"THERE IS NO KNOWLEDGE THAT IS NOT POWER."

RALPH WALDO EMERSON, (1803-1882)

LEAH EDELSTEIN-KESHET

DIFFERENTIAL CALCULUS FOR THE LIFE SCIENCES

THE AUTHOR AT THE UNIVERSITY OF BRITISH COLUMBIA

PUBLISHED BY THE AUTHOR AT THE UNIVERSITY OF BRITISH COLUMBIA

http://www.math.ubc.ca/~keshet/OpenBook.pdf

Printing as of September 2017

Contents

List of Figures

List of Tables

This book is dedicated to students who have a genuine desire to learn about the beauty and usefulness of mathematics and to many colleagues who have helped to shape my interest and philosophy of teaching life-science calculus.

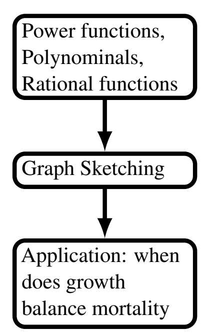
Preface

This preface describes the main philosophy of the course, and serves as a guide to the student and to the instructor. It outlines reasons for the way that topics are organized, and how this organization is intended to help introduce first year students to the major concepts and applications of differential calculus. The material for this book was collected during two decades of teaching calculus at the University of British Columbia, and benefitted greatly from insights and ideas of colleagues, as well as questions, interest, and enthusiasm of students and instructors.

Introduction to this book

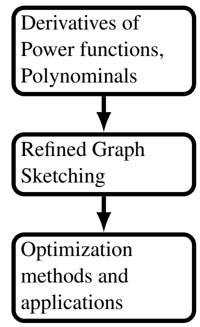
Calculus arose as a tool for solving practical scientific problems through the centuries. However, it is often taught as a technical subject with rules and formulas (and occasionally theorems), devoid of its connection to applications. In this course, the applications form an important focal point, with emphasis on life sciences. This places the techniques and concepts into practical context, as well as motivating quantitative approaches to biology taught to undergraduates. While many of the examples have a biological flavour, the level of biology needed to understand those examples is kept at a minimum. The problems are motivated with enough detail to follow the assumptions, but are simplified for the purpose of pedagogy.

The mathematical philosophy is as follows:

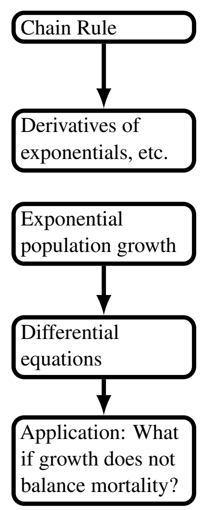


We start with basic observations about functions and graphs, with an emphasis on power functions and polynomials. We use elementary properties of a function to sketch its graph and to understand its shape, even before discussing derivatives; later we refine such graph-sketching skills. We consider useful ideas with biological implications even in this basic context. In fact, we discuss several examples in which two processes (such as growth and predation or nutrient uptake and consumption) are at balance. We show how setting up the relevant algebraic problem revels when such a balance can exist.

We introduce the derivative in three complementary ways: (1) As a rate of change, (2) as the slope we see when we zoom into the graph of a function, and (3) as a computational quantity that can be approximated by a finite difference. We discuss (1) by first defining an average rate of change over a finite time interval. We use actual data to do so, but then by refining the time interval, we show how this average rate of change approaches the instantaneous rate, i.e. the derivative. This helps to make the idea of the limit more intuitive, and not simply a formal calculation. We illustrate (2) using a sequence of graphs or interactive graphs with increasing magnification. We illustrate (3) using simple computation that can be carried out on a spreadsheet. The actual formal definition of the derivative (while presented and used) takes a back-seat to this discussion.

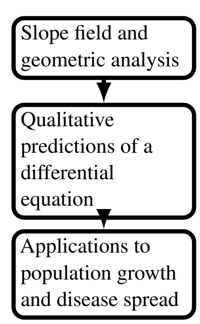


The next philosophical aspect of the course is that we develop all the ideas and applications of calculus using simple functions (power and polynomials) *first*, before introducing the more elaborate technical calculations. The aim is to highlight the usefulness of derivatives for understanding functions (sketching and interpreting their behaviour), and for optimization problems, before having to grapple with the chain rule and more intricate computation of derivatives. This helps to illustrate what calculus can achieve, and decrease the focus on rote mechanical calculations.



Once this entire “tour” of calculus is complete, we introduce the chain rule and its applications, and then the transcendental functions (exponentials and trigonometric). Both are used to illustrate biological phenomena (population growth and decay, then, later on, cyclic processes). Both allow a repeated exposure to the basic ideas of calculus - curve sketching, optimization, and applications to related rates. This means that the important concepts picked up earlier in the context of simpler functions can be reinforced again. At this point, it is time to practice and apply the chain rule, and to compute more technically involved derivatives. But, even more than that, both these topics allow us to informally introduce a powerful new idea, that of a differential equation.

By making the link between the exponential function and the differential equation $dy/dx = ky$, we open the door to a host of biological applications where we seek to understand how a system changes: how a population size grow? how does the mass of a cell change as nutrients are taken up and consumed? By revisiting our initial discussions, we identify the “balance points” as steady states, and we develop arguments to predict what changes with time would be observed. The idea of a deviation away from steady state also leads us to find the behaviour of solutions to the differential equation $dy/dt = a - by$. This leads to many useful applications. including the temperature of a cooling object, the level of drug in the bloodstream, simple chemical reactions, and many more.



Ultimately, a first semester calculus course is all about the applications of a derivative. We use this fact to explore nonlinear differential equations of the first order, using qualitative sketches of the direction field and the state space of the equation. Even though some of the (integration) methods for solving a differential equation are developed only in a second semester calculus course, we provide here the background to understanding what such equations are saying, and what they imply. These simple yet powerful qualitative methods allow us to get intuition to the behaviour of more realistic biological models, including density-dependent (logistic) growth and even the spread of disease. Many of the ideas here are geometric, and we return to interpreting the meaning of graphs and slopes yet again in this context.

The idea of a computational approach is introduced and practiced in several places, as appropriate. We use simple examples to motivate linear approximation and Newton's method for finding zeros of a function. Later, we use Euler's method to solve a simple differential equation computationally. All these methods are based on the derivative, and most introduce the idea of an iterated (repeated) process that is ideally handled by computer or calculator. The exposure to these computational methods, while novel and sometimes daunting, provides an important set of examples of how properly understanding the mathematics can be used directly for effective design of computational algorithms.

Acknowledgements

I am particularly grateful to Prof. William Casselman (emeritus, UBC) and Prof David Austin (Grand Valley State University, and formerly UBC) for the initial years of development of the approach and emphasis, as well as overall philosophy of this calculus course. Much of the plan described above had its beginnings while we were co-instructors in the 1990's at UBC. I am also grateful to recent colleagues, who have provided inspiration, input, and advice, and who have helped to shape the biological applications in the book. These include, most especially, Prof. Eric Cytrynbaum (UBC). Other notable influences came from Wes Maciejewski (San José State University) who first suggested the idea of videos. I wish to thank Dr. Sophie Burrill for the formatting and improved figures. Many students contributed to proofreading, finding typos, or adding to the collection of problems. A partial list includes Glen Pugh, Carol Shao-Shan Huang and Lu Fan. I am also grateful to Will Engle at the Center for Teaching and Learning Technology (CTLT) at UBC for encouragement to make this work an open resource and to the UBC Studios for assistance and support during preparation of videos.

1

Power functions as building blocks

Like tall architectural marvels that are made of simple units (beams, bricks, and tiles), many interesting functions can be constructed from simpler building blocks. In this chapter, we study a family of simple functions, the power functions - those of the form $f(x) = x^n$.

Mastered Material Check

1. Can you define **function**?
2. Give an example of a polynomial function; a rational function.

Our first task is to understand properties of the members of this “family”. We will see that basic observations of power functions such as x^2, x^3 leads to insights into a biological problem of why the size of living cells is limited. Later, we use power functions as “ building blocks” to construct polynomials, and rational functions . We then develop important approaches to sketch the shapes of the resulting graphs.

1.1 Power functions

Section 1.1 Learning goals

1. Interpret the shapes of power functions relative to one another.
2. Justify that power functions with low powers dominate near the origin, and power functions with high powers dominate far away from the origin.
3. Identify the points of intersection of two power functions.

Let us consider the power functions, that is functions of the form

$$y = f(x) = x^n,$$

Click on this link and then adjust the slider on this interactive **desmos** graph to see how the power n affects the shape of a power function in the first quadrant.

where n is a positive integer. Power functions are among the most elementary and “elegant” functions - we only need multiplications to compute their value at any point. They are thus easy to calculate, very predictable and smooth, and, from the point of view of calculus, very easy to handle.

From Figure 1.1, we see that the power functions ($y = x^n$ for powers $n = 2, \ldots, 5$) intersect at $x = 0$ and $x = 1$. This is true for all positive integer powers. The same figure also demonstrates another fact helpful for curve-sketching: the greater the power n, the *flatter* the graph near the origin and the *steeper* the graph beyond $x > 1$. This can be restated in terms of the relative size of the power functions. We say that *close to the origin, the functions with lower powers dominate, while far from the origin, the higher powers dominate*.

More generally, a power function has the form

$$y = f(x) = K \cdot x^n$$

where n is a positive integer and K, sometimes called the **coefficient**, is a constant. So far, we have compared power functions whose coefficient is $K = 1$. We can extend our discussion to a more general case as well.

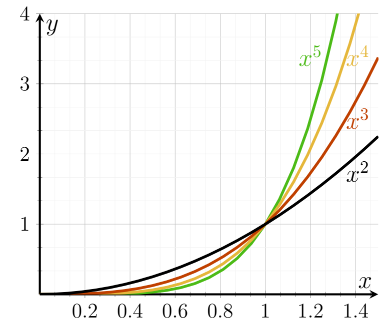


Figure 1.1: Graphs of a few power functions $y = x^n$. All intersect at $x = 0, 1$. As the power n increases, the graphs become flatter close to the origin, $(0,0)$, and steeper at large x-values.

Mastered Material Check

3. Use Figure 1.1 to approximate when $x^5 = 2$.
4. What is the first quadrant?

Example 1.1 *Find points of intersection and compare the sizes of the two power functions*

$$y_1 = ax^n, \quad \textit{and} \quad y_2 = bx^m.$$

where a and b are constants. You may assume that both a and b are positive.

Solution. This comparison is a slight generalization of the previous discussion. First, we note that the coefficients a and b merely scale the vertical behaviour (i.e. stretch the graph along the y axis). It is still true that the two functions intersect at $x = 0$; further, as before, the higher the power, the flatter the graph close to $x = 0$, and the steeper for large positive or negative values of x. However, now another point of intersection of the graphs occur when

$$ax^n = bx^m \quad \Rightarrow \quad x^{n-m} = (b/a).$$

We can solve this further to obtain a solution in the first quadrant

$$x = (b/a)^{1/(n-m)}. \tag{1.1}$$

This is shown in Figure 1.2 for the specific example of $y_1 = 5x^2, y_2 = 2x^3$. Close to the origin, the quadratic power function has a larger value, whereas for large x, the cubic function has larger values. The functions intersect when $5x^2 = 2x^3$, which holds for $x = 0$ or $x = \frac{5}{2} = 2.5$. ♢

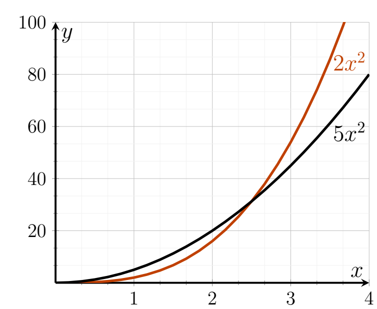


Figure 1.2: Graphs of two power functions, $y = 5x^2$ and $y = 2x^3$.

If b/a is positive, then in general the value given in (1.1) is a real number.

Example 1.2 *Determine points of intersection for the following pairs of functions:*

(a) $y_1 = 3x^4$ *and* $y_2 = 27x^2$,

(b) $y_1 = \left(\frac{4}{3}\right)\pi x^3$ *and* $y_2 = 4\pi x^2$.

Solution.

(a) Intersections occur at $x = 0$ and at $\pm(27/3)^{1/(4-2)} = \pm\sqrt{9} = \pm 3$.

(b) These functions intersect at $x = 0, 3$ but there are no other intersections at negative values of x. ♢

Note that in many cases, the points of intersection are irrational numbers whose decimal approximations can only be obtained by a scientific calculator or by some approximation method (such as **Newton's Method**, studied in Section 5.4)

Mastered Material Check

5. What is an irrational number?

With only these observations we can examine a biological problem related to the size of cells. By applying these ideas, we can gain insight into why cells have a size limitation, as discussed in the next section.

1.2 How big can a cell be? A model for nutrient balance

A summary of the cell size model. We discuss what cell size is consistent with a balance between nutrient absorption and consumption in a cell.

Section 1.2 Learning goals

1. Describe the derivation of a mathematical model for cell nutrient absorption and consumption.
2. Use parameters (k_1, k_2) rather than specific numbers in mathematical expressions.
3. Demonstrate the link between power functions in Section 1.1 and cell nutrient balance in the model.
4. Interpret the results of the model.

Consider the following biologically motivated questions:

- What physical and biological constraints determine the size of a cell?
- Why do some size limitations exist?
- Why should animals be made of millions of tiny cells, instead of a just a few large ones?

We already have enough mathematical prowess to address these questions - particularly if we assume a cell is spherical. Of course, this is often not the case. The shapes of living cells uniquely suit their functions. Many have long appendages, cylindrical parts, or branch-like structures. But here, we neglect all these beautiful complexities and look at a simple spherical cell because it suffices to answer our questions. Such mathematical simplifications can be very illuminating: they allow us to form a **mathematical model**. A mathematical model is just a representation of a real situation which simplifies things by representing the most important aspects, and neglecting or idealizing complicating details.

In this section, we follow a reasonable set of assumptions and mathematical facts to explore how nutrient balance can affect and limit cell size.

Building the model

In order to build the model we make some simplifying assumptions and then restate them mathematically. We base the model on the following

assumptions:

1. The cell is roughly spherical (See Figure 1.3).
2. The cell absorbs oxygen and nutrients through its surface. The larger the surface area, S, the faster the total rate of absorption. We assume that the rate at which nutrients (or oxygen) are absorbed is **proportional** to the surface area of the cell.
3. The rate at which nutrients are consumed (i.e., used up) in metabolism is proportional the volume, V, of the cell. The bigger the volume, the more nutrients are needed to keep the cell alive.

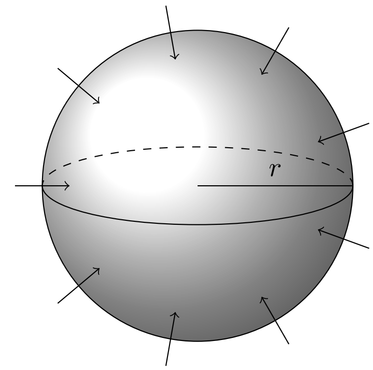


Figure 1.3: An assumed spherical cell absorbs nutrients at a rate proportional to its surface area S, but consumes nutrients at a rate proportional to its volume V.

Mastered Material Check

6. What does "A is proportional to B" mean?
7. What might the units be for quantities A, C, V, S and r?
8. Given your choices for 7., what are the units associated with k_1, k_2?

We define the following quantities for our model of a single cell:

$$\begin{aligned}
A &= \text{net rate of absorption of nutrients per unit time,}\\
C &= \text{net rate of consumption of nutrients per unit time,}\\
V &= \text{cell volume,}\\
S &= \text{cell surface area,}\\
r &= \text{ radius of the cell.}
\end{aligned}$$

We now rephrase the assumptions mathematically. By Assumption 2, the absorption rate, A, is proportional to S: this means that

$$A = k_1 S,$$

where k_1 is a **constant of proportionality**. Since absorption and surface area are positive quantities, only positive values of the proportionality constant make sense, so k_1 must be positive. The value of this constant depends on properties of the cell membrane such as its permeability or how many pores it contains to permit passage of nutrients. *By using a generic constant - called a **parameter** - to represent this proportionality constant, we keep the model general enough to apply to many different cell types.*

By Assumption 3, the rate of nutrient consumption, C is proportional to V, so that

$$C = k_2 V,$$

where $k_2 > 0$ is a second (positive) proportionality constant. The value of k_2 depend son the cell metabolism, i.e. how quickly it consumes nutrients in carrying out its activities.

By Assumption 1, the cell is spherical, thus its surface area, S, and volume, V, are:

$$S = 4\pi r^2, \quad V = \frac{4}{3}\pi r^3. \tag{1.2}$$

Putting these facts together leads to the following relationships between nutrient absorption A, consumption C, and cell radius r:

$$A = k_1(4\pi r^2) = (4\pi k_1)r^2,$$
$$C = k_2\left(\frac{4}{3}\pi r^3\right) = \left(\frac{4}{3}\pi k_2\right)r^3.$$

Rewriting this relationship as

$$A(r) = (4\pi k_1)r^2, \quad \text{and} \quad C(r) = \left(\frac{4}{3}\pi k_2\right)r^3. \qquad (1.3)$$

we observe that A, C are simply *power functions* of the cell radius, r, that is

$$A(r) = ar^2, \quad C(r) = cr^3.$$

Note: the powers are $n = 3$ for consumption and $n = 2$ for absorption.

The discussion of power functions in Section 1.1 now contributes to our analysis of how nutrient balance depends on cell size.

Mastered Material Check

9. What are constants a and c in terms of k_1 and k_2?
10. Why are we considering different values of r in Example 1.3?

Nutrient balance depends on cell size

In our discussion of cell size, we found two power functions that depend on the cell radius r, namely the nutrient absorption $A(r)$ and consumption $C(r)$ given in Eqns. (1.3). We first ask whether absorption or consumption of nutrients dominates for small, medium, or large cells.

Example 1.3 (A fine balance) *For what cell size is the consumption rate exactly balanced by the absorption rate? Which rate (consumption or absorption) dominates for small cells? For large cells?*

Solution.

The two rates "balance" (and their graphs intersect) when

$$A(r) = C(r) \quad \Rightarrow \quad \left(\frac{4}{3}\pi k_2\right)r^3 = (4\pi k_1)r^2.$$

A trivial solution to this equation is $r = 0$.

Note: this solution is not interesting biologically, but we should not forget it in mathematical analysis of such problems.

If $r \neq 0$, then, canceling a factor of r^2 from both sides gives:

$$r = 3\frac{k_1}{k_2}.$$

This means absorption and consumption rates are equal for cells of this size. For small r, the power function with the smaller power of r (namely $A(r)$) dominates, but for very large values of r, the power function with the higher power of r (namely $C(r)$) dominates. It follows that for smaller cells, absorption $A \approx r^2$ is the dominant process, while for larger cells, consumption rate $C \approx r^3$ dominates. *We conclude that cells larger than the critical size* $r = 3k_1/k_2$ *are unable to keep up with the nutrient demand, and cannot survive since consumption overtakes absorption of nutrients.* $\Diamond$

Using the above simple geometric argument, we deduced that cell size has strong implications on its ability to absorb nutrients or oxygen quickly enough to feed itself. For these reasons, cells larger than some maximal size (roughly 1mm in diameter) rarely occur.

A similar strategy also allows us consider the energy balance and sustainability of life on Earth - as seen next, in Section 1.3.

1.3 Sustainability and energy balance on Earth

Section 1.3 Learning goals

1. Justify the given mathematical model that describes the energy input and output on Planet Earth.
2. Use the given model to determine the energy equlibrium of the planet.

The sustainability of life on Planet Earth depends on a fine balance between the temperature of its oceans and land masses and the ability of life forms to tolerate climate change. As a follow-up to our model for nutrient balance, we introduce a simple energy balance model to track incoming and outgoing energy and determine a rough estimate for the Earth's temperature. We use the following basic assumptions:

1. Energy input from the sun, given the Earth's radius r, can be approximated as
$$E_{in} = (1-a)S\pi r^2, \tag{1.4}$$
where S is incoming radiation energy per unit area (also called the **solar constant**) and $0 \le a \le 1$ is the fraction of that energy reflected; a is also called the **albedo**, and depends on cloud cover, and other planet characteristics (such as percent forest, snow, desert, and ocean).

2. Energy lost from Earth due to radiation into space depends on the current temperature of the Earth T, and is approximated as
$$E_{out} = 4\pi r^2 \varepsilon \sigma T^4, \tag{1.5}$$
where ε is the **emissivity** of the Earth's atmosphere, which represents the Earth's tendency to emit radiation energy. This constant depends on cloud cover, water vapour, as well as on **greenhouse gas** concentration in the atmosphere; σ is a physical constant (the Stephan-Bolzmann constant) which is fixed for the purpose of our discussion.

Mastered Material Check

11. Do you think E_{in} is proportional to Earth's surface area or volume?

Notice there are several different symbols in Eqns. (1.4) and (1.5). Being clear about which are constants and which are variables is critical to using any mathematical model. As the next example points out, sometimes you have a choice to make.

Example 1.4 (Energy expressions are power functions) *Explain in what sense the two forms of energy above can be viewed as power functions, and what types of power functions they represent.*

Solution. Both E_{in} and E_{out} depend on Earth's radius as the power $\sim r^2$. However, since this radius is a constant, it is not fruitful to consider it as an interesting variable for this problem (unlike the cell size example in Section 1.2). However, we note that E_{out} depends on temperatureas $\sim T^4$. (We might also select the albedo as a variable and in that case, we note that E_{in} depends linearly on the albedo a.) ◊

Example 1.5 (Energy equilibrium for the Earth) *Explain how the assumptions above can be used to determine the equilibrium temperature of the Earth, that is, the temperature at which the incoming and outgoing radiation energies are balanced.*

Solution. The Earth is at equilibrium when

$$E_{in} = E_{out} \quad \Rightarrow \quad (1-a)S\pi r^2 = 4\pi r^2 \varepsilon \sigma T^4.$$

◊

We observe that the factors πr^2 cancel, and we obtain an equation that can be solved for the temperature T. It is instructive to examine how this temperature depends on the constants in the problem, and how it is affected by cloud cover and greenhouse gas level. This is also explored in Exercise 21

1.4 *First steps in graph sketching*

Section 1.4 Learning goals

1. Identify even and odd functions based on their graph.
2. Determine algebraically whether a function is even or odd.
3. Sketch the graph of a simple polynomial of the form $y = ax^n + bx^m$.
4. Sketch a rational function such as $y = Ax^n/(b + x^m)$.

Even and odd power functions

Adjust the slider to see how the even and odd power functions behave as their power increases.

So far, we have considered power functions $y = x^n$ with $x > 0$. But in mathematical generality, there is no reason to restrict the independent variable x to

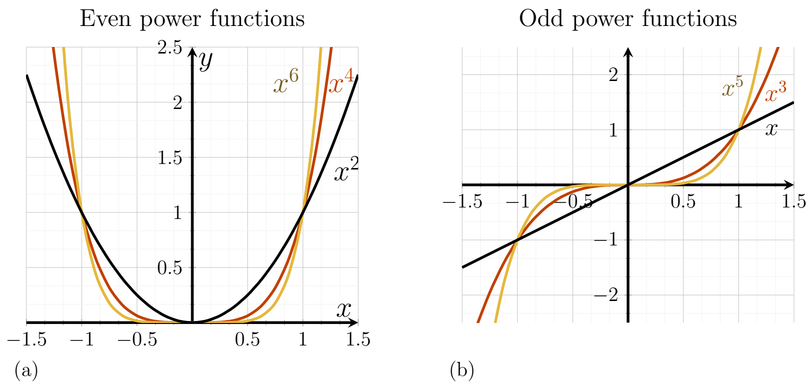


Figure 1.4: Graphs of power functions. (a) A few even power functions: $y = x^2$, $y = x^4$ and $y = x^6$. (b) Some odd power functions: $y = x$, $y = x^3$ and $y = x^5$. Note the symmetry properties.

positive values. Thus we expand the discussion to consider all real values of x. We examine now some symmetry properties that arise.

In Figure 1.4 (a) we see that power functions with an even power, such as $y = x^2$, $y = x^4$, and $y = x^6$, are symmetric about the y-axis. In Figure 1.4(b) we notice that power functions with an odd power, such as $y = x$, $y = x^3$ and $y = x^5$ are symmetric when rotated 180° about the origin. We adopt the term **even function** and **odd function** to describe such symmetry properties. More formally,

$$f(-x) = f(x) \quad \Rightarrow \quad \text{f is an even function,}$$
$$f(-x) = -f(x) \quad \Rightarrow \quad \text{f is an odd function}$$

Many functions are not symmetric at all, and are neither even nor odd. See Appendix C for further details.

Example 1.6 *Show that the function $y = g(x) = x^2 - 3x^4$ is an even function*

Solution. For g to be an even function, it should satisfy $g(-x) = g(x)$. Let us calculate $g(-x)$ and see if this requirement holds. We find that

$$g(-x) = (-x)^2 - 3(-x)^4 = x^2 - 3x^4 = g(x).$$

Here we have used the fact that $(-x)^n = (-1)^n x^n$, and that when n is even, $(-1)^n = 1$. ◊

All power functions are continuous and **unbounded**: for $x \to \infty$ both even and odd power functions satisfy $y = x^n \to \infty$. For $x \to -\infty$, odd power functions tend to $-\infty$. Odd power functions are **one-to-one**: that is, each value of y is obtained from a unique value of x and vice versa. This is not true for even power functions. From Fig 1.4 we see that all power functions go through the point $(0,0)$. Even power functions have a **local minimum** at the origin whereas odd power functions do not.

Mastered Material Check

12. Highlight the y-axes and circle the origins in Fig 1.4.
13. Consider Figure 1.4: where do even power functions intersect? Odd?
14. Show that $f(a) = a^5 - 3a$ is an odd function.
15. Give an example of a function which is **bounded**.
16. Verify $y = x^2$ is **not** one-to-one.
17. What graphical property do one-to-one functions share?

Definition 1.1 (Local Minimum) *A local minimum of a function $f(x)$ is a point x_{min} such that the value of f is larger at all sufficiently close points. Formally, $f(x_{min} \pm \varepsilon) > f(x_{min})$ for ε small enough.*

Sketching a simple (two-term) polynomial

Based on our familiarity with power functions, we now discuss functions made up of such components. In particular, we extend the discussion to **polynomials** (sums of power functions) and **rational functions** (ratios of such functions). We also develop skills in sketching graphs of these functions.

Example 1.7 (Sketching a simple cubic polynomial) *Sketch a graph of the polynomial*

$$y = p(x) = x^3 + ax. \tag{1.6}$$

How would the sketch change if the constant a changes from positive to negative?

Adjust the slider to see how positive and negative values of the coefficient a affect the shape of this simple polynomial.

Solution. The polynomial in Eqn. (1.6) has two terms, each one a power function. Let us consider their effects individually. Near the origin, for $x \approx 0$ the term ax dominates so that, close to $x = 0$, the function behaves as

$$y \approx ax.$$

This is a straight line with slope a. Hence, near the origin, if $a > 0$ we would see a line with positive slope, whereas if $a < 0$ the slope of the line should be negative. Far away from the origin, the cubic term dominates, so

$$y \approx x^3$$

at large (positive or negative) x values. Figure 1.5 illustrates these ideas.

In the first row we see the behaviour of $y = p(x) = x^3 + ax$ for large x, in the second for small x. The last row shows the graph for an intermediate range. We might notice that for $a < 0$, the graph has a local minimum as well as a local maximum. Such an argument already leads to a fairly reasonable sketch of the function in Eqn. (1.6). We can add further details using algebra to find **zeros** - that is where $y = p(x) = 0$. ◊

Mastered Material Check

17. Justify why the linear term dominates near the origin, while the cubic term dominates further out.
18. Sketch the graph of *any* function with horizontal asymptote $y = 2$.

Example 1.8 (Zeros) *Find the places at which the polynomial Eqn. (1.6) crosses the x axis, that is, find the* **zeros** *of the function $y = x^3 + ax$.*

Solution. The zeros of the polynomial can be found by setting

$$y = p(x) = 0 \quad \Rightarrow \quad x^3 + ax = 0 \quad \Rightarrow \quad x^3 = -ax.$$

The above equation always has a solution $x = 0$, but if $x \neq 0$, we can cancel and obtain

$$x^2 = -a.$$

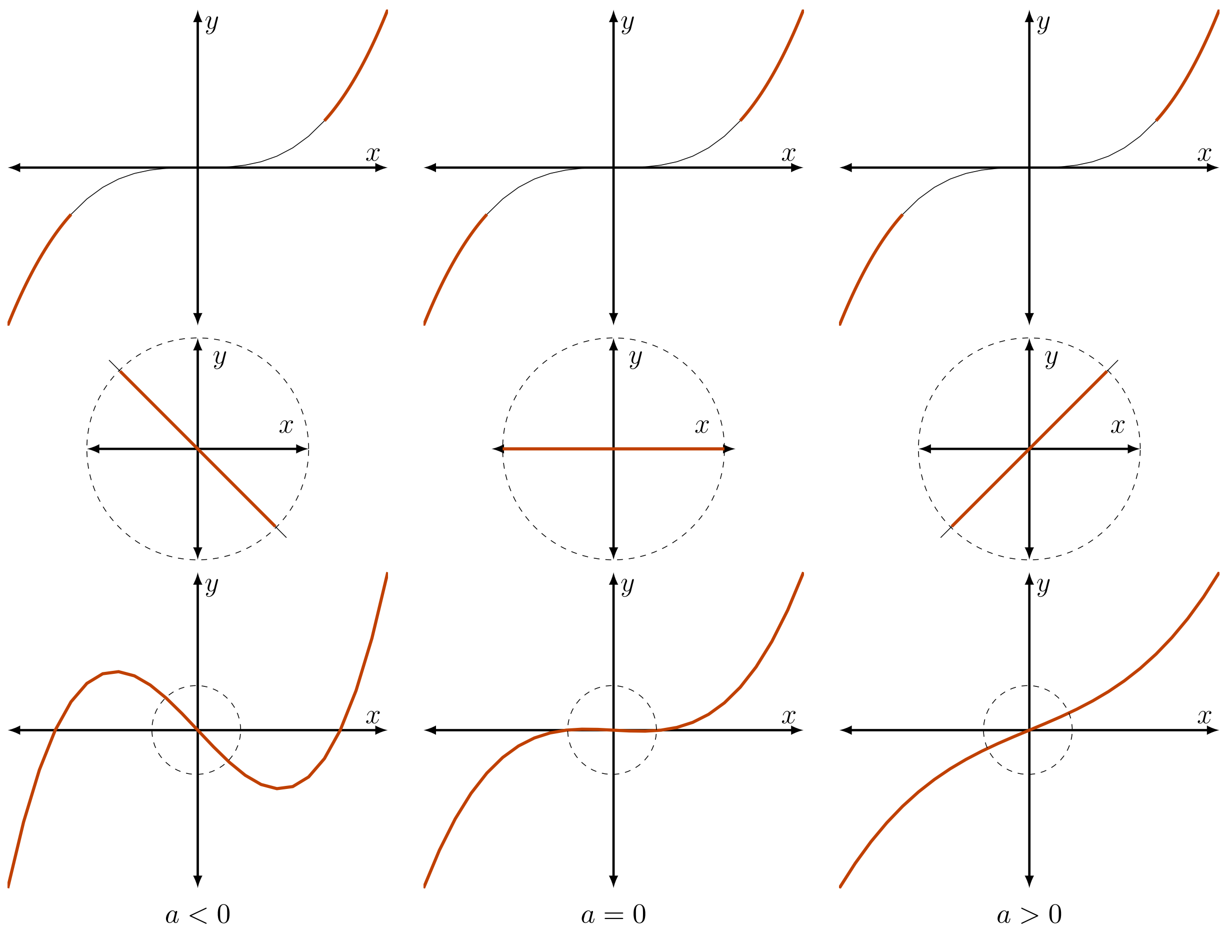


Figure 1.5: The graph of the polynomial $y = p(x) = x^3 + ax$ can be obtained by combining its two power function components. The cubic "arms" $y \approx x^3$ (top row) dominate for large x (far from the origin), while the linear part $y \approx ax$ (middle row) dominates near the origin. When these are smoothly connected (bottom row) we obtain a sketch of the desired polynomial. Shown here are three possibilities, for $a < 0, a = 0, a > 0$, left to right. The value of a determines the slope of the curve near $x = 0$ and thus also affects presence of a local maximum and minimum (for $a < 0$).

This would have no solutions if a is a positive number, so that in that case, the graph crosses the x axis only once, at $x = 0$, as shown in Figure 1.5. If a is negative, then the minus signs cancel, so the equation can be written in the form

$$x^2 = |a|$$

and we would have two new zeros at

$$x = \pm\sqrt{|a|}.$$

For example, if $a = -1$ then the function $y = x^3 - x$ has zeros at $x = 0, 1, -1$. ◊

Example 1.9 (A more general case) *Explain how you would use the ideas of Example 1.7 to sketch the polynomial $y = p(x) = ax^n + bx^m$. Without loss of generality, you may assume that $n > m \geq 1$ are integers.*

Solution. As in Example 1.7, this polynomial has two terms that dominate at different ranges of the independent variable. Close to the origin, $y \approx bx^m$

Mastered Material Check

19. Find the zeros of $y = x^3 + 3x$.

(since m is the lower power) whereas for large x, $y \approx ax^n$. The full behaviour is obtained by smoothly connecting these pieces of the graph. Finding zeros can refine the graph. ◊

A step back. The reasoning used here is an important first step in sketching the graph of a polynomial. In the ensuing chapters, we develop specialized methods to find zeros of more complicated functions (using an approximation technique called **Newton's method**). We also apply calculus tools to determine points at which the function attains local maxima or minima (called **critical points**), and how it behaves for very large positive or negative values of x. That said, the elementary steps described here remain useful as a quick approach for visualizing the overall shape of a graph.

Sketching a simple rational function

We apply similar reasoning to consider the graphs of simple rational functions. A **rational function** is a function that can be written as

$$y = \frac{p_1(x)}{p_2(x)}, \quad \text{where} \quad p_1(x) \text{ and } p_2(x) \text{ are polynomials.}$$

Example 1.10 (A rational function) *Sketch the graph of the rational function*

$$y = \frac{Ax^n}{a^n + x^n}, \quad x \geq 0. \tag{1.7}$$

What properties of your sketch depend on the power n? What would the graph look like for $n = 1, 2, 3$?

Adjust the sliders to see how the values of n, A, and a affect the shape of the rational function in (1.7).

Solution. We can break up the process of sketching this function into the following steps:

- The graph of the function in Eqn. (1.7) goes through the origin (at $x = 0$, we see that $y = 0$).
- For very small x, (i.e., $x << a$) we can approximate the denominator by the constant term $a^n + x^n \approx a^n$, since x^n is negligible by comparison, so that

  $$y = \frac{Ax^n}{a^n + x^n} \approx \frac{Ax^n}{a^n} = \left(\frac{A}{a^n}\right)x^n \quad \text{for small } x.$$

  This means that near the origin, the graph looks like a power function, $y \approx Cx^n$ (where $C = A/a^n$).
- For large x, i.e. $x >> a$, we have $a^n + x^n \approx x^n$ since x overtakes and dominates over the constant a, so that

  $$y = \frac{Ax^n}{a^n + x^n} \approx \frac{Ax^n}{x^n} = A \quad \text{for large } x.$$

  This reveals that the graph has a horizontal asymptote $y = A$ at large values of x.

Mastered Material Check

20. Why is a^n a constant?
21. Sketch the graph of *any* function with horizontal asymptote $y = 2$.

- Since the function behaves like a simple power function close to the origin, we conclude directly that the higher the value of n, the flatter is its graph near 0. Further, large n means sharper rise to the eventual asymptote.

The results are displayed in Figure 1.6. ◊

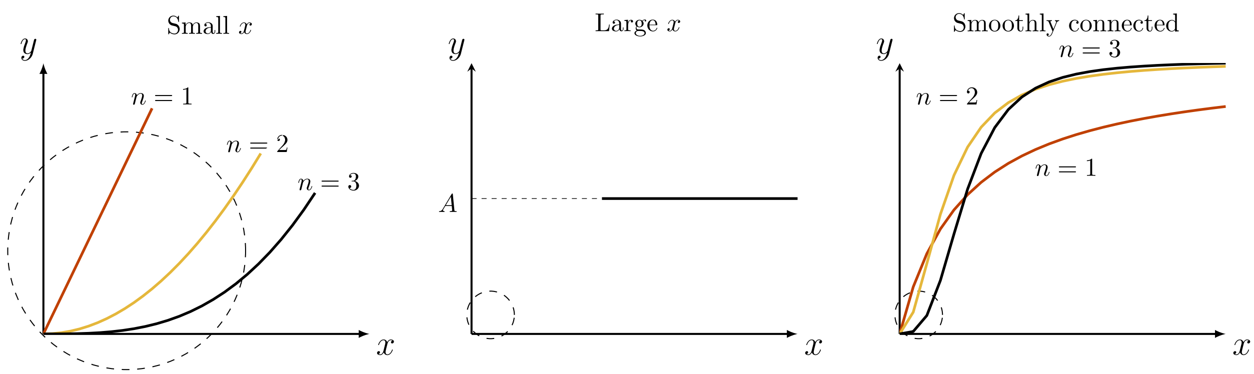


Figure 1.6: The rational functions Eqns.(1.7) with $n = 1, 2, 3$ are compared on this graph. Close to the origin, the function behaves like a power function, whereas for large x there is a horizontal asymptote at $y = A$. As n increases, the graph becomes flatter close to the origin, and steeper in its rise to the asymptote.

1.5 *Rate of an enzyme-catalyzed reaction*

Section 1.5 Learning goals

1. Describe the connection between Michaelis-Menten kinetics in biochemistry and rational functions described in Section 1.4.
2. Interpret properties of a graph such as Figure 1.8 in terms of properties of an enzyme-catalyzed reactions.

Rational functions introduced in Example 1.10 often play a role in biochemistry. Here we discuss two such examples and the contexts in which they appear. In both cases, we consider the initial rise of the function as well as its eventual saturation.

Saturation and Michaelis-Menten kinetics

Biochemical reactions are often based on the action of proteins known as **enzymes** that catalyze reactions in living cells. Fig. 1.7 depicts an enzyme E binding to its **substrate** S to form a **complex** C. The complex breaks apart into a **product**, P, and the original enzyme that can act once more. Substrate is usually plentiful relative to the enzyme.

In the context of this example, x represents the concentration of substrate in the reaction mixture. The speed of the reaction, v, (namely the rate at which product is formed) depends on x. When you actually graph the speed of the reaction as a function of the concentration, you see that it is not linear: Figure 1.8 is typical. This relationships, known as **Michaelis-Menten**

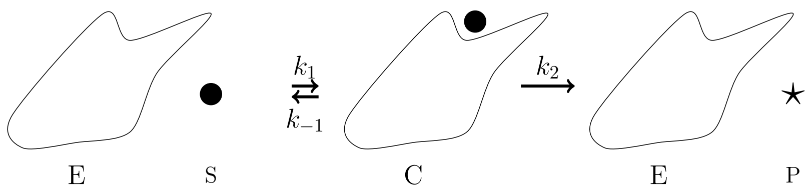


Figure 1.7: An enzyme (catalytic protein) is shown binding to a substrate molecule (circular dot) and then processing it into a product (star shaped molecule).

kinetics, has the mathematical form

$$\text{speed of reaction } = v = \frac{Kx}{k_n + x}, \tag{1.8}$$

where $K, k_n > 0$ are constants specific to the enzyme and the experimental conditions.

Equation (1.8) is a rational function. Since x is a concentration, it must be a positive quantity, so we restrict attention to $x \geq 0$. The expression in Eqn. (1.8) is a special case of the rational functions explored in Example 1.10, where $n = 1, A = K, a = k_n$. In Figure 1.8, we used plot this function for specific values of K, k_n. The following observations can be made

1. The graph of Eqn. (1.8) goes through the origin. Indeed, when $x = 0$ we have $v = 0$.

2. Close to the origin, the initial rise of the graph "looks like" a straight line. We can see this by considering values of x that are much smaller than k_n. Then the denominator $(k_n + x)$ is well approximated by the constant k_n. Thus, for small x, $v \approx (K/k_n)x$, so that the graph resembles a straight line through the origin with slope (K/k_n).

3. For large x, there is a horizontal asymptote. A similar argument for $x \gg k_n$, verifies that v is approximately constant at large enough x.

Michaelis-Menten kinetics represents one relationship in which **saturation** occurs: the speed of the reaction at first increases as substrate concentration x is raised, but the enzymes saturate and operate at a fixed constant speed K as more and more substrate is added.

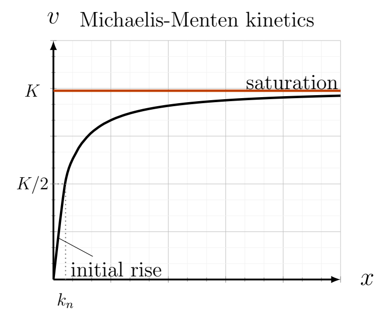


Figure 1.8: The graph of reaction speed, v, versus substrate concentration, x in an enzyme-catalyzed reaction, as in Eqn. 1.8. This behaviour is called Michaelis-Menten kinetics. Note that the graph at first rises almost like a straight line, but then it curves and approaches a horizontal asymptote. This graph tells us that the speed of the enzyme cannot exceed some fixed level, i.e. it cannot be faster than K.

	units	**example**
x	concentration	"nano Molar" , $nM \equiv 10^{-9}$ Moles per litre
v	concentration over time	nM min^{-1}
k_n		
K		

Table 1.1: Units for Michaelis-Menten kinetics, $v = \frac{Kx}{k_n + x}$.

Mastered Material Check

22. Complete Table 1.1.

Units. It is worth considering the units in Eqn. (1.8). Given that only quantities with identical units can be added or compared, and that the units of the two sides of the relationship *must balance*, fill Table 1.1.

Featured Problem 1.1 (Fish population growth 1) *The Beverton-Holt model relates the number of salmon in a population this year N_1 to the number of salmon that were present last year N_0, according to the relationship*

$$N_1 = k_1 \frac{N_0}{(1 + k_2 N_0)}, \quad k_1, k_2 > 0 \tag{1.9}$$

Sketch N_1 as a function of N_0 and explain how the constants k_1 and k_2 affect the shape of the graph you obtain. Is there a population level N_0 that would be exactly the same from one year to the next? Are there any restrictions on k_1 or k_2 for this kind of static ("steady state") population to be possible?

Hill functions

The Michaelis-Menten kinetics we discussed above fit into a broader class of **Hill functions**, which are rational functions of the form shown in Eqn. (1.7) with $n > 1$ and $A, a > 0$. This function is often referred to in the life sciences as *a Hill function with coefficient n*, (although the "coefficient" is actually a power in the terminology used in this chapter). Hill functions occur when an enzyme-catalyzed reaction benefits from **cooperativity** of a multi-step process. For example, the binding of the first substrate molecule may enhance the binding of a second.

Michaelis-Menten kinetics coincides with a Hill function for $n = 1$. In biochemistry, expressions of the form of Eqn. (1.7) with $n > 1$ are often denoted "sigmoidal" kinetics. Several such functions are plotted in Figure 1.9. We examined the shapes of these functions in Example 1.10.

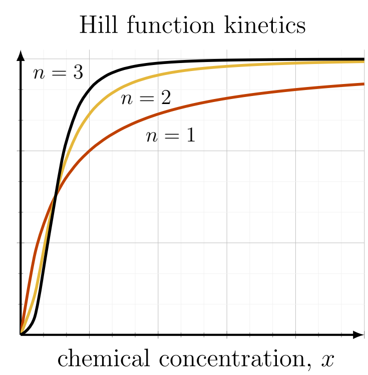


Figure 1.9: Hill function kinetics, from Eqn. (1.7), with $A = 3, a = 1$ and Hill coefficient $n = 1, 2, 3$. See also Fig 1.6 for an analysis of the shape of this graph.

All Hill functions have a horizontal asymptote $y = A$ at large values of x. If y is the speed of a chemical reaction (analogous to the variable we called v), then A is the "maximal rate" or "maximal speed" of the reaction. Since the Hill function behaves like a simple power function close to the origin, the higher the value of n, the flatter is its graph near 0, and the sharper the rise to the eventual asymptote. Hill functions with large n are often used to represent "switch-like" behaviour in genetic networks or biochemical signal transduction pathways.

The constant a is sometimes called the "half-maximal activation level" for the following reason: when $x = a$ then

$$v = \frac{Aa^n}{a^n + a^n} = \frac{Aa^2}{2a^2} = \frac{A}{2}.$$

This shows that the level $x = a$ leads to a reaction speed of $A/2$ which is half of the maximal possible rate.

Featured Problem 1.2 **Lineweaver-Burk plots.** *Hill functions can be transformed to a linear relationship through a change of variables. Consider the Hill function*

$$y = \frac{Ax^3}{a^3 + x^3}.$$

define $y = 1/Y$, $X = 1/x^3$. *Show that* Y *and* X *satisfy a linear relationship. Because we take the reciprocals of* x *and* y, X *and* Y *are sometimes called* reciprocal coordinates.

1.6 Predator Response

Interactions of predators and prey are often studied in ecology. Professor C.S. ("Buzz") Holling, (a former Director of the Institute of Animal Resource Ecology at the University of British Columbia) described three types of predators, termed "Type I", "Type II" and "Type III", according to their ability to consume prey as the prey density increases. The three Holling "predator functional responses" are shown in Fig. 1.10.

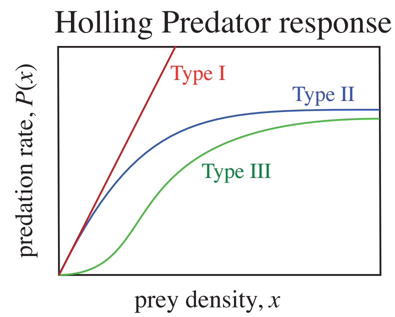


Figure 1.10: Holling's Type I, II, and III predator response. The predation rate $P(x)$ is the number of prey eaten by a predator per unit time. Note that the predation rate depends on the prey density x.

Quick Concept Checks

1. Match the predator responses shown in Fig. 1.10 with the descriptions given below

   1. As a predator, I get satiated and cannot keep eating more and more prey.
   2. I can hardly find the prey when the prey density is low, but I also get satiated at high prey density.
   3. The more prey there is, the more I can eat.

Based on Fig. 1.10, match the predator responses to functions shown below.

$$P_1(x) = kx,$$
$$P_2(x) = K\frac{x}{a+x},$$
$$P_3(x) = Kx^n, \quad n \geq 2$$
$$P_4(x) = K\frac{x^n}{a^n + x^n}, \quad n \geq 2$$

Hint: One of the curves "looks like a straight line" (so which function here is linear?). One of the choices is a power function. (Will it fit any of the other curves? why or why not?). Now consider the saturating curves and use our description of rational functions in Section 1.5 to select appropriate formulae for these functions.

The generality of mathematics allows us to adapt concepts we studied in one setting (enzyme biochemistry) to an apparently new topic (behaviour of predators).

See this short video explanation of the ladybug Type III predator response to its aphid prey.

1.6.1 A ladybug eating aphids

Here we use ideas developed so far to address a problem in population growth and biological control.

Featured Problem 1.3 (A balance of predation and aphid population growth) *Ladybugs are predators that love to eat aphids (their prey). Fig. 1.11 provides data* [1] *that supports the idea that ladybugs are type 3 predators.*

Let x= the number of aphids in some unit area (i.e., the density of the prey). Then the number of aphids eaten by a ladybug per unit time in that unit area will be called the **predation rate** *and denoted $P(x)$. The predation rate usually depends on the prey density, and we approximate that dependence by*

$$P(x) = K\frac{x^n}{a^n + x^n}, \quad \text{where} \quad K, a > 0. \tag{1.10}$$

Here we consider the case that $n = 2$. The aphids reproduce at a rate proportional to their number, so that the growth rate of the aphid population G (number of new aphids per hour) is

$$G(x) = rx \quad \text{where} \quad r > 0. \tag{1.11}$$

(a) *For what aphid population density x does the predation rate exactly balance the aphid population growth rate?*

(b) *Are there situations where the predation rate cannot match the growth rate? Explain your results in terms of the constants K, a, r.*

[1] MP Hassell, JH Lawton, and JR Beddington. Sigmoid functional responses by invertebrate predators and parasitoids. *The Journal of Animal Ecology*, 46(1):249–262, 1977

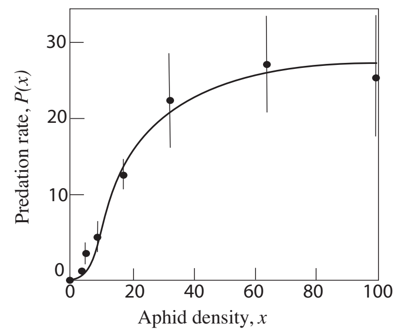


Figure 1.11: The predation rate of a ladybug depends on its aphid (prey) density.

Hints and partial solution

(a) The wording "the predation rate exactly balances the reproduction rate" means that the two functions $P(x)$ and $G(x)$ are exactly equal. Hence, to solve this problem, equate $P(x) = G(x)$ and determine the value of x (i.e., the number of aphids) at which this equality holds. You will find that one solution to this equation is $x = 0$. But if $x \neq 0$, you can cancel one factor of x from both sides and rearrange the equation to obtain a quadratic equation whose solution can be written down (in terms of the positive constants K, r, a).

(b) The solution you find in (a) is only a real number (i.e. a real solution exists) if the **discriminant** (quantity inside the square-root) is positive. Determine when this situation can occur and interpret your answer in terms of the aphid and ladybugs.

The solution to this problem is based on solving a quadratic equation, and so, relies on the fact that we chose the value $n = 2$ in the predation rate. What happens if $n > 2$? How do we solve the same kind of problem if $n = 3, 4$ etc? We return to this issue, and develop an approximate technique (Newton's method) in a later chapter.

Use the sliders to manipulate the predation constants K, a and the aphid growth rate parameter r. How many solutions are there to $P(x) = R(x)$? Show that for some parameter values, there is only a trivial solution at $x = 0$. Make a connection between this observation and part (b) of Example 1.3.

Hint: Recall that a quadratic equation $ax^2 + bx + c = 0$ has roots $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$. These roots are real provided $b^2 - 4ac \geq 0$.

1.7 Summary

1. Functions of the form $f(x) = K \cdot x^n$ (n a positive integer) are called *power functions* with coefficient K.
2. Power functions with larger powers of n form graphs that are flatter near the origin and steeper for $x > 1$.
3. An even function satisfies $f(-x) = f(x)$; an odd function satisfies $f(-x) = -f(x)$. Identifying even and odd functions can aid in graph sketching.
4. The zeros of a function $f(x)$ are roots of the equation $f(x) = 0$. Identifying the root(s) of a function helps in sketching its graph.
5. Polynomials are sums of power functions. Rational functions are ratios of polynomials. By examining the behaviour of terms that dominate near and far from the origin, we can obtain a rough sketch of such functions.
6. Mathematical models can be used to describe scientific phenomenon. Making reasonable assumptions and observations are necessary for building a successful model. Translating these assumptions and observations into mathematics is the key.
7. Hill functions can be transformed into a linear relationship using a change of variables; the plots that result are called Lineweaver-Burk plots.
8. The mathematical models explored in this chapter concerned:

   (a) cell size, based on nutrient balance;
   (b) energy balance on Earth;
   (c) biochemical reactions and Michaelis-Menten kinetics; and
   (d) enzyme-catalyzed reactions and Hill functions.

9. Units, while often suppressed in math texts, can be immensely useful in solving application problems. Only quantities with identical units can be added, or compared. Two sides of an equation must have identical units.

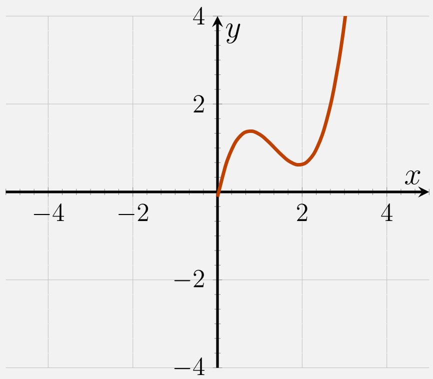
Quick Concept Checks

1. When is $x^2 > x^{10}$?

   (a) never (b) always (c) for small x (d) for large x

2. Why do we make assumptions when we build mathematical models?

3. Complete the sketch of the following graph, given that it is



   (a) an even function

   (b) an odd function

4. What is the relationship between Michaelis-Menten kinetics and Hill functions?

Exercises

1.1. **Power functions.** Consider the power function

$$y = ax^n, \quad -\infty < x < \infty.$$

Explain, possibly using a sketch, how the shape of the function changes when the coefficient a increases or decreases (for fixed n). How is this change in shape different from the shape change that results from changing the power n?

1.2. **Transformations.** Consider the graphs of the simple functions $y = x$, $y = x^2$, and $y = x^3$. Describe what happens to each of these graphs when the functions are *transformed* as follows:

(a) $y = Ax$, $y = Ax^2$, and $y = Ax^3$ where $A > 1$ is some constant.

(b) $y = x + a$, $y = x^2 + a$, and $y = x^3 + a$ where $a > 0$ is some constant.

(c) $y = (x - b)^2$, and $y = (x - b)^3$ where $b > 0$ is some constant.

1.3. **Sketching transformations.** Sketch the graphs of the following functions:

(a) $y = x^2$,

(b) $y = (x + 4)^2$,

(c) $y = a(x - b)^2 + c$ for the case $a > 0$, $b > 0$, $c > 0$.

(d) Comment on the effects of the constants a, b, c on the properties of the graph of $y = a(x - b)^2 + c$.

1.4. **Sketching polynomials.** Use arguments from Section 1.4 to sketch graphs of the following polynomials:

(a) $y = 2x^5 - 3x^2$,

(b) $y = x^3 - 4x^5$.

1.5. **Finding points of intersection.**

(a) Consider the two functions $f(x) = 3x^2$ and $g(x) = 2x^5$. Find all points of intersection of these functions.

(b) Repeat for functions $f(x) = x^3$ and $g(x) = 4x^5$.

Note: finding these points of intersection is equivalent to calculating the **zeros** of the functions in Exercise 4.

1.6. **Qualitative sketching skills.**

(a) Sketch the graph of the function $y = \rho x - x^5$ for positive and negative values of the constant ρ. Comment on behaviour close to zero and far away from zero.

(b) What are the zeros of this function and how does this depend on ρ?

(c) For what values of ρ would you expect that this function would have a local maximum ("peak") and a local minimum ("valley")?

1.7. **Finding points of intersection.** Consider functions $f(x) = Ax^n$ and $g(x) = Bx^m$. Suppose $m > n > 1$ are integers, and $A, B > 0$. Determine the values of x at which the the functions are the same - i.e. they intersect. Are there two places of intersection or three? How does this depend on the integer $m - n$?

Note: The point $(0,0)$ is always an intersection point. Thus, we are asking: when is there only *one* more and when there are *two* more intersection points? See Exercise 5 for an example of both types.

1.8. **More intersection points.** Find the intersection of each pair of functions.

(a) $y = \sqrt{x}$, $y = x^2$,

(b) $y = -\sqrt{x}$, $y = x^2$,

(c) $y = x^2 - 1$, $\frac{x^2}{4} + y^2 = 1$.

1.9. **Crossing the x-axis.** Answer the following by solving for x in each case. Find all values of x for which the following functions cross the x-axis (equivalently: the **zeros** of the function, or **roots** of the equation $f(x) = 0$.)

(a) $f(x) = I - \gamma x$, where I, γ are positive constants.

(b) $f(x) = I - \gamma x + \varepsilon x^2$, where I, γ, ε are positive constants. Are there cases where this function does not cross the x axis?

(c) In the case where the root(s) exist in part (b), are they positive, negative or of mixed signs?

1.10. **Crossing the x-axis.** Answer Exercise 9 by sketching a rough graph of each of the functions in parts (a-b) and using these sketches to determine how many real roots there are and where they are located (positive vs. negative x-axis).

Note: this exercise provides qualitative analysis skills that are helpful in later applications.

1.11. **Power functions.** Consider the functions $y = x^n$, $y = x^{1/n}$, $y = x^{-n}$, where n is an integer $n = 1, 2, \ldots$).

(a) Which of these functions increases most steeply for values of x greater than 1?

(b) Which decreases for large values of x?

(c) Which functions are not defined for negative x values?

(d) Compare the values of these functions for $0 < x < 1$.

(e) Which of these functions are not defined at $x = 0$?

1.12. **Roots of a quadratic.** Find the range m such that the equation $x^2 - 2x - m = 0$ has two unequal roots.

1.13. **Rational Functions.** Describe the shape of the graph of the function $y = Ax^n/(b + x^m)$ in two cases:

(a) $n > m$ and

(b) $m > n$.

1.14. **Power functions with negative powers.** Consider the function

$$f(x) = \frac{A}{x^a}$$

where $A > 0, a > 1$, with a an integer. This is the same as the function $f(x) = Ax^{-a}$, which is a power function with a negative power.

(a) Sketch a rough graph of this function for $x > 0$.

(b) How does the function change if A is increased?

(c) How does the function change if a is increased?

1.15. **Intersections of functions with negative powers.** Consider two functions of the form

$$f(x) = \frac{A}{x^a}, \quad g(x) = \frac{B}{x^b}.$$

Suppose that $A, B > 0,\ a, b > 1$ and that $A > B$. Determine where these functions intersect for positive x values.

1.16. **Zeros of polynomials.** Find all real zeros of the following polynomials:

(a) $x^3 - 2x^2 - 3x$,

(b) $x^5 - 1$,

(c) $3x^2 + 5x - 2$.

(d) Find the points of intersection of the functions $y = x^3 + x^2 - 2x + 1$ and $y = x^3$.

1.17. **Inverse functions.** The functions $y = x^3$ and $y = x^{1/3}$ are *inverse functions* (see Section 10.3 for a discussion of inverse functions).

(a) Sketch both functions on the same graph for $-2 < x < 2$ showing clearly where they intersect.

(b) The tangent line to the curve $y = x^3$ at the point $(1, 1)$ has slope $m = 3$, whereas the tangent line to $y = x^{1/3}$ at the point $(1, 1)$ has slope $m = 1/3$. Explain the relationship of the two slopes.

1.18. **Properties of a cube.** The volume V and surface area S of a cube whose sides have length a are given by the formulae

$$V = a^3, \quad S = 6a^2.$$

Note that these relationships are expressed in terms of power functions. The independent variable is a, not x. We say that "V is a function of a" (and also "S is a function of a").

(a) Sketch V as a function of a and S as a function of a on the same set of axes. Which one grows faster as a increases?

(b) What is the ratio of the volume to the surface area; that is, what is $\frac{V}{S}$ in terms of a? Sketch a graph of $\frac{V}{S}$ as a function of a.

(c) The formulae above tell us the volume and the area of a cube of a given side length. Suppose we are given either the volume or the surface area and asked to find the side.

(i) Find the length of the side as a function of the volume (i.e. express a in terms of V).

(ii) Find the side as a function of the surface area.

(iii) Use your results to find the side of a cubic tank whose volume is 1 litre.

(iv) Find the side of a cubic tank whose surface area is 10 cm^2.

Units.
Note that 1 litre = 10^3 cm^3.

1.19. **Properties of a sphere.** The volume V and surface area S of a sphere of radius r are given by the formulae

$$V = \frac{4\pi}{3} r^3, \quad S = 4\pi r^2.$$

Note that these relationships are expressed in terms of power functions with constant multiples such as 4π. The independent variable is r, not x. We say that "V is a function of r" (and also "S is a function of r").

(a) Sketch V as a function of r and S as a function of r on the same set of axes. Which one grows faster as r increases?

(b) What is the ratio of the volume to the surface area; that is, what is $\frac{V}{S}$ in terms of r? Sketch a graph of $\frac{V}{S}$ as a function of r.

(c) The formulae above tell us the volume and the area of a sphere of a given radius. But suppose we are given either the volume or the surface area and asked to find the radius.

(i) Find the radius as a function of the volume (i.e. express r in terms of V).

(i) Find the radius as a function of the surface area.

(i) Use your results to find the radius of a balloon whose volume is 1 litre.

(i) Find the radius of a balloon whose surface area is 10 cm^2

1.20. **The size of cell.** Consider a cell in the shape of a thin cylinder (length L and radius r). Assume that the cell absorbs nutrient through its surface at rate $k_1 S$ and consumes nutrients at rate $k_2 V$ where S, V are the surface area and volume of the cylinder. Here we assume that $k_1 = 12\mu\text{M}\ \mu\text{m}^{-2}$ per min and $k_2 = 2\mu\text{M}\ \mu\text{m}^{-3}$ per min.

Units.
Note that μM is 10^{-6} moles and μm is 10^{-6} meters.

(a) Use the fact that a cylinder (without end-caps) has surface area $S = 2\pi rL$ and volume $V = \pi r^2 L$ to determine the cell radius such that the rate of consumption exactly balances the rate of absorption.

(b) What do you expect happens to cells with a bigger or smaller radius?

(c) How does the length of the cylinder affect this nutrient balance?

1.21. **Energy equilibrium for Earth.** This exercise focuses on Earth's temperature, climate change, and sustainability.

(a) Complete the calculation for Example 1.5 by solving for the temperature T of the Earth at which incoming and outgoing radiation energies balance.

(b) Assume that greenhouse gases decrease the emissivity ε of the Earth's atmosphere. Explain how this would affect the Earth's temperature.

(c) Explain how the size of the Earth affects its energy balance according to the model.

(d) Explain how the albedo a affects the Earth's temperature.

1.22. **Allometric relationship.** Properties of animals are often related to their physical size or mass. For example, the metabolic rate of the animal (R), and its pulse rate (P) may be related to its body mass m by the approximate formulae $R = Am^b$ and $P = Cm^d$, where A, C, b, d are positive constants. Such relationships are known as *allometric* relationships.

(a) Use these formulae to derive a relationship between the metabolic rate and the pulse rate (*hint:* eliminate m).

(b) A similar process can be used to relate the Volume $V = (4/3)\pi r^3$ and surface area $S = 4\pi r^2$ of a sphere to one another. Eliminate r to find the corresponding relationship between volume and surface area for a sphere.

1.23. **Rate of a very simple chemical reaction.** We consider a chemical reaction that does not saturate, and a simple linear relationship between reaction speed and reactant concentration.

A chemical is being added to a mixture and is used up in a reaction. The rate of change of the chemical, (also called "the rate of the reaction") v M/sec is observed to follow a relationship

Units.
Note that M stands for Molar, which is the number of moles per litre.

$$v = a - bc$$

where c is the reactant concentration (in units of M) and a, b are positive constants.
Note: v is considered to be a function of c, and moreover, the relationship between v and c is assumed to be linear.

(a) What units should a and b have to make this equation consistent?

Note: in an equation such as $v = a - bc$, each of the three terms *must have* the same units. Otherwise, the equation would not make sense.

(b) Use the information in the graph shown in Figure 1.12 to find the values of a and b (*hint:* find the equation of the line in the figure, and compare it to the relationship $v = a - bc$).

(c) What is the rate of the reaction when $c = 0.005$ M?

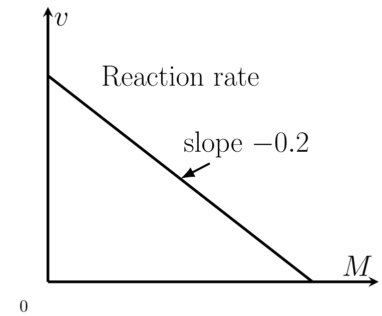


Figure 1.12: Figure for Exercise 23; rate of a chemical reaction.

1.24. **Michaelis-Menten kinetics.** Consider the Michaelis-Menten kinetics where the speed of an enzyme-catalyzed reaction is given by $v = Kx/(k_n + x)$.

(a) Explain the statement that "when x is large there is a horizontal asymptote" and find the value of v to which that asymptote approaches.

(b) Determine the reaction speed when $x = k_n$ and explain why the constant k_n is sometimes called the "half-max" concentration.

1.25. **A polymerization reaction.** Consider the speed of a polymerization reaction shown in Figure 1.13. Here the rate of the reaction is plotted as a function of the substrate concentration; this experiment concerned the polymerization of actin, an important structural component of cells; data from [Rohatgi et al., 2001]. The experimental points are shown as dots, and a Michaelis-Menten curve has been drawn to best fit these points. Use the data in the figure to determine approximate values of K and k_n in the two treatments shown.

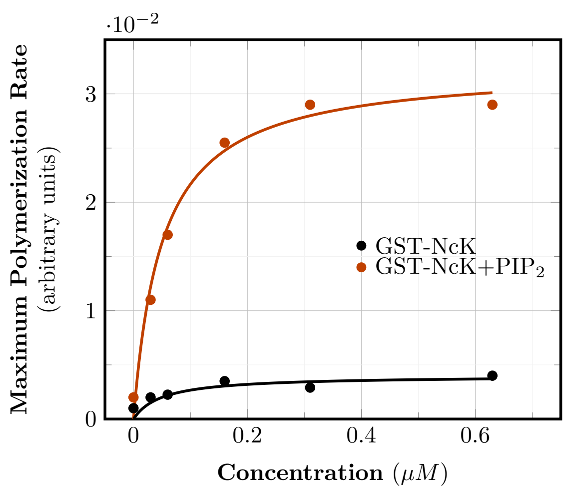


Figure 1.13: Figure for Exercise. 25; speed of polymerization.

1.26. **Hill functions.** Hill functions are sometimes used to represent a biochemical "switch," that is a rapid transition from one state to another. Consider the Hill functions

$$y_1 = \frac{x^2}{1+x^2}, \quad y_2 = \frac{x^5}{1+x^5},$$

(a) Where do these functions intersect?

(b) What are the asymptotes of these functions?

(c) Which of these functions increases fastest near the origin?

(d) Which is the sharpest "switch" and why?

1.27. **Transforming a Hill function to a linear relationship.** A Hill function is a nonlinear function - but if we redefine variables, we can transform it into a linear relationship. The process is analogous to transforming Michaelis-Menten kinetics into a Lineweaver-Burk plot, as discussed in Appendix G.1.

(a) Determine how to define appropriate variables X and Y (in terms of the original variables x and y) so that the Hill function $y = Ax^3/(a^3+x^3)$ is turned into a linear relationship between X and Y.

(b) Indicate how the slope and intercept of the line are related to the original constants A, a in the Hill function.

1.28. **Hill function and sigmoidal chemical kinetics.** It is known that the *rate* v at which a certain chemical reaction proceeds depends on the *concentration* of the reactant c according to the formula

$$v = \frac{Kc^2}{a^2+c^2},$$

where K, a are some constants. When the chemist plots the values of the quantity $1/v$ (on the "y" axis) versus the values of $1/c^2$ (on the "x axis"), she finds that the points are best described by a straight line with y-intercept 2 and slope 8. Use this result to find the values of the constants K and a.

1.29. **Lineweaver-Burk plots.** Shown in the Figure 1.14(a) and (b) are two Lineweaver-Burk plots (see Appendix G.1). By noting properties of these figures comment on the comparison between the following two enzymes:

(a) Enzyme (1) and (2).

(b) Enzyme (1) and (3).

1.30. **Michaelis-Menten enzyme kinetics.** The rate of an enzymatic reaction according to the *Michaelis-Menten kinetics* assumption is

$$v = \frac{Kc}{k_n+c},$$

where c is concentration of substrate (shown on the x-axis) and v is the reaction speed (given on the y-axis). Consider the data points given in Table 1.2.

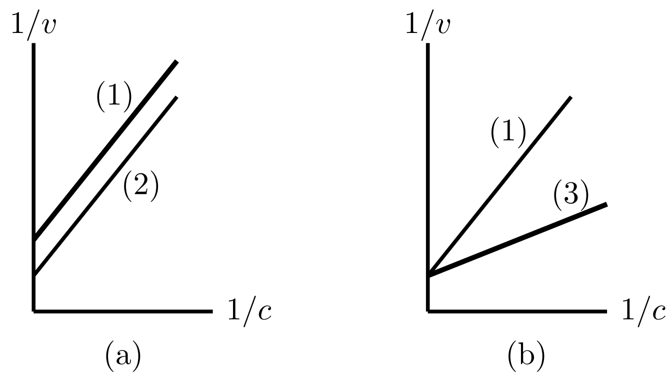


Figure 1.14: Figure for Exercise 29: Lineweaver-Burk plots.

Substrate concentration	nM	c	5	10	20	40	50	100
Reaction speed	nM/min	v	0.068	0.126	0.218	0.345	0.39	0.529

Table 1.2: Chemical reaction speed data.

(a) Convert this data to a Lineweaver-Burk (linear) relationship (see Appendix G.1 for discussion).

(b) Plot the transformed data values on a graph or spreadsheet, and estimate the slope and y-intercept of the line you get.

(c) Use these results to find the best estimates for K and k_n.

1.31. **Spacing in a school of fish** According to the biologist Breder [Breder, 1951], two fish in a school prefer to stay some specific distance apart. Breder suggested that the fish that are a distance x apart are attracted to one another by a force $F_A(x) = A/x^a$ and repelled by a second force $F_R(x) = R/x^r$, to keep from getting too close. He found the preferred spacing distance (also called the *individual distance*) by determining the value of x at which the repulsion and the attraction exactly balance.

Find the *individual distance* in terms of the quantities A, R, a, r (all assumed to be positive constants.)

2

Average rates of change, average velocity and the secant line

A physicist might study the motion of a falling ball by taking strobe images at fixed time intervals, and gluing them side by side to get a record of position of the ball over time. In a similar manner, cell biologists study the motion of proteins inside living cells. First, the proteins are labeled by fluorescent "tags" (this makes them visible in microscopic images). Then images of some thin strip of the cell are made at fixed time intervals, in regions through which the "glowing" (fluorescent) proteins move. Finally, those thin strips are "glued" together to form a record of the protein position over time, as shown in each panel of Fig 2.1. Biologists refer to such images as **kymographs**.

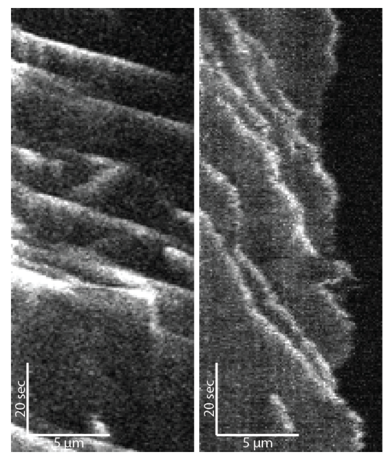


Figure 2.1: Cell biologists track the motion of proteins inside cells on images like this kymograph. Data courtesy of Anna Akhmanova's lab shows the position of proteins that track growing ends of long biopolymers called microtubules as they get longer or shrink inside cells. Displacement is shown horizontally (scale bar 5 μm) and time vertically (scale bar 20 sec).

The "streaks of light" in such kymographs allow us to determine the locations of the labeled proteins over time. as well as their velocity in the cell. But how fast were these proteins moving? Why are there zigzags in the left panel? And what happened in the treated cells (right panel) that made the streaks look different from those in the "normal cell" (left panel)[1]?

[1] Benjamin P Bouchet, Ivar Noordstra, Miranda van Amersfoort, Eugene A Katrukha, York-Christoph Ammon, Natalie D Ter Hoeve, Louis Hodgson, Marileen Dogterom, Patrick WB Derksen, and Anna Akhmanova. Mesenchymal cell invasion requires cooperative regulation of persistent microtubule growth by slain2 and clasp1. *Developmental cell*, 39(6):708–723, 2016

In this chapter we develop the tools to address some of these questions, and to characterize what we mean by velocity. As a first step, we introduce average rate of change. To motivate the idea, we examine data for common processes: changes in temperature, and motion of a falling object. Simple experiments are described in each case, and some features of the data are discussed. Based on each example, we calculate net change over some time interval and then define the **average rate of change** and average velocity. This concept generalizes to functions of any variable (not only time). We interpret this idea geometrically, in terms of the slope of a **secant line**.

In both cases, we ask how to use average rate of change (over a given interval) to find better and better approximations of the rate of change at a single instant, (i.e. at a point). We find that one way to arrive at this abstract concept entails refining the dataset - collecting data at closer and closer time points. A second - more abstract - way is to use a limit. Eventually, this procedure allows us to arrive at the definition of the **derivative**, which is the **instantaneous rate of change**.

2.1 *Time-dependent data and rates of change*

Section 2.1 Learning goals

1. Use (your favorite) graphical software package (spreadsheet, graphics calculator, online tools, etc.) to plot data points such as those in Table 2.1.
2. Describe the trends seen in such data using words such as 'increasing', 'decreasing', 'linear', 'nonlinear', 'shallow', 'steep', 'changes', etc.

In this section we consider time dependent processes and develop the idea of rates of change. We also use graphical software to represent the data.

Milk temperature and a recipe for yoghurt

To make yoghurt, heat milk to 190°F to kill off bacteria, then cool to 110°F. Add a spoonful of "live" pre-made yoghurt and keep the mixture at 110°F for 7-8 hours. This promotes growth of the microorganism *Lactobacillus*, that turns milk into yoghurt.

time (min)	Temperature (°F)
0.0	44.3
0.5	61
1.0	77
1.5	92
2.0	108
2.5	122
3.0	135.3
3.5	149.2
4.0	161.9
4.5	174.2
5.0	186

Table 2.1: Heating milk temperature data.

time (min)	Temperature (°F)
0	190
2	176
4	164.6
6	155.4
8	148
10	140.9
14	131
18	123
22	116
26	111.2

Table 2.2: Cooling milk temperature data.

Example 2.1 (Heating and cooling milk) *Shown in Tables 2.1 and 2.2 are sets of temperature measurements over time. Use your favorite software to plot the data and describe the trends you see in each graph.*

Solution. The data is plotted in Figure 2.2, and points are connected with line segments. The heating phase is shown on the left. (Temperature increases at a nearly linear rate.) On the right, the milk is cooling and the temperature decreases, but the slope of the graph becomes shallower with time. ◊

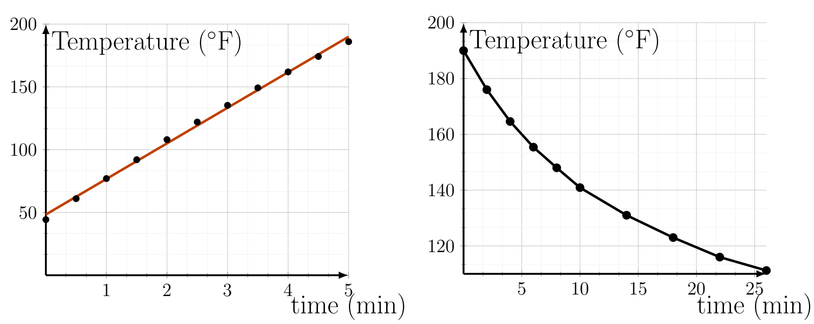


Figure 2.2: Plot of temperatures of milk being heated (left) and cooled (right).

In this chapter, we will be concerned with describing the rates of change of similar processes, that is in quantifying what we mean by "how fast is the temperature changing" in examples of this sort. Before answering, we introduce two other examples of time dependent data.

Mastered Material Check

1. Reproduce one of the graphs in Figure 2.2 using Celsius.

Data for swimming tuna

Example 2.2 (Bluefin tuna swimming distances) *The tuna fishing industry is of great economic value, but the danger of overfishing is recognized. Prof Molly Lutcavage*[2] *studied the swimming behaviour of Atlantic bluefin tuna (*Thunnus thynnus L.*) in the Gulf of Maine. She recorded their position over a period of 1-2 days. Some of her approximate data is given in Table 2.3. Plot the data points and describe the trends these display.*

[2] ME Lutcavage, RW Brill, GB Skomal, BC Chase, JL Goldstein, and J Tutein. Tracking adult north atlantic bluefin tuna (thunnus thynnus) in the northwestern atlantic using ultrasonic telemetry. *Marine Biology*, 137(2):347–358, 2000

Solution. As shown in Figure 2.3, distance traveled by Tuna 1 is roughly proportional to time spent, since its graph is roughly linear (almost a straight line). This linear relationship between distance travelled and time spent is called **uniform motion**.

Tuna 2 started with similar uniform motion, but later sped up. During $15 \leq t \leq 20$h, it was swimming much faster. ◊

time (hr)	**distance** Tuna 1 (km)	**distance** Tuna 2 (km)
0	0	0
5	29	32
10	51	55
15	78	80
20	140	111
25	160	125
30	182	150
35	218	180

Table 2.3: Data for tuna swimming distance collected by Prof. Molly Lutcavage in the Gulf of Maine.

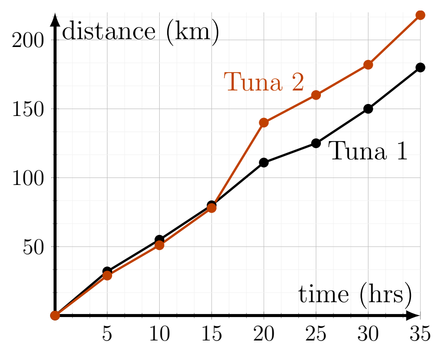


Figure 2.3: Distance travelled by two bluefin tuna over 35 hrs

Distance of a falling object

Long ago, Galileo devised some ingenious experiments to track the position of a falling object. He used his measurements to quantify the relationship between the total distance fallen over a given time. Although Galileo did not have formulae nor graph-paper in his day - and was thus forced to express this relationship in a cumbersome verbal way - what he had discovered was quite remarkable.

Example 2.3 (Gallileo's formula for height of a falling object) *Galileo discovered that the distance fallen, $y(t)$, is proportional to the square of the time t, that is*

$$y(t) = ct^2, \tag{2.1}$$

where c is a constant. When distance is measured in meters (m) and time in seconds (s) the constant is found to be $c = 4.9\text{m}/\text{s}^2$. Using Eqn. (2.1), plot a graph of the distance fallen $y(t)$ versus time t for $0 \leq t \leq 2$ seconds at intervals of 0.1s. Connect the data points and comment on the shape of the graph.

Mastered Material Check

2. Use Figure 2.3 to approximate how far each tuna travelled after 18 hours.
3. Does an object dropped from a height of 15m hit the ground in 2 seconds? 1 second?

Solution. The graph is shown in Figure 2.4. We recognize this as a parabola, resulting from the quadratic relationship of y and t. (In fact, the relationship is that of a simple power function with a constant coefficient.) ◊

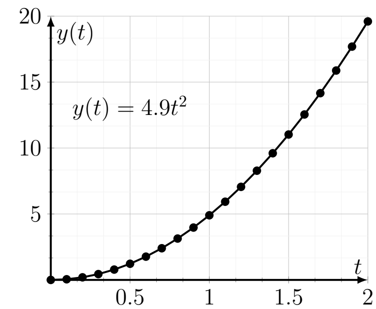


Figure 2.4: The distance y (in meters) of an object falling under the force of gravity, versus time t (in seconds).

Having looked at three examples of data for time-dependent processes, we now turn to quantifying the rate at which change occurs in each process. We start with the notion of average rate of change, and eventually refine and idealize this idea it to develop rates of change at an instant in time.

2.2 *The slope of a straight line is a rate of change*

Section 2.2 Learning goals

1. Define rate of change for a linear relationship.
2. Compute the rate of change for a linear relationship.

Mastered Material Check

4. What does it mean for two variables to have a linear relationship?
5. Why are we 'idealizing'?
6. Compute the rate of change for a linear relationship which goes through points $(1,4)$ and $(2,2)$.

In the examples discussed so far, we have plotted data and used words to describe trends. Our goal now is to formalize the idea of **change** and **rate of change**. Let us consider the simplest case where a variable of interest, y depends linearly on time, t. This was approximately true in some examples seen previously (Figure 2.2a, parts of Figure 2.3). We can describe this kind of relationship by the idealized equation

$$y(t) = mt + b. \tag{2.2}$$

A graph of y versus t is then a **straight line** with **slope** m and **intercept** b.

The equation of a straight line (2.2) specifies the slope m and the y intercept b of the line, as shown by manipulating the sliders on this interactive graph.

Definition 2.1 (Rate of change for a linear relationship) *For a straight line, we define the rate of change of y with respect to time t as the ratio:*

$$\frac{\textit{Change in } y}{\textit{Change in } t}.$$

We now make an important observation.

Observation ("Theorem"): The slope m of the straight line in Eqn. (2.2) corresponds to definition 2.1 of the rate of change of a linear relationship.

Hint: A "Theorem" is just a mathematical statement that can be established rigorously by an argument called a "proof". While we will not use such terminology often here, it is a staple of mathematics.

Proof: Taking **any** two points (t_1, y_1) and (t_2, y_2) on that line, and using the notation $\Delta y, \Delta t$ to represent the change in y and t we compute the ratio $\Delta y/\Delta t$and simplify algebraically to find:

$$\begin{aligned}\frac{\text{Change in } y}{\text{Change in } t} = \frac{\Delta y}{\Delta t} &= \frac{y_2 - y_1}{t_2 - t_1} \\ &= \frac{(mt_2 + b) - (mt_1 + b)}{t_2 - t_1} = \frac{mt_2 - mt_1}{t_2 - t_1} = m.\end{aligned}$$

Thus, the slope m corresponds exactly to the notion of change of y **per unit time** which we call henceforth the **rate of change of y with respect to time**. It is important to notice that this calculation leads to the same result *no matter which two points we pick on the graph of the straight line.* ◊

Featured Problem 2.1 (Velocity of growing microtubule tips) *Shown in Fig. 2.5 is a part of the kymograph from Fig. 2.1, but with a more "conventional" view of position y (in µm) on the vertical axis versus time t (in seconds) on the horizontal axis. The bright streaks are the tips of microtubules (MT) at various positions as they grow inside a cell.*

1. *Use the image and the superimposed grid to estimate the average velocities of microtubule tips (up to one significant digit). This is best done on the lower panel, where the position versus time graphs can be most easily viewed.*
2. *Based on the lines in the graph, explain whether the three microtubule tips shown are moving at similar or at quite different speeds.*
3. *Compare the normal and treated cells shown in two panels of Fig. 2.1, carefully noting the fact that the coordinate system differs from that of Fig. 2.5. How does the treatment affect the speed of microtubule tips?*
4. *Explain what could account for the apparent "zigzags" and "curves" (not straight lines) seen in both panels of Fig. 2.1.*

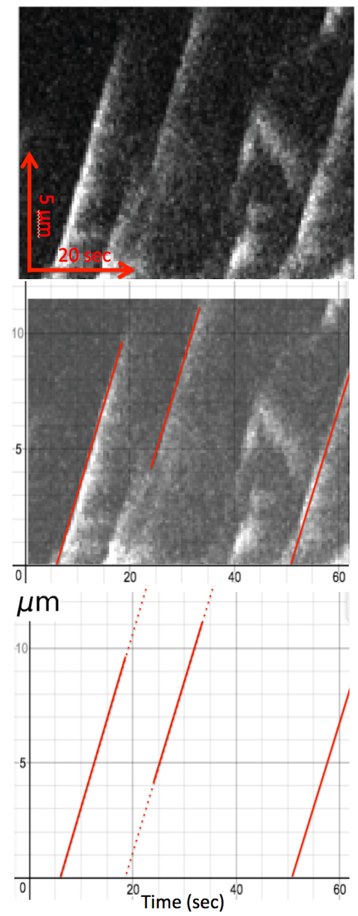


Figure 2.5: Part of the image from Fig. 2.5 is shown (top panel) on a yt (distance in μm vs time in seconds) coordinate system, with a grid superimposed. The red lines represent positions of proteins that track the tips of growing microtubules over time. In the lower panel, the same lines are shown on the grid alone; we can use this to estimate average velocity.

2.3 *The slope of a secant line is the average rate of change*

Section 2.3 Learning goals

1. Define of average rate of change; explain its connection with the slope of a secant line.
2. Compute the average rate of change using time-dependent data over a given time interval.
3. Given two points on the graph of a function, or two discrete data points, find both the slope and the equation of a secant line through those points.

We generalize the ideas in Section 2.2 to consider rates of change for relationships other than linear. Let $y = f(t)$ describe some relationship between time t and a variable of interest y. (This could be a set of discrete data points as in Figure 2.2, or a formula, as in Eqn. (2.1).)

Pick any two points $(a, f(a))$, and $(b, f(b))$ satisfying $y = f(t)$, and connect the points with a straight line. We refer to this line as the **secant line**, and we call its slope an average rate of change over the interval $a \le t \le b$. Formally, we define

Mastered Material Check

7. Sketch some nonlinear function and two different secant lines.
8. Is there any reason why we must draw secant lines between pairs of *successive* data points?

Definition 2.2 (Secant Line) *A secant line is a straight line connecting any two specific points on the graph of a function.*

Definition 2.3 (Average rate of change) *The average rate of change of $y = f(t)$ over the time interval $a \le t \le b$ is the slope of the secant line through the two points $(a, f(a))$, and $(b, f(b))$.*

Based on the above definition, we compute the average rate of change of f over the time interval $a \le t \le b$ as

$$\text{Average rate of change } = \frac{\text{Change in } f}{\text{Change in } t} = \frac{\Delta f}{\Delta t} = \frac{f(b) - f(a)}{b - a}.$$

Observe that the average rate of changes in general *depending on which two points we select*, in contrast to the linear case. (See left panel in Figure 2.6.) The word "average" sometimes causes confusion. One often speaks in a different context about the average value of a set of numbers (e.g. the average of $\{7, 1, 3, 5\}$ is $(7 + 1 + 3 + 5)/4 = 4$.) However the term *average rate of change* always means the slope of the straight line joining a pair of points.

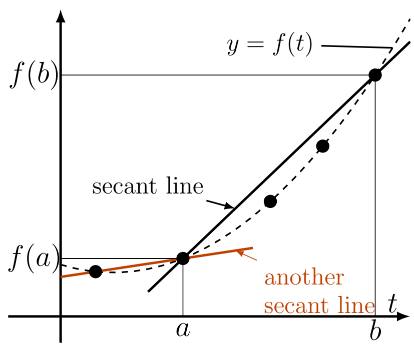


Figure 2.6: A set of time dependent data points (black circles) or smooth function (dashed curve) $f(t)$ showing a secant line through the points $(a, f(a))$, and $(b, f(b))$. Another secant line is drawn through $(a, f(a))$ and a different point to show the dependence on the points we select.

For example, the average rate of change of milk temperature discussed in Example 2.1 is

$$\frac{\text{Change in temperature}}{\text{Time taken}} = \frac{\Delta T}{\Delta t}.$$

Featured Problem 2.2 (Average rate of change of milk temperature) *1. Use the data in Tables 2.1 and 2.2 to show that the average rates of change of the temperature over the time interval $2 \le t \le 4$ min for the cooling phase is $-5.7°/$min. Repeat the calculation over a similar interval for the heating phase, and show that you get $26.95°/$min.*

2. *Write the* **equation of the secant line** *through the data points $t = 2$ and $t = 4$ for the heating phase.*

Definition 2.4 (Average velocity) *For a moving body, the average velocity over a time interval $a \le t \le b$ is the average rate of change of distance over the given time interval.*

Example 2.4 (Swimming velocity of Bluefin tuna) *Use the tuna swimming data in Figure 2.3 to answer the following questions:*

(a) Determine the average velocity of each of these two fish over the 35h shown in the figure.

(b) What is the fastest average velocity shown in this figure, and over what time interval and for which fish did it occur?

Solution.

(a) We find that Tuna 1 swam 180 km, whereas Tuna 2 swam 218 km over the course of 35 hr. Thus, the average velocity of Tuna 1 was $\bar{v} = 180/35 \approx 5.14$ km/h, whereas for Tuna 2 it was 6.23 km/h.

(b) The fastest average velocity corresponds to the segment of the graph that has the greatest slope. This occurs for Tuna 2 during the time interval $15 \leq t \leq 20$. Indeed, over that 5 hr interval the tuna has a **displacement** (net distance covered) of 140-78=62km. Its average velocity over that time interval was thus $62/5 = 12.4$km/h.

◊

Mastered Material Check

9. What was the average velocity of Tuna 1 during $15 \leq t \leq 20$?

Example 2.5 (Equation of secant line 2) *Find the equation of the secant line connecting the first and last data points for the Tuna 1 swimming distances in Figure 2.3.*

Solution. We defined the distance as 0 at time $t = 0$, so that the y intercept of the secant line is 0. We have already computed the slope of the secant line (average rate of change) as 5.14 km/h. Hence the equation of the secant line is

$$y_S = 5.14t.$$

◊

A secant line between two points, x_0 and $x_0 + h$ on the graph of a function $f(x)$ is shown in this link. You can change the base point x_0, the distance between the x coordinates, h, or you can input your own function for $f(x)$. The slope of the secant line is the average rate of change of f over the interval $x_0 \leq x \leq x_0 + h$

We can extend the definition of average rate of change to any function $f(x)$.

Definition 2.5 (Average rate of change of a function) *Suppose $y = f(x)$ is a function of some arbitrary variable x. The* average rate of change *of f between two points x_0 and $x_0 + h$ is given by*

$$\frac{\textit{Change in } y}{\textit{Change in } x} = \frac{\Delta y}{\Delta x} = \frac{[f(x_0+h) - f(x_0)]}{(x_0+h) - x_0} = \frac{[f(x_0+h) - f(x_0)]}{h}.$$

Here h is the difference of the x coordinates. The above ratio is the slope of the secant line shown in Figure 2.7.

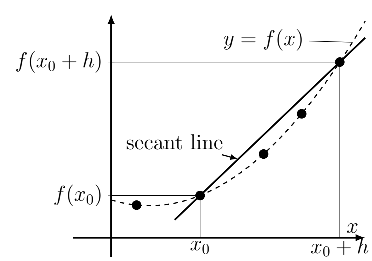


Figure 2.7: The graph of some arbitrary function $f(x)$ (dashed line) with a secant line through the points $(x_0, f(x_0))$ and $(x_0 + h, f(x_0 + h))$. The slope of the secant line is the average rate of change of f over the given interval.

Example 2.6 (Average velocity of a falling object) *Consider a falling object. Suppose that the total distance fallen at time t is given by Eqn. (2.1), $y(t) = ct^2$. Find the average velocity $\bar{v}$, of the object over the time interval $t_0 \leq t \leq t_0 + h$.*

Solution. In Figure 2.8, we reproduce the data for the falling object from Figure 2.4 and superimpose a secant line connecting two points labeled t_0 and $t_0 + h$.

We compute the average velocity as follows:

$$\begin{aligned}
\bar{v} &= \frac{y(t_0+h)-y(t_0)}{h} && \leftarrow \text{(definition of average velocity)} \\
&= \frac{c(t_0+h)^2 - c(t_0)^2}{h} && \leftarrow \text{the function of interest} \\
&= c\left(\frac{(t_0^2+2ht_0+h^2)-(t_0^2)}{h}\right) && \text{some algebra} \\
&= c\left(\frac{2ht_0+h^2}{h}\right) && \text{simplifying the expression} \\
&= c(2t_0+h). && (2.3)
\end{aligned}$$

The average velocity over the time interval $t_0 < t < t_0 + h$ is $\bar{v} = c(2t_0 + h)$. ◊

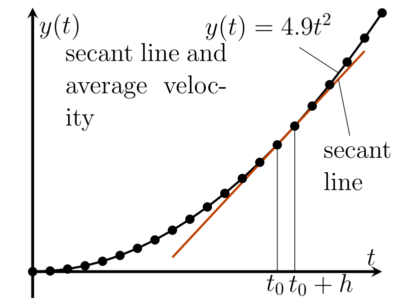


Figure 2.8: A secant line through two points on the graph of distance versus time for an object falling under the force of gravity.

2.4 *From average to instantaneous rate of change*

Section 2.4 Learning goals

1. Demonstrate that a data set with more frequent measurements corresponds to smaller time intervals Δt between data points.
2. Describe the connection between average rate of change over a very small time interval and instantaneous rate of change at a single point.

This section could also be titled "Shrinking the time-steps between measurements." So far, the average rates of change were computed over finite intervals. Our ultimate goal is to refine this idea and define a rate of change at each point, i.e. an **instantaneous rate of change**. But to do so, we first consider how a data set can be **refined** by making more frequent measurements to improve the notion of a rate of change *close to a given point*. We discuss two examples below.

Refined temperature data

We can refine the original data of temperature $T(t)$ for cooling milk from Figure 2.2 by taking more closely spaced time points. Table 2.4 provides a sample of the refined data.

The leftmost plot in Figure 2.9 shows the original data set with measurements every $\Delta t = 2$ min. The second and third plots have successively more refined measurements with shorter intervals between time points ($\Delta t = 1$ min and $\Delta t = 0.5$ min). After 10 minutes, fewer points were collected in each case.

Example 2.7 (Refined average rate of change) *Use the data in Table 2.4 to compute the average rate of change of the temperature over the time intervals*

time	Temp	time	Temp	time	Temp
0	190	0	190	0	190
2	176	1	182	0.5	185.5
4	164.6	2	176	1	182
6	155.4	3	169.5	1.5	179.2
8	148	4	164.6	2	176
10	140.9	5	159.8	2.5	172.9

Table 2.4: Partial data for temperature in °F for the three graphs shown in Figure 2.9. The pairs of columns indicate that the data has been collected at more and more frequent intervals $h = \Delta t$.

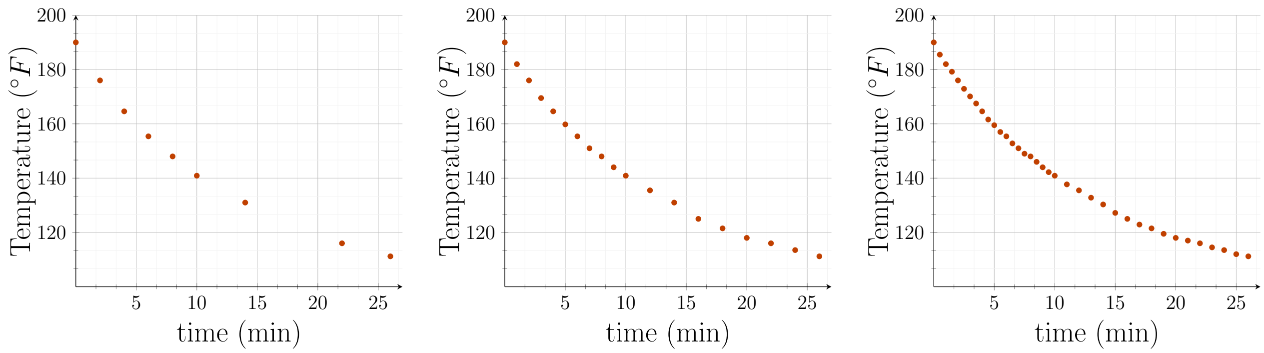


Figure 2.9: Three graphs of temperature versus time for cooling milk.

$2 \le t \le 2+h$ *where* $h = \Delta t = 2, 1, 0.5$ *min, respectively. Which calculation most accurately describes the behaviour "close to"* $t = 2$*min?*

Solution. Computing the ratio $\Delta T/\Delta t$, we obtain, for $\Delta t = 2, 1, 0.5$ the following average rates of change (in °F per min):

$$\Delta t = 2: \quad \frac{\Delta T}{\Delta t} = \frac{T(2+2) - T(2)}{4-2} = \frac{(164.6 - 176)}{(4-2)} = -5.7,$$

$$\Delta t = 1: \quad \frac{\Delta T}{\Delta t} = \frac{T(2+1) - T(2)}{3-2} = \frac{(169.5 - 176)}{(3-2)} = -6.5,$$

$$\Delta t = 0.5: \quad \frac{\Delta T}{\Delta t} = \frac{T(2+0.5) - T(2)}{2.5-2} = \frac{(172.9 - 176)}{(2.5-2)} = -6.2.$$

The last of these has been calculated over the smallest time interval, and most closely represents the rate of change of temperature close to the time $t = 2$ min. Exercise 2(b) leads to a similar comparison of this sort close to $t = 0$, and results in a similar set of finer values for the average rate of change "near" the initial data point. ◊

Mastered Material Check

10. At what time is the temperature approximately 120°F?

Refined data for the height of a falling object

We examine increasingly refined data for the height of a falling object in Figure 2.10.

Figure 2.10(a) shows three stroboscopic images, each giving successive vertical positions of an object falling from a height of 20 m over a 2 s time

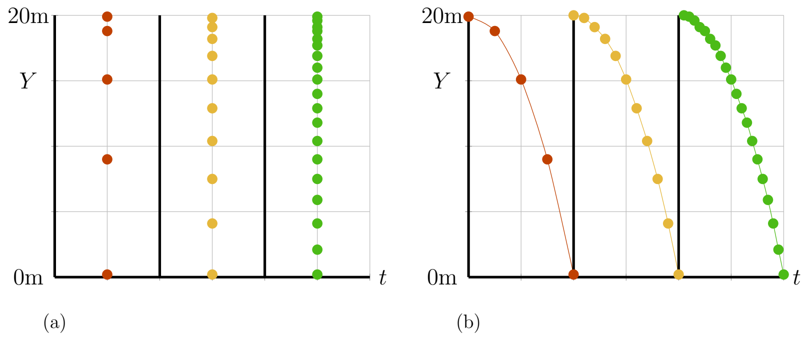


Figure 2.10: Data for the height Y of a falling object starting with $t = 0$ at the top, and following the object until $t = 2$ at its lowest point. The data set is refined ($\Delta t = 0.5, 0.2, 0.1$) to get more and more accurate tracking of the object. (a) Stroboscopic images. (b) Graphs of Y versus t.

period. The location of the ball is given first at intervals of $\Delta t = 0.5$ seconds, then at intervals of $\Delta t = 0.1$ and finally $\Delta t = 0.05$ s. In Figure 2.10(b), we graph the height $Y = Y_0 - ct^2$ against time t. (The distance fallen is still described by the function $y(t) = ct^2$, as in Example 2.3.)

By collecting data at finer time points, we can determine the "velocity" of the object with greater accuracy. Indeed, taking smaller and smaller time steps leads us to define **instantaneous velocity**.

Mastered Material Check

11. From what height was the object dropped in Figure 2.10?
12. If you wanted 50 equally spaced data points over a 2 s interval, what would Δt be?

Instantaneous velocity

A brief summary of average and instantaneous velocity in the example of a falling ball.

We know that the velocity over an interval can be calculated by finding the slope of a secant line connecting the endpoints of that interval. The slopes of the secant lines in Figure 2.10(b) are steeper at the end of the time interval than at the beginning - lending justification to what we intuitively know: a falling object's velocity increases as time passes.

With this in mind, to define an instantaneous velocity at some time t_0, we compute average velocities over decreasing time intervals $t_0 \leq t \leq t_0 + h$, allowing h to get smaller.

Note: we use the notation $h \to 0$ to denote the shrinking the time interval.

For example, we make the strobe flash faster so that $\Delta t = (t_0 + h) - t_0 = h \to 0$. At each stage, we calculate an average velocity, $\bar{v}$ for the interval $t_0 \leq t \leq t_0 + h$. As we continue to refine the measurements in this way, we arrive at a value for the velocity that we denote the **instantaneous velocity**. This number represents "the velocity of the ball at the very instant $t = t_0$".

Definition 2.6 (Instantaneous velocity) *The instantaneous velocity at time t_0, denoted $v(t_0)$ is defined as*

$$v(t_0) = \lim_{h \to 0} \bar{v}$$

where $\bar{v}$ is the average velocity over the time interval $t_0 \leq t \leq t_0 + h$. In other words,

$$v(t_0) = \lim_{h \to 0} \frac{y(t_0 + h) - y(t_0)}{h}.$$

We shall be more explicit about the meaning of the notation $\lim_{h\to 0}$ *in the next chapter.*

Example 2.8 (Computing an instantaneous velocity) *Use Gallileo's formula for the distance fallen, Eqn. (2.1), $y(t) = ct^2$, to compute the instantaneous velocity of a falling object at time t_0.*

Solution. We have already found the average velocity of the falling object over a time interval $t_0 \le t \le t_0 + h$ in Example 2.6, obtaining Eqn. (2.3),

$$\bar{v} = c(2t_0 + h).$$

Then, by Definition 2.6,

$$v(t_0) = \lim_{h\to 0} \bar{v} = \lim_{h\to 0} \ c(2t_0 + h) = 2ct_0.$$

Here we have used the fact that the expression $c(2t_0 + h)$ approaches $2ct_0$ as h shrinks to zero. This result holds for any time t_0. More generally, we could write that at time t, the instantaneous velocity is $v(t) = 2ct$. For example, for $c = 4.9\text{m/s}^2$, the velocity of an object at time $t = 1$ s after it is released is $v(1) = 9.8$ m/s. ♢

2.5 Introduction to the derivative

Section 2.5 Learning goals

1. Explain the first examples of calculation of the derivative.
2. Describe how the derivative is obtained from an average rate of change.
3. Compute the derivative of very simple functions such as $y = x^2$, and $y = Ax + B$.

We are ready for the the definition of the derivative.

Definition 2.7 (The derivative) *The derivative of a function $y = f(x)$ at a point x_0 is the same as the instantaneous rate of change of f at x_0. It is denoted $\frac{dy}{dx}\big|_{x_0}$ or $f'(x_0)$ and defined as*

$$\left.\frac{dy}{dx}\right|_{x_0} = f'(x_0) = \lim_{h\to 0} \frac{[f(x_0+h) - f(x_0)]}{h}.$$

As $h \to 0$, the secant line approaches a **tangent line**. Use the slider for h to show this trend, and note that the slope of the secant line (average velocity) approaches the slope of the tangent line (instantaneous velocity) at the point x_0.

We can use this to update our definition of instantaneous velocity:

Definition 2.8 (Velocity) *If $y = f(t)$ is the position of an object at time t then the derivative $f'(t)$ at time t_0 is the instantaneous velocity, also simply called the* **velocity** *of the object at that time.*

Example 2.9 (Formal calculation of velocity) *Use Gallileo's formula to set up and calculate the derivative of Eqn.* (2.1)*, and show that it corresponds to the instantaneous velocity obtained in Example 2.8.*

Solution. We set up the calculation using limit notation, recalling Gallileo's formula states $y(t) = ct^2$. We compute

$$\begin{aligned} v(t_0) &= \lim_{h\to 0} \frac{y(t_0+h)-y(t_0)}{h} \\ &= \lim_{h\to 0} \frac{c(t_0+h)^2 - c(t_0)^2}{h} \\ &= \lim_{h\to 0} c\left(\frac{(t_0^2+2ht_0+h^2)-(t_0^2)}{h}\right) \\ &= \lim_{h\to 0} c\left(\frac{2ht_0+h^2}{h}\right) = \lim_{h\to 0} c(2t_0+h) = 2ct_0. \end{aligned} \tag{2.4}$$

All steps but the last are similar to the calculation (and algebraic simplification) of average velocity (compare with Example 2.6). In the last step, we formally allow the time increment h to shrink, which is equivalent to taking $\lim_{h\to 0}$. ◊

Example 2.10 (Calculating the derivative of a function) *Compute the derivative of the function $f(x) = Cx^2$ at some point $x = x_0$.*

Solution. In the previous example, we calculated the derivative of the function $y = f(t) = ct^2$ with respect to t. Here we merely have a similar (quadratic) function of x. Thus, we have already solved this problem. By switching notation ($t_0 \to x_0$ and $c \to C$) we can write down the answer, $2Cx_0$ at once. However, as practice, we rewrite the steps in the case of the general point x

For $y = f(x) = Cx^2$ we have

$$\begin{aligned} \frac{dy}{dx} &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \\ &= \lim_{h\to 0} \frac{C(x+h)^2 - Cx^2}{h} \\ &= \lim_{h\to 0} C\frac{(x^2+2xh+h^2)-x^2}{h} \\ &= \lim_{h\to 0} C\frac{(2xh+h^2)}{h} = \lim_{h\to 0} C(2x+h) = 2Cx. \end{aligned}$$

Evaluating this result for $x = x_0$ we obtain the answer $2Cx_0$. ◊

We recognize from this definition that the derivative is obtained by starting with the slope of a secant line (average rate of change of f over the interval $x_0 \le x \le x_0 + h$) and proceeds to shrink the interval ($\lim_{h\to 0}$) so that it approaches a single point (x_0). In later chapters, the resultant line is called the **tangent line** and the value obtained identified as the **instantaneous rate of**

change of the function with respect to the variable x at the point of interest, x_0. We explore properties and meanings of this concept in the next chapter.

Note: we have used different notations to denote the derivative of $f(x) = y$. Further others exist. Each of the following may be used interchangeably:

$$f'(x), \quad \frac{df}{dx}, \quad \frac{d}{dx}f(x), \quad \frac{dy}{dx}, \quad y', \quad Df(x), \quad \text{and} \quad Dy.$$

These notations evolved for historical reasons and are used interchangeabley in science.

2.6 Summary

1. Graphs of time-dependent data are helpful to visualize trends such as increasing and decreasing values, steepness, linearity, and so on.
2. An average rate of change is the ratio of change in a dependent variable (y) over a range of the independent variable (x), often denoted $\Delta y/\Delta x$.
3. A secant line is a straight line through any two points on the graph of a function. The slope of a secant line is the average rate of change of the function over interval between the x coordinates of the two points.
4. The average rate of change of a function $y = f(x)$ on $x_0 \leq x \leq x_0 + h$ can be computed by the ratio

$$\frac{\Delta y}{\Delta x} = \frac{\text{Change in } y}{\text{Change in } x} = \frac{[f(x_0 + h) - f(x_0)]}{h}$$

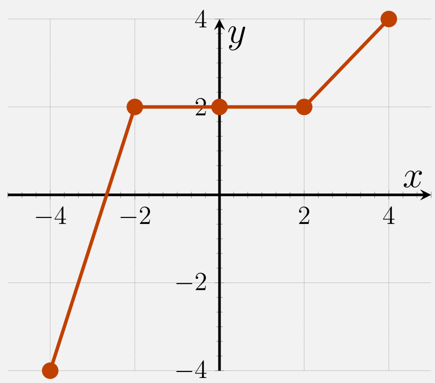
5. The instantaneous rate of change of a function $f(x)$ at x_0 can be found by taking the limit as $h \to 0$ of the average rate of change on $x_0 \leq x \leq x_0 + h$.
6. The derivative of a function $y = f(x)$ at a point x_0 is the same as the instantaneous rate of change of f at x_0.
7. In the case of time-dependent data, refining the data can lead to a better and better approximation of instantaneous rates of change.
8. This chapter explored datat related to the following time-dependent processes:

   (a) height of a falling object;
   (b) temperature of heating/cooling milk; and
   (c) swimming velocity of Bluefin tuna.

Quick Concept Checks

1. How do we calculate average rate of change of a time dependent process over a given interval?

2. Over what interval does the function depicted in the graph below have the greatest average rate of change? Smallest average rate of change?



3. Given the function defined by $\{(1,3),(2,5),(3,7)\}$, how many **different** secant lines can be formed?

4. Use the definition to calculate the derivative of $f(x) = 4x^2 + 3$ at $x_0 = 1$.

Exercises

2.1. **Heating milk.** Consider the data gathered for heating milk in Table 2.1and Figure 2.2 (left)

(a) Estimate the slope and intercept of the straight line shown in the figure and use to write the equation of this line. According to this approximate straight line relationship, what is the average rate of change of the temperature over the 5 min interval shown?

(b) Find a pair of points such that the average rate of change of the temperature is *smaller* than your result in part (a).

(c) Find a pair of points such that the average rate of change of the temperature is *greater* than your result in part (a).

(d) Milk boils at 212°F, and the recipe for yoghurt calls for avoiding a temperature this high. Use your common knowledge to explain why the data for heating milk is not actually linear.

2.2. **Refining the data.** Table 2.4 shows some of the data for cooling milk that was collected and plotted in Figure 2.9. Answer the following questions.

(a) Use the table to determine the average rate of change of the temperature over the first 10 min.

(b) Compute the average rate of change of the temperature over the intervals $0 \leq t \leq 2$, $0 \leq t \leq 1$ and $0 \leq t \leq 0.5$.

(c) Which of your results in (b) would be closest to the “instantaneous" rate of change of the temperature at $t = 0$?

2.3. **Height and distance dropped.** We have defined the variable $Y(t) =$ height of the object at time t and the variable $y(t)$ as the distance dropped by time t.

(a) State the connection between these two variables for a ball whose initial height is Y_0.

(b) How is the displacement over some time interval $a \leq t \leq b$ related between these two ways of describing the motion? (Assume that the ball is in the air throughout this time interval).

2.4. **Falling ball.** A ball is dropped from height $Y_0 = 490$ meters above the ground. Its height, Y, at time t is known to follow the relationship $Y(t) = Y_0 - \frac{1}{2}gt^2$ where $g = 9.8$ m /s^2.

(a) Find the average velocity of the falling ball between $t = 1$ and $t = 2$ seconds.

(b) Find the average velocity between t sec and $t + \varepsilon$ sec where $0 \leq \varepsilon \leq 1$ is some small time increment (assume that the ball is in the air during this time interval).

(c) Determine the time at which the ball hits the ground.

2.5. **Tuna average velocity.** Find the average velocity of Tuna 1 over each of the time intervals shown in Table 2.3, that is for $0 \leq t \leq 5$ hr, $5 \leq t \leq 10$ hr, etc.

2.6. **Average velocity and secant line.** The two points on Figure 2.8 through which the secant line is drawn are $(1.3, 8.2810)$ and $(1.4, 9.6040)$. Find the average velocity over this time interval and then give the equation of the secant line.

2.7. **Human Population Growth.** Table 2.5 gives data for the human population (in billions) over recorded history (with some estimates where data was not available).

Note: human population growth is further studied in Chapter 11.

(a) Plot the human population (in billions) versus time (in years) using graphing software of your choice.

(b) Determine the average rate of change of the human population over the successive time intervals.

(c) Plot the average rate of change versus time (in years) and determine over what time interval that average rate of change was greatest.

(d) Over what period (i.e. time interval) was this average rate of change *increasing most rapidly*? (*hint:* you should be able to answer this question either by looking at the graph you have drawn or by calculation)

year	**human population** (billions)
1	0.2
1000	0.275
1500	0.45
1650	0.5
1750	0.7
1804	1
1850	1.2
1900	1.6
1927	2
1950	2.55
1960	3
1980	4.5
1987	5
1999	6
2011	7
2020	7.7

Table 2.5: The human population (billions) over the years AD 1 to AD 2020.

2.8. **Average velocity at time t.** A ball is thrown from the top of a building of height Y_0. The height of the ball at time t is given by

$$Y(t) = Y_0 + v_0 t - \frac{1}{2} g t^2$$

where h_0, v_0, g are positive constants. Find the average velocity of the ball for the time interval $0 \leq t \leq 1$ assuming that it is in the air during this whole time interval. Express your answer in terms of the constants given.

2.9. **Average rate of change.** A certain function takes values given in Table 2.6. Find the average rate of change of the function over the intervals.

(a) $0 \leq t \leq 0.5$,

(b) $0 \leq t \leq 1.0$,

(c) $0.5 \leq t \leq 1.5$,

(d) $1.0 \leq t \leq 2.0$.

t	0	0.5	1.0	1.5	2.0
$f(t)$	0	1	0	-1	0

Table 2.6: Function values for Exercise 9.

2.10. **Average rate of change.** Find the average rate of change for each of the following functions over the given interval.

(a) $y = f(x) = 3x - 2$ from $x = 3.3$ to $x = 3.5$.

(b) $y = f(x) = x^2 + 4x$ over $[0.7, 0.85]$.

(c) $y = -\frac{4}{x}$ and x changes from 0.75 to 0.5.

2.11. **Trig mini-review.** Consider the table of values of the trigonometric functions $\sin(x)$ and $\cos(x)$ found in Table 2.7.

For the following, express your answer in terms of square roots and π. Do not compute decimal expressions.

(a) Find the average rate of change of $\sin(x)$ over $0 \le x \le \pi/4$.

(b) Find the average rate of change of $\cos(x)$ over $\pi/4 \le x \le \pi/3$.

(c) Is there an interval over which the functions $\sin(x)$ and $\cos(x)$ have the same average rate of change? (*hint:* consider the graphs of these functions over one whole cycle, e.g. for $0 \le x \le 2\pi$. Where do they intersect?)

Note: trigonometry is reviewed in Appendix F and studied further in Chapters 14 and 15.

x	$\sin(x)$	$\cos(x)$
0	0	1
$\frac{\pi}{6}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$
$\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$
$\frac{\pi}{3}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$
$\frac{\pi}{2}$	1	0

Table 2.7: Table of sine and cosine values

2.12. **Secant and tangent lines.** Let $y = f(x) = 1 + x^2$ and consider the point $(1,2)$ on its graph and some point nearby, for example

$$(1+h, 1+(1+h)^2).$$

(a) Find the slope of a secant line connecting these two points.

(b) The slope of a tangent line to $y = f(x)$ is the derivative $f'(x)$. Use the slope you calculated in (a) to determine what the slope of the tangent line to the curve at $(1,2)$ would be.

(c) Find the equation of the tangent line through the point $(1,2)$.

2.13. **Secant and tangent lines.** Given the function $y = f(x) = 2x^3 + x^2 - 4$,

(a) find the slope of the secant line joining the points $(4, f(4))$ and $(4 + h, f(4+h))$ on its graph, where h is a small positive number, then

(b) find the slope of the tangent line to the curve at $(4, f(4))$.

2.14. **Average rate of change.** Consider the function $f(x) = x^2 - 4x$ and the point $x_0 = 1$.

(a) Sketch the graph of the function.

(b) Find the average rate of change over the intervals $[1,3]$, $[-1,1]$, $[1,1.1]$, $[0.9,1]$ and $[1-h,1]$, where h is some small positive number.

(c) Find $f'(1)$.

2.15. **Approximation using a tangent line.** Let $y = f(x) = x^2 - 2x + 3$.

(a) Find the average rate of change over the interval $[2, 2+h]$.

(b) Find $f'(2)$.

(c) Using only the information from (a), (b) and $f(2) = 3$, approximate the value of y when $x = 1.99$, without substituting $x = 1.99$ into $f(x)$.

2.16. **Average rate of change.** For the following, express your answer in terms of square roots and π. Do not compute the decimal expressions. Note that

$$\tan(x) = \frac{\sin(x)}{\cos(x)}, \quad \cot(x) = \frac{\cos(x)}{\sin(x)}$$

(a) Find the average rate of change of $\tan(x)$ over $0 \leq x \leq \frac{\pi}{4}$.

(b) Find the average rate of change of $\cot(x)$ over $\frac{\pi}{4} \leq x \leq \frac{\pi}{3}$.

2.17. **Secant and tangent lines.**

(a) Find the slope of the secant line to the graph of $y = 2/x$ between the points $x = 1$ and $x = 2$.

(b) Find the average rate of change of y between $x = 1$ and $x = 1 + \varepsilon$ where $\varepsilon > 0$ is some positive constant.

(c) What happens to this slope as $\varepsilon \to 0$?

(d) Find the equation of the tangent line to the curve $y = 2/x$ at the point $x = 1$.

2.18. **Velocity and average velocity.** For each of the following motions where s is measured in meters and t is measured in seconds, find the velocity at time $t = 2$ and the average velocity over the given interval.

(a) $s = 3t^2 + 5$ and t changes from 2 to $3s$.

(b) $s = t^3 - 3t^2$ from $t = 3s$ to $t = 5s$.

(c) $s = 2t^2 + 5t - 3$ on $[1, 2]$.

2.19. **Acceleration.** The velocity v of an object attached to a spring is given by $v = -A\omega \sin(\omega t + \delta)$, where A, ω and δ are constants. Find the average change in velocity ("acceleration") of the object for the time interval $0 \leq t \leq \frac{2\pi}{\omega}$.

Note: acceleration is further explored in Chapter 4.

2.20. **Definition of the derivative.** Use the definition of derivative to calculate the derivative of the function

$$f(x) = \frac{1}{x+1}.$$

Note: intermediate steps are required.

3

Three faces of the derivative: geometric, analytic, and computational

In Chapter 2 we bridged two concepts: the **average rate of change** (slope of secant line) and the **instantaneous rate of change** (the derivative). We arrived at a technique for calculating a derivative algebraically. As a result, we introduced limits - a concept that merits further discussion. One goal of this chapter is to consider the technical aspects of limits - a requirement if we are to use the definition of the derivative to determine derivatives of common functions.

But first, we consider a distinct approach which is *geometric* in flavour. Namely, we show that the local behaviour of a continuous function is described by a tangent line at a point on its graph: we can visualize the tangent line by zooming into the graph of the function. This duality - the geometric (graphical) and analytic (algebraic calculation) views - form themes throughout the discussions to follow. They are two complementary, but closely related approaches to calculus.

3.1 The geometric view: zooming into the graph of a function

Section 3.1 Learning goals

1. Describe the link between the local behaviour of a function (seen by zooming into the graph) at a point and the tangent line to the graph of the function at that point.
2. Sketch a function's derivative given its graph.

Locally, the graph of a function looks like a straight line

In this section we consider well-behaved functions whose graphs are "smooth", as opposed to the discrete data points of Chapter 2. We link the derivative to the local shape of the graph of the function. By **local** behaviour we mean the shape we see when we zoom into a point on the graph. Imagine using a microscope where the center of the field of vision is some point of interest. As we zoom in, the graph looks flatter, until we observe a straight line, as shown in Figure 3.1.

Definition 3.1 (Tangent line) *The straight line that we see when we zoom into the graph of a smooth function at some point x_0 is called the* **tangent line** *at x_0.*

Definition 3.2 (Geometric definition of the derivative) *The* **slope of the tangent line** *at the point x is denoted as the* **derivative** *of the function at the given point.*

Mastered Material Check

1. Write down the equation of a generic straight line.
2. Identify the slope of that straight line.
3. How many zeros does the function $f(x) = x^3 - x$ have?

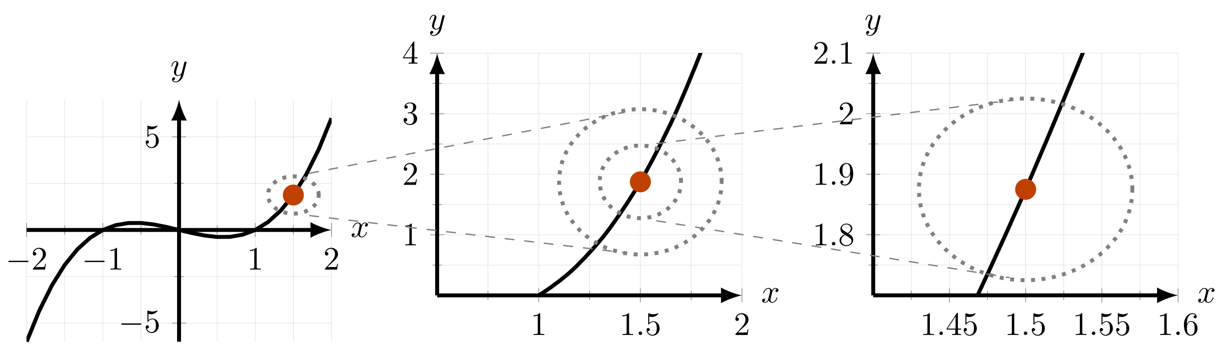


Figure 3.1: Zooming in on the graph of $y = f(x) = x^3 - x$ at the point $x = 1.5$ makes the graph "look like" a straight line - the tangent line. The slope of that tangent line is the derivative of the function at $x = 1.5$.

Example 3.1 (Zooming into a polynomial) *Consider the function shown in Figure 3.1, $y = f(x) = x^3 - x$, and the point $x = 1.5$. Find the tangent line to the graph of this function by zooming in at the given point.*

Solution. The graph of the function is shown in each panel of Figure 3.1, where we have indicated the point of interest with a red dot. Now zooming in on the given point. Locally, the graph resembles a straight line. This is the tangent line to $f(x)$ at $x = 1.5$. ◊

Example 3.2 (Zooming into the sine graph at the origin) *Determine the derivative of the function $y = \sin(x)$ at $x = 0$ by zooming into the origin on its graph. Write down the equation of the tangent line at that point.*

Solution. In Figure 3.2 we zoom into $x = 0$ on the graph of the function

$$y = \sin(x).$$

The sequence of zooms leads to a straight line (far right panel) that we identify once more as the tangent line to the function at $x = 0$. From the graph, the slope of this tangent line is 1. We say that the derivative of the function $y = f(x) = \sin(x)$ at $x = 0$ is 1, and write $f'(0) = 1$ to denote this. As this line goes through $(0,0)$ and has slope 1, its equation is $y = x$. We can also say that close to $x = 0$ the graph of $y = \sin(x)$ looks a lot like the line $y = x$. ◊

Zooming in to the graph of a function.
(1) Click on the + ("zoom in") button on this graph of $\sin(x)$ to see how, close to the point $(0,0)$, the graph looks like a straight line of slope 1.
(2) Now change the function to $f(x) = |x|$ and note that when zooming into the cusp at $x = 0$, we do not see a (single) tangent line. In that case we say that the derivative does not exist at the given point.

Mastered Material Check

4. Using the far right panel of Figure 3.2, perform the calculations that verify the $y = \sin(x)$ looks a lot like $y = x$ near $x = 0$.
5. Give an example of a function with a discontinuity.

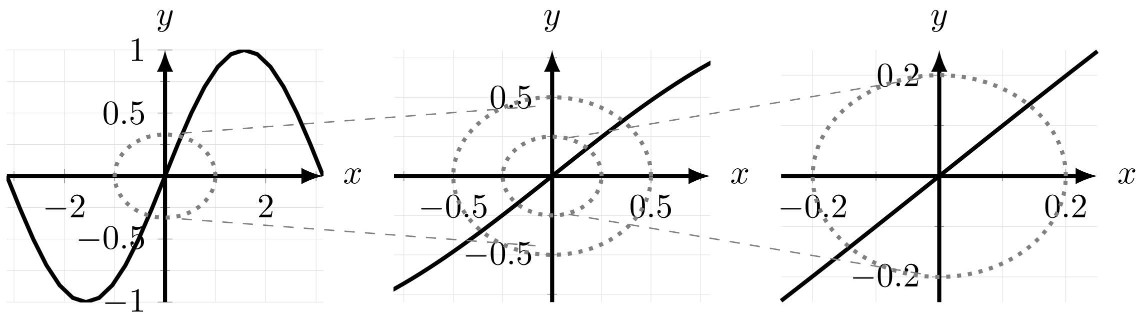


Figure 3.2: Zooming into the graph of the function $y = f(x) = \sin(x)$ at the point $x = 0$. Eventually, the graph resembles a line of slope 1. This is the tangent line at $x = 0$ and its slope, the derivative of $y = \sin(x)$ at $x = 0$ is 1.

At a cusp or a discontinuity, the derivative is not defined

If we zoom into a function at a cusp or a discontinuity, there is no single straight line that describes the local behaviour. For example, in Figure 3.3, we see two distinct lines meeting at a sharp "corner".

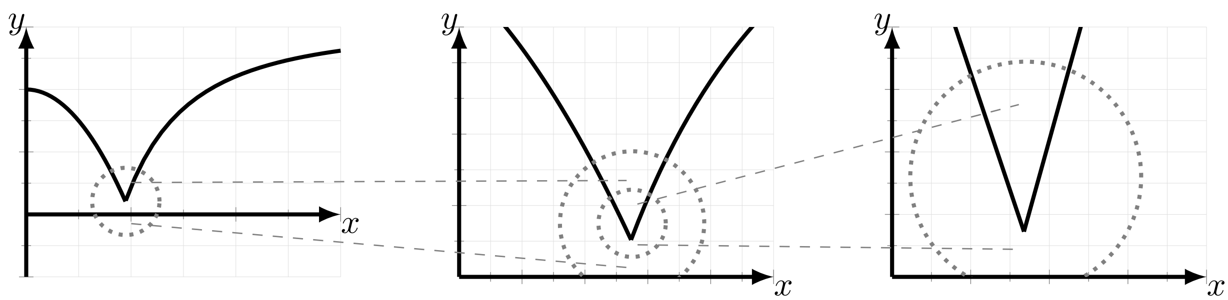


Figure 3.3: A function has no tangent line at a a cusp and the derivative is not defined at that point.

Zooming into a discontinuity presents another problem: there is no line at all, as in Figure 3.4. Finally, a function like $1/x$ has a singularity at $x = 0$ which shows up as a vertical line whose slope is infinite. In all such cases, we say that the function has no tangent line its derivative is not defined at the cusp, discontinuity, or singularity point.

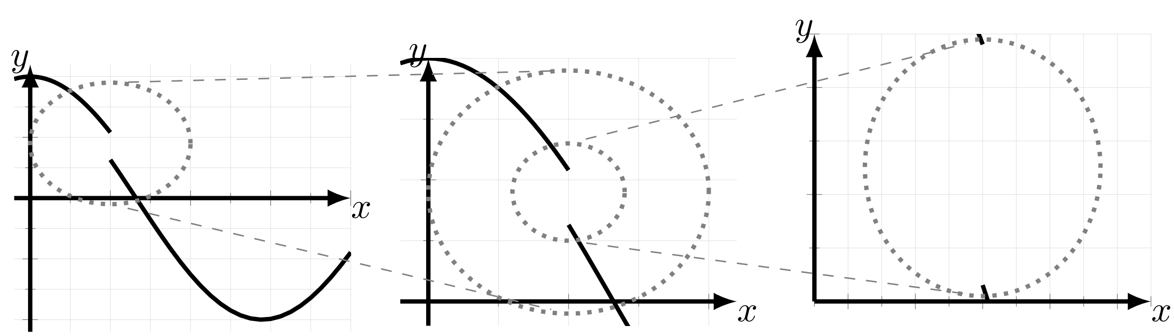


Figure 3.4: Zooming into a function at a discontinuity shows no single distinct line that might be a tangent line.

Mastered Material Check

6. Sketch the "zooming in" graph of the function $y = f(x) = \sin(x)$ at $x = 1$.
7. How many local minima are depicted in Figure 3.5?

From a function to a sketch of its derivative

The tangent line to the graph of a function varies from point to point along the graph of the function - what we see when zooming in depends on the location of the zoom. This means that *the derivative* $f'(x)$ *is, itself, also a function*. Here we consider the connection between these two functions by using the graph of one to sketch the graph of other. The hand-sketch is approximate, but preserves important elements.

Example 3.3 *Consider the function in Figure 3.5. Reason about the tangent lines at various points along to sketch the derivative* $f'(x)$.

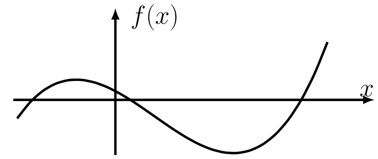


Figure 3.5: The graph of a function. We sketch its derivative.

Solution. In Figure 3.6 we first sketch a few tangent lines along the graph of $f(x)$. Focus on the slopes (rather than height, length, or other properties) of

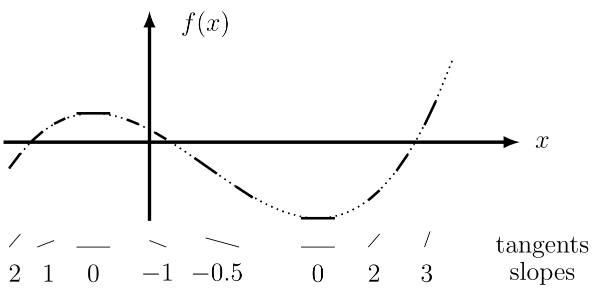


Figure 3.6: A few tangent lines of a function.

the dashes. Copying these lines in a row below the graph, we estimate their slopes roughly (approximate numerical values shown).

By convention, we "read" a graph from left to right. Slopes in Fig. 3.6 are first positive, then zero, then negative, increase again through zero, and then positive. There are two locations with zero slope (horizontal dashes). Next, in Figure 3.7, we these rough values for slopes. Only a few points have been plotted for $f'(x)$, but these trends are clear: the derivative function has two **zeros**, and it dips below the axis between these places. In Figure 3.7 we emphasize how the original function lines up with its derivative $f'(x)$.

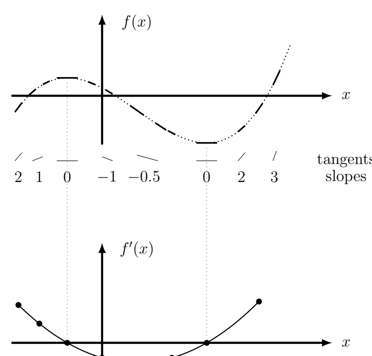


Figure 3.7: Sketching the derivative of a function.

We have aligned these graphs so that the slope of $f(x)$ matches the value of $f'(x)$ shown directly below. ◊

Example 3.4 *Sketch the derivative of the function shown in Figure 3.8.*

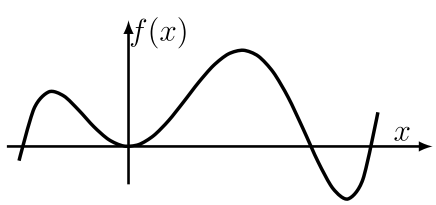


Figure 3.8: In Example 3.4, we sketch the derivative of this function.

Solution. We can get a good rough sketch by simply noting where the slopes are positive, negative, or zero. See Figure 3.9 for the entire process. The thin vertical lines demonstrate that $f'(x) = 0$ coincides with tops of hills or bottom of valleys on the graph of the function $f(x)$.

◊

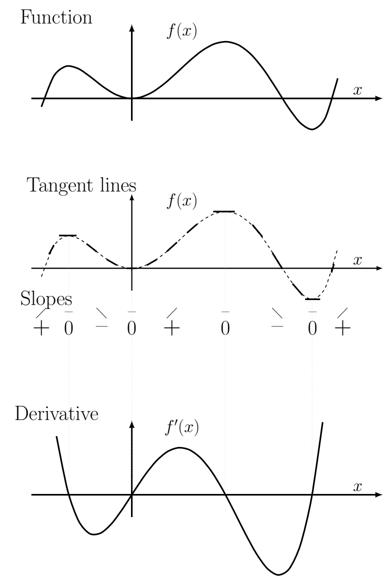


Figure 3.9: Sketching the derivative of a function for Example 3.4

Constant and linear functions and their derivatives

Example 3.5 (Derivative of $y = C$) *Use a geometric argument to determine the derivative of the function $y = f(x) = C$ at any point x_0 on its graph.*

Solution. This function is a horizontal straight line, whose slope is zero everywhere. Thus "zooming in" at any point x, leads to the same result, so the derivative is 0 everywhere. ◊

Example 3.6 (Derivative of $y = Bx$) *Use a geometric argument to determine the derivative of the function $y = f(x) = Bx$ at any point x_0 on its graph.*

Solution. The function $y = Bx$ is a straight line of slope B. At any point on its graph, it has the same slope, B. Thus the derivative is equal to B at any point on the graph of this function. ◊

Notice that in the above two examples we have found the derivative for the two power functions, $y = x^0$ and $y = x^1$. We summarize:

Mastered Material Check

8. If the equation of the tangent line to $f(x)$ at $x = 1$ is $y = mx + b$, what is $f'(1)$?
9. Sketch the graph of the derivative of the function $y = C$. (What slope do you see at each point when you "zoom in"?

The derivative of any constant function is zero. The derivative of the function $y = x$ is 1. The derivative of the function $y = k \cdot x$ is k.

Molecular motors

Microtubules (MT) are long, rod-like cellular structures (introduced in Chapter 2) with both structural and transportation roles in living cells. Human nerve cells can be up to 1 meter in length, which makes for a challenge to move material from the cell body - where it is made - to the cell ends where it is needed for repair or metabolism. Microtubules act like highways for "molecular motors", proteins that "drive" along these routes, transporting the necessary cargo.

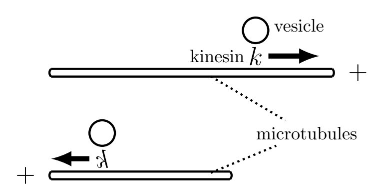


Figure 3.10: The molecular motor kinesin walks towards a microtubule "plus" end. It can detach and reattach to another microtubule.

Microtubules have distinct ends (called "plus" and "minus" ends). Some motors specialize in moving towards the + end, while others move towards the - end. Figure 3.10 is a schematic diagram of **kinesin** (represented by the letter k), a plus-end directed motor. As shown in the figure, kinesin can hop off one MT and onto another MT pointing in the opposite direction.

In Example 3.7, we study a sample vesicle track (displacement, y over time t) and decipher the sequence of motor events that caused that motion.

Example 3.7 (Motion of molecular motors) *The displacement $y(t)$ of a vesicle is shown in Figure 3.10(a).*

(a) Sketch the corresponding instantaneous velocity $v(t)$ for the vesicle.

(b) Use your sketch to explain what was happening to the kinesin carrying that vesicle.

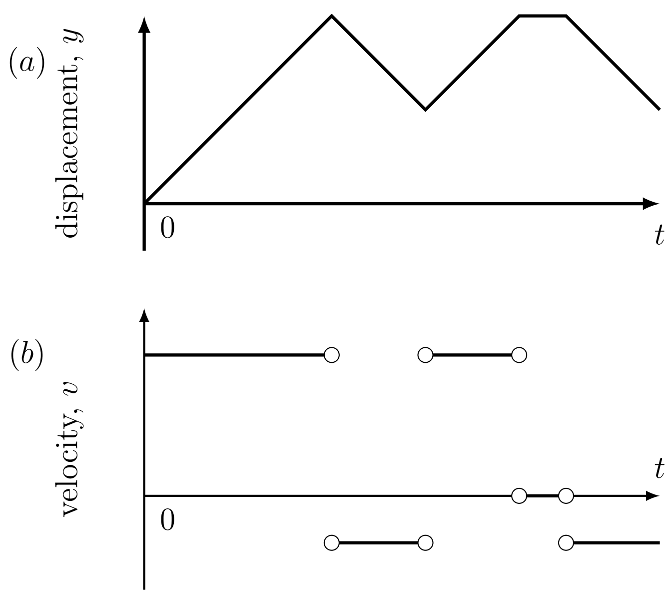


Figure 3.11: Molecular motors: (a) The displacement of the vesicle. (b) The velocity of the vesicle.

Solution.

(a) The plot in Figure 3.11(a) consists of straight line segments with sharp corners (cusps). Over each of these line segments, the slope dy/dt (which corresponds to the instantaneous velocity, $v(t)$) is constant. Segments with positive slope correspond to motion towards the right (as in the top microtubule track in Figure 3.10. Over times where the slope is negative, the motion is to the left. Where the slope is zero (flat graph), the vesicle was stationary.

In Figure 3.11(b), we sketch the graph of the instantaneous velocity, $v(t)$. Observe that $v(t)$, which is the derivative of $y(t)$, is not defined at the points where $y(t)$ has "corners".

(b) Based on Figure 3.11(b), the kinesin motor was moving on a right-facing microtubule, then hopped onto a left-facing microtubule, and then hopped back to a right -facing microtubule. For a brief time it was either stuck or detached from the microtubule tracks (stationary part). Finally, it hopped onto a left-moving microtubule.

◊

3.2 *The analytic view: calculating the derivative*

Section 3.2 Learning goals

1. Explain the definition of a continuous function.
2. Identify functions with various types of discontinuities.
3. Evaluate simple limits of rational functions.
4. Calculate the derivative of a simple function using the definition of the derivative, 2.7.

Technical matters: continuous functions and limits

In Chapter 2, we saw functions made up of discrete data points. In Chapter 1 we studied several continuous functions: power, polynomial or rational. Intuitively, on the graph of a continuous function, every point is connected to neighbouring points. For example, the power function Eqn. (2.1), $y(t) = ct^2$ is continuous for all values of t, whereas a function such as $y = \sqrt{x}$ is continuous and defined as a real number only for $x \geq 0$. The function $y = 1/(x+1)$ is defined and continuous for $x \neq -1$ (since division by zero is undefined).

We now make the intuitive discussion more precise with a formal definition, based on the concept of a limit. We first define what it means for a function to be continuous, and then show how limits are computed to test that definition.

Mastered Material Check

10. Use your favorite graphing software to verify these statements about the continuity of $y = \sqrt{x}$ and $y = 1/(x+1)$.

Definition 3.3 (Continuous function) *We say that $y = f(x)$ is continuous at a point $x = a$ in its domain if*

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

By this we mean that the function is defined at $x = a$, that the above limit exists, and that it matches with the value that the function takes at the given point.

This definition has two important parts. First, the function should be defined at the point of interest, and second, the value assigned by the function has to "fit the local behaviour" in the sense of the limit. This rules out a "jump" or "break" in the graph. When the above is not true at some point x_s, we say that the function is **discontinuous** at x_s. We give a few examples to demonstrate some different types of discontinuities that exist. At the same time we illustrate how limits are calculated.

Function with a hole in its graph. Consider a function of the form

$$f(x) = \frac{(x-a)^2}{(x-a)}.$$

Then if $x \neq a$, we can cancel a common factor, and obtain $(x-a)$. At $x=a$, the function is not defined because $\left(\frac{0}{0}\right)$ - which is undefined - results. In short, we have

$$f(x) = \frac{(x-a)^2}{(x-a)} = \begin{cases} x-a & x \neq a \\ \text{undefined} & x = a. \end{cases}$$

Even though the function is not defined at $x=a$, we can still evaluate the limit of f as x approaches a. We write

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

and say that "the limit as x approaches a" **exists** and is equal to 0. We also say that the function has a **removable discontinuity**. If we add the point $(a,0)$ to the set of points at which the function is defined then we obtain a continuous function identical to the function $x-a$. See also Appendix D.

Mastered Material Check

11. Sketch the graph of the function $f(x) = \frac{(x-1)^2}{(x-1)}$.

Function with jump discontinuity. Consider the function

$$f(x) = \begin{cases} -1 & x \leq a, \\ 1 & x > a. \end{cases}$$

We say that the function has a **jump discontinuity** at $x=a$. As we approach the point of discontinuity we observe that the function has two distinct values, depending on the direction of approach. We formally capture this observation using **right and left hand** limits,

$$\lim_{x\to a-} f(x) = -1, \quad \lim_{x\to a+} f(x) = 1.$$

Notice we use $\lim_{x\to a-}$ to denote approaching a from the left, and $\lim_{x\to a+}$ to denote approaching a from the right. Since the left and right limits are unequal, we say that "the limit **does not exist**" (abbreviated DNE).

Function with blow up discontinuity. Consider the function

$$f(x) = \frac{1}{x-a}.$$

Then as x approaches a, the denominator approaches 0, and the value of the function goes to $\pm\infty$. We say that the function "**blows up**" at $x=a$ and that the limit, $\lim_{x\to a} f(x)$, does not exist.

Figure 3.12 illustrates the differences between functions that are continuous everywhere, those that have a hole in their graph, and those that have a jump discontinuity or a blow up at some point a.

Mastered Material Check

12. What type of discontinuity does $\frac{x^2+4x+4}{x+2}$ have?
13. What type of discontinuity does $\frac{x^2+4x+4}{x+4}$ have?

Examples of limits. We now examine several examples of computations of limits. More details about properties of limits are provided in Appendix D.

By Definition 3.3, to calculate the limit of any function at a point of continuity, we simply **evaluate** the function at the given point.

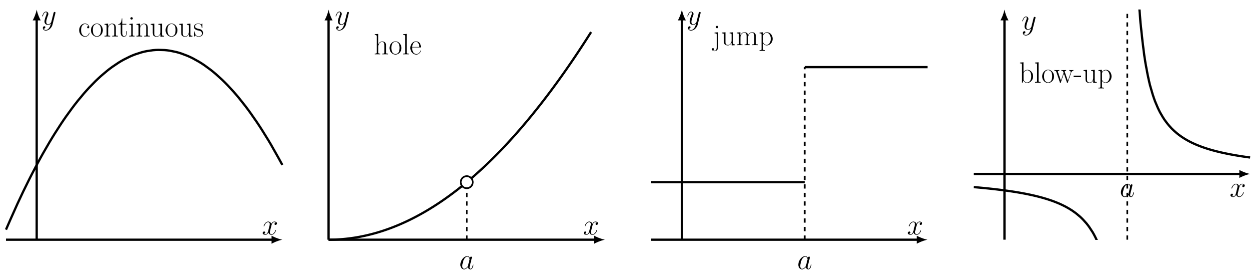


Figure 3.12: Left to right: a continuous function, a function with a removable discontinuity, a jump discontinuity, and a function with blow up discontinuity.

Example 3.8 (Simple limit of a continuous function) *Find the following limits:*

(a) $\lim_{x\to 3} x^2+2$ *(b)* $\lim_{x\to 1} \frac{1}{x+1}$ *(c)* $\lim_{x\to 10} \frac{x}{1+x}$.

Solution. In each case, the function is continuous at the point of interest (at $x = 3, 1, 10$, respectively). Thus, we simply "plug in" the values of x in each case to obtain

$$\text{(a)} \lim_{x\to 3} x^2+2 = 3^2+2 = 11 \quad \text{(b)} \lim_{x\to 1} \frac{1}{x+1} = \frac{1}{2} \quad \text{(c)} \lim_{x\to 10} \frac{x}{1+x} = \frac{10}{11}.$$

◊

Example 3.9 (Hole in graph limits) *Calculate the limits of the following functions. Note that each has a removable discontinuity ("a hole in its graph").*

(a) $\lim_{x\to 3} \frac{x^2-6x+9}{x-3}$, *(b)* $\lim_{x\to -1} \frac{x^2+3x+2}{x+1}$.

Solution. We first simplify algebraically by factoring the numerator, and then evaluate the limit. *Note that the simplification is possible so long as we evaluate the limit, rather than the actual function*, at the point of discontinuity.

$$\text{(a)} \lim_{x\to 3} \frac{x^2-6x+9}{x-3} = \lim_{x\to 3} \frac{(x-3)^2}{(x-3)} = \lim_{x\to 3}(x-3) = 0.$$

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

◊

Mastered Material Check

14. How do these examples change if the limit approaches a different value of x?

Example 3.10 (Limit involving $\sin(x)$**)** *Use the observation made in Example 3.2 to arrive at the value of* $\lim_{x\to 0} \frac{\sin(x)}{x}$.

Solution. Example 3.2 illustrated the fact that close to $x = 0$ the function $\sin(x)$ has the following behaviour:

$$\sin(x) \approx x, \quad \text{or} \quad \frac{\sin(x)}{x} \approx 1.$$

This is equivalent to the result

$$\lim_{x\to 0}\frac{\sin(x)}{x}=1. \tag{3.1}$$

We read this "as x approaches zero, the limit of $\sin(x)/x$ is 1." This limit is used in later calculations involving derivatives of trigonometric functions. ◊

Computing the derivative

As discussed in Section 2.5, calculating a derivative requires the use of limits. To summarize:

1. For the secant line connecting the points x and $x+h$ on the graph of a function, in the limit $h\to 0$, those points get closer together, leading to a tangent line.

2. The slope of a secant line is an average rate of change, but in the limit ($h\to 0$), we obtain the derivative, which is the slope of the tangent line.

We illustrate how to use the definition of the derivative to compute a few derivatives.

Example 3.11 (Derivative of a linear function) *Using Definition 2.7 of the derivative, compute the derivative of the function $y=f(x)=Bx+C$.*

Solution. We have already used a geometric approach to find the derivative of related functions in Examples 3.5-3.6. Here we do the formal calculation as follows:

$$\begin{aligned}
f'(x) &= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} && \text{(start with the definition)}\\
&= \lim_{h\to 0}\frac{[B(x+h)+C]-[Bx+C]}{h} && \text{(apply it to the function)}\\
&= \lim_{h\to 0}\frac{Bx+Bh+C-Bx-C}{h} && \text{(expand the numerator)}\\
&= \lim_{h\to 0}\frac{Bh}{h} && \text{(simplify)}\\
&= \lim_{h\to 0}B && \text{(cancel a factor of } h\text{)}\\
&= B && \text{(evaluate the limit).} \qquad (3.2)
\end{aligned}$$

Mastered Material Check

15. What other notations express the derivative of $f(x)$?

Hence, we confirmed that the derivative of $f(x)=Bx+C$ is $f'(x)=B$. This agrees with the sum of the derivatives of the two parts, Bx and C found in Examples 3.5-3.6. (Indeed, as we establish shortly, the derivative of the sum of two functions is the same as the sum of their derivatives.) ◊

Example 3.12 (Derivative of the cubic power function) *Compute the derivative of the function $y=f(x)=Kx^3$.*

See the steps in a related calculation of the derivative of $y = f(x) = Kx^n$.

Solution. For $y = f(x) = Kx^3$ we have

$$\begin{aligned}
\frac{dy}{dx} &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \\
&= \lim_{h\to 0} \frac{K(x+h)^3 - Kx^3}{h} \\
&= \lim_{h\to 0} K\frac{(x^3+3x^2h+3xh^2+h^3)-x^3}{h} \\
&= \lim_{h\to 0} K\frac{(3x^2h+3xh^2+h^3)}{h} \\
&= \lim_{h\to 0} K(3x^2+3xh+h^2) \\
&= K(3x^2) = 3Kx^2. \qquad (3.3)
\end{aligned}$$

Thus the derivative of $f(x) = Kx^3$ is $f'(x) = 3Kx^2$. ◊

Example 3.13 *Use the definition of the derivative to compute $f'(x)$ for the function $y = f(x) = 1/x$ at the point $x = 1$.*

See the calculation of the derivative of $y = f(x) = 1/x$.

Solution. We write down the formula for this calculation at any point x and then simplify algebraically, using common denominators to combine fractions, and then, in the final step, calculate the limit formally. Lastly we substitute the value $x = 1$ to find $f'(1)$.

$$\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} && \text{(the definition)} \\
&= \lim_{h\to 0} \frac{\frac{1}{(x+h)} - \frac{1}{x}}{h} && \text{(applied to the function)} \\
&= \lim_{h\to 0} \frac{\frac{[x-(x+h)]}{x(x+h)}}{h} && \text{(common denominator)} \\
&= \lim_{h\to 0} \frac{-h}{hx(x+h)} && \text{(algebraic simplification)} \\
&= \lim_{h\to 0} \frac{-1}{x(x+h)} && \text{(cancel factor of } h\text{)} \\
&= -\frac{1}{x^2} && \text{(limit evaluated)} \qquad (3.4)
\end{aligned}$$

Thus, the derivative of $f(x) = 1/x$ is $f'(x) = -1/x^2$ and at the point $x = 1$ it takes the value $f'(1) = -1$. ◊

In Exercise 3. 8 we apply similar techniques to the derivative of the square-root function to show that

$$y = f(x) = \sqrt{x} \quad \Rightarrow \quad f'(x) = \frac{1}{2\sqrt{x}}. \qquad (3.5)$$

In the next chapter, we formalize some observations about derivatives of power functions and rules of differentiation. This allows us to simplify some of the calculations involved in finding derivatives.

3.3 The computational view: software to the rescue!

Section 3.3 Learning goals

1. Use software to numerically compute an approximation to the derivative.
2. Explain that the approximation replaces a (true) tangent line with an (approximating) secant line.
3. Explain using words how the derivative shape is connected with the shape of the original function.
4. Interpret the differences between two types of biochemical kinetics: Michaelis-Menten and Hill function.

Mastered Material Check

16. Describe how to find the derivative of a function $f(x)$ at $x = x_0$ analytically.
17. Describe how to find the derivative of a function $f(x)$ at $x = x_1$ geometrically.

We have explored geometric and analytic aspects of the derivative. *Here we show a third aspect of the derivative: its numerical implementation* using a simple spreadsheet. The ideas introduced here reappear in a variety of problems where repetitive calculations are needed to arrive at a solution.

Definition 3.4 (Numerical derivative) *A numerical derivative is an approximation to the value of the derivative, obtained by using a finite value of* Δx*,*

$$f'(x)_{numerical} \approx \frac{\Delta f}{\Delta x} \quad \text{rather than the actual value } f'(x)_{actual} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}.$$

The numerical derivative would be a *good* approximation of the true derivative provided that Δx is "small enough", that is, a step size for which the function $f(x)$ does not changes dramatically. Since Δf is the difference of two values of f, ($\Delta f = f(x+\Delta x) - f(x)$) it follows that *the numerical derivative is the same as the slope of a secant line*. This important realization, associated with the second learning goal in this section, means that *a secant line is often used to approximate a tangent line, and the slope of a secant line is used to approximate a derivative in numerical computations*. We see this idea again in several contexts.

Derivative of Michaelis-Menten and Hill functions

A spreadsheet can be used to numerically approximate derivatives. We illustrate this using as examples the reaction speeds for Michaelis-Menten Eqn. (1.8) and Hill function kinetics, Eqn. (1.7) (see Section 1.5), repeated below:

$$v_{MM} = f_1(c) = \frac{Kc}{a+c}, \tag{3.6}$$

$$v_{Hill} = f_2(c) = \frac{Kc^n}{a^n + c^n}. \tag{3.7}$$

Hint: Why would we ever use a spreadsheet, when there is other software for graphing? Basically, spreadsheets are powerful tools for computation, for manipulating data and for eliminating repetitive calculations. We gradually pick up some important skills in this course.

(Both $f_1(c), f_2(c)$ are shown as functions of c in Figure 3.13a.) Our goal is to use a spreadsheet to compute a numerical approximation of the derivative of these two reaction speeds with respect to the chemical concentration c.

Featured Problem 3.1 (The derivative on a spreadsheet) *Use a spreadsheet (or your favorite software) to plot the derivatives of the functions* $v_{MM} = f_1(c), v_{Hill} = f_2(c)$.

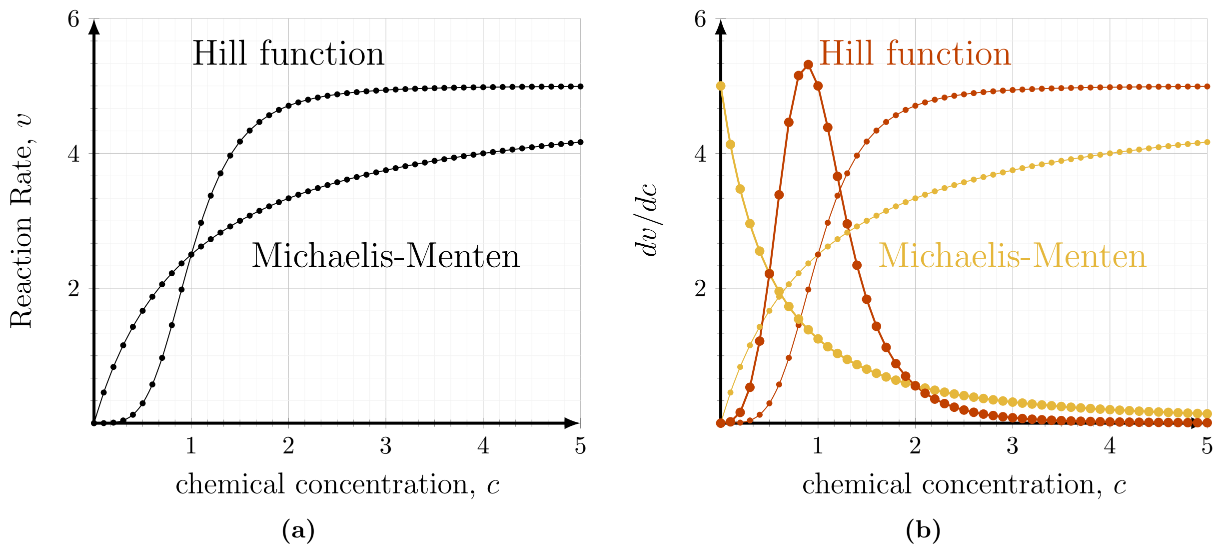


Figure 3.13: (a) A plot of $f_1(c), f_2(c)$ produced on a spreadsheet. (b) A plot of both the functions and their (approximate) derivatives.

Solution. Figure 3.13 shows the results of the spreadsheet calculation, but here we use only the Hill function to demonstrate typical spreadsheet manipulations.

(A) Open your favorite spreadsheet and label columns where data will be kept. In the link we display a spreadsheet with columns for the step size Δc, concentration c, and values of the function $f_2(c)$, with $K = 5, a = 1, n = 4$. The last column contains the approximation for the derivative $\Delta f / \Delta c$.

  (1) We input the desired step size Δc, here set to 0.1 (cell **A2**)

  (2) Input the value of c at which to start the calculations. Here we used $c = 0$ as the left endpoint (input 0 into cell **B2**). Then let the spreadsheet create the entire set of c values by inputting $= B2 + \$A\$2$ in cell **B3**, and dragging the "fill handle" (small square dot on bottom right hand corner) down the column. Note that the symbols \$ are universally used in spreadsheets to denote an absolute reference to a particular cell, whereas all other references are relative.

  (3) In cell **C2** type $= 5 * B2 \wedge 4/(1 + B2 \wedge 4)$ to create the formula for the desired function. The symbol $\wedge$ denotes a power. This will generate the

Link to Google Sheets. This spreadsheet shows how to create an approximation for the derivative of the Hill function in (3.7). Fig 3.13 was produced by a similar set of calculations, see Feature Problem 3.1. You can view this sheet and copy it elsewhere. You cannot edit it as is.

first point on the graph to be plotted. Similarly drag the fill handle to generate the rest of the points $f_2(c)$ corresponding to the all c values.

(B) Next, we compute the desired numerical approximations of the derivatives.

(1) Use column **D** for the numerical derivative of $f_2(c)$. To do so, *approximate* the actual derivative with a **finite difference**,

$$\frac{\Delta f_2}{\Delta c} \approx \frac{df_2}{dc}.$$

Note: importantly, the two expressions are *not equal*. However, for sufficiently small Δc, they approximate one another well.

(2) Dragging the fill handle down the column generates the desired list of values of the numerical derivative.

(C) Create a chart and plot the results. The x axis is the set of values of c.

Results of the above process (but modified for the two functions and their two derivatives) lead to the graphs shown on the right panel of Figure 3.13. ♢

Example 3.14 *Interpret the graphs of the derivatives in Figure 3.13b in terms of the way that reaction speed increases as the chemical concentration is increased in each of Michaelis-Menten and Hill function kinetics*

Mastered Material Check

18. Given time dependent data, can an exact derivative ever be determined?

Solution. Both derivatives are positive everywhere, since both $f_1(c)$ and $f_2(c)$ are increasing functions. For Michaelis-Menten, the derivative is always decreasing. This agrees with the observation that $f_1(c)$ (thin yellow curve) gradually levels off and flattens as c increases. While the reaction rate v_{MM} increases with c, the rate of increase, dv/dc, slows due to saturation at higher c values.

In contrast, the Hill function derivative *starts at zero*, increases sharply, and only then decreases to zero. Correspondingly, the Hill function (thin red curve) is flat at first, then becomes steeply increasing, and finally flattens to an asymptote. We can summarize this biochemically by saying that the initial reaction rate v_{Hill} is small and hardly changes near $c \sim 0$. For intermediate range of c, the reaction rate depends sensitively on c (evidenced by large dv/dc). As c increases to higher values, saturation slows down the rate of reaction, leading to the drop in dv/dc. ♢

3.4 Summary

1. If we "zoom in" enough on a point x_0 on a graph of a function (with "smooth" behaviour), we see a straight line. This straight line is the tangent line at that point. The slope of this line is the derivative (instantaneous rate of change) at that point, x_0.
2. Given the graph of a function $f(x)$, the derivative $f'(x)$ can be sketched by approximating the slopes of the tangent lines of $f(x)$, and plotting those slopes as points.

3. A function is continuous at $x = a$ in its domain if $\lim_{x \to a} f(x) = f(a)$. Discontinuous functions might have a hole (removable discontinuity), a jump, or blow up.
4. Computing derivatives requires the use of limits:

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

Limits are detailed further in Appendix D. In the absence of analytical methods, or in the presence of only data, a numerical derivative calculus can be used to approximate:

$$f'(x)_{numerical} \approx \frac{\Delta f}{\Delta x}.$$

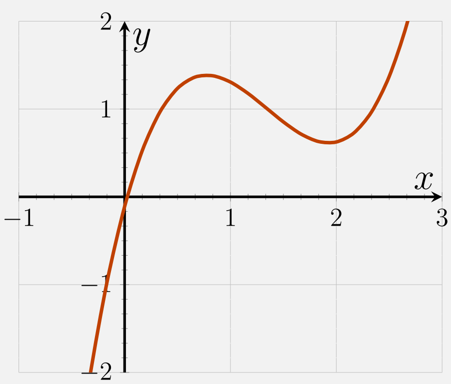
5. Derivatives that were computed in this chapters are summarized in Table 3.1.
6. The applications we encountered in this chapter included:
   (a) molecular motors and vesicle transport; and
   (b) Michaelis-Menten and Hill function kinetics for reaction speeds.

f(x)	**f′(x)**
C	0
Bx	B
$Bx+C$	B
Kx^3	$3Kx^2$
$\frac{1}{x}$	$-\frac{1}{x^2}$
$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$

Table 3.1: Some computed derivatives.

Quick Concept Checks

1. What geometric characteristic might stop a function from having a derivative?
2. Find the derivative of $y = 2x + 1$ using a
   (a) geometric argument, and
   (b) algebraic argument.
3. Draw the derivative of the function $f(x)$ at $x = 1$, depicted on the graph below.



4. Define a continuous function.

Exercises

3.1. **Sketching the derivative (geometric view).** Shown in Figure 3.14 are four functions. Sketch the derivative of each of these functions.

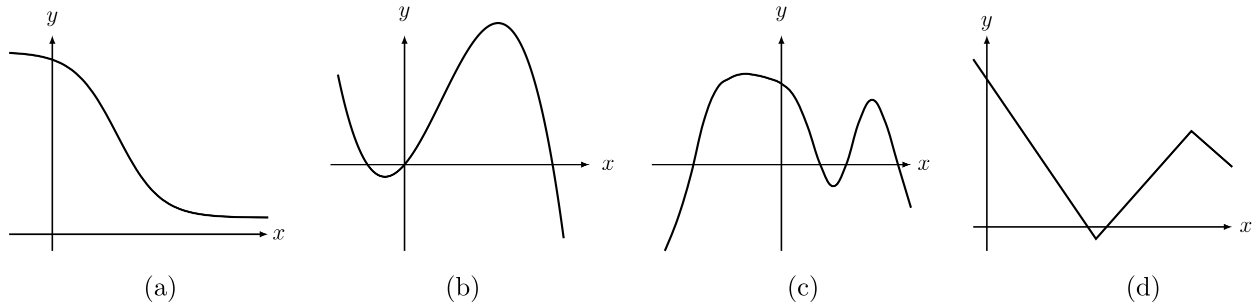


Figure 3.14: Figure for Exercise 1; graphs of four different functions.

3.2. **Sketching the function given its derivative.** Given the information in Table 3.2 about the the values of the derivative of a function, $g(x)$, sketch a (very rough) graph the function for $-3 \le x \le 3$.

3.3. **What the sign of the derivative tells us.** Given the information about the signs of the derivative of a function, $f(x)$ found in Table 3.2, sketch a (very rough) graph of the function for $-3 \le x \le 3$.

3.4. **Shallower or steeper rise.** Shown in Figure 3.15 are two similar functions, both increasing from 0 to 1 but at distinct rates. Sketch the derivatives of each one. Then comment on what your sketch would look like for a discontinuous "step function", defined as follows:

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & x \ge 0. \end{cases}$$

x	$g'(x)$	$f'(x)$
-3	-1	0
-2	0	+
-1	2	0
0	1	-
1	0	0
2	-1	+
3	-2	0

Table 3.2: Derivative data for two different functions: $g'(x)$ (Exercise 2) and $f'(x)$ (Exercise 3).

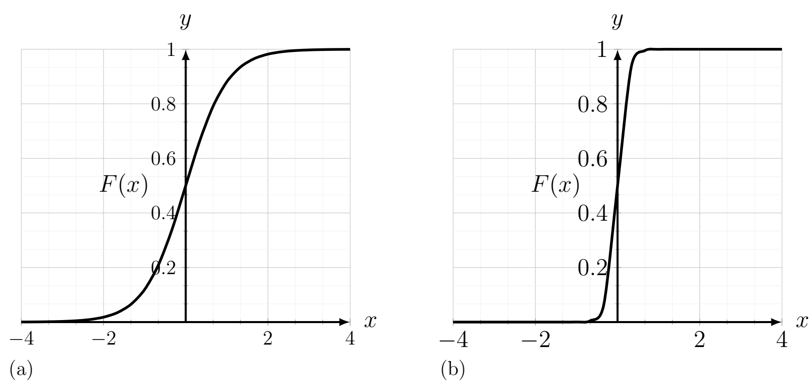


Figure 3.15: Figure for Exercise 4; two similar functions increasing at distinct rates.

3.5. **Introduction to velocity and acceleration.**The acceleration of a particle is the derivative of its velocity. Shown in Figure 3.16 is the graph of the velocity of a particle moving in one dimension. Indicate directly on the graph any time(s) at which the particle's acceleration is zero.

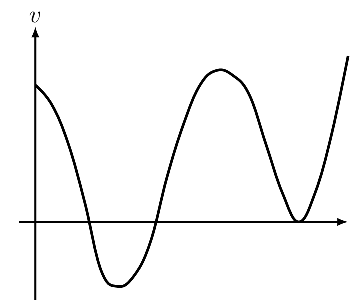


Figure 3.16: Figure for Exercise 5; velocity of a particle moving in one direction.

3.6. **Velocity, continued.** The vertical height of a ball, d (in meters) at time t (seconds) after it was thrown upwards was found to satisfy $d(t) = 14.7t - 4.9t^2$ for the first 3 seconds of its motion.

(a) What is the initial velocity of the ball (i.e. the instantaneous velocity at $t = 0$)?

(b) What is the instantaneous velocity of the ball at $t = 2$ seconds?

3.7. **Geometric view, continued.** Consider Figure 3.17.

(a) Given the function in Figure 3.17(a), graph its derivative.

(b) Given the function in Figure 3.17(b), graph its derivative

(c) Given the derivative $f'(x)$ shown in Figure 3.17(c) graph the function $f(x)$.

(d) Given the derivative $f'(x)$ shown in Figure 3.17(d) graph the function $f(x)$.

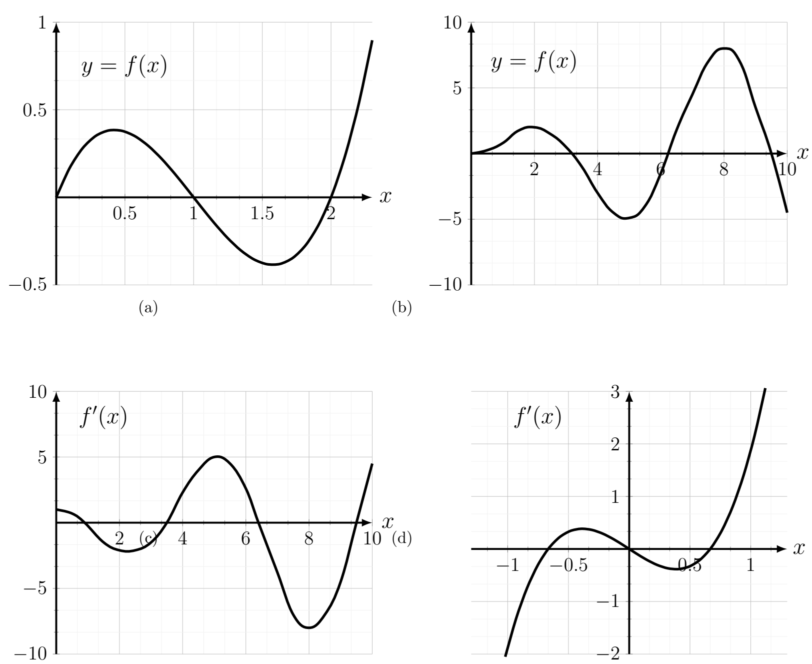


Figure 3.17: Figures of functions and derivatives for Exercise 7.

3.8. **Computing the derivative of square-root (from the definition).** Consider the function

$$y = f(x) = \sqrt{x}.$$

(a) Use the definition of the derivative to calculate $f'(x)$. Consider using the following algebraic simplification:

$$\frac{(\sqrt{a}-\sqrt{b})(\sqrt{a}+\sqrt{b})}{(\sqrt{a}+\sqrt{b})}=\frac{a-b}{(\sqrt{a}+\sqrt{b})}.$$

(b) Find the slope of the function at the point $x=4$.

(c) Find the equation of the tangent line to the graph at this point.

3.9. **Computing the derivative.** Use the definition of the derivative to compute the derivative of the function $y=f(x)=C/(x+a)$ where C and a are arbitrary constants. Show that your result is $f'(x)=-C/(x+a)^2$.

3.10. **Computing the derivative.** Consider the function

$$y=f(x)=\frac{x}{(x+a)}.$$

(a) Show that this function can be written as $f(x)=1-\dfrac{a}{(x+a)}$.

(b) Use the results of Exercise 9 to determine the derivative of this function (note: you do not need to use the definition of the derivative to do this computation). Show that you get $f'(x)=\frac{a}{(x+a)^2}$.

3.11. **Molecular motors.**

(a) Figure 3.18 (a) shows the displacement of a vesicle carried by a molecular motor. The motor can either walk right (R), left (L) along one of the microtubules or it can unbind (U) and be stationary, then rebind again to a microtubule. Sketch a rough graph of the velocity of the vesicle $v(t)$ and explain the sequence of events (using the letters R, L, U) that resulted in this motion.

(b) Figure 3.18 (b) shows the velocity $v(t)$ of another vesicle. Sketch a rough graph of its displacement starting from $y(0)=0$.

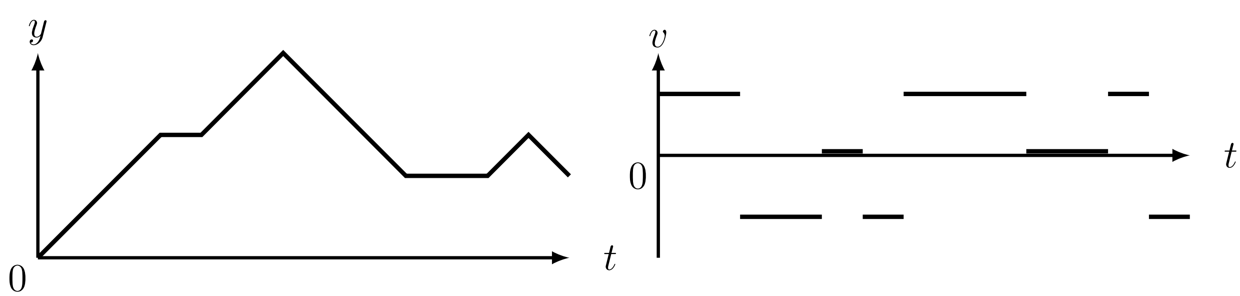


Figure 3.18: Figure for Exercise 11, molecular motors.

3.12. **Concentration gradient.** Certain types of tissues - epithelia - are made up of thin sheets of cells. Substances are taken up on one side of the sheet by some active transport mechanism, and then diffuse down a concentration gradient by a mechanism called facilitated diffusion on the opposite side.

Shown in Figure 3.19 is the concentration profile $c(x)$ of some substance across the width of the sheet (x represents distance). Sketch the corresponding concentration gradient, i.e. sketch $c'(x)$, the derivative of the concentration with respect to x.

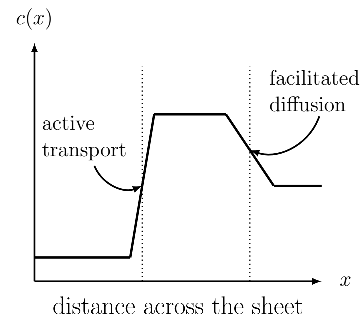


Figure 3.19: Figure for Exercise 12: concentration profile of a substance.

3.13. **Tangent line to a simple function.** What is the slope of the tangent line to the function $y = f(x) = 5x+2$ when $x = 2$? when $x = 4$? How would this slope change if a negative value of x was used? Why?

3.14. **Slope of the tangent line.** Use the definition of the derivative to compute the slope of the tangent line to the graph of the function $y = 3t^2 - t + 2$ at the point $t = 1$.

3.15. **Tangent line.** Find the equation of the tangent line to the graph of $y = f(x) = x^3 - x$ at the point $x = 1.5$ shown in Figure 3.1. You may use the fact that the tangent line goes through $(1.7, 1.47)$ as well as the point of tangency.

3.16. **Numerically computed derivative.** Consider the two Hill functions

$$H_1(x) = \frac{x^2}{0.01+x^2}, \quad H_2(x) = \frac{x^4}{0.01+x^4}$$

(a) Sketch a rough graph of these two functions on the same plot and/or describe in words what the two graphs would look like.

(b) On a second plot, sketch a rough graph of both derivatives of these functions and/or describe in words what the two derivatives would look like.

(c) Using a spreadsheet or your favourite software, plot the two functions over the range $0 \leq x \leq 1$.

(d) Use the spreadsheet to calculate an approximation for the derivatives $H_1'(x), H_2'(x)$ and plot these two functions together.

Note: in order to have a reasonably accurate set of graphs, you must select a small step size of $\Delta x \approx 0.01$.

3.17. **More numerically computed derivatives.** As we see in Chapter 14, trigonometric functions such as $\sin(t)$ and $\cos(t)$ can be used to describe biorhythms of various types. Here we numerically compute the first and second derivative of $y = \sin(t)$ and show the relationships between the trigonometric functions and their derivatives. We use only numerical methods (e.g. a spreadsheet), but in Chapter 14, we also study the analytical calculation of the same derivatives.

(a) Use a spreadsheet (or your favourite software) to plot, on the same graph the two functions

$$y_1 = \sin(t), y_2 = \cos(t), \quad 0 \leq t \leq 2\pi \approx 6.28.$$

Note that you should use a fairly small step size, e.g. $\Delta t = 0.01$ to get a reasonably accurate approximation of the derivatives.

(b) Use the same spreadsheet to (numerically) calculate (an approximate) derivative $y_1'(t)$ and add it to your graph.

(c) Now calculate $y_1''(t)$, that is (an approximation to) the derivative of the derivative of the sine function and add this to your graph.

4

Differentiation rules, simple antiderivatives and applications

In Chapter 2 we defined the derivative of a function, $y = f(x)$ by

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

Using this formula, we calculated derivatives of a few power functions. Here, we gather results so far, and observe a pattern, the **power rule**. This rule allows us to compute higher derivatives (e.g. second derivative etc.), to differentiate polynomials, and even to find *antiderivatives* by applying the rule "in reverse" (finding a function that has a given derivative). All these calculations are useful in common applications, including accelerated motion. These are investigated later in this chapter. We round out the technical material by stating several other rules of differentiation (product and quotient), allowing us to easily calculate derivatives of rational functions.

4.1 Rules of differentiation

Learning goals for Section 4.1

1. Express the power rule (Table 4.1) and be prepared to apply it to both derivatives and antiderivatives of power functions and polynomials.
2. Explain what is meant by the statement that "the derivative is a linear operation".
3. Describe the concept of an antiderivative and why it is defined only up to some constant.
4. Express the product and quotient rules and be able to apply these to calculating derivatives of products and of rational functions.

The derivative of power functions: the power rule

We have already computed the derivatives of several of the power functions. See Example 3.5 for $y = x^0 = 1$ and Example 3.6 for $y = x^1$. See also Example 2.10 for $y = x^2$ and Example 3.12 for $y = x^3$. We gather these results in Table 4.1. From the table, we observe that the derivative of a power function is also a power function: the original power becomes a coefficient and the new power is reduced by 1. We refer to this pattern as the **power rule** of differentiation.

We can show that this rule applies for any power function of the form $y = f(x) = x^n$ where n is an integer power. The calculation is essentially the same as examples illustrated in a previous chapter, but the step of expanding the binomial $(x+h)^n$ entails lengthier algebra. (Such expansion contains terms of the form $x^{n-k}h^k$ multiplied by **binomial coefficients**, and we include the details in Appendix E.) From now on, we simply use this result for any power function with integer powers (rather than calculating the derivative using its definition).

Function $f(x)$	**Derivative** $f'(x)$
1	0
x	1
x^2	$2x$
x^3	$3x^2$
$\vdots$	$\vdots$
x^n	nx^{n-1}
$x^{n/m}$	$(n/m)x^{(n/m)-1}$

Table 4.1: The **Power Rule** of differentiation states that the derivative of the power function $y = x^n$ is nx^{n-1}. For now, we have established this result for integer n. Later, we find this result holds for other powers that are not integer.

Example 4.1 *Find the equation of the tangent line to the graph of the power function $y = f(x) = 4x^5$ at $x = 1$, and determine the y-intercept of that tangent line.*

Solution. Using the power rule, the derivative of the function is

$$f'(x) = 20x^4.$$

At the point $x = 1$, we have $dy/dx = f'(1) = 20$ and $y = f(1) = 4$. This means that the tangent line goes through the point $(1,4)$ and has slope 20. Thus, its equation is

$$\frac{y-4}{x-1} = 20 \quad \Rightarrow \quad y = 4 + 20(x-1) = 20x - 16.$$

Letting $x = 0$ in the equation of the tangent line, we find that the y-intercept of line is $y = -16$. ♢

Mastered Material Check

1. Verify that the point $(1,4)$ satisfies the equation $y = 20x - 16$.

Example 4.2 (Energy loss and Earth's temperature) *In Section 1.3, we studied the energy balance on Earth. According to Eqn. (1.5), the rate of loss of energy from the surface of the Earth depends on its temperature according to the rule*

$$E_{out}(T) = 4\pi r^2 \varepsilon \sigma T^4.$$

Calculate the rate of change of this outgoing energy with respect to the temperature T.

Solution. The quantities π, ε, r are constants for this problem. Hence the rate of change ('derivative') of energy with respect to T, denoted $E'_{out}(T)$ is

$$E'_{out}(T) = (4\pi r^2 \varepsilon \sigma) \cdot 4T^3 = (16\pi r^2 \varepsilon \sigma)T^3.$$

◊

Next, we find that the result for derivatives of power functions can be extended to derivatives of polynomials, using simple properties of the derivative.

The derivative is a linear operation

The derivative satisfies several convenient properties, among them:

1. the derivative of a sum of two functions is the same as the sum of the derivatives; and

2. a constant multiplying a function can be brought outside the derivative.

We summarize these rules:

The derivative is a **linear operation**, that is:

$$\frac{d}{dx}(f(x)+g(x)) = \frac{df}{dx} + \frac{dg}{dx}, \tag{4.1}$$

$$\frac{d}{dx}Cf(x) = C\frac{df}{dx}. \tag{4.2}$$

Mastered Material Check

2. Verify Eqns. (4.1) and (4.2) hold for $f(x) = x^3$, $g(x) = x^2$ and $C = 4$.
3. Give an example which shows squaring is **not** a linear operation.

In general, a linear operation L is a rule or process that satisfies two properties:

1. $L[f+g] = L[f] + L[g]$ and

2. $L[cf] = cL[f]$,

where f, g are objects (such as functions, vectors, etc.) on which L acts, and c is a constant multiplier. We refer to Eqns. (4.1) and (4.2) as the **linearity** properties of the derivative.

The derivative of a polynomial

Using the linearity of the derivative, we can extend our differentiation power rule to compute the derivative of any polynomial. Recall that polynomials are sums of power functions multiplied by constants. A polynomial of **degree** n has the form

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots a_1 x + a_0, \tag{4.3}$$

where the coefficients, a_i are constant and n is an integer. Hence, the derivative of a polynomial is just the sum of derivatives of power functions (multiplied by constants). Formally, the derivative of Eqn. (4.3) is

$$p'(x) = \frac{dy}{dx} = a_n \cdot nx^{n-1} + a_{n-1}\cdot(n-1)x^{n-2} + \cdots + a_1. \tag{4.4}$$

(Observe that each term consists of the coefficient times the derivative of a power function. The constant term a_0 has disappeared since the derivative of any constant is zero.) The derivative, $p'(x)$, is a function in its own right, and a polynomial as well. Its degree, $n-1$, is one less than that of $p(x)$. In view of this observation, we could ask: what is the derivative of the derivative?

Notation: we henceforth refer to the "derivative of the derivative" as a **second derivative**, written in the notation $p''(x)$ or, equivalently $\frac{d^2p}{dx^2}$.

Using the same rules, we can compute, obtaining

$$p''(x) = \frac{d^2p}{dx^2} = a_n n(n-1)x^{n-2} + a_{n-1}(n-1)(n-2)x^{n-3} + \cdots + a_2. \quad (4.5)$$

The following examples should be used for practice.

Mastered Material Check

4. Are their other notations for the second derivative of $p(x)$ that you might expect?
5. Check that Eqn. (4.7) is an antiderivative of Eqn. (4.6) by differentiating Eqn. (4.7).

Example 4.3 *Find the first and second derivatives of the function*

(a) $y = f(x) = 2x^5 + 3x^4 + x^3 - 5x^2 + x - 2$ *with respect to x and*

(b) $y = f(t) = At^3 + Bt^2 + Ct + D$ *with respect to t.*

Solution. We obtain the results

(a) $f'(x) = 10x^4 + 12x^3 + 3x^2 - 10x + 1$ and $f''(x) = 40x^3 + 36x^2 + 6x - 10$.

(b) $f'(t) = 3At^2 + 2Bt + C$ and $f''(t) = 6At + 2B$.

In (b) the independent variable is t, but, of course, the rules of differentiation are the same. ◊

Antiderivatives of power functions and polynomials

Given a derivative, we can ask "what function was differentiated to lead to this result?" This reverse process is termed **antidifferentiation**, and the function we seek is then called an **antiderivative**. Antidifferentiation reverses the operation of differentiation. We ask, for example, which function has as its derivative

$$y'(t) = At^n? \quad (4.6)$$

The original function, $y(t)$, should have a power higher by 1 (of the form t^{n+1}), but the "guess" $y_{guess} = At^{n+1}$ is not quite right, since differentiation results in $A(n+1)t^n$. To fix this, we revise the "guess" to

$$y(t) = A\frac{1}{(n+1)}t^{n+1}. \quad (4.7)$$

Question. Is this the only function that has the desired property? No, there are other functions whose derivatives are the same. For example, consider adding an arbitrary constant C to the function in Eqn. (4.7) and note that we obtain the same derivative (since the derivative of the constant is zero). We summarize our findings:

$$\text{The antiderivative of } y'(t) = At^n \quad \text{is} \quad y(t) = A\frac{1}{(n+1)}t^{n+1} + C. \tag{4.8}$$

We also note a similar result that holds for functions in general:

Given a function, $f(x)$ we can only determine its antiderivative **up to some (additive) constant**.

Mastered Material Check

6. What is an example of a constant that is **not** additive?
7. Verify the solution to Example 4.4 by differentiating.
8. In Example 4.5, identify all:
   (a) constants,
   (b) dependent variables, and
   (c) independent variables.

We can extend the same ideas to finding the antiderivative of a polynomial.

Example 4.4 (Antiderivative of a polynomial) *Find an antiderivative of the polynomial $y'(t) = At^2 + Bt + C$.*

Solution. Since differentiation is a linear operation, we can construct the antiderivative by antidifferentiating each of the component power functions. Applying Eqn. (4.8) to the components we get,

$$y(t) = A\frac{1}{3}t^3 + B\frac{1}{2}t^2 + Ct + D,$$

where D is an arbitrary constant. We see that *the antiderivative of a polynomial is another polynomial whose degree is higher by 1*. ♢

Example 4.5 *The* second *derivative of some function is $y''(t) = c_1t + c_2$. Find a function $y(t)$ for which this is true*

Solution. The above polynomial has degree 1. Evidently, this function resulted by taking the derivative of $y'(t)$, which had to be a polynomial of degree 2. We can check that either $y'(t) = \frac{c_1}{2}t^2 + c_2t$, or $y'(t) = \frac{c_1}{2}t^2 + c_2t + c_3$ (for any constant c_3) could work. In turn, the function $y(t)$ had to be a polynomial of degree 3. One such function is

$$y(t) = \frac{c_1}{6}t^3 + \frac{c_2}{2}t^2 + c_3t + c_4,$$

where c_4 is any constant. In short, the relationship is:

$$\text{for differentiation} \quad y(t) \to y'(t) \to y''(t),$$

whereas

$$\text{for antidifferentiation} \quad y''(t) \to y'(t) \to y(t).$$

♢

These results are used in applications to acceleration, velocity, and displacement of a moving object in Section 4.2.

Product and quotient rules for derivatives

So far, using the power rule, and linearity of the derivative (Section 4.1), we calculated derivatives of polynomials. Here we state without proof (see Appendix E), two other rules of differentiation.

The product rule: If $f(x)$ and $g(x)$ are two functions, each differentiable in the domain of interest, then

$$\frac{d[f(x)g(x)]}{dx} = \frac{df(x)}{dx}g(x) + \frac{dg(x)}{dx}f(x).$$

Another notation for this rule is

$$[f(x)g(x)]' = f'(x)g(x) + g'(x)f(x).$$

Mastered Material Check

9. What does 'domain of interest' mean?
10. What is the domain of the function $f(x) = \frac{1}{x}$?
11. Verify Example 4.6 by noting $f(x)g(x) = x(1+x) = x+x^2$ and differentiating.

Example 4.6 *Find the derivative of the product of the two functions $f(x) = x$ and $g(x) = 1+x$.*

Solution. Using the product rule leads to

$$\begin{aligned}\frac{d[f(x)g(x)]}{dx} = \frac{d[x(1+x)]}{dx} &= \frac{d[x]}{dx}\cdot(1+x) + \frac{d[(1+x)]}{dx}\cdot x \\ &= 1\cdot(1+x)+1\cdot x = 2x+1.\end{aligned}$$

◊

The quotient rule: If $f(x)$ and $g(x)$ are two functions, each differentiable in the domain of interest, then

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{\frac{df(x)}{dx}g(x) - \frac{dg(x)}{dx}f(x)}{[g(x)]^2}.$$

We can also write this in the form

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

Example 4.7 *Find the derivative of the function $y = ax^{-n} = a/x^n$ where a is a constant and n is a positive integer.*

Solution. We can rewrite this as the quotient of the two functions $f(x) = a$ and $g(x) = x^n$. Then $y = f(x)/g(x)$ so, using the quotient rule leads to the derivative

$$\begin{aligned}\frac{dy}{dx} = \frac{f'(x)g(x) - g'(x)f(x)}{[g(x)]^2} &= \frac{0\cdot x^n - (nx^{n-1})\cdot a}{(x^n)^2} \\ &= -\frac{anx^{n-1}}{x^{2n}} = a(-n)x^{-n-1}.\end{aligned}$$

◊

This calculation shows that **the power rule of differentiation holds for negative integer powers.**

Featured Problem 4.1 (Dynamics of actin in the cell) *Actin is a structural protein that forms long filaments and networks in living cells. The actin network is continually assembling from small components (actin monomers) and disassembling back again. To study this process, scientists attach fluorescent markers to actin, and watch the* **fluorescence intensity** *change over time. In one experiment, both red and green fluorescent labels were used. The green label fluoresces only after it is activated by a pulse of light, whereas the red fluorescent protein is continually active.*

It was noted that the red and green fluorescence intensities (R, G) satisfied the following relationships:

Such relationships between a function of time and its own derivative are examples of **differential equations**, a topic we revisit in Chapter 11.

$$\frac{dR}{dt} = (a-b)R, \quad \frac{dG}{dt} = -bG$$

where a, b are constants that characterize the rate of assembly and disassembly ("breakup") of actin. Show that the derivative of the ratio, $\frac{d(R/G)}{dt}$, can be expressed in terms of the ratio R/G.

Featured Problem 4.2 (Derivative of the Beverton-Holt function) *The Beverton-Holt model for fish population growth was discussed in Featured Problem 1.1. A function relating the population of fish this year, y to the population of fish in the previous year x was*

$$y = f(x) = k_1 \frac{x}{(1+k_2 x)}$$

(where we have simplified the notation, $x = N_0, y = N_1$ of Eqn. 1.9. How sensitive is this year's population to slight changes in last year's population? Compute the derivative dy/dx to answer this question.

A preview of the chain rule

We give a brief preview of the **chain rule**, an important rule discussed in detail in Chapter 8. This rule extends the rules of differentiation to **composite functions**, that is, functions made up of applying a sequence of operations one after the other. For example, the two functions

$$f(u) = u^{10}, \quad u = g(x) = 3x^2 + 1, \tag{4.9}$$

could be applied one after the other to lead to the new composite function

$$f(g(x)) = \left(3x^2+1\right)^{10}.$$

The chain rule states that the derivative of this new function with respect to x is the product of derivatives of the individual functions.

The chain rule
If $y = f(z)$ and $z = g(x)$ are two functions, then the derivative of the composite function $f(g(x))$ is

$$\frac{dy}{dx} = \frac{df}{dz}\frac{dz}{dx}.$$

Example 4.8 *Use the chain rule to differentiate the composite function* $y = f(g(x)) = \left(3x^2+1\right)^{10}$.

Solution. The functions being composed are the same as in Eqn. (4.9). Applying the chain rule gives:

$$\frac{dy}{dx} = \frac{df}{du}\frac{du}{dx} = 10u^9 \cdot 6x = 60(3x^2+1)^9 x.$$

◊

The details of how to use the chain rule, and many applications are postponed to Chapter 8.

The power rule for fractional powers

Using the definition of the derivative, we have already shown that the derivative of $\sqrt{x}$ is $y'(x) = \frac{1}{2\sqrt{x}}$ (see Exercise 2.8). We restate this result using fractional power notation. Recall that $\sqrt{x} = x^{1/2}$.

The derivative of $y = \sqrt{x}$ is $y'(x) = \frac{1}{2}x^{-1/2}$.

This idea can be generalized to any fractional power. Indeed, we state here a result (to be demonstrated in Chapter 9).

Derivative of fractional-power function:

The derivative of $y = f(x) = x^{m/n}$ is $\displaystyle \frac{dy}{dx} = \frac{m}{n}x^{\left(\frac{m}{n}-1\right)}$.

Notice that this appears in the last row of Table 4.1.

Example 4.9 (Energy loss and Earth's temperature, revisited) *In Example 4.2, we calculated the rate of change of energy lost per unit change in the Earth's temperature based on Eqn. (1.5). Find the rate of change of Earth's temperature per unit energy loss based on the same equation.*

Solution. We are asked to find dT/dE_{out}. We first rewrite the relationship to express T as a function of E_{out}. To do so, we solve for T in Eqn. (1.5), obtaining

$$T = \left(\frac{E_{out}}{4\pi r^2 \varepsilon\sigma}\right)^{1/4} = \left(\frac{1}{4\pi r^2 \varepsilon\sigma}\right)^{1/4} E_{out}^{1/4} = KE_{out}^{1/4}.$$

The first term is a constant, and we use the rule for a derivative of a fractional power to compute that

$$\frac{dT}{dE_{out}} = \left(\frac{1}{4\pi r^2 \varepsilon\sigma}\right)^{1/4} \cdot \frac{1}{4} E_{out}^{(1/4)-1} = \left(\frac{1}{16\pi r^2 \varepsilon\sigma}\right)^{1/4} E_{out}^{-3/4}.$$

Note: in Chapter 9 we will show that **implicit differentiation** can be used to find the desired derivative without first solving for the variable of interest. ◊

4.2 *Application of the second derivative to acceleration*

Section 4.2 Learning goals

1. Recognize that velocity and acceleration are first and second derivatives of position with respect to time (and that velocity and position are first and second antiderivatives of acceleration.)
2. Given constant acceleration, be able to find the velocity and displacement of the moving object.

We have already defined the "derivative of a derivative" as the **second derivative**. Here we provide a natural example of a second derivative, the **acceleration** of an object: the rate of change of the velocity (which is, as we have seen, the rate of change of a displacement). We restate these relationships in terms of antiderivatives below.

Position, velocity, and acceleration

Consider an object falling under the force of gravity. Let $y(t)$ denote the position of the object at time t. From now on, we refer to the instantaneous velocity at time t simply as **the velocity**, $v(t)$.

Definition 4.1 (The velocity) *Given the position of some particle as a function of time, $y(t)$, we define the velocity as the rate of change of the position, i.e. the derivative of $y(t)$:*

$$v(t) = \frac{dy}{dt} = y'(t).$$

In general, v may depend on time, which we indicated by writing $v(t)$.

Definition 4.2 (The acceleration) *We define the acceleration as the (instantaneous) rate of change of the velocity, i.e. as the derivative of $v(t)$.*

$$a(t) = \frac{dv}{dt} = v'(t)$$

(acceleration could also depend on time, hence $a(t)$).

Mastered Material Check

12. Give three different examples of possible units for velocity.
13. Give three different examples of possible units for acceleration.

Since the acceleration is the derivative of the derivative of the original function, we also use the notation

$$a(t) = \frac{d}{dt}\left(\frac{dy}{dt}\right) = \frac{d^2y}{dt^2} = y''(t).$$

Note: we have used three equivalent ways of writing a second derivative. This notation evolved for historical reasons, and is used interchangeably in science. Acceleration is a second derivative of the position.

Given $a(t)$, the acceleration as a function of t, we can use antidifferentiation to obtain the velocity $v(t)$. Similarly, we can use the velocity $v(t)$ to determine the position $y(t)$ (up to some constant). The constants must be obtained from other information, as the next examples illustrate.

Example 4.10 (Uniformly accelerated motion of a falling object) *Suppose that the acceleration of an object is constant in time, i.e.* $a(t) = -g =$ *constant. Use antidifferentiation to determine the velocity and the position of the object as functions of time.*

Video summary of how to compute the velocity and displacement given uniform acceleration.

Note: here we have a coordinate system in which the positive direction is "upwards", and so the acceleration, which is in the opposite direction, is negative. On Earth, $g = 9.8$ m /s^2.

Solution. First, what function of time $v(t)$ has the property that

$$a(t) = v'(t) = -g = \text{constant?}$$

The function $a(t) = -g$ is a polynomial of degree 0 in the variable t, so the velocity, which is its antiderivative, has to be a polynomial of degree 1, such as $v(t) = -gt$. This is *one* antiderivative of the acceleration. Other functions, such as

$$v(t) = -gt + c, \tag{4.10}$$

would work for any constant c.

How would we pick a value for the constant c? We need additional information. Suppose we are told that $v(0) = v_0$ is the known **initial velocity**. Later we refer to the statement $v(0) = v_0$ as an "initial condition," since it specifies how fast the object was moving initially at $t = 0$. Then, substituting $t = 0$ into Eqn. (4.10), we find that $c = v_0$. Thus in general,

$$v(t) = -gt + v_0$$

where v_0 is the initial velocity of the object.

Next, let us determine the position of the object as a function of t, that is, $y(t)$. Recall that $v(t) = y'(t)$. Thus, antidifferentiation leads to a polynomial of degree 2,

$$y(t) = -\frac{1}{2}gt^2 + v_0 t + k, \tag{4.11}$$

where, as before we allow for some additive constant k. Reasoning as before, we can determine the value of the constant k from a known initial position of the object $y(0) = y_0$. As before (plugging $t = 0$ into Eqn. (4.12)), we find that $k = y_0$, so that

$$y(t) = -\frac{1}{2}gt^2 + v_0 t + y_0. \qquad (4.12)$$

Here we assumed that the acceleration is due to gravity, $-g$, but any other motion with constant acceleration would be treated in the same way. ♢

Mastered Material Check

14. Verify that the derivative of $y(t) = -12gt^2 + v_0 t + k$ is the given expression for $v(t)$.

Summary, uniformly accelerated motion: If an object moves with constant acceleration $-g$, then given its initial velocity v_0 and initial position y_0 at time $t = 0$, the position at any later time is given by:

$$y(t) = -\frac{1}{2}gt^2 + v_0 t + y_0.$$

This powerful conclusion is a direct result of

1. the assumption that the acceleration is constant,
2. the elementary rules of calculus, and
3. the definitions of velocity and acceleration as first and second derivatives of the position.

We further explore the relationships between position, velocity and acceleration of a falling object in the examples below.

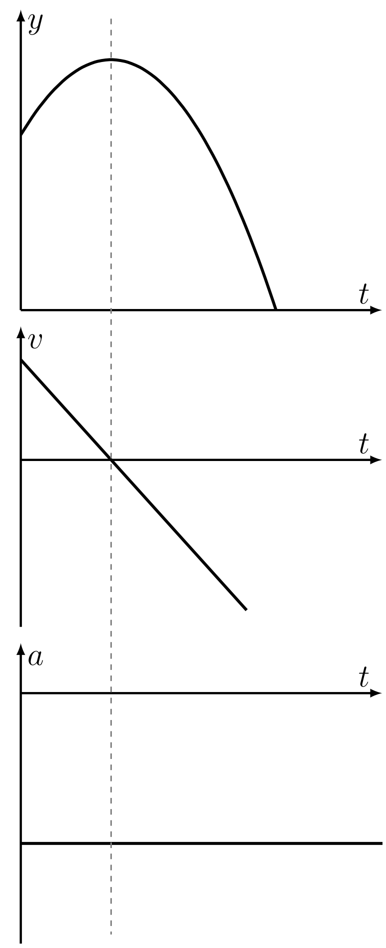


Figure 4.1: The position, velocity, and acceleration of an object that is thrown upwards and falls under the force of gravity.

Example 4.11 *Determine when the object reaches its highest point, and its velocity at that time.*

Solution. When the object reaches its highest point, its velocity has decreased to zero. From then on, the velocity becomes negative and the object falls back down. We solve for t in the equation $v(t) = 0$:

$$v(t) = v_0 - gt = 0 \quad \Rightarrow \quad t_{top} = \frac{v_0}{g}.$$

♢

Example 4.12 *When does the object hit the ground and with what velocity?*

Solution. Since y is height above ground, the object hits the ground when $y = 0$. Then we must solve for t in the equation $y(t) = 0$. This turns out to be a quadratic equation:

$$y(t) = \frac{1}{2}gt^2 - v_0 t - y_0 = 0, \quad \Rightarrow \quad t_{ground} = \frac{v_0 \pm \sqrt{v_0^2 + 2gy_0}}{g}.$$

We are interested in a solution with $t \geq 0$, so, rejecting the negative root,

Use the sliders to see how the initial velocity v_0 and initial height y_0 affect the time to reach the top and the time to reach the ground on this interactive graph.

Video summary of how we find the time that the object reaches its highest point and the time that it hits the ground.

$$t_{\text{ground}} = \frac{v_0}{g} + \frac{\sqrt{v_0^2 + 2gy_0}}{g}.$$

The velocity of the object when it hits the ground, $v(t_{\text{ground}})$, is then

$$v(t_{\text{ground}}) = v_0 - gt_{\text{ground}} = v_0 - g\left(\frac{v_0}{g} + \frac{\sqrt{v_0^2 + 2gy_0}}{g}\right) = -\sqrt{v_0^2 + 2gy_0}.$$

We observe that this velocity is negative, indicating (as expected) that the object is falling *down*. Figure 4.1 illustrates the relationship between the displacement, velocity and acceleration as functions of time. ◊

Mastered Material Check

15. Answer Examples 4.11 and 4.12 when $g = 5$ m/s^2, $v_0 = 20$ m/s and $y_0 = 100$ m.
16. What is the dashed line in Figure 4.1 indicating?

Featured Problem 4.3 (How to outrun a cheetah) *A cheetah can run at top speed v_c but it has to slow down (decelerate) to keep from getting too hot. Assume that the cheetah has a constant rate of (negative) acceleration $a = -a_c$. A gazelle runs at a slower speed v_g but it can maintain that constant speed for a long time. If the gazelle is initially a distance d from the cheetah, and running away, when would it be caught? Is there a (large enough) distance d such that it never gets caught?*

A cheetah chasing a gazelle. Who will win the race?

4.3 *Sketching the first and second derivative and the anti-derivatives*

Section 4.3 Learning goals

1. Given a sketch of a function, sketch its first and second derivatives.
2. Given the sketch of a function, sketch its first and second antiderivatives.

We continue to practice sketching derivatives (first *and* second) as well as antiderivatives, given a sketch of the original function. Here, we aim for *qualitative features*, where only the most important aspects of the graphs (locations of key points such as peaks and troughs) are indicated.

Example 4.13 (Sketching the derivative from the original function)
Sketch the first and second derivatives of the functions in Fig. 4.2.

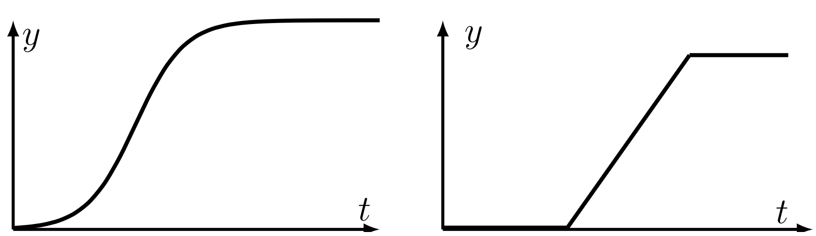


Figure 4.2: Function graphs for Example 4.13; one smooth, the other with cusps.

Solution. In Figure 4.3 we show the functions $y(t)$ (top row), their first derivatives $y'(t)$ (middle row), and the second derivatives $y''(t)$ (bottom row). In each case, we determined the slopes of tangent lines as a first step.

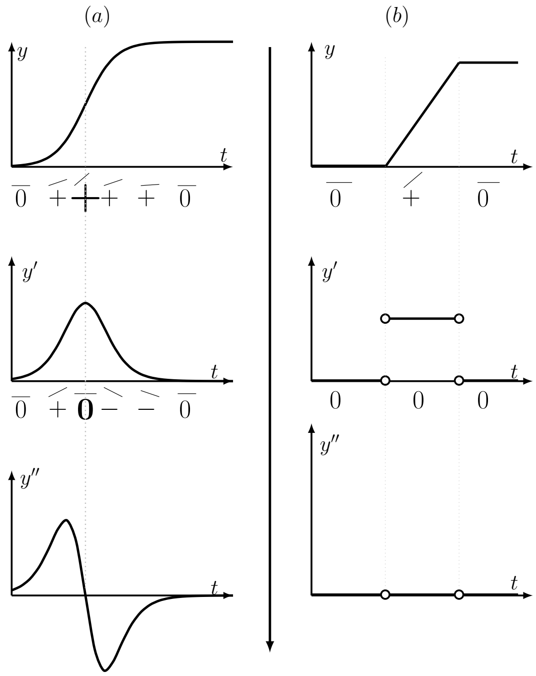


Figure 4.3: Solutions for Example 4.13: a sketch of first and second derivatives.

Along flat parts of the graph), the derivative is zero. This is indicated at several places in Fig. 4.3. In (b), there are cusps at which derivatives are not defined. ♢

Mastered Material Check

17. Identify any cusps in Figure 4.2.
18. What is each dotted line indicating in Figure 4.3?
19. What is the arrow in the middle of Figure 4.3 indicating?

Example 4.14 (Sketching a function from a sketch of its derivative)
Sketch the antiderivatives $y(x)$ for each derivative $y'(x)$ shown in Figure 4.4.

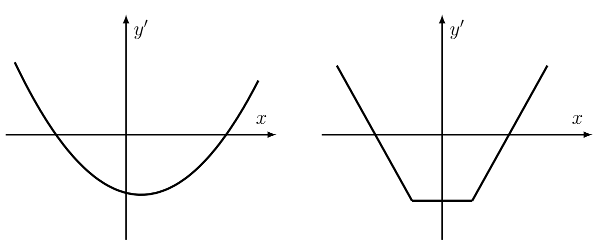


Figure 4.4: Derivatives $y'(x)$ for Example 4.14.

Solution. An antiderivative is only defined up to some (additive) constant. In the bottom panels of Figure 4.5 we show sample antiderivatives for each case. If we are given *additional* information, for example that $y(0) = 0$, we could then select one specific curve out of this family of solutions. A second point to observe is that antidifferentiation smoothes a function. Even though $y'(x)$ has cusps in (b), we find that $y(x)$ is smooth. We later see that the points at which $y'(x)$ has a cusp correspond to places where the concavity of $y(x)$ changes abruptly.

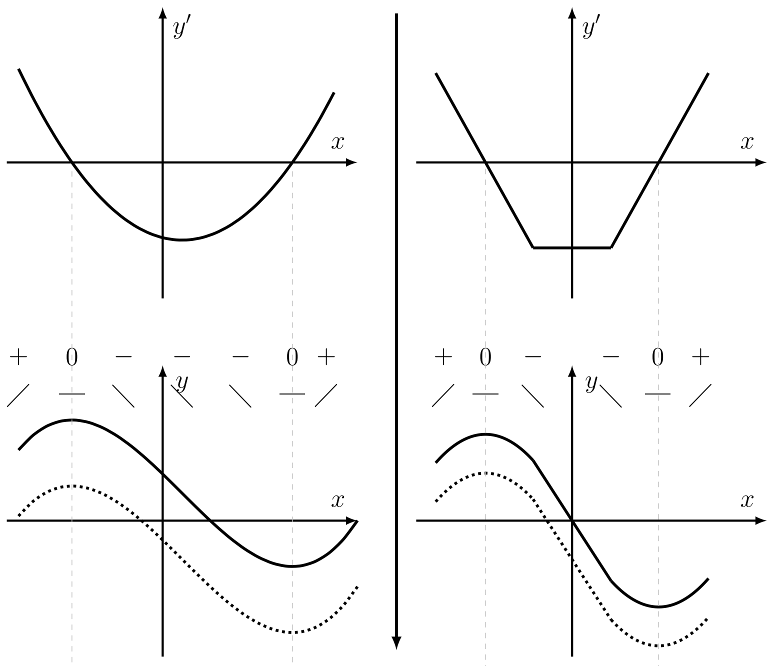


Figure 4.5: Using the sketch of two functions $y'(x)$ to sketch their antiderivative, $y(x)$.

◊

4.4 Summary

1. The derivative is a linear operation, which means that the derivative can be distributed over a sum of functions or exchanged with a constant multiplying a function. The derivative of a polynomial is thus a sum of derivatives of power functions.

2. A "second derivatives" is the derivate of a derivative.

3. Antiderivatives reverse the process of differentiation. Antidifferentiation is, however, not unique. Given a function $f(x)$, we can only determine its antiderivative up to some arbitrary constant.

4. The product, quotient and chain rules are,

$$[f(x)g(x)]' = f'(x)g(x) + g'(x)f(x),$$

$$\left[\frac{f(x)}{g(x)}\right]' = \frac{f'(x)g(x) - g'(x)f(x)}{[g(x)]^2},$$

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

These allow us to compute derivatives of functions made up of simpler components for which we have already established differentiation rules.

5. Given the graph of a function, we can use qualitative features - signs, zeros, peaks, valleys, and tangent lines slopes to sketch both its derivatives and antiderivatives.

6. The applications examined in this chapter included:

   (a) energy loss and Earth's temperature; and
   (b) position, velocity and acceleration of an object.

   In particular, these are related via differentiation and antidifferentiation. Given an object's position $y(t)$, its velocity is $v(t) = y'(t)$ and its acceleration is $a(t) = v'(t) = y''(t)$.

Quick Concept Checks

1. Which of the following operations are linear?

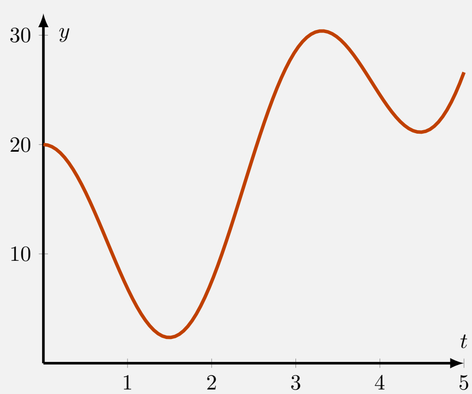
   (a) division
   (b) exponentiation
   (c) composition
   (d) squaring

2. Why are the product and chain rules useful?

3. Suppose an object has acceleration $a(t) = 10\text{m/s}^2$. What can you say about its:

   (a) accelleration at $t = 5$s?
   (b) velocity at time t?
   (c) velocity at time $t = 5$s?
   (d) position at time t?
   (e) position at time $t = 5$s?

4. Consider the following graph which describes the position y of an object at time t:



   Where is the object's velocity minimal? Maximal?

Exercises

4.1. **First derivatives.** Find the first derivative for each of the following functions.

(a) $f(x) = (2x^2 - 3x)(6x+5)$,

(b) $f(x) = (x^3+1)(1-3x)$,

(c) $g(x) = (x-8)(x^2+1)(x+2)$,

(d) $f(x) = (x-1)(x^2+x+1)$,

(e) $f(x) = \dfrac{x^2-9}{x^2+9}$,

(f) $f(x) = \dfrac{2-x^3}{1-3x}$,

(g) $f(b) = \dfrac{b^3}{2-b^{\frac{2}{3}}}$,

(h) $f(m) = \dfrac{m^2}{3m-1} - (m-2)(2m-1)$,

(i) $f(x) = \dfrac{(x^2+1)(x^2-2)}{3x+2}$.

4.2. **Logistic growth rate.** In logistic growth, the rate of growth of a population, R depends on the population size N as follows:

$$R = rN\left(1 - \frac{N}{K}\right),$$

where r and K are positive constants. Find the rate of change of the growth rate with respect to the population size, that is compute dR/dN.

Note: Logistic growth rate is further explored in Chapter 7.

4.3. **Michaelis-Menten and Hill kinetics.** Compute the derivatives of the following functions:

(a) The Michaelis Menten kinetics of Eqn. (1.8),

$$v = \frac{Kx}{k_n + x}.$$

(b) The Hill function of Eqn. (1.7), that is

$$y = \frac{Ax^n}{a^n + x^n}.$$

4.4. **Volume, surface area and radius of a sphere.** The volume and surface area of a sphere both depend on its radius:

$$V = \frac{4}{3}\pi r^3, \quad S = 4\pi r^2.$$

(a) Find the rate of change of the volume with respect to the radius and the rate of change of the surface area with respect to the radius.

(b) Find the rate of change of the surface area to volume ratio S/V with respect to the radius.

4.5. **Derivative of volume with respect to surface area.** Consider the volume and surface area of a sphere. (See Exercise 4 for the formulae.)

(a) Eliminate the radius and express V as a function of S.

(b) Find the rate of change of the volume with respect to the surface area.

4.6. **Surface area and volume of a cylinder.** The volume of a cylinder and the surface area of a cylinder with two flat end-caps are

$$V = \pi r^2 L, \quad S = 2\pi r L + 2\pi r^2$$

where L is the length and r the radius of the cylinder.

(a) Find the rate of change of the volume and surface area with respect to the radius, assuming that the length L is held fixed.

(b) Find the rate of change of the surface area to volume ratio S/V with respect to the radius assuming that the length L is held fixed.

4.7. **Growing circular colony.** A bacterial colony has the shape of a circular disk with radius $r(t) = 2 + t/2$ where t is time in hours and r is in units of mm. Express the area of the colony as a function of time and then determine the rate of change of area with respect to time at $t = 2$hr.

4.8. **Rate of change of energy during foraging.** When a bee forages for nectar in a patch of flowers, it gains energy. Suppose that the amount of energy gained during a foraging time span t is

$$f(t) = \frac{Et}{k+t}, \quad \text{where } E, k > 0 \text{ are constants.}$$

(a) If the bee stays in the patch for a very long time, how much energy can it gain?

(b) Use the quotient rule to calculate the rate of energy gain while foraging in the flower patch.

Note: foraging is returned to in Chapter 7.

4.9. **Ratio of two species.** In a certain lake it is found that the rate of change of the population size of each of two species $(N_1(t), N_2(t))$ is proportional to the given population size. That is

$$\frac{dN_1}{dt} = k_1 N_1, \quad \frac{dN_2}{dt} = k_2 N_2,$$

where k_1 and k_2 are constants. Find the rate of change of the ratio of population sizes (N_1/N_2) with respect to time $\frac{d(N_1/N_2)}{dt}$. Your answer should be in terms of k_1, k_2 and the ratio N_1/N_2.

4.10. **Invasive species and sustainability.** An invasive species is one that can out-compete and grow faster than the native species, resulting in takeover and displacement of the local ecosystem. Consider the two-lake system of Exercise 9. Suppose that initially, the ratio of the native species N_1 to the invasive species N_2 is very large. Under what condition (on the constants k_1, k_2) does that ratio decrease with time, i.e. does the invasive species take over?

4.11. **Numerical derivatives.** Consider the function

$$y(x) = 5x^3, \quad 0 \le x \le 1.$$

(a) Use a spreadsheet (or your favourite software) to compute an approximation of the derivative of this function over the given interval for $\Delta x = 0.25$ and compare to the true derivative, using the power rule. Comment on the comparison.

(b) Recompute the approximation to the derivative using $\Delta x = 0.05$ and comment on the results.

4.12. **Antiderivatives.** Find antiderivatives of the following functions, that is find $y(t)$.

(a) $y'(t) = t^4 + 3t^2 - t + 3$.

(b) $y'(x) = -x + \sqrt{2}$.

(c) $y' = |x|$.

4.13. **Motion of a particle.** The velocity of a particle is known to depend on time according to the relationship

$$v(t) = A - Bt^2, \quad A, B > 0 \text{ constants}$$

(a) Find the acceleration $a(t)$.

(b) Suppose that the initial position of the particle is $y(0) = 0$. Find the position at time t.

(c) At what time does the particle return to the origin?

(d) When is the particle farthest away from the origin?

(e) What is the largest velocity of the particle?

4.14. **Motion of a particle.** The position of a particle is given by the function $y = f(t) = t^3 + 3t^2$.

(a) Find the velocity and acceleration of the particle.

(b) A second particle has position given by the function $y = g(t) = at^4 + t^3$ where a is some constant and $a > 0$. At what time(s) are the particles in the same position?

(c) At what times do the particles have the same velocity?

(d) When do the particles have the same acceleration?

4.15. **Ball thrown from a tower.** A ball is thrown from a tower of height h_0. The height of the ball at time t is

$$h(t) = h_0 + v_0 t - (1/2)gt^2$$

where h_0, v_0, g are positive constants.

(a) When does the ball reach its highest point?

(b) How high is it at that point?

(c) What is the instantaneous velocity of the ball at its highest point ?

4.16. **Sketching a graph.** Sketch the graph of a function $f(x)$ whose derivative is shown in Figure 4.6. Is there only one way to draw this sketch? What difference might occur between the sketches drawn by two different students?

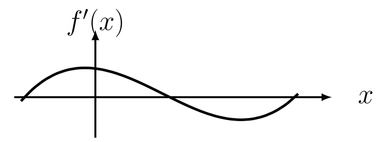


Figure 4.6: Figure for Exercise 16; sketching an antidertivative.

4.17. **Sketching the second derivative.** Given the derivative $f'(x)$ shown in Figure 4.7, graph the second derivative $f''(x)$.

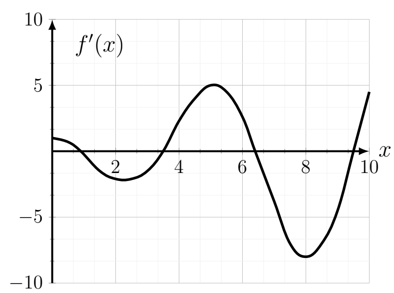
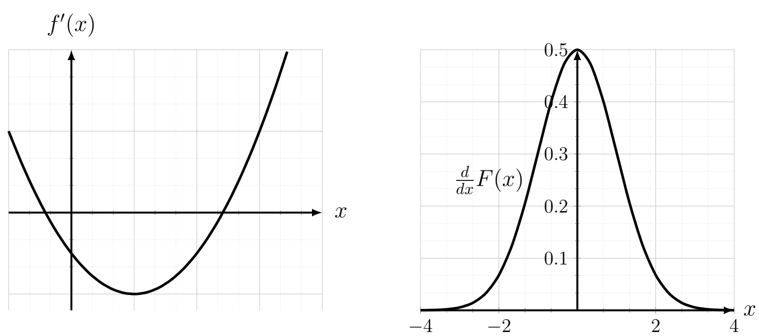


Figure 4.7: Figure for Exercise 17; sketching a derivative

4.18. **Sketching graphs.** Each of the graphs in Figure 4.8 depict the derivative of a function. Use these to sketch the corresponding antiderivatives and derivatives.



(a) (b)

Figure 4.8: Derivative graphs for Exercise 18

5

Tangent lines, linear approximation, and Newton's method

In Chapter 3, we defined the **tangent line** as the line we see when we zoom into the graph of a (continuous) function, $y = f(x)$, at some point. In much the same sense, the tangent line approximates the local behaviour of a function near the **point of tangency**., x_0. Given $x_0, y_0 = f(x_0)$, and the slope $m = f'(x_0)$ (the derivative), we can find the equation of the tangent line

$$\frac{\text{rise}}{\text{run}} = \frac{y - y_0}{x - x_0} = m = f'(x_0)$$

$$\Rightarrow \quad y = f(x_0) + f'(x_0)(x - x_0). \quad (5.1)$$

(See Appendix A for a review of straight lines.)

We use Eqn. (5.1) in several applications, including **linear approximation**, a method for estimating the value of a function near the point of tangency. A further application of the tangent line is **Newton's method** which locates zeros of a function (values of x for which $f(x) = 0$).

Mastered Material Check

1. What are the slope and y-intercept of the generic tangent line (given in Eqn. (5.1))?

5.1 The equation of a tangent line

Section 5.1 Learning goals

1. Given a simple function $y = f(x)$ and a point x, be able to find the equation of the tangent line to the graph at that point.
2. Graph both a function and its tangent line using a spreadsheet or your favorite software.

In the following examples, the equation of the tangent line is easily found.

Example 5.1 (Tangent to a parabola) *Find the equations of the tangent lines to the parabola $y = f(x) = x^2$ at the points:*

a) $x = 1$ *and* $x = 2$ *("Line 1" and "Line 2").*

b) *Determine whether these tangent lines intersect, and if so, where.*

Solution. The slopes of a tangent line is a derivative, which in this case is $f'(x) = 2x$.

a) This means $m_1 = f'(1) = 2 \cdot 1 = 2$ (for Line 1) and $m_2 = f'(2) = 2 \cdot 2 = 4$ (for Line 2). The points of tangency are on the curve (x, x^2). Thus these are $(1,1)$ for Line 1 and $(2,4)$ for Line 2. With the slope and a point for each line, we find that

$$\text{Line 1: } \frac{y-1}{x-1} = m_1 = 2, \quad \Rightarrow \quad y = 1 + 2(x-1) \quad \Rightarrow \quad y = 2x - 1,$$
$$\text{Line 2: } \frac{y-4}{x-2} = m_2 = 4 \quad \Rightarrow \quad y = 4 + 4(x-2) \quad \Rightarrow \quad y = 4x - 4.$$

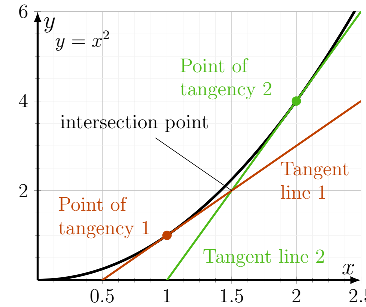


Figure 5.1: The graph of the parabola $y = f(x) = x^2$ and its tangent lines at $x = 1$ and $x = 2$. See Example 5.1 for the equations and point of intersection of these tangent lines.

b) Two lines intersect if their y values (and x values) are the same. Equating y values and solving for x, we get

$$2x - 1 = 4x - 4 \quad \Rightarrow \quad -2x = -3 \quad \Rightarrow \quad x = \frac{3}{2}.$$

Hence, the two tangent lines intersect at $x = 3/2$ as shown in Fig 5.1. ◊

Mastered Material Check

2. Under what circumstance do two lines *not* intersect?

The next example illustrates how a tangent line can be used to approximate the zero of a function. This idea is developed into a useful approximation method called **Newton's method** in Section 5.4.

Example 5.2 a) *Draw the graph of the function $y = f(x) = x^3 - x$ together with its tangent line at the point $x = 1.5$.*

b) *Where does that tangent line intersect the x-axis?*

c) *Compare that point of intersection with a zero of the function.*

Manipulate the slider to see the tangent line at various points on the graph of this function. Here $D(x)$ represents the derivative, $D(x) = df/dx$, and x_0 is the point of tangency. You can zoom in or out, change the range of the slider, or try a different function.

Solution.

a) The function is $f(x) = x^3 - x$, its derivative is $f'(x) = 3x^2 - 1$, and the point of interest is $(x, f(x)) = (1.5, 1.875)$. A tangent line at $x = 1.5$ has slope $f'(1.5) = 3(1.5)^2 - 1 = 5.75$, so its equation is

$$\frac{y - 1.875}{x - 1.5} = 5.75 \qquad \Rightarrow \qquad y = 1.875 + 5.75(x - 1.5)$$
$$\Rightarrow \qquad y = 5.75x - 6.75.$$

The function and this tangent line are shown in Figure 5.2.

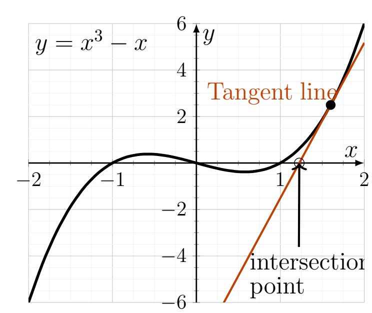


Figure 5.2: The graph of the function $y = f(x) = x^3 - x$ is shown in black, together with its tangent line at the point $x = 1.5$. The point at which the TL intersects the x-axis is a (rough) approximation of a nearby zero of the function.

b) The tangent line intersects the x-axis when $y = 0$, which occurs at

$$0 = 5.75x - 6.75 \quad \Rightarrow \quad x = \frac{6.75}{5.75} = 1.174.$$

c) A true zero of the function is a value of x for which $x^3 - x = 0$, and the one closest to the result in (b) is $x = 1$. Thus the point $(1.174, 0)$ where the tangent line intersects the x-axis is close but not quite equal to $(1,0)$. (But Newton's method will help us fix this gap.) ◊

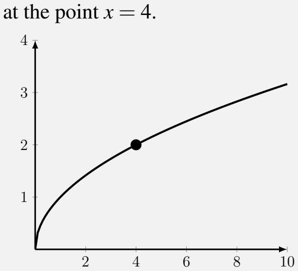
In many cases, directly solving for roots of functions (as in Example 5.2c) is not possible. In Section 5.4 we discuss how a repetition of this idea can be used to refine the approximation of a zero using **Newton's method**.

Example 5.3 **a)** *Find the equation of the tangent line to* $y = f(x) = x^3 - \rho x$ *for* $\rho > 0$ *constant, at the point* $x = 1$.

b) *Find where that tangent line intersects the x-axis.*

Mastered Material Check

3. How many zeros does the function depicted in Figure 5.2 have?
4. Check that the tangent line goes through the desired point and has the slope we found. One way to do this is to pick a simple value for ρ, e.g. $\rho = 1$ and do a quick check that the answer matches what we have found.
5. The following graph depicts $f(x) = \sqrt{x}$ on the interval $[0, 10]$. Draw its tangent line at the point $x = 4$.



Solution. This is the same type of calculation, but the constant, ρ - chosen to develop familiarity with alternative constant choices - makes the example slightly less straightforward.

a) The derivative of $f(x) = x^3 - \rho x$ is $f'(x) = 3x^2 - \rho$ so at $x = 1$, the slope is $m = f'(1) = 3 - \rho$. The point of tangency is $(1, f(1)) = (1, 1-\rho)$. Then, the equation of the tangent line is

$$\frac{y - (1-\rho)}{x - 1} = 3 - \rho. \quad \Rightarrow \quad y = (3-\rho)(x-1) + (1-\rho)$$
$$\Rightarrow \quad y = (3-\rho)x - 2.$$

b) To find the point of intersection, set $y = (3-\rho)x - 2 = 0$ and solve for x to obtain $x = 2/(3-\rho)$.

◊

Example 5.4 *Find the equation of the tangent line to the function* $y = f(x) = \sqrt{x}$ *at the point* $x = 4$.

Solution. By Eqn. (3.5), the derivative of $y = f(x) = \sqrt{x}$ is

$$f'(x) = \frac{1}{(2\sqrt{x})}.$$

At $x = 4$, the slope is $f'(4) = 1/(2\sqrt{4}) = 1/4$ and the point of tangency is $(4, \sqrt{4}) = (4, 2)$. Given this point and the slope, we calculate that the tangent line is:

$$\frac{y-2}{x-4} = 0.25 \quad \Rightarrow \quad y = 2 + 0.25(x-4).$$

◊

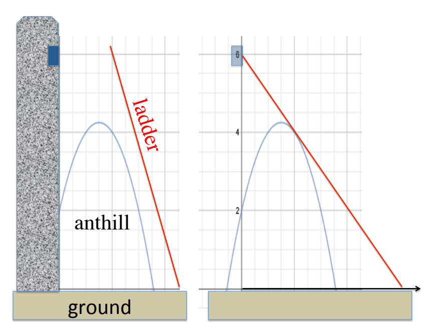


Figure 5.3: The ladder (red line) must reach the window at height 6m. What is the shortest ladder that avoids the anthill whose equation is $y = f(x) = -x^2 + 3x + 2$?

Featured Problem 5.1 (Shortest ladder) *In Fig. 5.3, what is the shortest ladder that you can use to reach the window at 6 meters height in the tower if there is an anthill in the way? Assume that the equation of the anthill is* $y = f(x) = -x^2 + 3x + 2$.

5.2 Generic tangent line equation and properties

Section 5.2 Learning goals

1. Explain the generic form of the tangent line equation (5.1) and be able to connect it to the geometry of the tangent line.
2. Find the coordinate of the point at which the tangent line intersects the x-axis (important for **Newton's Method** later on in Section 5.4).

Generic tangent line equation

We can find the general equation of a tangent line to an arbitrary function $f(x)$ at a point of tangency x_0. (The result is Eqn. (5.1).)

Shown in Figure 5.4 is a continuous function $y = f(x)$, assumed to be differentiable at some point x_0 where a tangent line is attached. We see:

1. The line goes through the point $(x_0, f(x_0))$.
2. The line has slope given by the derivative evaluated at x_0, that is, $m = f'(x_0)$.

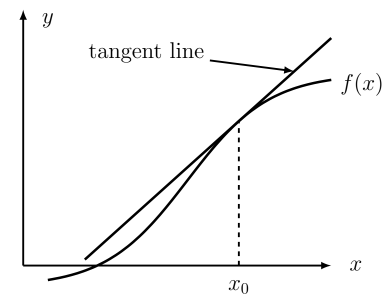


Figure 5.4: The graph of an arbitrary function $y = f(x)$ and a tangent line at $x = x_0$. The equation of this generic tangent line is Eqn. (5.2).

Mastered Material Check

6. Circle the point $(x_0, f(x_0))$ on Figure 5.4.
7. Circle where the tangent line depicted in Figure 5.4 would cross the x-axis.

Quick video with a derivation of the generic equation of a tangent line.

Then from the slope-point form of the equation of a straight line,

$$\frac{y - f(x_0)}{x - x_0} = m = f'(x_0).$$

Rearranging and eliminating the notation m, we have the desired result.

Summary, Tangent Line equation: The equation of a tangent line at $x = x_0$ to the graph of the differentiable function $y = f(x)$ is

$$y = f(x_0) + f'(x_0)(x - x_0). \quad (5.2)$$

Where a generic tangent line intersects the x-axis

From the generic tangent line equation (5.2) we can determine the (generic) coordinate at which it intersects the x-axis. The result is key to **Newton's method** for approximating the **zeros** of a function, explored in Section 5.4.

Example 5.5 *Let $y = f(x)$ be a smooth function, differentiable at x_0, and suppose that Eqn. (5.2) is the equation of the tangent line to the curve at x_0. Where does this tangent line intersect the x-axis?*

The calculations for Example 5.5. We show how to find the coordinate x_1 where a tangent line intersects the x axis.

Solution. At the intersection with the x-axis, we have $y = 0$. Plugging this into Eqn. (5.1) leads to

$$0 = f(x_0) + f'(x_0)(x - x_0) \quad \Rightarrow \quad (x - x_0) = -\frac{f(x_0)}{f'(x_0)}$$
$$\Rightarrow \quad x = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Thus the desired x coordinate, which we refer to as x_1 is

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}. \tag{5.3}$$

◊

Mastered Material Check

8. Where is x_1 (as described in Eqn. (5.3)) on Figure 5.4.?

5.3 *Approximating a function by its tangent line*

Section 5.3 Learning goals

1. Describe why a tangent line approximates the behaviour of a function close to the point of tangency.
2. Use a tangent line to find a linear approximation to a value of a given function at some point.
3. Determine whether a linear approximation overestimates or underestimates the value of the function.

We have seen that the tangent line approximates the local behaviour of a function, at least close enough to the point of tangency. Here we use this idea in a formal procedure called **linear approximation**. The idea is to chose a point (often called the **base point**) where the value of the function and its derivative are known, or are easy to calculate, and use the tangent line at that point to estimate values of the function in the vicinity. Specifically,

1. The generic equation of the tangent line to $y = f(x)$ at x_0 is given by Eqn. (5.2). That line approximates the behaviour of the function close to x_0, and leads to the so-called *linear approximation* of the function:

$$y = f(x_0) + f'(x_0)(x - x_0) \approx f(x)$$
$$\Rightarrow \quad f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

2. The approximation is exact at $x = x_0$, and holds well provided x is close to x_0. (The expression on the right hand side is precisely the value of y on the tangent line at $x = x_0$).

Example 5.6 *Use the fact that the derivative of the function* $f(x) = x^2$ *is* $f'(x) = 2x$ *to find a linear approximation for the value* $(10.03)^2$.

Solution. Without the aid of calculator, we know that value of $f(x)$ at the nearby point $x = 10$ is $10^2 = 100$. The derivative is $f'(x) = 2x$, so at $x = 10$ the slope of the tangent line is $f'(10) = 20$. The equation of the tangent line directly provides the linear approximation of the function.

$$\frac{y-100}{x-10} = 20 \quad \Rightarrow \quad y = 100 + 20(x-10),$$
$$\Rightarrow \quad f(x) \approx 100 + 20(x-10).$$

On the tangent line, the value of y corresponding to $x = 10.03$ is

$$f(10.03) \approx y = 100 + 20(10.03 - 10) = 100 + 20(0.03) = 100.6$$

which is our approximation to the value of the original function. This compares well with the calculator value $f(10.03)$ =100.6009. $\Diamond$

Mastered Material Check

9. Use the linear approximation for x^2 found in Example 5.6 to approximate $(9.97)^2$.
10. Use the linear approximation for $\sin(x)$ found in Example 5.7 to approximate $\sin(-0.05)$.
11. Can you think of an example of a function whose linear approximation is exact for all values?
12. Do you know what concave means?

Example 5.7 (Approximating the sine of a small angle) *Use a linear approximation to find a rough value for* $\sin(0.1)$.

Solution. We have not yet discussed finding derivatives of trigonometric functions, but recall from Example 3.2 that close to $x = 0$ the function $y = \sin(x)$ is well approximated by its tangent line, $y = x$. Hence, the linear approximation of $y = \sin(x)$ near $x = 0$ *is* $\sin(x) \approx x$ (provided x is in **radians**, to be discussed in Chapter 14). Thus, at $x = 0.1$ radians, we find that $\sin(0.1) \approx 0.1$, close to its true value of $\sin(0.1) = 0.09983$ (found using a calculator). $\Diamond$

Accuracy of the linear approximation

Example 5.8 (Over or underestimate?) *For each of Examples 5.6 and 5.7, determine whether the linear approximation over or underestimates the true value of the function.*

Solution. In Figure 5.5(a,b), we show the functions and their linear approximations. In (a) we see that the tangent line to $y = x^2$ at $x = 10$ is always underneath the graph of the function, so a linear approximation underestimates the true value of the function.

In (b), we see that the tangent line to $y = \sin(x)$ at $x = 0$ is above the graph for $x > 0$ and below the graph for $x < 0$. This shows that the linear approximation is larger than (overestimates) the function for $x > 0$ and smaller than (underestimates) the function for $x < 0$. $\Diamond$

In Chapter 6, we associate these properties with the **concavity** of the function, that is, whether the graph is locally concave up or down.

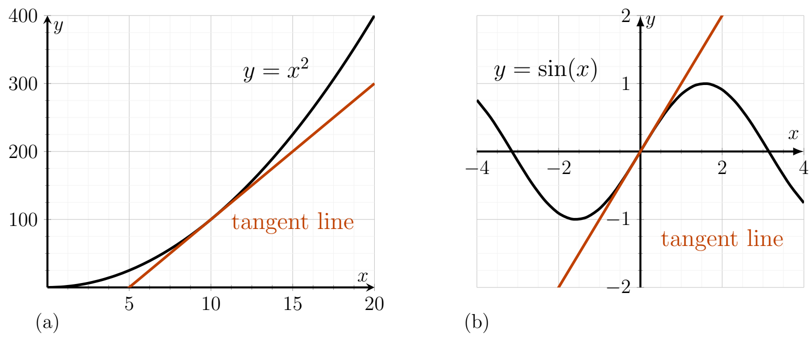


Figure 5.5: Functions (black curves) and their linear approximations (red) for Examples 5.6 and 5.7. Whenever the tangent line is below (above) the curve, we say that the linear approximation under (over)-estimates the value of the function.

Calculations for Example 5.9.

Example 5.9 **a)** *Use linear approximation to estimate the value of* $\sqrt{6}$.

b) *Determine whether the linear approximation underestimates or overestimates the function.*

Solution. The derivative of $y = f(x) = \sqrt{x} = x^{1/2}$ is

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

Both the function and its derivative require evaluation of a square root. Some numbers ("**perfect squares**") have convenient square roots. One such number, $x = 4$, is nearby, so we use it as the "base point" for a linear approximation.

Mastered Material Check

13. A perfect square is an integer m of the form $m = n^2$ where n is also an integer. List several different perfect squares.

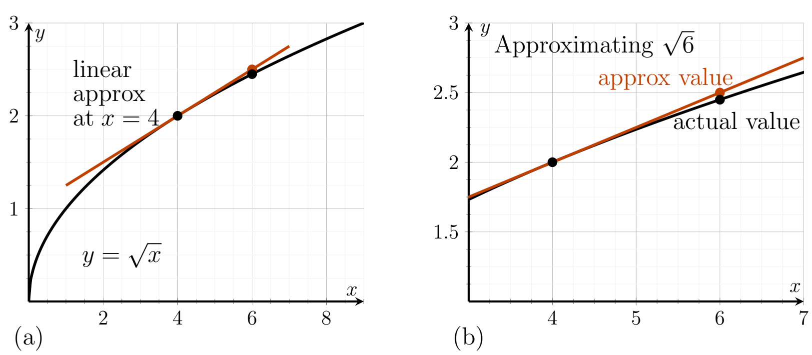


Figure 5.6: Linear approximation based at $x = 4$ to the function $y = f(x) = \sqrt{(x)}$. The panel on the right is a magnified view.

a) The slope of the tangent line at $x = 4$ is $f'(4) = 1/(2\sqrt{4}) = 1/4 = 0.25$ so the linear approximation of $\sqrt{6}$ is obtained as

$$y = f(4) + f'(4)(x-4) \quad \Rightarrow \quad y = 2 + 0.25(x-4)$$
$$\Rightarrow \quad \sqrt{6} \approx 2 + 0.25(6-4) = 2.5.$$

b) A graph of the function and its tangent line in Figure 5.6(a) and a zoomed portion in Figure 5.6(b) compares the true and approximated values of $\sqrt{6}$.

The tangent line is above the graph of the function, so the linear approximation *overestimates* values of the function. ◊

The discrepancy between true and approximated values is called the **error**. *The closer we are to the base point, the smaller the error in the approximation*. This is demonstrated by comparing the values in Table 5.1, computed using a spreadsheet with base point $x_0 = 4$.

x	**exact value** $f(x) = \sqrt{x}$	**approx. value** $y = f(x_0) + f'(x_0)(x - x_0)$
0.0000	0.0000	1.0000
2.0000	1.4142	1.5000
4.0000	2.0000	2.0000
6.0000	2.4495	2.5000
8.0000	2.8284	3.0000
10.0000	3.1623	3.5000
12.0000	3.4641	4.0000
14.0000	3.7417	4.5000
16.0000	4.0000	5.0000

Table 5.1: Linear approximation to $\sqrt{x}$ at the base point $x = 4$. The exact value is recorded in column 2 and the linear approximation in column 3. The approximation is reasonably good close to the base point.

Mastered Material Check

14. Find the error (true value minus approximate value) in the linear approximation for $\sqrt{6}$ found in Example 5.9 .
15. Determine the error of the approximation for $\sqrt{8}$ as given in Table 5.1.
16. Which approximation has larger error? Was this to be expected or not?

5.4 *Tangent lines for finding zeros of a function: Newton's method*

Section 5.4 Learning goals

1. Describe the geometry on which Newton's method is based (Figure 5.7).
2. Given $f(x)$ and initial guess x_0, use Newton's method to find improved values x_1, x_2, etc., for the zero of $f(x)$ (value of x such that $f(x) = 0$).
3. For a given function $f(x)$, select a suitable initial guess x_0 for Newton's method from which to start the iteration.

Definition 5.1 (Zero) *Given a function $f(x)$, we say that x^* is a* **zero** *of f if $f(x^*) = 0$. We also say that "x^* is a* **root** *of the equation $f(x) = 0$".*

In many cases, it is impossible to compute a value of a zero, x^* analytically. Based on tangent line approximations, we now explore **Newton's method**, an approximation that does the job.

Newton's method

Consider the function $y = f(x)$ shown in Figure 5.7. We have already found that a tangent line approximates the behaviour of a function close to a point of tangency. It can also be used to build up and refine an approximation of the zeros of the function.

Our goal is to find a decimal approximation for the value x such that $f(x) = 0$. (In Fig. 5.7, this value is denoted by x^*.) Newton's method is an **iterated scheme** (a procedure that gets repeated). It is applied several times, to generate a decimal expansion of the desired zero to any level of accuracy.

A starting value, x_0, initiates the method. This can be a rough first guess for the zero we seek, found graphically, for example. Newton's method, applied a number of times, will generate better and better approximations of the true zero, x^*.

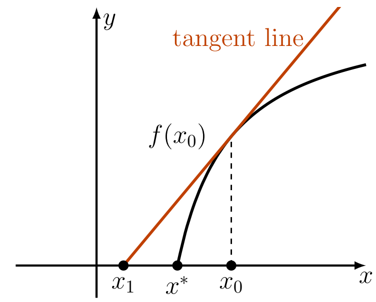


Figure 5.7: In Newton's method, we seek a decimal approximation for x^*, a zero of $y = f(x)$. A rough initial guess, x_0, is refined by "sliding down the tangent line" (glued to the curve at x_0). This brings us to an improved guess x_1. Repeating this again and again allows us to find the root to any desired accuracy.

Gluing a tangent line at x_0, we follow it down to its x-axis intersection. In a previous section, we have already computed that intersection point in (5.3), to be

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Then usually x_1 is closer to x^*, improving upon the initial guess.

Now use x_1 as the (improved) guess, and repeat the process. This generates values

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}.$$

$$\vdots$$

The values x_2, x_3 rapidly approach the desired root x^*. The same 'recipe' is repeated. In practice, when it works, Newton's method **converges** quite rapidly, that is, it approaches the root with excellent accuracy, after very few repetitions (**iterations**). To summarize,

Newton's method: Given an approximation x_k for the root of the equation $f(x) = 0$, we can improve the accuracy of that approximation with another iteration using

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Example 5.10 *Find zeros of the function $y = f(x) = x^3 - x - 3$ starting with the initial guess $x_0 = 1$.*

Newton's Method applied to the function $f(x) = x^3 - x - 3$.
Move the slider or the initial point x_0 to see how the first approximation x_1 and the second approximation x_2 change.
Compute the approximation x_3 on the same graph.

Solution. The derivative is $f'(x) = 3x^2 - 1$, so Newton's method for the improved guess is

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = x_0 - \frac{x_0^3 - x_0 - 3}{3x_0^2 - 1} = 1 - \frac{1^3 - 1 - 3}{3 \cdot 1^2 - 1}.$$

So, starting with $x_0 = 1$, we obtain

$$x_1 = 1.727272727, x_2 = 1.673691174, x_3 = 1.67170257, x_4 = 1.671699882.$$

The iterates converge to the result $x \approx 1.6717$. ◊

Mastered Material Check

15. Perform the calculations to verify the x_1, x_2, x_3 and x_4 found in Example 5.10.

Example 5.11 *Use Newton's method to find a decimal approximation of the square root of* 6.

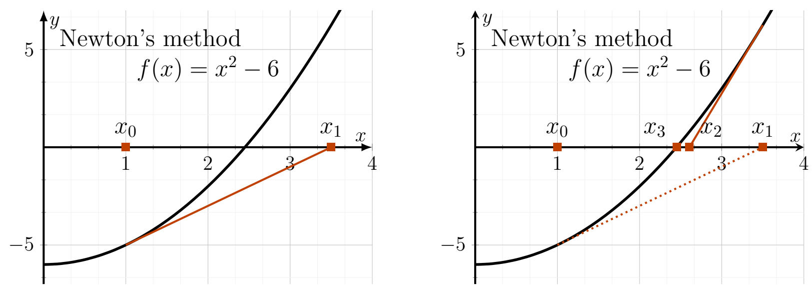


Figure 5.8: Newton's method applied to solving $y = f(x) = x^2 - 6 = 0$.

Solution. Because Newton's method finds zeros of a function, it is first necessary to restate the problem in the form "find a value of x such that a certain function $f(x) = 0$." Clearly, one function that would accomplish this is

$$f(x) = x^2 - 6$$

since $f(x) = 0$ corresponds to $x^2 - 6 = 0$, i.e. $x = \sqrt{6}$. Then the derivative of $f(x)$ is $f'(x) = 2x$, and the recipe to repeat is

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = x_0 - \frac{x_0^2 - 6}{2x_0}.$$

Starting with an initial guess $x_0 = 1$ (not very close to the value of the root), we show in Figure 5.8 how Newton's method applies a tangent line to determination x_1. In the right panel, we see how the value of x_1 is then used to obtain x_2 by repeating the calculation.

Mastered Material Check

16. Note that other functions have this property. For example, verify that a root of $f(x) = x^4 - 36$ is also $\sqrt{6}$.
17. Give another example of a function $f(x)$ for which $f(x) = 0$ has the root $\sqrt{6}$.

A spreadsheet is ideal for carrying out the repetitive calculations, as shown in Table 5.2. For example, we compute the following set of values using our spreadsheet. Observe that the fourth column contains the computed (Newton's method) values, x_1, x_2, etc. These values are then copied onto the first column to be used as new "initial guesses". After several repetitions, the numbers calculated converge to 2.4495, and no longer change to that level of accuracy displayed. This signals that we have obtained the root to 5 significant figures of accuracy. ◊

Link to Google Sheets. This spreadsheet implements Newton's method for Example 5.11. You can view the formulae by clicking on a cell in the sheet but you cannot edit the sheet here.

k	x_k	$f(x_k)$	$f'(x_k)$	x_{k+1}
0	1.00	-5.00	2.00	3.5
1	3.5	6.250	7.00	2.6071
2	2.6071	0.7972	5.2143	2.4543
3	2.4543	0.0234	4.9085	2.4495
4	2.4495	0.000	4.8990	2.4495

Table 5.2: Newton's method applied to Example 5.11. We start with $x_0 = 1$ as our initial approximation and refine it four times.

Note: it is possible that Newton's method fails to find a root - something we do not explore further here. This might happen if our initial guess is too poor.

5.5 *Aphids and Ladybugs, revisited*

In Example 1.3, we asked when predation (by a ladybug) and growth rate exactly match for an aphid population. We did so by solving an equation of the form $P(x) = G(x)$ for x the aphid density and $G(x) = rx$ the aphid growth rate ($r > 0$), and $P(x)$ the rate of predation of aphids by a ladybug. Our solution relied on the quadratic formula. Now consider the case that the predation rate is

$$P(x) = K\frac{x^3}{a^3 + x^3}, \quad \text{where} \quad K, a > 0. \tag{5.4}$$

In this case, steps shown in Example 1.3 lead to a cubic equation for x, which is not easy to solve by pen and paper. This is a classic situation where Newton's method proves useful.

Why do we need to use Newton's Method if we already solved the aphid -ladybug predation problem in Section 1.6.1?

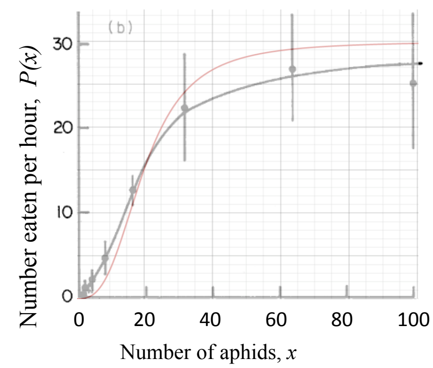


Figure 5.9: The function $P(x)$ superimposed on a graph of data from [Hassell et al., 1977].

Example 5.12 (Using Newton's Method to solve the aphid-ladybug problem) *Set up the problem for obtaining the number of aphids at which predation by a ladybug and population growth of aphids balance. Convert the equation to a form for which Newton's method is suitable. Then use Newton's method to solve your problem. Assume that $K = 30$ aphids eaten per hour, $a = 20$ aphids, and $r = 0.5$ per hour. To get a reasonable initial guess, plot $P(x)$ and $G(x)$ on the same graph and determine roughly where they intersect.*

Solution. The problem to be solved (assuming $x \neq 0$ is

$$P(x) = G(x) \quad \Rightarrow \quad K\frac{x^3}{a^3 + x^3} = rx \quad \Rightarrow \quad K\frac{x^2}{a^3 + x^3} = r$$

simplifying algebraically leads to the equation

$$Kx^2 = r(a^3 + x^3) \quad \Rightarrow \quad rx^3 - Kx^2 + ra^3 \equiv f(x) = 0.$$

Examine this graph of the predation rate $P(x)$ and the population growth rate $G(x)$ to find reasonable initial guess(es) for points of intersection. (We look only for positive values, since x represents the number of aphids.) These will be used as value(s) for x_0 in Newton's method.

Having converted the problem into the form $f(x) = 0$, we can apply Newton's method. We need the function and its derivative for Newton's method formula,

$$f(x) = rx^3 - Kx^2 + ra^3$$
$$f'(x) = 3rx^2 - 2Kx$$

Using the numerical values for the constants, and examining the graph of the two functions $P(x)$ and $G(x)$, we find intersections at $x = 0$ and $x_0 \approx 10$. There is another intersection at $x_0 \approx 60$. To implement the method, we apply Newton's formula,

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = x_0 - \frac{(rx_0^3 - Kx_0^2 + ra^3)}{3rx_0^2 - 2Kx_0}$$

with $x_0 = 10$. Table 5.3 summarizes the convergence to the root $x = 13.05407289$ after four rounds of improvement using Newton's method.

k	x_k	$f(x_k)$	$f'(x_k)$	x_{k+1}
0	10	1500.00	-450.00	13.33333333
1	13.33333333	-148.15	-533.33	13.05555556
2	13.05555556	-0.78	-527.66	13.05407294
3	13.05407294	0.00	-527.63	13.05407294
4	13.05407289	0.00	-527.63	13.05407294

Table 5.3: Newton's method applied to Example 5.12. We start with $x_0 = 10$ as the initial approximation.

Since x measures a density of aphids (e.g. average number per cm^2), it is reasonable to find a real valued solution (rather than an integer "number of aphids").

We conclude that at this aphid density, the aphid population would have a growth and a predation rate that exactly match. (Hence, we also expect that the aphid population would neither increase nor decrease.) What happens if growth and predation rates *do not match?* In such cases, we expect change to take place. How to analyze such situations will be the topic of a later chapter.

A short video discussion of the Aphid-ladybug problem and how to use Newton's Method to solve it on a spreadsheet, as discussed in Example 5.12.

5.6 *Harder problems: finding the point of tangency*

Section 5.4 Learning goals

1. Find a tangent line to a function that goes through some point (not necessarily on the graph of the function).
2. Determine tangent lines to functions that contain unspecified parameters.

In this section, we present a sample of problems in which the path to a solution is more subtle. In some of these, finding the point of tangency is part of the question. We must use clues about the function to solve for that point, as well as construct the tangent line equation from information supplied. In other cases, the problem involves a parameter whose value is not specified initially. Such examples are meant to hone problem solving skills.

Example 5.13 *Find any value(s) of the constant a such that the line $y = ax$ is tangent to the curve $y = f(x) = -x^2 + 3x - 2$.*

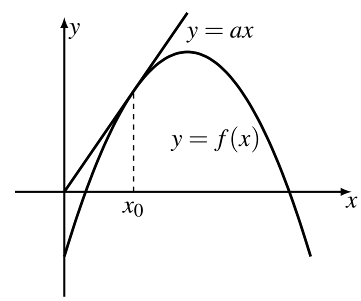


Figure 5.10: Figure for Example 5.13: finding the point of tangency.

Mastered Material Check

18. Factor $f(x) = -x^2 + 3x - 2$. Does the shape depicted in Figure 5.10 make sense?

Solution. We do not know the coordinate of any such point, but we label it x_0 in Figure 5.10 to denote that it is a definite (as yet to be determined) value. Finding x_0 is part of the problem. We collect the information to be used:

- The tangent line $y = ax$ intersects the graph of the function $y = f(x) = -x^2 + 3x - 2$ at $(x_0, f(x_0))$.
- Equating y values of the tangent line and the curve $y = f(x)$ at x_0 we get:

$$f(x_0) = -x_0^2 + 3x_0 - 2 = ax_0.$$

- The equation of the tangent line is $y = ax$. Its slope is a, which is also the derivative of $f(x)$ at x_0. Equating slopes gives:

$$f'(x_0) = -2x_0 + 3 = a.$$

We have two equations with two unknowns, (a and x_0). We can solve this system by substituting the value of a from the first equation into the second, getting

$$-x_0^2 + 3x_0 - 2 = (-2x_0 + 3)x_0.$$

Simplifying:

$$-x_0^2+3x_0-2=-2x_0^2+3x_0 \quad \Rightarrow \quad x_0^2-2=0, \quad x_0=\pm\sqrt{2}.$$

Thus, there are two possible points of tangency, as shown in Figure 5.11.

Finally, we find a using $a=-2x_0+3$. We get:

$$x_0=\sqrt{2} \quad \Rightarrow \quad a=-2\sqrt{2}+3, \quad \text{and} \quad x_0=-\sqrt{2} \quad \Rightarrow \quad a=2\sqrt{2}+3.$$

◊

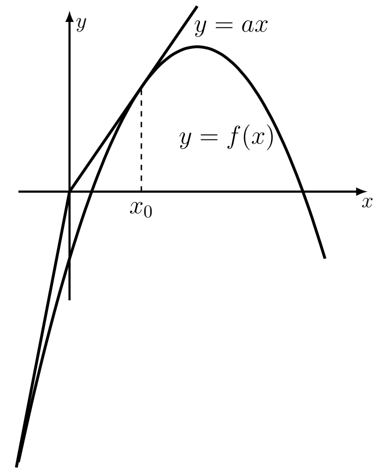


Figure 5.11: Two points of tangency in the solution to Example 5.13

The solution to Example 5.13 was set up by

- listing of information provided,
- deducing a set of equations based on that information, and
- following a chain of reasoning to arrive at the final solution.

Practicing such multi-step problems is a critical part of training for many fields, including science, medicine, engineering, etc.

Example 5.14 *Find the equation of the tangent line to the curve* $y=f(x)=1-x^2$ *that goes through the point* $(1,1)$.

Solution. Finding the point of tangency x_0 is part of the problem. We use the following facts:

- The tangent line goes through the point $(x_0, f(x_0))$ on the graph of the function and has slope $f'(x_0)$.
- Consequently, its equation has the form Eqn. (5.2): $y=f(x_0)=f'(x_0)(x-x_0)$.

For the given function and point of tangency x_0, we have

$$f(x_0)=1-x_0^2, \quad f'(x_0)=-2x_0.$$

Hence the tangent line equation is

$$y=f(x_0)+f'(x_0)(x-x_0)=(1-x_0^2)-2x_0(x-x_0).$$

We are told that this line goes through the point $(x,y)=(1,1)$ so that

$$1=(1-x_0^2)-2x_0(1-x_0), \quad \Rightarrow \quad 0=x_0^2-2x_0, \quad \Rightarrow \quad x_0^2=2x_0.$$

Thus, there are two possible points of tangency, $x_0=0,2$ and two tangent lines that satisfy the given condition. Plugging in these two values of x_0 into the generic equation for y leads to the two tangent line equations

1. $y=1$, and
2. $y=(1-2^2)-2\cdot 2(x-2)=-3-4(x-2)$. ◊

We can also find points of tangency for functions that contain general constants, as the next example illustrates.

Mastered Material Check

19. Identify the slope of the line $y=(1-x_0^2)-2x_0(x-x_0)$.
20. Determine the y-intercept of the line $y=(1-x_0^2)-2x_0(x-x_0)$.
21. Check that both lines found as solutions to Example 5.14 go through $(1,1)$ as desired.

Example 5.15 *Shown in Figure 5.12 is the function*

$$f(x) = C\frac{x}{x+a}$$

together with one of its tangent lines. The tangent line goes through a point $(-d, 0)$. *Find the equation of the tangent line.*

Solution. Finding the point of tangency x_0 is part of the problem in this case too. We use the same approach, and employ facts (1) and (2) from Example 5.14. We also use, for the specific function in this example,

$$f(x_0) = C\frac{x_0}{x_0+a} \quad \Rightarrow \quad f'(x_0) = C\frac{a}{(x_0+a)^2}.$$

(See Exercise 10 in Chapter 3). Hence, the equation of the tangent line is

$$y = f(x_0) + f'(x_0)(x - x_0) = C\frac{x_0}{x_0+a} + C\frac{a}{(x_0+a)^2}(x - x_0).$$

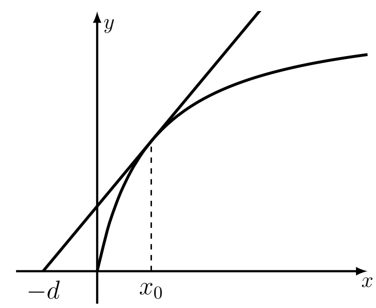


Figure 5.12: The graph of a function and its tangent line for Example 5.15.

We can simplify this equation by factoring to obtain:

$$y = \frac{C}{x_0+a}\left(x_0(x_0+a) + a(x - x_0)\right) = \frac{C}{x_0+a}\left(x_0^2 + ax\right).$$

It is important to realize that in this equation, x_0, C and a represent fixed (known) constants, and only x, y are variables. This means that the equation expresses a linear relationship between x and y, as appropriate for a straight line.

We know that the point $(-d, 0)$ is on this line, so that (plugging in $x = -d, y = 0$), we obtain

$$0 = \frac{C}{x_0+a}\left(x_0^2 - ad\right).$$

Solving for x_0 leads to $x_0 = \sqrt{ad}$. Moreover, we can now find the equation of the tangent line in terms of these parameters.

$$y = \frac{C}{\sqrt{ad}+a}(ad + ax) \quad \Rightarrow \quad y = \frac{C}{\sqrt{(d/a)}+1}(d + x)$$

where we have simplified by factoring a from numerator and denominator. We can easily see that when $x = -d$, we get $y = 0$, as required. This forms one check that our calculations are correct. ◊

5.7 Summary

1. The equation of a tangent line to $f(x)$ at x_0 is given by

$$y = f(x_0) + f'(x_0)(x - x_0).$$

2. If $L(x)$ is the tangent line to a function $f(x)$ at x_0, then $L(x)$ forms a linear approximation to $f(x)$ near the point x_0.

3. In some circumstances, the zero of a tangent line to a function $f(x)$ at a point x_0 can form an initial approximation to the zero of $f(x)$.
4. Newton's method is based on the property of tangent lines. Newton's method can solve a problem of the form $f(x) = 0$. Given an initial guess x_0, the method generates successive decimal approximations to the zeros of the function to *any* desired accuracy. The iteration scheme is:

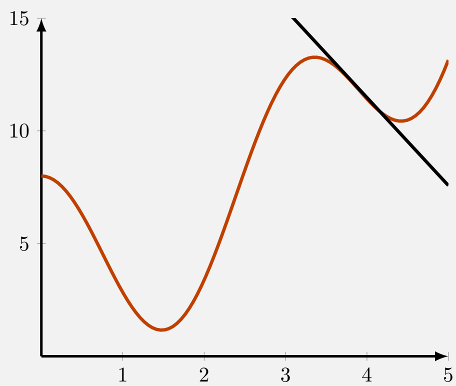
$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Quick Concept Checks

1. Is it possible for two different tangent lines of the same function to be parallel?

2. When would a tangent line **not** intersect the x-axis?

3. Consider the graph of the following function, and its tangent line at $x = 1$.



   (a) When would the linear (tangent line) approximation result in an overestimate? Under-estimate?

   (b) What is a reasonable interval on which to use this tangent line for approximation?

4. Why might Newton's method **not** work?

Exercises

5.1. **Equation of tangent line.** Find the equation of the tangent line to the function $y = f(x) = |x+1|$ at:

(a) $x = -1$,

(b) $x = -2$,

(c) $x = 0$.

If there is a problem finding a tangent line at one of these points, indicate what the problem is.

5.2. **Equation of tangent line.** A function $f(x)$ satisfies $f(1) = -1$ and $f'(1) = 2$. What is the equation of the tangent line of $f(x)$ at $x = 1$?

5.3. **Point of tangency.** Shown in Figure 5.13 is the graph of $y = x^2$ with one of its tangent lines.

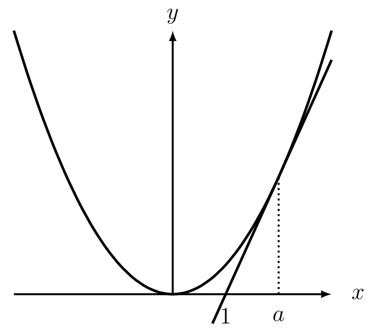


Figure 5.13: Figure for Exercise 3; $y = x^2$ and a point of tangency.

(a) Show that the slope of the tangent to the curve $y = x^2$ at the point $x = a$ is $2a$.

(b) Suppose that the tangent line intersects the x axis at the point $(1,0)$. Find the coordinate, a, of the point of tangency.

5.4. **Approximation with a tangent line.** Shown in Figure 5.14 is the function $f(x) = 1/x^4$ together with its tangent line at $x = 1$.

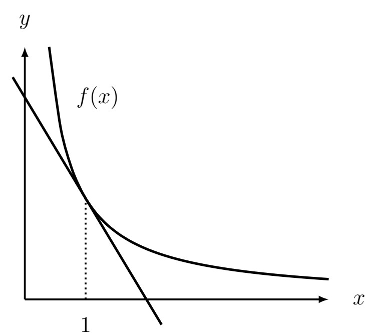


Figure 5.14: Figure for Exercise 4; $f(x) = \frac{1}{x^4}$ and a point of tangency.

(a) Find the equation of the tangent line.

(b) Determine the points of intersection of the tangent line with the x and the y-axes.

(c) Use the tangent line to obtain a linear approximation to the value of $f(1.1)$. Is this approximation larger or smaller than the actual value of the function at $x = 1.1$?

5.5. **Linear approximation.** Shown in Figure 5.15 is the function $f(x) = x^3$ with a tangent line at the point $(1,1)$.

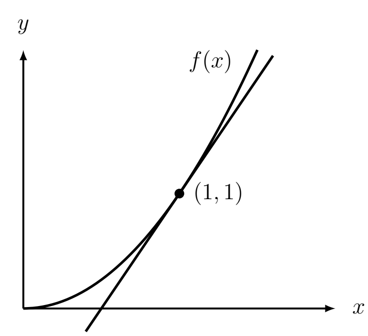


Figure 5.15: Figure for Exercise 5; $f(x) = x^3$ and a point of tangency.

(a) Find the equation of the tangent line.

(b) Determine the point at which the tangent line intersects the x axis.

(c) Compute the value of the function at $x = 1.1$. Compare this with the value of y on the tangent line at $x = 1.1$.

This latter value is the *linear approximation* of the function at the desired point based on its known value and known derivative at the nearby point $x = 1$.

5.6. **Generic tangent line.** Shown in Figure 5.16 is the graph of a function and its tangent line at the point x_0.

(a) Find the equation of the tangent line expressed in terms of x_0, $f(x_0)$ and $f'(x_0)$.

(b) Find the coordinate x_1 at which the tangent line intersects the x-axis.

5.7. **Estimating a square root.** Use Newton's method to find an approximate value for $\sqrt{8}$ (*hint:* first think of a function, $f(x)$, such that $f(x) = 0$ has the solution $x = \sqrt{8}$).

5.8. **Finding points of intersection.** Find the point(s) of intersection of: $y_1 = 8x^3 - 10x^2 + x + 2$ and $y_2 = x^3 + 15x^2 - x - 4$ (*hint:* an intersection point exists between $x = 3$ and $x = 4$).

5.9. **Roots of cubic equations.** Find the roots for each of the following cubic equations using Newton's method:

(a) $x^3 + 3x - 1 = 0$

(b) $x^3 + x^2 + x - 2 = 0$

(c) $x^3 + 5x^2 - 2 = 0$ (*hint:* find an approximation to a first root a using Newton's method, then divide the left hand side of the equation by $(x - a)$ to obtain a quadratic equation, which can be solved by the quadratic formula).

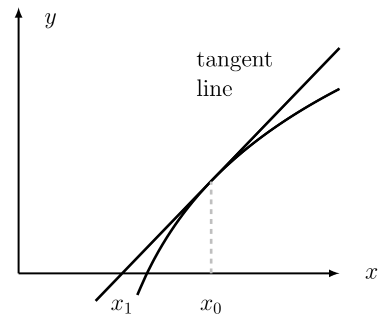


Figure 5.16: Figure for Exercise 6; generic function and tangent line.

5.10. **Intersecting tangent lines.** The parabola $y = x^2$ has two tangent lines that intersect at the point $(2,3)$. These are shown as the dark lines in Figure 5.17. Find the coordinates of the two points at which the lines are tangent to the parabola.

Note: note that the point $(2,3)$ is not on the parabola.

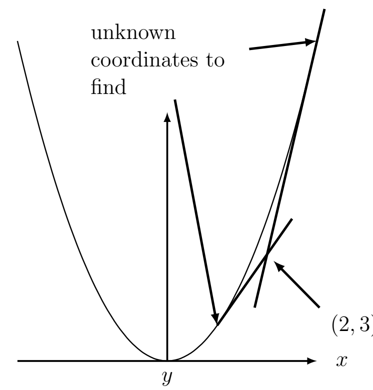


Figure 5.17: Figure for Exercise 10; $y = x^2$

5.11. **An approximation for the square root.** Use a linear approximation to find a rough estimate of the following functions at the indicated points.

(a) $y = \sqrt{x}$ at $x = 10$ (use the fact that $\sqrt{9} = 3$).

(b) $y = 5x - 2$ at $x = 1$.

5.12. **An approximation for the cube root.** Use the method of linear approximation to find the cube root of

(a) 0.065 (*hint:* $\sqrt[3]{0.064} = 0.4$)

(b) 215 (*hint:* $\sqrt[3]{216} = 6$)

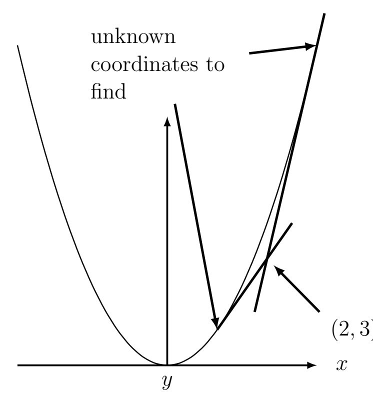


Figure 5.18: Figure for Exercise 13; using a graph to approximate.

5.13. **Approximating from a graph.** Use the data in the graph in Figure 5.18 to make the best approximation you can to $f(2.01)$.

5.14. **Linear approximation.** Approximate the value of $f(x) = x^3 - 2x^2 + 3x - 5$ at $x = 1.001$ using the method of linear approximation.

5.15. **Approximating cube volume.** Approximate the volume of a cube whose length of each side is 10.1 cm.

5.16. **Using Newton's method to find a critical point.** Consider the function

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

Critical points of a function are defined as values of x for which $g'(x) = 0$. However, for this fifth order polynomial, it is not easy to find such points analytically (i.e., using pencil and paper).

(a) Use Newton's Method to find a critical point for positive values of x. Find an initial approximation for the critical point by plotting the function, but use a spreadsheet and explain how you set up the calculations. Provide an answer accurate to 8 decimal points.

(b) Explain why a starting value of $x_0 = 1$ for Newton's Method does not lead to the positive critical point. You may support your argument with a graph.

Note: in Section 6.2 we study critical points in greater depth.

6

Sketching the graph of a function using calculus tools

The derivative of a function contains important information about the original function. In this chapter we focus on how properties of the first and second derivative can be used to help up refine curve-sketching techniques. The mathematics we develop in this chapter is used in a variety of applications, many found in Chapter 7.

6.1 Overall shape of the graph of a function

Section 6.1 Learning goals

1. Identify that the sign of the first derivative corresponds to the increasing or decreasing trend of a function.
2. Recognize that the sign of the second derivative corresponds to the concavity (curvature) of a function.

Increasing and decreasing functions

Consider a function given by $y = f(x)$. We first make the following observations:

1. If $f'(x) > 0$ then $f(x)$ is **increasing**.
2. If $f'(x) < 0$ then $f(x)$ is **decreasing**.

Mastered Material Check

1. Draw the graphs of two different functions that are increasing on the interval $[0,1]$.
2. Without using the derivative, how do you identify when a function is decreasing?

By convention, we read graphs from left to right, i.e. in the direction of the positive x-axis, so when we say "increasing" we mean that, as we move from left to right, the value of the function gets larger.

We can use the same ideas to relate the second derivative to the first derivative.

1. If $f''(x) > 0$ then $f'(x)$ is **increasing**. This means that the slope of the original function is getting steeper (more positive, from left to right). The function curves upwards: we say that it is **concave up**. See Figure 6.1(a).

2. If $f''(x) < 0$ then $f'(x)$ is **decreasing**. This means that the slope of the original function is getting shallower (more negative or less positive from left to right). The function curves downwards: we say that it is **concave down**. See Figure 6.1(b).

See video summarizing the connection between the shape of the graph of $y = f(x)$, and the derivatives of the function, $f'(x), f''(x)$ of the function.

Concavity and points of inflection

The second derivative of a function provides information about the **curvature** of the graph of the function, also called the **concavity** of the function.

- In Figure 6.1(a), $f(x)$ is **concave up**, so its second derivative is positive.
- In Figure 6.1(b), $f(x)$ is **concave down**, so second derivative is negative.

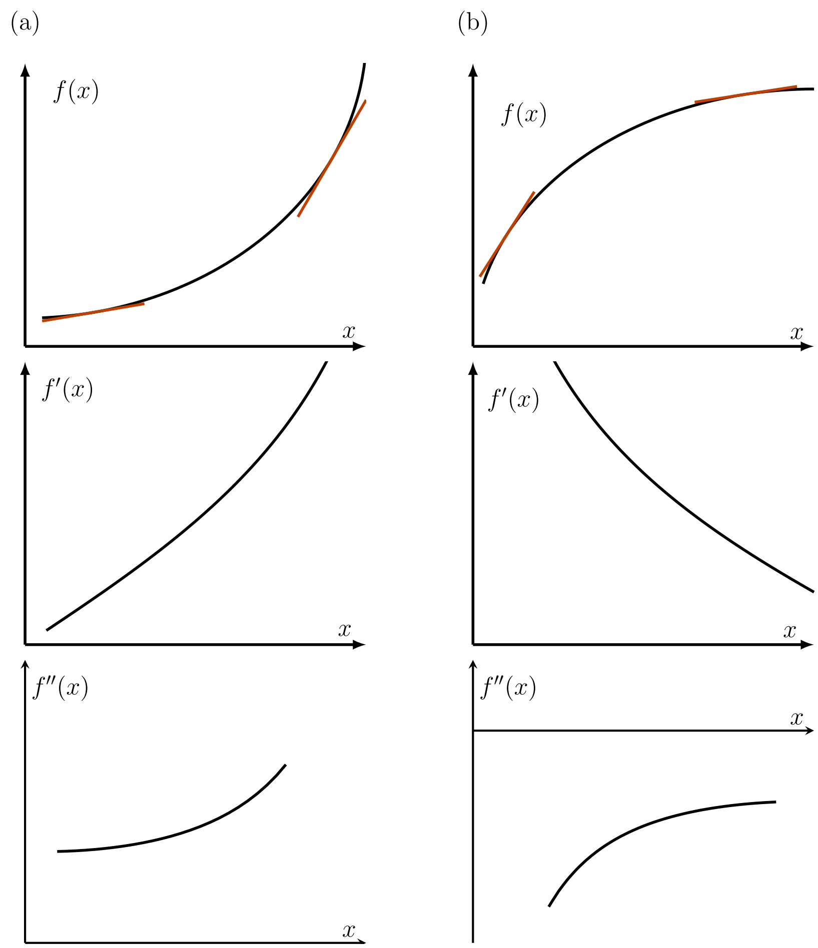


Figure 6.1: In (a) the function is concave up, and its derivative increases (in the positive direction). In (b), the function is concave down, and its derivative decreases.

Definition 6.1 *A* **point of inflection** *of a function $f(x)$ is a point x at which the concavity of the function changes (Figure 6.2).*

We can deduce from the definition and previous remarks that *at a point of inflection the second derivative changes sign*. This is illustrated in Figure 6.2.

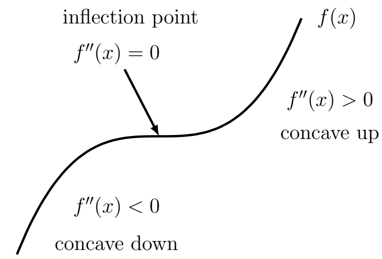


Figure 6.2: An inflection point is a place where the concavity of a function changes.

Note: *It is not enough to show that $f''(x) = 0$ to conclude that x is an inflection point.* We summarize the one-way nature of this relationship in the box and discuss it further in Example 6.1.

Inflection points

1. If the function $y = f(x)$ has a point of inflection at x_0 then $f''(x_0) = 0$.
2. If the function $y = f(x)$ satisfies $f''(x_0) = 0$, we **cannot conclude** that it has a point of inflection at x_0. We must actually check that $f''(x)$ changes sign at x_0.

Example 6.1 *Consider the the functions (a) $f_1(x) = x^3$ and (b) $f_2(x) = x^4$. Show that for both functions, the second derivative is zero at the origin ($f''(0) = 0$) but that only one of these functions actually has an inflection point at $x = 0$.*

A summary of Example 6.1, reinforcing the fact that $f''(x) = 0$ is not enough to guarantee an inflection point! We have to check that $f''(x)$ changes sign.

Solution. The functions are

$$\text{(a) } f_1(x) = x^3, \quad \text{(b) } f_2(x) = x^4.$$

The first derivatives are

$$\text{(a) } f_1'(x) = 3x^2, \quad \text{(b) } f_2'(x) = 4x^3.$$

and the second derivatives are:

$$\text{(a) } f_1''(x) = 6x, \quad \text{(b) } f_2''(x) = 12x^2.$$

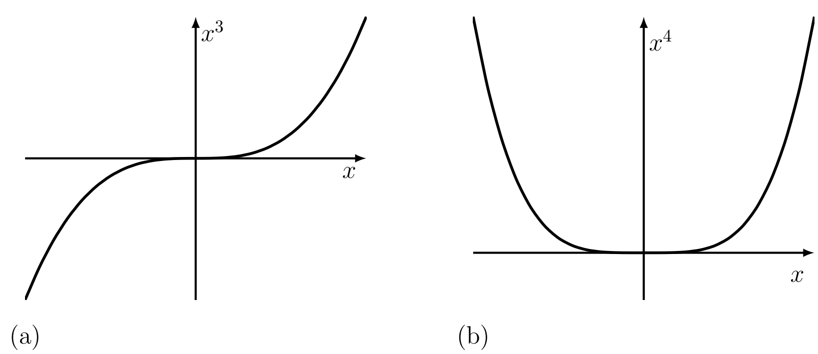


Figure 6.3: The functions (a) $f_1(x) = x^3$ and (b) $f_2(x) = x^4$ both satisfy $f''(0) = 0$. However, only x^3 has an inflection point at $x = 0$, whereas x^4 has a local minimum at that point. This is not a contradiction, since $f_2''(x)$ does not change sign at $x = 0$.

Thus, at $x = 0$ we have $f_1''(0) = 0, f_2''(0) = 0$. However, $x = 0$ is **NOT** an inflection point of x^4. In fact, it is a local minimum, as shown in Figure 6.3. ♢

Mastered Material Check

3. Check that $f(x) = x^4$ does not change sign at $x = 0$ by comparing the signs of $f(-0.1)$ and $f(0.1)$.

Determining whether $f''(x)$ changes sign

We defined an inflection point as a point on the graph of a function where the second derivative changes sign. But how do we detect if this sign change occurs at a given point?

We first state the following result

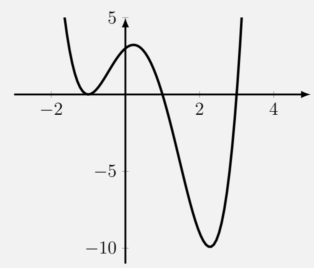
How we check where $f''(x)$ changes sign (to identify inflection points).

Sign change in a product of factors:
If an expression is a product of factors, e.g. $g(x) = (x-a_1)^{n_1}(x-a_2)^{n_2}\dots(x-a_m)^{n_m}$, then

1. The expression can be zero only at the points $x = a_1, a_2, \dots, a_m$.
2. The expression changes sign only at points $x = a_i$ for which n_i is an odd integer power.

Mastered Material Check

4. How might you determine sign changes for a function which is *not* given as product of factors?
5. Circle the places where the function depicted below changes sign.



Example 6.2 *Determine where $g(x) = x^2(x+2)(x-3)^4$ changes sign.*

Solution. The zeros of $g(x)$ are $x = 0, -2, 3$. At $x = 0$ and at $x = 3$ there is no sign change as the terms x^2 and $(x-3)^4$ are always positive. The factor $(x+2)$ goes from negative through zero to positive when x goes from $x < -2$ to $x = -2$ to $x > -2$. Hence, $g(x)$ only changes sign at $x = -2$. ◊

Example 6.3 *Where does the function $f(x) = (2/5)x^6 - x^4 + x$ have inflection point(s)?*

Solution. The derivatives of the function are

$$\begin{aligned} f'(x) &= (12/5)x^5 - 4x^3 + 1, \\ f''(x) &= 12x^4 - 14x^2 = 12x^2(x^2-1) = 12x^2(x+1)(x-1). \end{aligned}$$

Here we have completely factored the second derivative. Sign changes can only occur when there are factors with odd powers, such as $(x+1)$ and $(x-1)$. These change sign at $x = -1, 1$, respectively - making both inflection points. There is NO sign change at $x = 0$, since the factor x^2 is always positive. ◊

6.2 *Special points on the graph of a function*

Section 6.2 Learning goals

1. Define a zero of a function and be able to identify zeros for simple functions (factorizable polynomials).
2. Explain that a function $f(x)$ can have various types of critical points (maxima, minima, and other types) at which $f'(x) = 0$.
3. Find critical points for a given function.
4. Using first or second derivative tests, classify a given critical point as a maximum, minimum (or neither).

In this section we use tools of algebra and calculus to identify special points on the graph of a function. We first consider the **zeros** of a function, and then its **critical points**.

Zeros of a function

Example 6.4 (Factoring) *For the function $y = f(x) = x^2 - 5x + 6$, find zeros by factoring.*

Solution. This polynomial has factors $f(x) = (x-3)(x-2)$. Zeros are values of x satisfying $0 = (x-3)(x-2)$, so either $(x-3) = 0$ or $(x-2) = 0$. Hence, there are two zeros, $x = 2, 3$. ♢

Mastered Material Check

5. What is another term used for *zeros* of a function?

Example 6.5 *Find the zeros of the function $y = f(x) = x^3 - 3x^2 + x$.*

Solution. We can factor this function into $f(x) = x(x^2 - 3x + 1)$. From this we see that $x = 0$ is one of the desired zeros of f. To find the others, we apply the quadratic formula to the second factor, obtaining

$$x_{1,2} = \frac{1}{2}(3 \pm \sqrt{3^2 - 4}) = \frac{1}{2}(3 \pm \sqrt{5}).$$

Thus, there are a total of three zeros in this case:

$$x = 0, \quad \frac{1}{2}(3 + \sqrt{5}), \quad \frac{1}{2}(3 - \sqrt{5}).$$

♢

Example 6.6 *Find zeros of the function $y = f(x) = x^3 - x - 3$.*

Solution. This polynomial does not factor into integers, nor is it easy to apply a cubic formula (analogous to the quadratic formula). However, as we saw in Example 5.10, Newton's method leads to an accurate approximation for the only zero of this function, ($x \approx 1.6717$) ♢

Critical points

Definition 6.2 *A* **critical point** *of the function $f(x)$ is any point x at which the first derivative is zero, i.e. $f'(x) = 0$.*

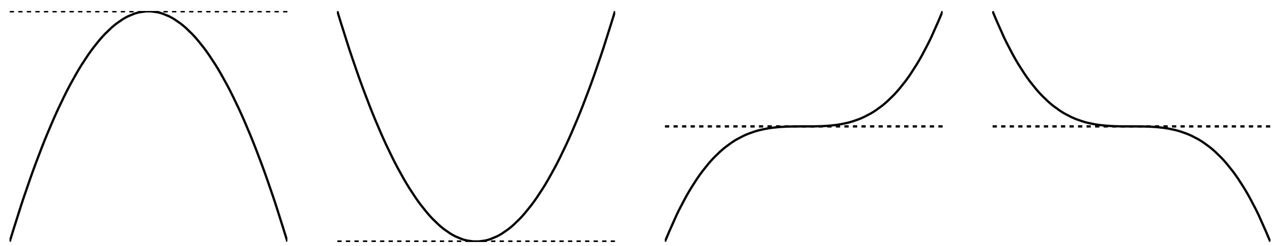


Figure 6.4: A critical point (place where $f'(x) = 0$) can be a local maximum, local minimum, or neither.

Clearly, a critical point occurs whenever the slope of the tangent line to the graph of the function is zero, i.e. the tangent line is horizontal. Figure 6.4 shows several possible shapes of the graph of a function close to a critical point. From left to right, these are a **local maximum**, a **local minimum** and two **inflection points**.

A summary of types of critical points and how to tell them apart.

Critical points play an important role in many scientific applications, as described in Chapter 7. Hence, we seek criteria to determine whether a critical point is a local maximum, minimum, or neither.

Example 6.7 *Consider the function $y = f(x) = x^3 + 3x^2 + ax + 1$. For what range of values of a are there no critical points for this function?*

Solution. We compute the first derivative $f'(x) = 3x^2 + 6x + a$. A critical point would occur whenever $0 = f'(x)$, which implies $0 = 3x^2 + 6x + a$. This is a quadratic equation whose solutions are

$$x_{1,2} = \frac{-6 \pm \sqrt{36 - 4a \cdot 3}}{6}.$$

There two real solutions provided the discriminant is positive, $36 - 12a > 0$. However, when $36 - 12a < 0$ (which corresponds to $a > 3$), there are no real solutions and consequently no critical points. ◊

Mastered Material Check

6. Why are there no critical points when $36 - 12a < 0$ for the function $f(x?) = x^3 + 3x^2 + ax + 1$

What happens close to a critical point

In Figure 6.5 we contrast the behaviour of two functions (top row), each with a different type of critical point. We compare their first and second derivatives close to that point (second and third rows, respectively). In each case, the first derivative $f'(x) = 0$ at the critical point.

Near the local maximum (moving from left to right), the slope of $f(x)$ transitions from positive to zero (at the critical point) to negative. This means that $f'(x)$ is a *decreasing* function, as indicated in Figure 6.5. Since changes in the first derivative are measured by *its* derivative $f''(x)$, we find that the

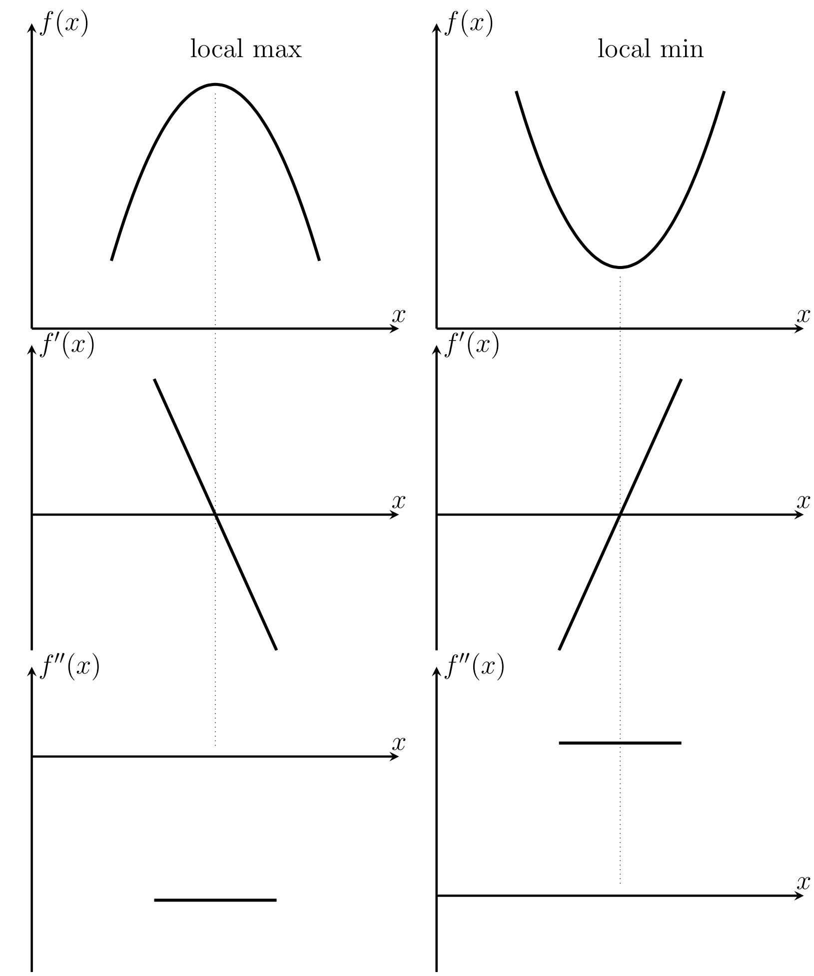


Figure 6.5: Close to a local maximum, $f(x)$ is concave down, $f'(x)$ is decreasing, so that $f''(x)$ is negative. Close to a local minimum, $f(x)$ is concave up, $f'(x)$ is increasing, so that $f''(x)$ is positive.

second derivative is negative at a local maximum. The converse is true near a local minimum (right column in Figure 6.5).

We collect and summarize conclusions about derivatives below.

Summary: first derivative

$f'(x) < 0$	$f'(x_0) = 0$	$f'(x) > 0$
decreasing function	critical point at x_0	increasing function

- **The first derivative test**: depends on changes in the *sign* of the first derivative close to a critical point, x_0.

  Near a **local maximum**, the sign pattern is:

1. $x < x_0$, $f'(x) > 0$;
2. $x = x_0$, $f'(x_0) = 0$;

3. $x > x_0$, $f'(x_0) < 0$.

Near a **local minimum**, the sign pattern is:

1. $x < x_0$, $f'(x) < 0$;
2. $x = x_0$, $f'(x_0) = 0$;
3. $x > x_0$, $f'(x) > 0$.

Summary: second derivative

$f''(x) < 0$	$f''(x_0) = 0$	$f''(x) > 0$
curve concave down	inflection point at x_0 only if f'' changes sign	curve concave up

- **The second derivative test**: is based on the sign of $f''(x_0)$ for x_0 a critical point (satisfying $f'(x_0) = 0$).

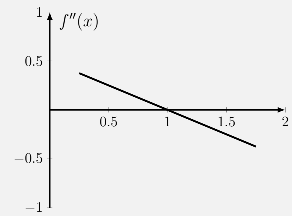
  If $f''(x_0) < 0$, x_0 is a **local maximum**.

  If $f''(x_0) = 0$, test is inconclusive about max/min at x_0. If f'' also changes sign at x_0, then x_0 is an **inflection point**.

  If $f''(x_0) > 0$, x_0 is a **local minimum**.

Mastered Material Check

7. Given the following graph of a function's *second* derivative, can you say anything about the concavity of the original function?



6.3 *Sketching the graph of a function*

Section 6.3 Learning goals

1. Given a function (polynomial, rational, etc.) be able to find its zeros, critical points, inflection points, and determine where it is increasing or decreasing, concave up or down.
2. Using a combination of the above techniques, together with methods of Section 1.4, assemble a reasonably accurate sketch of the graph of a function.
3. Using these techniques, identify all local as well as global **extrema** (minima and maxima) of a function $f(x)$ on an interval $a \leq x \leq b$.

In Section 1.4, we used elementary reasoning about power functions to sketch the graph of simple polynomials. Now that we have learned more advanced calculus techniques, we can hone such methods to produce more accurate sketches of the graph of a function. We devote this section to illustrating some examples.

Example 6.8 *Sketch the graph of the function $B(x) = C(x^2 - x^3)$. Assume that $C > 0$ is constant.*

Steps in the calculations of Example 6.8.

Mastered Material Check

8. What is the independent variable in Example 6.8? The dependent variable?
9. Given our discussion on "considering the powers" for Example 6.8, add a reasonable scale to each of x-axes for the graphs in Figure 6.6.

Solution. To prepare, we compute the derivatives:

$$B'(x) = C(2x - 3x^2), \quad B''(x) = C(2 - 6x).$$

1. **Zeros.** We start by finding the zeros of the function. Factoring makes this easy. Solving $B(x) = 0$, we find

$$0 = C(x^2 - x^3) = Cx^2(1 - x), \quad \Rightarrow \quad \Rightarrow \quad x = 0, 1.$$

2. **Consider the powers.** Reasoning about the powers as in Chapter 1, we surmise that close to the origin, x^2 dominates (producing a parabolic shape) whereas, far away, $-x^3$ dominates (producing the shape of an inverted cubic). We show this in a preliminary sketch, Figure 6.6.

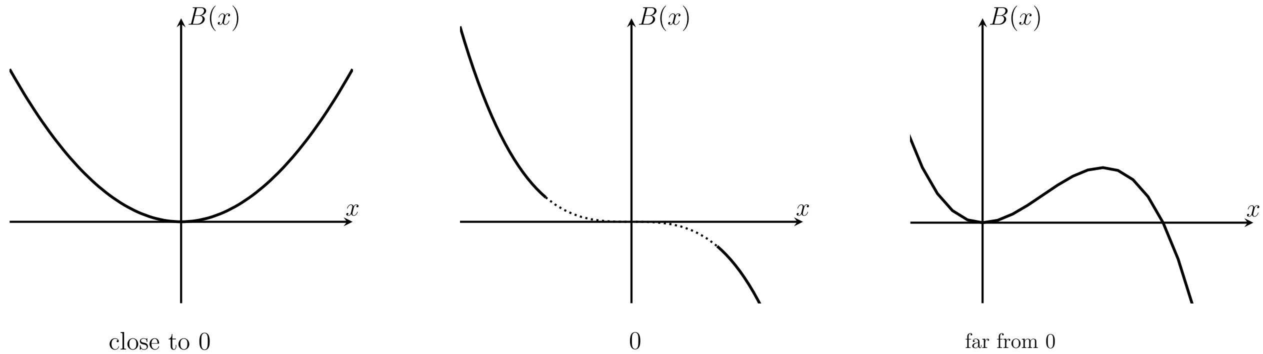


Figure 6.6: Figure for the function $B(x) = C(x^2 - x^3)$ in Example 6.8 showing which power dominates.

3. **First derivative.** To find critical points, set $B'(x) = 0$, obtaining

$$B'(x) = C(2x - 3x^2) = 0, \quad \Rightarrow \quad 0 = 2x - 3x^2 = x(2 - 3x),$$
$$\Rightarrow \quad x = 0, \frac{2}{3}.$$

By considering the sketch in Figure 6.6, we can see that $x = 0$ is a local minimum, and $x = 2/3$ a local maximum. Confirmation of this comes from the second derivative.

4. **Second derivative.** From the second derivative, $B''(0) = 2 > 0$, confirming that $x = 0$ is a local minimum. Further, $B''(2/3) = 2 - 6 \cdot (2/3) = -2 < 0$ so $x = 2/3$ is a local maximum, as expected.

5. **Classifying the critical points.** Now identifying where $B''(x) = 0$, we find that

$$B''(x) = C(2 - 6x) = 0, \quad \text{when} \quad 2 - 6x = 0 \quad \Rightarrow \quad x = \frac{2}{6} = \frac{1}{3}$$

we also note that the second derivative changes sign here: i.e. for $x < 1/3$, $B''(x) > 0$ and for $x > 1/3$, $B''(x) < 0$. We may conclude that there is an inflection point at $x = 1/3$. The final sketch, now labeled, is given in Figure 6.7.

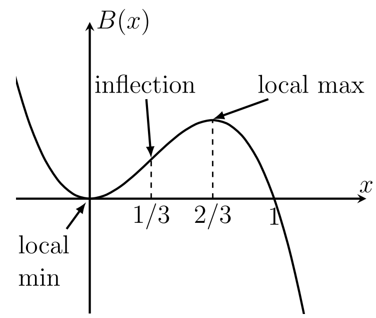


Figure 6.7: A sketch of $B(x) = C(x^2 - x^3)$ in Example 6.8.

◊

Example 6.9 *Sketch the graph of the function* $y = f(x) = 8x^5 + 5x^4 - 20x^3$

Solution. This example is more challenging, but similar ideas apply.

1. **Zeros.** Factoring the expression for y and then using the quadratic formula leads to

$$y = x^3(8x^2 + 5x - 20). \quad \Rightarrow \quad x = 0, -\frac{5}{16} \pm \frac{1}{16}\sqrt{665}.$$

In decimal form, these are approximately $x = 0, 1.3, -1.92$.

2. **Consider the powers.** The highest power is $8x^5$, so far from the origin we expect typical positive odd function behaviour. The lowest power is $-20x^3$, which means that close to zero, we expect to see a negative cubic. This implies that the function "turns around", creating local maxima and minima. We draw a rough sketch in Figure 6.8.

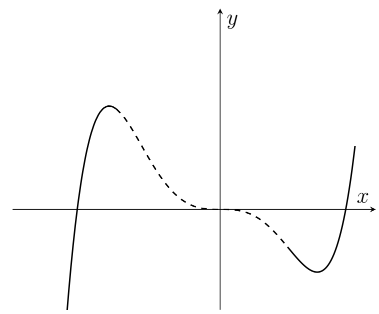


Figure 6.8: $y = f(x) = 8x^5 + 5x^4 - 20x^3$. From Example 6.9, this function behaves roughly like the negative cubic near the origin, and like $8x^5$ for large x.

3. **First derivative.** Calculating the derivative of $f(x)$ and then factoring leads to

$$\frac{dy}{dx} = f'(x) = 40x^4 + 20x^3 - 60x^2 = 20x^2(2x+3)(x-1)$$

so this derivative is zero at: $x = 0, 1, -3/2$. We expect critical points at these places.

4. **Second derivative.** We calculate the second derivative and factor to obtain

$$\frac{d^2y}{dx^2} = f''(x) = 160x^3 + 60x^2 - 120x = 20x(8x^2 + 3x - 6).$$

Thus, the second derivative is zero at

$$x = 0, -\frac{3}{16} + \frac{1}{16}\sqrt{201}, -\frac{3}{16} - \frac{1}{16}\sqrt{201}.$$

The values of these roots can be approximated by: $x = 0, 0.69, -1.07$.

Mastered Material Check

(a) Verify the given zeros of the second derivative of Example 6.9.

10. **Classifying the critical points.** To identify the types of critical points, we use the second derivative test.

- $f''(0) = 0$ so the test is inconclusive at $x = 0$.
- $f''(1) = 20(8 + 3 - 6) > 0$ implies a local minimum $x = 1$, and
- $f''(-3/2) = -225 < 0$ implies a local maximum at $x = -3/2$.

We summarize the results in Table 6.1.

The shape of the function, its first and second derivatives are shown in Figure 6.9. ◊

Mastered Material Check

11. Use software to verify the plot of the function $y = f(x)$ in Example 6.9. Then plot $2y$. What changes?

Table 6.1: $y = x^3(8x^2 + 5x^2 - 20)$ and its detailed behaviour.

$x =$	-1.92	-1.5	-1.07	0	0.69	1	1.3
$f(x) =$	0	32.0		0		-7	0
$f'(x) =$		0		0		0	
$f''(x) =$		<0	0	0	0	>0	
characteristic:	zero	max	inflection		inflection	min	zero

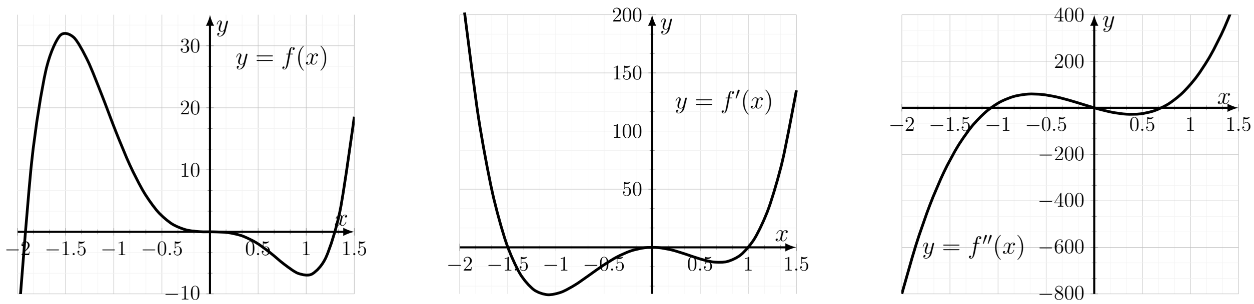


Figure 6.9: The function $y = f(x) = 8x^5 + 5x^4 - 20x^3$, and its first and second derivatives, $f'(x)$ and $f''(x)$.

Global maxima and minima, endpoints of an interval

A **global maximum** (also denoted **absolute maximum**) of a function over some interval is the largest value that the function attains on that interval. Similarly a **global minimum** (or **absolute minimum**) is the smallest value.

Note. For a function defined on a *closed* interval, we must check critical points and endpoints of the interval to determine where global maxima and minima occur, as illustrated in Example 6.10.

Steps in Example 6.10: Finding the absolute minimum and maximum of a function on a given interval.

Example 6.10 *Consider* $y = f(x) = \frac{2}{x} + x^2$ *on the interval* $0.1 \le x \le 4$. *Find the absolute maximum and minimum.*

Solution. We first compute the derivatives:

$$f'(x) = -2\frac{1}{x^2} + 2x, \quad f''(x) = 4\frac{1}{x^3} + 2.$$

Solving for critical points by setting $f'(x) = 0$, we find

$$-2\frac{1}{x^2} + 2x = 0, \quad \Rightarrow \quad -2\frac{1}{x^2} = 2x \quad \Rightarrow \quad x^3 = 1$$

which implies a critical point at $x = 1$. The second derivative at this point is

$$f''(1) = 4\frac{1}{1^3} + 2 = 6 > 0,$$

so that $x = 1$ is a local minimum.

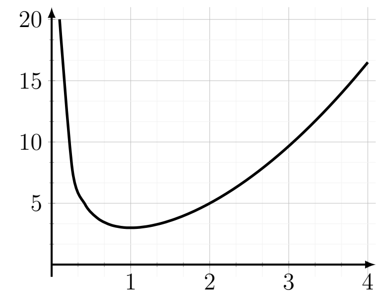


Figure 6.10: The function $f(x) = \frac{2}{x} + x^2$ on the interval $0.1 < x < 4$ has no local maximum.

We now calculate the value of the function at the endpoints $x = 0.1$ and $x = 4$ and at the critical point $x = 1$ to determine where global and local minima and/or maxima occur:

- $f(0.1) = \frac{2}{0.1} + 0.1^2 = 20.01$;
- $f(1) = \frac{2}{1} + 1^2 = 3$;
- $f(4) = \frac{2}{4} + 4^2 = 16.5$.

Consequently, $x = 1$ is both a local minimum and the global minimum on the given interval. There are no local maxima. The global maximum occurs at the left endpoint, $x = 4$. Figure 6.10 confirms our conclusions. ◊

6.4 *Summary*

1. We can improve the accuracy of a sketch of a function $f(x)$ by examining its derivatives. For example, if $f'(x) > 0$ then $f(x)$ is increasing, and if $f'(x) < 0$, then $f(x)$ is decreasing. Further, if $f''(x) > 0$ then this corresponds to $f'(x)$ increasing, which means $f(x)$ is concave up. Similarly, $f''(x) < 0$ corresponds to $f(x)$ being concave down.
2. If $f''(x_0) = 0$ **and** $f''(x)$ changes sign at x_0, then x_0 is a point of inflection of $f(x)$: a point where its concavity changes.
3. If $f'(x_0) = 0$, then x_0 is a critical point of the function $f(x)$. A critical point needs to be further tested to determine whether it is a local minimum, local maximum or neither.
4. Given a critical point x_0 (where $f'(x_0) = 0$), the first derivative test examines the sign of $f'(x)$ **near** x_0. If the sign pattern of $f'(x)$ is $+, 0, -$ (from left to right as we cross x_0), then x_0 is a local maximum. If the sign pattern is $-, 0, +$, then x_0 is a local minimum.
5. Given a critical point x_0 (where $f'(x_0) = 0$), the second derivative test examines the sign of $f''(x_0)$. If $f''(x_0) < 0$, then x_0 is a local maximum. If $f''(x_0) > 0$, then x_0 is a local minimum. If $f''(x_0) = 0$, the second derivative test is inconclusive.
6. A **global** maximum and minimum can only be identified after comparing the values of the function at all critical points and endpoints of the interval. (Remark: For discontinuous or non-smooth functions, we must also examine values of f at points of discontinuity and cusps, not discussed in this chapter).

Quick Concept Checks

1. Consider the function depicted below:

   (a) Circle all zeros.

   (b) Where are the local maxima?

   (c) How many local minima are there?

   (d) Where does the function change sign?

2. Draw a graph to justify that if $f'(x) > 0$, then $f(x)$ is increasing.

3. Where does the function $f(a) = a^3(a-\pi)(a+\rho)^5(a-2)^2$ change signs?

4. Suppose you are given that γ is a critical point of some function g. What would you ask to learn more about the shape of g?

Exercises

6.1. **Zeros, local minima and maxima.** A zero of a function is a place where $f(x) = 0$.

(a) Find the zeros, local maxima, and minima of the polynomial $y = f(x) = x^3 - 3x$

(b) Find the local minima and maxima of the polynomial $y = f(x) = (2/3)x^3 - 3x^2 + 4x$.

(c) Determine whether each of the polynomials given in parts (a) and (b) have an inflection point.

6.2. **Classifying critical points.** Find critical points, zeros, and inflection points of the function $y = f(x) = x^3 - ax$. Then classify the types of critical points that you have found.

6.3. **Sketching graphs.** For each of the following functions, sketch the graph for $-1 < x < 1$, find $f'(0),\ f'(1),\ f'(-1)$ and identify any local minima and maxima.

(a) $y = x^2$,

(b) $y = -x^3$,

(c) $y = -x^4$

(d) Using your observations, when can you conclude that a function whose derivative is zero at some point has a local maximum at that point?

6.4. **Sketching a graph.** Sketch a graph of the function $y = f(x) = x^4 - 2x^3$, using both calculus and methods of Chapter 1.

6.5. **Global maxima and minima.** Find the global maxima and minima for the function in Exercise 4 on the interval $0 \leq x \leq 3$.

6.6. **Absolute maximum and minimum.** Find the absolute maximum and minimum values on the given interval:

(a) $y = 2x^2$ on $-3 \leq x \leq 3$

(b) $y = (x-5)^2$ on $0 \leq x \leq 6$

(c) $y = x^2 - x - 6$ on $1 \leq x \leq 3$

(d) $y = \dfrac{1}{x} + x$ on $-4 \leq x \leq -\dfrac{1}{2}$.

6.7. **Local vs. absolute.** A function $f(x)$ has as its derivative $f'(x) = 2x^2 - 3x$

(a) In what regions is f increasing or decreasing?

(b) Find any local maxima or minima.

(c) Is there an absolute maximum or minimum value for this function?

6.8. **Minimum value.** Sketch the graph of $x^4 - x^2 + 1$ in the range -3 to 3. Find its minimum value.

6.9. **Critical points.** Identify all the critical points of the following function.

$$y = x^3 - 27$$

6.10. **Critical and inflection points.** Consider the function $g(x) = x^4 - 2x^3 + x^2$. Determine locations of critical points and inflection points.

6.11. **No critical points.** Consider the polynomial $y = x^3 + 3x^2 + ax + 1$. Show that when $a > 3$ this polynomial has no critical points.

6.12. **Critical points and generic parabola.** Find the values of a, b, and c if the parabola $y = ax^2 + bx + c$ is tangent to the line $y = -2x + 3$ at $(2, -1)$ and has a critical point when $x = 3$.

6.13. **Double wells and physics.** In physics, a function such as

$$f(x) = x^4 - 2x^2$$

is often called a *double well potential*. Physicists like to think of this as a "landscape" with hills and valleys. They imagine a ball rolling along such a landscape: with friction, the ball eventually comes to rest at the bottom of one of the valleys in this potential. Sketch a picture of this landscape and use information about the derivative of this function to predict where the ball might be found, i.e. where the valley bottoms are located.

6.14. **Function concavity.** Find the first and second derivatives of the function

$$y = f(x) = \frac{x^3}{1 - x^2}.$$

Use information about the derivatives to determine any local maxima and minima, regions where the curve is concave up or down, and any inflection points.

6.15. **Classifying critical points.** Find all the critical points of the function

$$y = f(x) = 2x^3 + 3ax^2 - 12a^2x + 1$$

and determine what kind of critical point each one is. Your answer should be given in terms of the constant a, and you may assume that $a > 0$.

6.16. **Describing a function.** The function $f(x)$ is given by

$$y = f(x) = x^5 - 10kx^4 + 25k^2x^3$$

where k is a positive constant.

(a) Find all the intervals on which f is either increasing or decreasing. Determine all local maxima and minima.

(b) Determine intervals on which the graph is either concave up or concave down. What are the inflection points of $f(x)$?

6.17. **Muscle shortening.** In 1938 Av Hill proposed a mathematical model for the rate of shortening of a muscle, v, (in cm/sec) when it is working against a load p (in gms). His so called force-velocity curve is given by the relationship

$$(p+a)v = b(p_0 - p)$$

where a, b, p_0 are positive constants.

(a) Sketch the shortening velocity versus the load, i.e., v as a function of p.

Note: the best way to do this is to find the intercepts of the two axes, i.e. find the value of v corresponding to $p = 0$ and vice versa.

(b) Find the rate of change of the shortening velocity with respect to the load, i.e. calculate dv/dp.

(c) What is the largest load for which the muscle contracts? (*hint:* a contracting muscle has a positive shortening velocity, whereas a muscle with a very heavy load stretches, rather than contracts, i.e. has a negative value of v).

6.18. **Reaction kinetics.** Chemists often describe the rate of a saturating chemical reaction using Michaelis-Menten (R_m) or sigmoidal (R_s) kinetics

$$R_m(c) = \frac{Kc}{k_n + c}, \quad R_s(c) = \frac{Kc^2}{k_n^2 + c^2}$$

where c is the concentration of the reactant, $K > 0$, $k_n > 0$ are constants. $R(c)$ is the speed of the reaction as a function of the concentration of reactant.

(a) Sketch the two curves. To do this, you should analyze the behaviour for $c = 0$, for small c, and for very large c. You will find a horizontal asymptote in both cases. We refer to that asymptote as the "maximal reaction speed". What is the "maximal reaction speed" for each of the functions R_m, R_s ?

Note: express your answer in terms of the constants K, k_n.

(b) Show that the value $c = k_n$ leads to a half-maximal reaction speed.

For the questions below, you may assume that $K = 1$ and $k_n = 1$.

(c) Show that sigmoidal kinetics, but not Michaelis-Menten kinetics has an inflection point.

(d) Explain how these curves would change if K is increased; if k_n is increased.

6.19. **Checking the endpoints.** Find the absolute maximum and minimum values of the function

$$f(x) = x^2 + \frac{1}{x^2}$$

on the interval $[\frac{1}{2}, 2]$. Be sure to verify if any critical points are maxima or minima and to check the endpoints of the interval.

7
Optimization

Calculus was developed to solve practical problems. In this chapter, we concentrate on optimization problems, where finding "the largest," "the smallest," or "the best" answer is the goal. We apply some of the techniques developed in earlier chapters to find local and global maxima and minima. A new challenge in this chapter is translating a word-problem into a mathematical problem. We start with elementary examples, and work to more complex situations with biological motivation.

Mastered Material Check

1. How do you find the critical points of a function $f(x)$?

7.1 Simple biological optimization problems

Section 7.1 Learning goals

1. Given a function, find the derivative of that function and identify all critical points.
2. Using a combination of sketching and tests for critical points developed in Section 6.2, diagnose the type of critical point.

In the first examples, the function to optimize is specified, making the problem simply one of carefully applying calculus methods.

Density dependent (logistic) growth in a population

Biologists often notice that the growth rate of a population depends not only on the size of the population, but also on how crowded it is. Constant growth is not sustainable. When individuals have to compete for resources, nesting sites, mates, or food, they cannot invest time nor energy in reproduction, leading to a decline in the rate of growth of the population. Such population growth is called **density dependent growth**.

One common example of density dependent growth is called the **logistic growth** law. Here it is assumed that the growth rate of the population, G

depends on the density of the population, N, as follows:

$$G(N) = rN\left(\frac{K-N}{K}\right).$$

Here N is the **independent variable**, and $G(N)$ is the function of interest. All other quantities are constant:

- $r > 0$ is a constant, called the **intrinsic growth rate**, and
- $K > 0$ is a constant, called the **carrying capacity**. It represents the population density that a given environment can sustain.

Importantly, when differentiating G, we treat r and K as "numbers". A generic sketch of G as a function of N is shown in Figure 7.1.

Mastered Material Check

2. Give an example of units for N.
3. What units might G carry?

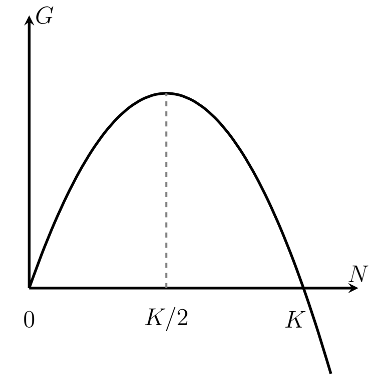


Figure 7.1: In logistic growth, the population growth rate rate G depends on population size N as shown here.

Example 7.1 (Logistic growth rate) *Answer the following questions:*

a) *Find the population density N that leads to the maximal growth rate $G(N)$.*

b) *Find the value of the maximal growth in terms of r, K.*

c) *For what population size is the growth rate zero?*

Solution. We can expand $G(N)$:

$$G(N) = rN\left(\frac{K-N}{K}\right) = rN - \frac{r}{K}N^2,$$

from which it is apparent that $G(N)$ is a polynomial in powers of N, with constant coefficients r and r/K.

a) To find critical points of $G(N)$, we find N such that $G'(N) = 0$, and then test for maxima:

$$G'(N) = r - 2\frac{r}{K}N = 0. \quad \Rightarrow \quad r = 2\frac{r}{K}N \quad \Rightarrow \quad N = \frac{K}{2}.$$

Hence, $N = K/2$ is a critical point, but is it a maximum? We check this in one of several ways. First, a sketch in Figure 7.1 reveals a downwards-opening parabola. This confirms a local maximum. Alternately, we can apply a tool from Section 6.2 such as the **second derivative test**:

$$G''(N) = -2\frac{r}{K} < 0 \quad \Rightarrow \quad G(N) \text{ concave down}$$
$$\Rightarrow \quad N = \frac{K}{2} \text{ is a local maximum}$$

Thus, the population density with the greatest growth rate is $K/2$.

b) The maximal growth rate is found by evaluating the function G at the critical point, $N = K/2$,

$$G\left(\frac{K}{2}\right) = r\left(\frac{K}{2}\right)\left(\frac{K-\frac{K}{2}}{K}\right) = r\frac{K}{2}\cdot\frac{1}{2} = \frac{rK}{4}.$$

c) To find the population size at which the growth rate is zero, we set $G = 0$ and solve for N:

$$G(N) = rN\left(\frac{K-N}{K}\right) = 0.$$

There are two solutions. One is trivial: $N = 0$. (This is biologically interesting in the sense that it rules out the ancient idea of **spontaneous generation** - a defunct theory that held that life can arise on its own, from dust or air. If $N = 0$, the growth rate is also 0, so no population spontaneously arises according to logistic growth.) The second solution, $N = K$ means that the population is at its "carrying capacity".

$\diamondsuit$

We return to this type of growth in Chapter 13.

Cell size for maximal nutrient accumulation rate

According to the model of Section 1.2, the nutrient absorption and consumption rates, $A(r), C(r)$, of a simple spherical cell of radius r are:

$$A(r) = k_1 S = 4k_1\pi r^2, \quad C(r) = k_2 V = \frac{4}{3}\pi k_2 r^3,$$

for $k_1, k_2 > 0$ constants. The **net rate of increase** of nutrients, which is the difference of the two is:

$$N(r) = A(r) - C(r) = 4k_1\pi r^2 - \frac{4}{3}\pi k_2 r^3. \tag{7.1}$$

This quantity is a function of the radius r of the cell.

Mastered Material Check

4. Give example units for each of $A(r)$, $C(r)$ and $N(r)$. Are there restrictions in place?

Example 7.2 *Determine the radius of the cell for which the net rate of increase of nutrients $N(r)$ is largest.*

Solution. We are asked to maximize $N(r)$ with respect to r. We first find critical points of $N(r)$, keeping in mind that $8k_1\pi$ and $4k_2\pi$ are constant for the purpose of differentiation. Critical points occur when $N'(r) = 0$, i.e.

$$N'(r) = 8k_1\pi r - 4k_2\pi r^2 = 0 \quad \Rightarrow \quad 4\pi r(2k_1 - k_2 r) = 0$$
$$\Rightarrow \quad r = 0, \quad 2\frac{k_1}{k_2}.$$

To identify the type of critical point, we use the second derivative test

$$N''(r) = 8k_1\pi - 8k_2\pi r = 8\pi(k_1 - k_2 r).$$

Substituting in $r = 2k_1/k_2$, we find that

$$N''\left(2\frac{k_1}{k_2}\right) = 8\pi\left(k_1 - k_2\frac{2k_1}{k_2}\right) = -8\pi k_1 < 0.$$

Thus, the second derivative is negative at at $r = 2k_1/k_2$, verifying that this is a local maximum. Hence the net rate of nutrient uptake is greatest for cells of radius $r = 2k_1/k_2$. $\diamondsuit$

7.2 *Optimization with a constraint*

Section 7.2 Learning goals

1. Set up an optimization word problem involving formulae for volume and surface area of geometric solids.

2. Identify a constraint in an optimization problem.

3. Use the constraint to eliminate one of the independent variables, and find a desired critical point. (As before, this includes classifying the critical point as a local minimum, maximum or neither.)

In the next examples, identifying the function to optimize is part of the challenge. We also consider cases with more than one independent variable, where a constraint is used to eliminate all but one.

A cylindrical cell with minimal surface area

Not all cells are spherical. Some are cylindrical or sausage shaped. We explore how minimization of surface area would determine the overall shape of a cylindrical cell with a circular cross-section.

The volume of the cell is assumed to be fixed, because the cytoplasm in its interior cannot be "compressed". However, suppose that the cell has a "rubbery" membrane that tends to take on the smallest surface area possible. In physical language, the **elastic energy** stored in the membrane tends to a minimum. We want to find the proportions of the cylinder (that is, the ratio of length to radius) so that the cell has **minimal surface area**.

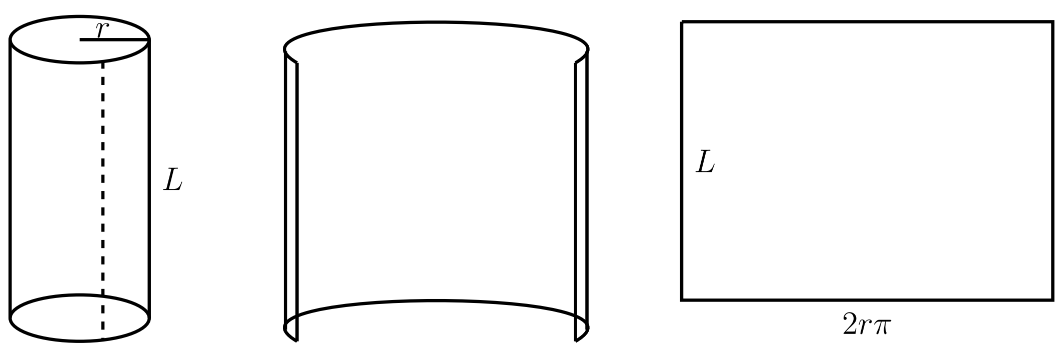


Figure 7.2: Properties of a cylinder

Recall the following properties for a cylinder of length L and radius r:

- The volume of a cylinder is the product of its base area, A, and its height, L. That is, $V = AL$. For a cylinder with circular cross-section, $V = \pi r^2 L$.

- As in Figure 7.2, a cylinder can be "cut and unrolled" into a rectangle with side lengths L and $2\pi r$, where r is the radius of the circular cross-section. The surface area is the product of these side lengths, $S_{\text{side}} = 2\pi rL$.

- If the "ends" of the cylinder are two flat circular caps then the sum of the areas of these two ends is $S_{\text{ends}} = 2\pi r^2$. While in a real cell, the end caps would not be actually flat, for simplicity, we assume flat, circular ends.
- The total surface area of the cylinder with flat ends is then

$$S = 2\pi rL + 2\pi r^2.$$

Mastered Material Check

5. If a cylindrical cell has volume $100\mu m^3$ and length $10\mu m$, what is its radius?
6. What is the surface area of a cylindrical cell with volume $100\mu m^3$ and length $10\mu m$?

Mathematically, our problem can be restated as follows

Example 7.3 *Minimize the surface area $S = 2\pi rL + 2\pi r^2$ of the cell, given that its volume $V = \pi r^2 L = K$ is constant*[1].

[1] I would like to thank Prof Nima Geffen (Tel Aviv University) for providing the inspiration for this example.

Solution. The shape of the cell depends on both the length, L, and the radius, r, of the cylinder. However, these are not independent. They are related to one another because the volume of the cell has to be constant. This is an example of an optimization problem with **a constraint**, i.e., a condition that has to be satisfied. The constraint is *"the volume is fixed"*, i.e.,

$$V = \pi r^2 L = K,$$

where $K > 0$ is a constant. This constraint allows us to eliminate one variables. For example, solving for L, we have

$$L = \frac{K}{\pi r^2}. \tag{7.2}$$

substituting this into the function S yields

$$S = 2\pi rL + 2\pi r^2. \quad \Rightarrow \quad S(r) = 2\pi r\frac{K}{\pi r^2} + 2\pi r^2 \quad \Rightarrow \quad S(r) = 2\frac{K}{r} + 2\pi r^2,$$

where S is now a function of a single independent variable, r (K and π are constants).

We have formulated the mathematical problem: **find the minimum of** $S(r)$.

We compute derivatives to find and classify the critical points:

$$S'(r) = -2\frac{K}{r^2} + 4\pi r, \quad S''(r) = 4\frac{K}{r^3} + 4\pi.$$

Since $K, r > 0$, the second derivative is always positive, so $S(r)$ is concave up. Any critical point we find is thus automatically a minimum. (In Exercise 7 we also consider the first derivative test as practice.) Setting $S'(r) = 0$:

$$S'(r) = -2\frac{K}{r^2} + 4\pi r = 0.$$

Solving for r, we obtain

$$2\frac{K}{r^2} = 4\pi r \quad \Rightarrow \quad r^3 = \frac{K}{2\pi} \quad \Rightarrow \quad r = \left(\frac{K}{2\pi}\right)^{1/3}.$$

We can also find the length of this cell from Eqn. 7.2.

$$L = \left(\frac{4K}{\pi}\right)^{1/3}.$$

(Details are left for Exercise 7).

We can finally characterize the shape of the cell. One way is to specify the ratio of its radius to its length. Based on our previous results, we find that ratio to be:

$$\frac{L}{r} = 2$$

(Exercise 7). This implies that $L = 2r$ which coincides with the length of a diameter of the same circle.

This means the "cylindrical cell" with rubbery membrane would be short and fat - something almost like a sphere of radius r with flattened ends. Some cells do grow as long cylindrical filaments, but such growth is inconsistent with an elastic membrane or a minimal surface area. (In fact, filamentous growth is one way for cells to maximize their surface area, and reduce the challenge of absorbing nutrients.) ◊

Mastered Material Check

7. Redo Example 7.3 for a cell with fixed volume $100 \mu m^3$.

Wine for Kepler's wedding

In 1613, Kepler set out to purchase a few barrels of wine for his wedding party. To compute the cost, the merchant would plunge a measuring rod through the tap hole, as shown in Figure 7.3 and measure the length L of the "wet" part of rod. The cost would be set at a value proportional to L.

Kepler noticed that barrels come in different shapes. Some are tall and skinny, while others are squat and fat. He conjectured that some shapes would contain larger volumes for a given length L, i.e. would contain more wine for the same price. Knowing mathematics, he set out to determine which barrel shape would be the best bargain for his wedding.

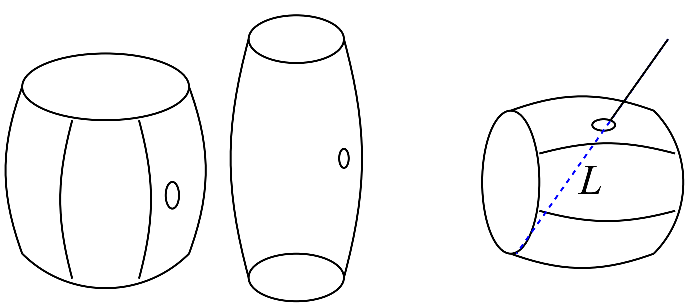


Figure 7.3: Barrels come in various shapes. But the cost of a barrel of wine was determined by the length L (dashed blue line segment) of the wet portion of the rod inserted into the tap hole. Kepler figured out which barrels contain the most wine for a given price.

Kepler sought the wine barrel that contains the most wine for a given cost. This is equivalent to asking *which cylinder has the largest volume* *for a fixed (constant) length L.* Below, we solve this optimization problem. An alternate approach is to seek the wine barrel that costs least for a given volume (Exercise 14), which leads to the same result.

Example 7.4 *Find the proportions (height:radius) of the cylinder with largest volume for a fixed length L (dashed line segment in Figure 7.3).*

Solution. We make the following assumptions:

1. the barrel is a simple cylinder, as shown in Figure 7.4,
2. the tap-hole (normally sealed to avoid leaks) is half-way up the height of the barrel, and
3. the barrel is full to the top with wine.

Mastered Material Check

8. Give two different examples of barrel dimensions which would both yeild a volume of 160L.

Let r, h denote the radius and height of the barrel. These two variables uniquely determine the shape as well as the volume of the barrel. Note that because the barrel is assumed to be full, the volume of the cylinder is the same as the volume of wine, namely

$$V = \text{base area} \times \text{height}. \quad \Rightarrow \quad V = \pi r^2 h. \tag{7.3}$$

The rod used to "measure" the amount of wine (and hence determine the cost of the barrel) is shown as the diagonal of length L in Figure 7.4. Because the cylinder walls are perpendicular to its base, the length L is the hypotenuse of a right-angle triangle whose other sides have lengths $2r$ and $h/2$. (This follows from the assumption that the tap hole is half-way up the side.) Thus, by the Pythagorean theorem,

$$L^2 = (2r)^2 + \left(\frac{h}{2}\right)^2. \tag{7.4}$$

The problem can now be stated mathematically: maximize V in Eqn. (7.3) subject to a fixed value of L in Eqn. (7.4). The fact that L is fixed means that we have a **constraint**, as before, that we use to reduce the number of variables in the problem.

Expanding the squares in the constraint and solving for r^2 leads to

$$L^2 = 4r^2 + \frac{h^2}{4} \quad \Rightarrow \quad r^2 = \frac{1}{4}\left(L^2 - \frac{h^2}{4}\right).$$

When we use this to eliminate r from the expression for V, we obtain

$$V = \pi r^2 h = \frac{\pi}{4}\left(L^2 - \frac{h^2}{4}\right)h = \frac{\pi}{4}\left(L^2 h - \frac{1}{4}h^3\right).$$

The mathematical problem to solve is now: find h that maximizes

$$V(h) = \frac{\pi}{4}\left(L^2 h - \frac{1}{4}h^3\right).$$

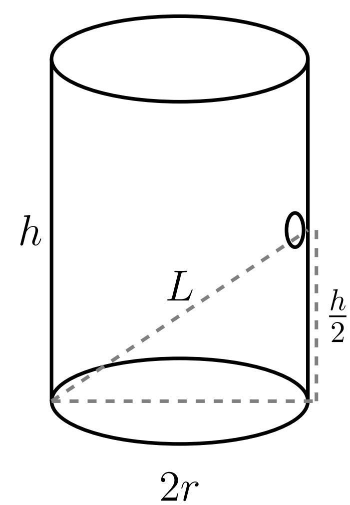


Figure 7.4: We simplify the problem to a cylindrical barrel with diameter $2r$ and height h. We assumed that the height of the tap-hole is $h/2$. Length L denotes the "wet" portion of the merchant's rod, used to determine the cost. We observe a Pythagorean triangle formed by the dashed line segments.

The function $V(h)$ is positive for h in the range $0 \leq h \leq 2L$, and $V = 0$ at the two endpoints of the interval. We can restrict attention to this interval since otherwise $V < 0$, which makes no physical sense. Since $V(h)$ is a smooth function, we anticipate that somewhere inside this range of values there should be a maximal volume.

Computing first and second derivatives, we find

$$V'(h) = \frac{\pi}{4}\left(L^2 - \frac{3}{4}h^2\right), \quad V''(h) = \frac{\pi}{4}\left(0 - 2\cdot\frac{3}{4}h\right) = -\frac{3}{8}\pi h < 0.$$

Setting $V'(h) = 0$ to find critical points, we then solve for h:

$$\begin{aligned} V'(h) = 0 \quad &\Rightarrow \quad L^2 - \frac{3}{4}h^2 = 0 \quad \Rightarrow \quad 3h^2 = 4L^2 \\ &\Rightarrow \quad h^2 = 4\frac{L^2}{3} \quad \Rightarrow \quad h = 2\frac{L}{\sqrt{3}}. \end{aligned}$$

We verify that this solution is a local *maximum* by the following reasoning.

The second derivative $V''(h) = -\frac{3}{8}\pi h < 0$ is always negative for any positive value of h, so $V(h)$ is concave down for $h > 0$, which confirms a local maximum. We also noted that $V(r)$ is smooth, positive within the range of interest and zero at the endpoints. As there is only one critical point in that range, it must be a local maximum.

Finally, we find the radius of the barrel by plugging the optimal h into the constraint equation, i.e. using

$$r^2 = \frac{1}{4}\left(L^2 - \frac{h^2}{4}\right) = \frac{1}{4}\left(L^2 - \frac{L^2}{3}\right) = \frac{1}{4}\left(\frac{2}{3}L^2\right)$$

$$\Rightarrow \quad r = \frac{1}{\sqrt{3}\sqrt{2}}L.$$

The shape of the optimal barrel can now be characterized. One way to do so is to specify the ratio of its height to its radius. (Tall skinny barrels have a largeh/r ratio, and squat fat ones have a low ratio.) By the above reasoning, the ratio of h/r for the optimal barrel is

$$\frac{h}{r} = \frac{2\frac{L}{\sqrt{3}}}{\frac{1}{\sqrt{3}\sqrt{2}}L} = 2\sqrt{2}. \tag{7.5}$$

Hence, for greatest economy, Kepler would have purchased barrels with height to radius ratio of $2\sqrt{2} = 2.82 \approx 3$. ◊

Mastered Material Check

9. If all barrels had a radius of 25cm, given the result Example 7.4, what would be the best barrel height?
10. What would the volume of such a barrel be?
11. Consider a barrel with radius 25cm and height 100cm. What is this barrel's volume?

7.3 Checking endpoints

Section 7.3 Learning goals

1. Recognize the distinction between local and global extrema.
2. Find the global minimum or maximum in a given word problem.

In some cases, the optimal value of a function does not occur at any of its local maxima, but rather at one of the endpoints of an interval. Here we consider such an example.

Example 7.5 (Maximal perimeter) *The area of a rectangle with side lengths x and y is $A = xy$. Suppose that the variable x is only allowed to take on values in the range $0.5 \le x \le 4$. Find the dimensions of the rectangle having* largest perimeter *whose fixed area is $A = 1$.*

Solution. The perimeter of a rectangle whose sides are length x, y is

$$P = x + y + x + y = 2x + 2y.$$

We are to maximize this quantity subject to the area of the rectangle being fixed, $A = xy = 1$. This is the constraint. We use it to solve for and to eliminate y from P.

$$y = \frac{1}{x}, \quad \Rightarrow \quad P(x) = 2x + \frac{2}{x}.$$

We look for x that maximizes $P(x)$. Computing the derivatives,

$$P'(x) = 2\left(1 - \frac{1}{x^2}\right), \quad P''(x) = \frac{4}{x^3} > 0.$$

Setting $P'(x) = 0$, we find critical points satisfying $x^2 = 1$ or $x = \pm 1$. We reject the negative root as irrelevant. We have found that $P''(x) > 0$ for all $x > 0$, so the critical point is a local *minimum*! This is clearly not the maximum we were looking for. This example reinforces the importance of diagnostic tests for the type of critical point.

Next, checking the endpoints of the interval, we evaluate $P(4) = 8.5$ and $P(0.5) = 5$. The largest perimeter for the rectangle thus occurs when $x = 4$, at the right endpoint of the domain, as shown in Figure 7.5. ♢

In Appendix G.4, we provide further examples of optimization in the context of geometric solids.

Mastered Material Check

12. How do you calculate the perimeter of a rectangle?
13. Why can a negative root of $P'(x) = 0$ in Example 7.5 be rejected as irrelevant?

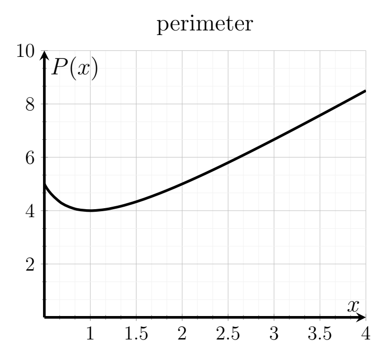


Figure 7.5: In Example 7.5, the critical point is a local minimum. The maximum occurs at the right end point of the interval $0.5 \le x \le 4$.

7.4 Optimal foraging

Section 7.4 Learning goals

1. Explain the development of a simple model for an animal foraging (collecting food to gain energy) in a food patch.
2. Interpret graphs of the rate of energy gain in various food patches, and explain the distinctions between types of food patches.
3. Determine how long to spend foraging in a food patch in order to optimize the average rate of energy gain per unit time.

Mastered Material Check

14. Use Figure 7.5 to estimate the side length x when $P(x) = 6$.
15. Verify your estimate algebraically.

Animals spend much of their time **foraging** - searching for food. Time is limited, since when the sun goes down, the risk of becoming food (to a predator) increases, and the likelihood of finding food decreases. Individuals who are most successful at finding food over that limited time have the greatest chance of surviving. It is argued by biologists that *evolution* tends to optimize animal behaviour by selecting those that are faster, stronger, or more fit, or - in this case - most efficient at finding food.

In this section, we investigate a model for optimal foraging. We follow the basic principles of [Stephens and Krebs, 1986] and [Charnov, 1976].

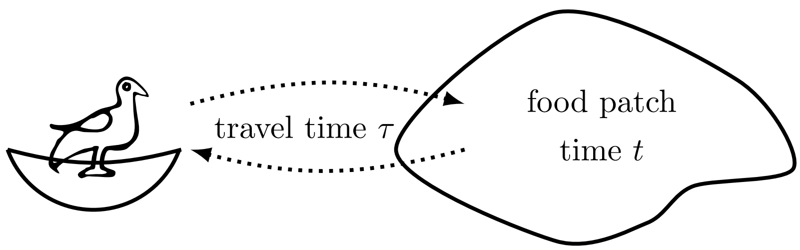


Figure 7.6: A bird travels daily to forage in a food patch. We want to determine how long it should stay in the patch to optimize its overall average energy gain per unit time.

Notation. We define the following notation:

- τ= travel time between nest and **food patch** (this is considered to be time that is unavoidably wasted).
- t = **residence time** in the patch (i.e. how long to spend foraging in one patch), also called **foraging time**,
- $f(t)$ = total energy gained by foraging in a patch for time t.

Energy gain in food patches. In some patches, food is ample and found quickly, while in others, it takes time and effort to obtain. The typical time needed to find food is reflected by various energy gain functions $f(t)$ shown in Figure 7.7.

Mastered Material Check

16. Which of the energy gain functions in Figure 7.7 are strictly increasing?

Example 7.6 (Energy gain versus patch residence time) *For each panel in Figure 7.7, explain what the graph of the total energy gain $f(t)$ is saying about the type of food patch: how easy or hard is it to find food?*

Solution. The types of food patches are as follows:

1. The energy gain is linearly proportional to time spent in the patch. In this case, the patch has so much food that it is never depleted. It would make sense to stay in such a patch for as long as possible.
2. Energy gain is independent of time spent. The animal gets the full quantity as soon as it gets to the patch.

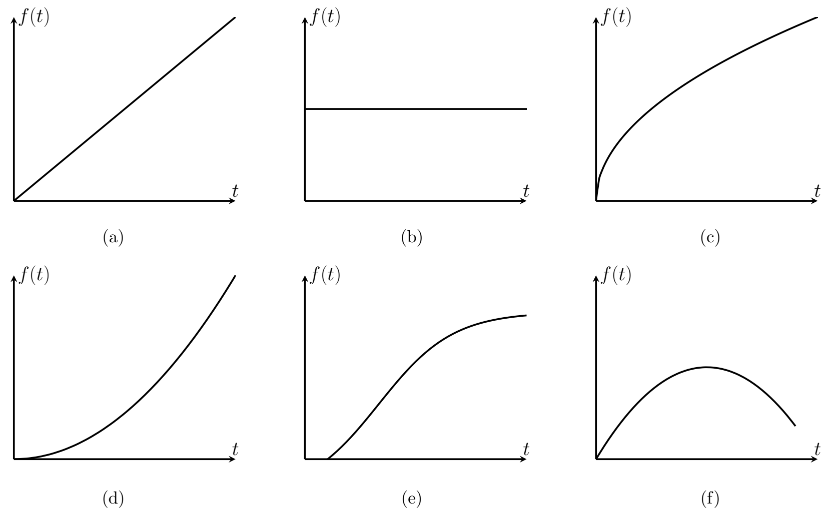


Figure 7.7: Examples of various total energy gain $f(t)$ for a given foraging time t. The shapes of these functions determine how hard or easy it is to extract food from a food patch.

3. Food is gradually depleted, (the total energy gain levels off to some constant as t increases). There is "diminishing return" for staying longer, suggesting that it is best not to stay too long.

4. The reward for staying longer in this patch increases: the net energy gain is concave up ($f''(t) > 0$), so its slope is increasing.

5. It takes time to begin to gain energy. After some time, the gain increases, but eventually, the patch is depleted.

6. Staying too long in such a patch is disadvantageous, resulting in net loss of energy. It is important to leave this patch early enough to avoid that loss.

◊

Mastered Material Check

17. Which model(s) can you automatically dismiss as not very biologically realistic?

Example 7.7 *Consider the hypothetical patch energy function*

$$f(t) = \frac{E_{max} t}{k+t} \quad \textit{where} \quad E_{max}, k > 0, \textit{ are constants.} \qquad (7.6)$$

a) *Match this function to one of the panels in Figure 7.7.*

b) *Interpret the meanings of the constants E_{max}, k.*

Solution.

a) The function resembles Michaelis-Menten kinetics (Figure 1.8). In Figure 7.7, Panel (3) is the closest match.

b) From Chapter 1, E_{max} is the horizontal asymptote, corresponding to an upper bound for the total amount of energy that can be extracted from the patch. The parameter k has units of time and controls the steepness of the function. Foraging for a time $t = k$, leads the animal to obtain half of the total available energy, since $f(k) = E_{max}/2$ (Exercise 27(a)). ◊

Example 7.8 (Currency to optimize) *We can assume that animals try to maximize the* average energy gain per unit time*, defined by the ratio:*

$$R(t) = \frac{\text{Total energy gained}}{\text{total time spent}},$$

Write down $R(t)$ for the assumed patch energy function Eqn. 7.6.

Solution. The 'total time spent' is a sum of the fixed amount of time τ traveling, and time t foraging. The 'total energy gained' is $f(t)$. Thus, for the patch function $f(t)$ assumed in Eqn. (7.6),

$$R(t) = \frac{f(t)}{(\tau+t)} = \frac{E_{max}t}{(k+t)(\tau+t)}. \tag{7.7}$$

◊

Mastered Material Check

18. What units might be used in the function $R(t)$?

We can now state the mathematical problem:

Find the time t that maximizes $R(t)$.

In finding such a t we are determining **the optimal residence time**.

Example 7.9 *Use tools of calculus and curve-sketching to find and classify the critical points of $R(t)$ in Eqn. (7.7).*

Solution. We first sketch $R(t)$, focusing on $t > 0$ for biological relevance.

- For $t \approx 0$, we have $R(t) \approx (E_{max}/k\tau)t$, which is a linearly increasing function.
- As $t \to \infty$, $R(t) \to E_{max}t/t^2 \to 0$, so the graph eventually decreases to zero.

These two conclusions are shown in Figure 7.8 (left panel), and strongly suggest that there should be a local maximum in the range $0 < t < \infty$, as shown in the right panel of Fig 7.8. Since the function is continuous for $t > 0$, this

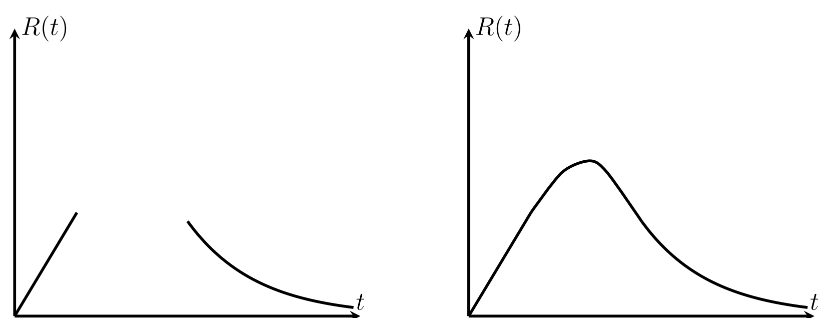


Figure 7.8: In Example 7.9 we first compose a rough sketch of the average rate of energy gain $R(t)$ in Eqn. (7.7). The graph is linear near the origin, and decays to zero at large t.

sketch verifies that there is a local maximum for some positive t value.

To find a local maximum, we compute $R'(t)$ using the quotient rule (Exercise 27c), and set $R'(t) = 0$:

$$R'(t) = E_{max} \frac{k\tau - t^2}{(k+t)^2(\tau+t)^2} = 0. \tag{7.8}$$

This can only be satisfied if the numerator is zero, that is

$$k\tau - t^2 = 0 \quad \Rightarrow \quad t_{1,2} = \pm\sqrt{k\tau}.$$

Rejecting the (irrelevant) negative root, we deduce that the critical point of the function $R(t)$ is $t_{crit} = \sqrt{k\tau}$. The sketch in Figure 7.8, verifies that this critical point is a local maximum. ♢

Example 7.10 *For practice, use one of the calculus tests for critical points to show that $t_{crit} = \sqrt{k\tau}$ is a local maximum for the function $R(t)$ in Eqn. (7.7).*

Solution. $R(t)$ is a rational function, so a second derivative is messy. Instead, we apply the first derivative test (Section 6.2), that is, we check the sign of $R'(t)$ on both sides of the critical point.

- Eqn. (7.8) gives $R'(t)$. Its denominator is positive, so the sign of $R'(t)$ is determined by its numerator, $(k\tau - t^2)$.
- Thus, $R'(t) > 0$ for $t < t_{crit}$, and $R'(t) < 0$ for $t > t_{crit}$.

This confirms that the function increases up to the critical point and decreases afterwards, so the critical point is a local maximum, henceforth denoted t_{max}. ♢

To optimize the average rate of energy gain, $R(t)$, we found that the animal should stay in the patch for a duration of $t = t_{max} = \sqrt{k\tau}$.

Mastered Material Check

19. Given t_{max} is the duration of time an animal should stay in a patch, and τ is travelling time, explain why the constant k is also in units of time.

Example 7.11 *Determine the average rate of energy gain at this optimal patch residence time, i.e. find the maximal average rate of energy gain.*

Solution. Computing $R(t)$ for $t = t_{max} = \sqrt{k\tau}$, we find that

$$R(t_{max}) = \frac{E_{max}t_{max}}{(k+t_{max})(\tau+t_{max})} = \frac{E_{max}}{\tau}\frac{1}{(1+\sqrt{k/\tau})^2}. \tag{7.9}$$

The reader is asked to fill in the steps for this calculation in Exercise 27(d). ♢

In Appendix G.5, we extend this example by considering a more general problem. Geometric arguments play a key role in that discussion.

7.5 Summary

1. Optimization is a process of finding critical points, and identifying local and global maxima/minima.
2. A scientific problem that address “biggest/smallest, best, most efficient” is often reducible to an optimization problem.

3. As with all mathematical models, translating scientific observations and reasonable assumptions into mathematical terms is an important first step.
4. The following applications were considered:

   (a) Density dependent population growth. Using a given logistic growth law, the following parameters were considered:

   - population growth rate (to be maximized),
   - population density,
   - intrinsic growth rate (constant),
   - carrying capacity (constant).

   (b) Nutrient absorption in a cell. Using the model developed in Section 1.2 for a spherical cell, we considered:

   - nutrient absorption rate,
   - nutrient consumption rate,
   - cell radius,
   - proportionality constants (determined based on context).

   We maximized the net rate of increase of nutrients - a difference between absorption and consumption rates.

   (c) Surface area of a cylindrical cell, which tends to be minimized do to energy conditions. The parameters we used were:

   - cell length,
   - cell radius,
   - cell volume (constant),
   - cell surface area (to be minimized).

   (d) Wine for Kepler's wedding, seeking the largest barrel volume for a fixed diagonal length. The following parameters were considered:

   - barrel volume, (to be maximized)
   - barrel height,
   - barrel radius,
   - length of the diagonal (constant).

   (e) Foraging time for an animal collecting food. We considered:

   - travel time between nest and food patch,
   - foraging time in the patch,
   - energy gained by foraging in a patch for various time durations.

Quick Concept Checks

1. If the growth rate of a population follows the following logistic equation:

$$G(N) = 1.2N\left(\frac{50000 - N}{50000}\right),$$

where N is the density of the population, under what circumstances is the population growing fastest?

2. When finding a global maximum, why is always imperative to check the endpoints?

3. Demonstrate the variability of barrel dimensions by giving two different height and radius pairs which lead to a volume of 50L.

4. Would the answer to Kepler's wine barrel problem have changed if we had solved for h^2 instead of r^2?

Exercises

7.1. **Find the numbers.** The sum of two positive number is 20. Find the numbers

(a) if their product is a maximum,

(b) if the sum of their squares is a minimum,

(c) if the product of the square of one and the cube of the other is a maximum.

7.2. **Distance, velocity and acceleration.** A tram ride departs from its starting place at $t = 0$ and travels to the end of its route and back. Its distance from the terminal at time t can be approximately described by the expression

$$S(t) = 4t^3(10-t)$$

where t is in minutes, $0 < t < 10$, and S is distance in meters.

(a) Find the velocity as a function of time.

(b) When is the tram moving at the fastest rate?

(c) At what time does it get to the furthest point away from its starting position?

(d) Sketch the acceleration, the velocity, and the position of the tram on the same set of axes.

7.3. **Distance of two cars.** At 9*A.M.*, car B is 25 km west of car A. Car A then travels to the south at 30 km/h and car B travels east at 40 km/h. When are they closest to each other and what is this distance?

7.4. **Cannonball movement.** A cannonball is shot vertically upwards from the ground with initial velocity $v_0 = 15\text{m/sec}$. The height of the ball, y (in meters), as a function of the time, t (in sec) is given by

$$y = v_0 t - 4.9t^2$$

Determine the following:

(a) the time at which the cannonball reaches its highest point,

(b) the velocity and acceleration of the cannonball at $t = 0.5$ s, and $t = 1.5$ s, and

(c) the time at which the cannonball hits the ground.

7.5. **Net nutrient increase rate.** In Example 7.2, we considered the net rate of increase of nutrients in a spherical cell of radius r. Here we further explore this problem.

(a) Draw a sketch of $N(r)$ based on Eqn. (7.1). Use your sketch to verify that this function has a local maximum.

(b) Use the first derivative test to show that the critical point $r = 2k_1/k_2$ is a local maximum.

7.6. **Nutrient increase in cylindrical cell.** Consider a long skinny cell in the shape of a cylinder with radius r and a fixed length L. The volume and surface area of such a cell (neglecting endcaps) are $V = \pi r^2 L = K$ and $S = 2\pi r L$.

(a) Adapt the formula for net rate of increase of nutrients $N(t)$ for a spherical cell Eqn. (7.1) to the case of a cylindrical cell.

(b) Find the radius of the cylindrical cell that maximizes $N(t)$. Be sure to verify that you have found a local maximum.

7.7. **Cylinder of minimal surface area.** In this exercise we continue to explore Example 7.3.

(a) Reason that the surface area of the cylinder, $S(r) = 2\frac{K}{r} + 2\pi r^2$ is a function that has a local minimum using curve-sketching.

(b) Use the first derivative test to show that $r = \left(\frac{K}{2\pi}\right)^{1/3}$ is a local minimum for $S(r)$.

(c) Show the algebra required to find the value of L corresponding to this r value and show that $L/r = 2$.

7.8. **Dimensions of a box.** A closed 3-dimensional box is to be constructed in such a way that its volume is 4500 cm^3. It is also specified that the length of the base is 3 times the width of the base.

Determine the dimensions of the box which satisfy these conditions and have the minimum possible surface area. Justify your answer.

7.9. **Dimensions of a box.** A box with a square base is to be made so that its diagonal has length 1; see Figure 7.9.

(a) What height y would make the volume maximal?

(b) What is the maximal volume? (*hint:* a box having side lengths ℓ, w, h has diagonal length D where $D^2 = \ell^2 + w^2 + h^2$ and volume $V = \ell w h$).

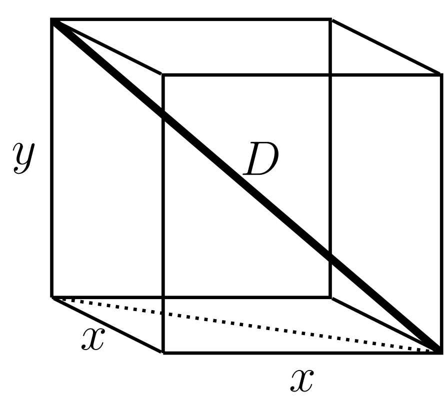


Figure 7.9: Figure for Exercise 9; box with a square base.

7.10. **Minimum distance.** Find the minimum distance from a point on the positive x-axis $(a,0)$ to the parabola $y^2 = 8x$.

7.11. **The largest garden.** You are building a fence to completely enclose part of your backyard for a vegetable garden. You have already purchased material for a fence of length 100 ft.

What is the largest rectangular area that this fence can enclose?

7.12. **Two gardens.** A fence of length 100 ft is to be used to enclose two gardens. One garden is to have a circular shape, and the other to be square.

Determine how the fence should be cut so that the sum of the areas inside both gardens is as large as possible.

7.13. **Dimensions of an open box.** A rectangular piece of cardboard with dimension 12 cm by 24 cm is to be made into an open box (i.e., no lid) by cutting out squares from the corners and then turning up the sides.

Find the size of the squares that should be cut out if the volume of the box is to be a maximum.

7.14. **Alternate solution to Kepler's wine barrel.** In this exercise we follow an alternate approach to the most economical wine barrel problem posed by Kepler (as in Example 7.4).

Through this approach, we find the proportions (height:radius) of the cylinder that minimizes the length L of the wet rod in Figure 7.3 for a fixed volume.

(a) Explain why minimizing L is equivalent to minimizing L^2 in Eqn. (7.4)

(b) Explain how Eqn. (7.3) can be used to specify a constraint for this problem. (*hint:* consider the volume, V to be fixed and show that you can solve for r^2).

(c) Use your result in (c) to eliminate r from the formula for L^2. Now $L^2(h)$ depends only on the height of the cylindrical wine barrel.

(d) Use calculus to find any local minima for $L^2(h)$. Be sure to verify that your result is a minimum.

(e) Find the corresponding value of r using your result in (b).

(f) Find the ratio h/r. You should obtain the same result as in Eqn. (7.5).

7.15. **Rectangle with largest area.** Find the side lengths, x and y, of the rectangle with largest area whose diameter L is given (*hint:* eliminate one variable using the constraint. To simplify the derivative, consider that critical points of A would also be critical points of A^2, where $A =$

xy is the area of the rectangle. If you have already learned the chain rule, you can use it in the differentiation).

7.16. **Shortest path.** Find the shortest path that would take a milk-maid from her house at $(10,10)$ to fetch water at the river located along the x-axis and then to the thirsty cow at $(3,5)$.

7.17. **Water and ice.** Why does ice float on water? Because the density of ice is lower! In fact, water is the only common liquid whose maximal density occurs above its freezing temperature. This phenomenon favours the survival of aquatic life by preventing ice from forming at the bottoms of lakes. According to the *Handbook of Chemistry and Physics,* a mass of water that occupies one liter at 0°C occupies a volume (in liters) of

$$V = -aT^3 + bT^2 - cT + 1$$

at T°C where $0 \leq T \leq 30$ and where the coefficients are

$$a = 6.79 \times 10^{-8}, \;\; b = 8.51 \times 10^{-6}, \;\; c = 6.42 \times 10^{-5}.$$

Find the temperature between 0°C and 30°C at which the density of water is the greatest. (*hint:* maximizing the density is equivalent to minimizing the volume. Why is this?).

7.18. **Drug doses and sensitivity.** The *reaction* $R(x)$ of a patient to a drug dose of size x depends on the type of drug. For a certain drug, it was determined that a good description of the relationship is:

$$R(x) = Ax^2(B - x)$$

where A and B are positive constants. The *sensitivity* of the patient's body to the drug is defined to be $R'(x)$.

(a) For what value of x is the reaction a maximum, and what is that maximum reaction value?

(b) For what value of x is the sensitivity a maximum? What is the maximum sensitivity?

7.19. **Thermoregulation in a swarm of bees.** In the winter, honeybees sometimes escape the hive and form a tight swarm in a tree, where, by shivering, they can produce heat and keep the swarm temperature elevated.

Heat energy is lost through the surface of the swarm at a rate proportional to the surface area ($k_1 S$ where $k_1 > 0$ is a constant). Heat energy is produced inside the swarm at a rate proportional to the mass of the swarm (which you may take to be a constant times the volume). We assume that the heat production is $k_2 V$ where $k_2 > 0$ is constant.

Swarms that are not large enough may lose more heat than they can produce, and then they die. The heat depletion rate is the loss rate minus the production rate. Assume that the swarm is spherical.

Find the size of the swarm for which the rate of depletion of heat energy is greatest.

7.20. **Cylinder inside a sphere.** Work through the steps for the calculations and classification of critical point(s) in Example G.2, that is, find the dimensions of the largest cylinder that would fit in a sphere of radius R.

7.21. **Circular cone circumscribed about a sphere.** A right circular cone is circumscribed about a sphere of radius 5. Find the dimension of this cone if its volume is to be a minimum.

Note: this is a rather challenging geometric problem.

7.22. **Optimal reproductive strategy.** Animals that can produce many healthy babies that survive to the next generation are at an evolutionary advantage over other, competing, species. However, too many young produce a heavy burden on the parents (who must feed and care for them). If this causes the parents to die, the advantage is lost. Further, competition of the young with one another for food and parental attention jeopardizes the survival of these babies.

Suppose that the evolutionary **Advantage** A to the parents of having litter size x is

$$A(x) = ax - bx^2.$$

Suppose that the **Cost** C to the parents of having litter size x is

$$C(x) = mx + e.$$

The **Net Reproductive Gain** G is defined as

$$G = A - C.$$

(a) Explain the expressions for A, C and G.

(b) At what litter size is the advantage, A, greatest?

(c) At what litter size is there least cost to the parents?

(d) At what litter size is the Net Reproductive Gain greatest?.

7.23. **Behavioural Ecology.** Social animals that live in groups can spend less time scanning for predators than solitary individuals. However, they waste time fighting with the other group members over the available food. There is some group size at which the net benefit is greatest because the animals spend the least time on these unproductive activities - and thus can spend time on feeding, mating, etc.

Assume that for a group of size x, the fraction of time spent scanning for predators is

$$S(x) = A\frac{1}{(x+1)}$$

and the fraction of time spent fighting with other animals over food is

$$F(x) = B(x+1)^2$$

where A, B are constants.

Find the size of the group for which the time wasted on scanning and fighting is smallest.

7.24. **Logistic growth.** Consider a fish population whose density (individuals per unit area) is N, and suppose this fish population grows **logistically**, so that the rate of growth R satisfies

$$R(N) = rN(1 - N/K)$$

where r and K are positive constants.

(a) Sketch R as a function of N or explain Figure 7.10.

(b) Use a first derivative test to justify the claim that $N = K/2$ is a local maximum for the function $G(N)$.

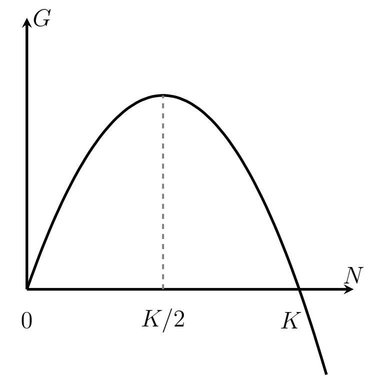


Figure 7.10: In logistic growth, the population growth rate G depends on population size N as shown here.

7.25. **Logistic growth with harvesting.** Consider a fish population of density N growing logistically, i.e. with rate of growth $R(N) = rN(1 - N/K)$ where r and K are positive constants. The rate of harvesting (i.e. removal) of the population is

$$h(N) = qEN$$

where E, the effort of the fishermen, and q, the catchability of this type of fish, are positive constants.

At what density of fish does the growth rate exactly balance the harvesting rate? This density is called the maximal sustainable yield: MSY.

7.26. **Conservation of a harvested population.** Conservationists insist that the density of fish should never be allowed to go below a level at which growth rate of the fish exactly balances with the harvesting rate. At this level, the harvesting is at its maximal sustainable yield. If more fish are taken, the population keeps dropping and the fish eventually go extinct.

What level of fishing effort should be used to lead to the greatest harvest at this maximal sustainable yield?

Note: you should first complete the Exercise 25.

7.27. **Optimal foraging.** Consider Example 7.7 for the optimal foraging model.

(a) Show that the parameter k in Eqn. (7.6) is the time at which $f(t) = E_{max}/2$.

(b) Consider panel (5) of Figure 7.7. Show that a function such as a Hill function would have the shape shown in that sketch. Interpret any parameters in that function.

(c) Use the quotient rule to calculate the derivative of the function $R(t)$ given by Eqn. (7.7) and show that you get Eqn. (7.8).

(d) Fill in the missing steps in the calculation in Eqn. (7.9) to find the optimal value of $R(t)$.

7.28. **Rate of net energy gain while foraging and traveling.** Animals spend energy in traveling and foraging. In some environments this energy loss is a significant portion of the energy budget. In such cases, it is customary to assume that to survive, an individual would optimize the rate of *net* energy gain, defined as

$$Q(t) = \frac{\text{Net energy gained}}{\text{total time spent}} = \frac{\text{Energy gained} - \text{Energy lost}}{\text{total time spent}} \qquad (7.10)$$

Assume that the animal spends p energy units per unit time in all activities (including foraging and traveling). Assume that the energy gain in the patch ("patch energy function") is given by Eqn. (7.6).

Find the optimal patch time, that is the time at which $Q(t)$ is maximized in this scenario.

7.29. **Maximizing net energy gain:** Suppose that the situation requires an animal to maximize its net energy gained $E(t)$ defined as

$$\begin{aligned} E(t) \quad &= \quad \text{energy gained while foraging} \\ &\qquad -\text{energy spent while foraging and traveling.} \end{aligned}$$

(This means that $E(t) = f(t) - r(t+\tau)$ where r is the rate of energy spent per unit time and τ is the fixed travel time).

Assume as before that the energy gained by foraging for a time t in the food patch is $f(t) = E_{max}t/(k+t)$.

(a) Find the amount of time t spent foraging that maximizes $E(t)$.

(b) Indicate a condition of the form $k < \boxed{?}$ that is required for existence of this critical point.

8
Introducing the chain rule

So far, examples were purposefully chosen to focus on power, polynomial, and rational functions that are each relatively easy to differentiate. We now introduce the differentiation rule that opens up our repertoire to more elaborate examples involving **composite functions**. This allows us to model more biological processes. We dedicate this chapter to the **chain rule** and its applications.

Mastered Material Check

1. Give an example each of:
   (a) a power function,
   (b) a polynomial function, and
   (c) a rational function.

8.1 The chain rule

Section 8.1 Learning goals

1. Summarize function composition and express a composite function in terms of the underlying composed functions.
2. Produce the chain rule of differentiation and apply it to find the derivative of a composite function.

Function composition

Consider Figure 8.1 which depicts a function composition: an independent variable, x, is used to evaluate a function, and the result, $u = f(x)$ then acts as an input to a second function, g. The final value is $y = g(u) = g(f(x))$.

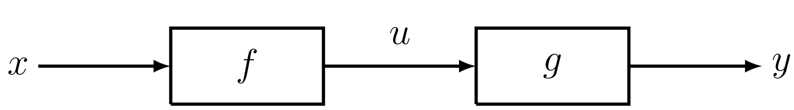


Figure 8.1: Function composition.

We refer to this two-step function operation as **function composition**.

Example 8.1 *Consider the functions* $f(x) = \sqrt{x}$ *and* $g(x) = x^2 + 1$. *Determine the functions obtained by composing these:*

a) $h_1(x) = g(f(x))$

b) $h_2(x) = f(g(x))$

Solution.

a) For h_1 we apply f first, followed by g, so $h_1(x) = (\sqrt{x})^2 + 1 = x + 1$ (provided $x \geq 0$.)

b) For h_2, the functions are applied in the reversed order so that $h_2(x) = \sqrt{x^2 + 1}$ (for any real x).

We note that the domains of the two functions are slightly different: h_1 is only defined for $x \geq 0$ since $f(x)$ is not defined for negative x, whereas h_2 is defined for all x. ◊

Mastered Material Check

2. Can you define the domain of a function?
3. What are the domains of $f(x)$ and $g(x)$ found in Example 8.2?

Example 8.2 *Express the function $h(x) = 5(x^3 - x^2)^{10}$ as the composition of two simpler functions.*

Solution. We can write this in terms of the two functions $f(x) = x^3 - x^2$ and $g(x) = 5x^{10}$. Then $h(x) = g(f(x))$. ◊

The chain rule of differentiation

Given a composite function $y = f(g(x))$, we require a rule for differentiating y with respect to x.

If $y = g(u)$ and $u = f(x)$ are both differentiable functions and $y = g(f(x))$ is the composite function, then the **chain rule** of differentiation states that

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.$$

Informally, the chain rule states that the change in y with respect to x is a product of two rates of change:

1. the rate of change of y with respect to its immediate input u, and
2. the rate of change of u with respect to its input, x.

Why does it work this way? Although the derivative is not a simple quotient, we gain an intuitive grasp of the chain rule by writing

$$\frac{dy}{dx} \approx \frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u}\frac{\Delta u}{\Delta x}$$

then it is apparent that the "cancellation" of terms Δu in numerator and denominator lead to the correct fraction on the left. The proof of the chain rule uses this essential idea, but care is taken to ensure that the quantity Δu is nonzero, to avoid the embarrassment of dealing with the nonsensical ratio $0/0$. The proof of the chain rule is found in Appendix E.4.

Example 8.3 *Compute the derivative of the function* $h(x) = 5(x^3 - x^2)^{10}$.

Solution. We express the function as $y = h(x) = 5u^{10}$ where $u = (x^3 - x^2)$ and apply the chain rule. Then

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \left(\frac{d(5u^{10})}{du}\right)\left(\frac{d(x^3 - x^2)}{dx}\right) = 50u^9(3x^2 - 2x).$$

Substituting for u leads to $dy/dx = 50(x^3 - x^2)^9(3x^2 - 2x)$. ◊

Example 8.4 *Compute the derivative of the function* $y = f(x) = \sqrt{x^2 + a^2}$, *where a is some positive real number.*

A quick summary of the first and second derivatives of the function in Example 8.4 for the case $a = 1$.

Solution. This function can be considered as the composition of $g(u) = \sqrt{u} = u^{1/2}$ and $u(x) = x^2 + a^2$, that is, we can write $f(x) = g(h(x))$. Then using the chain rule, we obtain

$$\frac{dy}{dx} = \frac{1}{2} \cdot (x^2 + d^2)^{-1/2} \cdot 2x = \frac{x}{(x^2 + d^2)^{1/2}} = \frac{x}{\sqrt{x^2 + d^2}}.$$

◊

Example 8.5 *Compute the derivative of the function*

$$y = f(x) = \frac{x}{\sqrt{x^2 + d^2}},$$

where d is some positive real number.

Solution. We use both the quotient rule and the chain rule for this calculation.

$$\frac{dy}{dx} = \frac{[x]' \cdot \sqrt{x^2 + d^2} - [\sqrt{x^2 + d^2}]' \cdot x}{(\sqrt{x^2 + d^2})^2}.$$

Here the $[\ldots]'$ denotes differentiation. Then

$$\frac{dy}{dx} = \frac{1 \cdot \sqrt{x^2 + d^2} - [\frac{1}{2} \cdot 2x \cdot (x^2 + d^2)^{-1/2}] \cdot x}{(x^2 + d^2)}.$$

We simplify algebraically by multiplying top and bottom by $(x^2 + d^2)^{1/2}$ and cancelling factors of 2 to obtain

$$\frac{dy}{dx} = \frac{x^2 + d^2 - x^2}{(x^2 + d^2)^{1/2}(x^2 + d^2)} = \frac{d^2}{(x^2 + d^2)^{3/2}}.$$

◊

Interpreting the chain rule

While the chain rule is not rigorously proved here (see Appendix E.4), we hope to extend our intuition about where it comes from. The following intuitive examples may help to motivate why the chain rule is based on a product of two rates of change.

Example 8.6 (Pollution level in a lake) *A species of fish is sensitive to pollutants in a lake. As humans populate the area around the lake, the fish population declines due to increased pollution levels. Quantify the rate at which the pollution level changes with time based on the pollution produced per human and the rate of increase of the human population.*

Solution. The rate of change (decline in this case) of the fish population depends on:

- the rate of change of the human population, and
- the rate of change in the pollution created per person.

If either increases, the effect on the fish population increases.

The chain rule implies that the net effect is a product of the two rates. Formally, for t time in years, $x = f(t)$ the number of people at the lake in year t, and $p = g(x)$ the pollution created by x people, the rate of change of the pollution p over time is a product of $g'(x)$ and $f'(t)$:

$$\frac{dp}{dt} = \frac{dp}{dx}\frac{dx}{dt} = g'(x)f'(t).$$

◊

Mastered Material Check

4. If pollution is measured in micrograms per cubic meter ($\mu g/m^3$), give units for p and dp/dt.

Example 8.7 (Population of carnivores, prey, and vegetation) *The population of large carnivores, C, on the African savannah depends on the population of gazelles that are their prey, P. The gazelle population, in turn, depends on the abundance of vegetation V, which depends on the amount of rain in a given year, r. Quantify the rate of change of the carnivore population with respect to rainfall.*

Solution. We can express these dependencies through hypothetical functions such as $V = g(r)$, $P = f(V)$ and $C = h(P)$, each depicting a relationship in the food chain.

A drought that decreases rainfall also decreases the abundance of vegetation. This then decreases the gazelle population, and eventually affect the population of carnivores. The rate of change in the carnivores population with respect to the rainfall, dC/dr, according to the chain rule, would be

$$\frac{dC}{dr} = \frac{dC}{dP} \cdot \frac{dP}{dV} \cdot \frac{dV}{dr}.$$

◊

Example 8.8 (Population of carnivores) *Consider Example 8.7, with the specific example of the relationships of carnivores and prey, $C = h(P) = P^2$, prey on vegetation, $P = f(V) = 2V$, and vegetation on rainfall, $V = g(r) = r^{1/2}$ (Figure 8.2). Quantify the rate of change of the carnivore population with respect to rainfall.*

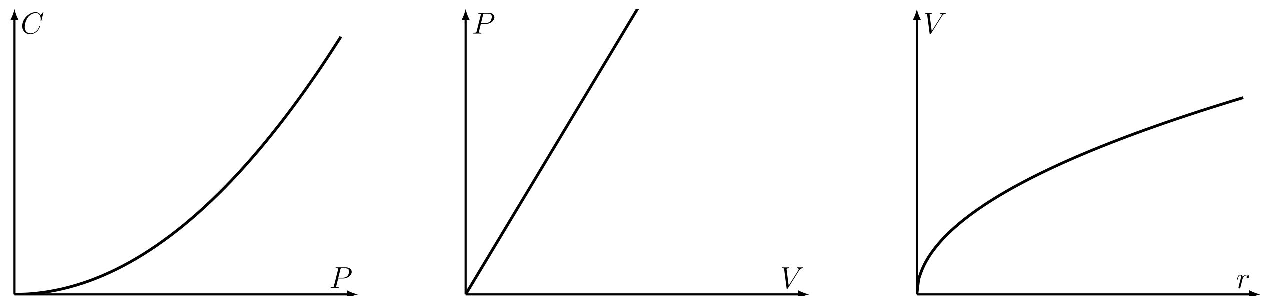


Figure 8.2: An example in which the population of carnivores, depends on prey P acording to $C = h(P) = P^2$, while the prey depend on vegetation, $P = f(V) = 2V$, and the vegetation depends on rainfall $V = g(r) = r^{1/2}$.

Solution. Computing all necessary derivatives,

$$\frac{dV}{dr} = \frac{1}{2}r^{-1/2}, \quad \frac{dP}{dV} = 2, \quad \frac{dC}{dP} = 2P,$$

so that, from Example 8.7, the rate of change in the carnivores with respect to the rainfall is

$$\frac{dC}{dr} = \frac{dC}{dP}\frac{dP}{dV}\frac{dV}{dr} = \frac{1}{2}r^{-1/2}(2)(2P) = \frac{2P}{r^{1/2}}.$$

Using the fact that $V = r^{1/2}$ and $P = 2V$, we obtain

$$\frac{dC}{dr} = \frac{2P}{V} = \frac{2(2V)}{V} = 4.$$

◊

Mastered Material Check

5. Take an alternative approach to Example 8.8 by expressing the number of carnivores C explicitly in terms of rainfall r, and then differentiating. Verify that both approaches yield the same solution.
6. While both approaches work in this case, why might they not in general?

Example 8.9 (Budget for coffee) *Your budget for coffee depends on the number of cups consumed per day and on the price per cup. The total budget changes if either price or the consumption goes up. Define appropriate variables and quantify the rate at which the coffee budget changes if both consumption and price change.*

Solution. The total rate of change of the coffee budget is a product of the change in the price and the change in the consumption. For t time in days, $x = f(t)$ the number of cups of coffee consumed, and $y = g(x)$ the price for x cups of coffee, we obtain

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} = g'(x)f'(t).$$

◊

Example 8.10 (Earth's temperature and greenhouse gases) *In Exercise 21 of Chapter 1, we found that the temperature of the Earth depends on the albedo a (fraction of incoming radiation energy reflected) according to the formula*

$$T = \left(\frac{(1-a)S}{\varepsilon\sigma}\right)^{1/4}. \tag{8.1}$$

Assume that the albedo a is the only quantity that depends on the level of greenhouse gases G, and that da/dG is known.Determine how the temperature changes as the level of greenhouse gases G increases.

Solution. Temperature T depends on the level of greenhouse gases G through the albedo a, so we write $T(a(G))$. Here S, ε, σ are all constants, so it simplifies calculation to rewrite T as

$$T(a) = \left(\frac{S}{\varepsilon\sigma}\right)^{1/4}(1-a)^{1/4}.$$

According to the chain rule,

$$\frac{dT}{dG} = \frac{dT}{da}\frac{da}{dG}.$$

Then

$$\frac{dT}{dG} = \left(\frac{S}{\varepsilon\sigma}\right)^{1/4}\frac{d}{da}\left[(1-a)^{1/4}\right]\frac{da}{dG} = \left(\frac{S}{\varepsilon\sigma}\right)^{1/4}\frac{1}{4}(1-a)^{(1/4)-1}\cdot(-1)\frac{da}{dG}.$$

Rearranging leads to

$$\frac{dT}{dG} = -\frac{1}{4}\left(\frac{S}{\varepsilon\sigma}\right)^{1/4}(1-a)^{-3/4}\frac{da}{dG}.$$

In general, greenhouse gasses affect both the Earth's albedo a and its emissivity ε. We generalize our results in Exercise 3. ◊

Mastered Material Check

7. List all constants in Example 8.10.
8. List all variables in Example 8.10.

8.2 *The chain rule applied to optimization problems*

Section 8.2 Learning goals

1. Read and interpret the derivation of each optimization model.
2. Carry out the calculations of derivatives appearing in the problems (using the chain rule).
3. Using optimization, find each critical point and identify its type.
4. Explain the interpretation of the mathematical results.

Armed with the chain rule, we can now differentiate a wider variety of functions, and address problems that were not tractable with the power, product, or quotient rules alone. We return to optimization problems where derivatives require use of the chain rule.

Shortest path from food to nest

Ants are good mathematicians! They can find the shortest route connecting their nest to a food source. But how do they do it? Each ant secretes a chemical **pheromone** that other ants tend to follow. This marks up the trail that they use and recruits nest-mates to food sources. The **pheromone** (chemical message for marking a route) evaporates after a while, so that, for a fixed given number of foraging ants, a longer trail has a less concentrated chemical marking than a shorter trail. This means that whenever a shorter route is found, the ants favour it. After some time, this leads to selection of the shortest possible trail.

Shown in Figure 8.3 is a common laboratory test scenario: ants in an artificial nest are offered two equivalent food sources. We ask: what is the shortest total path connecting nest and sources? This is a simplified version of the problem that the ants are solving.

Example 8.11 (Minimizing the total path length) *Use the diagram to determine the length of the shortest path that connects the nest to both food sources. Assume that* $d << D$*.*

Solution. We first consider two possibile paths connecting nest to food:

1. a **V**-shaped path and
2. a **T**-shaped path.

The length of a V-shaped path is $2\sqrt{D^2+d^2}$, whereas the length of a T-shaped path is $D+2d$.

Now consider a third possibility: a **Y**-shaped path, where the ants first walk straight ahead and then veer off to the left and right. All three possibilities are shown in Figure 8.3.

Mastered Material Check

9. If $D = 2$m and $d = 20$cm, how long is the V-shaped path? The T-shaped path?

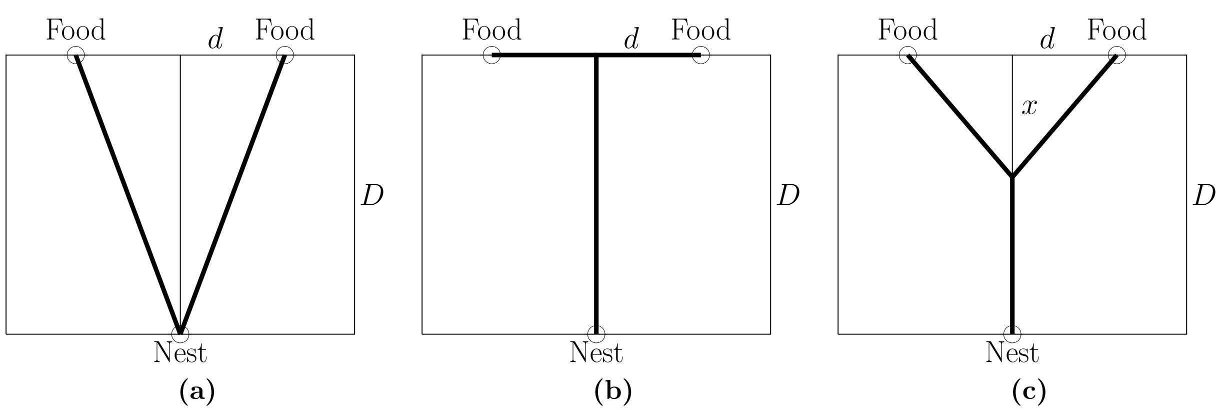


Figure 8.3: Three ways to connect the ants' nest to two food sources, showing (a) V-shaped, (b) T-shaped, and (c) Y-shaped paths.

Calculations are easiest if we denote the distance from the nest to the Y-junction as $D-x$, so that x is distance shown in the diagram. The length of the Y-shaped path is then

$$L_Y = L(x) = (D - x) + 2\sqrt{d^2 + x^2}. \quad (8.2)$$

Observe that when $x = 0$, then $L_T = D + 2d$, which corresponds to the T-shaped path length, whereas when $x = D$, then $L_V = 2\sqrt{d^2 + D^2}$ which V-shaped path length. Thus, $0 \le x \le D$ is the appropriate domain, and we have determined the values of L at the two domain endpoints.

To find the minimal path length, we look for critical points of the function $L(x)$. Differentiating (see Examples 8.4 and 8.5), we find

$$L'(x) = -1 + 2\frac{x}{\sqrt{x^2 + d^2}}, \quad L''(x) = 2\frac{d^2}{(x^2 + d^2)^{-3/2}} > 0.$$

Then critical points occur when

$$L'(x) = 0 \quad \Rightarrow \quad -1 + 2\frac{x}{\sqrt{x^2 + d^2}} = 0.$$

Simplifying leads to

$$\sqrt{x^2 + d^2} = 2x \quad \Rightarrow \quad x^2 + d^2 = 4x^2 \quad \Rightarrow \quad 3x^2 = d^2 \quad \Rightarrow \quad x = \frac{d}{\sqrt{3}}.$$

To determine the type of critical point, we note that the second derivative is positive and so the critical point is a local minimum.

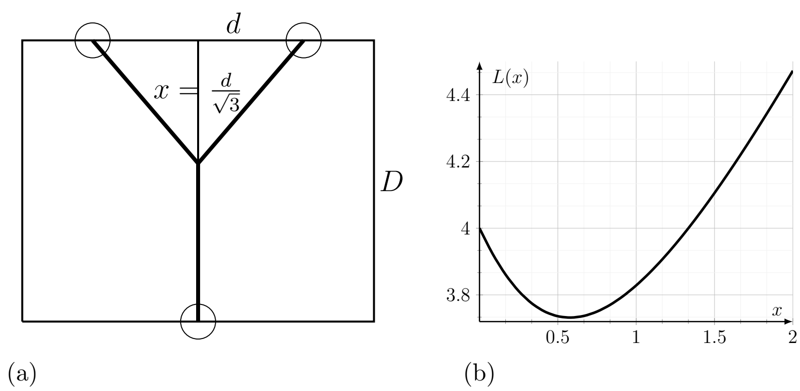


Figure 8.4: (a) In the configuration for the shortest path we found that $x = d/\sqrt{3}$. (b) The total length of the path $L(x)$ as a function of x for $D = 2, d = 1$. The minimal path occurs when $x = 1/\sqrt{3} \approx 0.577$. The length of the shortest path is then $L = D + \sqrt{3}d = 2 + \sqrt{3} \approx 3.73$.

Mastered Material Check

10. For $D = 2$m and $d = 20$cm, what is the shortest path length for the ants?

To determine the actual length of the path, we substitute $x = d/\sqrt{3}$ into the function $L(x)$ and obtain (after simplification, see Exercise 4)

$$L = L(x) = D + \sqrt{3}d.$$

The final result is summarized in Figure 8.4. The shortest path is Y-shaped, with $x = d/\sqrt{3}$. The ants march straight for a distance $D - (d/\sqrt{3})$, and then their trail branches to the right and left towards the food sources.

◊

Featured Problem 8.1 (Most economical wooden beam) *A cylindrical tree trunk is to be cut into a rectangular wooden beam. What is the most economical way to cut the beam so as to waste the least amount of material?*

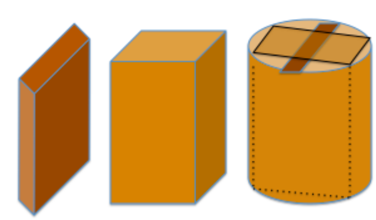


Figure 8.5: A cylindrical trunk can produce a variety of wooden beams. Which one uses the most wood and wastes as little as possible?

Food choice and attention

The following example is based on research about animal behaviour. As before, we consider how animals should spend their time, but now a choice is provided between two types of food.

Paying attention. Many types of food are **cryptic** - hidden in the environment - and so, require time and attention to find. Some types of food are more easily detected, but other foods might provide greater nourishment. Behavioural ecologist Reuven Dukas (McMaster U) studied how blue jays allocate attention to two food types [Dukas and Ellner, 1993, Dukas and Kamil, 2000, 2001]. The goal in the following problem is to find the optimal subdivision of time and attention between the food types so as to maximizes the total energy gain.

Setting up the model. Suppose that there are *two* available food types. Define notation as follows:

x	=	attention devoted to finding food of type 1
$P(x)$	=	probability of finding the food given attention x

We assume that $0 \le x \le 1$, with $x = 0$ representing no attention and $x = 1$ meaning full attention is devoted to finding food type 1.

Moreover, P is a probability which mean $0 \le P \le 1$. We assume that $P(0) = 0$ which means that when no attention is paid ($x = 0$) the probability of finding food is zero ($P = 0$). We also assume for simplicity that $P(1) = 1$, so when full attention paid $x = 1$, there is always success ($P = 1$).

Figure 8.6 displays hypothetical examples of $P(x)$. The horizontal axis is attention $0 \le x \le 1$, and the vertical axis, is the probability of success, $0 \le P \le 1$. All these curves share the assumed properties of full success with full attention, and no success with no attention. However, the curves differ in overall shape.

Questions.

1. What is the difference between foods of type 1 and 4?
2. Which food is easier to find, type 3 or type 4?
3. Compare the above to the case that $P(x) = x$ and explain what this new case implies.
4. What role is played by the concavity of the curve?

Observe that concave down curves such as 3 and 4 rise rapidly at small x, indicating that the probability of finding food increases a lot just by increasing the attention by a little: these represent foods that are relatively easy to find. Other curves (1 and 2) are concave up, indicting that much more attention is needed to gain appreciable increase in the probability of success: these represent foods that are harder to find. The concavity of the curves carries

Mastered Material Check

11. If $x = 1$ and full attention is devoted to finding food type 1, is any attention devoted to finding food type 2?
12. If $x = 0.5$, how much attention is being paid to finding food of type 1?

Mastered Material Check

13. Use Figure 8.6 to estimate the attention x needed to have a 50% probability of finding food type 1. That is, roughly estimate x such that $P(x) = 0.5$ for each of the probability curves.
14. If we fully divide attention between food types 1 and 2, and we spend 0.25 of our attention on finding food type 1, how much attention is given to food type 2? Converting the probability to a percentage may help with understanding

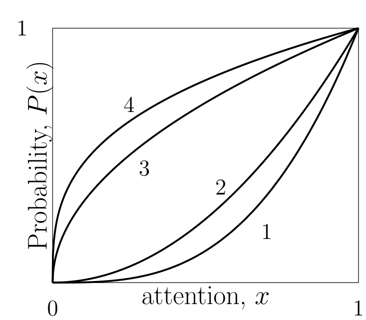


Figure 8.6: The probability, $P(x)$, of finding a food depends on the level of attention x devoted to finding that food. Here $0 \le x \le 1$, with $x = 1$ being “full attention”. We show possible curves for four types of foods, some easier to find than others.

this important information about the relative ease or difficulty of finding a given food type.

Let x and y denote the attention divided between food types 1 and 2, and suppose that type 2 is N-times more nutritious than type 1. We use $P_1(x)$ and $P_2(y)$ to denote the probabilities of finding food of type 1 and 2. Then the total nutritional value gained by splitting the attention is:

$$V(x) = P_1(x) + N\,P_2(y) = P_1(x) + N\,P_2(1-x).$$

The goal is to optimize nutritional value $V(x)$.

Adjust the slider to see how the relative nutritional value N of food type 2 affects the total value $V(x)$ as a function of the attention x. What type of critical point do we find?

Example 8.12 (**P_1 and P_2 as power function with integer powers**) *Consider the case that the probability of finding the food types is given by the simple power functions,*

$$P_1(x) = x^2, \quad P_2(y) = y^3.$$

See an explanation of Example 8.12.

(These functions satisfy $P(0) = 0, P(1) = 1$, in accordance with Figure 8.6.) Further, suppose that both foods are equally nutritious, so $N = 1$. Find the optimal $V(x)$.

Solution. The total nutritional value in this case is

$$V(x) = P_1(x) + N\,P_2(1-x) = x^2 + (1-x)^3.$$

We look for a maximum value of V. Using the chain rule to differentiate, we find that

$$\begin{aligned} V'(x) &= 2x + 3(1-x)^2(-1), \\ V''(x) &= 2 - 3(2)(1-x)(-1) = 2 + 6(1-x). \end{aligned}$$

Mastered Material Check

15. Within Example 8.12, calculate how much nutritional value is gained by the animal devoting $x = 0.4514$ of its attention to food type 1.
16. Consult Figure 8.7 to verify your result.

We observe that a negative factor (-1) comes from applying the chain rule to $(1-x)^3$. Setting $V'(x) = 0$ we get

$$\begin{aligned} 2x + 3(1-x)^2(-1) = 0. \quad &\Rightarrow \quad -3x^2 + 8x - 3 = 0 \\ &\Rightarrow \quad x = \frac{4 \pm \sqrt{7}}{3} \approx 0.4514, 2.21. \end{aligned}$$

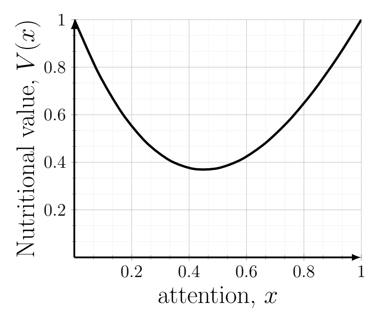


Figure 8.7: Figure for Example 8.12. The probabilities of finding foods of types 1 and 2 are concave up power functions. There is no local maximum.

Since attention takes on values in $0 \le x \le 1$, we reject the second root. The first root suggests that the animal should spend $\approx 0.45\%$ of its attention on food type 1 and the rest on type 2. However, to confirm such speculation, we must check whether the critical point is a maximum.

The second derivative is *positive* for all values of x in $0 \le xle1$, signifying a *local minimum*! The animal gains *least* by splitting its attention between two food typess in this case. Indeed, from Figure 8.7, we see that the most gain occurs at either $x = 0$ (only food type 2 sought) or $x = 1$ (only food type 1 sought).This example reemphasizes the importance of checking the type of critical point before drawing hasty conclusions. ◊

Example 8.13 (Fractional-power functions for P_1, P_2) *Now consider the case that the probability of finding two food types is given by the concave down power functions,*

$$P_1(x) = x^{1/2}, \quad P_2(y) = y^{1/3}$$

and both foods are equally nutritious ($N = 1$). Find the optimal food value $V(x)$.

Change the powers in the interactive graph to conform to Example 8.13. You can adjust the relative nutritional value, N. What type of critical point do we find?

See a brief recap of Example 8.13, and why we expect to find a local maximum in this case.

Solution. These functions also satisfy $P(0) = 0, P(1) = 1$, in accordance with the sketches shown in Figure 8.6. Then

$$V(x) = P_1(x) + P_2(1-x) = x^{1/2} + (1-x)^{(1/3)},$$

$$V'(x) = \frac{1}{2x^{1/2}} - \frac{1}{3(1.0-x)^{(2/3)}},$$

$$V''(x) = -\frac{1}{4x^{(3/2)}} - \frac{2}{9(1.0-x)^{(5/3)}}.$$

We must solve $V'(x) = 0$ to find the critical point. Unfortunately, this problem, turns out to be algebraically nasty. However, we can look for an approximate solution to the problem, using Newton's Method.

A plotting the graph of $V(x)$ in Figure 8.8 demonstrates that there is a maximum inside the interval $0 \leq x \leq 1$, i.e., for attention split between finding both foods. We further see from $V''(x)$ that the second derivative is negative for all values of x in the interval, indicating a local maximum, as expected. ◊

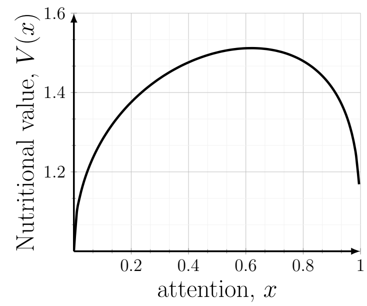


Figure 8.8: Figure for Example 8.13. The probabilities of finding foods of types 1 and 2 are both function that are concave down. As a result there is a local maximum for the nutritional value.

Applying Newton's method to find the critical point.

Example 8.14 *Use Newton's Method to find the critical point for the function $V(x)$ in Example 8.13.*

Solution. Finding the critical point of $V(x)$ reduces to solving $V'(x) = 0$. Let $f(x) = V'(x)$ - we must solve $f(x) = 0$ using Newton's Method (Recall Section 5.4).

Since the interval of interest is $0 \leq x \leq 1$, we start with an initial "guess" for the critical point at $x_0 = 0.5$, midway along this interval. Then, according to Newton's method, the improved guess would be

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

and, repeating this, at the k'th stage,

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

To use this method, carefully note that

$$f(x) = V'(x) = \frac{1}{2x^{1/2}} - \frac{1}{3(1.0-x)^{(2/3)}},$$

Mastered Material Check

17. Justify algebraically why $V''(x)$ in Example 8.13 is negative on the interval $0 \leq x \leq 1$.

18. Using Figure 8.8, what is the largest possible nutritional value?

$$f'(x) = V''(x) = -\frac{1}{4x^{(3/2)}} - \frac{2}{9(1.0-x)^{(5/3)}}.$$

Thus, we might use a spreadsheet in which cells A1 stores our initial guess, whereas B1, C1, and D1 store the values of $f(x), f'(x)$ and $x_0 - f(x_0)/f'(x_0)$. In the typical syntax of spreadsheets, this might read something like the following:

```
A1 0.5
B1 =(1/(2*(A1)^(1/2))-1/(3*(1-A1)^(2/3)))
C1 =(-1/(4*A1^(3/2))-2/(9*(1-A1)^(5/2)))
D1 =A1-B1/C1
```

Link to Google Sheets. This spreadsheet implements Newton's method for Example 8.14. You can view the formulae by clicking on a cell in the sheet but you cannot edit the sheet here.

This idea is implemented on a spreadsheet resulting in values shown in Table 8.1 starting from $x_0 = 0.5$. We see that the values converge to the location of the critical point, $x = 0.61977$ (and $y = 1-x = 0.38022$) within the interval of interest.

k	x_k	$f(x_k)$	$f'(x_k)$	x_{k+1}
0	0.50000	0.17797	-1.96419	0.59061
1	0.59061	0.04603	-2.62304	0.60816
2	0.60816	0.01866	-2.83923	0.61473
3	0.61473	0.00816	-2.93065	0.61751
4	0.61751	0.00367	-2.97130	0.61875
5	0.61875	0.00167	-2.98970	0.61931
6	0.61931	0.00076	-2.99809	0.61956

Table 8.1: Newton's method applied to Example 8.14. We start with $x_0 = 0.5$.

Epilogue. While the conclusions drawn above were disappointing in one specific case, it is not always true that concentrating all one's attention on one type is optimal. We can examine the problem in more generality to find when the opposite conclusion might be satisfied.

In general, the value gained is

$$V(x) = P_1(x) + N\,P_2(1-x).$$

A critical point occurs when

$$V'(x) = \frac{d}{dx}[P_1(x) + N\,P_2(1-x)] = P_1'(x) + NP_2'(1-x)(-1) = 0.$$

Mastered Material Check

19. Can you see where the (-1) comes from in $V'(x)$?

Suppose we have found a value of x in $0 < x < 1$ where this is satisfied. We then examine the second derivative:

$$V''(x) = \frac{d}{dx}[V'(x)] = \frac{d}{dx}[P_1'(x) - NP_2'(1-x)]$$
$$= P_1''(x) - NP_2''(1-x)(-1) = P_1''(x) + NP_2''(1-x).$$

The concavity of the function V is thus related to the concavity of the two functions $P_1(x)$ and $P_2(1-x)$. If these are concave down (e.g. as in food types 3 or 4 in Figure 8.6), then $V''(x) < 0$ and a local maximum occurs at any critical point found by our differentiation.

Another way of stating this observation is: if both food types are relatively easy to find, one can gain most benefit by splitting up the attention between the two. Otherwise, if both are hard to find, then it is best to look for only one at a time.

8.3 Summary

1. In this chapter we reviewed function composition, in which one function acts as the input to another function. The order of composition is important: $f(g(x)) \neq g(f(x))$.
2. The chain rule can be used to differentiate composite functions. If $y = g(u)$ and $u = f(x)$ are both differentiable, then

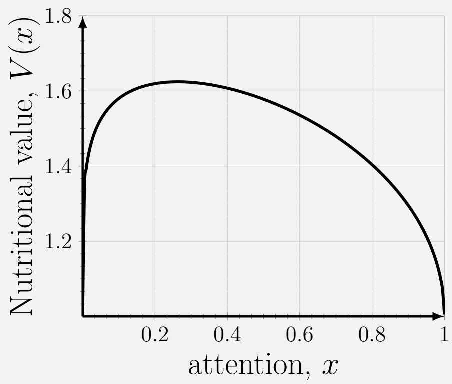
$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.$$

3. Applications seen in this chapter include:

   (a) pollution levels in a lake (depending on human population, and pollution created per person);
   (b) populations of carnivores (depending on populations of prey, vegetation, and rainfall);
   (c) budget for coffee (depending on amount consumed, and price per cup);
   (d) Earth's temperature (depending on the albedo, and green house gases);
   (e) ants' path to food (depending on locations of nest and food sources); and
   (f) food choice and attention (depending on probabilities of success, nutrition gained).

Quick Concept Checks

1. Write $f(x) = \sqrt{x^2(3x^2-1)^3}$ as the composition of two different functions in two different ways. Differentiate using the chain rule.
2. Write $f(x)$ as a composition of three different functions. Differentiate using the chain rule. Verify your solution matches that of **1.**
3. If an animal needs to divide its attention x between 4 food types, and $P_1(x) = 0.2$, $P_2(x) = 0.1$ and $P_3(x) = 0.5$, what is $P_4(x)$?
4. Consider the following graph depicting the nutritional value gained when x attention is applied to food type 1



(a) What is the maximum nutritional value that can be attained?

(b) How much attention should be paid to food type 1 in order to achieve this?

(c) Assuming there are only two different food types, how much attention should be paid to food type 2?

Exercises

8.1. **Practicing the chain rule.** Use the chain rule to calculate the following derivatives

(a) $y = f(x) = (x+5)^5$,

(b) $y = f(x) = 4(x^2+5x-1)^8$,

(c) $y = f(x) = (\sqrt{x}+2x)^3$.

8.2. **Growth curve.** An example of a growth curve in population biology is called the *Bertalanffy growth curve*, after Canadian biologist Ludwig von Bertalanffy. This curve is defineed by the equation

$$N = (a - b2^{-kt})^3,$$

where the constants a, b and k are positive and $a > b$; N denotes the size of the population and t denotes elapsed time. Find the growth rate dN/dt of the population.

Note: if $f(x) = 2^{ax}$, then we give that $f'(x) = 0.6931 \cdot a2^{ax}$. Derivatives of such *exponential functions* are studied in Chapter 10.

8.3. **Earth's temperature.** We expand and generalize the results of Example 8.10. As before, let G denote the level of greenhouse gases on Earth, and consider the relationship of temperature of the earth to the albedo a and the emissivity ε given by Eqn. (8.1).

(a) Suppose that a is constant, but ε depends on G. Assume that $d\varepsilon/dG$ is given. Determine the rate of change of temperature with respect to the level of greenhouse gasses in this case.

(b) Suppose that both a and ε depend on G. Find dT/dG in this more general case (*hint:* the quotient rule as well as the chain rule are needed).

8.4. **Shortest path from nest to food sources.**

(a) Use the first derivative test to verify that the value $x = \frac{d}{\sqrt{3}}$ is a local minimum of the function $L(x)$ given by Eqn (8.2)

(b) Show that the shortest path is $L = D + \sqrt{3}d$.

(c) In Section 8.2 we assumed that $d << D$, so that the food sources were close together relative to the distance from the nest. Now suppose that $D = d/2$. How would this change the solution?

8.5. **Geometry of the shortest ants' path.** Use the results of Section 8.2 to show that in the shortest path, the angles between the branches of the Y-shaped path are all $120°$. Recall that $\sin(30) = 1/2, \sin(60) = \sqrt{3}/2$.

8.6. **More about the ant trail.** Consider the lengths of the V and T-shaped paths in the ant trail example of Section 8.2. We refer to these as L_V and L_T; each depend on the distances d and D in Figure 8.3.

(a) Write down the expressions for each of these functions.

(b) Suppose the distance D is fixed. How do the two lengths L_V, L_T depend on the distance d? Use your sketching skills to draw a rough sketch of $L_v(d), L_T(d)$.

(c) Use you sketch to determine whether there is a value of d for which the lengths L_V and L_T are the same.

8.7. **Divided attention.** A bird in its natural habitat feeds on two kinds of seeds whose nutritional values are

- 5 calories per seed of type 1, and
- 3 calories per seed of type 2.

Both kinds of seeds are hidden among litter on the forest floor and have to be found. If the bird splits its attention into x_1 (a fraction of 1 - its whole attention) searching for seed type 1 and x_2 (also a fraction of 1) searching for seed type 2, then its probability of finding 100 seeds of the given type is

$$P_1(x_1) = (x_1)^3, \quad P_2(x_2) = (x_2)^5.$$

Assume that the bird pays full attention to searching for seeds so that $x_1 + x_2 = 1$ where $0 \leq x_1 \leq 1$ and $0 \leq x_2 \leq 1$.

(a) Give an expression for the total nutritional value V gained by the bird when it splits its attention. Use the constraint on x_1, x_2 to eliminate one of these two variables (for example, let $x = x_1$ and write x_2 in terms of x_1.)

(b) Find critical points of $V(x)$ and classify those points.

(c) Find absolute minima and maxima of $V(x)$ and use your results to explain the bird's optimal strategy for maximizing the nutritional value of the seeds it can find.

9

Chain rule applied to related rates and implicit differentiation

9.1 Applications of the chain rule to "related rates"

See A brief introduction.

Section 9.1 Learning goals

1. Given a geometric relationship and a rate of change of one of the variables, use the chain rule to find the rate of change of a related variable.
2. Use descriptive information about rates of change to set up the required relationships, and to solve a word problem involving an application of the chain rule ("related rates problem").

In many applications of the chain rule, we are interested in processes that take place over time. We ask how the relationships between certain geometric (or physical) variables affects the rates at which they change over time. Many of these examples are given as word problems, and we must assemble the required relationships to solve the problem. Some useful geometric relationships are presented in Table 9.1.

Volume of sphere	$V = \frac{4}{3}\pi r^3$
Surface area of sphere	$S = 4\pi r^2$
Area of circle	$A = \pi r^2$
Perimeter of circle	$P = 2\pi r$
Volume of cylinder	$V = \pi r^2 h$
Volume of cone	$V = \frac{1}{3}\pi r^2 h$
Area of rectangle	$A = xy$
Perimeter of rectangle	$P = 2x + 2y$
Volume of box	$V = xyz$
Sides of right triangle	$c^2 = a^2 + b^2$

Table 9.1: Geometric relationships used in various related rates problems.

Example 9.1 (Tumor growth) *The radius of a solid tumor expands at a constant rate, k. Determine the rate of growth of the volume of the tumor when the radius is $r = 1$cm. Assume that the tumor is approximately spherical as depicted in Figure 9.1.*

Tumor growth example: See the calculation in action.

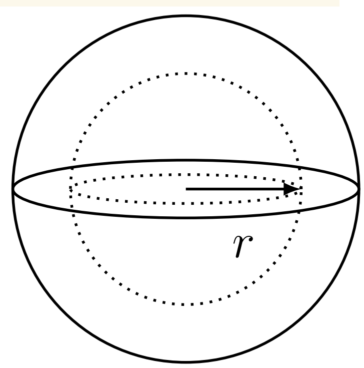


Figure 9.1: Growth of a spherical tumor. Since the radius changes with time, the volume, too, changes with time. We use the chain rule to link dV/dt to dr/dt.

Solution. The volume of a sphere of radius r, is $V(r) = (4/3)\pi r^3$. Here, r changes with time, so V changes with time. We indicate this chain of dependencies with the notation $r(t)$ and $V(r(t))$. Then function composition is apparent:

$$V(r(t)) = \frac{4}{3}\pi[r(t)]^3.$$

Then, using the chain rule,

$$\frac{d}{dt}V(r(t)) = \frac{dV}{dr}\frac{dr}{dt} = \frac{d}{dr}\left(\frac{4}{3}\pi r^3\right)\frac{dr}{dt} = \frac{4}{3}\pi \cdot 3r^2\frac{dr}{dt} = 4\pi r^2\frac{dr}{dt}.$$

But we are told that the radius expands at a constant rate, k, so that

$$\frac{dr}{dt} = k. \quad \Rightarrow \quad \frac{dV}{dt} = 4\pi r^2 k.$$

Hence, the rate of growth of the volume is proportional to the square of the radius; in fact, it is proportional to the surface area of the sphere. At the instant that $r = 1$ cm,

$$\frac{dV}{dt} = 4\pi k.$$

Importantly, the numerical value $r = 1$ cm holds only at one instant and is used at the end of the calculation, after the differentiation and simplification steps are completed. ◊

Mastered Material Check

1. What is dV/dt when the radius is $r = 2$cm?
2. What are the units of dV/dt?

Featured Problem 9.1 (Growth of a cell) *The mass of a cell is $m = \rho V$ where V is cell volume and ρ is cell density. (Usually, cell density is constant and close to that of water, $\rho \approx 1$g/cm^3.) Relate the cell rate of change of mass to rate of change of volume and to rate of change of radius. Assume that the cell is spherical.*

Example 9.2 (Convergent extension) *Most animals are longer head to tail than side to side. To obtain relative elongation along one axis, an embryo undergoes a process called* **convergent extension** *whereby a block of tissue elongates (extends) along one axis and narrows (converges) along the other axis as shown in Figure 9.2. Here we consider this process.*

Suppose that a rectangular block of tissue, with dimensions $L = w = 10$mm and thickness $\tau = 1$mm, extends at the rate of 1mm per day, while the volume V and thickness τ remain fixed. At what rate is the width w changing when the length is $L = 20$mm?

Mastered Material Check

3. How wide is the tissue when $L = 20$mm if thickness τ and volume V remain fixed?

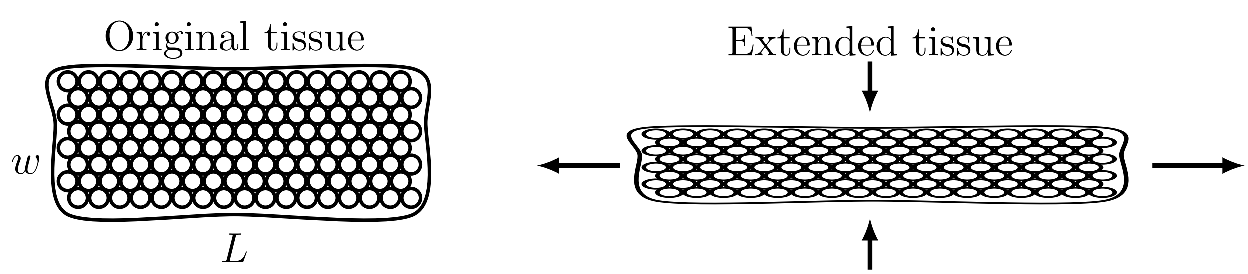


Figure 9.2: Convergent extension of tissue in embryonic development. Cells elongate along one axis (which increases L) while contracting along the other axis (decreasing w). Since the volume and thickness remain fixed, the changes in L can be related to changes in w.

Solution. We are told that the volume V and the thickness τ remain constant. We find, using the initial length, width and thickness, that the volume is $V = 10 \cdot 10 \cdot 1\text{mm}^3$. Further, at any given time t, the volume of the rectangular block is

$$V = L(t) \cdot w(t) \cdot \tau.$$

V depends on L and w, both of which depend on time. Hence, there is a chain of dependencies $t \to L, w, \to V$, Differentiating both sides with respect to t

leads to

$$\frac{dV}{dt} = \frac{d}{dt}\left(L(t)\cdot w(t)\tau\right) \quad \Rightarrow \quad 0 = \left(L'(t)\cdot w(t) + L(t)\cdot w'(t)\right)\tau.$$

(Here we have used the product rule to differentiate $L(t)\cdot w(t)$ with respect to t. We also used the fact that V is constant so its derivative is zero, and τ is constant, so it multiplies the derivative of $L(t)w(t)$ as would any multiplicative constant.) Consequently, canceling the constant factor and solving for $w'(t)$ results in

$$L'(t)w(t) + L(t)w'(t) = 0 \quad \Rightarrow \quad w'(t) = -\frac{L'(t)w(t)}{L(t)}.$$

At the instant that $L(t) = 20$, $w(t) = V/(L(t)\tau) = 100/20 = 5$. Hence we find that

$$w'(t) = -\frac{L'(t)w(t)}{L(t)} = -\frac{1\text{mm/day} \cdot 5\text{mm}}{20\text{mm}} = -0.25\text{mm/day}.$$

The negative sign indicates that w is decreasing while L is increasing. ♢

Example 9.3 (A spider's thread) *A spider moves horizontally across the ground at a constant rate, k, pulling a thin silk thread with it. One end of the thread is tethered to a vertical wall at height h above ground and does not move. The other end moves with the spider. Determine the rate of elongation of the thread.*

Spider silk example: See the calculation in action.

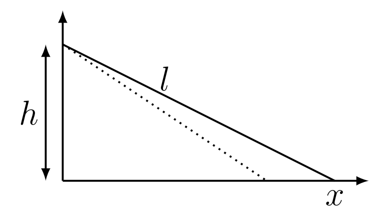


Figure 9.3: The length of a spider's thread.

Solution. Figure 9.3 illustrates the geometry, where x is the distance of the spider from the wall. We use the Pythagorean Theorem to relate the height of the tether point h, the spider's location x, and the length of the thread ℓ:

$$\ell^2 = h^2 + x^2.$$

Here, h is constant, while x, ℓ change with time, so that

$$[\ell(t)]^2 = h^2 + [x(t)]^2.$$

Differentiating with respect to t leads to

$$\frac{d}{dt}\left([\ell(t)]^2\right) = \frac{d}{dt}\left(h^2 + [x(t)]^2\right),$$

$$2\ell\frac{d\ell}{dt} = 0 + 2x\frac{dx}{dt} \quad \Rightarrow \quad \frac{d\ell}{dt} = \frac{2x}{2\ell}\frac{dx}{dt}.$$

Simplifying and using the fact that

$$\frac{dx}{dt} = k,$$

leads to

$$\frac{d\ell}{dt} = \frac{x}{\ell}k = k\frac{x}{\sqrt{h^2+x^2}}.$$

♢

Mastered Material Check

4. Repeat Example 9.3 given that the thread is tethered 0.5m above the ground and the spider is walking at a constant rate of 30cm/min.

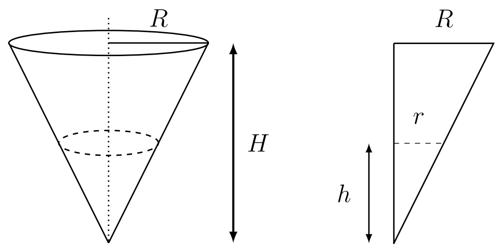


Figure 9.4: The geometry of a conical cup of (constant) height H and (constant) radius R. The surface of the water is indicated with dashed line. The water occupies a conical volume of height $h(t)$ and radius $r(t)$, which both get smaller as waters leaks out of the cup. The diagram on the left strongly suggests that similar triangles would be helpful in finding a relationship between the variables.

Example 9.4 (A conical cup) *Water is leaking at a constant rate out of a conical cup of height H and radius R. Find the rate of change of the height of water in the cup at the instant that the cup is full, if the volume is decreasing at a constant rate, k.*

Solution. Let us define h and r as the height and radius of water inside the cone. Then we know that the volume of this (conically shaped) water in the cone is

$$V = \frac{1}{3}\pi r^2 h,$$

or, in terms of functions of time,

$$V(t) = \frac{1}{3}\pi [r(t)]^2 h(t).$$

We are told that

$$\frac{dV}{dt} = -k,$$

where the negative sign indicates that volume is decreasing. By similar triangles, we note that

$$\frac{r}{h} = \frac{R}{H} \quad \Rightarrow \quad r = \frac{R}{H}h,$$

so we use this substitution to write the volume in terms of the height alone:

$$V(t) = \frac{1}{3}\pi \left[\frac{R}{H}\right]^2 [h(t)]^3.$$

Then the chain rule leads to

$$\frac{dV}{dt} = \frac{1}{3}\pi \left[\frac{R}{H}\right]^2 \cdot 3[h(t)]^2 \frac{dh}{dt}.$$

Now using the fact that volume decreases at a constant rate, we get

$$-k = \pi \left[\frac{R}{H}\right]^2 [h(t)]^2 \frac{dh}{dt} \quad \Rightarrow \quad \frac{dh}{dt} = \frac{-kH^2}{\pi R^2 h^2}.$$

The rate computed above holds at any time as the water leaks out of the container. At the instant that the cup is full, $h(t) = H$ and $r(t) = R$, so that

$$\frac{dh}{dt} = \frac{-kH^2}{\pi R^2 H^2} = \frac{-k}{\pi R^2}.$$

Draining Cone Example: See the calculation in action.

Mastered Material Check

5. What volume of water can be contained in a cone of height 5cm and radius 3cm?
6. What are the units of the constant k?
7. What is meant by similar triangles?

For example, for a cone of height $H = 4$ and radius $R = 3$,

$$\frac{dh}{dt} = \frac{-k}{9\pi}.$$

It is important to use information about a specific instant only after derivatives are computed. ◊

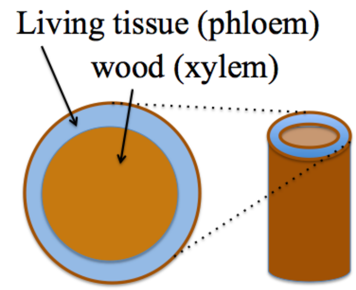


Figure 9.5: A thin shell of living tissue (phloem) surrounds the dead wooden part of a tree trunk (xylem). The fraction F of living tissue changes as the tree grows.

Featured Problem 9.2 (Growth of a Tree Trunk) *Consider a cylindrical tree trunk of radius R. Living cells occupy a thin shell (thickness d) just inside the tree bark. The interior of the trunk consists of dead cells that have turned into wood.*

1. *What fraction F of the trunk volume is living tissue?*
2. *How does the fraction F change with time as the tree grows? Assume that the radius of the trunk grows at a constant rate, and that the thickness d does not change. Compute the rate of change of F at the instant that the radius is 5 times the thickness d.*

Hints and setting up the problem

An explanation of how to set up the tree trunk problem.

- Assume that the wooden interior is cylindrical, as is the trunk. Find the volume of the shell by subtracting the volumes of these two cylinders, and now write down the fraction F.

  You should get
  $$F = \frac{\pi h\left[R^2 - (R-d)^2\right]}{\pi h R^2}.$$
  Simplify this expression.

- Compute the derivative dF/dt, remembering to use the Chain Rule. You may want to use the quotient rule for practice, or to first simplify the expression as much as possible and then compute a derivative.

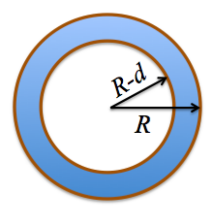


Figure 9.6: The radius of the trunk is R and the thickness of the phloem (living part of the trunk) is d. We assume a cylindrical geometry. F = fraction of volume in the thin (blue) cylindrical shell.

Some further steps, and an outline of what we are asked to do.

9.2 *Implicit differentiation*

Section 9.2 Learning goals

1. Identify the distinction between a function that is defined explicitly and one that is defined implicitly.
2. Describe implicit differentiation geometrically.
3. Compute the slope of a curve at a given point using implicit differentiation, find tangent line equations, and solve problems based on such ideas.

Implicit and explicit definition of a function

A review of the definition of a function (e.g. Appendix C) reminds us that for a given x value, only one y value is permitted. For example, for

$$y = x^2$$

any value of x leads to a single y value (Figure 9.7a). Geometrically, this means that the graph of this function satisfies the **vertical line property**: a given x value can have at most one corresponding y value. Not all curves satisfy this property. The elliptical curve in Figure 9.7b clearly fails this, intersecting some vertical lines twice. This simply means that, while we can write down an equation for such a curve, e.g.

$$\frac{(x-1)^2}{4} + (y-1)^2 = 1,$$

we cannot solve for a simple function that describes the entire curve. Nevertheless, the idea of a tangent line to such a curve - and consequently the slope of such a tangent line - is perfectly reasonable.

In order to make sense of this idea, we restrict attention to a local part of the curve, close to some point of interest (Figure 9.7c). Then *near this point*, the equation of the curve defines an **implicit function**, that is, close enough to the point of interest, a value of x leads to a unique value of y. We refer to this value as $y(x)$ to remind us of the relationship between the two variables.

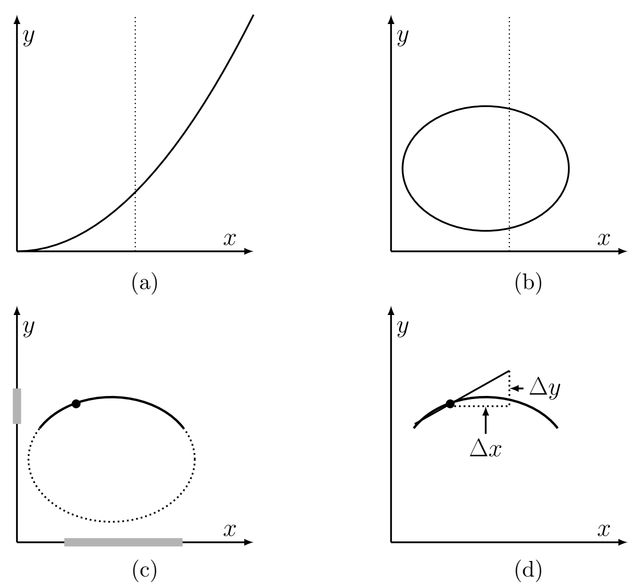


Figure 9.7: (a) A function has to satisfy the vertical line property. Hence, the curve shown in (b) cannot be a function. We can write down an equation for the curve, but we cannot solve for y explicitly. (c) However, close to a given point on the curve (dark point), we can think of how changing the x coordinate of the point (shaded interval on x axis) leads to a change in the corresponding y coordinate on the same curve (shaded interval on y axis). (d) We can also ask what is the slope of the curve at the given point. This corresponds to $\lim_{\Delta x \to 0} \Delta y / \Delta x$. Implicit differentiation can be used to compute that derivative.

How can we generalize the notion of a derivative to implicit functions? We observe from (Figure 9.7d) that a small change in x leads to a small change in y. Without writing down an explicit expression for y versus x, we can still determine these small changes, and form a ratio $\Delta y/\Delta x$ which is a (secant line) slope. Now let $\Delta x \to 0$ to arrive at the slope of a tangent line as before, dy/dx. In the next section we show how to do this using **implicit differentiation**, an application of the chain rule.

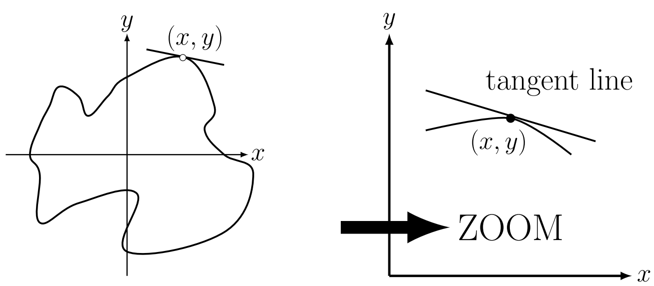
Slope of a tangent line at the point on a curve

A brief introduction to implicit differentiation and slope of a tangent line to a circle.

We now compute the tangent line at a point in several examples where it is inconvenient or impossible, to isolate y as a function of x (see Figure 9.8).

First, consider the simple example of a circle. We aim to find the slope of the tangent line at some point. The equation of a circle of radius 1 and centre at the origin $(0,0)$ is

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



(a) (b)

Figure 9.8: The curve in (a) is not a function and hence it can only be described implicitly. However, if we zoom in to a point in (b), we can define the derivative as the slope of the tangent line to the curve at the point of interest.

Here, the two variables are linked in a symmetric relationship. We can solve for y, obtaining not one but *two* functions.

$$\text{top of circle:} \quad y = f_1(x) = \sqrt{1-x^2},$$
$$\text{and bottom of circle:} \quad y = f_2(x) = -\sqrt{1-x^2}.$$

Mastered Material Check

8. In Example 9.5, why do we need to specify “in the first quadrant"? What are other possible point with $x = 1/2$ on the circle?

However, this makes the work of differentiation more complicated than necessary. Instead, we use implicit differentiation.

Example 9.5 (Tangent to a circle) a) *Use implicit differentiation to find the slope of the tangent line to the point $x = 1/2$ in the first quadrant on a circle of radius 1 and centre at $(0,0)$.*

b) *Find the second derivative d^2y/dx^2 at the same point.*

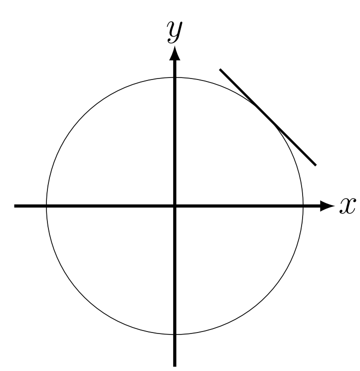


Figure 9.9: Tangent line to a circle by implicit differentiation.

Solution.

a) When $x = 1/2$ then $y = \pm\sqrt{1-(1/2)^2} = \pm\sqrt{3}/2$. The point in the first quadrant has y coordinate $y = +\sqrt{3}/2$.

Viewing x as the independent variable locally, (and y depending on x on the curve) we write

$$x^2 + [y(x)]^2 = 1.$$

Differentiating each side with respect to x:

$$\frac{d}{dx}\left(x^2 + [y(x)]^2\right) = \frac{d}{dx}1 = 0 \quad \Rightarrow \quad \left(\frac{dx^2}{dx} + \frac{d}{dx}[y(x)]^2\right) = 0.$$

Notice that in the second term, the value of x determines y which in turn determines y^2. Applying the chain rule, we obtain

$$\left(\frac{dx^2}{dx} + \frac{dy^2}{dy}\frac{dy}{dx}\right) = 0. \quad \Rightarrow \quad 2x + 2y\frac{dy}{dx} = 0.$$

Thus

$$2y\frac{dy}{dx} = -2x \quad \Rightarrow \quad \frac{dy}{dx} = -\frac{2x}{2y} = -\frac{x}{y}.$$

At the point of interest, $x = 1/2$, $y = \sqrt{3}/2$. Thus the slope of the tangent line is

$$y' = \frac{dy}{dx} = -\frac{x}{y} = -\frac{1/2}{\sqrt{3}/2} = \frac{-1}{\sqrt{3}} = \frac{-\sqrt{3}}{3}.$$

b) The second derivative can be computed using the quotient rule

$$y' = \frac{dy}{dx} = -\frac{x}{y} \quad \Rightarrow \quad \frac{d^2y}{dx^2} = \frac{d}{dx}\left(-\frac{x}{y}\right)$$

$$\frac{d^2y}{dx^2} = -\frac{1 \cdot y - x \cdot y'}{y^2} = -\frac{y - x\frac{-x}{y}}{y^2} = -\frac{y^2 + x^2}{y^3} = -\frac{1}{y^3}.$$

Substituting $y = \sqrt{3}/2$ from part (a) yields

$$\frac{d^2y}{dx^2} = -\frac{1}{(\sqrt{3}/2)^3} = -\frac{8}{3(3/2)}.$$

We used the equation of the circle, and our result for the first derivative to simplify the above. ♢

Note: we can see from the last expression that the second derivative is negative for $y > 0$, i.e. for the top semi-circle, indicating that this part of the curve is concave down (as expected). Indeed, as in the case of simple functions, the second derivative can help identify concavity of curves.

Mastered Material Check

9. Verify the result of Example 9.5(a) by differentiating the explicit function for the top half of a circle of radius 1, centered at the origin: $f(x) = \sqrt{1 - x^2}$.
10. Similarly, verify the result of Example 9.5(b).
11. The second derivative is positive for $y < 0$. What does this say about the bottom part of the circle?

Example 9.6 (Energy loss and Earth's temperature) *Redo Example 4.9 using implicit differentiation, that is: find the rate of change of Earth's temperature per unit energy loss based on Eqn. (1.5): $E_{out} = 4\pi r^2 \varepsilon \sigma T^4$.*

Solution. We rewrite the equation in the form

$$E_{out}(T) = (4\pi r^2 \varepsilon \sigma)T^4$$

and observe that the term in braces is constant. We then differentiate both sides with respect to E_{out}. We find,

$$\frac{dE_{out}}{dE_{out}} = (4\pi r^2 \varepsilon\sigma)\frac{dT^4}{dT}\frac{dT}{dE_{out}} \quad \Rightarrow \quad 1 = (4\pi r^2 \varepsilon\sigma)\cdot 4T^3 \frac{dT}{dE_{out}}.$$

The calculation is completed by rearranging this result. Thus

$$\frac{dT}{dE_{out}} = \frac{1}{16\pi r^2 \varepsilon\sigma}\frac{1}{T^3}$$

is the rate of change of the Earth's temperature per unit energy loss. ◊

Mastered Material Check

12. Does this result agree with that of Example 4.9,

$$\frac{dT}{dE_{out}} = \left(\frac{1}{16\pi r^2 \varepsilon\sigma}\right)^{1/4} E_{out}^{-3/4}?$$

Justify algebraically.

13. Can you define an inverse function?

9.3 *The power rule for fractional powers*

Implicit differentiation is a useful technique for finding derivatives of inverse functions. Here we use the known power rule for $y = x^2$ to find the derivative of its inverse function, $y = \sqrt{x} = x^{1/2}$. This general idea recurs in later chapters when we introduce new functions and their inverses.

Using implicit differentiation to compute the derivative of $y = \sqrt{x}$.

Example 9.7 (Derivative of $\sqrt{x}$) *Consider the function $y = \sqrt{x} = x^{1/2}$. Use implicit differentiation to compute the derivative of this function.*

Solution. Let us rewrite the relationship $y = \sqrt{x}$ in the form $y^2 = x$, but consider y as the dependent variable, i.e. when we differentiate, we remember that y depends on x:

$$[y(x)]^2 = x.$$

Taking derivatives of both sides leads to

$$\frac{d}{dx}\left([y(x)]^2\right) = \frac{d}{dx}(x) \quad \Rightarrow \quad 2[y(x)]\frac{dy}{dx} = 1 \quad \Rightarrow \quad \frac{dy}{dx} = \frac{1}{2y}.$$

We eliminate y by substituting $y = \sqrt{x}$. Then

$$\frac{dy}{dx} = \frac{1}{2\sqrt{x}} = \frac{1}{2}x^{-1/2}.$$

This verifies the power law for the above example. ◊

A similar procedure can be applied to any power function with fractional power. When we apply similar steps, we get the following rule:

Derivative of fractional-power function: The derivative of

$$y = f(x) = x^{m/n}$$

is

$$\frac{dy}{dx} = \frac{m}{n}x^{\left(\frac{m}{n}-1\right)}.$$

Mastered Material Check

14. Justify the derivative of the fractional power rule by actually carrying out an implicit differentiation calculation.

Example 9.8 (The astroid) *The curve*

$$x^{2/3} + y^{2/3} = 2^{2/3}$$

has the shape of an **astroid**. *It describes the shape (Figure 9.10) generated by a the path of a point on the perimeter of a disk of radius* $\frac{1}{2}$ *rolling inside the perimeter of a circle of radius* 2.

Find the slope of the tangent line to a point on the astroid.

See Demo created by David Austin of the astroid.

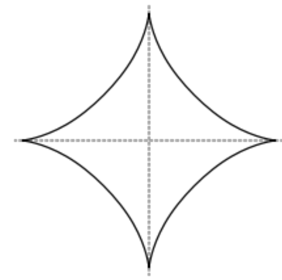


Figure 9.10: The astroid is generated by a disk or radius $1/2$ rolling inside a circle of radius 2.

Solution. Considering y as the dependent variable, we use implicit differentiation as follows:

$$\begin{aligned}
\frac{d}{dx}\left(x^{2/3} + [y(x)]^{2/3}\right) = \frac{d}{dx}2^{2/3} \quad &\Rightarrow \quad \frac{2}{3}x^{-1/3} + \frac{d}{dy}(y^{2/3})\frac{dy}{dx} = 0 \\
&\Rightarrow \quad \frac{2}{3}x^{-1/3} + \frac{2}{3}y^{-1/3}\frac{dy}{dx} = 0 \\
&\Rightarrow \quad x^{-1/3} + y^{-1/3}\frac{dy}{dx} = 0 \\
&\Rightarrow \quad \frac{dy}{dx} = -\frac{x^{-1/3}}{y^{-1/3}}.
\end{aligned}$$

The derivative fails to exist at $x = 0$ (where $x^{-1/3}$ is undefined) and at $y = 0$ (where $y^{-1/3}$ is undefined). ◊

Mastered Material Check

15. Point out places on Fig. 9.10 at which the derivative fails to exist, and explain the properties of those specific points on the astroid.

Example 9.9 (Horizontal tangent and concavity on a rotated ellipse) *Find the highest point on the (rotated) ellipse* $x^2 + 3y^2 - xy = 1$.

Using implicit differentiation to find the points on the top and bottom of the ellipse in Example 9.9.

Solution. The highest point on the ellipse has a horizontal tangent line, so we should look for the points on this curve at which $dy/dx = 0$.

1. **Finding the slope of the tangent line:** By implicit differentiation,

$$\frac{d}{dx}[x^2 + 3y^2 - xy] = \frac{d}{dx}1 \quad \Rightarrow \quad \frac{d(x^2)}{dx} + \frac{d(3y^2)}{dx} - \frac{d(xy)}{dx} = 0.$$

   We must use the product rule to compute the derivative of the last term xy:

$$2x + 6y\frac{dy}{dx} - \left(x\frac{dy}{dx} + \frac{dx}{dx}y\right) = 0 \quad \Rightarrow \quad 2x + 6y\frac{dy}{dx} - x\frac{dy}{dx} - 1y = 0.$$

   Grouping terms, we have

$$(6y - x)\frac{dy}{dx} + (2x - y) = 0 \quad \Rightarrow \quad \frac{dy}{dx} = \frac{(y - 2x)}{(6y - x)}.$$

   Setting $dy/dx = 0$, we obtain $y - 2x = 0$ so that $y = 2x$ at the point of interest. Next, we find the coordinates of the point.

2. **Determining the coordinates of the point we want:** We look for a point that satisfies the equation of the curve as well as the condition $y = 2x$. There

are two equations and two unknowns. Plugging $y = 2x$ into the original equation of the ellipse, we get:

$$x^2 + 3y^2 - xy = 1 \quad \Rightarrow \quad x^2 + 3(2x)^2 - x(2x) = 1.$$

After simplifying, this equation becomes $11x^2 = 1$, leading to the two possibilities

$$x = \pm\frac{1}{\sqrt{11}}, \quad y = \pm\frac{2}{\sqrt{11}}.$$

Which of these two points is at the top? The rotated ellipse is depicted in Figure 9.11 which gives strong indication it is the positive solution - but we can confirm this analytically.

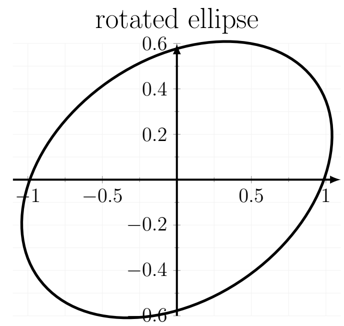


Figure 9.11: A rotated ellipse. In Example 9.9, we find the point at the top of the ellipse using implicit differentiation.

3. **Identifying the point at the top:** The top point on the ellipse is located at a point where the curve is concave down. Concavity can be determined using the second derivative, computed (from the first derivative) using the quotient rule:

$$\begin{aligned} y'' &= \frac{[y-2x]'(6y-x) - [6y-x]'(y-2x)}{(6y-x)^2} \\ &= \frac{[y'-2](6y-x) - [6y'-1](y-2x)}{(6y-x)^2}. \end{aligned}$$

4. **Plugging in information about the point:** Now that we have set down the form of this derivative, we make some important observations about the specific point of interest. (This is done as a final step, only after all derivatives have been calculated)

- We are only concerned with the sign of the second derivative. The denominator is always positive (since it is squared) and so does not affect the sign.
- At the top of ellipse, $y' = 0$, simplifying some of the terms above.
- At the top of ellipse, $y = 2x$ so the term $(y - 2x) = 0$.

We can thus simplify the expression for the y'' to obtain

$$y''(x) = \frac{[-2](6y-x) - [-x](0)}{(6y-x)^2} = \frac{[-2](6y-x)}{(6y-x)^2} = \frac{-2}{(6y-x)}.$$

Using the fact that $y = 2x$, we get the final form

$$y''(x) = \frac{-2}{(6(2x)-x)} = \frac{-2}{11x}.$$

Consequently, the second derivative is negative (implying concave down curve) whenever x is positive. This tells us that at the point with positive x value ($x = 1/\sqrt{11}$), we are at the top of the ellipse. A graph of this curve is shown in Figure 9.11. ♢

Mastered Material Check

16. Verify by hand that $x = \pm\frac{1}{\sqrt{11}}$ are two possible solutions to Example 9.9.
17. Repeat Example 9.9 looking for the *lowest* point on the rotated ellipse.

Featured Problem 9.3 (Tangent to an ellipse) *Find the equation of a line through the origin that is tangent to the ellipse*

$$(x-a)^2 + \frac{y^2}{s^2} = 1.$$

This ellipse has its center at $(a, 0)$ *and has axes* 1 *and* s.

9.4 Summary

1. The chain rule can be used to relate the changes in variables that depend on one other in some "chain of relationships". We use the term "related rates" to describe such problems.
2. Curves that fail the vertical line property cannot be describe by a single function $y = f(x)$, even if we can represent some of those curves by equation(s).
3. Zooming in on such a curve, we can define an implicit function that describes some local piece of the curve.
4. When we use implicit differentiation in two variables, we treat one variable as independent and the other as dependent. This allows us to differentiate the equation with respect to the independent variable using the chain rule.
5. Through implicit differentiation we showed that the derivative of $y = x^{m/n}$ is $\frac{dy}{dx} = \frac{m}{n}x^{(\frac{m}{n}-1)}$.
6. Applications addressed in this chapter included:

   (a) tumor growth (volume depends on radius which depends on a growth rate);
   (b) convergent extension in tissue of an embryo (relationship between the length and width of the growing tissue);
   (c) growth of a cell (the relationship between volume and radius).
   (d) spider's thread (length of thread depends on the spider's position, which depends on time);
   (e) growth of a tree trunk (determining the fraction of the trunk that is living tissue as the tree grows)
   (f) conical cup leaking water (height of water depends on volume, which depends on time); and
   (g) rate of Earth's temperature change per unit energy loss.

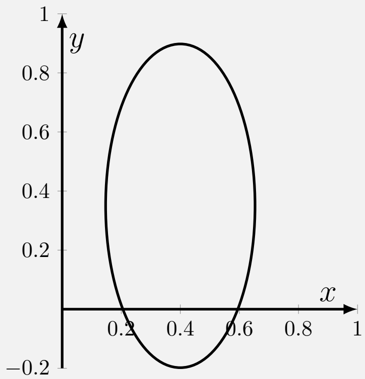
Quick Concept Checks

1. Let $a = b^{2/9}$. Determine $\frac{da}{db}$.

2. In Example 9.4, suppose that the cone does not leak, but that it is being filled with water at a constant rates. How would your work change?

3. Use implicit differentiation to find the slope of the tangent line to the circle $x^2 + y^2 = 1$ at the point $x = -1$? How does your result relate to the orientation of the tangent line to the circle at that point?

4. Consider the following curve.



Draw both tangent lines to this curve at $x = 0.5$.

Exercises

9.1. **Growth of a cell.** Consider the growth of a cell, assumed spherical in shape. Suppose that the radius of the cell increases at a constant rate $k > 0$ per unit time.

(a) At what rate would the volume, V, increase ?

(b) At what rate would the surface area, S, increase ?

(c) At what rate would the ratio of surface area to volume S/V change? Would this ratio increase or decrease as the cell grows?

Note: note that answers are expressed in terms of the radius of the cell.

9.2. **Growth of a circular fungal colony.** A fungal colony grows on a flat surface starting with a single spore. The shape of the colony edge is circular with the initial site of the spore at the centre of the circle. Suppose the radius of the colony increases at a constant rate C per unit time.

(a) At what rate does the area covered by the colony change ?

(b) The biomass of the colony is proportional to the area it occupies. Let α be the factor of proportionality. At what rate does the biomass increase?

9.3. **Limb development.** During early development, the limb of a fetus increases in size, but has constant proportions. Suppose that the limb is roughly a circular cylinder with radius r and length l in proportion

$$l/r = C$$

where C is a positive constant. It is noted that during the initial phase of growth, the radius increases at an approximately constant rate, i.e. that

$$dr/dt = a.$$

At what rate does the mass of the limb change during this time?

Note: assume that the density of the limb is 1 gm/cm^3 and recall that the volume of a cylinder is

$$V = Al$$

where A is the base area (in this case of a circle) and l is length.

9.4. **Pouring water in a trough.** A rectangular trough is 2 meter long, 0.5 meter across the top and 1 meter deep. At what rate must water be poured into the trough such that the depth of the water is increasing at 1 m/min when the depth of the water is 0.7 m?

9.5. **Spherical balloon.** Gas is being pumped into a spherical balloon at the rate of 3 cm^3/s.

(a) How fast is the radius increasing when the radius is 15 *cm*?

(b) Without using the result from (a), find the rate at which the surface area of the balloon is increasing when the radius is 15 *cm*.

9.6. **Ice melting.** A spherical piece of ice melts so that its surface area decreases at a rate of 1 cm^2/min. Find the rate that the diameter decreases when the diameter is 5 cm.

9.7. **Point moving on a parabola.** A point moves along the parabola $y = \frac{1}{4}x^2$ in such a way that at $x = 2$ the x-coordinate is increasing at the rate of 5 cm/s. Find the rate of change of y at this instant.

9.8. **Boyle's Law.** In chemistry, Boyle's Law describes the behaviour of an ideal gas: this law relates the volume V occupied by the gas to the temperature T and the pressure P as follows:

$$PV = nRT$$

where n, R are positive constants.

(a) Suppose that pressure is kept fixed by allowing the gas to expand as the temperature is increased. Relate the rate of change of volume to the rate of change of temperature.

(b) Suppose that the temperature is held fixed and the pressure is decreased gradually. Relate the rate of change of the volume to the rate of change of pressure.

9.9. **Spread of a population.** In 1905 a Bohemian farmer accidentally allowed several muskrats to escape an enclosure. Their population grew and spread, occupying increasingly larger areas throughout Europe. In a classical paper in ecology, it was shown by the scientist Skellam (1951) that the square root of the occupied area increased at a constant rate, k.

Determine the rate of change of the distance (from the site of release) that the muskrats had spread. Assume that the expanding area of occupation is circular.

9.10. **A convex lens.** A particular convex lens has a focal length of $f =$ 10 cm. Let p be the distance between an object and the lens, and q the distance between its image and the lens. These distances are related to the focal length f by the equation:

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q}.$$

Consider an object which is 30 cm away from the lens and moving away at 4 cm/sec.

How fast is its image moving and in which direction?

9.11. **A conical cup.** Water is leaking out of a small hole at the tip of a

Formula.
Note that the volume of a cone is $V = (\pi/3)r^2h$.

conical paper cup at the rate of 1 cm^3/min. The cup has height 8 cm and radius 6 cm, and is initially full up to the top.

Find the rate of change of the height of water in the cup when the cup just begins to leak.

9.12. **Conical tank.** Water is leaking out of the bottom of an inverted conical tank at the rate of $\frac{1}{10}$ m^3/min, and at the same time is being pumped in the top at a constant rate of k m^3/min. The tank has height 6 m and the radius at the top is 2 m.

Determine the constant k if the water level is rising at the rate of $\frac{1}{5}$ m/min when the height of the water is 2 m.

9.13. **The gravel pile.** Gravel is being dumped from a conveyor belt at the rate of 30 ft^3/min in such a way that the gravel forms a conical pile whose base diameter and height are always equal.

How fast is the height of the pile increasing when the height is 10 ft?

9.14. **The sand pile.** Sand is piled onto a conical pile at the rate of 10 m^3/min. The sand keeps spilling to the base of the cone so that the shape always has the same proportions: that is, the height of the cone is equal to the radius of the base.

Find the rate at which the height of the sandpile increases when the height is 5 m.

9.15. **Conical water reservoir.** Water is flowing into a conical reservoir at a rate of 4 m^3 / min. The reservoir is 3 m in radius and 12 m deep.

(a) How fast is the radius of the water surface increasing when the depth of the water is 8 m?

(b) In (a), how fast is the surface rising?

9.16. **Sliding ladder.** A ladder 10 meters long leans against a vertical wall. The foot of the ladder starts to slide away from the wall at a rate of 3 m/s.

(a) Find the rate at which the top of the ladder is moving downward when its foot is 8 meters away from the wall.

(b) In (a), find the rate of change of the slope of the ladder.

9.17. **Sliding ladder.** A ladder 5 m long rests against a vertical wall. If the bottom of the ladder slides away from the wall at the rate of 0.5 m/min how fast is the top of the ladder sliding down the wall when the base of the ladder is 1 m away from the wall?

9.18. **Species diversity in an area.** Ecologists are often interested in the relationship between the area of a region (A) and the number of different species S that can inhabit that region. Hopkins (1955)suggested a relationship of the form [Hopkins, 1955]

$$S = a\ln(1 + bA)$$

where a and b are positive constants.

Find the rate of change of the number of species with respect to the area. Does this function have a maximum?

9.19. **The burning candle.** A candle is placed a distance l_1 from a thin block of wood of height H. The block is a distance l_2 from a wall as shown in Figure 9.12. The candle burns down so that the height of the flame, h_1 decreases at the rate of 3 cm/hr. Find the rate at which the length of the shadow y cast by the block on the wall increases.

Note: your answer should be in terms of the constants l_1 and l_2. This is a challenging problem.

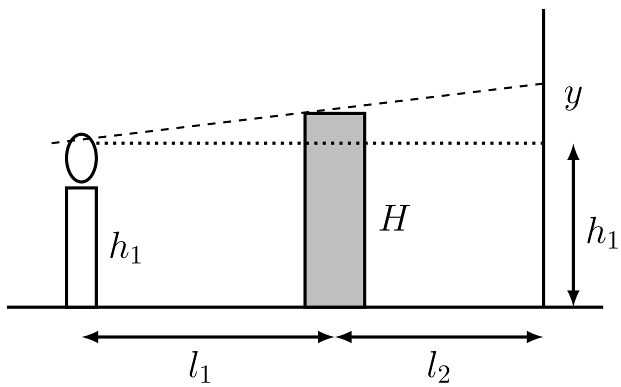


Figure 9.12: Figure for Exercise 19; shadow cast by burning candle.

9.20. **Implicit differentiation.** Use implicit differentiation to show that the derivative of the function

$$y = x^{1/3}$$

is

$$y' = (1/3)x^{-2/3}.$$

First write the relationship in the form $y^3 = x$, and then find dy/dx.

9.21. **Generalizing the Power Law.**

(a) Use implicit differentiation to calculate the derivative of the function

$$y = f(x) = x^{n/m}$$

where m and n are integers (*hint:* rewrite the equation in the form $y^m = x^n$ first).

(b) Use your result to derive the formulas for the derivatives of the functions $y = \sqrt{x}$ and $y = x^{-1/3}$.

9.22. **Tangent lines to a circle.** The equation of a circle with radius r and centre at the origin is

$$x^2 + y^2 = r^2$$

(a) Use implicit differentiation to find the slope of a tangent line to the circle at some point (x, y).

(b) Use this result to find the equations of the tangent lines of the circle at the points whose x coordinate is $x = r/\sqrt{3}$.

(c) Use the same result to show that the tangent line at any point on the circle is perpendicular to the radial line drawn from that point to the centre of the circle

Note: Two lines are perpendicular if their slopes are negative reciprocals.

9.23. **Implicit differentiation.** For each of the following, find the derivative of y with respect to x.

(a) $y^6 + 3y - 2x - 7x^3 = 0$

(b) $e^y + 2xy = \sqrt{3}$

9.24. **Tangent line to a circle.** The equation of a circle with radius 5 and centre at $(1, 1)$ is

$$(x-1)^2 + (y-1)^2 = 25$$

(a) Find the slope of the tangent line to this curve at the point $(4, 5)$.

(b) Find the equation of the tangent line.

9.25. **Tangent to a hyperbola.** The curve

$$x^2 - y^2 = 1$$

is a hyperbola. Use implicit differentiation to show that for large x and y values, the slope dy/dx of the curve is approximately 1.

9.26. **An ellipse.** Use implicit differentiation to find the points on the ellipse

$$\frac{x^2}{4} + \frac{y^2}{9} = 1$$

at which the slope is $-1/2$.

9.27. **Motion of a cell.** In the study of cell motility, biologists often investigate a type of cell called a keratocyte, an epidermal cell that is found in the scales of fish. This flat, elliptical cell crawls on a flat surface, and is known to be important in healing wounds. The 2D outline of the cell can be approximated by the ellipse

$$x^2/100 + y^2/25 = 1$$

where x and y are distances in μm. When the motion of the cell is filmed, points on the "leading edge" (top arc of the ellipse) move in a direction perpendicular to the edge.

Determine the direction of motion of the point (x_p, y_p) on the leading edge.

Units.
Note that 1μm, often called "1 micron", is 10^{-6} meters.

9.28. **The Folium of Descartes.** A famous curve (see Figure 9.13) that was studied historically by many mathematicians (including Descartes) is

$$x^3 + y^3 = 3axy$$

Assume that a is a positive constant.

(a) Explain why this curve cannot be described by a function such as $y = f(x)$ over the domain $-\infty < x < \infty$.

(b) Use implicit differentiation to find the slope of this curve at a point (x, y).

(c) Determine whether the curve has a horizontal tangent line anywhere, and if so, find the x coordinate of the points at which this occurs.

(d) Does implicit differentiation allow you to find the slope of this curve at the point $(0,0)$?

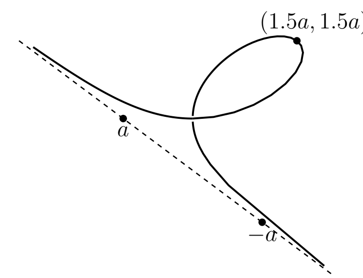


Figure 9.13: The Folium of Descartes in Exercise 28

9.29. **Isotherms in the Van-der Waal's equation.** In thermodynamics, the Van der Waal's equation relates the mean pressure, p of a substance to its molar volume v at some temperature T as follows:

$$(p + \frac{a}{v^2})(v - b) = RT,$$

where a, b, R are constants. Chemists are interested in the curves described by this equation when the temperature is held fixed.

Note: these curves are called isotherms.

(a) Find the slope, dp/dv, of the isotherms at a given point (v, p).

(b) Determine where points occur on the isotherms at which the slope is horizontal.

9.30. **The circle and parabola:** A circle of radius 1 is made to fit inside the parabola $y = x^2$ as shown in figure 9.14. Find the coordinates of the centre of this circle, i.e. find the value of the unknown constant c (*hint:* set up conditions on the points of intersection of the circle and the parabola which are labeled (a, b) in the figure. What must be true about the tangent lines at these points?).

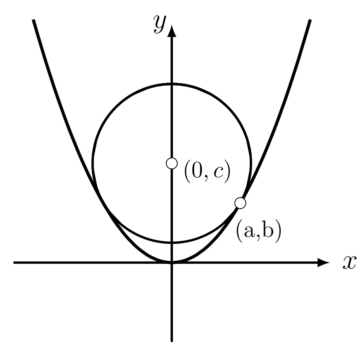


Figure 9.14: Figure for Exercise 30; circle inside a parabola.

9.31. **Equation of a tangent line.** Consider the curve whose equation is

$$x^3 + y^3 + 2xy = 4, \; y = 1 \text{ when } x = 1.$$

(a) Find the equation of the tangent line to the curve when $x = 1$.

(b) Find y'' at $x = 1$.

(c) Is the graph of $y = f(x)$ concave up or concave down near $x = 1$? (*hint:* differentiate the equation $x^3 + y^3 + 2xy = 4$ twice with respect to x).

10
Exponential functions

“The mathematics of uncontrolled growth are frightening. A single cell of the bacterium E. coli would, under ideal circumstances, divide every twenty minutes. That is not particularly disturbing until you think about it, but the fact is that bacteria multiply geometrically: one becomes two, two become four, four become eight, and so on. In this way it can be shown that in a single day, one cell of E. coli could produce a super-colony equal in size and weight to the entire planet Earth."

Michael Crichton, The Andromeda Strain, p. 247 [Crichton, 1969]

In this chapter, we introduce the exponential functions. We first describe the discrete process of **population doubling**, represented by powers of 2, namely, 2^n, where n is some integer. We generalize to a continuous function 2^x where x is any real number. We can then attach meaning to the notion of the derivative of an exponential function. In doing so, we encounter a specially convenient base denote e, leading to the most useful member of this class of functions, $y = e^x$. We discuss applications to unlimited growth in a population.

Mastered Material Check

1. If a population has size P, what do we mean by a doubled poulation size??
2. How large would the population be it it doubled twice?

10.1 Unlimited growth and doubling

Section 10.1 Learning goals

1. Explain the link between population doubling and integer powers of the base 2.
2. Given information about the doubling time of a population and its initial size, determine the size of that population at some later time.
3. Appreciate the connection between 2^n for integer values of n and 2^x for a real number x.

The Andromeda Strain

The Andromeda Strain scenario (described by Crichton in the opening quotation) motivates our investigation of population doubling and uncontrolled growth. Consider 2^n where $n = 1, 2\dots$ is an integer. We will study values of this discrete function as the "variable", n in the exponent changes. We list some values and display a graph of 2^n versus n in Figure 10.1. Notice that an initially "gentle" growth becomes extremely steep in just a few steps, as shown in the accompanying graph.

Note: properties of 2^n (and related expressions) are reviewed in Appendix B.1 where common manipulations are illustrated. We assume the reader is familiar with this material.

A screencast summary of population doubling and the Andromeda Strain. Edu.Cr.

Mastered Material Check

3. Compare the function $f(n) = 2^n$ and $g(n) = n^2$ for $n = 1, 2, \dots, 5$. How do these differ?

n	2^n
0	1
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024

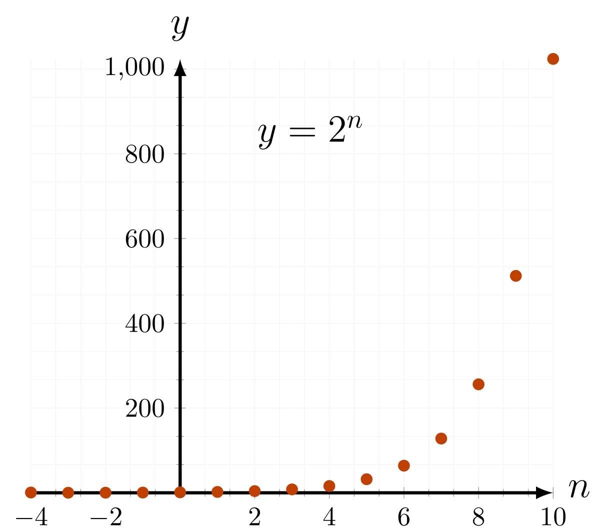


Figure 10.1: Powers of 2 including both negative and positive integers: here we show 2^n for $-4 < n < 10$.

The function 2^n first grows slowly, but then grows faster and faster as n increases. As a side remark, the fact that $2^{10} \approx 1000 = 10^3$, will prove useful for simple approximations. With this preparation, we can now check the accuracy of Crichton's statement about bacterial growth.

Mastered Material Check

4. Why would the approximation $2^{10} \approx 10^3$ be helpful?

Example 10.1 (Growth of E. coli) *Use the following facts to check the assertion made by Crichton's statement at the beginning of this chapter.*

- *Mass of* 1 *E. coli cell :* 1 *nanogram* = 10^{-9}*gm* = 10^{-12}*kg.*
- *Mass of Planet Earth :* $6 \cdot 10^{24}$ *kg.*

Mastered Material Check

5. How many cells of E. coli would there be after 20 minutes? 1 hour? 2 hours?

Solution. Based on the above two facts, we surmise that the size of an E. coli colony (number of cells, m) that together form a mass equal to Planet Earth would be

$$m = \frac{6 \cdot 10^{24}\ \text{kg}}{10^{-12}\ \text{kg}} = 6 \cdot 10^{36}.$$

Each hour corresponds to 3 twenty-minute generations. In a period of 24 hours, there are $24 \times 3 = 72$ generations, each doubling the colony size. After 1 day of uncontrolled growth, the number of cells would be 2^{72}. We can find a decimal approximation using the observation that $2^{10} \approx 10^3$:

$$2^{72} = 2^2 \cdot 2^{70} = 4 \cdot (2^{10})^7 \approx 4 \cdot (10^3)^7 = 4 \cdot 10^{21}.$$

Using a scientific calculator, the value is found to be $4.7 \cdot 10^{21}$, so the approximation is relatively good. $\Diamond$

Apparently, the estimate made by Crichton is not quite accurate. However it can be shown that it takes less than 2 days to produce a number far in excess of the "size of Planet Earth". The exact number of generations is left as an exercise for the reader and is discussed in Example 10.12.

Mastered Material Check

6. Verify that it takes less than 2 days to produce a number far in excess of the size of Planet Earth.

The function 2^x and its "relatives"

We would like to generalize the function 2^n to a continuous function, so that the tools of calculus - such as derivatives - can be used. To this end, we start with values that can be calculated based on previous mathematical experience, and then "fill in gaps".

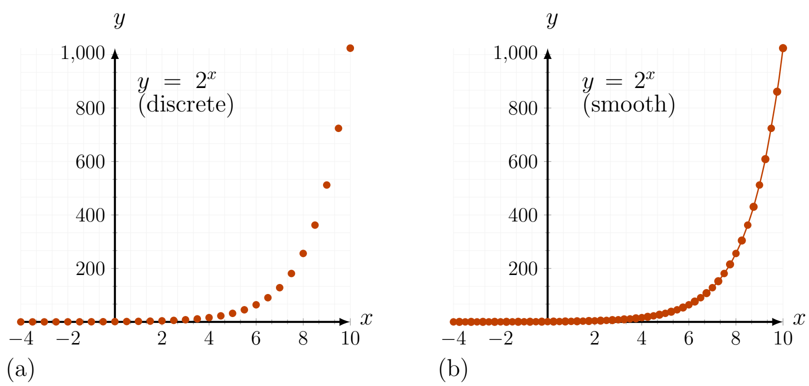


Figure 10.2: (a) Values of the function 2^x for discrete value of x. We can compute many values (e.g. for $x = 0, \pm 1, \pm 2$, by simple arithmetical operations, and for $x = \pm 1/2, \pm 3/2$ by evaluating square roots). (b) The function 2^x is connected smoothly to form a continuous curve.

From previous familiarity with power functions such as $y = x^2$ (not to be confused with 2^x), we know the value of

$$2^{1/2} = \sqrt{2} \approx 1.41421\ldots$$

We can use this value to compute

$$2^{3/2} = (\sqrt{2})^3, \quad 2^{5/2} = (\sqrt{2})^5,$$

and all other fractional exponents that are multiples of $1/2$. We can add these to the graph of our previous powers of 2 to fill in additional points. This is shown on Figure 10.2(a).

Similarly, we could also calculate exponents that are multiples of $1/4$ since

$$2^{1/4} = \sqrt{\sqrt{2}}$$

is a value that we can obtain. Adding these values leads to an even finer set of points. By continuing in the same way, we "fill in" the graph of the emerging function. Connecting the dots smoothly allows us to define a value for any real x, of a new continuous function,

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

Here x is no longer restricted to an integer, as shown by the smooth curve in Figure 10.2(b).

Mastered Material Check

7. Given $2^{1/2} \approx 1.41421$, find $2^{3/2}$ and $2^{5/2}$ without using fractional powers.
8. What method might you use to determine a decimal approximation of $2^{1/4}$ without computing fractional powers?
9. Why do we need to assume that $a > 0$ for the exponential function $y = a^x$?

Example 10.2 (Generalization to other bases) *Plot "relatives" of 2^x that have other bases, such as $y = 3^x$, $y = 4^x$ and $y = 10^x$ and comment about the function $y = a^x$ where $a > 0$ is a constant (called the* **base***).*

Solution. We first form the discrete function a^n for integer values of n, simply by multiplying a by itself n times. This is analogous to Figure 10.1. So long as a is positive, we can "fill in" values of a^x when x is rational (in the same way as we did for 2^x), and we can smoothly connect the points to lead to the continuous function a^x for any real x. Given some positive constant a, we define the new function $f(x) = a^x$ as the exponential function with base a. Shown in Figure 10.3 are the functions $y = 2^x$, $y = 3^x$, $y = 4^x$ and $y = 10^x$. ♢

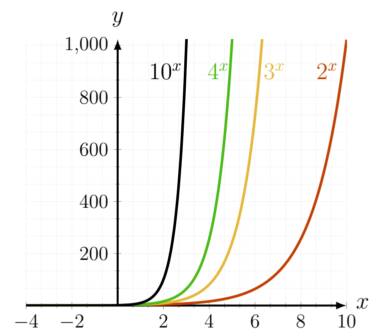


Figure 10.3: The function $y = f(x) = a^x$ is shown here for a variety of bases, $a = 2, 3, 4$, and 10.

10.2 Derivatives of exponential functions and the function e^x

Section 10.2 Learning goals

1. Using the definition of the derivative, calculate the derivative of the function $y = a^x$ for an arbitrary base $a > 0$.
2. Describe the significance of the special base e.
3. Summarize the properties of the function e^x, its derivatives, and how to manipulate it algebraically.
4. Recall the fact that the function $y = e^{kx}$ has a derivative that is proportional to the same function ($y = e^{kx}$).

Calculating the derivative of a^x

A screencast with the calculations for this section and motivation for the natural base e. Edu.Cr.

In this section we show how to compute the derivative of the exponential function. Rather then restricting attention to the special case $y = 2^x$, we consider an arbitrary positive constant a as the base. Note that the base has to be positive to ensure that the function is defined for all real x. For $a > 0$ let

$$y = f(x) = a^x.$$

Then, using the definition of the derivative,

$$\begin{aligned}
\frac{da^x}{dx} &= \lim_{h\to 0} \frac{(a^{x+h} - a^x)}{h} \\
&= \lim_{h\to 0} \frac{(a^x a^h - a^x)}{h} \\
&= \lim_{h\to 0} a^x \frac{(a^h - 1)}{h} \\
&= a^x \left[\lim_{h\to 0} \frac{a^h - 1}{h}\right].
\end{aligned}$$

The variable x appears only in the common factor a^x that can be factored out. The limit applies to h, not x. The terms inside square brackets depend only on the base a and on h, but once the limit is evaluated, that term is some constant (independent of x) that we denote by C_a. To summarize, we have found that

The **derivative of an exponential function** a^x is of the form $C_a a^x$ where C_a is a constant that depends only on the base a.

We now examine this in more detail with bases 2 and 10.

Example 10.3 (Derivative of 2^x) *Write down the derivative of $y = 2^x$ using the above result.*

Solution. For base $a = 2$, we have

$$\frac{d2^x}{dx} = C_2 \cdot 2^x,$$

where

$$C_2(h) = \lim_{h\to 0} \frac{2^h - 1}{h} \approx \frac{2^h - 1}{h} \quad \text{for small } h.$$

The decimal expansion value of C_2 is determined in the next example. ◊

Example 10.4 (The value of C_2) *Find an approximation for the value of the constant C_2 in Example 10.3 by calculating the value of the ratio $(2^h - 1)/h$ for small (finite) values of h, e.g., $h = 0.1, 0.01$, etc. Do these successive approximations for C_2 value approach a fixed real number?*

Solution. We take these successively smaller values of h and compute the value of $C_2 = (2^h - 1)/h$ on a spreadsheet.

The results are shown in Table 10.1, where we find that $C_2 \approx 0.6931$. (The actual value has an infinitely long decimal expansion that we here represent by its first few digits.) Thus, the derivative of 2^x is

$$\frac{d2^x}{dx} = C_2 \cdot 2^x \approx (0.6931) \cdot 2^x.$$

◊

Mastered Material Check

10. Describe geometrically the derivative of a^x.

Link to Google Sheets. The constant C_a in the derivative of a^x is calculated on this spreadsheet for $a = 2$. You can copy and paste this to our own spreadsheet and experiment with the value of the base a. Try to find a value of a between 2 and 3 for which C_a is close to 1.0.

h	C_2
0.1	0.717735
0.01	0.695555
0.001	0.693387
0.0001	0.693171
0.00001	0.693150
0.000001	0.693147
0.0000001	0.693147

Table 10.1: The constant C_2 in Example 10.4 is found by letting h get smaller and smaller. The value converges to $C_2 = 0.693147$.

Example 10.5 (The base 10 and the derivative of 10^x) *Determine the derivative of $y = f(x) = 10^x$.*

Solution. For base 10 we have

$$C_{10}(h) \approx \frac{10^h - 1}{h} \quad \text{for small } h.$$

We find, by similar approximation (Table 10.2), that $C_{10} \approx 2.3026$, so that

$$\frac{d10^x}{dx} = C_{10} \cdot 10^x \approx (2.3026) \cdot 10^x.$$

◊

Thus, the derivative of $y = a^x$ is proportional to itself, but the constant of proportionality (C_a) depends on the base.

h	C_{10}
0.1	2.589254
0.01	2.329299
0.001	2.305238
0.0001	2.302850
0.00001	2.302612
0.000001	2.302588
0.0000001	2.302585

Table 10.2: As in Table 10.1 but for the constant C_{10} in Example 10.5. (The advantage of using a spreadsheet is that we only need to change one cell to obtain this new set of values.)

Mastered Material Check

11. What does it mean for a function $f(x)$ to be proportional to itself?

The natural base e is convenient for calculus

In Examples 10.3-10.5, we found that the derivative of a^x is $C_a a^x$, where the constant C_a depends on the base. These constants are somewhat inconvenient, but unavoidable if we use an arbitrary base. Here we ask:

Does there exists a convenient base (to be called "e") for which the constant is particularly simple, namely such that $C_e = 1$?

This is the property of the **natural base** that we next identify.

We can determine such a hypothetical base using only the property that

$$C_e = \lim_{h \to 0} \frac{e^h - 1}{h} = 1.$$

This means that for small h

$$\frac{e^h - 1}{h} \approx 1,$$

so that

$$e^h - 1 \approx h \quad \Rightarrow \quad e^h \approx h + 1 \quad \Rightarrow \quad e \approx (1+h)^{1/h}.$$

More formally,

$$e = \lim_{h \to 0} (1+h)^{1/h}. \tag{10.1}$$

We can find an approximate decimal expansion for e by calculating the ratio in Eqn. (10.1) for some very small (but finite value) of h on a spreadsheet. Results are shown in Table 10.3. We find (e.g. for $h = 0.00001$) that

$$e \approx (1.00001)^{100000} \approx 2.71826.$$

Link to Google Sheets. The calculation of a decimal approximation to base e as shown in Table 10.3 .

h	approximation to e
0.1	2.5937425
0.01	2.7048138
0.001	2.7169239
0.0001	2.7181459
0.00001	2.7182682

Table 10.3: We can use a spreadsheet to find a decimal approximation to the natural base e using Eqn. (10.1) and letting h approach zero.

To summarize, we have found that for the special base, e, we have the following property:

The derivative of the function e^x is e^x.

The value of base e is obtained from the limit in Eqn. (10.1). This can be written in either of two equivalent forms.

The base of the natural exponential function is the real number defined as follows:

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

Mastered Material Check

12. Why can't we simply plug in $h = 0$ into Eqn. (10.1) evaluate the limit?
13. Let $h = \frac{1}{n}$ and rewrite Eqn (10.1).
14. Explain why each of Properties 1. → 8. hold for the function e^x.

Properties of the function e^x

We list below some of the key features of the function $y = e^x$. Note that all stem from basic manipulations of exponents as reviewed in Appendix B.1.

Use the slider to adjust the value of the base a in the function $y = a^x$; Compare your result with the function $y = e^x$. Explain what you see for $a > 1$, $a = 1$, $0 < a < 1$ and $a = 0$.

1. $e^a e^b = e^{a+b}$ as with all similar exponent manipulations.
2. $(e^a)^b = e^{ab}$ also stems from simple rules for manipulating exponents.
3. e^x is a function that is defined, continuous, and differentiable for all real numbers x.
4. $e^x > 0$ for all values of x.
5. $e^0 = 1$, and $e^1 = e$.
6. $e^x \to 0$ for increasing negative values of x.
7. $e^x \to \infty$ for increasing positive values of x.
8. The derivative of e^x is e^x (shown in this chapter).

Example 10.6 a) *Find the derivative of e^x at $x = 0$.*

b) *Show that the tangent line at that point is the line $y = x + 1$.*

Review: On this graph of $f(x) = e^x$ add a generic tangent line at any point x_0. (See Sections 5.1-5.2). Adjust a slider for x_0 to get the configuration shown in Fig. 10.4.

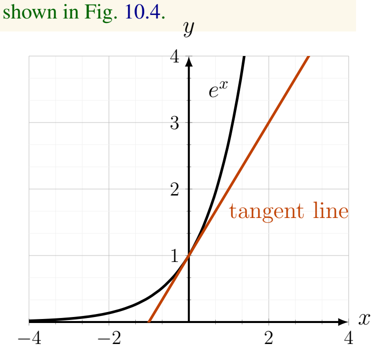


Figure 10.4: The function $y = e^x$ has the property that its tangent line at $x = 0$ has slope 1.

Solution.

a) The derivative of e^x is e^x. At $x = 0$, $e^0 = 1$.

b) The slope of the tangent line at $x = 0$ is therefore 1. The tangent line goes through $(0, e^0) = (0, 1)$ so it has a y-intercept of 1. Thus the tangent line at $x = 0$ with slope 1 is $y = x + 1$. This is shown in Figure 10.4. ◊

Composite derivatives involving exponentials

Using the derivative of e^x and the chain rule, we can now differentiate composite functions in which the exponential function appears.

Example 10.7 *Find the derivative of $y = e^{kx}$.*

Solution. Letting $u = kx$ gives $y = e^u$. Applying the simple chain rule leads to,

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

but

$$\frac{du}{dx} = k \quad \text{so} \quad \frac{dy}{dx} = e^u k = ke^{kx}.$$

◊

We highlight this result for future use:

The derivative of $y = e^{kx}$ is

$$\frac{dy}{dx} = ke^{kx}.$$

Mastered Material Check

15. Let $y = e^{5x}$. What is $\frac{dy}{dx}$?
16. Let $y = e^{\pi x}$. What is $\frac{dy}{dx}$?
17. List all constants in Example 10.8.
18. List all variables in Example 10.8.

Example 10.8 (Chemical reactions) *According to the collision theory of bimolecular gas reactions, a reaction between two molecules occurs when the molecules collide with energy greater than some activation energy, E_a, referred to as the Arrhenius activation energy. $E_a > 0$ is constant for the given substance. The fraction of bimolecular reactions in which this collision energy is achieved is*

$$F = e^{-(E_a/RT)},$$

where T is temperature (in degrees Kelvin) and $R > 0$ is the gas constant. Suppose that the temperature T increases at some constant rate, C, per unit time.

Determine the rate of change of the fraction F of collisions that result in a successful reaction.

Solution. This is a related rates problem involving an exponential function that depends on the temperature, which depends on time, $F = e^{-(E_a/RT(t))}$. We are asked to find the derivative of F with respect to time when the temperature increases.

We are given that $dT/dt = C$. Let $u = -E_a/RT$. Then $F = e^u$. Using the chain rule,

$$\frac{dF}{dt} = \frac{dF}{du}\frac{du}{dT}\frac{dT}{dt}.$$

Further, we have E_a, R, C are all constants, so

$$\frac{dF}{du} = e^u \quad \text{and} \quad \frac{du}{dT} = \frac{E_a}{RT^2}.$$

Assembling these parts, we have

$$\frac{dF}{dt} = e^u \frac{E_a}{RT^2} C = C\frac{E_a}{R} T^{-2} e^{-(E_a/RT)} = \frac{CE_a}{RT^2} e^{-(E_a/RT)}.$$

Thus, the rate of change of the fraction F of collisions that result in a successful reaction is given by the expression above. ◊

Featured Problem 10.1 (Ricker model for fish population growth) *Salmon are fish with non-overlapping generations. The adults lay eggs that are fertilized by males before the entire population dies. The eggs hatch to form a new generation. In Featured Problem 1.1, we considered one model for fish populations. Here we discuss a second model, the Ricker Equation, wherein the fish population this year, N_1, is related to the population last year, N_0, by the rule*

$$N_1 = N_0 e^{r\left(1 - \frac{N_0}{K}\right)}, \quad r, K > 0. \tag{10.2}$$

Here r is called an intrinsic growth rate, and K is the carrying capacity of the population. We investigate the following questions.

(a) *Is there a population level N_0 that would stay constant from one year to the next?*

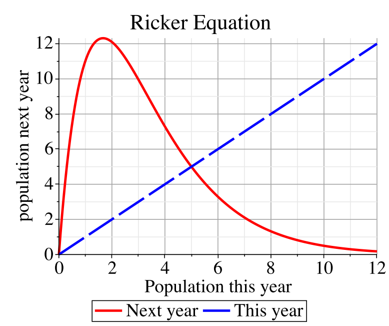


Figure 10.5: The functional form of the Ricker equation.

Adjust the sliders to observe how the parameters K and r affect the Ricker equation 10.2. What is special about the intersection of the two curves shown?

(b) *Simplify the notation by setting $x = N_0, y = N_1$. Compute the derivative dy/dx and interpret its meaning.*

(c) *What population level this year would result in the greatest possible population next year?*

The function e^x satisfies a new kind of equation

We divert our attention momentarily to an interesting observation. We have seen that the function

$$y = f(x) = e^x$$

satisfies the relationship

$$\frac{dy}{dx} = f'(x) = f(x) = y.$$

In other words, when differentiating, we get the same function back again. We summarize this observation:

The function $y = f(x) = e^x$ is equal to its own derivative. It hence satisfies the equation

$$\frac{dy}{dx} = y.$$

An equation linking a function and its derivative(s) is called a **differential equation**.

This is a new type of equation, unlike others previously seen in this course. In Chapters 11-13 we show that these differential equations have many applications to biology, physics, chemistry, and science in general.

10.3 Inverse functions and logarithms

Section 10.3 Learning goals

1. Explain the concept of inverse function from both algebraic and geometric points of view: given a function, determine whether (and for what restricted domain) an inverse function can be defined and sketch that inverse function.

2. Describe the relationship between the domain and range of a function and the range and domain of its inverse function. (Review Appendix C.5).

3. Apply these ideas to the logarithm, which is the inverse of an exponential function.

4. Reproduce the calculation of the derivative of $\ln(x)$ using implicit differentiation.

In this chapter we defined the new function e^x and computed its derivative. Paired with this newcomer is an inverse function, the natural logarithm, $\ln(x)$. Recall the following key ideas:

- Given a function $y = f(x)$, its inverse function, denoted $f^{-1}(x)$ satisfies
$$f(f^{-1}(x)) = x, \quad \text{and} \quad f^{-1}(f(x)) = x.$$
- The range of $f(x)$ is the domain of $f^{-1}(x)$ (and vice versa), which implies that in many cases, the relationship holds only on some subset of the original domains of the functions.
- The functions $f(x) = x^n$ and $g(x) = x^{1/n}$ are inverses of one another for all x when n is odd.
- The domain of a function (such as $y = x^2$ or other even powers) must be restricted (e.g. to $x \geq 0$) so that its inverse function ($y = \sqrt{x}$) is defined.
- On that restricted domain, the graphs of f and f^{-1} are mirror images of one another about the line $y = x$. Essentially, this stems from the fact that the roles of x and y are interchanged.

Mastered Material Check

19. Are $f(x) = x^n$ and $g(x) = x^{1/n}$ also inverses of one another for even integer n? Is this true for all x?
20. What is the inverse function for $y = x$? Over what range of values is the inverse defined?
21. What is the inverse function to $y = x^{2/3}$ and over what domain are the two functions inverses of one another?

Note symmetry about the line $y = x$ for this graph of $f(x) = x^n$ and $g(x) = x^{1/n}$. Adjust the slider for n to see how even and odd powers behave. What do you notice about the domain over which $g(x)$ is defined? Adjust the slider for a to observe "corresponding points" on the two graphs.

The natural logarithm is an inverse function for e^x

For $y = f(x) = e^x$ we define an inverse function, shown on Figure 10.6. We call this function the logarithm (base e), and write it as
$$y = f^{-1}(x) = \ln(x).$$

We have the following connection: $y = e^x$ implies $x = \ln(y)$. The fact that the functions are inverses also implies that
$$e^{\ln(x)} = x \quad \text{and} \quad \ln(e^x) = x.$$
The domain of e^x is $-\infty < x < \infty$, and its range is $x > 0$. For the inverse function, this domain and range are interchanged, meaning that $\ln(x)$ is only defined for $x > 0$ (its domain) and returns values in $-\infty < x < \infty$ (its range). As shown in Figure 10.6, the functions e^x and $\ln(x)$ are reflections of one another about the line $y = x$.

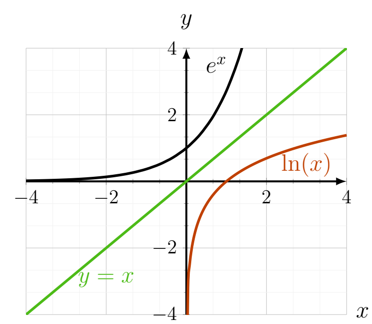


Figure 10.6: The function $y = e^x$ is shown with its inverse, $y = \ln x$.

Properties of the logarithm stem directly from properties of the exponential function. A review of these is provided in Appendix B.2. Briefly,

1. $\ln(ab) = \ln(a) + \ln(b)$,
2. $\ln(a^b) = b\ln(a)$,
3. $\ln(1/a) = \ln(a^{-1}) = -\ln(a)$.

Mastered Material Check

22. Give algebraic justification of the three properties of logarithms.

Featured Problem 10.2 (Agroforestry) *In agroforestry, the farming of crops is integrated with growing of trees to benefit productivity and maintain the health of an ecosystem. A tree can provide advantage to nearby plants by creating better soil permeability, higher water retention, and more stable*

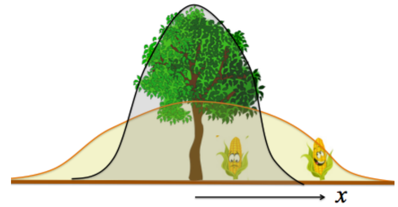


Figure 10.7: Too close to a tree, shading (grey) $S(x)$ interferes with crop growth. Just beyond this region, the advantage $A(x)$ to crop growth outweighs any disadvantage due to shading. We seek to find the optimal distance x for planting the crops.

temperatures. At the same time, trees produce shade and increased competition for nutrients. Both the advantage $A(x)$ and the shading $S(x)$ depend on distance from the tree, with shading a dominant negative effect right under the tree. Suppose that at a distance x from a given tree species, the net benefit B to a crop plant can be expressed as the difference

$$B(x) = A(x) - S(x), \quad \textit{where } \ A(x) = \alpha e^{-x^2/a^2}, \quad S(x) = \beta e^{-x^2/b^2}, \alpha\beta, a, b > 0$$

(a) How far away from the tree will the two influences break even? (b) Find the optimal distance to plan crops so that they derive maximal benefit from the nearby tree.

The Advantage $A(x)$, the shading effect $S(x)$, and the net benefit $B(x)$ for a crop as functions of distance x from a tree are shown here. Move the sliders to see how the spatial range a and the magnitude β affect the graphs.

Derivative of $\ln(x)$ *by implicit differentiation*

Implicit differentiation is helpful whenever an inverse function appears. Knowing the derivative of the original function allows us to compute the derivative of its inverse by using their relationship. We use implicit differentiation to find the derivative of $y = \ln(x)$.

First, restate the relationship in the inverse form, but consider y as the dependent variable - that is think of y as a quantity that depends on x:

$$y = \ln(x) \quad \Rightarrow \quad e^y = x \quad \Rightarrow \quad \frac{d}{dx}e^{y(x)} = \frac{d}{dx}x.$$

Applying the chain rule to the left hand side,

$$\frac{de^y}{dy}\frac{dy}{dx} = 1 \quad \Rightarrow \quad e^y\frac{dy}{dx} = 1 \quad \Rightarrow \quad \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}.$$

We have thus shown the following:

The derivative of $\ln(x)$ is $1/x$:

$$\frac{d\ln(x)}{dx} = \frac{1}{x}.$$

Inverse functions are mirror images of one another about the line $y = x$, since the role of independent and dependent variables are switched. Their tangent lines are also mirror images about the same line.

Tangent lines to the graphs of $y = e^x$ and $y = \ln(x)$ at **corresponding points** are mirror images about the line $y = x$. Adjust the slider to see the tangent lines at various points along the curves. What do we mean by “corresponding points”?

10.4 Applications of the logarithm

Section 10.4 Learning goals

1. Describe the relationships between properties of e^x and properties of its inverse $\ln(x)$, and master manipulations of expressions involving both.
2. Use logarithms for base conversions.
3. Use logarithms to solve equations involving the exponential function, i.e. solve $A = e^{bt}$ for t.)
4. Given a relationship such as $y = ax^b$, show that $\ln(y)$ is related linearly to $\ln(x)$, and use data points for (x, y) to determine the values of a and b.

Using the logarithm for base conversion

The logarithm is helpful in changing an exponential function from one base to another. We give some examples here.

Mastered Material Check

23. Why might one base be preferred over another?

Example 10.9 *Rewrite $y = 2^x$ in terms of base e.*

Solution. We apply ln and then exponentiate the result. Manipulations of exponents and logarithms lead to the desired results as follows:

$$y = 2^x \quad \Rightarrow \quad \ln(y) = \ln(2^x) = x\ln(2).$$

$$e^{\ln(y)} = e^{x\ln(2)} \quad \Rightarrow \quad y = e^{x\ln(2)}.$$

We find (using a calculator) that $\ln(2) = 0.6931\ldots$ This coincides with the value we computed earlier for C_2 in Example 10.4, so we have

$$y = e^{kx} \quad \text{where} \quad k = \ln(2) = 0.6931\ldots$$

◊

Example 10.10 *Find the derivative of $y = 2^x$.*

Solution. In Example 10.9 we expressed this function in the alternate form

$$y = 2^x = e^{kx} \quad \text{with} \quad k = \ln(2).$$

From Example 10.7 we have

$$\frac{dy}{dx} = ke^{kx} = \ln(2)e^{\ln(2)x} = \ln(2)2^x.$$

Through the above base conversion and chain rule, we relate the constant C_2 in Example 10.4 to the natural logarithm of 2: $C_2 = \ln(2)$. ◊

The logarithm helps to solve exponential equations

Equations involving the exponential function can sometimes be simplified and solved using the logarithm. We provide a few examples of this kind.

Example 10.11 *Find zeros of the function* $y = f(x) = e^{2x} - e^{5x^2}$.

Solution. We seek values of x for which $f(x) = e^{2x} - e^{5x^2} = 0$. We write

$$e^{2x} - e^{5x^2} = 0 \quad \Rightarrow \quad e^{2x} = e^{5x^2} \quad \Rightarrow \quad \frac{e^{5x^2}}{e^{2x}} = 1 \quad \Rightarrow \quad e^{5x^2-2x} = 1.$$

Taking logarithm of both sides, and using the facts that $\ln(e^{5x^2-2x}) = 5x^2 - 2x$ and $\ln(1) = 0$, we obtain

$$e^{5x^2-2x} = 1 \quad \Rightarrow \quad 5x^2 - 2x = 0 \quad \Rightarrow \quad x = 0, \frac{5}{2}.$$

We see that the logarithm is useful in the last step of isolating x, after simplifying the exponential expressions appearing in the equation. ♢

Andromeda Strain, revisited. In Section 10.1 we posed the question: how long does it take for the Andromeda strain population to attain a size of $6 \cdot 10^{36}$ cells, i.e. to grow to an Earth-sized colony? We now solve this problem using the continuous exponential function and the logarithm.

Recall that the bacterial doubling time is 20 min. If time is measured in minutes, the number, $B(t)$ of bacteria at time t could be described by the smooth function:

$$B(t) = 2^{t/20}.$$

Mastered Material Check

24. Verify that $B(t)$ agrees with Figure 10.1 and give powers of 2 at $t = 20, 40, 60, 80, \ldots$ minutes.
25. When, in general, will $B(t)$ give a power of 2?

Example 10.12 (The Andromeda strain) *Starting from a single cell, how long does it take for an E. coli colony to reach size of* $6 \cdot 10^{36}$ *cells by doubling every 20 minutes?*

Solution. We compute the time it takes by solving for t in $B(t) = 6 \cdot 10^{36}$, as shown below.

$$6 \cdot 10^{36} = 2^{t/20} \quad \Rightarrow \quad \ln(6 \cdot 10^{36}) = \ln(2^{t/20})$$

$$\ln(6) + 36\ln(10) = \frac{t}{20}\ln(2).$$

Solving for t,

$$t = 20\frac{\ln(6) + 36\ln(10)}{\ln(2)} = 20\frac{1.79 + 36(2.3)}{0.693} = 2441.27 \text{ min} = \frac{2441.27}{60} \text{ hr}.$$

Hence, it takes nearly 41 hours (but less than 2 days) for the colony to "grow to the size of planet Earth" (assuming the implausible scenario of unlimited growth). ♢

Example 10.13 (Using base e) *Express the number of bacteria in terms of base e (for practice with base conversions).*

Solution. Given $B(t) = 2^{t/20}$ is the number of bacteria at time t, we proceed as follows:

$$B(t) = 2^{t/20} \quad \Rightarrow \quad \ln(B(t)) = \frac{t}{20}\ln(2),$$

$$e^{\ln(B(t))} = e^{\frac{t}{20}\ln(2)} \quad \Rightarrow \quad B(t) = e^{kt} \quad \text{where} \quad k = \frac{\ln(2)}{20} \text{ per min.}$$

◊

The constant k has units of 1/time. We refer to k as the growth rate of the bacteria. We observe that this constant can be written as:

$$k = \frac{\ln(2)}{\text{doubling time}}.$$

As we see next, this approach is helpful in scientific applications.

Logarithms help plot data that varies on large scale

Living organisms come in a variety of sizes, from the tiniest cells to the largest whales. Comparing attributes across species of vastly different sizes poses a challenge, as visualizing such data on a simple graph obscures both extremes.

Suppose we wish to compare the physiology of organisms of various sizes, from that of a mouse to that of an elephant. An example of such data for metabolic rate versus mass of the animal is shown in Table 10.4.

animal	**body weight** M (gm)	**basal metabolic rate** (BMR)
mouse	25	1580
rat	226	873
rabbit	2200	466
dog	11700	318
man	70000	202
horse	700000	106

Table 10.4: Animals of various sizes (mass M in gm) have widely different basal metabolic rates (BMR, generally measured in terms of oxygen consumption rate, i.e. ml O_2 consumed per hr). A log-log plot of this data is shown in Figure 10.8.

It would be hard to see all data points clearly on a regular graph. For this reason, it is helpful to use logarithmic scaling for either or both variables. We show an example of this kind of **log-log plot**, where both axes use logarithmic scales, in Figure 10.8.

In allometry, it is conjectured that such data fits some power function of the form

$$y \approx ax^b, \text{ where } a, b > 0. \tag{10.3}$$

Note: this is not an exponential function, but a power function with power b and coefficient a.

Finding the **allometric constants** a and b using the graph in Fig 10.8 is now explained.

Mastered Material Check

26. Use software to plot the data given in Table 10.4. Why is it so hard to plot on a regular graph?

Example 10.14 (Log transformation) *Define $Y = \ln(y)$ and $X = \ln(x)$. Show that (10.3) can be rewritten as a linear relationship between Y and X.*

Solution. We have

$$Y = \ln(y) = \ln(ax^b) = \ln(a) + \ln(x^b) = \ln(a) + b\ln(x) = A + bX,$$

where $A = \ln(a)$. Thus, we have shown that X and Y are related linearly:

$$Y = A + bX, \quad \text{where} \quad A = \ln(a).$$

This is the equation of a straight line with slope b and Y intercept A. ◊

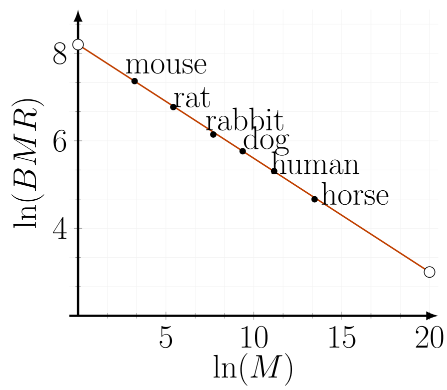


Figure 10.8: A log-log plot of the data in Table 10.4, showing ln(BMR) versus $\ln(M)$.

Example 10.15 (Finding the constants) *Use the straight line superimposed on the data in Figure 10.8 to estimate the values of the constants a and b.*

Solution. We use the straight line that has been fitted to the data in Figure 10.8. The Y intercept is roughly 8.2. The line goes approximately through $(20,3)$ and $(0,8.2)$ (open dots on plot) so its slope is $\approx (3-8.2)/20 = -0.26$. According to the relationship we found in Example 10.14,

$$8.2 = A = \ln(a) \quad \Rightarrow \quad a = e^{8.2} = 3640, \quad \text{and} \quad b = -0.26.$$

Thus, reverting to the original allometric relationship leads to

$$y = ax^b = 3640x^{-0.26} = \frac{3640}{x^{0.26}}.$$

From this we see that the metabolic rate y decreases with the size x of the animal, as indicated by the data in Table 10.4.

10.5 Summary

1. We reviewed exponential functions of the form $y = a^x$, where $a > 0$, the base, is constant.
2. The function $y = e^x$ is its own derivative, that is $\frac{dy}{dx} = e^x$. This function satisfies $\frac{dy}{dx} = y$, which is an example of a **differential equation**.
3. If $y = f(x)$, its inverse function is denoted $f^{-1}(x)$ and satisfies $f(f^{-1}(x)) = x$ and $f^{-1}(f(x)) = x$. The graph of f^{-1} is the same as the graph of f reflected across the line $y = x$. The domain of a function may have to be restricted so that its inverse function exists.
4. Let $f(x) = e^x$. The inverse of this function is $f^{-1}(x) = \ln(x)$. The derivative of $\ln(x)$ is $\frac{1}{x}$.
5. We can transform exponential relationships into linear relationships using logarithms. Such transformations allow for more meaningful plots, and can aid us in finding unknown constants in exponential relationships.
6. The applications in this chapter included:

(a) the Andromeda strain of E. coli (a bacterium) and its doubling;
(b) the Ricker equation for fish population growth from one year to the next;
(c) chemical reactions: the fraction which result in a successful reaction;
(d) how the advantage and disadvantage of plants growing near a tree depend on distance from the tree; and
(e) allometry: the relationship between body weight and basal metaboloci rate.

Quick Concept Checks

1. Instead of 1 E. coli cell, suppose we began with 2 which also doubled every 20 min. How long would it take for the population to grow to the size of the earth?

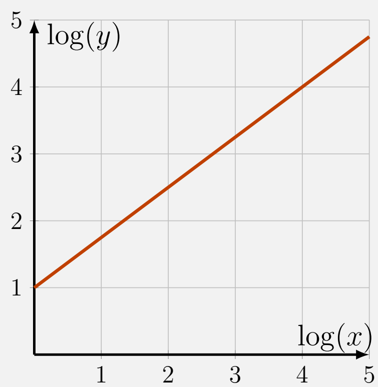
2. Given $\sqrt{3} \approx 1.74205$, compute without taking square roots:

   (a) $3^{3/2}$,

   (b) $3^{5/2}$,

3. Let $x = e^{\rho a}$. Determine $\frac{dx}{da}$.

4. Consider the following log-log plot



   (a) Let $Y = \log(y)$ and $X = \log(x)$. Find constants A and B such that $Y = AX + B$.

   (b) Determine constants a and b such that $y = ax^b$.

Exercises

10.1. **Polymerase Chain reaction (PCR).** The polymerase chain reaction (PCR) was invented by Mullis in 1983 to amplify DNA. It is based on the fact that each strand of (double-stranded) DNA can act as a template for the synthesis of the second ("complementary") strand. The method consists of repeated cycles of heating (which separates the DNA strands) and cooling (allowing for new DNA to be assembled on each strand). The reaction mixture includes the original DNA to be amplified, plus enzymes and nucleotides, the components needed to form the new DNA). Each cycle doubles the amount of DNA.

A particular PCR experiment consisted of 35 cycles.

(a) By what factor was the original DNA amplified? Give your answer both in terms of powers of 2 and in approximate decimal (powers of ten) notations.

(b) Use the approximation in the caption of Table 10.1 (rather than a scientific calculator) to find the decimal approximation.

10.2. **Invention of the game of chess.** According to some legends, the inventor of the game of chess (who lived in India thousands of years ago) so pleased his ruler, that he was asked to chose his reward.

"I would be content with grains of wheat. Let one grain be placed on the first square of my chess board, and double that number on the second, double that on the third, and so on," said the inventor. The ruler gladly agreed.

A chessboard has 8×8 squares. How many grains of wheat would be required for the last square on that board? Give your answer in decimal notation.

Note: in the original wheat and chessboard problem, we are asked to find the total number of wheat grains on all squares. This requires summing a geometric series, and is a problem ideal for early 2nd term calculus.

10.3. **Computing powers of** 2**.** In order to produce the graph of the continuous function 2^x in Figure 10.2, it was desirable to generate many points on that graph using simple calculations. Suppose you have an ordinary calculator with the operations $+, -, \times, /$. You also know that $\sqrt{2} \approx 1.414$.

How would you compute 2^x for the values $x = 7/2$, $x = -1/2$, and $x = -5$?

10.4. **Exponential base requirement.** Explain the requirement that a must be positive in the exponential function $y = a^x$. What could go wrong if a was a negative base?

10.5. **Derivative of 3^x.** Find the derivative of $y = 3^x$. What is the value of the multiplicative constant C_3 that shows up in your calculation?

10.6. **Graphing functions.** Graph the following functions:

(a) $f(x) = x^2 e^{-x}$,

(b) $f(x) = \ln(e^{2x})$.

10.7. **Changing bases.** Express the following in terms of base e:

(a) $y = 3^x$,

(b) $y = \frac{1}{7^x}$,

(c) $y = 15^{x^2+2}$.

Express the following in terms of base 2:

(d) $y = 9^x$,

(e) $y = 8^x$,

(f) $y = -e^{x^2+3}$.

Express the following in terms of base 10:

(g) $y = 21^x$,

(h) $y = 1000^{-10x}$,

(i) $y = 50^{x^2-1}$.

10.8. **Comparing numbers expressed using exponents.** Compare the values of each pair of numbers (i.e. indicate which is larger):

(a) $5^{0.75}, 5^{0.65}$

(b) $0.4^{-0.2}, 0.4^{0.2}$

(c) $1.001^2, 1.001^3$

(d) $0.999^{1.5}, 0.999^{2.3}$

10.9. **Logarithms.** Rewrite each of the following equations in logarithmic form:

(a) $3^4 = 81$,

(b) $3^{-2} = \frac{1}{9}$,

(c) $27^{-\frac{1}{3}} = \dfrac{1}{3}$.

10.10. **Equations with logarithms.** Solve the following equations for x:

(a) $\ln x = 2\ln a + 3\ln b$,

(b) $\log_a x = \log_a b - \frac{2}{3}\log_a c$.

10.11. **Reflections and transformations.** What is the relationship between the graph of $y = 3^x$ and the graph of each of the following functions?

(a) $y = -3^x$,

(b) $y = 3^{-x}$,

(c) $y = 3^{1-x}$,

(d) $y = 3^{|x|}$,

(e) $y = 2\cdot 3^x$,

(f) $y = \log_3 x$.

10.12. **Equations with exponents and logarithms.** Solve the following equations for x:

(a) $e^{3-2x} = 5$,

(b) $\ln(3x-1) = 4$,

(c) $\ln(\ln(x)) = 2$,

(d) $e^{ax} = Ce^{bx}$, where $a \neq b$ and $C > 0$.

10.13. **Derivative of exponential and logarithmic functions.** Find the first derivative for each of the following functions:

(a) $y = \ln(2x+3)^3$,

(b) $y = \ln^3(2x+3)$,

(c) $y = \ln(\cos \frac{1}{2}x)$,

(d) $y = \log_a(x^3 - 2x)$

(e) $y = e^{3x^2}$,

(f) $y = a^{-\frac{1}{2}x}$,

(g) $y = x^3 \cdot 2^x$,

(h) $y = e^{e^x}$,

(i) $y = \dfrac{e^t - e^{-t}}{e^t + e^{-t}}$.

Formula.
$\frac{d}{dx}(\log_a x) = \frac{1}{x \ln a}$

10.14. **Maximum, minimum and inflection points.** Find the maximum and minimum points as well as all inflection points of the following functions:

(a) $f(x) = x(x^2-4)$,

(b) $f(x) = x^3 - \ln(x), x > 0$,

(c) $f(x) = xe^{-x}$,

(d) $f(x) = \frac{1}{1-x} + \frac{1}{1+x}, -1 < x < 1$,

(e) $f(x) = x - 3\sqrt[3]{x}$,

(f) $f(x) = e^{-2x} - e^{-x}$.

10.15. **Using graph information.** Shown in Figure 15 is the graph of $y = Ce^{kt}$ for some constants C, k, and a tangent line. Use data from the graph to determine C and k.

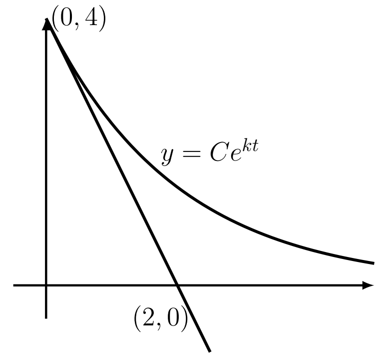


Figure 10.9: Figure for Exercise 15; $y = Ce^{kt}$ and a tangent line.

10.16. **Comparing exponential functions.** Consider the two functions

1. $y_1(t) = 10e^{-0.1t}$,
2. $y_2(t) = 10e^{0.1t}$.

Answer the following:

(a) Which one is decreasing and which one is increasing?

(b) In each case, find the value of the function at $t = 0$.

(c) Find the time at which the increasing function has doubled from this initial value.

(d) Find the time at which the decreasing function has fallen to half of its initial value.

Note: these values of t are called the doubling time, and half-life, respectively

10.17. **Invasive species.** An ecosystem with mature trees has a relatively constant population of beetles (species 1) - around 10^9. At $t = 0$, a single reproducing invasive beetle (species 2) is introduced accidentally.

If this population grows at the exponential rate

$$N_2(t) = e^{rt}, \quad \text{where } r = 0.5 \text{ per month},$$

how long does it take for species 2 to overtake the population of the resident species 1? Assume exponential growth for the entire duration.

10.18. **Human population growth.** It is sometimes said that the population of humans on Earth is growing exponentially. This means

$$P(t) = Ce^{rt}, \quad \text{where } r > 0.$$

We investigate this claim. To this end, we consider the human population beginning in year 1800 ($t = 0$). Hence, we ask whether the data in Table 10.5 fits the relationship

$$P(t) = Ce^{r(t-1800)}, \quad \text{where } t \text{ is time in years and } r > 0?$$

(a) Show that the above relationship implies that $\ln(P)$ is a linear function of time, and that r is the slope of the linear relationship (*hint:* take the natural logarithm of both sides of the relationship and simplify).

(b) Use the data from Table 10.5 for the years 1800 to 2020 to investigate whether $P(t)$ fits an exponential relationship (*hint:* plot $\ln(P)$, where P is human population (in billions) against time t in years - we refer to this process as "transforming the data".

(c) A spreadsheet can be used to fit a straight line through the transformed data you produced in (b).

(i) Find the best fit for the growth rate parameter r using that option.

(ii) What are the units of r?

(iii) What is the best fit value of C?

(d) Based on your plot of $\ln(P)$ versus t and the best fit values of r and C, over what time interval was the population growing more slowly than the overall trend, and when was it growing more rapidly than this same overall trend?

(e) Under what circumstances could an exponentially growing population be **sustainable**?

year	**human population** (billions)
1	0.2
1000	0.275
1500	0.45
1650	0.5
1750	0.7
1804	1
1850	1.2
1900	1.6
1927	2
1950	2.55
1960	3
1980	4.5
1987	5
1999	6
2011	7
2020	7.7

Table 10.5: The human population (billions) over the years AD 1 to AD 2020.

10.19. **A sum of exponentials.** Researchers that investigated the molecular motor dynein found that the number of motors $N(t)$ remaining attached to their microtubule tracks at time t (in sec) after a pulse of activation was well described by a double exponential of the form

$$N(t) = C_1 e^{-r_1 t} + C_2 e^{-r_2 t}, \quad t \geq 0.$$

They found that $r_1 = 0.1, r_2 = 0.01$ per second, and $C_1 = 75, C_2 = 25$ percent.

(a) Plot this relationship for $0 < t < 8$ min. Which of the two exponential terms governs the behaviour over the first minute? Which dominates in the later phase?

(b) Now consider a plot of $\ln(N(t))$ versus t. Explain what you see and what the slopes and other aspects of the graph represent.

10.20. **Exponential Peeling.** The data in Table 10.6 is claimed to have been generated by a double exponential function of the form

$$N(t) = C_1 e^{-r_1 t} + C_2 e^{-r_2 t}, \quad t \geq 0.$$

Use the data to determine the values of the constants r_1, r_2, C_1, and C_2.

time	$N(t)$
0.0000	100.0000
0.1000	57.6926
0.2000	42.5766
0.3000	35.8549
0.4000	31.8481
0.5000	28.8296
2.5000	4.7430
4.5000	0.7840
6.0000	0.2032
8.0000	0.0336

Table 10.6: Table for Exercise 20; data to be fit to a function of the form $N(t) = C_1 e^{-r_1 t} + C_2 e^{-r_2 t}, \quad t \geq 0$.

10.21. **Shannon Entropy.** In a recent application of information theory to the field of genomics, a function called the Shannon entropy, H, was considered. In it, a given gene is represented as a binary device: it can be either “on” or “off” (i.e. being expressed or not).

If x is the probability that the gene is “on” and y is the probability that it is “off”, the Shannon entropy function for the gene is defined as

$$H = -x\log(x) - y\log(y)$$

Note that

- x and y being probabilities just means that they satisfy $0 < x \leq 1$, and $0 < y \leq 1$ and
- the gene can only be in one of these two states, so $x + y = 1$.

Use these facts to show that the Shannon entropy for the gene is greatest when the two states are equally probable, i.e. for $x = y = 0.5$.

10.22. **A threshold function.** The response of a regulatory gene to inputs that affect it is not simply linear. Often, the following so-called “squashing function” or “threshold function” is used to link the input x to the output y of the gene:

$$y = f(x) = \frac{1}{1 + e^{(ax+b)}},$$

where a, b are constants.

(a) Show that $0 < y < 1$.

(b) For $b = 0$ and‘$a = 1$ sketch the shape of this function.

(c) How does the shape of the graph change as a increases?

10.23. **Graph sketching.** Sketch the graph of the function $y = e^{-t}\sin \pi t$.

10.24. **The Mexican Hat.** Consider the function

$$y = f(x) = 2e^{-x^2} - e^{-x^2/3}$$

(a) Find the critical points of f.

(b) Determine the value of f at those critical points.

(c) Use these results and the fact that for very large x, $f \to 0$ to draw a rough sketch of the graph of this function.

(d) Comment on why this function might be called “a Mexican Hat”.

Note: The second derivative is not very informative here, and we do not ask you to use it for determining concavity in this example. However, you may wish to calculate it for practice with the chain rule.

10.25. **The Ricker Equation.** In studying salmon populations, a model often used is the Ricker equation which relates the size of a fish population this year, x to the expected size next year y. The Ricker equation is

$$y = \alpha x e^{-\beta x}$$

where $\alpha, \beta > 0$.

(a) Find the value of the current population which maximizes the salmon population next year according to this model.

(b) Find the value of the current population which would be exactly maintained in the next generation.

(c) Explain why a very large population is not sustainable.

Note: these populations do not actually change continuously, since all the parents die before the eggs are hatched.

10.26. **Spacing in a fish school.** Life in a social group has advantages and disadvantages: protection from predators is one advantage. Disadvantages include competition for food or resources. Spacing of individuals in a school of fish or a flock of birds is determined by the mutual attraction and repulsion of neighbours from one another: each individual does not want to stray too far from others, nor get too close.

Suppose that when two fish are at distance $x > 0$ from one another, they are attracted with “force” F_a and repelled with “force” F_r given by:

$$F_a = Ae^{-x/a}$$

$$F_r = Re^{-x/r}$$

where A, R, a, r are positive constants.

Note: A, R are related to the magnitudes of the forces, while a, r to the spatial range of these effects.

(a) Show that at distance $x = a$, the first function has fallen to $(1/e)$ times its value at the origin. (Recall $e \approx 2.7$.)

(b) For what value of x does the second function fall to $(1/e)$ times its value at the origin? Note that this is the reason why a, r are called spatial ranges of the forces.

(c) It is generally assumed that $R > A$ and $r < a$. Interpret what this mean about the comparative effects of the forces.

(d) Sketch a graph showing the two functions on the same set of axes.

(e) Find the distance at which the forces exactly balance. This is called the comfortable distance for the two individuals.

(f) If either A or R changes so that the ratio R/A decreases, does the comfortable distance increase or decrease? Justify your response.

(g) Similarly comment on what happens to the comfortable distance if a increases or r decreases.

10.27. **Seed distribution.** The density of seeds at a distance x from a parent tree is observed to be

$$D(x) = D_0 e^{-x^2/a^2},$$

where $a > 0, D_0 > 0$ are positive constants. Insects that eat these seeds tend to congregate near the tree so that the fraction of seeds that get eaten is

$$F(x) = e^{-x^2/b^2}$$

where $b > 0$.

Note: These functions are called Gaussian or Normal distributions. The parameters a, b are related to the "width" of these bell-shaped curves.

The number of seeds that survive (i.e. are produced and not eaten by insects) is

$$S(x) = D(x)(1 - F(x))$$

Determine the distance x from the tree at which the greatest number of seeds survive.

10.28. **Euler's '*e*'.** In 1748, Euler wrote a classic book on calculus, "Introductio in Analysin Infinitorum" [Euler, 1748]in which he showed that the function e^x could be written in an expanded form similar to an (infinitely long) polynomial:

$$e^x = 1 + x + \frac{x^2}{1 \cdot 2} + \frac{x^3}{1 \cdot 2 \cdot 3} + \ldots$$

Use as many terms as necessary to find an approximate value for the number e and for $1/e$ to 5 decimal places.

Note: in other mathematics courses we see that such expansions, called power series, are central to approximations of many functions.

11
Differential equations for exponential growth and decay

In Section 10.2 we made an observation about exponential functions and a new kind of equation - a **differential equation** - that such functions satisfy. In this chapter we explore this observation in more detail. At first, this link is based on the simple relationship between an exponential function and its derivatives. Later, this expands into a more encompassing discussion of

1. how differential equations arise in scientific problems,
2. how we study their predictions, and
3. what their solutions can tell us about the natural world.

We begin by reintroducing these equations.

11.1 Introducing a new kind of equation

A screencast summary of the introduction: differential equations for exponential growth and decay. Edu.Cr.

Section 11.1 Learning goals

1. Explain that the exponential function and its derivative are proportional to one another, and thereby satisfy a relationship of the form $dy/dx = ky$.
2. Give the definitions of a differential equation and of a solution to a differential equation.
3. Explain that $y = e^{kt}$ is a solution to the differential equation $dy/dt = ky$.

Observations about the exponential function

In Chapter 10, we introduced the exponential function $y = f(x) = e^x$, and noted that it satisfies the relationship

$$\frac{de^x}{dx} = e^x, \quad \Rightarrow \quad \frac{dy}{dx} = y.$$

The equation on the right (linking a function to its own derivative) is a new kind of equation called a **differential equation** (abbreviated DE). We say that

$f(x) = e^x$ is a function that "satisfies" the equation, and we call this a **solution to the differential equation**.

Note: The solution to an algebraic equation is a number, whereas the solution to a differential equation is a function.

We call this a **differential equation** because it connects (one or more) derivatives of a function with the function itself.

Mastered Material Check

1. For whar constant C does $y = Ce^x$ satisfy the differential equation $dy/dx = y$?
2. What function satisfies the DE $dy/dz = y$?

Definition 11.1 (Differential equation) *A differential equation is a mathematical equation that relates one or more derivatives of some function to the function itself. Solving the differential equation is the process of identifying the function(s) that satisfies the given relationship.*

We will be interested in applications in which a system or process varies over time. For this reason, we will henceforth use the independent variable t, for **time** in place of the former generic "x".

Observations.

1. Consider the function of time: $y = f(t) = e^t$. Show that this function satisfies the differential equation

$$\frac{dy}{dt} = y.$$

Hint: Notice that we merely changed the notation very slightly. Now the derivative is "with respect to" t rather than x.

2. The functions $y = e^{kt}$ (for k constant) satisfy the differential equation

$$\frac{dy}{dt} = ky. \tag{11.1}$$

We can verify by differentiating $y = e^{kt}$, using the chair rule. Setting $u = kt$, and $y = e^u$, we have

$$\frac{dy}{dt} = \frac{dy}{du}\frac{du}{dt} = e^u \cdot k = ke^{kt} = ky \quad \Rightarrow \quad \frac{dy}{dt} = ky$$

Hence, we have established that $y = e^{kt}$ satisfies the DE (11.1).

It is interesting to ask: *I*s this is the only function that satisfies the differential equation 11.1? Are there other possible solutions? What about a function such as $y = 2e^{kt}$ or $y = 400e^{kt}$?

The reader should show that for any constant C, the function $y = Ce^{kt}$ is a solution to the DE (11.1). To do so, differentiate the function and plug into (11.1). Verifying that the two sides of the equation are then the same establishes the result. While we do not prove it here, it turns out that $y = Ce^{kt}$ are the *only* functions that satisfy Eqn. (11.1).

Hint: Notice that the constant C in front will appear in both the derivative and the function, and so will not change the equation.

Let us summarize what we have found out so far:

Solutions to the differential equation

$$\frac{dy}{dt} = ky \tag{11.2}$$

are the functions

$$y = Ce^{kt} \tag{11.3}$$

for C an arbitrary constant.

A few comments are in order. First, unlike *algebraic* equations - whose solutions are numbers - **differential equations** have solutions that are *functions*. Second, the constant k that appears in Eqn. (11.2), is the same as the constant k in e^{kt}. Depending on the sign of k, we get either

a) *exponential growth* for $k > 0$, as illustrated in Figure 11.1(a), or

b) *exponential decay* for $k < 0$, as illustrated in Figure 11.1(b).

Third, since e^{kt} is always positive, the constant C determines the sign of the function as a whole - whether its graph lies above or below the t axis.

A few curves of each type ($C > 0, C < 0$) are shown in each panel of Figure 11.1. The collection of curves in a panel is called a **family** of solution curves. The family shares the same value of k, but each member has a distinct value of C. Next, we ask how to specify a particular member of the family as *the* solution.

Mastered Material Check

3. Give an example of an algebraic equation and its solution.
4. Verify that $y = 3e^{-t}$ satisfies differential equation $\frac{dy}{dt} = -y$.
5. Why is e^{kt} always positive?
6. Plot, using software, $y = Ce^{t}$ for each of $C = -4, -2, 2$ and 4.
7. Plot, using software, $y = Ce^{-t}$ for each of $C = -4, -2, 2$, and 4.

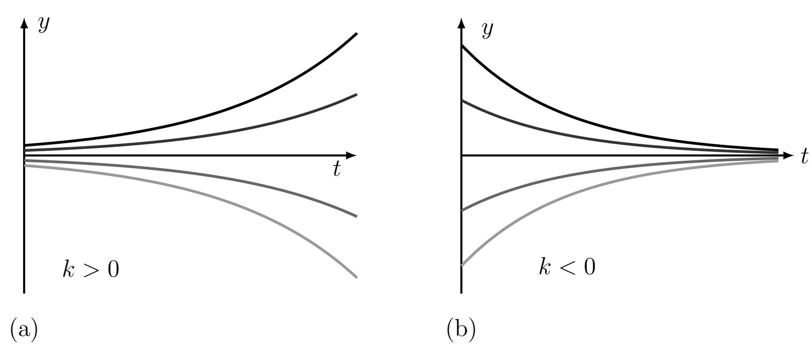


Figure 11.1: (a) A family of solutions to the differential equation (DE) (11.2). These are functions of the form $y = Ce^{kt}$ for $k > 0$ and arbitrary constant C. (b) Ampther family of solutions of a DE of the form (11.2), but for $k < 0$.

The solution to a differential equation

Definition 11.2 (Solution to a differential equation) *By a* **solution** *to a differential equation, we mean a function that satisfies that equation.*

We often refer to "solution curves" - the graphs of the family of solutions of a differential equation, as shown, for example in the panels of Figure 11.1.

So far, we found that "many" functions can be valid solutions of the differential equation (11.2), since we can chose the constant C arbitrarily in the family of solutions $y = Ce^{kt}$. Hence, in order to distinguish one specific solution of interest, we need additional information. This additional information is called an **initial value**, or **initial condition**, and it specifies one point belonging to the solution curve of interest. A common way to set an initial value is to specify a fixed value of the function (say $y = y_0$) at time $t = 0$.

Definition 11.3 (Initial value) *An* **initial value** *for a differential equation is a specified, known value of the solution at some specific time point (usually at time $t = 0$).*

Example 11.1 *Given the differential Eqn. (11.2) and the initial value*

$$y(0) = y_0,$$

find the value of C for the solution in Eqn. (11.3).

Solution. We proceed as follows:

$$y(t) = Ce^{kt}, \quad \text{so} \quad y(0) = Ce^{k \cdot 0} = Ce^0 = C \cdot 1 = C.$$

But, by the initial condition, $y(0) = y_0$. So,

$$C = y_0$$

and we have established that

$$y(t) = y_0 e^{kt}, \quad \text{where } y_0 \text{ is the initial value.}$$

◊

For example, in Figure 11.1, the initial value specifies that the solution we want passes through a specific point in the ty-plane - namely, the point $(0, y_0)$. Only one curve in the family of curves has that property. Hence, the initial value picks out a unique solution.

Example 11.2 *Find the solution to the differential equation*

$$\frac{dy}{dt} = -0.5y$$

that satisfies the initial condition $y(0) = 3$. Describe the behaviour of the solution you have found.

Solution. The DE indicates that $k = -0.5$, so solutions are exponential functions $y = Ce^{-0.5t}$. The initial condition sets the value of C. From previous discussion, we know that $C = y(0) = 3$. Hence, the solution is $y = 3e^{-0.5t}$. This is a decaying exponential. ◊

Adjust the sliders in this interactive graph to see how the values of k and C affect the shape of the graph of the function $y = Ce^{kt}$ as well as its initial value $y(0) = y_0$.
Note the transitions that take place when k changes from positive to negative.

Mastered Material Check

8. Given differential Eqn. (11.2) and the initial value $y(0) = 1$, find C for the solution in Eqn (11.3).
9. Repeat the above but for the initial value $y(0) = 10$.
10. Draw the ty-plane with the points $(0, y_0)$ for $y_0 = 1, 10$.
11. Use differentiation to verify that the unction $y = 3e^{-0.5t}$ in Example 11.2 is a solution to $dy/dt = -0.5y$ with initial condition $y(0) = 3$.

11.2 Differential equation for unlimited population growth

Section 11.2 Learning goals

1. Recall the derivation of a model for human population growth and describe how it leads to a differential equation.
2. Identify that the solution to that equation is an exponential function.
3. Define per capita birth rates and rates of mortality, and explain the process of estimating their values from assumptions about the population.
4. Compute the doubling time of a population from its growth rate and vice versa.

A screencast summary of the model for (unlimited) human population growth.

Differential equations are important because they turn up in the study of many natural processes that vary continuously. In this section we examine the way that a simple differential equation arises when we study continuous uncontrolled population growth.

Here we set up a mathematical model for population growth. Let $N(t)$ be the number of individuals in a population at time t. The population changes with time due to births and mortality. (Here we ignore migration). Consider the changes that take place in the population size between time t and $t+h$, where $\Delta t = h$ is a small time increment. Then

$$N(t+h) - N(t) = \begin{bmatrix} \text{Change} \\ \text{in } N \end{bmatrix} = \begin{bmatrix} \text{Number of} \\ \text{births} \end{bmatrix} - \begin{bmatrix} \text{Number of} \\ \text{deaths} \end{bmatrix} \quad (11.4)$$

Mastered Material Check

12. What is the dependent variable in this model? The independent variable?
13. What are the units associated with each variable in this model?
14. What does "x is proportional to y" mean?

Eqn. (11.4) is just a "book-keeping" equation that keeps track of people entering and leaving the population. It is sometimes called a **balance equation**. We use it to derive a differential equation linking the *derivative* of N to the *value* of N at the given time.

Notice that dividing each term by the time interval h, we obtain

$$\frac{N(t+h)-N(t)}{h} = \left[\frac{\text{Number of births}}{h}\right] - \left[\frac{\text{Number of deaths}}{h}\right].$$

The term on the left "looks familiar". If we shrink the time interval, $h \to 0$, this term is a derivative dN/dt, so

$$\frac{dN}{dt} = \begin{bmatrix} \text{Rate of} \\ \text{change of N} \\ \text{per unit time} \end{bmatrix} = \begin{bmatrix} \text{Number of} \\ \text{births per} \\ \text{unit time} \end{bmatrix} - \begin{bmatrix} \text{Number of} \\ \text{deaths per} \\ \text{unit time} \end{bmatrix}$$

For simplicity, we assume that all individuals are identical and that the number of births per unit time is proportional to the population size. Denote by r the constant of proportionality. Similarly, we assume that the number of deaths per unit time is proportional to population size with m the constant of proportionality.

Both r and m have meanings: r is the average **per capita birth rate**, and m is the average **per capita mortality rate** . Here, both are assumed to be fixed positive constants that carry units of $1/\text{time}$. This is required to make the units match for every term in Eqn. (11.4). Then

$$r = \text{per capita birth rate} = \frac{\text{number births per unit time}}{\text{population size}},$$

$$m = \text{per capita mortality rate} = \frac{\text{number deaths per unit time}}{\text{population size}}.$$

Consequently, we have

$$\text{Number of births per unit time} = rN,$$
$$\text{Number of deaths per unit time} = mN.$$

We refer to constants such as r, m as **parameters**. In general, for a given population, these would have specific numerical values that could be found through experiment, by collecting data, or by making simple assumptions. In Section 11.2, we show how some elementary assumptions about birth and mortality could help to estimate approximate values of r and m.

Taking the assumptions and the form of the balance equation (11.4) together we have arrived at:

$$\frac{dN}{dt} = rN - mN = (r-m)N. \quad (11.5)$$

This is a differential equation: it links the derivative of $N(t)$ to the function $N(t)$. By solving the equation (i.e. identifying its solution), we are be able to make a projection about how fast a population is growing.

Define the constant $k = r - m$. Then k is the **net growth rate**, of the population, so

$$\frac{dN}{dt} = kN, \quad \text{for } k = (r-m).$$

Suppose we also know that at time $t = 0$, the population size is N_0. Then:

- The function that describes population over time is (by previous results),

$$N(t) = N_0 e^{kt} = N_0 e^{(r-m)t}. \quad (11.6)$$

  (The result is identical to what we saw previously, but with N rather than y as the time-dependent function. We can easily check by differentiation that this function satisfies Eqn. (11.5).)

- Since $N(t)$ represents a population size, it has to be non-negative to have biological relevance. This is true so long as $N_0 \geq 0$.

- The initial condition $N(0) = N_0$, allows us to specify the (otherwise arbitrary) constant multiplying the exponential function.

Mastered Material Check

15. If there are 10 births/year in a population of size 1000, what is the birth rate r? Give units.
16. If there are 11 deaths/year in a population of size 1000, what is the mortality rate m? Give units.
17. Given the above conditions, what is the net growth rate k for such a population? Give units. Is the population growing or shrinking?

- The population grows provided $k > 0$ which happens when $r - m > 0$ i.e. when birth rate exceeds mortality rate.
- If $k < 0$, or equivalently, $r < m$ then more people die on average than are born, so that the population shrinks and (eventually) go extinct.

A simple model for human population growth

The differential equation (11.5) and its initial condition led us to predict that a population grows or decays exponentially in time, according to Eqn. (11.6). We can make this prediction quantitative by estimating the values of parameters r and m. To this end, let us consider the example of a human population and make further simplifying assumptions. We measure time in years.

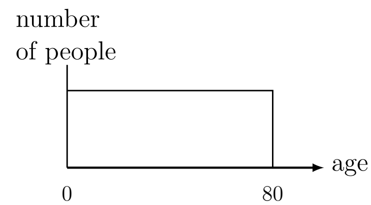


Figure 11.2: We assume a uniform age distribution to to determine the fraction of people who are fertile (and can give birth) or who are old (and likely to die). While slightly silly, this simplification helps estimate the desired parameters.

Assumptions.

- The age distribution of the population is “flat”, i.e. there are as many 10 year-olds as 70 year olds. Of course, this is quite inaccurate, but a good place to start since it is easy to estimate some of the quantities we need. Figure 11.2 shows such a **uniform age distribution**.
- The sex ratio is roughly 50%. This means that half of the population is female and half male.
- Women are fertile and can have babies only during part of their lives: we assume that the fertile years are between age 15 and age 55, as shown in Figure 11.3.
- A lifetime lasts 80 years. This means that for half of that time a given woman can contribute to the birth rate, or that $\frac{(55-15)}{80} = 50\%$ of women alive at any time are able to give birth.
- During a woman’s fertile years, we assume that on average, she has one baby every 10 years.
- We assume that deaths occur only from old age (i.e. we ignore disease, war, famine, and child mortality.)
- We assume that everyone lives precisely to age 80, and then dies instantly.

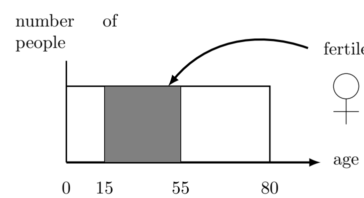


Figure 11.3: We assume that only women between the ages of 15 and 55 years old are fertile and can give birth. Then, according to our uniform age distribution assumption, half of all women are between these ages and hence fertile.

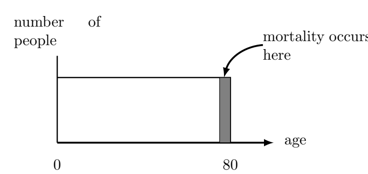


Figure 11.4: We assume that the people in the age bracket 79-80 years old all die each year, and that those are the only deaths. This, too, is a silly assumption, but makes it easy to estimate mortality in the population.

Based on the above assumptions, we can estimate the birthrate parameter r as follows:

$$r = \frac{\text{number women}}{\text{population}} \cdot \frac{\text{years fertile}}{\text{years of life}} \cdot \frac{\text{number babies per woman}}{\text{number of years}}$$

Thus we compute that

$$r = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{10} = 0.025 \text{ births per person per year.}$$

Note that this value is now a rate per person per year, averaged over the entire population (male and female, of all ages). We need such an average rate since our model of Eqn. (11.5) assumes that individuals "are identical". We now have an approximate value for the average human per capita birth rate, $r \approx 0.025$ per year.

Next, using our assumptions, we estimate the mortality parameter, m. With the flat age distribution shown in Figure 11.2, there would be a fraction of $1/80$ of the population who are precisely removed by mortality every year (i.e. only those in their 80^{th} year.) In this case, we can estimate that the per capita mortality is:

$$m = \frac{1}{80} = 0.0125 \text{ deaths per person per year.}$$

The net per capita growth rate is $k = r - m = 0.025 - 0.0125 = 0.0125$ per person per year. We often refer to the constant k as a **growth rate constant** and we also say that the population grows at the rate of 1.25% per year.

Example 11.3 *Using the results of this section, find a prediction for the population size $N(t)$ as a function of time t.*

Solution. We have found that our population satisfies the equation

$$\frac{dN}{dt} = (r - m)N = kN = 0.0125N,$$

so that

$$N(t) = N_0 e^{0.0125t}, \tag{11.7}$$

where N_0 is the starting population size. Figure 11.5 illustrates how this function behaves, using a starting value of $N(0) = N_0 = 7$ billion. ♢

Example 11.4 (Human population in 100 years) *Given the initial condition $N(0) = 7$ billion, determine the size of the human population at $t = 100$ years predicted by the model.*

Solution. At time $t = 0$, the population is $N(0) = N_0 = 7$ billion. Then in billions,

$$N(t) = 7e^{0.0125t}$$

so that when $t = 100$ we would have

$$N(100) = 7e^{0.0125 \cdot 100} = 7e^{1.25} = 7 \cdot 3.49 = 24.43.$$

Thus, with a starting population of 7 billion, there would be about 24.4 billion after 100 years based on the uncontrolled continuous growth model. ♢

A critique. Before leaving our population model, we should remember that our projections hold only so long as some rather restrictive assumptions are made. We have made many simplifications, and ignored many features that would seriously affect these results. These include (among others),

Mastered Material Check

18. Under these assumptions, for a population size of 800, how many male 35 year-olds would you expect? Women in their 60's?
19. Is the fertility assumption reasonable? Why or why not?
20. Explain the units attached to the birthrate parameter r.

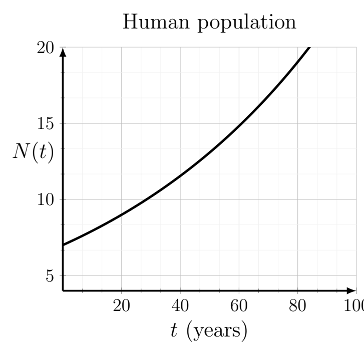


Figure 11.5: Projected world population (in billions) over 100 years, based on the model in Eqn. (11.7) and assuming that the initial population is ≈ 7 billion.

Mastered Material Check

21. Based on Figure 11.5, when would we expect the human population to reach 15 billion?

- variations in birth and mortality rates that stem from competition for resources and,
- epidemics that take hold when crowding occurs, and
- uneven distributions of resources or space.

We have also assumed that the age distribution is uniform (flat), but that is not accurate: the population grows only by adding new infants, and this would skew the distribution even if it is initially uniform. All these factors suggest that some "healthy skepticism" should be applied to any model predictions. Predictions may cease to be valid if model assumptions are not satisfied. This caveat will lead us to think about more realistic models for population growth. Certainly, the uncontrolled exponential growth would not be sustainable in the long run. That said, such a model is a good starting point for a first description of population growth, later to be adjusted.

Growth and doubling

In Chapter 10, we used base 2 to launch our discussion of exponential growth and population doublings. We later discovered that base e is more convenient for calculus, having a more elegant derivative. We also saw in Chapter 10, that bases of exponents can be inter-converted. These skills are helpful in our discussion of doubling times below.

The doubling time. How long would it take a population to double, given that it is growing exponentially with growth rate k? We seek a time t such that $N(t) = 2N_0$. Then

$$N(t) = 2N_0 \quad \text{and} \quad N(t) = N_0 e^{kt},$$

implies that the population has doubled when t satisfies

$$2N_0 = N_0 e^{kt}, \quad \Rightarrow \quad 2 = e^{kt} \quad \Rightarrow \quad \ln(2) = \ln(e^{kt}) = kt.$$

We solve for t. Thus, the **doubling time**, denoted τ is:

$$\tau = \frac{\ln(2)}{k}.$$

Mastered Material Check

22. What are the units associated with τ?
23. The human population hit 3 billion in 1959. How does this fit with our (imperfect) model?

Example 11.5 (Human population doubling time) *Determine the doubling time for the human population based on the results of our approximate growth model.*

Solution. We have found a growth rate of roughly $k = 0.0125$ per year for the human population. Based on this, it would take

$$\tau = \frac{\ln(2)}{0.0125} = 55.45 \text{ years}$$

for the population to double. Compare this with the graph of Fig 11.5, and note that over this time span, the population increases from 6 to 12 billion. ◊

Note: the observant student may notice that we are simply converting back from base *e* to base 2 when we compute the doubling time.

We summarize an important observation:

In general, an equation of the form

$$\frac{dy}{dt} = ky$$

that represents an exponential growth has a **doubling time** of

$$\tau = \frac{\ln(2)}{k}.$$

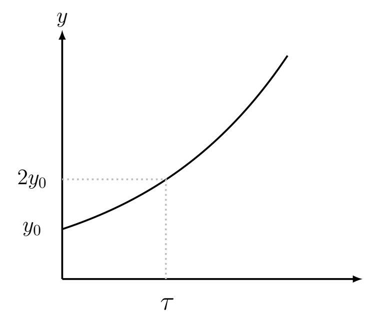


Figure 11.6: Doubling time for exponential growth.

This is shown in Figure 11.6. We have discovered that based on the uncontrolled growth model, the population doubles *every 55 years*! After 110 years, for example, there have been two doublings, or a quadrupling of the population.

Example 11.6 (A ten year doubling time) *Suppose we are told that some animal population doubles every* 10 *years. What growth rate would lead to such a trend?*

Solution. In this case, $\tau = 10$ years. Rearranging

$$\tau = \frac{\ln(2)}{k},$$

we obtain

$$k = \frac{\ln(2)}{\tau} = \frac{0.6931}{10} \approx 0.07 \text{ per year}.$$

Thus, a growth rate of 7% leads to doubling roughly every 10 years. ◊

11.3 Radioactive decay

Section 11.3 Learning goals

1. Describe the model for the number of radioactive atoms and explain how this leads to a differential equation.
2. Determine the solution of the resulting differential equation.
3. Given the initial amount, determine the amount of radioactivity remaining at a future time.
4. Describe the link between half-life of the radioactive material and its decay rate; given the value of one, be able to find the value of the other.

A radioactive material consists of atoms that undergo a spontaneous change. Every so often, some radioactive atom emits a particle, and decays into an inert form. We call this a process of **radioactive decay**. For any one atom, it is impossible to predict when this event would occur exactly, but based on the behaviour of a large number of atoms decaying spontaneously, we can assign a **probability** k of decay per unit time.

In this section, we use the same kind of book-keeping (keeping track of the number of radioactive atoms remaining) as in the population growth example, to arrive at a differential equation that describes the process. Once we have the equation, we determine its solution and make a long-term prediction about the amount of radioactivity remaining at a future time.

Deriving the model

We start by letting $N(t)$ be the number of radioactive atoms at time t. Generally, we would know $N(0)$, the number present initially. Our goal is to make simple assumptions about the process of decay that allows us to arrive at a mathematical model to predict values of $N(t)$ at any later time $t > 0$.

Assumptions.

(1) The process of radioactive decay is random, but on average, the probability of decay for a given radioactive atom is k per unit time where $k > 0$ is some constant.

(2) During each (small) time interval of length $\Delta t = h$, a radioactive atom has probability kh of decaying. This is merely a restatement of **(1)**.

Suppose that at some time t, there are $N(t)$ radioactive atoms. Then, according to our assumptions, during the time period $t \le t \le t+h$, on average $khN(t_0)$ atoms would decay. How many are there at time $t+h$? We can write the following balance-equation:

$$\begin{bmatrix} \text{Amount left} \\ \text{at time} \\ t+h \end{bmatrix} = \begin{bmatrix} \text{Amount present} \\ \text{at time} \\ t \end{bmatrix} - \begin{bmatrix} \text{Amount decayed} \\ \text{during time interval} \\ t \le t \le t+h \end{bmatrix}$$

or, restated:

$$N(t+h) = N(t) - khN(t). \tag{11.8}$$

Here we have assumed that h is a small time period. Rearranging Eqn. (11.8) leads to

$$\frac{N(t+h)-N(t)}{h} = -kN(t).$$

Considering the left hand side of this equation, we let h get smaller and smaller ($h \to 0$) and recall that

$$\lim_{h \to 0} \frac{N(t+h)-N(t)}{h} = \frac{dN}{dt} = N'(t)$$

Mastered Material Check

24. Suppose a given atom has a 1% chance of decay per 24 hours. What is this atom's probability of decay per week? Per hour?

where we have used the notation for a derivative of N with respect to t. We have thus shown that a description of the population of radioactive atoms reduces to

$$\frac{dN}{dt} = -kN. \tag{11.9}$$

We have, once more, arrived at a differential equation that provides a link between a function of time $N(t)$ and its own rate of change dN/dt. Indeed, this equation specifies that dN/dt is proportional to N, but with a negative constant of proportionality which implies decay.

Above we formulated the entire model in terms of the **number** of radioactive atoms. However, as shown below, the same equation holds regardless of the system of units used measure the amount of radioactivity

Example 11.7 *Define the number of moles of radioactive material by $y(t) = N(t)/A$ where A is* **Avogadro's number** *(the number of molecules in 1 mole: $\approx 6.022 \times 10^{23}$ - a dimensionless quantity, i.e. just a number with no associated units). Determine the differential equation satisfied by $y(t)$.*

Solution. We write $y(t) = N(t)/A$ in the form $N(t) = Ay(t)$ and substitute this expression for $N(t)$ in Eqn. (11.9). We use the fact that A is a constant to simplify the derivative. Then

$$\frac{dN}{dt} = -kN \quad \Rightarrow \quad \frac{Ady(t)}{dt} = -k(Ay(t)) \quad \Rightarrow \quad A\frac{dy(t)}{dt} = A(-ky(t))$$

cancelling the constant A from both sides of the equations leads to

$$\frac{dy(t)}{dt} = -ky(t), \quad \text{or simply} \quad \frac{dy}{dt} = -ky. \tag{11.10}$$

Thus $y(t)$ satisfies the same kind of differential equation (with the same negative proportionality constant) between the derivative and the original function. We will refer to (11.10) as the **decay equation**. ♢

Solution to the decay equation (11.10)

Suppose that initially, there was an amount y_0. Then, together, the differential equation and initial condition are

$$\frac{dy}{dt} = -ky, \quad y(0) = y_0. \tag{11.11}$$

We often refer to this pairing between a differential equation and an initial condition as an **initial value problem**. Next, we show that an exponential function is an appropriate solution to this problem

Example 11.8 (Checking a solution) *Show that the function*

$$y(t) = y_0 e^{-kt}. \tag{11.12}$$

is a solution to initial value problem (11.11).

Solution. Ee compute the derivative of the candidate function (11.12), and rearrange, obtaining

$$\frac{dy(t)}{dt} = \frac{d}{dt}[y_0 e^{-kt}] = y_0 \frac{de^{-kt}}{dt} = -ky_0 e^{-kt} = -ky(t).$$

This verifies that for the derivative of the function is $-k$ times the original function, so satisfies the DE in (11.11). We can also check that the initial condition is satisfied:

$$y(0) = y_0 e^{-k \cdot 0} = y_0 e^0 = y_0 \cdot 1 = y_0.$$

Hence, Eqn. (11.12) is the solution to the initial value problem for radioactive decay. For $k > 0$ a constant, this is a decreasing function of time that we refer to as **exponential decay**. ◊

The half life

Given a process of exponential decay, how long would it take for half of the original amount to remain? Let us recall that the "original amount" (at time $t = 0$) is y_0. Then we are looking for the time t such that $y_0/2$ remains. We must solve for t in

$$y(t) = \frac{y_0}{2}.$$

We refer to the value of t that satisfies this as the **half life**.

Example 11.9 (Half life) *Determine the half life in the exponential decay described by Eqn.* (11.12).

Solution. We compute:

$$\frac{y_0}{2} = y_0 e^{-kt} \quad \Rightarrow \quad \frac{1}{2} = e^{-kt}.$$

Now taking reciprocals:

$$2 = \frac{1}{e^{-kt}} = e^{kt}.$$

Thus we find the same result as in our calculation for doubling times, namely,

$$\ln(2) = \ln(e^{kt}) = kt,$$

so that the half life is

$$\tau = \frac{\ln(2)}{k}.$$

This is shown in Figure 11.7.

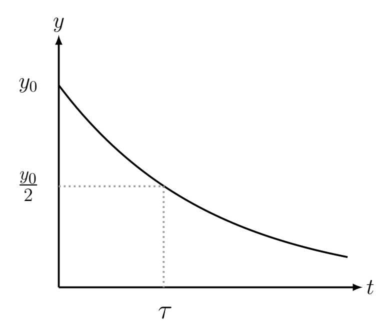


Figure 11.7: Half-life in an exponentially decreasing process.

Example 11.10 (Chernobyl: April 1986) *In 1986 the Chernobyl nuclear power plant exploded, and scattered radioactive material over Europe. The radioactive element iodine-131 (I^{131}) has half-life of 8 days whereas cesium-137 (Cs^{137}) has half life of 30 years. Use the model for radioactive decay to predict how much of this material would remain over time.*

Solution. We first determine the decay constants for each of these two elements, by noting that

$$k = \frac{\ln(2)}{\tau},$$

and recalling that $\ln(2) \approx 0.693$. Then for I^{131} we have

$$k = \frac{\ln(2)}{\tau} = \frac{\ln(2)}{8} = 0.0866 \text{ per day.}$$

Then the amount of I^{131} left at time t (in days) would be

$$y_I(t) = y_0 e^{-0.0866t}.$$

For Cs^{137}

$$k = \frac{\ln(2)}{30} = 0.023 \text{ per year.}$$

so that for T in years,

$$y_C(T) = y_0 e^{-0.023T}.$$

Note: we have used T rather than t to emphasize that units are different in the two calculations done in this example.

Example 11.11 (Decay to 0.1% **of the initial level)** *How long it would take for I^{131} to decay to 0.1 % of its initial level? Assume that the initial level occurred just after the explosion at Chernobyl.*

Solution. We must calculate the time t such that $y_I = 0.001 y_0$:

$$0.001 y_0 = y_0 e^{-0.0866t} \quad \Rightarrow \quad 0.001 = e^{-0.0866t} \quad \Rightarrow \quad \ln(0.001) = -0.0866t.$$

Therefore,

$$t = \frac{\ln(0.001)}{-0.0866} = \frac{-6.9}{-0.0866} = 79.7 \text{ days.}$$

Thus it would take about 80 days for the level of Iodine-131 to decay to 0.1% of its initial level. ◊

Mastered Material Check

25. Repeat the calculation in Example 11.11 for Cesium.
26. Convert the Cesium decay time units to days and repeat the calculation of Example 11.10 with the new time units.
27. If the decay rate of a substance is 10% per day, what is its half-life?

11.4 Deriving a differential equation for the growth of cell mass

Derivation of a differential equation that describes cell growth resulting from absorption and consumption of nutrients.

In Section 1.2, we asked how the size of a living cell influences the balance between the rates of nutrient absorption (called A) and consumption (denoted C). But what if the two processes do not balance? What happens to the cell if the rates are unequal?

If a cell absorb nutrients faster than nutrients are consumed ($A > C$), some of the excess nutrients accumulate, and this buildup of nutrient mass can be converted into cell mass. This can result in growth (increase of cell mass). Conversely, if the consumption rate exceeds the rate of absorption of nutrients ($C > A$), the cell has a shortage of metabolic “fuel”, and needs to convert some of its own mass into energy reserves that can power its metabolism - this would lead to loss in cell mass.

We can keep track of such changes in cell mass by using a simple "balance equation". The balance equation states that "the rate of change of cell mass is the difference between the rate of nutrient (mass) coming in (A) and the rate of nutrient (mass) being consumed (C), i.e.

$$\frac{dm}{dt} = A - C. \tag{11.13}$$

Each term in this equation must have the same units, mass of nutrient per unit time. A contributes positively to mass increase, whereas C is a rate of depletion that makes a negative contribution (hence the signs associated with terms in the equation). It also makes sense to adopt the assumptions previously made in Section 1.2 (and Featured Problem 9.1) that

$$A = k_1 S, \quad C = k_2 V, \quad m = \rho V,$$

where S, V, ρ are the surface area, volume, and density of the cell, and k_1, k_2, ρ are positive constants. Then Eqn. (11.13) becomes

Mastered Material Check What are the units of k_1, k_2, ρ?

$$\frac{dm}{dt} = A - C \quad \Rightarrow \quad \frac{d(\rho V)}{dt} = k_1 S - k_2 V. \tag{11.14}$$

The above equation is rather general, and does not depend on cell shape.

Now consider the special case of a spherical cell for which $V = (4/3)\pi r^3$, $S = 4\pi r^2$. This simplification will permit us to convert the balance equation into a differential equation that describes changes in cell radius over time. Now Eqn. (11.14) can be rewritten as

$$\frac{d[\rho \cdot (4/3)\pi r^3]}{dt} = k_1(4\pi r^2) - k_2(4/3)\pi r^3. \tag{11.15}$$

We can simplify the derivative on the right hand side using the chain rule, as done in Featured Problem 9.1, obtaining

$$\rho \frac{4\pi}{3}\pi(3r^2)\frac{dr}{dt} = k_1(4\pi r^2) - k_2(4/3)\pi r^3. \tag{11.16}$$

What does this tell us about cell radius?

One way to satisfy Eqn. (11.16) is to set $r = 0$ in each term. While this is a "solution" to the equation, it is not biologically interesting. (It merely describes a "cell" of zero radius that never changes.) Suppose $r \neq 0$. In that case, we can cancel out a factor of r^2 from both sides of the equation. (We can also cancel out 4π.) After some simplification, we arrive at

$$\rho \frac{dr}{dt} = k_1 - \frac{k_2}{3} r, \quad \Rightarrow \quad \frac{dr}{dt} = \frac{1}{\rho}\left(k_1 - \frac{k_2}{3} r\right).$$

With appropriate units and taking into account typical cell size and density, this equation might look something like

$$\frac{dr}{dt} = (1 - 0.1\, r). \tag{11.17}$$

Hint: If we use units of μm (=10^{-6}m) for cell radius, pg (=10^{-12}gm) for mass, and measure time in hours, then approximate values of the constants are
$\rho = 1$pg μm^{-3},
$k_1 = 1$pg μm^{-2} hr^{-1}, and
$k_2 = 0.3$pg μm^{-3} hr^{-1}.
In that case, the equation for cell radius is $dr/dt = (1 - 0.1 \cdot r)$.

From a statement about how cell mass changes, we have arrived at a resultant prediction about the rate of change of the cell radius. The equation we obtained is a differential equation that tells us something about a growing cell. In an upcoming chapter, we will build tools to be able to understand what this equation says, how to solve it for the cell radius $r(t)$ as a function of time t, and what such analysis predicts about the dynamics of cells with different initial sizes.

11.5 Summary

1. A differential equation is a statement linking the rate of change of some state variable with current values of that variable. An example is the simplest population growth model: if $N(t)$ is population size at time t:

$$\frac{dN}{dt} = kN.$$

2. A solution to a differential equation is a function that satisfies the equation. For instance, the function $N(t) = Ce^{kt}$ (for any constant C) is a solution to the unlimited population growth model (we check this by the appropriate differentiation). Graphs of such solutions (e.g. N versus t) are called solution curves.

3. To select a specific solution, more information (an initial condition) is needed. Given this information, e.g. $N(0) = N_0$, we can fully characterize the desired solution.

4. The **decay equation** is one representative of the same class of problems, and has an exponentially decaying solution.

$$\frac{dy}{dt} = -ky, \quad y(0) = y_0 \quad \Rightarrow \quad \text{Solution: } y(t) = y_0 e^{-kt}. \qquad (11.18)$$

5. So far, we have seen simple differential equations with simple (exponential) functions for their solutions. In general, it may be quite challenging to make the connection between the differential equation (stemming from some application or model) with the solution (which we want in order to understand and predict the behaviour of the system.)

In this chapter, we saw examples in which a natural phenomenon (population growth, radioactive decay, cell growth) motivated a mathematical model that led to a differential equation. In both cases, that equation was derived by making a statement that tracked the amount or number or mass of a system over time. Numerous simplifications were made to derive each differential equation. For example, we assumed that the birth and mortality rates stay fixed even as the population grows to huge sizes.

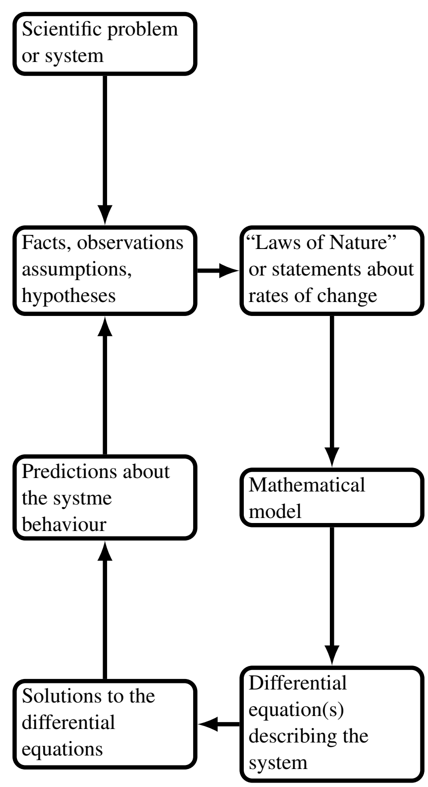


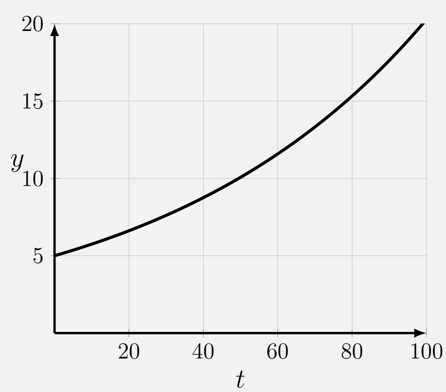
Figure 11.8: A "flow chart" showing how differential equations originate from scientific problems.

With regard to a larger context.

- Our purpose was to illustrate how a simple model is created, and what such models can predict.
- In general, differential equation models are often based on physical laws ("$F = ma$") or conservation statements ("rate in minus rate out equals net rate of change", or "total energy = constant").
- In biology, where the laws governing biochemical events are less formal, the models are often based on some mix of speculation and reasonable assumptions.
- In Figure 11.8 we illustrate how the scientific method leads to a cycle between the mathematical models and their test and validation using observations about the natural world.

Quick Concept Checks

1. Identify each of the following with either exponential growth or exponential decay:
   (a) $y = 20e^{3t}$;
   (b) $y = 5e^{-3t}$;
   (c) $\frac{dy}{dt} = 3t$;
   (d) $\frac{dy}{dx} = -5x$.
2. Determine the doubling time of the exponential growth function $N(t) = 500e^{2t}$.
3. Determine the half life of the of the exponential decay function $N(t) = 500e^{-2t}$.
4. Consider the following figure depicting exponential growth:



   What is the doubling time of this function?

Exercises

11.1. **Checking solutions of differential equations.** A differential equation is an equation in which some function is related to its own derivative(s).

For each of the following functions, calculate the appropriate derivative, and show that the function satisfies the indicated *differential equation*

(a) $f(x) = 2e^{-3x}, \quad f'(x) = -3f(x)$

(b) $f(t) = Ce^{kt}, \quad f'(t) = kf(t)$

(c) $f(t) = 1 - e^{-t}, \quad f'(t) = 1 - f(t)$

11.2. **Linear differential equations.** Consider the function $y = f(t) = Ce^{kt}$ where C and k are constants. For what value(s) of these constants does this function satisfy the equation

(a) $\frac{dy}{dt} = -5y$,

(b) $\frac{dy}{dt} = 3y$.

Note: an equation which involves a function and its derivative is called a differential equation.

11.3. **Checking initial value solution to a differential equation.** Check that the function (11.6) satisfies the differential equation (11.2) and the initial condition $N(0) = N_0$.

11.4. **Solving linear differential equations.** Find a function that satisfies each of the following *differential equations*.

Note: all your answers should be exponential functions, but they may have different dependent and independent variables.

(a) $\dfrac{dy}{dt} = -y$,

(b) $\dfrac{dc}{dx} = -0.1c$ and $c(0) = 20$,

(c) $\dfrac{dz}{dt} = 3z$ and $z(0) = 5$.

11.5. **Andromeda strain, revisited.** In Chapter 10 we discussed the growth of bacteria, starting from a single cell. The doubling time of the bacteria was given as 20 min.

Find the appropriate differential equation that describes this growth, the appropriate initial condition, and the exponential function (with base e) that is the solution to that differential equation. Use units of hours for time t.

11.6. **Population growth in developed and developing countries.** In Canada, women have only about 2 children during their 40 years of fertility, and people live to age 80. In underdeveloped countries, people on average live to age 60 and women have a child roughly every 4 years between ages 13 and 45.

Compare the per capita birth and mortality rates and the predicted population growth or decay in each of these scenarios, using arguments analogous to those of Section 11.2.

Find the growth rate k in percent per year and the doubling time for the growing population.

11.7. **Population growth and doubling.** A population of animals has a per-capita birth rate of $b = 0.08$ per year and a per-capita death rate of $m = 0.01$ per year. The population density, $P(t)$ is found to satisfy the differential equation

$$\frac{dP(t)}{dt} = bP(t) - mP(t)$$

(a) If the population is initially $P(0) = 1000$, find how big the population is in 5 years.

(b) When does the population double?

11.8. **Rodent population.** The per capita birthrate of one species of rodent is 0.05 newborns per day. This means that, on average, each member of the population results in 5 newborn rodents every 100 days. Suppose that over the period of 1000 days there are no deaths, and that the initial population of rodents is 250.

(a) Write a differential equation for the population size $N(t)$ at time t (in days).

(b) Write down the initial condition that N satisfies.

(c) Find the solution, i.e. express N as some function of time t that satisfies your differential equation and initial condition.

(d) How many rodents are there after 1 year ?

11.9. **Growth and extinction of microorganisms.**

(a) The population $y(t)$ of a certain microorganism grows continuously and follows an exponential behaviour over time. Its doubling time is found to be 0.27 hours. What differential equation would you use to describe its growth ?

Note: you must find the value of the rate constant, k, using the doubling time.

(b) With exposure to ultra-violet radiation, the population ceases to grow, and the microorganisms continuously die off. It is found that the half-life is then 0.1 hours. What differential equation would now describe the population?

11.10. **A bacterial population.** A bacterial population grows at a rate proportional to the population size at time t. Let $y(t)$ be the population size at time t. By experiment it is determined that the population at $t = 10$ min is $15,000$ and at $t = 30$ min it is $20,000$.

(a) What was the initial population?

(b) What is the population at time $t = 60$min?

11.11. **Antibiotic treatment.** A colony of bacteria is treated with a mild antibiotic agent so that the bacteria start to die. It is observed that the density of bacteria as a function of time follows the approximate relationship $b(t) = 85e^{-0.5t}$ where t is time in hours.

Determine the time it takes for half of the bacteria to disappear; this is called the *half-life*.

Find how long it takes for 99% of the bacteria to die.

11.12. **Two populations.** Two populations are studied. Population **1** is found to obey the differential equation

$$dy_1/dt = 0.2y_1$$

and population **2** obeys

$$dy_2/dt = -0.3y_2$$

where t is time in years.

(a) Which population is growing and which is declining?

(b) Find the doubling time (respectively half-life) associated with the given population.

(c) If the initial levels of the two populations were $y_1(0) = 100$ and $y_2(0) = 10,000$, how big would each population be at time t?

(d) At what time would the two populations be exactly equal?

11.13. **The human population.** The human population on Earth doubles roughly every 50 years. In October 2000 there were 6.1 billion humans on earth.

(a) Determine what the human population would be 500 years later under the uncontrolled growth scenario.

(b) How many people would have to inhabit each square kilometer of the planet for this population to fit on earth? (Take the circumference of the earth to be $40,000$ km for the purpose of computing its surface area and assume that the oceans have dried up.)

11.14. **Fish in two lakes.** Two lakes have populations of fish, but the conditions are quite different in these lakes. In the first lake, the fish population is growing and satisfies the differential equation

$$\frac{dy}{dt} = 0.2y$$

where t is time in years. At time $t = 0$ there were 500 fish in this lake. In the second lake, the population is dying due to pollution. Its population satisfies the differential equation

$$\frac{dy}{dt} = -0.1y,$$

and initially there were 4000 fish in this lake.

At what time are the fish populations in the two lakes identical?

11.15. **First order chemical kinetics.** When chemists say that a chemical reaction follows "first order kinetics", they mean that the concentration of the reactant at time t, i.e. $c(t)$, satisfies an equation of the form $\frac{dc}{dt} = -rc$ where r is a rate constant, here assumed to be positive. Suppose the reaction mixture initially has concentration 1M ("1 molar") and that after 1 hour there is half this amount.

(a) Find the "half life" of the reactant.

(b) Find the value of the rate constant r.

(c) Determine how much is left after 2 hours.

(d) When is only 10% of the initial amount be left?

11.16. **Chemical breakdown.** In a chemical reaction, a substance S is broken down. The concentration of the substance is observed to change at a rate proportional to the current concentration. It was observed that 1 Mole/liter of S decreased to 0.5 Moles/liter in 10 minutes.

(a) How long does it take until only 0.25 Moles per liter remain?

(b) How long does it take until only 1% of the original concentration remains?

11.17. **Half-life.** If 10% of a radioactive substance remains after one year, find its half-life.

11.18. **Carbon 14.** Carbon 14, or ^{14}C, has a half-life of 5730 years. This means that after 5730 years, a sample of Carbon 14, which is a radioactive isotope of carbon, has lost one half of its original radioactivity.

(a) Estimate how long it takes for the sample to fall to roughly 0.001 of its original level of radioactivity.

(b) Each gram of ^{14}C has an activity given here in units of 12 decays per minute. After some time, the amount of radioactivity decreases. For example, a sample 5730 years old has only one half the original activity level, i.e. 6 decays per minute. If a 1 gm sample of material is found to have 45 decays per hour, approximately how old is it?

Note: ^{14}C is used in radiocarbon dating, a process by which the age of materials containing carbon can be estimated. W. Libby received the Nobel prize in chemistry in 1960 for developing this technique.

11.19. **Strontium-90.** Strontium-90 is a radioactive isotope with a half-life of 29 years. If you begin with a sample of 800 units, how long does it take for the amount of radioactivity of the strontium sample to be reduced to

(a) 400 units

(b) 200 units

(c) 1 unit

11.20. **More radioactivity.** The half-life of a radioactive material is 1620 years.

(a) What percentage of the radioactivity remains after 500 years?

(b) Cobalt 60 is a radioactive substance with half life 5.3 years. It is used in medical application (radiology). How long does it take for 80% of a sample of this substance to decay?

11.21. **Salt in a barrel.** A barrel initially contains 2 kg of salt dissolved in 20 L of water. If water flows in the rate of 0.4 L per minute and the well-mixed salt water solution flows out at the same rate, how much salt is present after 8 minutes?

11.22. **Atmospheric pressure.** Assume the atmospheric pressure y at a height x meters above the sea level satisfies the relation

$$\frac{dy}{dx} = kx.$$

If one day at a certain location the atmospheric pressures are 760 and 675 torr (unit for pressure) at sea level and at 1000 meters above sea level, respectively, find the value of the atmospheric pressure at 600 meters above sea level.

12
Solving differential equations

In Chapter 11, we introduced differential equations to keep track of continuous changes in the growth of a population or the decay of radioactivity. We encountered a differential equation that tracks changes in cell mass due to nutrient absorption and consumption. Finally, we learned that the solutions to a differential equation is a function. In applications studied, that function can be interpreted as predictions of the behaviour of the system or process over time.

In this chapter, we further develop some of these ideas. We explore several techniques for finding and verifying that a given function is a solution to a differential equation. We then examine a simple class of differential equations that have many applications to processes of production and decay, and find their solutions. Finally, we show how an approximation method provides for numerical solutions of such problems.

12.1 Verifying that a function is a solution

Section 12.1 Learning goals

1. Given a function, check whether that function does or does not satisfy a given differential equation.
2. Verify whether a given function does or does not satisfy an initial condition.

In this section we concentrate on analytic solutions to a differential equation. By **analytic solution**, we mean a "formula" such as $y = f(x)$ that satisfies the given differential equation. We saw in Chapter 11 that we can check whether a function satisfies a differential equation (e.g., Example 11.8) by simple differentiation. In this section, we further demonstrate this process.

Example 12.1 *Show that the function $y(t) = (2t+1)^{1/2}$ is a solution to the* *differential equation and initial condition*

$$\frac{dy}{dt} = \frac{1}{y}, \quad y(0) = 1.$$

Solution. First, we check the derivative, obtaining

$$\begin{aligned}\frac{dy(t)}{dt} = \frac{d(2t+1)^{1/2}}{dt} &= \frac{1}{2}(2t+1)^{-1/2} \cdot 2 \\ &= (2t+1)^{-1/2} = \frac{1}{(2t+1)^{1/2}} = \frac{1}{y}.\end{aligned}$$

Hence, the function satisfies the differential equation. We must also verify the initial condition. We find that $y(0) = (2 \cdot 0 + 1)^{1/2} = 1^{1/2} = 1$. Thus the initial condition is also satisfied, and $y(t)$ is indeed a solution. ◊

Example 12.2 *Consider the differential equation and initial condition*

$$\frac{dy}{dt} = 1 - y, \quad y(0) = y_0. \tag{12.1}$$

a) *Show that the function $y(t) = y_0 e^{-t}$ is* **not** *a solution to this differential equation.*

b) *Show that the function $y(t) = 1 - (1 - y_0)e^{-t}$* ***is*** *a solution.*

Solution.

a) To check whether $y(t) = y_0 e^{-t}$ is a solution to the differential equation (12.1), we substitute the function into each side ("left hand side", LHS; "right hand side". RHS) of the equation. We show the results in the columns of Table 12.1. After some steps in the simplification, we see that the two sides do not match, and conclude that the function is not a solution, as it fails to satisfy the equation

b) Similarly, we check the second function. The calculations

are shown in columns of Table 12.2. We find that RHS=LHS, so the differential equation is satisfied. Finally, let us show that the initial condition $y(0) = y_0$ is also satisfied. Plugging in $t = 0$ we have

$$y(0) = 1 - (1 - y_0)e^0 = 1 - (1 - y_0) \cdot 1 = 1 - (1 - y_0) = y_0.$$

Thus, both differential equation and initial condition are satisfied.

◊

LHS	RHS
$\frac{dy}{dt}$	$1 - y$
$\frac{d[y_0 e^{-t}]}{dt}$	$1 - y_0 e^{-t}$
$-y_0 e^{-t}$	☹

Table 12.1: The function $y(t) = y_0 e^{-t}$ **is not** a solution to the differential equation (12.1). Plugging the function into each side of the DE and simplifying (down the rows) leads to expressions that do not match.

LHS	RHS
$\frac{dy}{dt}$	$1 - y$
$\frac{d}{dt}[1 - (1 - y_0)e^{-t}]$	$1 - [1 - (1 - y_0)e^{-t}]$
$-(1 - y_0)\frac{de^{-t}}{dt}$	$(1 - y_0)e^{-t}$
$(1 - y_0)e^{-t}$	✓

Table 12.2: (b) The function $y(t) = 1 - (1 - y_0)e^{-t}$ is a solution to the differential equation (12.1). The expressions we get by evaluating each side of the differential equation do match.

Example 12.3 (Height of water draining out of a cylindrical container) *A cylindrical container with cross-sectional area A has a small hole of area a at its base, through which water leaks out. It can be shown that height of water $h(t)$ in the container satisfies the differential equation*

$$\frac{dh}{dt} = -k\sqrt{h}, \tag{12.2}$$

(where k is a constant that depends on the size and shape of the cylinder and its hole: $k = \frac{a}{A}\sqrt{2g} > 0$ *and g is acceleration due to gravity.) Show that the function*

$$h(t) = \left(\sqrt{h_0} - k\frac{t}{2}\right)^2 \tag{12.3}$$

is a solution to the differential equation (12.2) and initial condition $h(0) = h_0$.

Solution. We first easily verify that the initial condition is satisfied. Substitute $t = 0$ into the function (12.3). Then we find $h(0) = h_0$, verifying the initial conditions.

To show that the differential equation (12.2) is satisfied, we differentiate the function in Eqn. (12.3):

$$\begin{aligned} \frac{dh(t)}{dt} = \frac{d}{dt}\left(\sqrt{h_0} - k\frac{t}{2}\right)^2 &= 2\left(\sqrt{h_0} - k\frac{t}{2}\right) \cdot \left(\frac{-k}{2}\right) \\ &= -k\left(\sqrt{h_0} - k\frac{t}{2}\right) = -k\sqrt{h(t)}. \end{aligned}$$

Here we have used the power law and the chain rule, remembering that h_0, k are constants. Now we notice that, using Eqn. (12.3), the expression for $\sqrt{h(t)}$ exactly matches what we have computed for dh/dt. Thus, we have shown that the function in Eqn. (12.3) satisfies both the initial condition and the differential equation. ◊

Mastered Material Check

1. Draw a diagram of the system described in Example 12.3.
2. What set of units would be reasonable for each of the parameters in Example 12.3.
3. Create a table to organize the calculations for this example, similar to Tables 12.1 and 12.2.

As shown in Examples 12.1- 12.3, if we are told that a function is a solution to a differential equation, we can check the assertion and verify that it is correct or incorrect. A much more difficult task is to find the solution of a new differential equation from first principles.

In some cases, **integration**, learned in second semester calculus, can be used. In others, some transformation that changes the problem to a more familiar one is helpful - an example of this type is presented in Section 12.2. In many cases, particularly those of so-called non-linear differential equations, great expertise and familiarity with advanced mathematical methods are required to find the solution to such problems in an analytic form, i.e. as an explicit formula. In such cases, approximation and numerical methods are helpful.

12.2 *Equations of the form $y'(t) = a - by$*

Section 12.2 Learning goals

1. Define steady states of a differential equation, and be able to find such special solutions.
2. Starting with the differential equation (12.4) $y' = a - by$, find a new differential equation for the deviation away from a steady state, $z(t)$ and show that it is a simple decay equation.
3. Use the transformed (decay) equation to find the solution for $z(t)$ and, for $y(t)$ in the original equation, (12.4).
4. Explain Newton's Law of Cooling (NLC), and the differential equation of the same type, $y' = a - by$. Find its solution and explain what this solution means.
5. Use the solution to NLC to predict the temperature of a cooling or heating object over time.
6. Describe a variety of related examples, and use the same methods to solve and interpret these (examples include chemical production and decay, the velocity of a skydiver, the concentration of drug in the blood, and others).

In this section we introduce an important class of differential equations that have many applications in physics, chemistry, biology, and other applications. All share a similar structure, namely all are of the form

$$\frac{dy}{dt} = a - by, \quad y(0) = y_0. \tag{12.4}$$

First, we show how a solution to such equation can be found. Then, we examine a number of applications.

Special solutions: steady states

We first ask about "special solutions" to the differential equation (12.4) in which there is no change over time. That is, we ask whether there are values of y for which $dy/dt = 0$.

From (12.4), we find that such solutions would satisfy

$$\frac{dy}{dt} = 0 \quad \Rightarrow \quad a - by = 0 \quad \Rightarrow \quad y = \frac{a}{b}.$$

In other words, if we were to start with the initial value $y(0) = a/b$, then that value would not change, since it satisfies $dy/dt = 0$, so that the solution at all future times would be $y(t) = a/b$. (Of course, this is a perfectly good function; it is simply a function that is alway constant.)

We refer to such constant solutions as **Steady States**.

An explanation of the way we find solutions to equations of the form $\frac{dy}{dt} = a - by$, with $y(0) = y_0$.

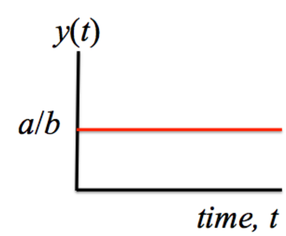


Figure 12.1: $y = a/b$ is a constant solution to the differential equation in (12.4). We call this type of solution a **steady state**.

Other solutions: away from steady state

What happens if we start with a value of y that is not exactly at the "special" steady state? Let us rewrite the DE in a more suggestive form,

$$\frac{dy}{dt} = a - by \quad \Rightarrow \quad \frac{dy}{dt} = -b\left(y - \frac{a}{b}\right),$$

(having factored out $-b$). The advantage is that we recognize the expression $(y - \frac{a}{b})$ as the difference, or **deviation** of y away from its steady state value. (That deviation could be either positive or negative, depending on whether y is larger or smaller than a/b.) We ask whether this deviation gets larger or smaller as time goes by, i.e., whether y gets further away or closer to its steady state value a/b.

Define $z(t)$ as that deviation, that is

$$z(t) = y(t) - \frac{a}{b},$$

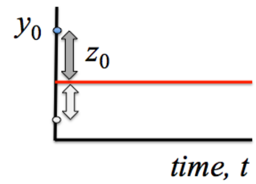


Figure 12.2: We define $z(t)$ as the deviation of y from its steady state value. Here we show two typical initial values of z, where $z_0 = y_0 - \frac{a}{b}$.

Then, since a, b are constants, we recognize that

$$\frac{dz}{dt} = \frac{dy}{dt}.$$

Second, the initial value of z follows simply from the initial value of y:

$$z(0) = y(0) - \frac{a}{b} = y_0 - \frac{a}{b}.$$

Now we can **transform** the equation (12.4) into a new differential equation for the variable z by using these two facts. We can replace the y derivative by the z derivative, and also, using Eqn. (12.4), find that

$$\frac{dz}{dt} = \frac{dy}{dt} = -b\left(y - \frac{a}{b}\right) = -bz.$$

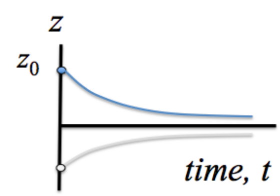


Figure 12.3: The deviation away from steady state (blue, grey curves) is $z(t) = y(t) - a/b$. We can solve the differential equation for $z(t)$ because it is a simple exponential decay equation. Here we show two typical solutions for z.

Hence, we have transformed the original DE and IC into the new problem

$$\frac{dz}{dt} = -bz, \quad z(0) = z_0, \qquad \left[\text{where } z_0 = y_0 - \frac{a}{b}\right].$$

But this is the familiar decay initial value problem that we have already solved before. So

$$z(t) = z_0 e^{-bt}.$$

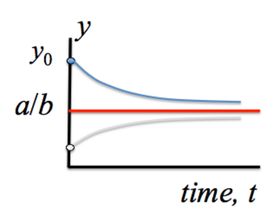


Figure 12.4: Finally, we can determine the solution $y(t)$.

We have arrived at the conclusion that the deviation from steady state **decays exponentially** with time, provided that $b > 0$. Hence, we already know that y should get closer to the constant value a/b as time goes by!

We can do even better than this, by transforming the solution we found for $z(t)$ into an expression for $y(t)$. To do so, use the definition once more, setting

$$z(t) = z_0 e^{-bt} \quad \Rightarrow \quad y(t) - \frac{a}{b} = \left(y_0 - \frac{a}{b}\right) e^{-bt}.$$

Solving for $y(t)$ then leads to

$$y(t) = \frac{a}{b} + \left(y_0 - \frac{a}{b}\right) e^{-bt}. \qquad (12.5)$$

Adjust the sliders to see how the parameters a and b and the initial value y_0 affect the shape of the function $y(t)$ in the formula (12.5).

Example 12.4 ($a = b = 1$) *Suppose we are given the differential equation and initial condition*

$$\frac{dy}{dt} = 1 - y, \quad y(0) = y_0. \tag{12.6}$$

Determine the solution to this differential equation.

Solution.

By substituting $a = 1, b = 1$ in the solution found above, we observe that

$$y(t) = 1 - (1 - y_0)e^{-t}.$$

Representative curves in this **family of solutions** are shown in Figure 12.5 for various initial values y_0. ♢

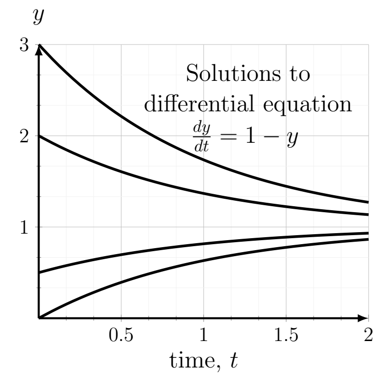


Figure 12.5: Solutions to Eqn. (12.6) are functions that approach $y = 1$.

Mastered Material Check

4. Find the steady state of Eqn. (12.6).
5. From Figure 12.5, determine what were the four different initial conditions used.
6. Rewrite these four initial conditions as the initial deviations away from steady state, that is, give the initial values, z_0 of the deviation.

We now apply the methods to a number of examples.

Featured Problem 12.1 (Predicting the size of a growing cell) *Find a solution to the differential equation (11.17) for the radius of a growing cell $r(t)$ (in units of $\mu m = 10^{-6}m$) as a function of time t (in hours), that is find $r(t)$ assuming that at time $t = 0$ the cell is $2\mu m$ in radius.*

By solving the above problem, we get a detailed prediction of cell growth based on assumed rates of nutrient intake and consumption.

Newton's law of cooling

Consider an object at temperature $T(t)$ in an environment whose ambient temperature is E. Depending on whether the object is cooler or warmer than the environment, it heats up or cools down. From common experience we know that, after a long time, the temperature of the object equilibrates with its environment.

Isaac Newton formulated a hypothesis to describe the rate of change of temperature of an object. He assumed that

> The rate of change of temperature T of an object is proportional to the difference between its temperature and the ambient temperature, E.

To rephrase this statement mathematically, we write

$$\frac{dT}{dt} \text{ is proportional to } (T(t) - E).$$

This implies that the derivative dT/dt is some constant multiple of the term $(T(t) - E)$. However, the sign of that constant requires some discussion. Denote the constant of proportionality by α temporarily, and suppose $\alpha \geq 0$. Let us check whether the differential equation

$$\frac{dT}{dt} = \alpha(T(t) - E),$$

makes physical sense.

Mastered Material Check

7. What can we say about the units of T and E?

Suppose the object is warmer than its environment ($T(t) > E$). Then $T(t) - E) > 0$ and $\alpha \geq 0$ implies that $dT/dt > 0$ which says that the temperature of the object should get *warmer*! But this does not agree with our everyday experience: a hot cup of coffee cools off in a chilly room. Hence $\alpha \geq 0$ cannot be correct. Based on this, we conclude that Newton's Law of Cooling, written in the form of a differential equation, should read:

$$\frac{dT}{dt} = k(E - T(t)), \quad \text{where} \quad k > 0. \tag{12.7}$$

Note: the sign of the term in braces has been switched.

Typically, given the temperature at some initial time $T(0) = T_0$, we want to predict $T(t)$ for later time.

Example 12.5 *Consider the temperature $T(t)$ as a function of time. Solve the differential equation for Newton's law of cooling*

$$\frac{dT}{dt} = k(E - T),$$

together with the initial condition $T(0) = T_0$.

Solution. As before, we transform the variable to reduce the differential equation to one that we know how to solve. This time, we select the new variable to be $z(t) = E - T(t)$. Then, by steps similar to previous examples, we find that

$$\frac{dz(t)}{dt} = -kz.$$

We also rewrite the initial condition in terms of z, leading to $z(0) = E - T(0) = E - T_0$. After carrying out **Steps 1-3** as before, we find the solution for $T(t)$,

$$T(t) = E + (T_0 - E)e^{-kt}. \tag{12.8}$$

In Figure 12.6 we show a family of curves of the form of Eqn. (12.8) for five different initial temperature values (we have set $E = 10$ and $k = 0.2$ for all these curves). ♢

Next, we interpret the behaviour of these solutions.

Mastered Material Check

8. Fill in the details for Example 12.5.
9. In Figure 12.6, what are the five different initial temperatures, T_0 corresponding to each solution curve?
10. In Figure 12.6, how many curves represent a heating object and how many a cooling object?

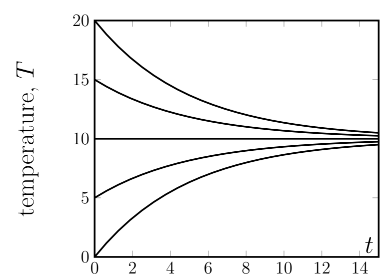


Figure 12.6: Temperature versus time, $T(t)$, for a cooling object.

Example 12.6 *Explain (in words) what the form of the solution in Eqn. (12.8) of Newton's law of cooling implies about the temperature of an object as it warms or cools.*

Solution. We make the following remarks

- It is straightforward to verify that the initial temperature is $T(0) = T_0$ (substitute $t = 0$ into the solution of Eqn. (12.8)). Now examine the time dependence. Only one term, e^{-kt} depends on time. Since $k > 0$, this is an exponentially decaying function, whose magnitude shrinks with time. The whole term that it multiplies, $(T_0 - E)e^{-kt}$, continually shrinks. Hence,

$$T(t) = E + \quad (T_0 - E)e^{-kt} \quad \Rightarrow \quad \text{as } t \to \infty, \quad e^{-kt} \to 0,$$
$$\text{so } T(t) \to E.$$

Thus the temperature of the object always approaches the ambient temperature. This is evident in the solution curves shown in Figure 12.6.

- We also observe that the direction of approach (decreasing or increasing) depends on the sign of the constant $(T_0 - E)$. If $T_0 > E$, the temperature approaches E from above, whereas if $T_0 < E$, the temperature approaches E from below.

- In the specific case that $T_0 = E$, there is no change at all. $T = E$ satisfies $dT/dt = 0$, and corresponds to a **steady state** of the differential equation, as previously defined.

  Steady states are studied in more detail in Chapter 13.

Mastered Material Check

11. Consider three cups of coffee left in a 20°C room. If one is iced, another is piping hot, and the third is room temperature, which cup will not change temperature? Which, thus, represents a steady state?
12. Convert the temperatures in Example 12.7 to Fahrenheit and repeat.

Using Newton's law of cooling to solve a mystery

Now that we have a detailed solution to the differential equation representing Newton's law of cooling, we can apply it to making exact determinations of temperature over time, or of time at which a certain temperature was attained. The following example illustrates an application of this idea.

Example 12.7 (Murder mystery) *It is a dark clear night. The air temperature is* 10° *C. A body is discovered at midnight. Its temperature is* 27° *C. One hour later, the body has cooled to* 24° *C. Use Newton's law of cooling to determine the time of death.*

Details of the calculations for Example 12.7.

Solution. We assume that body-temperature just before death was 37° C (normal human body temperature). Let $t = 0$ be the time of death. Then the initial temperature is $T(0) = T_0 = 37°$ C. We want to find the time elapsed until the body was found, i.e. time t at which the temperature of the body had cooled down to 27° C. We assume that the ambient temperature, $E = 10$, was constant. From Newton's law of cooling, the body temperature satisfies

$$\frac{dT}{dt} = k(10 - T).$$

From previous work and Eqn. (12.8), the solution to this DE is

$$T(t) = 10 + (37 - 10)e^{-kt}.$$

We do not know the value of the constant k, but we have enough information to find it. First, at discovery, the body's temperature was 27°. Hence at time t

$$27 = 10 + 27e^{-kt} \quad \Rightarrow \quad 17 = 27e^{-kt}.$$

Also at $t + 1$ (one hour after discovery), the temperature was 24° C, so

$$T(t+1) = 10 + (37 - 10)e^{-k(t+1)} = 24, \quad \Rightarrow \quad 24 = 10 + 27e^{-k(t+1)}.$$

Thus,

$$14 = 27e^{-k(t+1)}.$$

We have two equations for the two unknowns t and k. To solve for k, take a ratio of the sides of the equations. Then

$$\frac{14}{17} = \frac{27e^{-k(t+1)}}{27e^{-kt}} = e^{-k} \quad \Rightarrow \quad -k = \ln\left(\frac{14}{17}\right) = -0.194.$$

This is the constant that describes the rate of cooling of the body.

To find the time of death, t, use

$$17 = 27e^{-kt} \quad \Rightarrow \quad -kt = \ln\left(\frac{17}{27}\right) = -0.4626$$

finally, solving for t, we get

$$t = \frac{0.4626}{k} = \frac{0.4626}{0.194} = 2.384 \text{ hours.}$$

◊

Mastered Material Check

13. Give the concluding sentence for Example 12.7. Be sure to include an actual time of death, given that the body was discovered at midnight.
14. Use a plotting program to graph $T(t)$ for Example 12.7.
15. Use your plot to estimate how long it took for the body to cool off to 33°C.

Related applications and further examples

Having gained familiarity with specific examples, we now return to the general case and summarize the results.

The differential equation and initial condition

$$\frac{dy}{dt} = a - by, \quad y(0) = y_0 \tag{12.9}$$

has the solution

$$y(t) = \frac{a}{b} - \left(\frac{a}{b} - y_0\right) e^{-bt}. \tag{12.10}$$

Suppose that $a, b > 0$ in Eqn. (12.9). Then we can summarize the behaviour of the solutions (12.10) as follows:

- The time dependence of Eqn. (12.10) is contained in the term e^{-bt}, which (for $b > 0$) is exponentially decreasing. As time increases, $t \to \infty$, the exponential term becomes negligibly small, so $y \to a/b$.
- If initially $y(0) = y_0 > a/b$, then $y(t)$ approaches a/b from above, whereas if $y_0 < a/b$, it approaches a/b from below.
- If initially $y_0 = a/b$, there is no change at all ($dy/dt = 0$). Thus $y = a/b$ is a **steady state** of the DE in Eqn. (12.9).

Recognizing such general structure means that we can avoid repeating similar calculations from scratch in related examples. Newton's law of cooling is one representative of the class of differential equations of the form Eqn. (12.9). If we set $a = kE, b = k$ and $T = y$ in Eqn. (12.9), we get back to

Eqn. (12.7). As expected from the general case, T approaches $a/b = E$, the ambient temperature, which corresponds to a steady state of NLC.

Next, we describe other examples that share this structure, and hence similar dynamic behaviour.

Friction and terminal velocity A falling object accelerates under the force of gravity, but friction slows down this acceleration. The differential equation satisfied by the velocity $v(t)$ of the falling object with friction is

$$\frac{dv}{dt} = g - kv \tag{12.11}$$

where $g > 0$ is acceleration due to gravity and $k > 0$ is a constant representing the effect of air resistance. (In contrast to the "upwards pointing" coordinate system used in Example 4.10, here we focus on how the magnitude of the velocity changes with time.) Usually, a frictional force is assumed to be proportional to the velocity of the object, and to act in a direction that slows it down. (This accounts for the negative sign in Eqn. (12.11).) Parachutes operate on the principle of enhancing that frictional force to damp out the acceleration of a skydiver. Hence, Eqn. (12.11) is often called the **skydiver equation**.

Note. Eqn. (12.11) comes from a simple force balance:

$$ma = F_{gravity} - F_{drag},$$

and from the assumption that $F_{drag} = \mu v$, where $\mu > 0$ is the "drag coefficient". Dividing both sides by m and replacing a by dv/dt leads to this equation, with $k = \mu/m$.

Example 12.8 *Use the general results for Eqn.* (12.9) *to write down the solution to the differential equation* (12.11) *for the velocity of a skydiver given the initial condition* $v(0) = v_0$*. Interpret your results in a simple description of what happens over time.*

Solution. Eqn. (12.11) is of the same form as Eqn. (12.9), and has the same type of solutions. We merely have to adjust the notation, by identifying

$$v(t) \to y(t), \quad g \to a, \quad k \to b, \quad v_0 \to y_0.$$

Hence, without further calculation, we can conclude that the solution of (12.11) together with its initial condition is:

$$v(t) = \frac{g}{k} - \left(\frac{g}{k} - v_0\right) e^{-kt}. \tag{12.12}$$

The velocity is initially v_0, and eventually approaches g/k which is the **steady state** or **terminal velocity** for the object. Depending on the initial speed, the object either slows down (if $v_0 > g/k$) or speed up (if $v_0 < g/k$) as it approaches the terminal velocity. ◊

Mastered Material Check

16. Assign appropriate units to each of the parameters in Example 12.8.
17. When a sky-diver steps into the void, her initial vertical velocity is zero. Write down her velocity $v(t)$ based on results of Example 12.8 .

Chemical production and decay. A chemical reaction inside a fixed reaction volume produces a substance at a constant rate K_{in}. A second reaction results in decay of that substance at a rate proportional to its concentration. Let $c(t)$ denote the time-dependent concentration of the substance, and assume that time is measured in units of hours. Then, writing down a balance equation leads to a differential equation of the form

$$\frac{dc}{dt} = K_{\text{in}} - \gamma c. \tag{12.13}$$

Here, the first term is the rate of production and the second term is the rate of decay. The net rate of change of the chemical concentration is then the difference of the two. The constants $K_{\text{in}} > 0, \gamma > 0$ represent the rate of production and decay - recall that the *units of each term in any equation have to match.*For example, if the concentration c is measured in units of milli-Molar (mM), then dc/dt has units of mM/hr, and hence K_{in} must have units of mM/h and γ must have units of 1/hr.

Example 12.9 *Write down the solution to the DE (12.13) given the initial condition $c(0) = c_0$. Determine the steady state chemical concentration.*

Solution. Translating notation from the general case to this example,

$$c(t) \to y(t), \qquad K_{\text{in}} \to a, \qquad \gamma \to b.$$

Then we can immediately write down the solution:

$$c(t) = \frac{K_{\text{in}}}{\gamma} - \left(\frac{K_{\text{in}}}{\gamma} - c_0\right) e^{-\gamma t}. \tag{12.14}$$

Regardless of its initial condition, the chemical concentration will approach a steady state concentration is $c = K_{\text{in}}/\gamma$. ◊

In this section we have seen that the behaviour found in the general case of the differential equation (12.4), can be reinterpreted in each specific situation of interest. This points to one powerful aspect of mathematics, namely the ability to use results in abstract general cases to solve a variety of seemingly unrelated scientific problems that share the same mathematical structure.

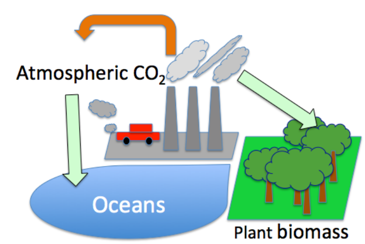


Figure 12.7: CO_2 is produced by emissions from burning fossil fuel and other human activities (orange arrow). The oceans and plant biomass are both sinks that absorbs CO_2 (light green arrows).

Featured Problem 12.2 (Greenhouse Gasses and atmospheric CO_2)
Climate change has been attributed partly to the accumulation of greenhouse gasses (such as carbon dioxide and methane) in the atmosphere.

Here we consider a simplified illustrative model for the carbon cycle that tracks the sources and sinks of CO_2 in the atmosphere. Consider $C(t)$ as the level of atmospheric carbon dioxide. Define the production rate of $C0_2$ due to utilization of fossil fuel and other human activity to be E_{FF}, and let the rate of absorption of CO_2 by the oceans be S_{OCEAN}. We will also assume that living plants absorb CO_2 at a rate proportional to their biomass and to the CO_2 level.

1. *Explain the following differential equation for atmospheric CO_2:*

$$\frac{dC}{dt} = E_{FF} - S_{OCEAN} - \gamma PC. \tag{12.15}$$

2. *Assuming that $E_{FF}, S_{OCEAN}, \gamma, P$ are constants, find the steady state level of CO_2 in terms of these parameters.*

3. *Find $C(t)$, that is, predict the amount of CO_2 over time, assuming that $C(0) = C_0$.*

Hint: CO_2 is usually given in units of "parts per million", ppm ($=10^{-6}$) ,1 ppm = 2.1 GtC. (1GtC= 1 gigaton carbon = 10^9 tons.) Time is typically given in years, so rates are "per year" (yr^{-1}).
Approximate parameter values:
$E_{FF} \approx 10$ GtC yr^{-1},
$S_{OCEAN} \approx 3$ GtC yr^{-1},
$P \approx 560$ Gt plant biomass,
$\gamma \approx 1.35 \cdot 10^{-5}$ yr^{-1} Gt^{-1}.

4. *Graph the function $C(t)$ for parameter values given in the problem, assuming that $C_0 = 400$ppm= 840 GtC.*

5. *How big an effect would be produced on the CO_2 level in 50 years if 15% of the plant biomass is removed to deforestation just prior to $t = 0$?*

Note: Information for Problem 12.2 is adapted from [Le Quéré et al., 2016], and may reflect many simplifications and approximations. In actual fact, most "constants" in the problem are time-dependent, making the real problem of predicting CO_2 levels much more challenging.

12.3 Euler's Method and numerical solutions

Section 12.3 Learning goals

1. Explain the idea of a numerical solution to a differential equation and how this compares with an exact or analytic solution.
2. Describe how Euler's method is based on approximating the derivative by the slope of a secant line.
3. Use Euler's method to calculate a numerical solution (using a spreadsheet) to a given initial value problem.

So far, we have explored ways of understanding the behaviour predicted by a differential equation in the form of an **analytic solution**, namely an explicit formula for the solution as a function of time. However, in reality this is typically difficult without extensive training, and occasionally, impossible even for experts. Even if we can find such a solution, it may be inconvenient to determine its numerical values at arbitrary times, or to interpret its behaviour.

For this reason, we sometimes need a method for computing an approximation for the desired solution. We refer to that approximation as a **numerical solution**. The idea is to harness a computational device - computer, laptop, or calculator - to find numerical values of points along the solution curve, rather than attempting to determine the formula for the solution as a function of time. We illustrate this process using a technique called **Euler's method**, which is based on an approximation of a derivative by the slope of a secant line.

Below, we describe how Euler's method is used to approximate the solution to a general initial value problem (differential equation together with initial condition) of the form

$$\frac{dy}{dt} = f(y), \quad y(0) = y_0.$$

Set up. We first must pick a "step size," Δt, and subdivide the t axis into discrete steps of that size. We thus have a set of time points $t_1, t_2, \ldots$, spaced

Δt apart as shown in Figure 12.8. Our procedure starts with the known initial value $y(0) = y_0$, and uses it to generate an approximate value at the next time point (y_1), then the next (y_2), and so on. We denote by y_k the value of the independent variable generated at the k'th time step by Euler's method as an approximation to the (unknown) true solution $y(t_k)$.

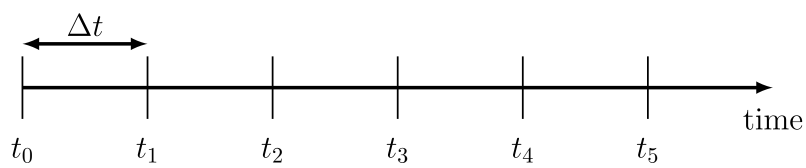


Figure 12.8: The time axis is subdivided into steps of size Δt.

Mastered Material Check

18. If $\Delta t = 0.1$ and $t_0 = 0$, what are t_1, t_2 and t_3?
19. Explain the difference between the value y_1 and the true solution $y(t_1)$.
20. If Δt is not sufficiently small, why might Euler's method give a bad approximation to the solution?

Method. We approximate the differential equation by a **finite difference equation**

$$\frac{dy}{dt} = f(y) \quad \text{approximated by} \quad \frac{y_{k+1} - y_k}{\Delta t} = f(y_k).$$

This approximation is reasonable only when Δt, the time step size, is small. Rearranging this equation leads to a process (also called **recurrence relation**) for linking values of the solution at successive time points,

$$\frac{y_{k+1} - y_k}{\Delta t} = f(y_k), \quad \Rightarrow \quad y_{k+1} = y_k + \Delta t \cdot f(y_k). \tag{12.16}$$

Application. We start with the known initial value, y_0. Then (setting the index to $k = 0$ in Eqn. (12.16)) we obtain

$$y_1 = y_0 + f(y_0)\Delta t.$$

The quantities on the right are known, so we can compute the value of y_1, which is the approximation to the solution $y(t_1)$ at the time point t_1. We can then continue to generate the value at the next time point in the same way, by approximating the derivative again as a secant slope. This leads to

$$y_2 = y_1 + f(y_1)\Delta t.$$

The approximation so generated, leading to values $y_1, y_2, \ldots$ is called **Euler's method**.

Applying this approximation repeatedly, leads to an **iteration method**, that is, the repeated computation

$$\begin{aligned} y_1 &= y_0 + f(y_0)\Delta t, \\ y_2 &= y_1 + f(y_1)\Delta t, \\ &\vdots \\ y_{k+1} &= y_k + f(y_k)\Delta t. \end{aligned}$$

Mastered Material Check

21. In Euler's method, can you determine t_2 directly? That is, without first computing t_1?
22. In Euler's method, can you determine y_2 directly? That is, without first computing y_1?

From this iteration, we obtain the approximate values of the function $y_k \approx y(t_k)$ for as many time steps as desired starting from $t = 0$ in increments of Δt up to some final time T of interest.

It is customary to use the following notations:

- t_0 : the initial time point, usually at $t = 0$.
- $h = \Delta t$: common notations for the step size, i.e. the distance between the points along the t axis.
- t_k : the k'th time point. Note that since the points are at multiples of the step size that we have picked, $t_k = k\Delta t = kh$.
- $y(t)$: the actual value of the solution to the differential equation at time t. This is usually not known, but in the examples discussed in this chapter, we can solve the differential equation exactly, so we have a formula for the function $y(t)$. In most hard scientific problems, no such formula is known in advance.
- $y(t_k)$: the actual value of the solution to the differential equation at one of the discrete time points, t_k (again, not usually known).
- y_k : the approximate value of the solution obtained by Euler's method. We hope that this approximate value is fairly close to the true value, i.e. that $y_k \approx y(t_k)$, but there is always some error in the approximation. More advanced methods that are specifically designed to reduce such errors are discussed in courses on numerical analysis.

Euler's method applied to population growth

We illustrate how Euler's method is used in a familiar example, that of unlimited population growth.

Example 12.10 *Apply Euler's method to approximating solutions for the simple exponential growth model that was studied in Chapter 11,*

$$\frac{dy}{dt} = ay, \quad y(0) = y_0$$

where a is a constant (see Eqn 11.2).

Solution. Subdivide the t axis into steps of size Δt, starting with $t_0 = 0$, and $t_1 = \Delta t, t_2 = 2\Delta t, \ldots$ The first value of y is known from the initial condition,

$$y_0 = y(0) = y_0.$$

We replace the differential equation by the approximation

$$\frac{y_{k+1} - y_k}{\Delta t} = ay_k \quad \Rightarrow \quad y_{k+1} = y_k + a\Delta t y_k, \quad k = 1, 2, \ldots$$

Mastered Material Check

23. Carry our Example 12.10 with $\Delta t = 0.1$, $a = 1$, and $y_0 = 1$.
24. Plot the first 5 points you determine. Compare with the true solution.
25. Solve the initial value problem in Example 12.11 analytically. Compare the points $(0, 100)$, $(0.1, 95)$, $(0.2, 90.25)$ and $(0.3, 85.7375)$ with the true solution at the corresponding t values.

In particular,

$$\begin{aligned} y_1 &= y_0 + a\Delta t y_0 = y_0(1+a\Delta t), \\ y_2 &= y_1(1+a\Delta t), \\ y_3 &= y_2(1+a\Delta t), \end{aligned}$$

and so on. At every stage, the quantity on the right hand side depends only on value of y_k that as already known from the step before. ♢

The next example demonstrates Euler's method applied to a specific differential equation.

Example 12.11 *Use Euler's method to find the solution to*

$$\frac{dy}{dt} = -0.5y, \quad y(0) = 100.$$

Use step size $\Delta t = 0.1$ *to approximate the solution for the first two time steps.*

Link to Google Sheets. This spreadsheet implements Euler's method for Example 12.11. You can view the formulae by clicking on a cell in the sheet but you cannot edit the sheet here.

Solution. Euler's method applied to this example would lead to

$$\begin{aligned} y_0 &= 100. \\ y_1 &= y_0(1+a\Delta t) = 100(1+(-0.5)(0.1)) = 95, \quad \text{etc.} \end{aligned}$$

We show the first five values in Table 12.3. Clearly, these kinds of repeated calculations are best handled on a spreadsheet or similar computer software.

k	t_k	y_k
0	0	100.00
1	0.1	95.00
2	0.2	90.25
3	0.3	85.74
4	0.4	81.45
5	0.5	77.38

Table 12.3: Euler's method applied to Example 12.11.

Euler's method applied to Newton's law of cooling

We apply Euler's method to Newton's law of cooling. Upon completion, we can directly compare the approximate numerical solution generated by Euler's method to the true (analytic) solution, (12.8), that we determined earlier in this chapter.

Example 12.12 (Newton's law of cooling) *Consider the temperature of an object* $T(t)$ *in an ambient temperature of* $E = 10°$. *Assume that* $k = 0.2$*/min. Use the initial value problem*

$$\frac{dT}{dt} = k(E-T), \quad T(0) = T_0$$

to write the exact solution to Eqn. (12.8) *in terms of the initial value* T_0.

Solution. In this case, the differential equation has the form

$$\frac{dT}{dt} = 0.2(10-T),$$

and its analytic solution, from Eqn. (12.8), is

$$T(t) = 10 + (T_0 - 10)e^{-0.2t}. \tag{12.17}$$

♢

Below, we use Euler's method to compute a solution from each of several initial conditions, $T(0) = 0, 5, 15, 20$ degrees.

Example 12.13 (Euler's method applied to Newton's law of cooling) *Write the Euler's method procedure for the approximate solution to the problem in Example 12.12.*

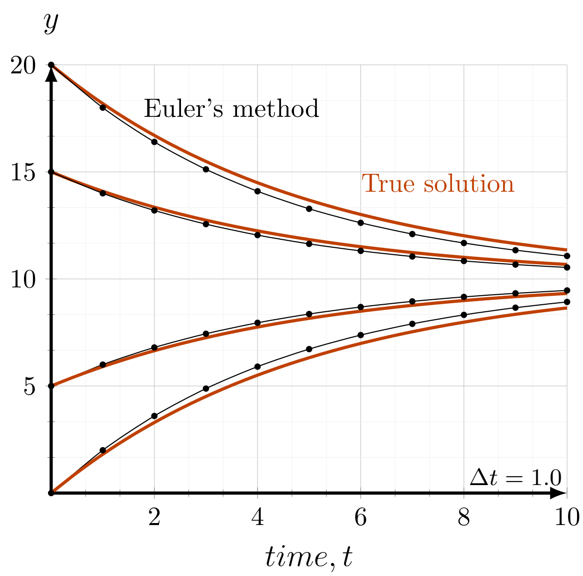
Solution. Euler's method approximates the differential equation by

$$\frac{T_{k+1} - T_k}{\Delta t} = 0.2(10 - T_k).$$

or, in simplified form,

$$T_{k+1} = T_k + 0.2(10 - T_k)\Delta t.$$

◊



time t_k	approx solution T_k	exact soln $T(t)$
0.0000	0.0000	0.0000
1.0000	2.0000	1.8127
2.0000	3.6000	3.2968
3.0000	4.8800	4.5119
4.0000	5.9040	5.5067
5.0000	6.7232	6.3212
6.0000	7.3786	6.9881
7.0000	7.9028	7.5340
8.0000	8.3223	7.9810

Figure 12.9: Euler's method applied to Newton's law of cooling. The graph shows the true solution (red) and the approximate solution (black).

Example 12.14 *Use Euler's method from Example 12.13 and time steps of size $\Delta t = 1.0$ to find a numerical solution to the the cooling problem. Use a spreadsheet for the calculations. Note that $\Delta t = 1.0$ is not a "small step;" we use it here for illustration purposes.*

Solution. The procedure to implement is

$$T_{k+1} = T_k + 0.2(10 - T_k)\Delta t.$$

In Figure 12.9 we show a typical example of the method with initial value $T(0) = T_0$ and with the time step size $\Delta t = 1.0$. Black dots represent the discrete values generated by the Euler method, starting from initial conditions, $T_0 = 0, 5, 15, 20$. Notice that the black curve is simply made up of line segments linking points obtained by the numerical solution. On the same graph, we also show the analytic solution (red curves) given by Eqn. (12.17) with the

Mastered Material Check

26. What change would you make in the process set up in Example 12.14 to improve the approximation made by Euler's method?

same four initial temperatures. We see that the black and red curves start out at the same points (since they both satisfy the same initial conditions). However, the approximate solution obtained with Euler's method is not identical to the true solution. The difference between the two (gap between the red and black curves) is the **numerical error** in the approximation.

12.4 Summary

1. Given a function, we can check whether it is a solution to a differential equation by performing the appropriate differentiation and algebraic simplification.
2. Solutions to differential equations in which there is no change at all ("constant solutions") are referred to as steady states.
3. The differential equations

   $$\frac{dy}{dt} = a - by, \qquad y(0) = y_0$$

   has a steady state solution $y = a/b$.
4. If we define the deviation from steady state, $z(t) = y(t) - \frac{a}{b}$, we get a decay equation for $z(t)$ that has exponentially decreasing solutions provided $b > 0$. This says that the eviation from steady state always decrease over time.
5. The resulting solution for $y(t)$ is

   $$y(t) = \frac{a}{b} - \left(\frac{a}{b} - y_0\right) e^{-bt}.$$

6. For some differential equations, it is not always possible to determine an analytic solution (explicit formula). Numerical solutions can be found using Euler's method, and serve as an approximate solution.
7. Euler's method takes a known initial value y_0 and uses the iteration scheme:

   $$y_{k+1} = y_k + f(y_k)\Delta t.$$

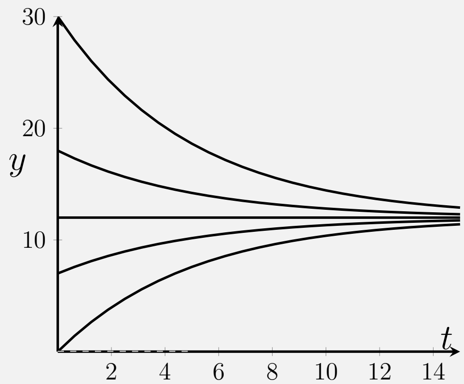
   to generate successive values of y_k that approximate the solution at time points $t_k = k\Delta t$
8. Applications considered in this chapter included:

   (a) height of water draining out of a cylindrical container (verifying a solution to a differential equation);
   (b) Newton's law of cooling (described by a linear differential equation);
   (c) growth of the radius of a cell;
   (d) the accumulation of greenhouse gasses in the atmosphere;
   (e) friction and terminal velocity; and
   (f) chemical production and decay.

Quick Concept Checks

1. Explain why an object at room temperature is at a steady state for Newton's law of cooling.

2. The following graph depicts solution curves to a particular differential equation of the form $dy/dt = a - by$.



   (a) Estimate the value that these solution curves are approaching.

   (b) Which solutions are approaching from above? From below?

3. Consider the following initial value problem:

$$\frac{dy}{dt} = 2 - 4y, \quad y(0) = 4,$$

   (a) What value does its solution curve approach?

   (b) Does its solution approach from above or below?

4. Why is a large value of Δt not a good idea when using Euler's method?

Exercises

12.1. **Water draining form a container.** In Example 12.3, we verified that the function $h(t) = \left(\sqrt{h_0} - k\frac{t}{2}\right)^2$ is a solution to the differential equation (12.2). Based on the meaning of the problem, for how long does this solution remain valid?

12.2. **Verifying a solution.** Verify that the function $y(t) = 1 - (1 - y_0)e^{-t}$ satisfies the initial value problem (differential equation and initial condition) (12.6).

12.3. **Linear differential equation.** Consider the differential equation

$$\frac{dy}{dt} = a - by$$

where a, b are constants.

(a) Show that the function

$$y(t) = \frac{a}{b} - Ce^{-bt}$$

satisfies the above differential equation for any constant C.

(b) Show that by setting

$$C = \frac{a}{b} - y_0$$

we also satisfy the initial condition

$$y(0) = y_0.$$

Remark: you have shown that the function

$$y(t) = \left(y_0 - \frac{a}{b}\right)e^{-bt} + \frac{a}{b}$$

is a solution to the *initial value problem* (i.e differential equation plus initial condition)

$$\frac{dy}{dt} = a - by, \quad y(0) = y_0.$$

12.4. **Steps in an example.** Complete the algebraic steps in Example **??** to show that the solution to Eqn. (12.4) can be obtained by the substitution $z(t) = a - by(t)$.

12.5. **Verifying a solution.** Show that the function

$$y(t) = \frac{1}{1-t}$$

is a solution to the differential equation and initial condition

$$\frac{dy}{dt} = y^2, \quad y(0) = 1.$$

Comment on what happens to this solution as t approaches 1.

12.6. **Verifying solutions.** For each of the following, show the given function y is a solution to the given differential equation.

(a) $t \cdot \dfrac{dy}{dt} = 3y$, $y = 2t^3$.

(b) $\dfrac{d^2y}{dt^2} + y = 0$, $y = -2\sin t + 3\cos t$.

(c) $\dfrac{d^2y}{dt^2} - 2\dfrac{dy}{dt} + y = 6e^t$, $y = 3t^2e^t$.

12.7. **Verifying a solution.** Show the function determined by the equation $2x^2 + xy - y^2 = C$, where C is a constant and $2y \neq x$, is a solution to the differential equation $(x-2y)\dfrac{dy}{dx} = -4x - y$.

12.8. **Determining the constant.** Find the constant C that satisfies the given initial conditions.

(a) $2x^2 - 3y^2 = C$, $y|_{x=0} = 2$.

(b) $y = C_1e^{5t} + C_2te^{5t}$, $y|_{t=0} = 1$ and $\frac{dy}{dt}|_{t=0} = 0$.

(c) $y = C_1\cos(t - C_2)$, $y|_{t=\frac{\pi}{2}} = 0$ and $\frac{dy}{dt}|_{t=\frac{\pi}{2}} = 1$.

12.9. **Friction and terminal velocity.** The velocity of a falling object changes due to the acceleration of gravity, but friction has an effect of slowing down this acceleration. The differential equation satisfied by the velocity $v(t)$ of the falling object is

$$\frac{dv}{dt} = g - kv$$

where g is acceleration due to gravity and k is a constant that represents the effect of friction. An object is dropped from rest from a plane.

(a) Find the function $v(t)$ that represents its velocity over time.

(b) What happens to the velocity after the object has been falling for a long time (but before it has hit the ground)?

12.10. **Alcohol level.** Alcohol enters the blood stream at a constant rate k gm per unit time during a drinking session. The liver gradually converts the alcohol to other, non-toxic byproducts. The rate of conversion per unit time is proportional to the current blood alcohol level, so that the differential equation satisfied by the blood alcohol level is

$$\frac{dc}{dt} = k - sc$$

where k, s are positive constants. Suppose initially there is no alcohol in the blood.

Find the blood alcohol level $c(t)$ as a function of time from $t = 0$, when the drinking started.

12.11. **Checking a solution.** Check that the differential equation (12.7) has the right sign, so that a hot object cools off in a colder environment.

12.12. **Details of Newtons Law of Cooling.** Fill in the missing steps in the solution to Newton's Law of Cooling in Example 12.5.

12.13. **Newton's Law of Cooling.** Newton's Law of Cooling states that the rate of change of the temperature of an object is proportional to the difference between the temperature of the object, T, and the ambient (environmental) temperature, E. This leads to the *differential equation*

$$\frac{dT}{dt} = k(E - T)$$

where $k > 0$ is a constant that represents the material properties and, E is the ambient temperature. (We assume that E is also constant.)

(a) Show that the function

$$T(t) = E + (T_0 - E)e^{-kt}$$

which represents the temperature at time t satisfies this equation.

(b) The time of death of a murder victim can be estimated from the temperature of the body if it is discovered early enough after the crime has occurred.

Suppose that in a room whose ambient temperature is $E = 20°$ C, the temperature of the body upon discovery is $T = 30°$ C, and that a second measurement, one hour later is $T = 25°$ C.

Determine the approximate time of death.

Remark: use the fact that just prior to death, the temperature of the victim was 37°C.

12.14. **A cup of coffee.** The temperature of a cup of coffee is initially 100 degrees C. Five minutes later, ($t = 5$) it is 50 degrees C. If the ambient temperature is $A = 20$ degrees C, determine how long it takes for the temperature of the coffee to reach 30 degrees C.

12.15. **Newton's Law of Cooling applied to data.** The data presented in Table 12.4 was gathered in producing Figure 2.2 for cooling milk during yoghurt production. According to Newton's Law of Cooling, this data can be described by the formula

$$T = E + (T(0) - E)e^{-kt}.$$

where $T(t)$ is the temperature of the milk (in degrees Fahrenheit) at time t (in min), E is the ambient temperature, and k is some constant that we determine in this exercise.

(a) Rewrite this relationship in terms of the quantity $Y(t) = \ln(T(t) - E)$, and show that $Y(t)$ is related linearly to the time t.

(b) Explain how the constant k could be found from this converted form of the relationship.

time (min)	**Temp**
0.0	190.0
0.5	185.5
1.0	182.0
1.5	179.2
2.0	176.0
2.5	172.9
3.0	169.5
3.5	167.0
4.0	164.6
4.5	162.2
5.0	159.8

Table 12.4: Cooling milk data for Exercise 15.

(c) Use the data in the table and your favourite spreadsheet (or similar software) to show that the data so transformed appears to be close to linear. Assume that the ambient temperature was $E = 20°$F.

(d) Use the same software to determine the constant k by fitting a line to the transformed data.

12.16. **Infant weight gain.** During the first year of its life, the weight of a baby is given by

$$y(t) = \sqrt{3t + 64}$$

where t is measured in some convenient unit.

(a) Show that y satisfies the differential equation

$$\frac{dy}{dt} = \frac{k}{y}$$

where k is some positive constant.

(b) What is the value for k?

(c) Suppose we adopt this differential equation as a model for human growth. State concisely (that is, in one sentence) one feature about this differential equation which makes it a reasonable model. State one feature which makes it unreasonable.

12.17. **Lake Fishing.** Fish Unlimited is a company that manages the fish population in a private lake. They restock the lake at constant rate (to restock means to add fish to the lake): N fishers are allowed to fish in the lake per day. The population of fish in the lake, $F(t)$ is found to satisfy the differential equation

$$\frac{dF}{dt} = I - \alpha NF \qquad (12.18)$$

(a) At what rate are fish added per day according to Eqn. (12.18)? Give both value and units.

(b) What is the average number of fish caught by one fisher? Give both the value and units.

(c) What is being assumed about the fish birth and mortality rates in Eqn. (12.18)?

(d) If the fish input and number of fishers are constant, what is the steady state level of the fish population in the lake?

(e) At time $t = 0$ the company stops restocking the lake with fish. Give the revised form of the differential equation (12.18) that takes this into account, assuming the same level of fishing as before. How long would it take for the fish to fall to 25% of their initial level?

(f) When the fish population drops to the level F_{low}, fishing is stopped and the lake is restocked with fish at the same constant rate

(Eqn (12.18), with $\alpha = 0$.) Write down the revised version of Eqn. (12.18) that takes this into account. How long would it take for the fish population to double?

12.18. **Tissue culture.** Cells in a tissue culture produce a cytokine (a chemical that controls the growth of other cells) at a constant rate of 10 nano-Moles per hour (nM/h). The chemical has a half-life of 20 hours.

Give a differential equation (DE) that describes this chemical production and decay. Solve this DE assuming that at $t = 0$ there is no cytokine. [1nM=10^{-9}M].

12.19. **Glucose solution in a tank.** A tank that holds 1 liter is initially full of plain water. A concentrated solution of glucose, containing 0.25 gm/cm^3 is pumped into the tank continuously, at the rate 10 cm^3/min and the mixture (which is continuously stirred to keep it uniform) is pumped out at the same rate.

How much glucose is in the tank after 30 minutes? After a long time? (*hint:* write a differential equation for c, the concentration of glucose in the tank by considering the rate at which glucose enters and the rate at which glucose leaves the tank.)

12.20. **Pollutant in a lake.** A lake of constant volume V gallons contains $Q(t)$ pounds of pollutant at time t evenly distributed throughout the lake. Water containing a concentration of k pounds per gallon of pollutant enters the lake at a rate of r gallons per minute, and the well-mixed solution leaves at the same rate.

(a) Set up a differential equation that describes the way that the amount of pollutant in the lake changes.

(b) Determine what happens to the pollutant level after a long time if this process continues.

(c) If $k = 0$ find the time T for the amount of pollutant to be reduced to one half of its initial value.

12.21. **A sugar solution.** Sugar dissolves in water at a rate proportional to the amount of sugar not yet in solution. Let $Q(t)$ be the amount of sugar undissolved at time t. The initial amount is 100 kg and after 4 hours the amount undissolved is 70 kg.

(a) Find a differential equation for $Q(t)$ and solve it.

(b) How long does it take for 50 kg to dissolve?

12.22. **Leaking water tank.** A cylindrical tank with cross-sectional area A has a small hole through which water drains. The height of the water in the tank $y(t)$ at time t is given by:

$$y(t) = (\sqrt{y_0} - \frac{kt}{2A})^2$$

where k, y_0 are constants.

(a) Show that the height of the water, $y(t)$, satisfies the differential equation

$$\frac{dy}{dt} = -\frac{k}{A}\sqrt{y}.$$

(b) What is the initial height of the water in the tank at time $t = 0$?

(c) At what time is the tank be empty ?

(d) At what rate is the **volume** of the water in the tank changing when $t = 0$?

12.23. **Determining constants.** Find those constants a, b so that $y = e^x$ and $y = e^{-x}$ are both solutions of the differential equation

$$y'' + ay' + by = 0.$$

12.24. **Euler's method.** Solve the decay equation in Example (12.11) analytically, that is, find the formula for the solution in terms of a decaying exponential, and then compare your values to the approximate solution values y_1 and, y_2 computed with Euler's method.

12.25. **Comparing approximate and true solutions:**

(a) Use Euler's method to find an approximate solution to the differential equation

$$\frac{dy}{dx} = y$$

with $y(0) = 1$. Use a step size $h = 0.1$ and find the values of y up to $x = 0.5$. Compare the value you have calculated for $y(0.5)$ using Euler's method with the true solution of this differential equation. What is the **error** i.e. the difference between the true solution and the approximation?

(b) Now use Euler's method on the differential equation

$$\frac{dy}{dx} = -y$$

with $y(0) = 1$. Use a step size $h = 0.1$ again and find the values of y up to $x = 0.5$. Compare the value you have calculated for $y(0.5)$ using Euler's method with the true solution of this differential equation. What is the error this time?

12.26. **Beginning Euler's method.** Give the first 3 steps of Euler's method for the problem in Example 12.13.

12.27. **Euler's method and a spreadsheet.** Use the spreadsheet and Euler's method to solve the differential equation shown below:

$$dy/dt = 0.5y(2 - y)$$

Use a step size of $h = 0.1$ and show (on the same graph) solutions for the following four initial values:

$$y(0) = 0.5,\; y(0) = 1,\; y(0) = 1.5,\; y(0) = 2.25$$

For full credit, include a short explanation your process (e.g. 1-2 sentences and whatever equations you implemented on the spreadsheet.)

13
Qualitative methods for differential equations

Mastered Material Check

1. What is meant by an analytic solution to a differential equation?
2. What other kind of solutions are possible?
3. Give an example of a nonlinear function $f(y)$.

Not all differential equations are easily solved analytically. Furthermore, even when we find the analytic solution, it is not necessarily easy to interpret, graph, or understand. This situation motivates qualitative methods that promote an overall understanding of behaviour - directly from information in the differential equation - without the challenge of finding a full functional form of the solution.

In this chapter we expand our familiarity with differential equations and assemble new, qualitative techniques for understanding them. We consider differential equations in which the expression on one side, $f(y)$, is **nonlinear**, i.e. equations of the form

$$\frac{dy}{dt} = f(y)$$

in which f is more complicated than the form $a - by$. Geometric techniques, rather than algebraic calculations form the core of the concepts we discuss.

13.1 Linear and nonlinear differential equations

Section 13.1 Learning goals

1. Identify the distinction between unlimited and density-dependent population growth. Be able to explain terms in the logistic equation in its original version, Eqn. (13.1), and its rescaled version, Eqn. (13.3).
2. State the definition of a linear differential equation.
3. Explain the law of mass action, and derive simple differential equations for interacting species based on this law.

In the model for population growth in Chapter 11, we encountered the differential equation

$$\frac{dN}{dt} = kN,$$

where $N(t)$ is population size at time t and k is a constant per capita growth rate. We showed that this differential equation has exponential solutions. It means that two behaviours are generically obtained: **explosive growth** if $k > 0$ or **extinction** if $k < 0$.

The case of $k > 0$ is unrealistic, since real populations cannot keep growing indefinitely in an explosive, exponential way. Eventually running out of space or resources, the population growth dwindles, and the population attains some static level rather than expanding forever. This motivates a revision of our previous model to depict **density-dependent growth**.

Mastered Material Check

4. What happens in the case that $k = 0$? Explain under what conditions this might arise and what happens to the population $N(t)$ in this case.

The logistic equation for population growth

Let $N(t)$ represent the size of a population at time t, as before. Consider the differential equation

$$\frac{dN}{dt} = rN\frac{(K-N)}{K}. \tag{13.1}$$

We call this differential equation the **logistic equation**.The logistic equation has a long history in modelling population growth of humans, microorganisms, and animals. Here the parameter r is the **intrinsic growth rate** and K is the **carrying capacity**. Both r, K are assumed to be positive constants for a given population in a given environment.

In the form written above, we could interpret the logistic equation as

$$\frac{dN}{dt} = R(N)\cdot N, \quad \text{where } R(N) = \left[r\frac{(K-N)}{K}\right].$$

The term $R(N)$ is a function of N that replaces the constant rate of growth k (found in the unrealistic, unlimited population growth model). R is called the **density dependent growth rate**.

Linear versus nonlinear

The logistic equation introduces the first example of a **nonlinear differential equation**. We explain the distinction between linear and nonlinear differential equations and why it matters.

Definition 13.1 (Linear differential equation) *A first order differential equation is said to be linear if it is a linear combination of terms of the form*

$$\frac{dy}{dt}, \quad y, \quad 1$$

that is, it can be written in the form

$$\alpha\frac{dy}{dt} + \beta y + \gamma = 0 \tag{13.2}$$

where α, β, γ do not depend on y. Note that "first order" means that only the first derivative (or no derivative at all) may occur in the equation.

Mastered Material Check

5. Can the differential equation $\frac{dy}{dt} = a - by$ be written in the form (13.2)? If so, what are the values of α, β, γ?

So far, we have seen several examples of this type with constant coefficients α, β, γ. For example, $\alpha = 1, \beta = -k$, and $\gamma = 0$ in Eqn. 11.2 whereas $\alpha = 1, \gamma = -a$, and $\beta = b$ in Eqn. (12.4). A differential equation that is not of this form is said to be nonlinear.

Example 13.1 (Linear versus nonlinear differential equations) *Which of the following differential equations are linear and which are nonlinear?*

$$(a)\ \frac{dy}{dt} = y^2, \quad (b)\ \frac{dy}{dt} - y = 5, \quad (c)\ y\frac{dy}{dt} = -1.$$

Solution. Any term of the form y^2, $\sqrt{y}$, $1/y$, etc. is nonlinear in y. A product such as $y\frac{dy}{dt}$ is also nonlinear in the independent variable. Hence equations (a), (c) are nonlinear, while (b) is linear. ◊

The significance of the distinction between linear and nonlinear differential equations is that nonlinearities make it much harder to systematically find a solution to the given differential equation by "analytic" methods. Most linear differential equations have solutions that are made of exponential functions or expressions involving such functions. This is not true for nonlinear equations.

However, as we see shortly, geometric methods are very helpful in understanding the behaviour of such nonlinear differential equations.

Mastered Material Check

6. For what values of α, β and γ can Example 13.1(b) be put into the form (13.2)?

Law of Mass Action

Nonlinear terms in differential equations arise naturally in various ways. One common source comes from describing interactions between individuals, as the following example illustrates.

In a chemical reaction, molecules of types A and B bind and react to form product P. Let $a(t), b(t)$ denote the concentrations of A and B. These concentrations depend on time because the chemical reaction uses up both types in producing the product.

The reaction only occurs when A and B molecules "collide" and stick to one another. Collisions occur randomly, but if concentrations are larger, more collisions take place, and the reaction is faster. If either the concentration a or b is doubled, then the reaction rate doubles. But if both a and b are doubled, then the reaction rate should be four times faster, based on the higher chances of collisions between A and B. The simplest assumption that captures this dependence is

$$\text{rate of reaction is proportional to } a \cdot b \quad \Rightarrow \quad \text{rate of reaction } = k \cdot a \cdot b$$

where k is some constant that represents the reactivity of the molecules.

We can formally state this result, known as the **Law of Mass Action** as follows:

Mastered Material Check

7. If the concentration of A is tripled, and that of B is doubled, how much faster would we expect the reaction rate to be?
8. Why does the product $a \cdot b$, rather than the sum $a + b$ appear in the Law of Mass Action ?

The Law of Mass Action: The rate of a chemical reaction involving an interaction of two or more chemical species is proportional to the *product of the concentrations* of the given species.

Example 13.2 (Differential equation for interacting chemicals) *Substance A is added at a constant rate of I moles per hour to a 1-litre vessel. Pairs of molecules of A interact chemically to form a product P. Write down a differential equation that keeps track of the concentration of A, denoted $y(t)$.*

Solution. First consider the case that there is no reaction. Then, the addition of A to the reactor at a constant rate leads to changing $y(t)$, described by the differential equation

$$\frac{dy}{dt} = I.$$

When the chemical reaction takes place, the depletion of A depends on interactions of pairs of molecules. By the law of mass action, the rate of reaction is of the form $k \cdot y \cdot y = ky^2$, and as it reduces the concentration, it appear with a minus sign in the DE. Hence

$$\frac{dy}{dt} = I - ky^2.$$

This is a nonlinear differential equation - it contains a term of the form y^2. ◊

Mastered Material Check

9. In each of Examples 13.2 and 13.3, clearly identify the constant quantities.

Example 13.3 (Logistic equation reinterpreted) *Rewrite the logistic equation in the form*

$$\frac{dN}{dt} = rN - bN^2$$

(where $b = r/K$ is a positive quantity).

a) *Interpret the meaning of this restated form of the equation by explaining what each of the terms on the right hand side could represent.*

b) *Which of the two terms dominates for small versus large population levels?*

Solution.

a) This form of the equation has growth term rN proportional to population size, as encountered previously in unlimited population growth. However, there is also a quadratic (nonlinear) rate of loss (note the minus sign) $-bN^2$. This term could describe interactions between individuals that lead to mortality, e.g. through fighting or competition.

b) From familiarity with power functions (in this case, the functions of N that form the two terms, rN and bN^2) we can deduce that the second, quadratic term dominates for larger values of N, and this means that when the population is crowded, the loss of individuals is greater than the rate of reproduction. ◊

Scaling the logistic equation

Consider units involved in the logistic equation (13.1):

$$\frac{dN}{dt} = rN\frac{(K-N)}{K}.$$

This equation has two parameters, r and K. Since units on each side of an equation must balance, and must be the same for terms that are added or subtracted, we can infer that K has the same units as N, and thus it is a population density. When $N = K$, the population growth rate is zero ($dN/dt = 0$).

It turns out that we can understand the behaviour of the logistic equation by converting it to a "generic" form that does not depend on the constant K. We do so by transforming variables, which amounts to choosing a convenient way to measure the population size.

Example 13.4 (Rescaling) *Define a new variable*

$$y(t) = \frac{N(t)}{K},$$

with $N(t)$ and K as in the logistic equation. Then $N(t) = Ky(t)$.

a) *Interpret what the transformed variable y represents.*

b) *Rewrite the logistic equation in terms of this variable.*

Mastered Material Check

10. Suppose an environment can sustain 2000 aphids per plant, and the current population size on a given plant is 1700. What is K, N and y based on this information?
11. This population is at what percent of its carrying capacity?

Solution.

a) The variable, $y(t)$ represents a scaled version of the population density. Instead of measuring the population in some arbitrary units - such as number of individuals per acre, or number of bacteria per ml - $y(t)$ measures the population in "multiples of the carrying capacity."

For example, if the environment can sustain 1000 aphids per plant (so $K = 1000$ individuals per plant), and the current population size on a given plant is $N = 950$ then the value of the scaled variable is $y = 950/1000 = 0.95$. We would say that "the aphid population is at 95% of its carrying capacity on the plant."

b) Since K is assumed constant, it follows that

$$N(t) = Ky(t) \quad \Rightarrow \quad \frac{dN}{dt} = K\frac{dy}{dt}.$$

Using this, we can simplify the logistic equation as follows:

$$\begin{aligned} \frac{dN}{dt} = rN\frac{(K-N)}{K}, \quad &\Rightarrow \quad K\frac{dy}{dt} = r(Ky)\frac{(K-Ky)}{K}, \\ &\Rightarrow \quad \frac{dy}{dt} = ry(1-y). \end{aligned} \tag{13.3}$$

◊

Eqn. (13.3) "looks simpler" than Eqn. (13.1) since it depends on only one parameter, r. Moreover, by understanding this equation, and transforming back to the original logistic in terms of $N(t) = Ky(t)$, we can interpret results for the original model. While we do not go further with transforming variables at present, it turns out that one can also further reduce the scaled logistic to an equation in which $r = 1$ by "rescaling time units".

Mastered Material Check

12. What are the units of the parameter r?
13. How might we use the parameter r to define a time-scale?

13.2 The geometry of change

Section 13.2 Learning goals

1. Explain what is a **slope field** of a differential equation. Given a differential equation (linear or nonlinear), construct such a diagram and use it to sketch solution curves.
2. Describe what a **state-space diagram** is; construct such a diagram and use it to interpret the behaviour of solution curves to a given differential equation.
3. Identify the relationships between a slope field, a state-space diagram, and a family of solution curves to a given differential equation.
4. Identify steady states of a differential equation and determine whether they are stable or unstable.
5. Given a differential equation and initial condition, predict the behaviour of the solution for $t > 0$.

In this section, we introduce a new method for understanding differential equations using graphical and geometric arguments. Such methods circumvent the solutions that we expressed in terms of analytic formulae. We resort to concepts learned much earlier - for example, the derivative as a slope of a tangent line - in order to use the differential equation itself to assemble a sketch of the behaviour that it predicts. That is, rather than writing down $y = F(t)$ as a solution to the differential equation (and then graphing that function) we sketch the qualitative behaviour of such solution curves directly from information contained in the differential equation.

Slope fields

Here we discuss a geometric way of understanding what a differential equation is saying using a **slope field**, also called a **direction field**. We have already seen that solutions to a differential equation of the form

$$\frac{dy}{dt} = f(y)$$

are curves in the (y,t)-plane that describe how $y(t)$ changes over time (thus, these curves are graphs of functions of time). Each initial condition $y(0) = y_0$ is associated with one of these curves, so that together, these curves form a *family* of solutions.

What do these curves have in common, geometrically?

- the slope of the tangent line (dy/dt) at any point on any of the curves is related to the value of the y-coordinate of that point - as stated in the differential equation.
- at any point $(t, y(t))$ on a solution curve, the tangent line must have slope $f(y)$, which depends only on the y value, and not on the time t.

Note: in more general cases, the expression $f(y)$ that appears in the differential equation might depend on t as well as y. For our purposes, we do not consider such examples in detail.

By sketching slopes at various values of y, *we obtain the slope field* through which we can get a reasonable idea of the behaviour of the solutions to the differential equation.

Example 13.5 *Consider the differential equation*

$$\frac{dy}{dt} = 2y. \tag{13.4}$$

Compute some of the slopes for various values of y and use this to sketch a slope field for this differential equation.

Mastered Material Check

14. Solve Differential Eqn. (13.4) analytically.

Solution. Equation (13.4) states that if a solution curve passes through a point (t,y), then its tangent line at that point has a slope $2y$, regardless of the value of t. This example is simple enough that we can state the following: for positive values of y, the slope is positive; for negative values of y, the slope is negative; and for $y = 0$, the slope is zero.

We provide some tabulated values of y indicating the values of the slope $f(y)$, its sign, and what this implies about the local behaviour of the solution and its direction. Then, in Figure 13.1 we combine this information to

y	$f(y)$	slope of tangent line	behaviour of y	direction of arrow
-2	-4	-ve	decreasing	↘
-1	-2	-ve	decreasing	↘
0	0	0	no change	→
1	2	+ve	increasing	↗
2	4	+ve	increasing	↗

Table 13.1: Table for the slope field diagram of differential equation (13.4), $\frac{dy}{dt} = 2y$, described in Example 13.5.

generate the direction field and the corresponding solution curves. Note that the direction of the arrows (rather than their absolute magnitude) provides the most important qualitative tendency for the slope field sketch. ◊

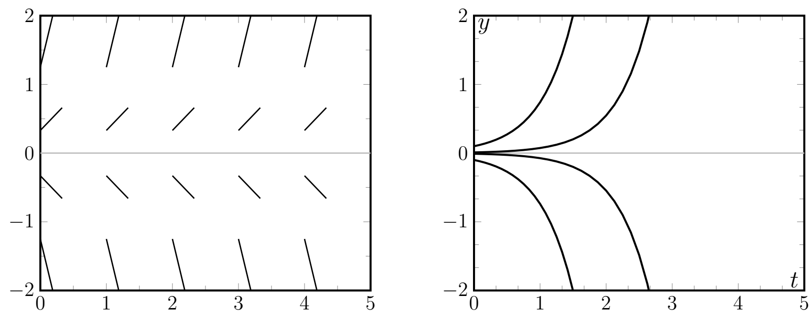


Figure 13.1: Direction field and solution curves for differential equation, $\frac{dy}{dt} = 2y$ described in Example 13.5.

In constructing the slope field and solution curves, the following basic rules should be followed:

1. By convention, time flows from left to right along the t axis in our graphs, so the direction of all arrows (not usually indicated explicitly on the slope field) is always from left to right.
2. According to the differential equation, for any given value of the variable y, the slope is given by the expression $f(y)$ in the differential equation. The sign of that quantity is particularly important in determining whether the solution is locally increasing, decreasing, or neither. In the tables, we indicate this in the last column with the notation $\nearrow$, $\searrow$, or $\rightarrow$.
3. There is a *single* arrow at any point in the ty-plane, and consequently solution curves cannot intersect anywhere (although they can get arbitrarily close to one another).

We see some implications of these rules in our examples.

Example 13.6 *Consider the differential equation*

$$\frac{dy}{dt} = f(y) = y - y^3. \tag{13.5}$$

Create a slope field diagram for this differential equation.

A summary of steps in creating the slope field for Example 13.6.

Solution. Based on the last example, we focus on the sign, rather than the value of the derivative $f(y)$, since that sign determines whether the solutions increase, decrease, or stay constant. Recall that factoring helps to find zeros, and to identify where an expression changes sign. For example,

$$\frac{dy}{dt} = f(y) = y - y^3 = y(1-y^2) = y(1+y)(1-y).$$

The sign of f depends on the signs of the factors $y, (1+y), (1-y)$. For $y < -1$, two factors, $y, (1+y)$, are negative, whereas $(1-y)$ is positive, so that the product is positive overall. The sign of $f(y)$ changes at each of the three points $y = 0, \pm 1$ where one or another of the three factors changes sign,

Mastered Material Check

15. Graph the function $f(y) = y(1+y)(1-y)$ and indiate where it changes sign.
16. Repeat the process for the function $f(y) = y^2(1+y)^2(1-y)$.

as shown in Table 13.2. Eventually, to the right of all three (when $y > 1$), the sign is negative. We summarize these observations in Table 13.2 and show the slopes field and solution curves in Figure 13.2. ◊

y	sign of $f(y)$	behaviour of y	direction of arrow
$y < -1$	+ve	increasing	↗
-1	0	no change	→
-0.5	-ve	decreasing	↘
0	0	no change	→
0.5	+ve	increasing	↗
1	0	no change	→
$y > 1$	-ve	decreasing	↘

Table 13.2: Table for the slope field diagram of the DE (13.5) described in Example 13.6.

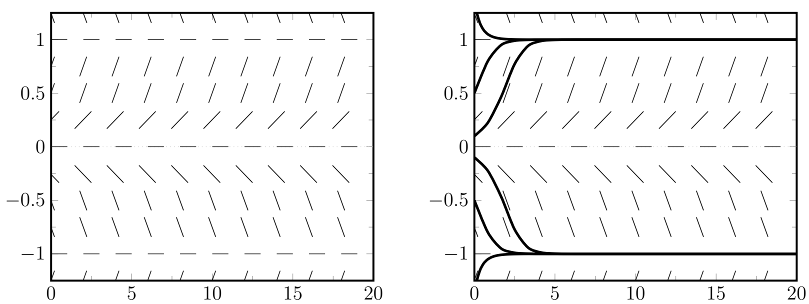


Figure 13.2: Direction field and solution curves for differential equation (13.5) described in Example 13.6.

Example 13.7 *Sketch a slope field and solution curves for the problem of a cooling object, and specifically for*

$$\frac{dT}{dt} = f(T) = 0.2(10 - T). \quad (13.6)$$

Solution. The family of curves shown in Figure 13.3 (also Figure 12.6) are solutions to (13.6). The function $f(T) = 0.2(10 - T)$ corresponds to the slopes of tangent lines to these curves. We indicate the sign of $f(T)$ and thereby the behaviour of $T(t)$ in Table 13.3. Note that there is only one

T	sign of $f(T)$	behaviour of T	direction of arrow
$T < 10$	+ve	increasing	↗
$T = 10$	0	no change	→
$T > 10$	-ve	decreasing	↘

Table 13.3: Table for the slope field diagram of $\frac{dT}{dt} = 0.2(10 - T)$ described in Example 13.7.

change of sign, at $T = 10$. For smaller T, the solution is always increasing and for larger T, the solution is always decreasing. The slope field and solution curves are shown in Figure 13.3. In the slope field, one particular value of t is coloured to emphasize the associated changes in T, as in Table 13.3. ◊

Mastered Material Check

17. Indicate the regions Figure 13.3 where T is increasing.
18. Where is T not changing in Figure 13.3?

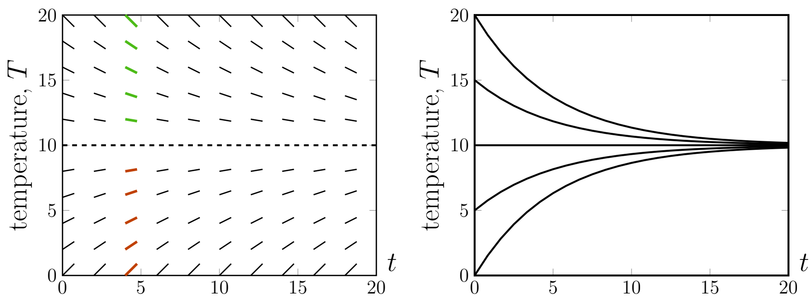


Figure 13.3: Slope field and solution curves for a cooling object that satisfies the differential equation (13.6) in Example 13.7.

We observe an agreement between the detailed solutions found analytically (Example 12.5), found using Euler's method (Example 12.13), and those sketched using the new qualitative arguments (Example 13.7).

State-space diagrams

In Examples 13.5-13.7, we saw that we can understand qualitative features of solutions to the differential equation

$$\frac{dy}{dt} = f(y), \tag{13.7}$$

by examining the expression $f(y)$. We used the sign of $f(y)$ to assemble a slope field diagram and sketch solution curves. The slope field informed us about which initial values of y would increase, decrease or stay constant. We next show another way of determining the same information.

First, let us define a **state space**, also called **phase line**, which is essentially the y-axis with superimposed arrows representing the direction of flow.

Definition 13.2 (State space (or phase line)) *A line representing the dependent variable (y) together with arrows to describe the flow along that line (increasing, decreasing, or stationary y) satisfying Eqn.* (13.7) *is called the* **state space** *diagram or the* **phase line** *diagram for the differential equation.*

Rather than tabulating signs for $f(y)$, we can arrive at similar conclusions by sketching $f(y)$ and observing where this function is positive (implying that y increases) or negative (y decreases). Places where $f(y) = 0$ ("zeros of f") are important since these represent **steady states** ("static solutions", where there is no change in y). Along the y axis (which is now on the horizontal axis of the sketch) increasing y means motion to the right, decreasing y means motion to the left.

As we shall see, the information contained in this type of diagram provides a qualitative description of solutions to the differential equation, but with the explicit time behaviour suppressed. This is illustrated by Figure 13.4, where we show the connection between the *slope field diagram* and the *state space diagram* for a typical differential equation.

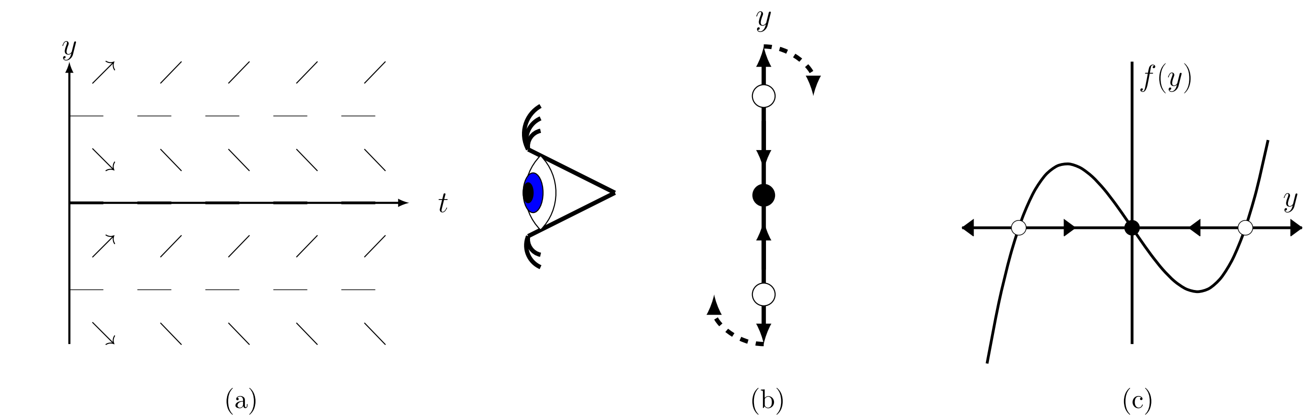


Figure 13.4: The relationship of the slope field and state space diagrams. (a) A typical slope field. A few arrows have been added to indicate the direction of time flow along the tangent vectors. Now consider "looking down the time axis" as shown by the "eye" in this diagram. Then the t axis points towards us, and we see only the y-axis as in (b). Arrows on the y-axis indicate the directions of flow for various values of y as determined in (a). Now "rotate" the y axis so it is horizontal, as shown in (c). The direction of the arrows exactly correspond to places where $f(y)$, in (c), is *positive* (which implies increasing y, $\rightarrow$), or *negative* (which implies decreasing y, $\leftarrow$). The state space diagram is the y-axis in (b) or (c).

Example 13.8 *Consider the differential equation*

$$\frac{dy}{dt} = f(y) = y - y^3. \tag{13.8}$$

Sketch $f(y)$ versus y and use your sketch to determine where y is static, and where y increases or decreases. Then describe what this predicts starting from each of the three initial conditions:

(i) $y(0) = -0.5$,

(ii) $y(0) = 0.3$, *or*

(iii) $y(0) = 2$.

Solution. From Example 13.6, we know that $f(y) = 0$ at $y = -1, 0, 1$. This means that y does not change at these steady state values, so, if we start a system off with $y(0) = 0$, or $y(0) = \pm 1$, the value of y is static. The three places at which this happens are marked by heavy dots in Figure 13.5(a).

Video explanation of the steps in the solution to Example 13.8.

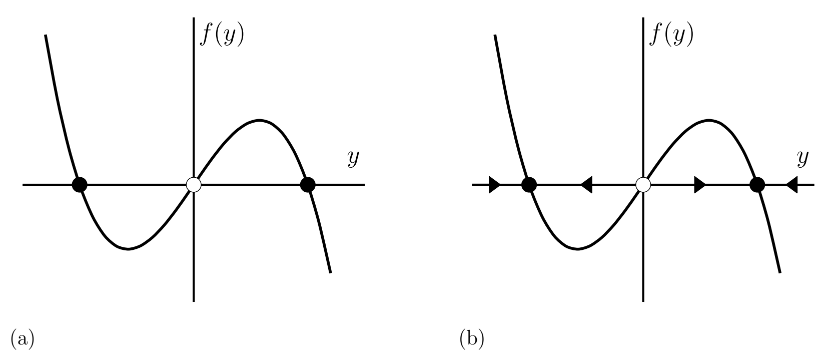


Figure 13.5: Steady states (dots) and intervals for which y increases or decreases for the differential equation (13.8). See Example 13.8.

We also see that $f(y) < 0$ for $-1 < y < 0$ and for $y > 1$. In these intervals, $y(t)$ must be a decreasing function of time ($dy/dt < 0$). On the other hand, for

$0 < y < 1$ or for $y < -1$, we have $f(y) > 0$, so $y(t)$ is increasing. See arrows on Figure 13.5(b). We see from this figure that there is a tendency for y to move away from the steady state value $y = 0$ and to approach either of the steady states at 1 or -1. Starting from the initial values given above, we have

(i) $y(0) = -0.5$ results in $y \to -1$,

(ii) $y(0) = 0.3$ leads to $y \to 1$, and

(iii) $y(0) = 2$ implies $y \to \infty$. ◊

Example 13.9 (A cooling object) *Sketch the same type of diagram for the problem of a cooling object and interpret its meaning.*

Solution. Here, the differential equation is

$$\frac{dT}{dt} = f(T) = 0.2(10 - T). \tag{13.9}$$

A sketch of the rate of change, $f(T)$ versus the temperature T is shown in Figure 13.6. We deduce the direction of the flow directly form this sketch. ◊

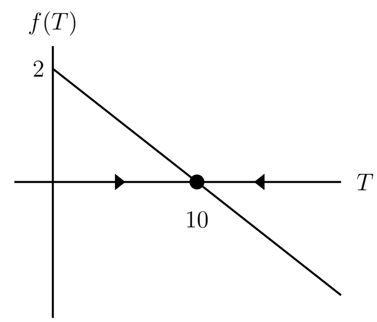


Figure 13.6: Figure for Example 13.9, the differential equation (13.9).

Mastered Material Check

19. In Figures 13.6 and 13.7, where is the function positive?
20. Consider Eqn. (13.10) analytically: what value does y approach?

Example 13.10 *Create a similar qualitative sketch for the more general form of linear differential equation*

$$\frac{dy}{dt} = f(y) = a - by. \tag{13.10}$$

For what values of y would there be no change?

Solution. The rate of change of y is given by the function $f(y) = a - by$. This is shown in Figure 13.7. The steady state at which $f(y) = 0$ is at $y = a/b$. Starting from an initial condition $y(0) = a/b$, there would be no change. We also see from this figure that y approaches this value over time. After a long time, the value of y will be approximately a/b. ◊

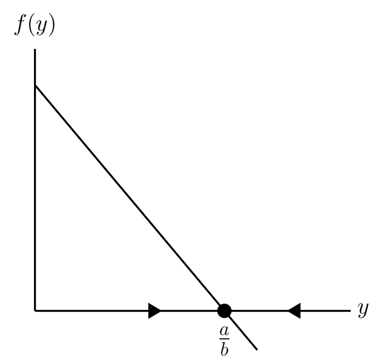


Figure 13.7: Qualitative sketch for Eqn. (13.10) in Example 13.10.

Steady states and stability

From the last few figures, we observe that wherever the function f on the right hand side of the differential equations crosses the horizontal axis (satisfies $f = 0$) there is a steady state. For example, in Figure 13.6 this takes place at $T = 10$. At that temperature the differential equation specifies that $dT/dt = 0$ and so, $T = 10$ is a steady state, a concept we first encountered in Chapter 12.

Definition 13.3 (Steady state) *A steady state is a state in which a system is not changing.*

Example 13.11 *Identify steady states of Eqn. (13.8),*

$$\frac{dy}{dt} = y^3 - y.$$

Solution. Steady states are points that satisfy $f(y) = 0$. We already found those to be $y = 0$ and $y = \pm 1$ in Example 13.8. ◊

From Figure 13.5, we see that solutions starting *close to* $y = 1$ tend to get closer and closer to this value. We refer to this behaviour as **stability** of the steady state.

Definition 13.4 (Stability) *We say that a steady state is* **stable** *if states that are initially close enough to that steady state will get closer to it with time. We say that a steady state is* **unstable***, if states that are initially very close to it eventually move away from that steady state.*

Example 13.12 *Determine the stability of steady states of Eqn.* (13.8)*:*

$$\frac{dy}{dt} = y - y^3.$$

Mastered Material Check

21. In the state space diagram in Figure 13.4, identify the stable steady states.

Solution. From any starting value of $y > 0$ in this example, we see that *after a long time*, the solution curves tend to approach the value $y = 1$. States close to $y = 1$ get closer to it, so this is a stable steady state. For the steady state $y = 0$, we see that initial conditions near $y = 0$ move away over time. Thus, this steady state is unstable. Similarly, the steady state at $y = -1$ is stable. In Figure 13.5 we show the stable steady states with black dots and the unstable steady state with an open dot. ◊

13.3 Applying qualitative analysis to biological models

Section 13.3 Learning goals

1. Practice the techniques of slope field, state-space diagram, and steady state analysis on the logistic equation.
2. Explain the derivation of a model for interacting (healthy, infected) individuals based on a set of assumptions.
3. Identify that the resulting set of two ODEs can be reduced to a single ODE. Use qualitative methods to analyse the model behaviour and to interpret the results.

The qualitative ideas developed so far will now be applied to to problems from biology. In the following sections we first use these methods to obtain a thorough understanding of **logistic population growth**. We then derive a model for the spread of a disease, and use qualitative arguments to analyze the predictions of that differential equation model.

Qualitative analysis of the logistic equation

We apply the new methods to the logistic equation.

Example 13.13 *Find the steady states of the logistic equation, Eqn. (13.1):*

$$\frac{dN}{dt} = rN\frac{(K-N)}{K}.$$

Solution. To determine the steady states of Eqn. (13.1), i.e. the level of population that would not change over time, we look for values of N such that

$$\frac{dN}{dt} = 0.$$

This leads to

$$rN\frac{(K-N)}{K} = 0,$$

which has solutions $N = 0$ (no population at all) or $N = K$ (the population is at its carrying capacity). ♢

We could similarly find steady states of the scaled form of the logistic equation, Eqn. (13.3). Setting $dy/dt = 0$ leads to

$$0 = \frac{dy}{dt} = ry(1-y) \quad \Rightarrow \quad y = 0, \text{ or } y = 1.$$

This comes as no surprise since these values of y correspond to the values $N = 0$ and $N = K$.

The scaled logistic equation, its slope field, and steady state values are discussed here.

A second way to analyze the scaled logistic equation, using the phase line approach, and its connection to the slope field method as described in Example 13.14.

Example 13.14 *Draw a plot of the rate of change dy/dt versus the value of y for the scaled logistic equation,Eqn. (13.3):*

$$\frac{dy}{dt} = ry(1-y).$$

Solution. In the plot of Figure 13.8 only $y \geq 0$ is relevant. In the interval $0 < y < 1$, the rate of change is positive, so that y increases, whereas for $y > 1$, the rate of change is negative, so y decreases. Since y refers to population size, we need not concern ourselves with behaviour for $y < 0$.

From Figure 13.8 we deduce that solutions that start with a positive y value approach $y = 1$ with time. Solutions starting at either steady state $y = 0$ or $y = 1$ would not change. Restated in terms of the variable $N(t)$, any initial population should approach its carrying capacity K with time. ♢

We now look at the same equation from the perspective of the slope field.

Mastered Material Check

22. Circle the steady states in Figure 13.8 and identify which one is stable.
23. Why is $y < 0$ not relevant in Example 13.14?

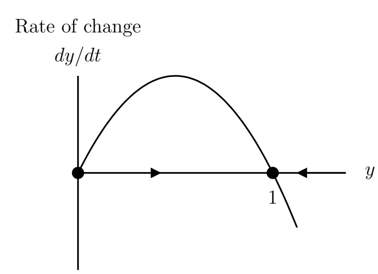


Figure 13.8: Plot of dy/dt versus y for the the scaled logistic equation (13.3).

Example 13.15 *Draw a slope field for the scaled logistic equation with $r = 0.5$, that is for*

$$\frac{dy}{dt} = f(y) = 0.5 \cdot y(1-y). \tag{13.11}$$

Solution. We generate slopes for various values of y in Table 13.4 and plot the slope field in Figure 13.9(a). ♢

Finally, we practice Euler's method to graph the numerical solution to Eqn. (13.11) from several initial conditions.

Example 13.16 (Numerical solutions to the logistic equation) *Use Euler's method to approximate the solutions to the logistic equation (13.11).*

y	**sign of $f(y)$**	**behaviour of y**	**direction of arrow**
0	0	no change	$\rightarrow$
$0 < y < 1$	+ve	increasing	$\nearrow$
1	0	no change	$\rightarrow$
$y > 1$	-ve	decreasing	$\searrow$

Table 13.4: Table for slope field for the logistic equation (13.11). See Fig 13.9(a) for the resulting diagram.

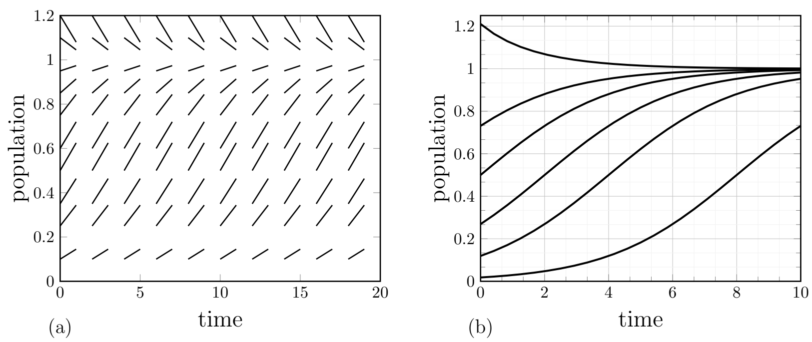


Figure 13.9: (a) Slope field and (b) solution curves for the logistic equation (13.11), $\frac{dy}{dt} = 0.5 \cdot y(1-y)$

Solution. In Figure 13.9(b) we show a set of solution curves, obtained by solving the equation numerically using Euler's method. To obtain these solutions, a value of $h = \Delta t = 0.1$ was used. The solution is plotted for various initial conditions $y(0) = y_0$. The successive values of y were calculated according to

$$y_{k+1} = y_k + 0.5 y_k (1 - y_k) h, \qquad k = 0, \dots 100.$$

From Figure 13.9(b), we see that solution curves approach the steady state $y = 1$, meaning that the population $N(t)$ approaches the carrying capacity K for all positive starting values. A link to the spreadsheet that implements Euler's method is included. $\diamond$

Mastered Material Check

24. What initial values y_0 were used in drawing the different solution curves depicted in Figure 13.9(b)?

Link to Google Sheets. This spreadsheet implements Euler's method for Example 13.16. A chart showing solutions from four initial conditions is included.

Example 13.17 (Inflection points) *Some of the curves shown in Figure 13.9(b) have an inflection point, but others do not. Use the differential equation to determine which of the solution curves have an inflection point.*

Solution. We have already established that all initial values in the range $0 < y_0 < 1$ are associated with increasing solutions $y(t)$. Now we consider the concavity of those solutions. The logistic equation has the form

$$\frac{dy}{dt} = ry(1-y) = ry - ry^2$$

Differentiate both sides using the chain rule and factor, to get

$$\frac{d^2y}{dt^2} = r\frac{dy}{dt} - 2ry\frac{dy}{dt} = r\frac{dy}{dt}(1-2y).$$

Mastered Material Check

25. How do we know that initial conditions in the range $0 < y_0 < 1$ lead to increasing solutions?

An inflection point would occur at places where the second derivative changes sign. This is possible for $dy/dt = 0$ or for $(1-2y) = 0$. We have already dismissed the first possibility because we argued that the rate of change is nonzero in the interval of interest. Thus we conclude that an inflection point would occur whenever $y = 1/2$. Any initial condition satisfying $0 < y_0 < 1/2$ would eventually pass through $y = 1/2$ on its way to the steady state level at $y = 1$, and in so doing, would have an inflection point. $\diamond$

A changing aphid population

In Chapters 1 and 5, we investigated a situation when predation and growth rates of an aphid population exactly balanced. But what happens if these two rates do not balance? We are now ready to tackle this question.

Featured Problem 13.1 (aphids) *Consider the aphid-ladybug problem (Example 1.3) with aphid density x, growth rate $G(x) = rx$, and predation rate by a ladybug $P(x)$ as in* (1.10). *(a) Write down a differential equation for the aphid population. (b) Use your equation, and a sketch of the two functions to answer the following question: What happens to the aphid population starting from various initial population sizes?*

Hint: Growth rate (number of aphids born per unit time) contributes positively, whereas predation rate (number of aphids eaten per unit time) contributes negatively to the rate of change of aphids with respect to time (dx/dt).

13.3.1 The radius of a growing cell

In Section 11.4 we examined a cell in which nutrient absorption and consumption each contribute to changing the mass balance of the cell. We first wrote down a differential equation of the form

$$\frac{dm}{dt} = A - C.$$

Assuming the cell was spherical, we showed that this equation results in the differential equation for the cell radius $r(t)$:

$$\frac{dr}{dt} = \frac{1}{\rho}\left(k_1 - \frac{k_2}{3} r\right), \quad k_1, k_2, \rho > 0 \tag{13.12}$$

Using tools in this chapter, we can now understand what this implies about cell size growth.

Featured Problem 13.2 (How cell radius changes) *Apply qualitative methods to Eqn.* (13.12) *so as to determine what happens to cells starting from various initial sizes. Is there a steady state cell size? How do your results compare to our findings in Section 1.2?*

A model for the spread of a disease

In the era of human immunodeficiency virus (HIV), Severe Acute Respiratory Syndrome (SARS), Avian influenza ("bird flu") and similar emerging infectious diseases, it is prudent to consider how infection spreads, and how it could be controlled or suppressed. This motivates the following example.

For a given disease, let us subdivide the population into two classes: healthy individuals who are susceptible to catching the infection, and those that are currently infected and able to transmit the infection to others. We consider an infection that is mild enough that individuals recover at some constant rate, and that they become susceptible once recovered.

Note: usually, recovery from an illness leads to partial temporary immunity. While this, too, can be modelled, we restrict attention to the simpler case which is tractable using mathematics we have just introduced.

The simplest case to understand is that of a fixed population (with no birth, death or migration during the timescale of interest). A goal is to predict whether the infection spreads and persists (becomes endemic) in the population or whether it runs its course and disappear. We use the following notation:

A video summary of the model for the spread of a disease, together with its analysis.

$$\begin{aligned} S(t) &= \text{size of population of susceptible (healthy) individuals,} \\ I(t) &= \text{size of population of infected individuals,} \\ N(t) = S(t) + I(t) &= \text{total population size.} \end{aligned}$$

We add a few simplifying assumptions.

1. The population mixes very well, so each individual is equally likely to contact and interact with any other individual. The contact is random.

2. Other than the state (S or I), individuals are “identical,” with the same rates of recovery and infectivity.

3. On the timescale of interest, there is no birth, death or migration, only exchange between S and I.

Example 13.18 *Suppose that the process can be represented by the scheme*

$$\begin{aligned} S + I &\to I + I, \\ I &\to S \end{aligned}$$

The first part, transmission of disease from I to S involves interaction. The second part is recovery. Use the assumptions above to track the two populations and to formulate a set of differential equations for $I(t)$ and $S(t)$.

Solution. The following balance equations keeps track of individuals

$$\begin{bmatrix} \text{Rate of} \\ \text{change of} \\ I(t) \end{bmatrix} = \begin{bmatrix} \text{Rate of gain} \\ \text{due to disease} \\ \text{transmission} \end{bmatrix} - \begin{bmatrix} \text{Rate of loss} \\ \text{due to} \\ \text{recovery} \end{bmatrix}$$

According to our assumption, recovery takes place at a constant rate per unit time, denoted by $\mu > 0$. By the law of mass action, the disease transmission rate should be proportional to the product of the populations, $(S \cdot I)$. Assigning $\beta > 0$ to be the constant of proportionality leads to the following differential equations for the infected population:

$$\frac{dI}{dt} = \beta SI - \mu I.$$

Similarly, we can write a balance equation that tracks the population of susceptible individuals:

$$\begin{bmatrix} \text{Rate of} \\ \text{change of} \\ S(t) \end{bmatrix} = - \begin{bmatrix} \text{Rate of Loss} \\ \text{due to disease} \\ \text{transmission} \end{bmatrix} + \begin{bmatrix} \text{Rate of gain} \\ \text{due to} \\ \text{recovery} \end{bmatrix}$$

Observe that loss from one group leads to (exactly balanced) gain in the other group. By similar logic, the differential equation for $S(t)$ is then

$$\frac{dS}{dt} = -\beta SI + \mu I.$$

We have arrived at a **system of equations** that describe the changes in each of the groups,

$$\frac{dI}{dt} = \beta SI - \mu I, \quad (13.13a)$$
$$\frac{dS}{dt} = -\beta SI + \mu I. \quad (13.13b)$$

◊

Mastered Material Check

26. Identify any constants in Eqns. (13.13)(a) and (b).
27. What are the units of those constants?
28. Why does the hint given in Example 13.19 help?

From Eqns. (13.13) it is clear that changes in one population depend on both, which means that the differential equations are **coupled** (linked to one another). Hence, we cannot "solve one" independently of the other. We must treat them as a pair. However, as we observe in the next examples, we can simplify this system of equations using the fact that the total population does not change.

Example 13.19 *Use Eqns.(13.13) to show that the total population does not change (*hint: *show that the derivative of $S(t) + I(t)$ is zero).*

Video showing that the population $N(t) = I(t) + S(t)$ is constant.

Solution. Add the equations to one another. Then we obtain

$$\frac{d}{dt}[I(t) + S(t)] = \frac{dI}{dt} + \frac{dS}{dt} = \beta SI - \mu I - \beta SI + \mu I = 0.$$

Hence

$$\frac{d}{dt}[I(t) + S(t)] = \frac{dN}{dt} = 0,$$

which mean that $N(t) = [I(t) + S(t)] = N$=constant, so the total population does not change. (In Eqn. (13.1), here N is a constant and $I(t), S(t)$ are the variables.) ◊

Example 13.20 *Use the fact that N is constant to express $S(t)$ in terms of $I(t)$ and N, and eliminate $S(t)$ from the differential equation for $I(t)$. Your equation should only contain the constants N, β, μ.*

Solution. Since $N = S(t) + I(t)$ is constant, we can write $S(t) = N - I(t)$. Then, plugging this into the differential equation for $I(t)$ we obtain

$$\frac{dI}{dt} = \beta SI - \mu I, \quad \Rightarrow \quad \frac{dI}{dt} = \beta(N-I)I - \mu I.$$

◊

Mastered Material Check

29. Redo Example 13.20 but eliminate $I(t)$ instead of $S(t)$.
30. Analyze the equation you get for $dS(t)/dt$ as done for dI/dt in Example 13.21.

Example 13.21 **a)** *Show that the above equation can be written in the form*

$$\frac{dI}{dt} = \beta I(K-I),$$

where K is a constant.

b) *Determine how this constant K depends on N, β, and μ.*

c) *Is the constant K positive or negative?*

Solution.

a) We rewrite the differential equation for $I(t)$ as follows:

$$\frac{dI}{dt} = \beta(N-I)I - \mu I = \beta I\left((N-I) - \frac{\mu}{\beta}\right) = \beta I\left(N - \frac{\mu}{\beta} - I\right).$$

b) We identify the constant,

$$K = \left(N - \frac{\mu}{\beta}\right).$$

c) Evidently, K could be *either positive or negative*, that is

$$\begin{cases} N \geq \frac{\mu}{\beta} & \Rightarrow K \geq 0, \\ N < \frac{\mu}{\beta} & \Rightarrow K < 0. \end{cases}$$

◊

Using the above process, we have reduced the system of two differential equations for the two variables $I(t)$, $S(t)$ to a *single* differential equation for $I(t)$, together with the statement $S(t) = N - I(t)$. We now examine implications of this result using the qualitative methods of this chapter.

Example 13.22 *Consider the differential equation for $I(t)$ given by*

$$\frac{dI}{dt} \equiv f(I) = \beta I(K-I), \quad \textit{where} \quad K = \left(N - \frac{\mu}{\beta}\right). \tag{13.14}$$

Find the steady states of the differential equation (13.14) and draw a state space diagram in each of the following cases:

(a) $K \geq 0$,

(b) $K < 0$.

Use your diagram to determine which steady state(s) are stable or unstable.

Solution. Steady states of Eqn. (13.14) satisfy $dI/dt = \beta I(K-I) = 0$. Hence, these steady states are $I = 0$ (no infected individuals) and $I = K$. The latter only makes sense if $K \geq 0$. We plot the function $f(I) = \beta I(K-I)$ in Eqn. (13.14) against the state variable I in Figure 13.10 (a) for $K \geq 0$ and (b) for $K < 0$. Since $f(I)$ is quadratic in I, its graph is a parabola and it opens downwards. We add arrows pointing right ($\rightarrow$) in the regions where $dI/dt > 0$ and arrows pointing left ($\leftarrow$) where $dI/dt < 0$.

Mastered Material Check

31. What is the significance of the grey shaded regions in Fig. 13.10.
32. Draw Fig. 13.10 for $K = 0$.
33. Why is $I = K$ not an admissible steady state if $K < 0$?

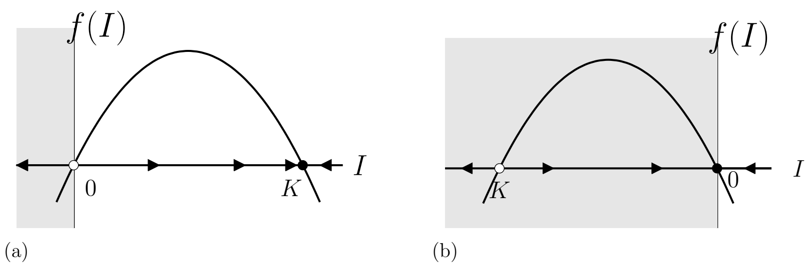


Figure 13.10: State-space diagrams for differential equation (13.14). Plots of $f(I)$ as a function of I in the cases (a) $K \geq 0$, and (b) $K < 0$. The grey regions are not biologically meaningful since I cannot be negative.

In case (a), when $K \geq 0$, we find that arrows point toward $I = K$, so this steady state is stable. Arrows point away from $I = 0$, so this represents an unstable steady state. In case (b), while we still have a parabolic graph with two steady states, the state $I = K$ is not admissible since K is negative. Hence only one steady state, at $I = 0$ is relevant biologically, and all initial conditions move towards this state. ◊

Example 13.23 *Interpret the results of the model in terms of the disease, assuming that initially most of the population is in the susceptible S group, and a small number of infected individuals are present at t* $= 0$.

Solution. In case (a), as long as the initial size of the infected group is positive ($I > 0$), with time it approaches K, that is, $I(t) \rightarrow K = N - \mu/\beta$. The rest of the population is in the susceptible group, that is $S(t) \rightarrow \mu/\beta$ (so that $S(t) + I(t) = N$ is always constant.) This first scenario holds provided $K > 0$ which is equivalent to $N > \mu/\beta$. There are then some infected and some healthy individuals in the population indefinitely, according to the model. In this case, we say that the disease becomes **endemic**.

In case (b), which corresponds to $N < \mu/\beta$, we see that $I(t) \rightarrow 0$ regardless of the initial size of the infected group. In that case, $S(t) \rightarrow N$ so with time, the infected group shrinks and the healthy group grows so that the whole population becomes healthy. From these two results, we conclude that the disease is wiped out in a small population, whereas in a sufficiently large population, it can spread until a steady state is attained where some fraction of the population is always infected. In fact we have identified a *threshold* that separates these two behaviours:

Mastered Material Check

34. In the case that $\beta = 0.001$per person per day and $\mu = 0.1$ per day, how large would the population have to be for the disease to become endemic?
35. Frequent hand-washing can be a protective measure that decreases the spread of disease. Which parameter of the model would this affect and in what way?

$$\frac{N\beta}{\mu} > 1 \quad \Rightarrow \quad \text{disease becomes endemic,}$$
$$\frac{N\beta}{\mu} < 1 \quad \Rightarrow \quad \text{disease is wiped out.}$$

The ratio of constants in these inequalities, $R_0 = N\beta/\mu$ is called the **basic reproduction number** for the disease. Many current and much more detailed models for disease transmission also have such threshold behaviour, and the ratio that determines whether the disease spreads or disappears, R_0 is of great interest in vaccination strategies. This ratio represents the number of infections that arise when 1 infected individual interacts with a population of N susceptible individuals.

A video summarizing the interpretation of the model and the meaning of the constant $R_0 = N\beta/\mu$.

13.4 Summary

1. A differential equation of the form $\alpha\frac{dy}{dt} + \beta y + \gamma = 0$ is linear (and "first order"). We encountered several examples of nonlinear DEs in this chapter.
2. A (possibly nonlinear) differential equation $\frac{dy}{dt} = f(y)$ can be analyzed qualitatively by observing where $f(y)$ is positive, negative or zero.
3. A slope field (or "direction field") is a collection of tangent vectors for solutions to a differential equation. Slope fields can be sketched from $f(y)$ without the need to solve the differential equation.
4. A solution curve drawn in a slope field corresponds to a single solution to a differential equation, with some initial y_0 value given.
5. A state space (or "phase line" diagram) for the differential equation is a y axis, together with arrows describing the flow (increasing/decreasing/stationary) along that axis. It can be obtained from a sketch of $f(y)$.
6. A steady state is stable if nearby states get closer. A steady state is unstable if nearby states get further away with time.
7. Creating/interpreting slope field and state space diagrams is helpful in understanding the behavior of solutions to differential equations.
8. Applications considered in this chapter included:

   (a) the logistic equations for population growth (a nonlinear differential equation, scaling, steady state and slope field demonstration);
   (b) the Law of Mass Action (a nonlinear differential equation);
   (c) a cooling object (state space and phase line diagram demonstration); and
   (d) disease spread model (an extensive exposition on qualitative differential equation methods).

Quick Concept Checks

1. Why is it helpful to rescale an equation?

2. Identify which of the following differential equations are linear:

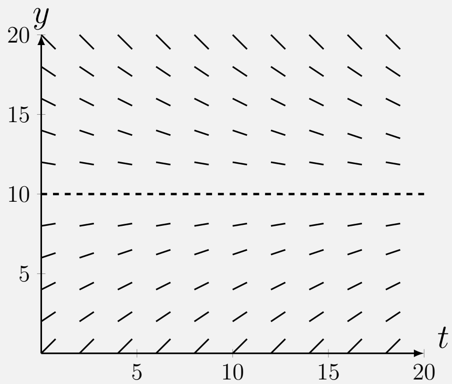
   (a) $5\frac{dy}{dt} - y = -0.5$

   (b) $\left(\frac{dy}{dt}\right)^2 + y + 1 = 0$

   (c) $\frac{dy}{dx} + \pi y + \rho = 3$

   (d) $\frac{dx}{dt} + x + 2 = -3x$

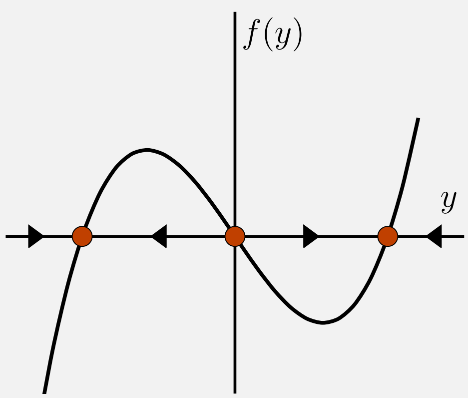
3. Consider the following slope field:



   (a) Where is y decreasing?

   (b) What is y approaching?

4. Circle the **stable** steady states in the following state space diagram



Exercises

13.1. **Explaining connections.** Explain the connection between Eqn. (13.2) and the equations Eqn. 11.2 and Eqn. (12.4).

13.2. **Slope fields.** Consider the differential equations given below. In each case, draw a slope field, determine the values of y for which no change takes place - such values are called steady states - and use your slope field to predict what would happen starting from an initial value $y(0) = 1$.

(a) $\dfrac{dy}{dt} = -0.5y$

(b) $\dfrac{dy}{dt} = 0.5y(2-y)$

(c) $\dfrac{dy}{dt} = y(2-y)(3-y)$

13.3. **Drawing slope fields.** Draw a slope field for each of the given differential equations:

(a) $\frac{dy}{dt} = 2+3y$

(b) $\frac{dy}{dt} = -y(2-y)$

(c) $\frac{dy}{dt} = 2-3y+y^2$

(d) $\frac{dy}{dt} = -2(3-y)^2$

(e) $\frac{dy}{dt} = y^2-y+1$

(f) $\frac{dy}{dt} = y^3-y$

(g) $\frac{dy}{dt} = \sqrt{y}(y-2)(y-3)^2$, $y \geq 0$.

13.4. **Linear or Nonlinear.** Identify which of the differential equations in Exercise 2 and 3 is linear and which nonlinear.

13.5. **Using slope fields.** For each of the differential equations (a) to (g) in Exercise 3, plot $\frac{dy}{dt}$ as a function of y, draw the motion along the y-axis, identify the steady state(s) and indicate if the motions are toward or away from the steady state(s).

13.6. **Direction field.** The direction field shown in the figure below corresponds to which differential equation?

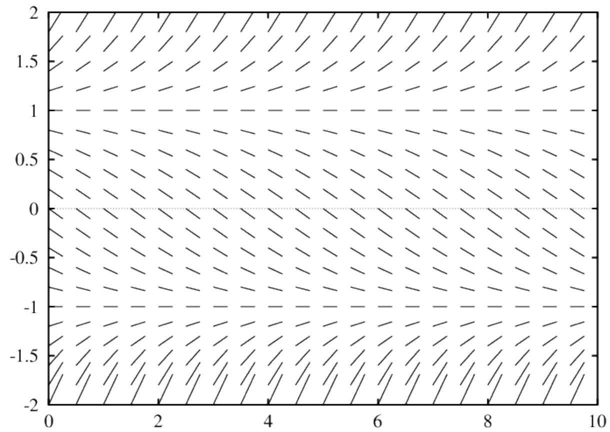
(A) $\frac{dy}{dt} = ry(y+1)$

(B) $\frac{dy}{dt} = r(y-1)(y+1)$

(C) $\frac{dy}{dt} = -r(y-1)(y+1)$

(D) $\frac{dy}{dt} = ry(y-1)$

(E) $\frac{dy}{dt} = -ry(y+1)$



13.7. **Differential equation.** Given the differential equation and initial condition

$$\frac{dy}{dt} = y^2(y-a), y(0) = 2a$$

where $a > 0$ is a constant, the value of the function $y(t)$ would

(A) approach $y = 0$;

(B) grow larger with time;

(C) approach $y = a$;

(D) stay the same;

(E) none of the above.

13.8. **There's a hole in the bucket.** Water flows into a bucket at constant rate I. There is a hole in the container. Explain the model

$$\frac{dh}{dt} = I - k\sqrt{h}.$$

Analyze the behaviour predicted. What would the height be after a long time? Is this result always valid, or is an additional assumption needed? (*hint:* recall Example 12.3.)

13.9. **Cubical crystal.** A crystal grows inside a medium in a cubical shape with side length x and volume V. The rate of change of the volume is given by

$$\frac{dV}{dt} = kx^2(V_0 - V)$$

where k and V_0 are positive constants.

(a) Rewrite this as a differential equation for $\frac{dx}{dt}$.

(b) Suppose that the crystal grows from a very small "seed." Show that its growth rate continually decreases.

(c) What happens to the size of the crystal after a very long time?

(d) What is its size (that is, what is either x or V) when it is growing at half its initial rate?

13.10. **The Law of Mass Action.** The Law of Mass Action in Section 13.1 led to the assumption that the rate of a reaction involving two types of molecules (A and B) is proportional to the product of their concentrations, $k \cdot a \cdot b$.

Explain why the sum of the concentrations, $k \cdot (a+b)$ would not make for a sensible assumption about the rate of the reaction.

13.11. **Biochemical reaction.** A biochemical reaction in which a substance S is both produced and consumed is investigated. The concentration $c(t)$ of S changes during the reaction, and is seen to follow the differential equation

$$\frac{dc}{dt} = K_{\max} \frac{c}{k+c} - rc$$

where $K_{\max}, k, r$ are positive constants with certain convenient units. The first term is a concentration-dependent production term and the second term represents consumption of the substance.

(a) What is the maximal rate at which the substance is produced? At what concentration is the production rate 50% of this maximal value?

(b) If the production is turned off, the substance decays. How long would it take for the concentration to drop by 50%?

(c) At what concentration does the production rate just balance the consumption rate?

13.12. **Logistic growth with proportional harvesting.** Consider a fish population of density $N(t)$ growing at rate $g(N)$, with harvesting, so that the population satisfies the differential equation

$$\frac{dN}{dt} = g(N) - h(N).$$

Now assume that the growth rate is logistic, so $g(N) = rN\frac{(K-N)}{K}$ where $r, K > 0$ are constant. Assume that the rate of harvesting is proportional to the population size, so that

$$h(N) = qEN$$

where E, the effort of the fishermen, and q, the catchability of this type of fish, are positive constants.

Use qualitative methods discussed in this chapter to analyze the behaviour of this equation. Under what conditions does this lead to a sustainable fishery?

13.13. **Logistic growth with constant number harvesting.** Consider the same fish population as in Exercise 12, but this time assume that the rate of harvesting is fixed, regardless of the population size, so that

$$h(N) = H$$

where H is a constant number of fish being caught and removed per unit time. Analyze this revised model and compare it to the previous results.

13.14. **Scaling time in the logistic equation.** Consider the scaled logistic equation (13.3). Recall that r has units of 1/time, so $1/r$ is a quantity with units of time. Now consider scaling the time variable in (13.3) by defining $t = s/r$. Then s carries no units (s is "dimensionless").

Substitute this expression for t in (13.3) and find the differential equation so obtained (for dy/ds).

13.15. **Euler's method applied to logistic growth.** Consider the logistic differential equation

$$\frac{dy}{dt} = ry(1-y).$$

Let $r = 1$. Use Euler's method to find a solution to this differential equation starting with $y(0) = 0.5$, and step size $h = 0.2$. Find the values of y up to time $t = 1.0$.

13.16. **Spread of infection.** In the model for the spread of a disease, we used the fact that the total population is constant ($S(t) + I(t) = N =$constant) to eliminate $S(t)$ and analyze a differential equation for $I(t)$ on its own.

Carry out a similar analysis, but eliminate $I(t)$. Then analyze the differential equation you get for $S(t)$ to find its steady states and behaviour, practicing the qualitative analysis discussed in this chapter.

13.17. **Vaccination strategy.** When an individual is vaccinated, he or she is "removed" from the susceptible population, effectively reducing the size of the population that can participate in the disease transmission. For example, if a fraction ϕ of the population is vaccinated, then only the remaining $(1-\phi)N$ individuals can be either susceptible or infected, so $S(t) + I(t) = (1-\phi)N$. When smallpox was an endemic disease, it had a basic reproductive number of $R_0 = 7$.

What fraction of the population would have had to be vaccinated to eradicate this disease?

13.18. **Social media.** Sally Sweetstone has invented a new social media App called HeadSpace, which instantly matches compatible mates according to their changing tastes and styles. Users hear about the App from one another by word of mouth and sign up for an account. The account expires randomly, with a half-life of 1 month. Suppose $y_1(t)$ are the number of individuals who are not subscribers and $y_2(t)$ are the number of are subscribers at time t. The following model has been suggested for the evolving subscriber population

$$\frac{dy_1}{dt} = by_2 - ay_1y_2,$$
$$\frac{dy_2}{dt} = ay_1y_2 - by_2.$$

(a) Explain the terms in the equation. What is the value of the constant b?

(b) Show that the total population $P = y_1(t) + y_2(t)$ is constant.

Note: this is a **conservation statement**.

(c) Use the conservation statement to eliminate y_1. Then analyze the differential equation you obtain for y_2.

(d) Use your model to determine whether this newly launched social media will be successful or whether it will go extinct.

13.19. **A bimolecular reaction.** Two molecules of A can react to form a new chemical, B. The reaction is **reversible** so that B also continually decays back into 2 molecules of A. The differential equation model proposed for this system is

$$\frac{da}{dt} = -\mu a^2 + 2\beta b$$
$$\frac{db}{dt} = \frac{\mu}{2}a^2 - \beta b,$$

where $a(t), b(t) > 0$ are the concentrations of the two chemicals.

(a) Explain the factor 2 that appears in the differential equations and the conservation statement. Show that the total mass $M = a(t) + 2b(t)$ is constant.

(b) Use the techniques in this chapter to investigate what happens in this chemical reaction, to find any steady states, and to explain the behaviour of the system

14
Periodic and trigonometric functions

Nature abounds with examples of cyclic processes. Perhaps most familiar is the continually repeating heartbeat that accompanies us through life. Electrically active muscles power the heart. That electrical activity can be recorded on the surface of the body by an electrocardiogram (ECG), as shown in Figure 14.1. In a normal healthy human at rest, the electrical activity pattern associated with a complete cycle repeats itself with roughly 1 heartbeat per second. At exercise, the beating heart pumps faster, so the pattern repeats more frequently.

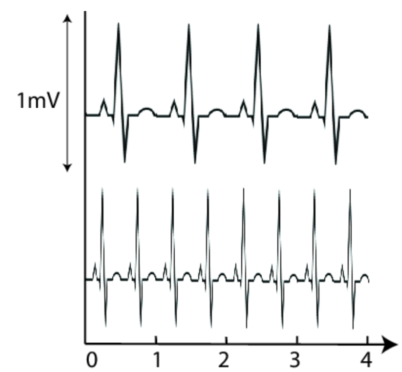


Figure 14.1: An electrocardiogram (ECG) shows the pattern of electrical activity of the heart. Vertical axis: millivolts, horizontal axis: time in seconds. At rest (top), the heart beats approximately 60 times per minute, but in exercise (lower trace), the heart rate increases.

To study the behaviour shown in Figure 14.1, we must first develop language that describes such periodic phenomena. For example, we need to quantify what is meant by “more frequent repetition" of a heart beat, “skipping a beat,” or other shifts in the pattern of this or of any other cycling system.

Before trying to understand intricate examples such as ECG’s, we begin with simple prototypes of periodic functions: the **trigonometric functions**, **sine** and **cosine**. Belonging to a wider class of periodic functions, these cases illustrate ideas of **amplitude**, **frequency**, **period**, and **phase**. Many cyclic phenomena can be described approximately by suitably adjusting such basic functions. Our study in this elementary context then aids in the goal of analyzing periodic functions in general - and periodicity of the ECG’s in particular (we return to this in Example 14.2).

As a second theme, we return to inverse functions and show that restrictions must be applied to ensure the existence of an inverse, particularly for trigonometric functions. Then, in Chapter 15, we calculate the derivatives of trigonometric functions and explore applications to rates of change of periodic phenomena or changing angles.

14.1 Basic trigonometry

Section 14.1 Learning goals

1. Define **radian** as a measure for angles.
2. Describe the correspondence between a point moving on a unit circle and the sine and cosine of the angle it forms at the origin.
3. Make correspondence between ratios of sides of a Pythagorean triangle and the trigonometric functions of one of its angles.
4. Review properties of the functions $\sin(x)$ and $\cos(x)$ and other trigonometric functions. State and apply the connections between these functions ("trigonometric identities").

Trigonometric functions are closely associated with angles and ratios of sides of a right-angle triangle. They are also connected to the motion of a point moving around a unit circle. Before we articulate these connections, we must agree on a universal way of measuring angles.

Angles and circles

Angles can be measured in a number of ways. One is to assign a value in **degrees**, with the convention that one complete revolution is represented by 360°. It turns out that this measure is not particularly convenient, and we instead replace it with a more universal quantity.

Our definition of angles will be based on the fact that circles of all sizes have one common geometric feature: they have the same ratio of **circumference**, to diameter, no matter what their size (or where in the universe they occur). We call that ratio π, that is

$$\pi = \frac{\text{Circumference of circle}}{\text{Diameter of circle}}.$$

By construction, the diameter D of a circle is a distance that corresponds to twice the radius R of that circle, so

$$D = 2R.$$

This leads to the familiar relationship of circumference C, to radius R,

$$C = 2\pi R.$$

This statement is merely a *definition* of the constant π.

As shown in Figure 14.3, an angle θ can be put into correspondence with an arc along the edge of a circle. For a circle of radius R and angle θ we define the arclength, S by the relation $S = R\theta$ where θ is measured in a

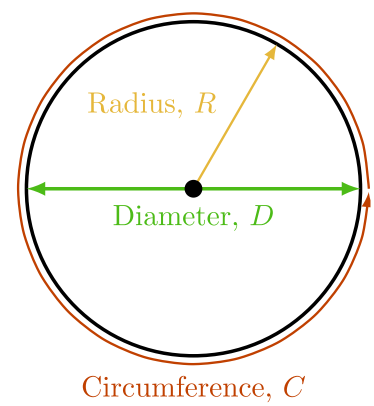


Figure 14.2: Circumference, diameter and radius of a circle.

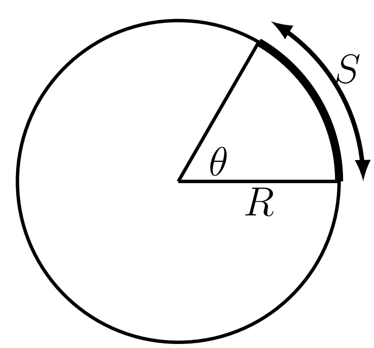


Figure 14.3: The angle θ in radians is related to the radius R of the circle, and the length of the arc S by the simple formula, $S = R\theta$.

convenient way that we now select. Both S and R carry units of “distance" or “length". But their ratio is,

$$\theta = \frac{S}{R},$$

so the units in numerator and denominator cancel, and the angle θ is **dimensionless** (carries no units).

Now consider a circle of radius $R = 1$ (called a **unit circle**) and denote by s a length of arc around the entire perimeter of this unit circle. Then

$$\theta = \frac{s}{1}.$$

In particular, for one complete revolution around the circle, the arclength is $s = 2\pi \cdot 1 = 2\pi$, which is just the circumference of the unit circle. In that case, it makes sense to consider the angle corresponding to one revolution as

$$\theta = \frac{2\pi}{1} = 2\pi.$$

This leads naturally to the definition of the **radian**: *we identify an angle of 2π radians with one complete revolution around the circle*. Note that (like degrees or other measures of angles), a radian is a number that carries no “units”.

We can now use this measure for angles to assign values to any fraction of a revolution, and thus, to any angle. For example, an angle of $90°$ corresponds to one quarter of a revolution around the perimeter of a unit circle, so we identify the angle $\pi/2$ radians with it. One degree is $1/360$ of a revolution, corresponding to $2\pi/360$ radians, and so on.

We summarize the properties of radians:

1. The length of an arc along the perimeter of a circle of radius R corresponding to an angle θ between two radii is $S = R\theta$ for θ in radians.
2. An angle in radians is the ratio of the arclength it subtends in a circle to the radius of that circle (and hence, a radian carries no units).
3. One complete revolution, or one full cycle corresponds to an angle of 2π radians.

We can convert between degrees and radians by remembering that $360°$ corresponds to 2π radians ($180°$ then corresponds to π radians, $90°$ to $\pi/2$ radians, etc.)

Mastered Material Check

1. What angle θ corresponds to a 1/6 revolution around the perimeter of a circle?
2. Sketch an angle of $\pi/4$ radians.
3. If the radius of a circle is 2 and an arc on its perimeter has length 0.5, what is the angle corresponding to that arc (in radians)?

Mastered Material Check

4. How many radians does $270°$ correspond to?
5. Label $\cos(t)$ and $\sin(t)$ where appropriate on Figure 14.4.

Defining the trigonometric functions $\sin(t)$ *and* $\cos(t)$

Consider a point (x, y) moving around the rim of a circle of radius 1, and let t be some angle (measured in radians) formed by the x-axis and the radius vector to the point (x, y) as shown in Figure 14.4.

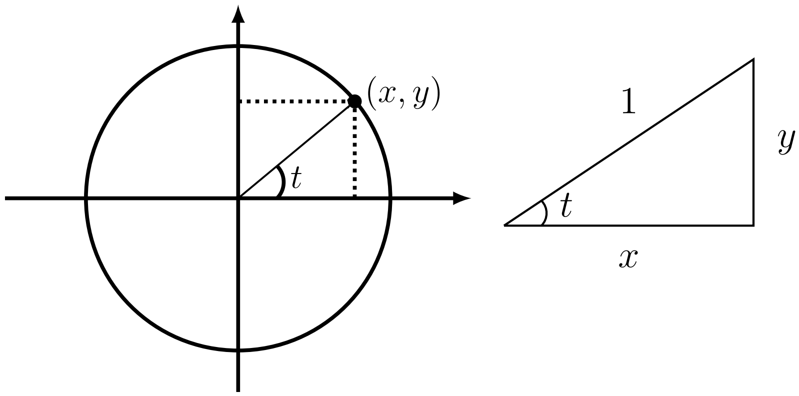


Figure 14.4: The equation of a circle of radius 1, with center at the origin is $x^2 + y^2 = 1$. The radius vector to the point (x, y) forms an angle t (radians) with the x-axis. In the triangle, shown enlarged on the right, lengths of sides are labeled. The trigonometric functions are just ratios of two sides of this triangle.

We define the functions sine and cosine, both dependent on the angle t (abbreviated $\sin(t)$ and $\cos(t)$) as follows:

$$\sin(t) = \frac{y}{1} = y, \quad \cos(t) = \frac{x}{1} = x$$

That is, the function sine tracks the y coordinate of the point as it moves around the unit circle, and the function cosine tracks its x coordinate.

Note: A review and definitions of trigonometric quantities is given in Figure F.1, Appendix F as ratios of sides in a right angle triangle. The hypotenuse in our diagram is simply the radius $r = 1$ of the circle.

A demonstration of the link between motion on a circle and the function $y = \sin(x)$. Click on the arrow left of the parameter a or shift the slider on a to see the moving point.

Featured Problem 14.1 (Cosine as motion around circle) *Adapt the interactive sine graph to represent the link between the graph of $y = \cos(x)$ and the x coordinate of a point moving around a circle.*

Properties of $\sin(t)$ *and* $\cos(t)$

We now explore the consequences of these definitions:

Values of sine and cosine.

- The radius of the unit circle is 1. This means that the x coordinate of any point (x, y) on the unit circle cannot be larger than 1 or smaller than -1. The same holds for the y coordinate. Thus, the functions $\sin(t)$ and $\cos(t)$ are always swinging between -1 and 1. ($-1 \leq \sin(t) \leq 1$ and $-1 \leq \cos(t) \leq 1$ for all angles t). The maximum value of each function is 1, the minimum is -1, and the average is 0.

- We adopt the convention that when the radius vector points along the x-axis, the angle is $t = 0$, and coordinates of the point are $x = 1, y = 0$. This implies that $\cos(0) = 1$, $\sin(0) = 0$.

- When the radius vector points up the y-axis, the angle is $\pi/2$ (corresponding to one quarter of a complete revolution), and coordinates of the point are $x = 0, y = 1$ so that $\cos(\pi/2) = 0, \sin(\pi/2) = 1$.

- Through geometry, we can also determine the lengths of all sides - and hence the ratios of the sides - of particular triangles, namely
  - equilateral triangles (in which all angles are 60°), and
  - right triangles (two equal angles of 45°).

  These types of calculations (omitted here) lead to some easily determined values for the sine and cosine of such special angles. These are shown in Table F.1 of the Appendix F.

Mastered Material Check

6. Review Appendix F and then use triangles to determine the x and y coordinates of angles of 60° and 45° in the unit circle.
7. Why does cosine have its largest value when the angle $t = 0$, at the beginning of the cycle?

Characteristics of sine and cosine.

- Both $\sin(t)$ and $\cos(t)$ go through the same values every time the angle t completes another cycle around the circle. We refer to such functions as **periodic** functions.
- The two functions, sine and cosine depict the same underlying motion, viewed from two perspectives: $\cos(t)$ represents the projection of the circularly moving point onto the x-axis, while $\sin(t)$ is the projection of the same point onto the y-axis. In this sense, the functions are "twins", and as such we expect many relationships connecting them.
- The cosine has its largest value at the beginning of the cycle, when $t = 0$ (since $\cos(0) = 1$), while sine has its peak value a little later $(\sin(\pi/2) = 1)$. Throughout their circular race, the sine function is $\pi/2$ radians ahead of the cosine, that is,
$$\cos(t) = \sin\left(t + \frac{\pi}{2}\right).$$
  See Figure 14.5 for graphs of both functions showing this shift by $\pi/2$.
- The **period** T of the sine function $\sin(t)$ is defined as the value of t for which one whole cycle (around the circle) has been completed. Accordingly, this period is $T = 2\pi$. Similarly the period of the cosine function $\cos(t)$ is also 2π. (See Figure 14.5.)
- The point $(x, y) = (\cos(t), \sin(t))$ is on a circle of radius 1, and, thus, its coordinates satisfy
$$x^2 + y^2 = 1.$$
  This implies that
$$\sin^2(t) + \cos^2(t) = 1 \qquad (14.1)$$
  for any angle t. This is an important relation, (also called a **trigonometric identity**), and one that is frequently used. See Appendix F for a review of other trigonometric identities.
- The sine and cosine functions have symmetries that we already encountered: $\sin(t)$ is an odd function (symmetric about the origin) and the $\cos(t)$ is an even function (symmetric about the y-axis). These symmetries also imply that $\sin(-t) = -\sin(t)$ and $\cos(-t) = \cos(t)$.

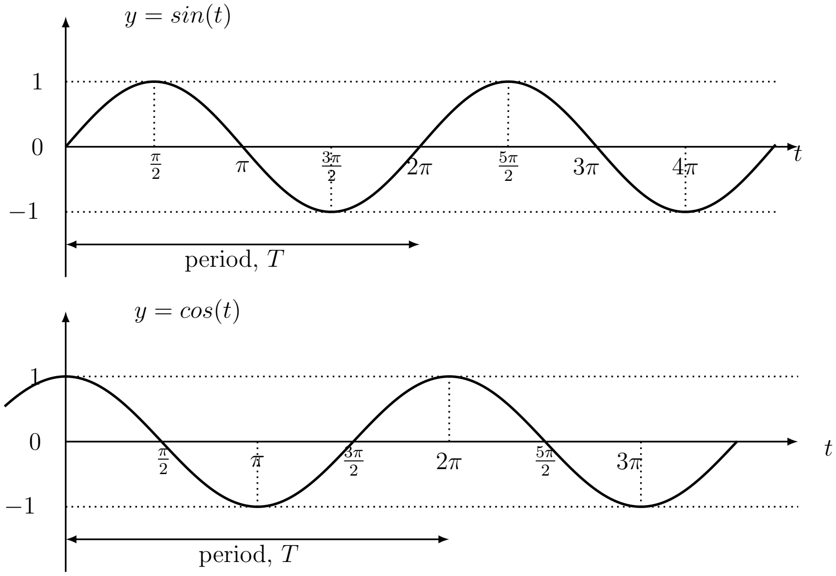


Figure 14.5: Periodicity of the sine and cosine. Note that the two curves are just shifted versions of one another.

Other trigonometric functions

Although we shall mostly be concerned with the two basic functions described above, several others are historically important and are encountered frequently in integral calculus. These include the following:

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

You can use this **desmos** graph to see all the trigonometric functions. Turn the graphs on or off by clicking on the (grey) circles to the left of the formulae. Notice the vertical asymptotes on some of these functions and think about where these asymptotes occur.

We review these and the identities that they satisfy in Appendix F. We also include the Law of Cosines in Eqn. (F.2), and angle-sum identities in the same appendix. Sine and cosine are the functions we focus on here.

14.2 Periodic functions

Section 14.2 Learning goals

1. Define a periodic function.
2. Given a periodic function, determine its **period**, **amplitude** and **phase**.
3. Given a graph or description of a periodic or rhythmic process, "fit" an approximate sine or cosine function with the correct period, amplitude and phase.

In Section 14.1, we identified the period of $\sin(t)$ and $\cos(t)$ as the value of t at which one full cycle is completed. Here we formalize the definition of a periodic function, define its period, frequency, and other properties.

Definition 14.1 (Periodic function) *A function is said to be* **periodic** *if*

$$f(t) = f(t+T),$$

where T is a constant that we call the **period** *of the function. Graphically, this means that if we shift the function by a constant "distance" along the horizontal axis, we see the same picture again.*

Example 14.1 *Show that the trigonometric functions are indeed periodic.*

Solution. The point (x,y) in Figure 14.4 repeats its trajectory every time a revolution around the circle is complete. This happens when the angle t completes one full cycle of 2π radians. Thus, as expected, the trigonometric functions are periodic, that is

$$\sin(t) = \sin(t+2\pi), \quad \text{and} \quad \cos(t) = \cos(t+2\pi).$$

Similarly

$$\tan(t) = \tan(t+2\pi), \quad \text{and} \quad \cot(t) = \cot(t+2\pi).$$

We say that the period of these functions is $T = 2\pi$ radians. The same applies to $\sec(t)$ and $\csc(t)$, that is, all six trigonometric functions are periodic. ◊

Mastered Material Check

8. If $\cos(\alpha) = \beta$, what is $\cos(\alpha + 2\pi)$? $\cos(\alpha - 6\pi)$?

Phase, amplitude, and frequency

In Appendix C we review how the appearance of any function changes when we **transform** variables. For example, replacing the independent variable x by $x-a$ (or αx) shifts (or scales) the function horizontally, multiplying f by a constant C scales the function vertically, etc. The same ideas apply to shapes of a trigonometric function when similar transformations are applied.

A function of the form

$$y = f(t) = A\sin(\omega t)$$

has both t and y-axes scaled, as shown in Figure 14.6(c). The the **amplitude** of the graph, A scales the y axis so that the oscillation swings between a minimum of $-A$ and a maximum of A. The **frequency** ω, scales the t-axis. This cycles are crowded together (if $\omega > 1$) or stretched out (if $\omega < 1$). One full cycle is completed when

$$\omega t = 2\pi,$$

and this occurs at time

$$t = \frac{2\pi}{\omega}.$$

Use the sliders on this **desmos** graph to see how the amplitude, A, frequency w, and phase ϕ affect the graph of the function $y = M + A\sin(w(t-\phi))$. You can also vary the mean value M.

Mastered Material Check

9. How many zeros are depicted in each panel of Figure 14.6?
10. How many local minima are depicted in each panel of Figure 14.6?
11. In each panel of Figure 14.6, identify where $y = 1$.
12. Indicate a single period on each panel of Figure 14.6.

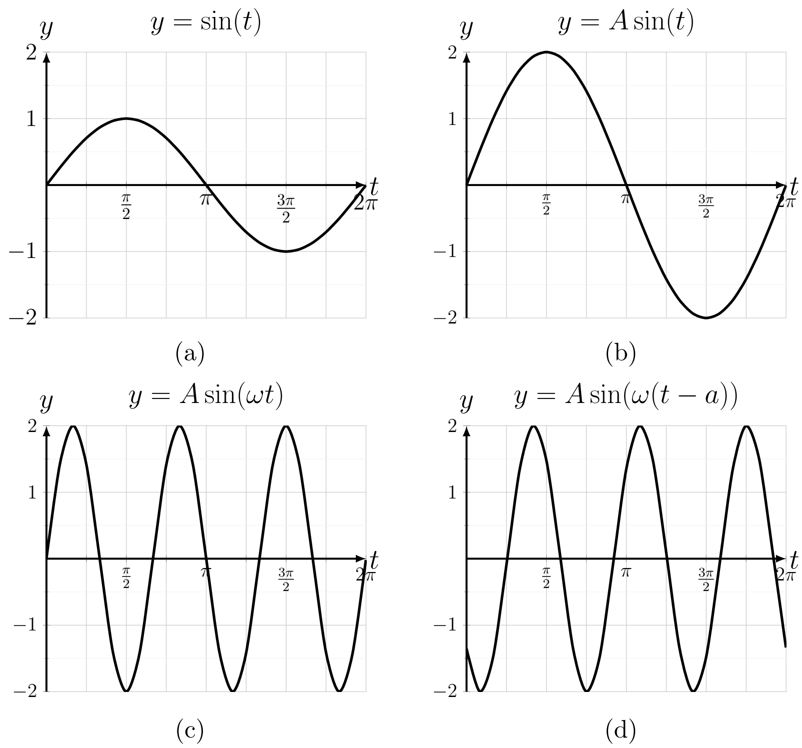


Figure 14.6: The sine function $y = \sin(t)$ (shown in (a)) is transformed in several ways. (b) Multiplying the function by a constant ($A = 2$) stretches the graph vertically. A is called the amplitude. (c) Multiplying the independent variable by a constant ($\omega = 3$) increases the frequency, i.e. the number of cycles per unit time. (d) Subtracting a constant ($a = 0.8$) from the independent variable shifts the graph horizontally to the right.

This time, called the *period* of the function is denoted by T. The connection between frequency and period is:

$$\omega = \frac{2\pi}{T}, \quad \Rightarrow \quad T = \frac{2\pi}{\omega}.$$

If we examine a graph of the function

$$y = f(t) = A\sin(\omega(t-a)),$$

we find that the basic sine graph has been shifted in the positive t direction by a, as in Figure 14.6(d). At time $t = a$, the value of the function is

$$y = f(t) = A\sin(\omega(a-a)) = A\sin(0) = 0,$$

so the cycle "starts" with a delay of $t = a$ relative to the basic sine function.

Another common variant of the same function can be written in the form

$$y = f(t) = A\sin(\omega t - \phi).$$

Here ϕ is called the **phase shift** of the oscillation. The above two forms are the same if we identify ϕ with ωa. The phase shift, ϕ has no units, whereas a has units of time, which is the same as the units of t. When $\phi = 2\pi$, (which

Mastered Material Check

13. What is the period of a trigonometric functions whose frequency is 5 cycles per min?
14. What is the frequency of a trigonometric functions whose period is 1 hr?

Mastered Material Check

15. Sketch a graph of $y = 3\sin(t - \frac{\pi}{2})$ and $y = 3\cos(t)$
16. Sketch a graph of $y = \cos(4(t-\pi))$

happens when $a = 2\pi/\omega$), the graph has been moved over to the right by one full period, making it identical to the original periodic graph.

Some of the scaled, shifted, sine functions described here are shown in Figure 14.6.

The periodic electrocardiogram

With the terminology of periodic function in place, we can now describe the ECG pattern for both normal resting individuals and those at exercise.

Recall that at rest, the heart beats approximately once per second. Consider the ECG trace on the left of Figure 14.7. This corresponds to a single heartbeat, and so, takes 1 second from start to finish. Suppose t represents time in seconds, and let $y = f(t)$ represent the electrical activity (in mV) at time t. Then, since this pattern repeats, the function f is periodic, with period $T = 1$ second. We can write

$$f(t) = f(t+1), \quad t \text{ in seconds.}$$

However, suppose the individual starts running. Then this relationship no longer holds, since heartbeats become more frequent, and the length of their period, T, decreases. This suggests a more natural way to mark off time - the amount of time it takes to complete a heartbeat cycle. Thus, rather than seconds, we define a such a variable, denoting it "the cycle time" and use the notation $\bar{\tau}$. Then the connection between clock time t and cycle time is

$$t = \text{time in seconds} = \text{number of cycles} \cdot \text{length of 1 cycle in sec} = \bar{\tau} \cdot T$$

Restated,

$$\bar{\tau} = \frac{t}{T}.$$

Note: the "number of cycles" need not be an integer - for example, $\bar{\tau} = 2.75$ means that we are $3/4$ way into the third electrical activity cycle.

Since f is a periodic function, we can "join up" its two ends and "wrap it around a circle", as shown in the schematic on Figure 14.7. Then successive heartbeats are depicted by traversing the circle over and over again. This suggests identifying the beginning of a cycle with 0 and the end of a cycle with 2π. To do so, we revise our cycle time clock as follows. Define

$$\tau = \text{number of radians traversed since time 0.}$$

Then every heartbeat corresponds to this clock adding an increment of 2π. The connection between this cycle clock and time t in seconds is

$$\tau = 2\pi\bar{\tau} = 2\pi\frac{t}{T} \tag{14.2}$$

(We check that when 1 cycle is complete, $t = T$ and $\tau = 2\pi$, as desired.)

We can now describe the periodicity of the ECG in terms of the cycle clock by the formula

$$f(\tau) = f(\tau + 2\pi). \tag{14.3}$$

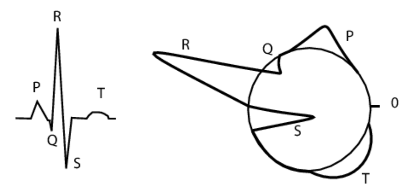


Figure 14.7: One full ECG cycle (left) has been "wrapped around a circle".

Mastered Material Check

17. If a runner's heart beats every 0.6 seconds, how many beat cycles ellapse in 10 seconds?
18. Suppose $t = 4$ sec and $\bar{\tau} = 5$. Determine T and interpret this situation.

As a check, we show in the next example that the relationship in Eqn. (14.3) reduces to the familiar period and frequency notation in terms of our original time t in seconds.

Example 14.2 (Period and frequency of heartbeat) *Use the formula we arrived at for τ and its connection to clock time t to transform back to time t in seconds. Express the periodicity of the function f both in terms of the period T and the frequency ω of heartbeats.*

Solution. Start with Eqn. (14.3),

$$f(\tau) = f(\tau + 2\pi).$$

Substitute $\tau = 2\pi t/T$ from Eqn. (14.3) and simplify by making a common denominator. Then

$$\begin{aligned} f\left(\frac{2\pi t}{T}\right) &= f\left(\frac{2\pi t}{T} + 2\pi\right) \\ &= f\left(\frac{2\pi}{T}(t+T)\right). \end{aligned}$$

We now rewrite this in terms of the frequency $\omega = 2\pi/T$ to arrive at

$$f(\omega t) = f(\omega(t+T)).$$

This relationship holds for any regular heartbeat, whether at rest or exercise where the frequency of the heartbeat, ω, is related to the period (duration of 1 beat cycle) by $\omega = 2\pi/T$. ◊

Rhythmic processes

Many natural phenomena are cyclic. It is sometimes convenient to represent such phenomena with a simple periodic functions, such as sine or cosine. Given some periodic process, we determine its frequency (or period), amplitude, and phase shift. We create a trigonometric function (sine or cosine) that approximates the desired behaviour.

To select a function, it helps to remember that (at $t = 0$) cosine starts at its peak, while sine starts at its average value of 0. A function that starts at the lowest point of the cycle is $-\cos(t)$. In most cases, the choice of sine or cosine to represent the cyclic phenomenon is arbitrary, they are related by a simple phase shift.

Next, pick a constant ω such that the trigonometric function $\sin(\omega t)$ (or $\cos(\omega t)$) has the correct period using the relationship $\omega = 2\pi/T$. We then select the amplitude, and horizontal and vertical shifts to complete the process. The examples below illustrate this process.

Example 14.3 (Daylight hours:) *In Vancouver, the shortest day (8 hours of light) occurs around December 22, and the longest day (16 hours of light) is around June 21. Approximate the cyclic changes of daylight through the season using the sine function.*

Solution. On Sept 21 and March 21 the lengths of day and night are equal, and then there are 12 hours of daylight (each of these days is called an **equinox**). Suppose we identify March 21 as the beginning of a yearly day-night length cycle. Let t be time in days beginning on March 21. One full cycle takes a year, i.e. 365 days. The period of the function we want is thus

$$T = 365$$

and its frequency (in units of per day) is

$$\omega = 2\pi/365.$$

Daylight shifts between the two extremes of 8 and 16 hours: i.e. 12 ± 4 hours. This means that the amplitude of the cycle is 4 hours. The oscillation take place about the average value of 12 hours. We have decided to start a cycle on a day for which the number of daylight hours is the average value (12). This means that the sine would be most appropriate, so the function that best describes the number of hours of daylight at different times of the year is:

$$D(t) = 12 + 4\sin\left(\frac{2\pi}{365}t\right)$$

where t is time in days and D the number of hours of light. $\Diamond$

Mastered Material Check

19. In August, the average number of daylight hours is 14. How does this fit with our model?
20. Repeat Example 14.3 using a cosine function.

Example 14.4 (Hormone levels) *The level of a certain hormone in the bloodstream fluctuates between undetectable concentration at t =7:00 and 100 ng/ml (nanograms per millilitre) at t =19:00 hours. Approximate the cyclic variations in this hormone level with the appropriate periodic trigonometric function. Let t represent time in hours from 0:00 hrs through the day.*

Solution. We first note that it takes one day (24 hours) to complete a cycle. This means that the period of oscillation is 24 hours, so that the frequency is

$$\omega = \frac{2\pi}{T} = \frac{2\pi}{24} = \frac{\pi}{12}.$$

The level of hormone varies between 0 and 100 ng/ml, which can be expressed as 50 ± 50 ng/ml. (The trigonometric functions are symmetric cycles, and we are finding both the average value about which cycles occur and the amplitude of the cycles.) We could consider the time midway between the low and high points, namely 13:00 hours as the point corresponding to the upswing at the start of a cycle of the sine function. (See Figure 14.8 for the sketch.) Thus, if we use a sine to represent the oscillation, we should shift it by 13 hrs to the left.

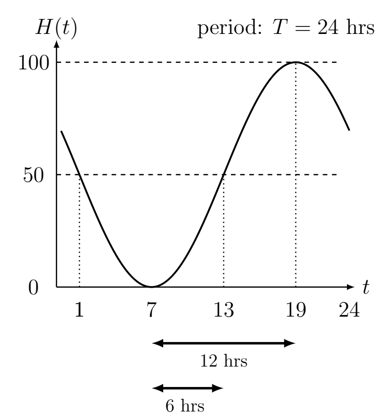


Figure 14.8: Hormonal cycles. The full cycle takes 24 hrs (hence the period is $T = 24$h and the frequency is $\omega = 2\pi/24$ per hour). The level $H(t)$ swings between 0 and 100 ng/ml. From the given information, we see that the average level is 50 ng/ml, and that the origin of a sine curve should be at $t = 13$ (i.e. 1/4 of the cycle which is 6 hrs past the time point $t = 7$).

Assembling these observations, we obtain the level of hormone, H at time t in hours:

$$H(t) = 50 + 50 \sin\left(\frac{\pi}{12}(t - 13)\right).$$

In the expression above, the number 13 represents a shift along the time axis, and carries units of time. We can express this same function in the form

$$H(t) = 50 + 50 \sin\left(\frac{\pi t}{12} - \frac{13\pi}{12}\right).$$

In this version, the quantity

$$\phi = \frac{13\pi}{12}$$

is the phase shift.

In selecting the periodic function to use for this example, we could have made other choices. For example, the same periodic can be represented by any of the functions listed below:

$$H_1(t) = 50 - 50 \sin\left(\frac{\pi}{12}(t - 1)\right),$$
$$H_2(t) = 50 + 50 \cos\left(\frac{\pi}{12}(t - 19)\right),$$
$$H_3(t) = 50 - 50 \cos\left(\frac{\pi}{12}(t - 7)\right).$$

All these functions have the same values, the same amplitudes, and the same periods. ◊

Mastered Material Check

21. Verify that H_1, H_2 and H_3 all have the same periods.

Example 14.5 (Phases of the moon) *A cycle of waxing and waning moon takes 29.5 days approximately. Construct a periodic function to describe the changing phases, starting with a "new moon" (totally dark) and ending one cycle later.*

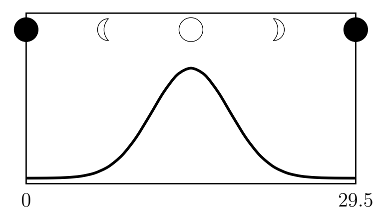


Figure 14.9: Periodic moon phases. The horizontal axis is time in days, and the vertical axis represents $P(t)$, the fraction of the moon that is visible from Earth on day t.

Solution. The period of the cycle is $T = 29.5$ days, so

$$\omega = \frac{2\pi}{T} = \frac{2\pi}{29.5}.$$

Let $P(t)$ be the fraction of the moon face on day t in the cycle. Then we should construct the function so that $0 < P < 1$, with $P = 1$ in mid cycle (see Figure 14.9). The cosine function swings between the values -1 and 1. To obtain a positive function in the desired range for $P(t)$, we add a constant and scale the cosine as follows:

$$\frac{1}{2}[1 + \cos(\omega t)].$$

This is still not quite right, since at $t = 0$ this function takes the value 1, rather than 0, as shown in Figure 14.9. To correct this we can either introduce a phase shift, i.e. set

$$P(t) = \frac{1}{2}[1 + \cos(\omega t + \pi)].$$

Then when $t = 0$, we get $P(t) = 0.5[1 + \cos\pi] = 0.5[1 - 1] = 0$. Or we can write

$$P(t) = \frac{1}{2}[1 - \cos(\omega t + \pi)],$$

which achieves the same result. ◊

Mastered Material Check

22. What fraction of the moon do you expect to be visible one week into the cycle?

14.3 *Inverse trigonometric functions*

Note: the material in this section can be omitted without loss of continuity in the next chapter. If this is done, merely skip Sections 15.1 and 15.3.

Section 14.3 Learning goals

1. Review inverse functions, and define the inverse trigonometric functions.
2. Explain why the domain of a periodic function must be restricted to define its inverse. Given any of the trigonometric functions, identify the suitably restricted domain on which an inverse function can be defined.
3. Simplify and/or interpret the meaning of expressions involving the trigonometric and inverse trigonometric functions.

Trigonometric functions provides another opportunity to illustrate the roles and properties of inverse functions. The inverse of a trigonometric function leads to exchange in the roles of the dependent and independent variables, as well as the the roles of the domain and range. Recall that geometrically, an inverse function is obtained by reflecting the original function about the line $y = x$. However, we must take care that the resulting graph represents a true function, i.e. satisfies all the properties required of a function.

The domains of $\sin(x)$ and $\cos(x)$ are both $-\infty < x < \infty$ while their ranges are $-1 \leq y \leq 1$. In the case of the function $\tan(x)$, the domain excludes values $\pm\pi/2$ as well as angles $2n\pi \pm \pi/2$ at which the function is undefined. The range of $\tan(x)$ is $-\infty < y < \infty$.

There is one difficulty in defining inverses for trigonometric functions: the fact that these functions repeat their values in a cyclic pattern means that a given y value is obtained from many possible values of x. For example, the values $x = \pi/2, 5\pi/2, 7\pi/2$, etc. all satisfy $\sin(x) = 1$. We say that these functions are not **one-to-one**. Geometrically, this means the graphs of the trigonometric functions intersect a horizontal line in numerous places.

When these graphs are reflected about the line $y = x$, they would intersect a *vertical* line in many places, and would fail to be functions: the function would have multiple y values corresponding to the same value of x, which is not allowed. See Appendix C where the inverse function for $y = x^2$ is discussed.

To avoid this difficulty, we restrict the domains of a trigonometric functions to a portion of their graphs that does not repeat. To do so, we select an interval over which the given trigonometric function is one-to-one, i.e. over which there is a unique correspondence between values of x and values of y. We then define the corresponding inverse function, as described below.

Arcine is the inverse of sine. The function $y = \sin(x)$ is one-to-one on the

Mastered Material Check

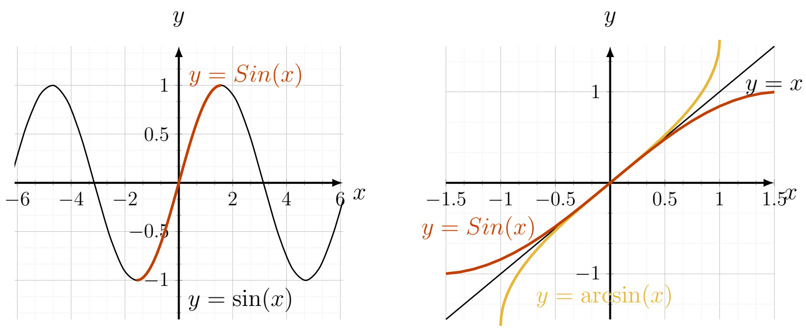
23. What property of a function might fail when we define an inverse function?
24. Give an example of two different functions which are not one-to-one.
25. Why is the range of $\tan(x)$ $-\infty < y < \infty$? Why are some points not in the domain?

The trigonometric and inverse trigonometric functions are shown. Turn the graphs on or off by clicking on the (grey) circles to the left of the formulae. Notice that each function and its inverse are reflections of one another about the line $y = x$. Also observe that the domain of the inverse functions is restricted, to avoid multiple y values for a given x value.

interval $-\pi/2 \leq x \leq \pi/2$. We define the associated function $y = Sin(x)$ (shown in red on Figures 14.10(a) and (b)) by restricting the domain of the sine function to $-\pi/2 \leq x \leq \pi/2$. On the given interval, we have $-1 \leq Sin(x) \leq 1$. We define the inverse function, called arcsine

$$y = \arcsin(x) \qquad -1 \leq x \leq 1$$

in the usual way, by reflection of $Sin(x)$ through the line $y = x$ as shown in Figure 14.10(a).



(a) (b)

Figure 14.10: (a) The original trigonometric function, $\sin(x)$, in black, as well as the portion restricted to a smaller domain, $Sin(x)$, in red. The red curve is shown again in part b. (b) Relationship between the functions $Sin(x)$, defined on $-\pi/2 \leq x \leq \pi/2$ (in red) and $\arcsin(x)$ defined on $-1 \leq x \leq 1$ (in orange). Note that one is the reflection of the other about the line $y = x$. The graphs in parts (a) and (b) are not on the same scale.

To interpret this function, we note that $\arcsin(x)$ is "the angle (in radians) whose sine is x". In Figure 14.11, we show a triangle in which $\theta = \arcsin(x)$. This follows from the observation that the sine of theta, opposite over hypotenuse ($\sin\theta = \frac{opp}{hyp}$ as reviewed in Appendix F) is $x/1 = x$. The length of the other side of the triangle is then $\sqrt{1-x^2}$ by the Pythagorean theorem.

For example $\arcsin(\sqrt{2}/2)$ is the angle whose sine is $\sqrt{2}/2$, namely $\pi/4$ (we can also see this by checking the values of trigonometric functions of standard angles shown in Table F.1)

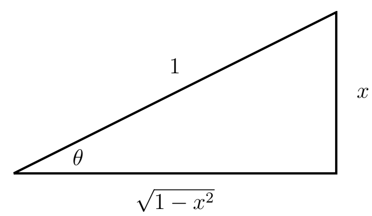


Figure 14.11: This triangle has been constructed so that θ is an angle such that $\sin(\theta) = sx/1 = x$. This means that $\theta = \arcsin(x)$.

The functions $\sin(x)$ and $\arcsin(x)$, reverse (or "invert") each other's effect, that is:

$$\arcsin(\sin(x)) = x \quad \text{for} \quad -\pi/2 \leq x \leq \pi/2,$$

$$\sin(\arcsin(x)) = x \quad \text{for} \quad -1 \leq x \leq 1.$$

Note: the allowable values of x that can be "substituted in" are not exactly the same for these two cases. In the first case, x is an angle whose sine we compute first, and then reverse the procedure. In the second case, x is a number whose arc-sine is an angle.

While we can evaluate $\arcsin(\sin(x))$ for any value of x, the result may not agree with the original value of x - unless we restrict attention to the interval $-\pi/2 \leq x \leq \pi/2$.

Mastered Material Check

26. What is the value of $\arcsin(\sin(x))$ for $x = \pi/2$ and for $5\pi/2$?
27. What is the value of $\sin(\arcsin(x))$ for $x = 1$?
28. Is $\arcsin(x)$ defined for $x = \pi/2$?

Example 14.6 *Let $x = \pi$. Compare x and* $\arcsin(\sin(x))$.

Solution. When $x = \pi$ we get that $\sin(x) = 0$. Thus $\arcsin(\sin(x)) = \arcsin(0) = 0$, which is not the same as $x = \pi$. ♢

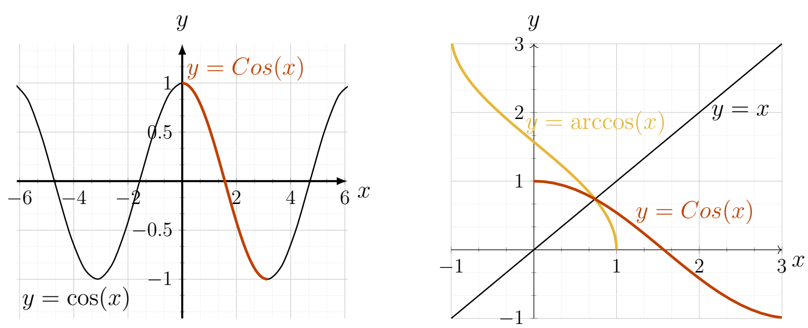
The other case also requires careful attention:

Example 14.7 *Let $x = 2$. Compare x and* $\sin(\arcsin(x))$.

Solution. Notice that $x = 2$ is outside of the interval $-1 \leq x \leq 1$. Thus $\arcsin(2)$ is simply not defined, and so neither is $\sin(\arcsin(x))$. ♢

Indeed, care must be taken in handling the inverse trigonometric functions.

Inverse cosine. We cannot use the same interval (as that for sine) to restrict the cosine function because $\cos(x)$ is an even function, symmetric about the y axis (and not one-to-one on the interval $\pi/2 \leq x \leq \pi/2$). Instead, the appropriate interval is $0 \leq x \leq \pi$. Let $y = Cos(x) = \cos(x)$ for $0 \leq x \leq \pi$ be



(a) (b)

Figure 14.12: (a) The original function $\cos(x)$, is shown in black; the restricted domain version, $Cos(x)$ is shown in red. The same red curve appears in part (b) on a slightly different scale. (b) Relationship between the functions $Cos(x)$ (in red) and $\arccos(x)$ (in orange). Note that one is the reflection of the other about the line $y = x$.

the restricted-domain version of cosine (red curve in Figure 14.12). On the interval $0 \leq x \leq \pi$, we have $1 \geq Cos(x) \geq -1$. We define arccosine

$$y = \arccos(x) \qquad -1 \leq x \leq 1$$

as the inverse of $Cos(x)$, as shown in orange in Figure 14.12(b).

The meaning of the expression $y = \arccos(x)$ is "the angle (in radians) whose cosine is x." For example, $\arccos(0.5) = \pi/3$ because $\pi/3$ is an angle whose cosine is $1/2 = 0.5$. The triangle in Figure 14.13, is constructed so that $\theta = \arccos(x)$. (This follows from the fact that $\cos(\theta)$ is adjacent over hypotenuse.) The length of the third side of the triangle is obtained using the Pythagorian theorem.

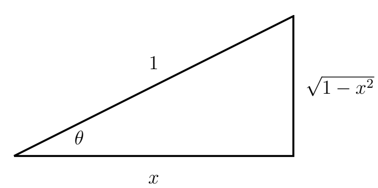


Figure 14.13: This triangle has been constructed so that θ is an angle such that $\cos(\theta) = x/1 = x$ implying that $\theta = \arccos(x)$.

The inverse relationship between the functions mean that

$$\arccos(\cos(x)) = x \quad \text{for} \quad 0 \leq x \leq \pi,$$

$$\cos(\arccos(x)) = x \quad \text{for} \quad -1 \leq x \leq 1.$$

The same subtleties apply as in the case of arcsine.

Inverse tangent. The function $y = \tan(x)$ is one-to-one on an interval $\pi/2 < x < \pi/2$. Unlike the case for $Sin(x)$, we must exclude the endpoints, where the function $\tan(x)$ is undefined. We therefore restrict the domain to $\pi/2 < x < \pi/2$, that is, we define,

$$y = Tan(x) = \tan(x) \quad \pi/2 < x < \pi/2.$$

Again in contrast to the sine function, as x approaches either endpoint of this interval, the value of $Tan(x)$ approaches $\pm\infty$, i.e. $-\infty < Tan(x) < \infty$. This means that the domain of the inverse function is $-\infty < x < \infty$, i.e. the inverse function is defined for all real values of x. We define the inverse tan function:

$$y = \arctan(x) \quad -\infty < x < \infty.$$

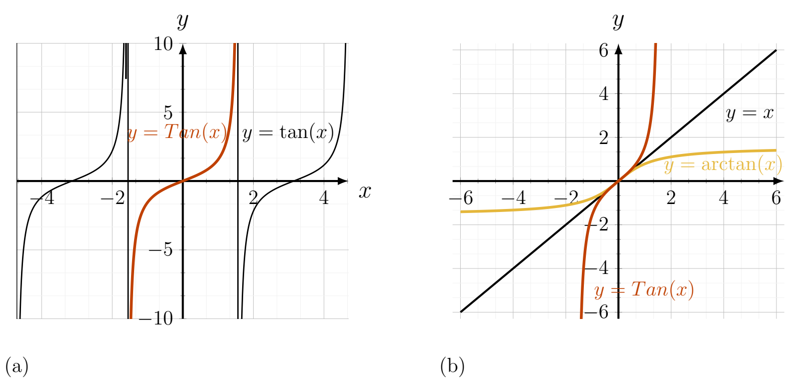


Figure 14.14: (a) The function $\tan(x)$, is shown in black, and $Tan(x)$ in red. The same red curve is repeated in part b. (b) Relationship between the functions $Tan(x)$ (in red) and $\arctan(x)$ (in orange). Note that one is the reflection of the other about the line $y = x$.

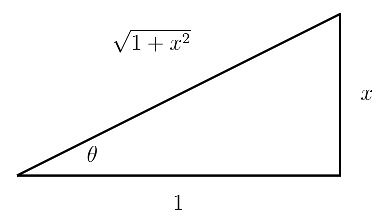


Figure 14.15: This triangle has been constructed so that θ is an angle such that $\tan(\theta) = x/1 = x$. This means that $\theta = \arctan(x)$.

As before, we can understand the meaning of the inverse tan function, by constructing a triangle in which $\theta = \arctan(x)$, shown in Figure 14.15.

The inverse tangent "inverts" the effect of the tangent on the relevant interval:

$$\arctan(\tan(x)) = x \quad \text{for} \quad -\pi/2 < x < \pi/2$$

$$\tan(\arctan(x)) = x \quad \text{for} \quad -\infty < x < \infty$$

The same comments hold in this case.

A summary of the above inverse trigonometric functions, showing their graphs on a single page is provided in Figure F.3 in Appendix F. Some of the standard angles allow us to define precise values for the inverse trig functions. A table of such standard values is given in the same Appendix (See Table F.2). For other values of x, one has to calculate the decimal approximation of the function using a scientific calculator.

Example 14.8 *Simplify the following expressions:*

a) $\arcsin(\sin(\pi/4))$,

b) $\arccos(\sin(-\pi/6))$

Solution.

a) $\arcsin(\sin(\pi/4)) = \pi/4$ since the functions are simple inverses of one another on the domain $-\pi/2 \le x \le \pi/2$.

b) We evaluate this expression piece by piece: First, note that $\sin(-\pi/6) = -1/2$. Then $\arccos(\sin(-\pi/6)) = \arccos(-1/2) = 2\pi/3$. The last equality is obtained from Table F.2.

◊

Example 14.9 *Simplify the expressions:*

a) $\tan(\arcsin(x))$,

b) $\cos(\arctan(x))$.

Solution.

a) Consider first the expression $\arcsin(x)$, and note that this represents an angle (call it θ) whose sine is x, i.e. $\sin(\theta) = x$. Refer to Figure 14.11 for a sketch of a triangle in which this relationship holds. Now note that $\tan(\theta)$ in this same triangle is the ratio of the opposite side to the adjacent side, i.e.

$$\tan(\arcsin(x)) = \frac{x}{\sqrt{1-x^2}}.$$

b) Figure 14.15 shows a triangle that captures the relationship $\tan(\theta) = x$ or $\theta = \arctan(x)$. The cosine of this angle is the ratio of the adjacent side to the hypotenuse, so that

$$\cos(\arctan(x)) = \frac{1}{\sqrt{x^2+1}}.$$

◊

Mastered Material Check

29. Redraw and further label triangles to aid in the solutions to Example 14.9.
30. Evaluste the expression $\arccos(\sin(\pi/4))$.
31. Simplify the expression $\sin(\arccos(x))$.

14.4 Summary

1. This chapter introduced and reviewed angles, cyclic processes, trigonometric, and periodic functions
2. amplitude, period, frequency, and phase were defined and identified with graphical properties
3. the functions cosine and sine correspond to (x, y) coordinates of a point moving around a circle of radius 1. $\tan(x) = \sin(x)/\cos(x)$ is their ratio.
4. rhythmic processes can be approximated by a sine or a cosine graph once the period, amplitude, mean, and phase shift are identified.
5. to define an inverse trigonometric function, the domain of the original trig function has to be restricted to make it one-to-one (no repeated y values).

6. Applications addressed in this chapter included:
   (a) electrocardiograms detecting the electrical activity of the heart;
   (b) daylight hours fluctuating with period of one year;
   (c) hormone levels that change on a daily rhythm; and
   (d) phases of the moon, with a 29.5 day period.

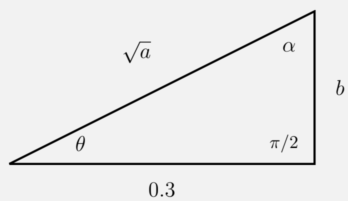
Quick Concept Checks

1. What is the range of the function $y = 8\sin(2t)$?

2. Does a phase shift change the period of a trigonometric function?

3. If, in a 1-minute interval, a heart beats 50 times, what is the length of a heart beat cycle?

4. Using the following right angle triangle, determine:



   (a) $\tan(\alpha)$;
   (b) $\arccos(\sin(\theta))$;
   (c) $\cos(\arccos(\sin(\alpha)$.

Exercises

14.1. **Radians, degrees and right tirangles.** Convert the following expressions in radians to degrees:

(a) π,

(b) $5\pi/3$,

(c) $21\pi/23$,

(d) 24π.

Convert the following expressions in degrees to radians:

(e) 100^o,

(f) 8^o,

(g) 450^o,

(h) 90^o.

Using a right triangle, evaluate each of the following:

(i) $\cos(\pi/3)$,

(j) $\sin(\pi/4)$,

(k) $\tan(\pi/6)$.

14.2. **Graphing functions.** Graph the following functions over the indicated ranges:

(a) $y = x\sin(x)$ for $-2\pi < x < 2\pi$,

(b) $y = e^x\cos(x)$ for $0 < x < 4\pi$.

14.3. **Sketching the graph.** Sketch the graph for each of the following functions:

(a) $y = \frac{1}{2}\sin 3\left(x - \frac{\pi}{4}\right)$,

(b) $y = 2 - \sin x$,

(c) $y = 3\cos 2x$,

(d) $y = 2\cos\left(\frac{1}{2}x + \frac{\pi}{4}\right)$.

14.4. **Converting angles.** The radian is an important unit associated with angles. One revolution about a circle is equivalent to 360 degrees or 2π radians.

(a) Convert the following angles (in degrees) to angles in radians. Express each as multiples of π, not as decimal expansions:

(i) 45 degrees

(ii) 30 degrees

(iii) 60 degrees

(iv) 270 degrees.

(b) Find the sine and the cosine of each of the above angles.

14.5. **Trigonometric functions and rhythmic functions.** Find the appropriate trigonometric function to describe the following rhythmic processes:

(a) Daily variations in the body temperature $T(t)$ of an individual over a single day, with the maximum of 37.5^{o}C at 8:00 am and a minimum of 36.7^o C 12 hours later.

(b) Sleep-wake cycles with peak wakefulness ($W = 1$) at 8:00 am and 8:00pm and peak sleepiness ($W = 0$) at 2:00pm and 2:00 am.

In both cases, express t as time in hours with $t = 0$ taken at 0:00 am.

14.6. **Trigonometric functions and rhythmic functions.** Find the appropriate trigonometric function to describe the following rhythmic processes:

(a) The displacement S cm of a block on a spring from its equilibrium position, with a maximum displacement 3 cm and minimum displacement -3 cm, a period of $\frac{2\pi}{\sqrt{g/l}}$ and at $t = 0$, $S = 3$.

(b) The vertical displacement y of a boat that is rocking up and down on a lake, with y measured relative to the bottom of the lake. It has a maximum displacement of 12 meters and a minimum of 8 meters, a period of 3 seconds, and an initial displacement of 11 meters when measurement was first started (i.e., $t = 0$).

14.7. **Sunspot cycles.** The number of sunspots (solar storms on the sun) fluctuates with roughly 11-year cycles with a high of 120 and a low of 0 sunspots detected. A peak of 120 sunspots was detected in the year 2000.

Which of the following trigonometric functions could be used to approximate this cycle?

(A) $N = 60 + 120\sin\left(\frac{2\pi}{11}(t - 2000) + \frac{\pi}{2}\right)$

(B) $N = 60 + 60\sin\left(\frac{11}{2\pi}(t + 2000)\right)$

(C) $N = 60 + 60\cos\left(\frac{11}{2\pi}(t + 2000)\right)$

(D) $N = 60 + 60\sin\left(\frac{2\pi}{11}(t - 2000)\right)$

(E) $N = 60 + 60\cos\left(\frac{2\pi}{11}(t - 2000)\right)$

14.8. **Inverse trigonometric functions.** As seen in Section 14.3, the inverse trigonometric function $\arctan(x)$ (also written $\tan^{-1}(x)$) means the angle θ where $-\pi/2 < \theta < \pi/2$ whose tan is x. Thus $\cos(\arctan(x)$

(or $\cos(\tan^{-1}(x))$ is the cosine of that same angle. By using a right triangle whose sides have length $1, x$ and $\sqrt{1+x^2}$ we can verify that

$$\cos(\arctan(x)) = 1/\sqrt{1+x^2}.$$

Use a similar geometric argument to arrive at a simplification of the following functions:

(a) $\sin(\arcsin(x))$,

(b) $\tan(\arcsin(x)$,

(c) $\sin(\arccos(x)$.

14.9. **Inverse trigonometric functions.** The value of $\tan(\arccos(x))$ is which of the following?

(A) $1-x^2$,

(B) x ,

(C) $1+x^2$,

(D) $\frac{\sqrt{1-x^2}}{x}$,

(E) $\frac{\sqrt{1+x^2}}{x}$,

14.10. **Inverse trigonometric functions.** The function $y = \tan(\arctan(x))$ has which of the following for its domain and range?

(A) Domain $0 \le x \le \pi$; Range $-\infty \le y \le \infty$

(B) Domain $-\infty \le x \le \infty$; Range $-\infty \le y \le \infty$

(C) Domain $-\pi \le x \le \pi$; Range $-\pi \le y \le \pi$;

(D) Domain $-\pi/2 \le x \le \pi/2$; Range $-\pi/2 \le y \le \pi/2$;

(E) Domain$-\infty \le x \le \infty$; Range $0 \le y \le \pi$

14.11. **Simplify trigonometric identity.**

(a) Use a double-angle trigonometric identity to simplify the following expression as much as possible:

$$y = \cos(2\arcsin(x)).$$

(b) For what values of x is this simplification possible?

15
Cycles, periods, and rates of change

Having acquainted ourselves with properties of the trigonometric functions and their inverses in Chapter 14, we are ready to compute their derivatives and apply our results to understanding rates of change of these periodic functions. We compute derivatives, and then use these results in a medley of problems on optimization, related rates, and differential equations.

15.1 Derivatives of trigonometric functions

Section 15.1 Learning goals

1. Use the definition of the derivative to calculate the derivatives of $\sin(x)$ and $\cos(x)$.
2. Using the quotient rule, compute derivatives of $\tan(x)$, $\sec(x)$, $\csc(x)$, and $\cot(x)$.
3. Using properties of the inverse trigonometric functions and implicit differentiation, calculate derivatives of $\arcsin(x)$, $\arccos(x)$, and $\arctan(x)$.

Limits of trigonometric functions

In Chapter 3, we zoomed in on the graph of the sine function close to the origin (Figure 3.2). Through this, we reasoned that

$$\sin(x) \approx x, \quad \text{for small } x.$$

Restated, with h replacing the variable x

$$\sin(h) \approx h, \quad \text{for small } h \quad \Rightarrow \quad \frac{\sin(h)}{h} \approx 1 \text{ for small } h$$

Observe the behaviour of the two limits, (15.1) and (15.2) by zooming into the graphs for h close to zero. (Click on the $+$ tab to zoom in.)

The smaller is h, the better this "tangent line" approximation becomes. In more formal limit notation, we say that

$$\lim_{h\to 0} \frac{\sin(h)}{h} = 1. \tag{15.1}$$

(See Eqn. (3.1).) This is a very important limit, and we apply it directly in computing the derivative of trigonometric functions using the definition of the derivative.

A similar analysis of the graph of the cosine function, leads to a second important limit:

$$\lim_{h\to 0} \frac{\cos(h)-1}{h} = 0. \tag{15.2}$$

We can now apply these to computing derivatives of both the sine and the cosine functions.

Mastered Material Check

1. Perform the calculation to verify that $\frac{d\cos(x)}{dx} = -\sin(x)$.
2. Based on properties of the sine function, estimate the value of $\sin(x)$ for $x = 0.01$ radians and for $x = 0.01$ degrees.
3. Repeat this for estimates of $\cos(x)$ for these two values of x.

Derivatives of sine, cosine, and other trigonometric functions

Let $y = f(x) = \sin(x)$ be the function to differentiate, where x is now the independent variable (previously t). We use the definition of the derivative to compute the derivative of this function.

Example 15.1 (Derivative of $\sin(x)$**)** *Compute the derivative of* $y = \sin(x)$ *using the definition of the derivative.*

Explanation of the calculation of the derivative of $\sin(x)$ using the definition of the derivative.

Solution. We apply the definition of the derivative as follows:

$$\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \\
\frac{d\sin(x)}{dx} &= \lim_{h\to 0} \frac{\sin(x+h)-\sin(x)}{h} \\
&= \lim_{h\to 0} \frac{\sin(x)\cos(h)+\sin(h)\cos(x)-\sin(x)}{h} \\
&= \lim_{h\to 0} \left(\sin(x)\frac{\cos(h)-1}{h} + \cos(x)\frac{\sin(h)}{h}\right) \\
&= \sin(x)\left(\lim_{h\to 0}\frac{\cos(h)-1}{h}\right) + \cos(x)\left(\lim_{h\to 0}\frac{\sin(h)}{h}\right) \\
&= \cos(x).
\end{aligned}$$

Observe that a trigonometric identity (for the sum of angles - see Eqn. (F.3)) and Limits (15.1) and (15.2) were used to obtain the final result. ◊

A similar calculation using the function $\cos(x)$ leads to the result

$$\frac{d\cos(x)}{dx} = -\sin(x).$$

Note: the same two limits appear in this calculation, as well as the trigonometric identity Eqn. (F.4).

We can now calculate the derivative of the any of the other trigonometric functions using the quotient rule.

Example 15.2 (Derivative of the function $\tan(x)$**)** *Compute the derivative of* $y = \tan(x)$ *using the quotient rule.*

Using the quotient rule to compute the derivative of $y = \tan(x)$.

Solution. We apply the quotient rule:

$$\frac{d\tan(x)}{dx} = \frac{[\sin(x)]'\cos(x) - [\cos(x)]'\sin(x)}{\cos^2(x)}.$$

Using the recently found derivatives for the sine and cosine, we have

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

But the numerator of the above can be simplified using the trigonometric identity Eqn. (14.1), leading to

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

◊

Mastered Material Check

4. Verify one or more of the derivatives of $\csc(x)$, $\sec(x)$ or $\cot(x)$ using the quotient rule.
5. For what ranges of values of x and y are the two statements $y = \arctan(x)$ as $\tan(y) = x$ equivalent?
6. For what range(s) of values of y are these two functions not inverses of one another?

Derivatives of the six trigonometric functions are given in Table 15.1. The first three are frequently encountered in practical applications and worth committing to memory.

$y = f(x)$	$f'(x)$
$\sin(x)$	$\cos(x)$
$\cos(x)$	$-\sin(x)$
$\tan(x)$	$\sec^2(x)$
$\csc(x)$	$-\csc(x)\cot(x)$
$\sec(x)$	$\sec(x)\tan(x)$
$\cot(x)$	$-\csc^2(x)$

Table 15.1: Derivatives of the trigonometric functions.

Featured Problem 15.1 (Lung volume) *Breathing is a rhythmic process. The volume of air in the lungs can be modelled by a function of the form*

$$V(t) = C + A\sin(\omega t + \phi),$$

where V is the volume in millilitres (ml) and t is time in seconds. Suppose that the minimum and maximum volumes are 1400 and 3400 ml, respectively, and that the maximum rate of change of V is 1200 ml/sec. What is the period of $V(t)$?

Derivatives of the inverse trigonometric functions

Implicit differentiation - introduced in Chapter 9 - can be used to determine the derivatives of the inverse trigonometric functions, explored in Section 14.3. As an example, we demonstrate how to compute the derivative of $\arctan(x)$. To do so, we need to recall that the derivative of the function $\tan(x)$ is $\sec^2(x)$. We also use the identity $\tan^2(x) + 1 = \sec^2(x)$. (See Eqn. (F.1).)

Let $y = \arctan(x)$. Then on the appropriate interval, we can replace this relationship with the equivalent one:

$$\tan(y) = x.$$

Differentiating implicitly with respect to x on both sides, we obtain

$$\sec^2(y)\frac{dy}{dx} = 1,$$

$$\frac{dy}{dx} = \frac{1}{\sec^2(y)} = \frac{1}{\tan^2(y) + 1}.$$

Now using again the relationship $\tan(y) = x$, we obtain

$$\frac{d\arctan(x)}{dx} = \frac{1}{x^2+1}.$$

This expression is used frequently in integral calculus. The derivatives of the important inverse trigonometric functions are given in Table 15.2.

$y = f(x)$	$f'(x)$
$\arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$
$\arccos(x)$	$-\frac{1}{\sqrt{1-x^2}}$
$\arctan(x)$	$\frac{1}{x^2+1}$.

Table 15.2: Derivatives of the inverse trigonometric functions.

15.2 Changing angles and related rates

Section 15.2 Learning goals

1. Explain how the chain rule is applied to geometric problems with angles that are change over time ("related rates").
2. Given a description of the geometry and/or rate of change of angle or side of a triangle, set up the mathematical problem and solve it using geometry and/or properties of the trigonometric functions.

The examples in this section provide practice with chain rule applications based on trigonometric functions. We discuss a number of problems, and show how the basic properties of these functions, together with some geometry, are used to arrive at desired results.

Example 15.3 (A point on a circle) *A point moves around the rim of a circle of radius* 1 *so that the angle θ between the radius vector and the x axis changes at a constant rate,*

$$\theta = \omega t,$$

where t is time. Determine the rate of change of the x and y coordinates of that point.

Solution. We have $\theta(t)$, $x(t)$, and $y(t)$ all functions of t. (The geometry is captured by Figure 14.4, but the angle has been renamed θ, and we consider it to be time-dependent.) The fact that θ is proportional to t means that

$$\frac{d\theta}{dt} = \omega.$$

The x and y coordinates of the point are related to the angle by

$$x(t) = \cos(\theta(t)) = \cos(\omega t),$$

$$y(t) = \sin(\theta(t)) = \sin(\omega t).$$

Then, by the chain rule,

$$\frac{dx}{dt} = \frac{d\cos(\theta)}{d\theta}\frac{d\theta}{dt},$$
$$\frac{dy}{dt} = \frac{d\sin(\theta)}{d\theta}\frac{d\theta}{dt}.$$

Performing the required calculations, we have

$$\frac{dx}{dt} = -\sin(\theta)\omega,$$

$$\frac{dy}{dt} = \cos(\theta)\omega.$$

◊

Mastered Material Check

7. Write a concluding sentence for Example 15.3.
8. If $\omega = 2$ per min, what is the rate of change of the x and y coordinates when $\theta = \pi/2$?

Example 15.4 (Runners on a circular track) *Two runners start at the same position on a circular race track of length* 400 *meters. Joe Runner takes* 50 *sec, while Michael Johnson takes* 43.18 *sec to complete the* 400 *meter race. Find the rate of change of the angle formed between the two runners and the center of the track, assuming that they run at a constant rate.*

Solution. We are told that the track is 400 meters in length (total). However, this information does not actually enter into the solution. Joe completes one cycle around the track (2π radians) in 50 sec, while Michael completes a cycle in 43.18 sec. This means that

- Joe has a period of $T_1 = 50$ sec, and a frequency of $\omega_1 = 2\pi/T_1 = 2\pi/50$ radians per sec.
- Michael has period is $T_2 = 43.18$ sec, and a frequency of $\omega_2 = 2\pi/T_2 = 2\pi/43.18$ radians per sec.

Let θ_J, θ_M be the angles subtended between one of the runners and the starting line. We take the x-axis as that starting line, by convention, as in Figure 14.4. From this, we find that

$$\frac{d\theta_J}{dt} = \frac{2\pi}{50} = 0.125 \text{ radians per sec},$$

$$\frac{d\theta_M}{dt} = \frac{2\pi}{43.18} = 0.145 \text{ radians per sec}.$$

Thus, the angle between the runners, $\theta_M - \theta_J$ changes at the rate

$$\frac{d(\theta_M - \theta_J)}{dt} = \frac{d\theta_M}{dt} - \frac{d\theta_J}{dt} = 0.145 - 0.125 = 0.02 \text{ radians per sec}.$$

◊

Example 15.5 (Simple law of cosines) *The law of cosines applies to an arbitrary triangle, as reviewed in Appendix F (see Eqn (F.2)). Consider the triangle shown in Figure 15.1. Suppose that the angle θ increases at a constant rate, $d\theta/dt = k$. If the sides $a = 3$, $b = 4$, are of constant length, determine the rate of change of the length c opposite this angle at the instant that $c = 5$.*

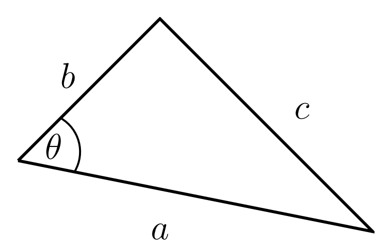


Figure 15.1: Law of cosines states that $c^2 = a^2 + b^2 - 2ab\cos(\theta)$.

Solution. Let a, b, c be the lengths of the three sides, with c the length of the side opposite angle θ. The law of cosines states that

$$c^2 = a^2 + b^2 - 2ab\cos(\theta).$$

We identify the changing quantities by writing this relation in the form

$$c^2(t) = a^2 + b^2 - 2ab\cos(\theta(t))$$

so it is evident that only c and θ vary with time, while a, b remain constant. Differentiating with respect to t and using the chain rule leads to:

$$2c\frac{dc}{dt} = -2ab\frac{d\cos(\theta)}{d\theta}\frac{d\theta}{dt}.$$

But $d\cos(\theta)/d\theta = -\sin(\theta)$ and $d\theta/dt = k$, so that

$$\frac{dc}{dt} = -\frac{ab}{c}(-\sin(\theta))\frac{d\theta}{dt} = \frac{ab}{c}k \cdot \sin(\theta).$$

At the instant in question, $a = 3$, $b = 4$, and $c = 5$, forming a Pythagorean triangle in which the angle opposite c is $\theta = \pi/2$. We can see this fact using the law of cosines, and noting that

$$c^2 = a^2 + b^2 - 2ab\cos(\theta), \quad 25 = 9 + 16 - 24\cos(\theta).$$

This implies that $0 = -24\cos(\theta)$, $\cos(\theta) = 0$ so that $\theta = \pi/2$. Substituting these into our result for the rate of change of the length c leads to

$$\frac{dc}{dt} = \frac{ab}{c}k = \frac{3 \cdot 4}{5}k.$$

◊

Mastered Material Check

9. Verify that a triangle formed by sides of length 3, 4 and 5 is a Pythagorean triangle.
10. Redo Example 15.5 with a triangle in which $a = 5, b = 12$ and c is the third side, at the instant when the tringle so formed is Pythagorean.
11. Write a concluding sentence for Example 15.5.

Example 15.6 (Clocks) *Find the rate of change of the angle between the minute hand and hour hand on a clock.*

Solution. We call θ_1 the angle that the minute hand subtends with the x-axis (horizontal direction) and θ_2 the angle that the hour hand makes with the same axis as depicted in Figure 15.2(a).

If our clock is working properly, each hand moves around at a constant rate. The hour hand traces one complete revolution (2π radians) every 12 hours, while the minute hand completes a revolution every hour. Both hands move in a clockwise direction, which (by convention) is towards negative angles. This means that

$$\frac{d\theta_1}{dt} = -2\pi \text{ radians per hour,}$$

$$\frac{d\theta_2}{dt} = -\frac{2\pi}{12} \text{ radians per hour.}$$

The angle between the two hands is the difference of the two angles, i.e.

$$\theta = \theta_1 - \theta_2.$$

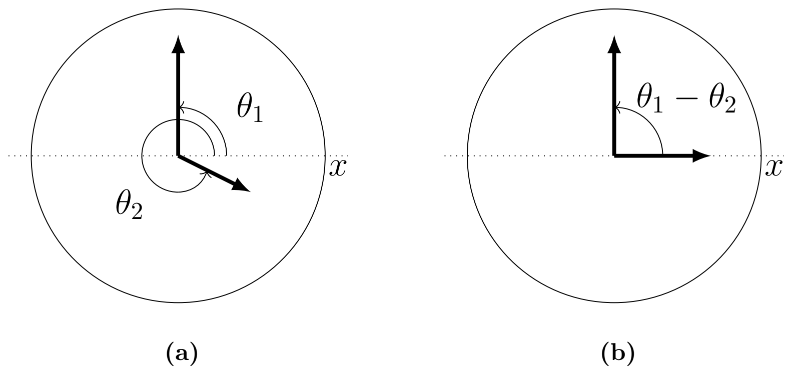


Figure 15.2: Figure for Examples 15.6 and 15.7; θ_1 and θ_2 are the angles the minute and hour hands, respectively, form with the x-axis.

Thus,

$$\frac{d\theta}{dt} = \frac{d}{dt}(\theta_1 - \theta_2) = \frac{d\theta_1}{dt} - \frac{d\theta_2}{dt} = -2\pi + \frac{2\pi}{12}.$$

We find that the rate of change of the angle between the hands is

$$\frac{d\theta}{dt} = -\pi\frac{11}{6}.$$

◊

Mastered Material Check

12. Highlight $\theta_1 - \theta_2$ on Figure 15.2(a).
13. What property of the derivative allows us to simplify the subtraction $\frac{d\theta_1}{dt} - \frac{d\theta_2}{dt}$?
14. Explain the significance of the minus sign in the solution to Example 15.7.
15. Redo Examples 15.6 and 15.7 for a clock whose two hands have lengths 12 and 5 cm. (The longer hand is the minute hand.)

Example 15.7 (Clocks, continued) *Suppose that the length of the minute hand is* 4 *cm and the length of the hour hand is* 3 *cm. At what rate is the* distance *between the ends of the hands changing when it is 3:00 o'clock?*

Solution. We use the law of cosines on the triangle shown in Figure 15.2. Side lengths are $a = 3$, $b = 4$, and $c(t)$ opposite the angle $\theta(t)$. From the previous example, we have

$$\frac{dc}{dt} = \frac{ab}{c}\sin(\theta)\frac{d\theta}{dt}.$$

At precisely 3:00 o'clock, the angle in question is $\theta = \pi/2$ and it can also be seen that the Pythagorean triangle abc leads to

$$c^2 = a^2 + b^2 = 3^2 + 4^2 = 9 + 16 = 25$$

so that $c = 5$. We found from our previous analysis that $d\theta/dt = \frac{11}{6}\pi$. Using this information leads to:

$$\frac{dc}{dt} = \frac{3\cdot 4}{5}\sin\left(\frac{\pi}{2}\right)\left(-\frac{11}{6}\pi\right) = -\frac{22}{5}\pi \text{ cm/hr}.$$

◊

15.3 The Zebra danio's escape responses

Section 15.3 Learning goals

1. Describe the geometry of a **visual angle**, and determine how that angle changes as the distance to the viewed object (or the size of the object) changes (an application of "related rates").
2. Determine how the rate of change of the visual angle of a prey fish (**zebra danio**) changes as a predator of a given size approaches it at some speed.
3. Explain the link between the rate of change of the visual angle and the triggering of an escape response.
4. Using the results of the analysis, explain in words under what circumstances the prey does (or does not) manage to escape from its predator.

We consider an example involving trigonometry and related rates with a biological application. We first consider the geometry on its own, and then link it to the biology of predator avoidance and escape responses.

Visual angles

Example 15.8 (Visual angle) *In the triangle shown in Figure 15.3, an object of height s is moving towards an observer. Its distance from the observer at some instant is labeled $x(t)$ and it approaches at some constant speed, v. Determine the rate of change of the angle $\theta(t)$ and how it depends on speed, size, and distance of the object. Often θ is called a visual angle, since it represents the angle that an image subtends on the retina of the observer.*

Demo of the changing visual angle. You can see how the angle depends on the size of the object s. Notice that the angle hardly changes when the object is far away, and changes dramatically as the object gets closer.

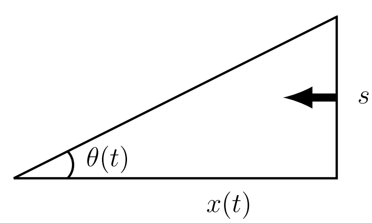


Figure 15.3: A visual angle θ would change as the distance x decreases. The size s is assumed constant. See Example 15.8.

Solution. The object approaches at some constant speed, v so that

$$\frac{dx}{dt} = -v.$$

where the minus sign means that the distance x is decreasing. Using the trigonometric relations, we see that

$$\tan(\theta) = \frac{s}{x}.$$

If the size, s, of the object is constant, then the changes with time imply that

$$\tan(\theta(t)) = \frac{s}{x(t)}.$$

We differentiate both sides of this equation with respect to t, and obtain

$$\frac{d\tan(\theta)}{d\theta}\frac{d\theta}{dt} = \frac{d}{dt}\left(\frac{s}{x(t)}\right),$$

$$\sec^2(\theta)\frac{d\theta}{dt} = -s\frac{1}{x^2}\frac{dx}{dt},$$

so that
$$\frac{d\theta}{dt} = -s\frac{1}{\sec^2(\theta)}\frac{1}{x^2}\frac{dx}{dt}.$$
We can use the trigonometric identity
$$\sec^2(\theta) = 1 + \tan^2(\theta)$$
to express our answer in terms of the size, s, the distance of the object, x and the speed v:
$$\sec^2(\theta) = 1 + \left(\frac{s}{x}\right)^2 = \frac{x^2+s^2}{x^2}$$
so
$$\frac{d\theta}{dt} = -s\left(\frac{x^2}{x^2+s^2}\right)\frac{1}{x^2}\frac{dx}{dt} = \frac{s}{x^2+s^2}v.$$
Thus, the rate of change of the visual angle is
$$sv/(x^2+s^2).$$
The angle thus changes at a rate proportional to the speed of the object. ◊

Mastered Material Check

16. Where did the minus sign go in the final result of $\frac{d\theta}{dt}$ for Example 15.8?

The dependence on the size of the object is more involved. This is explored next.

The Zebra danio and a looming predator

Visual angles are important to predator avoidance. We use the ideas of Example 15.8 to consider a problem in biology, studied by Larry Dill, a biologist at **Simon Fraser University** in Burnaby, BC.

The Zebra danio is a small tropical fish, with many predators. In order to survive, it must sense danger quickly enough to be able to escape from a pair of hungry jaws. At the same time, over-reacting to every perceived motion could be counter-productive, wasting energy and time better spent on foraging. Here, we investigate the visual basis of an escape response, based on a hypothesis by Dill [Dill, 1974a,b].

Figure 15.5 shows the relation between the angle subtended at the danio's eye and the profile size S of an approaching predator, currently located at distance x away. Suppose that the predator approaches at constant speed, v so that the distance x is decreasing. Then
$$\frac{dx}{dt} = -v.$$
Using this information and geometry, we characterize the rate of change of the angle α.

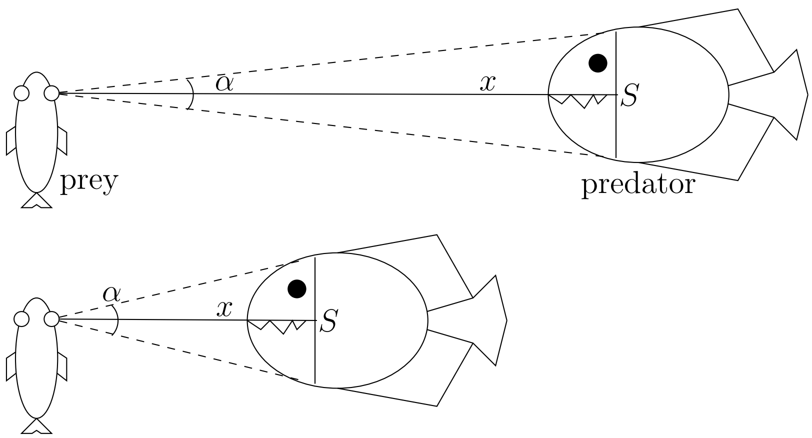


Figure 15.4: A cartoon showing the visual angle, $\alpha(t)$ and how it changes as a predator approaches its prey, the Zebra danio.

Example 15.9 (Danio's visual angle) *Use the above geometry to express the rate of change of angle α in terms of the size S and speed v of the approaching predator, and its distance away, x.*

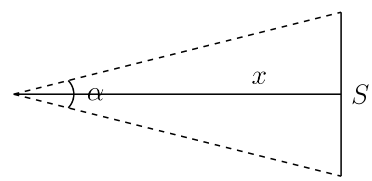


Figure 15.5: The geometry of the escape response problem.

Solution. If we consider the top half of the triangle shown in Figure 15.5 we find a Pythagorean triangle identical to that of Example 15.8, but with $\theta = \alpha/2$ and $s = S/2$. The side labeled x is identical in both pictures. Thus, the trigonometric relation that holds is:

$$\tan\left(\frac{\alpha}{2}\right) = \frac{(S/2)}{x}. \tag{15.3}$$

Furthermore, based on the results of Example 15.8, we know that $d\alpha/dt$ can be written as

$$\frac{d\alpha}{dt} = 2\frac{S/2}{x^2+(S/2)^2}v = \frac{Sv}{x^2+(S^2/4)}. \tag{15.4}$$

◊

Mastered Material Check

17. Verify the calculations for the solution of Example 15.9.
18. Write a concluding sentence for Example 15.9; address units.

Featured Problem 15.2 (Demo of looming predator) *Set up an interactive (Desmos) graph that represents an elliptical "predator" centered at $x = a$ that approaches its prey along the x axis. Draw the visual angle that this predator creates at $(0,0)$. Assume that the equation of the predator is given by the ellipse $(x-a)^2 + y^2/s^2 = 1$. By animating the parameter a, you should be able to make the predator move. (Refer to Featured Problem 9.3 for the geometry.)*

Example 15.10 (Distance-dependence) *Use the relationship in Eqn. (15.4) to sketch a rough graph of the rate of change of the visual angle α versus the distance x of the predator.*

Solution. We are asked to sketch $d\alpha/dt$ versus x. Let us denote by $f(x)$ the function of x that we want to graph. Then from Eqn. (15.4),

$$f(x) \equiv \frac{Sv}{x^2 + (S^2/4)}.$$

We first make three observations about this function.

- When $x = 0$, (and the predator has reached its prey),

$$f(0) = \frac{Sv}{0 + (S^2/4)} = \frac{4v}{S}.$$

  This provides the "y-intercept" of the graph.

- Since x appears in the denominator, the function $f(x)$ is always decreasing.

- For $x \to \infty$ (the predator is very far away), the denominator is very large, so

$$f(x) \to 0.$$

The sketch of $f(x)$ (the rate of change of the visual angle, $d\alpha/dt$) versus distance x of the predator is shown in Figure 15.6.

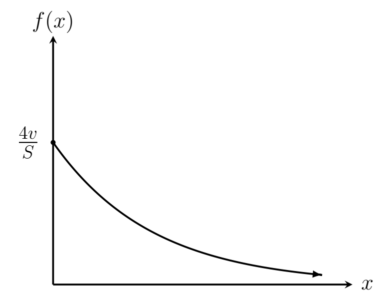


Figure 15.6: The function $f(x)$ plotted against x. This graph shows that the rate of change of visual angle $d\alpha/dt \equiv f(x)$ is small when the distance to the predator x is large.

◊

When to escape? What sort of visual input should the danio respond to, if it is to be efficient at avoiding the predator? In principle, we would like to consider a response that has the following features:

- If the predator is too far away, or moving slowly, (or moving further away), it is likely harmless and should not trigger an escape response.

- If the predator is moving quickly towards the danio, it is likely a threat, and should trigger the escape response.

Accordingly, the hypothesis proposed by Dill is that:

> The escape response is triggered when the predator approaches so quickly, that the rate of change of the visual angle is greater than some critical value.

Then the critical value, denoted $K_{\text{crit}} > 0$, is a constant that depends on the "skittishness of the prey or level of perceived danger of its environment. Hence, the danio's escape response is triggered when

$$\frac{d\alpha}{dt} = K_{\text{crit}}.$$

◊

Example 15.11 (Finding the predator's distance) *How far away is the predator when the escape response is triggered?*

Solution. Rewrite the above condition using the dependence of $d\alpha/dt$ on the geometric quantities in the problem. Then, we must solve for x in

$$\frac{Sv}{x^2 + (S^2/4)} = K_{\text{crit}}. \tag{15.5}$$

Figure 15.7(a) illustrates a geometric solution, showing the line $y = K_{\text{crit}}$ and the curve $y = Sv/(x^2 + (S^2/4))$ superimposed on the same graph.

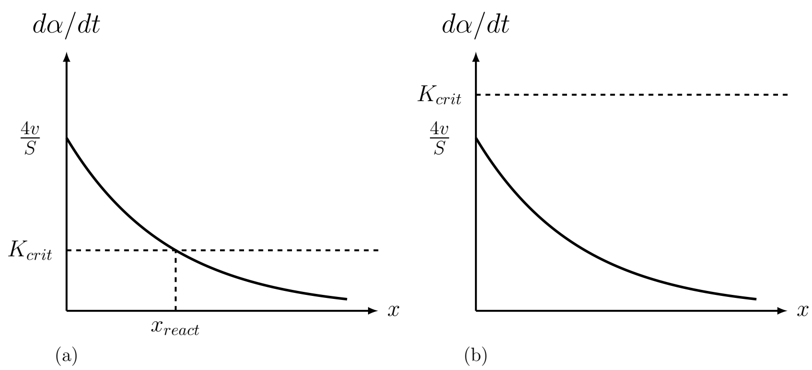


Figure 15.7: The rate of change of the visual angle $d\alpha/dt$ in two cases, when the quantity $4v/S$ is above (a) and below (b) some critical value.

The value $x = x_{\text{react}}$ is the distance of the predator at the instant that the escape response is triggered, termed the **reaction distance**. Solving for x, (Exercise 11) leads to:

$$x_{react} = \sqrt{\frac{Sv}{K_{\text{crit}}} - \frac{S^2}{4}} = \sqrt{S\left(\frac{v}{K_{\text{crit}}} - \frac{S}{4}\right)}. \tag{15.6}$$

◊

Example 15.12 (Lunch) *Interpret the reaction distance x_{react}. Are there ever cases in which the prey does not notice a predator in time to escape?*

Solution. Figure 15.7(b) illustrates a case where $K_{\text{crit}} = Sv/(x^2 + (S^2/4))$ is never satisfied. This could happen if either the danio has a very high threshold of alert (large K_{crit}), or if $4v/S$ is too low. That happens either if S is very large (big predator) or if v is small (predator slowly "stalks" its prey). From this scenario, we find that in some situations, the fate of the danio would be sealed in the jaws of its pursuer. ◊

Large slow predators beat danio's escape response. Notice the reaction distance, x_{react}, of the danio with reaction threshold K_{crit} is largest for certain sizes of predators. In Figure 15.8, we plot the reaction distance x_{react} versus the predator size S. We see that for very small predators ($S \approx 0$) or large predators ($S \approx 4v/K_{\text{crit}}$) the distance at which escape response is triggered is very small. This means that the danio may miss noticing such predators until they are too close for a comfortable escape, resulting in calamity.

On the other hand, some predators are detected when they are very far away- they have a large x_{react}. This is explored in Example 15.13.

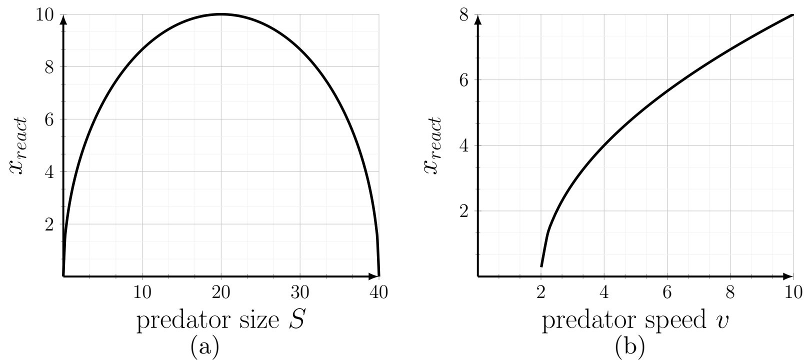


Figure 15.8: The reaction distance x_{react} is shown as a function of the predator size S in (a) and as a function of the predator speed v in (b). A very small or very large predator may fail to be detected until it is too close to escape. A very slow predator might never be detected.

Mastered Material Check

19. Which panel of Figure 15.8 would you use to determine the distance at which a danio would react to a predator of fixed size?
20. According to Figure 15.8 what size predator is easiest to detect farthest away? what size is undetectable until it is too close?
21. According to Figure 15.8 what is the slowest moving predator that is just barely detected? How far away is a predator detected if its speed is 10 distance units per unit time?

Example 15.13 (Bad design for a predator) *Some predators are more easily detected than others. Find the size of predator for which the reaction distance is maximal, and interpret your finding.*

Solution. We solve this problem using differentiation (Exercise 12) and find that x_{react} has a critical point at $S = 2v/K_{\text{crit}}$. From Figure 15.8(a), we see that this critical point is a local maximum. We can also reason based on Eqn. (15.6): x_{react} cannot be negative. However, we see that $x_{react} = 0$ at $S = 0$ and at $S = 4v/K_{\text{crit}}$. Hence, x_{react} has a local maximum for some predator size between these two values. In short, a predator of size $S = 2v/K_{\text{crit}}$ would be detected as far away as possible (largest possible x_{react}), giving the prey a good chance to escape. ◊

Observe that at sizes $S > 4v/K_{\text{crit}}$, the reaction distance is not defined. We also see this Figure 15.7(b): when $K_{\text{crit}} > 4v/S$, the straight line and the curve fail to intersect, and there is no solution.

Figure 15.8(b) illustrates the dependence of the reaction distance x_{react} on the speed v of the predator. We find that for small values of v, x_{react} is not defined: the danio would not notice the threat posed by predators that swim very slowly. See Exercise 13 for the largest velocity that fails to trigger the escape response.

Alternate approach involving inverse trig functions

Note: this section is optional and can be skipped or left for independent study.

In Section 15.3, we studied the escape response of the Zebra danio and showed that the connection between the visual angle and distance to predator satisfies

$$\tan\left(\frac{\alpha}{2}\right) = \frac{(S/2)}{x}. \qquad (15.7)$$

We also computed the rate of change of the visual angle per unit time using implicit differentiation and related rates. Here we illustrate an alternate approach using inverse trigonometric functions. (This provides practice with differentiation of inverse trigonometric functions.)

Example 15.14 *Use the inverse function arctan to restate the angle α in Eqn. 15.7 as a function of x. Then differentiate that function using the chain rule to compute $d\alpha/dt$.*

Solution. Restate the relationship using the function arctan:

$$\frac{\alpha}{2} = \arctan\left(\frac{S}{2x}\right).$$

Both the angle α and the distance x change with time; we indicate this by writing

$$\alpha(t) = 2\arctan\left(\frac{S}{2x(t)}\right).$$

Applying the chain rule, let $u = S/2x$. Recall that S is a constant, so $\alpha(u) = 2\arctan(u)$. Using the derivative of the inverse trigonometric function,

$$\frac{d\arctan(u)}{du} = \frac{1}{u^2+1}$$

and the chain rule lead to

$$\frac{d\alpha(t)}{dt} = \frac{d\arctan(u)}{du}\frac{du}{dx}\frac{dx}{dt} = \frac{1}{u^2+1}\left(-\frac{S}{2x^2(t)}\right)(-v).$$

Simplifying leads to the same result,

$$\frac{d\alpha}{dt} = \frac{Sv}{x^2+(S^2/4)}.$$

This rate of change of the visual angle agrees with Example 15.8. ◊

Mastered Material Check

22. Verify the simplification to the final solution in Example 15.14.

15.4 Summary

1. We used two limits

$$\lim_{h\to 0}\frac{\sin(h)}{h} = 1, \qquad \lim_{h\to 0}\frac{\cos(h)-1}{h} = 0$$

to computed the derivatives of two trigonometric functions

$$\frac{d\sin(x)}{dx} = \cos(x), \qquad \frac{d\cos(x)}{dx} = -\sin(x)$$

2. Implicit differentiation was used to compute the derivatives of the inverse trigonometric functions.

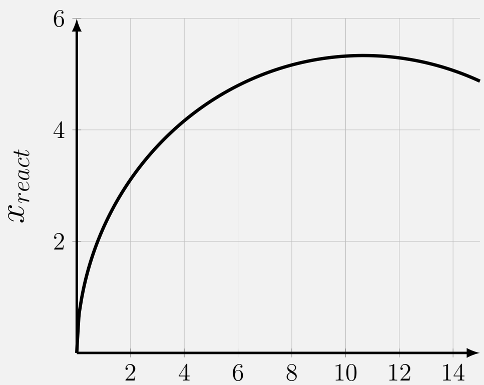
3. Trigonometric related rates, and implicit differentiation were discussed.
4. Applications addressed in this chapter included:
   (a) runners on a circular track (changing angle between them);
   (b) clock hands (changing angle between hands); and
   (c) Zebra danio's escape response (changing visual angle).

Quick Concept Checks

1. What is the derivative of $y = \cos(x) + \sin(x)$?

2. Rewrite the reaction distance formula twice: first clearly identifying the size, and second the speed, of the predator as the independent variable.

3. Consider the following graph in which x_{react} is the dependent variable:



Is size or speed the independent variable here? Why?

Exercises

15.1. **First derivatives.** Calculate the first derivative for the following functions:

(a) $y = \sin x^2$,

(b) $y = \sin^2 x$,

(c) $y = \cot^2 \sqrt[3]{x}$,

(d) $y = \sec(x - 3x^2)$,

(e) $y = 2x^3 \tan x$,

(f) $y = \frac{x}{\cos x}$,

(g) $y = x \cos x$,

(h) $y = e^{-\sin^2 \frac{1}{x}}$,

(i) $y = (2\tan 3x + 3\cos x)^2$,

(j) $y = \cos(\sin x) + \cos x \sin x$.

15.2. **Derivatives.** Take the derivative of the following functions.

(a) $f(x) = \cos(\ln(x^4 + 5x^2 + 3))$,

(b) $f(x) = \sin(\sqrt{\cos^2(x) + x^3})$,

(c) $f(x) = 2x^3 + \log_3(x)$,

(d) $f(x) = (x^2 e^x + \tan(3x))^4$,

(e) $f(x) = x^2\sqrt{\sin^3(x) + \cos^3(x)}$.

15.3. **Point moving on a circle.** A point is moving on the perimeter of a circle of radius 1 at the rate of 0.1 radians per second.

(a) How fast is its x coordinate changing when $x = 0.5$?

(b) How fast is its y coordinate changing at that time?

15.4. **Graphing trigonometric functions.** The derivatives of the two important trig functions are $[\sin(x)]' = \cos(x)$ and $[\cos(x)]' = -\sin(x)$. Use these derivatives to answer the following questions.

Let $f(x) = \sin(x) + \cos(x),\ \ 0 \le x \le 2\pi$

(a) Find all intervals where $f(x)$ is increasing.

(b) Find all intervals where $f(x)$ is concave up.

(c) Locate all inflection points.

(d) Graph $f(x)$.

15.5. **Tangent lines.** Find all points on the graph of $y = \tan(2x),\ -\frac{\pi}{4} < x < \frac{\pi}{4}$, where the slope of the tangent line is 4.

15.6. **Bird formation.** A "V" shaped formation of birds forms a symmetric structure in which the distance from the leader to the last birds in the V is $r = 10$m, the distance between those trailing birds is $D = 6$m and the angle formed by the V is θ, as shown in Figure 15.9.

Suppose that the shape is gradually changing: the trailing birds start to get closer so that their distance apart shrinks at a constant rate $dD/dt = -0.2$m/min while maintaining the same distance from the leader. Assume that the structure is always in the shape of a V as the other birds adjust their positions to stay aligned in the flock.

What is the rate of change of the angle θ?

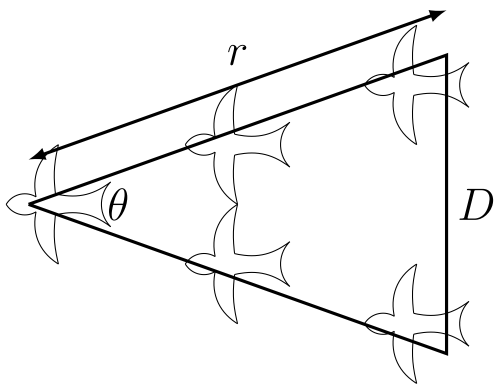


Figure 15.9: Figure for Exercise 6; bird flock formation.

15.7. **Hot air balloon.** A hot air balloon on the ground is 200 meters away from an observer. It starts rising vertically at a rate of 50 meters per minute. Find the rate of change of the angle of elevation of the observer when the balloon is 200 meters above the ground.

15.8. **Cannon-ball.** A cannon-ball fired by a cannon at ground level at angle θ to the horizon ($0 \le \theta \le \pi/2$) travels a horizontal distance (called the **range**, R) given by the formula below:

$$R = \frac{1}{16} v_0^2 \sin\theta \cos\theta.$$

Here $v_0 > 0$, the initial velocity of the cannon-ball, is a fixed constant and air resistance is neglected (see Figure 15.10).

What is the maximum possible range?

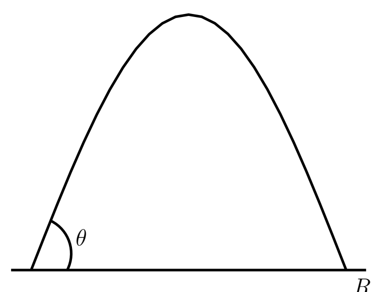


Figure 15.10: Figure for Exercise 8; cannon-ball trajectory.

15.9. **Leaning ladder.** A ladder of length L is leaning against a wall so that its point of contact with the ground is a distance x from the wall, and its point of contact with the wall is at height y. The ladder slips away from the wall at a constant rate C.

(a) Find an expression for the rate of change of the height y.

(b) Find an expression for the rate of change of the angle θ formed between the ladder and the wall.

15.10. **Cycloid curve.** A wheel of radius 1 meter rolls on a flat surface without slipping. The wheel moves from left to right, rotating clockwise at a constant rate of 2 revolutions per second.

Stuck to the rim of the wheel is a piece of gum, (labeled G); as the wheel rolls along, the gum follows a path shown by the wide arc (called a "cycloid curve") in Figure 15.11. The (x, y) coordinates of the gum (G) are related to the wheel's angle of rotation θ by the formulae

$$x = \theta - \sin\theta,$$
$$y = 1 - \cos\theta,$$

where $0 \le \theta \le 2\pi$.

(a) How fast is the gum moving horizontally at the instant that it reaches its highest point?

(b) How fast is it moving vertically at that same instant?

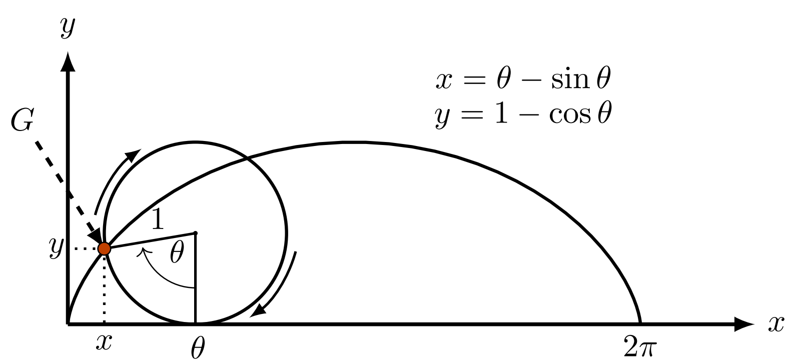


Figure 15.11: Figure for Exercise 10; a cycloid curve.

15.11. **Zebra danio's reaction distance.** Solve Eqn. (15.5) for x and show that you get the reaction distance $x \equiv x_{react}$ given in Eqn (15.6).

15.12. **Bad design for a predator.** Some predators are more easily detected than others. Use Eqn. (15.6) to find the size of predator for which the reaction distance is maximal.

15.13. **Sneaking up on the prey.**

(a) Use Eqn. (15.6) to show that a predator moving "slowly enough" can sneak up on the prey without being detected.

(b) What is the largest velocity for which a predator of size S is not detected by a prey that responds to a visual sighting when the rate of change of the visual angle exceeds the threshold K_{crit}?

15.14. **Inverse trigonometric derivatives.** Find the first derivative of the following functions.

(a) $y = \arcsin x^{\frac{1}{3}}$,

(b) $y = (\arcsin x)^{\frac{1}{3}}$,

(c) $\theta = \arctan(2r+1)$,

(d) $y = x \operatorname{arcsec} \frac{1}{x}$,

(e) $y = \frac{x}{a}\sqrt{a^2 - x^2} - \arcsin \frac{x}{a}$, $a > 0$,

(f) $y = \arccos \frac{2t}{1+t^2}$.

15.15. **Rotating wheel.** In Figure 15.12, the point P is connected to the point O by a rod 3 cm long. The wheel rotates around O in the clockwise direction at a constant speed, making 5 revolutions per second. The point Q, which is connected to the point P by a rod 5 cm long, moves along the horizontal line through O.

How fast and in what direction is Q moving when P lies directly above O?

Note: recall the law of cosines: $c^2 = a^2 + b^2 - 2ab\cos\theta$.

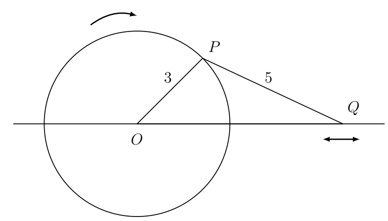


Figure 15.12: Figure for Exercise 15; rotating wheel.

15.16. **Sailing ship.** A ship sails away from a harbor at a constant speed v. The total height of the ship including its mast is h. See Figure 15.13.

(a) At what distance away does the ship disappear below the horizon?

(b) At what rate does the top of the mast appear to drop toward the horizon just before this?

Note: in ancient times this effect lead people to conjecture that the earth is round (radius R), a fact which you need to take into account.

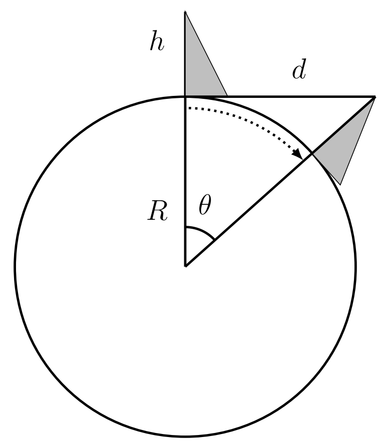


Figure 15.13: Figure for Exercise 16; ship sailing away.

15.17. **Implicit differentiation.** Find $\frac{dy}{dx}$ using implicit differentiation.

(a) $y = 2\tan(2x+y)$,

(b) $\sin y = -2\cos x$,

(c) $x\sin y + y\sin x = 1$.

15.18. **Equation of a tangent line.** Use implicit differentiation to find the equation of the tangent line to the following curve at the point $(1,1)$:

$$x\sin(xy - y^2) = x^2 - 1$$

15.19. **Implicit differentiation and** arcsin**.** The function $y = \arcsin(ax)$ is a so-called *inverse trigonometric function*. It expresses the same relationship as does the equation $ax = \sin(y)$.

Note: this function is defined only for values of x between $1/a$ and $-1/a$.

Use implicit differentiation to find y'.

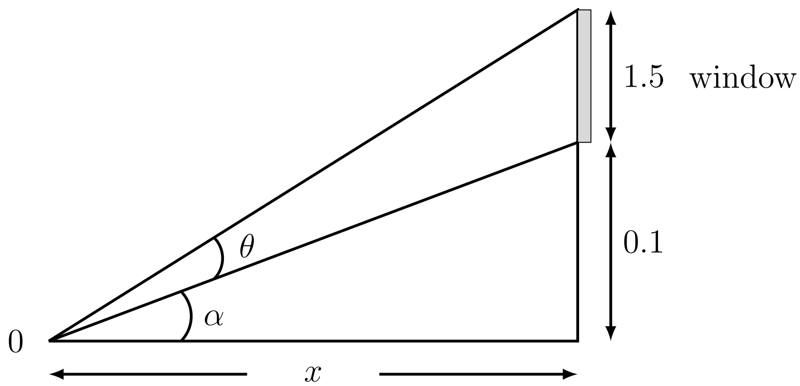


Figure 15.14: Figure for Exercise 20; distance from a window.

15.20. **Best view.**Your room has a window whose height is 1.5m. The bottom edge of the window is 10cm above your eye level, as depicted in Figure 15.14. How far away from the window should you stand to get the best view?

Note: "best view" means the largest visual angle, i.e. angle between the lines of sight to the bottom and to the top of the window.

15.21. **Fireworks.** You are directly below English Bay during a summer fireworks event and looking straight up. A single fireworks explosion occurs directly overhead at a height of 500m as depicted in Figure 15.15. The rate of change of the radius of the flare is 100m/sec.

Assuming that the flare is a circular disk parallel to the ground (with its centre directly overhead), what is the rate of change of the visual angle at the eye of an observer on the ground at the instant that the radius of the disk is $r = 100$ meters?

Note: the visual angle is the angle between the vertical direction and the line between the edge of the disk and the observer.

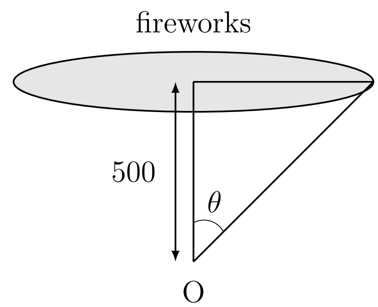


Figure 15.15: Figure for Exercise 21; fireworks overhead.

15.22. **Differential equations and their solutions.** Match the differential equations given in parts (i-iv) with the functions in (a-f) which are solutions for them.

Differential equations:

(i) $d^2y/dt^2 = 4y$

(ii) $d^2y/dt^2 = -4y$

(iii) $dy/dt = 4y$

(iv) $dy/dt = -4y$

Solutions:

(a) $y(t) = 4\cos(t)$

(b) $y(t) = 2\cos(2t)$

(c) $y(t) = 4e^{-2t}$

(d) $y(t) = 5e^{2t}$

(e) $y(t) = \sin(2t) - \cos(2t)$,

(f) $y(t) = 2e^{-4t}$.

Note: each differential equation may have more than one solution

15.23. **Periodic motion.**

(a) Show that the function $y(t) = A\cos(wt)$ satisfies the differential equation

$$\frac{d^2y}{dt^2} = -w^2 y$$

where $w > 0$ is a constant, and A is an arbitrary constant.

Note: w corresponds to the *frequency* and A to the *amplitude* of an oscillation represented by the cosine function.

(b) It can be shown using Newton's laws of motion that the motion of a pendulum is governed by a differential equation of the form

$$\frac{d^2y}{dt^2} = -\frac{g}{L}\sin(y),$$

where L is the length of the string, g is the acceleration due to gravity (both positive constants), and $y(t)$ is displacement of the pendulum from the vertical.

What property of the sine function is used when this equation is approximated by the Linear Pendulum Equation:

$$\frac{d^2y}{dt^2} = -\frac{g}{L}y.$$

(c) Based on this Linear Pendulum Equation, what function would represent the oscillations? What would be the frequency of the oscillations?

(d) What happens to the frequency of the oscillations if the length of the string is doubled?

15.24. **Jack and Jill.** Jack and Jill have an on-again off-again love affair. The sum of their love for one another is given by the function

$$y(t) = \sin(2t) + \cos(2t).$$

(a) Find the times when their total love is at a maximum.

(b) Find the times when they dislike each other the most.

15.25. **Differential equations and critical points.** Let

$$y = f(t) = e^{-t}\sin t, \ -\infty < t < \infty.$$

(a) Show that y satisfies the differential equation $y'' + 2y' + 2y = 0$.

(b) Find all critical points of $f(t)$.

16
Additional exercises

Exercises

16.1. **Multiple choice.**

(1) The equation of the tangent line to the function $y = f(x)$ at the point x_0 is

(a) $y = f'(x_0) + f(x_0)(x - x_0)$

(b) $y = x_0 + f(x_0)/f'(x_0)$

(c) $y = f(x) - f'(x)(x - x_0)$

(d) $y = f(x_0) + f'(x_0)(x - x_0)$

(e) $y = f(x_0) - f'(x_0)(x - x_0)$

(2) The functions $f(x) = x^2$ and $g(x) = x^3$ are equal at $x = 0$ and at $x = 1$. Between $x = 0$ and at $x = 1$, for what value of x are their graphs furthest apart?

(a) $x = 1/2$

(b) $x = 2/3$

(c) $x = 1/3$

(d) $x = 1/4$

(e) $x = 3/4$

(3) Consider a point in the first quadrant on the hyperbola $x^2 - y^2 = 1$ with $x = 2$. The slope of the tangent line at that point is

(a) $2/\sqrt{3}$

(b) $2/\sqrt{5}$

(c) $1/\sqrt{3}$

(d) $\sqrt{5}/2$

(e) $2/3$

(4) For $a, b > 0$, solving the equation $\ln(x) = 2\ln(a) - 3\ln(b)$ for x leads to

(a) $x = e^{2a-3b}$

(b) $x = 2a - 3b$

(c) $x = a^2/b^3$

(d) $x = a^2b^3$

(e) $x = (a/b)^6$

(5) The function $y = f(x) = \arctan(x) - (x/2)$ has local maXima (LX), local minima (LM) and inflection points(IP) as follows:

(a) LX: $x = 1$, LM: $x = -1$, IP: $x = 0$.

(b) LX: $x = -1$, LM: $x = 1$, IP: $x = 0$.

(c) LX: $x = -1$, LM: $x = 1$, IP: none.

(d) LX: $x = \sqrt{3}$, LM: $x = -\sqrt{3}$, IP: $x = 0$.

(e) LX: $x = -\sqrt{3}$, LM: $x = \sqrt{3}$, IP: $x = 0$.

(6) Consider the function $y = f(x) = 3e^{-2x} - 5e^{-4x}$.

(a) The function has a local maximum at $x = (1/2)\ln(10/3)$.

(b) The function has a local minimum at $x = (1/2)\ln(10/3)$.

(c) The function has a local maximum at $x = (-1/2)\ln(3/5)$.

(d) The function has a local minimum at $x = (1/2)\ln(3/5)$.

(e) The function has a local maximum at $x = (-1/2)\ln(3/20)$.

(7) Let m_1 be the slope of the function $y = 3^x$ at the point $x = 0$ and let m_2 be the slope of the function $y = \log_3 x$ at $x = 1$ Then

(a) $m_1 = \ln(3)m_2$

(b) $m_1 = m_2$

(c) $m_1 = -m_2$

(d) $m_1 = 1/m_2$

(e) $m_1 = m_2/\ln(3)$

(8) Consider the curve whose equation is $x^4 + y^4 + 3xy = 5$. The slope of the tangent line, dy/dx, at the point $(1,1)$ is

(a) 1

(b) -1

(c) 0

(d) -4/7

(e) 1/7

(9) Two kinds of bacteria are found in a sample of tainted food. It is found that the population size of type 1, N_1 and of type 2, N_2 satisfy the equations

$$\frac{dN_1}{dt} = -0.2N_1, \qquad N_1(0) = 1000,$$
$$\frac{dN_2}{dt} = 0.8N_2, \qquad N_2(0) = 10.$$

Then the population sizes are equal $N_1 = N_2$ at the following time:

(a) $t = \ln(40)$

(b) $t = \ln(60)$

(c) $t = \ln(80)$

(d) $t = \ln(90)$

(e) $t = \ln(100)$

(10) In a conical pile of sand the ratio of the height to the base radius is always $r/h = 3$. If the volume is increasing at rate 3 m^3/min, how fast (in m/min) is the height changing when $h = 2$m?

(a) $1/(12\pi)$

(b) $(1/\pi)^{1/3}$

(c) $27/(4\pi)$

(d) $1/(4\pi)$

(e) $1/(36\pi)$

Formula.
Note that the volume of a cone with height h and radius r is $V = (\pi/3)r^2 h$.)

(11) Shown in Figure 16.1 is a function and its tangent line at $x = x_0$. The tangent line intersects the x axis at the point $x = x_1$. Based on this figure, the coordinate of the point x_1 is

(a) $x_1 = x_0 + \frac{f(x_0)}{f'(x_0)}$,

(b) $x_1 = x_0 - f'(x_0)(x - x_0)$,

(c) $x_1 = x_0 - \frac{f'(x_1)}{f(x_1)}$

(d) $x_1 = x_0 + \frac{f'(x_1)}{f(x_1)}$,

(e) $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$

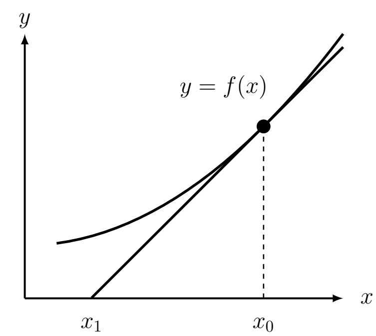


Figure 16.1: Graph for Exercise 1; a function and its tangent line.

(12) Consider Euler's method. For the differential equation and initial condition

$$\frac{dy}{dt} = (2 - y), \quad y(0) = 1$$

using one time step of size $\Delta t = 0.1$ leads to which value of the solution at time $t = 0.1$?

(a) $y(0.1) = 2$,

(b) $y(0.1) = 2.1$,

(c) $y(0.1) = 2.2$,

(d) $y(0.1) = 1.2$,

(e) $y(0.1) = 1.1$,

(13) Consider the function $y = \cos(x)$ and its tangent line to this function at the point $x = \pi/2$. Using that tangent line as a linear approximation of the function would lead to

(a) Overestimating the value of the actual function for any nearby x.

(b) Underestimating the value of the actual function for any nearby x.

(c) Overestimating the function when $x > \pi/2$ and underestimating the function when $x < \pi/2$.

(d) Overestimating the function when $x < \pi/2$ and underestimating the function when $x > \pi/2$.

(e) Overestimating the function when $x < 0$ and underestimating the function when $x > 0$.

16.2. **Related Rates.** Two spherical balloons are connected so that one inflates as the other deflates, the sum of their volumes remaining constant. When the first balloon has radius 10 cm and its radius is increasing at 3 cm/sec, the second balloon has radius 20 cm. What is the rate of change of the radius of the second balloon?

Formula.
Note that the volume of a sphere of radius r is $V = (4/3)\pi r^3$.

16.3. **Particle velocity.** A particle is moving along the x axis so that its distance from the origin at time t is given by

$$x(t) = (t+2)^3 + \lambda t$$

where λ is a constant

(a) Determine the velocity $v(t)$ and the acceleration $a(t)$.

(b) Determine the minimum velocity over all time.

16.4. **Motion.** A particle's motion is described by $y(t) = t^3 - 6t^2 + 9t$ where $y(t)$ is the *displacement* (in metres) t is time (in seconds) and $0 \le t \le 4$ seconds.

(a) During this time interval, when is the particle furthest from its initial position ?

(b) During this time interval, what is the greatest speed of the particle?

(c) What is the total *distance* (including both forward and backward directions) that the particle has travelled during this time interval?

16.5. **Falling object.** Consider an object thrown upwards with initial velocity $v_0 > 0$ and initial height $h_0 > 0$. Then the height of the object at time t is given by

$$y = f(t) = -\frac{1}{2}gt^2 + v_0 t + h_0.$$

Find critical points of $f(t)$ and use both the second and first derivative tests to establish that this is a local maximum.

16.6. **Linear approximation.** Find a linear approximation to the function $y = x^2$ at the point whose x coordinate is $x = 2$. Use your result to approximate the value of $(2.0001)^2$.

16.7. **HIV virus.** Initially, a patient has 1000 copies of the virus. How long does it take until the HIV infection is detectable? Assume that the number of virus particles y grows according to the equation

$$\frac{dy}{dt} = 0.05y$$

where t is time in days, and that the smallest detectable viral load is $350,000$ particles. Leave your answer in terms of logarithms.

16.8. **Fish generations.** In Fish River, the number of salmon (in thousands), x, in a given year is linked to the number of salmon (in thousands), y, in the following year by the function

$$y = Axe^{-bx}$$

where $A, b > 0$ are constants.

(a) For what number of salmon is there no change in the number from one year to the next?

(b) Find the number of salmon that would yield the largest number of salmon in the following year.

16.9. **Polynomial.** Find a polynomial of third degree that has a local maximum a $x = 1$, a zero and an inflection point at $x = 0$, and goes through the point $(1,2)$ (*hint*: assume $p(x) = ax^3 + bx^2 + cx + d$ and find the values of a, b, c, d).

16.10. **Critical points.**

(a) Find critical points for the function $y = e^x(1 - \ln(x))$ for $0.1 \leq x \leq 2$ and classify their types.

(b) The function $y = \ln(x) - e^x$ has a critical point in the interval $0.1 \leq x \leq 2$. It is not possible to solve for the value of x at that point, but it is possible to find out what kind of critical point that is. Determine whether that point is a local maximum, minimum, or inflection point.

16.11. **Lennard-Jones potential.** The Lennard-Jones potential, $V(x)$ is the potential energy associated with two uncharged molecules a distance x apart, and is given by the formula

$$V(x) = \frac{a}{x^{12}} - \frac{b}{x^6}$$

where $a, b > 0$. Molecules would tend to adjust their separation distance so as to minimize this potential. Find any local maxima or minima of this potential. Find the distance between the molecules, x, at which $V(x)$ is minimized and use the second derivative test to verify that this is a local minimum.

16.12. **Rectangle inscribed in a circle.** Find the dimensions of the largest rectangle that can fit exactly into a circle whose radius is r.

16.13. **Race track.** Figure 16.2 shows a 1 km race track with circular ends. Find the values of x and y that maximize the area of the rectangle.

16.14. **Leaf shape.** Now suppose that Figure 16.2 shows the shape of a leaf of some plant. If the plant grows so that x increases at the rate 2 cm/year and y increases at the rate 1 cm/year, at what rate is the leaf's entire area increasing?

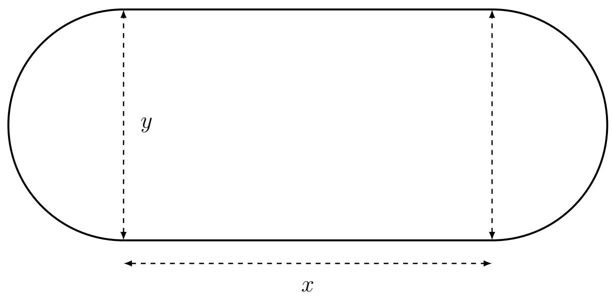


Figure 16.2: This shape is investigated in both Exercise 13 and 14; maximizing area or a racetrack, rate of growth of a leaf.

16.15. **Shape of E. coli.** A cell of the bacterium E. coli has the shape of a cylinder with two hemispherical caps, as shown in Figure 16.3. Consider this shape, with h the height of the cylinder, and r the radius of the cylinder and hemispheres.

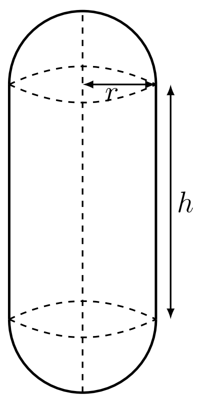


Figure 16.3: Shape of the cells described in Exercises 15 and 16

(a) Find the values of r and h that lead to the largest volume for a fixed constant surface area, S= constant.

(b) Describe or sketch the shape you found in (a).

(c) A typical E. coli cell has $h = 1\mu$m and $r = 0.5\mu$m. Based on your results in (a) and (b), would you agree that E. coli has a shape that maximizes its volume for a fixed surface area? (Explain your answer).

Formula.
Note that a hemisphere of radius r has volume $V = (2/3)\pi r^3$ and surface area $S = 2\pi r^2$.
For a cylinder: $V = \pi r^2 h$ and $S = 2\pi rh$.

16.16. **Changing cell shape.** If the cell shown in Figure 16.3 is growing so that the height increases twice as fast as the radius, and the radius is growing at 1 μm per day, at what rate does the volume of the cell increase? Leave your answer in terms of the height and radius of the cell.

16.17. **Minima and Maxima.**

(a) Consider the polynomial $y = 4x^5 - 15x^4$. Find all local minima maxima, and inflection points for this function.

(b) Find the global minimum and maximum for this function on the interval $[-1,1]$.

16.18. **Minima and Maxima.** Consider the polynomial $y = -x^5 - x^4 + 3x^3$. Use calculus to find all local minima maxima, and inflection points for this function.

16.19. **Growth of vine.** A vine grows up a tree in the form of a helix as shown on the left in Figure 16.4. If the length of the vine increases at a constant rate α cm/day, at what rate is the height of its growing tip increasing? Assume that the radius of the tree is r and the pitch of the helix (i.e. height increase for each complete turn of the helix) is p, a positive constant. Note that the right panel in Figure 16.4 shows the unwrapped cylinder, with the vine's location along it.

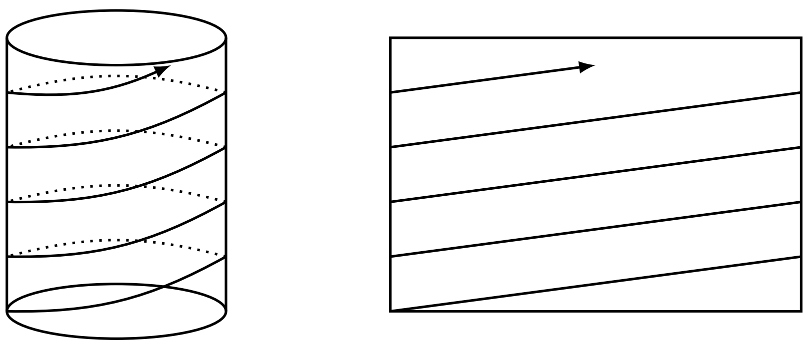


Figure 16.4: Growth of a vine in the shape of a spiral for Exercise 19.

16.20. **Newton's law of cooling.** Newton's Law of cooling leads to a differential equation that predicts the temperature $T(t)$ of an object whose initial temperature is T_0 in an environment whose temperature is E. The predicted temperature is given by $T(t) = E + (T_0 - E)e^{-kt}$ where t is time and k is a constant. Shown in Figure 16.5 is some data points plotted as $\ln(T(t) - E)$ versus time in minutes. The ambient temperature was $E = 22°$ C. Also shown on the graph is the line that best fits those 11 points. Find the value of the constant k.

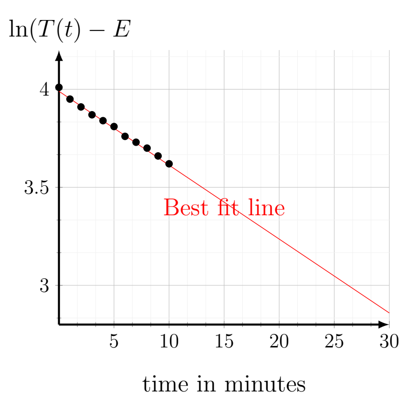


Figure 16.5: Figure for Exercise 20; Newton's law of cooling.

16.21. **Blood alcohol.** Blood alcohol level (BAL), the amount of alcohol in your blood stream (here represented by $B(t)$, is measured in milligrams of alcohol per 10 millilitres of blood. At the end of a party (time $t = 0$), a drinker is found to have $B(0) = 0.08$ (the legal level for driving impairment), and after that time, $B(t)$ satisfies the differential equation

$$\frac{dB}{dt} = -kB, \quad k > 0$$

where k is a constant that represents the rate of removal of alcohol form the blood stream by the liver.

(a) If the drinker had waited for 3 hrs before driving (until = 3), his BAL would have dropped to 0.04. Determine the value of the rate constant k (specifying appropriate units) for this drinker.

(b) According to the model, how much longer would it take for the BAL to drop to 0.01?

16.22. **Population with immigration.** An island has a bird population of density $P(t)$. New birds arrive continually with a constant colonization rate C birds per day. Each bird also has a constant probability per day, γ, of leaving the island. At time $t = 0$ the bird population is $P(0) = P_0$

(a) Write a differential equation that describes the rate of change of the bird population on the island.

(b) Find the steady state of that equation and interpret this in terms of the bird population.

(c) Give the solution of the differential equation you found in (b) and show that it satisfies the following two properties:

(i) the initial condition,

(ii) as $t \to \infty$ it approaches the steady state you found in (b).

(d) If the island has no birds on it at time $t = 0$, how long would it take for the bird population to grow to 80% of the steady state value?

16.23. **Learning.**

(a) It takes you 1 hrs (total) to travel to and from SFU every day to study Philosophy 101. The amount of new learning (in arbitrary units) that you can get by spending t hours at the university is given approximately by

$$L_P(t) = \frac{10t}{9+t}.$$

How long should you stay at SFU on a given day if you want to maximize your learning per time spent?

Note: time spent includes travel time.

(b) If you take Math 10000 instead of Philosophy, your learning at time t is

$$L_M(t) = t^2.$$

How long should you stay at SFU to maximize your learning in that case?

16.24. **Learning and forgetting.** Knowledge can be acquired by studying, but it is forgotten over time A simple model for learning represents the amount of knowledge, $y(t)$, that a person has at time t (in years) by a differential equation

$$\frac{dy}{dt} = S - fy,$$

where $S \geq 0$ is the rate of studying and $f \geq 0$ is the rate of forgetting. We assume that S and f are constants that are different for each person.

Note: your answers to the following questions will contain constants such as S or f.

(a) Mary never forgets anything. What does this imply about the constants S and f? Mary starts studying in school at time $t = 0$ with no knowledge at all. How much knowledge does she have after 4 years (i.e. at $t = 4$)?

(b) Tom learned so much in preschool that his knowledge when entering school at time $t = 0$ is $y = 100$. However, once Tom is in school, he stops studying completely. What does this imply about the constants S and f? How long does it take him to forget 75 % of what he knew?

(c) Jane studies at the rate of 10 units per year and forgets at rate of 0.2 per year. Sketch a "direction field" ("slope field") for the differential equation describing Jane's knowledge. Add a few curves $y(t)$ to show how Jane's knowledge changes with time.

16.25. **Least cost.** A rectangular plot of land has dimensions L by D as depicted in Figure 16.6. A pipe is to be built joining points A and C. The pipe can be above ground along the border of the plot (Section AB), but has to be buried underground along the segment BC. The cost per unit length of the underground portion is 3 times that of the cost of the above ground portion. Determine the distance y so that the cost of the pipe is as low as possible.

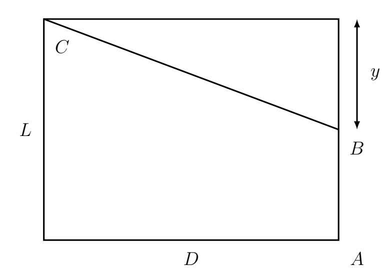


Figure 16.6: Figure for Exercise 25; least cost for installing a pipe.

16.26. **Least heat loss.** In an effort to increase sustainability, the university is aiming to use the shortest length of pipes to connect the buildings (labeled B) with the steam power plant (labeled P) in Figure 16.7 in order reduce loss of heat from the pipes to the surroundings.

(a) Set up a model and describe how you would solve this problem. Assume that the four sites are located at the corners of a square of side length L.

(b) How long is the shortest total pipe length connecting P to the 3 other corners?

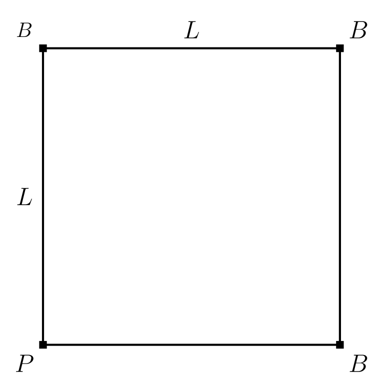


Figure 16.7: Figure for Exercise 26; least heat loss.

16.27. **Logistic equation and its solution.**

(a) Show that the function

$$y(t) = \frac{1}{1+e^{-t}}$$

satisfies the differential equation

$$\frac{dy}{dt} = y(1-y).$$

(b) What is the initial value of y at $t = 0$?

(c) For what value of y is the growth rate largest?

(d) What happens to y after a very long time?

16.28. **Human growth.** Given a population of 6 billion people on Planet Earth, and using the approximate growth rate of $r = 0.0125$ per year, how long ago was this population only 1 million? Assume that the growth has been the same throughout history (which is not actually true).

16.29. **Ducks in a row.** Graduate student Ryan Lukeman studies behaviour of duck flocks swimming near Canada Place in Vancouver, BC. This figure from his PhD thesis shows his photography set-up. Here $H = 10$ meters is the height from sea level up to his camera aperture at the observation point, $D = 2$ meters is the width of a pier (a stationary platform whose size is fixed), and x is the distance from the pier to the leading duck in the flock (in meters). A visual angle subtended at the camera is shown as α.

If the visual angle is increasing at the rate of $1/100$ radians per second, at what rate is the distance x changing at the instant that $x = 3$ meters?

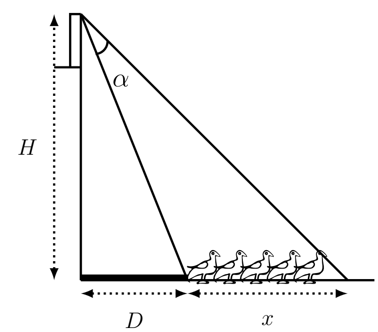


Figure 16.8: Figure for Exercise 29; ducks in a row.

16.30. **Circular race track.** Two runners are running around a circular race track whose length is 400 m, as shown in Figure 16.9(a). The first runner makes a full revolution every 100 s and the second runner every 150 s. They start at the same time at the start position, and the angles subtended by each runner with the radius of the start position are $\theta_1(t), \theta_2(t)$, respectively. As the runners go around the track both $\theta_1(t)$ and $\theta_2(t)$ change with time.

(a) At what rate it the angle $\phi = \theta_1 - \theta_2$ changing?

(b) What is the angle ϕ at $t = 25$s?

(c) What is the distance between the runners at $t = 25$s?

Note: here "distance" refers to the length of the straight line connecting the runners.

(d) At what rate is the distance between the runners changing at $t = 25$s?

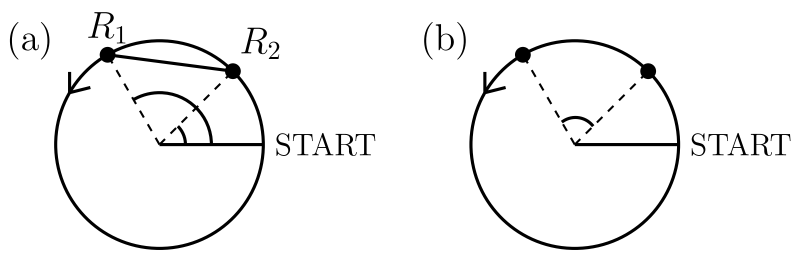


Figure 16.9: Figure for Exercises 30 and 31. The angles in (a) are $\theta_1(t), \theta_2(t)$. In (b), the angle between the runners is ϕ.

16.31. **Phase angle and synchrony.** Suppose that the same two runners as in Exercise 30 would speed up or slow down depending on the angle between them, ϕ (see Figure 16.9). Then $\phi = \phi(t)$ changes with time. We assume that the angle ϕ satisfies a differential equation of the form

$$\frac{d\phi}{dt} = A - B\sin(\phi)$$

where $A, B > 0$ are constants.

(a) What values of ϕ correspond to steady states (i.e. constant solutions) of this differential equation?

(b) What restriction should be placed on the constants A, B for these steady states to exist?

(c) Suppose $A = 1, B = 2$. Sketch the graph of $f(\phi) = A - B\sin(\phi)$ fo $-\pi \le \phi \le \pi$ and use it to determine what happens if the two runners start at the same point, ($\phi = 0$) at time $t = 0$.

16.32. **Tumor mass.** Figure 16.10 (not drawn to scale) shows a tumor mass containing a necrotic (dead) core (radius r_2), surrounded by a layer of actively dividing tumor cells. The entire tumor can be assumed to be spherical, and the core is also spherical.

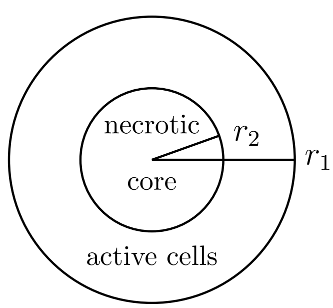


Figure 16.10: Figure for Exercise 32; tumor mass.

(a) If the necrotic core increases at the rate 3 cm^3/year and the volume of the active cells increases by 4 cm^3/year, at what rate is the outer radius of the tumor (r_1) changing when $r_1 = 1$ cm. (Leave your answer as a fraction in terms of π; indicate units with your answer.)

Formula.
Note that the volume and surface area of a sphere are $V = (4/3)\pi r^3$, $S = 4\pi r^2$.

(b) At what rate (in cm^2/yr) does the outer surface area of the tumor increase when $r_1 = 1\text{cm}$?

16.33. **Blood vessel branching.** Shown in Figure 16.11 is a major artery, (radius R) and one of its branches (radius r). A labeled schematic diagram is also shown (right). The length $0A$ is L, and the distance between 0 and P is d, where $0P$ is perpendicular to $0A$. The location of the branch point (B) is to be determined so that the total resistance to blood flow in the path ABP is as small as possible. (R, r, d, L are positive constants, and $R > r$.)

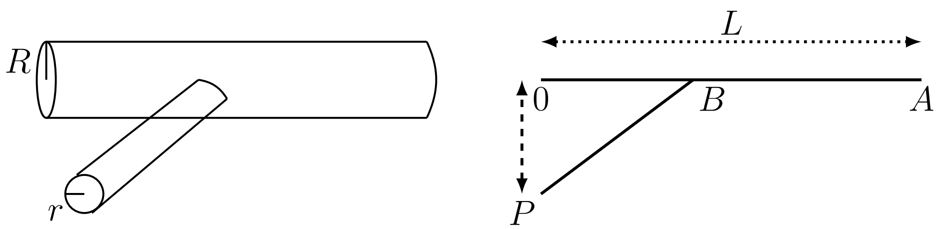


Figure 16.11: Figure for Exercise 33; blood vessel branching.

(a) Let the distance between 0 and B be x. What is the length of the segment BA and what is the length of the segment BP?

(b) *The resistance of any blood vessel is proportional to its length and inversely proportional to its radius to the fourth power* Based on this fact, what is the resistance, T_1, of segment BA and what is the resistance, T_2, of the segment BP?

(c) Find the value of the variable x for which the total resistance, $T(x) = T_1 + T_2$ is a minimum.

Formula.
Note that "z is inversely proportional to y" means that $z = k/y$ for some constant k.

16.34. **Implicit differentiation.** A surface that looks like an "egg carton" as depicted in Figure 16.12(a) can be described by the function

$$z = \sin(x)\cos(y).$$

The intersection of this surface with the plane $z = 1/2$ is a curve (also called a **level curve**). One such level curve is shown in Figure 16.12(b).

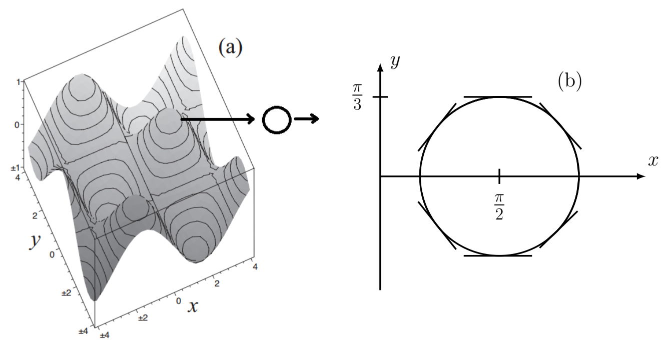


Figure 16.12: (a) The surface $\sin(x)\cos(y) = z$. (b) One level curve for this surface is shown on its own, and then enlarged, with some tangent lines. The tangent line to this level curve at the point $(\pi/2, \pi/3)$ is horizontal.

(a) Use implicit differentiation to find the slope of the tangent line to a point on such a curve.

(b) Find the slope of the tangent line to the same level curve at the point $x = \frac{\pi}{2}$.

(c) Find the slope of the tangent line to the same level curve at the point $x = \frac{\pi}{4}$.

Appendices

A

A review of Straight Lines

A.1 Geometric ideas: lines, slopes, equations

Straight lines have some important geometric properties, namely:

The slope of a straight line is the same everywhere along its length.

Slope of a straight line. We define the slope of a straight line as follows:

$$\text{Slope} = \frac{\Delta y}{\Delta x}$$

where Δy means "change in the y value" and Δx means "change in the x value" between two points. See Figure A.1 for what this notation represents.

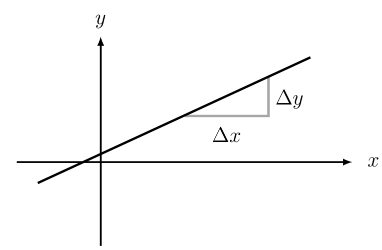


Figure A.1: The slope of a line (usually given the symbol m) is the ratio of the change in the y value, Δy to the change in the x value, Δx.

Equation of a straight line. Using this basic geometric property, we can find the equation of a straight line given any of the following information about the line:

- The y-intercept, b, and the slope, m:

$$y = mx + b.$$

- A point (x_0, y_0) on the line, and the slope, m, of the line:

$$\frac{y - y_0}{x - x_0} = m$$

- Two points on the line, say (x_1, y_1) and (x_2, y_2):

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1}$$

Note: any of these can be rearranged or simplified to produce the standard form $y = mx + b$, as discussed in the problem set.

The following examples serve as a refresher on finding the equation of the line that satisfies each of the given conditions.

Example A.1 *In each case write the equation of the straight line that satisfies the given statements.*

Note: *you should also be able to easily sketch the line in each case.*

(a) *The line has slope* 2 *and* y*-intercept* 4.

(b) *The line goes through the points* $(1,1)$ *and* $(3,-2)$.

(c) *The line has* y*-intercept -1 and* x*-intercept* 3.

(d) *The line has slope* -1 *and goes through the point* $(-2,-5)$.

Solution.

(a) We can use the standard form of the equation of a straight line, $y = mx + b$ where m is the slope and b is the y-intercept to obtain the equation: $y = 2x + 4$.

(b) The line goes through the points $(1,1)$ and $(3,-2)$. We use the fact that the slope is the same all along the line. Thus,

$$\frac{(y-y_0)}{(x-x_0)} = \frac{(y_1-y_0)}{(x_1-x_0)} = m.$$

Substituting in the values $(x_0, y_0) = (1,1)$ and $(x_1, y_1) = (3,-2)$,

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

This tells us that the slope is $m = -3/2$. We find that

$$y - 1 = -\frac{3}{2}(x-1) = -\frac{3}{2}x + \frac{3}{2}, \quad \Rightarrow \quad y = -\frac{3}{2}x + \frac{5}{2}.$$

(c) The line has y-intercept -1 and x-intercept 3, i.e. goes through the points $(0,-1)$ and $(3,0)$. We can use the method in (b) to get

$$y = \frac{1}{3}x - 1$$

Alternately, as a shortcut, we could find the slope,

$$m = \frac{\Delta y}{\Delta x} = \frac{1}{3}.$$

Note: Δ means "change in the value", i.e. $\Delta y = y_1 - y_0$.

Thus $m = 1/3$ and $b = -1$ (y-intercept), leading to the same result.

(d) The line has slope -1 and goes through the point $(-2,-5)$. Then,

$$\frac{(y+5)}{(x+2)} = -1, \quad \Rightarrow \quad y + 5 = -1(x+2) = -x - 2, \quad \Rightarrow \quad y = -x - 7.$$

B

A precalculus review

B.1 Manipulating exponents

Recall: $2^n = 2 \cdot 2 \ldots 2$ (with n factors of 2). This means:

$$2^n \cdot 2^m = \underbrace{(2 \cdot 2 \ldots 2)}_{n \text{ factors}} \cdot \underbrace{(2 \cdot 2 \ldots 2)}_{m \text{ factors}} = \underbrace{2 \cdot 2 \cdot \ldots \cdots 2}_{n+m \text{ factors}} = 2^{n+m}.$$

Similarly, we can derive many properties of manipulations of exponents. A list of these appears below, and holds for any positive base a.

1. $2^a 2^b = 2^{a+b}$ as with all similar exponent manipulations.
2. $(2^a)^b = 2^{ab}$ also stems from simple rules for manipulating exponents.
3. 2^x is a function that is defined, continuous, and differentiable for all real values x.
4. $2^x > 0$ for all values of x.
5. We define $2^0 = 1$, and we also have that $2^1 = 2$.
6. $2^x \to 0$ for increasingly negative values of x.
7. $2^x \to \infty$ for increasing positive values of x.

B.2 Manipulating logarithms

The following properties hold for logarithms of *any* base (we used base 2 in our previous section and keep the same base here). Properties of the logarithm stem directly from properties of the exponential function, and include the following:

1. $\log_2(ab) = \log_2(a) + \log_2(b)$.
2. $\log_2(a^b) = b \log_2(a)$.
3. $\log_2(1/a) = \log_2(a^{-1}) = -\log_2(a)$.

C
A Review of Simple Functions

We review a few basic concepts related to functions.

C.1 What is a function?

A function is just a way of expressing a special relationship between a value we consider as the input ("x") value and an associated output ("y") value. We write this relationship in the form

$$y = f(x)$$

to indicate that y depends on x. The only constraint on this relationship is that, for every value of x we can get *at most one* value of y. This is equivalent to the *"vertical line property"*: the graph of a function can intersect a vertical line at most at one point. The set of all allowable x values is called the *domain* of the function, and the set of all resulting values of y are the *range*.

Naturally, we do not always use the symbols x and y to represent independent and dependent variables. For example, the relationship

$$V = \frac{4}{3}\pi r^3$$

expresses a functional connection between the radius, r, and the volume, V, of a sphere. We say in such a case that "V is a function of r".

All the sketches shown in Figure C.2 are valid functions. The first is merely a collection of points, x values and associated y values. The second a histogram. The third sketch is meant to represent the collection of smooth continuous functions - those of most interest to usnin the study of calculus. On the other hand, the example shown in Figure C.1 is not the graph of a function. We see that a vertical line intersects this curve at more than one point. This is not permitted since a given value of x may have at most one corresponding value of y.

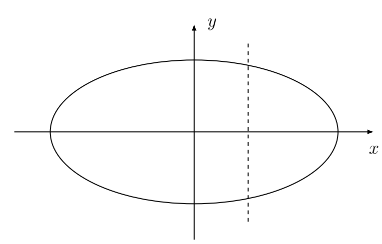


Figure C.1: The above elliptical curve cannot be the graph of a function. The dashed vertical line intersects the graph at more than one point: this means that a given value of x corresponds to "too many" values of y. If we restrict ourselves to the top part of the ellipse only (or the bottom part only), then we can create a function which has the corresponding graph.

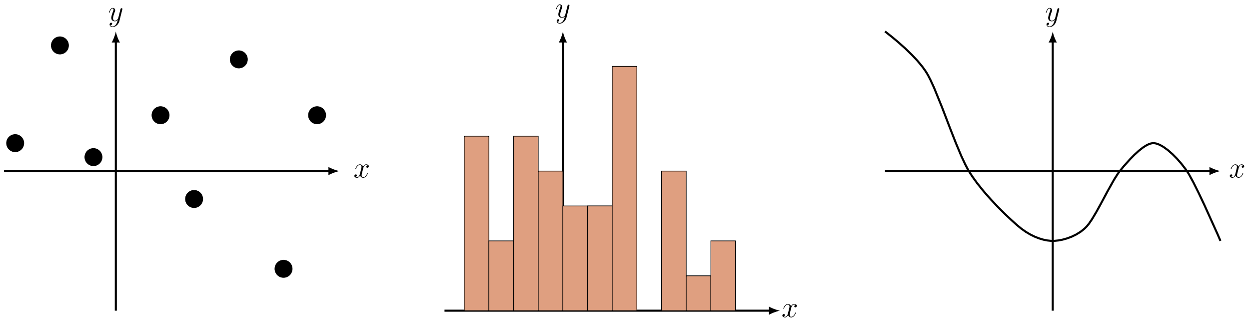
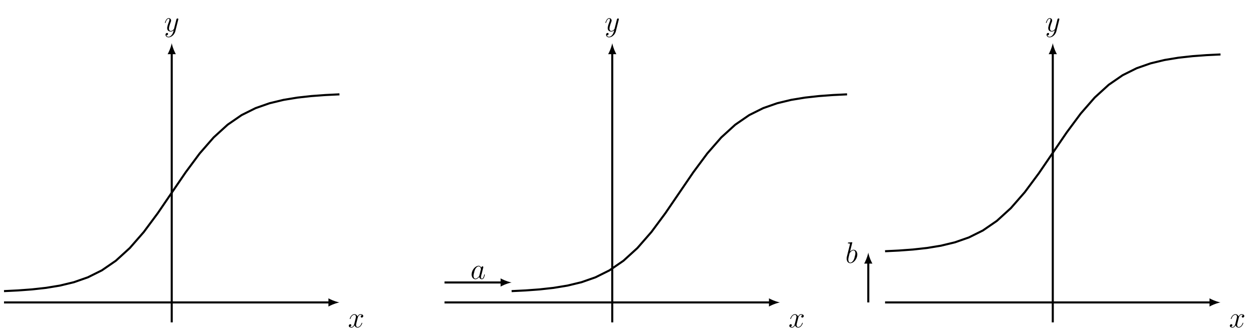


Figure C.2: All the examples above represent functions.

C.2 Geometric transformations

It is helpful to recognize what happens to the graph of a function when the relationship between the variables is changed slightly. Often this is called *applying a transformation*. In Figure C.3 we illustrate (a) an original function $f(x)$, (b) the function $f(x-a)$ which shifts f to the right along the positive x-axis by a distance a, and (c) the function $f(x)+b$ shifting f up the y-axis by height b.



(a) $y = f(x)$ (b) $y = f(x-a)$ (c) $y = f(x)+b$

Figure C.3: Shifting the graph of a function horizontally and vertically

Figure C.4 illustrates what happens to a function when shifts, scaling, or reflections occur: a function $y = f(x)$ is shown with a black solid line. On the same graph are superimposed the reflections of this graph about the x axis, $y = -f(x)$ (dashed black), about the y axis $y = f(-x)$ (red), and about the y and the x-axis, $y = -f(-x)$ (red dashed). The latter is equivalent to a rotation of the original graph about the origin.

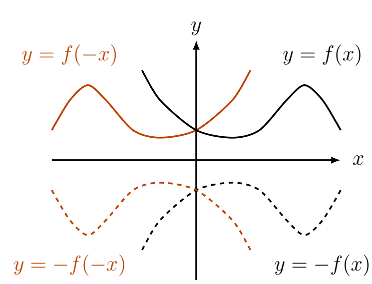


Figure C.4: A function and its reflections about x and y-axes.

C.3 Classifying

While life offers amazing complexity, one way to study living things is to classify them into related groups. A biologist looking at animals might group them according to certain functional properties - being warm blooded, being mammals, having fur or claws, or having some other interesting characteristic. In the same way, mathematicians often classify the objects they study (e.g. functions) into related groups.

An example of way to group functions into very broad classes is also shown in Figure C.5.

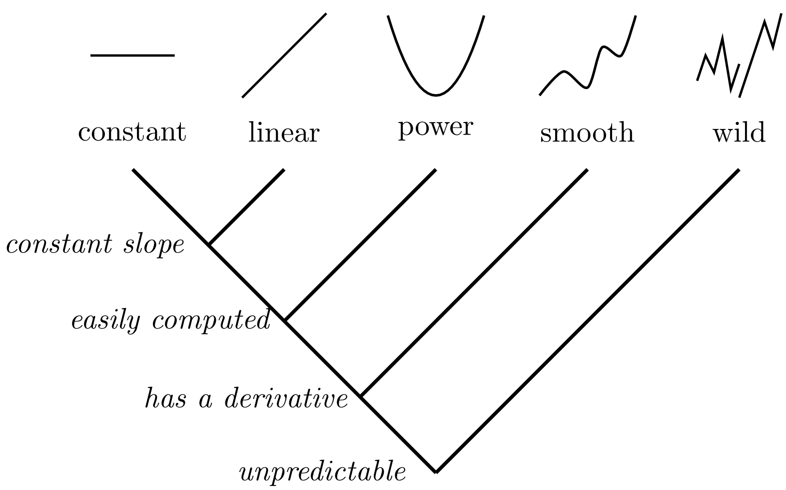


Figure C.5: Classifying functions according to their properties.

From left to right, the complexity of behaviour in this chart grows: at left, we see constant and linear functions; these are "most convenient" or simplest to describe: one or two parameters suffice (e.g. intercept or slope). Further to the right are smooth and continuous functions, while rightmost some more irregular, discontinuous function represents those that are outside the group of the "well-behaved".

In Section C.4 we study examples along this spectrum. Towards this end, we describe properties they share, properties they inherit form their "cousins," and new characteristics that appear at distinct branches.

C.4 Power functions and symmetry

In this section we list some features of each family of power functions.

Even integer powers. For $n = 2, 4, 6, 8\ldots$ the shape of the graph of $y = x^n$ is as shown in Figure 1.4(a). Note the following characteristics of these graphs:

1. The graphs of all the even power functions intersect at $x = 0$ and at at $x = \pm 1$. The value of y corresponding to both of these is $y = +1$.

Mastered Material Check

1. What are the coordinates of the three intersection points of even power functions?

2. All graphs have a lowest point - a *minimum value* - at $x = 0$.

3. As $x \to \pm\infty$, $y \to \infty$. We equivalently say that these functions are "unbounded from above."

4. The graphs are all symmetric about the y-axis. This special type of symmetry is of interest in other types of functions (not just power functions). A function with this property is called an **even function**.

Odd integer powers. For $n = 1, 3, 5, 7, \ldots$ and other odd powers, the graphs have shapes shown in Figure 1.4(b) and the following characteristics:

Mastered Material Check

2. What are the coordinates of the three intersection points of odd power functions?

1. The graphs of the odd power functions intersect at $x = 0$ and at $x = \pm 1$.

2. None of the odd power functions have a minimum value.

3. As $x \to +\infty$, $y \to +\infty$. As $x \to -\infty$, $y \to -\infty$. The functions are "unbounded from above and below."

4. The graphs are all symmetric about the origin. This special type of symmetry is of interest in other types of functions (not just power functions). A function with this type of symmetry is called an **odd function**.

Further properties of intersections

Consider the even and odd functions graphed in Figure C.6. Notice that a horizontal line intersects the graph only once for the odd power but possibly twice for the even power.

Note: we must allow for the case when the line does not intersect at all, or that it intersects precisely at the minimum point.

These observations holds in the case of general even and odd power functions as well.

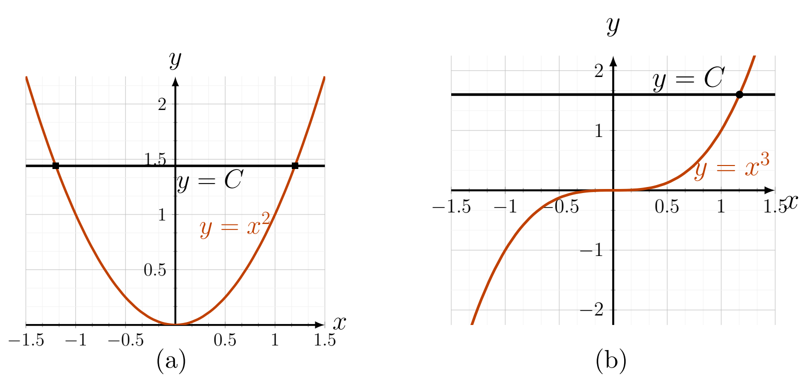


Figure C.6: The even power functions intersect a horizontal line in up to two places, while the odd power functions intersect such a line in only one place.

A horizontal line has an equation of the form $y = C$ where C is some constant. To find where it intersects the graph of a power function $y = x^n$, we solve an equation of the form

$$x^n = C. \tag{C.1}$$

To do so, take n^{th} root of both sides:

$$(x^n)^{1/n} = C^{1/n}.$$

Simplifying, using algebraic operations on powers leads to

$$(x^n)^{1/n} = x^{n/n} = x^1 = x = C^{1/n}.$$

However, we must allow that there may be more than one solution to Eqn. (C.1), as shown for some $C > 0$ in Figure C.6. This demonstrates a distinction between odd and even power functions. If n is even then the solutions to Eqn. (C.1) are

$$x = \pm C^{1/n},$$

whereas if n is odd, there is a single solution,

$$x = C^{1/n}.$$

In general, we can define even and odd *functions*.

Definition C.1 (Even and odd functions) *A function that is symmetric about the y-axis is said to be an* even *function. A function that is symmetric about the origin is said to be an* odd *function.*

Even functions satisfy the relationship

$$f(x) = f(-x).$$

Odd functions satisfy the relationship

$$f(x) = -f(-x).$$

Examples of even functions include $y = \cos(x)$, $y = -x^8$, $y = |x|$ which are all their own mirror images when reflected about the y-axis. Examples of odd functions are $y = \sin(x)$, $y = -x^3$, $y = x$. Each of these functions is its own double-reflection (about y and then x-axes).

In a later calculus course, when we compute integrals, taking these symmetries into account can help to simplify (or even avoid) calculations.

Optional: Combining even and odd functions

Not every function is either odd or even. However, if we start with symmetric functions, some manipulations can either preserve or reverse the symmetry.

Example C.1 *Show that the product of an even and an odd function is an odd function.*

Solution. Let $f(x)$ be even. Then

$$f(x) = f(-x).$$

Let $g(x)$ be an odd function. Then $g(x) = -g(-x)$. We define $h(x)$ to be the product of these two functions,

$$h(x) = f(x)g(x).$$

Using the properties of f and g,

$$f(x)g(x) = f(-x)[-g(-x)],$$

so, rearranging, we get

$$h(x) = f(x)g(x) = f(-x)[-g(-x)] = -[f(-x)g(-x)].$$

but this is just the same as $-h(-x)$. We have established that

$$h(x) = -h(-x),$$

so the new function is odd. ◊

A function is not always even or odd. Many functions are neither even nor odd. However, it *is* possible to show that given any function, $y = f(x)$, we can write it as a sum of an even and an odd function. This is left as a challenge for the reader, with the following *hint:*

Hint: suppose $f(x)$ is not an even nor an odd function. Consider defining the two associated functions:

$$f_e(x) = \frac{1}{2}(f(x) + f(-x)),$$

and

$$f_0(x) = \frac{1}{2}(f(x) - f(-x)).$$

Can you draw a sketch of what these would look like for the function given in Figure C.3(a)? Show that $f_e(x)$ is even and that $f_0(x)$ is odd. Now show that

$$f(x) = f_e(x) + f_0(x).$$

C.5 Inverse functions and fractional powers

Suppose we are given a function expressed in the form

$$y = f(x).$$

This implies that x is the independent variable, and y is obtained from it by evaluating a function, i.e. by using the "rule" or operation specified by that function. This mathematical statement expresses a certain relationship between the two variables, x and y, in which the roles are distinct: x is a value we pick, and y is calculated from it.

However, sometimes we can express a relationship in more than one way: as an example, if the connection between x and y is simple squaring, then provided $x > 0$, we might write either

$$y = x^2 \qquad \text{or} \qquad x = y^{1/2} = \sqrt{y}$$

to express the same relationship. In other words

$$y = x^2 \Leftrightarrow x = \sqrt{y}.$$

We have used two distinct functions to describe the relationship from two points of view: one function involves squaring and the other takes a square root. We may also notice that for $x > 0$,

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

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

i.e. that these two functions invert each other's effect.

Functions that satisfy

$$y = f(x) \Leftrightarrow x = g(y)$$

are said to be *inverse functions*. We often use the notation $f^{-1}(x)$ to denote the function that acts as an inverse function to $f(x)$.

Graphical property of inverse functions

The graph of an inverse function $y = f^{-1}(x)$ is geometrically related to the graph of the original function: it is a reflection of $y = f(x)$ about the 45° line, $y = x$. This relationship is shown in Figure C.7(a) for a pair of functions f and f^{-1}.

But why should this be true? The idea is as follows: suppose that (a,b) is any point on the graph of $y = f(x)$. This means that $b = f(a)$. That, in turn, implies that $a = f^{-1}(b)$, which then tells us that (b,a) must be a point on the graph of $f^{-1}(x)$. But the points (a,b) and (b,a) are related by reflection about the line $y = x$. This is true for any arbitrary point, and so must be true for *all* points on the graphs of the two functions.

Restricting the domain

The above argument establishes that, given the graph of a function, its inverse is obtained by reflecting the graph in an imaginary mirror placed along a line $y = x$.

However, a difficulty could arise. In particular, for the function

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

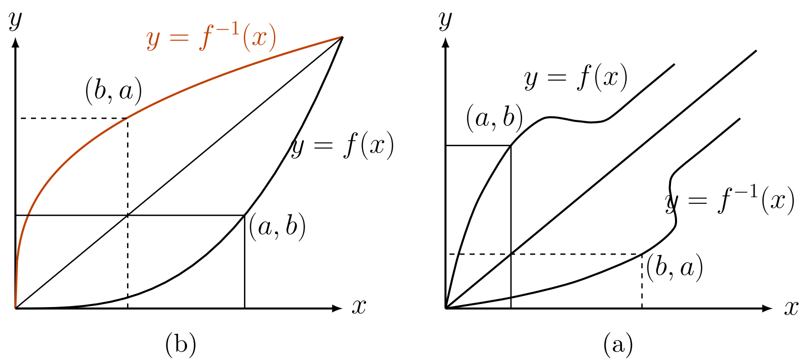


Figure C.7: The point (a,b) is on the graph of $y = f(x)$. If the roles of x and y are interchanged, this point becomes (b,a). Geometrically, this point is the reflection of (a,b) about the line $y = x$. Thus, the graph of the inverse function $y = f^{-1}(x)$ is related to the graph of the original function by reflection about the line $y = x$. In (b), the inverse is not a function, as it does not satisfy the vertical line property. In (a), both f and its reflection satisfy that property, and thus the inverse, f^{-1} is a true function.

a reflection of this type would lead to a curve that cannot be a function, as shown in Figure C.8.

Note: the sideways parabola would not be a function if we included both its branches, since a given value of x would have two associated y values.

To fix such problems, we simply restrict the domain to $x > 0$, i.e. to the solid parts of the curves shown in Figure C.8. For this subset of the x-axis, we have no problem defining the inverse function.

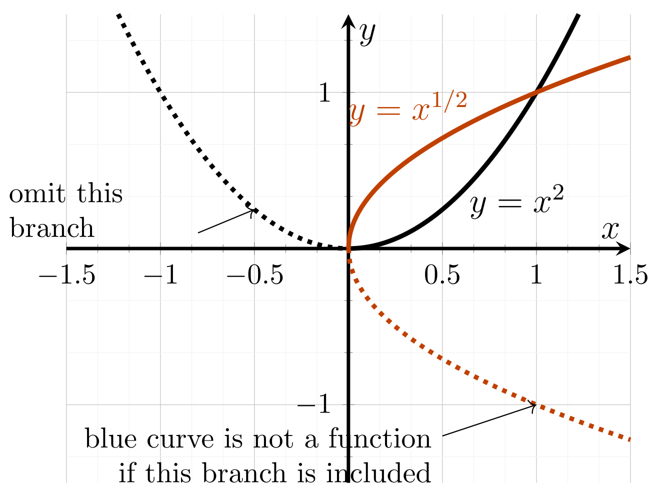


Figure C.8: The graph of $y = f(x) = x^2$ (black) and of its inverse function (red). We cannot define the inverse for all x, because the red parabola does not satisfy the vertical line property. However, if we restrict the original function to positive x values, this problem is circumvented.

Observe that the problem described above would be encountered for any of the even power functions by virtue of their symmetry about the y-axis but not by the odd power functions. For example,

$$y = f(x) = x^3, \qquad y = f^{-1}(x) = x^{1/3}$$

are inverse functions for all x values: when we reflect the graph of x^3 about the line $y = x$ we do not encounter the problem of multiple y values.

This follows directly from the horizontal line properties that we saw in Figure C.6. When we reflect the graphs shown in Figure C.6 about the line $y = x$, the horizontal lines are reflected onto vertical lines. Odd power functions have inverses that intersect a vertical line exactly once, i.e. they satisfy the "vertical line property" discussed earlier.

C.6 Polynomials

Recall that a polynomial is a function of the form

$$y = p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

This form is sometimes referred to as *superposition* (i.e. simple addition) of the basic power functions with integer powers. The constants a_k are called coefficients. In practice, some of these may be zero. We restrict attention to the case where all these coefficients are real numbers. The highest power n (whose coefficient is not zero) is called the *degree* of the polynomial.

We are interested in these functions for several reasons. Primarily, we find that computations involving polynomials are particularly easy, since operations include only the basic addition and multiplication.

Features of polynomials

- **Zeros of a polynomial:** values of x such that

  $$y = p(x) = 0.$$

  If $p(x)$ is quadratic (a polynomial of degree 2) then the quadratic formula gives a simple way of finding roots of this equation/zeros of the polynomial. Generally, for most polynomials of degree higher than 5, there is no analytical recipe for finding zeros. Geometrically, zeros are places where the graph of the function $y = p(x)$ crosses the x-axis. This fact is exploited in Chapter 5 to approximate the values of the zeros using *Newton's Method*.

- **Critical Points:** places on the graph where the value of the function is locally larger than those nearby (local maxima) or smaller than those nearby (local minima) are of interest to us. Calculus is one of the main tools for detecting and identifying such places.

- **Behaviour for very large x:** all polynomials are unbounded as $x \to \infty$ and as $x \to -\infty$. For large enough values of x, the power function $y = f(x) = x^n$ with the largest power, n, dominates over other power functions with smaller powers, as seen in Chapter 1. For

  $$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

the first (highest power) term *dominates* for large x. Thus for large x (whether positive or negative),

$$p(x) \approx a_n x^n \quad \text{for large } x.$$

- **Behaviour for small x:** close to the origin, power functions with smallest powers dominate (see Chapter 1). Thus, for $x \approx 0$ the polynomial is governed by the behaviour of the smallest (non-zero coefficient) power, i.e.,

$$p(x) \approx a_1 x + a_0 \quad \text{for small } x.$$

D

Limits

We introduced notation involving limits without carefully defining what was meant. Here, such technical matters are briefly discussed.

The concept of a **limit** helps us to describe the behaviour of a function close to some point of interest. This is useful in the case of functions that are either *not continuous*, or *not defined* somewhere. We use the notation

$$\lim_{x \to a} f(x)$$

to denote the value the function f approaches as x gets closer and closer to the value a.

D.1 Limits for continuous functions

If $x = a$ is a point at which the function is defined and continuous (informally: has no "breaks in its graph") the value of the limit and the value of the function at a point are the same, i.e.

If f is continuous at $x = a$ then

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

Example D.1 *Find* $\lim_{x \to 0} f(x)$ *for the function* $y = f(x) = 10$.

Solution. This function is continuous (and constant) everywhere. In fact, the value of the function is independent of x. We conclude immediately that

$$\lim_{x \to 0} f(x) = \lim_{x \to 0} 10 = 10.$$

◊

Example D.2 *Find* $\lim_{x \to 0} f(x)$ *for the function* $y = f(x) = \sin(x)$.

Solution. This function is a continuous trigonometric function, and has the value $\sin(0) = 0$ at the origin. Thus

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

◊

Power functions are continuous everywhere. This motivates the next example.

Example D.3 *Compute the limit* $\lim_{x \to 0} x^n$ *where n is a positive integer.*

Solution. The function in question, $f(x) = x^n$, is a simple power function that is continuous everywhere. Further, $f(0) = 0$ for any n a positive integer. Hence the limit as $x \to 0$ coincides with the value of the function at that point, so

$$\lim_{x \to 0} x^n = 0.$$

◊

D.2 Properties of limits

Suppose we are given two functions, $f(x)$ and $g(x)$. we also assume that both functions have (finite) limits at the point $x = a$. Then the following statements follow.

1. $\lim_{x \to a} (f(x) + g(x)) = \lim_{x \to a} f(x) + \lim_{x \to a} g(x)$
2. $\lim_{x \to a} (cf(x)) = c \lim_{x \to a} f(x)$
3. $\lim_{x \to a} (f(x) \cdot g(x)) = \left(\lim_{x \to a} f(x)\right) \cdot \left(\lim_{x \to a} g(x)\right)$
4. Provided that $\lim_{x \to a} g(x) \neq 0$, we also have that

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

The first two statements are equivalent to linearity of the process of computing a limit.

Example D.4 *Find* $\lim_{x \to 2} f(x)$ *for the function* $y = f(x) = 2x^2 - x^3$.

Solution. Since this function is a polynomial, and so continuous everywhere, we can simply plug in the relevant value of x, i.e.

$$\lim_{x \to 2} \left(2x^2 - x^3\right) = 2 \cdot 2^2 - 2^3 = 0.$$

Thus when x gets closer to 2, the value of the function gets closer to 0. ◊

Note: when the function is continuous, the value of the limit is the same as the value of the function at the given point.

D.3 Limits of rational functions

Case 1: Denominator nonzero

We first consider functions that are the quotient of two polynomials, $y = f(x)/g(x)$ at points were $g(x) \neq 0$. This allows us to apply Property 4 of limits together with what we have learned about the properties of power functions and polynomials. Much of this discussion is related to the properties of power functions and dominance of lower (higher) powers at small (large) values of x, as discussed in Chapter 1. In the examples below, we consider both limits at the origin (at $x = 0$) and at infinity (for $x \to \infty$). The latter means "very large x". See Section 1.4for examples of the informal version of the same reasoning used to reach the same conclusions.

Example D.5 *Find the limit as $x \to 0$ and as $x \to \infty$ of the quotients*

$$\text{(a) } \frac{Kx}{k_n + x}, \quad \text{(b)} \frac{Ax^n}{a^n + x^n}.$$

Solution. We recognize (a) as an example of the Michaelis-Menten kinetics, found in (1.8) and (b) as a Hill function in (1.7) of Chapter 1. We now compute, first for $x \to 0$,

$$\text{(a) } \lim_{x\to 0} \frac{Kx}{k_n + x} = 0, \quad \text{(b)} \lim_{x\to 0} \frac{Ax^n}{a^n + x^n} = 0.$$

This follows from the fact that, provided $a, k_n \neq 0$, both functions are continuous at $x = 0$, so that their limits are the same as the actual values attained by the functions. Now for $x \to \infty$

$$\text{(a) } \lim_{x\to\infty} \frac{Kx}{k_n + x} = \lim_{x\to\infty} \frac{Kx}{x} = K, \quad \text{(b)} \lim_{x\to\infty} \frac{Ax^n}{a^n + x^n} = \lim_{x\to\infty} \frac{Ax^n}{x^n} = A.$$

This follows from the fact that the constants k_n, a^n are always "swapped out" by the value of x as $x \to \infty$, allowing us to obtain the result. Other than the formal limit notation, there is nothing new here that we have not already discussed in Sections 1.5. ♢

Below we apply similar reasoning to other examples of rational functions.

Example D.6 *Find the limit as $x \to 0$ and as $x \to \infty$ of the quotients*

$$\text{(a) } \frac{3x^2}{9 + x^2}, \quad \text{(b)} \frac{1 + x}{1 + x^3}.$$

Solution. For part (a) we note that as $x \to \infty$, the quotient approaches $3x^2/x^2 = 3$. As $x \to 0$, both numerator and denominator are defined and the denominator is nonzero, so we can use the 4th property of limits. We thus find that

$$\text{(a) } \lim_{x\to\infty} \frac{3x^2}{9 + x^2} = 3, \quad \lim_{x\to 0} \frac{3x^2}{9 + x^2} = 0,$$

For part (b), we use the fact that as $x \to \infty$, the limit approaches $x/x^3 = x^{-2} \to 0$. As $x \to 0$ we can apply property 4 yet again to compute the (finite) limit, so that

$$\text{(b)} \lim_{x\to\infty} \frac{1+x}{1+x^3}, \quad \lim_{x\to 0} \frac{1+x}{1+x^3}.$$

◊

Example D.7 *Find the limits of the following function at* 0 *and* ∞

$$y = \frac{x^4 - 3x^2 + x - 1}{x^5 + x}.$$

Solution. For $x \to \infty$ powers with the largest power dominate, whereas for $x \to 0$, smaller powers dominate. Hence, we find

$$\lim_{x\to\infty} \frac{x^4 - 3x^2 + x - 1}{x^5 + x} = \lim_{x\to\infty} \frac{x^4}{x^5} = \lim_{x\to\infty} \frac{1}{x} = 0.$$

$$\lim_{x\to 0} \frac{x^4 - 3x^2 + x - 1}{x^5 + x} = \lim_{x\to 0} \frac{-1}{x} = -\lim_{x\to 0} \frac{1}{x} = \infty$$

So in the latter case, the limit does not exist. ◊

Case 2: zero in the denominator and "holes" in a graph

In the previous examples, evaluating the limit, where it existed, was as simple as plugging the appropriate value of x into the function itself. The next example shows that this is not always possible.

Example D.8 *Compute the limit as* $x \to 4$ *of the function* $f(x) = 1/(x-4)$

Solution. This function has a vertical asymptote at $x = 4$. Indeed, the value of the function shoots off to $+\infty$ if we approach $x = 4$ from above, and $-\infty$ if we approach the same point from below. We say that the limit **does not exist** in this case. ◊

Example D.9 *Compute the limit as* $x \to -1$ *of the function* $f(x) = x/(x^2 - 1)$

Solution. We compute

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

It is evident (even before factoring as we have done) that this function has a vertical asymptote at $x = -1$ where the denominator approaches zero. Hence, the limit does not exist. ◊

Next, we describe an extremely important example where the function has a "hole" in its graph, but where a finite limit exists. This kind of limit plays a huge role in the definition of a derivative.

Example D.10 *Find* $\lim_{x\to 2} f(x)$ *for the function* $y = (x-2)/(x^2 - 4)$.

Solution. This function is a quotient of two rational expressions $f(x)/g(x)$ but we note that $\lim_{x\to 2} g(x) = \lim_{x\to 2}(x^2-4) = 0$. Thus we cannot use property 4 directly. However, we can simplify the quotient by observing that for $x \neq 2$ the function $y = (x-2)/(x^2-4) = (x-2)/(x-2)(x+2)$ takes on the same values as the expression $1/(x+2)$. At the point $x = 2$, the function itself is not defined, since we are not allowed division by zero. However, the limit of this function does exist:

$$\lim_{x\to 2} f(x) = \lim_{x\to 2} \frac{(x-2)}{(x^2-4)}.$$

Provided $x \neq 2$ we can factor the denominator and cancel:

$$\lim_{x\to 2} \frac{(x-2)}{(x^2-4)} = \lim_{x\to 2} \frac{(x-2)}{(x-2)(x+2)} = \lim_{x\to 2} \frac{1}{(x+2)}$$

Now we can substitute $x = 2$ to obtain

$$\lim_{x\to 2} f(x) = \frac{1}{(2+2)} = \frac{1}{4}$$

$\Diamond$

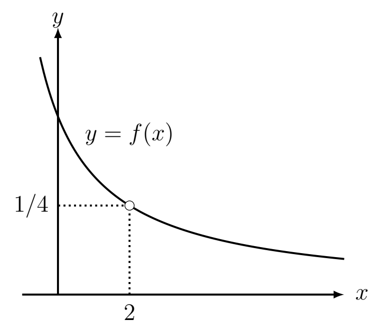


Figure D.1: The function $y = \frac{(x-2)}{(x^2-4)}$ has a "hole" in its graph at $x = 2$. The limit of the function as x approaches 2 does exist, and "supplies the missing point": $\lim_{x\to 2} f(x) = \frac{1}{4}$.

Example D.11 *Compute the limit*

$$\lim_{h\to 0} \frac{K(x+h)^2 - Kx^2}{h}.$$

Solution. This is a calculation we would perform to compute the derivative of the function $y = Kx^2$ from the definition of the derivative. Details have already been displayed in Example 2.10. The essential idea is that we expand the numerator and simplify algebraically as follows:

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

Even though the quotient is not defined at the value $h = 0$ (as the denominator is zero there), the limit exists, and hence the derivative can be defined. $\Diamond$

See also Example 3.12 for a similar calculation for the function Kx^3.

D.4 Right and left sided limits

Some functions are discontinuous at a point, but we may still be able to define a limit that the function attains as we approach that point from the right or from the left. (This is equivalent to gradually decreasing or gradually increasing x as we get closer to the point of interest.

Consider the function

$$f(x) = \begin{cases} 0 & \text{if } x < 0; \\ 1 & \text{if } x > 0. \end{cases}$$

This is a step function, whose values is 0 for negative real numbers, and 1 for positive real numbers. The function is not even defined at the point $x = 0$ and has a jump in its graph. However, we can still define a right and a left limit as follows:

$$\lim_{x \to^+ 0} f(x) = 0, \quad \lim_{x \to^- 0} f(x) = 1.$$

That is, the limit as we approach from the right is 0 whereas from the left it is 1. We also state the following result:

If $f(x)$ has a right and a left limit at a point $x = a$ and if those limits are equal, then we say that the limit at $x = a$ exists, and we write

$$\lim_{x \to^+ a} f(x) = \lim_{x \to^- a} f(x) = \lim_{x \to a} f(x)$$

Example D.12 *Find* $\lim_{x \to \pi/2} f(x)$ *for the function* $y = f(x) = \tan(x)$.

Solution. The function $\tan(x) = \sin(x)/\cos(x)$ cannot be continuous at $x = \pi/2$ because $\cos(x)$ in the denominator takes on the value of zero at the point $x = \pi/2$. Moreover, the value of this function becomes unbounded (grows without a limit) as $x \to \pi/2$. We say in this case that "the limit does not exist". We sometimes use the notation

$$\lim_{x \to \pi/2} \tan(x) = \pm\infty.$$

(We can distinguish the fact that the function approaches $+\infty$ as x approaches $\pi/2$ from below, and $-\infty$ as x approaches $\pi/2$ from higher values.)

◊

D.5 Limits at infinity

We can also describe the behaviour "at infinity" i.e. the trend displayed by a function for very large (positive or negative) values of x. We consider a few examples of this sort below.

Example D.13 *Find* $\lim_{x \to \infty} f(x)$ *for the function* $y = f(x) = x^3 - x^5 + x$.

Solution. All polynomials grow in an unbounded way as x tends to very large values. We can determine whether the function approaches positive or negative unbounded values by looking at the coefficient of the highest power of x, since that power dominates at large x values. In this example, we find that the term $-x^5$ is that highest power. Since this has a negative coefficient, the function approaches unbounded negative values as x gets larger in the positive direction, i.e.

$$\lim_{x \to \infty} x^3 - x^5 + x = \lim_{x \to \infty} -x^5 = -\infty.$$

◊

Example D.14 *Determine the following two limits:*

$$\text{(a) } \lim_{x\to\infty} e^{-2x}, \quad \text{(b) } \lim_{x\to-\infty} e^{5x},$$

Solution. The function $y = e^{-2x}$ becomes arbitrarily small as $x \to \infty$. The function $y = e^{5x}$ becomes arbitrarily small as $x \to -\infty$. Thus we have

$$\text{(a) } \lim_{x\to\infty} e^{-2x} = 0, \quad \text{(b) } \lim_{x\to-\infty} e^{5x} = 0.$$

◊

Example D.15 *Find the limits below:*

$$\text{(a) } \lim_{x\to\infty} x^2 e^{-2x}, \quad \text{(b) } \lim_{x\to 0} \frac{1}{x} e^{-x},$$

Solution. For part (a) we state here the fact that as $x \to \infty$, the exponential function with negative exponent decays to zero faster than any power function increases. For part (b) we note that for the quotient e^{-x}/x we have that as $x \to 0$ the top satisfies $e^{-x} \to e^0 = 1$, while the denominator has $x \to 0$. Thus the limit at $x \to 0$ cannot exist. We find that

$$\text{(a) } \lim_{x\to\infty} x^2 e^{-2x} = 0, \quad \text{(b) } \lim_{x\to 0} \frac{1}{x} e^{-x} = \infty,$$

◊

D.6 Summary of special limits

As a reference, in the table below, we collect some of the special limits that are useful in a variety of situations.

Function	$x \to$	Limit notation	value
e^{-ax}, $a > 0$	∞	$\lim_{x\to\infty} e^{-ax}$	0
e^{-ax}, $a > 0$	$-\infty$	$\lim_{x\to-\infty} e^{-ax}$	∞
e^{ax}, $a > 0$	∞	$\lim_{x\to\infty} e^{ax}$	∞
e^{kx}	0	$\lim_{x\to 0} e^{kx}$	1
$x^n e^{-ax}$, $a > 0$	∞	$\lim_{x\to\infty} x^n e^{-ax}$	0
$\ln(ax)$, $a > 0$	∞	$\lim_{x\to\infty} \ln(ax)$	∞
$\ln(ax)$, $a > 0$	1	$\lim_{x\to 1} \ln(ax)$	0
$\ln(ax)$, $a > 0$	0	$\lim_{x\to 0} \ln(ax)$	$-\infty$
$x\ln(ax)$, $a > 0$	0	$\lim_{x\to 0} x\ln(ax)$	0
$\frac{\ln(ax)}{x}$, $a > 0$	∞	$\lim_{x\to\infty} \frac{\ln(ax)}{x}$	0
$\frac{\sin(x)}{x}$	0	$\lim_{x\to 0} \frac{\sin(x)}{x}$	1

Table D.1: A collection of useful limits.

We can summarize the information in this table informally as follows:

1. The exponential function e^x grows faster than any power function as x increases, and conversely the function $e^{-x} = 1/e^x$ decreases faster than any power of $(1/x)$ as x grows. The same is true for e^{ax} provided $a > 0$.

2. The logarithm $\ln(x)$ is an increasing function that keeps growing without bound as x increases, but it does not grow as rapidly as the function $y = x$. The same is true for $\ln(ax)$ provided $a > 0$. The logarithm is not defined for negative values of its argument and as x approaches zero, this function becomes unbounded and negative. However, it approaches $-\infty$ more slowly than x approaches 0. For this reason, the expression $x\ln(x)$ has a limit of 0 as $x \to 0$.

E
Proofs

This Appendix was written by Dr. Sophie Burrill.

E.1 Proof of the power rule

We present a proof for the power rule. Recall from Section 4.1:

The **power rule** states that the derivative of the power function $f(x) = x^n$ is nx^{n-1}.

Proof. We begin with the definition of the derivative:

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

As mentioned in Section 4.1, the binomial $(x+h)^n$ entails lenthy algebra. We employ the **binomial theorem** (which we do not prove):

Binomial theorem. If n is a positive integer, then

$$\begin{aligned}(x+y)^n &= x^n + nx^{n-1}y + \frac{n(n-1)}{2\cdot 1}x^{n-2}y^2 + \frac{n(n-1)(n-2)}{3\cdot 2\cdot 1}x^{n-3}y^3 \\ &\quad + \ldots + \frac{n(n-1)\ldots(n-k+1)}{k(k-1)\ldots 2\cdot 1}x^{n-k}y^k + \ldots + nxy^{n-1} + y^n\end{aligned}$$

Note that this means the expansion of $(x+y)^n$ is a sum of terms of the form

$$c_k x^{n-k} y^k, \qquad k = 0, 1, \ldots, n$$

where c_k (called a "binomial coefficient") is a coefficient that depends on **both** n and k. The exact form of c_k is not necessary for the proof of the power rule - except for the terms $c_0 = 1$ and $c_1 = n$, the coefficients of x^n and $x^{n-1}y$.

Let us use the binomial theorem and expand the numerator in the definition of the derivative:

$$\begin{aligned} f(x+h)-f(x) &= (x+h)^n - x^n \\ &= (c_0x^n + c_1x^{n-1}h + c_2x^{n-2}h^2 + c_3x^{n-3}h^3 \\ &\quad + \ldots + c_{n-1}xh^{n-1} + c_nh^n) - x^n \end{aligned}$$

We can rewrite using the fact that $c_0 = 1$ and note that the terms x^n cancel:

$$(x+h)^n - x^n = (x^n + c_1x^{n-1}h + \ldots + c_nh^n) - x^n = c_1x^{n-1}h + \ldots + c_nh^n.$$

All of the remaining terms have h as a factor, which we can factor out:

$$(x+h)^n - x^n = h[c_1x^{n-1} + c_2x^{n-2}h + \cdots + c_nh^{n-1}].$$

Substituting this into the definition of the derivative we achieve:

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \lim_{h\to 0} \frac{(x+h)^n - x^n}{h} \\ &= \lim_{h\to 0} \frac{h[c_1x^{n-1} + c_2x^{n-2}h + \cdots + c_nh^{n-1}]}{h} \\ &= \lim_{h\to 0} [c_1x^{n-1} + c_2x^{n-2}h + \cdots + c_nh^{n-1}] \end{aligned}$$

Note that all terms except for the first have h as a factor and so tend to 0 as $h \to 0$. This gives that

$$f'(x) = c_1x^{n-1},$$

and we have already noted that $c_1 = n$, so

$$f'(x) = nx^{n-1},$$

which proves the power rule. □

E.2 *Proof of the product rule*

We proof the product rule. Recall from Section 4.1:

The product rule: If $f(x)$ and $g(x)$ are two functions, each differentiable in the domain of interest, then

$$\frac{d[f(x)g(x)]}{dx} = \frac{df(x)}{dx}g(x) + \frac{dg(x)}{dx}f(x).$$

Another notation for this rule is

$$[f(x)g(x)]' = f'(x)g(x) + g'(x)f(x).$$

Proof. Let $k(x) = f(x)g(x)$, the product of the two functions. We use the definition of the derivative:

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

Adding $0 = f(x)g(x+h) - f(x)g(x+h)$ allows us to perform some helpful factoring:

$$\begin{aligned} k'(x) &= \lim_{h\to 0} \frac{f(x+h)g(x+h)-f(x)g(x+h)+f(x)g(x+h)-f(x)g(x)}{h} \\ &= \lim_{h\to 0} \frac{[f(x+h)-f(x)]\cdot g(x+h)+f(x)\cdot[g(x+h)-g(x)]}{h} \\ &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} g(x+h) + f(x)\frac{g(x+h)-g(x)}{h} \end{aligned}$$

Due to properties of limits (see Appendix D.2) we can distribute the limit and recognize familiar derivatives:

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

Thus we have proved the power rule, that

$$[f(x)g(x)]' = f'(x)g(x)+f(x)g'(x).$$

□

E.3 Proof of the quotient rule

We provide a proof the quotient rule. Recall from Section 4.1:

The quotient rule: If $f(x)$ and $g(x)$ are two functions, each differentiable in the domain of interest, then

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{\frac{df(x)}{dx}g(x) - \frac{dg(x)}{dx}f(x)}{[g(x)]^2}.$$

We can also write this in the form

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

Proof. This proof also follows from the definition of the derivative; it contains some careful arithmetic. Let $k(x) = \frac{f(x)}{g(x)}$. Using the definition of the derivative we get:

$$k'(x) = \lim_{h\to 0} \frac{k(x+h)-k(x)}{h} = \lim_{h\to 0} \frac{\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}}{h} = \lim_{h\to 0} \frac{1}{h}\left[\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}\right].$$

Finding a common denominator and then adding $0 = g(x+h)f(x+h) - g(x+h)f(x+h)$ in the numerator we proceed:

$$\begin{aligned} k'(x) &= \lim_{h\to 0} \frac{1}{h}\left[\frac{f(x+h)g(x) - f(x)g(x+h)}{g(x+h)g(x)}\right] \\ &= \lim_{h\to 0} \frac{1}{h}\left[\frac{f(x+h)g(x) - f(x+h)g(x+h) + f(x+h)g(x+h) - f(x)g(x+h)}{g(x+h)g(x)}\right] \\ &= \lim_{h\to 0} \frac{1}{h}\left[\frac{-f(x+h)[g(x+h) - g(x)] + g(x+h)[f(x+h) - f(x)]}{g(x+h)g(x)}\right] \end{aligned}$$

Using properties of limits and identifying the definition of the derivative for $g'(x)$ and $f'(x)$ leads us to:

$$\begin{aligned} k'(x) &= \lim_{h\to 0}\left[\frac{-f(x+h)g'(x)}{g(x+h)g(x)} + \frac{g(x+h)f'(x)}{g(x+h)g(x)}\right] \\ &= \frac{-f(x)g'(x) + g(x)f'(x)}{[g(x)]^2} = \frac{f'(x)g(x) - g'(x)f(x)}{[g(x)]^2}. \end{aligned}$$

We have thus proved the quotient rule. Despite the arithmetic required, hopefully the fact that the definition of the derivative is all that is required provides the reader with some comfort. □

E.4 Proof of the chain rule

We present a plausibility argument for the chain rule. Recall from Section 8.1:

If $y = g(u)$ and $u = f(x)$ are both differentiable functions and $y = g(f(x))$ is the composite function, then the **chain rule** of differentiation states that

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.$$

Proof. We first note that if a function is differentiable, it is also continuous. Because of this continuity, when x changes a very little, u can change only by a little - there are no abrupt jumps. Thus, using our notation, if $\Delta x \to 0$ then $\Delta u \to 0$.

Now consider the definition of the derivative dy/du:

$$\frac{dy}{du} = \lim_{\Delta u\to 0} \frac{\Delta y}{\Delta u}.$$

This means that for any (finite) Δu,

$$\frac{\Delta y}{\Delta u} = \frac{dy}{du} + \varepsilon,$$

where $\varepsilon \to 0$ as $\Delta u \to 0$. Then

$$\Delta y = \frac{dy}{du}\Delta u + \varepsilon \Delta u.$$

Now divide both sides by some (nonzero) Δx:

$$\frac{\Delta y}{\Delta x} = \frac{dy}{du}\frac{\Delta u}{\Delta x} + \varepsilon\frac{\Delta u}{\Delta x}.$$

Taking $\Delta x \to 0$ we get $\Delta u \to 0$, (by continuity) and hence also $\varepsilon \to 0$ so that as desired,

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.$$

□

F

Trigonometry review

The definition of trigonometric functions in terms of the angle θ in a right triangle are reviewed in Figure. F.1.

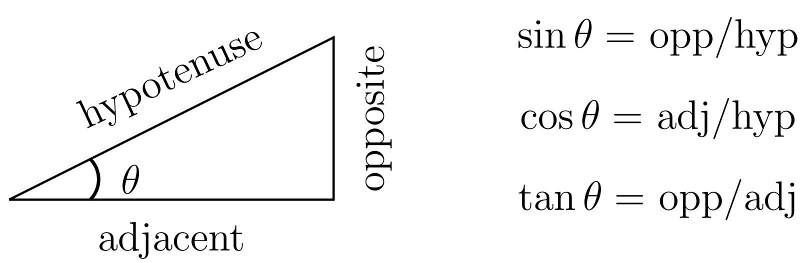


Figure F.1: Review of the relation between ratios of side lengths (in a right triangle) and trigonometric functions of the associated angle.

Based on these definitions, certain angles' sine and cosine can be found explicitly - and similarly $\tan(\theta) = \sin(\theta)/\cos(\theta)$. This is shown in Table F.1.

degrees	**radians**	$\sin(t)$	$\cos(t)$	$\tan(t)$
0	0	0	1	0
30	$\frac{\pi}{6}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{3}}$
45	$\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	1
60	$\frac{\pi}{3}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$
90	$\frac{\pi}{2}$	1	0	∞

Table F.1: Values of the sines, cosines, and tangent for standard angles.

We also define the other trigonometric functions as follows:

$$\tan(t) = \frac{\sin(t)}{\cos(t)}, \quad \cot(t) = \frac{1}{\tan(t)},$$

$$\sec(t) = \frac{1}{\cos(t)}, \quad \csc(t) = \frac{1}{\sin(t)}.$$

Sine and cosine are related by the identity

$$\cos(t) = \sin(t + \frac{\pi}{2}).$$

This identity then leads to two others of similar form. Dividing each side of the above relation by $\cos^2(t)$ yields

$$\tan^2(t) + 1 = \sec^2(t), \tag{F.1}$$

whereas division by $\sin^2(t)$ gives us

$$1 + \cot^2(t) = \csc^2(t).$$

These aid in simplifying expressions involving the trigonometric functions, as we shall see.

Law of cosines. This law relates the cosine of an angle to the lengths of sides formed in a triangle (see Figure F.2).

$$c^2 = a^2 + b^2 - 2ab\cos(\theta) \tag{F.2}$$

where the side of length c is opposite the angle θ.

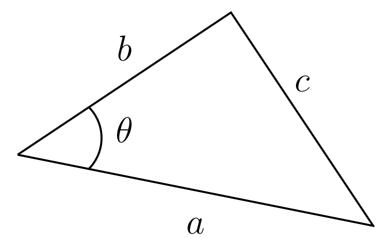


Figure F.2: Law of cosines states that $c^2 = a^2 + b^2 - 2ab\cos(\theta)$.

There are other important relations between the trigonometric functions (called *trigonometric identities*). These should be remembered.

Angle sum identities. The trigonometric functions are nonlinear. This means that, for example, the sine of the sum of two angles is *not* just the sum of the two sines. One can use the law of cosines and other geometric ideas to establish the following two relationships:

$$\sin(A+B) = \sin(A)\cos(B) + \sin(B)\cos(A) \tag{F.3}$$

$$\cos(A+B) = \cos(A)\cos(B) - \sin(A)\sin(B) \tag{F.4}$$

These two identities appear in many calculations and aid in computing derivatives of basic trigonometric formulae.

Related identities. The identities for the sum of angles can be used to derive a number of related formulae. For example, by replacing B by $-B$ we get the angle difference identities:

$$\sin(A-B) = \sin(A)\cos(B) - \sin(B)\cos(A)$$
$$\cos(A-B) = \cos(A)\cos(B) + \sin(A)\sin(B)$$

By setting $\theta = A = B$, we find the subsidiary double angle formulae:

$$\sin(2\theta) = 2\sin(\theta)\cos(\theta)$$
$$\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$$

and these can also be written in the form

$$2\cos^2(\theta) = 1 + \cos(2\theta)$$

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

F.1 Summary of the inverse trigonometric functions

In Table F.2 we show the table of standard values of functions $\arcsin(x)$ and $\arccos(x)$. In Figure F.3 we summarize the the relationships between the original trigonometric functions and their inverses.

x	$\arcsin(x)$	$\arccos(x)$
-1	$-\pi/2$	π
$-\sqrt{3}/2$	$-\pi/3$	$5\pi/6$
$-\sqrt{2}/2$	$-\pi/4$	$3\pi/4$
$-1/2$	$-\pi/6$	$2\pi/3$
0	0	$\pi/2$
$1/2$	$\pi/6$	$\pi/3$
$\sqrt{2}/2$	$\pi/4$	$\pi/4$
$\sqrt{3}/2$	$\pi/3$	$\pi/6$
1	$\pi/2$	0

Table F.2: Standard values of the inverse trigonometric functions.

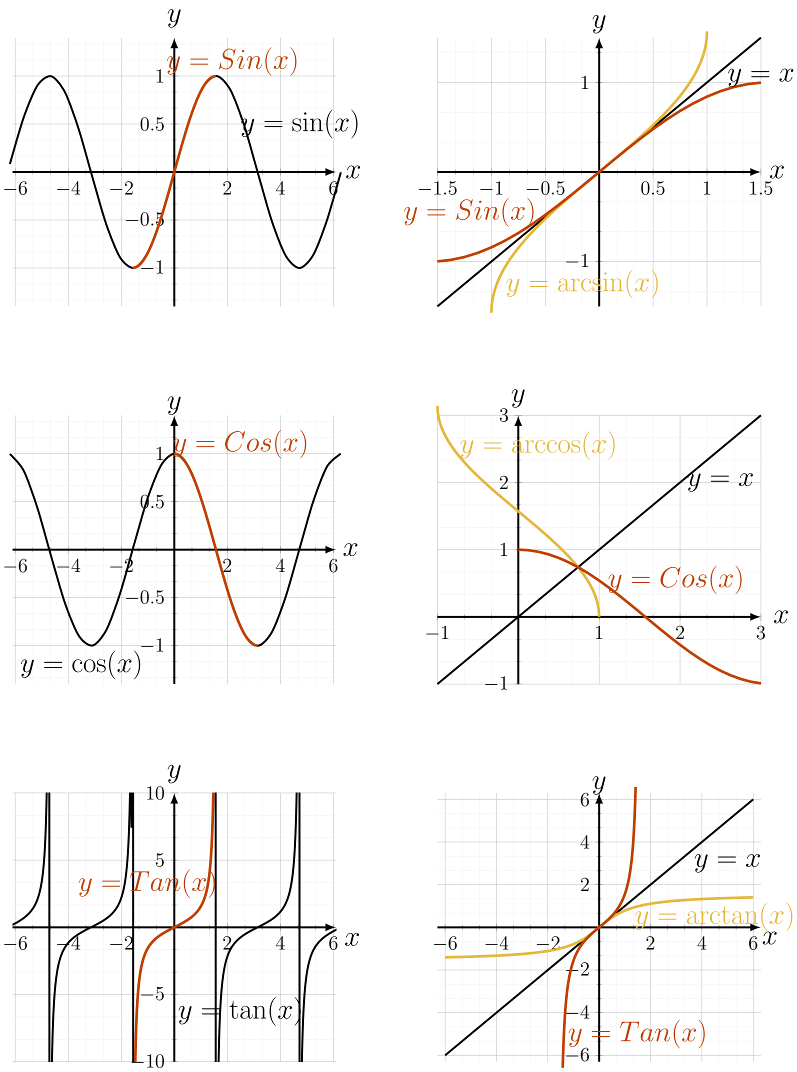


Figure F.3: A summary of the trigonometric functions and their inverses. (a) *Sin*(x) (b) $\arcsin(x)$, (b) *Cos*(x) (d) $\arccos(x)$, (e) *Tan*(x) (f) $\arctan(x)$. The red curves are the restricted domain portions of the original trig functions. The gold curves are the inverse functions.

G
For further study

In Sections G.1 and G.2 we suggest topics that are related to the material in Chapter 1. Section G.3 supplements biological examples in Chapters 3 and 4. The material in G.4 supplements examples in Chapter 7 and provides additional practice with optimization. The optimal foraging time for a *specific* patch function studied in Chapter 7 is generalized in Section G.5. Section G.6 extends the study of trigonometric functions and their derivatives as seen in Chapter 15 to differential equations.

G.1 Michaelis-Menten transformed to a linear relationship

Michaelis-Menten kinetics explored in Eqn. (1.8) is a nonlinear saturating function in which the concentration x is the independent variable on which the reaction velocity, v depends:

$$v = \frac{Kx}{k_n + x}.$$

As discussed in Section 1.5, the constants K and k_n depend on the enzyme and are often quantified in a biochemical assay of enzyme action. Historically, a convenient way to estimate the values of K and k_n was to measure v for many different values of the initial substrate concentration. Without nonlinear fitting software widely available, Eqn. (1.8) was transformed (meaning that it was rewritten) as a linear relationship.

We can do so as well with algebra. We begin by taking reciprocals of our equation and expanding:

$$\begin{aligned} \frac{1}{v} &= \frac{k_n + x}{Kx}, \\ &= \frac{k_n}{Kx} + \frac{x}{Kx} \\ &= \left(\frac{k_n}{K}\right)\frac{1}{x} + \left(\frac{1}{K}\right) \end{aligned}$$

This suggests defining the two constants:

$$m = \frac{k_n}{K}, \quad b = \frac{1}{K}.$$

In which case, the relationship between $1/v$ and $1/x$ becomes linear:

$$\left[\frac{1}{v}\right] = m\left[\frac{1}{x}\right] + b. \quad \text{(G.1)}$$

Both the slope, m and intercept b of the straight line provide information about the parameters. The relationship in Eqn. (G.1), which is a disguised variant of Michaelian kinetics, is called the Lineweaver-Burk relationship. In Exercise 30 this is used to estimate the values of K and k_n from biochemical data about an enzyme.

G.2 Spacing of fish in a school

Many animals live or function best when they are in a group. Social groups include herds of wildebeest, flocks of birds, and schools of fish, as well as swarms of insects. Life in a group can affect the way that individuals forage (search for food), their success at detecting or avoiding being eaten by a predator, and other functions such as mating, protection of the young, etc. Biologists are interested in the ecological implications of groups on their own members or on other species with whom they interact, and how individual behaviour, combined with environmental factors and random effects affect the shape, spacing and function of the groups.

In many social groups, the spacing between individuals is relatively constant from one part of the formation to another, because animals that get too close start to move away from one another, whereas those that get too far apart are attracted back. These spacing distances can be observed in a variety of groups, and were described in many biological publications. For example, Emlen [Emlen Jr, 1952] studied flocks of birds and found that gulls are spaced at about one body length apart. Similarly, Conder [J., 1949] observed a $2-3$ body lengths spacing distance in tufted ducks while Miller [R.S. and J.D., 1966] found that in the flock he observed, sandhill cranes try to keep about 5.8 ft apart.

To explain why certain spacing is maintained in a group of animals, it was proposed that there are mutual attraction and repulsion interactions (effectively acting like simple forces) between individuals. Breder [Breder, 1954] followed a number of species of fish that school, and measured the individual spacing in units of the fish body length, showing that individuals are separated by $0.16-0.25$ body length units. He suggested that the effective forces between individuals were similar to *inverse power laws* for repulsion and attraction. Breder considered a quantity he called *cohesiveness*, defined as:

$$c = \frac{A}{x^m} - \frac{R}{x^n}, \quad \text{(G.2)}$$

where A, R are magnitudes of attraction and repulsion, x is the distance between individuals, and m, n are integer powers that govern how quickly the interactions fall off with distance. We could re-express the formula in

Eqn. (G.2) as

$$c = Ax^{-m} - Rx^{-n}$$

Thus, the function shown in Breder's cohesiveness formula is related to our power functions, but the powers are negative integers. A specific case considered by Breder was $m = 0, n = 2$, i.e. constant attraction and inverse square law repulsion,

$$c = A - (R/x^2)$$

Breder specifically considered the "point of neutrality", where $c = 0$. The distance at which this occurs is:

$$x = (R/A)^{1/2}$$

where attraction and repulsion are balanced. This is the distance at which two fish would be most comfortable: neither tending to move apart, nor get closer together.

G.3 A biological speed machine

Lysteria monocytogenes is a parasite that lives inside cells of the host, causing a nasty infection. It has been studied by cellular biologists for its amazingly fast propulsion, which uses the host's actin filaments as "rocket fuel". Actin is part of the structural component of all animal cells, and is known to play a major role in cell motility. Lysteria manages to "hijack" this cellular mechanism, assembling it into its own comet tail, which it uses to propel inside the cell and pass from one cell to the next. Figure G.1 illustrates part of these curious traits.

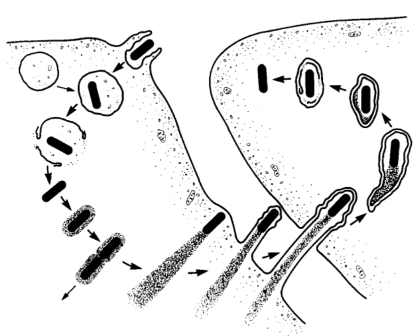


Figure G.1: The parasite *Lysteria* lives inside a host cell. It assembles a "rocket-like" tail made up of actin, and uses this assembly to move around the cell, and to pass from one host cell to another. Figure from [Tilney and Portnoy, 1989].

Researchers in cell biology use Lysteria to learn about motility at the cellular level. It has been discovered that certain proteins on the external surface of this parasite (ActA) are responsible for the ability of Lysteria to assemble an actin filament tail. Surprisingly, even small plastic beads artificially coated in Lysteria's ActA proteins can perform the same "trick": they assemble an actin tail which pushes the bead like a tiny rocket.

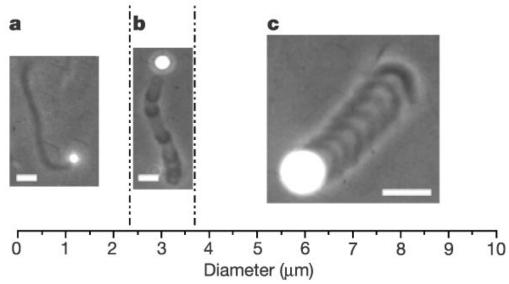


Figure G.2: Small spherical beads coated with part of Lysteria's special actin-assembly kit also gain the ability to swim around. Figure from [Bernheim-Groswasser et al., 2002].

In a recent papr, Bernheim-Groswasser et al. [Bernheim-Groswasser et al., 2002] describe the motion of these beads, shown in Figure G.2. When the position of the bead is plotted on a graph with time as the horizontal axis, (see Figure G.3) we find that the trajectory is not a simple one: it appears that the bead slows down periodically, and then accelerates.

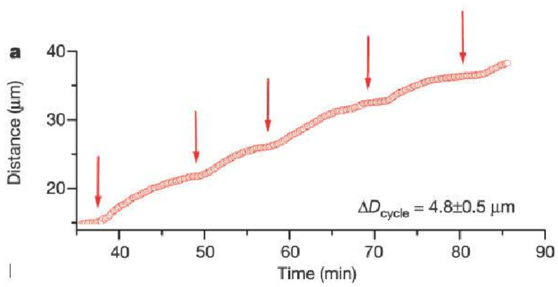


Figure G.3: The distance traveled by a little bead is shown as a function of time. The arrows point to times when the particle slowed down or stopped. We can use this data to analyze the velocity of the particles. Figure from [Bernheim-Groswasser et al., 2002].

With the techniques developed in Chapters 3 and 4, we can analyze the experimental data shown in Figure G.3 to determine both the average velocity of the beads, and the instantaneous velocity over the course of the motion.

Average velocity of the bead. We can get a rough idea of how fast the micro-beads are moving by computing an average velocity over the time interval shown on the graph. We can use two (approximate) data points $(t, D(t))$, at the beginning and end of the run, for example $(45, 20)$ and $(80, 35)$: the average velocity is

$$\bar{v} = \frac{\Delta D}{\Delta t}$$

$$\bar{v} = \frac{35-20}{80-45} \approx 0.43\mu \text{ min}^{-1}$$

so the beads move with average velocity 0.43 microns per minute.

Units.
One micron is 10^{-6} meters.

The changing instantaneous velocity. Because the actual data points are taken at finite time increments, the curve shown in Figure G.3 is not smooth. We smoothen it, as shown in Figure G.4 for a simpler treatment. In Figure G.5 we sketch this curve together with a collection of lines that represent the slopes of tangents along the curve. A horizontal tangent has slope zero:

this means that at all such points (also indicated by the arrows for emphasis), the velocity of the beads is zero. Between these spots, the bead has picked up speed and moved forward until the next time in which it stops.

We show the velocity $v(t)$, which is the derivative of the original function $D(t)$ in Figure G.6. As shown here, the velocity has periodic increases and decreases.

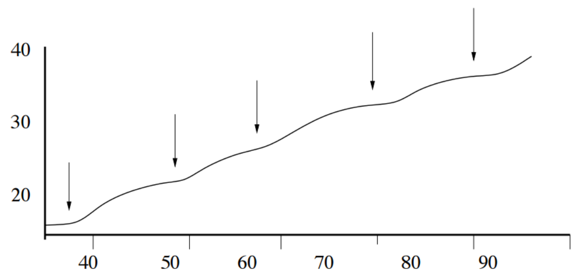


Figure G.4: The (slightly smoothened) bead trajectory is shown here.

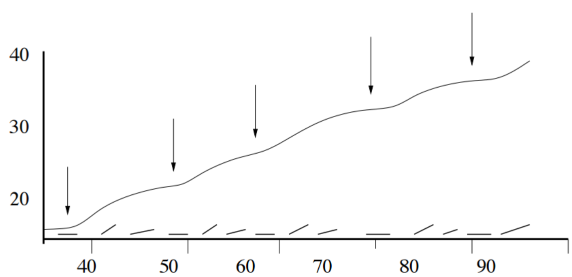


Figure G.5: We inserted a sketch of the tangent line configurations along the trajectory from beginning to end. We observe that some of these tangent lines are horizontal, implying a zero derivative, and, thus, a zero instantaneous velocity at that time.

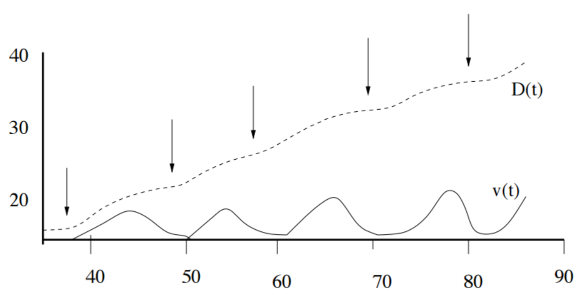


Figure G.6: Here we sketched the velocity on the same graph.

G.4 *Additional examples of geometric optimization*

Rectangular box with largest surface area

We consider two examples of optimization where volumes, lengths, and/or surface areas are considered.

Example G.1 (Wrapping a rectangular box) *A box with square base and arbitrary height has string tied around each of its perimeter. The total length of string so used is* 10 *inches. Find the dimensions of the box with largest surface area, i.e. determine the largest amount of wrapping paper needed to wrap this box.*

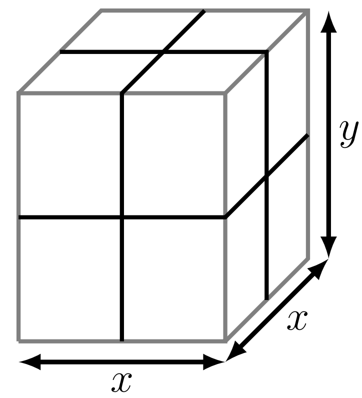


Figure G.7: A rectangular box is to be wrapped with paper.

Solution. The total length of string is shown in Figure G.7. It consists of three perimeters of the box is as follows:

$$L = 2(x+x)+2(x+y)+2(x+y) = 8x+4y = 10.$$

This total length is to be kept constant, so this equation is the constraint in this problem. This means that x and y are related to one another. We use this fact to eliminate one of them from the formula for surface area.

The surface area of the box is

$$S = 4(xy)+2x^2$$

since there are two faces (top and bottom) which are squares (area x^2) and four rectangular faces with area xy. At the moment, the total surface area S is expressed in terms of both variables. Suppose we eliminate y from S by rewriting the constraint in the form:

$$y = \frac{5}{2} - 2x.$$

Then

$$S(x) = 4x\left(\frac{5}{2} - 2x\right) + 2x^2 = 10x - 8x^2 + 2x^2 = 10x - 6x^2.$$

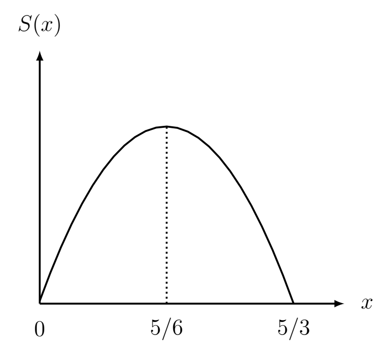


Figure G.8: Figure for Example G.1; surface area of a box.

We show the shape of this function in Figure G.8. Note that $S(x) = 0$ at $x = 0$ and at $10 - 6x = 0$ (which occurs at $x = 5/3$).

Since S is now expressed as a function of one variable, we can find its critical points by setting $S'(x) = 0$, i.e., solving

$$S'(x) = 10 - 12x = 0$$

for x: we get $x = 10/12 = 5/6$. To find the corresponding value of y we substitute our result back into the constraint. This results in

$$y = \frac{5}{2} - 2\left(\frac{5}{6}\right) = \frac{15-10}{6} = \frac{5}{6}.$$

Thus the dimensions of the box of interest are all the same, i.e. it is a cube with side length $5/6$.

We can verify that

$$S''(x) = -12 < 0,$$

(indeed this holds for all x), which means that $x = 5/6$ is a local maximum.

Further, we can find that

$$S = 4\left(\frac{5}{6}\right)\left(\frac{5}{6}\right) + 2\left(\frac{5}{6}\right)^2 = \frac{25}{6}$$

square inches, the maximum surface area of a box with such a constraint. Figure G.8 shows how the surface area varies as the dimension x of the box is varied.

A cylinder in a sphere

Example G.2 (Fitting a cylinder inside a sphere) *Find the cylinder of maximal volume that would fit inside a sphere of radius R.*

Solution. We sketch a cylinder inside a sphere as in Figure G.9.

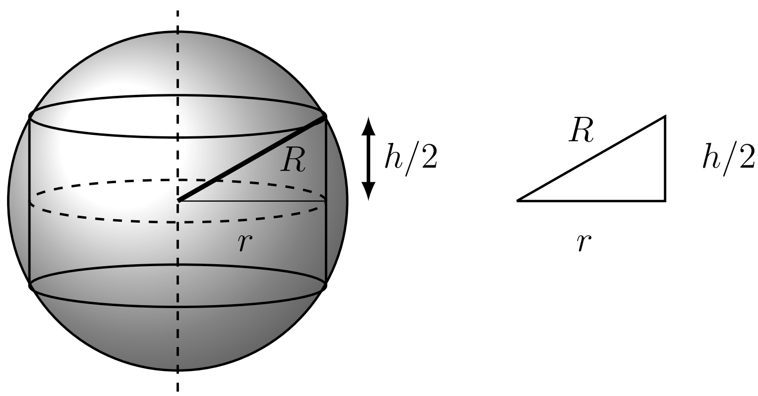


Figure G.9: Definition of variables and geometry to consider.

It is helpful to add the radius of the sphere and of the cylinder. We define the following:

$$h = \text{height of cylinder,}$$
$$r = \text{radius of cylinder,}$$
$$R = \text{radius of sphere.}$$

Then R is assumed a given fixed positive constant, and r and h are dimensions of the cylinder to be determined.

From Figure G.9 we see that the cylinder fits if the top and bottom rims touch the circle. When this occurs, the dark line in Figure G.9 is a radius of the sphere, and so has length R.

The connection between the variables (our constraint) is given from Pythagoras' theorem by:

$$R^2 = r^2 + \left(\frac{h}{2}\right)^2.$$

We maximize the volume of the cylinder,

$$V = \pi r^2 h$$

subject to the above constraint.

Eliminating r^2 using the Pythagoras theorem leads to

$$V(h) = \pi \left(R^2 - \frac{h^2}{4} \right) h.$$

We see that the problem is very similar to the previous discussion. The reader can show by working out the steps that

$$V'(h) = 0$$

occurs at the critical point

$$h = \frac{2}{\sqrt{3}} R$$

and this is a local maximum.

G.5 Optimal foraging with other patch functions

In Section 7.4, we computed an optimal foraging time for a specific patch function $f(t)$ given by Eqn. (7.6). However, we can gain insight and obtain an interesting result without making this assumption. We now consider a similar analysis with a more general example.

Example G.3 *Carry out the calculations for the optimal value patch residence time for a general patch energy function $f(t)$, without using the formula Eqn.* (7.6).

Solution. We use the expression for $R(t)$ given by Eqn. (7.7). Differentiating, we find the first derivative,

$$R'(t) = \frac{f'(t)(\tau + t) - f(t)}{(\tau + t)^2} = \frac{G(t)}{H(t)}$$

where

$$G(t) = f'(t)(\tau + t) - f(t), \quad H(t) = (\tau + t)^2.$$

(The calculation is easier with this notation.) To maximize $R(t)$ we set

$$R'(t) = 0$$

which can occur only when the numerator of the above equation is zero, i.e.

$$G(t) = 0.$$

This means that

$$f'(t)(\tau + t) - f(t) = 0$$

so that, after simplifying algebraically,

$$f'(t) = \frac{f(t)}{\tau + t}. \tag{G.3}$$

A geometric argument. In practice, we need to specify a function for $f(t)$ in order to solve for the optimal time t. However, we can also solve this problem using a **geometric argument**.

Eqn (G.3) equates two quantities that can be interpreted as slopes. On the right is the slope of a tangent line. On the left is the slope (rise over run) of some right triangle whose height is $f(t)$ and whose base length is $\tau + t$. In Figure G.10, we show each slope on its own: in the left panel, $f'(t)$ is the slope of the tangent line to the graph of $f(t)$. In the central panel, we have constructed some triangle with the property that its hypotenuse has slope $f(t)/[\tau + t]$. On the right panel we have superimposed both, selecting a value of t for which the slope of the triangle is the same as the slope of the tangent line.

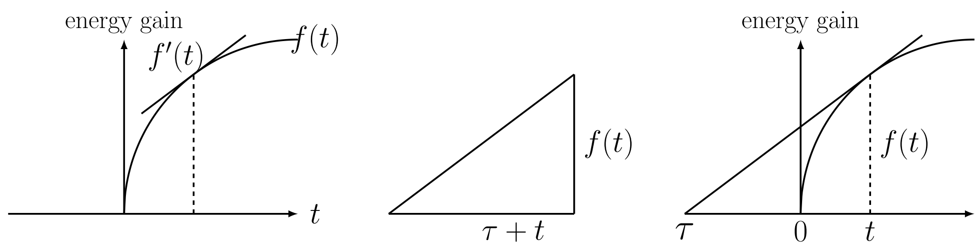


Figure G.10: The solution to the optimal foraging problem can be expressed geometrically in the form shown in this figure. The tangent line at the (optimal) time t should have the same slope as the hypotenuse of the right triangle shown above. The diagram on the far right is sometimes termed the "rooted tangent" diagram.

Notice that in order to fit the triangle on the same diagram, we had to place its tip at the point $-\tau$ along the horizontal axis. When these slopes coincide, it means that we have satisfied Eqn. (G.3), and found the desired time t for optimal foraging.

We can use this observation to come up with the following steps to solve an optimal foraging problem in general:

1. A biologist conducts some field experiments to determine the mean travel time from food to nest, τ, and the shape of the energy gain function $f(t)$.

   Note: this may require capturing the animal and examining the contents of its stomach. We leave this task to our biological colleagues.

2. We draw a sketch of $f(t)$ as shown in rightmost panel of Figure G.10 and extend the t axis in the negative direction. At the point $-\tau$ we draw a line that just touches the curve $f(t)$ at some point (i.e. a tangent line). The slope of this line is $f'(t)$ for some value of t.

3. The value of t at the point of tangency is the optimal time to spend in the patch!

The diagram drawn in our geometric solution (right panel in Figure G.10) is often called a "rooted tangent".

We have shown that the point labeled t indeed satisfies the condition that we derived above for $R'(t) = 0$, and hence is a critical point.

Checking the type of critical point. We still need to show that this solution leads to a maximum efficiency, (rather than, say a minimum or some other critical point). We do this by examining $R''(t)$.

Recall that

$$R'(t) = \frac{G(t)}{H(t)}$$

in terms of the notation used above. Then

$$R''(t) = \frac{G'(t)H(t) - G(t)H'(t)}{H^2(t)}.$$

But, according to our remark above, at the patch time of interest (the candidate for optimal time),

$$G(t) = 0$$

so that

$$R''(t) = \frac{G'(t)H(t)}{H^2(t)} = \frac{G'(t)}{H(t)}.$$

We substitute the derivative of $G'(t), H(t)$ into this ratio:

$$G(t) = f'(t)(\tau + t) - f(t) \quad \Rightarrow \quad G'(t) = f''(t)(\tau+t) + f'(t) - f'(t) = f''(t)(\tau+t)$$

We find that

$$R''(t) = \frac{f''(t)(\tau+t)}{(\tau+t)^2} = \frac{f''(t)}{(\tau+t)}.$$

The denominator of this expression is always positive, so the sign of $R''(t)$ is the same as the sign of $f''(t)$. But in order to have a maximum efficiency at some residence time, we need $R''(t) < 0$. This tells us that the gain function has to have the property that $f''(t) < 0$, i.e. has to be concave down at the optimal residence time.

Returning to some of the shapes of the function $f(t)$ that we saw in Figure 7.7, we see that only some of these lead to an optimal solution. In cases (1), (2), (4) the function $f(t)$ has *no* points of downwards concavity on its graph. This means that in such cases there is no local maximum. The optimal efficiency would then be attained by spending as much time as possible in just one patch, or as little time as possible in any patch, i.e. it would be attained at the endpoints.

G.6 Trigonometric functions and differential equations

As we saw in Chapter 15, the functions $\sin(t)$ and $\cos(t)$ are related to one another via differentiation: one is the derivative of the other (with a multiple of the factor (-1)):

$$\frac{d\sin(t)}{dt} = \cos(t), \qquad \frac{d\cos(t)}{dt} = -\sin(t).$$

The connection becomes even clearer when we examine the second derivatives of these functions:

$$\frac{d^2\sin(t)}{dt^2} = \frac{d\cos(t)}{dt} = -\sin(t), \quad \frac{d^2\cos(t)}{dt^2} = -\frac{d\sin(t)}{dt} = -\cos(t).$$

Thus, for each of the functions $y = \sin(t), y = \cos(t)$, we find that the function and its second derivative are related to one another by the **differential equation (DE)** $d^2y/dt^2 = -y$. Here the highest derivative is a second derivative, and we denote this a **second order** DE.

More generally, we make the following observations:

The functions

$$x(t) = \cos(\omega t), \quad y(t) = \sin(\omega t)$$

satisfy a pair of differential equations,

$$\frac{dx}{dt} = -\omega y, \quad \frac{dy}{dt} = \omega x.$$

The functions

$$x(t) = \cos(\omega t), \quad y(t) = \sin(\omega t)$$

also satisfy a related differential equation with a second derivative

$$\frac{d^2x}{dt^2} = -\omega^2 x.$$

These follow by the same reasoning, where the chain rule is applied in differentiation.

Students of physics may recognize the equation that governs the behaviour of a **harmonic oscillator**, and see the connection between the circular motion of our point on the circle, and the differential equation for periodic motion.

Bibliography

Anne Bernheim-Groswasser, Sebastian Wiesner, Roy M Golsteyn, Marie-France Carlier, and Cécile Sykes. The dynamics of actin-based motility depend on surface parameters. *Nature*, 417(6886):308–311, 2002.

Benjamin P Bouchet, Ivar Noordstra, Miranda van Amersfoort, Eugene A Katrukha, York-Christoph Ammon, Natalie D Ter Hoeve, Louis Hodgson, Marileen Dogterom, Patrick WB Derksen, and Anna Akhmanova. Mesenchymal cell invasion requires cooperative regulation of persistent microtubule growth by slain2 and clasp1. *Developmental cell*, 39(6):708–723, 2016.

C.M. Breder. Structure of a fish school. *Bull. Amer. Mus. Nat. Hist.*, 98:1–27, 1951.

C.M. Breder. Equations descriptive of fish schools and other animal aggregations. *Ecology*, pages 361–370, 1954.

Eric L Charnov. Optimal foraging, the marginal value theorem. *Theoretical population biology*, 9(2):129–136, 1976.

Michael Crichton. *The Andromeda Strain*. Knopf, Dell, NY, 1969.

Lawrence M Dill. The escape response of the zebra danio (Brachydanio rerio) I. The stimulus for escape. *Animal Behaviour*, 22(3):711–722, 1974a.

Lawrence M Dill. The escape response of the zebra danio (Brachydanio rerio) II. The effect of experience. *Animal Behaviour*, 22(3):723–730, 1974b.

Reuven Dukas and Stephen Ellner. Information processing and prey detection. *Ecology*, 74(5):1337–1346, 1993.

Reuven Dukas and Alan C Kamil. The cost of limited attention in blue jays. *Behavioral Ecology*, 11(5):502–506, 2000.

Reuven Dukas and Alan C Kamil. Limited attention: the constraint underlying search image. *Behavioral Ecology*, 12(2):192–199, 2001.

John T Emlen Jr. Flocking behavior in birds. *The Auk*, 1952.

Leonhard Euler. *Introductio in analysin infinitorum, auctore Leonhardo Eulero...* apud Marcum-Michaelem Bousquet, 1748.

MP Hassell, JH Lawton, and JR Beddington. Sigmoid functional responses by invertebrate predators and parasitoids. *The Journal of Animal Ecology*, 46(1):249–262, 1977.

Brian Hopkins. The species-area relations of plant communities. *The Journal of Ecology*, pages 409–426, 1955.

Conder P. J. Individual distance. *Ibis*, 91:649–655, 1949.

Corinne Le Quéré, Robbie M Andrew, Josep G Canadell, Stephen Sitch, Jan Ivar Korsbakken, Glen P Peters, Andrew C Manning, Thomas A Boden, Pieter P Tans, Richard A Houghton, et al. Global carbon budget 2016. *Earth System Science Data*, 8(2):605, 2016.

ME Lutcavage, RW Brill, GB Skomal, BC Chase, JL Goldstein, and J Tutein. Tracking adult north atlantic bluefin tuna (thunnus thynnus) in the northwestern atlantic using ultrasonic telemetry. *Marine Biology*, 137(2):347–358, 2000.

Rajat Rohatgi, Peter Nollau, Hsin-Yi Henry Ho, Marc W Kirschner, and Bruce J Mayer. Nck and phosphatidylinositol 4, 5-bisphosphate synergistically activate actin polymerization through the N-WASP-Arp2/3 pathway. *Journal of Biological Chemistry*, 276(28):26448–26452, 2001.

Miller R.S. and Stephen W. J.D. Spatial relationships in flocks of sandhill cranes. *Ecology*, 47(2):323–327, 1966.

D. W. Stephens and J. R. Krebs. *Foraging theory*. Princeton University Press, Princeton NJ, 1986.

Lewis G Tilney and Daniel A Portnoy. Actin filaments and the growth, movement, and spread of the intracellular bacterial parasite, listeria monocytogenes. *The Journal of cell biology*, 109(4):1597–1608, 1989.

Index

Index of Applications