A first course in calculus

Fok-Shuen Leung

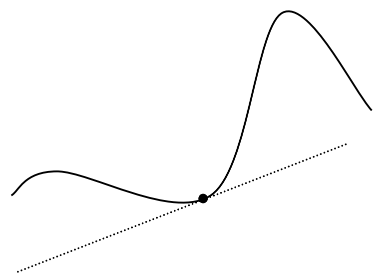
This text is based on the two-term Mathematics sequence in the University of British Columbia's Vantage One Science program. The sequence was first taught in this form in 2015-2016, where Chapters 1-10 were covered in the first term, and Chapters 11-18 were covered in the second term. It has greatly benefited from the input of the other instructors in the course: Patricia Angkiriwang, Linda Au, Stefan Dawydiak, Robert Fraser, Rachel Han, Tingting Huan, Hudson Lynn, Megan Nantel, Vanessa Radzimski, Pam Sargent, Joshua Scurll, Saifuddin Syed, Emily Tyhurst, and Dee Wang.

Contents

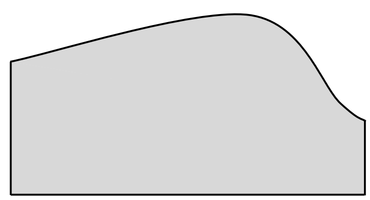
Limits and infinite limits

A. Motivation

Calculus is motivated by two problems, both leading to the same idea. The first problem, which motivates this course, is the *tangent line problem*: given a curve, how do we describe the tangent line at a particular point on the curve?

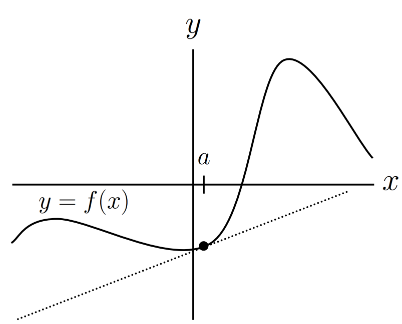


The second problem, which motivates the next course, in integral calculus, is the *area problem*: given a shape, how do we find its area?

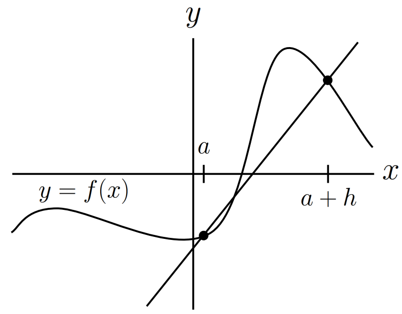


The tangent line problem was said by Descartes in *La Géométrie* to be "the most useful and general problem in geometry".

Here, we generally restrict curves to functions of one variable. The question is then: how do we find the slope of the tangent line? (The question has a physical interpretation: given a graph of distance *vs* time, how do we find instantaneous velocity?)



We can find the slope of the secant line between $(a, f(a))$ and $(a+h, f(a+h))$: it is $\frac{f(a+h)-f(a)}{h}$.



In order to find the slope of the tangent line at a, we consider the *limit* of the slopes of the secant lines: as h gets smaller, do the slopes get closer to some number? This idea of the limit is the unifying concept of calculus. In fact, this course and the next — indeed, all of calculus and analysis — may be rightly considered to be a disquisition on that concept.

We say that

$$\lim_{h \to 0} \frac{f(a+h)-f(a)}{h} = L,$$

or

$$\textit{the limit of } \frac{f(a+h)-f(a)}{h} \textit{ as } h \textit{ approaches } 0 \textit{ is } L,$$

if $\frac{f(a+h)-f(a)}{h}$ is arbitrarily close to L provided h is sufficiently close to 0. Note the peculiarity of the definition: the limit is simultaneously an *object* (L) and a *process* (the "convergence" of the slopes of secant lines).

B. The definition of limit

1. *Definition.* We say that

$$\lim_{x \to a} f(x) = L, \tag{1}$$

or

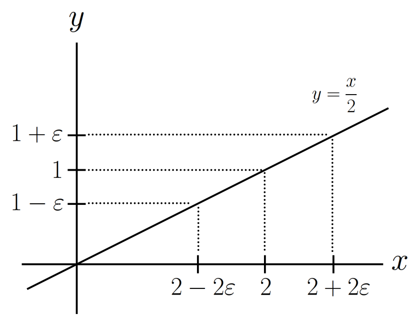
$$\textit{the limit of } f(x) \textit{ as } x \textit{ approaches } a \textit{ is } L,$$

if $f(x)$ is arbitrarily close to L provided x is sufficiently close to a. In that case, we also say the limit exists. If no such number L exists such that the statement (1) is true, then we say the limit does not exist.

Note that, in the definition above, we take the phrase "x is sufficiently close to a" to exclude the case $x = a$; that is, the value of $f(a)$ — or indeed whether $f(a)$ is even defined — is not relevant to the limit statement.

2. *Example.* $\displaystyle\lim_{x \to 2} \frac{x}{2} = 1.$

Is $\frac{x}{2}$ arbitrarily close to 1 provided x is sufficiently close to 2? That is, let $\varepsilon > 0$ be given; can we guarantee that $\frac{x}{2}$ is within ε of 1 provided x is sufficiently close to 2?



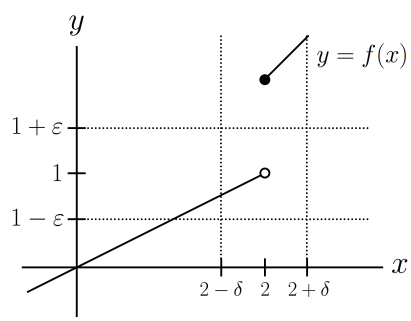
We can. Indeed, $\frac{x}{2}$ is guaranteed to be within ε of 1 provided x is within 2ε of 1. (Note that $\frac{x}{2}$ is also guaranteed to be within ε of 1 provided x is within $\frac{3}{2}\varepsilon$ of 1; 2ε is simply the largest "sufficient closeness".)

Note that ε here is an *arbitrary* closeness: it can be any positive number.

3. *Example.* Let $f(x) = \begin{cases} \frac{x}{2} & \text{if } x < 2 \\ x & \text{if } x \geq 2 \end{cases}$. Then $\lim_{x \to 2} f(x) \neq 1$.

It is crucial to appeal to the definition of limit. Is $f(x)$ arbitrarily close to 1 provided x is sufficiently close to 2? That is, let $\varepsilon > 0$ be given; can we guarantee that $f(x)$ is within ε of 1 provided x is sufficiently close to 2?

For small ε, we cannot. For example, let $\varepsilon = \frac{1}{2}$.



No matter how small an interval $(2-\delta, 2+\delta)$ to which we restrict x, there exist x-values in that interval — in particular, in the right half of the interval — such that $f(x)$ is not within ε of 1.

In contrast to the previous example, note that it is enough to show here that the guarantee cannot be made for a *particular* ε: this means it certainly cannot be made for arbitrary ε. However, $\varepsilon = \frac{1}{2}$ is not the only choice; in fact, any ε smaller than 1 will do.

4. *Infinite limits.* There are innumerable ways for a limit not to exist. However, one is common enough that it is described more concisely. We say that

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

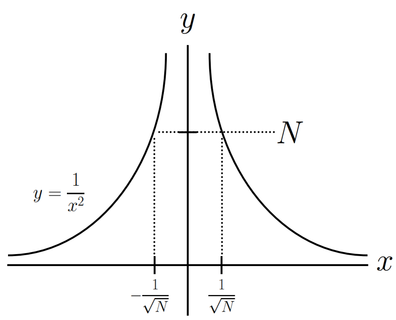
or

$$\textit{the limit of } f(x) \textit{ as } x \textit{ approaches } a \textit{ is equal to } \infty,$$

if $f(x)$ is arbitrarily large and positive provided x is sufficiently close to a. In that case, we say that the graph of $f(x)$ has a *vertical asymptote* $x = a$. (In fact, we say this if the limit of $f(x)$ as x approaches a from either side is equal to infinity; we define what this means in Exercise 1, below.)

5. *Example.* $\lim_{x\to 0} \dfrac{1}{x^2} = \infty$.

Is $\frac{1}{x^2}$ arbitrarily large and positive provided x is sufficiently close to 0? That is, let $N > 0$ be given; can we guarantee that $\frac{1}{x^2} > N$ provided x is sufficiently close to 2?



We can. $\frac{1}{x^2} > N$ is guaranteed provided x is within $\frac{1}{\sqrt{N}}$ of 0.

C. Exercises

1. Using the definition of limit, write down definitions and sketch illustrations of the following statements.
   (a) $\lim_{x\to a^+} f(x) = L$.
   (b) $\lim_{x\to a^-} f(x) = -\infty$.
2. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *If $f(a)$ is not defined, $\lim_{x\to a} f(x)$ does not exist.*
   (b) *If $f(1) = 2$, then $\lim_{x\to 1} f(x) = 2$.*
   (c) *The graph of a function may have more than one vertical asymptote.*
3. Prove that $\lim_{x\to 0} 2 = 2$.
4. Prove that $\lim_{x\to -3} x = -3$.
5. Let $f(x) = \begin{cases} x & \text{if } x \leq 1 \\ 2x - 1 & \text{if } x > 1 \end{cases}$. Prove that $\lim_{x\to 1} f(x) = 1$.
6. Let $f(x) = \begin{cases} 1 & \text{if } x = 1, \frac{1}{10}, \frac{1}{100}, \ldots \\ 0 & \text{otherwise} \end{cases}$. Prove that $\lim_{x\to 0} f(x) \neq 0$.
7. Let $f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$. Prove that $\lim_{x\to 0} f(x) \neq 0$.
8. Prove that $f(x) = \frac{1}{x-2}$ has a vertical asymptote.

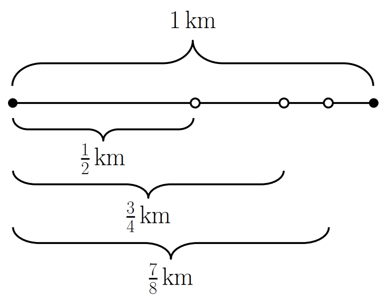
Limits at infinity and sequences

A. Motivation

Recall the tangent line problem, described by Descartes as "the most useful and general problem in geometry". The solution to this problem is to define the slope of a tangent line in terms of the limit of the slopes of secant lines.

While the precise definition of limit came after Descartes (it was formalized in the 19^{th} century by Cauchy and Weierstrass), the concept of limit preceded Descartes by many centuries.

In the 5^{th} century B.C., Zeno of Elea considered a person walking a certain distance — say 1 km. He must first walk half the distance, or $\frac{1}{2}$ km. He must then walk half the remaining distance, or a total of $\frac{3}{4}$ km; then half the remaining distance, or a total of $\frac{7}{8}$ km, and so on. In order to walk the full distance, an infinite number of increasingly shorter walks must be taken; this seemingly impossible description of an evidently possible task is called the Dichotomy Paradox.



The terms $\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \ldots$ form a *sequence* $\left\{\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \ldots\right\}$, which we also denote $\left\{1 - \frac{1}{2^n}\right\}$. We say that

$$\lim_{n\to\infty}\left(1 - \frac{1}{2^n}\right) = 1,$$

or

$$\left\{1 - \tfrac{1}{2^n}\right\} \textit{ converges to } 1,$$

if $1 - \frac{1}{2^n}$ is arbitrarily close to 1 provided n is sufficiently large. Note that this is easily interpreted as a definition in terms of functions, and it is there that we begin.

B. Limits at infinity

1. *Definition.* We say that

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

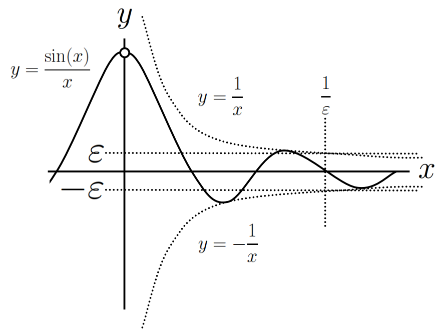
   or

   $$\textit{the limit of } f(x) \textit{ as } x \textit{ approaches } \infty \textit{ is } L,$$

   if $f(x)$ is arbitrarily close to L provided x is sufficiently large and positive. In that case, we say that the graph of $f(x)$ has a *horizontal asymptote* $y = L$. (In fact, we also say this if the limit of $f(x)$ as x approaches $-\infty$ exists; we define what this means in Exercise 1, below.)

2. *Example.* $\lim_{x\to\infty} \dfrac{\sin(x)}{x} = 0.$

Is $\frac{\sin(x)}{x}$ arbitrarily close to 0 provided x is sufficiently large and positive? That is, let $\varepsilon > 0$ be given; can we guarantee that $\frac{\sin(x)}{x}$ is within ε of 0 provided x is sufficiently large and positive?



We can. Indeed, $\frac{\sin(x)}{x}$ can be guaranteed to be within ε of 0 provided $x > \frac{1}{\varepsilon}$, for then $0 < \frac{1}{x} < \varepsilon$; and since $-1 \leq \sin(x) \leq 1$, $-\varepsilon \leq \frac{\sin(x)}{x} \leq \varepsilon$.

C. Exercises

1. Write down definitions and give examples of the following statements.
   (a) $\lim_{x\to-\infty} f(x) = L.$
   (b) $\lim_{x\to\infty} f(x) = -\infty.$
2. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *The graph of a function cannot intersect a horizontal asymptote of the function.*
   (b) *The graph of a function may have no more than two horizontal asymptotes.*
   (c) *If $f(x)$ is defined everywhere and the graph of $f(x)$ has no horizontal asymptotes, then $\lim_{x\to\infty} f(x) = \infty$ or $\lim_{x\to\infty} f(x) = -\infty$.*
3. Prove that the graph of $f(x) = \frac{1}{x}$ has a horizontal asymptote $y = 0$.
4. Determine all the horizontal asymptotes for the graphs of the following functions.
   (a) $f(x) = \frac{3x^2-x-3}{5x^2+x+5}$.
   (b) $f(x) = \sqrt{x^2+2} - x$.
   (c) $f(x) = \sin\left(\frac{1}{x}\right)$.

D. Sequences

1. *Definition.* A *sequence* is an ordered list with a first element but no last element. The notation $\{a_n\}$ denotes the sequence $\{a_1, a_2, a_3, \ldots\}$. (A different first element may be chosen; for example, $\{a_n\}_{n\geq 4}$ denotes the sequence $\{a_4, a_5, a_6, \ldots\}$.) We say that

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

or

$$\{a_n\} \textit{ converges to } L,$$

if a_n is arbitrarily close to L provided n is sufficiently large. A sequence that does not converge is said to *diverge* (see Exercise 1 below for more details).

2. *Examples.* The following are all examples of sequences.
   (a) $\left\{1 - \frac{1}{2^n}\right\} = \left\{\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \ldots\right\}$.
   (b) $\{(-1)^n n\} = \{-1, 2, -3, 4, \ldots\}$.
   (c) $\{1, 1, 2, 3, 5, 8 \ldots\}$. This is the *Fibonacci sequence*, defined recursively by setting $a_1 = a_2 = 1$ and $a_{n+2} = a_n + a_{n+1}$ for $n \geq 1$.

3. *The Bounded Monotone Convergence Theorem.* Suppose we know only that a sequence is increasing and bounded above by 2. The following result allows us to conclude that the sequence converges to some number less than or equal to 2.

   The Bounded Monotone Convergence Theorem. If $\{a_n\}$ is a bounded monotone sequence, then $\{a_n\}$ converges.

   Proof. Suppose $\{a_n\}$ is increasing and bounded. Let L be its least upper bound. We claim that $\{a_n\}$ converges to L. In particular, let $\varepsilon > 0$ be given. Can we guarantee that a_n is within ε of L provided n is sufficiently large? We can: since $\{a_n\}$ is increasing, either *all* a_n eventually satisfy $L - \varepsilon < a_n < L$, or *none* do, in which case $L - \varepsilon$ is an upper bound for $\{a_n\}$, contradicting the fact that L is the sequence's least upper bound. $\square$

   The Bounded Monotone Convergence Theorem is a fundamental result in that it depends explicitly, in the second sentence of the proof, on the *least upper bound property* of real numbers, a property equivalent to the *completeness axiom* for real numbers.

E. Exercises

1. Write down definitions and give examples of the following statements.
   (a) $\{a_n\}$ *diverges to* ∞.
   (b) $\{a_n\}$ *is bounded.*
   (c) $\{a_n\}$ *is monotonic.*
2. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *Bounded sequences converge.*
   (b) *Convergent sequences are bounded.*
   (c) *If* $\{a_n\}$ *diverges to* $-\infty$ *and* $\{b_n\}$ *converges to* 1, *then* $\{a_n b_n\}$ *diverges.*
3. For what values of r is the sequence $\{r^n\}$ convergent?
4. Use the Bounded Monotone Convergence Theorem to prove that the following sequences converge. Then determine what they converge to.
   (a) The sequence defined by $a_1 = 2$ and $a_{n+1} = \frac{1}{2}(a_n + 6)$ for $n \geq 1$.
   (b) The sequence defined by $a_1 = \sqrt{2}$ and $a_{n+1} = \sqrt{2a_n}$ for $n \geq 1$.

Series

A. Motivation

Recall the sequence which arises from Zeno's Dichotomy Paradox,

$$\left\{1 - \frac{1}{2^n}\right\} = \left\{\frac{1}{2}, \frac{1}{2} + \left(\frac{1}{2}\right)^2, \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3, \ldots\right\}.$$

The closed form of the sequence on the left-hand side is derived as follows. Consider the *partial sum* of n terms

$$S_n = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots + \frac{1}{2^n}, \tag{2}$$

from which

$$\frac{1}{2}S_n = \frac{1}{2^2} + \frac{1}{2^3} + \cdots + \frac{1}{2^n} + \frac{1}{2^{n+1}}. \tag{3}$$

Subtracting (15) from (2) and solving for S_n yields

$$S_n = 2\left(\frac{1}{2} - \frac{1}{2^{n+1}}\right) = 1 - \frac{1}{2^n}.$$

S_n is called the n^{th} *partial sum* of the *series*

$$\sum_{n\geq 1}\left(\frac{1}{2}\right)^n = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots,$$

a formal mathematical object representing an "infinite sum". We say that

$$\sum_{n\geq 1}\left(\frac{1}{2}\right)^n = 1,$$

or

$$\sum_{n\geq 1}\left(\frac{1}{2}\right)^n \textit{ converges to } 1,$$

if the sequence of partial sums $\{S_n\}$ converges to 1.

B. Series

1. *Definition.* A *series* is a formal infinite sum of elements. We say that the series

$$\sum_{n\geq 1} a_n = a_1 + a_2 + a_3 + \cdots \textit{ converges to } L$$

if the sequence of partial sums

$$\{S_n\} = \{a_1, a_1 + a_2, a_1 + a_2 + a_3, \ldots\}$$

converges to L. A series that does not converge is said to *diverge*.

2. *Geometric series.* A series of the form $\sum_{n\geq 1} ar^{n-1}$, where a and r are nonzero, is called a *geometric series.* We follow our motivating example above to determine when it converges:

$$\begin{aligned} S_n &= a + ar + ar^2 + \cdots + ar^{n-1}, \\ rS_n &= ar + ar^2 + \cdots + ar^{n-1} + ar^n, \end{aligned}$$

and subtracting one from the other,

$$S_n = \frac{a\,(1 - r^n)}{1 - r}$$

provided $r \neq 1$. We conclude that the series converges to $\frac{a}{1-r}$ when $|r| < 1$, and diverges when $|r| \geq 1$. (That it diverges when $r = 1$ is apparent by considering its unmanipulated partial sums.)

3. *Harmonic series.* The series

$$\sum_{n\geq 1} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots$$

is called the *harmonic series.* In your VANT 140 class, you will have encountered a proof that this series diverges, using consecutively larger groupings of terms. Here we present a proof by contradiction in the same spirit, due to Goldmakher: suppose the harmonic series converges to L; then we have

$$L = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6}\right) + \left(\frac{1}{7} + \frac{1}{8}\right) + \cdots \geq 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = L + \frac{1}{2},$$

a contradiction.

(We will have access to a third proof after defining integrals.)

C. The Divergence Test, the Comparison Test and the Limit Comparison Test

1. *The Divergence Test.* Compare convergent geometric series to the harmonic series. In both cases, the terms of the series converge to 0. What differs is how "quickly" they converge: they converge more quickly in the case of the geometric series, and this appears to make the difference between the series converging to a number and diverging to ∞.

   We claim that for a series to converge, it is necessary but not sufficient for its terms to converge to 0.

   The Divergence Test. If $\sum_{n\geq 1} a_n$ converges, then $\{a_n\}$ converges to 0.

   Proof. Suppose the series converges to L. Note that $a_n = S_n - S_{n-1}$. Thus

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (S_n - S_{n-1}) = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1} = L - L = 0. \quad \square$$

   The contrapositive is occasionally helpful: if $\{a_n\}$ does not converge to 0, then $\sum_{n\geq 1} a_n$ does not converge. This is almost always the first test of convergence that should be applied.

2. *The Comparison Test.* Note the logical direction of the Divergence Test. It is *not* correct to conclude that a series converges because its terms converge to 0 (the harmonic series exemplifies this). The Divergence Test can only indicate that a series might converge; a different test is needed to prove that it does. The majority of such tests involve a comparison with a known series — often the geometric or harmonic series that act as touchstones. The following test is the most direct method of comparison.

The Comparison Test. Let $\sum_{n\geq 1} a_n$ and $\sum_{n\geq 1} b_n$ be series with all positive terms.

(a) If $\sum_{n\geq 1} b_n$ converges and $a_n \leq b_n$ for all n, then $\sum_{n\geq 1} a_n$ converges.

(b) If $\sum_{n\geq 1} b_n$ diverges and $a_n \geq b_n$ for all n, then $\sum_{n\geq 1} a_n$ diverges.

Proof. In the first case, suppose $\sum_{n\geq 1} b_n$ converges to L. Now the partial sums $\{S_n\}$ of $\sum_{n\geq 1} a_n$, and the partial sums $\{T_n\}$ of $\sum_{n\geq 1} b_n$, are both increasing sequences. Indeed, we have $S_n \leq T_n \leq L$. Thus $\{S_n\}$ is a bounded monotone sequence which converges by the Bounded Monotone Convergence Theorem.

Retaining the same notation for partial sums in the second case, we have $S_n \geq T_n$; and since $\{T_n\}$ diverges to ∞, so does $\{S_n\}$. $\square$

3. *Example.* $\sum_{n\geq 1} \dfrac{1}{3^n+2}$ converges.

This follows since $\dfrac{1}{3^n+2} < \dfrac{1}{3^n}$ and $\sum_{n\geq 1} \dfrac{1}{3^n}$ is a convergent geometric series.

4. *The Limit Comparison Test.* Consider $\sum_{n\geq 1} \dfrac{1}{3^n-2}$. This series may not be obviously vulnerable to the the Comparison Test (though it may be configured into a vulnerable form). However, it is plausible that the series converges by virtue of being "similar" to $\sum_{n\geq 1} \dfrac{1}{3^n}$.

The Limit Comparison Test. Let $\sum_{n\geq 1} a_n$ and $\sum_{n\geq 1} b_n$ be series with all positive terms. If

$$\lim_{n\to\infty} \frac{a_n}{b_n} = L > 0,$$

then both series converge or both series diverge.

Proof. Let $0 < C < L < D$. By the definition of limit, $\frac{a_n}{b_n}$ is arbitrarily close to L — in particular, $C < \frac{a_n}{b_n} < D$ — provided n is sufficiently large. Thus we have $Cb_n < a_n < Db_n$ for sufficiently large n, and the conclusion follows by the Comparison Test. $\square$

5. *Example.* $\displaystyle\sum_{n\geq 1} \frac{1}{3^n - 2}$ converges.

This follows since $\displaystyle\lim_{n\to\infty} \frac{\frac{1}{3^n-2}}{\frac{1}{3^n}} = \lim_{n\to\infty} \frac{3^n}{3^n - 2} = 1.$

D. Exercises

1. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *If* $\{a_n\}$ *converges,* $\displaystyle\sum_{n\geq 1} a_n$ *converges.*
   (b) *If* $\displaystyle\sum_{n\geq 1} a_n$ *and* $\displaystyle\sum_{n\geq 1} b_n$ *are convergent series with all positive terms, then* $\displaystyle\sum_{n\geq 1} a_n b_n$ *converges.*
   (c) $\displaystyle\sum_{n\geq 1} a_n$ *converges if, and only if,* $\displaystyle\sum_{n\geq k} a_n$ *converges for any positive integer* k.
2. Determine if the following series converge; and if so, what they converge to.
   (a) $-\frac{1}{3} + 5 - \frac{25}{3} + \frac{125}{9} - \frac{625}{27} + \cdots$.
   (b) $-3 + \frac{1}{5} - \frac{3}{25} + \frac{9}{125} - \frac{27}{625} + \cdots$.
   (c) $\displaystyle\sum_{n\geq 1} \left(\frac{1}{2^n} - \frac{1}{2^{n+1}}\right)$.
3. Determine if the following series converge.
   (a) $\displaystyle\sum_{n\geq 1} 1.1^{1/n}$.
   (b) $\displaystyle\sum_{n\geq 1} \frac{\log(3n)}{n}$.
   (c) $\displaystyle\sum_{n\geq 1} (-1)^n \pi^n$.
4. Determine if the following series converge.
   (a) $\displaystyle\sum_{n\geq 1} \frac{1 + 3^n}{4^n}$.
   (b) $\displaystyle\sum_{n\geq 1} \frac{n}{n^2 - \cos^2(n)}$.
   (c) $\displaystyle\sum_{n\geq 1} \frac{2n^2 + 4n}{\sqrt{5 + n^5}}$.
5. Determine if the following series converge.
   (a) $\displaystyle\sum_{n\geq 1} \frac{1 + (-1)^n}{\sqrt{n}}$.
   (b) $\displaystyle\sum_{n\geq 1} \frac{n! + 1}{(n+1)!}$.
   (c) $\displaystyle\sum_{n\geq 1} \frac{(2n)!}{(n!)^2}$.

6. The *Cantor set* is constructed as follows:
   (a) Begin with the interval $[0, 1]$.
   (b) Remove the open middle third of each remaining interval.
   (c) Return to step (b).

   (For example, in the first iteration, the open interval $\left(\frac{1}{3}, \frac{2}{3}\right)$ is removed.) What is the total length of all of the removed intervals? How many numbers are *not* removed?

7. Let $a_n = \begin{cases} \frac{1}{n} & \text{if } n \text{ does not contain the digit 7} \\ 0 & \text{otherwise} \end{cases}$. Prove that $\displaystyle\sum_{n\geq 1} a_n$ converges.

8. Our version of the Limit Comparison Test above may be extended slightly. Let $\displaystyle\sum_{n\geq 1} a_n$ and $\displaystyle\sum_{n\geq 1} b_n$ be series with all positive terms.
   (a) If $\displaystyle\lim_{n\to\infty} \frac{a_n}{b_n} = 0$ and $\displaystyle\sum_{n\geq 1} b_n$ converges, prove that $\displaystyle\sum_{n\geq 1} a_n$ converges.
   (b) If $\displaystyle\lim_{n\to\infty} \frac{a_n}{b_n} = \infty$ and $\displaystyle\sum_{n\geq 1} b_n$ diverges, prove that $\displaystyle\sum_{n\geq 1} a_n$ diverges.

E. The Ratio Test

1. *The Ratio Test.* Note that the proof of the Limit Comparison Test reveals that it is simply a disguised Comparison Test. Indeed, the theoretical power of *all* convergence tests derives from the Comparison Test. In the Ratio Test, we essentially compare series against geometric series.

   For example, consider $\displaystyle\sum_{n\geq 1} \frac{n^2}{2^n}$. The numerator is polynomial in n, and the denominator is exponential in n. We intuit that the terms of the series converge to 0 so quickly that it may be considered to be "roughly geometric". Indeed,
   $$\lim_{n\to\infty} \frac{\frac{(n+1)^2}{2^{n+1}}}{\frac{n^2}{2^n}} = \lim_{n\to\infty} \frac{1}{2}\left(\frac{n+1}{n}\right)^2 = \frac{1}{2},$$
   so each term is roughly half of the previous term. This strongly suggests the series converges.

   The Ratio Test. Let $\displaystyle\sum_{n\geq 1} a_n$ be a series with all positive terms.
   (a) If $\displaystyle\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L < 1$, then $\displaystyle\sum_{n\geq 1} a_n$ converges.
   (b) If $\displaystyle\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L > 1$ or $\displaystyle\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \infty$, then $\displaystyle\sum_{n\geq 1} a_n$ diverges.

   Proof. In the first case, suppose we have
   $$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L < r < 1.$$

By the definition of limit, $\frac{a_{n+1}}{a_n}$ is arbitrarily close to L — in particular, $\frac{a_{n+1}}{a_n} < r$ — provided n is sufficiently large. Thus we have, for some positive integer k,

$$\begin{array}{ccccc} a_{k+1} & < & a_k r & & \\ a_{k+2} & < & a_{k+1} r & < & a_k r^2 \\ a_{k+3} & < & a_{k+2} r & < & a_k r^3 \\ & \vdots & & & \end{array}$$

and the conclusion follows by applying the Comparison Test to $\sum_{n \geq 1} a_n$ and the convergent series

$$a_1 + \cdots + a_k + a_k r + a_k r^2 + a_k r^3 + \cdots .$$

In the second case, suppose we have

$$\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = L > 1.$$

By the definition of limit, we have $\frac{a_{n+1}}{a_n} > 1$, or equivalently, $a_{n+1} > a_n$, provided n is sufficiently large. Thus $\sum_{n \geq 1} a_n$ diverges by the Divergence Test. $\square$

2. *Example.* $\sum_{n \geq 1} \dfrac{3^n}{n!}$ converges.

   This follows since $\lim_{n \to \infty} \frac{\frac{3^{n+1}}{(n+1)!}}{\frac{3^n}{n!}} = \lim_{n \to \infty} \dfrac{3^{n+1}}{3^n} \cdot \dfrac{n!}{(n+1)!} = \lim_{n \to \infty} \dfrac{3}{n+1} = 0 < 1.$

3. Note that, as for all other convergence tests, the conditions of the Ratio Test must be met for the test to be valid; and even then, the test may be inconclusive.

   Example. $\sum_{n \geq 1} \dfrac{1}{n}$ diverges and $\sum_{n \geq 1} \dfrac{4}{n(n+2)}$ converges.

   The first series is the harmonic series. The second series is a telescoping series which may be shown to converge to 3. (With the tool of integration, it may also be shown to converge by comparison with the convergent p-series $\sum_{n \geq 1} \dfrac{1}{n^2}$, say.) However, the Ratio Test is inconclusive about both series.

F. The Alternating Series Test

1. *Definition.* So far, with the exception of geometric series, we have restricted ourselves to series with positive terms. We now turn our attention to series where that restriction is removed.

   An *alternating series* is one whose terms are alternately positive and negative.

2. *Examples.* $\sum_{n\geq 1}\left(-\frac{1}{2}\right)^{n-1} = 1 - \frac{1}{2} + \frac{1}{2^2} - \frac{1}{2^3} + \cdots$ is an alternating geometric series which converges.

$\sum_{n\geq 1}\frac{(-1)^{n-1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$ is the *alternating harmonic series*, which also converges by the following convergence test.

3. *The Alternating Series Test.* If the terms of an alternating series tend toward 0 in absolute value, the series converges.

The Alternating Series Test. Let $\sum_{n\geq 1}(-1)^{n-1}a_n$, where $a_n > 0$ for $n \geq 1$, be an alternating series satisfying the following criteria:
(a) $a_n \geq a_{n+1}$ for $n \geq 1$, and
(b) $\lim_{n\to\infty} a_n = 0$.
Then the series converges.

Proof. We consider the partial sums $\{S_n\}$. Note that

$$\begin{aligned} S_2 &= a_1 - a_2 \\ S_4 &= S_2 + (a_3 - a_4) \geq S_2 \\ &\vdots \\ S_{2n} &= S_{2n-2} + (a_{2n-1} - a_{2n}) \geq S_{2n-2}. \end{aligned}$$

Thus $\{S_{2n}\}$ is an monotonically increasing sequence. Moreover, since

$$S_{2n} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2n-2} - a_{2n-1}) - a_{2n} \leq a_1,$$

$\{S_{2n}\}$ converges, say to L, by the Bounded Monotone Convergence Theorem.

As for the subsequence $\{S_{2n-1}\}$ of odd partial sums, we have

$$\lim_{n\to\infty} S_{2n-1} = \lim_{n\to\infty}(S_{2n} - a_{2n}) = \lim_{n\to\infty} S_{2n} - \lim_{n\to\infty} a_{2n} = L.$$

Thus $\{S_n\}$ converges to L. □

G. Absolute and conditional convergence

1. *Definition.* We say that $\sum_{n\geq 1} a_n$ *converges absolutely* if $\sum_{n\geq 1}|a_n|$ converges. A series which converges but does not converge absolutely is said to *converge conditionally.*

2. *Examples.* $\sum_{n\geq 1}\left(-\frac{1}{2}\right)^{n-1}$ converges absolutely.

For that matter, $\sum_{n\geq 1}\left(\frac{1}{2}\right)^{n-1}$ also converges absolutely.

However, the alternating harmonic series $\sum_{n\geq 1} \frac{(-1)^{n-1}}{n}$ converges conditionally.

3. Absolute convergence is a stronger criterion than mere convergence. This is plausible, since we suspect that removing the absolute value signs would cause some "cancellation" of previously positive terms, making it easier for the series to converge.

   To prove this, we note first that
$$0 \leq a_n + |a_n| \leq 2\,|a_n|\,.$$
   Therefore if $\sum_{n\geq 1} a_n$ converges absolutely, the inequality above implies that $\sum_{n\geq 1}(a_n + |a_n|)$ converges by the Comparison Test. Thus
$$\sum_{n\geq 1}(a_n + |a_n|) - \sum_{n\geq 1}|a_n| = \sum_{n\geq 1} a_n$$
   also converges.

   We may therefore test for convergence by testing for absolute convergence.

4. *Example.* $\sum_{n\geq 1} \frac{(-1)^n \log(n)}{5^n}$ converges.

   We use the Limit Comparison Test:
$$\lim_{n\to\infty} \frac{\left|\frac{(-1)^n \log(n)}{5^n}\right|}{\left(\frac{3}{5}\right)^n} = \lim_{n\to\infty} \frac{\log(n)}{3^n} = 0.$$
   (The last equality follows from the inequality $\log(n) < 2^n$, for example.) This proves that $\sum_{n\geq 1} \frac{(-1)^n \log(n)}{5^n}$ converges absolutely, whence $\sum_{n\geq 1} \frac{(-1)^n \log(n)}{5^n}$ converges.

H. Exercises

1. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *If* $\sum_{n\geq 1} a_n$ *converges, then* $\sum_{n\geq 1}(-1)^n a_n$ *converges.*
   (b) *If* $\sum_{n\geq 1} a_n$ *converges absolutely, then* $\sum_{n\geq 1}(-1)^n a_n$ *converges absolutely.*
   (c) *If* $\lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right| = \infty$*, then* $\sum_{n\geq 1} a_n$ *converges absolutely.*
2. For each of the following convergence tests, find an example of a series whose convergence is confirmed or disconfirmed by the test, and write down one sentence describing the characteristics of series which are "good candidates" for that test.
   (a) The Divergence Test.

(b) The Comparison Test.
(c) The Limit Comparison Test.
(d) The Alternating Series Test.
(e) The Ratio Test.

3. Determine if the following series converge; and if so, whether they converge absolutely.
   (a) $\displaystyle\sum_{n\geq 1} \frac{(-1)^n 3n}{4n-1}$.
   (b) $\displaystyle\sum_{n\geq 1} \frac{\sin\left(\frac{n\pi}{2}\right)}{n}$.
   (c) $\displaystyle\sum_{n\geq 1} \frac{n}{2-n^2}$.

4. Determine if the following series converge.
   (a) $\displaystyle\sum_{n\geq 1} \frac{(3n)!3^n}{(4n)!}$.
   (b) $\displaystyle\sum_{n\geq 1} \frac{3n!3^n}{(4n)!}$. (Hint: use your answer from (a).)
   (c) $\displaystyle\sum_{n\geq 1} \frac{(4n)!}{n^{40}}$.

5. Determine if the following series converge; and if so, whether they converge absolutely.
   (a) $\displaystyle\sum_{n\geq 1} \frac{(-1)^n}{3^{n-3}}$.
   (b) $\displaystyle\sum_{n\geq 1} \frac{\cos(n\pi)}{\log(\log(n))}$.
   (c) $\displaystyle\sum_{n\geq 1} \frac{(-1)^n n^3}{4^n}$.

6. Determine all values of r for which the series $\displaystyle\sum_{n\geq 1} \frac{1}{n}\left(\frac{r+1}{r}\right)^n$ converges.

7. Let $a_n = \begin{cases} \frac{1}{n} & \text{if } n \text{ is an integer power of } 2 \\ -\frac{1}{n!} & \text{otherwise} \end{cases}$. Does $\displaystyle\sum_{n\geq 1} a_n$ converge?

8. The *Root Test* is similar in spirit to the Ratio Test, but less commonly applicable. Let $\displaystyle\sum_{n\geq 1} a_n$ be a series with all positive terms. The test is in two parts.
   (a) If $\displaystyle\lim_{n\to\infty} a_n^{1/n} = L < 1$, then $\displaystyle\sum_{n\geq 1} a_n$ converges.
   (b) If $\displaystyle\lim_{n\to\infty} a_n^{1/n} = L > 1$, then $\displaystyle\sum_{n\geq 1} a_n$ diverges.

   Prove the Root Test. (Hint: follow the proof of the Ratio Test, and find a suitable geometric series to which to apply the Comparison Test.

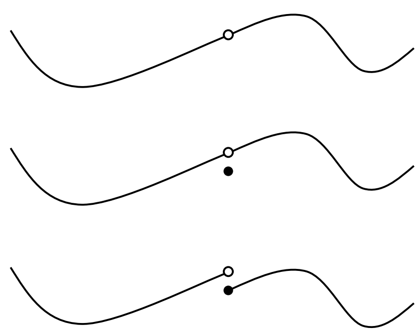
Continuity

A. Motivation

Recall the two motivating problems of calculus: the tangent line problem and the area problem. We defined limits in order to address the tangent line problem, which we interpreted as finding the limit of the slopes of certain secant lines on the graph of a function.

We now return to the tangent line problem proper. We wish to characterize functions for which the problem is solvable. We shall introduce two characteristics: *differentiability*, which is essentially solvability; and a weaker condition, *continuity*.

A continuous function is one whose graph has no holes, as in the top figure below; or jumps, as in the middle and bottom.



Once the definition is stated precisely, it may be used in a number of surprising settings.

B. The definition of continuity

1. *Definition.* We say that a function $f(x)$ is *continuous* at a if $\lim_{x\to x} f(x) = f(a)$. This entails three things:

   (a) that f be defined at a,

   (b) that $\lim_{x\to a} f(x)$ exist, and

   (c) that the two are equal.

   $f(x)$ is *continuous on an interval* if it is continuous at every point in that interval.

2. *Example.* $f(x) = |x|$ is continuous at 0.

   Note that $f(0) = 0$. It remains to show that $\lim_{x\to 0} f(x) = 0$. Is $f(x)$ arbitrarily close to 0 provided x is sufficiently close to 0? That is, let $\varepsilon > 0$ be given; can we guarantee that $f(x)$ is within ε of 0 provided x is sufficiently close to 0? We can. Indeed, $f(x)$ is guaranteed to be within ε of 0 provided x is within ε of 0.

3. Many of the functions encountered in this course are continuous. The following functions are continuous on their domain: polynomials, rational functions, rational powers, sums, products, quotients and composites of continuous functions, trigonometric functions and exponential functions. Proofs of their continuity, using the definition of limit, may be found on Spiderwire.

   (Spiderwire is a tool created by Emily Tyhurst. It provides proofs of all of the standard results, and many nonstandard results, in first-year differential and integral calculus courses, and represents them visually as edges and nodes on a directed graph. It is available online at spiderwire.math.ubc.ca.)

4. In the next section, we highlight two major consequences of continuity. Before doing that, we note here one consequence which is logically trivial but procedurally important. Where a function is continuous, we can take the limit of the function simply by "plugging in" the x-value — *this is not normally permissible.* For example,
$$\lim_{x\to 3} \frac{4x^2 - \sin(x)}{x+2} = \frac{4(3)^2 - \sin(3)}{3+2}.$$

C. Exercises

1. Write down definitions and given examples of the following statements.
   (a) $f(x)$ *is left continuous at* a.
   (b) $f(x)$ *is continuous on* $[l, r]$.
2. Give an example of a function which is continuous everywhere except at $x = 2$ and $x = 4$.
3. Find all value(s) of a which make the following function continuous everywhere:
$$f(x) = \begin{cases} \frac{x^2+x-6}{x+3} & \text{if } x \neq -3 \\ a & \text{if } x = -3 \end{cases}.$$
4. Explain where the following functions are not continuous.
   (a) $f(x) = \frac{x^2}{x}$.
   (b) $f(x) = \begin{cases} \frac{1}{x} & \text{if } x \neq 0 \\ 2 & \text{if } x = 0 \end{cases}$.
   (c) $f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$.

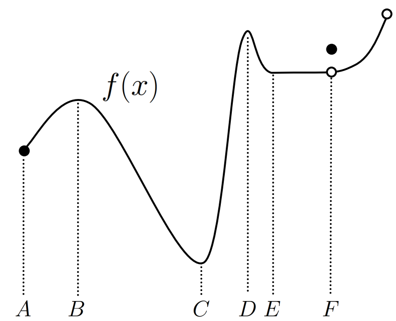
D. Consequences of continuity

1. Our careful definition of continuity is more abstract than necessary in most cases. The benefit of such a definition is its robustness: it allows us to deal with "pathological functions" like the Dirichlet function. The detriment is that it is not as intuitive as some other definitions.

   Two major consequences of continuity exemplify this concession. In both cases, the results seem obvious on first glance, but are not at all obvious on careful examination.

2. *Definitions.* Our first result requires some terminology. $f(x)$ has a *global maximum* (sometimes called an *absolute maximum*) at a if $f(a) \geq f(x)$ for all x in the domain of $f(x)$, and a *global minimum* (sometimes called an *absolute minimum*)at a if $f(a) \leq f(x)$ for all x in the domain of $f(x)$.

   $f(x)$ has a *local maximum* at a if $f(a) \geq f(x)$ provided x is sufficiently close to a, and a *local minimum* at a if $f(a) \leq f(x)$ provided x is sufficiently close to a.

   The function $f(x)$ pictured above has a local maximum at B, a global and local minimum at C, a local maximum at D, a local minimum at E, a local maximum *and* local minimum at every point in (E, F), and a local maximum at F. $f(x)$ has neither a global nor a local extremum at A.

3. *The Extreme Value Theorem.* It is helpful to be able to locate the extrema of a function. The Extreme Value Theorem does not allow this directly, but it guarantees the existence of extrema under the condition of continuity.

   The Extreme Value Theorem. Let $f(x)$ be continuous on $[l, r]$. Then $f(x)$ has global extrema on $[l, r]$.

   Proof. We first prove an intermediate result known as the Boundedness Theorem: $f(x)$ is bounded on $[l, r]$. We use the technique of bisection.

   Our proof is by contradiction. Assume that $f(x)$ is not bounded on $[l, r]$. In particular, assume that $f(x)$ has no upper bound.

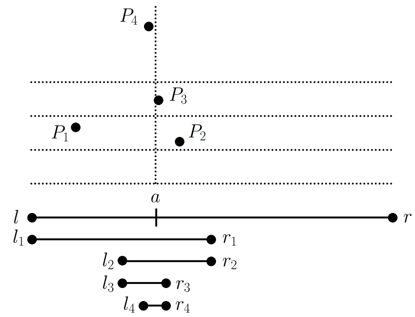
   Divide $[l, r]$ in half. Then $f(x)$ has no upper bound on (at least) one of those halves, which we shall call $[l_1, r_1]$. (If $f(x)$ has no upper bound on either half, we pick the left half.) There exists a point P_1 on the graph of the function on $[l_1, r_1]$ whose y-value is larger than 1.

   Again, divide $[l_1, r_1]$ in half, and select a half $[l_2, r_2]$ with no upper bound. There exists a point P_2 on the graph of the function on $[l_1, r_1]$ whose y-value is larger than 2.

   We repeat the process *ad infinitum*, getting a sequence of nested intervals

   $$[l, r] \supset [l_1, r_1] \supset [l_2, r_2] \supset [l_3, r_3] \supset \cdots$$

   whose lengths are halved each time, with a number P_n on the graph of the function on $[l_n, r_n]$ whose y-value is larger than n.



This yields a contradiction because we have constructed something like a vertical asymptote. The sequence of nested intervals converges to a single point a. We must have $\lim_{x\to a} f(x) = f(a)$ since $f(x)$ is continuous at a. But is $f(x)$ arbitrarily close to $f(a)$ provided x is sufficiently close to a? That is, let $\varepsilon > 0$ be given; can we guarantee that $f(x)$ is within ε of $f(a)$ provided x is sufficiently close to a? We cannot. For example, let $\varepsilon = 1$. No matter how small an interval $(a-\delta, a+\delta)$ to which we restrict x, there exist x-values in that interval — in particular, the ones corresponding to points $P_k, P_{k+1}, P_{k+2}, \ldots$ for some positive integer $k > f(a) + \varepsilon$ — whose corresponding y-values are too large.

If we assume that $f(x)$ has no lower bound, a similar contradiction may be derived.

This concludes the proof of the Boundedness Theorem. To prove the Extreme Value Theorem, we simply add the claim that $f(x)$ attains its least upper bound and its greatest lower bound. To see that it attains its least upper bound U, for example, assume it does not; then $\frac{1}{U-f(x)} < V$ for some $V > 0$ by the Boundedness Theorem. Hence $f(x) < U - \frac{1}{V}$, which contradicts the "leastness" of U. $\square$

4. *Example.* Given a fixed perimeter, there is a rectangle of maximum area of that perimeter.

   We fix the perimeter to be P and consider a rectangle of side lengths x and y. Now $2x + 2y = P$, whence $y = \frac{1}{2}(P - 2x)$. Thus the area of the rectangle is given by $A(x) = xy = \frac{1}{2}x(P - 2x)$, which is continuous and, by the Extreme Value Theorem, attains a global maximum on $\left[0, \frac{P}{2}\right]$.

   Later in this course, we develop related tools which may be used to determine the side lengths of the rectangle. This may also be done by completing the square of $A(x)$.

5. *The Intermediate Value Theorem.* The Intermediate Value Theorem guarantees an important naïve notion of continuity: if a continuous function starts below a line and ends above the line, it must cross the line. Like the Extreme Value Theorem, it is an *existence theorem*: it guarantees the existence of a "crossing point", but does not provide a method to find it.

   The Intermediate Value Theorem. Let $f(x)$ be continuous on $[l, r]$. Then for any number L between $f(l)$ and $f(r)$, there exists a number a in $[l, r]$ such that $f(a) = L$.

   The proof is left as an exercise. (Like the Boundedness Theorem above, the Intermediate Value Theorem may be proven by bisection: find an sequence of nested intervals

$$[l, r] \supset [l_1, r_1] \supset [l_2, r_2] \supset [l_3, r_3] \supset \cdots$$

   such that L is "bracketed" by $f(l_n)$ and $f(r_n)$.)

6. *Example.* $f(x) = x^3 - x^2 + x - 1$ has a root in $[0, 2]$.
   Since $f(x)$ is a polynomial, it is continuous. Moreover, since $f(0) = -1$ and $f(2) = 5$, by the Intermediate Value Theorem there exists a number a in $[0, 2]$ such that $f(a) = 0$.

E. Exercises

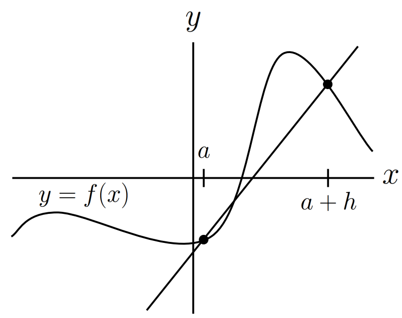
1. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *If $f(x)$ is defined everywhere, then $f(x)$ attains a global maximum on* $[0, 2]$.
   (b) *If $f(x)$ is continuous on $[l, r]$, then $f(x)$ attains a local maximum on* $[l, r]$.
   (c) *If $f(1) = 2$ and $f(3) = 4$, there exists a number $1 < a < 3$ such that* $f(a) = 3$.
2. Prove that $f(x) = x^3 - 15x + 1$ has three roots in $[-4, 4]$.
3. Determine where $f(x) = x^3 - 9x$ is positive and where it is negative, and explain how the Intermediate Value Theorem is applied, if you apply it.
4. Let $f(x)$ be continuous and $0 \leq f(x) \leq 1$ on $[0, 1]$. Prove that there exists a number a in $[0, 1]$ such that $f(a) = a$. Hint: apply the Intermediate Value Theorem to $g(x) = f(x) - x$.

Derivatives

A. Motivation

In our consideration of the tangent line problem, we define two characteristics of functions: differentiability and continuity. Having defined continuity, we now define the stronger condition of differentiability.

Roughly speaking, a differentiable function is one whose graph is smooth: as we zoom in on the graph, it looks "more linear" — but with what slope?



We use the mechanism introduced earlier. We denote the slope of the tangent line at a to be $f'(a)$, and define it to be the limit of the slope of the secant line between $(a, f(a))$ and $(a+h, f(a+h))$ as h approaches 0 —

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

— provided this limit exists.

B. The definition of derivative

1. *Definition.* The *derivative* of $f(x)$, or *the derivative of f with respect to x*, is another function

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

   where this limit exists. We say that a function $f(x)$ is *differentiable* at a if $f'(a)$ exists. $f(x)$ is *differentiable on an interval* if it is differentiable at every point in that interval.

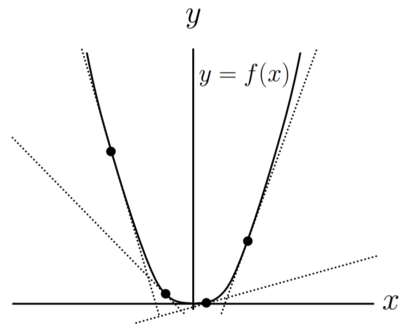
2. *Notation.* There are two common forms of notations for derivatives. $f'(x)$ is sometimes denoted $\frac{df}{dx}$. This may be extended as expected: the second, third and fourth derivatives of $f(x)$ may be respectively denoted $f''(x)$, $f^{(3)}(x)$ and $f^{(4)}(x)$; or $\frac{d^2f}{dx^2}$, $\frac{d^3f}{dx^3}$ and $\frac{d^4f}{dx^4}$. We use this notation interchangeably depending on which is more convenient.

3. *Example.* Let $f(x) = x^2$. Then $f'(x) = 2x$.

   By the definition of derivative,

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

This is plausible by looking at some slopes of tangent lines to the graph of $f(x)$: like $2x$, the slopes are negative when x is negative, and positive when x is positive.



4. *Example.* Let $f(x) = |x|$. Then $f(x)$ is not differentiable at 0.

   By the definition of derivative,

$$f'(0) = \lim_{h \to 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \to 0} \frac{|h|}{h}. \tag{4}$$

   However,

$$\frac{|h|}{h} = \begin{cases} -1 & \text{if } h < 0 \\ 1 & \text{if } h \geq 0 \end{cases},$$

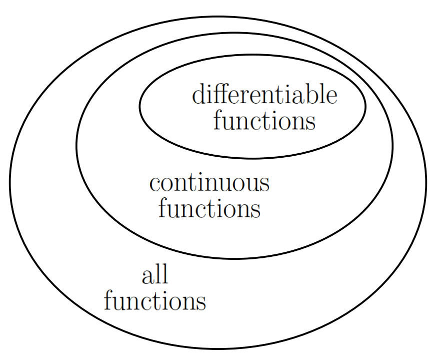
   hence the limit (4) does not exist.

   This is plausible by looking at the graph of $f(x)$. The slope of the secant line between $(0, 0)$ and $(h, |h|)$ is equal to -1 for negative h and 1 for positive h, and there is no "convergence" as h approaches 0.

5. From the point of view of problem solving, the definition of derivative — and even more fundamentally, the definition of limit — are default tools; they are often tedious to use, but they are broadly applicable. Just as continuity provides a fast way of calculating limits, there are shortcuts to calculating derivatives of certain classes of functions. However, the definition of derivative will always be a backup.

C. Differentiability and continuity

1. Differentiability is a stronger condition than continuity.



Certainly there are functions which are not continuous — for example, the greatest integer function $\lfloor x \rfloor$. The previous example, $|x|$, provides an instance of a function which is continuous but not differentiable. It remains to show that all differentiable functions are continuous.

2. **Claim.** Let $f(x)$ be differentiable at a. Then $f(x)$ is continuous at a.

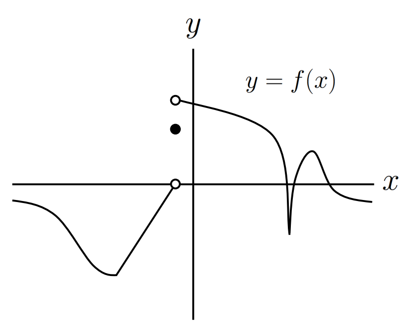
   Proof. $f(x)$ is continuous at a if $\lim_{x \to a} f(x) = f(a)$; or equivalently, if $\lim_{h \to 0} (f(a+h) - f(a)) = 0$. Now

$$\begin{aligned} \lim_{h \to 0} (f(a+h) - f(a)) &= \lim_{h \to 0} h \left(\frac{f(a+h) - f(a)}{h} \right) \\ &= \lim_{h \to 0} h \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \\ &= 0 \cdot f'(a) = 0. \quad \square \end{aligned}$$

   (The penultimate equality follows from the fact that the limit of a product is equal to the product of the limits if all of the limits exist; a proof of this is on spiderwire.math.ubc.ca.)

D. Exercises

1. Write down definitions and sketch illustrations of the following statements.
   (a) $f(x)$ is right-differentiable at a.
   (b) $f(x)$ has a vertical tangent line at a.
2. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *If $f(x)$ is differentiable, then $|f(x)|$ is differentiable.*
   (b) *If $|f(x)|$ is differentiable, then $f(x)$ is differentiable.*
   (c) *If $f(x)$ has a vertical asymptote $x = a$, then $f(x)$ is not differentiable at a.*
3. Let $f(x)$ have the graph below. Sketch the graph of $f'(x)$.



4. Sketch the graph, and come up with the expression for, a function which is continuous everywhere except at $x = 2$, and differentiable everywhere except at $x = 2$ and $x = 4$.
5. Let $f(x) = 2x + 3$. Confirm using the definition of derivative that $f'(x) = 2$. Why is this plausible?
6. Let $f(x) = \frac{1}{x}$. Calculate $f'(x)$ using the definition of derivative.
7. Find the equation of the line tangent to $y = 3x^2 + 2x$ at $x = 1$. Then sketch the curve and the tangent line.
8. There are two lines through $(-1, 1)$ tangent to $y = \frac{1}{x}$. Where do those lines intersect the curve?

The Power, Product and Quotient Rules

A. Motivation

In principle, we can now differentiate anything using the definition of derivative. But consider the tediousness of differentiating a function like $f(x) = \frac{(x^{40}-7x^5)(5x^3+x-2)}{x^2+1}$. It turns out that there are useful shortcuts.

B. Derivatives of polynomials

1. *The Power Rule.* Our immediate aim is to be able to differentiate *rational functions*, or quotients of polynomials, quickly. In order to differentiate polynomials themselves, we appeal to the Power Rule.

   The Power Rule. The derivative of x^n is nx^{n-1}.

   Proof. (Here, we prove only that the Power Rule holds for positive integers n. It is trivial to show that it holds for $n = 0$. We prove in a later exercise that it holds for negative integers n. After proving the Chain Rule, we will be able to prove the Power Rule for all rational numbers n. However, proving the Power Rule for irrational n — in fact, even defining x^n for irrational n — requires further definitions and results, which may be found on spiderwire.math.ubc.ca.)

   By the definition of derivative,

   $$\frac{d}{dx}\,x^n = \lim_{h\to 0}\frac{(x+h)^n - x^n}{h} = \lim_{h\to 0}\frac{x^n + nx^{n-1}h + h^2 f(x,h) - x^n}{h},$$

   where $f(x,h)$ is a polynomial in x and h. Thus

   $$\frac{d}{dx}\,x^n = \lim_{h\to 0}\big(nx^{n-1} + hf(x,h)\big) = nx^{n-1}. \quad \square$$

2. *Derivatives of constant multiples and sums.* In order to be able differentiate polynomials quickly, we require one more result.

   Let $f(x)$ and $g(x)$ be differentiable, and c and d be constants. Then by the definition of derivative,

   $$\begin{aligned}(cf(x)+dg(x))' &= \lim_{h\to 0}\frac{cf(x+h)+dg(x+h)-cf(x)-dg(x)}{h}\\ &= c\lim_{h\to 0}\frac{f(x+h)-f(x)}{h} + d\lim_{h\to 0}\frac{g(x+h)-g(x)}{h}\\ &= cf'(x)+dg'(x).\end{aligned}$$

   It follows by mathematical induction that for differentiable functions $f_1(x), f_2(x), \ldots, f_n(x)$ and constants $c_1, c_2, \ldots, c_n$,

   $$(c_1f_1(x) + c_2f_2(x) + \cdots + c_nf_n(x))' = c_1f_1'(x) + c_2f_2'(x) + \cdots + c_nf_n'(x).$$

   This, combined with the Power Rule, allows us to differentiate polynomials quickly.

3. *Examples.* $\frac{d}{dx}\left(x^{40}-7x^{5}\right)=40x^{39}-35x^{4}$, and $\frac{d}{dx}\left(5x^{3}+x-2\right)=15x^{2}+1$.

C. Derivatives of rational functions

1. *The Product Rule.* In order to be able to differentiate rational functions quickly, it remains only to find a technique for differentiating quotients. In fact, this is a corollary of a powerful technique for differentiating products.

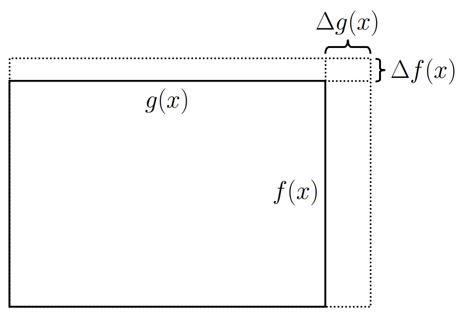
   The Product Rule. Let $f(x)$ and $g(x)$ be differentiable. Then $(f(x)g(x))'=f'(x)g(x)+f(x)g'(x)$.

   We provide an informal but insightful proof, as well as a formal one.

   Informal proof. Imagine $f(x)$ and $g(x)$ are lengths, and their product $h(x)$ is the area of the rectangle with side lengths $f(x)$ and $g(x)$. Suppose the side lengths change with respect to time x. We want to find the derivative $h'(x)$, which is the limit of the ratio

$$\frac{\text{change in } h(x)}{\text{change in } x}=\frac{\Delta h(x)}{\Delta x}$$

   as Δx approaches 0.



   In the picture, the change in the area of the rectangle is given by the area of three smaller rectangles:

$$\Delta f(x)g(x)+f(x)\Delta g(x)+\Delta f(x)\Delta g(x).$$

   Dividing through by Δx, we get

$$\left(\frac{\Delta f(x)}{\Delta x}\right)g(x)+f(x)\left(\frac{\Delta g(x)}{\Delta x}\right)+\left(\frac{\Delta f(x)}{\Delta x}\right)\Delta g(x),$$

   whose limit is simply

$$f'(x)g(x)+f(x)g'(x)+f'(x)\cdot 0=f'(x)g(x)+f(x)g'(x).\quad\Box$$

   Proof. By the definition of derivative,

$$
\begin{aligned}
(f(x)g(x))' &= \lim_{h\to 0}\frac{f(x+h)g(x+h)-f(x)g(x)}{h}\\
&= \lim_{h\to 0}\frac{f(x+h)g(x+h)-f(x)g(x+h)+f(x)g(x+h)-f(x)g(x)}{h}\\
&= \lim_{h\to 0}\left(\left(\frac{f(x+h)-f(x)}{h}\right)g(x+h)+f(x)\left(\frac{g(x+h)-g(x)}{h}\right)\right)\\
&= f'(x)g(x)+f(x)g'(x). \quad \square
\end{aligned}
$$

2. *Example.* $\frac{d}{dx}\left(\left(x^{40}-7x^5\right)\left(5x^3+x-2\right)\right) = \left(40x^{39}-35x^4\right)\left(5x^3+x-2\right)+\left(x^{40}-7x^5\right)\left(15x^2+1\right)$.

3. *The Quotient Rule.* The rule for differentiating quotients follows.

   The Quotient Rule. Let $f(x)$ and $g(x)$ be differentiable. Then $\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$.

   Proof. By the Product Rule,

$$\left(\frac{f(x)}{g(x)}\right)' = \left(f(x)\frac{1}{g(x)}\right)' = \frac{f'(x)}{g(x)} + f(x)\left(\frac{1}{g(x)}\right)'.$$

   By the definition of derivative,

$$
\begin{aligned}
\left(\tfrac{1}{g(x)}\right)' &= \lim_{h\to 0}\frac{\frac{1}{g(x+h)}-\frac{1}{g(x)}}{h}\\
&= \lim_{h\to 0}\frac{g(x)-g(x+h)}{hg(x+h)g(x)}\\
&= \lim_{h\to 0}\frac{-1}{g(x+h)g(x)}\cdot\frac{g(x+h)-g(x)}{h}\\
&= -\tfrac{g'(x)}{g(x)^2}.
\end{aligned}
$$

   Thus

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)}{g(x)} - f(x)\frac{g'(x)}{g(x)^2} = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}. \quad \square$$

4. *Example.*

$$
\begin{aligned}
&\tfrac{d}{dx}\left(\tfrac{\left(x^{40}-7x^5\right)\left(5x^3+x-2\right)}{x^2+1}\right)\\
&= \tfrac{\left(\left(40x^{39}-35x^4\right)\left(5x^3+x-2\right)+\left(x^{40}-7x^5\right)\left(15x^2+1\right)\right)(1+x)-\left(x^{40}-7x^5\right)\left(5x^3+x-2\right)(2x)}{(x^2+1)^2}.
\end{aligned}
$$

D. Exercises

1. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *If $f(x)$ and $g(x)$ are differentiable, then $(f(x)g(x))' = f'(x)g'(x)$.*
   (b) *Given a function $f(x)$ which is infinitely differentiable on its domain, repeated differentiation will eventually cause the function to go to zero; that is, there exists some n such that $f^{(n)}(x) = 0$.*

(c) *A rational function is differentiable everywhere on its domain.*

2. Differentiate the following functions.

   (a) $f(x) = g_1(x)g_2(x)g_3(x)$, where $g_1(x)$, $g_2(x)$ and $g_3(x)$ are differentiable.

   (b) $f(x) = \left(1 - x - x^2\right)\left(x^3 + \frac{3}{x^3}\right)$.

   (c) $f(x) = \dfrac{x^2}{2x + \frac{1}{3x+4}}$.

3. Determine where $f(x) = \left|x^3\right|$ is differentiable.

4. Find the equation of the line tangent to $y = \dfrac{2}{3 - 4x}$ at $y = -2$.

5. Find the equation of the horizontal line tangent to $y = \frac{x^2-1}{x^2+1}$.

6. Find the equation of the line tangent to $y = \frac{x-1}{x+1}$ which passes through $(-1, 0)$.

7. Let n be a positive integer.

   (a) Prove using the definition of derivative that the derivative of x^{-n} is $-nx^{-n-1}$.

   (b) Prove using the Quotient Rule that the derivative of x^{-n} is $-nx^{-n-1}$.

8. (a) Calculate the first few derivatives of $f(x) = \frac{1}{x}$, and make a conjecture for the n^{th} derivative of $f(x)$.

   (b) Confirm that your conjecture is true for $n = 1$.

   (c) Prove that if your conjecture is true for $n = k$, then it is true for $n = k + 1$.

   (d) Explain why this implies that your conjecture is true for all positive integers n.

The Chain Rule, implicit differentiation and related rates

A. Motivation

We are now able to differentiate sums, differences, products and quotients of differentiable functions. This week, we turn our attention to compositions of differentiable functions.

Functions like $f(x) = (3x + 2)^{10}$ may be differentiated using the Product Rule, but there is also a useful shortcut known as the *Chain Rule*. This rule has two major consequences in addition to its utility. First, it allows us to solve the tangent line problem for a large number of curves (which are not necessarily graphs of functions) using a technique known as *implicit differentiation*. Second, it allows us to consider a class of applied problems known as *related rates* problems, in which we wish to find the rate of change of a particular quantity given the rate of change of other quantities.

B. The Chain Rule

1. *Example.* We wish to calculate the derivative of $f(x) = (3x+2)^{10}$. Our initial guess is: $f'(x) = 10(3x+2)^9$, *since* $\frac{d}{dx}x^{10} = 10x^9$.

   However, we have not accounted for how the difference between $(3x+2)$ and x might affect the derivative. In particular, $(3x+2)$ is three times as "steep" as x; thus we "run through" x-values three times as quickly, and expect the derivative to be three times as big. Our revised guess is therefore: $f'(x) = 30(3x + 2)^9$, *since* $\frac{d}{dx}x^{10} = 10x^9$ *and the "inside function"* $(3x + 2)$ *is three times as "steep" as* x.

   This guess can be confirmed using the definition of derivative:

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{(3(x+h)+2)^{10} - (3x+2)^{10}}{h} \\ &= 3\lim_{h\to 0} \frac{(3x+2+3h)^{10} - (3x+2)^{10}}{3h} \\ &= 3\lim_{(3h)\to 0} \frac{(3x+2+(3h))^{10} - (3x+2)^{10}}{(3h)}. \end{aligned}$$

   Note that the limit in the third line is the definition of the derivative of X^{10} evaluated at $X = 3x + 2$; namely, $10(3x + 2)^9$. Therefore $f'(x) = 30(3x + 2)^9$.

2. **The Chain Rule.** Let $h(x) = f(g(x))$. Then $h'(x) = f'(g(x))g'(x)$, provided $f'(g(x))$ and $g'(x)$ are defined.

   Proof. By the definition of derivative,

$$h'(x) = \lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{h}. \tag{5}$$

   We wish to argue that this is equal to

$$\lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{g(x+h) - g(x)} \cdot \frac{g(x+h) - g(x)}{h} = \lim_{h\to 0} \frac{f(g(x)+H) - f(g(x))}{H} \cdot \frac{g(x+h) - g(x)}{h} \tag{6}$$

for $H = g(x+h) - g(x)$. This is a legitimate argument provided $g(x+h) - g(x) \neq 0$. Unfortunately, this is not guaranteed — for example, consider $g(x) = x^2 \sin\left(\frac{1}{x}\right)$.

The remainder of the proof is devoted to bypassing this obstacle. We do this by finding an expression for the denominator in (5) which "isolates" $g(x+h) - g(x)$ as in (6), but without possible division by 0. Let

$$E(H) = \begin{cases} \frac{f(g(x)+H)-f(g(x))}{H} - f'(g(x)) & \text{if } H \neq 0 \\ 0 & \text{if } H = 0 \end{cases}.$$

Regardless of whether $H = 0$,

$$f(g(x)+H) - f(g(x)) = (f'(g(x)) + E(H))\, H.$$

Taking $H = g(x+h) - g(x)$, we get

$$f(g(x+h)) - f(g(x)) = (f'(g(x)) + E(H))\,(g(x+h) - g(x)).$$

Substituting this into (5) yields

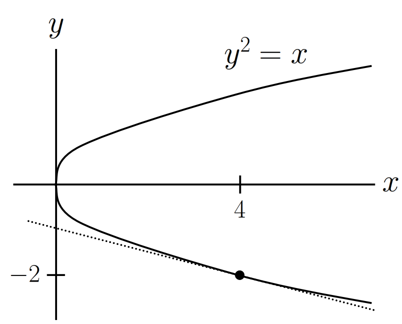
$$h'(x) = \lim_{h \to 0} (f'(g(x)) + E(H))\, \frac{g(x+h) - g(x)}{h} = (f'(g(x)) + 0)\, g'(x) = f'(g(x))g'(x). \quad \square$$

C. Exercises

1. Differentiate the following functions.
   (a) $f(x) = (x^3 + 2x^2 + 3)^{10}$.
   (b) $f(x) = (x^3 + 2x^2 + 3)^{10}(x + \frac{1}{x})^2$.
   (c) $f(x) = \left(\left(x^4 - \frac{5}{x^2}\right)^3 + x + 1\right)^2$.
2. Find the equation of the line tangent to $y = \frac{1}{(1+2x^2)^3}$ at $x = 1$.
3. Find the equation of the line tangent to $y = (ax+b)^3$ at $x = \frac{b}{a}$.
4. Let $h(x) = (f(x))^4$, and suppose $f(0) = \frac{1}{2}$ and $f'(0) = \frac{7}{3}$. Find the equation of the line tangent to $y = h(x)$ at $x = 0$.

D. Consequences of the Chain Rule

1. *Example.* We wish to find the slope of the line tangent to the curve $y^2 = x$ at the point $(4, -2)$.



The curve is not the graph of a function, but the top half and bottom half both are. One approach is to consider the bottom half as the function $f(x) = -\sqrt{x}$, and differentiate, getting

$$f'(x) = -\frac{1}{2x^{1/2}},$$

whence

$$f'(4) = -\frac{1}{2 \cdot 4^{1/2}} = -\frac{1}{4}.$$

Note that, at this point, this derivative must be calculated using the definition of derivative, as we have proven the Power Rule only for integer powers.

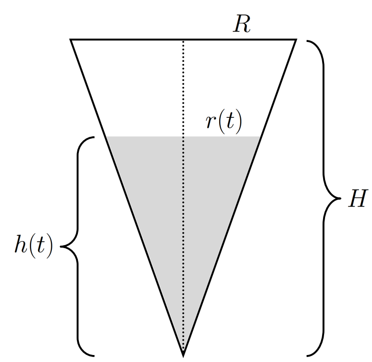
However, we can also simply differentiate both sides of the equation $y^2 = x$. The derivative of the right-hand side is 1. The derivative of the left-hand side may be calculated using the Power Rule and the Chain Rule to be $2yy'$. Solving $2yy' = 1$ for y' yields $y' = \frac{1}{2y}$, which at the point $(4, -2)$ is equal to $-\frac{1}{4}$.

2. *Implicit differentiation.* We call this technique of defining y as an implicit function of x in order to differentiate it *implicit differentiation.* It is particularly useful when it is difficult to identify curves as unions of functions.

3. *Related rates.* The Chain Rule and implicit differentiation are also used in *related rates* problems. In these problems, we encounter multiple quantities whose relationship is described by an equation. The equation can then be differentiated, usually with respect to time, to determine the relationship between the rates of change of the quantities.

4. *Example.* Suppose a small hole is cut in the bottom of a conical cup of height H and radius R, so that water empties from the cup at a constant rate. We wish to determine how quickly the water level is dropping when it is equal to half the height of the cup.

   Let $V(t)$, $r(t)$ and $h(t)$ be the volume, radius and height of water in the cup at time t, respectively. Let R and H be the radius and height of the cup, respectively.



   We are told that $V'(t)$ is constant, and we wish to find $h'(t)$ when $h(t) = \frac{H}{2}$.

   Now $V(t) = \frac{1}{3}\pi r(t)^2 h(t)$ (though we note that we are stating this without proof). For convenience we rewrite $r(t)$ in terms of $h(t)$: by similar triangles, we have $r(t) = \frac{R}{H}h(t)$, so

$$V(t) = \frac{\pi R^2}{3H^2}h(t)^3.$$

Differentiating, we get

$$V'(t) = \frac{\pi R^2}{H^2} h(t)^2 h'(t).$$

When $h(t) = \frac{H}{2}$, we have

$$h'(t) = \frac{4V'(t)}{\pi R^2}. \tag{7}$$

This is the rate at which the height of the water is changing.

We can check that this solution is at least plausible. First, the units on either side of (7) match. Second, we can “plug in” some extreme values for $V'(t)$ and check that our physical intuition is described accurately — for example, if $V'(t) = 0$ (no water is leaking), then $h'(t) = 0$ (the water level is not dropping).

5. In general, solutions to related rates problems will follow the same steps, which we enumerate here.
   (a) Draw and label a picture.
   (b) Identify what is known, and what is to be found.
   (c) Find an equation relating the quantities whose rates of change are known to the quantity whose rate of change is to be found.
   (d) Differentiate the equation implicitly, and solve.
   (e) Check the solution for plausibility.

   In step 3, geometric formulas — for example, similar triangles, the Pythagorean theorem, and area and volume formulas — are often helpful.

E. Exercises

1. Using the Power Rule for integer powers, prove the Power Rule for rational powers using implicit differentiation.
2. The curve $x^{2/3} + y^{2/3} = 1$ is called an astroid. Sketch the astroid and find the equations of all lines of slope -1 that are tangent to the curve.
3. The Bubble Nebula is an expanding sphere of gas and stellar ejecta in the constellation Cassiopeia. The radius of the Bubble Nebula is $3(10^{13})$ km, and it is expanding at a rate of $7(10^6)$ km/h. Determine how quickly the volume of the nebula is increasing.
4. Consider an aircraft flying north at 600 km/h, at an altitude of 4 km, passing directly overhead a car driving east at 100 km/h. Determine how fast the distance between them is changing one hour after the aircraft passes overhead the car.

Trigonometric and exponential derivatives

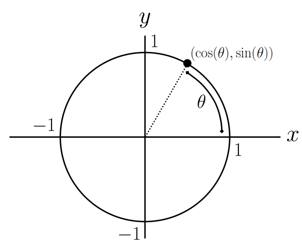
A. Motivation

One of the major uses of calculus is *mathematical modelling*, or using mathematics to describe and predict physical systems. It is useful to have a large toolbox of descriptive functions. Rational functions, on which we have be focussed so far, are very useful. To those, we now add trigonometric and exponential functions. Trigonometric functions are useful in describing periodic phenomena — for example, the position of a pendulum or the height of a tide. Exponential functions are useful in describing quantities where the rate of change of the quantity is proportional to the quantity itself — for example, populations, resources collecting compound interest, or chemical reactants.

B. Trigonometric functions and derivatives

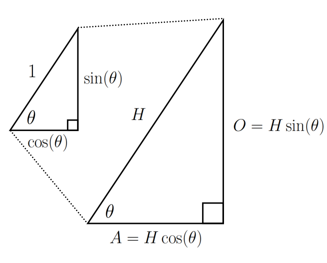
1. *Definitions.* We give three standard descriptions of the functions $\sin(\theta)$, $\cos(\theta)$ and $\tan(\theta)$ (we shall add a fourth description in terms of power series next term).

   Our first description is as follows. Given an angle, or equivalently, an arclength θ on the unit circle, $\sin(\theta)$ and $\cos(\theta)$ are the y- and x- values, respectively, of the point on the circle corresponding to that angle. (The identity $\sin^2(\theta) + \cos^2(\theta) = 1$ follows immediately.) We also define $\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)}$.



   Our second description is in terms the graphs of the functions. This is left as part of Exercise 1, below.

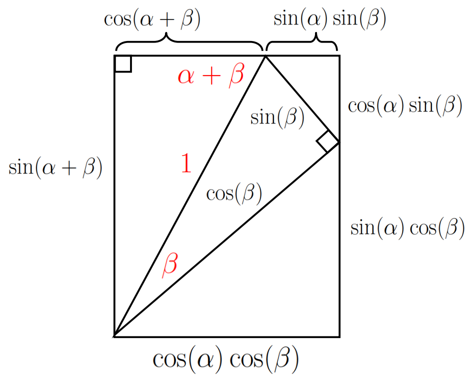
   Our third description is in terms of special triangles. In the figure below, the triangle on the left is the triangle inscribed in the unit circle. The triangle on the right is the similar triangle whose side lengths are H times bigger. The side lengths are labelled H for "hypotenuse", O for "opposite" and A for "adjacent", whence $\sin(\theta) = \frac{O}{H}$, $\cos(\theta) = \frac{A}{H}$ and $\tan(\theta) = \frac{O}{A}$.



2. *The derivative of* $\sin(\theta)$. By the definition of derivative,

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

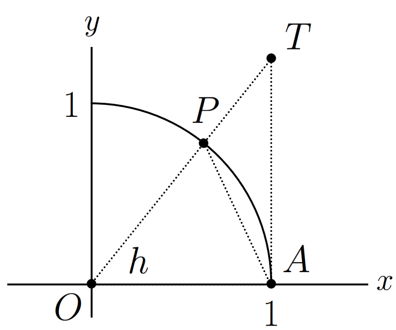
We derive an addition formula allowing us to expand $\sin(\theta+h)$. In the picture below, the information labelled in red implies the information labelled in black.



From it, we conclude that $\sin(\alpha+\beta) = \sin(\alpha)\cos(\beta)+\cos(\alpha)\sin(\beta)$. (Note that the picture suggests, but does not constitute, a full proof: for example, either α or β might be negative.) Thus

$$\frac{d}{d\theta}\sin(\theta) = \lim_{h\to 0}\frac{\sin(\theta)\cos(h)+\cos(\theta)\sin(h)-\sin(\theta)}{h} = \lim_{h\to 0}\left(\sin(\theta)\left(\frac{\cos(h)-1}{h}\right)+\cos(\theta)\left(\frac{\sin(h)}{h}\right)\right).$$

It remains to calculate $\lim\limits_{h\to 0}\dfrac{\sin(h)}{h}$ and $\lim\limits_{h\to 0}\dfrac{\cos(h)-1}{h}$.



The triangle ΔOPA has area $\frac{1}{2}\sin(h)$. The triangle ΔOTA has area

$$\frac{1}{2}|TA| = \frac{|TA|/|TO|}{2\,(1/|TO|)} = \frac{\sin(h)}{2\cos(h)}.$$

The area of the sector of the circle described by the angle h is $\frac{h}{2}$. This sector is larger than the triangle ΔOPA and smaller than the triangle ΔOTA; thus

$$\frac{\sin(h)}{2} \le \frac{h}{2} \le \frac{\sin(h)}{2\cos(h)},$$

whence

$$1 \leq \frac{h}{\sin(h)} \leq \frac{1}{\cos(h)}.$$

Taking the limit as h approaches 0 yields

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

Finally,

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

and we conclude that

$$\frac{d}{d\theta} \sin(\theta) = \lim_{h \to 0} \left(\sin(\theta) \left(\frac{\cos(h) - 1}{h} \right) + \cos(\theta) \left(\frac{\sin(h)}{h} \right) \right) = \cos(\theta).$$

3. *The derivative of* $\cos(\theta)$. Many of the same techniques are used. By the definition of derivative, the addition formula $\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$, and the limits derived above,

$$\begin{aligned}
\tfrac{d}{d\theta} \cos(\theta) &= \lim_{h \to 0} \frac{\cos(\theta + h) - \cos(\theta)}{h} \\
&= \lim_{h \to 0} \frac{\cos(\theta)\cos(h) - \sin(\theta)\sin(h) - \cos(\theta)}{h} \\
&= \lim_{h \to 0} \left(\cos(\theta) \left(\frac{\cos(h) - 1}{h} \right) - \sin(\theta) \left(\frac{\sin(h)}{h} \right) \right) \\
&= -\sin(\theta).
\end{aligned}$$

C. Exercises

1. (a) Sketch the graphs of $\sin(\theta)$, $\cos(\theta)$, $\tan(\theta)$, $\csc(\theta) = \frac{1}{\sin(\theta)}$, $\sec(\theta) = \frac{1}{\cos(\theta)}$ and $\cot(\theta) = \frac{1}{\tan(\theta)}$.
   (b) Calculate the derivatives of $\tan(\theta)$, $\csc(\theta)$, $\sec(\theta)$ and $\cot(\theta)$.
2. Differentiate the following functions:
   (a) $f(x) = \sin^3(x) + \tan^3(x)$.
   (b) $f(x) = \left(\frac{\cos(x)}{x^2 \sin(x)} \right)^4$.
   (c) $f(x) = x^2 \cos^2(\sin(x))$.
3. Find the equation of the line tangent to $y = \cos^2(x)$ at $x = \frac{\pi}{3}$.
4. Let

$$f(x) = \begin{cases} \tan(x) & \text{if } x < 0 \\ g(x) & \text{if } 0 \leq x \leq \pi \\ \cos(\sin(x)) & \text{if } x > \pi \end{cases}.$$

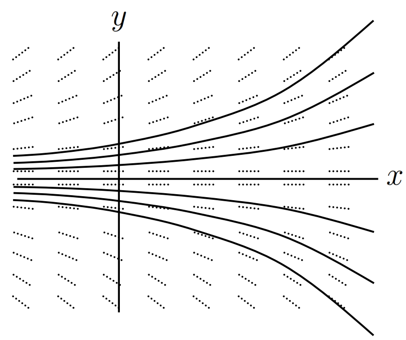
   Propose a function $g(x)$ that makes $f(x)$ differentiable for all $x > -\frac{\pi}{4}$.

D. Exponential functions and derivatives

1. We wish to find a function satisfying the differential equation

$$f'(x) = f(x). \tag{8}$$

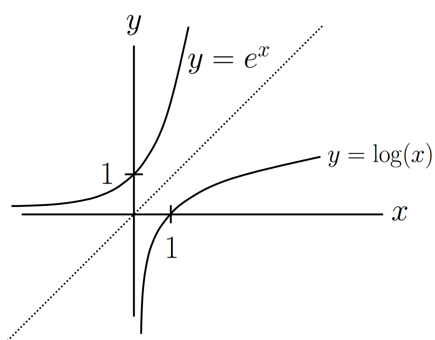
Suppose we have such a function. Then the slope of the tangent line to its graph at any particular y-value is equal to the y-value itself. For example, if the graph crosses the line $y = 1$, its tangent line at that point has slope 1. We can draw a collection of these tangent lines in a figure called a "direction field", and sketch a few possible functions which are "guided" by these tangent line segments.



We propose that the functions in the direction field are of the form a^x. Indeed, let $f(x) = a^x$. By the definition of derivative,

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

2. *Definitions.* Since we are in search of a function satisfying the differential equation (8) — that is, a function which is not merely proportional to its derivative, but equal to its derivative — we define e to be the number (around 2.718) such that the derivative of e^x evaluated at 0 is equal to 1. Then the $f(x) = e^x$ satisfies (8). We denote by $\log(x)$ (sometimes denoted $\ln(x)$ by non-mathematicians) the inverse of e^x.



(Note that, in making this definition, we make some leaps of logic — for example, how are we sure that such a number e exists? For more rigorous but less intuitive definitions of both e and $\log(x)$, see spiderwire.math.ubc.ca.)

3. *Derivatives of exponential functions.* By definition, $\frac{d}{dx}e^x = e^x$. To differentiate $\log(x)$, we set $y = \log(x)$, whence $e^y = x$, and differentiate implicitly: by the Chain Rule, $e^y y' = xy' = 1$; that is, $y' = \frac{1}{x}$.

E. Exercises

1. Come up with two functions which satisfy $f'(x) = 7f(x)$.
2. Differentiate the following functions:
   (a) $f(x) = \log\left(\frac{1}{x}\right)$.
   (b) $f(x) = x^{\cos(x)}$ at $x = \pi$.
   (c) $f(x) = \log(\log(\log(x)))$.
3. Titanium-44 has a half-life of 63 years. How long does it take for a sample to decay to one-third its original size?
4. In an ideal environment, a cell culture grows at a rate proportional to the number of cells present. Suppose a culture has 500 cells initially and 900 cells after 24 hours. How many cells will there be after an additional 10 hours?

The Mean Value Theorem and curve sketching

A. Motivation

The *Lennard-Jones potential* is a computationally simple model describing the potential energy of a diatomic molecule. In it, the energy is given by

$$V(r) = \varepsilon\left(\left(\frac{R}{r}\right)^{12} - 2\left(\frac{R}{r}\right)^{6}\right) \tag{9}$$

where r is the distance between the atoms, and R and ε are constants.

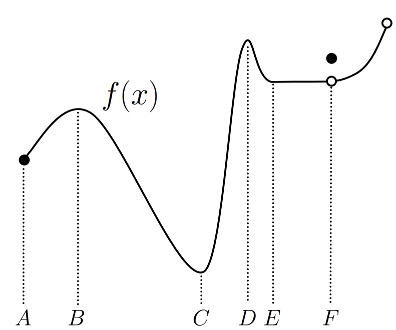
How might we describe this function?

We can use our physical intuition to make some guesses. For example, as the distance r increases from R, we expect $V(r)$ to increase as well. It would be useful to confirm this by describing the shape of the graph of $V(r)$ precisely: whether it has extrema, where it increases and where it decreases, what happens as r gets very large, and so on. It turns out that the derivative of $V(r)$ is very useful in getting this information.

B. Extrema and the Interior Extremum Theorem

1. Recall the Extreme Value Theorem: let $f(x)$ be continuous on $[l, r]$; then $f(x)$ has global extrema on $[l, r]$.

   Note that this theorem guarantees the existence of global extrema, but does not provide a way of locating them. However, we are now in a position to use derivatives to locate extrema.

   Recall the function $f(x)$ pictured above has a local maximum at B, a global and local minimum at C, a local maximum at D, a local minimum at E, a local maximum *and* local minimum at every point in (E, F), and a local maximum at F. In particular, the extrema are all located where the derivative vanishes or is undefined.

2. *Definition.* We say that $f(x)$ has a *critical point* at a if $f'(a) = 0$ or does not exist.

3. **The Interior Extremum Theorem.** If $f(x)$ has a local extremum at a, then $f(x)$ has a critical point at a.

   Proof. Suppose $f(x)$ has a local maximum at a. If $f(x)$ is not differentiable at a, we are done. If $f(x)$ is differentiable at a, then
$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$
   exists. Since $f(x)$ has a local maximum at a, the numerator is negative for sufficiently small values of h. The denominator takes the same sign as h. In particular, by considering $h < 0$, we have $f'(a) \geq 0$; and by considering $h > 0$, we have $f'(a) \leq 0$. Therefore we must have $f'(a) = 0$. If $f(x)$ has a local minimum at a, the proof is similar. $\square$

4. To find global extrema for continuous functions on closed intervals, we look at the function at the endpoints and at its critical points. The function *must* attain its global extrema at one or more of these points.

   To find local extrema for functions, we look at the function at its critical points. The function *may* attain its local extrema at one or more of these points. (To see that it may not, consider the function $f(x) = x^3$.)

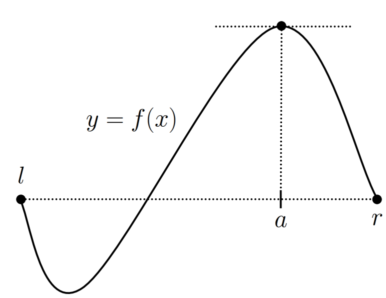
C. Monotonicity, concavity and the Mean Value Theorem

1. *Definition.* We say that $f(x)$ is *increasing* on an interval if for any pair of points $l < r$ in that interval, $f(l) < f(r)$. We say that $f(x)$ is *decreasing* on an interval if for any pair of points $l < r$ in that interval, $f(l) > f(r)$.

2. We may determine whether a function is increasing or decreasing on an interval by considering its derivative on that interval.

   To establish this connection, we appeal to a central result called the Mean Value Theorem. The Mean Value Theorem exemplifies the axiomatic structure of calculus in particular, and mathematics in general: it is a consequence of Rolle's Theorem, which is a consequence of the Extreme Value Theorem, which is a consequence of the fact that numbers are very "tightly packed", a condition called *completeness.* (This axiomatic structure is visualized on spiderwire.math.ubc.ca.)

3. **Rolle's Theorem.** Let $f(x)$ be continuous on $[l, r]$ and differentiable on (l, r), with $f(l) = f(r)$. Then there exists a number a in (l, r) such that $f'(a) = 0$.



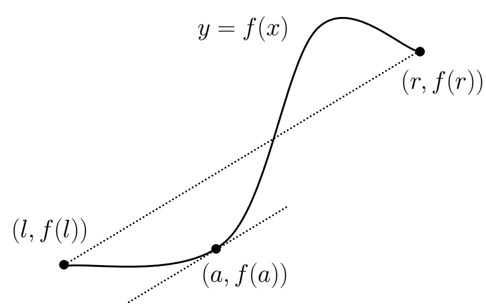
In the picture above, Rolle's theorem guarantees the existence (but says very little about the location) of the point a.

If $f(x)$ is a constant function, $f'(a) = 0$ for all numbers a in (l, r), and we are done.

If $f(x)$ is not a constant function, it must attain its global maximum or its global minimum (or both) in the interval (l, r). A global extremum in (l, r) must also be a local extremum, and by the Interior Extremum Theorem, the derivative must vanish there. $\square$

4. **The Mean Value Theorem.** Let $f(x)$ be continuous on $[l, r]$ and differentiable on (l, r). Then there exists a number a in (l, r) such that $f'(a) = \frac{f(r)-f(l)}{r-l}$.

   The picture below illustrates the conclusion of the theorem; namely, that there must be a number a in (l, r) such that $f'(a)$ is the average, or "mean", slope between the points $(l, f(l))$ and $(r, f(r))$.

   To prove the theorem, we appeal to Rolle's Theorem by considering the function

   $$g(x) = f(x) - \left(\frac{f(r) - f(l)}{r - l} (x - l) + f(l) \right),$$

   which describes the vertical distance between the $f(x)$ and the dotted "mean slope line". $g(x)$ satisfies the conditions of Rolle's Theorem, so we conclude that there exists a number a in (l, r) such that

   $$g'(a) = f'(a) - \frac{f(r) - f(l)}{r - l} = 0. \quad \square$$

5. We are now able to connect a function's derivative to whether it is increasing or decreasing. For example, suppose $f'(x) > 0$ on an interval. Then for any pair of points $l < r$ in that interval, by the Mean Value Theorem, there exists a number a in (l, r) such that

   $$\frac{f(l) - f(r)}{l - r} = f'(a) > 0;$$

   that is, $f(l) < f(r)$ — in other words, $f(x)$ is increasing on the interval.

6. *Definition.* If a differentiable function $f(x)$ has an increasing derivative on an interval, we say it is *concave up* on that interval. If it has a decreasing derivative on an interval, we say it is *concave down* on that interval. If $f(x)$ changes from concave up to concave down, or vice versa, at a, we say that $f(x)$ has an *inflection point* at a.

D. Exercises

1. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *Rolle's Theorem is a special case of the Mean Value Theorem.*
   (b) *Let $f(x)$ be defined on $[l, r]$ and differentiable on (l, r). Then there exists a number a in (l, r) such that $f'(a) = \frac{f(r)-f(l)}{r-l}$.*
   (c) *Let $f(x)$ be continuous on $[l, r]$. Then there exists a number a in (l, r) such that $f'(a) = \frac{f(l)-f(r)}{l-r}$.*
2. Give an example of a function which has the following properties.
   (a) $f(x)$ is defined everywhere.
   (b) $f(x)$ has a global maximum at $x = 1$, a local minimum (but not a global minimum) at $x = 2$, and no other extrema.
   (c) $f(x)$ is discontinuous at $x = 0$, $x = 1$ and $x = 2$, and continuous elsewhere.
3. Find all global extrema for the following functions in the given domains.
   (a) $f(x) = x^3 - 3x^2 - 9x - 3$ on $[-2, 1]$.
   (b) $f(x) = 2x - 3x^{2/3}$ on $[-1, 1]$.
   (c) $f(x) = |\sin(x)|$ in $[0, 10]$.
4. Let $f'(x) = 0$ on an interval. Prove using the Mean Value Theorem that $f(x)$ is constant on that interval.
5. Let $f'(x) < 0$ on an interval. Prove using the Mean Value Theorem that $f(x)$ is decreasing on that interval.
6. Let
$$f(x) = \begin{cases} x + 2x^2 \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}.$$
   Use the definition of derivative to show that $f'(0) = 1$, but then show that $f(x)$ is not increasing on any interval $(-\delta, \delta)$.
7. Give examples of functions with the following properties.
   (a) The function is concave up and increasing everywhere.
   (b) The function is concave down and increasing everywhere.
   (c) The function is concave up and decreasing everywhere.
   (d) The function is concave down and decreasing everywhere.
8. Suppose you wish to find the inflection points of a twice-differentiable function $f(x)$. Explain how you would do this. What assumptions are you making about the function?

E. Curve sketching and L'Hospital's Rule

1. Generally, we divide the curve sketching procedure into a few steps: we use the function to find the domain, intercepts and asymptotes; its derivative to find intervals of increase and decrease and extrema; and its second derivative to find intervals of concavity and inflection points.

2. *Example.* As an example, we return to the Lennard-Jones potential (9).

   Domain. $V(r)$ is defined on the intervals $(\infty, 0)$ and $(0, \infty)$, but its physical interpretation indicates that we should restrict the domain to be $(0, \infty)$.

Intercepts. There are no y-intercepts since $r = 0$ is not in the domain. To determine r-intercepts, we write

$$V(r) = \varepsilon \left(\frac{R^{12} - 2R^6 r^6}{r^{12}} \right)$$

and observe that $V(r) = 0$ when $r = \frac{1}{2^{1/6}} R$. Thus there is one r-intercept, $\left(\frac{1}{2^{1/6}} R, 0\right)$.

Asymptotes. To find asymptotes, we write

$$V(r) = \varepsilon \left(\frac{\frac{R^{12}}{r^6} - 2R^6}{r^6} \right),$$

and observe that $\lim_{r \to 0^+} V(r) = \infty$ and $\lim_{r \to \infty} V(r) = 0$. Thus $r = 0$ is a vertical asymptote and $y = 0$ is a horizontal asymptote.

Intervals of increase and decrease, and extrema. Next, we differentiate $V(r)$, getting

$$V'(r) = \varepsilon \left(-\frac{12R^{12}}{r^{13}} + \frac{12R^6}{r^7} \right) = -\frac{12\varepsilon R^6}{r^7} \left(\frac{R^6}{r^6} - 1 \right),$$

which vanishes when $r = R$. We determine the sign of $V'(r)$ on either side.

r	$(0, R)$	(R, ∞)
$V'(r)$	$-$	$+$
$V(r)$	decreasing	increasing

We conclude that $V(r)$ has a local minimum at $r = R$. In fact, since $V(r)$ is continuous, we may conclude that this is in fact a global minimum. $V(r)$ has no local nor global maxima.

Intervals of concavity and inflection points. Finally, we turn to the second derivative of $V(r)$. We have

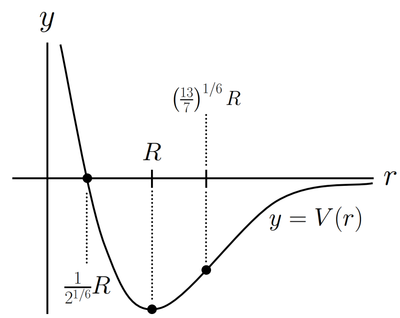
$$V''(r) = \varepsilon \left(\frac{12(13)R^{12}}{r^{14}} - \frac{12(7)R^6}{r^8} \right) = \frac{12\varepsilon R^6}{r^8} \left(\frac{13R^6}{r^6} - 7 \right),$$

which vanishes when $r = \left(\frac{13}{7}\right)^{1/6} R$. We determine the sign of $V''(r)$ on either side.

r	$\left(0, \left(\frac{13}{7}\right)^{1/6} R\right)$	$\left(\left(\frac{13}{7}\right)^{1/6} R, \infty\right)$
$V''(r)$	$+$	$-$
$V(r)$	concave up	concave down

We conclude that $V(r)$ has an inflection point at $r = \left(\frac{13}{7}\right)^{1/6} R$.

Summary. The information above may be summarized in the following sketch of the graph.



3. *Example.* In the Lennard-Jones potential example above, we were able to find the asymptotes by dividing the numerator and denominator of the function by another function r^6. We now consider an example where this kind of analysis is not possible.

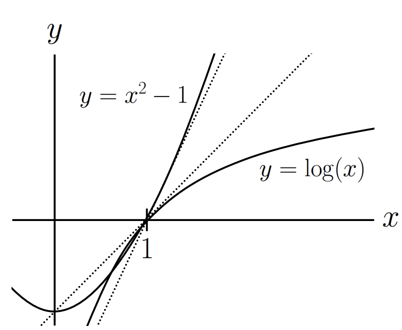
   In particular, suppose we wish to find whether the graph of $f(x) = \frac{\log(x)}{x^2-1}$ has any asymptotes.

   We have three cases to consider: a possible vertical asymptote $x = 0$ (where the numerator gets very large), a possible vertical asymptote $x = 1$ (where the denominator gets very small), and a possible horizontal asymptote as x gets very large.

   We can dispense with the first case easily. As x approaches 0 and is positive, the denominator approaches -1 while the numerator approaches $-\infty$. Thus $\lim_{x\to 0^+} f(x) = \infty$ and $x = 0$ is a vertical asymptote.

   In the second case, the numerator and denominator both approach 0. The question is: how "quickly" do they approach 0? For example, $\lim_{x\to 0} \frac{\sin(x)}{x} = 1$ implies that, as x approaches 0, $\sin(x)$ and x approach 0 at the same rate. On the other hand, $\lim_{x\to 0} \frac{2x^2}{x^2} = 2$ implies that, as x approaches 0, $2x^2$ approaches 0 at a rate twice as large as that of x^2. Both examples suggest that it makes sense to consider the *rate* at which the numerator and denominator approach 0.

   We propose that
$$\lim_{x\to 1} \frac{\log(x)}{x^2-1} = \lim_{x\to 1} \frac{\frac{d}{dx}\log(x)}{\frac{d}{dx}(x^2-1)} = \lim_{x\to 1} \frac{\frac{1}{x}}{2x} = \frac{1}{2}.$$



A picture supports this. Near $x = 1$, we can approximate $\log(x)$ and $x^2 - 1$ using their tangent lines at $x = 1$.

Similarly, in the third case, as x approaches ∞, both the numerator and the denominator approach ∞. Again, the question is: how "quickly" do they approach ∞? We propose that

$$\lim_{x\to\infty} \frac{\log(x)}{x^2-1} = \lim_{x\to\infty} \frac{\frac{d}{dx}\log(x)}{\frac{d}{dx}(x^2-1)} = \lim_{x\to\infty} \frac{\frac{1}{x}}{2x} = 0.$$

These results follow by L'Hospital's Rule.

4. **L'Hospital's Rule.** Let $f(x)$ and $g(x)$ be differentiable on an interval containing a, and let $g'(x) \neq 0$ on that interval (except possibly at a). Let $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$. Then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)},$$

provided the limit on the right-hand side exists or is ∞ or $-\infty$.

This statement of L'Hospital's Rule also holds if a is replaced by ∞ or $-\infty$, or if, in place of the second sentence, we have $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \infty$ or $-\infty$.

The proof of L'Hospital's Rule is rather technical, and we omit it here. (It may be found on spiderwire.math.ubc.ca.)

F. Exercises

1. Evaluate the following limits.
   (a) $\lim_{x\to\infty} \frac{x^n}{e^x}$, where n is a positive integer.
   (b) $\lim_{x\to 0} \frac{\cos(x) - \cos(2x)}{e^x - x - 1}$.
   (c) $\lim_{x\to 0^+} x^x$.
2. Explain why L'Hospital's Rule cannot be used to evaluate $\lim_{x\to 0} \frac{\sin(x)}{x}$.
3. Determine where the following functions are increasing, and where they are decreasing.
   (a) $f(x) = x^3 + 4x + 3$.
   (b) $f(x) = \frac{1}{1+x^2}$.
   (c) $f(x) = x + \sin(x)$.
4. Determine where the following functions are concave up, and where they are concave down.
   (a) $f(x) = x^2 + 2x + 1$.
   (b) $f(x) = \frac{\log(x^2)}{x}$.
   (c) $f(x) = xe^x$.
5. Sketch the graphs of the following functions.

(a) $f(x) = \dfrac{x^2 - 4}{x^2 - 1}$.

(b) $f(x) = xe^{-x^2}$.

(c) $f(x) = x + \sin(x)$ on $[0, 2\pi]$.

6. Sketch the graphs of the following functions.

   (a) $f(x) = (x^2 - 1)^{2/3}$.

   (b) $f(x) = x \log |x|$.

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

7. Sketch the graph of the curve $y^2 = x^3 - x$.

8. Give an example of a function $f(x)$ which has the following properties.

   (a) $f(1) = 0$.

   (b) $f(0) = 1$.

   (c) $f''(x) < 0$ for $x < 0$.

   (d) $f'(x) < 0$ for $x > 0$.

   (e) $\lim_{x \to 0^+} f(x) = \infty$.

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

Optimization

A. Motivation

Recall the Lennard-Jones potential describing the potential energy of a diatomic molecule:

$$V(r) = \varepsilon\left(\left(\frac{R}{r}\right)^{12} - 2\left(\frac{R}{r}\right)^{6}\right)$$

where r is the distance between the atoms, and R and ε are constants. What is the distance at which the potential energy is minimal?

This is an example of an optimization problem. Optimization problems are an application of curve sketching, which is itself an application of the Mean Value Theorem. It is appropriate that at the end of this course, we find ourselves situated at the end of long chain of implications stretching back to the notion of limit.

B. Optimization

1. *Example.* We begin with the question given above about the Lennard-Jones potential. The calculation here is essentially duplicated from the calculation required to sketch the graph of the function.

   We wish to minimize $V(r)$ on the domain $(0, \infty)$. We have

   $$V'(r) = \varepsilon\left(-\frac{12R^{12}}{r^{13}} + \frac{12R^6}{r^7}\right) = -\frac{12\varepsilon R^6}{r^7}\left(\frac{R^6}{r^6} - 1\right),$$

   which vanishes when $r = R$. We determine the sign of $V'(r)$ on either side.

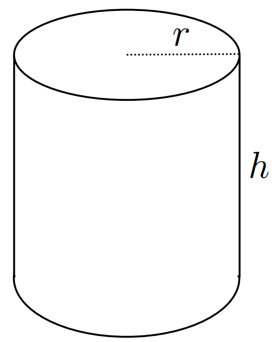
r	$(0, R)$	(R, ∞)
$V'(r)$	$-$	$+$
$V(r)$	decreasing	increasing

   We conclude that $V(r)$ has a local minimum at $r = R$. In fact, since $V(r)$ is continuous, we may conclude that this is in fact a global minimum. The potential energy of the molecule is minimized at $r = R$.

2. *Example.* In many optimization problems, the challenge is not to locate the extrema of a given function, but to establish what the function is in the first place. We now consider an example of such a problem.

   Suppose we wish to determine the most economical shape of a tin can. We shall frame the question as follows: *Given a fixed volume, what is the minimum surface area of a cylindrical can of that volume?*

   Let V, r and h denote the volume, radius and height of the can, respectively.



We wish to minimize $S = 2\pi r^2 + 2\pi rh$. We can use the formula $V = \pi r^2 h$ to reduce this to a function of one variable: $S(r) = 2\pi r^2 + \frac{2V}{r}$. We wish to minimize this function on the domain $(0, \infty)$. We have

$$S'(r) = 4\pi r - \frac{2V}{r^2} = \frac{2}{r^2}\left(2\pi r^3 - V\right),$$

which vanishes when $r = \left(\frac{V}{2\pi}\right)^{1/3}$. We determine the sign of $S'(r)$ on either side.

r	$\left(0, \left(\frac{V}{2\pi}\right)^{1/3}\right)$	$\left(\left(\frac{V}{2\pi}\right)^{1/3}, \infty\right)$
$S'(r)$	$-$	$+$
$S(r)$	decreasing	increasing

We conclude that $S(r)$ has a local minimum at $r = \left(\frac{V}{2\pi}\right)^{1/3}$. In fact, since $S(r)$ is continuous on its domain, this is a global minimum. The most economical shape of a can is one where the radius r is given by $r = \left(\frac{V}{2\pi}\right)^{1/3}$, or equivalently $r = \frac{h}{2}$.

Is this solution correct? This question can be interpreted at least two different ways. First, is the mathematics correct? An effective check is to solve the dual problem: *Given a fixed surface area, what is the maximum volume of a cylindrical can of that surface area?*

Second, is the conclusion plausible? The answer here is more subtle. Most cans are not as tall as they are wide. Why not? One reason is that, while can walls are rectangles which can be stamped from sheet metal with very little wasted material, can ends are circles which are stamped from sheet metal with some waste. Taking the cost of wasted sheet metal into consideration changes the conclusion.

C. Exercises

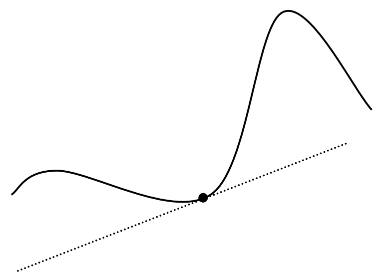
1. (a) Given a fixed surface area, what is the maximum volume of a cylindrical can of that surface area?
   (b) Explain why the following two problems are equivalent: (1) *Given a fixed volume, what is the minimum surface area of a cylindrical can of that volume?* and (2) *Given a fixed surface area, what is the maximum volume of a cylindrical can of that surface area?*
   (c) Given a fixed volume, what is the minimum surface area of a cylindrical can of that volume, assuming that the ends are circles which are stamped from squares of sheet metal?
2. Find two numbers whose difference is 10 and whose product is minimal.
3. What is the shortest distance between the curve $y - x^{3/2} = 1$ and the point $(8, 1)$?

4. (a) Of all rectangles of a given perimeter P, determine the dimensions of the rectangle of largest area.
   (b) Of all isosceles triangles of a given perimeter P, determine the dimensions of the triangle of largest area.
5. What is the area of the largest rectangle (with sides parallel to the axes) which may be inscribed in the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$?
6. Suppose a spherical cell has a rate of nutrient absorption which is proportional to its surface area, and a rate of nutrient consumption which is proportional to its volume. Let the constants of proportionality be a and b, respectively. What is the optimum radius of the cell?
7. Come up with a differentiable function whose graph has a tangent line of maximal slope at $x = 3$.
8. Suppose you wish to connect four points at the corners of a square. What is the total length of the shortest path? (The path may have branches.)

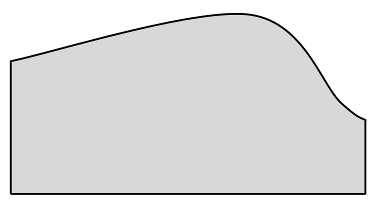
Riemann sums and integrability

A. Motivation

Recall that calculus is motivated by two problems: the *tangent line problem* (given a curve, how do we describe the tangent line at a particular point on the curve?)



and the *area problem* (given a shape, how do we find its area?).

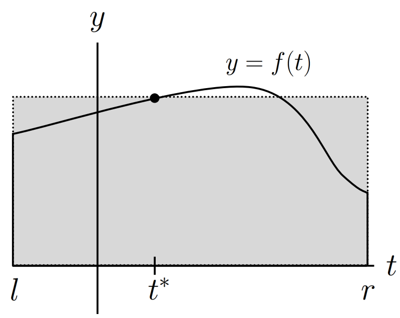


Our subject so far has been the tangent line problem, and the tools — in particular the derivative — needed to solve it. The subject of this course is the area problem.

As we shall see later, these two problems are intimately related. For now, we make the observation that they are related through the use of the limit.

Our first step is as it was in our efforts to solve the tangent line problem. There, we began by defining what we meant by "tangent line". Here, we begin by defining what we mean by "area".

We generally restrict ourselves to areas under graphs of functions of one variable. Our first approximation is to estimate the area A under the curve $y = f(t)$ on the interval $[l, r]$ to be equal to the area of the rectangle of width $r - l$ and height $f(t^*)$, where t^* is a "representative height".

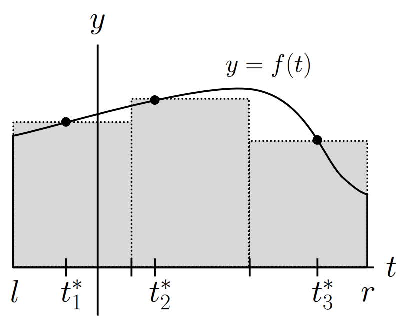


Thus

$$A \approx f\left(t^*\right)(r - l).$$

One benefit of doing this is that we use the “natural” vocabulary of areas of rectangles.

We can improve our approximation by partitioning $[l, r]$ into multiple subintervals of equal width, and selecting representative heights from each subinterval.



For example, with three subintervals, we would have

$$A \approx \sum_{i=1}^{3} f\left(t_i^*\right)\left(\frac{r-l}{3}\right).$$

As we refine our partition into more and more subintervals, we expect our approximation to get closer and closer to an “actual” area A. If it does, we define the limit to be the integral

$$A = \int_l^r f(t)dt.$$

B. The definition of integral and integrability

1. *Definition.* Let $f(t)$ be defined on $[l, r]$. We may partition the interval into n subintervals

$$l = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = r$$

of equal width Δt, and select

$$t_1^* \in [t_0, t_1], t_2^* \in [t_1, t_2], \ldots, t_n^* \in [t_{n-1}, t_n]$$

to be *sample points* in these subintervals. We say that $f(t)$ is *integrable* on $[l, r]$ if

$$\lim_{n\to\infty} \sum_{i=1}^{n} f\left(t_i^*\right) \Delta t$$

exists and is equal for all choices of sample points. The sum is called a *Riemann sum* (and in fact, "integrable" is shorthand for "Riemann integrable"; there are other kinds). If so, we use the notation

$$\int_l^r f(t) dt$$

to denote the limit, which we call the *integral of $f(t)$ from l to r*. If the limit does not exist, or is not equal for all choices of sample points, we say that $f(t)$ is not integrable on $[l, r]$.

2. *Example.* $\displaystyle\int_1^2 t\, dt = \frac{3}{2}$.

We partition $[1, 2]$ into n subintervals of width $\frac{1}{n}$. We wish to determine whether the limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} t_i^*$$

exists, where t_i^* is a sample point in the i^{th} subinterval $\left[1 + \frac{i-1}{n}, 1 + \frac{i}{n}\right]$. Since the function is increasing on every subinterval, we have $1 + \frac{i-1}{n} \leq t_i^* \leq 1 + \frac{i}{n}$ and

$$1 + \frac{1}{n^2} \sum_{i=1}^{n-1} i = \frac{1}{n} \sum_{i=1}^{n} \left(1 + \frac{i-1}{n}\right) \leq \frac{1}{n} \sum_{i=1}^{n} t_i^* \leq \frac{1}{n} \sum_{i=1}^{n} \left(1 + \frac{i}{n}\right) = 1 + \frac{1}{n^2} \sum_{i=1}^{n} i.$$

Using the fact that $1 + 2 + \cdots + k = \frac{k(k+1)}{2}$, we get the bounds

$$1 + \frac{n-1}{2n} \leq \frac{1}{n} \sum_{i=1}^{n} t_i^* \leq 1 + \frac{n+1}{2n}.$$

Taking the limit as $n \to \infty$ yields

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} t_i^* = \frac{3}{2}.$$

Note that, if the integral is indeed a reasonable description of the area under $y = t$ on $[1, 2]$, it ought to produce the same answer which may be recovered easily using classical geometry. This example helps confirm the plausibility of the description.

3. *Example.* $f(t) = \begin{cases} \frac{1}{t} & \text{if } t \neq 0 \\ 0 & \text{if } t = 0 \end{cases}$ is not integrable on $[0, r]$, where $r > 0$.

We partition $[0, r]$ into n subintervals of width $\frac{r}{n}$. We wish to determine whether the limit

$$\lim_{n\to\infty} \frac{r}{n} \sum_{i=1}^{n} \frac{1}{t_i^*} \tag{10}$$

exists, where t_i^* is a sample point in the i^{th} subinterval $\left[\frac{(i-1)r}{n}, \frac{ir}{n}\right]$. Now

$$\frac{r}{n}\sum_{i=2}^{n}\frac{1}{t_i^*} \geq \frac{r}{n}\sum_{i=2}^{n}\frac{n}{ir} = \sum_{i=2}^{n}\frac{1}{i}.$$

As $n \to \infty$, the sum on the right-hand side diverges, since the harmonic series diverges. Thus the limit (10) does not exist, and $\frac{1}{x}$ is not integrable on $[0, r]$.

Note that the existence of a vertical asymptote on an interval does not in itself guarantee that a function is not integrable on that interval. In fact, as question 1(c) below may suggest, and as we shall see next week, it turns out t^{-p} is integrable on $[0, r]$ for $r > 0$ if, and only if, $p < 1$.

4. *Integrability and continuity.* Which functions are integrable? It turns out that all continuous functions are. (The proof of this fact is available on spiderwire.math.ubc.ca. The idea of the proof is that continuity guarantees that "minor" refinements of a partition can change a Riemann sum by only a limited amount.) Of course discontinuous functions may also be integrable, as shown in the exercises below.

   The integrability of continuous functions guarantees the integrability of differentiable functions. In next week's topic, we explore a surprising and fundamental connection between integrability and the *antidifferentiability* of functions.

C. Exercises

1. Write down definitions of the following statements.
   (a) $\int_2^{-3} f(t)dt$.
   (b) $\int_l^{\infty} f(t)dt$.
   (c) $\lim_{l\to 0^+}\int_l^1 \frac{1}{\sqrt{t}}\,dt$.
2. Determine whether each of the following statements is true or false, and justify your answer.
   (a) $\int_l^r f(t)dt = -\int_r^l f(t)dt$ *if* $f(t)$ *is integrable on* $[l, r]$.
   (b) *If* $f(t)$ *is an odd function — that is,* $f(t) = -f(-t)$ *for all* t *— then* $\int_{-a}^{a} f(t)dt = 0$.
   (c) *If* $f(t)$ *and* $g(t)$ *are integrable and satisfy* $f(t) \leq g(t)$ *on* $[l, r]$*, then* $\int_l^r f(t)dt \leq \int_l^r g(t)dt$.
3. Prove that constant functions are integrable over any closed interval.
4. Prove that $f(t) = t^2$ is integrable over $[0, 1]$.
5. Let $f(t) = \begin{cases} 1 & \text{if } t = 0 \\ 0 & \text{if } t \neq 0 \end{cases}$. Prove that $f(t)$ is integrable over $[-a, a]$ for $a > 0$.
6. Let

   $$f(t) = \begin{cases} 1 & \text{if } t \text{ is rational} \\ 0 & \text{if } t \text{ is irrational} \end{cases}.$$

   Prove that $f(t)$ is not integrable over any closed interval.

7. Prove that $f(t) = \frac{1}{t^2}$ is integrable over $[1, \infty)$.
8. Let

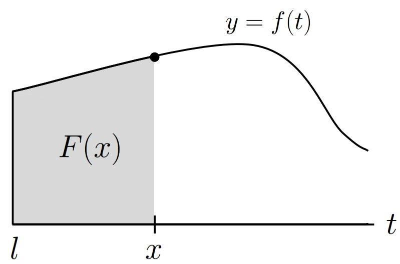
$$f(t) = \begin{cases} 1 & \text{if } t = \frac{1}{e}, \frac{1}{e^2}, \frac{1}{e^3}, \dots \\ 0 & \text{otherwise} \end{cases}.$$

   Prove that $f(t)$ is integrable over $[0, 1]$. (Hint: suppose there are n subintervals; how many can possibly include points of the form $\frac{1}{e^k}$ where k is a positive integer?)

The Fundamental Theorem of Calculus

A. Motivation

Let $f(t)$ be a continuous function, and $F(x) = \int_l^x f(t)dt$. What is $F'(x)$?



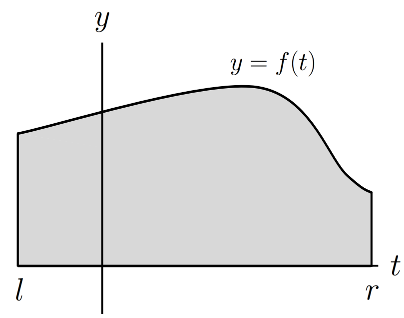
We frame this question in two ways. First: what is the rate of change of the area under the curve as the right endpoint of our interval increases? As x moves to the right, we expect the area $F(x)$ to increase more rapidly, because the function $f(x)$ gets bigger. Afterward, as we pass through the interval where $f(x)$ is decreasing, we expect $F(x)$ to increase since $f(x)$ is positive, but at a slower rate since $f(x)$ is decreasing. In other words, we expect $F'(x)$ to be proportional to $f(x)$.

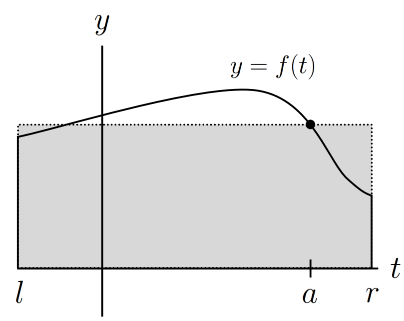
Second, we may also think of $f(t)$ as speed, and $F(x)$ as the distance travelled between times $t = l$ and $t = x$: what is the rate of change of the distance travelled over time? As x moves to the right, we expect the distance to increase more rapidly, because our speed increases. Afterward, we expect the distance to continue to increase, but at a slower rate since our speed has dropped. Again, we expect $F'(x)$ to be proportional to $f(x)$.

It turns out that $F'(x) = f(x)$. This is part of a central result known as the Fundamental Theorem of Calculus. This result allows us to calculate many integrals more efficiently, by establishing a startling connection between integrals and antiderivatives of functions.

B. The Fundamental Theorem of Calculus

1. *The Mean Value Theorem for Integrals.* We require a preliminary result to prove the Fundamental Theorem of Calculus. The Mean Value Theorem for Integrals states that there exists a number a such that the two shaded areas are identical.





This is plausible when framed in terms of speed and distance: the theorem states that at some point, our speed is exactly "average", in that travelling at that speed for the entire time would cover the same distance.

The Mean Value Theorem for Integrals. Let $f(t)$ be continuous on $[l, r]$. Then there exists a number a in $[l, r]$ such that

$$f(a)(r-l) = \int_l^r f(t)dt.$$

Proof. By the Extreme Value Theorem, $f(t)$ attains a global maximum $f(t_1) = U$ and global minimum $f(t_2) = L$ on $[l, r]$, from which we conclude that

$$L(r-l) \leq \int_l^r f(t)dt \leq U(r-l);$$

or equivalently,

$$L \leq \frac{1}{r-l}\int_l^r f(t)dt \leq U.$$

By the Intermediate Value Theorem, there exists a number a between t_1 and t_2 such that

$$f(a) = \frac{1}{r-l}\int_l^r f(t)dt. \quad \square$$

2. **The Fundamental Theorem of Calculus.** Let $f(t)$ be continuous on an interval I containing the

point l, and let

$$F(x) = \int_l^x f(t)dt.$$

Then $F(x)$ is differentiable on I, with $F'(x) = f(x)$. Furthermore, if $G(x)$ is any antiderivative of $f(x)$, then

$$\int_l^r f(t)dt = G(r) - G(l)$$

for all r in I.

Proof. By the definition of derivative,

$$F'(x) = \lim_{h \to 0} \frac{1}{h} \left(\int_l^{x+h} f(t)dt - \int_l^x f(t)dt \right) = \lim_{h \to 0} \frac{1}{h} \int_x^{x+h} f(t)dt.$$

By the Mean Value Theorem for Integrals, the integral is equal to $hf(a)$ for some a between x and $x+h$. Finally, we have

$$F'(x) = \lim_{h \to 0} f(a) = f(x)$$

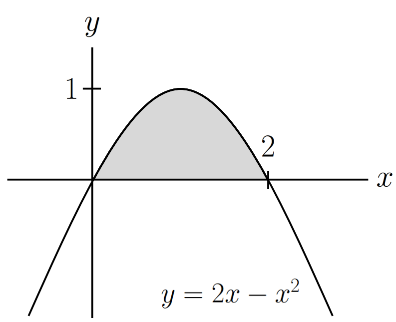
by continuity.

To prove the remainder of the Fundamental Theorem of Calculus, let $G(x)$ be any antiderivative of $f(x)$. We have just shown that $F(x)$ is also an antiderivative. It follows that $F(x) - G(x) = c$ for some constant c. In particular, we have

$$F(x) = \int_l^x f(t)dt = G(x) + c.$$

Setting $x = r$ yields $c = -G(l)$, and then setting $x = r$ then yields the last part of the theorem. □

3. *Example.* We wish to find the area above the x-axis and under the curve $y = 2x - x^2$.



This is equal to $\int_0^2 (2x - x^2)dx$. Now $x^2 - \frac{1}{3}x^3$ is an antiderivative of $2x - x^2$, so by the Fundamental Theorem of Calculus,

$$\int_0^2 (2x - x^2)dx = x^2 - \frac{1}{3}x^3 \Big|_0^2 = \left(2^2 - \frac{1}{3}2^3\right) - \left(0^2 - \frac{1}{3}0^3\right) = \frac{4}{3}.$$

C. Exercises

1. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *If $F(x)$ is an antiderivative of $f(x)$ and $G(x)$ is an antiderivative of $g(x)$, then $F(x)G(x)$ is an antiderivative of $f(x)g(x)$.*
   (b) *If $f(x)$ is continuous on $[l,r]$, then $\frac{d}{dx}\int_l^r f(x)dx = f(x)$.*
   (c) *All continuous functions have antiderivatives.*
2. Differentiate the following functions.
   (a) $\int_1^x e^{-t^2} dt$.
   (b) $\sin(x)\int_{x^2}^1 e^{-t^2} dt$.
   (c) $\sin(x)\int_{x^2}^{x^3} e^{-t^2} dt$.
3. Evaluate the following integrals.
   (a) $\int_0^b t^2 dt$.
   (b) $\int_{-2}^1 \left(t^3+3t-2\right) dt$.
   (c) $\int_0^\pi \sin(t) dt$.
4. (a) Evaluate $\int_0^\infty e^{-t} dt$.
   (b) Estimate $\int_0^\infty e^{-t^2} dt$ using the answer above.
5. (a) Prove that $\int_0^r t^{-p} dt$, where $r>0$, converges if, and only if, $p<1$.
   (b) Prove that $\int_l^\infty t^{-p} dt$, where $l>0$, converges if, and only if, $p>1$.
6. Solve the integral equation $f(x) = 1+4\int_3^x f(t)dt$. (Hint: differentiate the equation, and then solve the resulting differential equation.)
7. Find the global extrema of $f(x) = \int_0^{2x-x^2} \cos\left(\frac{1}{1+t^2}\right) dt$, if any exist.
8. Find the area of the region below $y = \frac{10}{x^2+1}$ and above $y=1$.

Techniques of integration

A. Motivation

The Fundamental Theorem of Calculus provides a startling way to calculate integrals by evaluating antiderivatives. To use the theorem effectively, it therefore makes sense to summarize techniques of determining antiderivatives. This roughly parallels the process in the first term of summarizing shortcuts which allow us to calculate derivatives without resorting directly to the definition of derivative.

Note that antidifferentiation is a means to integration; *they are not identical*, despite the unfortunate notation of "$\int f(t)dt$" sometimes being used to denote the antiderivative of $f(t)$.

We consider three techniques. Two of them — integration by substitution and integration by parts — "reverse" techniques of differentiation — the Chain Rule and the Product Rule, respectively. The third — integration by partial fractions — is an apparently simple technique which turns out to rely on a powerful theorem known as the Euclidean algorithm for polynomials.

B. Integration by substitution

1. *Example.* We wish to evaluate $\displaystyle\int_2^5 \frac{\cos\left(\sqrt{t-1}\right)}{\sqrt{t-1}}dt$.

   One approach is to find an antiderivative of

   $$f(t) = \frac{\cos\left(\sqrt{t-1}\right)}{\sqrt{t-1}}.$$

   We simplify this function using a change of variables. Set $u = \sqrt{t-1}$. Now it is true that

   $$f(t) = \frac{\cos(u)}{u}.$$

   However, our goal is to find an antiderivative; and if our antiderivative is in the variable u, then differentiating that antiderivative will trigger the Chain Rule, and we will end up with an extra term u'. The trick is to write $f(t)$ in terms of both u *and* u'. Instead of u', we use the troublesome but convenient notation

   $$du = \frac{1}{2\sqrt{t-1}}dt.$$

   (There is a formal justification for this notation using differentials.) Then

   $$f(t)dt = 2\cos(u)du.$$

   Now the antiderivative of $2\cos(u)$ is straightforward to calculate. However,

   $$\int_2^5 f(t)dt \neq \int_2^5 2\cos(u)du,$$

since changing the variable from t to u changes the interval over which we are integrating. Instead, we observe that the t-interval from 2 to 5 is equivalent to the u-interval from 1 to 2, and that

$$\int_2^5 f(t)dt = \int_1^2 2\cos(u)du = 2\sin(u)|_1^2 = 2\sin(2) - 2\sin(1).$$

(Procedurally, it is also worthwhile to check that $\sin(u)$ is indeed an antiderivative of $f(t)$.)

2. **The method of substitution.** Suppose $g(t)$ is differentiable on $[l, r]$, and suppose $f(t)$ is continuous on the range of $g(t)$. Then

$$\int_l^r f(g(t))\, g'(t)dt = \int_{g(l)}^{g(r)} f(u)du.$$

Proof. Let $F(t)$ be an antiderivative of $f(t)$. Then the derivative of $F(g(t))$ is

$$\frac{d}{dt}F(g(t)) = F'(g(t))\, g'(t) = f(g(t))\, g'(t),$$

which by the Fundamental Theorem of Calculus implies

$$\int_l^r f(g(t))\, g'(t)dt = F(g(r)) - F(g(l)) = \int_{g(r)}^{g(l)} f(u)du. \quad \square$$

3. *Example.* We wish to evaluate $\displaystyle\int_1^e \frac{\sin\left(\log(t^3)\right)}{t}\, dt.$

We make the substitution $u = \log(t^3) = 3\log(t)$, whence $du = \frac{3}{t}$ and

$$\int_1^e \frac{\sin\left(\log(t^3)\right)}{t}\, dt = \frac{1}{3}\int_0^3 \sin(u)\, du = \frac{1}{3}\left(-\cos(u)\right)\Big|_0^3 = \frac{1}{3}\left(1 - \cos(3)\right).$$

C. Exercises

1. Evaluate the following integrals.
   (a) $\displaystyle\int_0^2 \frac{t}{t^2+3}\, dt.$
   (b) $\displaystyle\int_{-5}^5 e^t\sqrt{e^t+1}\, dt.$
   (c) $\displaystyle\int_0^{\pi/6} \tan(t)dt.$
2. Evaluate the following integrals.
   (a) $\displaystyle\int_{-1}^1 \frac{2e^t - 2e^{-t}}{e^t + e^{-t}}\, dt.$
   (b) $\displaystyle\int_5^8 \frac{2t}{\sqrt{t-4}}\, dt.$

(c) $\displaystyle\int_2^3 \frac{1}{t^2\cos^2\left(\frac{1}{t}\right)}\,dt.$

3. Prove that $\displaystyle\int_l^r f(t+a)dt = \int_{l+a}^{r+a} f(t)dt$, and draw a picture illustrating the proof.

4. Evaluate the following integrals. (Hint: it is helpful to prove the identities $\cos^2(\theta) = \frac{1}{2}(1+\cos(2\theta))$ and $\sin^2(\theta) = \frac{1}{2}(1-\cos(2\theta))$, possibly by using the addition formulas to get expressions for $\cos(2\theta)$.)

   (a) $\displaystyle\int_0^{\pi/2} \sin(t)\cos^4(t)dt.$

   (b) $\displaystyle\int_0^{\pi/2} \sin^2(t)\cos^2(t)dt.$

   (c) $\displaystyle\int_0^{\pi/2} \cos^4(t)dt.$

5. In the previous exercise, you apply the method of substitution in a particular way, to integrals known as *trigonometric integrals*. Summarize your strategy for each of the following types of integrals.

   (a) $\displaystyle\int_l^r \sin^m(t)\cos^n(t)dt$, where m is odd.

   (b) $\displaystyle\int_l^r \sin^m(t)\cos^n(t)dt$, where n is odd.

   (c) $\displaystyle\int_l^r \sin^m(t)\cos^n(t)dt$, where both m and n are even.

6. Find the area of the ellipse described by $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

7. Evaluate the following integrals. (Hint: it is helpful to prove the identity $\tan^2(\theta) + 1 = \sec^2(\theta)$.)

   (a) $\displaystyle\int_1^2 \frac{1}{t^2\sqrt{t^2+1}}\,dt.$

   (b) $\displaystyle\int_0^1 \frac{t^3}{(4t^2+9)^{3/2}}\,dt.$

   (c) $\displaystyle\int_1^4 \sqrt{t^2-1}\,dt.$

8. In the previous two exercises, you apply the method of substitution in a particular way, sometimes known as *trigonometric substitution*. Summarize your strategy for each of the following types of integrals.

   (a) $\displaystyle\int_l^r f(t)dt$ where $f(t)$ contains the expression $\sqrt{a^2-t^2}$.

   (b) $\displaystyle\int_l^r f(t)dt$ where $f(t)$ contains the expression $\sqrt{a^2+t^2}$.

   (c) $\displaystyle\int_l^r f(t)dt$ where $f(t)$ contains the expression $\sqrt{t^2-a^2}$.

D. Integration by parts

1. *Example.* We wish to evaluate $\displaystyle\int_0^1 te^t dt$.

One advantage of this function is that its derivative contains a "copy" of itself: $\frac{d}{dt}\left(te^t\right) = e^t + te^t$; that is,

$$te^t = \frac{d}{dt}\left(te^t\right) - e^t.$$

The right-hand side is straightforward to antidifferentiate: an antiderivative of e^t is e^t, and even more simply, an antiderivative of $\frac{d}{dt}\left(te^t\right)$ is te^t. Thus

$$\int_0^1 te^t dt = \int_0^1 \left(\frac{d}{dt}te^t - e^t\right) dt = te^t\Big|_0^1 - \int_0^1 e^t dt = te^t\Big|_0^1 - e^t\Big|_0^1 = 1.$$

2. **The method of parts.** The method of parts is simply the process of "reversing the Product Rule". It can always be applied; the challenge is to apply it in a way that makes the integral more vulnerable to other techniques.

   The following mnemonic is a useful way to remember the method. By the Product Rule,

$$u(t)v'(t) = \frac{d}{dt}\left(u(t)v(t)\right) - u'(t)v(t).$$

   In the method of parts, we identify the integrand as $u(t)v'(t)$, replace it with the right-hand side, and then use the Fundamental Theorem of Calculus to evaluate the integral using antiderivatives. The mnemonic is as follows:

$$\int udv = uv - \int vdu.$$

3. *Example.* We wish to evaluate $\displaystyle\int_{\pi/4}^{\pi/3} t\cos(t)dt$.

   Let $u = t$ and $dv = \cos(t)dt$, whence $v = \sin(t)$, $du = dt$, and

$$\int_{\pi/4}^{\pi/3} t\cos(t)dt = t\sin(t)|_{\pi/4}^{\pi/3} - \int_{\pi/4}^{\pi/3} \sin(t)dt = t\sin(t) + \cos(t)|_{\pi/4}^{\pi/3} = \frac{\pi}{2\sqrt{3}} + \frac{1}{2} - \frac{\pi}{4\sqrt{2}} - \frac{1}{\sqrt{2}}.$$

4. Our aim when using the technique of parts is to replace our integrand with something which is easier to antidifferentiate. This means that we should generally pick "u" to be something that is "simpler" when differentiated, and "dv" to be something that is easy to antidifferentiate.

E. Exercises

1. Evaluate the following integrals.
   (a) $\displaystyle\int_1^e \frac{\log(t)}{\sqrt{t}}\,dt$.
   (b) $\displaystyle\int_1^e \log(t)dt$.
   (c) $\displaystyle\int_1^e \log(t)^2 dt$.

2. Evaluate the following integrals.
   (a) $\displaystyle\int_0^{\pi} e^t \cos(t)dt$.
   (b) $\displaystyle\int_0^{\pi} \cos(\log(t))dt$.
   (c) $\displaystyle\int_{-1}^{0} t^2 e^{-t} dt$.
3. (a) Let $f(t)$ be *continuously differentiable*; that is, let $f'(t)$ be differentiable. Prove that

$$\int_a^b f(t)f'(t)dt = \frac{1}{2}\left(f(b)^2 - f(a)^2\right).$$

   (b) Confirm the formula established in part (a) with a simple continuously differentiable function.
4. The integral

$$I_n = \int_0^1 t^n e^t dt$$

   may be calculated using the method of parts repeatedly. An alternate approach is to establish a recursive formula for integrals of this form called a *reduction formula*.
   (a) Prove the reduction formula $I_n = e - nI_{n-1}$.
   (b) Find I_4 by calculating I_0 directly and then using the reduction formula established in part (a).

F. Integration by partial fractions

1. *Example.* We wish to evaluate $\displaystyle\int_0^{\pi/4} \sec(t)dt$.

   This can be accomplished by substitution:

$$\int_0^{\pi/4} \sec(t)dt = \int_0^{\pi/4} \frac{\sec(t)\left(\sec(t)+\tan(t)\right)}{\sec(t)+\tan(t)}\, dt;$$

   then let $u = \sec(t) + \tan(t)$, whence $du = \left(\sec(t)\tan(t) + \sec^2(t)\right) dt$ and

$$\int_0^{\pi/4} \sec(t)dt = \int_1^{1+\sqrt{2}} \frac{1}{u}\, du = \log(1+\sqrt{2}).$$

   However, a less *ad hoc* approach is to observe that

$$\int_0^{\pi/4} \sec(t)dt = \int_0^{\pi/4} \frac{1}{\cos(t)}\, dt = \int_0^{\pi/4} \frac{\cos(t)}{1-\sin^2(t)}\, dt.$$

   We let $u = \sin(t)$, whence $du = \cos(t)dt$ and

$$\int_0^{\pi/4} \sec(t)dt = \int_0^{1/\sqrt{2}} \frac{1}{1-u^2}\, du.$$

Next, we simplify the integrand by writing

$$\frac{1}{1-u^2} = \frac{A}{1+u} + \frac{B}{1-u} \tag{11}$$

for constants A and B. By putting the right-hand side over a common denominator, we find that $A = B = \frac{1}{2}$; thus

$$\int_0^{\pi/4} \sec(t)dt = \frac{1}{2}\int_0^{1/\sqrt{2}} \left(\frac{1}{1+u} + \frac{1}{1-u}\right) du = \left(\log(1+u) - \log(1-u)\right)\Big|_0^{1/\sqrt{2}} = \log(1+\sqrt{2}).$$

2. The method, in (11), of rewriting a rational integrand as a sum of rational functions which are easier to antidifferentiate, is known as the method of *partial fractions.* Before presenting the general case, we consider a slightly more complicated example.

3. *Example.* We wish to evaluate $\displaystyle\int_3^4 \frac{4t^2-2t+3}{2t^3-4t^2+2t-4}\,dt.$

   The denominator may be factored as $2(t-2)\left(t^2+1\right)$. We write

$$\frac{4t^2-2t+3}{2t^3-4t^2+2t-4} = \frac{1}{2}\left(\frac{A}{t-2} + \frac{Bt+C}{t^2+1}\right),$$

   and solve for A, B and C (for example, by cross-multiplying and equating coefficients), getting $A = 3$, $B = 1$ and $C = 0$. Thus

$$\begin{aligned}
\int_3^4 \frac{4t^2-2t+3}{2t^3-4t^2+2t-4}\,dt &= \tfrac{1}{2}\int_3^4 \left(\frac{3}{t-2} + \frac{t}{t^2+1}\right) dt \\
&= \tfrac{1}{2}\int_3^4 \frac{3}{t-2}\,dt + \frac{1}{4}\int_{10}^{17} \frac{1}{u}\,du \\
&= \tfrac{3}{2}\log(t-2)\big|_3^4 + \tfrac{1}{4}\log(u)\big|_{10}^{17} \\
&= \tfrac{3}{2}\log(2) + \tfrac{1}{4}\log\left(\tfrac{17}{10}\right).
\end{aligned}$$

   (Note that the method of substitution is also used.)

4. **The method of partial fractions.** Let $f(x)$ and $g(x)$ be polynomials, with the degree of $f(x)$ strictly smaller than the degree of $g(x)$, and

$$g(x) = kL_1^{m_1}L_2^{m_2}\cdots L_i^{m_i}Q_1^{n_1}Q_2^{n_2}\cdots Q_j^{n_j}, \tag{12}$$

   where $L_1, L_2, \ldots, L_i$ are linear factors of the form $x-a$, and $Q_1, Q_2, \ldots, Q_j$ are irreducible quadratic factors of the form x^2+bx+c. Then $\frac{f(x)}{g(x)}$ may be rewritten as a sum of rational functions. This sum contains the terms

$$\frac{A_1}{L} + \frac{A_2}{L^2} + \cdots + \frac{A_m}{L^m}$$

   for each factor L^m in (12), and the terms

$$\frac{B_1x+C_1}{Q} + \frac{B_2x+C_2}{Q^2} + \cdots + \frac{B_nx+C_n}{Q^n}$$

for each factor Q^m in (12).

The proof of this theorem is a relatively straightforward application of the Euclidean algorithm for polynomials, but we omit it here. (It may be found on spiderwire.math.ubc.ca.)

5. Note that the presence of terms which are rational functions with irreducible quadratic denominators requires the use of the inverse trigonometric function $\arctan(x)$ which has been hitherto avoided. We consider examples involving this function in the exercises below.

G. Exercises

1. Evaluate the following integrals.
   (a) $\displaystyle\int_2^3 \frac{6}{t^2-1}\,dt.$
   (b) $\displaystyle\int_{-1}^1 \frac{5t}{t^2-t-6}\,dt.$
   (c) $\displaystyle\int_{-4}^{-3} \frac{6t^2}{t^4-5t^2+4}\,dt.$
2. Evaluate the following integrals.
   (a) $\displaystyle\int_0^2 \frac{6t^2+22t+18}{(t+1)(t+2)(t+3)}\,dt.$
   (b) $\displaystyle\int_{-1}^0 \frac{t^3+t^2-2t+1}{t-1}\,dt.$
   (c) $\displaystyle\int_0^1 \frac{2t^3+11t^2+28t+33}{t^2-t-6}\,dt.$
3. Write down definitions of the following inverse trigonometric functions, sketch their graphs, and then find their derivatives using implicit differentiation.
   (a) $\arcsin(x)$.
   (b) $\arccos(x)$.
   (c) $\arctan(x)$.
4. Evaluate the following integrals.
   (a) $\displaystyle\int_1^2 \frac{3t^2+2t+3}{t\,(t^2+1)}\,dt.$
   (b) $\displaystyle\int_2^3 \frac{3t^4-t^3+21t^2-4t+32}{t\,(t^2+4)^2}\,dt.$
   (c) $\displaystyle\int_0^1 \frac{5}{(t^2+1)^2}\,dt.$

Volumes

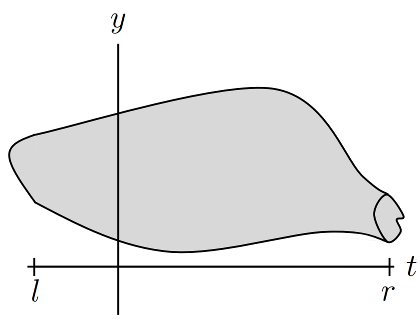
A. Motivation

Given a shape, how do we find its area? We saw at the beginning of this course that it is first necessary to define what we mean by "area". In our case, we used the intuitive notion of areas of rectangles to build up a formal notion of integrals under curves.

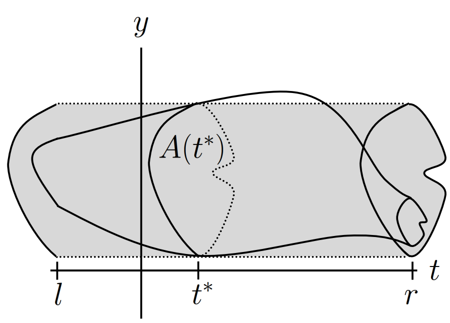
The idea of splitting up something large and difficult to understand into things which are infinitesimally small and, apart from their infinitesimal smallness, easy to understand is widely applicable elsewhere. (In fact, it is an idea that predates by centuries the formal calculus of Newton, Leibniz and later Weierstrass. Their crucial addition was rigour.) We shall consider two physical applications of the idea, to volumes and to work.

B. The definition of volume

1. Our definition of volume parallels our definition of area. Consider a three-dimensional solid on the interval $[l, r]$, as pictured below.



Our first approximation is to estimate the volume V to be equal to the area $A(t^*)$ of a "representative cross section" multiplied by the length $r - l$ — that is, to the volume of the cylinder of base area $A(t^*)$ and length $r - l$.

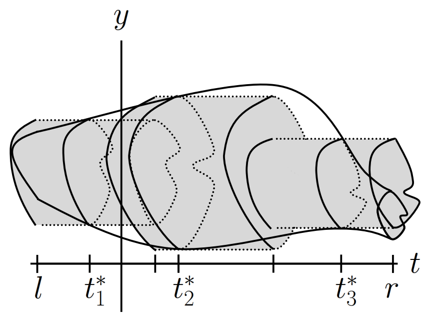


Thus

$$V \approx A(t^*)\,(r - l).$$

We can improve our approximation by partitioning $[l, r]$ into multiple subintervals of equal width, and selecting representative areas from each subinterval. For example, with three subintervals, we would have

$$V \approx \sum_{i=1}^{3} A\left(t_i^*\right)\left(\frac{r-l}{3}\right).$$



As we refine our partition into more and more subintervals, we expect our approximation to get closer and closer to the "actual" volume V.

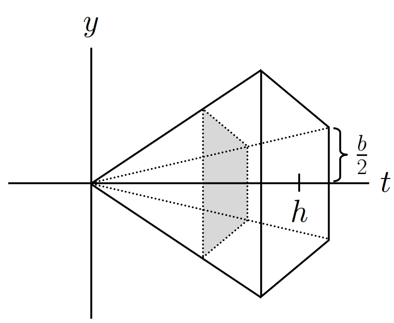
2. *Definition.* The volume of a solid on the interval $[l, r]$ having cross-sectional area $A(t)$ at position t is equal to

$$V = \int_l^r A(t)dt$$

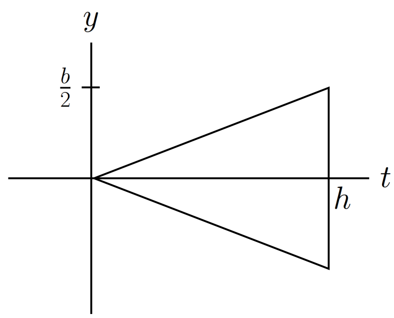
provided this integral exists.

3. *Example.* The volume of a square-based pyramid of height h and base side length b is $\frac{1}{3}b^2h$.

It is convenient to place the axes relative to the pyramid as follows.



We observe that the cross sections in that case are squares.



Indeed, note that, side-on, the top edge of the pyramid may be described by the line $y = \frac{b}{2h}\,t$. The cross-sectional area at t is therefore $\frac{b^2}{h^2}\,t^2$, and the volume is

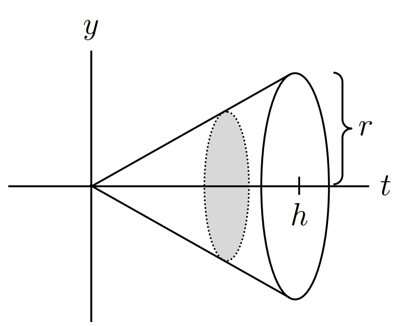
$$V = \int_0^h \frac{b^2}{h^2}\,t^2 dt = \frac{b^2}{h^2}\,\frac{t^3}{3}\bigg|_0^h = \frac{1}{3}b^2 h.$$

Note that, if our definition of volume is to be useful, we should be able to recover this answer by taking "slices" of the pyramid along *any* axis. (In most cases there will be a small number of "natural" axis choices.) We leave it as an exercise to calculate the volume of the pyramid using another axis choice.

C. Volumes by rotation and cylindrical shells

1. *Example.* A right circular cone of radius r and height h has volume $\frac{1}{3}\pi r^2 h$.

   As above, it is convenient to place the axes relative to the cone as follows.



   The cross-sectional area at t is $\pi\left(\frac{r}{h}\,t\right)^2$, and the volume is

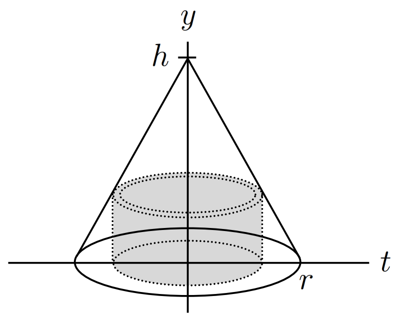
$$V = \int_0^h \pi\left(\frac{r}{h}\,t\right)^2 dt = \frac{\pi}{r^2}h^2\,\frac{t^3}{3}\bigg|_0^h = \frac{1}{3}\pi r^2 h.$$

2. *Calculating volumes by rotation.* The example above is an instance of calculating *volumes by rotation*; that is, the cross-sectional areas are circles. In general, if the region bounded by $t = l$, $t = r$ on the left

and right, and $y = 0$ and $y = f(t)$ on the bottom and top, is rotated about the t-axis, the resulting solid has a cross-sectional area at t of $\pi f(t)^2$, and a volume of

$$V = \int_l^r \pi f(t)^2 dt.$$

3. *Example.* We consider again the right circular cone of radius r and height h. However, this time we divide the solid into thin "cylindrical shells" rather than slices.



At t, the shell has radius t, height $h - \frac{h}{r}t$, and thickness Δt, hence volume $2\pi t\left(h - \frac{h}{r}t\right)\Delta t$. Summing the volumes of these shells and taking the limit yields

$$V = \int_0^r 2\pi t\left(h - \frac{h}{r}t\right)dt = \pi h t^2 - \frac{2h}{3r}\pi t^3\Big|_0^r = \frac{1}{3}\pi r^2 h.$$

4. *Calculating volumes by cylindrical shells.* The example above is an instance of calculating *volumes by cylindrical shells*. In general, if the region bounded by $t = l$, $t = r$ on the left and right, and $y = 0$ and $y = f(t)$ on the bottom and top, is rotated about the y-axis, the resulting solid has a cylindrical shell at t of area $2\pi t f(t)$, and a volume of

$$V = \int_l^r 2\pi t f(t) dt.$$

5. Note that it is not *a priori* obvious that calculating volumes by rotation and by cylindrical shells yields the same answer in all cases. In one of the questions below, we prove that the methods are consistent given certain highly constrained conditions. Rather more work is necessary to prove that they are consistent in general.

D. Exercises

1. Let R denote the finite region enclosed by $y = t$ and $y = t^2$. Calculate the volume of the solid obtained by rotating R about the following lines.
   (a) The t-axis.
   (b) The y-axis.
   (c) The line $t = 2$.
   (d) The line $y = t$. (Hint: write the volume as $\displaystyle\int_0^{\sqrt{2}} \pi r(z)^2 dz$ for a suitable $r(z)$.)

2. Calculate the volume of a sphere of radius r.
3. Calculate the volume of a tetrahedron of side length a.
4. Let R denote the region enclosed by the t-axis and $y = \frac{2\sqrt{t}}{4+t^2}$ in the half-plane $t \geq 0$. Find the volume of the solid obtained by rotating R about the t-axis.
5. Let R denote the region enclosed by $y = t^{12} - t^8 + 3t^6 + t^4 + t^3 + 16t - 4$ and $y = t^{12} - t^8 + 3t^6 + t^4 + 6t^2 + 8t - 4$ from $t = 0$ to $t = 4$. Find the volume of the solid obtained by rotating R about the y-axis.
6. Let R denote the region enclosed by $y = \sin(t)$ and $y = \cos(t)$ from $t = -\frac{\pi}{4}$ to $t = \frac{3\pi}{4}$. Find the volume of the solid obtained by rotating R about $x = -\frac{\pi}{4}$.
7. Let R denote the region enclosed by $y = t^3$ and $y = 1$ from $t = 0$ to $t = 1$. Let V denote the volume of the solid obtained by rotating R about the line $y = 1$. Find at least one other line on the plane such that the volume of the solid obtained by rotating R about that line is equal to V.
8. Let $f(t)$ be a continuously differentiable function (recall that a function is *continuously differentiable* if its derivative is continuous) which passes through the origin and is strictly increasing. Let R be the region enclosed by the t-axis and $y = f(t)$ from $t = 0$ to $t = a$. Let S denote the solid obtained by rotating R about the y-axis. Prove that the volume of S is equal regardless of whether it is calculated by rotation or by cylindrical shells.

Work

A. Motivation

Previously, we applied the mechanism of the integral — in general, the idea of splitting up a large object into infinitesimally small components — to volumes of solids. Our second application of the mechanism is to the physical concept of *work*.

B. The definition of work

1. Work is the energy identified with the action of a force. We define the work W done by a constant force F on a point that is displaced a distance d in the direction of the force to be

   $$W = Fd.$$

   It may be measured in newton-metres; that is, joules (J).

2. *Example.* A person lifting an object off the ground does positive work on the object: he applies an upward force on the object, which is displaced upward. The same person lowering the object back to the ground does negative work: he applies an upward force on the object (against the force of gravity), which is displaced in the opposite direction. The person does no work moving the object at a constant speed horizontally: the force applied is still upward against the force of gravity, but there is no force applied in the direction of displacement.

3. Defining work becomes rather more difficult if the force is not applied constantly.

   As in the case of volume, our definition of work parallels our definition of area. Suppose an object moves along the t-axis from $t = l$ to $t = r$, with a force of $f(t)$ acting in the same direction on the object at any point t.

   We attain our first approximation by assuming the force is constant over the interval $[l, r]$:

   $$W \approx f(t^*)(r - l),$$

   where $f(t^*)$ is a "representative" force.

   We can improve our approximation by partitioning $[l, r]$ into multiple subintervals of equal width, and selecting representative areas from each subinterval. For example, with three subintervals, we would have

   $$W \approx \sum_{i=1}^{3} f(t_i^*) \left(\frac{r - l}{3} \right).$$

   As we refine our partition into more and more subintervals, we expect our approximation to get closer and closer to the "actual" work W.

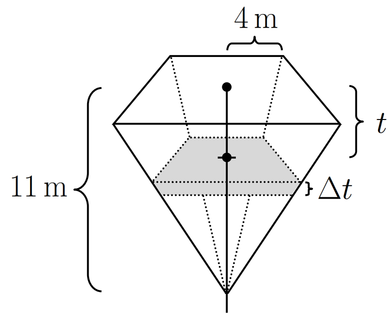
4. *Definition.* Suppose an object moves along the t-axis from $t = l$ to $t = r$, with a force of $f(t)$ acting in the same direction on the object at any point t. The work done on the object is equal to

$$W = \int_l^r f(t)dt$$

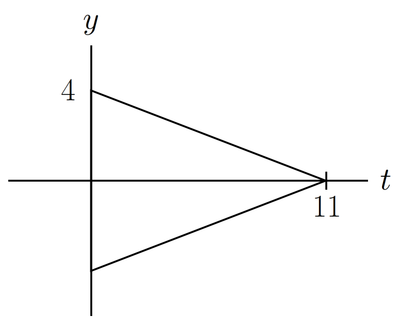
provided this integral exists.

5. *Example.* Suppose a tank in the shape of an inverted pyramid of height 11 m and base side length 8 m is filled with water. We wish to find the work required to empty the tank by pumping all the water to the top of the tank.

It is convenient to place the origin at the top of the tank and the t-axis pointing downward. We consider the water to be divided into small slices of thickness Δt.



At t, the cross-sectional area is $4\left(4 - \frac{4}{11}\,t\right)^2$ square metres. (To see this, consider the pyramid side-on, and observe that the top edge may be described by the line $y = 4 - \frac{4}{11}\,t$.)

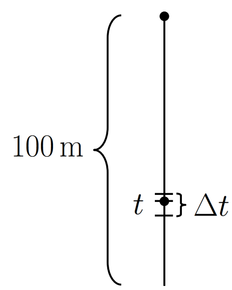


Thus the slice has volume $4\left(4 - \frac{4}{11}\,t\right)^2 \Delta t$, and the work done to raise this slice a distance t to the top of the tank is equal to $4\rho g t\left(4 - \frac{4}{11}\,t\right)^2 \Delta t$, where ρ is the density of water (which we take to be equal to $1000\,\mathrm{kg/m^3}$), and g is the acceleration due to gravity (which we take to be equal to $9.8\,\mathrm{m/s^2}$ on or near the surface of the Earth). Summing the work over all slices and taking the limit yields

$$\begin{aligned} W &= \int_0^{11} 4\rho g t \left(4 - \frac{4}{11}t\right)^2 dt \\ &= 4\rho g \int_0^{11} \left(16t - \frac{32}{11}t^2 + \frac{16}{121}t^3\right) dt \\ &= 4\rho g \left(8t^2 - \frac{32}{33}t^3 + \frac{4}{121}t^4\right)\Bigg|_0^{11} \\ &= 4\rho g \left(\frac{484}{3}\right) \\ &\approx 6330720. \end{aligned}$$

Thus approximately 6330720 J of work is done.

6. *Example.* Suppose one end of a 100 m steel rope weighing 91 kg is attached to the lip of the roof of a 120 m-tall building. We wish to calculate the work done in hoisting the rope up to the roof of the building.



It is convenient to place the origin at the top of the building and the t-axis pointing downward. We consider the rope to be divided into small sections of length Δt. At t, the section has length Δt and linear density $\frac{91}{100}$ kg/m. Thus the work done to raise this section a distance t to the roof of building is equal to $\frac{91}{100}gt\Delta t$. Summing the work over all sections and taking the limit yields

$$W = \int_0^{100} \frac{91}{100} gt\,dt = \frac{91}{200}gt^2\Bigg|_0^{100} = \frac{9100g}{2} = 44590.$$

Thus approximately 44590 J of work is done.

7. The main challenge in the work problems presented here and in the following section is to set up the problem in a useful way; the mathematics is straightforward.

C. Exercises

1. Consider an inverted right circular conical tank of radius 1 m and height 3 m filled with water. Calculate the work done in pumping all the water out the top of the tank.
2. Consider a spherical tank of radius 1 m filled with fluid of density 1240 kg/m^3. Calculate the work done in pumping all the fluid out the top of the tank.

3. Consider three tanks of equal height and volume, filled with water. The first tank is in the shape of a right circular cone. The second tank is the same shape, but inverted. The third tank is cylindrical. In all three cases, all the water is pumped out the top of the tank. In which case is the most work done? In which case is the least work done?

4. Suppose one end of a 100 m steel rope weighing 91 kg is attached to the lip of the roof of a 40 m-tall building. Calculate the work done in hoisting the other end of the rope up to the roof of the building.

5. (a) A 1 kg bucket on a rope of linear density 0.5 kg/m is drawn up a height of 40 m. Calculate the work done.
   (b) Calculate the work done in the previous scenario if the bucket leaks at a constant rate so that it is only half full by the time it reaches the top.

6. *Hooke's Law* states that the force required to hold a spring stretched or compressed a distance d beyond its natural length is equal to kd, where k is a positive constant associated with the spring. Suppose a spring has a natural length of 10 cm, and that a force of 4 N maintains it stretched to a length of 12 cm. Calculate the work done to stretch the spring from 10 cm to 15 cm.

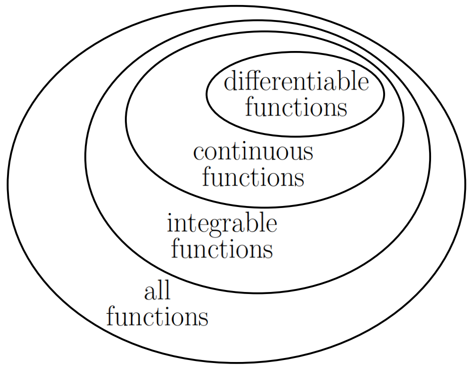
7. Suppose 3 J of work is done to compress the length of a spring by a factor of 2. Calculate the work done to compress the spring by a factor of 3.

8. The Earth exerts a gravitational force on an object of mass m (in kilograms) a height h (in metres) above the surface of the Earth of $\frac{km}{(r+h)^2}$, where k and r are positive constants (r is the radius of the Earth). Calculate the work done to raise a 1 kg object to a height of 10000 m.

Power series

A. Motivation

We now have definitions for what it means for a function to be differentiable, continuous or integrable.



Next, we consider how these definitions may be applied not just to functions as we know them, but to "series of functions".

As an example, recall the geometric series

$$\sum_{n\geq 1} \frac{1}{2^{n-1}} = 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots .$$

This series converges to 2. Indeed, the series

$$\sum_{n\geq 1} x^{n-1} = 1 + x + x^2 + x^3 + \cdots$$

converges to $\frac{1}{1-x}$, provided $|x| < 1$. The phrase "converges to" is deliberate; it is not *a priori* true that the series $1 + x + x^2 + x^3 + \cdots$ is "equal" to the function $\frac{1}{1-x}$, just as it is not true that the series $1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots$ is "equal" to 2. However, in what way *are* they equal? Can we do arithmetic on the series as we can on the function? Can we differentiate and integrate one as we can the other? It turns out that, under certain conditions, we can; and that dealing with series representations of functions is in many cases a more convenient, elegant, and even "natural" thing to do than dealing with the function itself.

B. Series revisited

1. Before proceeding, it will be useful to list the series convergence tests established earlier, and to introduce two new tests. We will make the most use of the Ratio Test to determine where series display absolute convergence.

2. **The Divergence Test.** If $\displaystyle\sum_{n\geq 1} a_n$ converges, then $\{a_n\}$ converges to 0.

3. **The Comparison Test.** Let $\displaystyle\sum_{n\geq 1} a_n$ and $\displaystyle\sum_{n\geq 1} b_n$ be series with all positive terms.

(a) If $\sum_{n\geq 1} b_n$ converges and $a_n \leq b_n$ for all n, then $\sum_{n\geq 1} a_n$ converges.

(b) If $\sum_{n\geq 1} b_n$ diverges and $a_n \geq b_n$ for all n, then $\sum_{n\geq 1} a_n$ diverges.

4. **The Limit Comparison Test.** Let $\sum_{n\geq 1} a_n$ and $\sum_{n\geq 1} b_n$ be series with all positive terms. If

$$\lim_{n\to\infty} \frac{a_n}{b_n} = L > 0,$$

then both series converge or both series diverge.

5. **The Ratio Test.** Let $\sum_{n\geq 1} a_n$ be a series with all positive terms.

(a) If $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L < 1$, then $\sum_{n\geq 1} a_n$ converges.

(b) If $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L > 1$ or $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \infty$, then $\sum_{n\geq 1} a_n$ diverges.

6. **The Root Test.** Let $\sum_{n\geq 1} a_n$ be a series with all positive terms.

(a) If $a_n^{1/n} = L < 1$, then $\sum_{n\geq 1} a_n$ converges.

(b) If $a_n^{1/n} = L > 1$, then $\sum_{n\geq 1} a_n$ diverges.

7. **The Alternating Series Test.** Let $\sum_{n\geq 1} (-1)^{n-1} a_n$, where $a_n > 0$ for $n \geq 1$, satisfy:

(a) $a_n \geq a_{n+1}$ for $n \geq 1$, and

(b) $\lim_{n\to 0} a_n = 0$.

Then the series converges.

8. **The Geometric Series Test.** $\sum_{n\geq 1} ar^{n-1}$ converges to $\frac{a}{1-r}$ if $|r| < 1$, and diverges if $|r| \geq 1$.

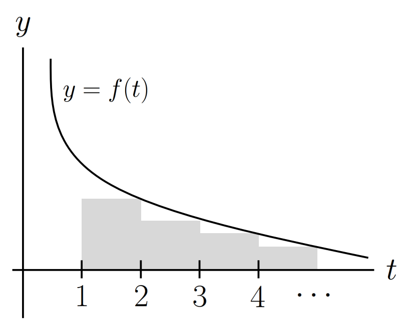
9. **The Integral Test.** Let $a_n = f(n)$, where $f(t)$ is continuous, positive, and non-increasing for $t \geq 1$.

(a) If $\int_1^\infty f(t)dt$ converges, then $\sum_{n\geq 1} a_n$ converges.

(b) If $\int_1^\infty f(t)dt$ diverges, then $\sum_{n\geq 1} a_n$ diverges.

Proof. Suppose $\int_1^\infty f(t)dt$ converges. We note that

$$a_2 + a_3 + \cdots + a_n \leq \int_1^n f(t)dt.$$



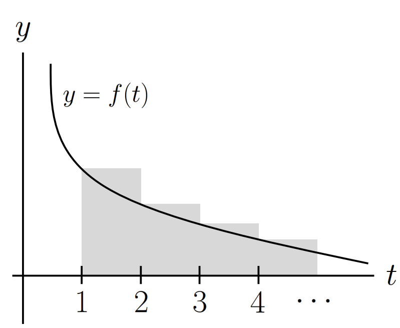
In particular, the partial sums of $\sum_{n\geq 1} a_n$ are bounded above —

$$a_1 + a_2 + \cdots + a_n \leq a_1 + \int_1^\infty f(t)dt$$

— and clearly increasing, hence converge.

Suppose now that $\int_1^\infty f(t)dt$ diverges. We note that

$$a_1 + a_2 + \cdots + a_n \geq \int_1^n f(t)dt.$$



Taking the limit of both sides as $n \to \infty$ shows that the partial sums diverge. $\square$

10. **The p-series Test.** $\sum_{n\geq 1} \frac{1}{n^p}$ converges if $p > 1$, and diverges if $p \leq 1$.

    Proof. This is a corollary of the Integral Test. $\square$

C. Power series

1. *Definition.* A series

$$\sum_{n\geq 0} a_n(x-c)^n = a_0 + a_1(x-c) + a_2(x-c)^2 + \cdots$$

is a *power series about* c. The constants $a_0, a_1, a_2, \ldots$ are the *coefficients* of the series; and c is the *centre of convergence*, so called since the series always converges at $x = c$.

2. *Examples.* We wish to determine for what values of x the following series converge.

(a) $\displaystyle\sum_{n\geq 0} n!x^n$.

(b) $\displaystyle\sum_{n\geq 0} \frac{x^n}{n!}$.

(c) $\displaystyle\sum_{n\geq 0} \frac{\left(\frac{2}{3}\right)^n (x+5)^n}{n+1}$.

We apply the Ratio Test to all three series. For the first, we have

$$\lim_{n\to\infty} \left| \frac{(n+1)!x^{n+1}}{n!x^n} \right| = \lim_{n\to\infty} (n+1)\,|x| = \infty$$

for all values of $x \neq 0$; thus the series converges only at the centre $x = 0$.

For the second, we have

$$\lim_{n\to\infty} \left| \frac{x^{n+1}}{x^n} \cdot \frac{n!}{x^n} \right| = \lim_{n\to\infty} \frac{|x|}{n+1} = 0$$

for all values of x; thus the series converges everywhere.

For the final series, we have

$$\lim_{n\to\infty} \left| \frac{\left(\frac{2}{3}\right)^{n+1} (x+5)^{n+1}}{n+2} \cdot \frac{n+1}{\left(\frac{2}{3}\right)^n (x+5)^n} \right| = \lim_{n\to\infty} \frac{2}{3} \left(\frac{n+1}{n+2} \right) |x+5| = \frac{2}{3}\,|x+5|\,, \tag{13}$$

which converges if

$$|x+5| < \frac{3}{2};$$

that is, if x is in the interval $\left(-5 - \frac{3}{2}, -5 + \frac{3}{2}\right)$. Now recall that the Ratio Test provides no information if the limit in (13) is equal to 1; that is, at the endpoints of the interval. We determine the convergence of the power series at the endpoints directly. If $x = -5 - \frac{3}{2}$, the series is equal to $\displaystyle\sum_{n\geq 0} \frac{(-1)^n}{n+1}$, which converges by the Alternating Series Test. On the other hand, if $x = -5 + \frac{3}{2}$, the series is equal to $\displaystyle\sum_{n\geq 0} \frac{1}{n+1}$, which is a divergent harmonic series. Thus in summary, the series converges if, and only if, x is in the interval $\left[-5 - \frac{3}{2}, -5 + \frac{3}{2}\right)$.

3. The examples above demonstrate the domain of possibilities for the convergence of power series.

   Theorem. For $\sum_{n\geq 0} a_n(x-c)^n$, exactly one of the following holds:

   (a) the series converges only at $x = c$,
   (b) the series converges everywhere, or
   (c) the series converges if $|x-c| < R$ and diverges if $|x-c| > R$ (R is called the series' *radius of convergence*).

   Moreover, in all three cases, the convergence is absolute. In the third case, the series may converge, but not necessarily absolutely, at either endpoint of the interval of convergence.

   Proof. It suffices to show that, if the series converges at $x_0 \neq c$, then it converges absolutely at any x that is closer to c; that is, at any x such that $|x-c| < |x_0-c|$.

   If $\sum_{n\geq 0} a_n(x_0-c)^n$ converges, by the Divergence Test we have $|a_n(x_0-c)^n| \leq a$ for some positive constant a. Thus

$$\sum_{n\geq 0} |a_n(x-c)^n| = \sum_{n\geq 0} |a_n(x_0-c)^n| \left|\frac{x-c}{x_0-c}\right|^n \leq \sum_{n\geq 0} a \left|\frac{x-c}{x_0-c}\right|^n,$$

   which converges by the Geometric Series Test. □

4. *Intervals of convergence.* Thus power series converge in *intervals of convergence* centred on the centre of convergence, and diverge elsewhere. It turns out that, where they converge, power series behave like functions: they can be added, subtracted, multiplied, divided, composed, raised to powers, differentiated and integrated.

D. Exercises

1. Determine if the following series converge.
   (a) $\sum_{n\geq 1} \frac{n}{n^2+5}$.
   (b) $\sum_{n\geq 1} \frac{n^3}{e^{n^4}}$.
   (c) $\sum_{n\geq 10} \frac{1}{n\log(n)}$.
2. Determine whether each of the following statements is true or false, and justify your answer.
   (a) If $\sum_{n\geq 0} a_n x^n$ diverges for $x = 3$, then it diverges for $x = 4$.
   (b) If $\sum_{n\geq 0} a_n x^n$ converges for all positive x, then it converges for all negative x.
   (c) If $\sum_{n\geq 0} a_n(x-c)^n$ has a radius of convergence $R > 0$, then the series $\sum_{n\geq 0} a_n$ converges.

3. Suppose $\sum_{n\geq 0} a_n x^n$ has radius of convergence R. Determine the radius of convergence of $\sum_{n\geq 0} a_n x^{n/2}$.

4. Find the interval of convergence for each of the following power series.
   (a) $\sum_{n\geq 1} \frac{(2x-3)^n}{n}$.
   (b) $\sum_{n\geq 1} \frac{(-x)^n}{n^4 4^n}$.
   (c) $\sum_{n\geq 0} \frac{(1+3^n)\, x^n}{n!}$.

5. Find the interval of convergence for each of the following power series.
   (a) $\sum_{n\geq 1} \frac{(x+2)^n}{n^n}$.
   (b) $\sum_{n\geq 1} \frac{x^{3n}}{n\log(n)}$.
   (c) $x + \frac{x^2}{3} + \frac{x^3}{5(3)} + \frac{x^4}{7(5)(3)} + \frac{x^5}{9(7)(5)(3)} + \cdots$.

6. (a) Write down a power series that has interval of convergence $[2, 8)$.
   (b) Write down a power series that has interval of convergence $[2, 8]$.

7. Determine the interval of convergence for the series

$$3 + x + 3x^2 + x^3 + 3x^4 + x^5 + \cdots ,$$

   as well as a formula for the function it converges to. (Hint: pair the terms.)

8. Determine the interval of convergence, as well as a formula for the function it converges to, for the series $\sum_{n\geq 0} a_n x^n$ where $a_{n+2} = a_n$ for all $n \geq 0$.

E. Operations on power series

1. It is straightforward to show that power series can be added and subtracted in the "natural" way inside their intervals of convergence: that is,

$$\sum_{n\geq 0} a_n (x-c)^n + \sum_{n\geq 0} b_n (x-c)^n = \sum_{n\geq 0} (a_n + b_n)\,(x-c)^n$$

   provided the series on the left-hand side converge. The radius of convergence of the series on the right-hand side is greater than or equal to the radii of convergence on the left-hand side.

2. Other arithmetic operations — multiplication, division, composition, exponentiation — are complicated to prove, and we defer them to textbooks in analysis.

3. *Differentiation and integration of power series.* On their intervals of convergence, except possibly at their endpoints, power series may be differentiated and integrated term-by-term. Thus we may think of them,

conveniently, as "infinite polynomials", and we may also think of them as *equal to* the functions they converge to. From this point onward, we say that

$$\sum_{n\geq 0} a_n(x-c)^n = f(x-c)$$

if the series converges to the function $f(x-c)$.

Theorem. Let

$$f(x) = \sum_{n\geq 0} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots$$

for $|x| < R$. Then for the same values of x, we have

$$f'(x) = \sum_{n\geq 1} n a_n x^{n-1} = a_1 + 2a_2 x + 3a_3 x^2 + \cdots$$

and

$$\int_0^x f(t)dt = \sum_{n\geq 0} \frac{a_n}{n+1} x^{n+1} = a_0 x + \frac{a_1}{2}x^2 + \frac{a_2}{3}x^3 + \cdots.$$

Proof. (Výborný) We start by showing $\sum_{n\geq 1} n a_n x^{n-1}$ converges. Let $H > 0$ be such that $|x| + H < R$. Then

$$\sum_{n\geq 1} \left|n a_n x^{n-1}\right| = \frac{1}{H}\sum_{n\geq 1} |a_n| \left(n|x|^{n-1}H\right) \leq \frac{1}{H}\sum_{n\geq 1} |a_n| \left(|x|+H\right)^n,$$

which converges since $|x| + H$ is in the interval of convergence.

Next, we show that the series $g(x) = \sum_{n\geq 1} n a_n x^{n-1}$ is in fact the derivative of $f(x) = \sum_{n\geq 0} a_n x^n$. Now

$$\left|\frac{f(x+h)-f(x)}{h} - g(x)\right| = \left|\sum_{n\geq 0} \frac{a_n(x+h)^n - a_n x^n - n a_n x^{n-1}h}{h}\right| \leq \sum_{n\geq 0} |a_n| \left|\frac{(x+h)^n - x^n - n x^{n-1}h}{h}\right|.$$

Note that x^n and $nx^{n-1}h$ are two terms in the expansion of $(x+h)^n$. Indeed, for sufficiently small h, we have, for appropriate binomial coefficients $c_2, c_3, \ldots, c_n$,

$$\begin{aligned}
\left|(x+h)^n - x^n - nx^{n-1}h\right| &\leq c_2|x|^{n-2}|h|^2 + c_3|x|^{n-3}|h|^3 + \cdots + c_n|h|^n \\
&\leq \frac{|h|^2}{H^2}\left(c_2|x|^{n-2}H^2 + c_3|x|^{n-3}H^3 + \cdots + c_n H^n\right) \\
&\leq \frac{|h|^2}{H^2}\left(|x|+H\right)^n.
\end{aligned}$$

Thus

$$\left|\frac{f(x+h)-f(x)}{h} - g(x)\right| \leq \frac{|h|}{H^2}\sum_{n\geq 0} |a_n| \left(|x|+H\right)^n.$$

Taking the limit as $h \to 0$ yields 0 on both sides, and by the definition of derivative, $f'(x) = g(x)$ as needed.

Finally, consider $h(x) = \sum_{n\geq 0} \frac{a_n}{n+1} x^{n+1}$. Since $\left|\frac{a_n}{n+1}\right| \leq |a_n|$, $h(x)$ converges for $|x| < R$ (indeed, the convergence is absolute, and $h(x)$ may in fact converge for more values of x). By the result just established, we can differentiate the series term-by-term:

$$h'(x) = \sum_{n \geq 0} a_n x^n = f(x).$$

By the Fundamental Theorem of Calculus, we have

$$\int_0^x f(t)dt = h(t)|_0^x = h(x),$$

which completes the proof. $\square$

4. *Example.* We wish to find power series representations for the functions $\frac{1}{(1-x)^2}$ and $\log(1+x)$.

We begin with the series

$$\frac{1}{1-x} = \sum_{n \geq 0} x^n = 1 + x + x^2 + x^3 + \cdots \tag{14}$$

for $-1 < x < 1$. Differentiating (14), we get

$$\frac{1}{(1-x)^2} = \sum_{n \geq 1} n x^{n-1} = 1 + 2x + 3x^2 + 4x^3 + \cdots$$

for $-1 < x < 1$. To get a power series for $\log(1+x)$, we observe first that we have

$$\frac{1}{1+t} = \sum_{n \geq 0} (-1)^n t^n = 1 - t + t^2 - t^3 + \cdots \tag{15}$$

for $-1 < t < 1$. We integrate (15):

$$\log(1+x) = \int_0^x \frac{1}{1+t}\,dt = \int_0^x \left(1 - t + t^2 - t^3 + \cdots\right) dt = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$$

for $-1 < x < 1$. In fact, checking the endpoints shows that we have convergence for $-1 < x \leq 1$. (The fact that, given convergence at $x = 1$, we have convergence to $\log(1+1)$, is due to a result known as Abel's Theorem, whose statement we will defer to textbooks in analysis.)

F. Exercises

1. Find power series representations for the following functions, along with their intervals of convergence.
   (a) $\frac{1}{2-x}$.
   (b) $\frac{1}{x^2+1}$.
   (c) $\frac{3}{1-2x^2}$.
2. Find power series representations for the following functions, along with their intervals of convergence.
   (a) $\frac{x}{x^2+1}$.
   (b) $\frac{1+x}{1-x}$.
   (c) $\frac{2x+3}{x^2+3x+2}$.

3. Find power series representations for the following functions, along with their intervals of convergence.
   (a) $\frac{1}{(1+x)^3}$.
   (b) $\frac{1}{(1-x)^3}$.
   (c) $\arctan(x)$.
4. Prove using your answer to part (c) in the previous question that $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$. What assumptions are you making without proof?
5. (a) Find a formula for the function that $\sum_{n \geq 1} n^2 x^n$ converges to.
   (b) Determine what $\sum_{n \geq 1} \frac{n^2}{3^n}$ converges to. (Hint: use part (a).)
6. Prove that $\sum_{n \geq 0} \frac{x^n}{n!}$ satisfies the differential equation $f'(x) = f(x)$.
7. Prove that $\sum_{n \geq 0} \frac{(-1)^n x^{2n}}{(2n)!}$ satisfies the differential equation $f''(x) = -f(x)$.
8. Find a power series that converges to $\frac{1}{1+x}$ but has centre 2. (Hint: set $t = x - 2$.)

Linear and higher degree approximations

A. Motivation

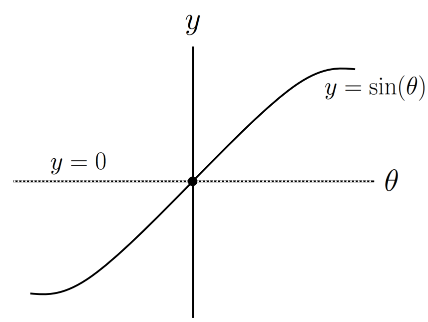
Physicists often approximate $\sin(\theta) \approx \theta$, $\cos(\theta) \approx 1$ or $\tan(\theta) \approx \theta$ for small values of θ. For example, in astronomy, this last approximation is used to derive the approximate formula

$$D = \frac{\theta d}{206265},$$

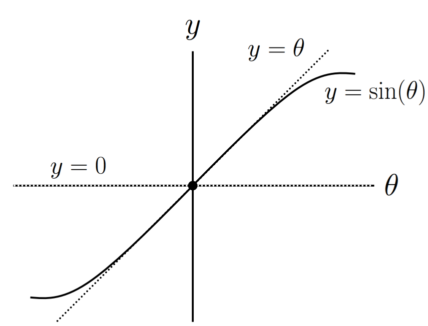
relating the linear size D of an object a long distance d away to its angular size θ, from the exact formula

$$D = d \tan\left(\frac{2\pi\theta}{1296000}\right).$$

What justifies these *small angle approximations*? Consider the function $\sin(\theta)$. For θ near 0, we may approximate $\sin(\theta) \approx \sin(0) = 0$: the approximation matches the function at 0.



However, given that $f(\theta)$ is differentiable, we may improve our approximation by also insisting that the *derivative* of the approximation match the derivative of the function at 0; namely, that $\sin(\theta) \approx \theta$. This is the *linear approximation* of $\sin(\theta)$ at 0.



In fact, since $\sin(\theta)$ is infinitely differentiable, we may extend this process indefinitely; we may insist on matching second derivatives, and third, and fourth. We shall see that, if we extend the process *ad infinitum*, we end up with a power series. This reverses the process introduced in the previous section: instead of starting with a power series and observing that we have a function, we start with a function and observe that we have a power series.

In both the linear approximation and the higher-degree — or indeed "infinite-degree" — approximations, the advantage is that the approximate form is often much easier and more natural to deal with.

B. Linear approximations

1. *Definition.* Let $f(x)$ be differentiable at a. The The *linear approximation of $f(x)$ at c* is

$$f(x) \approx f(c) + f'(c)(x - c).$$

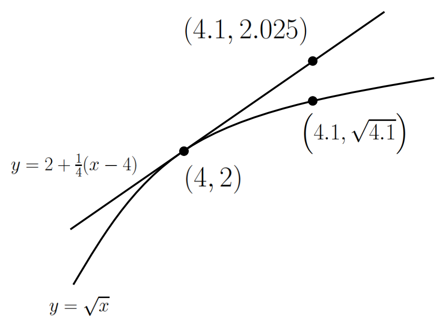
   We call the function $L(x) = f(c) + f'(c)(x - c)$ the *linearization* of $f(x)$ at c.

   Note that the linearization is simply the tangent line at c.

2. *Example.* We wish to estimate $\sqrt{4.1}$ using a linear approximation.

   We consider the linear approximation of $f(x) = \sqrt{x}$ at 4: $f(x) \approx 2+\frac{1}{4}(x-4)$. In particular, $f(4.1) \approx 2.025$.

   In addition, since $f(x)$ is concave down on an interval containing 4 and 4.1 — indeed everywhere on its domain — we may conclude that the approximation is greater than the actual value of $\sqrt{4.1}$. (This may also be determined directly in this case by squaring 2.025.) This is apparent from the picture below.



3. In the example above, note that the choice of both the function and the *centre* 4 may be changed. For example, it is equivalent, but perverse, to consider the linear approximation of $\sqrt{x+1}$ at 3. It is not as sensible to consider the linear approximation of $\sqrt{x}$ at 1. In general, to approximate $f(x_0)$, we choose the centre c to be point nearest x_0 at which the linearization may be easily calculated.

4. *The Generalized Mean Value Theorem.* While we may use concavity to determine whether our linear approximation is an overestimation or an underestimation, it is useful to be able to bound the size of the *error* — that is, the difference between the actual value and the approximate value of the function. We present below a theorem that does exactly that, and indeed recovers the information that concavity provides. First, however, we need a generalization of the Mean Value Theorem.

   The Generalized Mean Value Theorem. Let $f(x)$ and $g(x)$ be continuous on $[l, r]$ and differentiable on (l, r), and let $g'(x) \neq 0$ for all x in (l, r). Then there exists a number a in (l, r) such that

$$\frac{f'(a)}{g'(a)} = \frac{f(r) - f(l)}{g(r) - g(l)}.$$

Proof. We apply Rolle's Theorem to

$$h(x) = (f(x) - f(l))\,(g(r) - g(l)) - (f(r) - f(l))\,(g(x) - g(l))\,.$$

(While this function is somewhat mysterious and rather difficult to visualize, note that if we take $g(x)$ to be the identity function, we recover the function which was use in the same way to prove the Mean Value Theorem last term.) There exists a number a in (l, r) such that

$$h'(a) = f'(a)\,(g(r) - g(l)) - (f(r) - f(l))\,g'(a) = 0,$$

from which the result follows. $\square$

5. **Theorem.** Let $L(x)$ be the linearization of $f(x)$ at c, and let $E(x) = f(x) - L(x)$ be the error. Then there exists a number s between c and x such that

$$E(x) = \frac{f''(s)}{2}(x - c)^2,$$

provided $f(x)$ is twice-differentiable in an interval containing c and x.

Proof. Assume $x > c$ (the proof if $x < c$ is similar). We apply the Generalized Mean Value Theorem to the functions $E(t)$ and $(t - c)^2$ on the interval $[c, x]$, and conclude that there exists a number a in (c, x) such that

$$\frac{E'(a)}{2(a - c)} = \frac{E(x) - E(c)}{(x - c)^2 - (c - c)^2} = \frac{E(x)}{(x - c)^2}.$$

Now since $E(t) = f(t) - f(c) - f'(c)(t - c)$, we have $E'(t) = f'(t) - f'(c)$. Thus

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

But here we may apply the Mean Value Theorem to $f'(x)$ on $[c, a]$, and conclude that there exists a number s in (c, a) — that is, in (c, x) — such that

$$\frac{E(x)}{(x - c)^2} = \frac{1}{2}f''(s),$$

as needed. $\square$

We will generalize this theorem in section D.

6. *Example.* We wish to estimate the size of the error in the approximation $\sqrt{4.1} \approx 2.05$.

According to the theorem just proven, taking $f(x) = \sqrt{x}$, $a = 4$ and $L(x) = 2 + \frac{1}{2}(x - 4)$, the error $E(x) = f(x) - L(x)$ is given by

$$E(x) = \frac{1}{2}\left(-\frac{1}{4}s^{-3/2}\right)(x - 4)^2$$

for a number s between 4 and x. In particular, $E(4.1) = -\frac{1}{800}s^{-3/2}$ for a number s between 4 and 4.1. We may bound the error by taking $s = 4$, getting $|E(4.1)| \leq \frac{1}{6400} = 0.00015625$.

A calculator (but how does *that* calculate the error?) tells us that $\sqrt{4.1} - 2.025 \approx -0.000154$, which is indeed smaller in absolute value than the error bound predicted by the theorem.

C. Exercises

1. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *Other than at the centre, a function is never equal to its linear approximation.*
   (b) *All functions have linear approximations.*
   (c) *It is more accurate to approximate* $\sqrt{2}$ *using the function* $\sqrt{x}$ *and a linear approximation about* 4*, than about* 1*.*
2. Use a linear approximation to estimate the following numbers. Then state whether your approximation is an underestimate, an overestimate, or neither. If it is an underestimate or overestimate, bound the size of the error.
   (a) $e^{0.1}$.
   (b) $\log(0.9)$.
   (c) $\sqrt{24}$.
3. Repeat the process for Question 2 using a different function and centre, but obtaining the same result.
4. Repeat the process for Question 2 using a different function and centre, and obtaining a different result.
5. Use a linear approximation to estimate the following numbers. Then state whether your approximation is an underestimate, an overestimate, or neither. If it is an underestimate or overestimate, bound the size of the error.
   (a) $(1001)^{1/3}$.
   (b) $\frac{1}{3.99}$.
   (c) $\sqrt{25}$.
6. Use a linear approximation to estimate the following numbers. Then state whether your approximation is an underestimate, an overestimate, or neither. If it is an underestimate or overestimate, bound the size of the error.
   (a) $\tan(-0.1)$.
   (b) $\sin^2(0.1)$.
   (c) $\cos(46°)$.
7. Suppose $f(1) = 2$ and $f'(x) = e^{x^2}$ for all x. Use a linear approximation to estimate $f(1.1)$. Is your approximation is an underestimate, an overestimate, or neither? If it is an underestimate or overestimate, bound the size of the error.
8. Suppose the radius of a sphere is measured to be 10 cm, with a maximum error of 0.001 cm. Use an appropriate linear approximation to estimate the maximum error in the calculated volume of the sphere.

D. Taylor and Maclaurin series

1. *Example.* Let $f(x) = e^x$. In the previous section, we defined the linear approximation $T_1(x)$ to $f(x)$ at 0 to be the line tangent to $f(x)$ at 0 — that is, the linear function satisfying $T_1(0) = f(0)$ and $T_1'(0) = f'(0)$. The advantage of considering $T_1(x)$ instead of $f(x)$ is that, as a polynomial, $T_1(x)$ is easy to manipulate and understand.

   We may extend this process by considering a polynomial $T_n(x)$ of degree n, and insisting that the $0^{\text{th}}, 1^{\text{st}}, 2^{\text{nd}}, \ldots, n^{\text{th}}$ derivatives of $T_n(x)$ and $f(x)$ match at 0. Let $T_n(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$. Then we have the following.

$$
\begin{array}{rclcrclcrcl}
T_n(0) & = & a_0 & \text{and} & f(0) & = & 1, & \text{hence} & a_0 & = & 1. \\
T_n'(0) & = & a_1 & \text{and} & f'(0) & = & 1, & \text{hence} & a_1 & = & 1. \\
T_n^{(2)}(0) & = & 2a_2 & \text{and} & f^{(2)}(0) & = & 1, & \text{hence} & a_2 & = & \frac{1}{2!}. \\
T_n^{(3)}(0) & = & 3(2)a_3 & \text{and} & f^{(3)}(0) & = & 1, & \text{hence} & a_3 & = & \frac{1}{3!}. \\
 & \vdots & & & & \vdots & & & & \vdots & \\
T_n^{(n)}(0) & = & n!a_n & \text{and} & f^{(3)}(0) & = & 1, & \text{hence} & a_n & = & \frac{1}{n!}.
\end{array}
$$

Thus $T_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$.

Indeed, we may extend this process even further by considering an "infinite polynomial"; that is, a power series.

2. Let $f(x) = \sum_{n\geq 0} a_n(x-c)^n$. We wish to determine the relationship between the coefficients a_n and the function $f(x)$. Inside its interval of convergence, the power series is infinitely differentiable, with each differentiation yielding another series which is absolutely convergent in the interior of the same interval of convergence. In particular, we have the following.

$$
\begin{array}{rcllrcl}
f(x) & = & a_0 + a_1(x-c) + a_2(x-c)^2 + a_3(x-c)^3 + \cdots, & \text{hence} & f(c) & = & a_0. \\
f'(x) & = & a_1 + 2a_2(x-c) + 3a_3(x-c)^2 + 4a_4(x-c)^3 + \cdots, & \text{hence} & f'(c) & = & 1!a_1. \\
f^{(2)}(x) & = & 2a_2 + 3(2)a_3(x-c) + \cdots, & \text{hence} & f^{(2)}(c) & = & 2!a_2. \\
f^{(3)}(x) & = & 3(2)a_3 + 4(3)(2)a_4(x-c) + \cdots, & \text{hence} & f^{(3)}(c) & = & 3!a_2. \\
 & \vdots & & & & \vdots &
\end{array}
$$

Thus $a_n = \dfrac{f^{(n)}(c)}{n!}$, and $f(x) = \sum_{n\geq 0} \dfrac{f^{(n)}(c)}{n!}(x-c)^n$.

3. *Definition.* The series

$$\sum_{n\geq 0} \frac{f^{(n)}(c)}{n!}(x-c)^n = f(c) + \frac{f^{(1)}(c)}{1!}(x-c) + \frac{f^{(2)}(c)}{2!}(x-c)^2 + \cdots$$

is the *Taylor series of $f(x)$ about c*. If $c = 0$, we also call the series the *Maclaurin series of $f(x)$*.

4. *Example.* e^x has Maclaurin series $\sum_{n\geq 0} \dfrac{x^n}{n!}$. This series converges everywhere.

5. *Definition.* The degree n polynomial obtained by truncating the Taylor series of $f(x)$ about c is called the *degree n Taylor polynomial of $f(x)$ about c*.

6. Note that the degree 1 Taylor polynomial of $f(x)$ about c is simply the linear approximation of $f(x)$ about c. As for the linear approximation, it is useful to be able to estimate the error. Indeed, it is often crucial. In the previous example, the fact that the Maclaurin series of e^x converges everywhere indicates that, for all values of x, the function is equal to its Maclaurin series. In general, it is useful to determine where a function is equal to its Taylor series. One method is to determine the interval of convergence; another method is to show that the error converges to 0 as the degree of the Taylor polynomial increases to ∞.

7. **Taylor's Theorem (with integral remainder).** Let $f(x)$ be such that $f(x), f'(x), f''(x), \ldots, f^{(n+1)}(x)$ exist, and are continuous, on an interval containing c. Then for any x in that interval,

$$f(x) = f(c) + \frac{f'(c)}{1!}(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n + \frac{1}{n!}\int_c^x f^{(n+1)}(t)(x-t)^n dt.$$

Proof. The proof is left as an exercise. (The theorem may be proven by induction on n, using integration by parts in the inductive case.)

While this theorem gives an exact formula for the error, it is in practice often impossible to calculate. The following theorem, which generalizes the theorem in Section B giving the error for linear approximations, is more practical.

8. **Taylor's Theorem (with Lagrange remainder).** Let $f(x)$ be such that $f(x), f'(x), f''(x), \ldots, f^{(n+1)}(x)$ exist, and are continuous, on an interval containing c. Then for any x in that interval, there exists a number s between c and x such that

$$f(x) = f(c) + \frac{f'(c)}{1!}(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n + \frac{f^{(n+1)}(s)}{(n+1)!}(x-c)^{n+1}.$$

According to this theorem, the error is "roughly no worse than the next term in the Taylor series".

Proof. We use induction on n. The $n = 0$ case is exactly the Mean Value Theorem. (Though it is not necessary to prove another base case, we note that the $n = 1$ case was proven in section B.) Assume the theorem is true for $n = k-1$; that is, assume that, if $f(x)$ is a function such that $f(x), f'(x), f''(x), \ldots, f^{(k)}(x)$ exist, and are continuous, on an interval containing c, then there exists a number s between c and x such that the error for the $(k-1)^{\text{th}}$ Taylor polynomial is equal to

$$\frac{f^{(k)}(s)}{k!}(x-c)^k.$$

Assume $x > c$ (the proof if $x < c$ is similar). We apply the Generalized Mean Value Theorem to the functions $E(t) = f(t) - T_k(t)$ and $(t-c)^{k+1}$ on the interval $[c, x]$, and conclude that there exists a number a in (c, x) such that

$$\frac{E'(a)}{(k+1)(a-c)^k} = \frac{E(x) - E(c)}{(x-c)^{k+1} - (c-c)^{k+1}} = \frac{E(x)}{(x-c)^{k+1}}. \tag{16}$$

Now

$$\begin{aligned}
E(t) &= f(t) - \left(f(c) + \frac{f'(c)}{1!}(t-c) + \frac{f''(c)}{2!}(t-c)^2 + \cdots + \frac{f^{(k)}(c)}{k!}(t-c)^k\right), \\
E'(t) &= f'(t) - \left(f'(c) + f''(c)(t-c) + \frac{f^{(3)}(c)}{2!}(t-c)^2 + \cdots + \frac{f^{(k)}(c)}{(k-1)!}(t-c)^{k-1}\right) \\
E'(a) &= f'(a) - \left(f'(c) + f''(c)(a-c) + \frac{f^{(3)}(c)}{2!}(a-c)^2 + \cdots + \frac{f^{(k)}(c)}{(k-1)!}(a-c)^{k-1}\right).
\end{aligned}$$

This last expression is simply equal to the error for the degree $k-1$ Taylor polynomial for the function $f'(t)$ about c, evaluated at a. By the assumption, there exists a number s in (c, a) — that is, in (c, x) — such that $E'(a)$ is equal to

$$\frac{(f')^{(k)}(s)}{k!}(a-c)^k = \frac{f^{(k+1)}(s)}{k!}(a-c)^k.$$

Plugging this back into (16) yields

$$E(x) = \frac{f^{(k+1)}(s)}{(k+1)!}(x-c)^{k+1}$$

as needed. $\square$

9. *Example.* $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$ for all x.

Note that, in the example earlier in this section, we showed that the series on the right-hand side converges for all x; but *we did not show that it converges to e^x.*

To show this second fact, we use Taylor's Theorem (with Lagrange remainder). The degree n Maclaurin polynomial of e^x is $T_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$. The error $E_n(x) = e^x - T_n(x)$ satisfies

$$|E_n(x)| = \frac{e^s}{(n+1)!}|x|^{n+1} = e^s \frac{|x|^{n+1}}{(n+1)!}$$

for some number s between 0 and x. We wish to show that this error vanishes as n approaches ∞. Now since by the Ratio Test the series $\displaystyle\sum_{n\geq 0} \frac{|x|^n}{n!}$ converges for all x, we conclude by the Divergence Test that $\displaystyle\lim_{n\to\infty} \frac{|x|^n}{n!} = 0$; that is, $\displaystyle\lim_{n\to\infty} |E_n| = 0$.

E. Exercises

1. Determine whether each of the following statements is true or false, and justify your answer.
   (a) *A function may have a Taylor series representation but not a Maclaurin series representation.*
   (b) *A function which has a Taylor series representation is infinitely differentiable.*
   (c) *Given the same centres, a higher degree Taylor polynomial is a more accurate approximation to a function than a lower degree Taylor polynomial.*
2. Find the Maclaurin series and radius of convergence for each of the following functions.
   (a) $\sin(x)$.
   (b) $\cos(x)$.
   (c) e^x.
   (d) $\log(1+x)$.
   (e) $\arctan(x)$.
3. Find the Taylor series for each of the following functions about the given centre. Then determine the radius of convergence.
   (a) $\sin(x)$ about $\frac{\pi}{2}$.
   (b) e^{3x} about 1.
   (c) $\sqrt{x}$ about 25.
4. Find the Maclaurin series for the *error function* $\operatorname{erf}(x) = \frac{1}{\pi}\displaystyle\int_{-x}^{x} e^{-t^2}\,dt$. Why do you think the Maclaurin series is particularly useful for this function?
5. Find the Taylor series for $x\log(x)$ about 1.

6. Approximate $e^{0.1}$ using a degree 3 approximation, and bound the size of the error. What degree approximation is necessary to get within 10^{-10} of the actual value of $e^{0.1}$?
7. Approximate $\sin(-0.1)$ using a degree 5 approximation, and bound the size of the error. What degree approximation is necessary to get within 10^{-10} of the actual value of $\sin(-0.1)$?
8. Calculate $\lim\limits_{x \to 0} \dfrac{\left(e^{2x} - 1\right) \log\left(1 + x^2\right)}{\left(1 - \cos(3x)\right)^2}$.

Differential equations

A. Motivation

In ideal conditions, we expect the rate of growth of a population P to be proportional to P; that is, we expect

$$\frac{dP}{dt} = rP \tag{17}$$

where r is a positive constant describing the "growth rate".

However, conditions are generally not ideal. For example, the population's environment may have limited resources, and thus impose a limit on the size of the population that it may sustain. The *logistic differential equation* is one attempt at modelling this:

$$\frac{dP}{dt} = rP\left(1 - \frac{P}{K}\right) \tag{18}$$

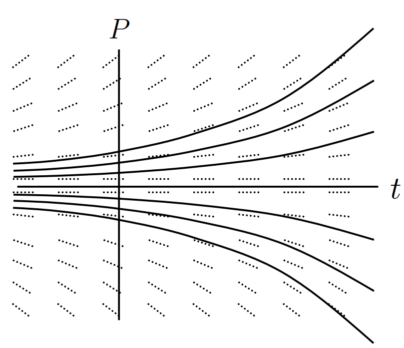
where r and K are positive constants, with K describing the environment's "carrying capacity".

(17) and (18) are examples of *differential equations*, equations relating functions and their derivatives. In this final section of the course, we explore the application of calculus to understanding or even solving differential equations. Differential equations exemplify the usefulness of calculus in describing and predicting the behaviour of dynamical systems. They are perhaps the most important application of the subject.

B. Direction fields and phase portraits

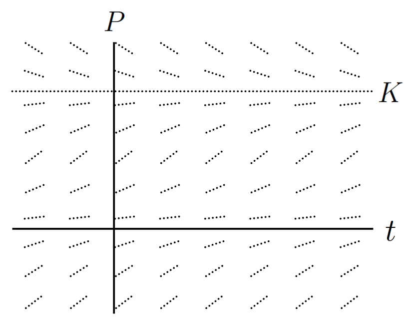
1. *Example.* Consider (17). We know already that the solutions are functions of the form ae^{rt} where a is a constant — in fact, one way of defining the constant e is via this differential equation.

   Recall how we defined the constant e last term. We considered the differential equation $\frac{dP}{dt} = P$. The slope of the line tangent to the graph of a solution at any particular y-value is equal to the y-value itself. For example, if the graph crosses the line $y = 1$, its tangent line at that point has slope 1. We drew a collection of these tangent lines in a figure called a *direction field*, and sketched the solutions "guided" by these tangent line segments.

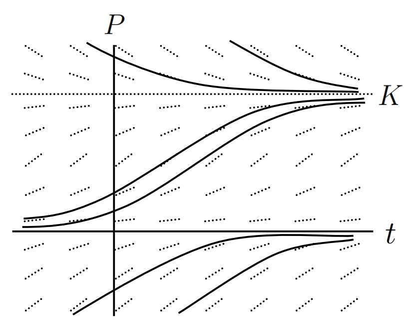
   

   The direction field does not yield analytic expressions of solutions, but it gives us a lot of information about solutions. For example, other than the trivial solution $P = 0$, we can see that all other solutions diverge to ∞ or $-\infty$ as $t \to \infty$. We can see that positive solutions are positive for all t, and negative solutions are negative for all t.

2. *Example.* Consider (18). There are two trivial solutions: $\frac{dP}{dt} = 0$ when $P = 0$ or $P = K$. Indeed, $\frac{dP}{dt}$ will be "close to" 0 the closer P is to one of these two trivial solutions. It remains only to observe that $\frac{dP}{dt} > 0$ when $P < 0$ (a physically impossible scenario) and $0 < P < K$, and $\frac{dP}{dt} < 0$ when $P > K$. The direction field is below.

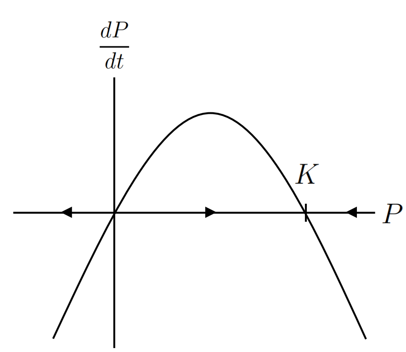


Some trajectories may be drawn as follows.



From the direction field, we infer that populations modelled by (18) move toward a stable equilibrium K. When a population is below the carrying capacity, the rate of growth also appears to be roughly exponential when the population is small.

3. *Example.* Much of the information inferred from a direction field may also be inferred from a *phase portrait* plotting a function's first derivative against itself. In the case of (18), we plot $\frac{dP}{dt}$ against P.



From the phase portrait, we may recover the information that $\frac{dP}{dt} = 0$ — that is, P is constant — when $P = 0$ or $P = K$; that $\frac{dP}{dt} > 0$ — that is, P is increasing — when $0 < P < K$; and that P is decreasing everywhere else. We also see that $\frac{dP}{dt}$ is maximal at $P = \frac{K}{2}$.

4. Whether we use a direction field, phase portrait, or one of the analytic methods demonstrated in the following sections, it is useful to check our information inferred from the differential equation against our intuition about the physical system being modelled, or against available data. In the case of (18), it does indeed make sense that a population above the carrying capacity of an environment should decrease, and that a population below the carrying capacity should increase.

C. Exercises

1. Sketch the direction field and some sample trajectories for $\dfrac{dy}{dx} = \sin(x)\sin(y)$.
2. Sketch the direction field and some sample trajectories for $\dfrac{dy}{dx} = x - y$. (Hint: begin by drawing some *isoclines*, lines where the derivative is constant.)
3. Sketch the direction field for $\dfrac{dy}{dx} = y + xy$. Describe in one or two sentences the solution that passes through the point $(-1, 1)$.
4. (a) Sketch the direction field and some sample trajectories for $\dfrac{dy}{dx} = (y-1)(y^2 - y - 2)$.
   (b) Sketch the phase portrait for the differential equation in part (a), and explain how the information from the direction field is recovered from the phase portrait.

D. Separation of variables

1. *Example.* Consider (18). We have

$$\frac{dP}{dt}\left(\frac{1}{P(1-P/K)}\right) = \frac{dP}{dt}\left(\frac{K}{P(K-P)}\right) = \frac{dP}{dt}\left(\frac{1}{P} + \frac{1}{K-P}\right) = r.$$

At this point, we treat "dP" and "dt" *as if they are functions*, and write

$$\left(\frac{1}{P} + \frac{1}{K-P}\right) dP = r\,dt.$$

We then take the antiderivative of both sides, indicating this by writing

$$\int \left(\frac{1}{P} + \frac{1}{K-P}\right) dP = \int r\,dt. \tag{19}$$

The antiderivative notation is troubling — it conflates antidifferentiation with integration — but useful in this context. The treatment of "dP" and "dt" as functions is actually an unjustified abuse of notation, but it is also useful, and we justify it below. We get

$$\begin{aligned} \log|P| - \log|K-P| &= rt + C \\ \log\left|\frac{K-P}{P}\right| &= -rt - C \\ \frac{K-P}{P} &= Ae^{-rt}, \end{aligned}$$

where $A = \pm e^{-C}$. It follows that $P = \frac{K}{1+Ae^{-rt}}$ where $A = \frac{K-P_0}{P_0}$, with P_0 denoting the initial population at $t = 0$.

This method of moving all of one of the variables — in this case P — to one side, and all of the other — in this case t — to the other side, and then antidifferentiating, is called *separation of variables*.

2. The method, and in particular the abuse of notation, may be justified using differentials. We may also use the Chain Rule: suppose we start with (19); then

$$\begin{aligned} \frac{d}{dt}\int\left(\frac{1}{P}+\frac{1}{K-P}\right)dP &= \frac{d}{dt}\int r\,dt \\ \frac{d}{dP}\left(\int\left(\frac{1}{P}+\frac{1}{K-P}\right)dP\right)\frac{dP}{dt} &= r \\ \left(\frac{1}{P}+\frac{1}{K-P}\right)\frac{dP}{dt} &= r, \end{aligned}$$

from which we recover the original differential equation (18). Finally, we note that the antiderivatives resulting from the abuse of notation may also be differentiated to justify the manipulation directly.

3. *Definition.* A *separable* differential equation is of the form

$$\frac{dy}{dx} = f(x)g(y).$$

It may be solved by separating the variables and then antidifferentiating: $\displaystyle\int \frac{dy}{g(y)} = \int f(x)dx.$

E. Exercises

1. Solve the following differential equations, given the initial conditions listed.
   (a) $\dfrac{dy}{dx} = \dfrac{\log(x)}{xy}$, $y(1) = 3$.
   (b) $\dfrac{dy}{dx} = e^y \sin(x)$, $y(0) = 1$.
   (c) $\dfrac{dy}{dx} = \sqrt{xy}$, $y(1) = 2$.
2. Solve the following integral equations.
   (a) $f(x) = 1 + \displaystyle\int_1^x t f(t) dt$.
   (b) $f(x) = 1 + \displaystyle\int_1^x t\,(1 - f(t))\,dt$.
   (c) $f(x) = 1 + \displaystyle\int_0^x \frac{f(t)^2}{1+t^2}\,dt$.
3. Find an infinite number of curves that intersect at right angles all ellipses of the form $x^2 + 2y^2 = a^2$, where a is a constant.
4. Suppose a reservoir contains 10000 L of polluted water, in which is dissolved 50 kg of pollutant. Suppose polluted water flows into the tank at a rate of 10 L/min, but each litre flowing in contains 2 grams of pollutant. The reservoir is constantly and thoroughly mixed, and the solution flows out at a rate of 10 L/min.

(a) Find a function describing the amount of pollutant in the reservoir at time t.

(b) Describe what happens as t gets very large. Does your mathematical model conform with your physical intuition?

(c) Repeat parts (a) and (b), assuming that the water flowing into the reservoir is pure.

F. Integrating factors

1. *Example.* Near sea level, the Earth exerts a downward force of mg on an object of mass m, where g is approximately $9.8\,\mathrm{m/s^2}$. If the object is dropped from rest, it encounters the force of air resistance proportional to its velocity v. We have

$$m\frac{dv}{dt} = mg - kv;$$

that is,

$$\frac{dv}{dt} + \frac{k}{m}v = g. \tag{20}$$

This is an example of a *linear differential equation.* We aim to antidifferentiate it with respect to t, and thus find an expression for v. Therefore it is helpful to be able to identify if any part of the equation is a known derivative.

The expression on the left-hand side of (20) resembles the derivative of a product of two functions: v and an unknown function. Our next step is to "force" the left-hand side to be exactly the derivative of a product of two functions. We wish to find a function I such that

$$\begin{aligned} \frac{d}{dt}(vI) &= \left(\frac{dv}{dt} + \frac{k}{m}v\right) I \\ \frac{dv}{dt}I + \frac{dI}{dt}v &= \frac{dv}{dt}I + \frac{k}{m}vI \\ \frac{dI}{dt} &= \frac{k}{m}I. \end{aligned}$$

At this point we have a separable differential equation. We get

$$\int \frac{dI}{I} = \int \frac{k}{m}\, dt.$$

One solution is

$$I = e^{\int \frac{k}{m}\, dt} = e^{kt/m}, \tag{21}$$

which we call the *integrating factor* of (20). Multiplying (20) by this factor yields

$$\begin{aligned} e^{kt/m}\frac{dv}{dt} + e^{kt/m}\frac{k}{m}v &= e^{kt/m}g \\ \int \left(\frac{d}{dt}\left(ve^{kt/m}\right)\right) dt &= \int \left(e^{kt/m}g\right) dt \\ ve^{kt/m} &= \frac{mg}{k}e^{kt/m} + c \end{aligned}$$

for some constant c; that is,

$$v = \frac{mg}{k} + \frac{c}{e^{kt/m}}.$$

We solve for c by recalling that the object is dropped from rest; that is, $v(0) = 0$, whence $c = -\frac{mg}{k}$ and

$$v = \frac{mg}{k}\left(1 - \frac{1}{e^{kt/m}}\right).$$

2. *Definition.* In general, given a differential equation of the form

$$\frac{dy}{dx} + P(x)y = Q(x),$$

we multiply both sides by the *integrating factor* $e^{\int P(x)dx}$ and then antidifferentiate.

We may justify this technique by an argument parallel to the example above: we have

$$\frac{dy}{dx}e^{\int P(x)dx} + P(x)ye^{\int P(x)dx} = \frac{d}{dx}\left(ye^{\int P(x)dx}\right),$$

which is easy to antidifferentiate by design. Note that there are no guarantees on whether the term on the right-hand side, $Q(x)e^{\int P(x)dx}$, is likewise antidifferentiable.

G. Exercises

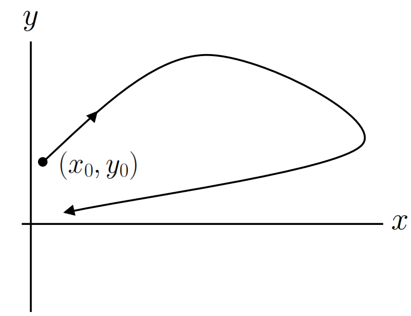
1. Solve the following differential equations, given the initial conditions listed.
   (a) $\frac{dy}{dx} - \frac{2y}{x} = 3x^2$, $y(1) = 1$.
   (b) $\frac{dy}{dx} + 2y = 4$, $y(0) = 2$.
   (c) $\frac{dy}{dx} - y = x$, $y(0) = 3$.
2. Solve the differential equation $x\frac{dy}{dx} + y + xy^2 = 0$. (Hint: use the substitution $u = \frac{1}{y}$, proposed by Bernoulli, to transform the equation into a linear differential equation.)
3. Suppose a reservoir contains 10000 L of polluted water, in which is dissolved 50 kg of pollutant. Suppose polluted water flows into the tank at a rate of 10 L/min, but each litre flowing in contains 2 g of pollutant. The reservoir is constantly and thoroughly mixed, and the solution flows out at a rate of 6 L/min. Find a function describing the amount of pollutant in the reservoir at time t.
4. In the first example of the previous section, it is assumed that air resistance is proportional to velocity. Suppose it is proportional to the square of velocity. In that case, what is the velocity of the object at time t?

H. Coupled differential equations

1. *Coupled differential equations.* Suppose we have two populations, x and y, whose rate of growth depends not only on x or y respectively, but on both. For instance, the populations might represent predators and prey, or parasites and hosts. The relationship between the populations and their rates of growth might be modelled by a *system of linear differential equations*, coupled differential equations of the form

$$\begin{aligned} \tfrac{dx}{dt} &= ax + by, \\ \tfrac{dy}{dt} &= cx + dy \end{aligned}$$

where a, b, c and d are constants. Given such a system, we wish to be able to graph trajectories on the (x, y)-plane. These trajectories indicate the relationship between x and y at any particular time, given initial populations x_0 and y_0. For example, in the trajectory below, we see that, starting with a small initial population x_0 compared to a larger initial population y_0, both populations initially increase; then y decreases while x continues to increase; and finally both x and y decrease.



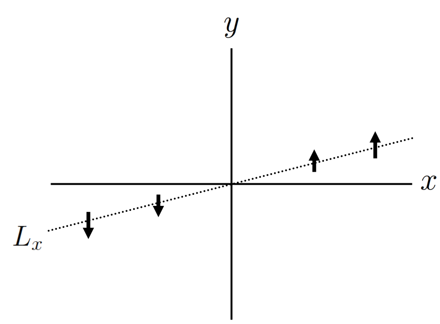
2. *Example.* Consider the system

$$\begin{aligned} \frac{dx}{dt} &= 3x - 12y, \\ \frac{dy}{dt} &= 5x + y. \end{aligned} \tag{22}$$

We may extract some information from this system of equations by considering when derivatives vanish. For example, $\frac{dx}{dt} = 0$ exactly on the line

$$L_x : y = \frac{1}{4}x \tag{23}$$

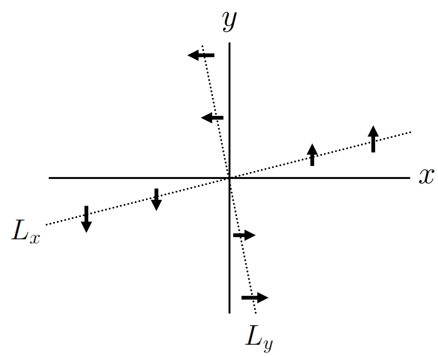
— this line is called the *x-nullcline*. If we sketch the (x, y)-plane and ask what the trajectory from a given initial condition (x_0, y_0) looks like, we observe that any trajectory must cross L_x vertically, since along that line, $\frac{dx}{dt} = 0$. Indeed, plugging (23) back into (22) indicates that, on L_x, we have $\frac{dy}{dt} = 5x + \frac{1}{4}x = \frac{21}{4}x$. In other words, $\frac{dy}{dt}$ is the same sign as x, and the trajectories cross L_x as indicated below.



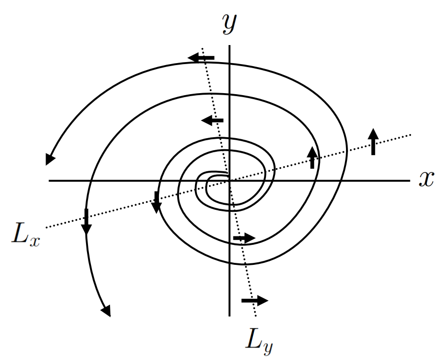
Similarly, we find that the y-nullcline is

$$L_y : y = -5x,$$

and that $\frac{dx}{dt} = 63x$ on L_y.



Beyond this, our analysis is somewhat improvised, but some sample trajectories are shown below.

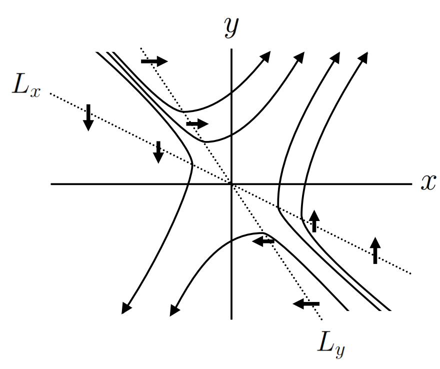


It is reasonable to conclude that the trajectories circle the origin. To conclude rigorously that they spiral outward (as opposed to inward or neither) requires tools beyond the scope of this course.

3. *Example.* Suppose we have the following system of equations.

$$\begin{aligned}\tfrac{dx}{dt} &= x+2y,\\ \tfrac{dy}{dt} &= 3x+2y.\end{aligned}$$

We may perform an analysis similar to the one in the previous example. We find that the x-nullcline is $L_x : y = -\frac{1}{2}x$, with $\frac{dy}{dt} = 2x$ on L_x; and that the y-nullcline is $L_y = -\frac{3}{2}x$, with $\frac{dx}{dt} = -2x$ on L_y. The nullclines and some trajectories are pictured below.



4. A number of interesting questions arise naturally from this improvised analysis. Are there any linear trajectories? Why do some systems lead to spiralling trajectories and others to trajectories of the sort in the second example? Is a similar analysis possible with more than two linear differential equations, or with nonlinear differential equations?

   In order to answer these questions, it is helpful to know some linear algebra. Linear algebra, and algebra in general, is in some sense the natural "companion" to calculus, and to analysis in general. It is appropriate that at the end of this course, we find ourselves on the border of another course and another subject.

I. Exercises

1. Dead trees decay naturally from seasonal and biological events. Decayed trees eventually become humus, which itself decays, but at a slower rate. Let x denote the biomass of humus in a forest at time t, and y denote the biomass of dead trees at time t. Then x and y satisfy the following system of differential equations:

$$\begin{aligned} \tfrac{dx}{dt} &= -x + 3y, \\ \tfrac{dy}{dt} &= -3y. \end{aligned}$$

   Sketch the nullclines and a few sample trajectories.
2. Sketch the nullclines and a few sample trajectories for the following system of differential equations.

$$\begin{aligned} \tfrac{dx}{dt} &= -5x + y, \\ \tfrac{dy}{dt} &= 4x - 2y. \end{aligned}$$

3. Sketch the nullclines and a few sample trajectories for the following system of differential equations.

$$\begin{aligned} \tfrac{dx}{dt} &= 7x + y, \\ \tfrac{dy}{dt} &= -4x + 3y. \end{aligned}$$

4. The following system of differential equations, called a *FitzHugh-Nagumo model*, describes the interaction between the membrane potential and action potential of a spiking neuron. Here, v represents the membrane potential and w represents the sodium gating variable:

$$\begin{aligned} \tfrac{dv}{dt} &= f(v + 0.07) - f(0.07) - w, \\ \tfrac{dw}{dt} &= 0.008(v - 2.54w) \end{aligned}$$

   where $f(v) = v(v - 0.139)(1 - v)$. Sketch the nullclines and a few sample trajectories.