

Contents

1 Review	1 ✓
1.1 Numbers and sets	1 ✓
✓1.1.1 Complex Numbers	3 ✓
1.2 Functions	6 ✓
✓1.2.1 Definitions	6 ✓
✓1.2.2 Continuity	9 ✓
✓1.2.3 Equivalence Classes of Functions	15 ✓
1.3 Linear Spaces, Linear Operators, and Linear Algebra	17 ✓
✓1.3.1 Linear operators	19 ✓
✓1.3.2 Linear algebra	25 ✓
1.4 Calculus	28 ✓
✓1.4.1 The derivative	29 ✓
✓1.4.2 Partial derivatives	34 ✓
✓1.4.3 Integration	35 ✓
1.5 Sequences and Series	37 ✓
✓1.5.1 Power Series	40 ✓
✓1.5.2 Taylor's Theorem	42 ✓
2 Review of Ordinary Differential Equations	49 ✓
✓2.1 Terminology	50 ✓
✓2.2 First-Order ODEs	53 ✓
✓2.2.1 The solution to separable first-order ODEs	55 ✓
✓2.2.2 Separation of variables in differential form	62 ✓
✓2.2.3 Non-separable linear first-order ODEs	65 ✓
✓2.3 Second-order ODEs	71 ✓
✓2.3.1 Directly Integrable Second-Order ODEs	71 ✓
✓2.3.2 Thinking about Solutions to Second-order ODEs	73 ✓
✓2.3.3 Linear Second-Order ODEs with constant Coefficients: Homogeneous Case	76 ✓
✓2.3.4 Nonhomogeneous Solutions for Second-Order ODEs with Constant Coefficients	87 ✓
3 Fourier Series Part I: Introductory Concepts	99 ✓
✓3.1 Terminology	99 ✓
✓3.2 Review: Power series	100 ✓
✓3.3 Review: Taylor series	102 ✓
✓3.4 Trigonometric Series	106 ✓
✓3.5 Fourier Series	107 ✓
✓3.5.1 Orthogonality Revisited	109 ✓

Au: Add
Problems to
TOC?

✓3.5.2 Fourier Sine Series	112 ✓
✓3.6 Fourier Cosine Series	116 ✓
✓3.7 Comparison of Basis Functions	121 ✓
✓3.8 The Spectrum	122 ✓
✓3.9 Change of Interval	126 ✓
✓3.10 Fourier series on symmetric intervals around zero	129 ✓
✓3.10.1 Even and Odd Functions	129 ✓
✓3.10.2 Fourier series on the interval $I = [-L, L]$	131 ✓
✓3.10.3 Periodic Continuation	138 ✓
✓3.11 Introduction to Convergence of Fourier Series	140 ✓
✓4 The Step and Delta Functions	147 ✓
✓4.1 Terminology	148 ✓
✓4.2 The Step and Delta Functions: The Basic Idea	150 ✓
✓4.2.1 Mathematical modeling of a point source	150 ✓
✓4.2.2 The path forward: Limiting behavior	152 ✓
✓4.3 A Construction for the Step and the Delta Functions	153 ✓
✓4.4 Delta Sequences	158 ✓
✓4.5 Properties of the delta function	159 ✓
✓4.6 An explanation of the sifting property for smooth functions	160 ✓
✓4.7 A computable example of the limit of a delta sequence and the sifting property	162 ✓
✓4.8 The step function and the derivative of the step function	164 ✓
✓4.9 Does the delta function have a Fourier Series Representation?	164 ✓
✓5 Introduction to Partial Differential Equations, Conservation Laws, and Constitutive Equations	169 ✓
✓5.1 Introduction	169 ✓
✓5.2 Terminology	169 ✓
✓5.3 Partial Derivatives and the Types of Partial Differential Equations	171 ✓
✓5.3.1 Partial derivatives	171 ✓
✓5.3.2 Characterization of linear PDE types	172 ✓
✓5.4 The Origins of Partial Differential Equations in Science and Engineering: The Axioms of Conservation	174 ✓
✓5.5 Derivation of Differential Conservation Equations	177 ✓
✓5.6 Constitutive Equations and Flux Laws	181 ✓
✓5.6.1 Convective Fluxes	182 ✓
✓5.6.2 Gradient Laws that Define Fluxes	184 ✓
✓5.7 Ancillary Conditions	187 ✓
✓5.7.1 Order of the PDE and determining necessary ancillary conditions	187 ✓
✓5.7.2 Types of ancillary conditions	188 ✓
✓5.7.3 Applications of ancillary conditions with one space and one time dimension	189 ✓
✓5.8 Transformations in the solutions to PDEs	193 ✓
✓5.8.1 Symmetries and transformations	193 ✓
✓5.8.2 An example transformation	193 ✓
✓5.9 Steady state problems	195 ✓
✓References	203 ✓

An: Perhaps this author is better known as just Danielle Steel?

"A bad review is like baking a cake with all the best ingredients and having someone sit on it."

Danielle Fernandes Dominique Schuelein-Steel, fourth bestselling fiction author of all time



Review

1.1 Numbers and sets

The concept of numbers is an intuitive and natural thing that we have all been using since we were little and learning to count out objects on our fingers. However, numbers themselves are an interesting topic; there is even a branch of mathematics called *number theory* that studies the integers. One might think that this is both boring and useless, but that turns out to be untrue (depending upon who you are, I suppose). For instance, number theory has been used to find very large prime numbers (numbers that are not the multiplication of two smaller numbers). Such numbers are of great value in use in public key encryption schemes and other kinds of computer security. In fact, there are several prizes currently offered (of about \$150K) to find a prime number with 10 or 100 million digits. The largest prime number currently known (as of July 2018) is $2^{77,232,917} - 1$, a number that has 23,249,425 digits.

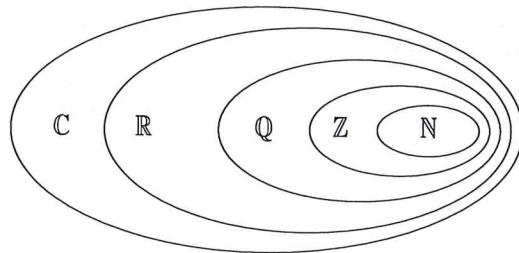
We will not make extensive explicit use of the concept of *sets*, but it is worthwhile understanding what they are since they are so prevalent in math and science generally. Plus, the concept relates to numbers in the sense that we can discuss different sets of numbers that are commonly referred to.

Definition 1.1 (Sets). A *set* is a collection of objects; more specifically for our purposes, a collection of mathematical objects. An *element* is simply one of the objects within the set. Sets can contain a finite number of objects, or an infinite number of objects.

The first kind of sets we will talk about are ones that you have probably encountered previously in your experience with math. These are the various convenient sets of numbers that are frequently referred to. The most frequently used sets of numbers are

- \mathbb{N} (The "natural" numbers). These are the set of positive integers $\{1, 2, 3, 4, \dots\}$. Some definitions include zero; to denote that, the convention is to use the symbol $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$.

✓ **Fig. 1.1** The number systems, from the perspective of sets. In this figure, each oval represents a set; an oval that is contains in another (larger) oval is a *subset* of the containing set.



- \mathbb{Z} (The integers). These are the set of all positive or negative integers, including $0 \{-2, -1, 0, 1, 2, \dots\}$

(1c) / 3/

- \mathbb{Q} (The rational numbers). The “ \mathbb{Q} ” here stands for *quotient*, or, in other words, the ratio of two numbers. The *rational* numbers are formed by the set of all possible ratios of all possible integers p and q in \mathbb{Z} so that all pairs for a ratio p/q which is a member of \mathbb{Q} . Note that there is a restriction; $q \neq 0$. Also, because we can have $q = 1$, of the integers are also rational numbers.

changes were ok, did not seem to make sense grammatically

- (1c) / 3/*
- \mathbb{P} (The irrational numbers). In short, the set of all numbers that cannot be expressed as a rational number, i.e., they are not elements of \mathbb{Q} .

(1c) / 3/

- \mathbb{R} (The real numbers). The real numbers are the set of all things that we might think of as conventional numbers; in other words, it contains all of the sets of numbers defined previously. For our purposes, the real numbers can be thought of most simply as the *union* (or the combining of) the rational and the irrational numbers (in mathematics this might be written $\mathbb{R} = \mathbb{Q} \cup \mathbb{P}$).

(1c) / 3/

- \mathbb{C} (The complex numbers). The complex numbers are a generalization of the concept of number. Most of us have been introduced to the idea of complex numbers early in our mathematical career; however, they frequently instill much unease and consternation until they can be better understood in the when the concept of a *field* is introduced; in short, a field is *a set* on which addition, subtraction, multiplication, and division are defined. Because there are few areas where complex numbers are important in engineering analysis, they will be described a little more detail in the material following.

A graphic indicating how these number systems (as sets) relate to one another is given in Fig. 1.1. Note that, for each of these number sets, the values $\pm\infty$ are not part of the set of numbers! There are number sets (e.g., the *extended* reals) that contain ∞ as a value, but these sets will not be used in this text.

These sets of numbers have interesting properties and histories of their own. For example, the irrational numbers were first discovered by the Greeks. In particular it is thought that a Pythagorean (i.e., a member of a Greek math cult following the tenants of Pythagoras), named Hippasus of Metapontum first proved this. Strangely, although it is the Pythagorean theorem that allowed Hippasus to show that irrational numbers exist, this went in stark contradiction to the Pythagorean belief that all mathematics could be expressed through ratios of integers. It is said that Hippasus discovered this startling fact while at sea, and his fellow Pythagoreans were so upset by the revelation that they threw him overboard! While the validity of this story is certainly suspect, what can be said is that revolutions in the understanding of numbers has frequently been met with substantial resistance.

The real numbers have also had their share of drama! Let's just say they are a bit weird and philosophically more challenging than you might actually think on first experience. Some of the most contentious

issues in defining modern mathematics have come from attempts to understand the set of real numbers. In particular are the notions of *limits* and *completeness*, topics that we will touch on (lightly) later. Before moving onto other topics, the following example illustrates how the concept of real numbers is more complex than it may appear on the surface.

ital Example 1.1 ($0.999\dots = 1$). We have all been exposed the idea of repeating decimal numbers, but the characteristics of such numbers can be elusive. For example, consider the sequence of numbers (we will define sequences more formally in §1.5) that approach the value 1 as follows, $s = \{0.9, 0.99, 0.999, 0.9999, \dots\}$ or more generally $s = \{0.(9)_n, n \in \mathbb{N}\}$. Clearly this sequence gets closer and closer to 1 as we add more repeats of the numeral 9 (i.e., increase n). Also, if we fix any small number, ε , then, no matter how small ε is, we can always take n large enough so that we get closer than within ε to the value 1.

$$1 - 0.(9)_n < \varepsilon \quad (1.1)$$

In some sense, then, is $0.(9)_n$ equal to 1? This seems like it would be a curious thing, but we can show that it is true. The following is an informal illustration. There are proper, concrete proofs for the somewhat *ad hoc* illustration below. Nonetheless, the point is made, and the result is correct, as odd as the result may seem!

$$\begin{aligned} x &= 0.9999\dots \\ 10x &= 9.999\dots \\ 10x &= 9 + 0.9999\dots \\ 10x &= 9 + x \\ 9x &= 9 \\ x &= 1 \end{aligned} \quad (1.2)$$

The explanation for this involves the idea of convergence of a sequence and completeness of a set to fully understand. For now, we will put this notion aside and simply recognize that real numbers (particularly those that are irrational or are repeating decimals) are trickier to understand than rational numbers are! There is even some more philosophical discussion about in what sense real numbers are (objectively) real. Because of the various concepts of infinity associated with the real numbers, they sometimes illustrate behavior that is counterintuitive (the fact that $0 \times \infty$ is not defined is one such example; it is easy to construct examples where 0 is equal to any number that you choose).

1.1.1 Complex Numbers

Complex numbers are purely mathematical objects. That does not mean that they are not useful or interesting, however. There are many examples of concepts that exist only mathematically, but are nonetheless useful for

many practical applications. Although the complex numbers involve some concepts that do not correspond directly to physical reality, they nonetheless have many useful applications.

The complex numbers are just an extension of the real numbers. Although almost everyone reading this text has probably encountered them previously, it is useful to recap their basic properties.

Definition 1.2. A *complex number* assumes that there exists a mathematical object, called the unit imaginary number, i , such that $i = \sqrt{-1}$, so that $i^2 = -1$. Every complex number consists of two parts, a *real* part, and an *imaginary* part that is proportional to i . The conventional form for a complex number is $a + bi$, where a and b are real numbers. The set of all complex numbers is usually denoted by \mathbb{C} .

The rules for addition, subtraction, and multiplication are slightly modified from those of real numbers as follows. First, addition and subtraction are done by adding and/or subtracting the real and imaginary parts of a complex number independently. Therefore

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

$$(a + bi) - (c + di) = (a - c) + (b - d)i$$

Multiplication of two complex numbers is defined as follows

$$(a + bi) \times (c + di) = ac + bci + adi - bd$$

or, equivalently, grouping terms

$$(a + bi) \times (c + di) = (ac - bd) + (bc + ad)i$$

The complex numbers are often represented as vectors on a plane, where the real part is plotted on the horizontal axis, and the complex part on the vertical axis. In this representation, a complex number would be represented by a pair of points, i.e., $x + yi \Leftrightarrow (x, iy)$. There are some advantages and disadvantages of this, but it generally improves interpretation for problems with physical significance. When this formalism is adopted, a complex number is frequently represented typographically as a vector, as in $\mathbf{z} = (x, iy)$. The representation of the complex numbers in this manner requires mildly re-defining the vector dot product. We define this in the following.

Definition 1.3 (The Complex Conjugate). For a complex number $\mathbf{z} = x + yi = (x, iy)$, the *complex conjugate* is defined by $\bar{\mathbf{z}} = x - yi = (x, -iy)$.

Definition 1.4 (The Complex Dot Product). For two vectors $\mathbf{z}_1 = (x_1, iy_1)$ and $\mathbf{z}_2 = (x_2, iy_2)$, the complex dot product between \mathbf{z}_1 and \mathbf{z}_2 is given by

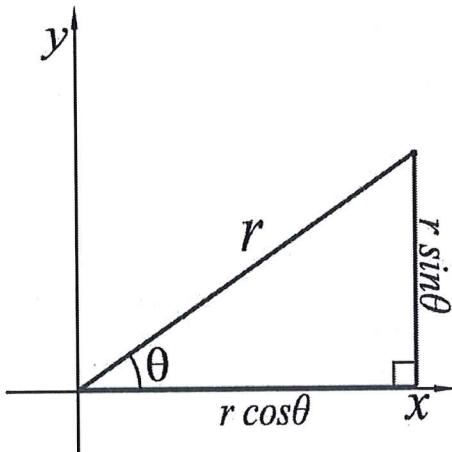
$$\mathbf{z}_1 \cdot \bar{\mathbf{z}}_2 = (x_1 x_2 + y_1 y_2)$$

which has the advantage of being a single real number. In particular, this means that a vector dotted with itself is defined by

$$\mathbf{z} \cdot \bar{\mathbf{z}} = (x, iy) \cdot (x, -iy) = x^2 + y^2$$

and the magnitude of a complex vector is defined by a rule that looks much like the rule for computing the length of a vector on a plane

✓**Fig. 1.2** The complex plane. Complex numbers can be given a meaningful interpretation as vectors on a plane.



$$\|z\| = \sqrt{z \cdot \bar{z}} = \sqrt{x^2 + y^2}$$

The recognition that complex numbers can be treated as vectors on a plane implies that they can be conveniently represented in polar coordinates. In fact, there are some significant reasons for doing so. To start, define the length of the vector by the number $r = \|z\| = \sqrt{x^2 + y^2}$. Then, we have the following relationships between the (x, y) and (r, θ) coordinate systems (Fig. 1.2).

$$r = \sqrt{x^2 + y^2} \quad x = r \cos(\theta) \quad iy = ir \sin(\theta) \quad (1.3)$$

or,

$$z = r(\cos(\theta), i \sin(\theta)) = r(\cos(\theta) + i \sin(\theta)) \quad (1.4)$$

You may have seen *Euler's formula* before; it is the formula that leads to the famous relationship $e^{i\pi} = -1$ (where, recall, e is the base for the natural logarithm). More generally, however, Euler's formula takes the form

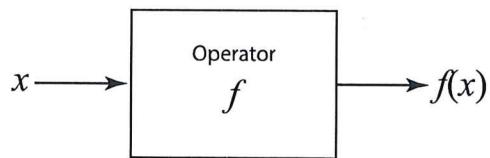
$$e^{i\theta} = \cos \theta + i \sin \theta \quad (1.5)$$

Although this formula looks truly remarkable, once we accept that the imaginary numbers are an acceptable extension of the reals, the proof of this result becomes quite simple (we will examine that further when we tackle infinite series). For now, we adopt the formula without proof. However, note that it allows a particularly simple representation of a complex number z as given by Eq. (1.4) in the form

$$z = re^{i\theta} \quad (1.6)$$

Although complex numbers have a helpful representation as vector quantities on the complex plane, they are technically just an extension of the real number system. Therefore, it is not common to adopt a bold-face type to represent them (as we have done above). In general, complex numbers are set in regular, italicized script, (e.g., z). Generally, the context prevents there from being any confusion. In future uses of complex

 **Fig. 1.3** An function, in conceptual form. A function can also be considered a type of operator.



 numbers, we will not use bold-faced script to represent them. Thus, the equation above would be more properly written

$$z = re^{i\theta} \quad (1.7)$$

1.2 Functions

1.2.1 Definitions

We all have been introduced to the concept of functions. When most people in college mathematics think of a function, the first thing that comes to mind is a relationship that looks something like

An: Fig. 1.5 is called
out before Figs. 1.3 + 1.4

Renumber?

In this description, we are given a *domain* (the numbers such that $-2 \leq x \leq 2$, recall, square brackets indicate *inclusive* interval notation), and a *range* for the function (the values $f(x)$ for the defined domain). A graph of this function is plotted in Fig. 1.5. Although this is how we typically think of a function, we can define it more generally.

Definition 1.5. A *function* (or sometimes *mapping*) is a relationship between two sets, A and B , such that each element of A is *uniquely* associated with an element of B .

The word *uniquely* appearing in our definition here is really important: a function is always single-valued. This definition needs a bit of additional explanation for our purposes. First of all, for us, the sets involved for a function are almost always intervals of the real number line. We even give these functions additional names and descriptions, as follows. We usually refer to functions as having *independent* and *dependent* variables. The independent variables are the *inputs* to the function (or more properly, the set that is input to the function; this is the set A in the definition). Conversely, the *dependent* variables are the set of values that are produced by the set of independent variables upon application of the function (this is the set B in the definition). Each member of the set of independent variables is mapped to *exactly one* member of the set of independent variables (this is the *uniquely* part of the definition). Fig. 1.3 gives a pictorial representation of the process.

The concept of functions that strictly increase or decrease is one that is used routinely in mathematics, so much so that they have a special name: *monotonic*. These functions are defined as follows.

Definition 1.6. A function $f(x)$ is said to be *monotonically increasing* if the following are true:

91
91
91
91
91
1. $f(x_1) \leq f(x_2)$ for all $x_1 < x_2$.

91
91
91
91
91
2. $f(x)$ is not the constant function ($f(x) = C$, C a constant).

91
Similarly, a function $f(x)$ is said to be *monotonically decreasing* if the following are true:

91
91
91
91
91
1. $f(x_1) \geq f(x_2)$ for all $x_1 < x_2$.

91
91
91
91
91
2. $f(x)$ is not the constant function.

Obviously, functions in general are neither monotonically increasing nor decreasing, but a combination of these two concepts over subdomains of their total domain. The constant function is neither monotonically increasing nor monotonically decreasing on any subdomain of its domain.

In science and engineering, we usually think of the sets involved as being intervals of the real number line. To put this in context, think of our example given by Eq. (1.8) above. Here, we can think of the set A as being $A = \{x : x \in [-2, 2]\}$ (note: this is read as follows “The set A is defined by the all of the numbers x such that x is between -2 and 2”). The set B we can think of as all of the numbers defined by the function f ; in set notation $B = \{f : f(x) = x^2 + 2 \text{ for all } x \in A\}$. Note also that these intervals are either *closed* or *open* intervals. The concepts of *closed* and *open* intervals (or, generally, sets) is actually much deeper and complex than one would assume. For our purposes, we can think of a closed interval as one that contains its stated endpoints, and an open interval as one that does not contain its stated end points. This can be made more clear by an example.

ital // *minus* // *3/* *91* *#1*
Example 1.2 (Domains of a Function as an Interval). The function

$$f(x) = \frac{\sin(x)}{x}, \quad x \in (0, \pi] \quad (1.9)$$

has the domain $D = (0, \pi]$. The function is monotonically decreasing on this domain. While the function may not seem particularly unusual, it actually does have some strange behavior at $x = 0$. A plot of the function is given in the figure below.

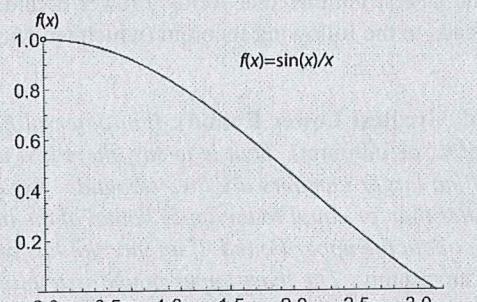


Fig. 1.4 The function $f(x) = (\sin x)/x$.

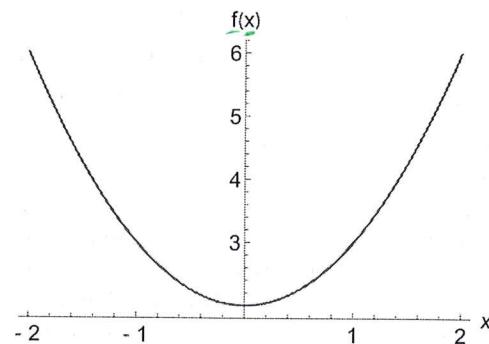
91

*Au: Fig. 1.4
not called out
in the text*

bf 91
91 *a+1*
91
This function is perfectly well-defined at all points between 0 and π (that is, $D = (0, \pi]$), but it is undefined at $x = 0$. Thus, the domain for this function can be stated as (and here we give three different statements, just for practice; they all indicate exactly the same thing!) *the following*

i.e.

✓ Fig. 1.5 A function.



The domain of this function is $0 < x \leq \pi$

The domain of this function is the interval $(0, 1]$

The domain is the set X , where $X = \{x : 0 < x \leq \pi\}$

One difficult concept arises in this kind of description, and this is encountered when you ask the following question: What is the *minimum* value of the set forming the domain? Technically, the domain of this function has no minimum value! The number 0 is not in the domain (hence, the “(” in the domain description). However, for any small number ϵ , say $\epsilon = 1 \times 10^{-100000000}$, the value of the function $(\sin x)/x$ is very near 1. It is only undefined at *exactly* $x = 0$. In cases like this, where we can't technically use *maximum* or *minimum*, we say instead that the number 0 is the *greatest lower bound* or *infimum* of the domain.

9/
5//

The example above suggests that identifying the smallest number in an open set is not generally possible. However, one can often identify the largest number that forms a lower bound for the set (in the example above, this is the number 0). This leads to the following theorem (which we will not prove!), which we will use primarily as a definition.

Theorem 1.1 (Least Upper Bound, Greatest Lower Bound). *If an interval has any lower bound, then it also has a greatest lower bound (g.l.b., or infimum). That is to say, there is a unique number, z_- , which is a lower bound for the interval such no larger numbers are lower bounds. Any numbers larger than z_- are either (i) in the interval, or (ii) greater than or equal to an upper bound of the interval.*

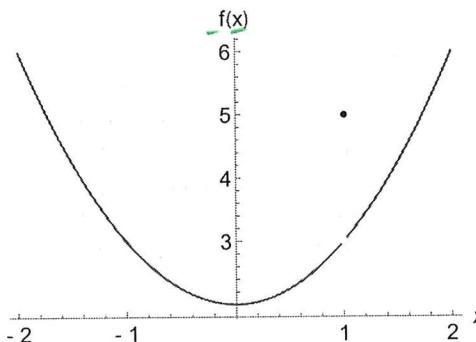
A similar statement can be made about the upper bound. If an interval has any upper bound, then it also has a least upper bound (l.u.b., or supremum). The least upper bound is a unique number, z_+ , which is an upper bound for the interval such no smaller numbers are upper bounds.

To help make this more concrete, consider the interval we examined above: $(0, 1]$. For this interval, the greatest lower bound is the number 0, even though the number 0 is not in the interval itself. The least upper bound is the number 1. No numbers, z_0 , smaller than 1 are, obviously, upper bounds for the interval, because $z_0 < 1$ and 1 is in the interval by definition.

that/
Y/
+/
that/
9//

9/

Fig. 1.6 A function. This function has a single point of discontinuity.



Gal x2

1.2.2 Continuity

Functions are not necessarily always continuous. For example, consider the following function. Assuming $x \in [-2, 2]$, define the function

$$f(x) = \begin{cases} 5 & x = 1 \\ x^2 + 2, & \text{otherwise} \end{cases} \quad (1.10)$$

This function is plotted in Fig. 1.6. Although it is true that this function requires some extra handling (for the point $x = 1$), the end result is not too dissimilar to what we are used to seeing. However, it does indicate that it might be useful to further characterize functions on the basis of how continuous and smooth they are. A few comments about the various labels that are applied to functions to describe how smooth they are are provided in the following.

A function $f(x)$ is *continuous* if it changes gradually as the independent variable x changes. For a one-dimensional function, one can think of this as being able to draw the function with pen and paper without taking one's pen off the paper. More formally, using the dreaded δ - ε arguments, one makes a more concrete statement. Unlike most δ - ε arguments, we are going to do this in slow steps. First, suppose we are interested in the continuity of a function at a point $x = a$. Now, we look at all of the points within a small distance, $\pm\delta$, around a . We can denote those points by all of the values for x such that $a - \delta < x < a + \delta$, or, equivalently, $|x - a| < \delta$. Now, for all of such points that we find on the x -axis (the domain), we can compute the absolute value $|f(x) - f(a)|$. Suppose we do this, and we find that there is a number $\varepsilon > 0$ such that

$$|f(x) - f(a)| < \varepsilon \text{ for all } x \text{ such that } |x - a| < \delta \quad (1.11)$$

Now, for a continuous function, we expect that as we squeeze the interval $a - \delta < x < a + \delta$ by making δ smaller, we should also be able to correspondingly find a new value for ε that meets the criterion given in (1.11). In other words, as $|x - a|$ gets smaller, so does $|f(x) - f(a)|$. This is essentially the way that continuity of a function is defined. Formally, the argument is turned inside-out (which is common in the presentation of definitions and proofs in mathematics), to read as follows.

Definition 1.7. A function, f , is *continuous* at a point $x = a$ if and only if (i) a is in an open interval of the domain of f , and (ii) we can always pick an ε (as small as we like), and still always find some interval around a (that is, $a - \delta < x < a + \delta$) such that $|f(x) - f(a)| < \varepsilon$, regardless of how small we have chosen ε to be.

Next

01

may

7/31
7/1

01

E901

9/1
9/1
i.e. 1

Definition 1.8 (Continuous Functions). A function is *continuous* if it is continuous for *each open sub-interval* of its domain.

31

In essence, the “for each open sub-interval” part of the definition is to avoid any possible problems with defining continuity at the least upper or greatest lower bound of the interval. In short, we don’t need to worry about continuity at the end points of our domain.

These definitions can be a bit tricky to apply, and they are not the only way to define continuity. Alternatively, one can insist that the following limits, taken from the left and right sides of a must be valid for each point in a domain for a function to be considered continuous.

$$\lim_{x^+ \rightarrow a} |f(x) - f(a)| \implies 0 \quad (1.12)$$

$$\lim_{x^- \rightarrow a} |f(x) - f(a)| \implies 0 \quad (1.13)$$

And sometimes this alternative is easier to use in applications. Note \swarrow The long double arrow (\implies) in mathematics should be interpreted as meaning “this implies” \circlearrowleft $\text{that } (1c)$ \oplus if

Regardless of which method is used, the results are the same. This is easiest to see with a simple example, as illustrated in Fig. 1.7.

Example 1.3 (Continuity of Functions).

In Fig. 1.7, it is apparent from inspection that the function $g(x) = \frac{1}{2}x^2$ (the orange line on the Fig. 1.7) is continuous. Suppose we consider the point $x = 1$. Clearly, we can pick a small number, say $\epsilon = 1/1000$, such that there is a small enough interval around $x = 1$ ($1 - \delta < x < 1 + \delta$) where $|f(x) - f(1)| < 1/1000$. In fact, we can compute what this interval is as follows \downarrow

$$|f(1 + \delta) - f(1)| < 1/1000$$

Accounting for the absolute value, there are two options \downarrow

$$\frac{1}{2}(1 + \delta)^2 - \frac{1}{2} < 1/1000 \quad \text{or}$$

$$\frac{1}{2} - \frac{1}{2}(1 - \delta)^2 < 1/1000$$

Solving these inequalities for δ , we find

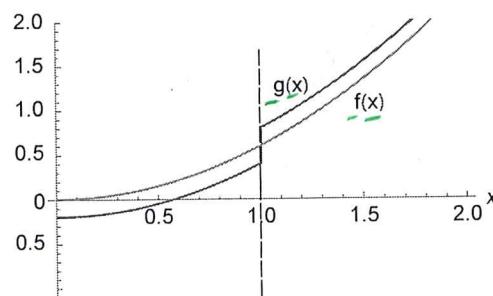
$$|\delta| < \sqrt{\frac{1002}{1000}} - 1$$

$$|\delta| < 1 - \sqrt{\frac{998}{1000}}$$

The first of these is the smallest, so it is safe to take δ equal to that value.

You can check this directly by noting that

Fig. 1.7 Two functions, f and g . The function g has a discontinuous jump in the value of $\frac{1}{2}$ at $x = 1$.



$$\left| f\left(1 + \sqrt{\frac{1002}{1000}} - 1\right) - f(1)\right| = \frac{1}{1000}$$

$$\left| f\left(1 - \sqrt{\frac{1002}{1000}} - 1\right) - f(1)\right| = 0.000999 < \frac{1}{1000}$$

No matter how small we make ε , we can always find a value of δ so that the inequality is valid.

Au: Change to match usage above.

Now consider the function f (the blue line on Fig. 1.7). Here, our scheme works reasonably well until we set ε to be less than about $1/2$. Because there is a jump of $1/2$ right at $x = 1$, we can easily find an ε that breaks our definition. For example, set $\varepsilon = 1/10$. Now we are looking for the values of x in an interval $a - \delta \leq x \leq a + \delta$, such that $|f(x) - f(a)| < 1/10$. Of course, we can see just by looking at the graph that there is no such interval. In any small interval around $x = a$, the value of $|f(x) - f(a)|$ is at least $\frac{1}{2}$, but definitely never smaller than that value.

The definition of continuous, then, aligns with our intuitive notion, even if the definition itself takes a bit of thinking to fully understand.

Functions that have a jump in them, such as the one in Fig. 1.7 are called *discontinuous*. When there are a *finite* number of discontinuities, sometimes the functions are called *piecewise continuous*. In other words, a piecewise continuous function is a function that has a finite number of jumps in it and doesn't blow up to $\pm\infty$ anywhere. This is an important enough concept that it deserves a specific definition.

Definition 1.9 (Piecewise Continuous Function). A *piecewise continuous function* is a function that is *continuous* everywhere, except at a *finite* number of points.

Within continuous functions, the functions are sometimes *smooth* by their derivatives. As an example, examine Fig. 1.8. This function is defined by

$$f(x) = \begin{cases} 2x + 2 & 0 \leq x < 2 \\ 4(x - \frac{1}{2}) & 2 \leq x < 3 \\ 10(x - 2) & 3 \leq x \leq 4 \end{cases} \quad (1.14)$$

In general, the word *smooth* is used to indicate a function that has a derivative at each point in its domain. Thus, the function given by Eq. (1.8) is smooth; the function given by Eq. (1.10) is technically *non-smooth*,

Fig. 1.8 A continuous function with discontinuous derivatives at two points. This kind of function is sometimes denoted C^0 to indicate that it is continuous, but it does not have derivatives at all points in the domain.

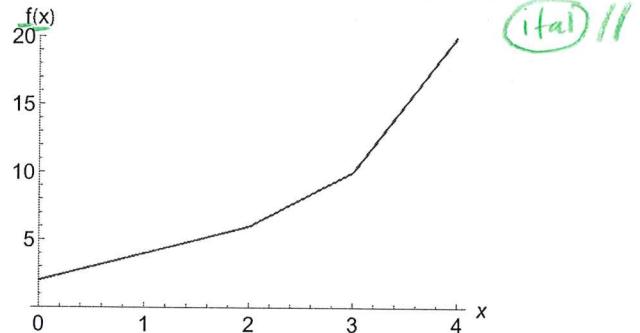
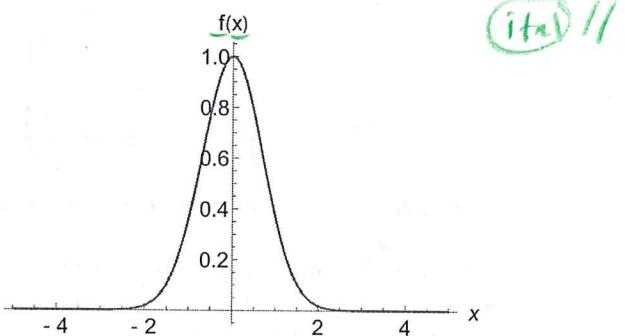


Fig. 1.9 A *really* continuous function. This function has derivatives of all orders; thus is called a C^∞ function.



but this arises because of the discontinuity imposed by a single point. Extending the idea of piecewise functions, we can call the function given by Eq.(1.17) *piecewise smooth*. Most of the functions that we study in this text will be of this kind.

91
Definition 1.10 (Piecewise Smooth Function). A *piecewise smooth function* is a function where the first derivative of the function is well-defined (i.e., it exists, and it is not infinite) everywhere, except at a *finite* number of points.

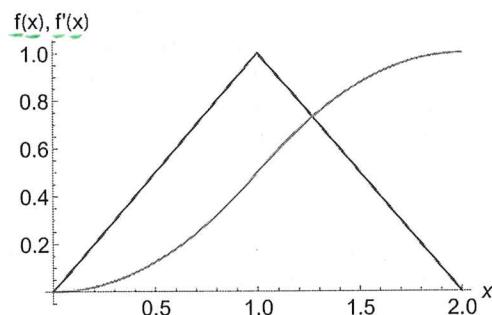
91
We can further characterize smoothness by how many derivatives a function has. If a function has an infinite number of derivatives that exist, then it is called C^∞ . An example here is the function $f(x) = \exp(-x^2)$. This function is smooth (it is plotted in Fig. 1.9), and is often called a *Gaussian* function. We can define particular C^∞ functions that are called *analytic* functions; however this will have to wait until after the review of derivatives and integration is presented.

31
Some functions have only a limited number of derivatives that exist. As an example here, consider the function

$$f(x) = \begin{cases} \frac{x^2}{2} & 0 \leq x < 1 \\ -\frac{x^2}{2} + 2x - 1 & 1 \leq x \leq 2 \end{cases} \quad (1.15)$$

201
This function looks innocuous enough—it is plotted as the orange curve in Fig. 1.10. However, its derivative (plotted in blue) has a cusp in it at the location $x = 1$. Therefore, although its first derivative exists and

Fig. 1.10 A continuous function (orange) with continuous first derivatives (blue). However, the second derivatives are not continuous at $x = 1$. This function is called a C^1 function to indicate that it has continuous first derivatives everywhere in the domain.



ital x4

is continuous, its second derivative is not defined at $x = 1$. To reflect this, functions with one continuous derivative are called C^1 . Although we have not yet reviewed derivatives, we note the following generalization of this idea.

rom cap X3
not subscript
91

Definition 1.11 (n^{th} -order continuous function). Suppose a function f is such that the first n derivatives (i.e., f' , f'' , $f''' \dots f^{(n)}$) are both bounded and exist (i.e., there are no points such that the derivative of order $(n-1)$ generates a cusp or other discontinuity). Then the function is continuous in the derivative up to order n , or, more simply, C^n .

51

There is a special class of functions called *analytic* that have a number of interesting properties. Many of the familiar functions that we know about are analytic. Examples include *the following*

1. All polynomials of finite degree.
2. The exponential and logarithmic functions.
3. The trigonometric functions.

An: Please fill in.
~1~
~m/
~an/
~0/
An: As meant?
neither/
res/

In the early days of mathematical evolution (say, through the early 1900s), analytic functions were synonymous with *functions*. Since that time, the notion of what constitutes a function has grown considerably; some of the unusual examples of functions that are not analytic in some part of their domain (or everywhere, in the case of the Thomae function given in Fig. ??) are presented in the material later in the text.

Definition 1.12 (Analytic Function—Definition 1). Suppose a function f is defined on a domain D . The function f is *analytic function* everywhere in D if for every closed interval K that is a subset of D ($K \subset D$) there exists a constant C such that for every point $x \in K$ and every non-negative integer k , the following bound holds

$$\left| \frac{d^k f}{dx^k}(x) \right| \leq C^{k+1} k! \quad (1.16)$$

While this definition does make a mathematical mouthful, it is relatively easy to understand intuitively. What the definition is trying to tell us is that, for any closed finite interval that is part of the domain of the function, that the function, nor any of its derivatives, go to infinity in that domain. There are some technical issues regarding the domain of the function (and the closed subsets that are selected from it) that will not be covered here. In general, however, we can think of a function as being analytic at a point in its domain if its value and the value of all of its derivatives do not tend toward infinity.

As a final note about the classification of functions, there is one additional classification that is useful to know about. There are, in one sense, two different kinds of functions that we use in common practice. There are polynomial functions, and the roots of polynomial functions to start with. A polynomial function is a polynomial (of any finite order) whose coefficients are also polynomials. For example, a polynomial function of order 4 is defined by

$$a_4(x)x^4 + a_3(x)x^3 + a_2(x)x^2 + a_1(x)x + a_0(x) = 0$$

For such a function, the roots generate new functions, $f(x)$, involving (rational) fractional powers (including negative powers). Such functions are called *algebraic* because they can be defined using only the rules of algebra applied in a *finite* algorithm. The following are all examples of algebraic functions.

$$\begin{aligned} f(x) &= x \\ f(x) &= x^{-1} = \frac{1}{x} \\ f(x) &= \sqrt{x} \\ f(x) &= \frac{\sqrt{1+x^3}}{x^{3/7} - \sqrt{7}x^{1/3}} \end{aligned}$$

The nice thing about these functions is that they can be entirely described mathematically simply by describing the polynomial that generates them. Algebraic functions are expressions involving only a finite number of terms, and using only the algebraic operations addition, subtraction, multiplication, division, and raising to a rational fractional power. In a sense, we fully “understand” these functions as long as we understand the operations that define them. This is literally constructive— for many examples of algebraic functions, we can give someone an algorithm (with a finite number of terms) that explains what the function is and how to compute its values. Importantly, all algebraic functions are given by the roots of some polynomial; however, for rational polynomials of degree 5 or higher, it is not true that all such polynomials have roots that are algebraic functions (the proof of this is known as the *Abel-Ruffini theorem*).

It turns out, however, that there are many interesting and useful functions that are not the roots of any polynomial equation. Familiar examples include the $\sin x$, $\cos x$, $\ln x$, and e^x functions. Such functions are called (and, yes, this is really the name) *transcendental functions*. The idea behind the name is that these functions transcend description by the discipline that we normally call algebra. Transcendental functions will show up later in our studies of differential equations, and some will prove to be essential for describing solutions to such equations. All transcendental functions are *analytic* (a term we will define in by the Taylor series in detail later), so they all have convergent Taylor series representations.

There are functions that are even more *interesting* to attempt to define because their structure begins to challenge the concept of function altogether. Consider the following function (sometimes called the modified Dirichlet function or the Thomae function)

$$f(x) = \begin{cases} 0 & x = 0 \\ \frac{1}{q} & x \in \mathbb{Q}, x = p/q \\ 0 & x \in \mathbb{P} \end{cases} \quad (1.17)$$

It is understood here that $x = p/q$ is expressed in fully reduced fraction form. This is a really unusual function, and clearly it has no simple, closed algebraic formula. A plot of this function appears as Fig. 1.7. In addition to its strange definition and look, it also has some other unusual characteristics. It turns out that

the real numbers are “denser” than the rational numbers. In a sense, there are more (many, many more) irrational numbers than rational ones. Thus, this function is continuous (and has a derivative) at each point where x is an irrational number. It is discontinuous at every point where x is rational. It is not terribly important to understand from this example all of the details of why this kind of behavior exists (although our discussion of intervals above is suggestive), but the essential idea is to build an intuition that the real (irrational) numbers constitute a much larger (denser) set than the rational numbers do.

There is one final class of functions that does not really fit naturally into the descriptions above. These are the *homogeneous* functions. They are defined as follows.

Definition 1.13 (Homogeneous function). a homogeneous function is one with multiplicative scaling behaviour: if all its arguments are multiplied by a factor, then its value is multiplied by some power of this factor. In other words, homogeneous functions display the following behavior

$$f(\alpha x) = \alpha^n f(x)$$

Here, n is a real number, and it called the degree of the function.

Example 1.4 (Homogeneous functions.). Below are a few examples of homogeneous functions.

1. All polynomials of the form $f(x) = x^n$ are homogeneous. To see this, just note the following: $f(\alpha x) = (\alpha x)^n = \alpha^n x^n$. In this case, the degree of the function is equal to the degree of the polynomial.
2. No functions with additive constants are homogeneous. This is also easy to demonstrate. Suppose $f(\alpha x) = \alpha^n f(x)$ is a homogeneous function. Now consider the function $g(x) = f(x) + c$. The function $g(x)$ cannot be homogeneous, because a multiplicative scalar α will not scale the constant c . Functions like $g(x)$ (i.e., ones that contain an additive constant) are called *affine*. This just means that the two functions are connected by a linear scaling of the coordinate (the independent variable) followed by a translation (the additive scalar).
3. The function $f(x) = 1/x^2$ is homogeneous. To see this, note

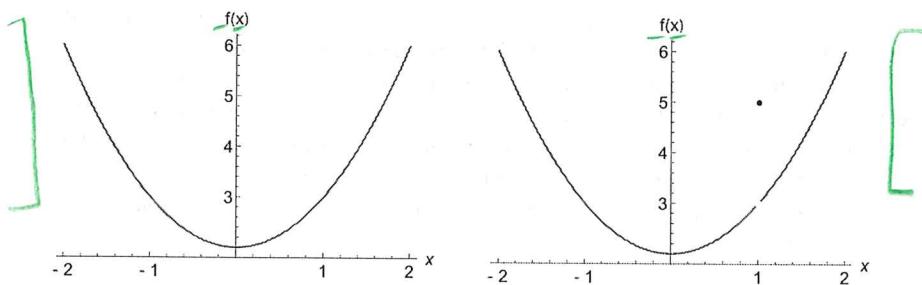
$$\begin{aligned} f(\alpha x) &= \frac{1}{(\alpha x)^2} \\ &= \frac{1}{\alpha^2} \frac{1}{x^2} \\ &= \alpha^{-2} x^{-2} \end{aligned}$$

So, this function is homogeneous, with degree equal to -2 .

1.2.3 Equivalence Classes of Functions

Later on, when we discuss the concept of Fourier series, it will be helpful to understand the notion of *equivalence classes* of functions. In short, two functions f and g are considered to belong to the same

An: Center
Fig. 1.11



b) 91

Fig. 1.11 Two functions that belong to the same equivalence class. The second differs from the first at only a single point.

equivalence class if they differ from each other by, at most, a finitely countable number of points. An example of two functions that are of the same equivalence class is given in Fig. 1.11. The mathematical representation of these two functions is given by

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

$$g(x) = \begin{cases} x^2 & \text{for } x \neq 1 \\ 5 & \text{for } x = 1 \end{cases}$$

Note that we have the following limits for g (the function on the right-hand side) near $x = 1$

$$\lim_{x \rightarrow 1^+} g(x) = 1$$

$$\lim_{x \rightarrow 1^-} g(x) = 1$$

where 1^+ indicates the limit coming from the right-hand side, and 1^- represents the limit coming from the left-hand side. Be sure to recall here that the limit exists in the sense of *approaching* the value $x = 1$, but not actually reaching that value. If you need a refresher on one-sided limits, your undergraduate calculus text will definitely cover this material. Alternatively, you can look on-line, for instance at "Paul's Online Notes" (e.g., for one-sided limits, <http://tutorial.math.lamar.edu/Classes/CalcI/OneSidedLimits.aspx>).

It turns out that, from the perspective of integration (including Riemann integration, see §1.4.3), the presence of a single discontinuous point does not affect the result of the integration. Speaking colloquially, a single point has no measure, so it does not add or subtract from the integral. The following definitions makes this more formal.

Definition 1.14 (Removable discontinuity). A function is said to have a set of removable discontinuities if (a) the number of discontinuous points is finite, and (b) the left and right-hand limits at each point of discontinuity are equal. The value of the function at the discontinuous points can be taken to be the value of the limit at that point.

Definition 1.15 (Equivalence class of functions). Two functions f and g are said to belong to the same equivalence class if (a) the two functions f and g have only removable discontinuities, and (b) $f - g = 0$ for

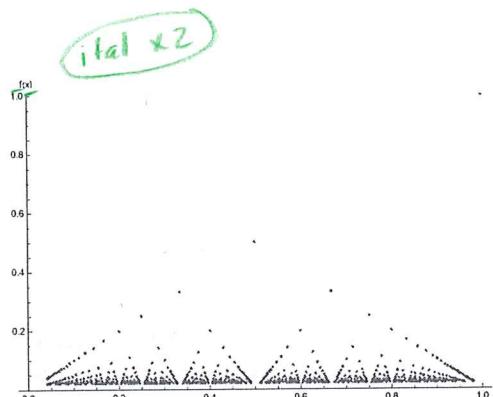
An: Previously we used
(i) and (ii) for inline
lists. Okay?

An; Fig. 1.12 not called out in the text. Also, it is very hard to read.

1.3 Linear Spaces, Linear Operators, and Linear Algebra

Fig. 1.12 A graph illustrating a collection of values for the Thomae function using $2 \leq p \leq 50, 2 \leq q \leq 50$. This may look super weird, but it meets the criteria for a function.

17



(a) all points that are not located at a discontinuity in either f or in g . Alternatively, (b) could be stated by " $f - g = 0$ after all discontinuities are removed".

1.3 Linear Spaces, Linear Operators, and Linear Algebra

We will have only a few occasions to use linear algebra in this material, but it is worth briefly reviewing, and it also allows us to discuss the concept of a *linear operator*. Many of the problems that we encounter in the material to come involve linear operators, so spending some time understanding what they are is worthwhile.

Operators are just what they sound like: They are mathematical constructs that operate on something (say, a function) to generate something else (for example, a different function) (Fig. 1.3). This is a very general concept, and it is difficult to give it a precise meaning. As an example, all functions are operators. To make this concrete, take a look at the following example.

Example 1.5 (Functions as operators).

We are all familiar with the idea of a *function*. In the notation that we have are probably most familiar with, a function is defined as a one-to-one (i.e., unique) mapping between two sets. In a more familiar setting, we think of a function as having a *range* (which is usually an interval on the real line, plotted on the horizontal axis by convention) and a *domain* (which is usually some portion of the real line, plotted on the vertical axis by convention). For example, a properly defined function would be any of the following:

$$f(x) = 5x \quad 0 < x < 10 \quad (1.18)$$

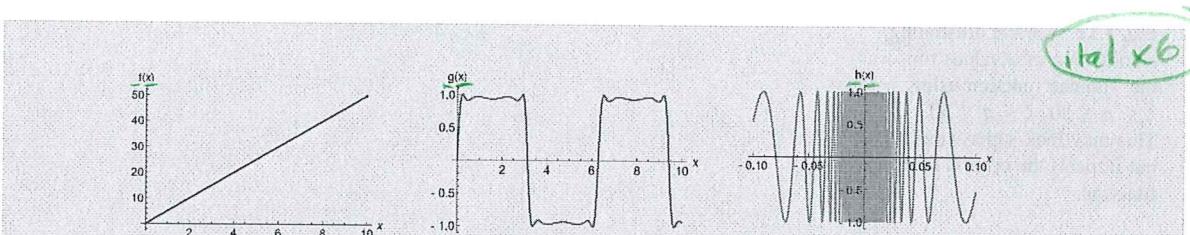
$$g(x) = \sin[2\sin(2\sin\{2\sin(x)\})] \quad \infty < x < \infty \quad (1.19)$$

$$h(x) = \sin[1/x] \quad -0.1 < x < 0.1 \quad (1.20)$$

These functions are plotted in Fig. 1.13.

and/or

An: Fig. 1.13 is very hard to read



ital x6

Fig. 1.13 Figure for Example 1.2. Three functions.

What is important to realize about these functions is that they are operations on their independent variables. In some sense, a function exists *separately* from their independent variables. Thus, once I have defined $f(x)$, as, for example, in Eq. (1.18), then what I have really done is to define an *operation* to be performed on one set (the independent variables) to generate another set (the functional value). Often when we think of functions as being operators, we consider (1) the function to define the operation, and (2) the operation is done on whatever is put into the function. The thing put into the function, when used in this sense, is often called the *argument*. The argument can be a simple independent variable defined over an interval (our conventional use of functions) or some other set.

For example, suppose we define the following set: $X = \{x : x = 1, 3, 4, 5, 42\}$. We would read this in words as "The set X is the set that contains values of x such that x is equal any of the values 1, 3, 4, 5, and 42". The set X has a finite number of values in it, but we can still use any of the functions above as operations on this set. Let's interpret $f(X)$ the following way:

$$f(X) = \{f(1), f(3), f(4), f(5), f(42)\}$$

or, computing the values using the expression $f(x) = 5x$

$$f(X) = \{5, 15, 20, 25, 210\}$$

An: This does not make sense. Please clarify.

We can even think about putting functions in other functions (creating a as arguments when we use the operator idea). As an example, consider the function $p(z) = z^2 + 1$. Now, can we make sense of the operations $f(p)$ and $h(p)$? Sure, we need only apply the operations to the argument of the function, regardless of what the argument is.

$$\begin{aligned} f(p) &= 5p(z) \\ &= 5(z^2 + 1) \\ &= 5z^2 + 5 \end{aligned}$$

$$\begin{aligned} h(p) &= \sin(1/p(z)) \\ &= \sin\left(\frac{1}{z^2+1}\right) \end{aligned}$$

When we interpret a function $f(x)$ this way, we sometimes write the function without the independent variable (with the idea that the independent variable can be anything), as in simply f . This is just a notational device used to be more compact, there is usually no deeper meaning associated with it.

1.3.1 Linear operators

Now that we understand the basic idea of an operator, we can consider what a *linear operator* is. This is actually pretty simple at this point.

Definition 1.16 (Linear Operator). A *linear operator*, \mathcal{L} , is an operator that subscribes to the properties of *additivity* and *homogeneity*. Specifically, this means *the following*

$$\text{additivity: } \mathcal{L}(f+g) = \mathcal{L}(f) + \mathcal{L}(g) \quad (1.21)$$

$$\text{homogeneity: } \mathcal{L}(\alpha f) = \alpha \mathcal{L}(f) \quad (1.22)$$

Linear operators are a subset of all possible operators. In fact, most interesting phenomena in engineering and physics are nonlinear in general. Linearity exists only as an approximation to the more general nonlinear behavior. In this text, we will be generally concerned with (although not exclusively!) linear operations and linear operators.

As discussed above, functions can themselves be thought of as operators. Thus, a linear function would subscribe to the properties of additivity and homogeneity above. As an example, the following function is linear

$$f(x) = 5x \quad (1.23)$$

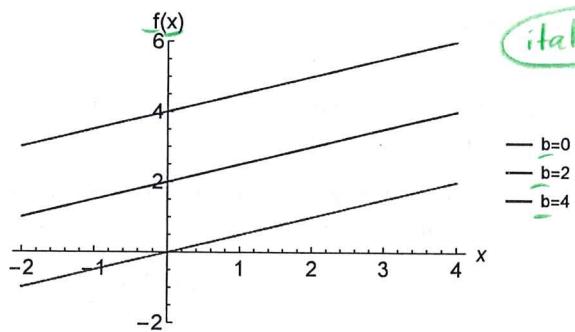
To show that it is linear, note that we can prove both additivity and homogeneity in one step. If we make the substitution $x \rightarrow \alpha y + \beta z$, then we find

$$\begin{aligned} f(\alpha y + \beta z) &= 5(\alpha y + \beta z) \\ &= \alpha 5y + \beta 5z \\ &= \alpha f(y) + \beta f(z) \quad (\because \text{linear}) \end{aligned} \quad (1.24)$$

illustrating that the function is indeed linear.

There is a particular feature of linearity that can create some confusion. We are used to calling functions such as $y = mx + b$ *linear*. However, it is not difficult to show that this function does not meet the properties of additivity and homogeneity (try it!). There is a subtle reason for this failure. In a very real sense, $y = mx + b$ is actually the *translation* of a more fundamental function $y = mx$. Take a look at Fig. 1.14. Each of the functions illustrated is a translation of the function $y = \frac{1}{2}x$. Such translations are called *affine transformations*. It turns out that when we call a function *linear*, we actually mean that the fundamental function itself (where $b = 0$) is linear. The constant term only represents an affine transformation of this more fundamental function. In fact, in each of the lines defined in Fig. 1.14, we could eliminate the constant term b by simply making a transformation of the coordinate system. For example, if we moved the coordinate system vertically by a distance of 2, then the blue line in the figure would then have $b = 0$. Thus, in a sense, all affine transformations of a function are the same as far as *linearity* is concerned. In order to assess linearity, the first step would be to first make an affine transformation of the coordinate system so that $b = 0$. In a more practical sense, we can essentially ignore constants when checking operators (including functions) for linearity.

✓
Fig. 1.14 Translations of the function $y = \frac{1}{2}x$.



(cap x4) / E / Example 1.6 (A familiar example: The derivative as an operation). The process of differentiation is one example of a familiar operation. Consider, for example, the following.

$$\mathcal{L}(f(x)) = \frac{d}{dx}(f(x))$$

9/ In this example, the operator is defined by derivative notation $\frac{d}{dx}$. When we say $\mathcal{L} \equiv \frac{d}{dx}$, we are just defining the fact that the “abstract” operator in this case is the derivative operation, denoted by $\frac{d}{dx}$.

(cap x3) / E / Example 1.7 (The derivative as a linear operator). Is the derivative operator a linear one? We can check. Recall the definition.

$$\mathcal{L}(\alpha f(x) + \beta g(x)) = \alpha \mathcal{L}(f(x)) + \beta \mathcal{L}(g(x)) \quad (1.25)$$

Since our operator in this case is just $\frac{d}{dx}$, we need only to check that the linearity identity is met. In other words, we need to evaluate

$$\begin{aligned} \mathcal{L}(\alpha f(x) + \beta g(x)) &= \frac{d}{dx}(\alpha f(x) + \beta g(x)) \\ &= \frac{d}{dx}(\alpha f(x)) + \frac{d}{dx}(\beta g(x)) \\ &= \alpha \frac{d}{dx}(f(x)) + \beta \frac{d}{dx}(g(x)) \end{aligned}$$

This is exactly the form that a linear operation must take as defined by Eq. (1.25), so the derivative is a linear operator.

~ 8 / A compound linear operator is just the sum of one or more linear operators. The easiest way to understand these is through some simple examples. Although we will not use operator notation extensively in this text, it is a useful notation to understand.

(cap x2) Example 1.8 (Compound linear operators). Compound linear operators act as follows.

$$(\mathcal{L}_1 + \mathcal{L}_2)(f(x)) \equiv \mathcal{L}_1(f(x)) + \mathcal{L}_2(f(x)) \quad (1.26)$$

It is important to recognize that, although the notation looks as though it indicates multiplication, it does not! Here is a more complex looking example that helps to understand the idea better.

$$\begin{aligned} \left(\frac{d}{dx} + \frac{d}{dy} \right) (x^2 y) &= \frac{d}{dx}(x^2 y) + \frac{d}{dy}(x^2 y) \\ &= 2xy + x^2 \end{aligned}$$

Note that the operators *do not* act like multiplication!

$$\begin{aligned} \left(\frac{d}{dx} + \frac{d}{dy} \right) (x^2 y + y^3) &= \frac{d}{dx}(x^2 y + y^3) + \frac{d}{dy}(x^2 y + y^3) \\ &= 2xy + x^2 + 3y^2 \end{aligned}$$

(3) Functions can be treated as operators also, but only *homogeneous* functions are linear operators.

(cap E) Example 1.9 (Functions as operators). Consider the functions defined in the previous example.

$$f(x) = 5x, \quad 0 < x < 10 \quad (1.27)$$

$$g(x) = \sin[2 \sin(2 \sin\{2 \sin(x)\})], \quad \infty < x < \infty \quad (1.28)$$

$$h(x) = \sin[1/x], \quad -0.1 < x < 0.1 \quad (1.29)$$

(Au: Please fill in.) Are they linear operators? To check this, we need only try the operations on the quantity $\alpha y + \beta z$ to see if the condition given by Eq. ?? is met. For the generic operator L , we will substitute our particular functional operators. Thus, for the example of f , we have the result

$$\begin{aligned} f(\alpha y + \beta z) &= 5(\alpha y + \beta z) \\ &= \alpha 5y + \beta 5z \\ &= \alpha f(y) + \beta f(z) \end{aligned}$$

(10) Which indicates that f is linear. For the function g , we find

$$\begin{aligned} g(\alpha y + \beta z) &= \sin[2 \sin(2 \sin\{2 \sin(\alpha y + \beta z)\})] \\ &\neq \alpha g(y) + \beta g(z) \end{aligned}$$

so g is not a linear operator (and also not a linear function!). The final function h is left to the reader to check.

There are two more concepts that are useful to introduce when discussing the (somewhat abstract) concept of linear operators. These are the *identity operator* and the *inverse operator*.

Definition 1.17 (Identity operator). The identity operator, \mathcal{I} , is any operator such that $\mathcal{I}(f) = f$ for all admissible objects in the domain of \mathcal{I} (e.g., numbers, functions, vectors, etc.) f .

Example 1.10 (Identity operators). Here are three common examples of identity operators that you have seen before.

1. The number "1" is the identity operator in arithmetic. For any number a , we have $1 \cdot a = a$.
2. We can define an identity function $g(f)$ in a domain $D = \{0 < x < 1\}$ as follows: $g(f(x)) = f(x)$ for all x in D . It is not too difficult to recognize that this operator is equivalent to multiplying a function by the number 1.
3. consider the following matrix multiplication (Note: linear algebra is covered in the next section).

$$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Now, for every possible vector \mathbf{a} of dimension two, we have

$$\mathbf{I} \cdot \mathbf{a} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

Therefore, \mathbf{I} is an identity operator for the domain of vectors of dimension two.

We won't be using the *theorem-proof* format very often in this material, but occasionally it is helpful, especially when the proofs are short and clever. The definition of the inverse operator can be done this way.

Theorem 1.2 (Inverse operator). A linear operator can have an inverse, \mathcal{L}^{-1} , only if $\mathcal{L}(x) = 0$ implies that $x = 0$.

Proof. If $\mathcal{L}(x) = y$ then the inverse of \mathcal{L} is the mapping which takes y back to x . (AN ASIDE: Here, it might be helpful to think of a conventional function, such that $\mathcal{L}(x) = y(x)$, and x is the domain of the horizontal axis. For example, $\mathcal{L}(x) = x^2$ for $0 < x < 1$ associates each independent value in the domain x with a unique value for the result (which we call $y(x)$) in the range.)

Suppose now that $\mathcal{L}(x_1) = y_0$ and $\mathcal{L}(x_2) = y_0$. Then by linearity, $\mathcal{L}(x_1 - x_2) = 0$. One of the following results must be true, either (a) $x_1 - x_2 = 0$ (i.e., x_1 and x_2 are the same value), or (b) $x_1 - x_2 \neq 0$, so that two values of x in the domain of $\mathcal{L}(x)$ that would be mapped to the same value y_0 in the range. For the inverse function, the roles of the domain and range are interchanged. Thus, for the inverse function, y forms the domain, and $\mathcal{L}^{-1}(y) = x$ is the range. However, by definition, a function can have only one value in the range associated with a value in the domain. Thus, the option where x_1 and x_2 are not equal is not possible

(or the inverse would not be a function). Thus we must have that $x_1 = x_2$, and this means $\mathcal{L}(x) = 0$ is only true for $x = 0$. ■

9/

Without getting overly technical, for our purposes an inverse linear operator \mathcal{L}^{-1} will exist when we can show

$$\mathcal{L}^{-1}(\mathcal{L}) = \mathcal{L}(\mathcal{L}^{-1}) = \mathcal{I} \quad (1.30)$$

and that \mathcal{L}^{-1} is never multiple-valued.

As usual, examples really help.

An: Are you able to fill
in this blank space?

Example 1.11 (Inverse operators.).

1. Suppose we have $f(x) = x^2$ for $\{x : 0 < x < 1\}$. Does this function have an inverse?
Try these functions.

$$y = f(x) = x^2 \quad (1.31)$$

$$x = g(y) = \sqrt{y} \quad (1.32)$$

Now, note that $f(g(y)) = f(\sqrt{y}) = y$. But, from Eq. (1.31), we have $y = x^2$; thus, we have $f(g(y)) = x^2$. Now, note that the identity operator is such that $\mathcal{I}(f(x)) = \mathcal{I}(x^2) \equiv x^2$. So, by definition, we must have that $f(g(x)) = x^2 \equiv \mathcal{I}(f(x))$. Therefore, $f(g(x))$ is the identity operator, and, by definition, g is the inverse of f .

Functions that are inverses of one another have an interesting graphical feature. Consider the two functions above; if we plot them both as functions of x (i.e., we plot the functions $y = x^2$ and $y = \sqrt{x}$, we obtain the plot below. Functions that are inverses of one another have reflective symmetry about the line $y = x$.

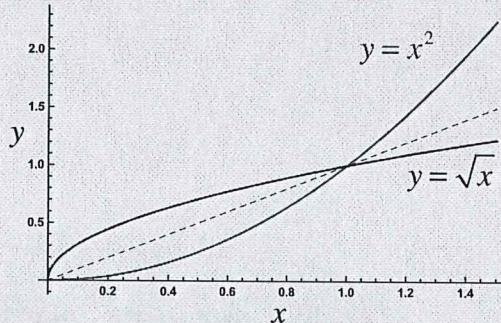


Fig. 1.15: Inverse functions.

Au: Fig. 1.15 is not called out in the text.

There is one caveat that needs to be attended to when discussing the idea of linear operators and linear equations. The idea of *linearity* is usually associated with only the *operator* part of the equation. In particular, what this means is that constants involved in a function are excluded. This is by convention, not necessity. Consider the following linear equation.

$$f(x) = 3x + 12 \quad (1.33)$$

Obviously the operator f must be linear, because $f(x)$ is a line! However, if we attempt our normal process for checking linearity, we encounter a problem.

$$f(\alpha y + \beta z) = 3(\alpha y + \beta z) + 12 = \alpha 3y + \beta 3z + 12 \quad (1.34)$$

$$\neq \alpha f(y) + \beta f(z) \quad (1.35)$$

So, what went wrong? Well, the answer here is a bit tricky. We should, technically, think of constants as not being part of the operator. In other words, suppose we define $L(x) = 3x$. Then, the equation above can be written as follows

$$f(x) = \mathcal{L}(x) + 12 \quad (1.36)$$

(cap) And the actual question can be posed as "is the operator \mathcal{L} linear"? And the answer to that question is obviously yes. The important thing to remember here is that the *linearity of the equation* has only to do with the operators involved, not with any particular constants that may be tacked on to the operators. Thus, a linear equation is any equation whose operator (which, by definition does not include additive constants) is linear.

(cap) 1.3.2 Linear algebra

Now that we know what a linear operator is, we can discuss linear algebra. Most people's experience with linear algebra leaves them without an intuitive notion of what is really going on. Because linear algebra is such a mathematically rich topic, it is the first place where more formal mathematical analysis can sensibly be done in terms of vector spaces; it tends to favor mathematical detail over process. At any rate, for the material covered here, a simple review of linear algebra is sufficient. To start, we will consider a set of linear functions; for the purposes of generating a concrete example, consider the following.

$$f_1(x, y, z) = 3x + 2y + z - 5 = 0 \quad (1.37)$$

$$f_2(x, y, z) = x + y - z - 1 = 0 \quad (1.38)$$

$$f_3(x, y, z) = 2x + 2y + 3z - 10 = 0 \quad (1.39)$$

Obviously this is a set of equations, and obviously they are linear. Do these equations have a solution? Well, we know that there must be if the three equations are *independent*. That is to say, no equation is a linear combination of the other two; or, there are no coefficients $a, b \neq 0$ such that $f_1 = af_2 + bf_3$. Note that this one condition is sufficient (why?). Assuming that the equations are linearly independent, and that there are as many equations as there are variables, then there is a solution to the set of equations (this is known from the *fundamental theorem of linear algebra*, which we will not prove!).

One of the reasons for matrix algebra in the first place is that writing all of these equations down is somewhat repetitive. Thus, we can compact the notation by defining matrix multiplication. To start, note that the set of equations above can be written out more compactly (eliminating the function labels, which are unnecessary).

$$3x + 2y + z = 5 \quad (1.40)$$

$$x + y - z = 1 \quad (1.41)$$

$$2x + 2y + 3z = 10 \quad (1.42)$$

Now, suppose we define four vectors as follows.

$$\begin{aligned}\mathbf{r}_1 &= (3, 2, 1) \\ \mathbf{r}_2 &= (1, 1, -1) \\ \mathbf{r}_3 &= (2, 2, 3) \\ \mathbf{x} &= (x, y, z)\end{aligned}$$

Recall that the dot product between two vectors is given by $\mathbf{a} \cdot \mathbf{x} = (a_1, a_2, a_3) \cdot (x, y, z) = a_1x + a_2y + a_3z$. Noting this, we can write our equation more compactly as follows.

$$\begin{aligned}\mathbf{r}_1 \cdot \mathbf{x} &= 5 \\ \mathbf{r}_2 \cdot \mathbf{x} &= 1 \\ \mathbf{r}_1 \cdot \mathbf{x} &= 10\end{aligned}$$

Or, noting $\mathbf{x} = (x, y, z)$, using conventional notation for vectors as follows.

$$\begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 10 \end{bmatrix} \quad (1.43)$$

This essentially *defines* matrix multiplication by the notation

$$\begin{bmatrix} 3 & 2 & 1 \\ 1 & 1 & -1 \\ 2 & 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 10 \end{bmatrix} \quad (1.44)$$

51

where to complete the multiplication, each row (taken as a vector) is dotted with the vector $\mathbf{x} = (x, y, z)$. This is the easiest way to remember how to do matrix multiplication!

To solve this problem, there are a number of ways that we can proceed. The easiest one is just to do what we would have done if we still had everything written out as three equations in three unknowns: eliminate variables from the equations simultaneously. In short, we can do any of the following operations:

1. Multiply any row by a constant.
2. Add *any* two rows, and replace either of those two by the result.
3. Interchange any two rows.

There is a caveat here—whatever we do to the rows of the matrix, we need to also do to the vector on the right-hand side (they are equations after all!). The simplest way to see this is to just do an example. It is possible to do this in a super-orderly, algorithmic fashion, but with small matrices, it is often better to look for easy opportunities for simplifications. To start, multiply row 2 by -2 , and add that to row 3, then divide row 3 by 5 . This gives

$$\begin{bmatrix} 3 & 2 & 1 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 8/5 \end{bmatrix} \quad (1.45)$$

9/and 1/n

Now, multiply row three by -1 and add to row 1; then add row 3 to row 2. Now we have

1.4.1 The derivative

Definition 1.20 (The Derivative of a Function). A derivative of a continuous (at least C^1 , as defined above) function f , at a point x in the domain of the function is defined by the limit

$$f'(x) = \lim_{t \rightarrow x} \frac{f(t) - f(x)}{t - x} \quad (1.53)$$

or, equivalently, we can define

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \quad (1.54)$$

There are at least two common notations for the derivative: f' and $\frac{df}{dx}$. For functions of a single variable, there is little chance for confusion. Sometimes, for purposes of clarity or presentation, one of the two is preferable to the other, especially when multiple functions of different dependent variables are considered.

Note that the definition of the derivative automatically provides the definition of higher-order derivatives. For example, consider the function $g(x) = f'(x)$. Then we have

$$g'(x) = (f'(x))' = \lim_{\Delta x \rightarrow 0} \frac{g(x + \Delta x) - g(x)}{\Delta x}$$

Or, establishing the notation for the second derivative

$$f''(x) = \frac{d^2 f}{dx^2} = \lim_{\Delta x \rightarrow 0} \frac{f'(x + \Delta x) - f'(x)}{\Delta x}$$

Or, applying the original definition of the derivative given by Eq. 1.54, we find

$$f''(x) = \frac{d^2 f}{dx^2} = \lim_{\Delta x \rightarrow 0} \frac{f(x + 2\Delta x) - 2f(x + \Delta x) + f(x)}{(\Delta x)^2}$$

This provides some explanation for why the second derivative is denoted by $\frac{d^2 f}{dx^2}$.

Generally, we do not derive the derivatives of functions from first principles, except perhaps in our introductory course on calculus. However, it is useful to recall how this is done.

Example 1.13 (Derivatives). Computing derivatives directly from the definition of the derivative is not tremendously difficult, but sometimes it does require a little creativity in determining the limit. As an example, let's look at how to find the derivative of the function $f(x) = x^2$.

$$\begin{aligned}
 f'(x) &= \lim_{t \rightarrow x} \frac{f(t) - f(x)}{t - x} \\
 &= \lim_{t \rightarrow x} \frac{t^2 - x^2}{t - x} \\
 &= \lim_{t \rightarrow x} \frac{(t - x)(t + x)}{t - x} \\
 &= \lim_{t \rightarrow x} (t + x) \\
 &= 2x
 \end{aligned}$$

Using the definition of the derivative, it is also easy to derive what the derivative of a product of two functions is. This is called the *product rule for differentiation* or the *Leibniz rule for differentiation*. The derivation is reasonably straightforward.

Theorem 1.3 (Product Rule for Differentiation).

$$\frac{d}{dx} [f(x)g(x)] = f'(x)g(x) + f(x)g'(x) \quad (1.55)$$

Proof. The proof for this is just a straightforward application of the definition of the derivative.

$$\begin{aligned}
 \frac{d}{dx} [f(x)g(x)] &= \lim_{t \rightarrow x} \frac{f(t)g(t) - f(x)g(x)}{t - x} \\
 &= \lim_{t \rightarrow x} \frac{f(t)g(t) + [f(x)g(t) - f(x)g(t)] - f(x)g(x)}{t - x} \\
 &= \lim_{t \rightarrow x} \frac{[f(t) - f(x)]g(t)}{t - x} + \lim_{t \rightarrow x} \frac{[g(t) - g(x)]f(x)}{t - x} \\
 &= \frac{df(x)}{dx}g(x) + f(x)\frac{dg(x)}{dx}
 \end{aligned}$$

Example 1.14 (Product Rule). Out of recurring spates of existential boredom, Elon Musk arranges to have a burrito inexplicably launched into space using the latest SpaceX rocket. After top scientists compute the appropriate stable orbit for the burrito, the burrito is deployed, and begins its endless orbit of the planet with a typical satellite orbiting period of about 2 hours.

Wanting more, Elon decides to up the ante from his previous energy-spewing device (a not-a-flamethrower sold by the Boring company, <https://www.boringcompany.com/not-a-flamethrower/>), he has a massive Earth-based laser built with a companion state-of-the-art space burrito tracking system, and begins a program of heating the burrito. Elon wants to maintain a nearly constant temperature as it orbits (a temperature which is just enough to keep the cheese melted, but not so hot that it would burn the roof of one's mouth). However, despite their best efforts, the temperature fluctuates with the same frequency as the orbit period. Generally, satellites in Earth orbit (such as a burrito) experience temperature swings between $\pm 100^\circ\text{C}$ to $\pm 200^\circ\text{C}$ around 0°C on each

7/

that!

chom chom
chom chom

orbit, so maintaining a constant temperature from an Earth-based laser is challenging under the best of circumstances.

The effective heat flux $Q(t)$ from the laser is related to the temperature $T(t)$ of the burrito through the effective heat transfer coefficient $k(t)$ by the expression $Q(t) = k(t)T(t)$. The heat flux laserered up to the burrito has to change over time as the burrito's temperature changes periodically as it enters and leaves the shadow of the earth. The effective heat transfer coefficient is a function of time in this situation because the distribution of temperatures within the burrito itself is not uniform, but rather varies with amount of heat flux being laserered into it. Months of experiments and ground-based remote sensing determine that the relationship between the laser heat flux and the burrito temperature is

$$T(t) = \frac{Q(t)}{k(t)} = \alpha_{\text{burrito}} \sin\left(\frac{\pi t}{2}\right) \left[k_0 + \beta_{\text{burrito}} \sin\left(\frac{\pi t}{2}\right) \right] - \gamma_{\text{burrito}}$$

where α_{burrito} , β_{burrito} , and γ_{burrito} are constants. In an attempt to understand better why the burrito will not stay at an exactly constant temperature, Elon wants to know first what the time rate of change of the temperature is currently. Please compute this for him.

Solution. Note first the definition

$$\frac{d}{dt} \sin(at) = a \cos(at)$$

Now, applying the product rule, we find

$$\begin{aligned} \frac{dT(t)}{dt} &= \frac{1}{2} \pi \alpha_{\text{burrito}} \cos\left(\frac{\pi t}{2}\right) \left(k_0 + \beta_{\text{burrito}} \sin\left(\frac{\pi t}{2}\right) \right) \\ &\quad + \frac{1}{2} \pi \alpha_{\text{burrito}} \beta_{\text{burrito}} \sin\left(\frac{\pi t}{2}\right) \cos\left(\frac{\pi t}{2}\right) \end{aligned}$$

There is another rule for differentiation that is essential; this is the *composition rule for derivatives*, more frequently called the *chain rule for derivatives*. This rule is handy when one has functions embedded in other functions; in other words, a *composite* function.

Definition 1.21 (Composite Function). A *composite function* is a function whose independent variable is also a function.

As an example, consider the relationship between position and velocity. Suppose that you have a position function (in one dimension) such that $x = x(t)$. Then the velocity, which depends on position and time, $v(x(t), t)$, is a composite function.

The chain rule provides a method of computing the derivative of composite functions. The proof of the chain rule is pretty complex, so it will not be presented here, but it is available in nearly every introductory text on calculus. The result, however, is important and will be used frequently in the material that follows.

Theorem 1.4 (Chain Rule for Differentiation). Suppose we have a function f whose argument is another function g . Assume that $g(t)$ has a derivative in the set of points $T = \{t : a < x < b\}$, and that $f(y)$ has a derivative in the set of points $Y = \{g(t)\}$. The composite function $f(g(t))$ is differentiable, and its derivative is given by

$$\frac{d}{dt}f(g(t)) = \frac{df(g)}{dg} \frac{dg(t)}{dt}$$

Keeping the functions g and f straight sometimes causes confusion. Some examples are helpful.

Example 1.15 (Product Rule, Velocity-Position Example).

Above, we mentioned the relationship between position and velocity as an example that could be considered a composite function problem. This is one example where one can compute derivatives both with and without the use of the chain rule. In essence, it provides a way to validate that the chain rule leads to correct results.

Suppose you go out to run along the street, but in a very strange way (why you decide to do this is open to discussion). As you run along, you run in a way such that your position is the cube of the time you have been running (obviously, you can't keep this up forever, but for a short time it is possible.) In particular, suppose your position, $x(t)$, is given by $x(t) = \frac{1}{3}t^3$. Now, we know that velocity is defined as the time rate of change of position, and acceleration is the time rate of change of velocity, so

$$\begin{aligned} v(t) &= \frac{dx(t)}{dt} \\ &= t^2 \end{aligned} \quad \begin{aligned} a(t) &= \frac{dv(t)}{dt} \\ &= 2t \end{aligned} \quad (A)$$

From here, we are going to do something that seems a bit unusual, but it is necessary so that we can validate the chain rule. First, note that we can express x as a function of v , as follows (noting $v = t^2$ implied $t = v^{\frac{1}{2}}$)

$$x(v(t)) = \frac{1}{3}t^3 = \frac{1}{3}v^{\frac{3}{2}} \quad (B)$$

Now, we have expressed x as a composite function, $x(v(t))$. We can compute the time derivative of this function by

$$\frac{dx(v(t))}{dt} = \frac{dx}{dv} \frac{dv}{dt} \quad (C)$$

We can compute dx/dv by

$$\frac{dx}{dv} = \frac{d}{dv} \left(\frac{1}{3}v^{\frac{3}{2}} \right) = \frac{1}{2}v^{\frac{1}{2}}$$

and dv/dt has already been computed above in Eq. (A). Combining these, we find

$$\begin{aligned} \frac{dx(v(t))}{dt} &= \frac{dx}{dv} \frac{dv}{dt} \\ &= \left(\frac{1}{2}v^{\frac{1}{2}} \right) (2t) \end{aligned}$$

and substituting $v = t^2$ gives the result

$$\frac{dx(v(t))}{dt} = t^2$$

which is identical to what is given in Eq. (A) above.

Example 1.16 (Product Rule) Some kinds of catalyst can be deactivated by chemicals produced during the catalytic reaction, or by external factors such as UV light. This process is sometimes called *catalytic poisoning*. Similar kinds of deactivation can happen to the enzymes in cell systems in biological reactors.

Suppose that an experiment is run, and it is determined that the amount of product begin produced, under conditions of deactivation (whether cells or catalyst), is given by

$$c(t) = c_0 \exp [-(k_0 - k_1 t)t]$$

In other words, an first-order-like rate process has an effective rate constant that is a function of time, (that is, $k_{\text{effective}} = k_0 - k_1 t$). Determine the rate of reaction, $c'(t)$.

Solution.

Using the product rule requires that we first identify the composite functions. Often, in practice, this is not done explicitly; rather, people just keep mental note of which function is which. However, it is instructive to explicitly identify the functions when there is any potential for confusion. For this problem, take

$$g(t) = -(k_0 - k_1 t)t$$

$$f(g) = c_0 \exp(g)$$

We need to recall the definition for the derivative of the exponential

$$\frac{d}{dg} \exp(g) = \exp(g)$$

The exponential is the only function (except the function 0) whose derivative is the same as the starting function! With this, we have all we need.

$$\frac{dg(t)}{dt} = -(k_0 - \frac{k_1 t}{2})$$

$$\frac{df(g)}{dg} = \exp(g)$$

So, the result is ...

$$\begin{aligned}\frac{d}{dt} f(g(t)) &= \frac{df(g)}{dg} \frac{dg(t)}{dt} \\ &= c_0 \exp(g) \left[-(k_0 - \frac{k_1 t}{2}) \right] \text{①}\end{aligned}$$

Substituting the function $g(x)$ (from above) into this result and rearranging gives the final result

$$\frac{d}{dt} f(g(t)) = -c_0(k_0 - \frac{k_1 t}{2}) \exp[-(k_0 - k_1 t)t] \text{②}$$

1.4.2 Partial derivatives

There is not much more to say regarding derivatives, except to explain the notion of a derivative when a function has multiple independent variables. So, with out delay, we will define the partial derivative of a function with two independent variables. The case of additional independent variables is identical, so no more than two is required for the definition.

Def 1.22 (Partial Derivatives). For a function with two independent variables, a partial derivative is the derivative of the function with respect to only one of the two variables (the other variable being held constant). Assume that $f(x, t)$ a continuous (at least C^1) function of x and t over a domain Ω (Ω could be an irregularly-shaped domain, so we will skip an effort to provide a more detailed set description of it). The partial derivative of f with respect to each variable is given by

$$\begin{aligned}\frac{\partial f(x, t)}{\partial x} &= \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, t) - f(x, t)}{\Delta x} \\ \frac{\partial f(x, t)}{\partial t} &= \lim_{\Delta t \rightarrow 0} \frac{f(x, t + \Delta t) - f(x, t)}{\Delta t}\end{aligned}$$

Similar results hold for functions of three or more variables, and the extension should be reasonably transparent based on the examples above.

Def 1.23 (The Chain Rule for Functions of Two or Three Variables). Recall that a *composite function* is a function whose independent variable is also a function. When a function is dependent upon two or more variables that are themselves functions, the *Chain Rule* allows us to determine the derivative in the following form.

$$\frac{d}{dt} f(x(t), y(t)) = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}$$

Note that in this definition, the derivatives of $x(t)$ and $y(t)$ are not *partial* derivatives, but conventional derivatives. This is because $x(t)$ and $y(t)$ are *functions of a single variable, t*. Therefore, the conventional derivative is the correct form of the derivative for those quantities.

71
71

1.4.3 Integration

Integration is a simple concept, but it turns out to be quite deep in actual applications. In introductory calculus, the concept of the Riemann integral is introduced. This integral applies to most functions that are encountered in science and engineering; in particular, it is useful for evaluating functions that are *piecewise smooth* as defined previously. This is not the only kind of integral that can be defined. For example, consider the following function (the Dirac function).

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is irrational on } x \in [0, 1] \\ 0 & \text{if } x \text{ is rational on } x \in [0, 1] \end{cases}$$

This function is not piecewise smooth because it has an infinite number of holes in it (i.e., it has a hole at every possible fraction between 0 and 1!). The conventional Riemann integral cannot be used for such a function. However, there are more general forms of the integral (such as the Lebesgue integral) that can be used to measure such functions. We will not explore these in any way in this text.

For our purposes, the development of the integral will not be reviewed in detail. Instead, a few important properties of the integral are presented. The first of these is a reminder about the fundamental theorem of calculus.

The fundamental theorem of calculus says some really important things. Primarily, it tells us the following.

Theorem 1.5. Suppose a smooth function, F , is defined on the interval $[a, b]$. Because the function is smooth, it has a derivative, $F' = f$. Then

$$\int_{t=a}^{t=b} f(t) dt = F(b) - F(a)$$

or, equivalently,

$$\int_{t=a}^{t=b} \frac{dF(t)}{dt} dt = F(b) - F(a)$$

71

The extension of this theorem to *piecewise continuous* functions is straightforward. It involves simply computing the integral over each of the (finite number) continuous intervals.

This is a very powerful theorem, and it essentially maps the problem of finding integrals on to the problem of finding derivatives. That is to say, if we are given a function f , and we happen to know a function F whose derivative is equal to f , then we can compute the integral of f with that knowledge. The function F is sometimes called the *antiderivative* of f for that reason. Although this sounds somewhat circular, it is not. Most of the “known” integrals that exist do so because we have identified the antiderivative for the function.

For many, a first course in calculus involves learning many “techniques” to find the antiderivative of functions. Most of this we will leave in the past, with the idea that we will all remember (or be able to look up) most of the fundamental integrals and derivatives that we encounter. However, there is one “technique” that is very useful in a number of applications, and it is one that we will have the opportunity to employ several times. This is the rule for integration by parts.

Theorem 1.6. Let f and g be smooth functions on an interval $x \in [a, b]$. Then

$$\int_{x=a}^{x=b} f(x) \frac{dg(x)}{dx} dx = f(x)g(x) \Big|_{x=a}^{x=b} - \int_{x=a}^{x=b} \frac{df(x)}{dx} g(x) dx$$

Frequently, this rule is written in the easy-to-remember form

$$\int_{x=a}^{x=b} u dv = uv \Big|_{x=a}^{x=b} - \int_{x=a}^{x=b} v du$$

Integration by parts is particularly useful under the following circumstances: (i) there is an integrand of the form of n th-order polynomial times some function that we know how to integrate n times, or (ii) there is an integrand that contains an n th-order derivative times some function that we know how to differentiate n times, and we would like to eliminate the derivative. This will be made clearer in the following example.

cap 18/1 rom 1 rom 1

Example 1.17 (Integration by parts). Integrate the following functions over the interval $x \in [0, 1]$.

(a) $f(x) = x \sin(\pi x)$

(b) $g(x) = \frac{df}{dx} \sin(\pi x)$ given $f(0) = 0$

Solution. For (a), let $u = x$, $dv = \sin(\pi x)dx$. Then $du = dx$ and $v = -\pi \cos(\pi x)$ (noting $d/dx(-\cos x) = \sin x$). Then

$$\begin{aligned} \int_0^1 x \sin(\pi x) dx &= \frac{x}{\pi} \cos(\pi x) \Big|_0^1 - \int_0^1 -\cos(\pi x) dx \\ &= \frac{1}{\pi} + \sin(\pi x) \Big|_0^1 \\ &= \frac{1}{\pi} \end{aligned}$$

For (b), let $u = \sin(\pi x)$, $dv = \frac{df}{dx} dx$. Then, $du = 1/\pi \cos(\pi x)$ and $v = f(x)$ (fundamental theorem of calculus). Then

$$\begin{aligned}\int_0^1 \frac{df}{dx} \sin(\pi x) dx &= \left. \frac{x}{\pi} \cos(\pi x) f(x) \right|_0^1 - \int_0^1 f(x) \frac{1}{\pi} \cos(\pi x) dx \\ &= - \int_0^1 f(x) \frac{1}{\pi} \cos(\pi x) dx\end{aligned}$$

In this latter example, we can go no further than this without knowing more about f ; however, we have *eliminated* the derivative, which is frequently a useful operation.

1.5 Sequences and Series

Infinite series are incredibly useful, and they are the true workhorses of applied mathematics. Many important functions (especially many of the transcendental ones that we use routinely, such as $\sin(x)$ and $\exp(x)$) *are computed* fundamentally by use of an infinite series.

Sequences are related to series. One way to think of a sequence is that it is just a indexed list of numbers or expressions. Thus, gathering each term in an infinite series together as a list would form a sequence. To make this concrete, we have the following definition.

Definition 1.24 (Finite Sequence). A *finite sequence* is a list of objects (elements) indexed by a consecutive subset of the natural numbers ($\mathbb{N} = \{1, 2, 3, 4, \dots, N\}$ or $\mathbb{N}_0 = \{1, 2, 3, 4, \dots, N\}$). In this text, a sequence of N (or $N+1$ if the first index is 0) terms is denoted $A = (a_1, a_2, a_3, \dots, a_N)$ (or $A = (a_0, a_1, a_2, \dots, a_N)$ if the first index is 0).

Definition 1.25 (Infinite Sequence). A *sequence* is a list of objects (elements) indexed by the natural numbers ($\mathbb{N} = \{1, 2, 3, 4, \dots\}$ or $\mathbb{N}_0 = \{1, 2, 3, 4, \dots\}$). This list may be specified by a functional rule, or by listing the elements explicitly. Sequences may have a finite or countably infinite number of terms. The element in the position i (where $i \in \mathbb{N}$ or \mathbb{N}_0) is called the i^{th} term of the sequence. Regardless of how the elements are defined, sequences are viewed as being *functions over the natural numbers*. A common notation for an infinite sequence is $A = (a_n)_{n=1}^{\infty}$ (or $A = (a_n)_{n=0}^{\infty}$ if the first index is zero).

Example 1.18 (Examples of sequences). Sequences are just lists of objects. Thus, they represent a fairly general mathematical concept. Here are a few examples.

1. The following is a somewhat boring finite sequence: $A = (1, 1, 1, 1, 3, 1)$. It has a finite number of terms (six), and each element is the same, except for the fifth term.

2. The following is an infinite sequence containing all of the even numbers greater than zero: $A = (2, 4, 6, 8, \dots)$.
3. Here are two other ways to denote the very same sequence of the previous example: $A = (2n)_{n=1}^{\infty}$ and $S = (2n)_{n \in \mathbb{N}}$.
4. Here is an infinite sequence that denotes a familiar irrational number: $A = (3, 1, 4, 1, 5, 9, 2, 6, 5, 3, 5, \dots)$
5. Here is a sequence whose terms converge to 1 as $n \rightarrow \infty$: $A = (1 + 1/n^2)_{n=1}^{\infty} = (2, \frac{5}{4}, \frac{10}{9}, \frac{17}{16}, \frac{26}{25}, \dots)$. We can compute the following limit for this sequence:
 $\lim_{n \rightarrow \infty} (a_n) = 1 + \lim_{n \rightarrow \infty} (\frac{1}{n^2}) = 1$.

In the last example above, we introduced the idea of *convergence* for infinite sequences. To be concrete, we need to establish what it means for a sequence to converge. The issue that we need to consider is how to show that an infinite sequence actually approaches some limit that we define. This is a somewhat tricky prospect. Suppose we have a sequence denoted by $A = (a_n)_{n \in \mathbb{N}}$. If you think about it, there is no actual number a_∞ ! All we can really ask about an infinite sequence is the question “what happens as n becomes an arbitrarily large integer?” This is *exactly* the notion behind a limit. While we have discussed limits approaching a finite value in the material previous to this section, we have not yet discussed infinite limits.

A *mathematical* discussion of limits inevitably an ε - δ -like argument. In this case, instead of having a δ interval in which some error measure must be small, we instead have an integer for which all values of the sequence indexed by this integer or higher are sufficiently close to a limit. To make this clear, here is the definition of the limit of a sequence.

Definition 1.26 (Limits of a sequence). Suppose $(a_n)_{n=1}^{\infty}$ is a sequence. Then, the limit L exists if the sequence gets arbitrarily close to L as n increases. In other words, for every $\varepsilon > 0$ that can be chosen, no matter how small the value of ε , then it is also true that there is an integer N , such that

$$\text{if } n \geq N, \text{ then } |a_n - L| < \varepsilon \quad (1.56)$$

In that case, we say that $(a_n)_{n=1}^{\infty}$ has a limit, L , and write

$$\lim_{n \rightarrow \infty} (a_n) = L \quad (1.57)$$

Definition 1.27 (Diverging sequences). Suppose $(a_n)_{n=1}^{\infty}$ is a sequence. Then, if a_n tends to infinity as n becomes arbitrarily large, the sequence is said to diverge. In other words, if for every number $M < \infty$, no matter how large M , then there is still always an integer N such that

$$\text{if } n \geq N, \text{ then } |a_n| > M \quad (1.58)$$

In that case, we say that $(a_n)_{n=1}^{\infty}$ diverges (i.e., the limit tends to ∞).

$$\lim_{n \rightarrow \infty} (a_n) = \infty \text{ or } \lim_{n \rightarrow \infty} (a_n) = -\infty \quad (1.59)$$

With the concept of sequences defined, it is relatively straightforward to define an *infinite series*. In short, an infinite series is just the sum of some infinite sequence. However, it is useful to think about them in the following sense.

Definition 1.28 (Series). Suppose $(a_n)_{n=1}^{\infty}$ is a sequence. Now, define the *partial sums* of the sequence by

$$S_N = \sum_{n=1}^{n=N} a_n$$

31

is,

$$\begin{aligned} S_1 &= a_1 \\ S_2 &= a_1 + a_2 \\ S_3 &= a_1 + a_2 + a_3 \\ &\dots \end{aligned}$$

Then, the S_n are known as the partial sums of the series. As we allow $N \rightarrow \infty$, the partial sums define the following infinite series

$$S = \sum_{n=1}^{\infty} a_n \quad (1.60)$$

is an infinite series.

There are two important things to note here.

1. As noted above, it is not necessary that the lower bound of a series start at 1. However, most series used routinely in applied mathematics start at either 0 or 1. If starting at some integer other than 1, the definitions above would be modified in the obvious way.
2. Not all series necessarily *converge*. In fact, it turns out that even some *non-convergent* series are useful in applied mathematics. In fact, if you have ever applied Stirling's approximation for the factorial

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e} \right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3} - \frac{571}{2488320n^4} + \dots \right) \quad (1.61)$$

31

then you have used a non-convergent series. Convergence of series is an exceptionally complex topic, and one that we will not review in detail here.

Although determining whether or not a series converges (and to what it converges to) will not be reviewed in detail here, we will revisit convergence in later chapters when it arises in various applications. We can define what it means for a series to converge, however (even if we do not explore the various tests for convergence at this juncture).

Definition 1.29 (Convergence of an Infinite Series). Suppose we define the partial sums of an infinite series, S_n , as we have above. Then, the series is said to *converge* if the partial sums tend toward a fixed limit L as $N \rightarrow \infty$.

$$S = \lim_{N \rightarrow \infty} s_N = \sum_{n=1}^{N} a_n = L \quad (1.62)$$

Although we will not be exploring series convergence in detail, there are a few convergence tests that will be useful later. We present only three here; however, these three tests alone are quite a powerful subset of the possible series convergence tests. These are presented, without proof, as follows.

cap 11

Theorem 1.7 (The integral test). Let $(a_n)_n = 1^\infty$ be a nonnegative (i.e., each term is positive or zero) sequence, and let f be a continuous, monotonically decreasing function on $[0, \infty)$ defined such that

$$f(n) = a_n \text{ for } n \geq 1 \quad (1.63)$$

Then the series

$$\sum_{n=1}^{n=\infty} a_n \quad (1.64)$$

converges if and only if the integral

$$\int_{x=1}^{\infty} f(x) dx \quad (1.65)$$

is finite.

Theorem 1.8 (Ratio Test). We can say that a series converges if

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} < 1 \quad (1.66)$$

This latter expression is called the (general) ratio test (which you may have learned in a course on calculus). The converse is also true; that is, a series is said to diverge if

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} > 1 \quad (1.67)$$

If the limit of this quantity is exactly equal to 1, then nothing can be said about convergence of the series.

cap 11

Theorem 1.9 (Leibniz convergence test).

Suppose the sequence $(|a_n|)_{n=1}^{n=\infty}$ is

an n

+/-

1. A monotonically decreasing sequence ($|a_{n+1}| < |a_n|$ for all values of n), and
2. $n \rightarrow \infty \lim a_n = 0$

- 1/

Then the alternating series

$$\sum_{n=0}^{\infty} (-1)^n a_n \quad (1.68)$$

$$\sum_{n=0}^{\infty} (-1)^{n+1} a_n \quad (1.69)$$

converge.

1.5.1 Power Series

A power series is just a label applied to a particular kind of series. A power series is defined by any series of the form

$$\sum_{n=1}^{\infty} c_n x^n \quad (1.70)$$
Q1

Note that, as with series in general, the lower index can start at any number, although it will generally be 0 or 1. Power series are one of the most frequently used devices in all of applied mathematics. In the next example, we provide the power series representation of a few familiar functions.

91 9

Example 1.19 (Power Series). Trigonometric functions are examples of transcendental functions (as introduced above). By definition, this means that there is no *finite* set of algebraic steps that one can use to generate such functions (e.g., there is no polynomial that will reproduce them exactly). If you think about trigonometric functions, you will realize that you have probably never been asked to compute the value of, say, the sine function yourself. This is because the sine function must be defined by an algorithm that has, in principle, an infinite number of steps!

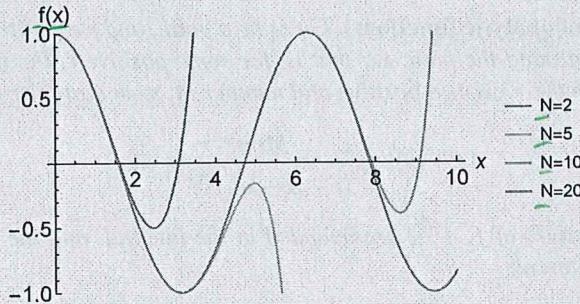
Although functions like $\sin(x)$ and $\cos(x)$ cannot be represented by a finite number of algebraic steps, they *can be represented by an infinite number of algebraic steps*. The following are the power series for these functions.

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \quad (1.71)$$

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \quad (1.72)$$

While it is true that, in principle, the value of these functions is defined by an infinite number of terms, it is often the case with series that the terms defined by increasingly large n become smaller and smaller in magnitude. In a practical sense, this means that it is often possible to get a good representation for a series at a point of interest by using a finite number of terms. As an example, in ~~the~~ Fig. 1.161 below we plot the approximation to the sine function using $N = 2, 5, 10$, and 20 terms. It is clear that even for $N = 2$, the power series actually gives a good approximation to the function in the interval of approximately $0 < x < 1.5$.

Fig. 1.161



bf 91 9

Fig. 1.161 The sine function approximated by its power series, using an increasing number of terms.

1.5.2 Taylor's Theorem

A Taylor series is actually just a power series. The genius of Taylor was to work out how one might compute the terms in a power series for any function in which one knows the derivatives. To start out, let's constructively define the *Taylor series* as follows.

Definition 1.30 (Taylor Series). For a function that has an infinite number of derivatives at a point $x = a$, that function has a Taylor series expansion that represents the value of the function at a location $x = (a + \Delta x)$ as follows.

$$f(a + \Delta x) = f(a) + \Delta x \frac{df}{dx} + \frac{1}{2!} (\Delta x)^2 \frac{d^2 f}{dx^2} + \frac{1}{3!} (\Delta x)^3 \frac{d^3 f}{dx^3} + \dots$$

where the factorial of an integer n is defined by $n! = n \cdot (n - 1) \cdot (n - 2) \dots 2 \cdot 1$. This definition says nothing about whether or not this series converges or not; as discussed above, we will consider convergence for specific cases that arise later. The Taylor series provides us with an intuitive tool to define *analytic* functions (which were defined previously in §5.1). This second definition is as follows.

Definition 1.31 (Analytic Functions—Definition 2). An *analytic function* is a function which (i) is defined (i.e., it has a computable value) on an open interval (I) of the real line (I is the set of values x , such that $x \in (a, b)$; or, in set notation $I = \{x : a < x < b\}$), (ii) its Taylor series converges to a definite value on the interval I and (iii) the definite value it converges to is equal to $f(x)$ for all values of x in the interval I .

Admittedly, that is quite a mouthful of requirements; however, many functions that are of interest in science and engineering are analytic. For example, all polynomials on a finite domain are analytic. The functions $\sin x$, $\cos x$, and $\ln x$ are analytic on appropriate domains. So, although the notion of being analytic seems stringent on the surface, most of the functions we commonly encounter are analytic on some portion of their domains.

One of the nice properties of all analytic functions is that *all* of the derivatives of an analytic function are bounded. In other words, no derivative of an analytic function can grow to infinity anywhere in the domain where the function is analytic. This is strongly suggested by the fact that all analytic functions have convergent Taylor series (i.e., any derivative that approached infinity would prevent the Taylor series from converging to a finite value). Nonetheless, the following theorem is stated without proof.

Theorem 1.10 (Derivatives of analytic functions). Let x_0 be a point and r a positive value. Let a function, f be analytic in an interval around the point x_0 ; that is, for some positive r , the function f is analytic in $\{I : x_0 - r < x < x_0 + r\}$. Then there exists a positive real number M , such that for every $n \in \mathbb{N}$

$$|f^{(n)}(x)| \leq \frac{Mrn!}{(r - |x - x_0|)^{(n+1)}}$$

In other words, the derivatives of f , $f^{(n)}$ are bounded in the interval, and the derivative cannot grow arbitrarily large within the interval.

The next example is one that shows that all polynomials are analytic functions.

↑
An: Are you able
to remove gray
from this page?

(cap) 18/ Example 1.20 (*Polynomials are Analytic*). Suppose we have a polynomial on an open interval other than $I = (-\infty, \infty)$. The polynomial is analytic everywhere in its domain.

3/ This is not a proof, but it could easily be turned into one. Instead, consider a concrete polynomial, for example $f(x) = x^3 + 2x^2 + 145x - 1$ defined on the *open* interval $0 < x < 1$. Suppose we want to determine the Taylor series for this polynomial around the point $a = 0$. We can expand this polynomial in a Taylor series around the point $x = a = 0$ to determine the value of the function at the point $x = (a + \Delta x)$. To do so, we first need to compute all of the derivatives of f . This seems as though it might be nearly impossible, until we note

$$f'(x) = 3x^2 + 4x + 145$$

$$f''(x) = 6x + 4$$

$$f'''(x) = 6$$

$$f^{(4)}(x) = 0$$

$$f^{(5)}(x) = 0 \dots$$

Now, according to the formula, the Taylor series is given by

$$f(a + \Delta x) = f(a) + \Delta x \frac{df}{dx} + \frac{1}{2!} (\Delta x)^2 \frac{d^2 f}{dx^2} + \frac{1}{3!} (\Delta x)^3 \frac{d^3 f}{dx^3} + \dots$$

Noting that, for this example, $\Delta x = x - a = x - 0 = x$, then we have

$$f(x) = f(0) + x \frac{df}{dx} + \frac{1}{2!} x^2 \frac{d^2 f}{dx^2} + \frac{1}{3!} x^3 \frac{d^3 f}{dx^3} + \dots$$

Substituting the derivatives above yields

$$f(x) = -1 + x(3x^2 + 4x + 145) + \frac{1}{2!} x^2(6x + 4) + \frac{1}{3!} x^3 6 + 0 + 0 + 0 + \dots$$

Simplifying terms, we find

$$f(x) = x^3 + 2x^2 + 145x - 1$$

3/ *9/* which was the original polynomial. Thus, this polynomial is its own Taylor series! This is actually true for all polynomials, and it is not hard to prove using the principle of induction. The ancillary conclusion that can be reached is that all polynomials defined on a (open, non-infinite) domain are analytic functions.

(cap) 4/ So, we have established that the Taylor series can represent all analytic functions, and we examined a particular example of a polynomial that illustrated that the series actually did what we hoped it would. One might, at this juncture, wonder: *Why does the Taylor series work?* It turns out, that this is reasonably easy to prove. Again, we will not do this as a formal proof, but rather as a construction that illustrates the method. A formal proof can be constructed from this outline.

To start, recall the fundamental theorem of calculus for the function f in the form

$$\int_{x=a}^{x=a+\Delta x} \frac{df}{d\xi} d\xi = f(a + \Delta x) - f(a) \quad (1.73)$$

where here we have used ξ as the variable of integration. This is already reasonably suggestive. Rearranging, note that we have

$$f(a + \Delta x) = f(a) + \int_{x=a}^{x=a+\Delta x} \frac{df}{d\xi} d\xi \quad (1.74)$$

Now, we use integration by parts once (setting $u = df/d\xi$, $dv = d\xi$) to give

$$f(a + \Delta x) = f(a) + (a + \Delta x) \frac{df}{dx} \Big|_{a+\Delta x} - a \frac{df}{dx} \Big|_a - \int_{x=a}^{x=a+\Delta x} \xi \frac{d^2 f}{d\xi^2} d\xi \quad (1.75)$$

Now note, also by the fundamental theorem of calculus,

$$\int_{x=a}^{x=a+\Delta x} (a + \Delta x) \frac{d^2 f}{d\xi^2} d\xi = (a + \Delta x) \frac{df}{dx} \Big|_{a+\Delta x} - (a + \Delta x) \frac{df}{dx} \Big|_a \quad (1.76)$$

Rewriting this as

$$(a + \Delta x) \frac{df}{dx} \Big|_{a+\Delta x} = \int_{x=a}^{x=a+\Delta x} a \frac{d^2 f}{d\xi^2} d\xi + (a + \Delta x) \frac{df}{dx} \Big|_a \quad (1.77)$$

Finally, substituting this into Eq. (1.75) gives

$$f(a + \Delta x) = f(a) + \Delta x \frac{df}{dx} \Big|_a - \int_{x=a}^{x=a+\Delta x} (\xi - a) \frac{d^2 f}{d\xi^2} d\xi \quad (1.78)$$

Repeating this process n times yields Taylor's formula to term n .

Problems

1. The determinant for a 2×2 matrix is pretty easy to find. Consider the matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

where a_{ij} indicates (by convention) the value of the i^{th} row and the j^{th} column.

Now, do the following. (i) Use row operations to make the matrix in upper triangular form, and (ii) compute the determinant from the product along the diagonal of the triangular matrix. The result should be $\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}$.

2. As long as we have computed the determinant for a 2×2 matrix, we may as well learn about the solution to a set of two equations in two unknowns. There is a method, known as *Cramer's rule* that allows you to find the solution to *any* set of linear equations as long as you can find the appropriate determinants. It turns out that for an $n \times n$ system of equations, you need to find n determinants. In general, this is not so easy to do; for small matrices, however, it can be very convenient.

For two equations in two unknowns, the coefficient matrix, \mathbf{A} is 2×2 . For a 2×2 system of the form

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

or

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Cramer's rule says that the solution can be found from the following determinants where the right-hand side has been substituted into the columns of \mathbf{A}

$$x = \frac{1}{\det(\mathbf{A})} \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} \quad y = \frac{1}{\det(\mathbf{A})} \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$$

where both $\det(\cdot)$ and $|\cdot|$ indicate the determinant. Suppose we have the set of equations

$$\begin{aligned} 2x + 4y &= 2 \\ 2x + 2y &= \frac{3}{2} \end{aligned}$$

or, in matrix form

$$\begin{bmatrix} 2 & 4 \\ 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ \frac{3}{2} \end{bmatrix}$$

Use Cramer's rule to solve for x and y .

3. Linearity is an important concept to be familiar with. For the following operators, $\mathcal{L}(\cdot)$, determine if they are linear or nonlinear.

a. $\mathcal{L}(x) = ax^2 + bx$

- b. $\mathcal{L}(x) = \sqrt{a^2 x}$
- c. $\mathcal{L} = \frac{d^2}{dx^2} + a \frac{d}{dx}$, so that $\mathcal{L}(f) = \frac{d^2 f}{dx^2} + a \frac{df}{dx}$
- d. $\mathcal{L}(f) = \int_x f(x) dx$
4. Integration is done one of two ways: either using an indefinite integral, or by using a definite integral. The fundamental theorem of calculus connects these for us. One way to help remember the fundamental theorem of calculus is to compute the area under a curve using the two methods. So, please try this. For the function $f(x) = x$ on the interval $x \in [1, 2]$ compute the area two ways.

An: I think (i) and (ii) are okay here, since they appear within a lettered list.

by / 9/ 3/

- a. Compute the indefinite integral. This will contain an unknown constant; call it C . The value of C is fixed when we specify which interval we want to compute the area for. So, suppose that we want to compute the area for the following two intervals: (i) $0 \leq x \leq 1$, and (ii) $0 \leq x \leq 2$. Compute the area under the curve $f(x) = x$ for intervals (i) and (ii) by drawing a sketch of the function, and then using simple geometry to compute the area of the associated regions.
- b. Now, compute the numerical value of the indefinite integral for this problem as follows. You know the proper areas that each interval represents. You have also computed the indefinite integral with unknown constant, C in part (a). For the intervals (i) and (ii), compute the value of the unspecified constant, C . It should be the same for the two intervals.
- c. Note that we can get the area of the interval $1 \leq x \leq 2 \Leftrightarrow x \in [1, 2]$ by subtraction using the two areas above. Draw this on a sketch, and also compute what the area is numerically (from your results in part (b)).
- d. Finally, compute the definite integral

$$\int_{x=1}^{x=2} x dx \quad (1.79)$$

You should (kind of obviously perhaps at this point) get exactly the same value. The fundamental theorem of calculus states that the area for some fixed interval is just the difference between two indefinite integrals, as you have shown.

5. Use integration by parts to compute the following integrals (i.e., find an answer in terms of functions no longer involving integrals) for $x \in [0, 1]$. Note: You may have to apply integration by parts more than once to get a final result!
- a. $f(x) = x \sin(\pi x)$
- b. $f(x) = x \cos(\pi x)$
- c. $f(x) = x^2 \sin(\pi x)$
- d. $f(x) = x^2 \cos(\pi x)$
6. For the functions f and g in Fig. 1.7, consider the domain $x \in [0, 1.5]$. What are the ranges of the two functions (approximate estimate as best you can)? Are either of the functions piecewise continuous? Are either of the functions piecewise smooth?
7. The exponential function has the unusual property that its Taylor series converges everywhere on the real line. Given that information, is the following function analytic on the domain $D = [0, 1]$?

$$f(x) = \exp(\pi(x-1)), \quad (1.80)$$

9/

Why or why not?

8. **Euler's rule.** Euler's rule, given by

$$e^{ix} = \cos x + i \sin x$$

3 /

is actually pretty easy to prove. What you need to know to do that is the following (and that $x^0 = 1$ and $0! = 1$ by definition)

Q1

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \frac{x^0}{0!} + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$$

$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \dots$$

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \dots$$

3 /

Use these series, in addition to the properties of the imaginary number $i = \sqrt{-1}$ to prove Euler's rule.

9. Consider the following function, defined piecewise over the interval $0 \leq x \leq 2$.

$$f(x) = \begin{cases} 0 & \text{for } 0 \leq x \leq 1 \\ x & \text{for } 1 < x \leq 2 \end{cases}$$

- a. Is this function continuous on the interval $x \in [0, 2]$?
- b. Is this function C^1 continuous on the interval $x \in [0, 2]$?
- c. Can this function be described as piecewise C^1 continuous? Why or why not?

Au : ?

[[[[[Do more of these]]]]]

10 //

4 /

10. **The Divide and Conquer method for finding the square root.** The ancient Babylonians had a method to compute the square root of any positive real number, a . It was an iterative algorithm that started with an initial *guess* of the root, and then provided the method to find the exact root. Mathematically, what they did was to develop the following *implicit* relationship for the square root.

$$x^2 = a$$

(by definition)

$$2x^2 = x^2 + a$$

(add x^2 both sides)

$$2x = x + \frac{a}{x}$$

(divide by x)

$$x = \frac{x + a/x}{2}$$

(solve for x , implicit equation)

rom y6 /

Here, a is the starting number, and x is the square root being sought.

Given an initial guess, x_1 , then a better estimate could be found by

$$x_2 = \frac{x_1 + a/x_1}{2} \quad (1.81)$$

or, in general,

$$x_{i+1} = \frac{x_i + a/x_i}{2}, \quad i = 1, 2, 3, \dots \text{ (or } i \in \mathbb{N}) \quad (1.82)$$

Use this method to develop a sequence of estimates for $\sqrt{2}$ (i.e., $a = 2$), starting with the (reasonably bad...) initial guess of $x_1 = 1$. You can use a calculator (or spreadsheet) to do the work, but please show the result at each step. Also compute the relative error at each estimate step i , $\epsilon_i = (\text{estimate}_i - \text{actual})/\text{actual}$, where the value of estimate_i is the value computed at the i^{th} step of the algorithm above, and the value of actual is the value computed by a calculator or spreadsheet. Stop the algorithm when the relative error is less than $1/100$.

11. Which of the following functions are *homogeneous*? Show your work.

- a. $f(x) = x$
- b. $f(x) = x^2 \sin(\pi x)$
- c. $f(x) = x^\pi$
- d. $f(x) = \sqrt{5x}$
- e. $f(x) = x^2 \sqrt{x}$

12. Does the following series converge? Which test for convergence can be used to establish the convergence or divergence of this series?

$$P(n) = \sum_{n=1}^{\infty} \frac{1}{n^2} \quad (1.83)$$