Calculus

Prof. Robert Ghrist

11/15/16

Contents

Part I

Single Variable

Chapter 1

Functions

1.1 Functions

A *function* can be visualized as a machine that takes in an input x and returns an output $f(x)$. The collection of all possible inputs is called the *domain*, and the collection of all possible outputs is called the *range*.

This course deals with functions whose domains and ranges are $\mathbb{R}$ or subsets of $\mathbb{R}$ (this is the notation for the real numbers).

Examples

1. Polynomials, e.g. $f(x) = x^3 - 5x^2 + x + 9$. Give the domain and range of f.
   Answer: The domain is $\mathbb{R}$, because we can plug in any real number into a polynomial. The range is $\mathbb{R}$, which we see by noting that this is a cubic function, so as $x \to -\infty$, $f(x) \to -\infty$, and as $x \to \infty$, $f(x) \to \infty$.

2. Trigonometric functions, e.g. $\sin$, $\cos$, $\tan$. Give the domain and range for each of these.
   Answer: For $\sin$ and $\cos$: domain is $\mathbb{R}$; range is $[-1, 1]$. For $\tan$, the domain is $\{x \in \mathbb{R} : x \neq \frac{\pi}{2} + k\pi\}$; range is $\mathbb{R}$.

3. The exponential function, e^x. Give the domain and range for the exponential.
   Answer: Domain is $\mathbb{R}$; range is $((0, \infty)$.

4. The natural logarithm function, $\ln x$. Recall that this is the inverse of the exponential function. Give the domain and range for $\ln x$.
   Answer: Domain is $(0, \infty)$; range is $\mathbb{R}$. Notice how the domain and range of the exponential relate to the domain and range of the natural logarithm.

5. Is $\sin^{-1}$ a function? If so, why? If not, is there a way to make it into a function?
   Answer: $\sin^{-1}$ is not a function, because one input has many outputs. For example, $\sin^{-1}(0) = 0, \pi, 2\pi, \ldots$. By restricting the range of $\sin^{-1}$ to $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, one gets the function $\arcsin$.

1.1.1 Operations on Functions

Composition

The *composition* of two functions, f and g, is defined to be the function that takes as its input x and returns as its output $g(x)$ fed into f.

$$f \circ g(x) = f(g(x))$$

Example:
$\sqrt{1 - x^2}$ can be thought of as the composition of two functions, f and g. If $g = 1 - x^2$, f would be the function that takes an input $g(x)$ and returns its square root.

Example:

Compute the composition $f \circ f$, i.e. the composition of f with itself, where $f(x) = \dfrac{1}{x+1}$.

Answer:

We find that

$$f \circ f(x) = f(f(x))$$
$$= f\left(\frac{1}{x+1}\right)$$
$$= \frac{1}{1/(x+1)+1}$$
$$= \frac{x+1}{1+x+1}$$
$$= \frac{x+1}{x+2}.$$

Inverse

The *inverse* is the function that undoes f. If you plug $f(x)$ into f^{-1}, you will get x. Notice that this function works both ways. If you plug $f^{-1}(x)$ into $f(x)$, you will get back x again.

$$f^{-1}(f(x)) = x$$

$$f(f^{-1}(x)) = x$$

NOTE: f^{-1} denotes the inverse, not the reciprocal. $f^{-1}(x) \neq \frac{1}{f(x)}$.

Example:

Let's consider $f(x) = x^3$. Its inverse is $f^{-1}(x) = x^{\frac{1}{3}}$.

$$f^{-1}(f(x)) = (x^3)^{\frac{1}{3}} = x$$

$$f(f^{-1}(x)) = (x^{\frac{1}{3}})^3 = x$$

Notice that the graphs of f and f^{-1} are always going to be symmetric about the line $y = x$. That is the line where the input and the output are the same.

1.1.2   Classes of Functions

Polynomials

A polynomial $P(x)$ is a function of the form

$$P(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n$$

The top power n is called the degree of the polynomial. We can also write a polynomial using a summation notation.

$$P(x) = \sum_{k=0}^{n} c_k x^k$$

Rational functions

Rational functions are functions of the form $\dfrac{P(x)}{Q(x)}$ where each is a polynomial.

Example:

$$\frac{3x-1}{x^2+x-6}$$

is a rational function. You have to be careful of the denominator. When the denominator takes a value of zero, the function may not be well-defined.

Powers

Power functions are functions of the form cx^n, where c and n are constant real numbers. Other powers besides those of positive integers are useful.

Example:

What is x^0? Answer: $x^0 = 1$

What is $x^{-\frac{1}{2}}$? Answer: Recall a fractional power denotes root. For example, $x^{\frac{1}{2}} = \sqrt{x}$. The negative sign in the exponent means that we take the reciprocal. So, $x^{-\frac{1}{2}} = \frac{1}{\sqrt{x}}$.

What is $x^{\frac{22}{7}}$? Answer: One can rewrite this as $\left(x^{22}\right)^{1/7}$. That means we take x to the 22nd power and then take the 7th root of the result. $x^{\frac{22}{7}} = \sqrt[7]{x^{22}}$

What is x^π? We are not yet equipped to handle this, but we will come back to it later.

Trigonometrics

You should be familiar with the basic trigonometric functions **sin**, **cos**. One fact to keep in mind is $\cos^2 \theta + \sin^2 \theta = 1$ for any θ. This is known as a *Pythagorean identity*, which is so named because of one of the ways to prove it:

$$\cos^2 \theta + \sin^2 \theta = 1$$
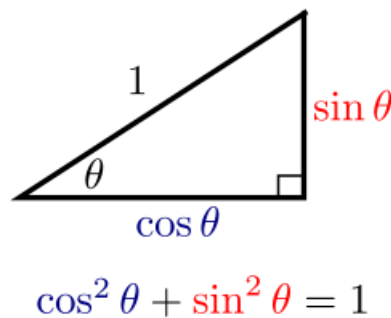
By looking at a right triangle with hypotenuse 1 and angle θ, and labeling the adjacent and opposite sides accordingly, one finds by using Pythagoras' Theorem that $\cos^2 \theta + \sin^2 \theta = 1$.

Another way to think about it is to embed the above triangle into a diagram for the unit circle where we see that $\cos \theta$ and $\sin \theta$ returns the x and y coordinates, respectively, of a point on the unit circle with angle θ to the x-axis:
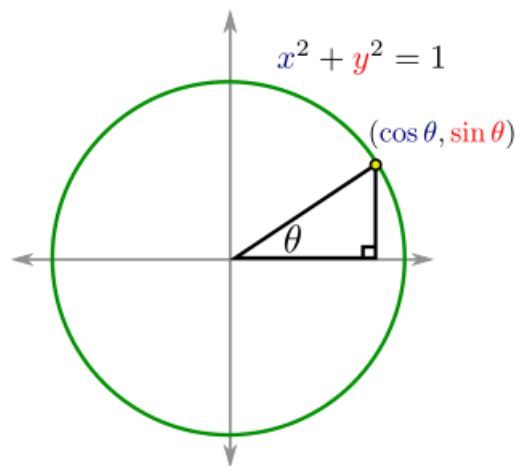
That explains the nature of the formula $\cos^2 \theta + \sin^2 \theta = 1$. It comes from the equation of the unit circle $x^2 + y^2 = 1$.

Others trigonometric functions:

$$\tan = \frac{\sin}{\cos}$$

$$\cot = \frac{\cos}{\sin} \text{ , the reciprocal of } \tan$$

$$\sec = \frac{1}{\cos} \text{ , the reciprocal of the } \cos$$

$$\mathbf{csc} = \frac{1}{\mathbf{sin}}$$, the reciprocal of the **sin**

All four of these have vertical asymptotes at the points where the denominator goes to zero.

Inverse Trigonometrics

We often write $\mathbf{sin}^{-1}$ to denote the inverse, but this can cause confusion. Be careful that $\mathbf{sin}^{-1} \neq \frac{1}{\mathbf{sin}}$.
To avoid the confusion, the terminology **arcsin** is recommended for the inverse of the **sin** function.

The **arcsin** function takes on values $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and has a restricted domain $[-1, 1]$.

The **arccos** function likewise has a restricted domain $[-1, 1]$, but it takes values $[0, \pi]$.

The **arctan** function has an unbounded domain, it is well defined for all inputs. But it has a restricted range $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

Exponentials

Exponential functions are of the form $\mathbf{c}^{\mathbf{x}}$, where $\mathbf{c}$ is some positive constant. The most common such function, referred to as *the* exponential, is $\mathbf{e}^{\mathbf{x}}$. This is the most common because of its nice integral and differential properties (below).

Algebraic properties of the exponential function:

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

$$(e^x)^y = e^{xy}$$

Differential/integral properties:

$$\frac{d}{dx} e^x = e^x$$

$$\int e^x dx = e^x + C$$

Recall the graph of $\mathbf{e}^{\mathbf{x}}$, plotted here alongside its inverse, $\mathbf{ln}\,\mathbf{x}$:
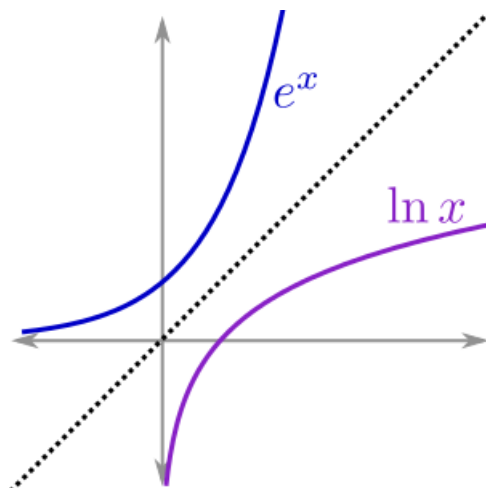
Note that the graphs are symmetric about the line $\mathbf{y} = \mathbf{x}$ (as is true of the graphs of a function and its inverse). Before continuing, one might ask, what is $\mathbf{e}$? There are several ways to define $\mathbf{e}$, which will be revealed soon. For now, it is an irrational number which is approximately 2.718281828.

Euler's Formula

To close this lesson, we give a wonderful formula, which for now we will just take as a fact:

Euler's Formula

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

The i in the exponent is the imaginary number $\sqrt{-1}$. It has the properties $i^2 = -1$. i is not a real number. That doesn't mean that it doesn't exist. It just means it is not on a real number line.

Euler's formula concerns the exponentiation of an imaginary variable. What exactly does that mean? How is this related to trigonometric functions? This will be covered in our next lesson.

Additional Examples

Example Find the domain of
$$f(x) = \frac{1}{\sqrt{x^2 - 3x + 2}}.$$

Note that the square root is only defined when its input is non-negative. Also, the denominator in a rational function cannot be 0. So we find that this function is well-defined if and only if $x^2 - 3x + 2 > 0$. Factoring gives
$$(x - 2)(x - 1) > 0.$$

By plotting the points $x = 1$ and $x = 2$ (where the denominator equals 0) and testing points between them, one finds that $x^2 - 3x + 2 > 0$ when $x < 1$ or $x > 2$:
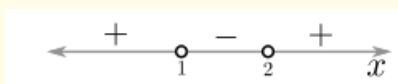
So the domain of f is $x < 1$ or $2 < x$. In interval notation, this is $(-\infty, 1) \cup (2, \infty)$.

Example Find the domain of
$$f(x) = \ln(x^3 - 6x^2 + 8x).$$

Since $\ln$ is only defined on the positive real numbers, we must have $x^3 - 6x^2 + 8x > 0$. Factoring gives
$$x(x^2 - 6x + 8) = x(x - 2)(x - 4) > 0$$

As in the above example, plotting the points where this equals 0 and then testing points, we find that the domain is $0 < x < 2$ and $4 < x$. In interval notation, this is $(0, 2) \cup (4, \infty)$.

1.2   The Exponential

This module deals with a very important function: the exponential. The first question one might ask is: what is the exponential function e^x? We know certain values of the function such as $e^0 = 1$, but what about an irrational input such as e^π, or an imaginary input e^i? Is it possible to make sense of these values?

The following definition answers these questions.

The Exponential e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \qquad = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

where
$$k! = k(k - 1)(k - 2)\cdots 3 \cdot 2 \cdot 1,$$

and $0! = 1$

One can now plug values for x into the above sum to compute e^x. When $x = 0$, for instance, one finds that $e^0 = 1$, (since all the terms with x disappear) as expected. By plugging in $x = 1$, the true value of e is found to be $e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots$.

1.2.1   A long polynomial

There are technical concerns when trying to add up an infinite number of things. These issues will be dealt with later in the modules on series. For now, treat the infinite sum above as a long polynomial (the actual term is the Taylor series about $x = 0$, which will be more formally dealt with in the next module). Polynomials are nice because they are easy to integrate and differentiate. Recall the power rule for differentiating and integrating a monomial x^k, where k is a constant:

$$\frac{d}{dx}x^k = kx^{k-1} \int x^k \, dx \qquad\qquad = \frac{1}{k+1}x^{k+1} + C \quad (k \neq -1)$$

1.2.2   Properties of e^x

Recall the following properties of the exponential function:

1. $e^{x+y} = e^x e^y$

2. $e^{x \cdot y} = (e^x)^y = (e^y)^x$

3. $\frac{d}{dx}e^x = e^x$

4. $\int e^x dx = e^x + C$.

Consider the last two properties in terms of the long polynomial. Taking the derivative of the long polynomial for e^x gives

$$\frac{d}{dx}(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots) = 0 + 1 + \frac{2x}{2!} + \frac{3x^2}{3!} + \frac{4x^3}{4!} + \cdots$$
$$= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

which is the original long polynomial. Integrating also gives (up to the constant of integration) the original long polynomial. This agrees with facts about the derivative and integral of e^x. Thus, the long polynomial for e^x captures two of the key features of e^x; namely, e^x is its own derivative and its own integral.

1.2.3   Euler's formula

Recall that the imaginary number i is defined by $i = \sqrt{-1}$. So $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, and this continues cyclically (for a review of complex/imaginary numbers, see wikipedia). Assume the following fact, known as Euler's formula, mentioned in the last module.

> Euler's formula
> $$e^{ix} = \cos x + i \sin x.$$

Consider what happens when ix is plugged into the long polynomial for e^x. By simplifying the powers of i, and grouping the result into its real and imaginary parts, one finds

$$e^{ix} = 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \cdots$$
$$= 1 + ix + \frac{i^2 x^2}{2!} + \frac{i^3 x^3}{3!} + \cdots$$
$$= 1 + ix - \frac{x^2}{2!} - i\frac{x^3}{3!} + \frac{x^4}{4!} + i\frac{x^5}{5!} + \cdots$$
$$= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\right).$$

If this is supposed to equal $\cos x + i \sin x$, then the real part must be $\cos x$, and the imaginary part must be $\sin x$. It follows that

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k}}{(2k)!}$$
$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

These formulas should be memorized, both in their long polynomial form and their more concise summation notation form.

Example Compute $1 - \frac{\pi^2}{2!} + \frac{\pi^4}{4!} - \cdots$.

> Note that this is the long polynomial for $\cos x$, evaluated at $x = \pi$. So the value is $\cos \pi = -1$.

Example Check that taking the derivative of the long polynomial for $\sin x$ gives the long polynomial for $\cos x$ (hence, verify that $\frac{d}{dx} \sin x = \cos x$.

> Computing the derivative term by term gives
>
> $$\frac{d}{dx} \sin(x) = \frac{d}{dx} \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots \right)$$
> $$= 1 - 3\frac{x^2}{3!} + 5\frac{x^4}{5!} - \cdots$$
> $$= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots,$$
>
> which is the long polynomial for $\cos x$, as desired.

Example Show that the long polynomial for e^x satisfies the first property above, namely $e^{x+y} = e^x e^y$. Hint: start with the long polynomials for e^x and e^y and multiply these together, and carefully collect like terms to show it equals the long polynomial for e^{x+y}.

> Beginning with $e^x \cdot e^y$, we find
>
> $$e^x \cdot e^y = \left(1 + x + \frac{x^2}{2!} + \cdots \right) \left(1 + y + \frac{y^2}{2!} + \cdots \right)$$
> $$= 1 + (x + y) + \left(\frac{x^2}{2!} + xy + \frac{y^2}{2!} \right) + \cdots$$
> $$= 1 + (x + y) + \frac{x^2 + 2xy + y^2}{2!} + \cdots$$
> $$= 1 + (x + y) + \frac{(x + y)^2}{2!} + \cdots,$$
>
> which is the long polynomial for e^{x+y}, as desired.

1.2.4 More on the long polynomial

The idea of a long polynomial is reasonable, because it actually comes from taking a sequence of polynomials with higher and higher degree:

$$f_0(x) = 1$$
$$f_1(x) = 1 + x$$
$$f_2(x) = 1 + x + \frac{x^2}{2}$$
$$f_3(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$
$$\vdots$$

Each polynomial in the sequence is, in a sense, the best approximation possible of that degree. Put another way, taking the first several terms of the long polynomial gives a good polynomial approximation of the function. The more terms included, the better the approximation. This is how calculators compute the exponential function (without having to add up infinitely many things).
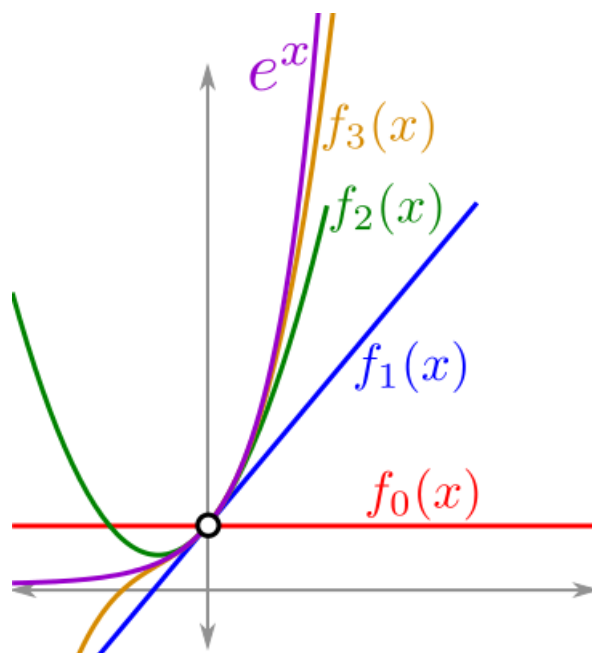
1.2.5   Exercises

- So, how good of an approximation is a polynomial truncation of e^x? Use a calculator to compare how close e is to the linear, quadratic, cubic, quartic, and quintic approximations. How many digits of accuracy do you seem to be gaining with each additional term in the series?

- Now, do the same thing with $1/e$ by plugging in $x = -1$ into the series. Do you have the same results? Are you surprised?

- Use the first three terms of the series for e^x to approximate $\sqrt[10]{e}$ and e^{10}. How accurate do you think these approximations are?

- Calculate the following sum using what you know:

$$\sum_{n=0}^{\infty} (-1)^n \frac{(\ln 3)^n}{n!}$$

- Write out the first four terms of the following series

$$\sum_{n=0}^{\infty} (-1)^n \frac{\pi^{2n}}{2^n n!}$$

- Write out the following series using summation notation:

$$1 - \frac{2}{3!} + \frac{4}{5!} - \frac{8}{7!} + \cdots$$

- Estimate $\sin(1/2)$ to three digits of accuracy. How many terms in the series did this take?

- We've seen that $i = e^{i\pi/2}$ via Euler's formula. Using this and some algebra, tell me what is i^i. Isn't that nice? Now, tell me, what is $((i^i)^i)$? Are you surprised? That's like, unreal!

- Practice your summation notation by rewriting the sum

$$\sum_{n=2}^{\infty} (-1)^n \frac{x^{n-2}}{n^3}$$

  as a sum over an index that goes from zero to infinity.

- Use the first two nonzero terms of the Taylor series for $\cos(x)$ to approximate $\cos(\frac{1}{10})$.

- Use Euler's formula to derive the double angle formulas $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$ and $\sin(2\theta) = 2\sin(\theta)\cos(\theta)$.

1.3 Taylor series

The long polynomial from the last module is actually called a Taylor series about $x = 0$ (this is referred to as a Maclaurin series in some textbooks, but this course will use the term Taylor series). The last module gave the Taylor series for e^x, $\sin x$, and $\cos x$. The logical next question is to ask whether every function has a Taylor series.

The answer is that most *reasonable* functions, and almost all of the functions encountered in this course, have a Taylor series. That is, every reasonable function f can be written as

$$f(x) = \sum_{k=0}^{\infty} c_k x^k = c_0 + c_1 x + c_2 x^2 + \cdots .$$

This module describes how to compute the coefficients c_k for a given function f.

1.3.1 The definition of a Taylor series at x=0

The definition of the Taylor series of f at $x = 0$ is

Taylor series at $x = 0$

$$f(x) = f(0) + \frac{f'(0)}{1!} x + \frac{f(0)}{2!} x^2 + \frac{f'(0)}{3!} x^3 + \cdots = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k,$$

where $f^{(k)}(0)$ is the kth derivative of f evaluated at 0. In other words, the coefficient c_k mentioned above is given by

$$c_k = \frac{f^{(k)}(0)}{k!} = \frac{1}{k!} \cdot \frac{d^k f}{dx^k}\bigg|_0$$

This seems circular, since the definition uses the function, and its derivatives, to write down the function. However, the definition only actually requires information about the function at a single point (in this case, 0). It is best to think of the Taylor series as a way of turning a function into a polynomial.

Example Compute the Taylor series for e^x using the above definition to see that it matches the given series from the last module.

Here, $f(x) = e^x$, and every derivative of e^x is e^x. Therefore, for all k we have

$$f^{(k)}(x) = e^x,$$

and so $f^{(k)}(0) = 1$ for all k. Plugging into the Taylor series formula gives

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$$

$$= \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

$$= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots ,$$

as claimed.

Example Compute the Taylor series for $f(x) = \sin x$ using the above definition, and verify it matches the series found using Euler's formula.

Computing the derivatives, and then evaluating at $x = 0$ gives the following table:

$$f(x) = \sin(x) \qquad\qquad f(0) = 0$$
$$f'(x) = \cos(x) \qquad\qquad f'(0) = 1$$
$$f''(x) = -\sin(x) \qquad\qquad f''(0) = 0$$
$$f'''(x) = -\cos(x) \qquad\qquad f'''(0) = -1$$
$$\vdots$$

Thus,

$$\sin(x) = 0 + \frac{1}{1!}x + \frac{0}{2!}x^2 + \frac{-1}{3!}x^3 + \cdots$$
$$= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots,$$

confirming what was found last time.

Example Compute the Taylor series for $f(x) = x^2 - 5x + 3$.

Again, by directly using the definition:

$$f(x) = x^2 - 5x + 3 \qquad\qquad f(0) = 3$$
$$f'(x) = 2x - 5 \qquad\qquad f'(0) = -5$$
$$f''(x) = 2 \qquad\qquad f''(0) = 2$$
$$f'''(x) = 0 \qquad\qquad f'''(0) = 0$$
$$\vdots$$

So it follows that

$$f(x) = 3 - 5x + \frac{2}{2!}x^2 = 3 - 5x + x^2$$

(since all the subsequent derivatives are 0), which is the original function. This should not be a surprise, since the Taylor series represents a function as a long polynomial (henceforth called by its proper name: *series*). If f was a polynomial to begin with, it stands to reason that the Taylor series for f should just be f itself.

1.3.2   Why Taylor series matter

The big idea of this module is that the Taylor series can be thought of as an operator (a machine) which turns a function into a series. This is a useful operator because some functions are hard (or even impossible) to express using combinations of familiar functions. Nevertheless, these functions can often be understood by computing their Taylor series.

   Example The *Bessel function*, denoted J_0, is best defined by its Taylor series:

$$J_0 = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2^{2k}(k!)^2}$$
$$= 1 - \frac{1}{2^2}x^2 + \frac{1}{2^4(2!)^2}x^4 - \frac{1}{2^6(3!)^2}x^6 + \cdots$$

This series has only the even powers of x, and it alternates, which is reminiscent of the series for cosine. One difference is that the denominator in the Bessel function grows more quickly than the denominator in the series for cosine. Thus, we might expect the graph to be a wave with a decreasing amplitude, which is exactly what we find:
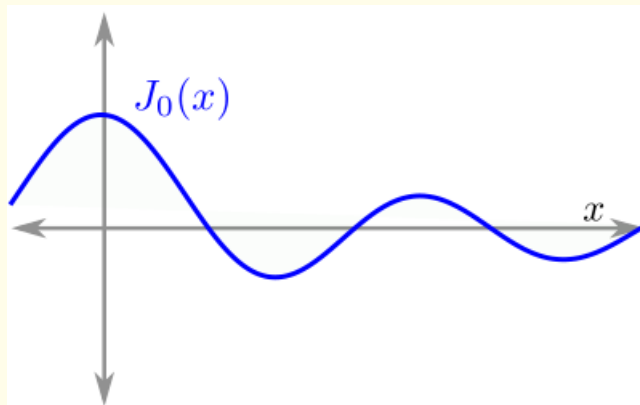
It turns out that the Bessel function describes many physical phenomena, including the shape of a hanging chain as it is rotated, and the shape of the waves formed after a stone is thrown into a pool of water.

1.3.3   Taylor series as polynomial approximants

The main reason Taylor series are useful is that they turn a potentially complicated function into something simple: a polynomial. Granted, this polynomial is infinitely long in general, but in practice it is only necessary to compute the first few terms to get a good, local approximation of the function. The more terms one includes, the better the polynomial approximates the function.

As an example, consider a particle on the number line with position function $p(t)$. At time 0, say its position is 5. Then one approximation of its position as a function of time is $p_0(t) = 5$. Given more information, say its velocity at time 0 is 3, the approximation becomes better. The next approximation as a function of time is $p_1(t) = 5 + 3t$. Now, suppose its acceleration at time 0 is -4. Then $p_2(t) = 5 + 3t - \frac{4}{2}t^2 = 5 + 3t - 2t^2$ is an even better polynomial approximation of the position function.
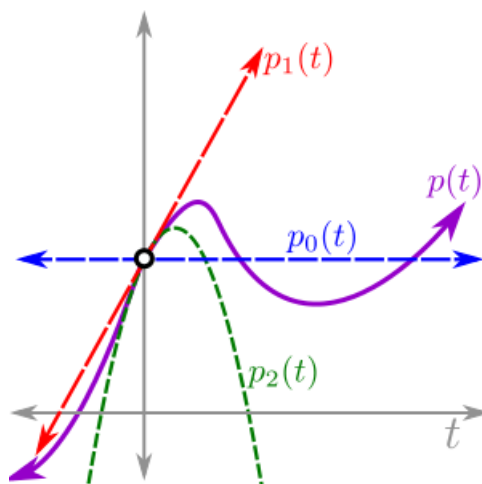
1.3.4   Exercises

- What is the Taylor series of $x^4 - 3x^3 + 2x^2 + 7x - 3$. This should be an easy one!

- What is the Taylor series of $(x - 2)^2(x - 3)$? This, also, should not be *too* hard...

- Compute a few derivatives and figure out the first few terms of the Taylor series of $\dfrac{1}{1 - x}$. Have you seen this series before?

- What are the first two nonzero terms in the Taylor series of $\sqrt[3]{1 - 2x}$?

- What is the coefficient of the cubic term in the Taylor series of e^{-3x}?

- Use what you know about Taylor series to determine the third derivative of $\sin^3(2x)\cos^2(3x)$ at $x = 0$. That's a *lot* easier than computing the derivatives!

- The ERF function is defined in terms of a difficult integral:

$$ERF(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt$$

- Even if you don't remember integrals all that well, you know how to integrate a polynomial, right? So, Taylor expand the integrand and integrate term by term to get the Taylor series for ERF.

- What is the third derivative of ERF(x) at zero?

- Why does a Taylor series have all those $(n!)$ terms in the denominator? Let's see. Compute the Taylor series of $f(x) = (1 + x)^5$ by (1) using the binomial theorem (or multiplication) to expand that power; then (2) by differentiating the function and using the Taylor series formula. What do you notice when you keep computing higher derivatives?

1.4   Computing Taylor series

The previous module gave the definition of the Taylor series for an arbitrary function. It turns out that this is not always the easiest way to compute a function's Taylor series. Just as functions can be added, subtracted, multiplied, and composed, so can their corresponding Taylor series.

Recall that the Taylor series for a function (f) is given by

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots .$$

Using the definition of the Taylor series involves taking a lot of derivatives, which could be a lot of work, especially if the function involves compositions and products of functions, e.g. $f(x) = \sin(x^2)e^{x^3}$.

This module will show how to compute the Taylor series of such functions more easily by using the Taylor series for functions we already know.

1.4.1   Substitution

Our first method, substitution, allows us to plug one function into the Taylor series of another. Consider the function

$$f(x) = \frac{1}{x}\sin(x^2).$$

Computing the Taylor series for f from the definition would involve the quotient rule, chain rule, and a lot of algebra. But by taking the series for $\sin x$ and substituting x^2 into this series, and then distributing the $\frac{1}{x}$, one finds

$$\frac{1}{x}\sin(x^2) = \frac{1}{x}\left((x^2) - \frac{1}{3!}(x^2)^3 + \frac{1}{5!}(x^2)^5 - \cdots \right)$$

$$= \frac{1}{x}\left(x^2 - \frac{1}{3!}x^6 + \frac{1}{5!}x^{10} - \cdots \right)$$

$$= x - \frac{1}{3!}x^5 + \frac{1}{5!}x^9 - \cdots .$$

Note that getting this many terms using the definition would involve taking nine derivatives of the original function, which would be a lot of work! To get a more complete description of the Taylor series, one can use the summation notation, and again substitute to find

$$\frac{1}{x}\sin(x^2) = \frac{1}{x}\sum_{k=0}^{\infty}(-1)^k \frac{(x^2)^{2k+1}}{(2k+1)!}$$

$$= \frac{1}{x}\sum_{k=0}^{\infty}(-1)^k \frac{x^{4k+2}}{(2k+1)!}$$

$$= \sum_{k=0}^{\infty}(-1)^k \frac{x^{4k+1}}{(2k+1)!}$$

Example Find the Taylor series for e^{x^3} by substitution.

Recall the series for e^x is

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

Substituting x^3 into the series for e^x gives

$$
\begin{aligned}
e^{x^3} &= 1 + x^3 + \frac{(x^3)^2}{2!} + \frac{(x^3)^3}{3!} + \cdots \\
&= 1 + x^3 + \frac{x^6}{2!} + \frac{x^9}{3!} + \cdots \\
&= \sum_{k=0}^{\infty} \frac{(x^3)^k}{k!} \\
&= \sum_{k=0}^{\infty} \frac{x^{3k}}{k!}
\end{aligned}
$$

1.4.2 Combining like terms

Another way to use previous knowledge of one Taylor series to find another is by combining like terms. This requires some careful algebra, but it is no more difficult than multiplying two polynomials together.

For example, consider the function

$$f(x) = \cos^2(x) = \cos(x) \cdot \cos(x).$$

Finding the series for a function multiplied by another function is the same as taking the series for each function and multiplying them together, and then collecting like terms. This is where some algebra is required.

$$
\begin{aligned}
\cos(x) \cdot \cos(x) &= \left(1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots\right)\left(1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots\right) \\
&= 1 + \left(-\frac{1}{2!} - \frac{1}{2!}\right)x^2 + \left(\frac{1}{4!} + \frac{1}{2!}\frac{1}{2!} + \frac{1}{4!}\right)x^4 + \cdots \\
&= 1 - x^2 + \frac{1}{3}x^4 + \cdots .
\end{aligned}
$$

To see where the coefficient of x^4 comes from, note that every x^4 term comes from some term from the left series multiplied together with some term from the right series whose powers add up to 4. There are three such pairs: 1 on the left paired with $\frac{1}{4!}x^4$ on the right; $-\frac{1}{2!}x^2$ on the left paired with $-\frac{1}{2!}x^2$ on the right; and $\frac{1}{4!}x^4$ on the left paired with 1 on the right. This is the same algebra one would do when multiplying two polynomials together; this is just a way of collecting like terms in a systematic way.

Example Use the trigonometric identity

$$\cos^2 x = \frac{1 + \cos(2x)}{2}$$

and substitution to find the series for $\cos^2 x$. Try to give the series in summation notation (other than the first term)

By the above identity,

$$\cos^2 x = \frac{1}{2}\left(1 + \cos(2x)\right)$$
$$= \frac{1}{2}\left(1 + \left(1 - \frac{(2x)^2}{2!} + \frac{(2x)^4}{4!} - \cdots\right)\right)$$
$$= \frac{1}{2}\left(2 - \frac{4x^2}{2} + \frac{16x^4}{24} - \cdots\right)$$
$$= 1 - x^2 + \frac{x^4}{3} - \cdots.$$

In summation notation,

$$\cos^2 x = \frac{1}{2}\left(1 + \sum_{k=0}^{\infty}(-1)^k\frac{(2x)^{2k}}{(2k)!}\right)$$
$$= \frac{1}{2} + \frac{1}{2}\sum_{k=0}^{\infty}(-1)^k\frac{(2x)^{2k}}{(2k)!}$$
$$= \frac{1}{2} + \sum_{k=0}^{\infty}(-1)^k\frac{2^{2k-1}x^{2k}}{(2k)!}$$
$$= 1 + \sum_{k=1}^{\infty}(-1)^k\frac{2^{2k-1}x^{2k}}{(2k)!}.$$

1.4.3   Hyperbolic trigonometric functions

The hyperbolic trigonometric functions $\sinh(x)$, $\cosh(x)$, and $\tanh(x)$ are defined by

$$\sinh(x) = \frac{e^x - e^{-x}}{2}$$
$$\cosh(x) = \frac{e^x + e^{-x}}{2}$$
$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{\sinh(x)}{\cosh(x)}.$$
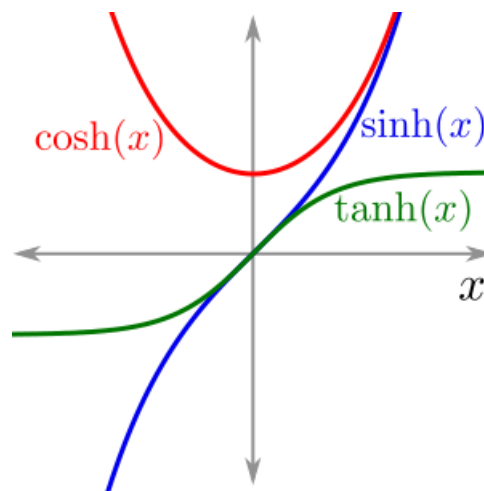
These hyperbolic trig functions, although graphically quite different from their traditional counterparts, have several similar algebraic properties, which is why they are so named. For example, the Pythagorean identity for cosine and sine has a version for hyperbolic cosine and sine:

$$\cosh^2(x) - \sinh^2(x) = 1.$$

One can verify this using the definitions and some algebra. But there is a geometric intuition for this relationship. Recall that cosine and sine give the x and y coordinates, respectively, for a point on

the unit circle $x^2 + y^2 = 1$. The hyperbolic cosine and hyperbolic sine give the x and (y) coordinates, respectively, for points on the hyperbola $x^2 - y^2 = 1$:
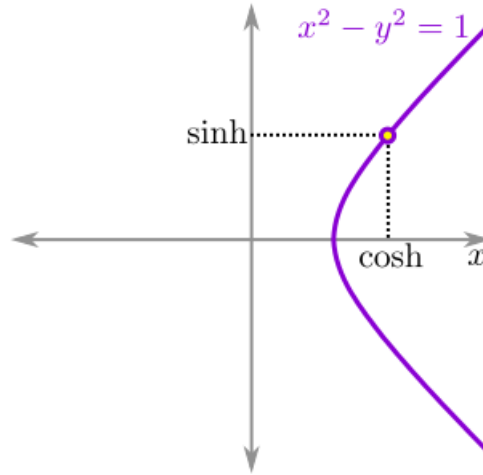
Example Using the Taylor series for e^x and substitution, show that the Taylor series for **cosh** and **sinh** are

$$\cosh(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}$$

$$\sinh(x) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}.$$

Note that these are almost the same as the series for cosine and sine, respectively, except they do not alternate. This gives another reason for the names of these functions.

$$
\begin{aligned}
\cosh(x) &= \frac{e^x + e^{-x}}{2} \\
&= \frac{1}{2}\left[(1 + x + \frac{x^2}{2!} + \cdots) + (1 - x + \frac{x^2}{2!} - \cdots)\right] \\
&= \frac{1}{2}\left[2 + 2\frac{x^2}{2!} + 2\frac{x^4}{4!} + \cdots\right] \\
&= 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots \\
&= \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}.
\end{aligned}
$$

$$
\begin{aligned}
\sinh(x) &= \frac{e^x - e^{-x}}{2} \\
&= \frac{1}{2}\left[(1 + x + \frac{x^2}{2!} + \cdots) - (1 - x + \frac{x^2}{2!} - \cdots)\right] \\
&= \frac{1}{2}\left[2x + 2\frac{x^3}{3!} + 2\frac{x^5}{5!} + \cdots\right] \\
&= x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots \\
&= \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}.
\end{aligned}
$$

1.4.4   Manipulating Taylor series

Another way of using one Taylor series to find another is through differentiation and integration. For instance, to find the Taylor series for the derivative of $\boldsymbol{f}$, one can differentiate the Taylor series for $\boldsymbol{f}$ term by term.

Example By differentiating the Taylor series for **sinh** and **cosh**, show that

$$\frac{d}{dx}\sinh x = \cosh x$$

$$\frac{d}{dx}\cosh x = \sinh x.$$

This is yet another relationship which is similar (though not identical) to the relationship between sine and cosine.

Differentiating hyperbolic sine gives

$$\frac{d}{dx}\sinh x = \frac{d}{dx}\sum_{k=0}^{\infty}\frac{x^{2k+1}}{(2k+1)!}$$

$$= \sum_{k=0}^{\infty}(2k+1)\frac{x^{2k}}{(2k+1)!}$$

$$= \sum_{k=0}^{\infty}\frac{x^{2k}}{(2k)!}$$

$$= \cosh x,$$

as desired. Similarly, differentiating hyperbolic cosine gives

$$\frac{d}{dx}\cosh x = \frac{d}{dx}\sum_{k=0}^{\infty}\frac{x^{2k}}{(2k)!}$$

$$= \sum_{k=0}^{\infty}(2k)\frac{x^{2k-1}}{(2k)!}$$

$$= \sum_{k=1}^{\infty}\frac{x^{2k-1}}{(2k-1)!}$$

$$= \sum_{k=0}^{\infty}\frac{x^{2k+1}}{(2k+1)!}.$$

There was a little bit of reindexing there, but by writing out a few terms of each series, one can see that all of the above equalities are true.

1.4.5   Higher Order Terms in Taylor Series

In some situations, it will be convenient only to write the first few terms of a Taylor series. This is particularly true when combining or composing more than one Taylor series. Up until now, an ellipsis has been used to indicate that there are more terms in the series that are being omitted.

There is another way, sometimes used in this course, of notating the omitted terms in a Taylor series. That is by referring to them as Higher Order Terms (or H.O.T. for short). Having the extra HOT in a series means that all the remaining terms in the series have a higher degree than the previous terms.

Example The function e^x can be written as

$$e^x = 1 + x + \frac{1}{2!}x^2 + \text{ HOT},$$

or it could also be written as

$$e^x = 1 + x + \text{ HOT}.$$

The point at which the higher order terms are cut-off is somewhat arbitrary and depends on the situation. There is a more formal way of keeping track of the higher order terms, called Big-O notation, which is presented in orders of growth.

Example Find the first two non-zero terms of the Taylor series for

$$f(x) = 1 - 2xe^{\sin x^2}.$$

Beginning with the innermost function, in this case $\sin x^2$, we find that

$$\sin x^2 = x^2 - \frac{1}{3!}(x^2)^3 + \text{HOT} = x^2 - \frac{1}{6}x^6 + \text{HOT}.$$

Then plugging this into the series for e^x gives

$$e^{\sin x^2} = 1 + \left(x^2 - \frac{1}{6}x^6 + \text{HOT}\right) + \frac{1}{2!}\left(x^2 + \text{HOT}\right)^2 + \frac{1}{3!}\left(x^2 + \text{HOT}\right)^3 + \text{HOT}$$

$$= 1 + x^2 + \frac{1}{2}x^4 + \left(-\frac{1}{6} + \frac{1}{6}\right)x^6 + \text{HOT}$$

$$= 1 + x^2 + \frac{1}{2}x^4 + \text{HOT}$$

Then to complete the answer, plug this into the original function to find

$$f(x) = 1 - 2x\left(1 + x^2 + \frac{1}{2}x^4 + \text{ HOT}\right)$$

$$= 1 - 2x - 2x^3 - x^5 + \text{ HOT}.$$

1.4.6   Extra examples

Example Compute the Taylor series (at 0) for $\sin^2 x$ up to and including terms of order 6. Try to give the full Taylor series in summation notation.

$$\sin^2 x = (x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots)(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots)$$

$$= x^2 + (-\frac{1}{3!} - \frac{1}{3!})x^4 + (\frac{1}{5!} + \frac{1}{3! \cdot 3!} + \frac{1}{5!})x^6 + \cdots$$

$$= x^2 - \frac{1}{3}x^4 + \frac{2}{45}x^6 - \cdots.$$

To get the full Taylor series, one can use the identity

$$\sin^2 x = \frac{1 - \cos(2x)}{2}$$

to find that

$$\sin^2 x = \frac{1 - \cos(2x)}{2}$$

$$= \frac{1}{2}\left(1 - \left(1 - \frac{(2x)^2}{2!} + \frac{(2x)^4}{4!} - \cdots\right)\right)$$

$$= \frac{1}{2}\left(\frac{(2x)^2}{2!} - \frac{(2x)^4}{4!} + \frac{(2x)^6}{6!} - \cdots\right)$$

$$= \frac{1}{2}\sum_{k=1}^{\infty}(-1)^{k-1}\frac{(2x)^{2k}}{(2k)!}.$$

Example Find the first three terms of the Taylor series for $\sqrt{f(x)}$, where

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots.$$

Let $g(x) = \sqrt{f(x)}$, where

$$g(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \cdots.$$

Then $g(x)^2 = f(x)$, and so the same holds for the Taylor series:

$$\left(b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \cdots\right)^2 = a_0 + a_1 x + a_2 x^2 + \cdots.$$

Multiplying out and collecting like terms gives

$$b_0^2 + (b_0 b_1 + b_1 b_0)x + (b_0 b_2 + b_1 b_1 + b_2 b_0)x^2 + \cdots = a_0 + a_1 x + a_2 x^2 + \cdots.$$

Now, equating coefficients of the monomials on the left and right gives the first few equations (of an infinite system of equations)

$$b_0^2 = a_0$$
$$2b_0 b_1 = a_1$$
$$2b_0 b_2 + b_1^2 = a_2.$$

Solving these equations gives the first three coefficients of g:

$$b_0 = \sqrt{a_0}$$
$$b_1 = \frac{a_1}{2\sqrt{a_0}}$$
$$b_2 = \frac{1}{2\sqrt{a_0}} \left(a_2 - \frac{a_1^2}{4a_0} \right).$$

Thus,

$$\sqrt{a_0 + a_1 x + a_2 x^2 + \cdots} = \sqrt{a_0} + \frac{a_1}{2\sqrt{a_0}} x + \frac{1}{2\sqrt{a_0}} \left(a_2 - \frac{a_1^2}{4a_0} \right) x^2 + \cdots.$$

1.4.7   Exercises

- Compute the Taylor series of $\cos(2x)\sin(3x)$ up to and including terms of degree 5. Don't try computing derivatives for this!

- Use a Taylor polynomial to give a cubic approximation to $2xe^{3x}$

- Compute the Taylor series of $e^{1-\cos t}$ in summation notation.

- Compute the Taylor series of $\cos(\sin(x))$ to fourth order.

- Compute the Taylor series of $\sin(\cos(x))$ to forth order. What happens that makes this different than the last problem? (Hint: $\cos(0) = 1$ but $\sin(0) = 0$...)

- Compute the first three nonvanishing terms in the Taylor series of $e^{2x}(\sinh 3x)/x$.

- Compute the Taylor series of $3x^2 e^{-x^2} \sin 2x^3$ up to and including terms of order eight! Wow, that means a lot of work, right? Think... which terms should you expand first?

- Compute the Taylor series of $\frac{1}{x} e^{-x^2} \sinh(2x)$ up to the fourth order term.

- What is the second derivative of the function $e^{x\cosh(x^2)}$ at $x = 0$?

- Compute the following limit $\lim_{x\to 0} (1 - e^x)\frac{\sin(x^2)}{x^3}$

1.5   Convergence

A Taylor series can be thought of as an infinite polynomial. Up until now, we have not worried about the issues that come up when adding up infinitely many things. This module deals with two main issues:

1. A function may not have a Taylor series at all;

2. A function's Taylor series may not converge everywhere, even within the function's domain.

1.5.1 Functions without a Taylor series

The first problem is that some functions cannot be expressed in the form

$$f(x) = \sum_{k=0}^{\infty} c_k x^k = c_0 + c_1 x + c_2 x^2 + \cdots$$

Examples include **tan**, which has vertical asymptotes, and **ln**, which is not defined for $x \leq 0$. Polynomials are not able to capture these sorts of discontinuities and asymptotes.

The geometric series

The geometric series is an example of a Taylor series which is well behaved for some values of x and nonsensical for other values of x. The claim is that

$$1 + x + x^2 + x^3 + x^4 + \cdots = \frac{1}{1-x},$$

for $|x| < 1$.

> **Note** This is not a formal proof, which would require a few tools and definitions we have not yet learned.
> Let $y = 1 + x + x^2 + x^3 + \cdots$. Multiplying both sides by x gives
>
> $$y = 1 + x + x^2 + x^3 + \cdots \qquad xy = x + x^2 + x^3 + x^4 + \cdots$$
>
> Now, subtracting the second equation from the first, all the terms other than 1 cancel on the right, leaving us with
> $$y(1-x) = 1.$$
> Dividing by $1 - x$ gives $y = \frac{1}{1-x}$.

Example Compute the Taylor series for $f(x) = \frac{1}{1-x}$ directly from the definition.

> $$f(x) = \frac{1}{1-x} \qquad\qquad f(0) = 1$$
> $$f'(x) = \frac{1}{(1-x)^2} \qquad\qquad f'(0) = 1$$
> $$f(x) = \frac{2}{(1-x)^3} \qquad\qquad f(0) = 2$$
> $$f(x) = \frac{6}{(1-x)^4} \qquad\qquad f(0) = 6.$$
>
> Notice the pattern that
> $$f^{(k)}(x) = \frac{k!}{(1-x)^{k+1}},$$
> at least for the first few k. To see that the pattern continues, assume it holds for some k, and show that it holds for $k+1$ (this is a proof technique known as mathematical induction). If $f^{(k)}(x) = \frac{k!}{(1-x)^{k+1}}$, then
> $$f^{(k+1)}(x) = \frac{(k+1)k!}{(1-x)^{k+2}} = \frac{(k+1)!}{(1-x)^{k+2}},$$
> as desired. Then $f^{(k)}(0) = k!$, so according to the definition of Taylor series, it follows that
> $$\frac{1}{1-x} = 0! + 1!x + \frac{2!}{2!}x^2 + \frac{3!}{3!}x^3 + \cdots \qquad = 1 + x + x^2 + x^3 + \cdots,$$
> which agrees with the above.

Note The geometric series only holds when $|x| < 1$. This makes sense, because if $|x| > 1$, the powers of x are getting bigger and bigger and so the series should not converge. If $x = 1$, then the series is adding 1 infinitely many times, which diverges. If $x = -1$, then the series oscillates between 1 and 0, and hence does not converge.

The takeaway is that every Taylor series has a convergence domain where the series is well-behaved, and outside that domain the series will not converge. For many functions, the domain is the whole real number line (e.g. the series for e^x, **sin**, **cos**, **cosh**, and **sinh** all converge everywhere), but be aware that there are functions whose Taylor series do not converge everywhere. This will be covered more formally in Series Convergence And Divergence.

Example A beam of light of intensity L hits a pane of glass. Half of the light is reflected, and a third of the light is transmitted; the rest is absorbed. When a beam of light of intensity L hits two parallel panes with an air gap between them, how much light is transmitted through both panes? (The following figure shows how the light gets reflected and rereflected. The first transmitted and reflected beams of light are labeled with their respective intensities. The question asks for the total of the beams of light emerging on the right side of the right pane of glass).
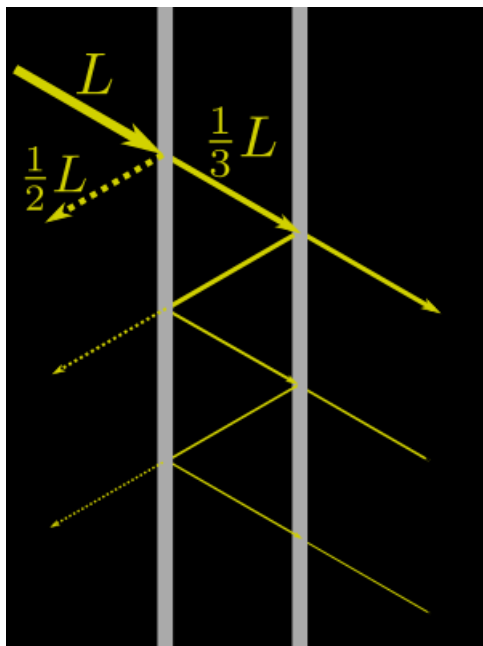
By labeling more of the transmitted and reflected beams of light, a pattern emerges among the beams of light on the right side of the right pane:
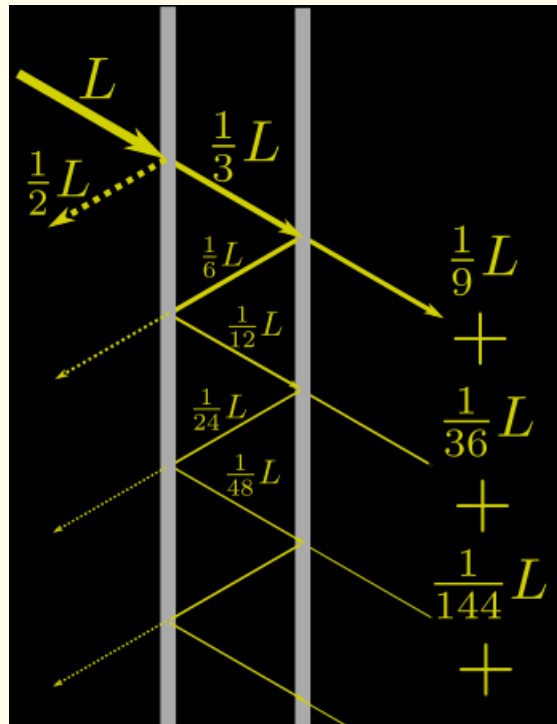
$\frac{1}{9}, \frac{1}{36}, \frac{1}{144}, \ldots$. Note that each beam is $\left(\frac{1}{4}\right)$ the previous beam. Thus, the total light emerging on the right side of the right pane of glass is

$$\frac{L}{9} + \frac{L}{36} + \frac{L}{144} + \cdots$$

$$= \frac{L}{9}\left(1 + \frac{1}{4} + \frac{1}{16} + \cdots\right)$$

$$= \frac{L}{9}\left(\frac{1}{1 - 1/4}\right)$$

$$= \frac{L}{9}\frac{4}{3}$$

$$= \frac{4L}{27},$$

by using the formula for the geometric series.

Example Use the Taylor series of $\frac{1}{1-x}$ to derive the Taylor series of $\ln(1 + x)$. Hint: recall that $\ln(1 + x) = \int \frac{1}{1+x}dx$.

Note that

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

$$= 1 - x + x^2 - x^3 + x^4 - \cdots.$$

Now, integrating gives $\int \frac{dx}{1+x} = \ln(1 + x) + C$ on the one hand, and

$$\int (1 - x + x^2 - x^3 + x^4 - \cdots)dx = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$$

$$= \sum_{k=1}^{\infty}(-1)^{k-1}\frac{x^k}{k},$$

on the other hand. Plugging in $x = 0$ shows that $C = 0$, and so

$$\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$$
$$= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}.$$
$$(|x| < 1)$$

Note that because this relied on the geometric series, which only holds for $|x| < 1$, the same restriction holds for the Taylor series for $\ln(1 + x)$.

Example Use the fact that

$$\arctan x = \int \frac{1}{1 + x^2} \, dx$$

to find the Taylor series for $\arctan x$.

Using the fact, and the geometric series, we find that

$$\arctan(x) = \int \frac{1}{1 + x^2} \, dx$$
$$= \int \frac{1}{1 - (-x^2)} \, dx$$
$$= \int \left(1 - x^2 + x^4 - x^6 + \cdots \right) \, dx$$
$$= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + C.$$
$$|x| < 1$$

Plugging in $x = 0$ gives that $C = 0$, since $\arctan 0 = 0$. Thus,

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots$$
$$= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k + 1}$$
$$|x| < 1.$$

So even though **arctan** is defined for all x, its Taylor series only converges for $|x| < 1$.

Example Another important function is the binomial series $(1 + x)^\alpha$, where α is some constant. Show that

$$(1 + x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha - 1)}{2!} x^2 + \frac{\alpha(\alpha - 1)(\alpha - 2)}{3!} x^3 + \cdots$$
$$= \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k,$$

where

$$\binom{\alpha}{k} = \frac{\alpha(\alpha - 1)(\alpha - 2) \cdots (\alpha - k + 1)}{k!}.$$

This series also only holds for $|x| < 1$.

For fixed α we have $f(x) = (1 + x)^\alpha$. Then proceeding from the definition of the Taylor series,

one computes

$$f(x) = (1+x)^\alpha \qquad\qquad f(0) = 1$$
$$f'(x) = \alpha(1+x)^{\alpha-1} \qquad\qquad f'(0) = \alpha$$
$$f(x) = \alpha(\alpha-1)(1+x)^{\alpha-2} \qquad\qquad f(0) = \alpha(\alpha-1)$$
$$f(x) = \alpha(\alpha-1)(\alpha-2)(1+x)^{\alpha-3} \qquad\qquad f(0) = \alpha(\alpha-1)(\alpha-2)$$
$$\vdots \qquad\qquad\qquad\qquad\qquad \vdots$$

One finds that, in general, $f^{(k)}(0) = \alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)$. Thus, the Taylor expansion for $(1+x)^\alpha$ is

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \cdots$$
$$= 1 + \binom{\alpha}{1}x + \binom{\alpha}{2}x^2 + \binom{\alpha}{3}x^3 + \cdots$$
$$= \sum_{k=0}^{\infty} \binom{\alpha}{k}x^k,$$

as claimed.

1.5.2 Summary

Here are all the series we have found so far. The following hold for all x:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$
$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$
$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$$
$$\cosh x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}$$
$$\sinh x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}.$$

The following hold for $|x| < 1$:

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$$
$$\ln(1+x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}$$
$$\arctan x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}$$
$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k}x^k.$$

1.5.3 Electrostatics example

Here we use the geometric series and the binomial series from above in an example from electrostatics. An electric dipole is a pair of equally and oppositely charged particles separated by a short distance. One question of interest in electrostatics is the electrostatic potential, which is the sum of the point charge potentials from each pole.

The point charge potential from a single particle with charge q, at a distance d from the particle, is

$$V = \frac{kq}{d},$$

where k is a constant called the Coulomb constant. Then a dipole with particles of charge q and $-q$ has net electrostatic potential

$$V = \frac{kq}{d_+} - \frac{kq}{d_-},$$

where d_+ is the distance to the positively charged particle, and d_- is the distance to the negatively charged particle:
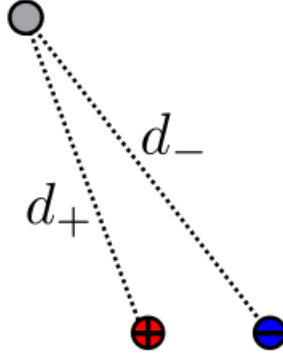
We will calculate the first order term for the electrostatic potential at two different locations: p_1 and p_2:
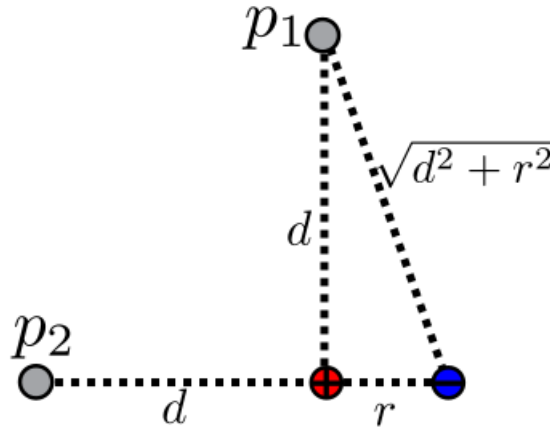
First consider p_1, located directly above and distance d from the positive particle. Let r be the distance between the charged particles. Then $d_+ = d$, and by the Pythagorean theorem, $d_- = \sqrt{d^2 + r^2}$. It follows that the electrostatic potential is

$$V = \frac{kq}{d_+} - \frac{kq}{d_-}$$

$$= \frac{kq}{d} - \frac{kq}{\sqrt{d^2 + r^2}}.$$

Now, factoring out $\frac{kq}{d}$, and applying the binomial series with $\alpha = -\frac{1}{2}$, we find

$$V = \frac{kq}{d}\left[1 - \frac{1}{\sqrt{1 + (r/d)^2}}\right]$$

$$= \frac{kq}{d}\left[1 - \left(1 + (r/d)^2\right)^{-1/2}\right]$$

$$= \frac{kq}{d}\left[1 - \left(1 - \frac{1}{2}(r/d)^2 + \text{ HOT}\right)\right]$$

$$= \frac{1}{2}\frac{kqr^2}{d^3} + \text{ HOT.}$$

At position p_2, which is directly left of and distance d from the positive particle, we have $d_+ = d$, and $d_- = d + r$, so we find that the electrostatic potential at p_2 is

$$V = \frac{kq}{d_+} - \frac{kq}{d_-}$$
$$= \frac{kq}{d} - \frac{kq}{d+r}.$$

Again, factoring out $\frac{kq}{d}$ and expanding using the geometric series gives

$$V = \frac{kq}{d}\left(1 - \frac{1}{1 + \frac{r}{d}}\right)$$
$$= \frac{kq}{d}\left(1 - \left(1 - \frac{r}{d} + \text{HOT}\right)\right)$$
$$= \frac{kqr}{d^2} + \text{HOT}.$$

1.5.4   Exercises

- Consider a snowman built from solid snowballs of radius 2^{-n}, for $n = 0, 1, 2, \ldots$, all stacked on top of one another. How many units tall is the snowman? How many cubic units of snow was required to build it?

- Compute the Taylor series about zero of

$$\ln\frac{1 + 3x}{1 - 3x}$$

- Compute the Taylor series about zero of

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

- Using your answer to the previous problem, compute the Taylor series about zero of $\arcsin x$, using termwise integration and the fact that

$$\arcsin x = \int \frac{dx}{\sqrt{1 - x^2}}$$

- For which values z is the Taylor series of $\sqrt[4]{3 - 2z^2}$ guaranteed to converge?

- Use the binomial series to give the Taylor expansion of $(1 + x)^3$. Now, do it with your head: easier, right? Recall, we have said that the binomial series only converges when $|x < 1|$, but, clearly, that cannot be a *sharp* constraint, since $(1 + x)^3$ is good for all x, right? Well, Horatio, there are more things... By the end of this course, we will learn when and how to bend some of these restrictions.

- Build a cylinder with radius 1 and height 3. Build a second cylinder with radius 1/2 and height 9, a third cylinder with radius 1/4, height 27, a fourth cylinder with radius 1/8 and height 81, and so on. What is the total volume of the cylinders?

- For which values of x does the Taylor series of $(\frac{1}{4} - 3x^2)^{1/4}$ converge?

1.6   Expansion points

Up until now, Taylor series expansions have all been at $x = 0$. The Taylor series at $x = 0$ gives a good approximation to the function near 0. But what if we want a good approximation to the function near a different point a? That is the topic of this module.

1.6.1   Expansion points

A function f has a Taylor series expansion about any point $x = a$ provided that f and all its derivatives exist at a. The definition of the Taylor series for f about $x = a$ is

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \cdots$$
$$= \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x - a)^k.$$

We say this is a series in $(x - a)$. A different way to view this series is by making the change of variables $x = a + h$. After cancellation, this yields

$$f(a + h) = f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \cdots$$
$$= \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}h^k.$$

1.6.2   Taylor polynomial for approximation

Recall that the first few terms of the Taylor series for f about $x = 0$ gives a polynomial (the Taylor polynomial) which is a good approximation for f near 0. Similarly, the Taylor polynomial for f about $x = a$ gives a polynomial which is a good approximation of f near $x = a$. Note, however, that as the input gets further away from the expansion point a, the approximation gets worse.

Example Find the Taylor series for $f(x) = 3x^2 - x + 4$ about $x = 2$.

Computing the derivatives, and evaluating at $(x = 2)$, one finds

$$\begin{aligned}
f(x) &= 3x^2 - x + 4 & f(2) &= 14 \\
f'(x) &= 6x - 1 & f'(2) &= 11 \\
f(x) &= 6 & f(2) &= 6 \\
f(x) &= 0 & f(2) &= 0.
\end{aligned}$$

And all the subsequent derivatives are 0. So from the definition, one finds that

$$f(x) = 14 + 11(x - 2) + \frac{6}{2!}(x - 2)^2$$
$$= 14 + 11(x - 2) + 3(x - 2)^2.$$

This appears to be different than the polynomial f with which we began. If one multiplies out this polynomial and collects like terms, however, the result is the original polynomial. This should not be surprising, since the best polynomial approximation to a polynomial is the polynomial itself, even factored into a slightly different form.

Example Compute the Taylor series expansion for $\ln(x)$ about $x = 1$.

Begin by computing the first few derivatives and evaluating at $x = 1$:

$$\begin{aligned}
f(x) &= \ln(x) & f(1) &= 0 \\
f'(x) &= x^{-1} & f'(1) &= 1 \\
f(x) &= -x^{-2} & f(1) &= -1 \\
f(x) &= 2x^{-3} & f(1) &= 2.
\end{aligned}$$

The pattern that emerges is $f^{(k)}(x) = (-1)^{k-1}(k - 1)!x^{-k}$. To see that the pattern holds, check that

$$f^{(k+1)}(x) = (-1)^{k-1}(-k)(k - 1)!x^{-k-1} = (-1)^k k!x^{-(k+1)},$$

as desired. So by induction, the pattern holds. It follows that $f^{(k)}(1) = (-1)^{k-1}(k - 1)!$ for

$k \geq 1$. Plugging in to the formula, one finds that

$$\ln(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}(k-1)!}{k!}(x-1)^k$$

$$= \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(x-1)^k}{k}$$

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

Note that with the change of variables $h = x - 1$ (and hence $x = h + 1$, we find that

$$\ln(1+h) = h - \frac{h^2}{2} + \frac{h^3}{3} - \frac{h^4}{4} + \cdots ,$$

which is the same series we found earlier for $\ln(1 + x)$.

Note that the Taylor polynomial is only a good approximation to the function on the domain of convergence. For functions whose domain of convergence is the entire number line, this is not a concern. But for functions such as $\ln x$, the Taylor polynomials will only be a good approximation within the domain of convergence, which is $0 < x < 2$. Outside of that domain, the Taylor polynomials diverge wildly from $\ln x$, as shown here:
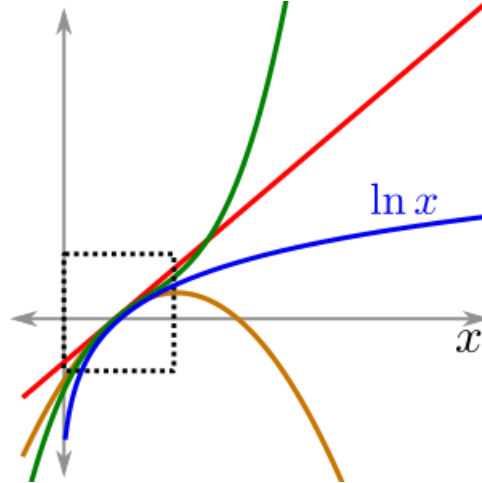
Even within a function's domain of convergence, a Taylor polynomial's approximation gets worse as the input gets further away from a. One way to improve an approximation is to include more and more terms of the Taylor series in the Taylor polynomial. However, this involves computing more and more derivatives. Another way to improve the approximation for $f(x)$ is to choose an expansion point a which is close to x.

Example Use the Taylor polynomial of degree 2 for $f(x) = \sqrt{x}$ about $x = 1$ to approximate $\sqrt{10}$. Then repeat the process about $x = 9$ and compare the results.

Using the definition, one finds

$$f(x) = \sqrt{x} \qquad\qquad f(1) = 1 \qquad\qquad f(9) = 3$$

$$f'(x) = \frac{1}{2\sqrt{x}} \qquad\qquad f'(1) = \frac{1}{2} \qquad\qquad f'(9) = \frac{1}{6}$$

$$f(x) = -\frac{1}{4x^{3/2}} \qquad\qquad f(1) = -\frac{1}{4} \qquad\qquad f''(9) = -\frac{1}{108}$$

Thus, the Taylor polynomial about $x = 1$ is

$$\sqrt{x} \approx 1 + \frac{1}{2}(x-1) - \frac{1/4}{2!}(x-1)^2$$

$$= 1 + \frac{1}{2}(x-1) - \frac{1}{8}(x-1)^2.$$

And the corresponding approximation is

$$\sqrt{10} \approx 1 + \frac{1}{2} \cdot 9 - \frac{1}{8} \cdot 9^2$$
$$\approx -4.6,$$

which is obviously quite far off the mark. On the other hand, the Taylor polynomial about $x = 9$ is

$$\sqrt{x} \approx 3 + \frac{1}{6}(x - 9) - \frac{1/108}{2!}(x - 9)^2$$
$$= 3 + \frac{1}{6}(x - 9) - \frac{1}{216}(x - 9)^2.$$

And the corresponding approximation is

$$\sqrt{10} \approx 3 + \frac{1}{6} \cdot 1 - \frac{1}{216} \cdot 1^2$$
$$\approx 3.1620,$$

which is quite a good approximation of $\sqrt{10} \approx 3.1623$.

1.6.3 Caveat for compositions

When computing the Taylor expansion for the composition $f \circ g$ about $x = a$, one must be careful of expansion points. In particular, one cannot simply take the series for g at $x = a$ and plug it into the series for f at $x = a$.

Example Consider the expansion for $e^{\cos(x)}$ about $x = 0$.

Although $\cos(x) = 1 - \frac{x^2}{2!} + \cdots$, and $e^x = 1 + x + \frac{x^2}{2!} + \cdots$, one will run into trouble trying to write

$$e^{\cos(x)} = 1 + (1 - \frac{x^2}{2!} + \cdots) + \frac{1}{2}(1 - \frac{x^2}{2!} + \cdots)^2 + \cdots.$$

The trouble is that collecting like terms requires adding up infinitely many things. For instance, the constant term above is $1 + 1 + \frac{1}{2} + \cdots$. The reason this is a problem is that Taylor series are supposed to give a good polynomial approximation of a function without requiring too much computation or information about the function.

Remember that $e^x = 1 + x + \frac{x^2}{2} + \cdots$ is a good approximation when x is near 0. However, when x is near 0, $\cos(x)$ is near 1. So plugging the series for $\cos(x)$ into the series for e^x does not give a good approximation.

To avoid this problem when computing the Taylor series for the composition $f \circ g$ at $x = a$, one should plug the Taylor expansion of g about $x = a$ into the expansion of f about $x = g(a)$. In the above example, the expansion of e^x about $x = 1$ is

$$e^x = e + e(x - 1) + \frac{e}{2!}(x - 1)^2 + \cdots,$$

so

$$e^{\cos(x)} = e + e\left[\left(1 - \frac{x^2}{2!} + \cdots\right) - 1\right] + \frac{e}{2!}\left[\left(1 - \frac{x^2}{2!} + \cdots\right) - 1\right]^2 + \cdots$$
$$= e + e(-\frac{x^2}{2} + \cdots) + \frac{e}{2}(-\frac{x^2}{2} + \cdots)^2 + \cdots$$
$$= e - \frac{e}{2}x^2 + \cdots.$$

1.6.4 Exercises

- Without using a calculator, find a decimal approximation to $\sqrt{83}$ by Taylor-expanding $\sqrt{x}$ about $a = 81$ and using the zero-th and first order terms.

- Without using a calculator, find a decimal approximation to $\sqrt[3]{124}$ using linear approximation. How close was your answer to truth?

- Without using a calculator, find a decimal approximation to $\cos(1)$ [in radians!] using linear approximation. How close was your answer to truth? (Hint: $\pi/3 \approx 1$...)

- Taylor expand $\sin x$ about $x = \pi$ and compute all the terms. Does what you get make sense?

- Use completing the square and the geometric series to get the Taylor expansion about $x = 2$ of $\frac{1}{x^2+4x+3}$

- Approximate $1004^{1/3}$ using the zeroth and first order terms of the Taylor series.

1.7 Limits

Having concluded our study of Taylor series, we now move on to limits. Some of the major topics of calculus (continuity, differentiation, and integration) can all be expressed using limits.

1.7.1 Definition of the limit

The limit formalizes the behavior of a function as its input approaches some value. The formal definition of the limit is

> **Limit** $\lim_{x \to a} f(x) = L$ if and only if for every $\epsilon > 0$ there exists $\delta > 0$ such that $|f(x) - L| < \epsilon$ whenever $0 < |x - a| < \delta$. If there is no such L, then the limit does not exist.

In words, this says that the limit of a function exists if, when the input to f is very close to a (but not equal to a), the output from f is very close to L. This can also be thought of in terms of tolerances: given a certain ϵ tolerance for the output (seen as the band around L in the graph below), one can find a tolerance δ on the input (the band around a) so that for inputs within the tolerance, the corresponding outputs stays within ϵ of the desired output:
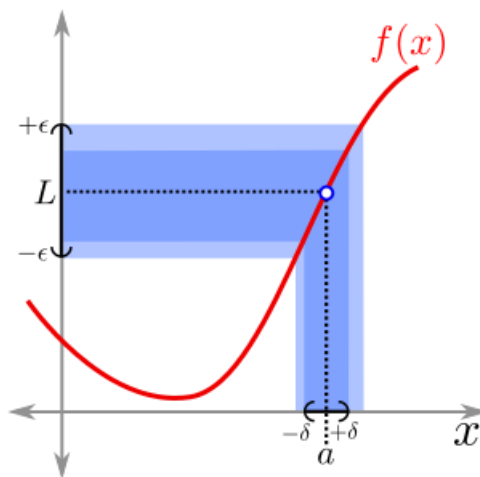
No matter how small ϵ is made, there must be some δ, which must depend on ϵ, generally. Actually finding δ often requires a little bit of work.

Example Using the definition of the limit, show that $\lim_{x \to 3} x^2 = 9$.

> **Note** This is rather technical, and is only a demonstration of the process required to prove a limit exists from the definition. This course deals almost exclusively with continuous functions, where such proofs are not necessary.
>
> We must show that for any given $\epsilon > 0$, there exists δ (which depends on ϵ) such that $0 < |x - 3| < \delta$ implies $|x^2 - 9| < \epsilon$. Let $\epsilon > 0$ be given. A little bit of algebra shows that
>
> $$|x^2 - 9| = |x - 3| \cdot |x + 3|.$$
>
> We get to control $|x - 3|$ with δ. We also have (by using the triangle inequality) that
>
> $$|x + 3| = |x - 3 + 6| \le |x - 3| + 6 < \delta + 6.$$

Thus,
$$|x^2 - 9| = |x - 3| \cdot |x + 3| < \delta \cdot (\delta + 6).$$

Now, if we pick δ to be the minimum of 1 and $\frac{\epsilon}{7}$, then we simultaneously guarantee that $\delta \leq \frac{\epsilon}{7}$ and $\delta + 6 \leq 7$, and so we find
$$|x^2 - 9| < \delta \cdot (\delta + 6) \leq \frac{\epsilon}{7} \cdot 7 = \epsilon,$$

as desired.

1.7.2   When limits may not exist

There are a few ways a limit might not exist:

1. A *discontinuity*, or jump, in the graph of the function. In this case, the limit does not exist because the limit from the left and the limit from the right are not equal.

2. A *blow-up*, when the function has a vertical asymptote.

3. An *oscillation*, where the graph of the function oscillates infinitely up and down as the input approaches a certain value.
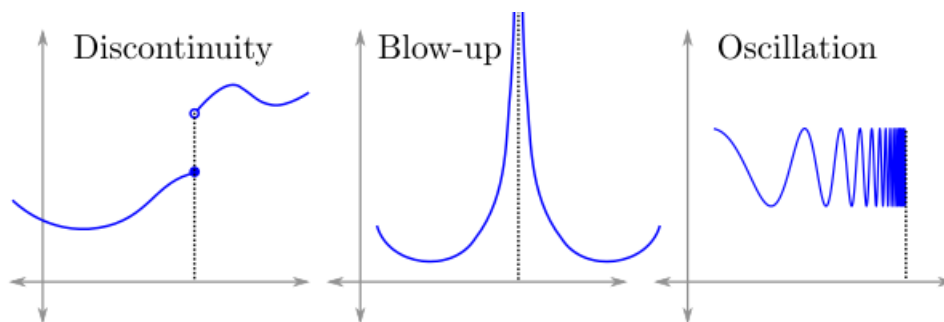
Most functions in this course will be well-behaved and will not have the above problems. The formal term for a well-behaved function is continuous.

1.7.3   Continuous functions

A function is continuous at the point (a) if the limit ($\lim_{x \to a} f(x)$) exists and ($\lim_{x \to a} f(x) = f(a)$). Intuitively, this says that there are no holes or jumps in the graph of (f) at (a).

Finally, a function is continuous if it is continuous at every point in its domain.

1.7.4   Rules for limits

There are rules for adding, multiplying, dividing, and composing limits. Suppose that $\lim\limits_{x \to a} f(x)$ and $\lim\limits_{x \to a} g(x)$ exist. Then

1. (Sum) $\lim\limits_{x \to a} (f + g)(x) = \lim\limits_{x \to a} f(x) + \lim\limits_{x \to a} g(x)$.

2. (Product) $\lim\limits_{x \to a} (f \cdot g)(x) = \left(\lim\limits_{x \to a} f(x) \right) \left(\lim\limits_{x \to a} g(x) \right)$.

3. (Quotient) $\lim\limits_{x \to a} \left(\dfrac{f}{g} \right)(x) = \dfrac{\lim_{x \to a} f(x)}{\lim_{x \to a} g(x)}$, provided that $\lim\limits_{x \to a} g(x) \neq 0$.

4. (Chain) $\lim\limits_{x \to a} (f \circ g)(x) = f \left(\lim\limits_{x \to a} g(x) \right)$, if f is continuous.

Almost all the functions encountered in this course are continuous, and so limits in most cases can be evaluated by simply plugging in the limiting input value into the function. The one case that sometimes gets complicated is the Quotient rule above when the limit of the denominator is 0.

Example Show that $\lim\limits_{x \to 0} \dfrac{\sin(x)}{x} = 1$.

There are several proofs of this limit (e.g. memorization, l'Hospital's rule), but the simplest method is to use the Taylor series. Because x is near 0, the Taylor series expansion for $\sin x$ applies, and so

$$
\begin{aligned}
\lim_{x \to 0} \frac{\sin(x)}{x} &= \lim_{x \to 0} \frac{x - \frac{x^3}{3!} + \cdots}{x} \\
&= \lim_{x \to 0} \frac{x\left(1 - \frac{x^2}{3!} + \cdots\right)}{x} \\
&= \lim_{x \to 0} 1 - \frac{x^2}{3!} + \cdots \\
&= 1.
\end{aligned}
$$

This works because all the terms involving x go to 0 as x goes to 0.

Example Find $\displaystyle\lim_{x \to 0} \frac{1 - \cos x}{x}$.

Replacing $\cos$ with its Taylor series (again, since x is near 0), we find

$$
\begin{aligned}
\lim_{x \to 0} \frac{1 - \cos x}{x} &= \lim_{x \to 0} \frac{1 - \left(1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots\right)}{x} \\
&= \lim_{x \to 0} \frac{\frac{1}{2!}x^2 - \frac{1}{4!}x^4 + \cdots}{x} \\
&= \lim_{x \to 0} \frac{1}{2!}x - \frac{1}{4!}x^3 + \cdots \\
&= 0.
\end{aligned}
$$

Example Compute $\displaystyle\lim_{x \to 0} \frac{\cos(x) - \sin(x) - 1}{e^x - 1}$.

Again, use the Taylor series about $x = 0$ for each function:

$$
\begin{aligned}
\lim_{x \to 0} \frac{\cos(x) - \sin(x) - 1}{e^x - 1} &= \lim_{x \to 0} \frac{\left(1 - \frac{x^2}{2!} + \cdots\right) - \left(x - \frac{x^3}{3!} + \cdots\right) - 1}{\left(1 + x + \frac{x^2}{2!} + \cdots\right) - 1} \\
&= \lim_{x \to 0} \frac{-x - \frac{x^2}{2!} + \cdots}{x + \cdots} \\
&= \lim_{x \to 0} \frac{x(-1 - \cdots)}{x(1 + \cdots)} \\
&= \lim_{x \to 0} \frac{-1 - \cdots}{1 + \cdots} \\
&= -1.
\end{aligned}
$$

Example Compute $\displaystyle\lim_{x \to 0} \frac{\sqrt[3]{1 + 4x} - 1}{\sqrt[5]{1 + 3x} - 1}$.

Here, we use the binomial series with $\alpha = \frac{1}{3}$ in the numerator, and $\alpha = \frac{1}{5}$ in the denominator.

We find

$$\lim_{x\to 0} \frac{(1+4x)^{1/3}-1}{(1+3x)^{1/5}-1} = \lim_{x\to 0} \frac{\left(1+\frac{1}{3}(4x)+ \text{ HOT}\right)-1}{\left(1+\frac{1}{5}(3x)+ \text{ HOT}\right)-1}$$

$$= \lim_{x\to 0} \frac{\frac{4}{3}x+ \text{ HOT}}{\frac{3}{5}x+ \text{ HOT}}$$

$$= \lim_{x\to 0} \frac{\frac{4}{3}+ \text{ HOT}}{\frac{3}{5}+ \text{ HOT}}$$

$$= \frac{\frac{4}{3}}{\frac{3}{5}}$$

$$= \frac{20}{9}$$

There are other methods for computing these types of limits, including memorization, algebraic tricks, and l'Hopital's rule (more on that in the next module). However, in many cases, these different methods can all be replaced by a simple application of Taylor series.

1.8   Exercises

Compute the following limits:

1. $\displaystyle\lim_{q\to 1} \frac{q^2+q+1}{q+3}$

2. $\displaystyle\lim_{x\to -2} \frac{x^2-4}{x+2}$

3. $\displaystyle\lim_{x\to 0} \frac{\sec x \tan x}{\sin x}$

4. $\displaystyle\lim_{x\to +\infty} \frac{6x^2-3x+1}{3x^2+4}$

5. $\displaystyle\lim_{x\to +\infty} \frac{x^2+x+1}{x^4-3x^2+2}$

6. $\displaystyle\lim_{y\to 0} \frac{\ln(1+2y)\sin y}{y^2\cos 2y}$

7. $\displaystyle\lim_{x\to 1} \frac{\ln x}{x^2}$

8. $\displaystyle\lim_{t\to 0} (3t^2+4t)\cot(t)$

9. $\displaystyle\lim_{z\to 0} \frac{z\cos(\sin(z))}{\sin(2z)}$

10. $\displaystyle\lim_{x\to 0} \frac{\ln(x+1)\arctan x}{x^2}$

11. $\displaystyle\lim_{x\to 0} \frac{\ln^2(\cos x)}{2x^4-x^5}$

12. $\displaystyle\lim_{s\to 0} \frac{e^s s \sin s}{1-\cos 2s}$

13. $\displaystyle\lim_{x\to 0^+} \frac{\sin(\arctan(\sin x))}{\sqrt{x}\sin 3x+x^2+\arctan 5x}$

14. $\displaystyle\lim_{x\to 0} \frac{\sin x-\cos x-1}{6xe^{2x}}$

15. $\displaystyle\lim_{x\to 0} \frac{\arctan x - 3\sin x + 2x}{3x^3}$

16. $\displaystyle\lim_{p\to 0} \frac{1 - p - \cos 3p}{p^3}$

17. $\displaystyle\lim_{x\to\infty} x^{1/x}$

1.9 L'Hopital's Rule

In previous modules, we saw that Taylor series are useful for computing certain limits of ratios. But sometimes, a fact known as L'Hopital's rule is easier to use than Taylor series. While L'Hopital's rule is commonly taught in a first calculus course, the justification for why it works is not usually taught. This module gives a justification for L'Hopital's rule, using Taylor series.

1.9.1 L'Hopital's rule

There are some limit situations where Taylor series are not particularly easy to use. For example, if the limit is being taken at a point about which the Taylor expansion is not already known, or the limit is at infinity, then using Taylor series is usually more work than it is worth. These are the situations where L'Hopital's rule can be helpful.

> **L'Hopital's Rule, $\frac{0}{0}$ case**
>
> If f and g are continuous functions such that $\displaystyle\lim_{x\to a} f(x) = 0$ and $\displaystyle\lim_{x\to a} g(x) = 0$, then $\displaystyle\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}$, provided this limit exists. If this is still of the form $\frac{0}{0}$, then derivatives may be taken again, and so on.

> The Taylor series for f and g about a are given by
>
> $$f(x) = f(a) + f'(a)(x-a) + \frac{f(a)}{2!}(x-a)^2 + \cdots$$
>
> $$g(x) = g(a) + g'(a)(x-a) + \frac{g(a)}{2!}(x-a)^2 + \cdots.$$
>
> Since, by hypothesis, $f(a) = g(a) = 0$, it follows that
>
> $$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f(a) + f'(a)(x-a) + \frac{1}{2}f(a)(x-a)^2 + \cdots}{g(a) + g'(a)(x-a) + \frac{1}{2}g(a)(x-a)^2 + \cdots}$$
>
> $$= \lim_{x\to a} \frac{f'(a)(x-a) + (1/2)f(a)(x-a)^2 + \cdots}{g'(a)(x-a) + (1/2)g(a)(x-a)^2 + \cdots}$$
>
> $$= \lim_{x\to a} \frac{(x-a)\left[f'(a) + (1/2)f(a)(x-a) + \cdots\right]}{(x-a)\left[g'(a) + (1/2)g(a)(x-a) + \cdots\right]}$$
>
> $$= \lim_{x\to a} \frac{f'(a) + (1/2)f(a)(x-a) + \cdots}{g'(a) + (1/2)g(a)(x-a) + \cdots}.$$
>
> Now, as $x \to a$, all the terms with $x - a$ go to 0, which leaves $\frac{f'(a)}{g'(a)}$. If this fraction is still $0/0$, then L'Hopital's rule says to take the derivative of the numerator and the denominator again. In terms of the Taylor series, this moves to the next leading terms in the numerator and denominator.

Example Using L'Hopital's rule, compute two of the limits from the last module:

$$\lim_{x\to 0} \frac{\sin x}{x}$$

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

These are both in the $\left(\frac{0}{0}\right)$ case, so differentiating numerator and denominator gives

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

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

Example Compute $\lim\limits_{x \to 0} \dfrac{\tan x}{\arcsin x}$.

Since $\tan 0 = \arcsin 0 = 0$, we are in the $\frac{0}{0}$ case of L'Hopital's rule. Recall that

$$\frac{d}{dx} \tan x = \sec^2 x$$
$$\frac{d}{dx} \arcsin x = \frac{1}{\sqrt{1 - x^2}}.$$

Thus, applying L'Hopital's rule gives

$$\lim_{x \to 0} \frac{\tan x}{\arcsin x} = \lim_{x \to 0} \frac{\sec^2 x}{\frac{1}{\sqrt{1-x^2}}}$$
$$= \frac{1}{1}$$
$$= 1.$$

Depending on the situation, it still might be easier to use Taylor series, especially if there are compositions and products of functions (assuming we know all the relevant Taylor series).

Example Compute $\lim\limits_{x \to 0} \dfrac{x^2 \ln(\cos x)}{\sin^2(3x^2)}$.

We know all the relevant Taylor series for the functions in this problem, so that should be an easier method. We find

$$\lim_{x \to 0} \frac{x^2 \ln(\cos x)}{\sin^2(3x^2)} = \lim_{x \to 0} \frac{x^2 \ln\left(1 - \frac{x^2}{2!} + \text{HOT}\right)}{(3x^2 + \text{HOT})^2}$$
$$= \lim_{x \to 0} \frac{x^2 \left(-\frac{x^2}{2} - \text{HOT}\right)}{9x^4 + \text{HOT}}$$
$$= \lim_{x \to 0} \frac{-\frac{x^4}{2} + \text{HOT}}{9x^4 + \text{HOT}}$$
$$= -\frac{1}{18}.$$

Using L'Hopital here would be quite a lot of work. It turns out that we would have to apply the rule four times, which involves a lot of product and chain rule.

L'Hopital's Rule, $\frac{\infty}{\infty}$ case

If $\lim\limits_{x \to a} f(x) = \lim\limits_{x \to a} g(x) = \infty$, then $\lim\limits_{x \to a} \dfrac{f(x)}{g(x)} = \lim\limits_{x \to a} \dfrac{f'(x)}{g'(x)}$, again provided this limit exists.

Example Compute $\lim\limits_{x \to 0} \dfrac{\ln x}{\frac{1}{x^2}}$.

Note that $\ln x \to -\infty$ as $x \to 0$, and $\frac{1}{x^2} \to \infty$ as $x \to 0$. Therefore, the $\frac{\infty}{\infty}$ case of L'Hopital's rule applies. Applying the rule,

$$\lim_{x \to 0} \frac{\ln x}{x^{-2}} = \lim_{x \to 0} \frac{\frac{1}{x}}{-2x^{-3}}$$

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

$$= 0.$$

Example Use L'Hopital's rule to compute $\displaystyle\lim_{x \to \pi} \frac{\sin(x)}{e^x \cos(x/2)}$.

Since $\sin, \cos, \exp$ are all continuous functions, and $\sin(\pi) = e^\pi \cos(\pi/2) = 0$, the hypotheses for L'Hopital's rule are met. So it follows that

$$\lim_{x \to \pi} \frac{\sin(x)}{e^x \cos(x/2)} = \lim_{x \to \pi} \frac{[\sin(x)]'}{[e^x \cos(x/2)]'}$$

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

$$= \frac{-1}{0 - (1/2)e^\pi \sin(\pi/2)}$$

$$= 2e^{-\pi}.$$

Note that although we know the Taylor series for these functions at $x = 0$, the limit here is as $x \to \pi$. Thus, we cannot use the Taylor series approach, because a Taylor series about $x = 0$ does not give a good approximation when x is not close to 0.

1.9.2 Other indeterminate forms

Some limits do not initially look like cases where L'Hopital's rule applies, but with some algebra they can be rearranged into one of the applicable cases. These are called indeterminate forms.

Case: $\infty - \infty$

First, consider $\displaystyle\lim_{x \to a} f(x) - g(x)$, where $\displaystyle\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = \infty$. Usually, one or both of f and g are ratios of other functions. In this case, getting a common denominator usually transforms the limit into one where L'Hospital's rule or a Taylor series approach applies.

Example Compute $\displaystyle\lim_{x \to 0} \frac{1}{\sin^2 x} - \frac{1}{x^2}$.

Getting a common denominator, and then Taylor expanding gives

$$\lim_{x \to 0} \frac{1}{\sin^2 x} - \frac{1}{x^2} = \lim_{x \to 0} \frac{x^2 - \sin^2 x}{x^2 \sin^2 x}$$

$$= \lim_{x \to 0} \frac{x^2 - \left(x - \frac{1}{3!}x^3 + \cdots\right)^2}{x^2 \left(x - \frac{1}{3!}x^3 + \cdots\right)^2}$$

$$= \lim_{x \to 0} \frac{x^2 - \left(x^2 - \frac{2}{3!}x^4 + \cdots\right)}{x^2 \left(x^2 - \frac{2}{3!}x^3 + \cdots\right)}$$

$$= \lim_{x \to 0} \frac{\frac{2}{6}x^4 + \cdots}{x^4 + \cdots}$$

$$= \frac{1}{3}.$$

Case: $\infty \cdot 0$

Next, consider $\lim\limits_{x \to a} f(x)g(x)$ where $\lim\limits_{x \to a} f(x) = \infty$ and $\lim\limits_{x \to a} g(x) = 0$. Since $\infty \cdot 0$ is not defined, it is not clear what this limit is. However, the product can be turned into one of the following ratios where L'Hopital's rule applies:

$$\lim_{x \to a} f(x)g(x) = \lim_{x \to a} \frac{g(x)}{1/f(x)}$$
$$= \lim_{x \to a} \frac{f(x)}{1/g(x)}$$

since dividing by the reciprocal of a number is the same as multiplying. Now, note that $\lim\limits_{x \to a} 1/f(x) = 0$, since $\lim\limits_{x \to a} f(x) = \infty$. Thus, $\lim\limits_{x \to a} \frac{g(x)}{1/f(x)}$ is now in the $\frac{0}{0}$ case of L'Hopital's rule.

Similarly, $\lim\limits_{x \to a} \frac{f(x)}{1/g(x)}$ is in the $\frac{\infty}{\infty}$ form of l'Hopital's rule, and so it can be applied here too.

Deciding which of the above forms to use depends on the situation, but in many situations either form will work.

Example Compute $\lim\limits_{x \to 0^+} x \ln x$. (Here we use the one-sided limit $x \to 0^+$ becuase $\ln$ is only defined on the positive real numbers).

This is of the form $0 \cdot (-\infty)$. In this case, it is easier to flip x into the denominator, because it is easier to take the derivative of x^{-1} than it is to take the derivative of $(\ln x)^{-1}$. So, we find

$$\lim_{x \to 0^+} x \ln x = \lim_{x \to 0^+} \frac{\ln x}{x^{-1}},$$

which is of the form $\frac{-\infty}{\infty}$, so applying L'Hopital's rule gives

$$\lim_{x \to 0^+} \frac{\ln x}{x^{-1}} = \lim_{x \to 0^+} \frac{x^{-1}}{-x^{-2}}$$
$$= \lim_{x \to 0^+} -x$$
$$= 0.$$

Case: ∞^0

Another indeterminate form arises when raising one function of x to a power which involves another function of x. Suppose $\lim\limits_{x \to a} f(x) = \infty$ and $\lim\limits_{x \to a} g(x) = 0$. Then what is $\lim\limits_{x \to a} f(x)^{g(x)}$?

On the one hand, it seems that raising ∞ to any power should be ∞. On the other hand, raising anything to the 0th power should be 1. To find what the answer actually is, let

$$y = \lim_{x \to a} f(x)^{g(x)},$$

and take the $\ln$ of both sides. Now, recall that $\ln$ is a continuous function. Therefore, from the rules of limits in the last module, taking $\ln$ of a limit is the same as the limit of the $\ln$:

$$\ln(y) = \ln\left(\lim_{x \to a} f(x)^{g(x)}\right)$$
$$= \lim_{x \to a} \ln\left(f(x)^{g(x)}\right)$$
$$= \lim_{x \to a} g(x) \ln(f(x))$$

(the last step uses the fact that $\ln(a^b) = b\ln(a)$). Now, this is of the form $0 \cdot \infty$, which was covered above. Note that when this limit is computed, it is $\ln(y)$ which has been found, and so the answer must be exponentiated to find y, the original limit.

Example Compute $\lim\limits_{x \to \infty} x^{1/x}$.

This is of the ∞^0 form. Let $y = \lim_{x \to \infty} x^{1/x}$. Then taking $\ln$ gives

$$
\begin{aligned}
\ln(y) &= \lim_{x \to \infty} \ln\left(x^{1/x}\right) \\
&= \lim_{x \to \infty} \frac{1}{x} \ln(x) \\
&= \lim_{x \to \infty} \frac{\ln x}{x} \\
&= \lim_{x \to \infty} \frac{1/x}{1} \\
&= 0
\end{aligned}
$$

(L'Hopital's rule was used in the second to last step).
So $\ln(y) = 0$, and so $y = 1$ is the answer.

Case: 0^0

Consider the limit $f(x)^{g(x)}$, where $\lim_{x \to a} f(x) = 0$ and $\lim_{x \to a} g(x) = 0$. Because 0^0 is not defined, this is another indeterminate form. It can be dealt with as the ∞^0 case, by first taking $\ln$, computing the resulting limit, and then exponentiating.

Example Compute $\lim_{x \to 0+} x^x$.

Letting $y = \lim_{x \to 0+} x^x$, and taking logarithms gives

$$
\begin{aligned}
\ln y &= \lim_{x \to 0+} x \ln x \\
&= 0,
\end{aligned}
$$

by an example above. Thus $y = e^0 = 1$.

1.9.3 Limits going to infinity

L'Hopital also works with limits going to infinity; the same hypothesis and conclusions hold. Before doing some examples, what does it mean for $\lim_{x \to \infty} f(x) = L$?

Limit at infinity $\lim_{x \to \infty} f(x) = L$ if and only if for every $\epsilon > 0$ there exists $M > 0$ such that $|f(x) - L| < \epsilon$ whenever $x > M$. If there is no such L, then the limit does not exist.

Example Compute $\lim_{x \to \infty} \frac{\ln x}{\sqrt{x}}$.

Both the numerator and denominator go to ∞ as $x \to \infty$, so applying L'Hopital's rule gives

$$
\begin{aligned}
\lim_{x \to \infty} \frac{\ln x}{\sqrt{x}} &= \lim_{x \to \infty} \frac{x^{-1}}{\frac{1}{2} x^{-1/2}} \\
&= \lim_{x \to \infty} \frac{2}{\sqrt{x}} \\
&= 0.
\end{aligned}
$$

Example Compute $\lim_{x \to \infty} \frac{e^x}{x^2}$.

The numerator and denominator both go to ∞ as $x \to \infty$, so this is the $\frac{\infty}{\infty}$ case of L'Hopital's

rule. It follows that

$$\lim_{x \to \infty} \frac{e^x}{x^2} = \lim_{x \to \infty} \frac{[e^x]'}{[x^2]'}$$
$$= \lim_{x \to \infty} \frac{e^x}{2x}$$
$$= \lim_{x \to \infty} \frac{e^x}{2}$$
$$= \infty.$$

Note that L'Hopital's rule was used twice here since $\lim_{x \to \infty} \frac{e^x}{2x}$ is still the $\frac{\infty}{\infty}$ case.

While L'Hopital often works, there are situations where it fails to give an answer, and a little extra thought must be employed.

Example Compute $\lim_{x \to \infty} \tanh x$.

Recall that $\tanh x = \frac{\sinh x}{\cosh x}$. Both $\sinh x \to \infty$ and $\cosh x \to \infty$ as $x \to \infty$. But applying L'Hopital's rule (and then applying it again) gives

$$\lim_{x \to \infty} \tanh x = \lim_{x \to \infty} \frac{\sinh x}{\cosh x}$$
$$= \lim_{x \to \infty} \frac{\cosh x}{\sinh x}$$
$$= \lim_{x \to \infty} \frac{\sinh x}{\cosh x},$$

so L'Hopital's rule clearly will not give us an answer here. Instead, writing out the definition of $\tanh x$ and doing a little algebra, we find

$$\lim_{x \to \infty} \tanh x = \lim_{x \to \infty} \frac{e^x - e^{-x}}{e^x + e^{-x}}$$
$$= \lim_{x \to \infty} \frac{e^x(1 - e^{-2x})}{e^x(1 + e^{-2x})}$$
$$= 1.$$

Example Compute $\lim_{x \to \infty} \frac{x \ln x}{\ln(\cosh x)}$.

Using L'Hopital's rule, we find

$$\lim_{x \to \infty} \frac{x \ln x}{\ln(\cosh x)} = \lim_{x \to \infty}$$
$$\frac{\frac{x}{x} + 1 \cdot \ln x}{\frac{\sinh x}{\cosh x}}$$
$$= \lim_{x \to \infty} \frac{1 + \ln x}{\tanh x}$$
$$= \infty,$$

since the denominator goes to 1 (from the previous example) and the numerator goes to infinity.

It is also possible to deal with limits going to infinity using Taylor series, but it involves some algebra. The idea is to use a substitution to turn the limit going to infinity into a limit going to zero. Symbolically, if $x \to \infty$, then let $z = 1/x$. It follows that $z \to 0$ as $x \to \infty$, and by replacing x with $1/z$ throughout, the limit is transformed.

$$\lim_{x \to \infty} f(x) = \lim_{z \to 0} f(1/z).$$

This process works when the limit at 0 exists. A more general technique would only look at the one-sided limit from the right-hand side:

$$\lim_{x \to \infty} f(x) = \lim_{z \to 0^+} f(1/z).$$

Example Compute $\displaystyle\lim_{x \to \infty} \frac{\sin(1/x)}{1/x}$.

Using the substitution $z = 1/x$, the limit becomes

$$\lim_{x \to \infty} \frac{\sin(1/x)}{1/x} = \lim_{z \to 0} \frac{\sin(z)}{z}$$
$$= 1,$$

as was shown in the last module.

1.9.4 Exercises

Compute the following limits. Should you use l'Hopital's rule or Taylor expansion?

1. $\displaystyle\lim_{x \to 2} \frac{x^3 + 2x^2 - 4x - 8}{x - 2}$

2. $\displaystyle\lim_{x \to \pi/3} \frac{1 - 2\cos x}{\pi - 3x}$

3. $\displaystyle\lim_{x \to \pi} \frac{4\sin x \cos x}{\pi - x}$

4. $\displaystyle\lim_{x \to 9} \frac{2x - 18}{\sqrt{x} - 3}$

5. $\displaystyle\lim_{x \to 0} \frac{e^x - \sin x - 1}{x^2 - x^3}$

6. $\displaystyle\lim_{x \to 1} \frac{\cos(\pi x/2)}{1 - \sqrt{x}}$

7. $\displaystyle\lim_{x \to 4} \frac{3 - \sqrt{5 + x}}{1 - \sqrt{5 - x}}$

8. $\displaystyle\lim_{x \to 0} \left(\frac{1}{x} - \frac{1}{\ln(x + 1)} \right)$

9. $\displaystyle\lim_{x \to \pi/2} \frac{\sin x \cos x}{e^x \cos 3x}$

10. $\displaystyle\lim_{x \to +\infty} \frac{\ln x}{e^x}$

11. $\displaystyle\lim_{x \to +\infty} x \ln \left(1 + \frac{3}{x} \right)$

12. $\displaystyle\lim_{x \to +\infty} \frac{(\ln x)(\sinh x)}{(x - 1)e^x}$

1.10 Orders of growth

When dealing with limits as $x \to 0$, it is the lowest order term (i.e. the term with the smallest power) which matters the most, since higher powers of x are very small when x is close to 0. On the other hand, as $x \to \infty$, what is known as the asymptotic growth of a function. In this case, it is the highest order term which matters the most. This module deals with limits of both types and provides a more formal notion of how quickly a function grows or shrinks.

1.10.1   Hierarchy of functions going to infinity

First, consider the monomial x^n, where n is a fixed, positive integer. Looking at the graphs of these monomials, it becomes clear that as $x \to \infty$, $x^{n+1} > x^n$:
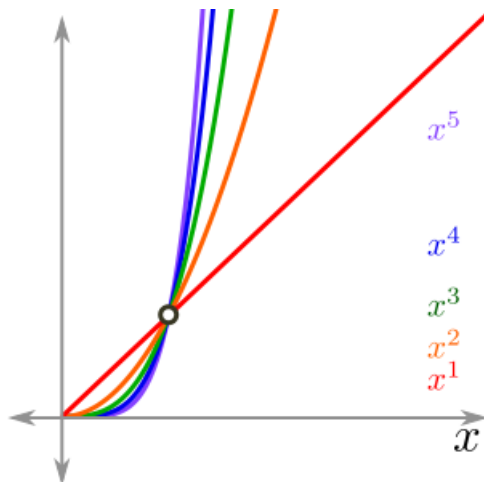
What happens when the functions involved are not polynomials? For example, how does the growth of the exponential compare to a polynomial? Or factorial and exponential?

In general, one can compare the asymptotic growth of the functions f and g by considering the limit $\lim_{x \to \infty} \frac{f(x)}{g(x)}$. If this limit is ∞, then f dominates. If the limit is 0, then g dominates. And if the limit is a constant, then f and g are considered equal (asymptotically).

Example Compare exponential growth and polynomial growth.

Fix an integer n, and compute the limit (using l'Hopital repeatedly):

$$
\begin{aligned}
\lim_{x \to \infty} \frac{x^n}{e^x} &= \lim_{x \to \infty} \frac{n x^{n-1}}{e^x} \\
&= \lim_{x \to \infty} \frac{n(n-1)x^{n-2}}{e^x} \\
&= \vdots \\
&= \lim_{x \to \infty} \frac{n!}{e^x} \\
&= 0
\end{aligned}
$$

(note that n is fixed, so $n!$ is a constant). Thus, the exponential dominates the monomial x^n (and thus any polynomial as well).

Example Compare the asymptotic behavior of $\ln^n x$ (for some fixed integer n) and x.

Again, using L'Hopital's rule repeatedly gives

$$
\begin{aligned}
\lim_{x \to \infty} \frac{\ln^n x}{x} &= \lim_{x \to \infty} \frac{n \ln^{n-1} x \frac{1}{x}}{1} \\
&= \lim_{x \to \infty} \frac{n \ln^{n-1} x}{x} \\
&= \vdots \\
&= \lim_{x \to \infty} \frac{n!}{x} \\
&= 0,
\end{aligned}
$$

since $n!$ is a constant here. This shows any polynomial beats any constant power of logarithm.

Example Compare the asymptotic growth of the factorial and the exponential.

First, when (x) is not an integer, the factorial is defined by

$$x! = \int_{t=0}^{\infty} t^x e^{-t} \, dt.$$

This definition shares properties with the traditional definition of factorial: $x! = x \cdot (x-1)!$, and $0! = 1$, and they coincide when x is an integer. It is not critical to know this integral definition, but know that $n! = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1$ for an integer n.

When x is an integer, note that e^x is e multiplied with itself x times. On the other hand, $x!$ has x factors, most of which are bigger than e (at least when x is bigger than, say, 5). So as x increases to $x+1$, e^x only gains another factor of e, but $x!$ gains a factor of $x+1$. This explains why $x!$ grows faster than e^x (and similarly for any exponential function).

Thus we have the following hierarchy of growth, from greatest to smallest:

1. Factorial: $x! = x(x-1)(x-2)\cdots 3 \cdot 2 \cdot 1$.

2. Exponential: c^x for any $c > 1$ (usually $c = e$).

3. Polynomial: x^k for any $k > 0$.

4. Logarithmic: $\ln(x)$ and other related functions.

When a pair of functions are of a similar type, such as e^x and $3^{\sqrt{x}}$, one must compare these using the limit of the ratio of the functions, as above.

1.10.2 Hierarchy of functions going to 0

As above, first consider monomials x^n. As $x \to 0$, the inequality for monomials is the reverse of what it was for $x \to \infty$. That is, as $x \to 0$, we find that $x^n > x^{n+1}$. Intuitively, small numbers become even smaller when you raise them to higher powers.

It is important to keep track of whether a limit is going to 0 or ∞, since in the first case, the lowest order terms dominate, and in the second case the highest order terms dominate.

Example Compute $\displaystyle\lim_{x\to 0} \frac{2x^3 - x^2 + x}{x^3 + 2x}$ and $\displaystyle\lim_{x\to\infty} \frac{2x^3 - x^2 + x + 1}{x^3 + 2x + 2}$.

As $x \to 0$, the higher powers of x go to 0 quickly, leaving the lowest order terms $\frac{x}{2x} = \frac{1}{2}$ in the limit. As $x \to \infty$, it is the highest order terms (the x^3 terms) which dominate, so ignoring the lower order terms leaves $\frac{2x^3}{x^3} = 2$ in the limit.

1.10.3 Big-O notation

When dealing with limits as $x \to 0$ or $x \to \infty$, it is best to have a formal notation for the approximations which result from dropping higher or lower order terms. Big-O notation, pronounced "big oh", provides this formality.

Big-O notation, $x \to 0$

The function $f(x)$ is in $O(x^n)$, as $x \to 0$ if

$$|f(x)| < C|x|^n$$

for some constant C and all x sufficiently close to 0. Put another way, a function $f(x)$ is in $O(x^n)$, for x close to 0, if $f(x)$ approaches 0 at least as fast as a constant multiple of x^n. More generally, a function $f(x)$ is in $O(g(x))$, as $x \to 0$ if

$$|f(x)| < C|g(x)|,$$

for some constant C and x sufficiently close to 0.

Big-O notation, as $x \to 0$, can be thought of as a more specific way of saying higher order terms. Just as Taylor series could include a different number of terms before indicating the rest is higher order terms (depending on the situation), the same is true for big-O.

Example Express $\arctan(x)$ using big-O notation as $x \to 0$.

By taking the Taylor series for $\arctan(x)$, we find that

$$\arctan(x) = x - \frac{x^3}{3} + O(x^5)$$

as $x \to 0$.
We could also say

$$\arctan(x) = x - O(x^3)$$

as $x \to 0$.

The definition for big-O as $x \to \infty$ is almost identical, except the bound needs to apply for all x sufficiently large.

Big-O notation, $x \to \infty$

The function $f(x)$ is in $O(x^n)$, as $x \to \infty$ if

$$|f(x)| < C|x|^n$$

for some constant C and all x sufficiently large. In other words, a function $f(x)$ is in $O(x^n)$, as $x \to \infty$, if $f(x)$ approaches infinity no faster than a constant multiple of x^n.
More generally, $f(x)$ is in $O(g(x))$, as $x \to \infty$ if

$$|f(x)| < C|g(x)|$$

for some constant C and all x sufficiently large.

Example The monomial x^n is in $O(e^x)$ as $x \to \infty$ for any (fixed) n. This is a restatement of the above fact that the exponential dominates polynomials as $x \to \infty$.

Example Show that $x\sqrt{x^2 + 3x + 5} = x^2 + \frac{3}{2}x + O(1)$ as $x \to \infty$. Hint: use the binomial series.

Recall that the binomial series

$$(1 + x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha - 1)}{2!}x^2 + \ldots$$

requires that $|x| < 1$. So we must do a little algebra to get the square root to be of this form.

Factoring out a x^2 will do the trick:

$$x\sqrt{x^2 + 3x + 5} = x\sqrt{x^2\left(1 + \frac{3}{x} + \frac{5}{x^2}\right)}$$

$$= x^2\left(1 + \frac{3}{x} + \frac{5}{x^2}\right)^{1/2}$$

$$= x^2\left(1 + \frac{1}{2}\left(\frac{3}{x} + \frac{5}{x^2}\right) + O\left(\frac{1}{x^2}\right)\right)$$

$$= x^2 + \frac{3}{2}x + O(1),$$

since all the terms involving $\frac{1}{x^2}$ become constants when multiplied by x^2.

Example Justify the following statements as $x \to 0$:

1. $5x + 3x^2$ is in $O(x)$ but is not in $O(x^2)$.

2. $\sin x$ is in $O(x)$ but is not in $O(x^2)$.

3. $\ln(1 + x) - x$ is in $O(x^2)$ but is not in $O(x^3)$.

4. $1 - \cos(x^2)$ is in $O(x^4)$ but is not in $O(x^5)$.

5. $\sqrt{x}$ is not in $O(x^n)$ for any $n \geq 1$.

6. e^{-1/x^2} is in $O(x^n)$ for all n.

The first four of these can be justified by looking at the Taylor series and then finding the monomial of the lowest power (remember that as $x \to 0$, the dominant term is that of the lowest power). To see that $\sqrt{x}$ is not in $O(x^n)$ for any $n \geq 1$, we must show that given any constant C and $\epsilon > 0$, there exists some $x < \epsilon$ such that

$$\sqrt{x} > Cx^n$$

(this is the negation of the definition of big oh). Solving the above inequality for x, one finds that if

$$x < \left(\frac{1}{C}\right)^{2/(2n-1)},$$

then $\sqrt{x} > Cx^n$, and so $\sqrt{x}$ is not in $O(x^n)$ for any n. Finally, to see that e^{-1/x^2} is in $O(x^n)$, it suffices to show that

$$e^{-1/x^2} < |x|^n$$

for all x sufficiently small. Taking natural log of both sides (this preserves the inequality since log is an increasing function) gives

$$-\frac{1}{x^2} < n \ln|x|.$$

Rearranging this (recall that multiplying an inequality by a negative number reverses the inequality) gives

$$\frac{1}{n} > -x^2 \ln|x|.$$

But recall from a previous module that $x \ln x \to 0$ as $x \to 0^+$. This ensures that no matter the (fixed) value of n, for sufficiently small x we will have

$$\frac{1}{n} > -x^2 \ln|x|,$$

and hence

$$e^{-1/x^2} < |x|^n,$$

as desired.

Example Justify the following statements as $x \to \infty$:

1. $\arctan x$ is in $O(1)$ as well as $O(x^n)$ for any $n \geq 0$.

2. $x\sqrt{1 + x^2}$ is in $O(x^2)$ but is not in $O(x^{3/2})$.

3. $\ln \sinh x$ is in $O(x)$ but is not in $O(\ln x)$.

4. $\cosh x$ is in $O(e^x)$ but is not in $O(x^n)$ for any $n \geq 0$.

5. $\ln(x^5)$ is in $O(\ln x)$ as well as $O(x^n)$ for all n.

6. x^x is in $O(e^{x^n})$ for all $n > 1$.

1. Note that as $x \to \infty$, $\arctan x \to \frac{\pi}{2}$. Thus $\arctan x$ is bounded, and so it is in $O(1)$.

2. Using the binomial series, as in a previous example, shows that

$$x\sqrt{1 + x^2} = x \left(x^2 \left(1 + \frac{1}{x^2} \right) \right)^{1/2} \qquad = x^2 \left(1 + \frac{1}{x^2} \right)^{1/2}$$
$$= x^2 \left(1 + \frac{1}{2}\frac{1}{x^2} + O\left(\frac{1}{x^4}\right) \right)$$
$$= x^2 + O(1).$$

Thus, $x\sqrt{1 + x^2}$ is in $O(x^2)$, but no power smaller than x^2.

3. For large x, e^{-x} is very small, and so $\sinh x \approx \frac{e^x}{2}$. Therefore,

$$\ln \sinh x \approx \ln \frac{e^x}{2} = x - \ln 2.$$

is in $O(x)$ but not in $O(\ln x)$, since logarithms are smaller than polynomials.

4. Similarly, $\cosh x \approx \frac{e^x}{2}$ for large x. Hence $\cosh$ is in $O(e^x)$ but cannot be bounded by any polynomial.

5. A handy property of logarithms tells us that

$$\ln(x^5) = 5\ln x,$$

which is in $O(\ln x)$ since it is itself a multiple of $\ln x$.

6. Fix $n > 1$. It suffices to show that for large enough x, we have

$$x^x < e^{x^n}.$$

Taking the logarithm of both sides gives

$$x \ln x < x^n \quad \Leftrightarrow \quad \ln x < x^{n-1}.$$

We saw earlier that any positive power of x beats the logarithm for large values of x. So (since $n > 1$) this inequality holds for large x, and so x^x is in $O(e^{x^n})$.

1.10.4   Application: Error Analysis

When approximating a function by the first few terms of its Taylor series, there is a trade-off between convenience (the ease of the computation) and accuracy. Big-O notation can help keep track of this error.

Example Consider the approximation, for x close to 0,

$$\sin(x^2)e^x = \left(x^2 + O(x^6) \right)\left(1 + O(x) \right)$$
$$= x^2 + x^2 \cdot O(x) + O(x^6) + O(x^6) \cdot O(x)$$
$$= x^2 + O(x^3).$$

So the error of approximating $\sin(x^2)e^x \approx x^2$ can be bounded by Cx^3, for some C when x is small. Determining a good C, in general, is tricky, and there is more on error bounds in Taylor Remainder Theorem.

1.10.5 Application: Computational Complexity

In computer science, an algorithm is a sequence of steps used to solve a problem. For example, there are algorithms to sort a list of numbers and algorithms to find the prime factorization of a number. Computational complexity is a measure of how efficient an algorithm is. Basically, a computer scientist wants to know roughly how the number of computations carried out by the computer will grow as the input (e.g. the length of the list to be sorted, or the size of the number to be factored) gets larger.

Example: Multiplication

Consider how much work is required to multiply two n-digit numbers using the usual grade-school method. There are two phases to working out the product: multiplication and addition.

First, multiply the first number by each of the digits of the second number. For each digit in the second number this requires n basic operations (multiplication of single digits) plus perhaps some "carries", so say a total of $2n$ operations for each digit in the second number. This means that the multiplication phase requires $n \cdot (2n)$ basic operations.

The addition phase requires repeatedly adding n digit numbers together a total of $n - 1$ times. If each addition requires at most $2n$ operations (including the carries), and there are $n - 1$ additions that must be made, it comes to a total of $(2n)(n - 1)$ operations in the addition phase.

Adding these totals up gives about $4n^2$ total operations. Thus, the total number of basic operations that must be performed in multiplying two n digit numbers is in $O(n^2)$ (since the constant coefficient does not matter).

The reason the constant coefficient does not really matter when thinking about computational complexity, is that a faster computer can only improve the speed of a computation by a constant factor. The only way to significantly improve a computation is to somehow drastically cut the number of operations required to perform the operation. The next example shows an example of how important algorithmic improvements can be on computational complexity.

Example: Sorting

In sorting algorithms, the most basic operation is the comparison. For a sorting algorithm, one wants to know how many comparisons of two numbers will be made, on average.

One common sorting algorithm, which is used by most people who are sorting items by hand, is called Insertion Sort. It turns out that the number of comparisons for Insertion Sort, on average, is $O(n^2)$ as $n \to \infty$, where n is the length of the list of numbers to be sorted.

A more sophisticated sorting algorithm, called Mergesort, uses $O(n \ln(n))$ comparisons on average. This may not seem like a significant improvement over Insertion Sort, but consider the number of comparisons used to sort a list of 1000000 integers: Insertion Sort would use on the order of $1000000^2 = 10^{12}$ comparisons on average, where as Mergesort uses on the order of 13×10^6 comparisons. To put that in perspective, if Mergesort took a half second to complete the computation, Insertion Sort would take over ten hours!

1.10.6 Big O in Other Areas of Mathematics

Stirling's formula

Stirling's formula gives an asymptotic approximation for $x!$:

$$\ln(x!) = x \ln x - x + O(\ln x).$$

In a slightly more precise form, it can be written

$$x! = \sqrt{2\pi x} \left(\frac{x}{e}\right)^x \left(1 + O\left(\frac{1}{x}\right)\right).$$

Prime number theorem

In number theory, one very important function, $\pi(x)$, is defined to be the number of primes less than or equal to x. The Prime Number Theorem says that

$$\pi(x) = \text{ number of primes} \leq x$$
$$= \frac{x}{\ln x}\left(1 + O\left(\frac{1}{\ln x}\right)\right).$$

1.10.7   Exercises

- Compute the following limits.

$$\lim_{x\to+\infty} \frac{e^{2x}}{x^3 + 3x^2 + 4}$$

$$\lim_{x\to+\infty} \frac{e^{3x}}{e^{x^2}}$$

$$\lim_{x\to+\infty} \frac{e^x(x-1)!}{x!}$$

$$\lim_{x\to+\infty} \frac{(x^2 - 3)(x^2 + 3)}{2x^4 - 2x^2 + 1}$$

$$\lim_{x\to+\infty} \frac{2^x + 1}{(x+1)!}$$

$$\lim_{x\to+\infty} \frac{(3\ln x)^n}{(2x)^n}$$

- Simplify the following asymptotic expression as $x \to 0$

$$f(x) = \left(x - x^2 + O(x^3)\right) \cdot \left(1 + 2x + O(x^3)\right)$$

- Simplify the following asymptotic expression as $x \to \infty$

$$f(x) = \left(x^3 + 2x^2 + O(x)\right) \cdot \left(1 + \frac{1}{x} + O\left(\frac{1}{x^2}\right)\right)$$

- Here are some rules for positive functions with $x \to \infty$:

$$O(f(x)) + O(g(x)) = O(f(x) + g(x))$$

$$O(f(x)) \cdot O(g(x)) = O(f(x) \cdot g(x))$$

Using these, show that

$$O\left(\frac{5}{x}\right) + O\left(\frac{\ln(x^2)}{4x}\right)$$

simplifies to

$$O\left(\frac{\ln x}{x}\right)$$

- Which of the following are in $O(x^2)$ as $x \to 0$?

$$x\ln(1 + x)$$

$$5x^2 + 6x + 1$$

$$1 - e^{-x}$$

$$x\sqrt{x^2 + 4x^3 + 5x^6}$$

$$x\sinh^2(3x)$$

$$\frac{x^2}{\ln(1 + x)}$$

Chapter 2

Differentiation

2.1 Derivatives

There are several definitions of the derivative of a function $f(x)$ at $x = a$. These definitions are all equivalent, but they are all important because they emphasize different aspects of the derivative.

Derivative (first definition)

$$f'(a) = \frac{df}{dx}\bigg|_{x=a} = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

If the limit does not exist, then the derivative is not defined at (a).

This first definition emphasizes that the derivative is the rate of change of the output with respect to the input. The next definition is similar.

Derivative (second definition)

$$f'(a) = \frac{df}{dx}\bigg|_{x=a} = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}.$$

If the limit does not exist, then the derivative is not defined at (a).

This definition can be interpreted as the change in output divided by the change in input, as the change in input goes to 0. One can see this is equivalent to the first definition by making the substitution $h = x - a$. The third definition looks quite different from the first two.

Derivative (third definition)

The derivative of $f(x)$ at $x = a$, $f'(a)$, is the constant C such that for any variation to the input h, the following holds.

$$f(a + h) = f(a) + Ch + O(h^2).$$

That is, $f'(a)$ is the first-order variation of the output. If no such C exists, then the derivative does not exist.

To show the equivalence, one can do a little algebra to see that

$$\frac{f(a + h) - f(a)}{h} = C + O(h).$$

Then taking the limit on both sides as ($h \to 0$) shows that ($C = f'(a)$).

Example Using the second and third definitions above, compute the derivative of $f(x) = x^n$, where n is a positive integer.

Using the binomial expansion and the above definition, one finds

$$
\begin{aligned}
f'(a) &= \lim_{h \to 0} \frac{(a+h)^n - a^n}{h} \\
&= \lim_{h \to 0} \frac{a^n + na^{n-1}h + O(h^2) - a^n}{h} \\
&= \lim_{h \to 0} \frac{na^{n-1}h + O(h^2)}{h} \\
&= \lim_{h \to 0} na^{n-1} + O(h) \\
&= na^{n-1}
\end{aligned}
$$

Using the third definition, and again the binomial expansion, one writes

$$
\begin{aligned}
f(a+h) &= (a+h)^n \\
&= a^n + na^{n-1}h + O(h^2),
\end{aligned}
$$

so $f'(a) = na^{n-1}$.

Example Find the derivative of e^x using the third definition.

Note that $e^{a+h} = e^a \cdot e^h$. Using our knowledge of the Taylor series for e^h, we have

$$
\begin{aligned}
e^{a+h} &= e^a e^h \\
&= e^a \left(1 + h + O(h^2)\right) \\
&= e^a + e^a h + O(h^2),
\end{aligned}
$$

and so the derivative of e^x, evaluated at $x = a$, is e^a.

Example Find the derivative of $\cos x$ using the third definition. Hint: use the identity

$$
\cos(a+h) = \cos(a)\cos(h) - \sin(a)\sin(h).
$$

Using the above identity and our knowledge of Taylor series, we find

$$
\begin{aligned}
\cos(a+h) &= \cos(a)\cos(h) - \sin(a)\sin(h) \\
&= \cos(a)\left(1 + O(h^2)\right) - \sin(a)\left(h + O(h^3)\right) \\
&= \cos(a) - \sin(a)h + O(h^2),
\end{aligned}
$$

so the derivative of $\cos x$, evaluated at $x = a$, is $-\sin(x)$.

Example Find the derivative of $f(x) = \sqrt{x}$ using the third definition.

First, write

$$
\begin{aligned}
f(a+h) &= \sqrt{a+h} \\
&= \sqrt{a}\sqrt{1 + \frac{h}{a}}.
\end{aligned}
$$

Now, recalling the binomial series $(1+x)^\alpha = 1 + \alpha x + O(x^2)$, we find

$$
\begin{aligned}
\sqrt{a}\sqrt{1 + \frac{h}{a}} &= \sqrt{a}\left(1 + \frac{1}{2}\frac{h}{a} + O(h^2)\right) \\
&= \sqrt{a} + \frac{1}{2\sqrt{a}}h + O(h^2),
\end{aligned}
$$

and so the derivative of $\sqrt{x}$, evaluated at $x = a$, is $\frac{1}{2\sqrt{a}}$.

2.1.1   Notation

There are several different notations for the derivative of $y = f(x)$. The best options are

$$\frac{df}{dx} \quad \text{or} \quad \frac{dy}{dx},$$

because they make it clear that the input is x and the output is $f(x)$ or y, respectively.

The next tier of options are fair, and have the advantage of requiring less writing, but they lose the benefit of knowing what the input variable is:

$$f' \quad \text{or} \quad \dot{y} \quad \text{or} \quad df.$$

The third option, df, is known as differential notation, which will be covered more in a later module.

Do not try to cancel the d's in the derivative. Do not write the d's in cursive, or replace the d's with Δ's (those notations have a different meaning).

2.1.2   Interpretations

The derivative is commonly interpreted as the slope of the tangent line to the graph of the function. This is fine when the function has one input and one output. But what happens in the (more realistic) situation of a function with more than one input and more than one output? How does one graph such a function? And if the units of the input and output are different, what is the unit of slope?

A better interpretation for the derivative is as the rate of change of output with respect to input. This interpretation makes sense with functions of many inputs and outputs. However, that will be covered in the multivariable sequel to this course.

2.1.3   Examples with respect to time

The most common use of the derivative is with respect to time. Here are several such examples from different areas.

Physics

Velocity $v(t)$ is the derivative of position $(x(t))$, with respect to time. Similarly, acceleration $a(t)$ is the derivative of velocity with respect to time:

$$v = \frac{dx}{dt} \quad \text{and} \quad a = \frac{dv}{dt}$$

Electromagnetism

Electric current, I, in a circuit is the rate of change of charge, Q, with respect to time:

$$I = \frac{dQ}{dt}.$$

Chemistry

The reaction rate for the product P, denoted r_P, in a chemical reaction is proportional to the rate of change of the concentration of P, denoted $[P]$, with respect to time:

$$r_P = k\frac{d[P]}{dt}.$$

2.1.4   Examples with respect to other variables

Spring constant

The spring constant k for a spring is the derivative of force with respect to deflection:

$$k = \frac{d(\text{force})}{d(\text{deflection})}.$$

Elasticity

The elasticity modulus λ of a material is the rate of change of stress with respect to strain:

$$\lambda = \frac{d(\text{stress})}{d(\text{strain})}.$$

Viscosity

The viscosity of a fluid μ is related to the shear stress by the equation

$$\text{shear stress} = \mu \frac{d(\text{velocity})}{d(\text{height})}.$$

Tax rates

The marginal tax rate is the rate of change of tax with respect to income:

$$\text{marginal tax rate} = \frac{d(\text{tax})}{d(\text{income})}.$$

2.1.5 Exercises

- A rock is dropped from the top of a 320-foot building. The height of the rock at time t is given as $s(t) = -8t^2 + 320$, where t is measured in seconds. Find the speed (that is, the absolute value of the velocity) of the rock when it hits the ground in feet per second. Round your answer to one decimal place.

- A very rough model of population size P for an ant species is $P(t) = 2\ln(t + 2)$, where t is time. What is the rate of change of the population at time $t = 2$?

- A particle's position, p, as a function of time, t, is represented by $p(t) = \frac{1}{3}t^3 - 3t^2 + 9t$. When is the particle at rest?

- Hooke's law states that the force F exerted by an "ideal" spring displaced a distance x from its equilibrium point is given by $F(x) = -kx$, where the constant k is called the "spring constant" and varies from one spring to another. In real life, many springs are nearly ideal for small displacements; however, for large displacements, they might deviate from what Hooke's law predicts. Much of the confusion between nearly-ideal and non-ideal springs is clarified by thinking in terms of series: for x near zero, $F(x) = -kx + O(x^2)$. Suppose you have a spring whose force follows the equation $F(x) = -2\tan 3x$. What is its spring constant?

- The profit, P, of a company that manufactures and sells N units of a certain product is modeled by the function $P(N) = R(N) - C(N)$. The revenue function, $R(N) = S \cdot N$, is the selling price S per unit times the number N of units sold. The company's cost, $C(N) = C_0 + C_{\text{op}}(N)$, is a sum of two terms. The first is a constant C_0 describing the initial investment needed to set up production. The other term, $C_{\text{op}}(N)$, varies depending on how many units the company produces, and represents the operating costs. Companies care not only about profit, but also "marginal profit", the rate of change of profit with respect to N. Assume that $S = \$50$, $C_0 = \$75,000$, $C_{\text{op}}(N) = \$50\sqrt{N}$, and that the company currently sells $N = 100$ units. Compute the marginal profit at this rate of production. Round your answer to one decimal place.

- In Economics, "physical capital" represents the buildings or machines used by a business to produce a product. The "marginal product of physical capital" represents the rate of change of output product with respect to physical capital (informally, if you increase the size of your factory a little, how much more product can you create?). A particular model tells us that the output product Y is given, as a function of capital K, by $Y = AK^\alpha L^{1-\alpha}$, where A is a constant, L is units of labor (assumed to be constant), and α is a constant between 0 and 1. Determine the marginal product of physical capital predicted by this model.

2.2   Differentiation rules

Recall from the previous module the third definition given of the derivative, that

$$f(x + h) = f(x) + \frac{df}{dx}h + O(h^2).$$

The derivative can be thought of as a multiplier: a change of h to the input gets amplified by a factor of $\frac{df}{dx}$ to become a change of $\frac{df}{dx}h$ to the output. This interpretation helps make sense of the following rules.

2.2.1   Differentiation rules

Suppose u and v are differentiable functions of x. Then the following rules (written using the shorthand differential notation) hold:

Linearity

$$d(u + v) = du + dv \quad \text{and} \quad d(c \cdot u) = c \cdot du,$$

where c is a constant.

Product

$$d(u \cdot v) = u \cdot dv + v \cdot du.$$

Chain

$$d(u \circ v) = du \cdot dv.$$

Linearity

Using the third definition of the derivative from the last module, we find

$$(u + v)(x + h) = u(x + h) + v(x + h)$$
$$= u(x) + \frac{du}{dx}h + O(h^2) + v(x) + \frac{dv}{dx}h + O(h^2)$$
$$= (u + v)(x) + \left(\frac{du}{dx} + \frac{dv}{dx} \right) h + O(h^2),$$

as desired. Similarly,

$$(c \cdot u)(x + h) = c \cdot u(x + h)$$
$$= c \left(u(x) + \frac{du}{dx}h + O(h^2) \right)$$
$$= (c \cdot u)(x) + \left(c\frac{du}{dx} \right) h + O(h^2).$$

Product

Again using the third definition of the derivative, we find

$$(u \cdot v)(x + h) = u(x + h) \cdot v(x + h)$$
$$= \left(u(x) + \frac{du}{dx}h + O(h^2) \right) \cdot \left(v(x) + \frac{dv}{dx}h + O(h^2) \right)$$
$$= u(x)v(x) + u(x)\frac{dv}{dx}h + v(x)\frac{du}{dx}h + O(h^2)$$
$$= (u \cdot v)(x) + \left(u(x)\frac{dv}{dx} + v(x)\frac{du}{dx} \right) h + O(h^2),$$

as desired.

Chain

The chain rule is justified similarly, with a little bit more algebra:

$$(u \circ v)(x + h) = u(v(x + h))$$
$$= u\left(v + \frac{dv}{dx}h + O(h^2)\right).$$

To simplify the notation temporarily, let $\tilde{h} = \frac{dv}{dx}h + O(h^2)$. Then

$$u(v + \tilde{h}) = u(v) + \frac{du}{dv}\tilde{h} + O(\tilde{h}^2)$$
$$= u(v) + \frac{du}{dv}\left(\frac{dv}{dx}h + O(h^2)\right) + O((\frac{dv}{dx}h + O(h^2))^2)$$
$$= u(v) + \frac{du}{dv} \cdot \frac{dv}{dx}h + O(h^2),$$

as desired.

Another common way to express the Chain rule, using the more traditional derivative notation, is

$$\frac{du}{dx} = \frac{du}{dv} \cdot \frac{dv}{dx}.$$

Caveat

Note that in the chain rule, the output of v is being plugged in as an input to u. Therefore, in the above, if the derivative $d(u \circ v)$ is being evaluated at $x = a$, then du is evaluated at $x = v(a)$ and dv is evaluated at $x = a$. More explicitly,

$$d(u \circ v)\Big|_{x=a} = du\Big|_{x=v(a)} \cdot dv\Big|_{x=a}.$$

Example Compute $\left(\frac{d}{dx}\left(e^{\sin x}\right)\right)$.

In the above notation, $u(x) = e^x$ and $v(x) = \sin x$, and the question asks for the derivative $d(u \circ v)$. Again, remembering to evaluate du at $v(x)$, one finds that

$$d\left(e^{\sin x}\right) = d(e^x)\Big|_{\sin(x)} d(\sin x)\Big|_x$$
$$= e^{\sin x} \cos x.$$

2.2.2   Other rules

There are a few other differentiation rules commonly taught in a first year calculus class, which can all be proven using the rules from above.

Reciprocal

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

Quotient

$$d\left(\frac{u}{v}\right) = \frac{vdu - udv}{v^2}.$$

Inverse

$$d(u^{-1}) = \frac{1}{du}\bigg|_{u^{-1}}$$

Note: u^{-1} is the inverse of u, not the reciprocal (which was covered above).

Reciprocal

One can think of this as an application of the chain rule, by writing

$$\frac{1}{v} = u \circ v$$

where $u(x) = \frac{1}{x}$. Or one can see it as a special case of the quotient rule.

Quotient

Let $w = \frac{u}{v}$, so that $u = w \cdot v$. By the product rule,

$$du = w \cdot dv + v \cdot dw.$$

Solving for dw, replacing w with $\frac{u}{v}$, and clearing fractions gives

$$
\begin{aligned}
dw &= \frac{du - w \cdot dv}{v} \\
&= \frac{du - \frac{u}{v}dv}{v} \\
&= \frac{v \cdot du - u \cdot dv}{v^2},
\end{aligned}
$$

as desired.

Inverse

Note that $x = u \circ u^{-1}$ by the definition of the inverse of a function. Differentiating both sides of this equation, using the chain rule on the right, gives

$$1 = du\bigg|_{u^{-1}} d(u^{-1})\bigg|_{x}$$

Then solving for $d(u^{-1})$ gives

$$d(u^{-1}) = \frac{1}{du}\bigg|_{u^{-1}}$$

as desired.

Example Show that

$$\frac{d}{dx}\sec x = \sec x \tan x,$$

using the reciprocal rule.

We find that

$$
\begin{aligned}
\frac{d}{dx}\sec x &= \frac{d}{dx}\left(\frac{1}{\cos x}\right) \\
&= -\frac{1}{\cos^2 x}(-\sin x) \\
&= \frac{1}{\cos x}\cdot\frac{\sin x}{\cos x} \\
&= \sec x \tan x.
\end{aligned}
$$

Example Show that

$$\frac{d}{dx}\tan x = \sec^2 x,$$

using the quotient rule.

We find that

$$\begin{aligned}
\frac{d}{dx}\tan x &= \frac{d}{dx}\frac{\sin x}{\cos x} \\
&= \frac{d(\sin x)\cos x - d(\cos x)\sin x}{\cos^2 x} \\
&= \frac{\sin^2 x + \cos^2 x}{\cos^2 x} \\
&= \frac{1}{\cos^2 x} \\
&= \sec^2 x.
\end{aligned}$$

Example Show that

$$\frac{d}{dx}\ln x = \frac{1}{x},$$

using the inverse derivative rule.

The inverse of the logarithm is the exponential, so $u = e^x$ in the inverse rule. Thus,

$$\begin{aligned}
\frac{d}{dx}\ln x &= \frac{1}{d(e^x)}\Big|_{\ln x} \\
&= \frac{1}{e^x}\Big|_{\ln x} \\
&= \frac{1}{x},
\end{aligned}$$

as desired.

2.2.3   Bonus

There are operators in other areas of mathematics which act similarly to the derivative. Finding such similarities in disparate fields is useful, because theorems from one field can often be carried over to the other and proved using similar techniques. These connections give a deeper understanding of both fields.

Boundary of spaces

Consider the *boundary* of a space. Loosely speaking, the boundary of a space is its outline, border, or edge. Think of the boundary as an operator, denoted ∂:
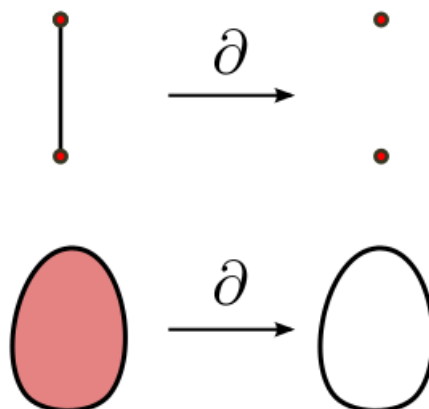
It turns out the the boundary operator ∂ acts similarly to a derivative. In particular, there is a product rule for ∂. But first, we need a notion of product for spaces.

The *Cartesian product* of two spaces X and Y, denoted $X \times Y$, is the set of ordered pairs (x, y) where $x \in X$ and $y \in Y$. This can be visualized by taking the first space and extruding it along the second space (easiest to visualize if the second space is a line segment), as in the following examples:
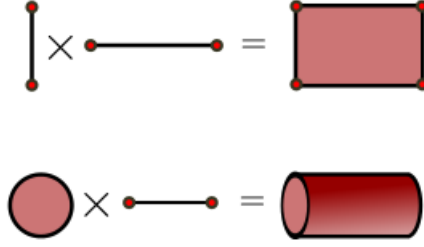
So the Cartesian product of two line segments is a filled in rectangle, and the Cartesian product of a disc and a line segment is a solid cylinder.

Now, consider the boundaries of these regions. The boundary of the filled in rectangle is its border, which can be thought of as the two vertical edges and the horizontal top and bottom edges. Note that this can be expressed as the boundary of the first segment times the second segment union the first segment times the boundary of the second segment:

Similarly, for the solid cylinder, the boundary is the union of the lateral area and the end caps. The lateral area is the Cartesian product of the circle with the line segment, which is the boundary of the disc times the line segment. The other piece (the end caps) is the disc times two points, which is the disc times the boundary of the line segment:


These examples suggest the (true) fact that for two spaces A and B, one has

$$\partial(A \times B) = \partial(A) \times B \cup A \times \partial(B).$$

Thinking of the boundary operator ∂ as the derivative, $\times$ as multiplication, and $\cup$ as addition, this is exactly like the product rule for functions given above. If you like this strange example, you may wish to take a course in Topology some day...

Lists

Consider a list of five distinct objects, all labeled with x. We might symbolize this by x^5. Now, consider the deletion operator D which deletes an object from the list. There are five different lists that might result (depending on which object was deleted), and this would logically be symbolized by $5x^4$, under the convention that a plus $+$ stands for "logical OR":
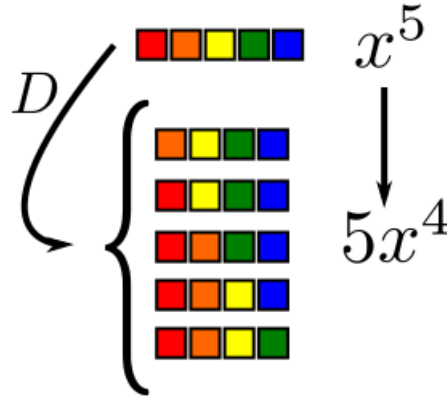
There is an entire calculus for lists and other grammatical constructs in Computer Science. As one simple example, how would we algebraicize the empty list? It has zero elements, so logically it should be expressed as $x^0 = 1$. Now, let $\mathcal{L}$ denote the collection of all finite lists. The trivial observation that any list is either empty or has a first element can be translated into an algebraic equation:

$$\mathcal{L} = 1 + x\mathcal{L}.$$

Here's the cool part...solve for $\mathcal{L}$ and you get:

$$\mathcal{L} = \frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \cdots,$$

the geometric series! Remember that plus mean "OR", so that this equation says any list is either the empty list, or the list of length 1, or the list of length 2, etc. If you find this sort of thing fun, you may wish to learn some Computer Science and some Analytic Combinatorics.

2.2.4   Exercises

- Find the derivative of $f(x) = \sqrt{x}(2x^2 - 4x)$.

- Find the derivative of $f(x) = 6x^4 - \dfrac{3}{x^2} - 2\pi$.

- Find the derivative of $f(x) = 7(x^3 + 4x)^5 \cos x$.

- Find the derivative of $f(x) = (e^x + \ln x) \sin x$.

- Find the derivative of $f(x) = \dfrac{\sqrt{x+3}}{x^2}$.

- Find the derivative of $f(x) = \dfrac{\ln x}{\cos x}$.

- Find the derivative of $f(x) = \dfrac{\sqrt[3]{x} - 4}{x^3}$.

- Find the derivative of $f(x) = \sin^3(x^3)$.

- Find the derivative of $f(x) = e^{-1/x^2}$.

- Use the information about functions f and g from the following table to compute the value of $\dfrac{d}{dx}\Big|_{x=1} g(f(x))$.

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

- Suppose that a certain quantity A is a function of another quantity B, which, in turn, depends on a third quantity C. We know that $B(C) = \sqrt{C}$. If the rate of change of A with respect to B is B^2, what is the rate of change of A with respect to C?

- This problem concerns the boundary operator ∂ from the bonus material. Denote by I the closed unit interval $[0, 1]$. Then, as observed, $\partial I = \{0\} \cup \{1\}$ is the union of two points. Let's get a little "creative". Denote by I^n the "n-cube", that is, the Cartesian product of n intervals: $I^n = I \times I \times \cdots \times I$. This is a well-defined and perfectly reasonable n-dimensional cube. (Just because you can't visualize doesn't mean it can't exist!) Note that $I^1 = I$ and I^0 is a single point (a zero-dimensional cube!). As a step towards building a "calculus of spaces", let us write $\partial I^1 = 2I^0 = 2$ as a way of saying that the boundary of an interval consists of two points and that $I^0 = 1$. The boundary of an n-dimensional cube consists of a certain number of $(n-1)$-dimensional cubes (called "faces"). For example, a square I^2 has four faces. Using what you know about derivatives, answer this: how many faces does I^n have?

2.3 Linearization

One of the main uses of the derivative is linearization, which uses the first two terms (the constant and linear term) of the Taylor series as an approximation. In many applications, this gives a very good approximation, as we will see in some examples.

2.3.1 Linear variation visualized

There are several geometric examples where it is possible to see the linear variation as the change in area as a parameter is changed by a small amount.

Square

The area of a square of side length x is given by $A(x) = x^2$. When that is varied by a small amount h, the result is

$$A(x + h) = (x + h)^2 = x^2 + 2xh + h^2.$$
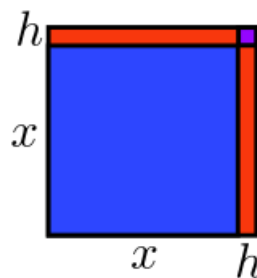
$$(x + h)^2 = x^2 + 2xh + h^2$$

The linear variation is $2xh$, which can be seen in the diagram as the rectangles along the right and top edges of the square. There are two of them, each with area xh. The final bit of area, the purple square in the diagram, has area h^2, which is higher order.

Triangle

The area of a right triangle with legs length x is $A(x) = \frac{1}{2}x^2$. When the leg is varied by h, the result is

$$A(x+h) = \frac{1}{2}x^2 + xh + \frac{1}{2}h^2.$$

$$\frac{1}{2}(x+h)^2 = \frac{1}{2}x^2 + xh + \frac{1}{2}h^2$$

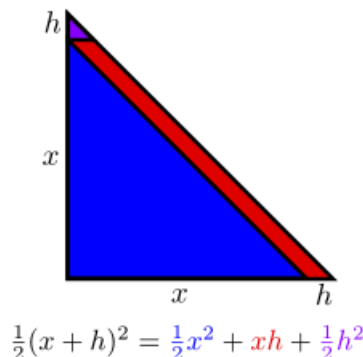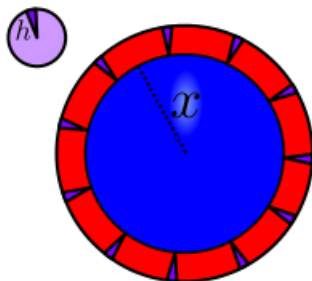Visually, the linear variation xh comes from the red parallelogram, of base h and height x, running along the hypotenuse. The higher order term $\frac{1}{2}h^2$ comes from the small purple triangle at the tip of the triangle.

Disc

The area of a disc of radius x is $A(x) = \pi x^2$. If the radius is increased by h, the result is

$$A(x+h) = \pi(x+h)^2 = \pi x^2 + 2\pi xh + \pi h^2.$$

$$\pi(x+h)^2 = \pi x^2 + 2\pi xh + \pi h^2$$

Visually breaking this into the linear variation and higher order variation is a little bit harder. The best way is to imagine taking the ring formed by the increased radius and breaking it into rectangles and wedges. In the limit, the wedges can be rearranged into a disc of radius h, and the rectangles can be arranged to form a strip of length $2\pi x$ (the circumference of the inner circle) and width h.

2.3.2　Linear approximation

The equation underlying any linear approximation should be familiar, since it is just the first order Taylor series about $x = a$, after making the substitution $h = x - a$:

$$f(a+h) \approx f(a) + f'(a)h.$$

This will be a good linear approximation provided that h is small, i.e., the point a is close to the input we are trying to approximate. In general, one wants to pick a to be an input where it is easy to compute $f(a)$ and $f'(a)$ which is as close to the desired input as possible.

Example Using a linear approximation, estimate $\sqrt{250}$.

The function here is $f(x) = \sqrt{x}$. Possible choices for a are perfect squares, because it is easy to compute the square root of squares. The nearest perfect square is $256 = 16^2$, so we choose $a = 256$. Thus, $h = x - a = -6$. Then

$$f(x) \approx f(a) + f'(a)h\sqrt{250}$$
$$\approx \sqrt{256} + \frac{1}{2\sqrt{256}}(-6)$$
$$= 16 - \frac{6}{32}$$
$$= 15\frac{13}{16}$$
$$\approx 15.8,$$

which is very close to the calculator's answer of $15.811\ldots$.

Example Using a linear approximation, estimate $\sqrt{104}$. Is this an over-approximation or an under-approximation?

As above, $a = 100$ is the closest point where it is easy to compute $\sqrt{x}$ and derivatives. Then $h = 4$, so the linear approximation is

$$f(a + h) \approx f(a) + f'(a)h$$
$$= \sqrt{100} + \frac{1}{2\sqrt{100}} \cdot 4$$
$$= 10 + \frac{4}{20}$$
$$= 10.2.$$

This is an over-approximation. One way to see why is to consider the graph of $\sqrt{x}$, which is concave down, so the linear approximation is above the true value. Another argument is to consider the next term of the Taylor series, which is negative (since the second derivative of $\sqrt{x}$ is negative). For comparison, the value according to a calculator is $\sqrt{104} \approx 10.198$.

Example Approximate π^{20}. Hint: $\pi^2 \approx 9.86$.

From the hint, $\pi^{20} = (\pi^2)^{10} \approx 9.86^{10}$. Thus, we are trying to approximate $f(x) = x^{10}$ at $x = 9.86$. The nearest easy input is $a = 10$, so we find

$$f(9.86) \approx f(10) + f'(10)(-.14)$$
$$= 10^{10} + 10(10)^9(-.14)$$
$$= 10^{10}(1 - .14)$$
$$= 8.6 \cdot 10^9.$$

The true answer is approximately $8.77 \cdot 10^9$, so this estimate is within 2%.

Example Approximate e^{30}. Hint: $e^3 \approx 20.1$, and $2^{10} \approx 1000$.

From the first hint, $e^{30} = (e^3)^{10} \approx (20.1)^{10}$, so consider the linear approximation for $f(x) = x^{10}$ near $a = 20$:

$$f(x) \approx f(20) + f'(20)(x - 20)$$
$$x^{10} \approx 20^{10} + 10 \cdot 20^9(x - 20)$$

So it follows that

$$e^{30} \approx (20.1)^{10}$$
$$\approx 20^{10} + 10 \cdot 20^9 \cdot (.1)$$
$$\approx 20^{10} \left[1 + (.5)(.1) \right]$$
$$\approx 2^{10} \cdot 10^{10} \cdot 1.05$$
$$\approx 1.05 \cdot 10^{13}$$

(the last step used the second hint that $2^{10} \approx 10^3$). The true answer is approximately $e^{30} \approx 1.068 \cdot 10^{13}$, so the error is less than 2%.

2.3.3   Newton's method

Another application of linearization gives a way of approximating the root of a function. This is called Newton's method.

Given a continuous, differentiable function f, the goal is to find a root (i.e. a value a such that $f(a) = 0$). Suppose x is an initial guess of a root. Then $x + h$ will be a root for some small value of h. Linearizing, and ignoring the higher order terms, gives

$$f(x + h) = f(x) + f'(x)h + O(h^2)$$
$$\approx f(x) + f'(x)h.$$

Since we supposed that $x + h$ was a root of f, it follows that $f(x + h) = 0$. Therefore, setting the above equal to 0 and solving for h gives

$$f(x) + f'(x)h = 0 \quad \Rightarrow \quad h = -\frac{f(x)}{f'(x)}.$$

Thus, an even closer guess of a root is

$$x + h = x - \frac{f(x)}{f'(x)}.$$

By taking this new guess of a root, and repeating the above process, one (hopefully) gets a better and better approximation of a root.

More formally, this is what is called a difference equation. Given an initial guess, called x_0, of a root of the function, one uses the update rule

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

to get x_1, and then x_2, and so on.

The resulting sequence hopefully converges to a root of f. Graphically, what is happening is as follows:

1. Pick a guess x_0.

2. Find the tangent line to f through the point $(x_0, f(x_0))$.

3. Let x_1 be the point where the tangent line intersects the x-axis.

4. Repeat steps 2 and 3 (see the figure).
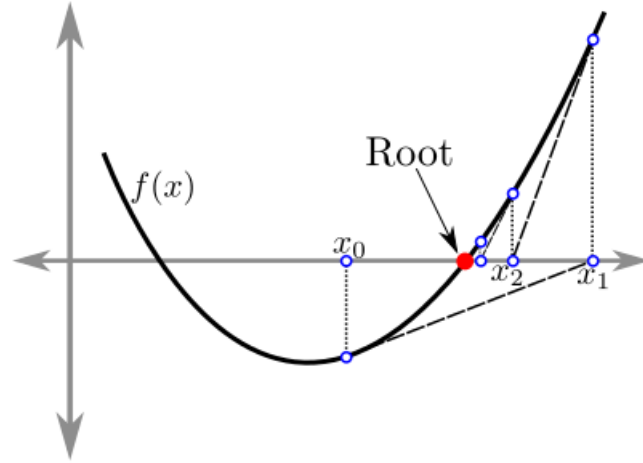
Caveat

This sequence is only defined if $f'(x_n)$ exists and is non-zero for every x_n in the sequence. Even if the sequence is defined, it may not converge to anything. But if the sequence is defined and it does converge, say to L, then L is a root of f.

Example Find the update rule for approximating $\frac{1}{a}$, the reciprocal of a number a.

First, we must find a function which has $\frac{1}{a}$ as a root. We might try $f(x) = x - \frac{1}{a}$, but unfortunately (after a little algebra) this leads to the update rule

$$x_{n+1} = \frac{1}{a},$$

which is not particularly helpful, since that is the quantity we are trying to approximate. Another try would be $f(x) = a - \frac{1}{x}$. This will work. Note that $f'(x) = \frac{1}{x^2}$, so the update rule is

$$\begin{aligned}
x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\
&= x_n - \frac{a - \frac{1}{x_n}}{\frac{1}{x_n^2}} \\
&= x_n - \left(x_n^2 a - x_n\right) \\
&= 2x_n - ax_n^2.
\end{aligned}$$

Note that this rule does not involve division, but only multiplication and subtraction.

Example Find the update rule for using Newton's method to approximate $\sqrt{a}$. Use the update rule twice with initial guess $x_0 = 3$ to estimate $\sqrt{11}$.

The first step is to find a function whose root is $\sqrt{a}$. A good choice is $f(x) = x^2 - a$. Then according to Newton's method, the update rule is

$$\begin{aligned}
x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\
&= x_n - \frac{x_n^2 - a}{2x_n} \\
&= x_n - \frac{x_n}{2} + \frac{a}{2x_n} \\
&= \frac{x_n}{2} + \frac{a}{2x_n}.
\end{aligned}$$

Using the update rule with $x_0 = 3$ and $a = 11$ gives

$$
\begin{aligned}
x_1 &= \frac{x_0}{2} + \frac{11}{2x_0} \\
&= \frac{3}{2} + \frac{11}{6} \\
&= \frac{20}{6} \\
&= \frac{10}{3}
\end{aligned}
$$

Updating one more time gives

$$
\begin{aligned}
x_2 &= \frac{x_1}{2} + \frac{11}{2x_1} \\
&= \frac{5}{3} + \frac{11}{\frac{20}{3}} \\
&= \frac{199}{60} \\
&\approx 3.3166
\end{aligned}
$$

The calculator gives that $\sqrt{11} \approx 3.3166$, so we get a good approximation with only a little bit of work. One could repeat the update rule several more times to get an even better approximation.

Example Find the update rule for finding $\sqrt[3]{a}$. Use the update rule once with the initial guess of $x_0 = 5$ to estimate $\sqrt[3]{100}$.

The function in this case is $f(x) = x^3 - a$. Then $f'(x) = 3x^2$, and so the update rule is

$$
\begin{aligned}
x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\
&= x_n - \frac{x_n^3 - a}{3x_n^2} \\
&= x_n - \frac{1}{3}x_n + \frac{a}{3x_n^2} \\
&= \frac{2}{3}x_n + \frac{a}{3x_n^2}.
\end{aligned}
$$

Using this to estimate $\sqrt[3]{100}$ with the initial guess $x_0 = 5$, one finds

$$
\begin{aligned}
x_1 &= \frac{2}{3} \cdot x_0 + \frac{100}{3x_0^2} \\
&= \frac{10}{3} + \frac{4}{3} \\
&= \frac{14}{3} \\
&\approx 4.666
\end{aligned}
$$

If we compare this to the answer from a calculator, we find that $\sqrt[3]{100} \approx 4.64$, and so even after just one step, we are within 1% of the true answer.

2.3.4   Exercises

- Use a linear approximation to estimate $\sqrt[3]{67}$. Round your answer to four decimal places. Hint: what is 4^3?

- Use a linear approximation to estimate the cosine of an angle of $66°$. Round your answer to four

decimal places. Hint: remember that $60° = \dfrac{\pi}{3}$, and hence $6° = \dfrac{\pi}{30}$.

- Use Newton's method to determine the intersection of e^{-x} and x.

- The golden ratio $\varphi = \dfrac{1 + \sqrt{5}}{2}$ is a root of the polynomial $x^2 - x - 1$. If you use Newton's method to estimate its value, what is the appropriate update rule for the sequence x_n ?

- To approximate $\sqrt{10}$ using Newton's method, what is the appropriate update rule for the sequence x_n?

- You want to build a square pen for your new chickens, with an area of $1200\,\text{ft}^2$. Not having a calculator handy, you decide to use Newton's method to approximate the length of one side of the fence. If your first guess is $30\,\text{ft}$, what is the next approximation you will get?

- You are in charge of designing packaging materials for your company's new product. The marketing department tells you that you must put them in a cube-shaped box. The engineering department says that you will need a box with a volume of $500\,\text{cm}^3$. What are the dimensions of the cubical box? Starting with a guess of $8\,\text{cm}$ for the length of the side of the cube, what approximation does one iteration of Newton's method give you? Round your answer to two decimal places.

- Without using a calculator, approximate 9.98^{98}. Here are some hints. First, 9.98 is close to 10, and $10^{98} = 1\,\text{E}\,98$ in scientific notation. What does linear approximation give as an estimate when we decrease from 10^{98} to 9.98^{98}?

- A diving-board of length L bends under the weight of a diver standing on its edge. The free end of the board moves down a distance $D = PL^3/3EI$ where P is the weight of the diver, E is a constant of elasticity (that depends on the material from which the board is manufactured), and I is a moment of inertia. (These last two quantities will again make an appearance in Lectures 13 and 41, but do not worry about what exactly they mean now...) Suppose our board has a length $L = 2\,\text{m}$, and that it takes a deflection of $D = 20\,\text{cm}$ under the weight of the diver. Use a linear approximation to estimate the deflection that it would take if its length was increased by $20\,\text{cm}$.

2.4 Higher derivatives

The nth derivative of a function $f(x)$, denoted $f^{(n)}(x)$ or $\dfrac{d^n}{dx^n}(f)$, is defined recursively by

$$f^{(n)}(x) = \frac{d}{dx} f^{(n-1)}(x).$$

In other words, the nth derivative is what one gets by taking the derivative n times. Note that in the $\frac{d}{dx}$ notation, the power n goes to the right of dx, to emphasize the fact that the nth derivative of f is achieved by iterating n times the operator $\frac{d}{dx}$. So

$$\left(\frac{d}{dx}\right)^n f = \frac{d^n}{dx^n} f.$$

2.4.1 Interpretations

Let $x(t)$ denote the position of a moving body as a function of time. Then the velocity $v(t)$ of the body is

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

The acceleration of an object is the second derivative of its position function (i.e. the derivative of its velocity):

$$a(t) = \frac{dv}{dt} = \frac{d}{dt}\left(\frac{dx}{dt}\right) = \frac{d^2 x}{dt^2}.$$

The jerk of an object is the third derivative of its position function (i.e. the derivative of its acceleration):

$$j(t) = \frac{d^3 x}{dt^3}.$$

The snap (or jounce) of an object is the fourth derivative of its position:

$$s(t) = \frac{d^4 x}{dt^4}.$$

Quadrotors

The maneuverability of nano quadrotors depends on controlling both the jerk and the snap (in addition to velocity and acceleration).

Curvature

In geometry, curvature can (informally) be thought of as how quickly the graph of the function curves. Consider the largest circle that can comfortably sit tangent to the graph of f at the point a. Intuitively, the larger the radius R of the circle that fits, the smaller the curvature. Here is a curve with several of these circles (called osculating circles) drawn it at different indicated points.
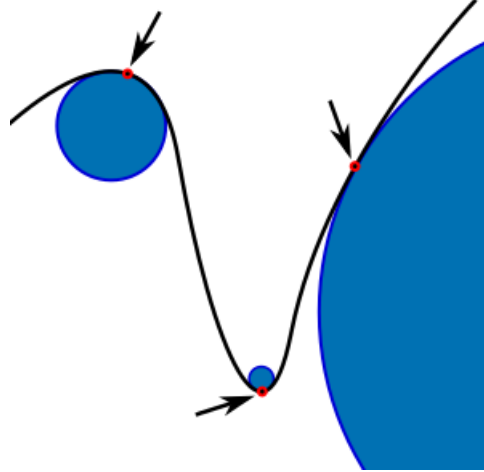

In fact, the curvature of a curve f, denoted κ, is defined by $\kappa = \frac{1}{R}$, where R is the radius of the largest circle which fits the curve to second order. With some algebra, one finds the following expression for curvature in terms of the first and second derivatives of f:

Curvature of a function f

$$\kappa = \frac{|f''|}{(1 + (f')^2)^{3/2}}.$$

For a given point on the curve, draw its osculating circle, say of radius R (right now, R is unknown to us; we will eventually find R). Then place the coordinate axes so that the origin is at the center of the circle (note that the curvature at a given point only depends on the radius of the osculating circle, which is independent of where the axes are placed):
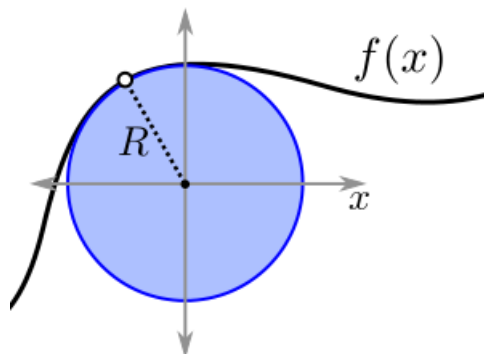
The equation of the osculating circle is $x^2 + y^2 = R^2$. Solving for y gives

$$y = \sqrt{R^2 - x^2} = (R^2 - x^2)^{1/2},$$

whose first and second derivatives should respectively match the first and second derivatives of the function f at the point (that is what it means for the circle to match the function up to second order).

Remember that the first derivative and second derivative of f at the given point are just constants. We will set the derivative and second derivative of the equation of the circle equal to these constants, respectively, and then solve for R. The first derivative of the equation of the circle is

$$\frac{d}{dx}(R^2 - x^2)^{1/2} = \frac{1}{2}(R^2 - x^2)^{-1/2}(-2x)$$
$$= -x(R^2 - x^2)^{-1/2}.$$

The second derivative of the equation of the circle (using the product rule) is

$$\frac{d}{dx}\left(-x(R^2 - x^2)^{-1/2}\right) = -(R^2 - x^2)^{-1/2} - x\left(-\frac{1}{2}\right)(R^2 - x^2)^{-3/2}(-2x)$$
$$= -(R^2 - x^2)^{-1/2} - x^2(R^2 - x^2)^{-3/2}$$
$$= -\frac{1}{(R^2 - x^2)^{1/2}} - \frac{x^2}{(R^2 - x^2)^{3/2}}$$
$$= -\frac{(R^2 - x^2)}{(R^2 - x^2)^{3/2}} - \frac{x^2}{(R^2 - x^2)^{3/2}}$$
$$= \frac{-R^2}{(R^2 - x^2)^{3/2}}$$
$$= -R^2(R^2 - x^2)^{-3/2}.$$

So setting the corresponding derivatives of f equal to the derivatives of the circle gives

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

Now we do some algebra to solve for R in terms of f' and f''. Squaring the first equation gives

$$(f')^2 = x^2(R^2 - x^2)^{-1}.$$

Solving this equation for x^2 gives

$$x^2 = \frac{(f')^2 R^2}{1 + (f')^2}.$$

Plugging this into the second equation, and doing some algebra gives

$$f'' = -R^2(R^2 - x^2)^{-3/2}$$
$$= -R^2\left(R^2 - \frac{(f')^2 R^2}{1 + (f')^2}\right)^{-3/2}$$
$$= -R^2\left[R^2\left(1 - \frac{(f')^2}{1 + (f')^2}\right)\right]^{-3/2}$$
$$= \frac{-R^2}{R^3}\left(\frac{1 + (f')^2 - (f')^2}{1 + (f')^2}\right)^{-3/2}$$
$$= -\frac{1}{R}\left(1 + (f')^2\right)^{3/2}.$$

Taking absolute values (since the radius of a circle should not be negative) gives

$$|f''| = \frac{1}{R}\left(1 + (f')^2\right)^{3/2}.$$

Now, solving for $\kappa = \frac{1}{R}$ gives

$$\kappa = \frac{1}{R} = \frac{|f''|}{(1 + (f')^2)^{3/2}},$$

as desired.

Note that for a straight line, the second derivative $f'' = 0$, and so $\kappa = 0$, which matches intuition. In this case, the osculating circle is infinite. Similarly, $\kappa = 0$ at inflection points of f.

Elasticity

Consider an elastic beam with uniform cross section and static load $q(x)$, where x is the location of the load along the beam. Then the deflection $u(x)$ (the amount the beam sags at location x) satisfies the equation

$$EI\frac{d^4u}{dx^4} = q(x),$$

where E and I are constants: E is the constant of elasticity (depends on the material), and I is the moment of inertia (depends on the shape of the beam).

Taylor series

As seen in previous modules, information about the derivatives of a function evaluated at a single point gives information about the function for inputs near that point via the Taylor series:

$$f(x) = f(a) + f'(a)(x - a) + \frac{f(a)}{2!}(x - a)^2 + \frac{f'(a)}{3!}(x - a)^3 + \cdots$$

Taking a few terms of this series gives a polynomial which is a good approximation of f near a. The more derivatives one knows, the more terms one can include in the series, and the better the approximation.

2.4.2   Bonus: Another look at Taylor series

Consider the alternative way to express the Taylor series, in terms of the distance h from the base point a:

$$\begin{aligned}
f(a + h) &= f(a) + f'(a)h + \frac{f(a)}{2!}h^2 + \frac{f'(a)}{3!}h^3 + \cdots \\
&= \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}h^k \\
&= \sum_{k=0}^{\infty} \frac{h^k}{k!}\left(\left.\frac{d}{dx}\right|_a\right)^k f \\
&= \sum_{k=0}^{\infty} \frac{1}{k!}\left(h\left.\frac{d}{dx}\right|_a\right)^k f.
\end{aligned}$$

Note that this resembles the Taylor series for the exponential e^x, where

$$x = h\left.\frac{d}{dx}\right|_a.$$

This may seem a little unusual. But the idea of exponentiating an operator to get another operator is a useful tool, which comes up in other areas of mathematics. In this notation, we can write

$$f(a + h) = e^{\left(h\left.\frac{d}{dx}\right|_a\right)} f.$$

Another way to think of this is that $e^{h\frac{d}{dx}}$ is the shift operator, which takes in the function $f(x)$ and gives back the function $f(x + h)$.

2.4.3   Exercises

- You are given the position, velocity and acceleration of a particle at time $t = 0$. The position is $p(0) = 2$, the velocity $v(0) = 4$, and the acceleration $a(0) = 3$. Using this information, which Taylor series should they use to approximate $p(t)$, and what is the estimated value of $p(4)$ using this approximation?

- If a particle moves according to the position function $s(t) = t^3 - 6t$, what are its position, velocity and acceleration at $t = 3$?

- If the position of a car at time t is given by the formula $p(t) = t^4 - 24t^2$, for which times t is its velocity decreasing?

- What is a formula for the second derivative of $f(t) = t^2 \sin 2t$? Use this formula to compute $f''(\pi/2)$.

- Use a Taylor series expansion to compute the third derivative of $f(x) = \sin^3(\ln(1+x))$ at zero.

- What is the curvature of the graph of the function $f(x) = -2\sin(x^2)$ at the point $(0,0)$?

2.5   Optimization

One of the most important applications of derivatives is optimization. In some introductory calculus classes these types of problems are called max/min problems: given a function, what is the maximum or minimum output subject to some constraints. This module will review how derivatives can be used in these problems and give some of the reasons why these methods work.

2.5.1   Critical points

First, observe that for a differentiable function f, if the derivative is not zero at a point, then that point cannot be a maximum or a minimum. For instance, if the derivative is positive, then the output is increasing with respect to the input, so by increasing the input, one can increase the output. Hence, the point is not a maximum. If the derivative is negative, then decreasing the input will increase the output, so that point cannot be a maximum.
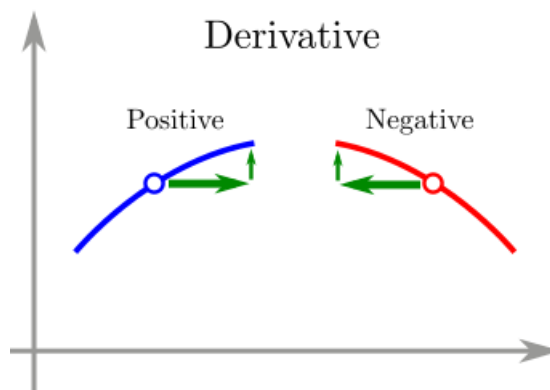
Similarly, a point cannot be a minimum if the derivative is not zero. Thus, the only possible inputs where a maximum or minimum can occur are those where the derivative is zero. This motivates the following definition

> **Critical point**
>
> A critical point of a function f is an input $x = a$ where either $f'(a) = 0$ or where the derivative is undefined.

Critical points include maximum and minimum points (called extrema) as well as inflection points; these are the points where the derivative is 0. Other critical points occur at corner points or discontinuities, where the derivative is undefined. The reason for including points where the derivative is not defined is that such a point could be a maximum or minimum:
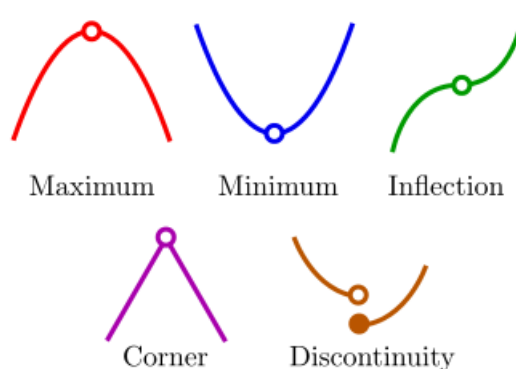
Example Compute the critical points of $f(x) = x^3 - 6x^2 + 9x - 5$.

The derivative $f'(x) = 3x^2 - 12x + 9$ is defined everywhere, so the critical points are where $f'(x) = 0$. Since

$$f'(x) = 3x^2 - 12x + 9$$
$$= 3(x^2 - 4x + 3)$$
$$= 3(x - 3)(x - 1),$$

the critical points are $x = 3$ and $x = 1$.

2.5.2　Classifying critical points

Once one has computed a critical points $x = a$, one can classify whether it is a maximum or minimum using the second derivative test:

Second Derivative Test

Suppose $x = a$ is a critical point of f where $f'(a) = 0$.

1. If $f''(a) > 0$, then f has a local minimum at a.

2. If $f''(a) < 0$, then f has a local maximum at a.

3. If $f''(a) = 0$, then the test fails.

In the third case, one can use the Taylor expansion about $x = a$ to determine the behavior of the function. In this case, $x = a$ could still be a local maximum, minimum, or inflection point.

The second derivative test is justified by considering the Taylor series for f about $x = a$:

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

since $f'(a) = 0$. Thus, when x is close to a, $f(x)$ behaves like a parabola centered at $x = a$. Recall that the sign of the coefficient of the square term in a parabola determines if the parabola opens up or down. A positive coefficient means the parabola opens upward, and a negative coefficient means the parabola opens downward.

Here, the coefficient of the $(x - a)^2$ is $\frac{1}{2}f(a)$. So if $f(a) > 0$, then the parabola opens upward, meaning $f(a)$ is a local minimum of f. If $f(a) < 0$, then the parabola opens downward, meaning $f(a)$ is a local maximum of f. If $f(a) = 0$, then one has to look at more terms of the Taylor series to determine f's behavior at a.

Example Use the Taylor series about $x = 0$ for

$$\sin^2 x \ln(\cos x)$$

to determine whether the function has a local maximum, local minimum, or inflection point at $x = 0$. (Take as a given that $x = 0$ is a critical point).

Expanding and multiplying the Taylor series gives

$$\sin^2 x \ln(\cos x) = \left(x + O(x^3)\right)^2 \ln\left(1 - \frac{1}{2!}x^2 + O(x^4)\right)$$
$$= (x + O(x^3))(x + O(x^3))\left(-\frac{1}{2}x^2 + O(x^4)\right)$$
$$= -\frac{1}{2}x^4 + O(x^6).$$

Thus, near $x = 0$ the function behaves like $-\frac{1}{2}x^4$, which is downward opening (because of the negative coefficient) and U shaped (because it is an even power). Therefore, the function has a local maximum at $x = 0$. Note that the second derivative test, besides being tricky to apply with all of the product rules and chain rules, would ultimately be inconclusive in this example.

Example Consider a square sheet of cardboard of side length L. By cutting equal sized squares of side length x from each corner of the sheet and folding up the flaps which are formed, one gets an open box:
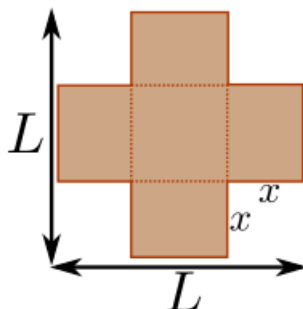
Note that as x gets bigger, the box gets taller but the area of the base of the box shrinks. As x gets smaller, the area of the base grows, but the height shrinks. Find the value of x which maximizes the volume of the resulting box.

The volume of the box is the area of the base times the height. The base is a square of side length $L - 2x$ (since x has been cut from both sides). The height of the box is x. Thus

$$V = (L - 2x)^2 \cdot x = 4x^3 - 4Lx^2 + L^2 x$$

Finding the critical points means taking the derivative with respect to x and setting equal to 0:

$$\frac{dV}{dx} = 12x^2 - 8Lx + L^2 = 0.$$

This factors as

$$(6x - L)(2x - L) = 0,$$

so the critical points are $x = \frac{L}{6}$ and $x = \frac{L}{2}$. To apply the second derivative test, we compute

$$\frac{d^2 V}{dx^2} = 24x - 8L$$

and evaluate at each critical point:

$$\left. \frac{d^2 V}{dx^2} \right|_{x=L/2} = 4L > 0 \quad \text{and} \quad \left. \frac{d^2 V}{dx^2} \right|_{x=L/6} = -4L < 0.$$

Thus, $x = \frac{L}{2}$ is a local minimum and $x = \frac{L}{6}$ is a local maximum. (Note also that for $x = \frac{L}{2}$ there is no cardboard left, since the removed corners have consumed the entire square!). The volume that results from $x = \frac{L}{6}$ is

$$V = \left(\frac{2L}{3} \right)^2 \frac{L}{6} = \frac{2L^3}{27}.$$

Example Classify the critical points of $f(x) = x^3 - 6x^2 + 9x - 5$.

As found in a previous example, the critical points of f are $x = 3$ and $x = 1$. The second derivative of f is $f(x) = 6x - 12$. Thus, $f(3) = 6 > 0$ and $f''(1) = -6 < 0$, and it follows from the second derivative test that 3 is a local minimum of f, and 1 is a local maximum of f.

Example Suppose a firm producing widgets expects to sell $3000 - 10p^2$ units (where p is the price of the unit). What price p should the firm set to maximize revenue (note that revenue here is just price times quantity sold)?

Revenue is $R(p) = (3000 - 10p^2)p = 3000p - 10p^3$. Taking the derivative gives $R'(p) = 3000 - 30p^2$, and setting equal to 0 gives

$$3000 - 30p^2 = 0$$
$$3000 = 30p^2$$
$$100 = p^2$$
$$p = \pm 10.$$

So R has critical point $p = 10$ (ignore $p < 0$ since price should be positive). The second derivative is $R(p) = -60p$. Thus $R(10) = -600 < 0$, and $p = 10$ is a local maximum. At this price, the revenue is
$$R(10) = 2000 \cdot 10 = 20000.$$

2.5.3   Global Extrema

While a local maximum or minimum is sometimes useful information, what is usually more important is the global maximum and minimum values of a function on a closed interval $[a, b]$ (or subject to some other constraint such as $x \geq 0$). These are called the global extrema, or absolute extrema, of a function.

Global extrema on the interval $[a, b]$ either occur at critical points of f or at the endpoints of the interval. So in addition to finding the critical points of f in the interval and checking their values, one must also evaluate f at the endpoints of the interval to find the global extrema.

Example Find the global extrema of $f(x) = x^3 - 6x^2 + 9x - 5$ on the interval $[2, 4]$.

From the prior examples, $x = 1$ and $x = 3$ are the critical points of f. But $x = 1$ is not in the interval $[2, 4]$, so disregard it. Then evaluate f at 2,3,4 to find the extreme values:

$$f(2) = 8 - 24 + 18 - 5 = -3$$
$$f(3) = 27 - 54 + 27 - 5 = -5$$
$$f(4) = 64 - 96 + 36 - 5 = -1.$$

Thus the absolute maximum is -1 and occurs when $x = 4$. The absolute minimum is -5 and occurs at $x = 3$.

2.5.4   Application: Statistics

In statistics, one often takes experimental data points of the form (x_i, y_i) and looks for a relationship. A very simple relationship is the linear relationship $y = mx$. The data may not follow this relationship perfectly, and there may be some slight experimental error or other noise, so one tries to find the value of m which best fits the data:
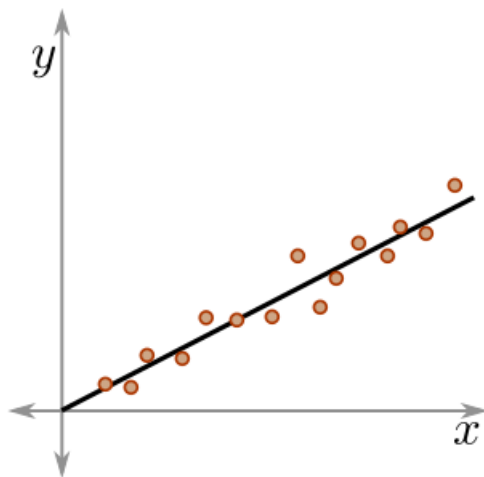
This process, called a linear regression, can be framed as an optimization problem. But what is the quantity being optimized?

There are several different linear regression models, depending on the quantity being minimized. These different quantities yield different best fit lines. One of the most common models is called ordinary least squares. This method seeks to minimize the sum of the squares of the residuals, which are the vertical distances from the points to the line:
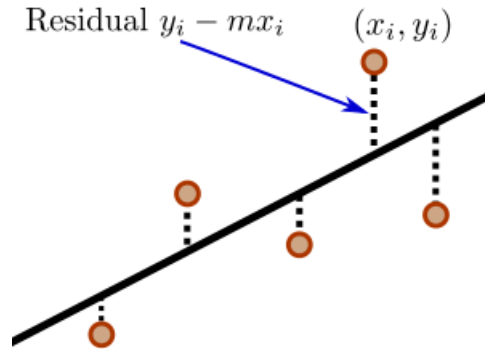
As shown above, the residual for a given point (x_i, y_i) is $y_i - mx_i$. Thus, the quantity being minimized is

$$S(m) = \sum_i (y_i - mx_i)^2.$$

Taking the derivative with respect to m gives

$$\frac{dS}{dm} = \sum_i 2(y_i - mx_i)(-x_i)$$

$$= \sum_i \left(-2x_i y_i + 2mx_i^2\right)$$

$$= -2\sum_i x_i y_i + 2m \sum_i x_i^2.$$

Setting this equal to 0 and solving for m gives

$$m = \frac{\displaystyle\sum_i x_i y_i}{\displaystyle\sum_i x_i^2}.$$

Applying the second derivative test, we compute

$$\frac{d^2 S}{dm^2} = 2\sum_i x_i^2 > 0,$$

so the above value of m minimizes the sum of squares (hence the least squares name).

Note To find the line of best fit of the form $y = mx + b$ requires methods of multivariable calculus (because there are two variables, m and b, which need to be optimized). Optimization with multiple variables is not much more difficult than for a single variable, but these methods are beyond the scope of this course.

2.5.5 Exercises

- Find all the local maxima and minima of the function $y = xe^{-x^2}$.

- Which type of critical point does the function $f(x) = e^{\sin(x^4)} \cos(x^2)$ have at zero ?

- Use a Taylor series about $x = 0$ to determine whether the function $f(x) = \sin^3(x^3)$ has a local maximum or local minimum at the origin.

- Find the location of the global maximum and minimum of $f(x) = x^3 - 6x^2 + 1$ on the interval $[-1, 7]$.

- Consider a stretch of highway in which cars are traveling at an average speed v. The "traffic density" u is the total amount of cars on our stretch of road divided by its length. These two quantities are related: the less cars on the road, the faster drivers are able to go. On the other hand, if traffic becomes heavy, drivers will naturally decrease their speed. The so-called "parabolic model" assumes that this relationship is dictated by the equation:

$$u = u_{\max} \left(1 - \frac{v}{v_{\max}} \right)$$

  where $u_{\max}$ represents the capacity of the road, and $v_{\max}$ the speed limit on it. The amount of cars passing through our road is called the "traffic flux" or "throughput", and is given by the product of the traffic density and the average speed: $F = uv$. Using the parabolic model, find out at what average speed v_* the flux through our road is maximized.

- A manufacturing company wants to know how many workers it should hire. If it employs too many people, the machines in the factory will be overutilized and the workers will have to wait until they are free, thus reducing the number of units each one will produce in a day's work. On the other hand, too few workers would leave the machines idle for long periods of time. A rough model for the relationship between the number n of workers and their productivity p is given by the equation

$$p = p_{\max} \left(1 - \frac{n}{n_{\max}} \right)$$

  where $p_{\max} = 10$ is the maximum number of units a worker can produce in a day and $n_{\max} = 100$ is the maximum number of workers the factory can accommodate. The amount of units U manufactured in the whole factory in one day is equal to the product of the number of workers and the number of units each one produces: $U = np$. How many workers should the company hire in order to maximize its production?

- A technology company has just invented a new gadget. In order to maximize the profit derived from its sale, the company must make a critical decision: at what price should it be sold? A market study suggests that the number N of units sold would approximately follow the equation $N = N_{\max} e^{-P/\lambda}$, where P is the sale price, $N_{\max} = 10,000,000$ is the number of units that would saturate the market, and $\lambda = \$50$. If it costs $\$250$ to manufacture one of these gadgets, at what price P_* would be profit of the company be maximized?

- The manufacturing process of a certain chemical substance is exothermic, that is, it releases heat. The amount of heat released, Q, depends on the temperature T at which the process is carried out, and it is given by the equation $Q = \alpha (T - T_0)^{-2} e^{(T-T_0)/\lambda}$, where $T_0 = 70^\circ F$ is the room temperature of the manufacturing plant, and $\alpha = 3000\, J\,(^\circ F)^2$ and $\lambda = 50^\circ F$. If the temperature T must be maintained above $100^\circ F$, at what temperature T_* would be the heat loss be minimized?

- Classify the critical point $x = 0$ of the function $f(x) = \frac{\sin^2(3x^2)\cos(x)}{x}$ using Taylor series.

- Construct a box without a top whose base is a square. The material cost for the bottom is $10 per square feet, the cost for the side is $5 per square feet. The box must have volume 8 cubic feet. Determine the dimension of the box that will minimize the cost.

2.6   Differentials

This module deals with differentials, e.g. dx or du. A formal treatment of differential forms is beyond the scope of this course. For now, the best way to think about the differential dx or du is to think of them as rates of change, and relate them with the chain rule:

$$du = \frac{du}{dx} dx.$$

In words, the rate of change of u equals the rate of change of u with respect to x times the rate of change of x. This is not a perfect interpretation, but it will serve our purposes for this course.

2.6.1 The differential as an operator

Think of the differential (d) as an operator which can be applied to an equation $f = g$ to give back $df = dg$. This process allows one to find the derivative of functions which are defined implicitly, i.e., functions which cannot be "solved for y" as $y = f(x)$. This method is called implicit differentiation.

Example Find ($\frac{dy}{dx}$) of the circle ($x^2 + y^2 = r^2$).

Taking the differential (remembering the chain rule) and doing some algebra gives

$$2x(dx) + 2y(dy) = 0$$
$$2y(dy) = -2x(dx)$$
$$\frac{dy}{dx} = \frac{-2x}{2y}$$
$$= -\frac{x}{y}.$$

So $\frac{dy}{dx} = -\frac{x}{y}$.

Example In economics, the marginal rate of substitution (MRS) of X for Y is the rate at which a consumer is willing to exchange good Y for good X to maintain an equal level of satisfaction (called utility in economics jargon).

Let $U(X, Y)$ denote a particular consumer's utility function. As a particular example, let X be coffee, in ounces, and Y be doughnuts. Then the curve $U(X, Y) = C$ represents all the different combinations of coffee and doughnuts where the consumer is equally happy. For example, if the utility function is

$$U(X, Y) = Y^2(X - 3),$$

then the following shows the graph of the curve $U(X, Y) = 4$. The two plotted points show that this consumer is equally satisfied with 4 ounces of coffee and 2 doughnuts as with 7 ounces of coffee and 1 doughnut.
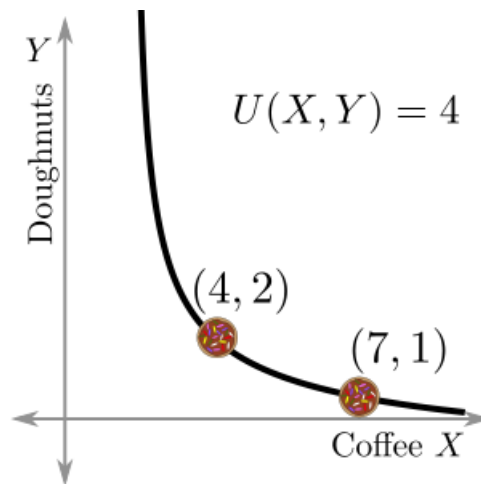
The MRS, then, is the rate of exchange of Y for X (doughnuts for coffee) so that utility stays the same (i.e. we stay on the curve). Mathematically, this is

$$\text{MRS} = -\frac{dY}{dX}.$$

Find MRS for the utility function given above, $U(X, Y) = Y^2(X - 3)$. Then calculate the MRS at the points $(4, 2)$ and $(7, 1)$ and interpret the results.

Using implicit differentiation, we find

$$Y^2(X - 3) = C$$
$$2Y\, dY\, (X - 3) + Y^2\, dX = 0.$$

Now solving for $-\frac{dY}{dX}$ we find

$$\text{MRS} = -\frac{dY}{dX} = \frac{Y^2}{2Y(X-3)} = \frac{Y}{2(X-3)}.$$

Evaluating at $(4, 2)$ gives a MRS of 1, which means that at that point the consumer is willing to substitute 1 doughnut for an ounce of coffee. At $(7, 1)$ the MRS is $\frac{1}{8}$, which means that the consumer is only willing to give up $\frac{1}{8}$ of a doughnut for an additional ounce of coffee.
In a sense the MRS is a measure of how a consumer reacts to the scarcity of one good relative to another and how that affects her willingness to exchange the goods.

Example Find the derivative of the function $y = f(x)$ defined implicitly by the equation

$$ye^x + x\ln(y) = e.$$

Taking the differential (remembering the product rule) gives

$$(dy)e^x + y(e^x dx) + (dx)\ln(y) + x\left(\frac{1}{y}dy\right) = 0.$$

Ultimately, the goal is to solve for $\frac{dy}{dx}$, so collecting all the terms with dy on the left and the terms with dx on the right gives

$$\left(e^x + \frac{x}{y}\right)dy = (-ye^x - \ln(y))\,dx.$$

Finally, dividing through gives
$$\frac{dy}{dx} = \frac{-ye^x - \ln(y)}{e^x + x/y}.$$

2.6.2   Related rates

The differential is often used in related rates problems. A related rates problem typically has a physical description and asks for the rate at which some quantity is changing. The description must be translated into an implicit relation between the variables involved, and then implicit differentiation is used to find the desired derivative.

Example Suppose a 10 foot ladder is leaning against a wall. The base of the ladder starts sliding away from the wall at a rate of 4 feet per second. At the moment when the base of the ladder is 6 feet away from the wall, at what rate is the top of the ladder sliding down the wall?
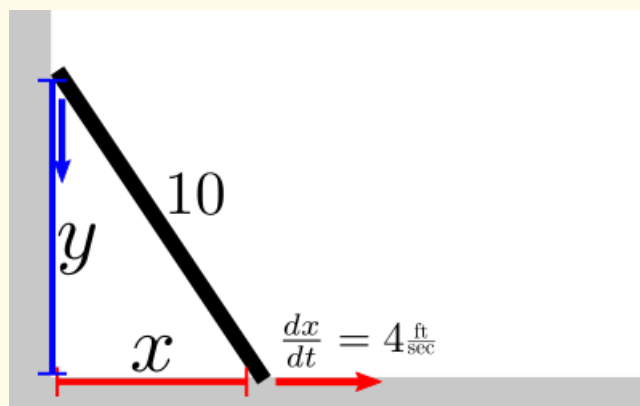
Let x be the distance from the base of the ladder to the wall and y be the distance from the top of the ladder to the floor. Then by the Pythagorean theorem, $x^2 + y^2 = 10^2$. At the moment in question, $x = 6$ so $y = 8$. The differential of this equation gives

$$2x\,dx + 2y\,dy = 0,$$

and solving for $\frac{dy}{dx}$ gives $\frac{dy}{dx} = -\frac{x}{y} = -\frac{6}{8} = -\frac{3}{4}$. But the question asked for $\frac{dy}{dt}$, not $\frac{dy}{dx}$. But by the Chain rule,

$$\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt}.$$

Since $\frac{dy}{dx} = -\frac{3}{4}$ and $\frac{dx}{dt} = 4$ (given in the problem), it follows that

$$\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt} = \left(-\frac{3}{4}\right) \cdot 4 = -3.$$

So the top of the ladder is sliding down the wall at a rate of 3 feet per second.

Example Consider the shape of a stream of water as it flows from a faucet. The stream has a circular cross-section which gets narrower lower in the stream, and the goal is to find how the radius of that cross-section is changing with respect to time.

Assume on the one hand that the water is flowing at a constant rate C. On the other hand, the area of the cross section times the velocity through that cross section equals the flow:

Flow rate is constant so
$$\text{area} \cdot \text{velocity} = C$$
$$\pi r^2 v = C$$
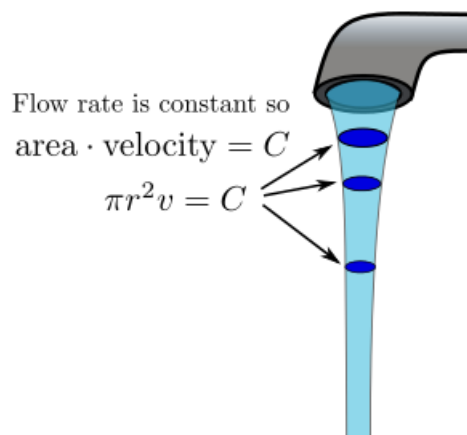
Dividing by π and taking the differential gives

$$r^2 v = C$$
$$2r \, dr \, v + r^2 \, dv = 0$$
$$dr = -\frac{r^2}{2rv} \, dv = -\frac{r}{2v} \, dv.$$

So

$$\frac{dr}{dv} = -\frac{r}{2v}.$$

Now, using the chain rule and a few facts from physics gives that

$$\begin{aligned}
\frac{dr}{dt} &= \frac{dr}{dv} \frac{dv}{dt} \\
&= -\frac{r}{2v} \frac{dv}{dt} \\
&= -\frac{rg}{2v} \\
&= -\frac{rg}{2(v_0 + gt)}.
\end{aligned}$$

This is known as a differential equation (in particular, a separable differential equation). One can solve this equation to get an explicit expression for r in terms of t; see the module on separable differential equations.

2.6.3   Relative rates of change

We can normalize the differential du by dividing by u. This gives $\frac{du}{u}$. This is known as the relative rate of change. Note that

$$\frac{du}{u} = d(\ln u).$$

where V is voltage across the resistor, I is the current, and R is the resistance of the resistor. If the voltage across a variable resistor is fixed, find the relationship between the relative rates change of resistance and current.

Beginning with Ohm's law and taking the differential gives

$$V = IR$$
$$dV = R\,dI + I\,dR.$$

Now, note that $dV = 0$ since voltage was assumed to be constant. Then dividing through by IR gives

$$R\,dI + I\,dR = 0$$
$$\frac{R\,dI}{IR} + \frac{I\,dR}{IR} = 0$$
$$\frac{dI}{I} + \frac{dR}{R} = 0$$
$$\frac{dI}{I} = -\frac{dR}{R}.$$

Thus, the relative rates of change of resistance and current are equal and opposite.

Example In a geometric solid (say, a sphere or a cube), how does the relative rate of change in volume compare to that of surface area?

For a sphere of radius r, the volume and area (and their differentials) are

$$V = \frac{4}{3}\pi r^3 \qquad\qquad\qquad dV = 4\pi r^2\,dr$$
$$A = 4\pi r^2 \qquad\qquad\qquad dA = 8\pi r\,dr.$$

Then the relative rates of change of volume and area are

$$\frac{dV}{V} = \frac{4\pi r^2\,dr}{\frac{4}{3}\pi r^3}$$
$$= 3\frac{dr}{r}.$$

and

$$\frac{dA}{A} = \frac{8\pi r\,dr}{4\pi r^2}$$
$$= 2\frac{dr}{r}.$$

Thus, the relative rate of change of volume is $\frac{3}{2}$ that of area. Written another way,

$$\frac{dV/V}{dA/A} = \frac{3}{2}.$$

For a cube of side length s, it turns out the same holds. The volume and area are

$$V = s^3 \qquad\qquad\qquad dV = 3s^2\,ds$$
$$A = 6s^2 \qquad\qquad\qquad dA = 12s\,ds.$$

The relative rates are $\frac{dV}{V} = 3\frac{ds}{s}$ and $\frac{dA}{A} = 2\frac{ds}{s}$, so once again

$$\frac{dV/V}{dA/A} = \frac{3}{2}.$$

Example In economics, the demand curve for a good is the quantity Q of the good that a consumer would purchase as a function of the price P of the good. The demand curve slopes downward since a consumer will typically buy less of a good if it is more expensive (the exception being a Giffen good).

The price elasticity of demand, E, for a good is the rate of change of relative quantity fluctuation with respect to relative price fluctuation. Informally, it can be thought of as the percent change in quantity resulting from a percent change in price. One can also think of E as a measure of how price sensitive a consumer is for that good at that price. Mathematically,

$$E = -\frac{dQ/Q}{dP/P}.$$

The negative sign is there to force elasticity to be positive (without it, $\frac{dQ/Q}{dP/P}$ would always be negative due to the downward slope of the demand curve).

A good is said to be elastic at a certain price if $E > 1$ (that is, a consumer is highly sensitive to price changes). A good is inelastic at a certain price if $E < 1$. An example of an elastic good is wine, since a small increase in the relative price of wine can result in a consumer substituting a different alcohol for it. An example of an inelastic good is toilet paper, since regardless of price changes, a consumer is likely to require about the same amount of toilet paper:
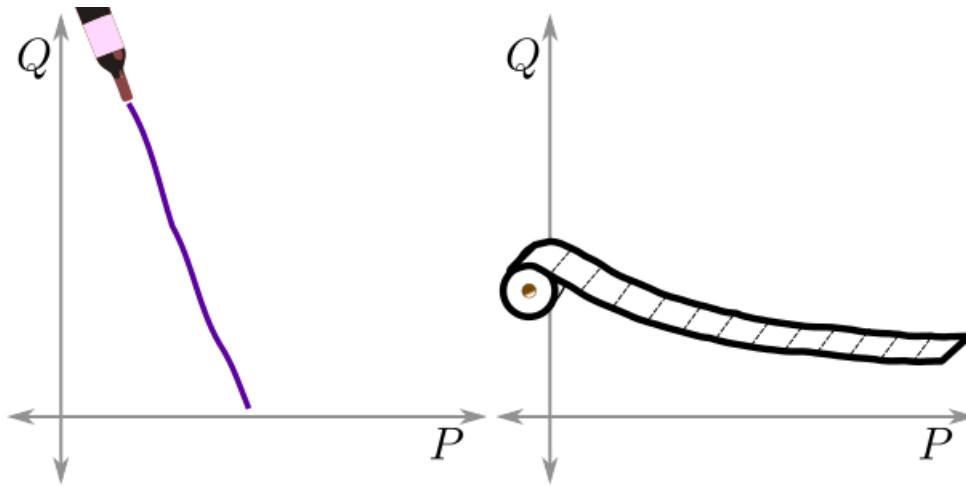
Revenue, R, is given by $R = P \cdot Q$. How does one maximize the relative revenue with respect to relative change in price?

2.6.4   Exercises

- Use implicit differentiation to find $\dfrac{dy}{dx}$ from the equation $y^2 - y = \sin 2x$.

- Find the derivative $\dfrac{dy}{dx}$ if x and y are related through $xy = e^y$.

- Use implicit differentiation to find $\dfrac{dy}{dx}$ if $\sin x = e^{-y\cos x}$.

- Find the derivative $\dfrac{dy}{dx}$ from the equation $x\tan y - y^2 \ln x = 4$.

- Model a hailstone as a round ball of radius R. As the hailstone falls from the sky, its radius increases at a constant rate C. At what rate does the volume V of the hailstone change?

- The volume of a cubic box of side-length L is $V = L^3$. How are the relative rates of change of L and V related?

- Consider a box of height h with a square base of side length L. Assume that L is increasing at a rate of 10% per day, but h is decreasing at a rate of 10% per day. Use a linear approximation to find at what (approximate) rate the volume of the box changing. "Hint:" consider the relative rate of change of the volume of the box.

- A large tank of oil is slowly leaking oil into a containment tank surrounding it. The oil tank is a vertical cylinder with a diameter of 10 meters. The containment tank has a square base with side length of 15 meters and tall vertical walls. The bottom of the oil tank and the bottom of the containment tank are concentric (the round base inside the square base). Denote by h_o the height of the oil inside of the oil tank, and by h_c the height of the oil in the containment tank. How are the rates of change of these two quantities related?

- The "stopping distance" D_{stop} is the distance traveled by a vehicle from the moment the driver becomes aware of an obstacle in the road until the car stops completely. This occurs in two phases. (1) The first one, the "reaction phase", spans from the moment the driver sees the obstacle until he or she has completely depressed the brake pedal. This entails taking the decision to stop the vehicle, lifting the foot from the gas pedal and onto the brake pedal, and pressing the latter down its full distance to obtain maximum braking power. The amount of time necessary to do all this is called the "reaction time" t_{react}, and is independent of the speed at which the vehicle was traveling. Although this quantity varies from driver to driver, it is typically between $1.5\,\mathrm{s}$ and $2.5\,\mathrm{s}$. For the purposes of this problem, we will use an average value of $2\,\mathrm{s}$. The distance traversed by the vehicle in this time is unsurprisingly called "reaction distance" D_{react} and is given by the formula $D_{react} = v t_{react}$, where v is the initial speed of the vehicle.
  (2) In the "braking phase", the vehicle decelerates and comes to a complete stop. The "braking distance" D_{brake} that the vehicle covers in this phase is proportional to the square of the initial speed of the vehicle: $D_{brake} = \alpha v^2$. The constant of proportionality α depends on the vehicle type and condition, as well as on the road conditions. Consider a typical value of $10^{-2}\,\mathrm{s^2/m}$.
  If the initial speed of the vehicle is $108\,\mathrm{km/h} = 30\,\mathrm{m/s}$, what is the ratio between the relative rate of change of the stopping distance and the relative rate of change of the initial speed?

- Assume that you possess equal amounts of a product X and Y, but you value them differently. Specifically, your "utility function" is of the form $U(X, Y) = CX^\alpha Y^\beta$ for α, β, and C positive constants. What is your marginal rate of substitution (MRS) of Y for X?

2.7 Differentiation as an operator

The sum, product, quotient, and chain rules make it possible to differentiate many functions. However, there are some more exotic functions which cannot be differentiated using these tools alone. For example, what is the derivative of 2^x or x^x or x^{x^x}?

The derivative should be interpreted as a rate of change, but what about the act of differentiation? Differentiation is an operator: it takes in a function and gives out another function. Other examples of operators include the logarithm, exponentiation, and integration. These (and other) operators can be applied to an entire equation to transform a hard problem to an easy problem and (once the solution is found) back again. This idea will allow us to compute the derivatives of the exotic functions above, and more.

2.7.1 Logarithmic differentiation

A common combination is called logarithmic differentiation, which consists of applying the logarithm operator followed by the differentiation operator. It is best demonstrated by example.

Example Find the derivative of e^x using logarthmic differentiation.

We already know the derivative of e^x, but suppose we did not. Let $y = e^x$. Then applying the logarithm operator to this equation gives

$$\ln y = \ln(e^x) = x.$$

Now applying the differentiation operator to the equation (and remembering the chain rule) gives

$$\frac{dy}{y} = dx,$$

so $\frac{dy}{dx} = y$, and since $y = e^x$ from our original definition, we have

$$\frac{dy}{dx} = e^x,$$

as expected.

Example Use logarithmic differentiation to show that $\frac{d}{dx}(a^x) = a^x \ln a$.

Similarly to the above example, let $y = a^x$. Then taking the logarithm gives

$$\ln y = \ln(a^x) = x \ln a.$$

Differentiating gives

$$\frac{dy}{y} = (\ln a)\, dx$$

And so

$$\frac{dy}{dx} = y(\ln a) = a^x \ln a,$$

as desired.

Example Find the derivative of x^x.

Let $y = x^x$. Taking the logarithm gives

$$\ln y = x \ln x.$$

Differentiating (using the product rule on the right) gives that

$$\frac{dy}{y} = x\frac{1}{x}\, dx + \ln x\, dx.$$

Next, factoring and solving for $\frac{dy}{dx}$ gives

$$\frac{dy}{dx} = y(1 + \ln x) = x^x(1 + \ln x)$$

2.7.2 Other operators

There are other operators which can be used prior to differentiation. Consider the following examples.

Example Compute the derivative of $(\ln x)$ by using exponentiation followed by differentiation.

Letting $y = \ln x$, we exponentiate the equation to find

$$e^y = x.$$

Now, differentiating gives

$$e^y\, dy = dx,$$

and solving for $\frac{dy}{dx}$ gives

$$\frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{e^{\ln x}} = \frac{1}{x},$$

as desired.

Note that $\ln x$ is the inverse of e^x, and so when we exponentiated the equation, it could be thought of as applying the inverse of $\ln x$. This same method works for many other inverse functions. In particular, applying a trigonometric function can be thought of as an operator as well. This method can be used to find the derivatives of the various inverse trigonometric functions.

Example Find the derivative of $\arcsin(x)$.

Letting $y = \arcsin x$, take the sine of the equation to find
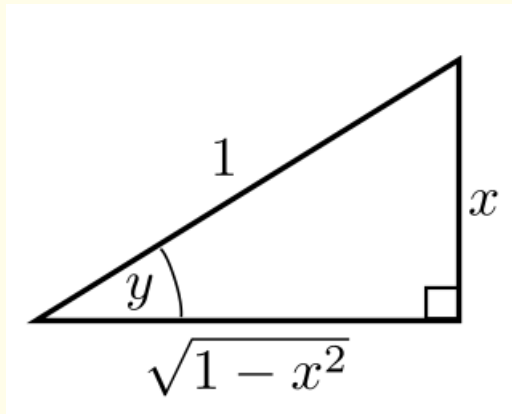
$$\sin y = x.$$

Now, differentiating gives

$$\cos y\, dy = dx.$$

Thus, we find that
$$\frac{dy}{dx} = \frac{1}{\cos y}.$$

Now, a little bit of trigonometry helps rewrite $\sec y$ in terms of x. Our original equation had $y = \arcsin x$. That means that y is the angle such that $\sin y = x$. Since sine is the opposite over the hypotenuse, we can express this relationship with the following right triangle:

where the adjacent leg comes from the Pythagorean theorem. It follows that $\cos y = \sqrt{1 - x^2}$, and so we find
$$\frac{dy}{dx} = \frac{1}{\cos y} = \frac{1}{\sqrt{1 - x^2}}.$$

Example Find the derivative of $\arctan(x)$.

Let $y = \arctan(x)$. Applying $\tan$ to the equation gives
$$\tan y = x,$$

and differentiating gives
$$\sec^2 y \, dy = dx.$$

Therefore,
$$\frac{dy}{dx} = \frac{1}{\sec^2 y}.$$

We can now do similar right triangle trig as in the previous example. Or we can recall that by the Pythagorean identity for tangent and secant, we have
$$\tan^2 y + 1 = \sec^2 y$$

Making this substitution gives
$$\frac{dy}{dx} = \frac{1}{\sec^2 y}$$
$$= \frac{1}{\tan^2 y + 1}$$
$$= \frac{1}{x^2 + 1}.$$

2.7.3   Operators in other contexts

Besides being useful in computing derivatives of exotic functions, operators (especially the logarithm) can also be useful in computing limits. The method is similar to the above method for derivatives.

Example Show that
$$\lim_{x \to \infty} \left(1 + \frac{a}{x}\right)^x = e^a.$$

This is a common limit which will come up again in the course.

Let y be the function given by
$$y = \left(1 + \frac{a}{x}\right)^x.$$

Take the logarithm of both sides, and use the power property of logarithms to see
$$\ln y = \ln \left(1 + \frac{a}{x}\right)^x$$
$$= x \ln \left(1 + \frac{a}{x}\right).$$

Now, taking the limit of both sides gives
$$\lim_{x \to \infty} \ln y = \lim_{x \to \infty} x \ln \left(1 + \frac{a}{x}\right).$$

Now, since $x \to \infty$, we have that $\frac{a}{x}$ is small, and so we can use our Taylor series for $\ln(1 + x)$, which gives us that
$$\lim_{x \to \infty} \ln y = \lim_{x \to \infty} x \ln \left(1 + \frac{a}{x}\right)$$
$$= \lim_{x \to \infty} x \left(\frac{a}{x} + O\left(\frac{1}{x^2}\right)\right)$$
$$= \lim_{x \to \infty} a + O\left(\frac{1}{x}\right)$$
$$= a.$$

Recall from our limit rules that the order of the logarithm and the limit can be switched since the logarithm is a continuous function. Thus
$$\ln \left(\lim_{x \to \infty} y\right) = a.$$

So, exponentiating both sides, we have that
$$\lim_{x \to \infty} y = e^a,$$

as desired.

2.7.4 Infinite Power Tower

Consider the infinite power tower
$$y = x^{x^{x^{x^{\cdots}}}}.$$

That is, x raised to the x raised to the x etc. This is certainly an unusual function. A better way to define this function is implicitly:
$$y = x^y.$$

To see that this makes intuitive sense, note that the exponent of the first x in the infinite tower is itself an infinite tower, so replacing the exponent of x with y is sensible.

Use logarithmic differentiation to find the derivative of this function (that is, $\frac{dy}{dx}$).

First, taking the logarithm gives
$$\ln y = y \ln x.$$

Now, differentiating the equation (implicitly) gives

$$\frac{dy}{y} = \ln x \, dy + y \frac{dx}{x}$$

Factoring and solving for $\frac{dy}{dx}$ gives

$$\left(\frac{1}{y} - \ln x \right) dy = y \frac{dx}{x}$$

$$\frac{dy}{dx} = \frac{y}{x \left(1/y - \ln x \right)}$$

$$= \frac{y^2}{x(1 - y \ln x)}.$$

This shows that although a function may be difficult to understand, it can nevertheless be fairly easy to find its derivative.

It turns out that this function is well-defined and differentiable on

$$e^{-e} < x < e^{1/e}$$

One can check that the (x, y) pairs $(e^{-e}, e^{-1}), (1, 1), (\sqrt{2}, 2), (e^{1/e}, e)$ all satisfy the above implicit equation.

2.7.5   Exercises

- Find the derivative of $f(x) = (\ln x)^x$

- Find the derivative of $f(x) = x^{\ln x}$

- Compute $\lim\limits_{x \to +\infty} \left(\frac{x + 2}{x + 3} \right)^{2x}$

- Compute $\lim\limits_{x \to 0+} [\ln(1 + x)]^x$

- Compute $\lim\limits_{x \to 0} \left(1 + \arctan \frac{x}{2} \right)^{2/x}$

- Compute $\lim\limits_{x \to 0+} \left(\frac{2}{x} \right)^{\sin x}$

- Let

$$\alpha = 1 + \frac{2}{2 + \frac{2}{2 + \frac{2}{2 + \cdots}}}$$

What is the value of α?

- Let

$$\varphi = \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \cdots}}}}}$$

What is the value of φ?

Chapter 3

Integration

3.1 Antidifferentiation

This module begins our study of integration. Integration, or anti-differentiation, can be thought of as running differentiation in reverse, or undoing the derivative.

This motivates the following definition:

> **The Indefinite Integral**
>
> The indefinite integral of $f(t)$, denoted $\int f(t)\,dt$, is the class of functions whose derivative is $f(t)$. $\int f(t)\,dt$ is also referred to as the anti-derivative of f. The act of taking the indefinite integral is an operator which is referred to as anti-differentiation or integration.

Note The indefinite integral of a function is only defined up to an added constant, called the constant of integration. In other words, if $F(x)$ is an anti-derivative of $f(x)$, then $G(x) = F(x) + C$, C a constant, is also an anti-derivative of f, because C disappears when differentiated. Conversely, any two indefinite integrals of $f(x)$ differ only by some constant.

Any of the known derivatives from the previous chapter can be rephrased as an integral. For example, just as there was a power rule for differentiating monomials, there is a corresponding power rule for integrating monomials. And any anti-derivative can easily be checked by taking the derivative and seeing that the result gives back the original function.

Example Give the integral of each of the following functions: x^n, $\frac{1}{x}$, $\sin x$, $\cos x$, e^x.

$$\int x^n dx = \frac{1}{n+1}x^{n+1} + C$$

$$\int \frac{1}{x}dx = \ln|x| + C$$

$$\int \sin x\, dx = -\cos x + C$$

$$\int \cos x\, dx = \sin x + C$$

$$\int e^x dx = e^x + C$$

(Don't forget the constant!)

There are other functions which are harder to integrate by merely using one of the derivatives we already know. Some of these can be integrated using other techniques from upcoming modules, but there are also functions whose anti-derivative cannot be expressed in terms of simple functions.

3.1.1 Differential equations

The motivating problem for the study of anti-differentiation is solving a differential equation. A differential equation is an equation involving a function and its derivative. In this course, we deal with ordinary differential equations, ODEs, which are differential equations involving only functions of one variable and

the derivative with respect to that variable (future courses deal with partial differential equations, which involve functions of several variables and partial derivatives).

Solving a differential equation means finding the function (or class of functions, usually) which satisfy the differential equation.

A Simple ODE

The simplest differential equation is of the form

$$\frac{dx}{dt} = f(t).$$

Here, the goal is to find the function $x(t)$ whose derivative with respect to t is $f(t)$. But this is precisely what the integral is. And so, the solution of the differential equation $\frac{dx}{dt} = f(t)$ is given by $x(t) = \int f(t)dt$.

Using the interpretation of the derivative as slope, one can think of the function $f(t)$ as describing the slope of the function $x(t)$:
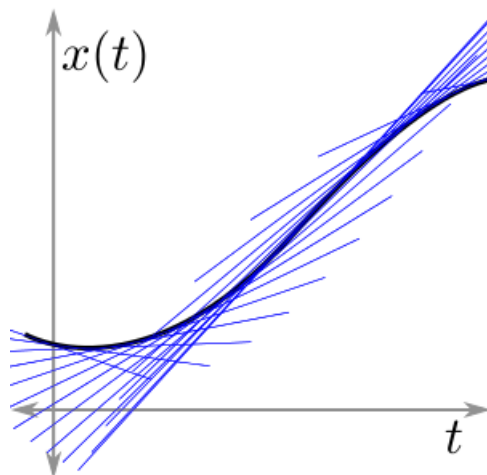
Thus, $x(t)$ is a function which fits the slopes prescribed by $f(t)$. Note that any constant vertical shift of a solution $x(t)$ will still have the same slope at each point. This is one interpretation of the integration constant: it represents the potential vertical shifts to a solution of the differential equation.

Example Consider a falling object. Let $x(t)$ be the height of the object at time t, $v(t)$ be the velocity of the object, and assume that acceleration is the constant $-g$ (negative because gravity pulls down). Express the height of the object as a function of t, v_0, and x_0; here, v_0 and x_0 are the velocity and height of the object, respectively, at time $t = 0$.

We know from an earlier module that

$$\frac{dv}{dt} = a = -g.$$

Beginning with the second of these equations, we find that

$$v(t) = \int (-g)\, dt = -gt + C.$$

We can determine C by plugging in $t = 0$. This cancels that $-gt$ and leaves us with $C = v(0) = v_0$, the initial velocity. Thus,

$$v(t) = -gt + v_0$$

Now, using the fact that

$$\frac{dx}{dt} = v,$$

we find that

$$x(t) = \int v(t)\, dt$$
$$= \int (-gt + v_0)\, dt$$
$$= -\frac{1}{2}gt^2 + v_0 t + C.$$

Again, we can find C by plugging in $t = 0$. This leaves us with $x(0) = C$, and so $C = x_0$, the initial height. Thus,

$$x(t) = -\frac{1}{2}gt^2 + v_0 t + x_0.$$

The Next Simplest ODE

Another slightly more complex ODE is of the form

$$\frac{dx}{dt} = f(x).$$

Before we discuss how to solve this in general, we consider a specific example, which is one of the most famous differential equations:

$$\frac{dx}{dt} = ax,$$

where a is a constant. We solve this differential equation in three different ways:

1. (Guess) Solve this differential equation by first observing that $x = Ce^t$ satisfies $\frac{dx}{dt} = x$ and then adjusting the exponent so that an extra factor of a comes out when differentiating. Hint: remember the chain rule.

Observe that $x = Ce^{at}$ will get an extra factor of a when differentiated by the chain rule. That is,

$$\frac{d}{dt}\left(Ce^{at}\right) = aCe^{at}.$$

And so $x(t) = Ce^{at}$ is a solution of the differential equation.

2. (Series) Solve the differential equation by assuming

$$x(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3 + \cdots$$

and then determining what the constants c_i must be to satisfy the differential equation.

Assuming that

$$x(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3 + \cdots,$$

and then taking the derivative of this series, term by term, we find

$$\frac{dx}{dt} = 0 + c_1 + 2c_2 t + 3c_3 t^2 + \cdots.$$

On the other hand, from the original differential equation we have

$$\frac{dx}{dt} = ax$$
$$= a\left(c_0 + c_1 t + c_2 t^2 + c_3 t^3 + \cdots\right)$$
$$= ac_0 + ac_1 t + ac_2 t^2 + ac_3 t^3 + \cdots.$$

Because these two series both equal $\frac{dx}{dt}$, they must be equal to each other. But two series are equal if and only if their corresponding coefficients are equal. Therefore,

$$c_1 = ac_0$$
$$2c_2 = ac_1$$
$$3c_3 = ac_2,$$

and so on. Solving these equations one by one gives

$$c_1 = ac_0$$
$$c_2 = \frac{1}{2}ac_1 = \frac{1}{2}a^2c_0$$
$$c_3 = \frac{1}{3}ac_2 = \frac{1}{6}a^3c_0$$

And, generally, $c_n = \frac{1}{n!}a^n c_0$ (this can be proven using a method called induction). Doing a little bit of factoring and grouping of factors, we find

$$x(t) = c_0 + ac_0 t + \frac{1}{2!}a^2 c_0 t^2 + \frac{1}{3!}a^3 c_0 t^3 + \cdots$$
$$= c_0\left(1 + (at) + \frac{1}{2!}(at)^2 + \frac{1}{3!}(at)^3 + \cdots\right)$$
$$= c_0 e^{at},$$

which is, again, of the form Ce^{at}.

3. (Integration) Rearrange the differential equation into the form

$$\frac{dx}{x} = a\,dt$$

and integrate both sides to solve the differential equation.

Using the chain rule and substituting according to the differential equation, we have

$$dx = \frac{dx}{dt}\,dt$$
$$dx = ax\,dt$$
$$\frac{dx}{x} = a\,dt.$$

Integrating both sides of the equation gives

$$\ln x = at + C$$

(only one constant of integration is necessary here, because a constant on the left side could be subtracted from both sides and absorbed into C). Now, exponentiating the equation gives $x = e^{at+C}$. By exponential rules, $e^{at+C} = e^{at}e^C$, and the e^C is often rewritten as a new constant, often written C again. Thus, the solution to $\frac{dx}{dt} = ax$ is $x(t) = Ce^{at}$, where C is any constant.

The differential equation from this example is sometimes used as a simple model of population growth. In words, the differential equation says that the growth of a population is proportional to the size of the population. As the solution above slows, this model implies the population has exponential growth. This is not a very good model for most populations because of competition for resources and overcrowding. But under certain conditions and for short periods of time, some populations (for instance, bacteria with an abundant food supply) do exhibit exponential growth. For more examples of exponential growth, see the next module.

Initial value problems

Although a general differential equation's solution often depends on a constant (sometimes several), an additional condition called an initial value or initial condition can specify a specific solution. This condition is usually of the form $y(t_0) = y_0$. A differential equation with such an initial condition is called an initial value problem. To solve such a problem, first find the general solution and then use the initial value to find the specific constant of integration which satisfies the initial condition.

In the context of population growth, the initial value is typically the size of the population at time 0. This is particularly nice in the exponential growth model, because the solution is of the form $P(t) = De^{At}$. So if $P(0) = P_0$ is given, then plugging this in gives $P(t) = P_0 e^{At}$.

3.1.2 Exercises

- $\int (4x^3 + 3x^2 + 2x + 1)\, dx =$

- $\dfrac{d}{dx} \int \ln \tan x \, dx =$

- $\int \left(\dfrac{d}{dx} e^{-x} \right) dx =$

- Find the general solution of the differential equation $\dfrac{dx}{dt} = t^2$

- Find the general solution of the differential equation $\dfrac{dx}{dt} = x^2$

- There is a large class of differential equations – the so-called "linear" ones – for which we can find solutions using the Taylor series method discussed in the Lecture. One such differential equation is
$$t\frac{d^2 x}{dt^2} + \frac{dx}{dt} + tx = 0$$
This is a particular case of the more general "Bessel differential equation", and one solution of it is given by the Bessel function $J_0(t)$ that we saw earlier. Notice that this involves not only the first derivative but also the second derivative. For this reason, it is said to be a "second order" differential equation.
In this problem we will content ourselves with finding a relationship (specifically, a "recurrence relation") on the coefficients of a Taylor series expansion about $t = 0$ of a solution to our equation.
Hence consider the Taylor series $x(t) = \displaystyle\sum_{k=0}^{\infty} c_k t^k$
Substituting this into the differential equation above will give you two conditions. The first one is $c_1 = 0$. What is the other one?
"Note:" this problem involves some nontrivial manipulation of indices in summation notation. Do not get discouraged if it feels more difficult than other problems: it is!

- Find the general solution of the differential equation $\dfrac{dx}{dt} = t^3 + x^2 t^3$

3.2 Exponential growth examples

Recall from the last module that the differential equation $\frac{dx}{dt} = ax$ has solution $x = Ce^{at}$, where C is some constant. The constant (C) can be thought of as an initial condition, the value of the function at time $t = 0$. When $a > 0$, the function has *exponential growth*. When $a < 0$, the function has *exponential decay*:
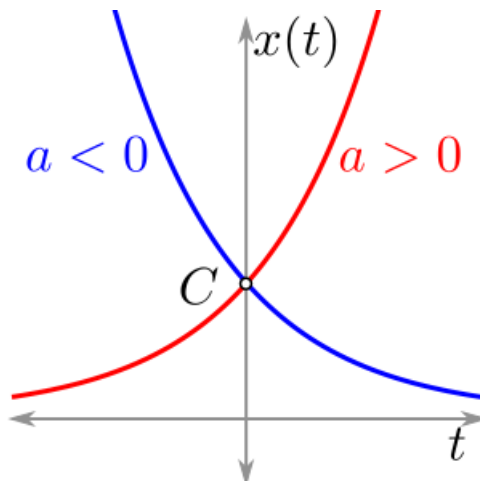
This module is devoted to several examples of exponential growth and decay.

3.3   More differential equations

3.3.1   Separable differential equations

3.4   ODE Linearization

3.5   Integration by Substitution

3.6   Integration by parts

3.7   Trigonometric substitution

3.8   Partial fractions

3.9   Definite integrals

3.10   Fundamental Theorem of Integral Calculus

3.11   Improper integrals

3.12   Trigonometric integrals

3.13   Tables and computers

Chapter 4

Applications

Chapter 5

Discretization

5.1 Series