometric series is one for which we can actually compute a general formula for $S_N$, from which we can tell if the series converges, and if so compute its sum.

What makes a series $\sum a_n$ geometric is that there exists a constant $r \in \mathbb{R} - \{0\}$ such that

$$\left(\forall n\right) \left[ \frac{a_{n+1}}{a_n} = r \right]. \quad (10.8)$$

In other words, such a series can be defined recursively by

$$a_{n+1} = r \cdot a_n. \quad (Note \ that \ this \ is \ equivalent \ to \ a_n = r \cdot a_{n-1}, \ so \ long \ as \ a_{n-1} \ is \ defined.)$$

Put more colloquially, a geometric series is one in which we get the next term by multiplying the present term by the same constant each time. Examples of geometric series follow:

- $\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots \quad (r = 1/2),$
- $\sum_{n=2}^{\infty} \frac{6}{5^n} = \frac{6}{25} + \frac{6}{125} + \frac{6}{625} + \cdots \quad (r = 1/5),$
- $\sum_{n=1}^{\infty} \frac{2(-1)^n}{3^n} = -\frac{2}{3} + \frac{2}{9} - \frac{2}{27} + \cdots \quad (r = -1/3),$
- $\sum_{n=1}^{\infty} \frac{1}{3^{2n}} = \frac{1}{9} + \frac{1}{81} + \frac{1}{729} + \cdots \quad (r = 1/9).$

Note that this last series can be rewritten $\sum_{n=1}^{\infty} \frac{1}{3^n}$, or even $\sum_{n=0}^{\infty} \left[ \frac{1}{3} \cdot \left(\frac{1}{3}\right)^n \right]$. In fact, unlike the telescoping series, every geometric series can be written in the same form, namely

$$\sum_{n=0}^{\infty} \alpha r^n = \alpha + \alpha r + \alpha r^2 + \alpha r^3 + \cdots, \quad (10.9)$$

where

$$\alpha \text{ is the first term of the series, and} \quad (10.10)$$

$$r \text{ is the constant ratio, } a_{n+1}/a_n. \quad (10.11)$$

In the examples above, the first terms are $\alpha = 1, 6/25, -2/3, 1/9$ respectively. Each of the series above can be rewritten in $\Sigma$-notation in the form (10.9), starting with $n = 0$. For instance, the third series above can be rewritten, using $\alpha = -2/3$ and $r = -1/3$, as

$$\sum_{n=1}^{\infty} \frac{2(-1)^n}{3^n} = -\frac{2}{3} + \frac{2}{9} - \frac{2}{27} + \cdots = \sum_{n=0}^{\infty} -\frac{2}{3} \left(\frac{-1}{3}\right)^n.$$

In fact, once we know a series is geometric (that is, that $a_{n+1} = r \cdot a_n$ for each $n$), all we need to do is to identify $\alpha$ and $r$, and we can write the series in the exact $\Sigma$-notation form (10.9).

\footnote{With geometric series, it is understood that $\alpha^0$ represents 1, even though technically this is only correct if $r > 0$. In each general setting in which we follow the convention that $r^0$ is defined to be 1 (regardless of the sign of $r$), we will remark on this point.}
Example 10.1.5 Write the series $4 + \frac{2}{9} + \frac{1}{27} + \frac{1}{81} + \cdots$ in the form (10.9).

Solution: Though perhaps not immediately obvious, in fact each successive term is $\frac{1}{6}$ times its immediate predecessor. The first term is 4. We translate these two facts as $\alpha = 4$ and $r = \frac{1}{6}$, and so this series is the same as the series

$$\sum_{n=0}^{\infty} 4 \cdot \left(\frac{1}{6}\right)^n.$$

As with telescoping series, a geometric series allows for a simple formula for $S_N$. To use the formula, however, we need to make two assumptions:

1. that the series is already written in the form $\sum_{n=0}^{\infty} \alpha r^n = \alpha + \alpha r + \alpha r^2 + \alpha r^3 + \cdots$, and

2. that $r \neq 1$.

As we have seen, the first requirement is easy enough to accomplish: we need only identify $\alpha$ (the first term in the geometric series) and $r$. The second requirement is for technical reasons we will encounter momentarily. We do not lose much in assuming $r \neq 1$, since in the case $r = 1$ the series is simply $\alpha + \alpha + \alpha + \cdots$, which is clearly a divergent series if $\alpha \neq 0$, and trivial if $\alpha = 0$.\(^3\) Now we state our theorem.

**Theorem 10.1.1** For a geometric series $\sum_{n=0}^{\infty} \alpha r^n = \alpha + \alpha r + \alpha r^2 + \alpha r^3 + \cdots$, assuming $r \neq 1$, we have

$$S_N = \frac{\alpha(1 - r^{N+1})}{1 - r}.$$  \hspace{1cm} (10.12)

**Proof:** The usual method of proof of (10.12) is to exploit the geometric nature of the series in the following way:

\[
\begin{align*}
S_N &= \alpha + \alpha r + \alpha r^2 + \cdots + \alpha r^N \\
\Rightarrow \quad r \cdot S_N &= \alpha r + \alpha r^2 + \alpha r^3 + \cdots + \alpha r^{N+1} \\
\Rightarrow \quad (1 - r)S_N &= \alpha + 0 + 0 + 0 - \alpha r^{N+1} \\
\end{align*}
\]

In the first line we wrote the definition of $S_N$. In the next line we multiplied that equation by $r$. In the third line, the second line is subtracted from the first. In doing so, the terms $\alpha r, \alpha r^2, \cdots, \alpha r^N$ cancel, leaving only $\alpha - \alpha r^{N+1}$ on the right-hand side. This gives us

$$(1 - r)S_N = \alpha \left(1 - r^{N+1}\right).$$

Since we are assuming $r \neq 1$, we can divide by $1 - r$ and get (10.12), as desired.

To utilize (10.12), one needs to know $\alpha$, $r$ and $N$. Note that $N$ is not the number of terms, but the highest power of $r$ which occurs. In fact there are $N+1$ terms added to arrive at $S_N$, since the first is $\alpha^0$.

\(^3\)We will not generally consider the case $\alpha = 0$ because it is trivial, and because we cannot identify a unique $r$. Indeed, if $\alpha = 0$, then any geometric recursion $a_{n+1} = r \cdot a_n$ is valid, but our original method of defining $r$, namely (10.8) on page 705, is undefined if $\alpha = 0$. 
10.1. SERIES AND PAR TIAL SUMS

Example 10.1.6 Consider the series \(1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots\). Find the sum of the first 9 terms.

Solution: What we are seeking here is \(S_9 = \frac{\alpha(1-r^{9+1})}{1-r}\), where \(\alpha = 1\) and \(r = \frac{1}{2}\). Thus

\[
S_9 = \frac{1 - \left(\frac{1}{2}\right)^9}{1 - \frac{1}{2}} = \frac{1 - \frac{1}{512}}{\frac{1}{2}} = \frac{511}{256} = \frac{512 - 1}{256} = \frac{511}{256} = 1.99609375
\]

The formula (10.12) also works when \(r < 0\).

Example 10.1.7 Consider the series \(1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \cdots\). Find the sum of the first 9 terms.

Solution: Again we want \(S_9\), but while \(\alpha = 1\) as before, here we have \(r = -1/2\).

\[
S_9 = \frac{1 - \left(-\frac{1}{2}\right)^9}{1 - \left(-\frac{1}{2}\right)} = \frac{1 - \frac{1}{512}}{\frac{3}{2}} = \frac{512}{512} = \frac{512 + 1}{256} = \frac{513}{256} = \frac{171}{256} = 0.66796875
\]

Example 10.1.8 Suppose one deposits into an account (without interest) one penny ($0.01) on the first day of a month, then deposits two pennies ($0.02) the next day, four pennies the next, and so on, each day depositing twice what was deposited the day before. How much money is in the account after the first week (7 payments), second week, third week, and thirty-first day?

Solution: This is the same as asking for partial sums of the series \(0.01 + 0.02 + 0.04 + 0.08 + \cdots\). This is a geometric series (10.9) with \(\alpha = 0.01\) and \(r = 2\). Here we have to be careful about \(N\), since after the first day \(N = 0\), after the second \(N = 1\), etc. Now we compute the total deposit after

- **1 week, i.e., 7 days, we have** \(N = 6\) and

\[
S_6 = \frac{0.01 \left[1 - 2^7\right]}{1 - 2} = \frac{0.01 \left[1 - 2^7\right]}{-1} = 0.01(2^7 - 1) = 0.01(127) = 1.27.
\]

- **2 weeks, i.e., 14 days, we have** \(N = 13\) and (continuing the pattern above)

\[
S_{13} = \frac{0.01 \left[1 - 2^{14}\right]}{1 - 2} = 0.01(2^{14} - 1) = 0.01(16383) = 163.83.
\]

- **3 weeks, i.e., 21 days, we have** \(N = 20\) and

\[
S_{20} = \cdots = 0.01(2^{21} - 1) = 0.01(2,097,151) = 20,971.51
\]

- **31 days, so we have** \(N = 30\), and

\[
S_{30} = \cdots = 0.01(2^{31} - 1) = 0.01(2,147,483,647) = 21,474,836.47.
\]

This latest example illustrates that, when \(r > 1\), the function \(N \mapsto S_N\) is essentially exponential. Indeed, as a function of \(N\),

\[
S_N = \frac{\alpha}{1 - r} \left[1 - r^{N+1}\right] = \frac{\alpha}{1 - r} \cdot \frac{-\alpha \cdot r^{N+1}}{1 - r} = \frac{\alpha}{1 - r} \left[\frac{\alpha}{r - 1} \cdot r^N\right] = A + Br^N,
\]

where \(A = \frac{\alpha}{r - 1}\) and \(B = \frac{\alpha r}{r - 1}\). Thus as a function of \(N\), \(S_N\) is basically a vertical translation of an exponential growth \(Br^N\), assuming again that \(r > 1\). This partially explains why some use the term “geometric growth” when referring to exponential growth.

\(^4\)If \(r \notin (0, 1)\) we get a translation of exponential decay; if \(r \in (-1, 0)\) we get a kind of “damped oscillation”; if \(r = -1\) we get steady oscillation; and if \(r < -1\) we get a growing oscillation. Details are left to the reader.
10.1.3 Convergence/Divergence in Geometric Series

Now we look at necessary and sufficient conditions for a geometric series to converge. If a given geometric series does converge, we compute its sum. Our result is the following:

**Theorem 10.1.2** For a geometric series \( \sum_{n=0}^{\infty} \alpha r^n = \alpha + \alpha r + \alpha r^2 + \alpha r^3 + \cdots \), where \( \alpha \neq 0 \),

1. the series **converges** if and only if \(|r| < 1\), i.e., \( r \in (-1, 1) \);

   equivalently, the series **diverges** if and only if \(|r| \geq 1\), i.e., \( r \in (-\infty, -1] \cup [1, \infty) \).

2. if \(|r| < 1\), then the series converges to \( \frac{\alpha}{1 - r} \).

Restated, the geometric series converges to \( \frac{\alpha}{1 - r} \) if \(|r| < 1\), and diverges otherwise.

**Proof:** The proof requires some care, as the various cases contain their own technicalities.

- Case \( r = 1 \). In such a case, it is not difficult to see (we just count the terms!) that
  \[
  S_N = \sum_{n=0}^{N} \alpha = (N+1)\alpha \to \infty \quad \text{as } N \to \infty.
  \]
  Thus \( r = 1 \) gives a divergent series.

- Case \( r = -1 \). In such a case, we have
  \[
  \sum_{n=0}^{\infty} \alpha (-1)^n = \alpha - \alpha + \alpha - \alpha + \cdots,
  \]
  and so
  \[
  S_N = \begin{cases} 
  \alpha, & \text{if } n \text{ is even,} \\
  0, & \text{if } n \text{ is odd.}
  \end{cases}
  \]
  In other words, \( \{S_N\}_{N=0}^{\infty} = \alpha, 0, \alpha, 0, \alpha, 0, \cdots \), which is clearly a divergent sequence, i.e., the series itself is divergent (by definition).\(^5\)

- Case \(|r| > 1\). Here we can use the formula for the partial sums:
  \[
  S_N = \frac{\alpha (1 - r^{N+1})}{1 - r}.
  \]
  Now there is only one term which is not a fixed constant, and so the convergence of this expression depends upon only the convergence that, \( r^{N+1} \)-term. Clearly if \(|r| > 1\), this is an exponential growth, and diverges. For the general case \(|r| > 1\), we get that\(^6\)
  \[
  r^{N+1} \text{ converges } \implies |r^{N+1}| \text{ converges } \iff |r|^{N+1} \text{ converges}. \quad (10.13)
  \]
  But for \(|r| > 1\), we have \(|r|^{N+1} \text{ diverges} \iff |r|^{N+1} \text{ diverges.} \quad (10.13)\]

\(^5\)Recall that the convergence of the series is defined by the convergence of the (sequence of) partial sums.

\(^6\)This follows from continuity of the function \( x \mapsto |x| \) giving us the \( \text{“} \implies \text{”} \). See Theorem 3.10.2, page 286.
10.1. SERIES AND PARTIAL SUMS

- Case $|r| < 1$. Again we look at the variable part of the formula for $S_N$. It is enough to show that $|r| < 1 \implies r^{N+1}$ converges. One method is to use the sandwich theorem. In the argument below, note that $|r| < 1 \implies |r| \in (0,1) \implies r^{N+1} \to 0$. The relevant sandwich theorem application is then (as $N \to \infty$):

$$-|r|^{N+1} = -|r^{N+1}| \leq r^{N+1} \leq |r|^{N+1}$$

Thus $|r| < 1 \implies r^{N+1} \to 0$ as $N \to \infty$. We can conclude that

$$|r| < 1 \implies S_N = \frac{\alpha (1 - r^{N+1})}{1 - r} \to \frac{\alpha (1 - 0)}{1 - r} = \frac{\alpha}{1 - r} \text{ (as } N \to \infty)$$

This completes the proof.

The implication above is worth repeating in a summarized form:

$$|r| < 1 \implies \sum_{n=0}^{\infty} \alpha r^n = \frac{\alpha}{1 - r},$$

(10.14)

Also worth mentioning:

$$|r| \geq 1, \alpha \neq 0 \implies \sum_{n=0}^{\infty} \alpha r^n \text{ diverges.}$$

(10.15)

**Example 10.1.9** Here are some series computations using the theorem and (10.14).

- $\sum_{n=0}^{\infty} 2 \left(\frac{1}{3}\right)^n = \frac{2}{1 - \frac{1}{3}} = \frac{3}{2} = 2 \cdot \frac{3}{2} = 3$. ($\alpha = 2, r = \frac{1}{3}$.)

- $\sum_{n=0}^{\infty} 0.99^n = \frac{1}{1 - 0.99} = \frac{1}{0.01} = 100$. ($\alpha = 1, r = 0.99$.)

- $\sum_{n=0}^{\infty} 1.01^n$ diverges. ($\alpha = 1, r = 1.01$ so $|r| > 1$, and the series diverges.)

- $\sum_{n=2}^{\infty} \frac{1}{3^n} = \frac{1}{3^2} + \frac{1}{3^3} + \frac{1}{3^4} + \frac{1}{3^5} + \cdots = \frac{1}{3} \left(\frac{1}{1 - \frac{1}{3}}\right) = \frac{1}{3} \left(\frac{3}{2}\right) = \frac{1}{2}$. (First term is $\alpha = \frac{1}{3}, r = \frac{1}{3}$.)

- $1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \cdots = \frac{1}{1 - \left(-\frac{1}{2}\right)} = \frac{1}{3} = \frac{2}{3}$. ($\alpha = 1, r = -\frac{1}{2}$.)

- $\sum_{n=1}^{\infty} e^{-n} = \frac{e}{1-e} = \frac{e}{1-e} \cdot \frac{e}{e} = \frac{e^2}{e - 1}$. ($\alpha = e, r = \frac{1}{e}$.)

Recall that for any $x \in \mathbb{R}$, we have $-|x| \leq x \leq |x|$. 

\[\text{\footnotesize } \]
• \( \sum_{n=1}^{\infty} \frac{5}{3^{2n}} = \sum_{n=1}^{\infty} \frac{5}{9^n} = \frac{5/9}{1 - \frac{1}{3}} = \frac{5}{8} \) \( (\alpha = \frac{5}{9}, r = \frac{1}{3}) \)

• \( \sum_{n=0}^{\infty} (-5)^n = \frac{1}{4} - \frac{5}{4^2} + \frac{5^2}{4^3} - \frac{5^3}{4^4} + \cdots = \frac{\frac{1}{4}}{1 + \frac{5}{16}} = \frac{16}{21} \) \( (\alpha = \frac{1}{4}, r = -\frac{5}{16}) \)

**Exercises**

1. Show that the following series can be written as a telescoping series, and discuss its convergence:

\[ \sum_{n=1}^{\infty} \ln \left( \frac{n}{n+1} \right) \]

2. Do the same with the series

\[ \sum_{n=1}^{\infty} \frac{1}{n^2 + 5n + 6} \]

3. Do the same with the series

\[ \sum_{n=1}^{\infty} \frac{1}{n(n+2)} \]

4. Do the same with the series

\[ \sum_{n=1}^{\infty} \left[ \frac{1}{n} - \frac{1}{n+3} \right] \]

5. Find the \( N \)th partial sum of

\[ \sum_{n=0}^{\infty} 3^n \]

and use it to determine if the series converges or diverges.

6. Do the same for the series

\[ \sum_{n=1}^{\infty} \frac{2}{3^n} \]

7. For each, determine if the series converges or diverges, and if it converges, what is its sum (that it converges to).

(a) \( 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots \)

(b) \( 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \cdots \)

(c) \( 100 + 10 + 1 + \frac{1}{10} + \frac{1}{100} + \cdots \)

(d) \( 100 - 10 + 1 - \frac{1}{10} + \frac{1}{100} - \cdots \)

(e) \( \frac{1}{1000} + \frac{1}{100} + \frac{1}{10} + 1 + 10 + 100 + \cdots \)

8. For each of the following geometric series, find the first term and the ratio. Also determine if it converges or diverges, and if the former, find its sum. (For some of these, it might help to write out a few terms.)

(a) \( \sum_{n=2}^{\infty} (-3)^n \)

(b) \( \sum_{n=0}^{\infty} \left[ (-1)^n \left( \frac{4}{5} \right)^n \right] \)

(c) \( \sum_{n=1}^{\infty} \frac{9}{2^n} \)

(d) \( \sum_{n=0}^{\infty} 5 \cdot \left( \frac{4}{5} \right)^{2n+1} \)

(e) \( \sum_{n=0}^{\infty} \frac{2^n}{3^{2n-1}} \)

(f) \( \sum_{n=2}^{\infty} \frac{3^n}{(-2)^{n+1}} \)

(g) \( \sum_{n=0}^{\infty} \frac{3^n}{2^{2n+1}} \)

9. Give an alternative proof of the formula (10.12) for the partial sums of geometric series. For this new proof, begin with the formula for \( S_N \) as in the original proof (page 706), and then multiply by \( (1 - r) \), noting how the right-hand side simplifies. (See also page 93.)
10.2 NTTFD and Integral Test

Because it is the exceptional case (e.g., geometric, telescoping) that we can actually find a compact formula for $S_N$, we have to develop other tests for the convergence or divergence of series. There will be several such tests, and which particular test or tests are expeditious and conclusive will vary from series to series. We explore the first of those tests in this section. We start with two series that are similar, though one converges and the other diverges.

Example 10.2.1 The following are facts regarding two particular series.

- $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.
- $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges.

These are not obvious at all, but need to be proven. We do so by comparing the terms in the series—represented by areas of rectangles—to certain improper integrals, as in Figure 10.1.

For $\sum_{n=1}^{\infty} \frac{1}{n}$, from the left-hand graph of Figure 10.1 we observe:

$$S_1 = \frac{1}{1} \geq \int_1^2 \frac{1}{x} \, dx$$
$$S_2 = \frac{1}{1} + \frac{1}{2} \geq \int_1^3 \frac{1}{x} \, dx$$
$$S_3 = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} \geq \int_1^4 \frac{1}{x} \, dx$$
$$\vdots$$
$$S_N = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N} \geq \int_1^{N+1} \frac{1}{x} \, dx.$$
Thus \( S_N \geq \int_1^{N+1} \frac{1}{x} \, dx = \ln(N+1) - \ln 1 = \ln(N+1) \to \infty \) as \( N \to \infty \).

We must conclude that \( S_N \to \infty \) as \( N \to \infty \), which implies that \( \sum_{n=1}^{\infty} \frac{1}{n} \) diverges (to infinity).

In fact for this example we have strict inequality “>” in each of the above, but it is enough that we have the non-strict “\( \geq \)”.

For \( \sum_{n=1}^{\infty} \frac{1}{n^2} \), instead we note from the right-hand graph of Figure 10.1 above that

\[
S_2 = 1 + \frac{1}{4} \leq 1 + \int_1^{2} \frac{1}{x^2} \, dx \\
S_3 = 1 + \frac{1}{4} + \frac{1}{9} \leq 1 + \int_1^{3} \frac{1}{x^2} \, dx \\
S_4 = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} \leq 1 + \int_1^{4} \frac{1}{x^2} \, dx \\
\vdots \\
S_N = 1 + \frac{1}{4} + \cdots + \frac{1}{N^2} \leq 1 + \int_1^{N} \frac{1}{x^2} \, dx.
\]

Now clearly \( \{S_N\}_{N=1}^{\infty} \) is an increasing sequence, and \( \int_1^{N} \frac{1}{x^2} \, dx \) is obviously increasing with \( N \). Furthermore, for all \( N \)

\[
S_N \leq 1 + \int_1^{N} \frac{1}{x^2} \, dx \leq 1 + \int_1^{\infty} \frac{1}{x^2} \, dx. \tag{10.16}
\]

We can compute this improper integral as follows:

\[
\int_1^{\infty} \frac{1}{x^2} \, dx = \lim_{\beta \to \infty} \int_1^{\beta} \frac{1}{x^2} \, dx = \lim_{\beta \to \infty} \left[ -\frac{1}{x} \right]_1^{\beta} = \lim_{\beta \to \infty} \left( \frac{-1}{\beta} + 1 \right) = 1. \tag{10.17}
\]

Putting computation (10.17) into (10.16), we get the upper bound

\[
S_N \leq 1 + 1 = 2.
\]

Thus \( \{S_N\}_{N=1}^{\infty} \) is a bounded, and obviously increasing sequence (since we are summing positive terms at each step), and therefore converges,\(^8\) so

\[
\sum_{n=1}^{\infty} \frac{1}{n^2} = \lim_{N \to \infty} S_N
\]

also converges, q.e.d. In fact, we even have an upper bound for the series:

\[
\sum_{n=1}^{\infty} \frac{1}{n^2} \leq 2.
\]

However, we do not know from these computations its actual value. We only know that it does converge.\(^9\)

\(^8\)Recall that any bounded, increasing (or decreasing) sequence must converge to a limit. See Section 3.11.

\(^9\)Note that we used an improper integral to investigate the convergence or divergence of the series, but we did not compute the series’ exact value. We can get an estimate, meaning that we find a bound for it, but we cannot compute the exact value with these methods. There are methods for doing so, but they are indirect and beyond the scope of this text. In fact \( \sum_{n=1}^{\infty} \frac{1}{n^2} = \pi^2/6 \), from Fourier Analysis, but it is a rather serendipitous computation that happens to give us that result, and not easily generalized. However, we can find how many terms must be added to be within a certain tolerance of the actual series’ value, as we will see in a later subsection.
10.2.1 First Divergence Test: NTTFD

When we compare the two series above, we see that both have terms that are shrinking to zero, but one series’ terms, $1/n^2$ shrink much faster than the other’s, namely $1/n$, just as $1/x^2$ shrinks much faster than $1/x$, hence the former’s improper integral on $[1, \infty)$ converges while the latter’s does not. In fact, according to our next theorem it is necessary for the terms to shrink if the series is to have any chance of converging. However, it is not enough that they shrink to zero; they have to shrink fast enough (assuming all terms are the same sign), or the series will diverge.

Thus it is necessary (even though not sufficient) that the terms of a series must shrink to zero if the series is to have any chance of converging. This we codify in a theorem.

**Theorem 10.2.1** Given a series $\sum_{n=k}^{\infty} a_n$. If the series converges, then $a_n \to 0$ as $n \to \infty$:

$$\left[ \sum_{n=k}^{\infty} a_n \text{ converges} \right] \implies \left[ \lim_{n \to \infty} a_n = 0 \right].$$ (10.18)

**Proof:** Suppose $\sum_{n=k}^{\infty} a_n$ converges, i.e., $\sum_{n=k}^{\infty} a_n = L$ for some $L \in \mathbb{R}$ (so in particular $L$ is finite). Then by definition

$$S_N \to L \text{ as } N \to \infty.$$

Now recall $S_n = a_n + S_{n-1}$, so that $a_n = S_n - S_{n-1}$. Taking $n \to \infty$ we get

$$a_n = S_n - S_{n-1} \to L - L = 0, \quad \text{q.e.d.}$$

This proof can seem “slick” to a novice mathematics student, but it is entirely correct. The proof notwithstanding, it should be intuitively clear that, if we are going to “add up” infinitely many terms, and have the sums approach a finite number, then the terms we are adding are going to have to shrink to zero, at least in the limit. The proof uses the fact that $S_n \to L \implies S_{n-1} \to L$, the latter limit occurring “one step behind” the former, but occurring nonetheless since $n \to \infty \implies n - 1 \to \infty \implies S_{n-1} \to L$.

Note that it was important that $L$ be finite in the limit computation above. For instance, if $S_n \to \infty$, we would have $a_n = S_n - S_{n-1}$ giving $\infty - \infty$-form (which is indeterminate) as $n \to \infty$.

Again, the intuition behind the theorem is that, in order to be able to add infinitely many terms—one at a time in the sense that we compute $S_k$, $S_{k+1}$, etc., and look for a trend in these sums towards $L$—the terms that we add, i.e., $a_k$, $a_{k+1}$, $a_{k+2}$, etc., have to shrink eventually if the partial sums are to get closer and closer to (“approach”) a finite number $L$.

In fact, the form of the theorem which we use is the contrapositive. Recall the logical equivalence $P \implies Q \iff (\sim Q) \implies (\sim P)$.\(^\text{10}\) In this case, $P$ is the statement that the series converges (to a finite number $L$), while $Q$ is the statement that $a_n \to 0$. The contrapositive for of Theorem 10.2.1 is our main result in this section, and we dub that result *nth term test for divergence*, or NTTFD:

\(^\text{10}\)To be sure, here $P \implies Q$ is read, “$P$ implies $Q$.” The symbol “$\sim$” is still the “not,” or logical negation, operator. The symbol “$\iff$” stands in for logical equivalence. Recall $P \implies Q$ and $(\sim Q) \implies (\sim P)$ are *contrapositives* of each other, and are logically equivalent.
**Theorem 10.2.2 (NTTFD)** If it is not the case that \( a_n \to 0 \), then \( \sum_{n=k}^{\infty} a_n \) diverges. Put symbolically,

\[
a_n \not\to 0 \implies \sum_{n=k}^{\infty} a_n \text{ diverges.}
\] (10.19)

**Proof:** It is enough to say that this is the contrapositive of Theorem 10.2.1, and therefore also true. One way to write this symbolically is the following.

\[
\sum_{n=k}^{\infty} a_n \text{ converges} \iff (a_n \to 0) \iff [\sim (a_n \to 0)] \iff \sum_{n=k}^{\infty} a_n \text{ diverges.}
\]

The statement on the right must be a tautology (always true), since it is equivalent to the statement on the left, which—being the statement of Theorem 10.2.1—is itself a tautology. The statement on the right being a tautology means that it can stand alone as a tautology, written as in (10.19), q.e.d.

This theorem is undoubtedly one of the most misunderstood and misapplied results in all of Calculus I and II. It is as important to understand what it does not say, as it is to understand what it says. The theorem says that if the terms of a series do not shrink to zero, then the series must diverge.

But it is not as comprehensive as one might think. After all, if the terms of a series do shrink to zero, the theorem is silent! (Therein lies the unfortunately very common mistake made by calculus students.) To emphasize this we look at the following examples.

**Example 10.2.2** Discuss what Theorem 10.2.2 has to say about the series

(a) \( \sum_{n=1}^{\infty} \frac{n}{2n+1} \);

(b) \( \sum_{n=1}^{\infty} \frac{1}{n+1} \)

(c) \( \sum_{n=1}^{\infty} \cos \frac{1}{n} \);

(d) \( \sum_{n=1}^{\infty} \sin \frac{1}{n} \);

(e) \( \sum_{n=1}^{\infty} \sin \frac{1}{n^2} \).

**Solution:**

(a) \( \frac{n}{2n+1} \to \frac{1}{2} \neq 0 \text{ NTTFD} \Rightarrow \sum_{n=1}^{\infty} \frac{n}{2n+1} \text{ diverges.} \)

(b) \( \frac{1}{n+1} \to 0 \). The NTTFD is inconclusive.

(c) \( \cos \frac{1}{n} \to \cos 0 = 1 \neq 0 \text{ NTTFD} \Rightarrow \sum_{n=1}^{\infty} \cos \frac{1}{n} \text{ diverges.} \)

(d) \( \sin \frac{1}{n} \to \sin 0 = 0 \). The NTTFD is inconclusive.
(e) $\sin \frac{1}{n^2} \to \sin 0 = 0$. The NTTFD is inconclusive.

Looking closely at the symbolic statement of NTTFD given in (10.19), we see that there is never an implication of convergence. Indeed, the test either concludes divergence, or is inconclusive. This is a very quick but incomplete test, which can only detect divergence in certain (still common) circumstances, namely that $a_n \neq 0$.

Indeed, in (a) and (c) above, NTTFD gave us divergence. However, it said nothing in (b), (d) and (e), as the “if” part of the theorem was not true. In fact, of these three in which NTTFD is silent—(b), (d) and (e)—it turns out that (b) and (d) are divergent, while (e) is convergent. The methods to see this are introduced in later sections. An example from the exercises in the last section gives a case where we can in fact show that it is possible that $a_n \to 0$, but the series diverges:

Example 10.2.3 Consider the series $\sum_{n=1}^{\infty} \ln \left( \frac{n}{n+1} \right)$. Determine if it converges or diverges.

Solution: First we note $\ln \left( \frac{n}{n+1} \right) \to \ln 1 = 0$ as $n \to \infty$. Thus NTTFD is inconclusive.\(^{11}\)

Looking closer at this series, we should eventually notice that it is telescoping. This becomes clear if we rewrite it:

$$\sum_{n=1}^{\infty} \ln \left( \frac{n}{n+1} \right) = \sum_{n=1}^{\infty} [\ln(n) - \ln(n+1)] = [\ln 1 - \ln 2] + [\ln 2 - \ln 3] + [\ln 3 - \ln 4] + \cdots.$$ 

It is not difficult to see that $S_N = \ln 1 - \ln N = -\ln N$, and thus

$$S_N = -\ln N \to -\infty \text{ as } N \to \infty.$$ 

Since the partial sums diverge, by definition so does the series.

10.2.2 Integral Test Proper (IT)

Of course there are series which converge. However the NTTFD is never powerful enough to ever prove it (but can sometimes detect divergence in a series). Proving actual convergence requires other tests. The integral test is one such test:

**Theorem 10.2.3** Suppose we have a series $\sum_{n=k}^{\infty} a_n$ such that

1. $a_n = f(n)$ for each $n \geq k$, where
2. $f(x) \geq 0$ is continuous and nonincreasing on $[k, \infty)$.

Then $\sum_{n=k}^{\infty} a_n$ and $\int_{k}^{\infty} f(x) \, dx$ both converge or both diverge. In other words,

$$\sum_{n=k}^{\infty} a_n \text{ converges } \iff \int_{k}^{\infty} f(x) \, dx \text{ converges}.$$ 

Equivalently,

$$\sum_{n=k}^{\infty} a_n \text{ diverges } \iff \int_{k}^{\infty} f(x) \, dx \text{ diverges}.$$ 

\(^{11}\)Some textbooks would write that the test “fails.” That seems a bit strong. It is merely inconclusive, so we need to look deeper at the particular series and perhaps employ some other test which will be conclusive.
Proof: We will not write the whole proof here, but just mention that it follows the same kind of reasoning we used to show \( \sum \frac{1}{n} \) diverges while \( \sum \frac{1}{n^2} \) converges. It is not too difficult to see—by drawings similar to those early examples—that

(a) If \( \int_k^\infty f(x) \, dx \) diverges, then it diverges to \( \infty \), and so as \( N \to \infty \) we have

\[
S_N \geq \int_k^{N+1} f(x) \, dx \to \infty,
\]

so \( \sum_{n=k}^{\infty} a_n \) diverges to infinity.

(b) If \( \int_k^\infty f(x) \, dx \) converges and is thus finite, then

\[
S_N \leq a_k + \int_k^N f(x) \, dx \leq a_k + \int_k^\infty f(x) \, dx < \infty,
\]

so \( S_N \) is a bounded, clearly nondecreasing sequence, so it must converge, implying (by definition) that \( \sum_{n=k}^{\infty} a_n \) converges.

Example 10.2.4 Consider any series \( \sum_{n=1}^{\infty} \frac{1}{n^p} \), where \( p > 0 \). Then \( f(x) = 1/x^p \) is clearly a decreasing function on \([1, \infty)\). Now

\[
\lim_{\beta \to \infty} \int_1^\beta x^{-p} \, dx = \begin{cases} 
\lim_{\beta \to \infty} \frac{x^{1-p} |_{1}^{\beta}}{1-p} = \frac{-1}{1-p} & \text{if } p > 1 \quad (1-p < 0) \\
\lim_{\beta \to \infty} \ln x |_{1}^{\beta} = \infty & \text{if } p = 1 \\
\lim_{\beta \to \infty} \frac{x^{1-p} |_{1}^{\beta}}{1-p} = \infty & \text{if } p < 1 \quad (1-p > 0)
\end{cases}
\]

from which we get that \( \int_1^\infty \frac{1}{x^p} \, dx \) converges if and only if \( p > 1 \), and diverges if and only if \( p \leq 1 \). This gives us the following as well:

Theorem 10.2.4 \( \sum_{n=1}^{\infty} \frac{1}{n^p} \) converges if and only if \( p > 1 \) (and diverges if and only if \( p \leq 1 \)).

Series of the form \( \sum_{n=1}^{\infty} \frac{1}{n^p} \) are called \( p \)-series. They are nearly as important as geometric series within the general theory of series. If one forgets about the two cases, the integral test, Theorem 10.2.3, page 715 makes deriving this last theorem fairly straightforward.

Example 10.2.5 Determine if the series \( \sum_{n=2}^{\infty} \frac{1}{n \ln n} \) converges.

Solution: Clearly this is a series of positive terms, which decrease monotonically. We can look therefore at \( \int_2^\infty \frac{1}{x \ln x} \, dx \), since \( f(x) = 1/(x \ln x) \) is also positive and decreasing on \([2, \infty)\). Now

\[
\lim_{\beta \to \infty} \int_2^\beta \frac{1}{x \ln x} \, dx = \lim_{\beta \to \infty} \ln | \ln x |_{\beta}^{2} = \lim_{\beta \to \infty} \ln(\ln(\beta) - \ln(2)) = \infty,
\]

12We can note that the denominator is increasing and positive, and so intuitively these fractions \( 1/x^p \) must shrink. On the other hand, we can also compute \( f'(x) = \frac{d}{dx} x^{-p} = -p x^{-p-1} = -p/(x^{p+1}) < 0 \), so the fact that \( f' < 0 \) on \([1, \infty)\) also shows that \( f(x) \) is decreasing there.
we conclude that $\int_2^{\infty} \frac{1}{x \ln x} \, dx$ diverges to infinity, and therefore so does the original series, by the Integral Test.

We will have other tests in subsequent sections. However, for series which satisfy the hypotheses of the Integral Test, it is the most sensitive, and can determine convergence or divergence for many series for which other tests are inconclusive. (It already has shown convergence or divergence for series for which the NTTFD is silent.) It is also interesting to note that the Integral Test, or its general reasoning buried in the proof, can determine which geometric series converge and which diverge, assuming $\alpha > 0$, $r > 0$, and $r \neq 1$ (or the series obviously diverges).

$$\lim_{\beta \to -\infty} \int_0^{\beta} \alpha r^x \, dx = \lim_{\beta \to -\infty} \frac{\alpha r^x}{\ln r} \bigg|_0^{\beta} = \begin{cases} \infty & \text{if } r > 1 \\ \frac{\alpha}{\ln r} & \text{if } r \in (0, 1). \end{cases}$$

Note that for $r \in (0, 1)$ we have $\ln r < 0$ so the integral is positive. So the Integral Test concludes that $\sum \alpha r^n$ converges if $r \in (0, 1)$ and diverges if $r \geq 1$. (If $\alpha < 0$ we just look at $\sum \alpha r^n = -\sum (-\alpha)r^n$, and if $\alpha = 0$ the series obviously converges.) Unfortunately the Integral Test can not be used here for $r < 0$, but fortunately we have a formula (10.12), page 706 for $S_N$ anyhow, from which we can judge convergence based upon what occurs if $N \to \infty$.

Example 10.2.6 Consider $\sum_{n=0}^{\infty} \frac{1}{n^2 + 1}$, and discuss if it converges or diverges.

Solution: First note that $a_n = f(n) > 0$, where $f(x) = \frac{1}{x^2 + 1}$ is a decreasing function on $[0, \infty)$. Again this is intuitive, but a quick check shows $f' < 0$ on $(0, \infty)$, which is enough to show $f$ is decreasing on $[0, \infty)$.\(^{13}\)

Next we check the relevant improper integral.

$$\int_0^{\infty} \frac{1}{x^2 + 1} \, dx = \lim_{\beta \to -\infty} \int_0^{\beta} \frac{1}{x^2 + 1} \, dx = \lim_{\beta \to -\infty} \left(\tan^{-1} \beta - \tan^{-1} 0\right) = \frac{\pi}{2} - 0 = \frac{\pi}{2}.$$

Since the improper integral converges, so does the series $\sum_{n=0}^{\infty} \frac{1}{n^2 + 1}$ converge.

One interesting question is how many terms of the series do we need to add in order to get a good approximation of the full series? We clearly can not add infinitely many, and this series is not telescoping and not geometric, so finding a simple formula for $S_N$ for which we can let $N \to \infty$ seems unlikely. So instead we look back to the integral that proves the series converges. While there seems to be a general formula, it is probably best practice (unless one does such problems repeatedly) to re-draw the situation when needed. Nonetheless, it is interesting to see the general rule:

Theorem 10.2.5 Suppose $\sum a_n$ satisfies the hypotheses of the Integral Test, in the sense that $a_n = f(n)$ for some $f(x)$ defined, nonnegative and nonincreasing on some interval $[k, \infty)$, and $\int_k^{\infty} f(x) \, dx < \infty$. Then the series converges to some value $\sum a_n = S$, and for $N > k$ we have

$$|S - S_N| = S - S_N \leq \int_N^{\infty} f(x) \, dx. \quad (10.20)$$

\(^{13}\)In fact, we really only need $f(x)$ to be decreasing on some set $[k, \infty)$, since we can always sum the finitely many terms that occur before $n = k$. It is the infinite “tail-end” sum of the series that determines convergence, i.e., we need to compute if $\sum_{n=k}^{\infty} a_n$ converges. We have the liberty to ignore what happens for any finite number of terms at the “front” of the series, assuming they are defined.
CHAPTER 10. SERIES OF CONSTANTS

For our previous example, if we wished to know \( \sum_{n=0}^{\infty} \frac{1}{n^2 + 1} \) to within 0.10, we can find \( N \) so that \( \int_{N}^{\infty} \frac{1}{x^2 + 1} \, dx \leq 0.10 \), and that will guarantee \( S - S_N \leq 0.10 \) as well. Now for any given \( N \), we have

\[
\int_{N}^{\infty} \frac{1}{x^2 + 1} \, dx = \lim_{\beta \to \infty} \int_{N}^{\beta} \frac{1}{x^2 + 1} \, dx = \lim_{\beta \to \infty} \left( \tan^{-1} \beta - \tan^{-1} N \right) = \frac{\pi}{2} - \tan^{-1} N.
\]

For \( \pi/2 - \tan^{-1} N \leq 0.01 \), we add \( \tan^{-1} N \) to both sides and see we need

\[
\tan^{-1} N \geq \frac{\pi}{2} - 0.10 \implies N \geq \tan \left( \frac{\pi}{2} - 0.10 \right) \approx 9.96644419,
\]

so we take \( N \geq 10 \) for such accuracy.

If instead we wanted accuracy to within 0.001, we have a similar calculation yielding

\[
\tan^{-1} N \geq \frac{\pi}{2} - 0.001 \implies N \geq \tan \left( \frac{\pi}{2} - 0.001 \right) \approx 999.99962177,
\]

so we would need to take \( N \geq 1000 \), and thus compute

\[
\sum_{n=0}^{1000} \frac{1}{n^2 + 1}
\]

to assure that we achieve such accuracy in using \( S_N \) to approximate the full series \( S \). Note that it quickly becomes more suitable for numerical (specifically, electronic) computational devices.\(^{14}\)

\(^{14}\)In computing the inequality, we used that \( \tan x \) is an increasing function on \((-\pi/2, \pi/2)\), so for all \( x_1, x_2 \in (-\pi/2, \pi/2) \), we have \( x_1 < x_2 \implies \tan x_1 < \tan x_2 \).
and since $1/n^2$ is summable, i.e., $\sum \frac{1}{n^2}$ converges ($p$-series, $p > 1$), the series of smaller terms must also. That is called the Direct Comparison Test, though there are still others.\footnote{As is often the case, we use “$\leq$” when we in fact have “$<$.” This is partly to show that “$\leq$” is sufficient here, though to be more precise we should use “$<$.”}

## Exercises

1. Use the NTTFD to determine which of the following series must diverge (based only upon that test). If NTTFD is inconclusive, so state.
   - (a) $\sum_{n=1}^{\infty} e^{1/n}$
   - (b) $\sum_{n=1}^{\infty} \frac{n}{2^n}$
   - (c) $\sum_{n=1}^{\infty} n \sin \frac{1}{n}$

2. Use the integral test to determine convergence or divergence of the following series.
   - (a) $\sum_{n=1}^{\infty} \frac{\ln n}{n}$
   - (b) $\sum_{n=2}^{\infty} \frac{1}{n \ln n}$
   - (c) $\sum_{n=2}^{\infty} \frac{1}{n (n \ln n)^2}$
   - (d) $\sum_{n=1}^{\infty} \frac{n}{e^n}$
   - (e) $\sum_{n=1}^{\infty} \frac{\sin \frac{1}{n}}{n^2}$
   - (f) $\sum_{n=1}^{\infty} \frac{\ln n}{n^{3/2}}$

3. It is known from other fields that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \pi^2/6 \approx 1.644 934 067$.
   - (a) Find $S_{10}$, and estimate how accurate that is using the integral test estimate (10.20).
   - (b) How large would $N$ need to be to ensure that $S_N$ is within 0.0001 of the full series?

4. The integral test requires that $f(x)$ be eventually monotonically decreasing. Here we give an example where the series and improper integral behave very differently, to show that we can not relax our hypotheses on $f$.
   - (a) Show that $\int_1^{\infty} \sin^2(\pi x) \, dx$ diverges (to infinity).
   - (b) Show that $\sum_{n=1}^{\infty} \sin^2(\pi \cdot n)$ converges.
   - (c) Graph $y = \sin^2(\pi x)$, for $x \geq 0$. Explain why this is not a proper function to choose when determining the convergence of $\sum \sin(n\pi)$ with the integral test.
10.3 Comparison Tests

Before proceeding to the topic of this section, we briefly review what we have developed so far regarding series.

To say that \( \sum a_n \) converges to some number \( S \in \mathbb{R} \), by definition, to say that \( S_N \to S \), where \( S_N \) is the \( N \)th partial sum of the series. There were two—namely the telescoping and geometric series—for which we were able to compute \( S_N \) and therefore its limit as \( N \to \infty \), and therefore \( \sum a_n \).

Next we noted that we could quickly detect divergence in some cases with the NTTFD, as

\[ a_n \neq 0 \implies \sum a_n \text{ diverges.} \]

Of course \( a_n \to 0 \neq \implies \sum a_n \text{ converges.} \)

Then we realized that there are many series which we can prove converge, as in cases where \( \{S_N\} \) is a bounded and increasing sequence, but we could not directly calculate the value of the full series. We proved these converged using integral tests, where we compared them to

\[ \int_1^\infty f(x) \, dx, \]

where \( a_n = f(n) \) for an appropriate function \( f \). The series converged if and only if the improper integral did, but we were careful that, (1) \( f(x) \) was eventually nonincreasing on some interval of the form \([k, \infty)\), and (2) we did not make the mistake of claiming that the series and the integral converged to the same number, but only that their behaviors were similar enough that convergence (or divergence) of one meant the same for the other.

After these developments, we made special note of two types of series for which we know immediately whether or not they converge (recall “iff” is short for “if and only if”):

1. geometric series: \( \sum_{n=0}^{\infty} \alpha r^n \) converges iff \( r \in (-1, 1) \), and diverges otherwise (here \( \alpha \neq 0 \));

2. \( p \)-series: \( \sum_{n=1}^{\infty} \frac{1}{n^p} \) converges iff \( p > 1 \), and diverges otherwise (i.e., for \( p \leq 1 \)).

For the \( p \)-series, we proved this using the Integral Test, and gave some special mention to the borderline case, \( p = 1 \), mentioning that the harmonic series \( \sum \frac{1}{n} \) diverges.

These two series types above will be very useful for both of our comparison tests developed in this section, especially the \( p \)-series. However, before we look at these we make another observation:

**Theorem 10.3.1** \( \sum_{n=1}^{\infty} a_n \) converges \( \iff \) \( (\forall M \in \{1, 2, 3, \ldots\}) \left[ \sum_{n=M}^{\infty} a_n \text{ converges} \right] \)

This is just the statement that the first \( M - 1 \) terms are not what determine convergence or divergence, no matter how large or small \( M \) happens to be. It is the series’ tail end, containing infinitely many terms to sum, which determines if the series converges or not. We can always add a finite number of real numbers and the result will be a finite real number, but when we attempt to somehow “add” an infinite number of terms, in truth we can not but instead appeal to the well-defined partial sums \( S_N \), and then use a limit argument to let \( N \to \infty \).

This theorem is useful because many tests for convergence or divergence in fact allow us to ignore a finite number of terms in the series, to focus instead on the crucial “tail end” of the series.

Recall that if a series \( \sum a_n \) is convergent, we say that the sequence of terms \( a_n \) is *summable*. (It is also common to say the convergent series itself is “summable.”) It is helpful to have such vocabulary at our disposal to streamline later arguments.
10.3. COMPARISON TESTS

10.3.1 Direct Comparison Test (DCT)

This test, like the Integral Test, is a test for series $\sum a_n$ and $\sum b_n$ of positive-term series. Here is its statement:

Theorem 10.3.2 Direct Comparison Test (DCT) for Positive-Term Series. For such series,

$0 \leq a_n \leq b_n$, $\sum_{n=1}^{\infty} b_n$ converges $\implies \sum_{n=1}^{\infty} a_n$ converges. \hfill (10.21)$

$0 \leq a_n \leq b_n$, $\sum_{n=1}^{\infty} a_n$ diverges $\implies \sum_{n=1}^{\infty} b_n$ diverges. \hfill (10.22)

Perhaps the easiest way to interpret this is as follows, keeping in mind we are only discussing series with nonnegative terms added:

- If the series with the larger terms converges, so must the series with the smaller terms.
- If the series with the smaller terms diverges, so must the series with the larger terms.

Again, to be precise, this is true assuming that all terms are nonnegative.\textsuperscript{16} To be even more precise, we should then also use “greater” for “larger,” and “lesser” for “smaller” above. See Figure 10.3.

Proof: First we show that $\sum b_n$ converges implies $\sum a_n$ converges. To show this, we note that if $S_N$ is the $N$th partial sum of $\sum a_n$, then clearly $S_N$ is nondecreasing (since $a_n \geq 0$), and furthermore

$$S_N = a_1 + a_2 + \cdots + a_N \leq b_1 + b_2 + \cdots + b_N \leq \sum_{n=1}^{\infty} b_n,$$

\textsuperscript{16}It is common practice to bend the language somewhat to use “positive-term series” to refer to series $\sum a_n$ where $a_n \geq 0$ for all $n$. A less concise but more precise term would be “nonnegative-term series.” It is important that we do not have the terms changing sign along the series, which is a situation we will deal with Section 10.4.
So \( S_N \leq \sum b_n \), i.e., each \( S_N \) is bounded by the whole series \( \sum b_n \). Thus \( \{S_N\}_{N=1}^{\infty} \) is a bounded, nondecreasing sequence, and must therefore converge. This proves (10.21). In fact (10.22) is just its contrapositive so it is also then proven, but it is also interesting to prove it separately.

So suppose instead that \( \sum a_n \) diverges. Being a nonnegative-term series it must therefore diverge to infinity. So let \( S_N \) be the \( N \)th partial sum of \( \sum a_n \) and \( \mathcal{S}_N \) be the \( N \)th partial sum of \( \sum b_n \). Then

\[
\mathcal{S}_N = b_1 + b_2 + \cdots + b_N \geq a_1 + a_2 + \cdots + a_N = S_N \longrightarrow \infty \quad \text{as} \quad N \to \infty,
\]

showing that \( \mathcal{S}_N \) diverges to infinity (\( \mathcal{S}_N \geq S_N \to \infty \)), i.e., \( \sum b_n \) diverges (to infinity), q.e.d.

Of course the same is true if the two series start somewhere besides \( n = 1 \). Furthermore, in light of Theorem 10.3.1 (page 720), in fact we only require \( 0 \leq a_n \leq b_n \) to be true “eventually,” i.e., for all \( n \geq N \) where \( N \) is some finite number. (Here we use \( N \) differently than in the partial sums \( S_N \).)

**Corollary 10.3.1** Suppose for some \( N \) we have \( n \geq N \implies 0 \leq a_n \leq b_n \). Then (10.21) and (10.22) still hold.

**Example 10.3.1** Consider the series \( \sum_{n=1}^{\infty} \frac{1}{\ln n + n^{3/2}} \). We note that for \( n \geq 1 \) we have \( \ln n \geq 0 \), so the denominator is greater than or equal to \( n^{3/2} \), and of course \( \sum_{n=1}^{\infty} \frac{1}{n^{3/2}} \) converges.

\[
0 < n^{3/2} \leq \ln n + n^{3/2} \implies 0 < \frac{1}{\ln n + n^{3/2}} \leq \frac{1}{n^{3/2}} \quad \text{summable}
\]

\[
\therefore \sum_{n=1}^{\infty} \frac{1}{n^{3/2}} \text{ converges} \implies \sum_{n=1}^{\infty} \frac{1}{\ln n + n^{3/2}} \text{ converges.}
\]

(The first inequalities “0 <” are just to indicate that all terms are positive, so the Direct Comparison Test, DCT, applies.) Since \( \sum_{n=1}^{\infty} \frac{1}{n^{3/2}} \) converges (\( p \)-series, \( p = 3/2 > 1 \)), we can conclude by the DCT that \( \sum_{n=1}^{\infty} \frac{1}{\ln n + n^{3/2}} \) also converges.

Note that we used the fact that, when all terms are positive, a larger denominator implies a smaller fraction.\(^{17}\) We also used that \( \ln n \geq 0 \) for \( n \geq 1 \) (though “eventually,” i.e., for all \( n \geq N \) for some \( N \), is enough).

**Example 10.3.2** Consider \( \sum_{n=1}^{\infty} \frac{\ln n}{n} \). While we could use an Integral Test on this series to determine convergence or divergence,\(^{18}\) it will be faster to note that, for large enough \( n \) (in

\(^{17}\)Recall that if \( A, B, C, D > 0 \), then

\[
A < B \implies \frac{A}{C} < \frac{B}{C} \quad \text{(larger numerator \( \implies \) larger fraction)}
\]

\[
C < D \implies \frac{A}{C} > \frac{A}{D} \quad \text{(larger denominator \( \implies \) smaller fraction)}
\]

\(^{18}\)We might also consider the NTTFD, Theorem 10.2.2 (page 714), but it does not apply since

\[
\lim_{n \to \infty} \frac{\ln n}{n} = \frac{\infty/\infty}{\ln R} \quad \lim_{n \to \infty} \frac{1}{1} \quad (1/\infty)/1 = 0,
\]

so it diverges.

or
particular, \( n > e^1 \) so for integers, \( n \geq 3 \) we have

\[
0 < \frac{1}{\sum_{n=1}^{\infty} \frac{1}{n}} < \frac{\ln n}{n}, \quad \text{not summable}
\]

\[
\therefore \sum_{n=1}^{\infty} \frac{1}{n} \quad \text{diverges} \quad \Rightarrow \quad \sum_{n=1}^{\infty} \frac{\ln n}{n} \quad \text{diverges.}
\]

The DCT is particularly useful when we can not (or can not easily) integrate the respective function, as in our first example (Example 10.3.1, page 722).

**Example 10.3.3** Discuss the convergence/divergence of \( \sum_{n=1}^{\infty} \frac{1}{\sqrt{n^5 + 1}} \).

**Solution:** We note that the terms in the series would be larger if the “+1” were absent in the denominators.

\[
0 \leq \frac{1}{\sqrt{n^5 + 1}} < \frac{1}{\sqrt{n^5}} = \frac{1}{n^{5/2}} \quad \text{summable}
\]

\[
\therefore \sum_{n=1}^{\infty} \frac{1}{n^{5/2}} \quad \text{converges} \quad \Rightarrow \quad \sum_{n=1}^{\infty} \frac{1}{\sqrt{n^5 + 1}} \quad \text{converges.}
\]

### 10.3.2 A Hierarchy of Functions and DCT

It is useful to note that different functions \( f(n) \) which grow to \( \infty \) as \( n \to \infty \) do so much faster than others.

**Theorem 10.3.3** In the list below, any function \( f(n) \) listed to the left of another function \( g(n) \) will grow so much more slowly than \( g(n) \) that \( \lim_{n \to \infty} f(n)/g(n) = 0 \).

\[
\ln(\ln n), \quad \ln n, \quad n^r (r > 0), \quad n^s (s > r), \quad a^n (a > 1), \quad b^n (b > 1), \quad n!, \quad n^n.
\]

This is not an exhaustive list, but offers some useful facts and intuition. It will take some effort and later methods to show why \( n! = 1 \cdot 2 \cdot 3 \cdots n \) fits into the hierarchy where it does. It is somewhat more intuitive to see that \( n^n \) is properly placed, at least in comparison to \( n! \):

\[
0! = 1 \quad \text{(by definition)}
\]

\[
1! = 1 \quad 1^1 = 1
\]

\[
2! = 1 \cdot 2 = 2 \quad 2^2 = 2 \cdot 2 = 4
\]

\[
3! = 1 \cdot 2 \cdot 3 = 6 \quad 3^3 = 3 \cdot 3 \cdot 3 = 27
\]

\[
4! = 1 \cdot 2 \cdot 3 \cdot 4 = 4 \cdot 3! = 24 \quad 4^4 = 4 \cdot 4 \cdot 4 \cdot 4 = 256
\]

\[
5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 5 \cdot 4! = 120 \quad 5^5 = 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5 = 3,125
\]

so the NTTFD is inconclusive. However it is also not difficult to observe that

\[
\int_1^{\infty} \frac{\ln x}{x} \, dx = \lim_{\beta \to \infty} \int_1^{\beta} \frac{\ln x}{x} \, dx = \lim_{\beta \to \infty} \frac{1}{\beta} \left( \ln \beta \right)^2 = \lim_{\beta \to \infty} \left[ \frac{1}{2} (\ln \beta)^2 - 0 \right] = \infty,
\]

which proves \( \sum_{n=1}^{\infty} \frac{\ln n}{n} \) diverges, by the Integral Test.
and so on. This hierarchy helps us to use the DCT to determine convergence or divergence of some series, and sometimes the NTTFD to determine divergence.

**Example 10.3.4** Consider $\sum_{n=2}^{\infty} \frac{1}{\ln n}$. We can direct-compare this to $\sum_{n=2}^{\infty} \frac{1}{n}$. (Note why we cannot start at $n = 1$.) For larger enough $n$ (actually for all $n > 0$), we have $\ln n < n$ (proven by the fact that $(\ln n)/n \to 0$, see original limit in Footnote 18, page 723), and so for large enough $n$ we have

$$0 < \frac{1}{n} \underset{\text{not summable}}{\downarrow} < \frac{1}{\ln n}$$

$$\therefore \sum_{n=2}^{\infty} \frac{1}{n}$$ diverges $\implies \sum_{n=2}^{\infty} \frac{1}{\ln n}$ diverges.

**Example 10.3.5** Consider $\sum_{n=1}^{\infty} \frac{1}{n!}$. Since for instance $n! > n^2$ for large enough $n$ (in particular, $n \geq 4$), we can write

$$0 < \frac{1}{n!} \underset{\text{summable}}{\downarrow} < \frac{1}{n^2}$$

$$\therefore \sum_{n=1}^{\infty} \frac{1}{n^2}$$ converges $\implies \sum_{n=1}^{\infty} \frac{1}{n!}$ converges.

Note that not all comparisons are useful. For instance, it is true that $1/n! < 1/n$, but $1/n$ is not summable, i.e., $\sum \frac{1}{n}$ diverges, so that comparison does not let us conclude convergence or divergence of the series $\sum \frac{1}{n!}$:

$$0 < \frac{1}{n!} \underset{\text{not summable}}{\downarrow} < \frac{1}{n}$$

(No Conclusion Possible!)

The problem is that we are using a series whose sums approach infinity, which is not a useful, or even proper, “bound” for anything. Any series we find, convergent or otherwise, can have its terms bounded by those of a divergent series. (Just take $b_n = a_n + 1$ or $b_n = a_n + n$ or similar.) Thus the inequality above is indeterminate as far as determining in and of itself the convergence or divergence of the series of the terms on the left of the inequality. Similarly, knowing the terms of a series are larger than those of a convergent series is useless.

$$0 < \frac{1}{n^2} \underset{\text{summable}}{\downarrow} < \frac{1}{n}$$

(No Conclusion Possible!)

In fact, note that $0 < \frac{1}{n \ln n} < \frac{1}{n}$, and $0 < \frac{1}{n(\ln n)^2} < \frac{1}{n}$, but $\sum \frac{1}{n} \ln n$ diverges while $\sum \frac{1}{n(\ln n)^2}$ converges, both by routine applications of the Integral Test, so knowing that a series’ terms are smaller than those of $\sum \frac{1}{n}$ (or any other positive-term divergent series) does not guarantee convergence or divergence.
Indeed, there are series that converge and series that diverge whose terms are smaller than those in the harmonic series \( \sum \frac{1}{n} \). However, there are no convergent series with terms greater than the harmonic series.

Note also we could have used the hierarchy of functions to conclude \( \sum \frac{1}{n!} \) converges, because \( \frac{1}{n} < \frac{1}{2^n} \), and \( \sum (1/2)^n \) converges (geometric, \( |r| = 1/2 < 1 \)), also giving us \( \sum \frac{1}{n!} \) converges by the DCT.

**Example 10.3.6** Consider \( \sum_{n=1}^{\infty} \frac{\cos^2 n}{\sqrt{n^3 + 5}} \). Since \( 0 \leq \cos^2 n \leq 1 \), we can write

\[
0 \leq \frac{\cos^2 n}{\sqrt{n^3 + 5}} \leq \frac{1}{\sqrt{n^3 + 5}} < \frac{1}{\sqrt{n^3}} = \frac{1}{n^{3/2}} \text{ summable},
\]

\[
\therefore \sum_{n=1}^{\infty} \frac{1}{n^{3/2}} \text{ converges } \Rightarrow \sum_{n=1}^{\infty} \frac{\cos^2 n}{\sqrt{n^3 + 5}} \text{ converges.}
\]

**Example 10.3.7** Consider \( \sum_{n=1}^{\infty} \frac{n}{n^4 + 1} \). Note that

\[
0 < \frac{n}{n^4 + 1} < \frac{n}{n^4} = \frac{1}{n^3} \text{ summable}
\]

\[
\therefore \sum_{n=1}^{\infty} \frac{1}{n^3} \text{ converges } \Rightarrow \sum_{n=1}^{\infty} \frac{n}{n^4 + 1} \text{ converges.}
\]

### 10.3.3 Limit Comparison Test (LCT)

At times, the DCT requires a bit more cleverness than necessary. For instance, if we take the previous example and make minor changes, considering instead \( \sum_{n=2}^{\infty} \frac{n}{n^4 - 1} \), our intuition is that for large enough \( n \), the terms in this series should be very similar to those of \( \sum_{n=2}^{\infty} \frac{1}{n^3} \), since the difference in the denominators becomes less and less significant for large \( n \). So our intuition is that this series probably converges as well, but finding a formula for \( S_N \), or a series which has larger terms but still converges, or even using an integral test, can be difficult (or in the case of finding a formula for \( S_N \), perhaps impossible). One comparison that does work—included here just for completeness—is the following:

\[
0 < \frac{n}{n^4 - 1} = \frac{n^2}{n(n^4 - 1)} < \frac{2(n^2 - 1)}{n(n^4 - 1)} = \frac{2}{n(n^2 + 1)} < \frac{2}{n^3} \text{ summable}
\]

However the first equality does not immediately seem well motivated, and the second inequality is not so obvious unless perhaps we write out several terms of each to see the pattern emerging.

And yet it seems like we should be able to argue that \( \frac{n^2}{n^4 - 1} \approx \frac{1}{n^2} \) implies that, since \( \sum \frac{1}{n^2} \) converges we can conclude that \( \sum \frac{n}{n^4 - 1} \) also converges. Indeed we can, if we can appropriately quantify what we mean by \( \frac{n}{n^4 - 1} \approx \frac{1}{n^2} \), according to the following theorem, if we can appropriately quantify what we mean by \( \frac{n}{n^4 - 1} \approx \frac{1}{n^2} \).
Theorem 10.3.4 Limit Comparison Test (LCT). Suppose \( a_n \geq 0, b_n > 0 \) for large enough \( n \), and that \( \lim_{n \to \infty} \frac{a_n}{b_n} = L \in [0, \infty] \).

1. If \( L \in (0, \infty) \), then
   \[
   \sum a_n \text{ converges } \iff \sum b_n \text{ converges, or equivalently } \sum a_n \text{ diverges } \iff \sum b_n \text{ diverges.}
   \]

2. If \( L = \infty \) and \( \sum b_n \) diverges, then \( \sum a_n \) diverges.

3. If \( L = 0 \) and \( \sum b_n \) converges, then \( \sum a_n \) converges.

The idea of the theorem can be more casually described as follows (think \( a_n \approx L \cdot b_n \)):

1. for large \( n \), we have \( a_n \approx L \cdot b_n \), and so if \( L \in (0, \infty) \), both series converge or both diverge, since the “tail end” terms of one series are approximately a constant multiple (\( L \neq 0, \infty \)) of the “tail end” terms of the other;

2. if \( L = \infty \), then \( a_n \) is a larger and larger multiple of \( b_n \), so if \( \sum b_n \) diverges, so does \( \sum a_n \); (If \( \sum a_n \) converges or \( \sum b_n \) diverges, this is indeterminate regarding convergence or divergence of the other);

3. if \( L = 0 \), then \( a_n \) is a smaller and smaller multiple of \( b_n \), so if \( \sum b_n \) converges, so does \( \sum a_n \). (If \( \sum a_n \) converges or \( \sum b_n \) diverges, this is indeterminate regarding convergence or divergence of the other.)

A proof would actually rely on the DCT, by showing that we can eventually make the correct comparison with an appropriate series to give us the conclusion. Without going through every case, we look at the proof that if \( \sum b_n \) converges, and \( a_n/b_n \to L \in (0, \infty) \), then we must conclude \( \sum a_n \) converges.

**Proof:** Assume \( \lim_{n \to \infty} \frac{a_n}{b_n} = L \in (0, \infty) \), and that \( \sum b_n \) converges. Then for large enough \( n \), we have \( a_n/b_n \approx L \), so \( a_n/b_n \in (L - \frac{1}{2}L, L + \frac{1}{2}L) = (\frac{1}{2}L, \frac{3}{2}L) \). Thus for large enough \( n \) we have \( 0 < a_n < \frac{3}{2} \cdot b_n \), so we can direct-compare \( \sum a_n \) to \( \sum \frac{3}{2} \cdot b_n \) (which converges, since \( \sum b_n \) does):

\[
\sum b_n \text{ converges } \implies \frac{3}{2} \cdot \sum b_n = \sum \left( \frac{3}{2} \cdot b_n \right) \text{ converges } \implies \sum a_n \text{ converges,}
\]

by the DCT. The other cases are proven similarly, with modifications.

The case proved above is the most commonly referenced case, and so we make the following definition:

**Definition 10.3.1** For two positive-term series \( \sum a_n \) and \( \sum b_n \), we call the two series limit-comparable if and only if

\[
\lim_{n \to \infty} \frac{a_n}{b_n} = L \in (0, \infty).
\]

\[19\] Technically, we have not defined what it means for \( \lim_{n \to \infty} \frac{a_n}{b_n} = \infty \), so we mean that the sequence \[ a_n/b_n \to \infty \] is null. For the curious reader, a definition of \( c_n \to \infty \) would be \( (\forall M > 0)(\exists N)[n > N \implies c_n > M] \).
Note we omit the cases $L \in \{0, \infty\}$. With this definition, part of the Limit Comparison Test (LCT) can now be phrased:

Two limit-comparable series will both converge, or both diverge.

**Example 10.3.8** Consider $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n^2 + 1}}$. We could use an integral test, but that would require either a complicated formula or trigonometric substitution to find the relevant antiderivative. The naïve Direct Comparison Test would have us note that $1/\sqrt{n^2 + 1} < 1/n$, but $\sum (1/n)$ is not summable, so $\sum 1/\sqrt{n^2 + 1}$ being a “smaller” series than a divergent series is inconclusive.

However, for large enough $n$, it seems $1/\sqrt{n^2 + 1} \approx 1/\sqrt{n^2} = 1/n$, so the series seems similar to the harmonic series $\sum (1/n)$, so we verify that the two are limit-comparable:

$$\lim_{n \to \infty} \frac{\frac{1}{\sqrt{n^2 + 1}}}{\frac{1}{n}} = 0/0 = \lim_{n \to \infty} \frac{\sqrt{n^2}}{\sqrt{n^2 + 1}} = \lim_{n \to \infty} \sqrt{\frac{n^2}{n^2 + 1}} = \frac{1}{\sqrt{1 + 0}} = 1 \in (0, \infty).$$

This shows that the original series is limit-comparable to $\sum \frac{1}{n}$, which diverges, so we can conclude $\sum \frac{1}{\sqrt{n^2 + 1}}$ also diverges.

**Example 10.3.9** Consider $\sum_{n=1}^{\infty} \frac{2n + 1}{n^3 + 5n^2 + 6n}$. While an integral test is do-able, as is a DCT argument, it seems more natural to note that $\frac{2n + 1}{n^3 + 5n^2 + 6n} \approx \frac{2n}{n^3} = \frac{2}{n^2}$ for large $n$. We can limit-compare the series to $\sum \frac{2}{n^2}$, or even $\sum \frac{1}{n^2}$. For this example we will do the latter.

$$\lim_{n \to \infty} \frac{\frac{2n + 1}{n^3 + 5n^2 + 6n}}{\frac{1}{n^2}} = \lim_{n \to \infty} \frac{2n + 1}{n^3 + 5n^2 + 6n} \approx \lim_{n \to \infty} \frac{2n^3 + n^2}{n^3 + 5n^2 + 6n} = 1.$$ 

Our original series is limit-comparable to $\sum \frac{1}{n^2}$, which converges, and so therefore must our original series.

In the conclusive cases where $(a_n)/(b_n) \to L \in \{0, \infty\}$, there is usually a DCT argument that would also get us our result. For instance, if we look at $\sum \frac{\ln n}{n}$, for the DCT we can show that $\frac{\ln n}{n} \geq \frac{1}{n}$, and since $\sum \frac{1}{n}$ diverges we know the same is true of $\sum \frac{\ln n}{n}$. But if we instead limit-compare the two series, we get

$$\lim_{n \to \infty} \frac{\frac{\ln n}{n}}{\frac{1}{n}} = \lim_{n \to \infty} \ln n = \infty,$$

and since $\sum \frac{1}{n}$ diverges we can safely say that so does $\sum \frac{\ln n}{n}$. (Recall the idea is that $\frac{\ln n}{n}$ is a larger and larger multiple of $\frac{1}{n}$, which is not summable.)

**Example 10.3.10** Consider $\sum_{n=1}^{\infty} \frac{\ln n}{n} e^{-n}$. For this we can use either DCT or LCT.

For the Direct Comparison Test, we note that since $\ln n < n$ for large $n$, we have $\frac{\ln n}{n} e^{-n} < 1 \cdot e^{-n} = (\frac{1}{e})^n$, which is summable (geometric series, $|r| = \frac{1}{e} < 1$). Hence so is $\sum \frac{\ln n}{n} e^{-n}$.
For the Limit Comparison Test, we limit-compare the original series to $\sum e^{-n}$, which again converges (see above). So we compute
\[
\lim_{n \to \infty} \frac{\ln n e^{-n}}{e^{-n}} = \lim_{n \to \infty} \frac{\ln n}{n} = 0,
\]
which we get from either one l'Hôpital's Rule step or from our hierarchy of functions. Since $\sum e^{-n}$ converges, and the limit above is zero, so does $\sum \frac{\ln n e^{-n}}{n}$ converge by a special case ($L = 0$) of the LCT. (Recall the idea there is that the given series terms are smaller and smaller multiples of the convergent series $\sum e^{-n}$.)

The limit comparison test can sometimes produce limits which are challenging to compute. Often such examples would be easier with a later technique, as in the next example (which would be more appropriate for a Ratio Test introduced in Section 10.5), but ultimately such tests rely for their proofs on these more primitive tests.

**Example 10.3.11** Consider the series $\sum_{n=1}^{\infty} \frac{n}{\ln n} e^{-n}$.

This has three functions from our hierarchy, listed here in the order they appear in the hierarchy: $\ln n, n$ and $e^n$ (in a denominator).

Now $\frac{n}{\ln n} e^{-n}$ may indeed shrink, but not as quickly as $e^{-n}$, since $\frac{n}{\ln n} \to \infty$. Note that $e \in (2, 3)$, so $\frac{1}{e} \in \left(\frac{1}{2}, \frac{1}{3}\right)$, so we will compare the original series to $\sum \left(\frac{1}{e}\right)^n$, whose terms shrink more slowly than $\sum e^{-n}$, so we will see how $\sum 2^{-n}$ compares to our original series in the limit.

\[
\lim_{n \to \infty} \frac{n}{\ln n} e^{-n} = \lim_{n \to \infty} \frac{n}{\ln n} \left(\frac{2}{e}\right)^n = \lim_{n \to \infty} \frac{n}{\ln n} \cdot \frac{1}{\ln n} = 0.
\]

The last computation is based on the hierarchy of functions again, with the polynomial power $n^1$ divided by an exponential $(e/2)^n$, since $e/2 > 1$. A LHR argument could also work. Noting that $2/e < 1 \implies \ln(2/e) < 0$, this limit can also be arrived at by defining $y = \frac{n}{\ln n} \left(\frac{2}{e}\right)^n$, and finding $\lim_{n \to \infty} y = -\infty$, implying $y \to 0$:

\[
\lim_{n \to \infty} \ln y = \lim_{n \to \infty} \left[ n \ln \frac{2}{e} + \ln n - \ln(\ln n) \right] = -\infty \implies \lim_{n \to \infty} y = \frac{e^{-\infty}}{\infty} = 0.
\]

While this limit on the left looks like an indeterminate case where we need to combine the expressions into one (large) fraction and use l'Hôpital's Rule, in fact if we look at the hierarchy of functions, Theorem 10.3.3 (page 723), we see that the first-degree polynomial term $(\ln \frac{2}{e}) n$ grows much faster in size than $\ln n$, and for that matter $\ln(\ln n)$, so its effect will dominate as $n \to \infty$. Furthermore, the limit above will be $-\infty$, since the coefficient of the dominating $n^1$-term is $\ln \frac{2}{e} < 0$. (Again, $2/e < 1 \implies \frac{2}{e} < 1 \implies \ln \frac{2}{e} < 0$, and in fact $\ln \frac{2}{e} \approx -0.3068528 < 0$.)

However we arrive at the limit, what we have is that
\[
\lim_{n \to \infty} \frac{n}{\ln n} e^{-n} = 0,
\]
and since $\sum \left(\frac{1}{e}\right)^n$ converges, so does $\sum \frac{n}{\ln n} e^{-n}$. (Again the idea was $\frac{n}{\ln n} e^{-n}$ is a shrinking multiple of $(1/2)^n$, which is summable.)
While $\sum \frac{1}{lnn} e^{-n}$ had terms bigger than the geometric series $\sum e^{-n}$, its terms were actually smaller than another geometric series $\sum \left(\frac{1}{2}\right)^n$, whose terms were slightly larger than those of $\sum e^{-n}$.

The argument in the previous example was quite sophisticated, and indeed was not really necessary, after we develop the tools of Section 10.5, but much can be done with these more primitive tests. (Also, as mentioned earlier, the more sophisticated tests ultimately rely on DCT for their proofs, so in principle any series using those tests could have convergence or divergence proven by DCT, though it is often much easier to use the computational machinery built into the later tests.)

10.3.4 Summary

The two tests developed here, namely DCT and LCT, rely on our being able to identify series to compare our given series to in a meaningful way to predict convergence or divergence of our given series, based upon convergence or divergence of a known series, typically a $p$-series or geometric series from before. This takes practice, because if we compare a series to another in an inconclusive (or indeterminate) way, we will be tempted to apply the two tests in invalid ways.

Briefly, the main theorems for positive-term series were

\[
\begin{align*}
\text{(DCT)} & \quad \begin{cases} 
0 \leq a_n \leq b_n, & \sum b_n \text{ converges} \implies \sum a_n \text{ converges}, \\
0 \leq a_n \leq b_n, & \sum a_n \text{ diverges} \implies \sum b_n \text{ diverges}, 
\end{cases} \\
\text{(LCT)} & \quad \lim_{n \to \infty} \frac{a_n}{b_n} = L \in (0, \infty) \implies \sum a_n, \sum b_n \text{ both converge or both diverge}. 
\end{align*}
\]

There were other cases of the LCT, namely $L = 0$ or $L = \infty$, but if we keep in mind that $a_n \approx L \cdot b_n$ for $n$ large, we can predict them as well.\(^{20}\)

\(^{20}\)We wrote down two cases, but each had contrapositives as well, so it can be rather confusing to list them all.
10.3.5 Elementary Series Theorems

In order to take full advantage of the Direct and Limit Comparison Tests (DCT and LCT) above, it is best if we have some elementary results regarding series at our disposal. These are intuitive, though as always some care must be taken to be clear about what they do and do not say.

**Theorem 10.3.5** Suppose \( k \neq 0 \). Then \( \sum a_n \) converges if and only if \( \sum (k \cdot a_n) \) converges, and furthermore
\[
\sum (k \cdot a_n) = k \cdot \sum a_n.
\]

**Proof:** This is just the observation that, assuming \( n \) starts at 1 in the sum \( \sum (k \cdot a_n) \), we have
\[
\sum_{N \downarrow} (k \cdot a_n) = ka_1 + ka_2 + \cdots + k a_N = k(a_1 + a_2 + \cdots + a_N).
\]

**Example 10.3.12** Discuss the convergence or divergence of \( \sum \frac{1}{5^n} \) and \( \sum n^2 \).

**Solution:** For each of these, we simply “factor” the multiplicative constants:
- \( \sum \frac{1}{5^n} = \frac{1}{5} \sum \frac{1}{n} \) diverges (since \( \sum \frac{1}{n} \) diverges).
- \( \sum \frac{2}{n^2} = 2 \sum \frac{1}{n^2} \) converges (since \( \sum \frac{1}{n^2} \) converges).

Now either of the above examples could have been solved using the LCT above, since the first is limit-comparable to \( \sum \frac{1}{n} \), and the other to \( \sum \frac{1}{n^2} \), but the arguments above are much more efficient for these examples.

We next look at combinations of two series, in some cases.

**Theorem 10.3.6** Given two series \( \sum a_n \) and \( \sum b_n \).

1. If \( \sum a_n \) and \( \sum b_n \) both converge, then
\[
\sum (a_n \pm b_n) = \sum a_n \pm \sum b_n.
\]

Furthermore, if one of the series, \( \sum a_n \) or \( \sum b_n \) converges but the other diverges, then \( \sum (a_n \pm b_n) \) diverges.

The idea of the proof is pretty simple. We start with the case where both converge and we are adding. Suppose \( S_N \) is the partial sum of the series \( \sum (a_n + b_n) \). Then
\[
\sum_{N \downarrow} (a_n + b_n) = a_1 + b_1 + a_2 + b_2 + \cdots + a_N + b_N = \sum a_n + \sum b_n.
\]

Thus \( \sum (a_n + b_n) \) is the sum of two convergent series. From the limit argument above as well, we can see that if one of the series, say \( \sum a_n \) converges but \( \sum b_n \) diverges, it is necessary that \( \sum (a_n + b_n) \) diverges. In that case we wouldn’t actually write (10.26), because it makes no sense to have an equation where a divergent series is an expression added to another quantity. Similarly one should not write, for instance
\[
\sum_{n=1}^{\infty} \left( \frac{1}{n} - \frac{1}{n+1} \right) = \sum_{n=1}^{\infty} \frac{1}{n} - \sum_{n=1}^{\infty} \frac{1}{n+1} \quad \text{(FALSE!)}
\]
10.3. COMPARISON TESTS

because the series on the left actually converges (as a telescoping series, or we can limit or direct compare it to \( \sum \frac{1}{n^2} \) since it can be combined to give \( \sum \frac{1}{n(n+1)} \)), where the two on the right do not.

The upshot of all this is that (10.25) and (10.26) are both true and make perfect sense, if the underlying series converge. When they don’t, there may or may not be something useful to say about the combined sum.
Figure 10.4: Illustration of the intuition for the Alternating Series Test. Here \(a_1, a_3, a_5, \cdots\) are all positive (represented by black arrows), while \(a_2, a_4, a_6, \cdots\) are all negative (represented by gray arrows), and for simplicity \(|a_1| > |a_2| > |a_3| > \cdots\), and \(|a_n| \to 0\) as \(n \to \infty\). This will force \(\{S_{2n}\}\) (even partial sums) to be an increasing sequence, and \(\{S_{2n-1}\}\) (odd partial sums) to be decreasing. Since these sequences of partial sums are “coming together,” meaning their distances from each other \(|S_N - S_{N-1}| = |a_N|\) is shrinking to zero as \(N \to \infty\), we see these must converge to some number \(S \in \mathbb{R}\). That is the essence of the Alternating Series Test (AST) below.

10.4 Alternating Series and Absolute Convergence

Unlike the previous two sections on positive-term series, in this section we look at series \(\sum a_n\) which alternate signs, between positive and negative terms. Such series are called, naturally enough, alternating series. It turns out that if the terms alternate, and their sizes shrink monotonically to zero, then that is enough for the series to be known to converge. Basically, even if the terms do not shrink to zero quickly, there is a recurring partial cancellation, somewhat like we saw in telescoping series, though here we are much less likely be able to find a simple formula for the partial sums whose limits we can compute.

Some alternating series rely on the alternation for convergence, and are thus called conditionally convergent, where other alternating series have the terms shrink fast enough that the alternation is not in fact necessary to ensure convergence. The latter series are called absolutely convergent, for reasons that will become clear later in this section.

10.4.1 Alternating Series Test (AST)

The main result in this section is the following (see Figure 10.4).

Theorem 10.4.1 Suppose \(\sum_{n=1}^{\infty} a_n\) satisfies the following three conditions:

1. The terms of the series alternate signs \((+, -, +, - , +, \cdots)\) i.e., \(a_{n+1} < 0\) for all \(n = 1, 2, 3, \cdots\).

2. \(|a_1| \geq |a_2| \geq |a_3| \geq \cdots\), i.e., \(|a_n|\) \(\to 0\) as \(n \to \infty\) is a decreasing sequence.

3. \(\lim_{n \to \infty} |a_n| = 0\).
Rephrased, we suppose the terms of \( \{a_n\} \) alternate signs and shrink in absolute size monotonically to zero. If this is the case, then \( \sum a_n \) converges.

Recall that \( \sum a_n \) converges means that the sequence \( \{S_N\}_{N=1}^{\infty} = \left\{ \sum_{n=1}^{N} a_n \right\} \) of partial sums converges to some finite number \( S \), i.e., \( S_N \to S \in \mathbb{R} \) as \( N \to \infty \). This is partially illustrated in Figure 10.4, page 732, where we see the partial sums form a sequence of terms \( S_N \) which oscillate left and right on the number line, but between tighter and tighter confines.

It should also be noted that it is only necessary for the series’ terms to show the alternation and shrinking-to-zero (monotonically) hypotheses of the AST “eventually,” i.e., for all \( n > N \), some \( N \geq 1 \). For simplicity many texts assume \( |a_1| > |a_2| > |a_3| > \cdots \) and \( |a_n| \to 0 \), but the weaker “\( \geq \)” suffices. Note that we can summarize quickly (2) and (3) above by writing

\[
|a_1| \geq |a_2| \geq |a_3| \geq \cdots \to 0.
\]

Along with the alternation, this implies convergence.\(^{21}\)

It is also interesting to note that if \( |a_n| \neq 0 \), then \( a_n \neq 0 \) implying \( \sum a_n \) diverges (NTTFD, Section 10.2).

**Example 10.4.1** Consider the series \( \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \).

If we write out a few terms of this series, we get

\[
\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots.
\]

Clearly the series’ terms alternate signs, and shrink monotonically in absolute value to zero:

\[
\lim_{n \to \infty} |a_n| = \lim_{n \to \infty} \left| \frac{(-1)^{n+1}}{n} \right| = \lim_{n \to \infty} \frac{|(-1)^{n+1}|}{|n|} = \lim_{n \to \infty} \frac{1}{n} = 0.
\]

If we wish to check for monotonicity of \( |a_n| \), we can note that the denominator is obviously increasing monotonically, and the numerator is constant, so the fractions which are \( |a_n| = \frac{1}{n} \) are clearly decreasing (to zero) monotonically. We could also note that if \( f(x) = \frac{1}{x} \), then \( f'(x) = -1/x^2 < 0 \) for \( x \geq 1 \) (or even all \( x \neq 0 \)), so \( |a_n| = f(n) \) is clearly decreasing monotonically for \( n \geq 1 \).

With all this, we know the series converges by the AST.

Usually a short inspection assures us that the terms in the series shrink in absolute size monotonically to zero, though sometimes sophisticated arguments are required to make this clear.

Note that the convergence would still apply if the alternation of signs began with a negative term, as in

\[
\sum_{n=1}^{\infty} \frac{(-1)^{n}}{n} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} + \frac{1}{6} - \cdots.
\]

\(^{21}\)Some texts define an alternating series by assuming \( \{a_n\} \) is a positive-term series, and then considering the series \( \sum (-1)^{n} a_n \) or \( \sum (-1)^{n+1} a_n \) or similar series. This arguably has some advantage later, but it is minor, if existent at all.
CHAPTER 10. SERIES OF CONSTANTS

In fact the series above would be the additive inverse (negative) of the series in the previous example above (so they would both converge or both diverge). Furthermore, it should be pointed out that this new series can also be written
\[\sum_{n=1}^{\infty} \frac{\cos(n\pi)}{n} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} + \frac{1}{6} - \cdots.\]

One interesting aspect of the convergent alternating series is that we can estimate \(|S - S_N|\) easily from an illustration such as Figure 10.4 at the beginning of the section (page 732).

**Theorem 10.4.2** Suppose \(\sum a_n\) is an alternating series which satisfies the hypotheses (1)–(3) of the AST, Theorem 10.4.1, page 733. Then
\[|S_N - S| \leq |a_{N+1}|. \quad (10.27)\]
Furthermore, if we can replace the inequality “\(\geq\)” with “\(>\)” in the hypotheses of the AST, we can replace the inequality “\(\leq\)” with “\(<\)” in (10.27).

Indeed, the distance between \(S_N\) and \(S\) is always (for such series) no more than \(|a_{N+1}|\) because by adding \(a_{N+1}\) to \(S_N\) we “overshoot” \(S\) in arriving at \(S_{N+1}\).

For a simple application of (10.4.2), suppose we wish to approximate the series \(S = \sum (-1)^n n \rightleftharpoons\) by taking \(S_N\) for \(N\) large enough that \(|S - S_N| < 0.001\). Then we can use this estimate to find \(N\) large enough to be sure \(S_N\) is indeed within 0.001 of the full series \(S\). We do this by inserting the inequality (10.27) within \(|S - S_N| < 0.001\):
\[|S - S_N| < |a_{N+1}| \quad \text{inserted} \quad \Rightarrow \quad \frac{1}{N+1} < 0.001\]
\[\Rightarrow \quad \frac{1}{0.001} < N + 1\]
\[\Rightarrow \quad 1000 < N + 1\]
\[\Rightarrow \quad 999 < N.\]
Thus we need \(N > 999\), or \(N \geq 1000\) to guarantee by (10.27) we have \(|S - S_N| < 0.001\).\(^{22}\)

Other series which we can quickly see converge by the AST follow:
\[\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}, \quad \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{\ln n}, \quad \sum_{n=1}^{\infty} \frac{(-1)^n}{n!}, \quad \sum_{n=1}^{\infty} (-1)^n \sin \left(\frac{1}{n}\right).\]
In fact the last series may warrant a check:
\[f(x) = \sin \frac{1}{x} \quad \Rightarrow \quad f'(x) = \left(\cos \frac{1}{x}\right) \cdot \left(-1 \times \frac{1}{x^2}\right) < 0\]
for large enough \(x\) that \(\cos \frac{1}{x} \approx \cos 0 = 1 > 0\). Furthermore, \(|(-1)^n \sin \frac{1}{n}| \to |\sin 0| = 0\) as \(n \to \infty\).

On the other hand, we must be careful to note that alternation alone is not enough to conclude a series converges.

\(^{22}\) Using methods from the next chapter, we can in fact show that \(\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \ln 2 \approx 0.693147181\). To guarantee this kind of accuracy (\(\pm 10^{-9}\)) using (10.27) we would, for this series, need to sum \(N = 10^9\) terms, which would require some care and skill even for a computerized computation.
Example 10.4.2 Consider \( \sum_{n=1}^{\infty} \frac{(-1)^n}{n+1} \).

Here we have alternation, but

\[
\lim_{n \to \infty} \left| \frac{(-1)^n}{n+1} \right| = \lim_{n \to \infty} \frac{n}{n+1} = 1 \neq 0,
\]

and so the series diverges. The last limit can be computed using either algebra (factor \( n \) from the numerator and denominator) or l'Hôpital’s Rule.

Similarly, \( \sum \frac{(-1)^n}{n \ln n} \) will not converge, since

\[
|\frac{(-1)^n}{n \ln n}| = \frac{n}{(\ln n)^2} \to \infty \text{ as } n \to \infty.
\]

Ultimately it is the NTTFD that lets us conclude divergence, since \( |a_n| \to 0 \iff a_n \to 0 \).

10.4.2 Absolutely and Conditionally Convergent Series

Here we point out that there is a convergence which is stronger (more stringent) than our previous definition that \( S_N \to S \in \mathbb{R} \). Before arriving at a definition of this stronger convergence criterion, however, we first look more closely at two similar but crucially different alternating series:

\[
\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots \tag{10.28}
\]

\[
\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} = 1 - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + \frac{1}{25} - \frac{1}{36} + \cdots \tag{10.29}
\]

Note that if we remove the alternation from the first series (10.28), it becomes the harmonic series and therefore diverges. That is not the case with the second series, which if we remove the alternation we get \( \sum \frac{1}{n} \), which is a \( p \)-series with \( p = 2 > 1 \) so it converges.

Putting this another way, the first series relies on the alternation to converge; the second series has terms which shrink fast enough that if we did not allow alternation (by removing the factor \( (-1)^{n+1} \) or other methods), we still get a finite number for our “infinite series.”

The manner in which we detect if the terms shrink fast enough that they do not require alternation is to insert absolute values around each term in the series, which makes each term nonnegative and therefore eliminates the partial cancellation which the AST relied upon for the intuition behind that theorem.

So we note that

\[
\sum_{n=1}^{\infty} \frac{|(-1)^{n+1}|}{n} = \sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \cdots \text{ (diverges)},
\]

\[
\sum_{n=1}^{\infty} \frac{|(-1)^{n+1}|}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \frac{1}{36} + \cdots \text{ (converges)}.
\]

So somehow the convergence of series (10.29) seems stronger than that of series (10.28), but we must make this more precise, as we do below.

Definition 10.4.1 We call a series \( \sum a_n \) absolutely convergent if and only if \( \sum |a_n| \) converges.

23This is not true for limits other than zero; for other limits, we only have \( \iff \). For instance, \( \lim_{n \to \infty} a_n = -3 \implies \lim_{n \to \infty} |a_n| = 3 \). In fact \( \lim_{n \to \infty} |a_n| = 3 \) is even true for the divergent sequence \( \{3(-1)^n\} \). However, \( |x| = 0 \iff x = 0 \), while \( |x| = 3 \iff x = 3 \).
So the term “absolutely” refers to the absolute values we inserted. However, while “absolutely convergent” seems in this context to mean the series with absolute values inserted does converge, in another context “absolutely convergent” seems to indicate a type of magnified “regular convergence” (in the sense $S_N \to S \in \mathbb{R}$), with “absolutely” as an adjective. In fact, both interpretations are correct, the second one following from the theorem below.

**Theorem 10.4.3** Suppose $\sum a_n$ is absolutely convergent, i.e., $\sum |a_n|$ converges. Then $\sum a_n$ also converges, in the sense that its partial sums form a convergent sequence $\{S_N\}$.

In other words, absolute convergence implies regular convergence:

$$\sum |a_n| \text{ converges } \implies \sum a_n \text{ converges.} \quad (10.30)$$

We will not prove this, since it is more appropriate for a course in real analysis. However it should have the ring of truth.

**Example 10.4.3** Consider $\sum_{n=0}^{\infty} \frac{(-1)^n}{n!}$.

While this series converges by the AST, we can also prove that it converges absolutely. In fact

$$\left| \frac{(-1)^n}{n!} \right| = \frac{1}{n!} < \frac{1}{n^2}$$

for large enough $n$. Since $\sum \frac{1}{n^2}$ converges, so must $\sum \frac{1}{n!}$. Thus $\sum_{n=0}^{\infty} \frac{(-1)^n}{n!}$ is absolutely convergent.

Note that the series in the above example is defined at $n = 0$ while $\sum \frac{1}{n^2}$ is not, but we only need to be sure the “tail ends” of the series are direct-comparable. Note also that once we insert absolute values, we are back to the positive-term series tests, such as the Integral Test, DCT (used above), and LCT.

It should also be pointed out that absolute convergence and convergence are the same for positive-term series. In fact they are also the same for negative-term-only series as well. Only when we know there is some alternation do the two concepts differ.

Finally, if the series has some alternation of signs, then $\sum |a_n| \neq \sum a_n$, and in fact $|\sum a_n| < \sum |a_n|$ since (among other reasons) the former allows cancellation and the latter does not.\(^{24}\)

So we have series that converge absolutely, and series which converge but not absolutely, and series which diverge. For the second type we have another name to identify them more precisely:

**Definition 10.4.2** If $\sum a_n$ converges but not absolutely, we call that series conditionally convergent.

In other words, $\sum a_n$ is conditionally convergent iff $\sum a_n$ converges but $\sum |a_n|$ diverges. In such a case we note that the convergence of the original series must have been due to some alternations of sign,\(^{25}\) and if we remove the alternation by inserting absolute values around each term, the terms of the series do not shrink fast enough to be summable, so convergence is conditioned on the alternation.

\(^{24}\)It is always important to distinguish between convergence of a series and what it actually converges to, which are two different questions. Many of the convergence tests do not pretend knowledge of the actual value of the series, though some hints regarding its value may be present in the logic of a given test, or its particular application.

\(^{25}\)though not necessarily a consistent $+ - + - + - \cdots$, since other patterns may similarly account for it, such as $+ - + - + - + - + - + - + -$ or similar.
Thus \( \sum \frac{(-1)^{n+1}}{n} \) converges (AST), but not absolutely, and is thus conditionally convergent.

So to restate a fact mentioned earlier, any given series is either absolutely convergent, conditionally convergent, or divergent. The union of the first two types is what we simply call convergence \((S_N \to S \in \mathbb{R})\). Next we list a few quick examples, some already considered, but which help put these concepts into context.

<table>
<thead>
<tr>
<th>series</th>
<th>converges?</th>
<th>absolutely?</th>
<th>conditionally?</th>
<th>remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sum \frac{(-1)^n}{n^2} )</td>
<td>yes</td>
<td>yes</td>
<td>—</td>
<td>( \sum \frac{</td>
</tr>
<tr>
<td>( \sum \frac{(-1)^n}{n} )</td>
<td>yes</td>
<td>—</td>
<td>yes</td>
<td>( \sum \frac{</td>
</tr>
<tr>
<td>( \sum \frac{(-1)^n \ln n}{n} )</td>
<td>yes</td>
<td>—</td>
<td>yes</td>
<td>( \sum \frac{</td>
</tr>
<tr>
<td>( \sum \frac{(-1)^n n}{n+1} )</td>
<td>no</td>
<td>—</td>
<td>—</td>
<td>( \frac{</td>
</tr>
<tr>
<td>( \sum \frac{(-2)^n}{3^{n+1}} )</td>
<td>yes</td>
<td>yes</td>
<td>—</td>
<td>( \sum \frac{</td>
</tr>
</tbody>
</table>

The series we get when we insert absolute values can be tested using, respectively, \( p \)-series \((p = 2)\), \( p \)-series \((p = 1)\), DCT or the Integral Test, NTTFD, and geometric series convergence criteria.

Note that absolute convergence and conditional convergence both imply “old-fashioned” convergence \((S_N \to S \in \mathbb{R})\), but otherwise are mutually exclusive: a series can not be both absolutely convergent and conditionally convergent. One way to illustrate this is with a “possibility tree” like given below.

```
Converges? Absolutely?

yes  yes  absolutely convergent

yes  no   conditionally convergent

no    no   divergent
```

In previous sections, we only concerned ourselves with the first question in the tree, regarding convergence or divergence. Now we get more specific, and ask what kind of convergence. To be sure, if the terms we attempt to sum all have the same sign, then convergence and absolute convergence are the same. It is when the terms to add have nonconstant sign that we ask whether the series diverges, or converges only because of the alternation, or would have converged even without the alternation.

The next section contains two tests which can only detect absolute convergence, or divergence, or, in many cases, neither. Indeed, as in most of our series tests, at times they are conclusive and at other times they are inconclusive.

### 10.4.3 One Last Remark Concerning Absolute Convergence

One interesting aspect of an absolutely convergent sequence is that it does not matter what order we use to sum the terms, as long as all are summed in the limit. In fact we can even pick...
out two or more “subseries” and sum them separately. So for instance, if \( \sum |a_n| \) converges, then
\[
\sum a_n = \sum a_{2n} + \sum a_{2n-1},
\]
i.e., we can add the even and odd terms separately and get the same result. We can not do this with a conditionally convergent, alternating series since both of these “subseries” will diverge.

We will not prove this remark, but upon some reflection it should have a ring of truth. It is similar to how we accumulate the areas of a convergent improper integral, even if it has multiple “improper” endpoints.

In fact, an elementary homework exercise in senior-level analysis is to show that if you choose any real number \( R \in \mathbb{R} \), then any conditionally convergent series can have its terms rearranged in such a way that the sum converges to \( R \), in the sense that its new (after the rearrangement) partial sums \( S_N \) do. (We can also rearrange the terms so the series diverges.) Thus the order in which we add terms matters very much in any conditionally convergent series. This is not the case with absolutely convergent series; order of addition in \( \sum a_n \) is not an issue if \( \sum |a_n| \) converges.

**Example 10.4.4** Consider the series \( \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \).

We cannot break this into two series of even and odd terms, because if we attempted to do so with this convergent series, we would get
\[
\sum_{n \text{ odd}} \frac{(-1)^{n+1}}{n} + \sum_{n \text{ even}} \frac{(-1)^{n+1}}{n} = \left(1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \cdots\right) + \left(-\frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \frac{1}{8} - \cdots\right)
\]
\[
= \sum_{n=1}^{\infty} \frac{1}{2n-1} - \sum_{n=1}^{\infty} \frac{1}{2n} \quad \text{(FALSE!)}
\]

The problem is that both series diverge to \( \infty \). Furthermore, there are many orders we could rearrange these terms. For demonstration purposes, we can argue that one method leaves us with a sum greater than 1, and another with a sum less than \(-1\). Here is how we can do that.

1. Add \( 1 + \frac{1}{9} + \frac{1}{5} = 1 + \frac{8}{15} > 1 + \frac{1}{2} \).
2. Add \( -\frac{1}{2} \) to this previous sum: \( 1 + \frac{1}{9} + \frac{1}{5} - \frac{1}{2} > 1 \).
3. Add \( \frac{1}{7} + \frac{1}{9} + \cdots \) to the sum until it is greater than \( 1 + \frac{1}{2} \). We can do this since \( \frac{1}{7} + \frac{1}{9} + \frac{1}{11} + \cdots \) diverges.
4. Add \( -\frac{1}{4} \), and the partial sum will still be greater than 1.
5. Add enough of the remaining odd terms of the original series until the sum is greater than \( 1 + \frac{1}{6} \).
6. Add \( -\frac{1}{6} \). The sum is still greater than 1.
7. Continue with this pattern forever. The sums will continue to be greater than 1, and we can always add more terms from the tail-end of the (divergent) odd series to overcome the next term to be added in the even series, and the sums will continue to be greater than 1.

A similar process can be achieved to keep the sum less than \(-1\), by choosing enough even terms to sum to a number less than \(-2\), then adding 1, then adding enough of the remaining
even terms until the sum is a number less than \(-1 - \frac{1}{3}\), and then adding \(\frac{1}{3}\), and summing enough of the remaining even terms until the total sum is less than \(-1 - \frac{1}{5}\), then adding \(\frac{1}{5}\), and so on.

Both procedures give us a rearrangement of the terms of the original alternating series, but the new partial sums will remain forever greater than 1 in the first case, and less than \(-1\) in the second. Both procedures ensure that every term in the original sum is included eventually in the new sum.

For an absolutely convergent series like \(\sum\frac{(-1)^{n+1}}{n^2}\), the above process is impossible because no partial series will diverge (see Steps 3, 5 above), and any way we add all of the terms (eventually with some limit process) will result in the same sum.
10.5 Ratio and Root Tests

The Ratio and Root Tests introduced here detect if the “infinite tail-end” behavior of a given series is comparable to the behavior of a geometric series, and if so, with what ratio. For many interesting series, these tests can lead us to easily determine if the given series converges absolutely, or diverges. (Conditional convergence and its detection was addressed in Section 10.4.) Which test is used depends upon which computation is easier for a given series, usually as determined by the algebra involved in certain limit computations.

The proofs of the tests rely upon knowledge of geometric series and the Direct Comparison Test, so in principle any of the problems here can be computed using a Direct Comparison Test, or sometimes using other tests (such as the Integral Test), but the tests here will often be much easier to use, depending upon the series.

As with other techniques, there will be many important series for which the Ratio Test or Root Test is appropriate, and some other important series for which they are both inconclusive and therefore useless, so these tests are not replacements for the previous tests. For instance, the tests here are inappropriate for analyzing $p$-series, as we will see, but many other series will be difficult to conclusively analyze any other way besides using one of these two tests given in this section.

Note that if a geometric series converges, it does so absolutely, since it is in fact $|r| = |a_{n+1}|/|a_n|$ which is the crucial quantity; $\sum \alpha r^n$ converges if and only if $\sum |\alpha r^n|$ converges since the absolute values of the ratios of these two series are both $|r|$. In other words, applying the absolute value to each term would not change convergence (though it would certainly change the value of the series).

For the ratio and root tests, we will define a quantity $\rho$ (the lower-case Greek letter “rho”) for a given series, and this $\rho$ will mimic $|r|$ from geometric series. There are two ways we define $\rho$, but they are usually equal (at least for examples found here), though we decline to prove it for this discussion. What we will prove is that they are equal to $|r|$ for a geometric series. Furthermore, when $\rho < 1$ (i.e., $\rho \in [0,1)$) the series will converge, and do so absolutely; when $\rho > 1$ the series will diverge, and in fact a well-informed version of the NTTFD would apply but it might not be obvious because the limit involved in the NTTFD ($a_n \not\to 0$) could be much harder to compute (or prove) than those that appear in the Ratio or Root Test.

Unfortunately, when $\rho = 1$ the tests will in fact be inconclusive, as we will show with examples, so there is not a perfect correlation with conclusions based upon $|r|$ from the geometric series, so we have to make note of that. What this tells us is that there is some room between the geometric series with $|r| = 1$ (which diverges), and series with $\rho = 1$, some of which converge and some of which diverge. In those cases we have to look to one of our previous techniques to attempt to find a conclusive test. As hinted previously, the $p$-series will all have $\rho = 1$, and of course some converge and others diverge. Many alternating series will also have $\rho = 1$, but some will converge (because $|a_n|$ shrinks to zero), while others will diverge (by NTTFD, since $|a_n| \not\to 0$).

For the Ratio Test (RAT) we will define

$$\rho_{\text{Ratio}} = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right|.$$  

Note that for a geometric series $\sum \alpha r^n$, with $\alpha, r \neq 0$, this becomes

$$\rho_{\text{Ratio}} = \lim_{n \to \infty} \left| \frac{\alpha r^{n+1}}{\alpha r^n} \right| = \lim_{n \to \infty} |r| = |r|.$$
Thus $\rho_{\text{Ratio}} = |r|$ for the case of a geometric series. For the Root Test (ROOT) we instead define

$$\rho_{\text{Root}} = \lim_{n \to \infty} \sqrt[n]{|a_n|} = \lim_{n \to \infty} |a_n|^{1/n}.$$ 

For the geometric series this becomes

$$\rho_{\text{Root}} = \lim_{n \to \infty} |\alpha r^n|^{1/n} = \lim_{n \to \infty} |\alpha|^{1/n} (|r|^n)^{1/n} = |\alpha|^{0} |r| = |r|.$$ 

Rather than distinguish between $\rho_{\text{Ratio}}$ and $\rho_{\text{Root}}$, we will simply refer to both as $\rho$, especially since they are usually the same number. However, for each computation it should be clear from the context which definition of $\rho$ is used.

If we calculate $\rho$ for a series $\sum a_n$, and $\rho < 1$, then we can interpret this to mean $\sum |a_n|$ behaves very much like a geometric series with ratio $\rho$, in the sense that they somehow converge similarly (and absolutely, but probably to different values). The same is true of series $\sum a_n$ where $\rho > 1$, diverging the same way such a geometric series would. As mentioned before, unfortunately the tests below are not comprehensive. If $\rho = 1$ or does not exist, we can not immediately decide from that fact alone whether or not the series converges (absolutely or otherwise), or diverges. However the tests are quite useful for numerous and important cases. These tests follow next.

**Theorem 10.5.1 Ratio Test (RAT):** Suppose for a series $\sum a_n$ the limit

$$\rho = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right|$$

exists, or is $\infty$ (i.e., the sequence $\{|a_{n+1}/a_n|\}$ diverges to $\infty$). Then

1. If $\rho \in [0, 1)$, i.e., $\rho < 1$ then $\sum a_n$ converges absolutely.
2. If $\rho > 1$ (including $\rho = \infty$), then $\sum a_n$ diverges.
3. If $\rho = 1$ then this test is inconclusive (and some other test must be used).

**Theorem 10.5.2 Root Test (ROOT):** Suppose for a series $\sum a_n$ the limit

$$\rho = \lim_{n \to \infty} \sqrt[n]{|a_n|} = \lim_{n \to \infty} |a_n|^{1/n}$$

exists, or is $\infty$. Then

1. If $\rho \in [0, 1)$, i.e., $\rho < 1$ then $\sum a_n$ converges absolutely.
2. If $\rho > 1$ (including $\rho = \infty$), i.e., $\rho > 1$ then $\sum a_n$ diverges.
3. If $\rho = 1$ then this test is inconclusive (and some other test must be used).

It is somewhat an art to decide which of the two tests (if any) is better suited for calculating $\rho$. The first is better for factorials, and the second usually better if $a_n$ is of the form $(f(n))^n$, but there are exceptions. Of the two tests the Ratio Test is more often used, but there are certainly cases where the Root Test is closer to ideal, as in cases where $|a_n|^{1/n}$ simplifies nicely.

---

26In fact, for most series found in calculus textbooks, $\rho_{\text{Ratio}} = \rho_{\text{Root}}$, so both are just referred to as $\rho$. However, if $\sum a_n$ had infinitely many zero terms, or nearly-zero terms between larger terms (an oscillating sequence of some kind), the Ratio Test might be more problematic than the Root Test, but where both $\rho$'s are defined they will coincide.
Example 10.5.1  Consider \( \sum_{n=0}^{\infty} \frac{1}{n!} \). Using the Ratio Test, we compute \( \rho \). Since computations involving factorials will become more and more important in this and the next chapter, we will write out this particular computation in some extra detail.

\[
\rho = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \to \infty} \frac{1}{n+1} \cdot \frac{n!}{(n+1)!} = \lim_{n \to \infty} \frac{n!}{(n+1)n!} = \lim_{n \to \infty} \frac{1}{n+1} \cdot \frac{1/\infty}{0} < 1.
\]

Therefore \( \sum_{n=0}^{\infty} \frac{1}{n!} \) converges (absolutely, but that is redundant here since all terms are positive).

We could have also used a Direct Comparison Test (DCT) along with our “Hierarchy of Functions” (Theorem 10.3.3, page 723) to just state that eventually \( 0 < \frac{1}{n} < \frac{1}{n^2} \), or even \( 0 < \frac{1}{n} < \frac{1}{2n} \), both of which are summable. However, we computed \( \rho = 0 \) which means that the terms added in the series shrink faster than any geometric series (which would have positive \( \rho \)), which is a useful insight in its own right.

Example 10.5.2  Consider the series \( \sum_{n=0}^{\infty} \frac{(-2)^n}{n!} \). We compute \( \rho \) for application of the Ratio Test again:

\[
\rho = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \to \infty} \left| \frac{(-2)^{n+1}}{(n+1)!} \right| \cdot \frac{n!}{(-2)^n n!} = \lim_{n \to \infty} \frac{2^{n+1}}{2^n} \cdot \frac{n!}{(n+1)n!} = \lim_{n \to \infty} \frac{2}{n+1} \cdot \frac{2/\infty}{0} < 1.
\]

Therefore \( \sum_{n=0}^{\infty} \frac{(-2)^n}{n!} \) converges absolutely. (Note that we also get \( \sum_{n=0}^{\infty} \frac{2^n}{n!} \) converges.)

In fact the example above gives us an elegant proof that \( \lim_{n \to \infty} \frac{2^n}{n!} = 0 \). The example shows, by the Ratio Test, that the series \( \sum_{n=0}^{\infty} \frac{2^n}{n!} \) converges, so (by the contrapositive of the NTTFD) the terms must shrink to zero, i.e.,

\[
\lim_{n \to \infty} \frac{2^n}{n!} = 0.
\]

The Ratio Test is clearly quite useful when there is much cancellation in \( |a_{n+1}/a_n| \), such as with factorials and some exponential functions.

Example 10.5.3  Consider the series \( \sum_{n=1}^{\infty} \frac{2^n}{n^n} \). A Ratio Test argument would be unwieldy (as the reader is invited to check), so we look instead to the Root Test:

\[
\lim_{n \to \infty} \frac{a_n}{n^m} = \lim_{n \to \infty} \left( \frac{2^n}{n^n} \right)^{1/n} = \lim_{n \to \infty} \frac{2}{n^{1/n}} = 0 < 1.
\]

Thus the series converges (absolutely, but that is redundant here).
10.5. RATIO AND ROOT TESTS

As noted after the previous example, we then easily get that \(2^n/n^n \to 0\) as \(n \to \infty\), giving another of our orders in our hierarchy of functions (Theorem 10.3.3, page 723). Replacing 2 by any other number will show that \(a^n\) grows more slowly than \(n^n\). Next we see a proof that \(n^n\) grows faster than \(n!\).

**Example 10.5.4** Consider the series \(\sum_{n=1}^{\infty} \frac{n!}{n^n}\).

The \(n!\) term seems better suited to the Ratio Test, where the \(n^n\) term indicates a Root Test. Since we can more easily deal with a ratio of powers than a root of a factorial, we will opt for the Ratio Test.

\[
\rho = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \to \infty} \frac{(n+1)!}{n^n} \frac{n^n}{(n+1)^{n+1}} = \lim_{n \to \infty} \frac{(n+1)!}{n!} \cdot \frac{n^n}{(n+1)^{n+1}} = \lim_{n \to \infty} \frac{(n+1) \cdot n^n}{(n+1)^{n+1}}
\]

For this \(1^\infty\) form limit, we need some logarithmic arguments. We will let \(y = \left(\frac{n}{n+1}\right)^n\) and find \(\lim_{n \to \infty} \ln y\):

\[
\lim_{n \to \infty} \ln y = \lim_{n \to \infty} n \ln \left(\frac{n}{n+1}\right) = \lim_{n \to \infty} \frac{\ln n - \ln (n+1)}{n-1} \cdot n \ln \left(\frac{n}{n+1}\right) = \lim_{n \to \infty} \frac{\ln \frac{n}{n+1}}{\frac{n}{n+1}} = \lim_{n \to \infty} \frac{n}{n+1}
\]

\[
\Rightarrow \quad \rho = \lim_{n \to \infty} y = \lim_{n \to \infty} e^{\ln y} = e^{\lim_{n \to \infty} \frac{\ln \frac{n}{n+1}}{n-1}} = e^{\frac{1}{e}} < 1.
\]

Thus \(\rho < 1\), and the series converges (absolutely, which is yet again redundant here). Note that we knew \(\ln n - \ln (n+1) \to 0\) because it is the same as \(\ln \frac{n}{n+1} \to 0 = \ln 1\).

We can again argue that because \(\sum \frac{n!}{n^n}\) converges, we must have \(\lim_{n \to \infty} \frac{n!}{n^n} = 0\).

We should also note that there is a Direct Comparison Test argument that this series should converge. Note that for large enough \(n\), we have

\[
0 < \frac{n!}{n^n} = \frac{1 \cdot 2 \cdot 3 \cdots (n-1) \cdot (n)}{n \cdot n \cdot n \cdots n \cdot n} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdots n-1 \cdot n}{n \cdot n \cdot n \cdots n \cdot n} < \frac{2 \cdot 1 \cdot 1 \cdots 1 \cdot 1}{n^2} = \frac{2}{n^2}.
\]
Thus \( 0 < \frac{a_n}{a_{n+1}} < \frac{1}{e} \) for large \( n \), and since \( \sum \frac{a_n}{a_{n+1}} \) converges, so does \( \sum a_{n+1} \).

This is a fairly common type of argument, showing that part of the fraction representing \( a_n \) is less than a certain size, or greater than some other size, with what is remaining representing a useful series for the DCT.

In fact, as mentioned previously, since the ratio and root tests are ultimately proved using a Direct Comparison Test (on \( \sum |a_n| \)), it is not surprising that there is a DCT argument which gives us the convergence result above. However it was worth considering the Ratio Test, because we are left with the knowledge that, not only is the series shrinking faster than \( \sum \frac{a_n}{a_{n+1}} \), it is in fact shrinking approximately geometrically in the infinite tail, with a ratio of \( 1/e \approx 0.367879441 \).

**Example 10.5.5** Consider the series \( \sum_{n=1}^{\infty} \left[ (-1)^{n+1} \cdot \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n-1)}{2^2 \cdot 4^2 \cdot 6^2 \cdot \ldots \cdot (2n)^2} \right] \).

Here we will use the Ratio Test. Note how we insert \( n+1 \) into the formula for \( a_n \) to find \( a_{n+1} \), but also how, when we set it up to see what cancels, we have to look at not only the last terms multiplied in our fractions, but also the terms just before the last terms.

\[
\rho = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \to \infty} \left| \frac{(-1)^{n+2} \cdot 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n+1)}{(-1)^{n+1} \cdot 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n)} \right|
\]

\[
= \lim_{n \to \infty} \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n+1)}{2^2 \cdot 4^2 \cdot 6^2 \cdot \ldots \cdot (2n+2)^2} \cdot \frac{2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2n)^2}{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n-1)}
\]

\[
= \lim_{n \to \infty} \frac{2n+1}{(2n+2)^2} = \lim_{n \to \infty} \frac{2n+1}{4n^2 + 8n + 4} = 0.
\]

Since \( \rho = 0 < 1 \), we have that the original series converges absolutely.

In the above, we had to notice that in the pattern \( 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n+1) \), the terms multiplied each differ by 2, and so the term before \( (2n+1) \) would be \( (2n-1) \), and thus this product is also \( 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n-1)(2n+1) \). Similarly for the \( 2^2 \cdot 4^2 \cdot 6^2 \cdot \ldots \cdot (2n)^2 \) term. Working backwards from the ends of patterns like these is pretty common when using the ratio test.

The next two numbered examples show how the Ratio and Root Tests are not always sufficient to determine the convergence or divergence of a series.

**Example 10.5.6** Consider the series \( \sum_{n=1}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n-1)}{2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2n)} \).

The terms in this series are somewhat similar to those in the previous example, though without alternation, which is irrelevant to the Ratio Test. So we compute \( \rho \) as before:

\[
\rho = \lim_{n \to \infty} \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n+1)}{2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2n+2)}
\]

\[
= \lim_{n \to \infty} \left[ \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n+1)}{2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2n+2)} \cdot \frac{2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2n)}{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n-1)} \right]
\]

\[
= \lim_{n \to \infty} \frac{2n+1}{2n+2} = 1.
\]
This is a case where \( \rho = 1 \), which is inconclusive, so we have to look elsewhere. In doing so, we will use a technique of pairing numerator and denominator factors, similar to the method in the remarks after Example 10.5.4, page 743. For this series we can write

\[
\frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)} = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2n \cdot 2 \cdot 4 \cdots (2n-2)}
\]

\[
= \frac{1}{2n} \cdot \frac{3 \cdot 5 \cdots 2n-1}{2 \cdot 4 \cdots 2n-2}
\]

\[
> \frac{1}{2n} > 0,
\]

and since \( \sum \frac{1}{2n} \) diverges, so does \( \sum \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)} \) by the DCT.

The reader is invited to try other pairings, but with the factors at hand in the numerator and denominator, it will be impossible to find a convergent, positive-term series with larger terms than our given one.

**Example 10.5.7** Let us attempt a ratio test for the series \( \sum \frac{1}{n} \) and \( \sum \frac{1}{n^2} \). Computing \( \rho \) for these two series in turn, we get

\[
\rho = \lim_{n \to \infty} \left| \frac{a_{n+1}/a_n}{n} \right| = \lim_{n \to \infty} \frac{n+1}{n} = \lim_{n \to \infty} \frac{n}{n+1} = 1,
\]

\[
\rho = \lim_{n \to \infty} \left| \frac{a_{n+1}/a_n}{n} \right| = \lim_{n \to \infty} \frac{(n+1)^2}{n^2} = \lim_{n \to \infty} \frac{n^2}{(n+1)^2} = 1.
\]

Of course the first series diverges, while the second series converges, though both have \( \rho = 1 \) in common. Thus, knowing \( \rho = 1 \) is by itself inconclusive. (Note that both the divergence of \( \sum \frac{1}{n} \) and the convergence of \( \sum \frac{1}{n^2} \) were ultimately proved by the Integral Test, as were the convergence or divergence of all \( p \)-series.)

The Ratio and Root Tests will be especially important in the next chapter, where we will need to know where series of nonconstant terms converge. These series will be of the form \( \sum a_n (x-a)^n \), where \( x \) is variable and \( a, a_0, a_1, a_2, \cdots \) are fixed constants. Clearly such a series will converge at \( x = a \), but how far from \( a \) can \( x \) wander and still have the series converge? (One might ask, what is the domain of a function given by \( f(x) = \sum a_n (x-a)^n \)?) It depends upon the terms \( a_0, a_1, a_2, \cdots \). Because the \( x^n \)-factor in each term is geometric, it is natural to use tests which probe for comparisons to geometric series. The fact that most of the functions studied in calculus can be written in such a manner attests to the importance of such series.

It is noteworthy that we have not proven either the RAT or the ROOT. A proof of either will require careful reading, compared to earlier proofs, but we will at least give an outline of a proof of the Ratio Test (RAT).

Suppose we have a series \( \sum a_n \) so that \( \rho < 1 \). We need to show that this implies \( \sum |a_n| \) converges, i.e., \( \sum a_n \) converges absolutely. Now \( |a_{n+1}/a_n| \to \rho \), so for large enough \( N \), we have \( n \geq N \) implies \( \frac{1}{2} \leq |a_{n+1}/a_n| < \frac{1}{2} \rho \), i.e., these ratios are close enough to \( \rho \) to be greater than or equal to \( \frac{1}{2} \rho \) but no larger than the number half-way between \( \rho \) and 1. If we define \( S_N \) to be the \( N \)th partial sum of \( \sum |a_n| \) (instead of the original series), then for \( N \geq N \) we have the
Inequality

\[ S_N = \sum_{n=1}^{N} |a_n| \]

\[ = \sum_{n=1}^{N} |a_n| + \sum_{n=N+1}^{N} |a_n| \]

\[ < \sum_{n=1}^{N} |a_n| + \sum_{n=N+1}^{N} \left[ a_N \left( \frac{\rho + 1}{2} \right)^{n-N} \right] \]

\[ < \sum_{n=1}^{N} |a_n| + \sum_{n=N+1}^{N} \left[ a_N \left( \frac{\rho + 1}{2} \right)^{n-N} \right] \]

We used that, for such \( n \) we have

\[ |a_{N+1}| < \frac{\rho + 1}{2} |a_N|, \]

\[ |a_{N+2}| < \frac{\rho + 1}{2} |a_{N+1}| < \left( \frac{\rho + 1}{2} \right)^2 |a_N|, \]

In general we get, for \( n > N \), that

\[ |a_n| < |a_N| \left( \frac{\rho + 1}{2} \right)^{N-n}, \]

as used in the inequality for \( S_N \) above. Since the final series written in that inequality is a convergent, geometric series \((r = (\rho + 1)/2 \in (-1, 1))\), we have \( \{S_N\} \) is a bounded, obviously increasing sequence, and therefore converges.

One could instead note a DCT argument with the second series being \( \sum_{n=1}^{\infty} |a_N| \left( \frac{\rho + 1}{2} \right)^{N+n} \), which will have terms larger than those of the original series eventually, and still has ratio \( \frac{\rho + 1}{2} \), which implies convergence.

A similar argument can be made, somewhat modified, to show that \( \rho > 1 \Rightarrow |a_n| \neq 0 \). With \( \rho = 1 \) there is no “wiggle room” to produce the inequalities we need. Also, the argument is a bit trickier to show the same are true for the Root Test (ROOT).
Chapter 11

Taylor Series

In Chapter 10 we explored series of constant terms \( \sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \cdots \). In this chapter we next analyze series with variable terms, i.e., terms which are functions of a variable such as \( x \). As we will see, perhaps the most naturally arising variable series are the **power series**:

**Definition 11.0.1** A **power series centered at** \( x = a \) is a series of the form

\[
P(x) = \sum_{n=0}^{\infty} a_n (x - a)^n = a_0 + a_1(x - a) + a_2(x - a)^2 + a_3(x - a)^3 + \cdots ,
\]

(11.1)

where \( a, a_0, a_1, a_2, \text{ etc.}, \) are constants.

Many familiar functions are in fact equal to infinite series of the form (11.1), at least where the series converge, and so such functions can be approximated to varying degrees by the partial sums of these series. The existence of these **polynomial** partial sums explains, for instance, how calculators and similar devices compute approximate values of these functions for a wide range of inputs. (After all, computing a polynomial’s value at a given input requires only a finite number of multiplications and additions, which could even be accomplished with just paper and a pencil.) It also allows for reasonable approximations in applications where the exact equations would be too difficult to solve with the actual functions, but may be simpler with approximations. Moreover, and perhaps surprisingly, there are numerous settings where it is actually easier to deal with such a function as represented by an infinite series than by its usual representation.

We will also see that there are functions which arise in applications and are easily given in the form (11.1), but which are not equal to anything from our usual catalog of known functions. Thus by including for consideration all functions which can be expressed by power series, whether or not they have other conventional representations, we greatly expand our function catalog.

The following are general questions which arise in studying series, along with a preview of the answers.

1. **If we are given a function \( f(x) \), how do we produce a power series (11.1) which also represents the function \( f(x) \)?**

   —In answering this, we will first look at **Taylor Polynomials**,\(^1\) which coincide with the partial sums of power series expansions when a function possesses such an expansion.

---

\(^1\)Named for English mathematician Brook Taylor (1685–1731). As is often the case with early calculus discoveries, there is some controversy over whom to give credit, since hints of the results were often present earlier, and better statements usually arise later. Nonetheless, apparently after a paper by in 1786 by Swiss mathematician Simon Antoine Jean Lhuilier (1750–1840) referred to “Taylor series,” the series and polynomials bear Taylor’s name. Lhuilier was also responsible for the “\( \lim \)” notation in limits, as well as left- and right-hand limits, and many other important aspects of our modern notation.
(2) Can we always do that?
—No, and we will eventually look at cases where functions have pathologies which do not allow us to represent them with power series, for example functions which must be defined piece-wise or have discontinuities where we wish to approximate them. However, the functions which do allow for power series representations is vast—too vast to ignore—and it takes some effort to write a function which does not, at least on small intervals, allow for series representations.

(3) How accurate are the partial sums of the power series as approximations of the function \( f(x) \) it represents?
—This will be approached intuitively, visually and by way of a “Remainder Theorem,” which can give bounds on the error, or “remainder,” when we use a partial sum to approximate the actual function. (This is Theorem 11.2.1, page 769.)

(4) Given a power series (11.1), for which values of \( x \) does it converge?
—For this we will rely mostly, but not exclusively, upon a slightly clever application of the Ratio Test. This is perhaps to be expected since power series have a strong resemblance to geometric series.\(^2\)

(5) Besides approximating given functions through their partial sums, what other computational uses do power series possess?
—This will be explored in some detail later in the chapter. In short, power series give a new context in which to explore relationships among functions, with some interesting derivative and integration applications, as well as a few “real-world” applications, particularly from physics.

11.1 Taylor Polynomials: Examples and Derivation

Taylor Polynomials are a very important theoretical and practical concept in calculus and higher mathematics. As such, the general form given below should be committed to memory, as often happens naturally as it is revisited repeatedly through examples and exercises. While we will eventually derive these polynomials from reasonable first principles at the end of this section, for now we simply define them.

**Definition 11.1.1** The \( N \)th order Taylor Polynomial for the function \( f(x) \) centered at the point \( a \), where \( f(a), f'(a), \ldots, f^{(N)}(a) \) all exist, is given by\(^3\)

\[
P_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)(x-a)^n}{n!} = f(a) + f'(a)(x-a) + \frac{f''(a)(x-a)^2}{2!} + \frac{f'''(a)(x-a)^3}{3!} + \cdots + \frac{f^{(N)}(a)(x-a)^N}{N!}.
\]  

\(^2\)This is meant in the sense that, if \( a_0, a_1, \text{etc.} \), in (11.1) were all the same number, the series would be geometric, with ratio \( r = (x-a) \).

\(^3\)We normally do not bother to write the factors \( \frac{1}{0!} \) and \( \frac{1}{1!} \) in the first two terms, since \( 0!, 1! = 1 \). We also use the convention that \( f^{(0)} = f, f^{(1)} = f', f^{(2)} = f'', \text{etc.} \)
11.1. TAYLOR POLYNOMIALS: EXAMPLES AND DERIVATION

The zeroth, first, second and third order Taylor Polynomials for a function \( f(x) \) and centered at \( x = a \) would be the following:

\[
\begin{align*}
P_0(x) &= f(a), \\
P_1(x) &= f(a) + f'(a)(x - a), \\
P_2(x) &= f(a) + f'(a)(x - a) + \frac{f''(a)(x - a)^2}{2!}, \\
P_3(x) &= f(a) + f'(a)(x - a) + \frac{f''(a)(x - a)^2}{2!} + \frac{f'''(a)(x - a)^3}{3!}.
\end{align*}
\]

A few notes are appropriate here.

1. The \( N \)th-order Taylor Polynomial with center \( x = a \) is the sum of the \((N - 1)\)st-order Taylor Polynomial with the same center, and the term \( \frac{1}{N!}f^{(N)}(a)(x - a)^N \), so we just add a single “term” to a Taylor Polynomial to arrive at the next-order Taylor Polynomial.

2. \( P_1(x) \) is the same as the linear approximation of \( f(x) \) centered at \( x = a \), so it is often called “the first-order approximation of \( f(x) \) at (or near) \( x = a \).” \( P_2(x) \) is then called the quadratic, or second-order approximation, \( P_3(x) \) the cubic, or third-order approximation, and so on.

Example 11.1.1 Find \( P_0(x), P_2(x), \cdots, P_5(x) \) at \( x = 0 \) for the function \( f(x) = e^x \).

**Solution:** First note that if we construct \( P_5(x) \), the first term will be \( P_0(x) \), the first two will comprise \( P_1(x) \), the first three terms will give \( P_2(x) \), and so on.

Anytime we need to construct a Taylor Polynomial of a function \( f(x) \), we first construct the chart of the function and its relevant derivatives at the center. For this example, we construct the following chart with \( a = 0 \).

\[
\begin{align*}
f(x) &= e^x \implies f(0) = 1 \\
f'(x) &= e^x \implies f'(0) = 1 \\
f''(x) &= e^x \implies f''(0) = 1 \\
f'''(x) &= e^x \implies f'''(0) = 1 \\
f^{(4)}(x) &= e^x \implies f^{(4)}(0) = 1 \\
f^{(5)}(x) &= e^x \implies f^{(5)}(0) = 1
\end{align*}
\]

Now, according to our definition (11.2),

\[
P_5(x) = \underbrace{f(0) + f'(0)(x - 0)}_{P_0(x)} + \underbrace{\frac{f''(0)(x - 0)^2}{2!}}_{P_1(x)} + \underbrace{\frac{f'''(0)(x - 0)^3}{3!}}_{P_2(x), \text{ etc.}} + \underbrace{\frac{f^{(4)}(0)(x - 0)^4}{4!}}_{P_3(x)} + \underbrace{\frac{f^{(5)}(0)(x - 0)^5}{5!}}_{P_4(x)}
\]

\[
= 1 + 1x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \frac{1}{5!}x^5.
\]
From the computation above we can also write

\[ P_0(x) = 1, \]
\[ P_1(x) = 1 + x, \]
\[ P_2(x) = 1 + x + \frac{x^2}{2!}, \]
\[ P_3(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}, \]
\[ P_4(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!}, \]
\[ P_5(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!}. \]

The polynomials \( P_0(x), \ldots, P_5(x) \) are graphed in Figure 11.1, page 751. A couple of observations from that figure are in order:

- Clearly, as we add more terms to get higher-order Taylor Polynomials, the curves tend to more closely follow the behavior of the function, at least near the center \( (x = 0 \text{ in the above example). This will be explained as we proceed.} \)

- In all cases, the highest-order nonzero term eventually dominates the polynomial’s behavior for large \( |x| \). For instance, for large \( |x| \), \( P_5(x) \) more clearly behaves like the degree-5 polynomial it is, and thus very differently from the original function \( f(x) = e^x \):

  1. \( |P_5(x)| \to \infty \) as \( x \to \pm \infty \), and in particular \( x \to -\infty \implies P_5(x) \to -\infty \), though \( e^x \to 0^+ \) as \( x \to -\infty \).

  2. As \( x \to \infty \), \( e^x - P_5(x) \to \infty \), i.e., the exponential will grow much faster than will any polynomial, including a degree-5 polynomial such as this particular \( P_5(x) \).

It is natural to ask why the Taylor Polynomials \( P_N(x) \) seem to give us better and better approximations of the function \( f(x) \) as we increase \( N \). The following observation gives some hint:

**Theorem 11.1.1** If \( f(x) \) is \( N \)-times differentiable at \( x = a \), then \( P_N(x) \), as defined by (11.2), satisfies:

\[
\begin{align*}
P_N(a) &= f(a) \\
P_N'(a) &= f'(a) \\
P_N''(a) &= f''(a) \\
&\vdots \\
P_N^{(N)}(a) &= f^{(N)}(a) \quad (\text{i.e., } P_N^{(N)}(x) \text{ is constant}) \\
P_N^{(m)}(x) &= 0 \quad \text{for all } m \in \{N + 1, N + 2, N + 3, \ldots \}
\end{align*}
\]
Figure 11.1: Graphs of $y = e^x$ and the Taylor Polynomial approximations $P_0(x) - P_5(x)$, plotted with dashed lines.
The upshot of this is that $P_N$ is the simplest polynomial such that $f, f', f'', \ldots, f^{(N)}$ respectively match $P_N, P_N', P_N'', \ldots, P_N^{(N)}$ at the center $x = a$:

- $P_0$ has the same height as $f$ at $x = a$;
- $P_1$ has the same height and slope as $f$ at $x = a$;
- $P_2$ has the same height, slope and second derivative as $f$ at $x = a$;

and so on. While it becomes difficult to visualize how matching higher derivatives with $f$ will continue the trend of better approximation, it should have the ring of truth. For instance, we can claim that the polynomial $P_3(x)$ matches the function $f(x)$ in height, slope, second derivative (concavity?), and the (instantaneous) rate of change in the second derivative at $x = a$. To go to the fourth-order approximation we note that how fast $f''$ is changing at the center $x = a$ will be “picked up” by $f^{(4)}$, at least in the instantaneous sense, and thus by $P_4(x)$ since it shares the height and first four derivatives with $f(x)$ at $x = a$. This type of reasoning will be addressed again in our error estimates for our approximations $P_N(x) \approx f(x)$, that is, estimates for the size of the errors $f(x) - P_N(x)$ in these approximations. It will also be addressed at the end of this section in our derivation of the Taylor Polynomials from some first principles.

So in our above Example 11.1.1, $P_3(x)$ matches the height and first five derivatives of $e^x$ at $x = 0$, which helps it to “fit” the curve of $y = e^x$ (i.e., approximate the behavior of $f(x)$) better than the lower-order approximations which do not match as many derivatives of $f(x) = e^x$ as does $P_3(x)$. Indeed, $P_3(x)$ is the simplest polynomial which matches the height, slope, “concavity,” third derivative, fourth derivative and fifth derivative of $e^x$ at $x = 0$. Higher-order Taylor Polynomials $P_5(x), P_7(x)$ and so on will match all that, and more.

A pattern clearly emerges for $P_N(x)$, centered at $a = 0$ for $f(x) = e^x$. If we desired $P_6(x)$, we would simply add $\frac{x^6}{6!}$, and if we desired $P_7(x)$ we would then further add $\frac{x^7}{7!}$, and so on. It would be a simple exercise to generate $P_{20}(x)$ or higher, and to compute its values using any rudimentary programming language.\footnote{The graphs here and throughout the book are generated with the Postscript language, which is more of a publishing language and far from being a first choice for intense, scientific computations, but is quite adequate here. One technique for making the computations more computer-friendly, and pencil and paper-friendly, is to rewrite the polynomial}

\[
P_5(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} = 1 + x \left( 1 + \frac{1}{2} x \left( 1 + \frac{1}{3} x \left( 1 + \frac{1}{4} x \left( 1 + \frac{1}{5} x \right) \right) \right) \right).
\]

With the second form, there are fewer multiplications (if we consider, say, $x^5$ and $5!$ as each comprising four multiplications), and we do not have to rely on the computer to compute powers of large numbers, divided by large factorials, and sum these. It is akin to the process known as synthetic division for computing polynomial values.
11.1. TAYLOR POLYNOMIALS: EXAMPLES AND DERIVATION

Proof: First we note how derivative the of a general nth-order term in our polynomial (11.2) simplifies, assuming \( n \geq 1 \):

\[
\frac{d}{dx} \left[ \frac{f^{(n)}(a)(x - a)^n}{n!} \right] = \frac{f^{(n)}(a)}{n!} \cdot n(x - a)^{n-1} = \frac{f^{(n)}(a)}{(n-1)!} \cdot n(x - a)^{n-1} = \frac{f^{(n)}(a)}{(n-1)!} \cdot (x - a)^{n-1}.
\]

We made use of the fact that \( a \), \( f^{(n)}(a) \) and \( n! \) are all constants in the computation above. We will also use the fact that any additive constants, i.e., terms of form \( "(x - a)^{n-1}" \) will have derivative zero. Finally note that any term with \( (x - a)^n \), where \( n \geq 1 \), will be zero at \( x = a \).

From these observations it is routine (if not totally transparent) that we can demonstrate the computations in Theorem 11.1.1. To make the pattern clear, we assume here that \( N > 3 \). In each of what follows, we first take derivatives at each line, and then evaluate at \( x = a \).

\[
P_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!} (x - a)^n \quad \Rightarrow \quad P_N(a) = \frac{f^{(0)}(a)}{0!} = f(a)
\]

\[
P'_N(x) = \sum_{n=1}^{N} \frac{f^{(n)}(a)}{(n-1)!} (x - a)^{n-1} \quad \Rightarrow \quad P'_N(a) = \frac{f^{(1)}(a)}{0!} = f'(a)
\]

\[
P''_N(x) = \sum_{n=2}^{N} \frac{f^{(n)}(a)}{(n-2)!} (x - a)^{n-2} \quad \Rightarrow \quad P''_N(a) = \frac{f^{(2)}(a)}{0!} = f''(a)
\]

\[
P'''_N(x) = \sum_{n=3}^{N} \frac{f^{(n)}(a)}{(n-3)!} (x - a)^{n-3} \quad \Rightarrow \quad P'''_N(a) = \frac{f^{(3)}(a)}{0!} = f'''(a)
\]

\[\vdots\]

\[
P^{(N-1)}_N(x) = \sum_{n=N-1}^{N} \frac{f^{(n)}(a)}{(n-(N-1))!} (x - a)^{n-(N-1)} \quad \Rightarrow \quad P^{(N-1)}_N(a) = f^{(N-1)}(a)
\]

\[
P^{(N)}_N(x) = f^{(N)}(a) \quad \Rightarrow \quad P^{(N)}_N(a) = f^{(N)}(a)
\]

\[
P^{(m)}_N(x) = 0, \quad m \in \{N+1, N+2, N+3, \ldots \}, \quad \text{q.e.d.}
\]

Example 11.1.2 There is a simple real-world motivation for this kind of approach. Suppose a passenger on a train wishes to know approximately where the train is. At some time \( t_0 \), he passes the engineer’s compartment and sees the mile marker \( s_0 \) out the front window. He also sees the speedometer reading \( v_0 \). If the train is not accelerating or decelerating noticeably, he can follow his watch and expect the train to move approximately \( v_0(t - t_0) \) in the time \( [t_0, t] \). In other words,

\[
s \approx s_0 + v_0(t - t_0).
\]

On the other hand, perhaps he feels some acceleration, as the train leaves an urban area, for instance. If the engineer has an acceleration indicator, and it reads \( a_0 \) at time \( t_0 \), then the
Again we construct a chart.

If our passenger can even compute how \( a = s'' \) is changing, then assuming that change is at a constant rate, i.e., that \( s'''(t) \approx s'''(t_0) \), we can go another order higher and claim:

\[
s \approx s_0 + v_0(t - t_0) + \frac{1}{2}a_0(t - t_0)^2 + \frac{1}{3} s'''(t_0)(t - t_0)^3.
\]

Indeed this will likely be the best estimate thus far when \( |t - t_0| \) is small (and \( s''' \) is still relatively constant). However, we have to be aware that this latest approximation is a degree-three polynomial, and will therefore act like one as \( |t| \) (and therefore \( |t - t_0| \)) gets large, so we have to always be aware of the range of \( t \) for which the approximation is accurate.

Next we look at some more examples.

**Example 11.1.3** Find \( P_2(x) \) for \( f(x) = \sqrt{1 + x^2} \) centered at \( a = 0 \).

**Solution:** We compute the first two derivatives, and evaluate them at 0:

\[
f(x) = \sqrt{1 + x^2} \quad \implies f(0) = 1
\]

\[
f'(x) = \frac{1}{2\sqrt{1 + x^2}} \cdot 2x = x(1 + x^2)^{-1/2} \quad \implies f'(0) = 0
\]

\[
f''(x) = x \cdot \frac{-1}{2} (1 + x^2)^{-3/2} + (1 + x^2)^{-1/2} \cdot 1 \implies f''(0) = 1.
\]

Thus

\[
P_2(x) = f(0) + f'(0)(x - 0) + \frac{1}{2} f''(0)(x - 0)^2
\]

\[= 1 + \frac{1}{2} x^2.
\]

See Figure 11.2, page 755 for the graphs of \( f(x) \) and \( P_2(x) \).

In most applications, one chooses a center \( x = a \) so that \( a, f(a), f'(a), f''(a) \) and so on are all “nice” numbers, though theoretically we could have found \( P_2(x) \) in Example 11.1.3 with \( a = \sqrt{3} \). On the other hand, if we can easily enough compute \( \sqrt{3} \) (for our \((x-a)^n \) terms), we probably could equally easily compute \( \sqrt{x^2 + 1} \).

In Section 11.5.3 we will see a pattern which will help us compute higher-order Taylor Series for functions such as this. Clearly the derivative computations needed to find \( f'''(0), f^{(4)}(0) \) and so on quickly become unwieldy, and so a shortcut will be welcome. For many physics-type problems, however, \( P_2(x) \) is a very useful approximation, particularly for \( x \in [-1, 1] \).

**Example 11.1.4** Find \( P_3(x) \) at \( a = 1 \) if \( f(x) = 2x^3 - 9x^2 + 5x + 11 \).

**Solution:** Again we construct a chart.

\[
f(x) = 2x^3 - 9x^2 + 5x + 11 \implies f(1) = 9
\]

\[
f'(x) = 6x^2 - 18x + 5 \implies f'(1) = -7
\]

\[
f''(x) = 12x - 18 \implies f''(1) = -6
\]

\[
f'''(x) = 12 \implies f'''(1) = 12
\]

\footnote{Notice that if \( f''' \) were truly constant, then (11.4) would be exact and not an approximation. Similarly, if \( f''' \) were truly constant, then (11.5) would be exact.}
Now
\[ P_4(x) = f(1) + f'(1)(x - 1) + \frac{f''(1)(x - 1)^2}{2!} + \frac{f'''(1)(x - 1)^3}{3!} + \frac{f^{(4)}(1)(x - 1)^4}{4!} \]
\[ = 9 - 7(x - 1) - \frac{6(x - 1)^2}{2!} + \frac{12(x - 1)^3}{3!} + \frac{12(x - 1)^4}{4!} \]
\[ = 9 - 7(x - 1) - 3(x - 1)^2 + 2(x - 1)^3. \]

This is a trivial, yet important kind of example, for if we expanded out the last line above in powers of \( x \) we would get back the original polynomial, which shows that the simplest polynomial matching this function and its first three derivatives at \( x = 1 \) is the polynomial itself. Furthermore, we can see from our chart, that \( f^{(4)}(x) = 0, f^{(5)}(x) = 0, \ldots \), and so \( P_3 = P_4 = P_5 = \cdots \).

We will enshrine this result in the following theorem:

**Theorem 11.1.2** Suppose \( f(x) \) is an \( N \)th-degree polynomial, i.e.,
\[ f(x) = A_Nx^N + A_{N-1}x^{N-1} + \cdots + A_1x + A_0. \] (11.6)

Then regardless of \( a \in \mathbb{R} \), we have \( (\forall m \geq N) [P_m(x) = f(x)] \).

In other words, a polynomial will be the same as its Taylor Polynomials of all orders which are at least as high as the degree of the polynomial, regardless of the center of the Taylor Polynomial.

The proof is interesting to read through, though the result is more important than the proof. We include the proof here for completeness.

**Proof:** We will prove this in stages.

1. An important general observation we will use repeatedly is the following:
\[ (\forall x \in \mathbb{R})[g'(x) = h'(x)] \iff (\exists C)[g(x) - h(x) = C]. \] (11.7)

In other words, if two functions have the same derivative functions, then the original two functions differ only by a constant. (This is also true if the functions and derivatives are only considered on single intervals.)
(2) Since \( f \) and \( P_N \) are both \( N \)th-degree polynomials, we have \( f^{(N)}(x) \) and \( P^{(N)}(x) \) are constants.

(3) By Theorem 11.1.1, page 750, we have \( f^{(N)}(a) = P^{(N)}(a) \).

(4) From (2) and (3), we have

\[
P^{(N)}(x) = P^{(N)}(a) = f^{(N)}(a) = f^{(N)}(x). \tag{11.8}
\]

Thus \( P^{(N)}(x) = f^{(N)}(x) \).

(5) By (1), we can thus conclude that \( P^{(N-1)}(x) \) and \( f^{(N-1)}(x) \) differ by a constant.

(6) Since \( P^{(N-1)}(a) = f^{(N-1)}(a) \), and (5), we must have \( P^{(N-1)}(x) = f^{(N-1)}(x) \).

In other words, since \( P^{(N-1)}(x) \) and \( f^{(N-1)}(x) \) differ by a constant, and since \( P^{(N-1)}(a) - f^{(N-1)}(a) = 0 \), the constant referred to in (5) must be zero.

(7) The argument above can be repeated to get \( P^{(N-2)}(x) = f^{(N-2)}(x) \), and so on, until finally we indeed get \( P'(x) = f'(x) \).

(8) The last step is the same. From (1), \( P \) and \( f \) differ by a constant, but since \( P(a) = f(a) \), that constant must be zero, so \( P(x) - f(x) = 0 \), i.e., \( P(x) = f(x) \).

It is important that the original function \( f(x) \) above was a polynomial, or else the conclusion is false.

The theorem is useful for both analytical and algebraic reasons. If we wish to expand an \( N \)th-degree polynomial (11.6) in powers of \( x - a \) (instead of the usual \( x = x - 0 \)), then we can just compute \( P_N(x) \) centered at \( x = a \). From the theorem, we can easily “re-center” any polynomial, meaning we can write it as a sum of powers of \( (x - a) \) instead of \( x \), the original “center” of course being zero.

**Example 11.1.5** Write the following polynomial in powers of \( x \): \( f(x) = (x + 5)^4 \).

**Solution:** We can use the binomial expansion (with Pascal’s Triangle, for instance) for this, but we can also use the Taylor Polynomial centered at \( a = 0 \):

\[
f(x) = (x + 5)^4 \quad \implies \quad f(0) = 625
\]

\[
f'(x) = 4(x + 5)^3 \quad \implies \quad f'(0) = 4 \cdot 5^3
\]

\[
f''(x) = 4 \cdot 3(x + 5)^2 \quad \implies \quad f''(0) = 4 \cdot 3 \cdot 5^2
\]

\[
f'''(x) = 4 \cdot 3 \cdot 2(x + 5) \quad \implies \quad f'''(0) = 4 \cdot 3 \cdot 2 \cdot 5
\]

\[
f^{(4)}(x) = 4 \cdot 3 \cdot 2 \cdot 1 \quad \implies \quad f^{(4)}(0) = 4 \cdot 3 \cdot 2 \cdot 1
\]

\[
f^{(m)}(x) = 0 \quad \text{any } m > 4
\]

\[
P_4(x) = f(0) + f'(0)x + \frac{f''(0)x^2}{2!} + \frac{f'''(0)x^3}{3!} + \frac{f^{(4)}(0)x^4}{4!}
\]

\[
= 5^4 + 4 \cdot 5^3 x + \frac{4 \cdot 3 \cdot 5^2 x^2}{2!} + \frac{4 \cdot 3 \cdot 2 \cdot 5 x^3}{3!} + \frac{4 \cdot 3 \cdot 2 \cdot 1 x^4}{4!}
\]

\[
= 625 + 500x + 150x^2 + 20x^3 + x^4.
\]

Because this is \( P_4(x) \) for a fourth-degree polynomial function, it equals that polynomial function, i.e.,

\[
(x + 5)^4 = 625 + 500x + 150x^2 + 20x^3 + x^4.
\]
11.1. TAYLOR POLYNOMIALS: EXAMPLES AND DERIVATION

Of course arguably the more interesting Taylor Polynomials do not involve polynomial approximations of polynomials. The relationship to ordinary polynomials explored above is nonetheless interesting. For the remainder here, we will look at examples where \( f(x) \) is not itself a polynomial.

**Example 11.1.6** Consider the function \( f(x) = \sqrt[3]{x} \), with \( a = 27 \).

a. Calculate \( P_1(x) \), \( P_2(x) \), \( P_3(x) \).

b. Use these to approximate \( \sqrt[3]{26} \).

c. Compare these to the actual value of \( \sqrt[3]{26} \), as determined by calculator.

**Solution:** We take these in turn.

a. First we will construct a chart.

\[
\begin{align*}
 f(x) &= x^{1/3} & f(27) &= 3 \\
 f'(x) &= \frac{1}{3}x^{-2/3} & f'(27) &= \frac{1}{3} \cdot \frac{1}{9} = \frac{1}{27} \\
 f''(x) &= -\frac{2}{9}x^{-5/3} & f''(27) &= -\frac{2}{9} \cdot \frac{1}{243} = -\frac{2}{2187} \\
 f'''(x) &= \frac{10}{27}x^{-8/3} & f'''(27) &= \frac{10}{27} \cdot \frac{1}{6561} = \frac{10}{177,147},
\end{align*}
\]

Thus,

\[
\begin{align*}
P_1(x) &= 3 + \frac{1}{27}(x - 27) \\
P_2(x) &= 3 + \frac{1}{27}(x - 27) - \frac{2}{2187} (x - 27)^2 \\
&= 3 + \frac{1}{27}(x - 27) - \frac{1}{4374} (x - 27)^2 \\
P_3(x) &= 3 + \frac{1}{27}(x - 27) - \frac{1}{4374} (x - 27)^2 + \frac{\left(\frac{10}{177,147}\right)}{3!} (x - 27)^3 \\
&= 3 + \frac{1}{27}(x - 27) - \frac{1}{4374} (x - 27)^2 + \frac{10}{1,062,882} (x - 27)^3.
\end{align*}
\]

b. From these we get

\[
\begin{align*}
P_1(26) &= 3 + \frac{1}{27}(26 - 27) = 3 + \frac{1}{27}(-1) = 3 - \frac{1}{27} = \frac{80}{27} \approx 2.9629630 \\
P_2(26) &= 3 + \frac{1}{27}(-1) + \frac{1}{4374}(-1)^2 = \frac{12,961}{4374} \approx 2.9627343 \\
P_3(26) &= P_2(26) + \frac{10}{1,062,882}(-1)^3 = \frac{314,951,300}{106,2882} \approx 2.9627249.
\end{align*}
\]

c. The actual value (to 8 digits) is \( \sqrt[3]{26} \approx 2.9624961 \). The errors \( R_1(26), R_2(26) \) and \( R_3(26) \), in each of the above approximations are respectively

\[
\begin{align*}
R_1(26) &= \sqrt[3]{26} - P_1(26) \approx 2.9624961 - 2.9629630 = -0.0004669 \\
R_2(26) &= \sqrt[3]{26} - P_2(26) \approx 2.9624961 - 2.9627343 = -0.0002382 \\
R_3(26) &= \sqrt[3]{26} - P_3(26) \approx 2.9624961 - 2.9627249 = -0.0002288.
\end{align*}
\]
Thus we see some improvement in these estimates. For other functions it can be more or less dramatic. In Section 11.2 we will state the form of the error, or remainder $R_N(x) = f(x) - P_N(x)$, and thus be able to explore the accuracy of $P_N(x)$.

**Example 11.1.7** Find $P_5(x)$ at $a = 0$ for $f(x) = \sin x$.

**Solution:** Again we construct the chart.

\[
\begin{align*}
  f(x) &= \sin x \quad \implies f(0) = 0 \\
  f'(x) &= \cos x \quad \implies f'(0) = 1 \\
  f''(x) &= -\sin x \quad \implies f''(0) = 0 \\
  f'''(x) &= -\cos x \quad \implies f'''(0) = 0 \\
  f^{(4)}(x) &= \sin x \quad \implies f^{(4)}(0) = 0 \\
  f^{(5)}(x) &= \cos x \quad \implies f^{(5)}(0) = 1,
\end{align*}
\]

from which we get

\[
P_5(x) = 0 + 1x + 0x^2 + \frac{-1x^3}{2!} + 0x^4 + \frac{1x^5}{5!} = x - \frac{x^3}{3!} + \frac{x^5}{5!}.
\]

From this chart we can see an obvious pattern where

\[
P_6(x) = P_5(x) + 0 = P_5(x),
\]

\[
P_8(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + 0 = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} = P_7(x),
\]

and so on.

This answers the question of how electronic calculators compute $\sin x$: by means of just such a Taylor Polynomial. It also hints at an answer for why physicists often simplify a problem by replacing $\sin x$ with $x$: that is the simplest polynomial which matches the height, slope and concavity of $\sin x$ at $x = 0$ is a very simple function indeed, namely $P_2(x) = x$.

See Figures 11.3 and 11.4, page 759 to compare $\sin x$ to $P_1(x), P_3(x), \ldots, P_{13}(x) = P_{14}(x)$. Clearly the polynomials are increasingly better at approximating $\sin x$ as we add more terms. On the other hand, as $|x|$ gets large these approximations eventually behave like the polynomials they are in the sense that $|P_n(x)| \to \infty$ as $|x| \to \infty$. This is not alarming, since it is the local behavior, in this case near $x = 0$ (more generally near $x = a$), that we exploit when we use polynomials to approximate functions. It is worth remembering, however, so that we do not attempt to use a Taylor Polynomial to approximate a function too far from the center, $x = a$, of the Taylor Polynomial.

**Example 11.1.8 (Application)** As already mentioned, physicists often take advantage of the second order approximation $\sin x \approx P_2(x) = 0 + x + 0x^2$, that is,

\[
\sin x \approx x \quad \text{for } |x| \text{ small.} \tag{11.9}
\]

\[\text{Note that when using Taylor Polynomials to compute a trigonometric function such as } \sin x, \text{ the calculus is greatly simplified when we assume } x \text{ is in radians (which are dimensionless). Therefore a calculator giving its approximation of, say, } \sin 57^\circ \text{ will convert the angle into radians first.}\]
11.1. TAYLOR POLYNOMIALS: EXAMPLES AND DERIVATION

Figure 11.3: \( \sin x, P_1(x), P_2(x) = x \) (gray), \( P_3(x), P_4(x) = x - \frac{x^3}{3!} \) (dots), and \( P_5(x), P_6(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} \) (dashed).

Figure 11.4: \( \sin x \) and \( P_1(x) = P_2(x), P_3(x) = P_4(x), \cdots, P_{13}(x) = P_{14}(x) \).
The classic example is the modeling of the simple pendulum. See the illustration below, which we then model mathematically.

Suppose a pendulum of mass \( m \) is hanging from an always taut and straight string of negligible weight. Let \( \theta \) be the angle the string makes with the downward vertical direction. We will take \( \theta > 0 \) if \( \theta \) represents a counterclockwise rotation, as is standard. Also \( g \) is the acceleration due to gravity, approximately 32 ft/sec\(^2\) or 9.8 m/sec\(^2\).

The component of velocity which is in the direction of motion of the pendulum is given by

\[
\frac{ds}{dt} = l \frac{d(\theta)}{dt} = l \frac{d^2\theta}{dt^2},
\]

and the acceleration by its derivative,

\[
\frac{d}{dt} \frac{ds}{dt} = l \frac{d^2\theta}{dt^2} = l \frac{d^2\theta}{dt^2}.
\]

Now the force in the direction of the motion has magnitude \( mg \sin \theta \), but is a restorative force, and is thus in the opposite direction of the angular displacement. It is not too difficult to see that this force is given by

\[
-F = -mg \sin \theta,
\]

for \( \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \). Thus, by equating the force and the acceleration in the angular direction, we get

\[
ml \frac{d^2\theta}{dt^2} = -mg \sin \theta \tag{11.10}
\]

which simplifies to

\[
\frac{d^2\theta}{dt^2} = -\frac{g}{l} \sin \theta. \tag{11.11}
\]

This is a relatively difficult differential equation\(^8\) to solve. However, if we assume \(|\theta|\) is small, we can use \( \sin \theta \approx \theta \) and instead solve the following equation which holds approximately true\(^9\):

\[
\frac{d^2\theta}{dt^2} = -\frac{g}{l} \theta \tag{11.12}
\]

\(^7\)For those familiar with moments of inertia, the analog of \( F = ma \) is

\[N = I \alpha,\]

where \( N \) is torque, \( I \) is the moment of inertia, and \( \alpha \) is the angular acceleration, in rad/sec\(^2\). Using the fact that, for this example, torque is also defined by \( N = F_{\text{tan}} l = -mg \sin \theta \), we get the equations

\[N = -mg \sin \theta = ml^2 \frac{d^2\theta}{dt^2},\]

giving equation (11.10) after dividing by \( l \).

\(^8\)A differential equation is an equation involving the derivatives of a function \( y \) (or \( \theta \) here). The goal in “solving” a differential equation is to find all functions \( y \) which satisfy the equation. Courses in differential equations assume the student has learned calculus for two or three semesters, though it is common for simple differential equations to be found in introductory calculus books.

\(^9\)We should point out here that (11.12) is an example of a simple harmonic oscillator, which is any physical system governed by an equation of the form

\[Q''(t) = -\kappa Q(t), \quad \kappa > 0\]

(kappa, the lower-case Greek letter kappa being a constant) which has solution

\[Q(t) = A \sin \sqrt{\kappa} t + B \cos \sqrt{\kappa} t,\]

and period \( 2\pi/\sqrt{\kappa} \). Examples include springs which are governed by Hooke’s Law \( F(s) = -ks \), where \( k > 0 \) and \( s = s(t) \). Recall \( F = m \frac{ds}{dt^2} \), so Hooke’s Law becomes \( \frac{d^2s}{dt^2} = -\frac{k}{m} s \), giving a simple harmonic oscillator.
The solution of (11.12) is
\[ \theta = A \sin \left( \sqrt{\frac{g}{l}} \cdot t \right) + B \cos \left( \sqrt{\frac{g}{l}} \cdot t \right). \] (11.13)

Here \( A \) and \( B \) are arbitrary constants depending on the initial \((t = 0)\) position and velocity of the pendulum. Notice that (11.13) is periodic, with a period \( \tau \) where \( \tau = 2\pi/\sqrt{g/l} \), i.e.,
\[ \tau = 2\pi \sqrt{\frac{l}{g}}. \] (11.14)

That is the formula found in most physics texts for the period of a pendulum. However, it is based upon an approximation, albeit quite a good one for \(|\theta|\) small. Still, the higher we allow the pendulum to swing, the less we can rely on this approximation of the period.

To a novice, it might not be terribly satisfying to resort to approximations when attempting to solve a problem, but “in the lab” and when designing practical applications, understanding how to approximate, and the limitations of the practice, are quite valuable, and usually better appreciated with more exposure to the possibilities.

**Example 11.1.9** Let us find \( P_6(x) \) where \( f(x) = \cos x \) and \( a = 0 \).

**Solution:** We construct the table again:
\[
\begin{align*}
  f(x) = \cos x & \implies f(0) = 1 \\
  f'(x) = -\sin x & \implies f'(0) = 0 \\
  f''(x) = -\cos x & \implies f''(0) = -1 \\
  f'''(x) = \sin x & \implies f'''(0) = 0 \\
  f^{(4)}(x) = \cos x & \implies f^{(4)}(0) = 1 \\
  f^{(5)}(x) = -\sin x & \implies f^{(5)}(0) = 0 \\
  f^{(6)}(x) = -\cos x & \implies f^{(6)}(0) = -1
\end{align*}
\]

Since the odd derivatives are zero at \( x = 0 \), only the even-order terms appear, and we have
\[
P_6(x) = 1 + \frac{-1(x - 0)^2}{2!} + \frac{1(x - 0)^4}{4!} + \frac{-1(x - 0)^6}{6!}
= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!}.
\]

From this a pattern clearly emerges, and we could easily calculate
\[
P_{14}(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \frac{x^{10}}{10!} + \frac{x^{12}}{12!} - \frac{x^{14}}{14!}.
\]

We might also point out that \( P_{15} \) would be the same, since the odd terms were all zero.
Example 11.1.10 Find $P_5$ for $f(x) = \ln x$ with center $a = 1$.

**Solution:** First, the table is constructed as usual by computing first $f^{(n)}(x)$ and then $f^{(n)}(a) = f^{(n)}(1)$.

\[
\begin{align*}
  f(x) &= \ln x \quad \Rightarrow \quad f(1) = 0 \\
  f'(x) &= x^{-1} \quad \Rightarrow \quad f'(1) = 1 \\
  f''(x) &= -1x^{-2} \quad \Rightarrow \quad f''(1) = -1 \\
  f'''(x) &= 2x^{-3} \quad \Rightarrow \quad f'''(1) = 2 \\
  f^{(4)}(x) &= -3 \cdot 2x^{-4} \quad \Rightarrow \quad f^{(4)}(1) = -3 \cdot 2 \\
  f^{(5)}(x) &= 4 \cdot 3 \cdot 2x^{-5} \quad \Rightarrow \quad f^{(5)}(1) = 4 \cdot 3 \cdot 2
\end{align*}
\]

Now we construct $P_5$:

\[
P_5(x) = 0 + 1(x - 1) + \frac{-1(x - 1)^2}{2!} + \frac{2(x - 1)^3}{3!} + \frac{-3 \cdot 2(x - 1)^4}{4!} + \frac{4 \cdot 3 \cdot 2(x - 1)^5}{5!}.
\]

Recalling the definition of factorials, in which $2! = 2 \cdot 1, \; 3! = 3 \cdot 2 \cdot 1, \; 4! = 4 \cdot 3 \cdot 2 \cdot 1, \; \text{and} \; 5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$, we see that the above simplifies to

\[
P_5(x) = 1(x - 1) - \frac{1}{2} (x - 1)^2 + \frac{1}{3} (x - 1)^3 - \frac{1}{4} (x - 1)^4 + \frac{1}{5} (x - 1)^5.
\]

It is not hard to see that $f^{(n)}(x) = (-1)^{n+1} (n-1)! x^{-n}$, and so $f^{(n)}(1) = (-1)^{n+1} (n-1)!$. The obvious pattern which appears in $P_5$ should continue for $P_6, P_7, \text{etc.}$ Thus we can calculate any $P_N(x)$ for this example:

\[
P_N(x) = \sum_{n=1}^{N} \frac{(-1)^{n+1} (x - 1)^n}{n}.
\]

If we wished to have $N \to \infty$, we get the full Taylor Series:

\[
\sum_{n=1}^{\infty} \frac{(-1)^{n+1} (x - 1)^n}{n}.
\]

However this might not converge for all $x$. Indeed, from the ratio test we get

\[
\rho = \lim_{n \to \infty} \left| \frac{(-1)^{n+2} (x - 1)^{n+1}}{n+1} \right| \left/ \left| \frac{(-1)^{n+1} (x - 1)^n}{n} \right| \right| = \lim_{n \to \infty} \frac{n|x - 1|}{n+1} = |x - 1|,
\]

and so $\rho < 1$ when $|x-1| < 1$, i.e., $x \in (0, 2)$, and $\rho > 1$ when $|x-1| > 1$, i.e., $x \in (-\infty, 0) \cup (2, \infty)$. When $x = 0$ we have a negative harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$ which diverges, and when $x = 2$ we have the conditionally convergent alternating series $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$. Summarizing, thus the series converges for $x \in (0, 2]$, and diverges elsewhere.

This use of the Ratio Test is the most common method for computing where such a series—namely one in which we have a formula for the $n$th-degree term—converges.

From many of the previous examples we see that the table many Taylor Polynomials have patterns which emerge easily from the derivative computations. However, we will see that this is not the case for many important series which we can nonetheless use other methods to derive the pattern. In fact those methods are often easier than attempting what we will later characterize as “brute force,” or “from scratch” method of construction here, which is deriving the $n$th term by computing $f^{(n)}(x)$ to compute $f^{(n)}(a)$ to construct $P_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)(x-a)^n}{n!}$. 

11.1. Taylor Polynomials: Examples and Derivation

Motivation for and Derivation of Taylor Polynomials’ Forms from First Principles

We end this Section 11.1 with a derivation of the Taylor Polynomials from nearly “first principles.” While arguably not crucial to a relative novice student of calculus, it is nonetheless valuable and interesting in its own right because of the insights that can be gleaned from the creative ideas it employs. That said, its placement here is mostly for completeness. It will likely be much better motivated after the subsequent sections in this chapter are mastered, so the reader should feel free to peruse casually here first, and revisit it after studying the rest of the chapter and having a better understanding of the complete context.

One development of $P_N(x)$, left for the exercises, is a derivation based upon the assumption that $P_N$ matches $f$ in all derivatives (including the “zeroth”) up to order $N$ at the center $a$, and then the coefficients $a_n$ of the $(x-a)^n$ terms are all found using methods similar to what we used to find coefficients of partial fraction decompositions.

However, the derivation here uses a different motivation, and is based upon reasonable integral approximations. We will continually refer to the special case where $a = 0$—for which $P_0(x)$ through $P_5(x)$ are graphed along with $f(x)$ in Figure 11.1, page 751—to illustrate the principles developed here.

In summary, $P_N(x)$ is the (necessarily polynomial) function we arrive at by deduction under the assumptions that

1. we only have the following data for $f$: $f(a)$, $f'(a)$, $f''(a)$, $f'''(a)$, · · · , $f^{(N)}(a)$, and
2. given no other data for $f$, we assume that its derivative $f^{(N)}(x)$ is approximately constant, for $x$ near $x = a$. That is,

$$f^{(N)}(x) \approx f^{(N)}(a), \quad \text{for } x \text{ near } a.$$

If a function did have constant $N$th derivative, it would be a polynomial of degree at most $N$, which we could see by integrating that derivative $N$ times.

The idea is to find simple polynomial approximations for a more complicated function given certain data regarding its behavior. In particular, if we know $f(a)$, $f'(a)$, $f''(a)$, and so on, then we should know something about how the function $f(x)$ behaves near $x = a$, and be able to produce a polynomial which mimics that behavior. In doing so we will make repeated use of the following lemma, which is useful in many other contexts as well:

**Lemma 11.1.1** Given any function $g$, with derivative $g'$ existing and continuous on the closed interval with endpoints $x$ and $a$ (i.e., $[a,x]$ or $[x,a]$, depending upon whether $x \leq a$ or $a \leq x$), the following equation holds:

$$g(x) = g(a) + \int_a^x g'(t) \, dt. \quad (11.15)$$

This is easy enough to verify. Since $g$ is clearly an antiderivative of $g'$, the Fundamental Theorem of Calculus gives

$$g(a) + \int_a^x g'(t) \, dt = g(a) + g(t) \bigg|_a^x = g(a) + g(x) - g(a) = g(x),$$

which is the equation (11.15) in reverse, q.e.d.

Two simple observations are worth making here.

1. It is interesting to verify this for the special case of (11.15) when $x = a$: $g(a) = g(a) + \int_a^a g'(t) \, dt = g(a) + 0 = g(a)$. Recall that such an integral as appears here over any interval of length zero is necessarily zero.
2. This easily lends itself to a simple physics application. If we replace \( g(x) \) by \( s(t) \), where \( s(t) \) is the position at time \( t \) and \( s_0 = s(t_0) \), we get \( s(t) = s_0 + \int_{t_0}^{t} s'(t) \, dt \). If \( s'(t) = v(t) \) is constant, we get \( s(t) = s_0 + v(t_f - t_0) \). If \( t_0 = 0 \), we get \( s = s_0 + vt \). Other cases also follow quickly from (11.15).

**Derivation of \( P_0(x) \)**

For a function \( f(x) \), if we would like to approximate the value of the function for \( x \) near \( a \), the simplest assumption is that the function is approximately constant near \( x = a \). The obvious choice for that constant is \( f(a) \) itself. Hence we might assume \( f(x) \approx f(a) \). (Note that \( f(a) \) is itself a constant.) The approximation of \( f(x) \) which assumes the function approximately constant is then \( P_0(x) \):

\[
P_0(x) = f(a).
\]

This is also called the *zeroth-order approximation* of \( f(x) \) centered at \( x = a \), and we can write \( f(x) \approx P_0(x) \) for \( x \) near \( a \), i.e., for \( |x - a| \) small. (See again Figure 11.1, page 751.) Summarizing, for \( x \) near \( a \),

\[
f(x) \approx f(a). \tag{11.17}
\]

A natural question then arises: how good is the approximation (11.17)? Later we will have a sophisticated estimate on the error in assuming \( f(x) \approx P_0(x) = f(a) \). For now we take the opportunity to foreshadow that result by attacking the question intuitively. The answer will depend upon the answers to two related questions, which can be paraphrased as the following.

(i) How good is the assumption that \( f \) is constant on the interval from \( a \) to \( x \)?

In other words, how fast is \( f \) changing on that interval?

(ii) How far is \( x \) from \( a \)?

These factors both contribute to the error. For instance if the interval from \( a \) to \( x \) is short, then a relatively slow change in \( f \) means small error \( f(x) - P_0(x) = f(x) - f(a) \) over such an interval. Slow change can, however, accumulate to create a large error if the interval from \( a \) to \( x \) is long. On the other hand, a small interval can still allow for large error if \( f \) changes quickly on the interval. The key to estimating how fast the function changes is, as always, the size of its derivative, assuming the derivative exists. Translating (i) and (ii) above into mathematical quantities, we say the bounds of the error will depend upon the following:

(a) the size of \(|f'(t)|\) as \( t \) ranges from \( a \) to \( x \) (assuming \( f'(t) \) exists for all such \( t \)), and

(b) the distance \(|x - a|\).

We will see similar factors accounting for error as we look at higher-order approximations \( P_1(x) \), \( P_2(x) \) and so on in this subsection, and the actual form of the general estimate for the error (also known as the *remainder*) in subsequent sections.

**Derivation of \( P_1(x) \)**

It was remarked in the last subsection that \( P_0 \) is not likely a good approximation for \( x \) very far from \( a \) if \( f' \) is large. In computing \( P_1(x) \), we will not assume \( f \) is approximately constant (as we did with \( P_0 \)), but instead assume that \( f' \) is approximately constant. To be clear, here are the assumptions from which \( P_1 \) is computed:
11.1. TAYLOR POLYNOMIALS: EXAMPLES AND DERIVATION

- We know \( f(a) \) and \( f'(a) \);
- \( f'(t) \) is approximately constant for \( t \) from \( a \) to \( x \).

For this derivation we will use the lemma from the beginning of this section (that is Lemma 11.1.1, page 763). Note that the following derivation uses the fact that \( f'(a) \) is a constant, and our assumption \( f'(t) \approx f'(a) \).

\[
f(x) = f(a) + \int_a^x f'(t) \, dt
\]

\[
\approx f(a) + \int_a^x f'(a) \, dt = f(a) + f'(a)\|t\| |_{a}^{x} = f(a) + f'(a)x - f'(a)a = f(a) + f'(a)(x - a). \]

Thus we define \( P_1(x) \), the first-order approximation of \( f(x) \) centered at \( x = a \) by

\[
P_1(x) = f(a) + f'(a)(x - a). \quad (11.18)
\]

This was also called the linear approximation of \( f(x) \) at \( a \) in Chapter 5 ((5.13), page 496).

From the graphs in Figure 11.1, page 751 we can see how \( P_0 \) and \( P_1 \) can differ. Because assuming constant derivative is often less risky, error-wise, than assuming constant height, \( P_1(x) \) is usually a better approximation for \( f(x) \) near \( x = a \), and indeed one can usually stray farther from \( x = a \) and have a reasonable approximation for \( f(x) \) if \( P_1(x) \) is used instead of \( P_0(x) \).\(^{10}\)

Again we ask how good is this newer approximation \( P_1(x) \), and again the intuitive response is that it depends upon answers two questions:

(i) How close is \( f'(t) \) to constant in the interval between \( a \) and \( x \)?

(ii) How far are we from \( x = a \)?

The first question can be translated into, “how fast is \( f' \) changing on the interval between \( a \) and \( x \)?” This can be measured by the size of \( f'' \) in that interval, if it exists there. Again translating (i) and (ii) into quantifiables, we get that the accuracy of \( P_1(x) \) depends upon

(a) the size of \( |f''(t)| \) as \( t \) ranges from \( a \) to \( x \) (assuming \( f''(t) \) exists for all such \( t \)), and

(b) the distance \( |x - a| \).

If \( f'' \) is relatively small, then \( f' \) is relatively constant, and then the computation we made giving \( f(x) \approx f(a) + f'(a)(x - a) \), i.e., \( f(x) \approx P_1(x) \), will be fairly accurate as long as \( |x - a| \) is not too large. See again Figure 11.1, page 751.

**Derivation of \( P_2(x) \)**

To better accommodate the change in \( f'' \), we next replace the assumption that \( f' \) is constant with the assumption that, rather than constant, it is changing at a constant rate. In other words, we assume that \( f'' \) is constant. So our assumptions in deriving \( P_2(x) \) are:

\(^{10}\)Note that in an example of motion, this is like choosing between an assumption of constant position, and of constant velocity. Intuitively the constant velocity assumption should yield a better approximation of position, for a while, than would a constant position assumption. However there are functions with very fast oscillations but low magnitude, for which the assumption of a constant height is less problematic than the assumption of a constant derivative, which may be quite large. Indeed a function with a very large derivative may stay surprisingly bounded, while a strictly bounded function can have large values for derivatives, so the value of these assumptions of some kind of constancy must be considered in context. Further consideration of these points is left to the reader.
• $f(a)$, $f'(a)$ and $f''(a)$ are known;

• $f''(t)$ is approximately constant from $t = a$ to $t = x$, i.e., $f''(t) \approx f''(a)$.

Again we use the lemma at the beginning of the section, except this time we use it twice: first, in approximating $f'$; and then integrating that approximation to approximate $f$.

$$f'(x) = f'(a) + (f'(x) - f'(a))$$
$$= f'(a) + \int_{x}^{x} f''(t) \, dt$$
$$\approx f'(a) + \int_{a}^{x} f''(a) \, dt$$
$$= f'(a) + f''(a)(x - a).$$

Note that the computation above was the same as from the previous section, except that the part of $f'$ there is played by $f''$ here, and the part of $f$ there is played by $f'$ here. We integrate again to approximate $f$. The second line below uses the approximation for $f'$ derived above.

$$f(x) = f(a) + \int_{a}^{x} f'(t) \, dt$$
$$\approx f(a) + \int_{a}^{x} \left[ f'(a) + f''(a)(t - a) \right] dt$$
$$= f(a) + f'(a)(x - a) + \left[ \frac{f''(a)}{2} (t - a)^2 \right]_{a}^{x}$$
$$= f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2 - \frac{1}{2} f''(a)(a - a)^2$$
$$= f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2.$$

Thus we define the second-order (or quadratic) approximation of $f(x)$ centered at $x = a$ by

$$P_2(x) = f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2. \quad (11.19)$$

Again, the accuracy depends upon (i) how close $f''(t)$ is to constant from $t = a$ to $t = x$, and (ii) how far we are from $x = a$. These can be quantified by the sizes of (a) $|f'''(t)|$ on the interval from $t = a$ to $t = x$, and (b) how large is $|x - a|$.

It is reasonable to take into account how fast $f'$ changes on the interval from $a$ to $x$. For $P_2$ we assume, not that $f'$ is approximately constant as we did with $P_1(x)$, but that the rate of change of $f'$ is constant on the interval, i.e., that $f''$ is constant (and equal to $f''(a)$) on the interval. In fact this tends to make $P_2(x)$ “hug” the graph of $f(x)$ better, since it accounts for the concavity. Figure 11.1, page 751 shows how $P_0(x)$, $P_1(x)$ and $P_2(x)$ can give progressively better approximations of $f(x)$ near $x = a$ (for the case $f(x) = e^x$ and $a = 0$). The extent to which we err in that assumption is the extent to which $f''$ (related to concavity) is non-constant, but at least near $x = a$, $P_2(x)$ accommodates concavity, as well as slope and height of the function $f(x)$.
Conclusion

The proof of the final formula for

\[ P_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)(x-a)^n}{n!} \]

would use an induction method, where one proves that

1. the formula holds true for the first, or first few cases, say \( N = 0, 1, 2 \) under their respective assumptions (that they match the function and its first \( N \) derivatives at \( x = a \), and that their degrees are at most \( N \)), and

2. the establishment of the formula for \( P_{N+1}(x) \) (regardless of \( N \in \mathbb{N} \cup \{0\} \)) implies its truth for the \((N+1)\)st case. That is,

\[ P_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)(x-a)^n}{n!} \implies P_{N+1}(x) = \sum_{n=0}^{N+1} \frac{f^{(n)}(a)(x-a)^n}{n!}. \]

Thus the establishment of the formula for \( P_0, P_1 \) and particularly \( P_2 \) implies it is also established for \( P_3 \), which in turn implies it is established for \( P_4 \), and so on, so that for instance its truth for \( P_{1000} \) is established because it is just a matter of following the implication in (2), also called the induction step, 998 times.

While we already proved (1), the proof of (2) is somewhat long and distracting, so we omit it here. However we will include in the exercises the case of computing \( P_3(x) \) from scratch, where one assumes knowledge of \( f(a), f'(a), f''(a), f'''(a) \) and assumes that \( f'''(x) \approx f'''(a) \), i.e., \( f''' \) is approximately constant, and integrating back to what that would imply for \( P_3(x) \), the function which is an at most degree-3 polynomial and conforms to those assumptions on its derivative, and is thus an approximation of \( f(x) \), at least near \( x = a \). By the time a student derives the formula for \( P_3(x) \) in that manner, it should seem quite reasonable that the pattern will continue for \( P_4(x) \), \( P_5(x) \) and so on.

**Exercises**

1. Given \( f(x) = \frac{1}{1-x} \), and \( a = 0 \),
   
   (a) show using (11.2), page 748 that
   
   \[ P_3(x) = 1 + x + x^2 + x^3 + x^4 + x^5. \]
   
   (b) What do you suppose is the general formula for \( P_N(x) \)?
   
   (c) Recalling facts about geometric series, for \( |x| < 1 \) what is the sum \( \sum_{n=0}^{\infty} x^n \)?

2. Find \( P_3(x) \) if \( f(x) = e^{2x} \) and \( a = 0 \).

3. Find \( P_3(x) \) if \( f(x) = e^{-3x} \) and \( a = 0 \).

4. If \( e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \) represents a series for \( f(x) = e^x \), then how would we expect to represent the following as series?
   
   (a) \( e^{2x} = \)
   
   (b) \( e^{-3x} = \)
   
   (c) \( e^{x^2} = \)
   
   (d) \( x^2 e^x = \)

5. Find \( P_3(x) \) where \( f(x) = \sin x \) and \( a = \pi \).

6. Find \( P_3(x) \) if \( f(x) = \tan x \) and \( a = \frac{\pi}{4} \).

7. Find \( P_2(x) \) if \( f(x) = \tan^{-1} x \) and \( a = 0 \).

8. Find \( P_2(x) \) if \( f(x) = \tan^{-1} x, a = 1 \).
9. Find \( P_2(x) \) if \( f(x) = \sqrt{1 + x^2}, \ a = 0 \).

10. Find \( P_3(x) \) if \( f(x) = x^3, \ a = 1 \).

11. Find a formula for \( P_N(x) \) if \( f(x) = \frac{1}{x}, \ a = 1 \).

12. Find a formula for \( P_N(x) \) if \( f(x) = \frac{1}{x}, \ a = -1 \).

13. Find \( P_3(x) \) if \( f(x) = \sin x, \ a = \frac{\pi}{2} \).

14. Find \( P_3(x) \) if \( f(x) = \sin x, \ a = -\frac{\pi}{6} \).

15. Show that (11.13) is indeed a solution to (11.12) by taking two time derivatives of each side of (11.12), remembering to employ the chain rule where appropriate.

16. If \( \alpha \in \mathbb{R} \), find \( P_5(x) \) for \( f(x) = (1 + x)^\alpha \) and \( a = 0 \).

17. Suppose at time \( t = 1 \) we know that \( s = 2, \ v = 5 \) and \( a = -7 \). What is likely to be our best approximation for \( s(t) \) near time \( t = 1 ? \)

18. Assuming \( f'''(x) = 6 \) for all \( x \), and \( f''(2) = 8, \ f'(2) = 7 \) and \( f(2) = 5 \), what is \( f(x) \)?

19. If we know \( f'''(0) = 12, \ f''(2) = 22, \ f'(4) = 92, \) and \( f(1) = 2 \), assuming \( f'''(x) \) is constant, what is \( f(x) \)?

For Exercises 20–22, use \( P_3(x) \) centered at \( a = 0 \) to approximate the given quantity. Compare that to the actual value (given by a calculator or similar device).

20. \( f(x) = \sin x \) at \( x = \pi/4 \).

21. \( f(x) = \cos x \) at \( x = \pi/4 \).

22. \( f(x) = e^x \) at \( x = 0.5 \).

23. Consider \( f(x) = \ln x \), and its Taylor Polynomials \( P_n(x) \) centered at \( a = 1 \).

(a) Compute \( P_0(x), P_1(x), \ldots, P_6(x) \).

(A pattern should become readily apparent.)

(b) Using a calculator or similar device find \( P_0(2), P_1(2), \ldots, P_6(2) \) as approximations of \( \ln 2 \). Compare these to \( \ln 2 \), and comment on the apparent efficiency of the approach \( P_n(2) \to \ln 2 \) as \( n \to \infty \).

(c) Repeat the above but with \( P_0(1/2), P_1(1/2), \ldots, P_6(1/2) \), as approximations of \( \ln(1/2) \).

(d) Note that \( \ln(1/2) = -\ln 2 \). Does this suggest a more efficient method of approximating \( \ln 2 \) using Taylor Polynomials? (Note the relative positions of 2, 1/2 and the center of your polynomials.)

(e) Repeat (b)–(c) to compute \( P_0(1/4), P_1(1/4), \ldots, P_6(1/4) \), compared to \( \ln(1/4) = -2 \ln 2 \).

(f) Is there any reason why we might not be interested in \( P_n(0) \)?

24. By applying \( \frac{d^2}{dx^2} \) to both sides of (11.13), show that \( \theta \) satisfies (11.12).

25. Compute \( P_3(x) \) in the general case by (1) listing the hypotheses from which \( P_3(x) \) arises as an approximation of \( f(x) \), and (2) performing the integration steps from those hypotheses.

(Read “Conclusion,” page 767.)
11.2 Accuracy of \( P_N(x) \)

All of this makes for lovely graphs, but one usually needs some certainty regarding just how accurate we can expect \( P_N(x) \) to be if it is to be used to approximate \( f(x) \). Fortunately, there is a way to estimate—here meaning to find an upper bound on the size of—the error arising from replacing \( f(x) \) with \( P_N(x) \). This difference \( f(x) - P_N(x) \) is also referred to as the remainder \( R_N(x) \):

\[
R_N(x) = f(x) - P_N(x).
\]  

(11.20)

Perhaps the name “remainder” makes more sense if we rewrite (11.20) in the form

\[
f(x) = P_N(x) + R_N(x) \tag{11.21}
\]

Of course if we knew the exact value of \( R_N(x) \), then by (11.21) we know \( f(x) \) since we can always calculate \( P_N(x) \) exactly, even with pencil and paper since, after all, it is just a polynomial. Often the best we can expect is to possibly have some estimate on the size of \( R_N(x) \). This can often be accomplished by knowing the rough form of \( R_N \), as is given in the following theorem.

**Theorem 11.2.1 (Remainder Theorem)** Suppose that \( f, f', f'', \ldots, f^{(N)} \) and \( f^{(N+1)} \) all exist and are continuous on the closed interval with endpoints both \( a \) and \( x \). Then

\[
R_N(x) = \frac{f^{(N+1)}(z)(x-a)^{N+1}}{(N + 1)!} \tag{11.22}
\]

where \( z \) is some (unknown) number between \( a \) and \( x \).

With this (11.21) could be rewritten

\[
f(x) = f(a) + \frac{f'(a)(x-a)}{1!} + \cdots + \frac{f^{(N)}(a)(x-a)^N}{N!} + \frac{f^{(N+1)}(z)(x-a)^{N+1}}{(N + 1)!} . \tag{11.23}
\]

Thus, the remainder looks just like the next term to be added to construct \( P_{N+1}(x) \), except that the term \( f^{(N+1)}(a) \) is replaced by the unknown quantity \( f^{(N+1)}(z) \).

A few examples of how the form (11.23) plays out are in order.

**Example 11.2.1** Write \( f(x) = e^x \) as the sum of \( P_4(x) \) and the remainder \( R_4(x) \), with center \( a = 0 \).

**Solution:** Since all derivatives of \( e^x \) are not only existing on all of \( \mathbb{R} \), but also simply \( e^x \), then of course \( f^{(n)}(0) = e^0 = 1 \) for all \( n = 0, 1, 2, \ldots \), we can write

\[
e^x = P_4(x) + R_4(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{f^{(5)}(z)(x-0)^5}{5!} \\
= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{e^z x^5}{5!} ,
\]

for some \( z \) between 0 and \( x \).

---

\textsuperscript{11}There are several remainder theorems addressing the size or form of the remainder \( R_N(x) \), including one offered by Taylor himself. This form (11.22) is due to Joseph-Louis Lagrange (1736–1813), an Italian-born mathematician and physicist whose importance to both fields—and to the understanding of their interconnectedness—cannot be overstated. However his work tends to deal in advanced topics which are not easily explained without the context of at least upper-division undergraduate mathematics and physics. The remainder theorem above is one exception.
Example 11.2.2 Write \( \sin x \) as the sum of \( P_3(x) \) and the remainder \( R_3(x) \).

Solution: Note that all derivatives of \( \sin x \) are of the form \( \pm \sin x \) or \( \pm \cos x \), which exist and are continuous on all of \( \mathbb{R} \). Now we constructed the chart for constructing up to \( P_5(x) \) for this function in Example 11.1.7, page 758, but we will do so again here but in a more summary form:

<table>
<thead>
<tr>
<th>( n )</th>
<th>( f^{(n)}(x) )</th>
<th>( 0 )</th>
<th>( 1 )</th>
<th>( 2 )</th>
<th>( 3 )</th>
<th>( 4 )</th>
<th>( 5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>( \sin x )</td>
<td>( \cos x )</td>
<td>( - \sin x )</td>
<td>( - \cos x )</td>
<td>( \sin x )</td>
<td>( \cos x )</td>
<td></td>
</tr>
</tbody>
</table>

From this we can write

\[
\sin x = P_3(x) + R_3(x) = x - \frac{x^3}{3!} + \frac{f^{(4)}(z)x^4}{4!} = x - \frac{x^3}{3!} - \frac{(\cos z)x^4}{4!},
\]

for some \( z \) between 0 and \( x \).

In fact we can write \( \sin x \) in any of the following ways:

\[
\begin{align*}
\sin x &= x + \frac{(\cos z)x^2}{2!}, & \text{for some } z \text{ between } 0 \text{ and } x, \\
\sin x &= x + \frac{(-\sin z)x^3}{3!}, & \text{for some } z \text{ between } 0 \text{ and } x, \\
\sin x &= x - \frac{x^3}{3!} + \frac{(-\cos z)x^4}{4!}, & \text{for some } z \text{ between } 0 \text{ and } x, \\
\sin x &= x - \frac{x^3}{3!} + \frac{(\sin z)x^5}{5!}, & \text{for some } z \text{ between } 0 \text{ and } x, \\
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{(\cos z)x^6}{6!}, & \text{for some } z \text{ between } 0 \text{ and } x,
\end{align*}
\]

and so on. The fact that the Taylor Polynomials for \( \sin x \), centered at \( a = 0 \) contain many “zero” terms means that we have a couple of choices for the remainder terms, for instance depending upon whether we wish to consider \( x - \frac{x^3}{3!} + \frac{x^5}{5!} \) to be \( P_3(x) \) or \( P_4(x) \), which are the same for this particular function \( \sin x \) with \( a = 0 \). Note that in each of the cases given above, the \( z \) will be between 0 and \( x \), but we should not expect to have the same value for \( z \) in each of the above, even if we choose the same value for \( x \).

A general proof of the Remainder Theorem is beyond the scope of this textbook. However, in the exercises the reader is invited to explore how the first case is simply the Mean Value Theorem (Theorem 5.3.1, page 488).

There are several cases where it is useful to find upper bounds (also called estimates) on the size of the remainders, which are after all the errors we incur by replacing functions with their Taylor Polynomial approximations.

Example 11.2.3 Suppose that \( |x| < 0.75 \). In other words, \(-0.75 < x < 0.75\). Then what is the possible error if we use the approximation \( \sin x \approx x - \frac{x^3}{3!} + \frac{x^5}{5!} ? \)

Solution: Notice that we are asking what is the remainder for the Taylor Polynomial \( P_6(x) \) (see Figures 11.3 and 11.4, page 759) where \( f(x) = \sin x \) and \( a = 0 \), if \( |x| < .75 \). (Recall that, for \( \sin x \), we have \( P_5 = P_6 \) when \( a = 0 \).) We will use the fact that \( |\sin z| \leq 1 \) and \( |\cos z| \leq 1 \) no matter what is the value of \( z \). Thus

\[
|R_6(x)| = \frac{|f^{(7)}(z)(x-0)^7|}{7!} = \frac{|-\cos z \cdot x^7|}{7!} = \frac{1}{7!} |\cos z| \cdot |x|^7 \leq \frac{1}{7!} \cdot 1 \cdot .75^7 = 0.00002648489.
\]
11.2. ACCURACY OF $P_N(X)$

This should be encouraging, since we have nearly five digits of accuracy from a polynomial with only three terms, when our angle is in the range $\pm0.75 \approx \pm43^\circ$.

A quick check shows that, to $\sin 0.75 \approx 0.681638760$, $P_6(0.75) \approx 0.6816650391$, and so the difference is $\sin 0.75 - P_6(0.75) \approx -0.000026279$, which is slightly less in absolute value than our error estimate of $0.00002648489$.

Example 11.2.4 Suppose we want to use the approximation $e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!}$.

a. How accurate is this if $|x| < 5$?

b. How accurate is this if $|x| < 2$?

c. What if $|x| < 1$?

Solution: Since the approximating polynomial is $P_4(x)$ with $a = 0$, we are looking for a bound for

$$|R_4(x)| = \left| \frac{f^{(5)}(z)x^5}{5!} \right| = \frac{e^zx^5}{5!} = \frac{1}{120}e^z|x|^5.$$  

a. $|x| < 5$: Now $z$ is between 0 and $x$, and since the exponential function is increasing, the worst possible case scenario is to have the greatest possible value for $z$ (which will be $x$ or 0, which ever is greater). Since the greatest $x$ can be is 5, it is safe to use $e^z < e^5$. Thus,

$$|R_4(x)| = \frac{1}{120}e^z|x|^5 < \frac{1}{120}e^5 \cdot 5^5 \approx 3865.$$  

Thus we see the exponential is not so well approximated by $P_4(x)$ for the whole range $|x| < 5$.

b. $|x| < 2$: Now we have $z$ between 0 and $x$, and $x$ between $−2$ and 2, so the the it is only safe to assume $z < 2$. Similar to the above, this gives

$$|R_4(x)| = \frac{1}{120}e^z|x|^5 < \frac{1}{120}e^2 \cdot 2^5 \approx 1.97.$$  

We see we have a much better approximation if $|x| < 2$.

c. $|x| < 1$: Here we can only assume $z < 1$:

$$|R_4(x)| = \frac{1}{120}e^z|x|^5 < \frac{1}{120}e^1 \cdot 1^5 \approx 0.02265.$$  

There are several remarks which should be made about this example.

1. Notice that we “begged the question,” since we used calculations of $e^5$, $e^2$ and $e^1$ to approximate the error. This is all correct, but perhaps a strange thing to do since such quantities are exactly what we are trying to approximate with the Taylor Polynomial. But even with this problem, the polynomial is useful because it can be quickly calculated for the whole range $|x| < 5$, 2 or 1 for some application, and the accuracy estimated using only $e^5$, $e^2$ or $e^1$, which are finitely many values.

One way to avoid this philosophical problem entirely is to use $x > 0 \implies e^x < 3^x$, since $3^x$ is easier to calculate for the integers we used. For example, $e^5 < 3^5$. However, we need to be somewhat careful, since $x < 0 \implies 3^x < e^x$. Here it would be fine to use $3^x$, since we were interested in a larger range of $x$ which included positive numbers. If only interested in $x \in (−5,0)$, for example, we might use $e^x < 2^x$ there.
2. Note that the error shrinks in a-c, that is as we restrain \( x \) so that \( |x| < 5, 2, 1 \) respectively for two reasons:
   
   (a) \( |f^{(5)}(z)| = e^z \) shrinks, since \( z \) is more constrained.
   
   (b) \(|x|^{5} \) shrinks, since the maximum possible value of \(|x|\) is smaller.

   We benefit from both these factors when we shrink \(|x|\).

3. If we truly needed more accuracy for \(|x| < 5\), we could take a higher-order Taylor Polynomial, such as \( P_{15}(x) \), giving

   \[
   |R_{15}(x)| = \frac{1}{15!}e^z|x|^{15} < \frac{1}{15!}e^55^{15} \approx 3.5
   \]

   This might still seem like a large error, but it is relatively small considering \( e^5 \approx 148 \). If the error is still too large, consider \( P_{20}(x) \), with

   \[
   |R_{20}(x)| = \frac{1}{21!}e^z|x|^{21} < \frac{1}{20!}e^55^{20} \approx 0.000277.
   \]

   When we increase the order of the Taylor Polynomial, we always have the benefit of a growing factorial term \( N! \) in the remainder’s denominator. As long as the term \( |f^{N+1}(z)| \) does not grow significantly, the factorial will dominate the exponential \(|x - a|^{N+1}\).

4. Finally, the exponential will always increase faster as \( x \to \infty \) than any polynomial (be it \( P_N(x) \) for a fixed \( N \) or any other polynomial), and “flatten out” like no polynomial can (excepting the zero polynomial) as \( x \to -\infty \), so it is really not a good candidate for approximation very far from zero.

A reasonable question to ask next is how large do we need to have \( N \) so that \( P_N(x) \) is within a tolerable size. The next examples consider that question.

**Example 11.2.5** Suppose we wish to find a Taylor Polynomial \( P_N(x) \) for \( f(x) = \cos x \) centered at \( x = 0 \) so that \( P_N(x) \) is within \( 10^{-7} \) of \( f(x) \) for \(|x| < \pi\). What is the range of \( N \) which assures this?

**Solution:** Here we will use the guaranteed, if seemingly crude, estimate for the size of the error \( |R_N(x)| \), in which we again note that \( f^{(n)}(z) \) will be of the form \( \pm \sin z \) or \( \pm \cos z \) regardless of \( n \), and thus \( |f^{(n)}(z)| \leq 1 \) regardless of \( z \). From this we get

\[
|R_N(x)| = \left| \frac{x^{N+1}f^{(N+1)}(z)}{(N+1)!} \right| \leq \frac{|x|^{N+1} \cdot 1}{(N+1)!} < \frac{\pi^{N+1}}{(N+1)!}.
\]

It is enough that this last term is at most \( 10^{-7} \), but solving such an inequality does not involve elementary algebraic manipulations. Instead we will need experiment with some numerical values, comparing \( N \) to \( \frac{\pi^{N+1}}{(N+1)!} \), the latter listed rounded upwards to assure correctness.

\[
\begin{array}{|c|c|c|c|c|c|c|}
\hline
N = & \cdots & 15 & 16 & 17 & 18 & 19 & 20 & \cdots \\
|R_N| \leq & \cdots & 5 \times 10^{-6} & 8 \times 10^{-7} & 2 \times 10^{-7} & 3 \times 10^{-8} & 4 \times 10^{-9} & 6 \times 10^{-10} & \cdots \\
\hline
\end{array}
\]

From the chart we see that \( N \geq 18 \) guarantees that \( P_N(x) \) is within \( 10^{-7} \) of \( \cos x \), for \( -\pi < x < \pi \).

We know that the size of the estimate will continue to decrease because with each increment we multiply it by a factor \( \pi/(N+1) \), which is less than 1 once \( N > 3 \).

It is common to use a “worst-case” estimate in computations such as the one above, in that case using \(|\pm \sin z|, |\pm \cos z| \leq 1\) and \(|x| < \pi\). It would be very difficult to find more precise bounds for that range of \( x \).
11.2. ACCURACY OF \( P_N(X) \)

Example 11.2.6 Find \( N \) so that \( P_N(x) \) as an approximation for \( f(x) = e^x \) is accurate to within \( 10^{-5} \) when \( |x| < 2 \).

Solution: Here we have \( f^{(n)}(z) = e^z \) regardless of \( n \), and so for some \( z \) between 0 and \( x \) (and thus \( z \in (-2, 2) \)) we have

\[
|R_N(x)| = \left| e^z |x|^{N+1} \right| \leq \frac{e^2 \cdot 2^{N+1}}{(N+1)!}.
\]

It is enough that this last quantity be smaller than \( 10^{-5} \). As in the example above, algebraic techniques will not yield an answer directly, and so we will need to perform some numerical experiments. Below we list some values of \( N \) and \( e^2 \cdot 2^{N+1}/(N+1)! \), the latter rounded upwards, except for one crucial value, namely \( N = 12 \).

\[
\begin{array}{cccccccc}
N = & 9 & 10 & 11 & 12 & 13 & 14 & \cdots \\
|R_N| \leq & 3 \times 10^{-3} & 4 \times 10^{-4} & 7 \times 10^{-5} & 9.8 \times 10^{-6} & 2 \times 10^{-6} & 2 \times 10^{-7} & \cdots \\
\end{array}
\]

We see from the chart, and the clear fact that these estimates will continue to decrease, that \( N \geq 12 \) suffices. Thus \( P_{12}(x) \) and higher ordered Taylor Polynomials centered at \( a = 0 \) will approximate \( f(x) = e^x \) within \( 10^{-5} \) for \( |x| < 2 \).

That the estimates on the error in the above example will continue to decrease is again seen by the fact that we can derive the \( N = m \) estimate by multiplying the previous estimate and \( 2/(m+1) \), which is less than 1 once \( m > 1 \), and so that next estimate will be smaller.

In the next example we can more directly compute \( N \) to give the error bound we desire.

Example 11.2.7 For \( f(x) = \ln x \), assuming \( |x - 1| < 0.5 \), find \( N \) which guarantees that \( P_N(x) \) centered at \( a = 1 \) is within \( 10^{-5} \) of \( \ln x \).

Solution: In Example 11.1.10, page 762 we saw that \( f^{(n)}(x) = (-1)^{n+1}(n-1)!x^{-n} \), for \( n = 1, 2, 3, \cdots \). (It is a simple enough computation but for space reasons we refer the reader to that example.) We also derived \( P_N(x) \) in that example, and can say that for \( x > 0 \)—that is, where all derivatives exist and are continuous (on an interval containing 1), the remainder theorem (Theorem 11.2.1, page 769) gives us

\[
\ln x = \sum_{n=1}^{N} \frac{(-1)^{n+1}(n-1)!}{n!} (x-1)^n + \frac{f^{(N+1)}(z)(x-1)^{N+1}}{(N+1)!} R_N(x)
\]

\[
= \sum_{n=1}^{N} \frac{(-1)^{n+1}(x-1)^n}{n} + \frac{((-1)^N + z)^N}{(N+1)!} (x-1)^{N+1}
\]

So we desire \( N \) such that \( |x - 1| < 0.5 \implies |R_N(x)| < 10^{-5} \). Note that \( |x - 1| < 0.5 \iff x \in (0.5, 1.5) \), and since \( z \) is between 1 and \( x \) we also have \( z \in (0.5, 1.5) \implies \frac{1}{2} \in (2/3, 2) \). Thus

\[
|R_N(x)| = \left| \frac{(-1)^N(x-1)^{N+1}}{(N+1)} \cdot \left( \frac{1}{2} \right)^{N+1} \right| < \frac{1}{N+1} \cdot (1/2)^{N+1} \cdot 2^{N+1} = \frac{1}{N+1}
\]

A sufficient condition that \( |R_N(x)| < 10^{-5} \) is then \( \frac{1}{N+1} \leq 10^{-5} \), which we can solve easily:

\[
\frac{1}{N+1} \leq \frac{1}{10^5} \iff 10^5 \leq N + 1 \iff 99,999 \leq N.
\]

Thus we can guarantee an error of less than \( 10^{-5} \) if \( N \geq 99,999 \), assuming \( |x - 1| < 0.5 \).
In the example above we were somewhat lucky that some factors in the remainder estimate canceled. Suppose instead we assume $|x - 1| < \frac{3}{4}$. This expands slightly our range of $x$, so that $-3/4 < x - 1 < 3/4$ and so $1/4 < x < 7/4$, and this has implications regarding our estimate.

If we were to assume $x \in [1, 7/4]$, then we have $z$ in the same range (between 1 and $x$, and therefore in $z \in [1, 7/4]$ as well). In such a case $x - 1 \in [0, 3/4]$ and $\frac{1}{z} \in (4/7, 1)$, giving our error estimate as

$$|R_N(x)| = \left| \frac{(-1)^N(x - 1)^{N+1}}{(N+1)} \cdot \left(\frac{1}{z}\right)^{N+1} \right| \leq \frac{\left(\frac{3}{4}\right)^{N+1}}{N+1} \cdot 1^{N+1} = \frac{1}{(N+1)} \cdot \left(\frac{3}{4}\right)^{N+1}.$$  

From that estimate we can see clearly that $|R_N(x)| \to 0$ as $N \to \infty$.

Unfortunately, if we have $x \in (1/4, 1]$, with $z$ in the same range, we get $x - 1 \in (-3/4, 0]$ and $\frac{1}{z} \in [1, 4)$. In this case our most obvious estimate becomes

$$|R_N(x)| = \left| \frac{(-1)^N(x - 1)^{N+1}}{(N+1)} \cdot \left(\frac{1}{z}\right)^{N+1} \right| \leq \frac{\left(\frac{3}{4}\right)^{N+1}}{N+1} \cdot 4^{N+1} = \frac{3^{N+1}}{N+1},$$

which will grow larger as $N$ grows, and a quick numerical experiment can show this estimate never achieves anything nearly as small as $10^{-5}$.

What went wrong in this second case was that our estimate was too crude: we looked at a worst case scenario with $x$ and $z$ separately, when clearly they are coupled. Using completely different techniques, we will see later that, for $x \in (0, 2]$, we will have

$$\ln x = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}(x - 1)^n}{n+1} = (x - 1) - \frac{(x - 1)^2}{2} + \frac{(x - 1)^3}{3} - \frac{(x - 1)^4}{4} + \cdots,$$

and so the remainder terms will shrink for a given $x$, just not “uniformly;” they will tend to shrink faster for $x$ closer to 1, and not in quite the same way for $x \in (0, 1)$ as for $x \in (1, 2]$.

If we take for granted that the above series expansion is correct for $x \in (0, 2]$, then we can use alternating series methods to find the bounds on errors when $x \in [1, 2]$. For $x \in (1/4, 1]$ we can use a direct comparison test to a geometric series. For instance, if $x = 1/4$, the series becomes

$$-\frac{3/4}{2} + \frac{(-3/4)^3}{3} - \frac{(-3/4)^4}{4} + \cdots = -\sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{3}{4}\right)^n.$$

If we call this series $\sum a_n$, then $|a_n| \leq (3/4)^n$, from which we can have a geometric series, and from which we have

$$|R_N(1/4)| = \sum_{n=N+1}^{\infty} \frac{1}{n} (3/4)^n < \sum_{n=N+1}^{\infty} (3/4)^n = \frac{(3/4)^{N+1}}{1 - \frac{1}{4}} = \frac{1}{4} \left(\frac{3}{4}\right)^{N+1}.$$  

If we would like to ensure $|R_N(1/4)| < 10^{-5}$, we would solve (noting that $\ln(3/4) < 0$):

$$\frac{1}{4} \left(\frac{3}{4}\right)^{N+1} < 10^{-5} \implies \left(\frac{3}{4}\right)^{N+1} < 4 \times 10^{-5} \implies (N + 1) \ln(3/4) < \ln(4 \times 10^{-5}) \implies N > \frac{\ln(4 \times 10^{-5})}{\ln(3/4)} - 1 \implies N > 34.2,$$

and so we would take $N \geq 35$ to ensure our error is within $10^{-5}$, in using $P_N(1/4)$ to approximate $\ln \frac{1}{4}$. 

11.2. ACCURACY OF $P_N(X)$

Exercises

For Exercises 1–6, write the function in the form $f(x) = P_N(x) + R_N(x)$, where $P_N(x)$ and $R_N(x)$ are written out explicitly (see Examples 11.2.1–11.2.2).

1. $f(x) = \sin x$, $a = \pi$, $N = 5$
2. $f(x) = \sqrt{x}$, $a = 1$, $N = 3$
3. $f(x) = \frac{1}{x}$, $a = 10$, $N = 4$
4. $f(x) = e^x$, $a = 0$, $N = 9$.
5. $f(x) = \sec x$, $a = \pi$, $N = 2$.
6. $f(x) = \ln x$, $a = e$, $N = 3$.

7. Explain why the series below converges, and to the limit claimed below. (Hint: apply a hierarchy of functions reasoning to $R_N(x)$.)

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = e^x.$$ 

8. Many physics problems take advantage of the approximation $\tan x \approx x$ for $|x|$ small.

(a) Conjecture on where this approximation comes from, from a purely mathematical standpoint.

(b) Estimate the error for each of the three cases $|x| < 1$, 0.1 and 0.01. (Feel free to use a calculator to find upper bounds for the error.)

9. Suppose we wanted to find a Taylor Polynomial for $f(x) = \sin x$, centered at $a = 0$, with accuracy $|R_N(x)| \leq 10^{-10}$ valid for $-2\pi \leq x \leq 2\pi$. Find $N$ for the lowest-order Taylor Polynomial $P_N(x)$ which guarantees that accuracy for that interval, based upon the remainder formula. (This may require some numerical experimentation with the estimates.)

10. Repeat the previous problem, but for $f(x) = e^x$ and the interval $|x| \leq 10$.

11. Show that the Remainder Theorem for $P_0(x)$ is really just the Mean Value Theorem, Theorem 5.3.1, page 488. (Hint: $z = \xi$.)
11.3 Taylor/Maclaurin Series

Now we come to the heart of the matter. Basically, the Taylor Series of a function \( f \) which has all derivatives \( f', f'', \ldots \) existing at \( a \), is the series we get when we let \( N \to \infty \) in the expression for \( P_N(x) \). The Taylor Series equals the function if and only if the remainder terms shrink to zero as \( N \to \infty \):

### 11.3.1 Checking Validity of Taylor Series

**Theorem 11.3.1** Recalling the definition of the remainder \( R_N(x) = f(x) - P_N(x) \), where \( P_N(x) \) is an \( N \)th-order Taylor Polynomial for \( f(x) \) centered at some number \( a \in \mathbb{R} \), we have

\[
\lim_{N \to \infty} R_N(x) = 0 \iff f(x) = \lim_{N \to \infty} P_N(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!},
\]

that is,

\[
f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!} \iff \lim_{N \to \infty} R_N(x) = 0.
\]

**Proof:** First we prove \((\iff)\). Assume \( f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!} \). Then

\[
R_N(x) = f(x) - P_N(x) = \sum_{n=N+1}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!} \to 0 \quad \text{as} \quad N \to \infty. \tag{11.24}
\]

Next we prove \((\implies)\). Assume \( R_N(x) \to 0 \) as \( N \to \infty \). Then

\[
N \to \infty \quad \implies \quad f(x) - R_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)(x-a)^n}{n!} \quad \text{and} \quad f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!},
\]

which shows \( f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!} \), q.e.d.

The series we get from Theorem 11.3.1 above has the following name:

**Definition 11.3.1** Supposing that all derivatives of \( f(x) \) exist at \( x = a \), the series given by

\[
\sum_{n=0}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!}
\]

is called the Taylor Series of \( f(x) \) centered at \( x = a \).

\[\text{12}^\text{Recall that the “tail end” } \sum_{n=N+1}^{\infty} b_n \text{ of a convergent series } \sum_{n=0}^{\infty} b_n \text{ shrinks to zero as } N \to \infty. \text{ This “tail end” is represented by } S - S_N, \text{ where } S \text{ is the full series and } S_N \text{ the } N \text{th partial sum. Recall } S_N \to S \implies S - S_N \to 0.\]
Thus, Theorem 11.3.1 can be restated that the Taylor Series will equal the function if and only if the remainders $R_N$ from the Taylor Polynomials shrink to zero as $N \to \infty$.

A special case of the Taylor Series is the case $a = 0$. This occurs often enough it is given its own name:

**Definition 11.3.2** If a Taylor Series is centered at $a = 0$, it is called a Maclaurin Series. In other words, if all derivatives of $f(x)$ exist at $x = 0$, the function’s Maclaurin Series is given by

$$
\sum_{n=0}^{\infty} \frac{f^{(n)}(0)x^n}{n!} \quad (11.26)
$$

The partial sums are sometimes called Maclaurin Polynomials. Note that both Taylor Series, and the special case of the Maclaurin Series, are in fact power series, introduced in (11.1), page 747.

In the following propositions, we will consider several Taylor and Maclaurin Series, and show where they converge based on Theorem 11.3.1 (which we restated in (11.24)) and other observations. Showing that $R_N \to 0$ in some cases will require creativity, but once we establish this fact for a series we will assume it from then on, as with those below:

**Proposition 11.3.1**

$$
e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \text{for all } x \in \mathbb{R}. \quad (11.27)
$$

**Proof:** Recall that $(\forall x \in \mathbb{R})(\forall n \in \{0, 1, 2, 3, \cdots\}) [f^{(n)}(x) = e^x]$. Thus, for any fixed $x \in \mathbb{R}$, we have

$$
R_N(x) = \frac{f^{(N+1)}(z)x^{N+1}}{(N+1)!} = e^z \frac{x^{N+1}}{(N+1)!}.
$$

Now $z$, while depending upon $N$ and $x$, is nonetheless between $x$ and 0, and so by the increasing nature of the exponential function we have $e^z < \max\{e^0, e^x\}$, and is thus bounded by $M = \max\{e^0, e^x\}$ (which depends only upon $x$, and not upon $N$ or $z$). Thus

$$
|R_N(x)| = e^z \frac{|x|^{N+1}}{(N+1)!} \leq M \cdot \frac{|x|^{N+1}}{(N+1)!} \to M \cdot 0 = 0 \quad \text{as } N \to \infty,
$$

since the numerator grows geometrically (or shrinks geometrically) as $N$ increases, while the denominator grows as a factorial. Recall that the factorial will dominate the exponential regardless of the base (in this case $|x|$) as $N \to \infty$. Since we showed $R_N(x) \to 0$ (by showing the equivalent statement $|R_N(x)| \to 0$) for any $x$, by Theorem 11.3.1, page 776, (11.27) follows, q.e.d.\footnote{A clever way to prove more directly that $M \cdot \frac{|x|^{N+1}}{(N+1)!} \to 0$ as $N \to \infty$ would be to show that the series $\sum_{n=1}^{\infty} M \cdot \frac{|x|^{n+1}}{(n+1)!}$ converges, which can be proved using a fairly straight-forward Ratio Test. This would show that the “$n$th term” approaches zero in the limit, since $\sum a_n$ converges $\implies a_n \to 0$, which is the contrapositive of the $n$th-term test for divergence (NTTTFD, Section 10.2).}  

13 Named for Colin Maclaurin, 1698–1746, a Scottish mathematician. He was apparently aware of Taylor Series, citing them in his work, but made much creative use of those centered at $a = 0$ and so eventually was honored to have the special case named for him.

14 It was important to notice that $e^x$ was bounded once $x$ was chosen, and that the bound is going to change with each $x$. The upshot is that for a given $x$, $R_N(x) \to 0$ but for two different $x$-values, this convergence of the remainder to zero—and thus the convergence of the Taylor series to the value $f(x)$—can occur at very different rates.
Also, absolute values were not needed around the $e^z$-term, since it will always be positive. Finally, to accommodate the case $x = 0$, we substituted the weaker “$\leq$” for the “$<$” in our inequality above. For the case $x = 0$, a careful look at the $P_N$ show $R_N(0) \equiv 0$. This is because 0 is where the series is centered. (Recall $P_N(a) = f(a)$.)

**Proposition 11.3.2**

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{n+1}}{(2n+1)!} \quad \text{for all } x \in \mathbb{R}. \quad (11.28)$$

**Proof:** Now $f^{(n)}(x)$ is of the form $\pm \sin x$ or $\pm \cos x$, which means it is bounded absolutely by 1, i.e., $|f^{(n)}(x)| \leq 1$. Thus for any given $x \in \mathbb{R}$ we have

$$|R_N(x)| = \left| \frac{f^{(N+1)}(z) x^{N+1}}{(N+1)!} \right| \leq 1 \cdot \frac{|x|^{N+1}}{(N+1)!} \to 1 \cdot 0 = 0 \text{ as } N \to \infty.$$

Again this is because the geometric term $|x|^{N+1}$ is a lower order of growth (and may even decay if $x \in (-1, 1)$) than the factorial $(N+1)!$. Thus, according to Theorem 11.3.1, (11.28) follows, q.e.d.

A nearly identical argument shows that

**Proposition 11.3.3**

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} \quad \text{for all } x \in \mathbb{R}. \quad (11.29)$$

Not all Taylor series converge to the function for all of $x \in \mathbb{R}$. Furthermore, it is often difficult to prove $R_N(x) \to 0$ when other techniques can give us that the Taylor Series in fact converges. For example, consider the following:

**Proposition 11.3.4**

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \cdots = \sum_{n=0}^{\infty} x^n. \quad \text{for all } x \in (-1, 1). \quad (11.30)$$

Though we can calculate the series directly (see Exercise 6, page 767), Equation (11.30) is obvious if we read it backwards, realizing that the series is geometric with first term $\alpha = 0$ and ratio $x$ (as in Theorem 10.1.1, page 706). Moreover, the series converges when $|x| < 1$ and diverges otherwise, from what we know of geometric series. From these observations, Proposition 11.3.4 is proved. We will see in Section 11.5 that many of the connections and manipulations we would like to make with Taylor/Maclaurin Series are legitimate. In fact, these methods are often much easier than computations from the Taylor Series definition. Consider Proposition 11.3.4. The actual remainder is of the form

$$R_N(x) = \frac{(N+1)!}{(1-z)^{(N+2)}} \left( \frac{x^{N+1}}{(N+1)!} \right) = \frac{x^{N+1}}{(1-z)^{N+2}}. \quad (11.31)$$

We know $z$ is between 0 and $x$, but without knowing more about where, it is not obvious that the numerator in our simplified $R_N$ will decrease in absolute size faster than the denominator. We will not belabor the point here, but just conclude that resorting to using facts about geometric series is a much simpler approach than attempting to prove $R_N(x) \to 0$ when $|x| < 1$. (See also the discussion after Example 11.2.7, page 773.)

Another interesting Taylor Series is the following:
Proposition 11.3.5 The following is the Taylor Series for $\ln x$ centered at $x = 1$:

$$\ln x = 1(x - 1) - \frac{1}{2}(x - 1)^2 + \frac{1}{3}(x - 1)^3 - \frac{1}{4}(x - 1)^4 + \cdots$$

(11.32)

$$= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}(x - 1)^n}{n}$$

for $|x - 1| < 1$, i.e., $x \in (0, 2)$.

We found $P_N$ in Example 11.1.10, page 762. A proof that (11.32) is valid for $(1/2, 2)$ in which one shows $R_N(x) \to 0$ in that interval is left as Exercise ???. The proof that the series is valid for all of $(0, 2)$ is left as an exercise in Section 11.4, and again in Section 11.5 after other methods are available. Finally, it is not difficult to show that the series also converges at $x = 2$ (by the Alternating Series Test) and so the series in fact converges for all of $(0, 2]$, so that by Abel’s Theorem, introduced later as Theorem 11.4.1, page 782, the series converges to $\ln x$ in all of $(0, 2]$.

11.3.2 Techniques for Writing Series using $\Sigma$-Notation

There are some standard tricks for writing formulas to achieve the correct terms in the summation. For instance, inserting a factor $(-1)^n$ or $(-1)^{n+1}$ to achieve the alternation of sign, depending upon whether the first term carries a “+” or “−.” We also pick up only the odd terms in the $\sin x$ expansion by using the $2n + 1$ factors, and get the evens in the $\cos x$ using the $2n$. Perhaps the best way to get comfortable with these manipulations is to write out a few terms of the summations on the right of (11.28), (11.29) and (11.32). For example, we can check the summation notation is consistent in (11.28) as follows (note we define $(-1)^0 = 1$ for simplicity):

$$\sum_{n=0}^{\infty} \frac{(-1)^{n}x^{2n+1}}{(2n+1)!} = \frac{(-1)^{0}x}{1!} + \frac{(-1)^{1}x^{3}}{3!} + \frac{(-1)^{2}x^{5}}{5!} + \frac{(-1)^{3}x^{7}}{7!} + \cdots$$

$$= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

However it would also be perfectly legitimate to instead write the above series as

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}x^{2n-1}}{(2n-1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots.$$
Example 11.3.1 Write the following in a compact Σ-notation.

a. \( \frac{x}{2} + \frac{x^2}{4} + \frac{x^3}{6} + \frac{x^4}{8} + \cdots \)

b. \( -\frac{x^3}{2} + \frac{x^5}{4} - \frac{x^7}{8} + \cdots \)

c. \( -\frac{x^2}{1 \cdot 3} + \frac{x^4}{1 \cdot 3 \cdot 5} - \frac{x^6}{1 \cdot 3 \cdot 5 \cdot 7} + \cdots \)

Solution:

a. We see the powers of \( x \) are increasing by 1, while the denominators are increasing by 2 with each new term added. The summations will appear different depending upon where the indices begin. Here are two possibilities, though the first is more obvious:

\[
\frac{x}{2} + \frac{x^2}{4} + \frac{x^3}{6} + \frac{x^4}{8} + \cdots = \sum_{n=1}^{\infty} \frac{x^n}{2n},
\]

\[
\frac{x}{2} + \frac{x^2}{4} + \frac{x^3}{6} + \frac{x^4}{8} + \cdots = \sum_{n=0}^{\infty} \frac{x^{n+1}}{2(n + 1)}. 
\]

b. Here we have only odd powers of \( x \). It is worth noting that therefore the powers of \( x \) are increasing by 2. We have alternating factors of \( \pm 1 \). In the denominator we have powers of 2. This can be written in the following ways (among others):

\[
x - \frac{x^3}{2} + \frac{x^5}{4} - \frac{x^7}{8} + \cdots = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^{2n-1}}{2n-1},
\]

\[
x - \frac{x^3}{2} + \frac{x^5}{4} - \frac{x^7}{8} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n + 1}. 
\]

c. The powers of \( x \) here are all even, hence increasing by 2 with each step. There is also alternation of signs. Finally the denominators are products of odd numbers, similar to a factorial but skipping the even factors. In a case like this, we allow for a more expanded writing of the pattern in the Σ-notation. We write the following:

\[
-\frac{x^2}{1} + \frac{x^4}{1 \cdot 3} - \frac{x^6}{1 \cdot 3 \cdot 5} + \frac{x^8}{1 \cdot 3 \cdot 5 \cdot 7} + \cdots = \sum_{n=1}^{\infty} \frac{(-1)^n x^{2n}}{1 \cdot 3 \cdot 5 \cdots (2n - 1)},
\]

\[
-\frac{x^2}{1} + \frac{x^4}{1 \cdot 3} - \frac{x^6}{1 \cdot 3 \cdot 5} + \frac{x^8}{1 \cdot 3 \cdot 5 \cdot 7} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^{n+1} x^{2n+2}}{1 \cdot 3 \cdot 5 \cdots (2n + 1)}. 
\]

If we had some compelling reason, we might even begin at \( n = 3 \), for instance:

\[
-\frac{x^2}{1} + \frac{x^4}{1 \cdot 3} - \frac{x^6}{1 \cdot 3 \cdot 5} + \frac{x^8}{1 \cdot 3 \cdot 5 \cdot 7} + \cdots = \sum_{n=3}^{\infty} \frac{(-1)^n x^{2n-4}}{1 \cdot 3 \cdot 5 \cdot (2n - 5)}. 
\]

It is understood that the denominator contains all the odd factors up to \( (2n - 1) \), \( (2n + 1) \) or \( (2n - 5) \), depending on the form chosen. Though the first two terms do not contain all of \( 1 \cdot 3 \cdot 5 \), we put in those three numbers into the Σ-form to establish the pattern, which is understood to terminate at \( (2n - 1) \) or \( (2n + 1) \) even if that means stopping before 3 or 5.
11.3. TAYLOR/MACLAURIN SERIES

Whenever there is alternation, expect \((-1)^n\) or \((-1)^{n+1}\) or similar factors to be present. An increase by 2 at each step is achieved by \((2n+k)\), where \(k\) is chosen to get the first term correct. An increase by 3 would require a \((3n+k)\). With some practice it is not difficult to translate a series written longhand, but with a clear pattern, into \(\Sigma\)-notation. (For series of constants, we also used \((-1)^n = \cos(n\pi)\).)

\[\text{Exercises}\]

For Exercises 1–4, show that the \(\Sigma\)-notation for the series below (namely those in (11.27), (11.29), (11.30), and (11.32)) expands to the respective series pattern given on the left.

1. \[1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.\]
2. \[-\frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}.\]
3. \[1 + x + x^2 + x^3 + \cdots = \sum_{n=0}^{\infty} x^n.\]
4. \[(x - 1) - \frac{1}{2} (x - 1)^2 + \frac{1}{3} (x - 1)^3 - \cdots = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} (x - 1)^n}{n}.\]

5. Rewrite the power series for \(\sin x\) centered at \(a = 0\), but in such a way that it starts with \(n = 1\).

6. Do the same for \(\cos x\).

For Exercises 7–11 write each series using \(\Sigma\)-notation: first starting with \(n = 1\), and then starting with \(n = 0\).

7. \[-\frac{x^2}{2} + \frac{x^4}{3} - \frac{x^6}{4} + \cdots\]
8. \[x^2 + \frac{x^4}{4} + \frac{x^6}{9} + \frac{x^8}{16} + \frac{x^{10}}{25} + \cdots\]
9. \[-\frac{x^2}{2} - \frac{x^4}{2 \cdot 4} + \frac{x^6}{2 \cdot 4 \cdot 6} - \frac{x^8}{2 \cdot 4 \cdot 6 \cdot 8} + \cdots\]
10. \[\frac{x}{1 \cdot 1} + \frac{x^3}{3 \cdot 1 \cdot 2} + \frac{x^5}{5 \cdot 1 \cdot 2 \cdot 3} + \frac{x^7}{7 \cdot 1 \cdot 2 \cdot 3 \cdot 4} + \cdots\]
11. \[-\frac{2}{4} \cdot \frac{4x}{7} + \frac{6x^2}{10} - \frac{8x^3}{13} + \cdots\]
13. Prove that the remainder
\[R_N(x) = \frac{x^{N+1}}{(1-z)^{N+2}}\]
from (11.31) does approach zero as \(N \to \infty\) for the case \(x \in (-1, 0)\). Note that it is enough to show \(|R_N(x)| \to 0\). (Hint: In what interval is \(1-z\) in this case?)
11.4 General Power Series and Interval of Convergence

11.4.1 Definition of General Power Series

While most of our familiar functions can be written as power series, meaning form (11.1), page 747 repeated here as

\[ f(x) = \sum_{n=0}^{\infty} a_n (x - a)^n, \]  

(11.33)

there are many functions which must be written as series (for instance, \( \int e^{x^2} \, dx \)). In some sense, there are more power series than “nice” functions (usual combinations of powers, trigonometric, logarithmic and exponential functions) which also have power series representations. It is therefore interesting to study power series without reference to functions they may or may not have been derived from.

When we are given such a function represented by a power series (11.33), it is clear that \( a_0 = f(a) \), but less clear that \( a_1 = f'(a) \), or \( a_2 = \frac{1}{2!} f''(a) \), etc., which is what happens with Taylor Series where we know the function \( f \) and how to compute its derivatives. Even finding \( f'(a) \) is somewhat difficult because, as we know from the definition of the derivative,

\[ f'(a) = \lim_{\Delta x \to 0} \frac{f(a + \Delta x) - f(a)}{\Delta x}, \]

and it is perhaps an understatement to note that it is not immediately clear how to compute that limit from (11.33). Note how \( f(x) \) must be defined—and continuous—in an open interval containing \( a \) or this limit which defines the derivative cannot exist. Fortunately, given any function defined by a power series (11.33), we are guaranteed to have only certain cases for its domain. We rely on the following very useful, and eventually intuitive, result.

11.4.2 Abel’s Theorem

**Theorem 11.4.1 (Abel’s Theorem\(^{15} \))**: A power series of form (11.33) will converge at \( x = a \) only and diverge elsewhere, or converge absolutely in an open interval \( x \in (a - R, a + R) \) and diverge outside the closed interval \( [a, b] \) with the same endpoints, i.e., diverge for \( x \in (-\infty, a - R) \cup (a + R, \infty) \). If the power series also converges at an endpoint \( a - R \) or \( a + R \), it will be continuous to the endpoint from the side interior to the interval.

**Definition 11.4.1** The number \( R \) above is called the **radius of convergence** of (11.33). We say \( R = 0 \) if the power series converges at \( a \) only. It is quite possible that \( R = \infty \), in which case the power series converges on all of \( \mathbb{R} \). Otherwise, \( R \in (0, \infty) \) is nonzero and finite and the power series

\begin{enumerate}
  \item converges for \( |x - a| < R \), and
  \item diverges for \( |x - a| > R \).
\end{enumerate}

\(^{15}\)Named for Niels Henrik Abel, 1802–1829, a Norwegian mathematician most notable for founding Group Theory, on the way to proving the impossibility of solving the general fifth-degree polynomial equations by a formula with radicals, unlike second-degree (quadratic), third-degree (cubic) or fourth-degree (quartic) equations, which do have formulas for their solutions. While solving the general quadratic equation (using the “quadratic formula”) is basic enough, the third-degree and fourth-degree “formulas” are much more involved, and Abel dispelled any hope that such formulas exist for higher-degree polynomial equations. Here we are interested in his (very different) theorem on the convergence of power series.
Because of this result above it makes sense to talk of the interval of convergence of a power series. Its form will be one of the following, depending upon the specific series:

\{a, (-\infty, \infty), (a - R, a + R), [a - R, a + R], (a - R, a + R), (a - R, a + R)\}.

For the Taylor Series for \(e^x\), \(\sin x\) and \(\cos x\), we know this interval of convergence is simply \((-\infty, \infty) = \mathbb{R}\), and so we say \(R = \infty\) in those cases. In contrast, the Taylor Series for \(\ln x\) centered at \(x = 1\) converges at least in \(|x - 1| < 1\), as shown in Proposition 11.3.5, page 779.

\[
\ln x = 1(x - 1) - \frac{1}{2}(x - 1)^2 + \frac{1}{3}(x - 1)^3 - \frac{1}{4}(x - 1)^4 + \cdots = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}(x - 1)^n}{n}
\]

for (at least) \(|x - 1| < 1\), i.e., \(x \in (0, 2)\).

While Abel’s Theorem does not state whether or not a series converges at the endpoints, it is not difficult to see that the series above converges for \(x \in (0, 2]\), i.e., converges at the right endpoint \(x = 2\) (by the Alternating Series test), and diverges at the left endpoint \(x = 0\) (since there it is the harmonic series). Abel’s theorem then does say that the series will then be left-continuous at \(x = 2\), and since so is \(\ln x\), they must agree at that point. Thus the series equals \(\ln x\) on all of \(x \in (0, 2]\).

### 11.4.3 Finding the Interval and Radius of Convergence

In most cases, the Ratio and Root Tests are the tools used to find the interval of convergence for a given power series. From there we usually observe the actual radius, as it is basically half the length of the interval, or equivalently, the distance from the center \(a\) to one of the endpoints \(a \pm R\). For most cases we will use the Ratio Test.

**Example 11.4.1** Find the interval of convergence for the series \(\sum_{n=0}^{\infty} \frac{x^n}{n!}\).

**Solution:** Actually we know this series, and that it converges to \(e^x\) for all \(x \in \mathbb{R}\), so the interval is \((-\infty, \infty) = \mathbb{R}\), and thus \(R = \infty\). We deduced this from the form of the remainder \(R_N = \frac{1}{(N+1)!}e^{xN+1}\).

But how would we determine where it converges without knowing the form of the remainder? The key here is to use the Ratio Test for an arbitrary \(x\). First we write

\[
f(x) = \sum_{n=0}^{\infty} a_n(x - a)^n = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \sum_{n=0}^{\infty} u_n.
\]

Most textbooks introduce \(u_n\) above for convenience in applying the Ratio Test. (The reader should feel free to skip that step where relevant.) Next we calculate

\[
\rho = \lim_{n \to \infty} \left| \frac{u_{n+1}}{u_n} \right| = \lim_{n \to \infty} \left| \frac{x^{n+1}}{(n+1)!} \cdot \frac{1}{x^n} \right| = \lim_{n \to \infty} \frac{x^{n+1}}{n!} \cdot \frac{n!}{(n+1)!} = \lim_{n \to \infty} |x| \cdot \frac{n!}{n!} = \lim_{n \to \infty} |x| \cdot \frac{1}{n+1} = 0 \quad \text{for every } x \in \mathbb{R}.
\]

Recall that the series will converge absolutely if \(\rho < 1\), and we in fact for this case have \(\rho = 0\) for every real \(x\). Since \(\rho = 0 < 1\) regardless of \(x \in \mathbb{R}\), the series converges absolutely on all of \(\mathbb{R} = (-\infty, \infty)\), which gives the interval of convergence. (Here we take the radius to be \(R = \infty\).)
It is arguably easier to find that the series for $e^x$ converges (absolutely) for all $x$ by using the Ratio Test as above, than using the form of the remainder $R_N(x)$ and showing $R_N(x) \to 0$ as $N \to \infty$. Indeed, the Ratio Test is usually the preferred method for finding where a given power series converges.

**Example 11.4.2** Find the interval and radius of convergence for the series $\sum_{n=0}^{\infty} \frac{2^n(x-5)^n}{2n-1}$.

**Solution:** Just as above,

$$\rho = \lim_{n \to \infty} \left| \frac{u_{n+1}}{u_n} \right| = \lim_{n \to \infty} \left| \frac{\left( \frac{2^{n+1}(x-5)^{n+1}}{2(n+1)-1} \right)}{\frac{2^n(x-5)^n}{2n-1}} \right| = \lim_{n \to \infty} \frac{2^{n+1}}{2^n} \cdot \frac{2n-1}{2(n+1)-1} \cdot \frac{(x-5)^{n+1}}{(x-5)^n} = \lim_{n \to \infty} 2 \cdot \frac{2n-1}{2n+1} \cdot |x-5| = 2 \cdot |x-5|.$$

Remember that the $x$ in the line above is a constant as far as the limit goes (since the limit is in $n$). To find the region where $\rho < 1$ we simply solve

$$\frac{2|x-5|}{\rho} < 1 \iff |x-5| < \frac{1}{2} \iff -1/2 < x - 5 < 1/2 \iff 9/2 < x < 11/2.$$

Thus we know for a fact that the series converges absolutely for $x \in (9/2, 11/2)$. A similar calculation gives us divergence in $(-\infty, 9/2) \cup (11/2, \infty)$, and we usually do not bother repeating the calculations to see this. The only question left is what happens at the two boundary points.

$x = 9/2$:

$$\sum_{n=0}^{\infty} \frac{2^n(9/2-5)^n}{2n-1} = \sum_{n=0}^{\infty} \frac{2^n(-1/2)^n}{2n-1} = \sum_{n=0}^{\infty} \frac{2^n \left( \frac{1}{2} \right)^n (-1)^n}{2n-1} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n-1}.$$

The resultant series converges by the Alternating Series Test (alternates, terms shrink in absolute size monotonically to zero). Thus the series does converge at the left endpoint $x = 9/2$.

$x = 11/2$:

$$\sum_{n=0}^{\infty} \frac{2^n(11/2-5)^n}{2n-1} = \sum_{n=0}^{\infty} \frac{2^n(1/2)^n}{2n-1} = \sum_{n=0}^{\infty} \frac{2^n \left( \frac{1}{2} \right)^n}{2n-1} = \sum_{n=0}^{\infty} \frac{1}{2n-1}.$$

This series diverges (limit-comparable to the harmonic series $\sum \frac{1}{n}$). Thus the power series diverges at this endpoint.

The conclusion is that the interval of convergence is $x \in [9/2, 11/2)$.

Note that the center of the interval is $\left( \frac{9}{2} + \frac{11}{2} \right) / 2 = 10/2 = 5$, and so the “center” being at $a = 5$ (which we can also read from the original summation notation), we see that the interval extends by $1/2$ to both right and left of the center, so $R = 1/2$. We could also find this by computing half of the length of the interval, i.e., $\left( \frac{11}{2} - \frac{9}{2} \right) / 2 = (2/2)/2 = 1/2$. 


Example 11.4.3 Find the radius and interval of convergence for \( \sum_{n=1}^{\infty} (nx)^n \).

Solution: This would be a difficult series to analyze with the Ratio Test (as the reader is invited to attempt), and the Root Test seems more appropriate. Here we use \( \rho = \rho_{\text{root}} \), and get
\[
\rho = \lim_{n \to \infty} \sqrt[n]{|nx|^{1/n}} = \lim_{n \to \infty} |nx| = \left\{ \begin{array}{ll}
0 & \text{if } x = 0, \\ \infty & \text{if } x \neq 0.
\end{array} \right.
\]
Thus the series diverges for \( x \neq 0 \), and the “interval” of convergence is simply \([0, 0] = \{0\}\), and the radius is simply \( R = 0 \).

In the example above, except at \( x = 0 \) the terms all increased in size rather than shrinking to zero. In effect, \((nx)^n\) is the product of a very rapidly growing \( n^n \) with an exponentially (or “geometrically”) growing \( x^n \) if \(|x| > 1\), and exponentially shrinking \( x^n \) if \(|x| < 1\). However, even the case of the exponential shrinkage cannot overcome the rapid growth of \( n^n \), which then dominates the behavior of \( n^n x^n = (nx)^n \). Cases where \( R = 0 \) are not the most commonly studied, but they do occur and anyone dealing with series has to be aware of them.

Also notable from this latest example is that there are cases where the Root Test is preferable to the Ratio Test. In fact, as we noted when these two tests were first introduced in Section 10.5, there is even some overlap. Recall that both tests were modeled on comparisons to the Geometric Series \( \sum a_0 r^n \).

It should therefore, upon reflection, be no surprise that the Ratio and Root Tests are called upon in many cases to determine where a power series converges. After all, such series \( \sum a_n x^n \) can be interpreted to be variations of geometric series.

11.4.4 Taylor/Power Series Connection

There is a nice connection between Taylor and Power Series centered at a given point \( a \). In short, they are the same, assuming there is an interval (of “wiggle room”), around the center of the series, on which the power series converges to the function. We introduce this connection here initially for the reader to note for future reference, and then greatly expand its scope and application in Section 11.5.

To see this connection we first need the following theorem, which we state without proof:

Theorem 11.4.2 Manipulations with Power Series: Suppose we are given a function defined by a power series
\[
f(x) = \sum_{n=0}^{\infty} a_n (x - a)^n
\]
which converges in some open interval \(|x - a| < R\), where \( R > 0 \).

\[
f^{(n)}(a) = n! a_n.
\]

Note that (11.35) is equivalent to \( a_n = \frac{f^{(n)}(a)}{n!} \), so the coefficients of the power series will be exactly the same as those of the Taylor Series, assuming the power series is valid in some open interval \(|x - a| < R\), some \( R > 0 \).

In advanced calculus, functions which can be represented in \(|x - a| < R\) by a convergent power series are given a special name:

Definition 11.4.2 A function \( f(x) \) which has a power series representation (11.34) converging in some open interval \(|x - a| < R\) (for some \( R > 0 \)) is called real-analytic in that interval.
Equivalently, a function is real-analytic on an open interval $|x - a| < R$ if and only if its Taylor Series converges to the function in the same interval.

There is a very rich and beautiful theory of real-analytic functions which is beyond the scope of this text. It is a theory which has a remarkably simple extension to functions of a complex variable

$$z \in \mathbb{C} = \{x + iy : x, y \in \mathbb{R}\}, \quad i = \sqrt{-1}.$$

This may seem a complication, but the theory is often simplified by this generality, after which the real-analytic results follow from the complex theory. In fact the term radius of convergence comes from the complex-analytic theory, where the complex values $z$ for which $\sum a_n(z - a)^n$ converges lie in a disk of radius $R$ inside the complex plane $\mathbb{C}$. Such are topics for advanced calculus or complex analysis courses, usually at the senior or graduate levels. However, we will explore some aspects of the theory suitable for this level of textbook in Section 11.6.

### Exercises

For Exercises 1–13, find

(a) the interval of convergence, including endpoints where applicable, and

(b) the radius of convergence.

1. $f(x) = \sum_{n=1}^{\infty} \frac{x^n}{2^n \sqrt{n}}$.

2. $f(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$. See Example 11.3.5, page 791.

3. $\ln x = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}(x - 1)^n}{n}$. See Proposition 11.3.5, page 779.

4. $f(x) = \sum_{n=0}^{\infty} nx^n$.

5. $f(x) = \sum_{n=0}^{\infty} n! x^n$.

6. $e^{x^2} = \sum_{n=0}^{\infty} \frac{x^{2n}}{n!}$. See Proposition 11.27, page 777.

7. $f(x) = \sum_{n=2}^{\infty} \frac{(x + 1)^n}{(\ln n)^n}$.

8. $f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n^n}$.

9. $f(x) = \sum_{n=0}^{\infty} \frac{(n!)^2(x - 5)^n}{(2n)!}$.

10. $f(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{n^2 \cdot 10^n}$.

11. $\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n + 1)!}$.

12. $f(x) = \sum_{n=1}^{\infty} \frac{(-1)^n n! (x - 3)^n}{n^n}$.

13. $f(x) = \sum_{n=2}^{\infty} \frac{3^n(x + 2)^n}{\ln n}$.

14. Assume for a moment that all our work with Taylor Series can be generalized to the complex plane $\mathbb{C}$. Note that $i = i$, $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, etc. Use all this and known Maclaurin Series to prove Euler’s Identity:

$$e^{i\theta} = \cos \theta + i \sin \theta. \quad (11.36)$$

Note that this implies that $e^{i\pi} = -1$, or more interestingly $e^{i\pi} + 1 = 0$ an often-cited, beautifully compact equation relating four of the most important numbers in mathematics.
15. Use (11.36) and the facts that 
\( \sin(-\theta) = -\sin \theta \) and \( \cos(-\theta) = \cos \theta \),
to show the following relationship between trigonometric and hyperbolic 
functions (see Exercise 6, page 796):

(a) \( \cos x = \cosh(ix) \);

(b) \( \sin x = \frac{1}{i} \sinh(ix) \).

16. Use Exercises 14 and 15 to prove the 
following trigonometric identities:

(a) \( \sin^2 x + \cos^2 x = 1 \);

(b) \( \sin 2x = 2 \sin x \cos x \);

(c) \( \cos 2x = \cos^2 x - \sin^2 x \);

(d) \( \sin(x + y) = \sin x \cos y + \cos x \sin y \).
11.5 Valid Manipulations with Taylor/Power Series

Taylor Series are very robust in the sense that most algebra and calculus-based methods for constructing functions from other functions translate to series. Some care must be taken to ensure a proper interval of convergence results, but even that consideration follows fairly easily from the process.

Here we will look at both algebraic and calculus-based manipulations of Taylor Series. In so doing, it should become clear that such methods are often preferable to brute-force computations from the definition of Taylor Series. Furthermore, some functions require us to use series representations rather than previous types of formulas, and such manipulations are sometimes quite helpful in finding representations from known functions.

11.5.1 Algebraic Manipulations

We begin this subsection with a bit of theory which is mostly straightforward, and somewhat interesting, but we will be somewhat brief with it here so it will not become a distraction. The main theorem is the following:

**Theorem 11.5.1** If there are two power series representations of a function $f(x)$ which are valid within an open interval surrounding the center $a$, i.e., if there exists $\delta > 0$ such that $x \in (a - \delta, a + \delta)$ implies

$$f(x) = \sum_{k=0}^{\infty} a_k (x - a)^k = \sum_{k=0}^{\infty} b_k (x - a)^k,$$

then $a_0 = b_0$, $a_1 = b_1$, $a_2 = b_2$, and so on.

The theorem is stating that any two power series representations (including a Taylor Series) of the same function with the same center must really be the same series. In other words, any power series representation for a function is unique at each point where it is valid. From (11.35) of Theorem 11.4.2, page 785 we then also get that any valid power series representation of a function within an open interval is also its Taylor Series with the same center.

**Example 11.5.1** Use the Maclaurin Series for $e^x$ to calculate the Maclaurin Series for $e^{x^2}$.

**Solution:** We simply replace $x$ with $x^2$ in the series for $e^x$.

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

$$e^{x^2} = 1 + (x^2) + \frac{(x^2)^2}{2!} + \frac{(x^2)^3}{3!} + \frac{(x^2)^4}{4!} = \sum_{k=0}^{\infty} \frac{(x^2)^k}{k!}$$

$$\iff e^{x^2} = 1 + x^2 + \frac{x^4}{2!} + \frac{x^6}{3!} + \frac{x^8}{4!} + \cdots = \sum_{k=0}^{\infty} \frac{x^{2k}}{k!},$$

valid for all $x \in \mathbb{R}$, since the original series was valid everywhere (and $x \in \mathbb{R} \implies x^2 \in \mathbb{R}$, and can therefore be inputted to the original series for $e^x$).

A few comments are in order regarding how the theory implies the validity of the series representation for $e^{x^2}$ above. Because the series $e^x = \sum \frac{x^n}{n!}$ is true for any $x \in \mathbb{R}$, we could also
use our abstract function notation to write
\[ e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + ( ) + \frac{(x)^2}{2!} + \frac{(x)^3}{3!} + \frac{(x)^4}{4!} + \cdots, \]
and any input \(( ) \in \mathbb{R}\) on the left can be equivalently input on the right, and the values of the outputs will be the same. That should also be true of the value of the output if the input is \(x^2\):
\[ e^{x^2} = \sum_{n=0}^{\infty} \frac{(x^2)^n}{n!} = 1 + (x^2) + \frac{(x^2)^2}{2!} + \frac{(x^2)^3}{3!} + \frac{(x^2)^4}{4!} + \cdots, \]
but then we can simplify each of the terms within the summation on the right-hand side, as we did in Example 11.5.1 above. Asking if a series representation makes sense at actual values, and observing the affirmative answer, helps us to see the validity of the new series. (We will argue similarly in subsequent examples.)

Furthermore, we can also dispel any doubt that this is superior to calculating such a Taylor Series from the original definition of Taylor Series. Recall that we would need formulas for \(f^{(n)}(0)\) to compute \(f^{(n)}(0)\) to compute the Taylor Coefficients. The first two are easy enough: \(f(x) = e^{x^2}\); \(f'(x) = 2xe^{x^2}\). For \(f''\), we need a product rule and another chain rule: \(f''(x) = 2x \left(2xe^{x^2}\right) + e^{x^2} = e^{x^2} (4x + 1)\). Next we would need another product rule and a chain rule to find \(f'''\), for which simplifying would be even more difficult. By then, we would likely conclude the algebraic method above is superior. Similarly it is not difficult to compute the following:

**Example 11.5.2** Find the Maclaurin Series for \(f(x) = x^3 \sin 2x\).

**Solution:** We will construct this series in stages, beginning with the series for \(\sin x\).

\[
\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!},
\]
\[
\Rightarrow \sin 2x = (2x) - \frac{(2x)^3}{3!} + \frac{(2x)^5}{5!} - \frac{(2x)^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n (2x)^{2n+1}}{(2n+1)!},
\]
\[
= 2x - \frac{8x^3}{3!} + \frac{32x^5}{5!} - \frac{128x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n 2^{2n+1} x^{2n+1}}{(2n+1)!},
\]
\[
\Rightarrow x^3 \sin 2x = x^3 \left(2x - \frac{8x^3}{3!} + \frac{32x^5}{5!} - \frac{128x^7}{7!} + \cdots\right) = x^3 \left(\sum_{n=0}^{\infty} \frac{(-1)^n 2^{2n+1} x^{2n+1}}{(2n+1)!}\right),
\]
\[
= 2x^4 - \frac{8x^6}{3!} + \frac{32x^8}{5!} - \frac{128x^{10}}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n 2^{2n+1} x^{2n+4}}{(2n+1)!}.
\]
valid for all \(x \in \mathbb{R}\).

Again it is not difficult to see that the series should be valid at any given value for \(x \in \mathbb{R}\), since we can place any value into the series for \(\sin( )\), including the value \(2x\) (which is defined regardless of our choice of \(x \in \mathbb{R}\), simplify each term, multiply the series by another “constant” such as \(x^3\) (only \(n\) has a range of values within given sum), and get the correct value for \(x^3 \sin 2x\). Since the correct power series centered at zero should be unique, it must be the one computed above.

This example also shows how \(\Sigma\)-notation can make shorter work of some series constructions.
11.5.2 Derivatives and Integrals

As has already been mentioned, many of the manipulations we would hope we can do with Taylor Series are in fact possible. For instance, we can take derivatives and integrals as expected:

**Theorem 11.5.2** Suppose that \( f(x) \) is given by some Taylor Series

\[
f(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + a_3(x-a)^3 + \cdots = \sum_{n=0}^{\infty} a_n(x-a)^n. \tag{11.37}
\]

1. (Also a theorem of Abel.) If the series converges in an open interval containing \( x \), then inside that interval, we can differentiate “term by term” to get

\[
f'(x) = a_1 + 2a_2(x-a) + 3a_3(x-a)^2 + \cdots = \sum_{n=1}^{\infty} na_n(x-a)^{n-1}. \tag{11.38}
\]

2. Furthermore, integrating (11.37) term by term we get

\[
\int f(x) \, dx = a_0(x-a) + a_1 \frac{(x-a)^2}{2 \cdot 1!} + a_2 \frac{(x-a)^3}{3 \cdot 2!} + \cdots + C
\]

\[
= \sum_{n=0}^{\infty} a_n \frac{(x-a)^{n+1}}{(n+1)!} + C, \tag{11.39}
\]

with the special case that, if the series converges on the closed interval with endpoints \( a \) and \( x \), we have

\[
\int_a^x f(t) \, dt = \sum_{n=0}^{\infty} a_n \frac{(t-a)^{n+1}}{(n+1)!} \bigg|_a^x = \sum_{n=0}^{\infty} a_n \frac{(x-a)^{n+1}}{(n+1)!}. \tag{11.40}
\]

A very simple demonstration of the derivative part of this theorem is the following:

**Example 11.5.3** We do the following calculation \( \frac{d}{dx} e^x = e^x \), but using series to show the reasonableness of the theorem above.

\[
\frac{de^x}{dx} = \frac{d}{dx} \left( 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \right) = 0 + 1 + \frac{2x}{2 \cdot 1} + \frac{3x^2}{3 \cdot 2 \cdot 1} + \frac{4x^3}{4 \cdot 3 \cdot 2 \cdot 1} + \cdots
\]

\[
= 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots
\]

\[
= e^x,
\]

as expected. Using \( \Sigma \)-notation, keeping in mind that the first \( n = 0 \) term differentiates to zero, we get

\[
\frac{d}{dx} \left( \sum_{n=0}^{\infty} \frac{1}{n!} x^n \right) = \sum_{n=0}^{\infty} \frac{1}{n!} \cdot n x^{n-1} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} x^{n-1} = \sum_{n=0}^{\infty} \frac{1}{n!} x^n = e^x.
\]

The step where we rewrite the new series to begin at \( n = 0 \) is clear if a few terms are written out in the expansions of each series.
The series for \( \frac{1}{1-x} \) was given in (11.30), page 778, but was shown easily remembered due to its relationship with a simple geometric series (see also (10.9), page 705):
\[
\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + x^4 + \cdots = \frac{1}{1-x},
\]
valid for \(|x| < 1\). We will use this in some examples below.

**Example 11.5.4** Compute the series for \( \frac{1}{(1-x)^2} = \frac{d}{dx} \left[ \frac{1}{1-x} \right] \) centered at \( a = 0 \).

**Solution:** This is a straightforward computation, either with the term-by-term expansion or with the \( \Sigma \)-notation, and an optional rewriting of the final summation. (Note how the first term vanishes in the derivative.)
\[
\frac{1}{(1-x)^2} = \frac{d}{dx} \left[ 1 + x + x^2 + x^3 + x^4 + \cdots \right] = 1 + 2x + 3x^2 + 4x^3 + 5x^4 + \cdots,
\]
\[
\frac{1}{(1-x)^2} = \frac{d}{dx} \left[ \sum_{n=0}^{\infty} x^n \right] = \sum_{n=1}^{\infty} (nx^{n-1}) = \sum_{n=0}^{\infty} (n+1)x^n.
\]

Since the original series was valid for \(|x| < 1\), so will be the new series. (The reader is welcome to perform a ratio test to confirm this.)

**Example 11.5.5** Use the series for \( \frac{1}{1-x} \) to derive a series for \( \frac{1}{1+x^2} \). Then use that series to find a series for \( \tan^{-1} x \).

**Solution:** We first replace \( x \) with \( -x^2 \) in that series, since \( \frac{1}{1+x^2} = \frac{1}{1-(-x^2)} \):
\[
\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \cdots = \sum_{n=0}^{\infty} x^n \quad \text{(valid for } |x| < 1) \]
\[
\Rightarrow \quad \frac{1}{1 + x^2} = \frac{1}{1 - (-x^2)} = 1 + (-x^2) + (-x^2)^2 + (-x^2)^3 + (-x^2)^4 + \cdots = \sum_{n=0}^{\infty} (-x^2)^n
\]
\[
= 1 - x^2 + x^4 - x^6 + x^8 + \cdots = \sum_{n=0}^{\infty} (-1)^n x^{2n}.
\]

This is valid wherever \(|x^2| < 1\), which it is not too difficult to see is again wherever \(|x| < 1\).\(^{16}\)

Next we use the fact that \( \tan^{-1} 0 = 0 \), so that
\[
\tan^{-1} x = \tan^{-1} x - \tan^{-1} 0 = \int_0^x \frac{1}{1+t^2} dt
\]
\[
= \int_0^x \sum_{n=0}^{\infty} (-1)^n t^{2n} dt = \sum_{n=0}^{\infty} (-1)^n \frac{t^{2n+1}}{2n+1} \bigg|_0^x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} - 0.
\]

\(^{16}\)Recall that \(|x^2| = |x|^2\). Also recall that the square root function is increasing on \([0, \infty)\), and so (by definition) preserves inequalities. Thus
\[
|x|^2 < 1 \iff \sqrt{|x^2|} < \sqrt{1} \iff |x| < 1.
\]
The antiderivative of \( x \) series are left-continuous as \( x \) converges by the Alternating Series Test at \( x = 1 \), and the series is continuous where it converges by Abel’s Theorem.

Thus
\[
\tan^{-1} x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n + 1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots. \tag{11.41}
\]

Alternatively, the final form in (11.41) can be had by the more expanded form:
\[
\tan^{-1} x = \int_0^x \frac{1}{1 + t^2} \, dt = \int_0^x \left[ 1 - t^2 + t^4 - t^6 + \cdots \right] \, dt \\
= \left[ t - \frac{1}{3} t^3 + \frac{1}{5} t^5 - \frac{1}{7} t^7 + \frac{1}{9} t^9 - \cdots \right] \bigg|_0^x = x - \frac{1}{3} x^3 + \frac{1}{5} x^5 - \frac{1}{7} x^7 + \frac{1}{9} x^9 - \cdots.
\]

Once again, this is valid where \( |x^2| < 1 \), i.e., where \( |x| < 1 \). However, we see that the series converges by the Alternating Series Test at \( x = 1 \), and so the interval of convergence is in fact \( x \in (-1, 1] \). We know that the series equals \( \tan^{-1} x \) even at \( x = 1 \) because both \( \tan^{-1} x \) and the series are left-continuous as \( x \to 1^− \), the former due to the fact \( \tan^{-1} x \) is continuous for \( x \in \mathbb{R} \), and the series is continuous where it converges by Abel’s Theorem.

In fact one valid, if not terribly efficient, method of computing \( \pi \) is from using
\[
\pi = 4 \cdot \frac{\pi}{4} = 4 \tan^{-1}(1) = 4 \sum_{n=0}^{\infty} \frac{(-1)^n}{2n + 1} = 4 \left( 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots \right). \tag{11.42}
\]

**Example 11.5.6** Find \( \int_0^x e^{t^2} \, dt \).

**Solution:** It is an interesting but futile exercise to try to find the antiderivatives of \( e^{x^2} \) using the usual tricks: substitution, integration by parts, etc. It is well-known that there is no “closed form” for this antiderivative, i.e., using the usual functions in the usual manners. It is also true that, since \( e^{x^2} \) is continuous on \( \mathbb{R} \), there must exist continuous antiderivatives.\(^{17}\) Our discussion here presents a strategy for calculating this integral: writing the integrand as a series, and integrating term by term. As before, we will write the steps and the solution in two ways: one method is to write out several terms of the series and declare a pattern; the other, done simultaneously, is to use the \( \Sigma \)-notation. Hopefully by now they are equally simple to deal with.

\[
e^t = 1 + \frac{t}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{t^n}{n!}
\]

\[
\Rightarrow e^{t^2} = 1 + \frac{t^2}{1!} + \frac{(t^2)^2}{2!} + \frac{(t^2)^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{(t^2)^n}{n!}
\]

\[
\Rightarrow e^t = 1 + \frac{t^4}{1!} + \frac{t^6}{2!} + \frac{t^8}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{t^{2n}}{n!}
\]

\[
\Rightarrow \int_0^x e^{t^2} \, dt = \int_0^x \left( 1 + \frac{t^2}{1!} + \frac{t^4}{2!} + \frac{t^6}{3!} + \cdots \right) \, dt = \int_0^x \left( \sum_{n=0}^{\infty} \frac{t^{2n}}{n!} \right) \, dt
\]

\[
= \left( t + \frac{t^3}{3 \cdot 1!} + \frac{t^5}{5 \cdot 2!} + \frac{t^7}{7 \cdot 3!} + \cdots \right) \bigg|_0^x = \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n + 1)n!} \bigg|_0^x
\]

\[
= \left( x + \frac{x^3}{3 \cdot 1!} + \frac{x^5}{5 \cdot 2!} + \frac{x^7}{7 \cdot 3!} + \cdots \right) = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n + 1)n!}.
\]

\(^{17}\)This comes from one of the statements of the Fundamental Theorem of Calculus.
Thus
\[ \int_0^\infty e^{t^2} dt = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)n!}. \]

We could also write the general antiderivative
\[ \int e^{x^2} dx = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)n!} + C. \]

Other antiderivatives which must be found this way are \( \int \sin x^2 dx \), \( \int \cos x^2 dx \).

### 11.5.3 The Binomial Series and an Application

The following series comes up in enough applications that it is worth some focus. It is the following:
\[
(1 + x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha - 1)x^2}{2!} + \frac{\alpha(\alpha - 1)(\alpha - 2)x^3}{3!} + \cdots \tag{11.43}
\]

This series (11.43) is known as the Binomial Series. It can also be written
\[
(1 + x)^\alpha = \sum_{n=0}^{\infty} \frac{\alpha(\alpha - 1) \cdots (\alpha - n + 1)x^n}{n!}.
\]

This series is valid for \(|x| < 1\), and sometimes also valid at one or both endpoints \(x = \pm 1\). It is not difficult to prove, and is a worthwhile exercise. In fact, for \(\alpha \in \{0, 1, 2, 3, \cdots\}\), the function is a polynomial and the series terminates (in the sense that all but finitely many terms are zero), simply giving an expansion of the polynomial, valid for all \(x\).

The derivation of (11.43) is straightforward. See Exercise 23. Here are some quick examples:

\[
\frac{1}{\sqrt{1 + x}} = 1 - \frac{1}{2} x + \frac{(-\frac{1}{2})(-\frac{3}{2})x^2}{2!} + \frac{(-\frac{1}{2})(-\frac{3}{2})(-\frac{5}{2})x^3}{3!} + \cdots \quad (\alpha = -\frac{1}{2})
\]

\[
\frac{1}{1 + x^2} = 1 - x^2 + \frac{(-1)(-2)x^4}{2!} + \frac{(-1)(-2)(-3)x^6}{3!} + \cdots \quad (\alpha = -1)
\]

\[
(1 + x)^3 = 1 + 3x + \frac{3 \cdot 2x^2}{2!} + \frac{3 \cdot 2 \cdot 1 x^3}{3!} + \frac{3 \cdot 2 \cdot 1 \cdot 0 x^4}{4!} + \frac{3 \cdot 2 \cdot 1 \cdot 0 \cdot (-1)x^5}{5!} + \cdots \quad (\alpha = 3)
\]

Actually, the last one is valid for all \(x\), and the one above it was found as a step in Example 11.5.5, page 791. Other algebraic manipulations can also sometimes put a function into a form suitable for applying the Binomial Series. Consider the following, with \(\alpha = 1/3\), we complete the square under the radical, and the natural center of the resulting series is \(a = -1\). (Note also that \(\sqrt[3]{-A} = \sqrt[3]{-1} \sqrt[3]{A} = -\sqrt[3]{A}, \) since 3 is odd.)

\[
\sqrt[3]{x^2 + 2x} = \sqrt{\frac{x^2 + 2x + 1}{x + 1} - 1} = \sqrt{(x + 1)^2 - 1} = -\sqrt{1 - (x + 1)^2}
\]

\[
= - \left[ 1 + \frac{1}{3}(x + 1)^2 + \frac{1}{3}(-2)(x + 1)^4 + \frac{1}{3}(-2)(-5)(x + 1)^6 + \cdots \right]
\]
Similarly, each of the following manipulations are valid though they yield different intervals of convergence:

\[
(3 - 8x)^{1/4} = (1 - (8x - 2))^{1/4} = 1 + \frac{\frac{1}{2}(-8(x - \frac{1}{2}))}{1!} + \frac{\frac{1}{2} \cdot \frac{3}{4}(-8(x - \frac{1}{2}))^2}{2!} + \frac{\frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6}(-8(x - \frac{1}{2}))^3}{3!} + \cdots,
\]

\[
(3 - 8x)^{1/4} = 3^{1/4} \left(1 - \frac{8}{3}\right)^{1/4} = 3^{1/4} \left[1 + \frac{\frac{1}{2}(\frac{-8}{3})}{1!} + \frac{\frac{1}{2} \cdot \frac{3}{4}(\frac{-8}{3})^2}{2!} + \frac{\frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6}(\frac{-8}{3})^3}{3!} + \cdots \right].
\]

The first representation is centered at \(x = 1/4\), and definitely valid where \(|8x - 2| \leq 1\), i.e., \(x \in (1/8, 3/8)\), while the second is centered at \(x = 0\) and definitely valid where \(|8x/3| < 1\), i.e., \(x \in (-3/8, 3/8)\).

While binomial series tend to be complicated to write, there are elegant applications. One particularly beautiful application relates Albert Einstein’s Special Relativity to Newtonian Mechanics. This application is given in the following example.

**Example 11.5.7 (Application)** According to Einstein, kinetic energy is that energy which is due to the motion of an object, and can be defined as \(E_k = E_{\text{total}} - E_{\text{rest}}\), this being a function of velocity for a given mass \(m\):

\[
E_k(v) = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} - mc^2 = mc^2 \left(1 - \frac{v^2}{c^2}\right)^{-1/2} - mc^2.
\]

Contained in the above is the very famous equation \(E_{\text{rest}} = mc^2\). Also notice that the total energy \(E_{\text{total}}\) blows up as \(v \to c^{-}\) or \(v \to -c^{-}\), i.e., as velocity approaches the speed of light. At \(v = \pm c\), we are dividing by zero in the total energy, and thus the theory that ordinary objects cannot achieve the speed of light (for it would take infinite energy to achieve it).

Now let us expand this expression of \(E_k(v)\) by applying the Binomial Series to \(\left(1 - \frac{v^2}{c^2}\right)^{-1/2}\), with \(\alpha = -1/2\) and replacing \(x\) with \(-v^2/c^2\). Thus \(E_k = mc^2 \left[(1 - v^2/c^2)^{-1/2} - 1\right]\) becomes

\[
E_k(v) = mc^2 \left(1 - \frac{1}{2} \left(-\frac{v^2}{c^2}\right) + \frac{(-\frac{1}{2}) \left(-\frac{3}{2}\right) \left(-\frac{v^2}{c^2}\right)^2}{2!} + \cdots \right) - mc^2
\]

\[
\approx mc^2 \left(1 + \frac{1}{2} \frac{v^2}{c^2}\right) - mc^2 \quad \text{when } \frac{v^2}{c^2} \text{ is small.} \tag{11.45}
\]

Multiplying this out, we see that

\[
E_k \approx mc^2 + mc^2 \cdot \frac{1}{2} \frac{v^2}{c^2} - mc^2 = \frac{1}{2} mv^2. \tag{11.46}
\]

**Summarizing,**

\[
E_k(v) \approx \frac{1}{2} mv^2 \quad \text{when } |v| << c. \tag{11.47}
\]
Here the notation $|v| < c$ means that $|v|$ is much smaller than $c$, giving us that $v^2/c^2$ is very small. So we see that Newton’s kinetic energy formula $E_k = \frac{1}{2}mv^2$ is just an approximation of Einstein’s, which is to be expected since Newton was not considering objects at such high speeds. In effect, Newton could not see the whole kinetic energy curve, where Einstein’s theories could detect more phenomena which governed the behavior of the curve of $E_k$ versus $v$ through a larger range of velocities $v$.

Exercises

The following are very useful exercises for students to attempt themselves. One should first attempt these using the written out expansion

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots,$$

and then using the $\Sigma$-notation if possible, comparing the results.

1. Use the series for $e^x$ to find a series expansion for the general antiderivative of $e^{2x}$. (You can try to find the antiderivative using non-series methods, but it cannot be written using the usual functions. It is interesting to attempt to use the old methods, to see why they fail.)

2. Use the Maclaurin series for $\sin x$ to do the following:
   (a) Write a series for $\sin 2x$.
   (b) Use the series above to prove that $\frac{d}{dx} \sin 2x = 2 \cos 2x$. (It may help to also write the series for $2 \cos 2x$ separately.)
   (c) Write a series for $\cos x^2$.
   (d) Use the series above, and the series for $-2x \sin x^2$, to prove that $\frac{d}{dx} \cos x^2 = -2x \sin x^2$.

3. Use the Maclaurin Series for $\sin x$ and $\cos x$ to show that
   $$\sin(-x) = -\sin x,$$
   $$\cos(-x) = \cos x.$$

   In each of the following, unless otherwise stated, leave your final answers in $\Sigma$-notation.

4. Find the Maclaurin series for $f(x) = \ln(x + 1)$ using (11.32). Where is this series valid?

5. Approximate $\int_0^{\sqrt{\pi}} \cos x^2 \, dx$ by computing the first five nonzero terms of the Maclaurin series for $\int \cos x^2 \, dx$.

6. The Hyperbolic Functions: The three most important hyperbolic functions are

   $$\sinh x = \frac{e^x - e^{-x}}{2} \quad (11.48)$$
   $$\cosh x = \frac{e^x + e^{-x}}{2} \quad (11.49)$$
   $$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (11.50)$$

   Though not immediately obvious, it is true that $\tanh x$ is invertible, and that its inverse has the property that

   $$\frac{d}{dx} \tanh^{-1} x = \frac{1}{1 - x^2}. \quad (11.51)$$

   Find the Maclaurin series for $f(x) = \tanh^{-1} x$ given that

   $$\tanh^{-1} x = \int_0^x \frac{1}{1 - t^2} \, dt. \quad (11.52)$$

   (See Example 11.5.5, page 791.) Where is this series valid? (Actually the integral in (11.52) can also be computed
with partial fractions, and the final answer written without resorting to series.)

7. (Proof of Proposition 11.3.5) Derive the Taylor Series for \( \ln x \) with \( a = 1 \) using the fact that

\[
\ln x = \int_1^x \frac{1}{t} \, dt
\]

for \( x > 0 \), and

\[
\frac{1}{t} = \frac{1}{1 - (1 - t)}.
\]

Where is this series guaranteed valid?

8. Evaluate \( \lim_{x \to 0} \frac{\sin x}{x} \) using a Taylor Series centered at \( a = 0 \).

9. Evaluate \( \lim_{x \to 0} \frac{1 - \cos x}{x^2} \) using a Taylor Series centered at \( a = 0 \).

10. Do the same for \( \lim_{x \to 0} \frac{1 - \cos x}{x^2} \).

11. Evaluate the integral \( \int_{0.4}^{0.4} \cos x^2 \, dx \) by using the Taylor Polynomial \( P_3(x) \) for \( \cos x \) centered at \( a = 0 \) (and therefore \( P_6(x) \) for \( \cos x^2 \)). This is called a Fresnel integral, which appears in studies of optics.

12. In (11.42), page 792 we see that

\[
\pi = 4 \left[ 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots \right].
\]

How many terms should we add in the series above to be assured that our sum is

(a) within 0.01 of \( \pi \)?

(b) within 0.00001 of \( \pi \)?

13. Use Maclaurin series for \( \sin x \) and \( \cos x \) to demonstrate the following:

(a) \( \frac{d}{dx} \sin x = \cos x \).

(b) \( \frac{d}{dx} \cos x = -\sin x \).

14. Use the fact that \( 1 + x + x^2 + x^3 + \cdots = \frac{1}{1 - x} \), and that \( \frac{d}{dx} \left( \frac{1}{1 - x} \right) = \frac{1}{(1 - x)^2} \) to find the Maclaurin Series expansion for

\[
f(x) = \frac{1}{(1 - x)^2}.
\]

15. Use the facts that \( \tan^{-1} x = \int_0^x \frac{1}{1 + t^2} \, dt \), and that \( \frac{1}{1 + t^2} = \frac{1}{1 - (-1t)} \) to compute the Maclaurin Series for \( \tan^{-1} x \).

16. Show that \( \frac{1}{e} = 1 - 1 + \frac{1}{3} - \frac{1}{5} + \frac{1}{7} \) by using the series for \( e^x \) centered at \( x = 0 \).

17. Starting from the series for \( e^x \), compute the Taylor Series for

(a) \( \sinh x = \frac{e^x - e^{-x}}{2} \)

(b) \( \cosh x = \frac{e^x + e^{-x}}{2} \)

For Exercises 18–22, approximate the definite integrals \( \int_a^b f(x) \, dx \) by replacing \( f(x) \) with an appropriate Taylor Polynomial \( P_n(x) \), centered at \( a = 0 \) if possible, and centered elsewhere if necessary. Also, compare your approximation to the exact value for each integral.

18. \( \int_0^{\pi/4} \sin x \, dx \)

19. \( \int_1^2 e^x \, dx \)

20. \( \int_1^3 \ln x \, dx \)

21. \( \int_1^3 \sqrt{1 + x} \, dx \)

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

23. Derive the series (11.43) using the formula for Taylor/Maclaurin Series where \( f(x) = (1 + x)^\alpha \) and \( \alpha = 0 \).
24. Find a series representation for the following functions using the binomial series (11.43). Do not attempt to use \( \Sigma \)-notation, but rather write out the first five terms of the series to establish the pattern.

(a) \( f(x) = (1 + x)^{3/2} \)
(b) \( f(x) = (1 - x)^{3/2} \)
(c) \( f(x) = \frac{1}{\sqrt{1 + x}} \)
(d) \( f(x) = \frac{1}{\sqrt{1 + x^2}} \)
(e) \( f(x) = \frac{x^3}{\sqrt{1 + x}} \)
(f) \( f(x) = \frac{1}{\sqrt{1 - x^2}}. \)

25. Find the series expansion for \( f(x) = \ln(1 + x^3) \) by using the fact that \( \ln(x^{2t}) = \int_0^x \frac{2t}{1 + t^2} dt. \)

26. Find a more general form of the binomial series by using (11.43) to derive a series for \( f(x) = (b + x)^\alpha \) (11.53) and determine for what values of \( x \) is it valid. (Hint: Use (11.43) after factoring out \( b^{\alpha} \) from \( f. \))

27. Complete the square and use the binomial series to write a series expansion for the following. Also determine an interval \( |x - a| < R \) where the series is guaranteed to be valid.

(a) \( f(x) = \frac{1}{\sqrt{x^2 - 6x + 10}} \)
(b) \( f(x) = \sqrt{4x^2 + 12x + 13} \)
(c) \( f(x) = (-2x^2 + 3x + 5)^{-2/3} \)

28. Using (11.44), page 794 to show that \( E_k(v) \geq \frac{1}{2}mv^2 \) for \( |v| < c \), with equality only occurring when \( v = 0 \). Thus (11.47) is always an underestimation unless \( v = 0 \). (Hint: Look at the signs of all the terms we ignore in the approximation.)

29. Approximate \( \int_0^1 \sqrt{1 + x^3} \, dx \) by using the the Binomial Series expansion for \( \sqrt{1 + x^3} = (1 + x^3)^{1/2} \), and using the first three nonzero terms of this expansion in your integral.

30. Consider \( f(x) = e^{x^2} \).

(a) Write the Maclaurin series for \( f(x) \).
(b) Find \( f^{(9)}(0) \).
(c) Find \( f^{(10)}(0) \).

31. It can be shown that
\[
\frac{\pi}{4} = \tan^{-1} \left( \frac{1}{3} \right) + \tan^{-1} \left( \frac{1}{2} \right).
\]
Use this fact to approximate \( \pi \) by using the Taylor Series for \( \tan^{-1} x \) centered at \( a = 0 \) and the approximation \( P_3(x) \).

32. Use the fact that
\[
\tan(\alpha + \beta) = \frac{\tan \alpha + \tan \beta}{1 - \tan \alpha \tan \beta}
\]
to prove the assertion at the beginning of the previous exercise.

33. As in Exercise 31, estimate \( \pi \) by using the fact that
\[
\frac{\pi}{4} = 4 \tan^{-1} \left( \frac{1}{5} \right) - \tan^{-1} \left( \frac{1}{239} \right).
\]

34. As in Exercise 31, estimate \( \pi \) by using the fact that
\[
\frac{\pi}{4} = \tan^{-1} \left( \frac{1}{2} \right) + \tan^{-1} \left( \frac{1}{5} \right) + \tan^{-1} \left( \frac{1}{8} \right).
\]

35. Write the Maclaurin series for \( f(x) = \frac{1}{2} \sin 2x \) by

(a) using the series for \( \sin x \).
(b) using instead the series for \( \sin x \) and \( \cos x \) and the fact (from the double angle formula) that \( f(x) = \sin x \cos x \).

(Just write out the first several terms of the product, being careful to distribute correctly, to verify the answer is the same as in part (a).)
11.6 Complications and the Role of Complex Numbers

A common engineering and science research technique is to assume there is a function which describes some relationship between two variables, and that the function has a Taylor Series representation. Then the researcher might look at data and attempt to find the best fitting polynomial of some specified degree that fits the data. Some limit on the degree has to be specified, since anytime we have $N$ data of the form $\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, where none of the $x_i$ are repeated with different $y_i$, we can fit an $(N - 1)$-degree polynomial to that data perfectly, but its predictive abilities may be little or nonexistent, since high-degree polynomials tend to have rather violent behavior, particularly as $|x| \to \infty$ or even between the data points.

Moreover, Taylor Series not only assume that all derivatives exist at the center, but by Abel’s Theorem and our ability to differentiate series, we expect the function in question to have all of its derivatives inside the interval of convergence. The contrapositive of that fact gives us that once we run into a problem with the function or one of its derivatives as we move from the center of the series, we cannot move any farther from the center and expect the series to be valid for the function. The upshot is that the researcher who assumes a Taylor Series expansion of a function must be careful to only use that assumption within intervals where the function and its derivatives should all be defined. Attempting to “fit” data to polynomials beyond that will likely have little or no predictive value.

We will first look at some cases where we can expect problems with our Taylor Series, in the sense that we cannot expect the given function to be equal to a Taylor Series. Most of those cases will upon reflection become pretty obvious, but some are more subtle.

As we have seen already, a Ratio Test can often give us the open part of an interval of convergence (with the endpoints usually checked separately), though we were able to avoid the Ratio Test for some of our series derived from, say, the geometric series. From the discussion above (further developed below) we can also see problems with assuming a valid Taylor Series when functions run into other difficulties, which a Ratio Test will not necessarily detect (the series may converge but not to the function). We will explore this in Subsection 11.6.1.

It turns out that the most natural place for series to “live” and be observed is not so much the real line $\mathbb{R}$ and its intervals, but the complex plane

$$\mathbb{C} = \left\{ x + iy \mid x, y \in \mathbb{R}, i = \sqrt{-1} \right\}$$

and its open discs, an open disc meaning the interior of a circle (not including the circle itself). This allows “wiggle room” in all directions from the center, which allows for things such as derivatives, where in $\mathbb{R}$ we only require “wiggle room” to the left and right. In fact it is from this complex context that the term radius of convergence comes to us. We will look into this further in Subsection 11.6.3. That discussion usually waits until students finish a 2–3 semester calculus sequence and proceed to a Differential Equations course, but it is included here to give some more context to Taylor Series.

11.6.1 Troubles stemming from continuity problems

Most of our familiar functions are analytic where they are defined, and so can be represented by Taylor Series for usefully large intervals. These functions include all polynomials, rational functions, exponentials, roots, logarithms and trigonometric functions, as well as combinations of these through addition, subtraction, multiplication, division and composition. We already mentioned that there are power series which are perfectly respectable functions, but which cannot be written as combinations of familiar functions. This may leave the student with the
11.6. COMPLICATIONS AND THE ROLE OF COMPLEX NUMBERS

incorrect impression that we can always find and manipulate Taylor Series for all functions with impunity.

However, there are many functions we encountered in Chapter 3 which had more pathological behaviors and will not always be analytic where defined. Therefore Taylor Series are often useless and inappropriate in dealing with such functions, at least if we wish to center a series at a problematic point of the function, or assume a series will be valid as we allow the input variable, say \( x \), have its value move “through” a problematic value.

The purpose of this section is to alert the student to situations in which Taylor Series—or even Taylor Polynomials—are not appropriate for approximation except possibly with careful modifications.

**Example 11.6.1** Consider \( f(x) = |x| \), which we graph below:

![Graph of \( f(x) = |x| \)]

The following problems arise with attempting to use a Taylor Series representation for \( f(x) = |x| \):

- **If we attempt to construct Taylor Polynomials at \( a = 0 \), we would have to stop at \( P_0(x) = 0 \) because there are no derivatives to compute at \( a = 0 \). Furthermore, \( P_0(x) \) is clearly a terrible approximation of \( f(x) \) as we stray from its center \( x = 0 \).**

- **If we were to construct a Taylor Series for \( f(x) \) at, say, \( a = 1 \) we would find that the series would terminate after the first-order term, because except at \( x = 0 \), locally this function is a line. Consider for instance the Taylor Series centered at \( a = 1 \), where we have, for \( N \geq 2 \),

\[
P_N(x) = f(1) + f'(1)(x - 1) + \frac{1}{2!} f''(1)(x - 1)^2 + \frac{1}{3!} f'''(1)(x - 1)^3 + \cdots + \frac{1}{N!} (x - 1)^N
\]

\[
= 1 + 1(x - 1) + 0 + 0 + 0 + \cdots + 0
\]

\[
= x.
\]

We would get the same series (letting \( N \to \infty \)) for any other center \( a > 0 \), which a direct computation would show. Furthermore, such a series would **not** be the same as the function for \( x < 0 \), since \( f(x) \neq x \) when \( x < 0 \).

- **Similarly for \( a = -1 \) we would have \( N \geq 2 \) \( \implies \)**

\[
P_N(x) = f(-1) + f'(-1)(x + 1) + \frac{1}{2!} f''(-1)(x + 1)^2 + \frac{1}{3!} f'''(-1)(x + 1)^3 + \cdots + \frac{1}{N!} (x + 1)^N
\]

\[
= -1 - 1(x + 1) + 0 + 0 + 0 + \cdots + 0
\]

\[
= -x.
\]

This is equal to \( f(x) \) for \( x \leq 0 \) but is incorrect for \( x > 0 \).
What ruins the above series’ chances of being the same as the function on all of \( \mathbb{R} \) is the fact that the absolute value function is not differentiable at \( x = 0 \). Anywhere else we can have a Taylor Series equal to the function locally, but not globally.

The coefficients of the Taylor Series follow from the local behavior of the function, not its global behavior. On \( x \in (0, \infty) \) we have \( f(x) = x \), with its obvious Taylor Series (which simplifies to just \( x \)), while on \( x \in (-\infty, 0) \) with its obvious Taylor Series (which simplifies to just \( -x \)). However, neither of these Taylor Series can equal the function on the other side of \( x = 0 \): a Taylor Series centered at \( a > 0 \) will be incorrect for \( x < 0 \), and a Taylor Series centered at \( a < 0 \) will be incorrect for \( x > 0 \).

While we can see the “kink” at \( x = 0 \) in the graph for \( f(x) = |x| \), which causes a major discontinuity in the derivatives there, sometimes the problem is more subtle, from the graphical perspective. It might not be so subtle from the functional definition perspective: piece-wise defined functions are often suspect. Recall that \( |x| \) is defined to be \( x \) on \([0, \infty)\) and \(-x\) on \((-\infty, 0)\).

**Example 11.6.2** Consider the function \( f(x) = \begin{cases} x^2 & \text{if } x \geq 0 \\ -x^2 & \text{if } x < 0 \end{cases} \) with graph and derivatives

\[
\begin{align*}
 f'(x) & = \begin{cases} 2x & \text{if } x \geq 0 \\ -2x & \text{if } x < 0 \end{cases} \\
 f''(x) & = \begin{cases} 2 & \text{if } x > 0 \\ -2 & \text{if } x < 0 \end{cases}
\end{align*}
\]

This function has similar complications as the previous, except they emerge in the next-higher-order Taylor Polynomials:

- At \( a = 0 \) we can construct \( P_1(x) = 0 \) (zero height and slope) but we cannot construct \( P_2(x) \) or higher because \( f''(0) \) does not exist.
- For a (positive) center \( a > 0 \) we can construct even the full Taylor Series, which will simplify to \( x^2 \), but not be equal to the function for \( x < 0 \).
- For a (negative) center \( a < 0 \) we can construct the full Taylor Series, which will simplify to \(-x^2\), but will not equal the function for \( x > 0 \).

**Example 11.6.3** Consider the function

\[
 f(x) = \begin{cases} e^{-1/x^2}, & \text{for } x \neq 0 \\ 0 & \text{for } x = 0. \end{cases}
\]

Clearly, as \( x \to 0 \) we have form \( e^{-\infty} \) and so \( x \to 0 \implies f(x) \to 0 \), and since \( f(0) = 0 \) we have continuity at \( x = 0 \). The function is also symmetric with respect to the y-axis. It is notable that \( f(x) \to 0 \) somewhat quickly as \( x \to 0 \) because of the growth in \( 1/x^2 \), and thus negative growth in \(-1/x^2\). Indeed the function is graphed below. Though it appears “flat” it is only zero at \( x = 0 \), which would take a much higher resolution graphic to verify.
It is an interesting exercise to show that all derivatives of \( f(x) \) exist everywhere, including at \( x = 0 \). Furthermore, after computing a few derivatives and some of the ensuing limits one can show that \((\forall n \in \{0, 1, 2, 3, \ldots\}) f^{(n)}(0) = 0\). Thus the Maclaurin Series for \( f(x) \) would be itself zero, as in
\[
\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} \cdot x^n = \sum_{n=0}^{\infty} 0x^n = 0,
\]
and so while this series would be equal to \( f(x) \) at \( x = 0 \), it would not be elsewhere, as \( f(x) = 0 \iff x = 0 \). So even though all derivatives exist for \( f \) at \( x = 0 \), and the function and its derivatives are all continuous on \( x \in \mathbb{R} \), the Taylor Series centered at \( a = 0 \) does not converge to the function except at the center of the series.

Such a function as in Example 11.6.3 above is certainly smooth at \( x = 0 \) and indeed all of \( \mathbb{R} \), as are all of its derivatives, but it is not real-analytic at \( x = 0 \) because it cannot be represented as a power series in an open interval containing \( x = 0 \). (Neither were the functions in the previous Examples 11.6.1, 11.6.2; see Definition 11.4.2, page 785.)

In fact the Taylor Series centered at any \( a \in \mathbb{R} \) would only be guaranteed to converge to the function at \( |x - a| < |a| \), because it could not extend to “the other side of zero,” and we know that it must converge within a certain “radius” of the center, and diverge once past that radius from the center. In the next subsection we will see that what is crucial is what happens inside the complex plane, where the term “radius of convergence” makes more sense.

### 11.6.2 The Complex Plane

Here we will look very briefly at the complex plane, which is the geometric interpretation of complex numbers \( z = x + iy \), where \( x, y \in \mathbb{R} \), and \( i = \sqrt{-1} \). We would call \( x \) the real part of \( z \), and \( iy \) to be the imaginary part of \( z \).

At first this seems preposterous because clearly \( \sqrt{-1} \notin \mathbb{R} \), since the square of any real number will not be negative. While it may seem easy to dismiss any number with an “imaginary” part \( iy \) as being a figment of the imagination and of no actual consequence, there nonetheless are many important physical phenomena best described using complex numbers, as their geometric properties (which we develop below) have many real-world analogs. Furthermore, complicated “real-number” phenomena are often most easily analyzed by lifting them into the complex plane, making observations there, and bringing these observations back into the real line.\(^{18}\)

So if we take as given that there is a number system which includes all the real numbers, but also a quantity \( i = \sqrt{-1} \), we get the following multiplication facts:

\[
\begin{align*}
    i^1 &= i \\
    i^2 &= -1 \\
    i^3 &= -i \\
    i^4 &= 1 \\
\end{align*}
\]

\[
\begin{align*}
    i^{4n+1} &= i \\
    i^{4n+2} &= -1 \\
    i^{4n+3} &= -i \\
    i^{4n+4} &= 1
\end{align*}
\]

\(^{18}\)It is akin to giving someone lost in a wilderness an aerial map, or a brief lift in a helicopter, so that they can glimpse their predicament from above. This could indeed be useful in finding a path out of the wilderness, even if the actual solution is still to be taken at ground level.
In fact the above pattern follows for \( n \in \mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\} \), but that requires a discussion of division for the negative exponents. Before discussing division, one has to first discuss multiplication, which has its own complications. Assuming in the discussion below that \( x_1, x_2, y_1, y_2 \in \mathbb{R} \), we have

\[
\begin{align*}
  z_1 &= x_1 + iy_1 \\
  z_2 &= x_2 + iy_2 \\
  \Rightarrow \\
  z_1z_2 &= (x_1 + iy_1)(x_2 + iy_2) \\
  &= (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1).
\end{align*}
\]

We see a kind of intermingling of the real and imaginary parts of \( z_1 \) and \( z_2 \) to form the real and imaginary parts of the product \( z_1z_2 \). While that may appear quite complicated and esoteric, in fact there is a geometric interpretation which is not all that difficult. For instance, multiplying by the \textit{imaginary unit} \( i \) has the same effect as a \( \pi/2 \) (90°) rotation in the \textit{complex plane}, where we graph \( z = x_1 + iy_1 \) the same way we graph \( (x_1, y_1) \) in what looks like the regular \( xy \)-plane, though here the horizontal axis is referred to as the \textit{real axis}, and the vertical axis is referred to as the \textit{imaginary axis}. In the diagrams below we do see how multiplying by \( i \) is indeed the same as rotating the point around the origin \( 0 = 0 + i \cdot 0 \) by \( \pi/2 \).

Note how in the first graph, each time we multiply by \( i \) we “travel” from \( i^0 = 1 \), to \( i^1 = i \), \( i^2 = -1 \), \( i^3 = -i \), back to \( i^4 = 1 \) and so on. In the second graph note the relative positions of \( 2 + 3i \) and \( i(2 + 3i) = -3 + 2i \): the latter is a \( \pi/2 \) rotation from the former.

This already hints at why complex numbers can be useful in the physical sciences: rotations in a plane can be modeled as multiplications by powers of \( i \).

The scope of this text would have to be greatly expanded to prove the validity of the following, but the reader should be assured by the presence of dozens of textbooks on the subject, that we are allowed to perform calculus in complex variables (properly understood), which allows us to accept, for instance, the following identity of Euler:\(^{19}\)

\[
e^{i\theta} = \cos \theta + i \sin \theta.
\]

\(^{19}\)Leonhard Euler (sounds like “oiler”), 1707–1783 was an extremely prolific Swiss mathematician and physicist. A student studying graduate level mathematics will read his name often, perhaps more often than that of any other historical figure. He had a particular talent for discovering facts ahead of the time in which they could actually be proved rigorously, such as his identity (11.56).
This follows from the Maclaurin series for $e^\theta$, $\cos \theta$ and $\sin \theta$, where (of course) $\theta$ is in radians:

$$e^{i\theta} = 1 + (i\theta) + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \frac{(i\theta)^6}{6!} + \frac{(i\theta)^7}{7!} + \cdots$$

$$= 1 + i\theta - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + i\frac{\theta^5}{5!} - \frac{\theta^6}{6!} - i\frac{\theta^7}{7!} + \cdots$$

$$= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots\right)$$

$$= \cos \theta + i \sin \theta.$$

This equation (11.56) is useful in many contexts. For instance, it can be used to find the most basic trigonometric identities that involve more than one angle, if we consider two expansions for $\exp[i(\alpha + \beta)]$:

$$e^{i(\alpha + \beta)} = \cos(\alpha + \beta) + i \sin(\alpha + \beta), \quad \text{and}$$

$$e^{i(\alpha + \beta)} = e^{i\alpha} e^{i\beta}$$

$$= e^{i\alpha} (\cos \beta + i \sin \beta)$$

$$= (\cos \alpha + i \sin \alpha)(\cos \beta + i \sin \beta)$$

$$= (\cos \alpha \cos \beta - \sin \alpha \sin \beta) + i(\sin \alpha \cos \beta + \cos \alpha \sin \beta),$$

$$\implies \cos(\alpha + \beta) + i \sin(\alpha + \beta) = (\cos \alpha \cos \beta - \sin \alpha \sin \beta) + i(\sin \alpha \cos \beta + \cos \alpha \sin \beta).$$

Now anytime we have $x_1 + iy_1 = x_2 + iy_2$, where $x_1, x_2, y_1, y_2 \in \mathbb{R}$, we must have $x_1 = x_2$ and $y_1 = y_2$; that is, the real parts $x_1, x_2$ must be the same and the imaginary parts $iy_1, iy_2$ must be the same. Setting the two different forms above for the real part equal, and doing the same for the imaginary parts (divided by $i$), we get

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta,$$

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

The pair of trigonometric identities above are proved geometrically in most trigonometric textbooks, but the proof using complex numbers and Euler’s identity as above is routine once one is comfortable with complex numbers. Many more trigonometric identities follow from these, and the facts that $\cos(-\theta) = \cos \theta$ and $\sin(-\theta) = -\sin \theta$ (both of which can be proved using their own Maclaurin Series). For instance, if we set $\alpha, \beta = \theta$ we have $\cos 2\theta = \cos^2 \theta - \sin^2 \theta$, and $\sin 2\theta = 2 \sin \theta \cos \theta$ from these.

This gives rise to further geometric aspects of complex numbers. Consider Figure 11.5. It is customary to define, for $z = x + iy$, the “absolute value” of $z$, given by

$$|z| = \sqrt{x^2 + y^2},$$

which is the distance from $z$ to the origin $0 = 0 + i \cdot 0$. (Similarly $|x|$ is the distance from $x$ to zero but on the real line.) We can also define an angle $\theta$ which the ray from 0 to $z$ makes with the positive real axis, measured counterclockwise. If we do so, it is not hard to see that $x = |z| \cos \theta$ and $y = |z| \sin \theta$. It is common to see $|z|$ replaced by the real variable $r$, so $r = \sqrt{x^2 + y^2}$ and

$$z = r \cos \theta + ir \sin \theta = r(\cos \theta + i \sin \theta) = re^{i\theta}.$$

$^{20}$This quantity $|z|$ has many other names such as the modulus, norm, magnitude, and length of $z$. 

11.6. COMPLICATIONS AND THE ROLE OF COMPLEX NUMBERS

803
Figure 11.5: A complex number \( z = x + iy \) written in polar form \( z = re^{i\theta} = r(\cos \theta + i \sin \theta) \).

This is called the polar form of the complex number \( z \). (A similar theme is developed with the usual Cartesian Plane, \( \mathbb{R}^2 \), in Chapter 12.)

This gives us some interesting aspects of complex multiplication. If \( z_1 = r_1e^{i\theta_1} \) and \( z_2 = r_2e^{i\theta_2} \), then
\[
z_1z_2 = r_1r_2e^{i(\theta_1+\theta_2)},
\]
so that when we multiply two complex numbers, in the product their lengths (absolute values) are multiplied, and their angles are added.

Besides giving further illumination on the idea that multiplying by \( i \) is the same as revolving the complex number 90° around the origin, this also lets us “work backwards” to solve some other interesting problems. For instance, what should be the square root of \( i \)? One problem with answering this is that there are actually two square roots of \( i \), namely \( i \) and \(-i\), and there are two square roots of 9, namely 3 and \(-3\). We usually choose one to be “the square root,” and so with the complex plane we might choose only those whose angles \( \theta \) are within \([0, \pi)\), though that is only one convention. In fact in most applications we are interested in all roots, so in the computations below we use quotation marks around the expressions for the roots. We also exploit the ambiguity regarding what exactly should be \( \theta \), since once we have a workable \( \theta \) we also have \( \theta + 2n\pi \) also legitimate, for \( n \in \mathbb{Z} \).

Example 11.6.4 Find all fourth roots of 16.

Solution: Here we write 16 in the form \(|z|e^{i\theta}\) using four consecutive legitimate values for \( \theta \), and then formally (or “naively”) apply the 1/4 power:

\[
\begin{align*}
16 &= 16e^{i0} & \Rightarrow & 16^{1/4} = 16^{1/4}e^{i\frac{0}{4}} = 2 (\cos 0 + i \sin 0) = 2 (1 + i \cdot 0) = 2,
16 &= 16e^{i2\pi} & \Rightarrow & 16^{1/4} = 16^{1/4}e^{i\frac{2\pi}{4}} = 2 \left( \cos \frac{\pi}{2} + i \sin \frac{\pi}{2} \right) = 2 (0 + i \cdot 1) = 2i,
16 &= 16e^{i4\pi} & \Rightarrow & 16^{1/4} = 16^{1/4}e^{i\frac{4\pi}{4}} = 2 (\cos \pi + i \sin \pi) = 2 (-1 + i \cdot 0) = -2,
16 &= 16e^{i6\pi} & \Rightarrow & 16^{1/4} = 16^{1/4}e^{i\frac{6\pi}{4}} = 2 \left( \cos \frac{3\pi}{2} + i \sin \frac{3\pi}{2} \right) = 2 (0 + i \cdot (-1)) = -2i.
\end{align*}
\]
If we were to write $16 = 16 e^{8\pi i}$, we would get the same root as we got in the first case above, namely “$16^{1/4} = 2 e^{i2\pi} = 2$” as before. Similarly with the other possible values of $\theta$: we would again get only the four previous fourth roots, namely $\pm 2, \pm 2i$.

These fourth roots of 16 can also be found by solving $x^4 = 16$ using high school algebra, but the technique above also allows us to find any roots of any number which we can write in the form $z = re^{i\theta}$.

Example 11.6.5 Find the square roots of $i$.

Solution: We proceed as above, noting that $i$ makes an angle of $90^\circ$ with the positive real axis. We will use $\theta = \frac{\pi}{2}$ and $\theta = 2\pi + \frac{5\pi}{2} = \frac{5\pi}{2}$ to find our two second roots.

\[
\begin{align*}
i & = 1 e^{i\pi/2} \implies i^{1/2} = 1^{1/2} e^{i\pi/4} = 1 \left( \cos \frac{\pi}{4} + i \sin \frac{\pi}{4} \right) = 1 + i, \\
i & = 1 e^{i5\pi/2} \implies i^{1/2} = 1^{1/2} e^{i\frac{5\pi}{4}} = 1 \left( \cos \frac{5\pi}{4} + i \sin \frac{5\pi}{4} \right) = -\frac{1}{\sqrt{2}} + i \frac{1}{\sqrt{2}} = -1 + i.
\end{align*}
\]

Thus the square roots of $i$ are $\pm (1 + i)/\sqrt{2}$. Note that these make 45° and 225° angles with the positive real axis, so when we square these—and thus double the angles—we arrive at angles of 90° and 450°, which are where we will find $i$. The lengths of either root are $\sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = 1$, so when we square these roots we get a complex number with length $1^2 = 1$. So our computed roots have the correct angle and the correct length when squared.

The reader can verify that adding another multiple of $2\pi$ to the original angle for $i$ will yield one of the same two square roots of $i$ in the process above.

Anytime we graph the nth roots of a number, on the complex plane these roots will always have the same absolute value (distance from the origin), and successive ones will make angles of $2\pi/n$ between them, because we write the original number with successive angles in increments of $2\pi$, so when we take the “$1/n$” power we get angles differing by $2\pi/n$. This also explains why there will be exactly $n$ such roots, after which the process’s outcomes are repeated.

### 11.6.3 The Complex Plane’s Role

While very useful and interesting in their own right, the main purpose of introducing complex numbers here is to show their importance in the theory of power series. In particular, Abel’s Theorem is actually a theorem about power series for complex numbers:

**Theorem 11.6.1** Any power series $\sum_{k=0}^{\infty} a_k (z - z_0)^k$, where $z_0, a_0, a_1, a_2, \cdots \in \mathbb{C}$, will converge absolutely either

(i) at $z_0$ only, or

(ii) on all of $\mathbb{C}$, or

(iii) within a circle where $|z - z_0| < R$ for some $R > 0$, and diverge where $|z - z_0| > R$

In each of these cases, the convergence will be absolute, meaning that $\sum |a_k (z - z_0)^k|$ will converge.
This theorem then applies to $\mathbb{R} \subseteq \mathbb{C}$, and we see that when we intersect the “open circles” of convergence in $\mathbb{C}$ for a series centered at some $a \in \mathbb{R}$, with the real line $\mathbb{R}$, we get open intervals in $\mathbb{R}$ of convergence centered at $a \in \mathbb{R}$. Like the previous statement of Abel’s Theorem, there is no mention of the boundary, which is the actual circle $|z - z_0| = R$ in $\mathbb{C}$.

The theorem can shed some light on why the Taylor Series for certain “well-behaved” functions—unlike those in Subsection 11.6.1—fail to converge on all of $\mathbb{R}$: they might not be so well behaved in $\mathbb{C}$.

Example 11.6.6 Consider the function $f(x) = 1/(x^2 + 1)$. This function and all of its derivatives exist on all of $\mathbb{R}$, as the reader can verify. Its Maclaurin Series is given by

$$\frac{1}{x^2 + 1} = \frac{1}{1 - (-x^2)} = \sum_{k=0}^{\infty} (-1)^k x^{2k},$$

which we get from geometric series methods. The interval of convergence is $x \in [-1, 1]$.

If instead we look at $f$ as a function of a complex variable $z$ with the same formula, we have

$$f(z) = \frac{1}{z^2 + 1}$$

which is undefined at $z = \pm i$, where the denominator would be zero. With Abel’s Theorem stating that outside of a circle of some radius $R$ the series representation

$$f(z) = \frac{1}{z^2 + 1} = \sum_{k=0}^{\infty} (-1)^k z^{2k}$$

will diverge, and converge inside the open disc bounded by the circle, it is reasonable that the series for $f(z)$ should converge for $|z| < 1$, and diverge for $|z| > 1$ since $|z| = 1$ is where the function first encounters any discontinuities (in this case, in the function and all its derivatives).

In the diagram at the left, the white region is $|z| < 1$, where the series converges (absolutely), and the gray region is $|z| > 1$, where it diverges. Note that $\pm i$ are on the boundaries of the dashed circle. These points $\pm i$ are precisely where $f(z) = \frac{1}{1 \pm z^2}$ has a discontinuity (dividing by zero), and so we should expect the Maclaurin series to be valid at most up to the circle, as per Abel’s Theorem, and this would imply absolute convergence of the real-variable series for $f(x)$ within $x \in (-1, 1)$. (One must test the boundary points separately.)

So when we attempt to determine the region in which a series expansion for a function is valid, the more complete context is $\mathbb{C}$. For instance, if we are only looking in $\mathbb{R}$ then the function $f(x) = \frac{1}{x^2 + 1}$ has no problems in the function itself or its derivatives anywhere in $\mathbb{R}$, but when we consider the context of $\mathbb{C}$, perhaps rewriting it as $f(z) = 1/(z^2 + 1)$, we can immediately detect a problem at $z = \pm i$. 
11.6. COMPLICATIONS AND THE ROLE OF COMPLEX NUMBERS

Example 11.6.7 Find the largest open interval of convergence for the Taylor Series representation of \( f(x) = \frac{1}{x^2+1} \) centered at \( x = 5 \). Do not write the actual series.

Solution: Again, in \( \mathbb{C} \) the only discontinuities in the function or its derivatives are at \( z = \pm i \), which are a distance \( \sqrt{5^2 + 1^2} = \sqrt{26} \) from the center \( a = 5 \), and so the largest open interval of convergence would be \( x \in (5 - \sqrt{26}, 5 + \sqrt{26}) \). The reader is encouraged to draw the open disc in \( \mathbb{C} \) as above, though it would be centered at \( z = 5 \) and would extend to its boundary which would contain \( z = \pm i \). (Note that the left endpoint of the real interval of convergence would be negative.) The actual series would be of the form \( \sum a_k(x - 5)^k \).

The technique above would be much easier than first finding the actual form of the series, and then using a Ratio Test technique to find the actual interval of convergence.

11.6.4 Summary

The reader might at this point be wondering how we know the series referred to in the above example would converge to the function in that interval, while the Maclaurin Series for the function in Example 11.6.3, page 800 does not, even though there are no troubles on all of \( \mathbb{R} \) with the function or derivative. The explanation is that the function \( e^{-1/2z} \) has some very violent behavior near \( z = 0 \) in the complex plane, behavior which does not occur anywhere along the real line.\(^2\)

The correct explanation, which again is not proved here due to the scope of this textbook, is that our usual functions found in this textbook, with the exception of those defined piecewise (including \( |x| \)), will have Taylor Series which converge in any open disc \( |z - z_0| < R \), where \( z_0 \) is the center and where \( R \) is the distance from \( z_0 \) to the nearest discontinuity. This was the analysis in Example 11.6.6, page 806 and the subsequent Example 11.6.7. This applies to all combinations of polynomial, root, trigonometric, arc-trigonometric, exponential and logarithmic functions using addition, subtraction, multiplication, division and functional composition (meaning the output of one function is fed as an input into another). It is also helpful to know (by a contrapositive-type argument using Abel’s Theorem) that any function with a Taylor Series which converges to that function on all of \( \mathbb{R} \) must have that series converge on all of \( \mathbb{C} \): if it did not converge on all of \( \mathbb{C} \), it could not on all of \( \mathbb{R} \) either, as a problem in \( \mathbb{C} \) would limit the size of a disc of convergence there, which could therefore not include all of \( \mathbb{R} \).

\(^2\)The point \( z = 0 \) is called, in complex function theory, an essential singularity. In fact, as we can see from the series for \( e^z \), we could write

\[ e^{-1/z^2} = 1 - \frac{1}{z^2} + \frac{1}{2! \cdot z^4} - \frac{1}{3! \cdot z^6} + \cdots, \]

we can expect more and more “singular” behavior as \( z \to 0 \) in \( \mathbb{C} \), meaning as \( 0 < |z| < \varepsilon \) for smaller and smaller \( \varepsilon > 0 \). Recall how \( z^{-n} = (1/z)^n \) will make angle \( n\theta \) from the positive real axis, where \( \theta \) is the angle made by \( 1/z \), and so these terms in the above series, until the factorials take over, can have some dramatic behavior in the partial sums. (That is not so much the case when \( \theta \in \{0, \pi\} \), i.e., when \( z \in \mathbb{R} \).)

A surprising and beautiful theorem of complex analysis says that any open disc containing an essential singularity \( z_0 \) will “map to” all of \( \mathbb{C} \) excepting perhaps a single value, so for such a function \( f \) we have the output from the function, with input from the disc, is all of \( \mathbb{C} \) or could possibly miss a single value in \( \mathbb{C} \). Thus

\[ \{ f(z) \mid 0 < |z - z_0| < \varepsilon \} = \mathbb{C}, \text{ or } \mathbb{C} - \{w_0\}, \]

where \( w_0 \in \mathbb{C} \) depending upon the function. Once a student of complex variables is aware of the nature of an essential singularity (having a series representation with infinitely many negative powers of \( z - z_0 \) being the signature of such functions and their singularities at \( z_0 \), detecting them is routine, and that student could use that knowledge to again help detect where a function can be represented by a convergent Taylor Series, and where that is impossible. In fact \( f(z) = e^{-1/z^2} \) can have a series in any disc that avoids the singularity, namely the origin. In fact the only value not in the range of the function is zero, though that value is approached as \( z \to 0^\pm \), that is, along the real axis. That is why we defined \( f(x) \) to be zero at \( x = 0 \) in Example 11.6.3, page 800.
We can conclude that we can find Taylor Series representations for most of the functions we encounter in this textbook, and that these series will be valid on intervals the limits of which might be easier to find by looking at the functions in the complex plane \( \mathbb{C} \) instead of in \( \mathbb{R} \). That was the case with \( f(x) = 1/(x^2 + 1) \), because we can see \( f(z) = 1/(z^2 + 1) \) has clear problems at \( z = \pm i \), but when we look instead at functions such as logarithms, the definitions of which are somewhat complicated in \( \mathbb{C} \), it is perhaps better to use real-number methods (such as the Ratio Test), though it should be noted that \( f(x) = \ln x \) has a discontinuity at \( x = 0 \), so we expect the same of \( f(z) = \ln z \) (whatever that means), and so the disc in \( \mathbb{C} \) in which a series centered at \( z = 1 \) cannot extend more than a distance of 1 in any direction, so clearly neither can the interval of convergence in \( \mathbb{R} \).

Piecewise-defined functions have the other difficulty discussed in this Section 11.6, in that a Taylor Series that works very well for the formula for one piece is unlikely to extend to the other pieces, which we expect to have different formulas for their definitions there.

With these two ideas in mind (being wary of piece-wise defined functions, and the possibility of looking into \( \mathbb{C} \) to find where a real Taylor Series converges), one can avoid some common mistakes of scientific researchers who assume a series expansion of a function in order to fit data to polynomials. That assumption is often correct, but not always, and it is important to be able to detect when function input values lie outside the interval where a Taylor Series is valid.

**Exercises**

1. Show by direct computation that if \( z = (1 + i)/\sqrt{2} \), then \( z^2 = i \).

2. Find the four fourth roots of \(-16\), using the technique in Example 11.6.4, page 11.6.4. Graph all the roots together.

3. Where will the Maclaurin Series for \( f(x) = 1/(x^4 + 16) \) be valid? Use two different methods for solving this:

   (a) using geometric series arguments, and

   (b) using the previous problem and a complex plane argument.

4. Consider the complex conjugate of a complex number \( z \in \mathbb{C} \) defined by \( \overline{z} \) as below:

   \[
   z = x + iy \\
   \overline{z} = x - iy.
   \] (11.57)

   This is also written \( x + iy = x - iy \).

   (a) Show that \( z\overline{z} = |z|^2 \).

   (b) Show how to use this with division, where

   \[
   \frac{a + bi}{c + di} = \frac{(ac + bd) + i(bc - ad)}{c^2 + d^2}.
   \]

5. Show that when we divide \( z_1 \) by \( z_2 \), the quotient \( z_1/z_2 \) has angle \( \theta = \theta_1 - \theta_2 \), where \( z_1, z_2 \) have angles \( \theta_1, \theta_2 \), respectively, in the sense of Figure 11.5, page 804.

6. Show that \( z \) and \( \frac{1}{z} \) will have angles whose terminal rays point in opposite directions, assuming \( z \neq 0 \).
Chapter 12

Topics in Two-Dimensional Analytic Geometry

In this chapter we look at topics in analytic geometry so we can use our calculus in many new settings. Most of the discussion will involve developing these settings, but once developed we will have some immediate input from our calculus.

Analytic geometry is the name we give to the combining of geometry with theories of equations. For instance, we place \( xy \)-coordinate axes in a plane, renaming the plane Euclidean or Cartesian 2-space, and immediately we can associate every line with an equation \( Ax + By = C \) for some \( A, B, C \in \mathbb{R} \). Thus we can associate geometric figures with equations involving \( x \) and \( y \), and vice-versa, which opens up the possibility of algebraic and calculus-based analysis of geometric figures, and geometric analysis of the equations. It is nearly impossible to exaggerate the importance of these connections between equations and geometric objects.

We will not always use \( x \) and \( y \), also known as rectangular coordinates, to describe points and equations. In particular, we will be interested in polar coordinates, which have many uses as well. We will also generalize the idea of a position \( (x, y) \) to the idea of a vector \( \overrightarrow{PQ} \).

In the next chapter, we will extend these ideas into three-dimensional space, with rectangular coordinates \( (x, y, z) \), as well as cylindrical and spherical coordinates. We will examine lines and planes in space, surfaces in space, three-dimensional vectors and their applications.
CHAPTER 12. TOPICS IN TWO-DIMENSIONAL ANALYTIC GEOMETRY

Figure 12.1: Congruent infinite right-circular cones (with the same axes of symmetry and vertices), with intersecting planes. The plane will intersect the cones in either a single point, a line, two lines or a conic section: i.e., a parabola, a circle, an ellipse, or a hyperbola. Graphic is from Wikimedia Commons.

12.1 Conic Sections

In short, conic sections are plane curves which are either circles, ellipses, parabolas or hyperbolas. By their natures, there are multiple ways of approaching these curves. We will only consider those with vertical or horizontal lines of symmetry.\textsuperscript{1} We will first look at the actually definitions of the conic sections, but these are only occasionally necessary in applications, particularly in optics and astronomy, so we will then look at how to draw their graphs based upon their equations without referencing their definitions. Finally we will see how the definitions give rise to the equations, which will help us to write the equations given different types of data for the curves, including the location of foci.

12.1.1 Conic Sections Defined Geometrically

A conic section is an intersection of a single plane with two stacked, congruent, infinite right circular cones. This is illustrated in Figure 12.1. From there we can see that there are several ways these geometric figures can intersect. Not included in the Figure 12.1 above are other possibilities, which are not of interest here except as “degenerate cases,” such as a single point, a single line or two lines meeting at the vertices of the cones (left to the imagination of the reader). The more interesting cases we deal with here are the parabola, ellipse (including the circle), and the hyperbola. When one uses the term “conic section,” it is usually in reference to these “more interesting” cases.

However, there are other definitions of these figures, based upon distances. The circle is the easiest: the set of all points in a plane which are the same, fixed distance, or radius from a fixed point, called the center, also in the plane. If the center is \((h, k)\) and the radius is \(r\), this definition quickly becomes (from the Pythagorean Theorem)

\[
\sqrt{(x - h)^2 + (y - k)^2} = r,
\]

\textsuperscript{1}A dedicated analytic geometry or a linear algebra course can be appropriate settings to consider equations of rotated conic sections with non-vertical or non-horizontal lines of symmetry.
or the more common form, which is equivalent since \( r > 0 \):

\[
(x - h)^2 + (y - k)^2 = r^2.
\]  \(12.1\)

For \( r = 0 \) our “circle” would be a single point, which is also a possible intersection of a plane with our cones above. Thus one can debate whether or not to consider a single point to somehow be a circle.

The ellipse is defined somewhat similarly, except that for the ellipse we take two fixed points, or foci (singular is “focus”) in the plane, and find all points whose distances sum to some fixed total distance. A special case of this can be seen in Figure 12.6, page 821. It will take some effort to derive the form of an equation for an ellipse from this definition, and we will do so later, in Subsection 12.1.3.

Next in complexity is the parabola, which is defined by all the points in the plane whose distance to a fixed point, or focus in the plane, is the same as the distance to a fixed line, or directrix lying in the plane but not containing the focus. This is illustrated in Figure 12.5, page 820. The derivation of an equation for a parabola is somewhat simpler, and will also be included later.

Finally, there is the hyperbola, which is defined by all points in a plane whose distances to two fixed points, or foci in the plane, differ (in absolute value) by a constant. This is illustrated in Figure 12.7, page 823.

While the definitions are interesting and illustrative, we can do much with these figures without resorting to these definitions, or even the foci or (in the case of the parabola) directrices. This is because we can get much general information regarding the positions and shapes of the figures from simplified equations in which a focus or directrix does not appear directly. For completeness we will return to these definitions in Subsection 12.1.3, and use information from those derivations in later computations. While we can draw, say, a parabola without knowing its focus or directrix, if we are interested in optics or acoustics these things are more important. For most calculus applications they are not so much.

### 12.1.2 Simplest Equations of Conic Sections

Most algebra students learn quickly that the graph of \( y = x^2 \) is a parabola. After plotting simple points, we see a line of symmetry \( x = 0 \), intersecting at a vertex \((0, 0)\). After a small amount of the graphical theory of functions, one usually then learns that any equation of the form

\[ f(x) = a(x - h)^2 + k, \quad a \neq 0 \]  \(12.2\)

will be similar, but with the vertex shifted to \((h, k)\), the axis of symmetry now being \( x = h \), and the parabola opening upward if \( a > 0 \), and downward if \( a < 0 \). Shortly after that, one is taught that any function of the form \( f(x) = ax^2 + bx + c \ (a \neq 0) \) is also a parabola, and we can use

\footnote{However, many algebra students make the mistake of believing that any “U-shaped,” or upside-down U-shaped curve is a parabola. Just as not every round “loop” is a circle, we should not expect every curve whose shape superficially resembles a parabola to actually be a parabola. The curve \( y = \sin x \), for \( x \in [0, \pi] \) is one such example, which resembles a parabola but is in fact not a piece of a parabola. Nor are either branches of a hyperbola “parabolic.” Similarly, not every “oval” is an ellipse.}
completing the square to find its form (12.2):\footnote{Since we have calculus here we can certainly use it to find that the one critical point where \( f'(x) = 0 \) is at \( x = -b/2a \), and that this must be the \textit{vertex} of the parabola.}

\[
f(x) = ax^2 + bx + c
= a \left( x^2 + \frac{b}{a}x \right) + c
= a \left[ x^2 + \frac{b}{a}x + \left( \frac{b}{2a} \right)^2 - \left( \frac{b}{2a} \right)^2 \right] + c
= a \left( x - \frac{b}{2a} \right)^2 - a \cdot \frac{b^2}{4a^2} + c
= a \left( x - \frac{b}{2a} \right)^2 + \left( c - \frac{b^2}{4a} \right).
\]

Thus \( y = f(x) \) has the same shape as \( y = ax^2 \) except it has been moved horizontally by \( h = \frac{b}{2a} \) and vertically by \( k = c - \frac{b^2}{4a} \).

**Example 12.1.1** Find the vertex and graph the parabola \( y = 3x^2 + 18x - 2 \).

**Solution:** We do as above, completing the square after first factoring the second and first-degree terms collectively. Recall that when completing a square of the form \( x^2 + \beta x \), we add and subtract \( (\beta/2)^2 \), to get \( x^2 + \beta x + \beta^2/4 - \beta^2/4 = \left( x + \frac{\beta}{2} \right)^2 - \beta^2/4 \).

\[
y = 3(x^2 + 6x) - 2 = 3(x^2 + 6x + 9 - 9) - 2
= 3(x + 3)^2 - 27 - 2 = 3(x + 3)^2 - 29.
\]

Thus we are asked to plot \( y = 3(x + 3)^2 - 29 \), which has a vertex at \((-3, -29)\). It is common practice to plot the vertex and two symmetric points, such as \( x = h \pm 1 \) if it is convenient. Here that would be the points \((-2, -26)\) and \((-4, -26)\). Because of the location of the curve, we omit the axes here:

Of course there are also those of the form

\[
x = a(y - k)^2 + h, \quad \text{or} \quad x - h = a(y - k)^2.
\]
These open horizontally, to the right (vertex on the left) if $a > 0$, and to the left if $a < 0$.

Now consider an equation which is claimed to represent an ellipse:

**Example 12.1.2** Consider the curve $\frac{x^2}{9} + \frac{y^2}{16} = 1$.

Note that this is a variation on the unit circle $x^2 + y^2 = 1$, except for scaling in the horizontal and vertical directions. In fact the above equation can be rewritten

$$\left(\frac{x}{3}\right)^2 + \left(\frac{y}{4}\right)^2 = 1.$$

If we look at ordered pairs $(X, Y)$ which satisfy $x^2 + y^2 = 1$, then for analogous points on the ellipse $(x/3)^2 + (y/4)^2 = 1$ we require $(3X, 4Y)$:

$$(3X/3)^2 + (4Y/4)^2 = 1 \iff X^2 + Y^2 = 1.$$  

This effectively stretches the graph by a factor of 3 in the horizontal direction, and by a factor of 4 in the vertical direction, giving us our ellipse. Below we give the graphs, equations, and one analogous point on each.

The example above, properly generalized, indicates a method of efficient plotting of ellipses.

If we generalize this to have different centers $(h, k)$ (not simply $(0, 0)$ as above), we get a form

$$\frac{(x - h)^2}{a^2} + \frac{(y - k)^2}{b^2} = 1. \quad (12.3)$$

Here we assume $a, b > 0$. For any such equation, graphing the ellipse is straight-forward:

- Identify the center $(h, k)$, and label it.
- From the center, move left and right by $a$ to find the points along the horizontal axis.
- From the center, move up and down by $b$ to find the points on the vertical axis.
- Graph the ellipse.
Figure 12.2: Ellipses with equations \((x - h)^2/a^2 + (y - k)^2/b^2 = 1\). The centers are \((h, k)\) in both cases. For the first ellipse, \(a < b\), so it has a vertical \textit{major axis} of length \(2a\), and horizontal \textit{minor axis} of length \(2b\). In the second ellipse, \(a > b\) so the major axis is horizontal with length \(2a\), and the minor axis is vertical with length \(2b\).

See Figure 12.2. There we see two axes: a horizontal axis of length \(2a\), and a vertical axis of length \(2b\). The longer of these axes is called the \textit{major axis}, and the shorter of these axes is called the \textit{minor axis}.

Example 12.1.3 Consider the ellipse \(\frac{(x + 2)^2}{9} + \frac{(y - 3)^2}{16} = 1\). The center is \((-2, 3)\), and \(a = 3\), \(b = 4\), so four points we can plot immediately are \((-2 \pm 3, 3)\), \((-2, 3 \pm 4)\), \(i.e.,\) points \((-5, 3)\), \((1, 3)\), \((-2, -1)\) and \((-2, 7)\):

Example 12.1.4 Find the center and four axis (major and minor) points of the ellipse

\[4x^2 - 12x + 5y^2 + 20y = 0.\]
Solution: The usual process of completing the square, and the final steps to have the constant 1 on the right-hand side, are as follow:

\[ 4x^2 - 12x + 5y^2 + 20y = 0 \]
\[ \iff 4(x^2 - 3x) + 5(y^2 + 4y) = 0 \]
\[ \iff 4 \left( x^2 - 3x + \frac{9}{4} - \frac{9}{4} \right) + 5(y^2 + 4y + 4 - 4) = 0 \]
\[ \iff 4(x - \frac{3}{2})^2 - 9 + 5(y + 2)^2 - 20 = 0 \]
\[ \iff 4(x - \frac{3}{2})^2 + 5(y + 2)^2 = 29 \]
\[ \iff \frac{4}{29}(x - \frac{3}{2})^2 + \frac{5}{29}(y + 2)^2 = 1 \]
\[ \iff \frac{(x - \frac{3}{2})^2}{\frac{29}{4}} + \frac{(y + 2)^2}{\frac{29}{5}} = 1. \]

From this we see the center is \((3/2, -2)\), and the points on the axes are

\[ \left( \frac{3}{2} \pm \frac{\sqrt{29}}{2}, -2 \right), \quad \left( \frac{3}{2}, -2 \pm \sqrt{\frac{29}{5}} \right). \]

More useful for graphing perhaps are decimal approximations of these points:

\((4.19, -2), (-1.19, -2), (1.5, .41), (1.5, -4.4)\).

A rough sketch of this ellipse is then relatively easy, either by plotting the four points above, or by using the distances

\[ a = \frac{1}{2} \sqrt{29} \approx 2.69, \quad b = \sqrt{\frac{29}{5}} \approx 2.41. \]

Since these are so similar in length, the ellipse is closer to “circular” than the previous examples.

For the hyperbola, we start with the simplest examples, namely \(x^2 - y^2 = 1\) and \(y^2 - x^2 = 1\). These are illustrated in Figure 12.3 Let us take these in turn, though we note that there is a clear analogy between the two curves, as \(x\) and \(y\) basically exchange roles.

\[ x^2 - y^2 = 1: \] For this curve, we first note that \(x^2 = 1 - y^2 \in [1, \infty)\), which requires \(|x| \in [1, \infty)\), i.e., \(x \in (-\infty, -1) \cup [1, \infty)\). Therefore \(x\) is somewhat limited in possible values, while \(y\) is not. In fact the “vertices” of the hyperbola will occur where \(x = \pm 1\),
and \( y = 0 \). From the equation, the graph is obviously symmetric with respect to both axes, since \((a, b)\) on the graph means all possible combinations of \((\pm a, \pm b)\) will also be on the graph. The line of symmetry passing through both vertices is called the *axis* of the hyperbola, which in this case is the \(x\)-axis. Finally, for large \( x \) we have

\[
y = \pm \sqrt{x^2 + 1} \approx \pm \sqrt{x^2} = \pm |x| = \pm x,
\]
giving us the two asymptotic lines \( y = x \) and \( y = -x \).

- \( y^2 - x^2 = 1 \): For this curve, we require \( y \in (-\infty, -1] \cup [1, \infty) \), but there is no restriction on \( x \). The vertices occur where \( y = \pm 1 \), and \( x = 0 \), and the graph is symmetric with respect to both axes, the \(y\)-axis being the \( x \)-axis of this particular hyperbola. Finally, for large \( x \) we have

\[
y = \pm \sqrt{x^2 + 1} \approx \pm \sqrt{x^2} = \pm |x| = \pm x,
\]
again giving us asymptotic lines \( y = x \) and \( y = -x \).

Note that both hyperbolas have a natural “center” at \((0, 0)\), which is both where the asymptotes intersect each other, and a point with respect to which the parabola is symmetric (in the “inverting lens” fashion).

We can now find our general equation of the hyperbola centered at \((h, k)\):

\[
\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1, \quad (12.4)
\]

\[
\frac{(y-h)^2}{b^2} - \frac{(x-h)^2}{a^2} = 1. \quad (12.5)
\]

For the case (12.4), the vertices are at \((h \pm a, k)\), and the axis is the line \( y = k \). For case (12.5), the vertices are at \((h, k \pm b)\), and the axis is the line \( x = h \). In both cases, we have center \((h, k)\), asymptotes through \((h, k)\) with slopes \(m = \pm \frac{b}{a}\):

\[
y - k \approx \pm \frac{b}{a}(x - h). \quad (12.6)
\]

In fact, when graphing these it is simpler to use the point \((h, k)\) and the slopes \(\pm b/a\), rather than using (12.6).
12.1. CONIC SECTIONS

Figure 12.4: Illustration of graphing technique for two hyperbolas, with the same center 
(h, k), and same values for a and b. They will have the same asymptotes, and we can use the
same “box” to guide our graphs for both. See Example 12.1.5

Example 12.1.5 Graph the hyperbola \( \frac{(x-2)^2}{9} - \frac{(y+1)^2}{16} = 1 \).

Solution: The equation puts restrictions on x, and the vertices will occur where \( (x-2)^2/9 = 1 \),
i.e., where \( x-2 = \pm 3 \), while \( (y+1)^2/16 = 0 \), so the vertices are at \( (2 \pm 3, -1) \). In fact, a
quick method of graphing this is to

1. first, make note of the center \( (2, -1) \);
2. next, move \( a = 3 \) units horizontally both left and right to find the vertices,
3. next, draw the asymptotes as lines through the center \( (2, -1) \) and with slopes \( \pm b/a = \pm 4/3 \),
   and
4. finally, draw a hyperbola through the vertices found above, approaching the asymptotes
   (usually rather quickly).
5. Alternatively, draw the (dotted) box below, and use it to find the vertices and asymptotes
to draw the curve, as explained below.

See Figure 12.4 for the graph illustrating the technique. The extra “dotted” line segments
are optional, and represent

(a) points obtained by moving from the center \( (h, k) = (2, -1) \) horizontally by \( \pm a = \pm 3 \),
   and vertically by \( \pm b = \pm 4 \), and will include the vertices, and
(b) the “corners” of a box through which the lines passing through the center and have
   slopes \( \pm b/a = \pm 4/3 \). These help to draw the asymptotes.

The lengths of each dotted segment are \( a = 3 \) for horizontal segments, and \( b = 4 \) for each
vertical segment.
It should be pointed out that the related hyperbola
\[
\frac{(y+1)^2}{16} - \frac{(x-2)^2}{9} = 1
\]
will have the same asymptotes, and the same “box” as above, though the vertices will be at 
\((h,k \pm b) = (-2,3), (-2,-5)\). The graph of this hyperbola is superimposed on the graph for the
hyperbola in Example 12.1.5, but in dashed, gray lines.

If the equation of the hyperbola is not given in the standard forms (12.4) or (12.5), we may
need to manipulate the given equation to achieve such form.

**Example 12.1.6** Graph the equation \(2x^2 - 6x = 3y^2 - 18y\).

**Solution:** We will complete both squares separately, and see what is left outside the squares
to decide which form the standard equation should have.

\[
\begin{align*}
2(x^2 - 3x) &= 3(y^2 - 6y) \\
\iff 2 \left( x^2 - 3x + \frac{9}{4} - \frac{9}{4} \right) &= 3(y^2 - 6y + 9 - 9) \\
\iff 2 \left( x - \frac{3}{2} \right)^2 - \frac{9}{2} &= 3(y - 3)^2 - 27 \\
\iff 27 - \frac{9}{2} &= 3(y - 3)^2 - 2 \left( x - \frac{3}{2} \right)^2 \\
\iff \frac{45}{2} &= 3(y - 3)^2 - 2 \left( x - \frac{3}{2} \right)^2 \\
\iff \frac{1}{15/2} &= \frac{(y - 3)^2}{45/2} - \frac{(x - \frac{3}{2})^2}{45/4}
\end{align*}
\]

If we wish to make the form more obvious, we could write
\[
\frac{(y - 3)^2}{\left( \sqrt{\frac{45}{2}} \right)^2} - \frac{(x - \frac{3}{2})^2}{\left( \frac{3\sqrt{5}}{2} \right)^2} = 1.
\]

Here, \(a = \frac{3}{2}\sqrt{5} \approx 3.3541\), and \(b = \sqrt{15/2} \approx 2.73861\), which will help us to graph the hyperbola
reasonably. Note that the slopes of the asymptotes are \(\pm b/a = \pm \frac{\sqrt{15/2} - \frac{2}{3\sqrt{5}}}{\sqrt{2/3}} \approx \pm 0.81650\).
Using the center \((3/2,3)\) we can graph the hyperbola, though if we sketch it by hand we can use
\(\pm b/a \approx \pm 4/5\), for instance (though the graph below is computer-generated).
12.1. CONIC SECTIONS

\[ a = \frac{3}{2} \sqrt{5} \approx 3.35410, \]
\[ b = \sqrt{\frac{15}{2}} \approx 2.73861. \]
12.1.3 Equations of Conic Sections from Definitions

Consider the definition of a parabola: the set of all points which are the same distance from a fixed point (the focus), and a fixed line (the directrix). The vertex will be the point on the parabola which is the shortest distance from both of these. We will only consider parabolas with vertical or horizontal axes (lines) of symmetry, which will pass through both the vertex and the focus (and will be perpendicular to the directrix). Our derivation will be for a vertically opening parabola (with a vertical axis of symmetry), from which the horizontally opening parabolas’ equations will follow.

We will take \((h,k)\) to be the vertex, and \(p\) to be the displacement, or “directed distance,” from the vertex to the focus. (See Figure 12.5.) Thus \(-p\) is the displacement between the vertex and the directrix.

Now consider any point \((x,y)\) on the parabola. Then the distance to the focus being equal to the distance to the directrix will give us

\[
\sqrt{(x-h)^2 + [y-(k+p)]^2} = |y-(k-p)|.
\]

We now square both sides, expand and cancel, though for reasons that will become clearer later, we will not expand the \((x-h)^2\) term. Note \(|y-(k-p)|^2 = [y-(k-p)]^2\).

\[
\begin{align*}
(x-h)^2 + [y-(k+p)]^2 &= [y-(k-p)]^2 \\
\iff (x-h)^2 + y^2 - 2(k+p)y + (k+p)^2 &= y^2 - 2(k-p)y + (k-p)^2 \\
\iff (x-h)^2 - 2ky - 2py + k^2 + 2kp + p^2 &= -2ky + 2py + k^2 - 2kp + p^2 \\
\iff (x-h)^2 &= 4py - 4pk \\
\iff (x-h)^2 &= 4p(y-k).
\end{align*}
\]

With this, and the analogous result for horizontally opening parabolas, also with vertex at \((h,k)\), we have the following:

- A vertical-axis parabola with vertex at \((h,k)\) and the focus at a vertical displacement \(p\)
from the vertex will have equation

$$(x - h)^2 = 4p(y - k),$$

focus: $$(h, k + p),$$

directrix: $$y = k - p.$$ \hfill (12.7)

- A horizontal-axis parabola with vertex at $$(h, k)$$ and the focus at a horizontal displacement $$p$$ from the vertex will have equation

$$(y - k)^2 = 4p(x - h),$$

focus: $$(h + p, k),$$

directrix: $$x = k - p.$$ \hfill (12.8)

For the ellipse, we choose some distance $$d$$ which is the sum of the distances from a point $$(x, y)$$ to the foci, which we will assume for simplicity are located at $$(\pm c, 0)$$. (We can easily adjust for more complicated cases later.) If we let $$d$$ be the sum of distances from $$(x, y)$$ to the two foci, we get

$$\sqrt{(x + c)^2 + y^2} + \sqrt{(x - c)^2 + y^2} = d.$$ 

One needs to square both sides to remove the radicals, but this can not be done in one step. It is usually algebraically simpler to have one of the radicals alone on one side before squaring.
Proceeding as before, we have

\[
\sqrt{(x+c)^2 + y^2} = d - \sqrt{(x-c)^2 + y^2}
\]

\[\implies \quad x^2 + 2cx + y^2 = d^2 - 2d\sqrt{(x-c)^2 + y^2} + x^2 + y^2 - 2cx + y^2
\]

\[\iff \quad 2d\sqrt{(x-c)^2 + y^2} = d^2 - 4cx
\]

\[\implies \quad 4d^2[x^2 - 2cx + c^2 + y^2] = d^4 - 8cd^2x + 16c^2x^2
\]

\[\iff 4d^2x^2 - 8cd^2x + 4c^2d^2 + 4d^2y^2 = d^4 - 8cd^2x + 16c^2x^2
\]

\[\iff (4d^2 - 16c^2)x^2 + 4d^2y^2 = d^4 - 4c^2d^2
\]

\[\iff 4(d^2 - 4c^2)x^2 + 4d^2y^2 = d^2(d^2 - 4c^2)
\]

\[\iff \frac{4}{d^2}x^2 + \frac{4}{d^2}y^2 = 1
\]

\[\iff \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.
\]

Now if we let \(a^2 = d^2/4\) and \(b^2 = (d^2 - 4c^2)/4\), we see that

\[a^2 - b^2 = \frac{d^2}{4} - \frac{d^2 - 4c^2}{4} = \frac{4c^2}{4} = c^2.
\]

Thus we have

\[\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,
\]

(12.9)

where

\[a^2 - b^2 = c^2,
\]

(12.10)

If one is interested, we also find \(d\) since \(d^2 = a^2/4\), though for an ellipse with the foci on a vertical axis we would replace \(a\) with \(b\). In fact, the longer (major) axis will be in the direction of the variable \(x\) if \(a > b\), and \(y\) if \(b > a\). However, a perhaps more easily recalled method for finding \(d\) might be to find the foci and one point on the ellipse, and simply compute the sum of distances to the foci from that point. Because \(d\) does not explicitly appear in the equation of an ellipse, it is usually not included in the discussion, beyond the derivation.

By considering vertical and horizontal translations of the equation above, and considering cases where the major axis is vertical, we conclude that the general equation for an ellipse centered at \((h, k)\), is given by

\[\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1,
\]

(12.11)

where

- if \(a > b\) the major axis (containing the foci) is horizontal, with endpoints \((h \pm a)\), with foci \((h \pm c, k)\), \(c^2 = a^2 - b^2\), and minor axis endpoints \((h, k \pm b)\);

- if \(b > a\), the major axis (containing the foci) is vertical, with endpoints \((h, k \pm b)\), with foci \((h, k \pm c)\), \(c^2 = b^2 - a^2\), and minor axis with endpoints \((h \pm a, k)\).

- In all cases, the center is at \((h, k)\), axes have endpoints \((h \pm a, k)\) and \((h, k \pm b)\), the foci are a distance \(c\) (i.e., a displacement \(\pm c\)) from the center \((h, k)\) along the major (longer) axis, where

\[c^2 = |a^2 - b^2|.
\]

(12.12)
Figure 12.7: A special case of the hyperbola, with foci at $(0, -c)$ and $(0, c)$. The distances from any point $(x, y)$ on the hyperbola to the foci will differ by a constant (in the sense that the absolute value of the differences will be constant). The large-$x$ and large-$y$ behavior of the hyperbola is asymptotically linear (see dotted lines).

For the hyperbola, we will first look at the case where the foci are located at $(0, \pm c)$, as in Figure 12.7. For any point $(x, y)$ on the hyperbola, this translates to

\[
\sqrt{x^2 + (y - c)^2} - \sqrt{x^2 + (y + c)^2} = d. 
\]

We begin with the case that the first distance is greater, and leave the other case as an exercise. (Note that for this case, $y < 0$.)

\[
\sqrt{x^2 + (y - c)^2} - \sqrt{x^2 + (y + c)^2} = d \\
\Leftrightarrow \sqrt{x^2 + (y - c)^2} = d + \sqrt{x^2 + (y + c)^2} \\
\Rightarrow \sqrt{x^2 + (y - c)^2} = d + \sqrt{x^2 + (y + c)^2} \\
\Rightarrow \sqrt{x^2 + (y - c)^2} = d + \sqrt{x^2 + (y + c)^2} \\
\Rightarrow 16c^2y^2 + 8cd^2y + d^4 = 4d^2(x^2 + y^2 + 2cy + c^2) \\
\Rightarrow 16c^2y^2 + 8\phi d^2y + d^4 = 4d^2x^2 + 4d^2y^2 + 8\phi d^2y + 4c^2d^2 \\
\Rightarrow (16c^2 - 4d^2)y^2 - 4d^2x^2 = 4c^2d^2 - d^4 \\
\Rightarrow 4(4c^2 - d^2)y^2 - 4d^2x^2 = 4c^2d^2 - d^4 \\
\Rightarrow \frac{4}{d^2}y^2 - \frac{4}{4c^2 - d^2}x^2 = 1 \\
\Rightarrow \frac{y^2}{d^2/4} - \frac{x^2}{(4c^2 - d^2)/4} = 1 \\
\]

Letting $b^2 = d^2/4$ and $a^2 = (4c^2 - d^2)/4$, this becomes

\[
\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1. 
\]
Note that \( b^2 + a^2 = 4c^2/4 \), i.e.,
\[
a^2 + b^2 = c^2. \tag{12.14}
\]

Note that if we solve for \( y \) in (12.13), we have
\[
\frac{y^2}{b^2} = 1 + \frac{x^2}{a^2} \\
\implies y^2 = b^2 + \frac{b^2}{a^2}x^2 \\
\iff y = \pm b \sqrt{\frac{a^2 + x^2}{a^2}} \approx \pm \frac{b}{a} \sqrt{x^2}, \tag{12.15}
\]
this last line being for large \( x \). Recall \( \sqrt{x^2} = |x| = \pm x \), depending upon whether \( x \geq 0 \) or \( x < 0 \).
Thus,
\[
x \text{ large } \implies y \approx \pm \frac{b}{a} x. \tag{12.16}
\]

We see that the large-\( x \) behavior is for \( y \approx \pm \frac{b}{a} x \), which gives two linear asymptotes: \( y = \frac{b}{a} x \) and \( y = -\frac{b}{a} x \). It is also notable that \( \frac{b^2}{a^2} \geq 1 \) gives us \( y^2 \geq b^2 \), requiring \( y \geq b \) or \( y \leq -b \), i.e., \( y \in (-\infty, -b] \cup [b, \infty) \), so \( y \) is somewhat limited in possible values, where \( x \) is not, as we can see within the equations (12.15). The vertices of the hyperbola are at \((0, \pm b)\). Through those vertices is the axis. Of course there are two lines of symmetry, the axis and the line through the center \((0,0)\) and perpendicular to the axis.

Looking at translations of our simplified model, we get a general form with translated features. The equation and the hyperbola’s features will be as follow:
\[
\frac{(y-k)^2}{b^2} - \frac{(x-h)^2}{a^2} = 1 \tag{12.17}
\]

- centered at \((h, k)\),
- vertical axis line \( x = h \)
- asymptotes \( y - k = \pm \frac{b}{a}(x - h) \), i.e., containing \((h, k)\) with slopes \( \pm \frac{b}{a} \),
- vertices at \((h, k \pm b)\),
- foci at vertical distances \( c \) from the center \((h, k)\), i.e., at \((h, k \pm c)\), where \( c^2 = a^2 + b^2 \)

If instead we have a horizontal axis, we have:
\[
\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \tag{12.18}
\]
- centered at \((h, k)\),
- horizontal axis line \( y = k \)
- asymptotes \( y - k = \pm \frac{b}{a}(x - h) \), i.e., containing \((h, k)\) with slopes \( \pm \frac{b}{a} \),
- vertices at \((h \pm a, k)\),
- foci at vertical distances \( c \) from the center \((h, k)\), i.e., at \((h \pm c, k)\), where \( c^2 = a^2 + b^2 \)
12.1.4 Eccentricity

Physicists often discuss a concept called eccentricity, associated with conic sections or, more loosely, how much a curved path deviates from a circle (which then has zero eccentricity). For most of our conic sections, it is the ratio of the distance from the “center” to a focus, divided by the distance from the center to a vertex. For the ellipse, this would be \( e = c/a \) if \( a > b \), or \( e = c/b \) if \( b > a \), so we can simply write \( e = c/\max\{a, b\} \). Since \( c^2 = |a^2 - b^2| \), we have \( c < \max\{a, b\} \), \( 0 < e < 1 \) for an ellipse. A circle is like an ellipse with both foci at the center \((c = 0)\), so \( e = 0 \). For a hyperbola, \( c^2 = a^2 + b^2 \) so \( e = c/\max\{a, b\} > 1 \).

For the parabola, it would seem strange to discuss a “center.” For the other conic sections, it seems that “center” could be the one point through which the conic section is symmetric. Recall that “symmetric with respect to a point” means that, were that point a “lens,” any part of the figure on one “side” of the point (“lens”) has a corresponding part which is the optic inversion of the first part through that point (“lens”). However the parabola has no such symmetry with respect to a point. (For the others, it is at their center \((h, k)\).) The solution is to instead define a “directrix” for each figure, and not just for the parabola, and then the eccentricity is the ratio of the distance from any point on the figure to the focus versus the distance to the directrix.\(^4\)

From the definition of the parabola—all points which are the same distance to the focus as the directrix—we trivially have \( e = 1 \).

**Exercises**

1. Compute the general formula (12.6) for the asymptotes of a hyperbola for both (12.4) and (12.5).

\(^4\)We will not pursue this in depth here, but mention one case here briefly. For an ellipse, the directrix is a line such that the ellipse can be defined as all points \((x, y)\) for which the distance \(d_1\) to the focus is proportional to the perpendicular distance \(d_2\) to the directrix. That proportionality constant is then \( e < 1 \). For the ellipse in Figure 12.6, a directrix would be \( x = a^2/c \), which is to the right of the ellipse since \( a^2/c = a \cdot \frac{a}{c} > a \). For the ellipse below, \( a = 2, b = 1 \), so \( c = \sqrt{a^2 - b^2} = \sqrt{3} \), and then \( e = c/a = \sqrt{3}/2 < 1 \).
12.2 Parametric Curves

The term **parametric curve** comes from the idea that both the $x$-coordinate and the $y$-coordinate of the curve will be functions of a third variable, called the **parameter**. This allows for many more types of curves than those given functionally by $y = f(x)$, and even includes some that would be difficult to give implicitly.

An example of a parametric curve in the plane can be

$$
\begin{align*}
x &= \cos t, \\
y &= \sin t.
\end{align*}
$$

Thus we follow the $x$-coordinate as $t$ varies, and separately (if we like) the $y$-coordinate as $t$ varies. When we graph this for the first time, we might make a chart and graph the points that occur, and attempt to deduce the shape of the graph. Choosing $t$-values with known sines and cosines, we might produce the following table and graph. Note $\sqrt{3}/2 \approx 0.866$, and $1/\sqrt{2} \approx 0.707$.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$x = \cos t$</th>
<th>$y = \sin t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$\pi/6$</td>
<td>$\sqrt{3}/2$</td>
<td>1/2</td>
</tr>
<tr>
<td>$\pi/4$</td>
<td>1/2</td>
<td>1/2</td>
</tr>
<tr>
<td>$\pi/3$</td>
<td>1/2</td>
<td>$\sqrt{3}/2$</td>
</tr>
<tr>
<td>$\pi/2$</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$2\pi/3$</td>
<td>$\sqrt{3}/2$</td>
<td>-1/2</td>
</tr>
<tr>
<td>$3\pi/4$</td>
<td>1/2</td>
<td>-1/2</td>
</tr>
<tr>
<td>$5\pi/6$</td>
<td>1/2</td>
<td>$-\sqrt{3}/2$</td>
</tr>
<tr>
<td>$\pi$</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>$7\pi/6$</td>
<td>$-\sqrt{3}/2$</td>
<td>-1/2</td>
</tr>
<tr>
<td>$5\pi/4$</td>
<td>$-1/\sqrt{2}$</td>
<td>$-1/\sqrt{2}$</td>
</tr>
<tr>
<td>$4\pi/3$</td>
<td>$-1/2$</td>
<td>$-\sqrt{3}/2$</td>
</tr>
<tr>
<td>$3\pi/2$</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>$5\pi/3$</td>
<td>1/2</td>
<td>$-\sqrt{3}/2$</td>
</tr>
<tr>
<td>$7\pi/4$</td>
<td>1/2</td>
<td>$-1/\sqrt{2}$</td>
</tr>
<tr>
<td>$11\pi/6$</td>
<td>$\sqrt{3}/2$</td>
<td>-1/2</td>
</tr>
</tbody>
</table>

12.3 Polar Coordinates

12.4 Calculus in Polar Coordinates
12.5. VECTORS IN $\mathbb{R}^2$; SCALAR (DOT) PRODUCTS

In this section we will look at vectors “in the plane,” particularly the familiar $xy$-plane, which is a two-dimensional space.\footnote{A two-dimensional space is one in which we require two variable to describe a point’s position.} Later in the text we will examine vectors in space, also sometimes known as $xyz$-space, which is a three-dimensional space.

The classical definition of a vector is a quantity with both magnitude and direction. We will see that there is an intentional ambiguity in this definition.

Another common definition of a vector is a directed line segment. This is even more problematic, as it removes exactly the ambiguity that we will see later is crucial to the usefulness of vectors. Indeed, describing a vector as a directed line segment is a bit like describing an angle as the union of two rays with the same origin. In the former you lose that the crucial elements are the length and direction (not the geometric position of the vector), while with angles we find it useful to think of them as rotations of well-defined amounts (and in particular directions) regardless of the pivot points.

Both definitions are somewhat geometric. As with derivatives and definite integrals, there is a geometric context in which vectors are easily visualized. However, there are other quanti-
ties which are “vector quantities,” and there is a purely algebraic definition of a vector which accommodates these as well.

To introduce the actual concept of vector, we will use one such example of a vector quantity, which closely mirrors the geometric intuitions. That is the concept of a vector as a displacement. As before, a displacement can be defined to be a net change in position. Suppose for instance we move from \((0, 0)\) to \((2, 3)\). This would be a change of +2 in the horizontal position, and +3 in the vertical.

Now suppose instead we move from \((-4, -1)\) to \((-2, 2)\). This would also be a change of +2 in the horizontal and +3 in the vertical directions, respectively. Both motions represent the same net displacement, which we signify by the vector \(\langle 2, 3 \rangle\). It does not matter what is our initial point, as long as our final point is right 2 and up 3 from our initial point. In all cases it is represented by the same vector \(\langle 2, 3 \rangle\). See Figure 12.8 at the beginning of this section.

While it is important that we realize that each of these displacements—of +2 in the horizontal and +3 in the vertical—is considered to be the same net displacement and therefore the same vector, for many purposes it is best to define a standard position for vectors, namely that the tail is fixed to the origin \((0, 0)\). Then the head of the vector \(\langle 2, 3 \rangle\) would lie at \((2, 3)\). The standard position of a vector allows us to easily explain these concepts of magnitude and direction. The magnitude is a measure of the vector’s size, and the direction is, of course, the direction it points. The length of the vector is given by the notation \(\|\langle 2, 3 \rangle\|\), and the Pythagorean Theorem gives it to us immediately: \(\|\langle 2, 3 \rangle\| = \sqrt{2^2 + 3^2} = \sqrt{13}\). More generally,

\[
\|\langle a, b \rangle\| = \sqrt{a^2 + b^2}. \tag{12.19}
\]

The length of the vector is also known as its magnitude, modulus, and sometimes called its absolute value.\(^7\) The direction \(\theta\) in which the vector points is measured off of the positive \(x\)-axis, just as is an angle in standard position. In general,

\[
\tan \theta = \frac{a}{b}, \tag{12.20}
\]

and does not have to be in any particular range. Some texts will have \(0 \leq \theta < 360^\circ\), while others will use \(-180^\circ < \theta \leq 180^\circ\), but all that is required is that we allow the angles available to describe all possible directions in which a vector can point. Note that we usually decline to define a direction for the zero vector \(\langle 0, 0 \rangle\), though it clearly has length \(\|\langle 0, 0 \rangle\| = \sqrt{0^2 + 0^2} = 0\). Note also that at the moment we are interested in the geometry of these vectors, not the calculus, so using radian measure for \(\theta\) is not yet necessary. Some texts defined the argument of the vector, \(\arg\langle a, b \rangle\) to be some particular angle \(\theta\), but we will usually just describe (albeit more verbosely) the angle in question.

When looking at vectors in the plane, it is only necessary to specify the two coordinates \(a\) and \(b\) of the endpoints \((a, b)\) where the head of vector \(\langle a, b \rangle\) points when \(\langle a, b \rangle\) is in standard position. Because \(a, b \in \mathbb{R}\), we often signify the set of all such possible vectors as

\[
\mathbb{R}^2 = \left\{ \langle a, b \rangle \mid a, b \in \mathbb{R} \right\}.
\]

Because it takes exactly two numbers to specify to identify the vector, \(\mathbb{R}^2\) is called a two-dimensional space.

---

\(^6\)This is similar to the “standard position” of an angle \(\theta\), which allows much of the analysis—especially of the trigonometric kind—of such objects.\(^7\)When the length of \(\langle a, b \rangle\) is called its it absolute value, the notation usually reflects this as well: \(\|\langle a, b \rangle\| = \sqrt{a^2 + b^2}\).
One of the aspects that makes vectors interesting is how they “add” and similarly combine. It is quite intuitive when put simply, and quite interesting when viewed geometrically. When we add two vectors in \( \mathbb{R}^2 \), we do so as follows:

\[
\langle a_1, b_1 \rangle + \langle a_2, b_2 \rangle = \langle a_1 + a_2, b_1 + b_2 \rangle.
\]

The new vector is then called the resultant vector, representing the net displacement. As we see from the equation above, to find the net displacement—when we first displace by \( \langle a_1, b_1 \rangle \), followed by another displacement of \( \langle a_2, b_2 \rangle \)—we look at the total horizontal displacement \( a_1 + a_2 \) and the total vertical displacement \( b_1 + b_2 \) to form our new vector. Algebraically this is very simple, but at times we will be interested in a geometric perspective. Geometric vector addition is traditionally described in the following two ways:

**Tail-to-Head:** where the origin (tail) of the second vector is placed at the head of the first;

**Parallelogram Rule:** where we form a parallelogram with the two vectors, in standard position, forming one corner, and the resultant vector coming from that corner to the diagonally opposed corner.
\[ \theta = \frac{2\pi}{6} \quad \theta = \frac{11\pi}{6} \]
12.6 Three-Dimensional Space
12.7 Vectors in $\mathbb{R}^3$; Vector (Cross) Products
12.8 Lines and Planes in $\mathbb{R}^3$
1 \oneinch + \hoffset  
2 \oneinch + \voffset  
3 \oddsidemargin = 28pt  
4 \topmargin = 23pt  
5 \headheight = 12pt  
6 \headsep = 18pt  
7 \textheight = 598pt  
8 \textwidth = 417pt  
9 \marginparsep = 7pt  
10 \marginparwidth = 0pt  
11 \footskip = 25pt  
\marginparpush = 5pt (not shown)  
\hoffset = 0pt  
\voffset = 36pt  
\paperwidth = 597pt  
\paperheight = 845pt
one inch + \hoffset
\evensidemargin = 28pt
\headheight = 12pt
\textwidth = 598pt
\marginparsep = 7pt
\textwidth = 417pt
\footskip = 25pt
\hoffset = 0pt
\paperwidth = 597pt
\topmargin = 23pt
\headsep = 18pt
\textwidth = 417pt
\marginparwidth = 0pt
\marginparpush = 5pt (not shown)
\paperheight = 845pt