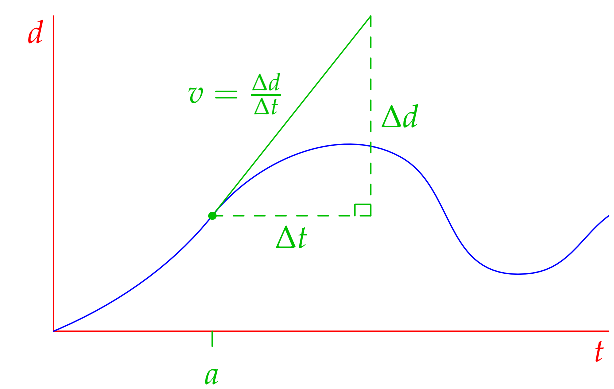
3 Differentiation

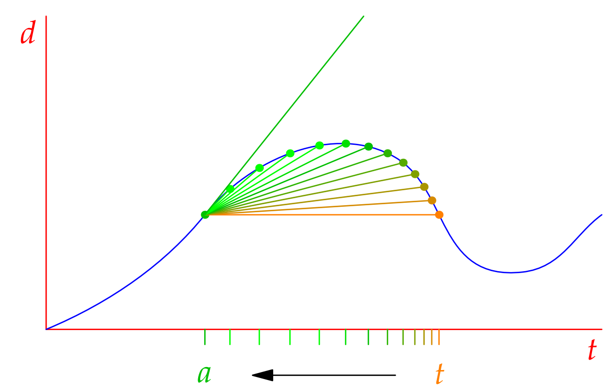
Differentiation grew out of the problem of *instantaneous velocity*. Velocity can only easily be measured as an *average* over a time interval:[13] if an object travels Δd meters in Δt seconds, then its average velocity is $v_{\text{av}} = \frac{\Delta d}{\Delta t}\,\text{ms}^{-1}$. An early 'definition' (dating to the 1300s) makes the instantaneous velocity equal to the constant velocity that would be observed if a body were to stop accelerating: while useless for the purposes of measurement, this is essentially Newton's first law regarding inertial motion (1687). We also see the concept of the *tangent line* beginning to appear. Indeed if one graphs position against time, intuition tells us:

- The graph of inertial (constant speed) motion is a straight line whose slope is the velocity.
- The tangent line to a curve has slope equal to the instantaneous velocity.

The problem of finding, defining and computing instantaneous velocity thus morphed into the consideration of tangent lines to curves. With the advent of analytic geometry in the early 1600s, mathematicians such as Fermat and Descartes pioneered versions of the familiar *secant ('cutting') line* method for computing tangents.



Instantaneous velocity equals constant velocity corresponding to tangent line



Secant lines approximate tangent line as $t \to a$

The average velocity of the particle over the time interval $[a, t]$ is the slope of the secant line, namely

$$v_{\text{av}}(a,t) = \frac{d(t) - d(a)}{t - a}$$

Since the secant lines approximate the tangent line as t approaches a, it seems reasonable that we should compute the instantaneous velocity in this manner:

$$v(a) = \lim_{t \to a} v_{\text{av}}(a,t) = \lim_{t \to a} \frac{d(t) - d(a)}{t - a}$$

This is, of course, the modern definition of the *derivative*.

[13] Even a modern technique such as Doppler-shift compares measurements separated by the extremely small period of a light or soundwave. These are still therefore *average* velocities, albeit taken over very small time intervals.

3.28 Basic Properties of the Derivative

Definition 3.1. Let $f : U \to \mathbb{R}$ and $a \in U^\circ$ an **interior point**. We say that f *is differentiable at* a if the following limit exists (is *finite*!)

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a}$$

We call this limit the *derivative of* f *at* a and denote its value by $\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x=a}$ or $f'(a)$.

If $f'(a)$ exists for all $a \in U$ then f is *differentiable (on* U*)*; the derivative becomes a function $f'(x) = \frac{\mathrm{d}f}{\mathrm{d}x}$.

The two notations are partly attributable to the primary founders of calculus: Issac Newton and Gottfried Leibniz. Each has its pros and cons and you should be comfortable with both.

One-sided derivatives Differentiability only makes sense at interior points of U since the defining limit is two-sided. *Left-* and *right-derivatives* may be defined using one-sided limits; differentiability is then equivalent to these being equal. All results in this section hold for one-sided derivatives with suitable (sometimes tedious) modifications. It is common, though strictly incorrect, to say that f is differentiable on $[a, b)$ if it is differentiable on the interior (a, b) and *right*-differentiable at a. In these notes we will strictly adhere to Definition 3.1: differentiable means *two-sided.*

Examples 3.2. Basic examples should be familiar from elementary calculus.

1. Let $f(x) = x^2 + 4x$. Then, for any $a \in \mathbb{R}$,

$$\begin{aligned}\lim_{x\to a} \frac{f(x) - f(a)}{x - a} &= \lim_{x\to a} \frac{x^2 + 4x - a^2 - 4a}{x - a} = \lim_{x\to a} \frac{(x-a)(x+a+4)}{x-a} \\ &= \lim_{x\to a}(x + a + 4) = 2a + 4\end{aligned}$$

   Note how the definition of $\lim\limits_{x\to a}$ allows us to cancel the $x - a$ terms from the numerator and denominator. We conclude that f is differentiable (on $\mathbb{R}$) and that $f'(x) = 2x + 4$.

2. Let $g(x) = \frac{x+1}{2x-3}$. Then, for any $a \neq \frac{3}{2}$,

$$\begin{aligned}\lim_{x\to a} \frac{f(x) - f(a)}{x - a} &= \lim_{x\to a} \frac{1}{x-a}\left[\frac{x+1}{2x-3} - \frac{a+1}{2a-3}\right] = \lim_{x\to a} \frac{5a - 5x}{(x-a)(2x-3)(2a-3)} \\ &= \lim_{x\to a} \frac{-5}{(2x-3)(2a-3)} = \frac{-5}{(2a-3)^2}\end{aligned}$$

   f is therefore differentiable on its domain $\mathbb{R} \setminus \{\frac{3}{2}\}$ with derivative $f'(x) = \frac{-5}{(2x-3)^2}$.

The familiar expressions

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

are equivalent to the original definition (Exercise 5). While seemingly simpler, they sometimes lead to nastier calculations: see what happens if you try the previous example in this language…

We now turn to perhaps the most well-known result of elementary calculus.

Theorem 3.3 (Power Law). *Let $r \in \mathbb{R}$. Then $f(x) = x^r$ is differentiable with $f'(x) = rx^{r-1}$.*

The domains of f and f' depend messily on r, but the formula holds at least on the interval $(0, \infty)$. We leave a complete proof to the exercises and instead consider a few generalizable examples.

Examples 3.4. 1. If $n \in \mathbb{N}$ and $a \in \mathbb{R}$, a simple factorization yields

$$\begin{aligned}\lim_{x\to a} \frac{x^n - a^n}{x-a} &= \lim_{x\to a} \frac{(x-a)(x^{n-1} + ax^{n-2} + \cdots + a^{n-2}x + a^{n-1})}{x-a} \\ &= \lim_{x\to a}(x^{n-1} + ax^{n-2} + \cdots + a^{n-2}x + a^{n-1}) = na^{n-1}\end{aligned} \tag{$*$}$$

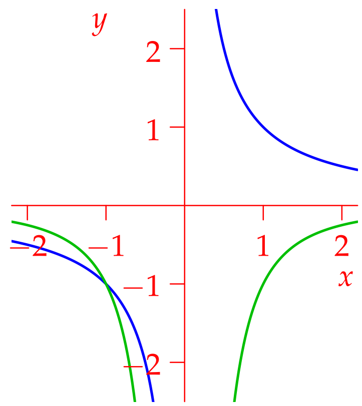
We conclude that $\frac{d}{dx}x^n = nx^{n-1}$.

2. If $f(x) = x^{-1}$ and $a \neq 0$, then

$$\lim_{x\to a} \frac{x^{-1} - a^{-1}}{x-a} = \lim_{x\to a} \frac{a-x}{ax(x-a)} = \lim_{x\to a} \frac{-1}{ax} = -\frac{1}{a^2}$$

from which we conclude that $f'(x) = -x^{-2}$.
A similar approach followed by the factorization $(*)$ proves the power law for all negative integer exponents:

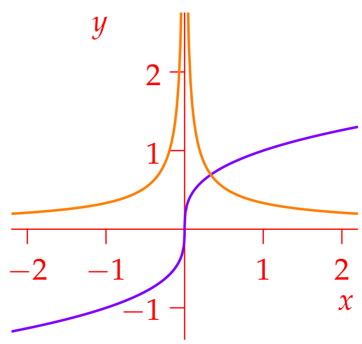
$$\frac{x^{-n} - a^{-n}}{x-a} = \frac{a^n - x^n}{a^n x^n (x-a)} = \cdots$$



3. To differentiate $x^{1/n}$, substitute $x = y^n$ and observe case 1. For instance, if $g(x) = x^{1/3}$ and $a \neq 0$, then $y = x^{1/3}$ and $b = a^{1/3}$ yield

$$\lim_{x\to a} \frac{x^{1/3} - a^{1/3}}{x-a} = \lim_{y\to b} \frac{y-b}{y^3 - b^3} = \frac{1}{3b^2} = \frac{1}{3}a^{-2/3}$$
$$\implies g'(x) = \frac{1}{3}x^{-2/3}$$

Note that g is *not* differentiable at $x = 0$!



We could similarly compute the derivative for all rational exponents, though it is much easier to wait for the chain rule. The power law for irrational exponents is somewhat more ticklish.

Corollary 3.5 (Basic Transcendental Functions). *Recalling our development of power series in Chapter 2, the power law (for positive integers!) is all we need to see that*

$$\frac{d}{dx}\exp(x) = \exp(x), \qquad \frac{d}{dx}\sin x = \cos x, \qquad \frac{d}{dx}\cos x = -\sin x$$

It is also possible to develop these results independently of power series (see e.g. Exercise 12).

Failure of differentiability

It is instructive to consider how a function might fail to be differentiable. Firstly, a familiar fact shows that functions are not differentiable at discontinuities.

Lemma 3.6. *If f is differentiable at a then f is continuous at a.*

Proof. Just take the limit (think carefully why this works!):

$$\lim_{x\to a} f(x) = \lim_{x\to a}\left[\frac{f(x)-f(a)}{x-a}(x-a)+f(a)\right] = f'(a)(0-0)+f(a) = f(a)$$ ∎

It remains to consider situations when a function is continuous but not differentiable.

Examples 3.7. The following exemplify all situations where a function is continuous on an interval and differentiable everywhere except at a single interior point. As with isolated discontinuities, these are classified by considering the three ways in which the derivative limit might not converge.

1. A *vertical tangent line* occurs when the limit is infinite. For instance, $g(x) = x^{1/3}$ at $x = 0$.

2. *Corners* occur when the one-sided limits are unequal (could be infinite). For instance, $f(x) = |x|$ is not differentiable at zero, with one-sided limits

   $$\lim_{x\to 0^+}\frac{|x|-|0|}{x-0} = \lim_{x\to 0^+}\frac{x}{x} = 1 \neq \lim_{x\to 0^-}\frac{|x|-|0|}{x-0} = \lim_{x\to 0^-}\frac{-x}{x} = -1$$

   Indeed f is differentiable everywhere except at zero, with

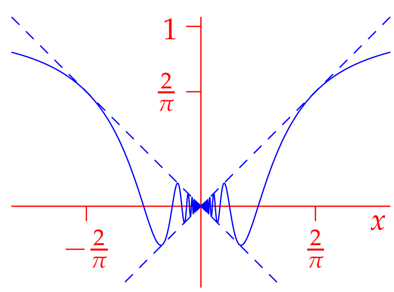
   $$f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

   A *cusp* describes the special case where the one-sided limits are $\infty \neq -\infty$.

3. A *singularity* is where left- and/or right-limits do not exist. The standard example is

   $$f(x) = \begin{cases} x\sin\frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

   which is continuous on $\mathbb{R}$ and differentiable everywhere except at zero: the details are in Exercise 10.



Singularities and vertical tangent lines can also prevent one-sided differentiability.

More esoteric examples of non-differentiability are possible:

- Utilizing series, we can create functions which are continuous on an interval but *nowhere differentiable*! For an example, see Exercise 15.
- It is also possible to construct a function which differentiable (and thus continuous) at precisely one point; can you think of an example?

The Basic Rules of Differentiation

Theorem 3.8. *Let f, g be differentiable and k, l be constants.*

1. *(Linearity) The function $kf + lg$ is differentiable with $(kf + lg)' = kf' + lg'$.*
2. *(Product rule) The function fg is differentiable with $(fg)' = f'g + fg'$.*
3. *(Inverse functions) Suppose f is bijective, $b = f^{-1}(a)$ is an interior point of $\operatorname{dom} f^{-1}$, and $f'(a) \neq 0$, then f^{-1} is differentiable at b and*

$$\left.\frac{\mathrm{d}}{\mathrm{d}y}\right|_{y=b} f^{-1}(y) = \frac{1}{f'(a)} = \frac{1}{f'\big(f^{-1}(b)\big)}$$

Proof. Parts 1 and 2 follow from the limit laws:

$$\lim_{x\to a} \frac{(kf+lg)(x)-(kf+lg)(a)}{x-a} = \lim_{x\to a}\left[k\frac{f(x)-f(a)}{x-a} + l\frac{g(x)-g(a)}{x-a}\right] = kf'(a) + lg'(a)$$

$$\lim_{x\to a} \frac{f(x)g(x)-f(a)g(a)}{x-a} = \lim_{x\to a}\left[\frac{f(x)-f(a)}{x-a}g(x) + f(a)\frac{g(x)-g(a)}{x-a}\right] = f'(a)g(a) + f(a)g'(a)$$

Note where we used the continuity of g in the second line ($\lim g(x) = g(a)$). Part 3 is an exercise. ∎

The inverse function rule should be intuitive: since the graphs of f and f^{-1} are related by reflection in the diagonal $y = x$, gradients at corresponding points are reciprocals. The result feels even more natural in Leibniz's notation: $\frac{\mathrm{d}x}{\mathrm{d}y} = \frac{1}{\mathrm{d}y/\mathrm{d}x}$.

Examples 3.9. 1. Linearity permits the differentiation of any polynomial: e.g.,

$$\frac{\mathrm{d}}{\mathrm{d}x}(7x^2 + 13x^4) = 7\frac{\mathrm{d}}{\mathrm{d}x}x^2 + 13\frac{\mathrm{d}}{\mathrm{d}x}x^4 = 14x + 52x^3$$

2. The product rule extends the reach of differentiation to include simple combinations: e.g.,

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^4 \sin x) = \left(\frac{\mathrm{d}}{\mathrm{d}x}x^4\right)\sin x + x^4\frac{\mathrm{d}}{\mathrm{d}x}\sin x = 4x^3\sin x - x^4\cos x$$

3. Inverse trigonometric functions can now be differentiated: e.g.,

$$y = \sin^{-1} x \implies \frac{\mathrm{d}}{\mathrm{d}x}\sin^{-1} x = \frac{\mathrm{d}y}{\mathrm{d}x} = \left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^{-1} = \frac{1}{\cos y} = \frac{1}{\sqrt{1-\sin^2 y}} = \frac{1}{\sqrt{1-x^2}}$$

4. Define the natural logarithm to be the inverse of the (bijective!) exponential function $\exp(x)$:

$$y = \ln x \iff x = \exp y$$

It follows that

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln x = \left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^{-1} = \frac{1}{\exp y} = \frac{1}{x}$$

The full details, and the justification that $\exp x = e^x$, are in Exercise 14.

Theorem 3.10 (Chain Rule). *If g is differentiable at a, and f is differentiable at $g(a)$, then $f \circ g$ is differentiable at a with derivative*

$$(f \circ g)'(a) = f'\big(g(a)\big)\, g'(a)$$

In Leibniz's notation, $\frac{\mathrm{d}(f \circ g)}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\mathrm{d}g}{\mathrm{d}x}$: this *looks* like a simple cancellation of the $\mathrm{d}g$ terms…[14]

Proof. Since f and g are differentiable, a is interior to $\mathrm{dom}(g)$ and $g(a)$ is interior to $\mathrm{dom}(f)$. Since g is continuous at a, there must exist some open interval $U \ni a$ for which $x \in U \implies g(x) \in \mathrm{dom}(f)$. Define $\gamma : \mathrm{dom}(f) \to \mathbb{R}$ via

$$\gamma(v) = \begin{cases} \frac{f(v) - f\left(g(a)\right)}{v - g(a)} & \text{if } v \neq g(a) \\ f'\big(g(a)\big) & \text{if } v = g(a) \end{cases} \tag{$*$}$$

Since f is differentiable at $g(a)$,we see that γ is continuous there: indeed $\lim\limits_{v \to g(a)} \gamma(v) = f'\big(g(a)\big)$. For any $x \in U \setminus \{a\}$, let $v = g(x)$ in $(*)$. Then

$$\frac{f\big(g(x)\big) - f\big(g(a)\big)}{x - a} = \gamma\big(g(x)\big)\frac{g(x) - g(a)}{x - a}$$

Take limits as $x \to a$ for the result. ∎

Corollary 3.11 (Quotient Rule). *Suppose f and g are differentiable. Then $\frac{f}{g}$ is differentiable whenever $g(x) \neq 0$. Moreover*

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$$

The proof is an exercise.

Examples 3.12. 1. By the quotient rule,

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan x = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\sin x}{\cos x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \sec^2 x$$

2. We can now differentiate highly involved combinations of elementary functions:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\tan(e^{4x^2}) - \frac{7x}{\sin x}\right] = 8xe^{4x^2}\sec^2(e^{4x^2}) - \frac{7\sin x - 7x\cos x}{\sin^2 x}$$

[14]This is completely unjustified since $\mathrm{d}g$ does not (for us) have independent meaning. The same problem appears in a famously flawed one-line 'proof' of the chain rule:

$$\lim_{x \to a}\frac{f\big(g(x)\big) - f\big(g(a)\big)}{x - a} \stackrel{?}{=} \lim_{x \to a}\frac{f\big(g(x)\big) - f\big(g(a)\big)}{g(x) - g(a)}\lim_{x \to a}\frac{g(x) - g(a)}{x - a}$$

The second limit doesn't make sense unless $g(x) \neq g(a)$ for all x on some punctured neighborhood of a: in particular, $g(x)$ cannot be *constant*! The faulty argument may be repaired by replacing this difference quotient with $f'\big(g(a)\big)$ whenever $g(x) = g(a)$, *before* taking the limit. This is precisely what $\gamma\big(g(x)\big)$ does in the correct proof.

Exercises 3.28. *Key concepts: Differentiability, Basic rules: linearity, power, product, chain, quotient*

1. Use Definition 3.1 to calculate the derivatives.

   (a) $f(x) = x^3$ at $x = 2$
   (b) $g(x) = x + 2$ at $x = a$
   (c) $f(x) = x^2 \cos x$ at $x = 0$
   (d) $r(x) = \frac{3x+4}{2x-1}$ at $x = 1$

2. Differentiate the function $f(x) = \cos\left(e^{x^5-3x}\right)$ using the chain and product rules.

3. (a) Prove the quotient rule (Corollary 3.11) by combining the chain and product rules.
   (b) Prove the inverse derivative rule (Theorem 3.8, part 3).
   (*Hint: You can't simply differentiate* $1 = \frac{\mathrm{d}x}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x} f\left(f^{-1}(x)\right)$ *using the chain rule; why not?*)

4. (a) Find the derivatives of secant, cosecant and cotangent using the quotient rule.
   (b) Why did we choose the positive square-root when computing $\frac{\mathrm{d}}{\mathrm{d}x} \sin^{-1} x$? What is the standard domain of arcsine, and what happens at $x = \pm 1$?
   (c) Find the derivatives of the inverse trigonometric functions using the inverse function rule.

5. Using the definition of the derivative, and supposing that f is differentiable at a, prove that
$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = \lim_{h \to 0} \frac{f(a+h) - f(a-h)}{2h}$$

6. Use induction to prove the power law $\frac{\mathrm{d}}{\mathrm{d}x} x^n = nx^{n-1}$ when $n \in \mathbb{N}$ using *only* the product rule and the fact that $\frac{\mathrm{d}}{\mathrm{d}x} x = 1$.

7. Prove that $f(x) = x\,|x|$ is differentiable everywhere and compute its derivative.

8. Show that $f(x) = x^{2/3}$ has a *cusp* (see Example 3.7.2) at $x = 0$.

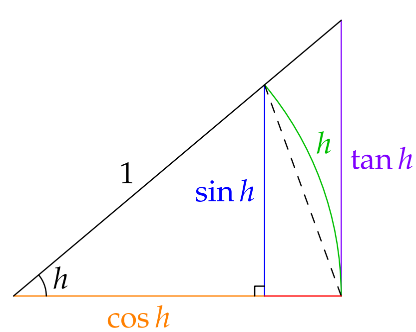
9. Show that following function is differentiable everywhere and compute its derivative:
$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$
   Moreover, prove that the derivative f' is *discontinuous* at $x = 0$.

10. Prove that the function in Example 3.7.3 is differentiable everywhere *except* at $x = 0$.

11. Suppose $f(x) = x^2$ whenever $x \in \mathbb{Q}$ and $f(x) = 0$ whenever $x \notin \mathbb{Q}$. At what values of x is f differentiable? Prove your assertion.

12. (a) Suppose $0 < h < \frac{\pi}{2}$. Use the picture to show that
$$0 < \frac{1 - \cos h}{h} < \sin \frac{h}{2} \quad \text{and} \quad \sin h < h < \tan h$$
   Hence conclude that $\lim_{h \to 0} \frac{\sin h}{h} = 1$ and $\lim_{h \to 0} \frac{1 - \cos h}{h} = 0$.

   (b) Use part (a) to prove that $\frac{\mathrm{d}}{\mathrm{d}x} \sin x = \cos x$



13. (Hard) Use induction to prove the Leibniz rule (general product rule):

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$$

Warning! The last two exercises are much longer and & tougher: have a go if you appreciate a challenge.

14. **The Exponential Function & the Power Law**

The ratio tests shows that the power series $\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}$ converges for all real x. *Define* $e := \exp(1)$. Certainly e^x makes sense whenever $x \in \mathbb{Q}$. If x is irrational, instead define

$$e^x := \sup\{e^q : q \in \mathbb{Q},\ q < x\}$$

The goal of this question is to *prove* that $\exp(x) = e^x$. As a nice bonus we recover Bernoulli's limit identity $e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$ and obtain a complete proof of the power law!

(a) For all $x, y \in \mathbb{R}$, prove that $\exp(x + y) = \exp(x)\exp(y)$

(*Hint: use the binomial theorem and change the order of summation*)

(b) Show that $\exp(x)$ is always positive, *even when* $x < 0$.

(c) Prove that $\exp : \mathbb{R} \to (0, \infty)$ is bijective.

(*Hint:* $x \geq 0 \implies \exp(x) \geq 1 + x$*; take limits then apply part (a)*)

(d) Prove that $e^x = \exp(x)$. Do this in three stages:

- If $x \in \mathbb{N}$, use part (a). Now check for $x \in \mathbb{Z}^-$.
- If $x = \frac{m}{n} \in \mathbb{Q}$, first compute $\left[\exp(\frac{m}{n})\right]^n$.
- If x is irrational, consider a sequence of rational numbers $q_n < x$ with $e^{q_n} \to e^x \ldots$

(e) Let $\ln : (0, \infty) \to \mathbb{R}$ be the inverse function of exp. Prove the logarithm laws:

$$\ln(xy) = \ln x + \ln y \quad \text{and} \quad \ln x^r = r \ln x$$

(*Just do this when* $r \in \mathbb{N}$*; in general, another argument like part (d) is required*)

(f) We've already seen that $\frac{\mathrm{d}}{\mathrm{d}y} \ln y = \frac{1}{y}$. Use the fact that

$$\frac{\mathrm{d}}{\mathrm{d}y} \ln y = \lim_{h\to 0} \frac{\ln(y + h) - \ln y}{h}$$

to prove that $\exp(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$, thus recovering Bernoulli's definition of e.

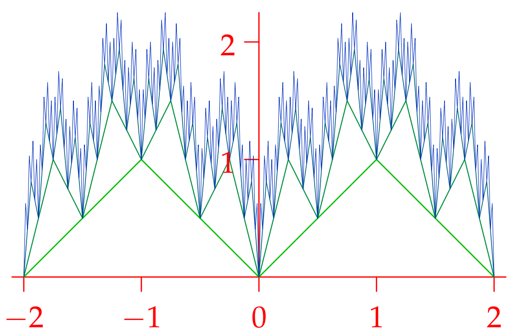
(g) For any $r \in \mathbb{R}$, *define* $x^r := \exp(r \ln x)$. Hence obtain the power law for any exponent.

15. **A Very Strange Function**

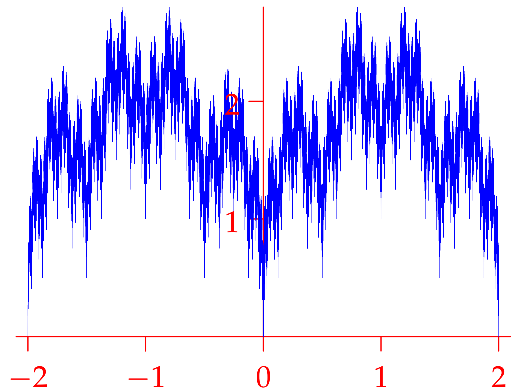
Here is a classic example of a continuous but nowhere-differentiable function!

Let f be the *sawtooth* function defined by $f(x) = |x|$ whenever $x \in [-1, 1]$ and extending periodically to $\mathbb{R}$ so that $f(x+2) = f(x)$. Now define $g : \mathbb{R} \to \mathbb{R}$ via

$$g(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n f(4^n x)$$



$f(x)$ and iterations to $n = 3$



$g(x)$ (really $n = 6$, but can you tell?!)

(a) Prove that g is well-defined and continuous on $\mathbb{R}$.

(b) Let $x \in \mathbb{R}$ and $m \in \mathbb{N}$ be fixed. Define $h_m = \pm\frac{1}{2} \cdot 4^{-m}$ where the sign is chosen so that no integers lie strictly between $4^m x$ and $4^m(x + h_m) = 4^m x \pm \frac{1}{2}$.

For each $n \in \mathbb{N}_0$, define

$$k_n = \frac{f\big(4^n(x + h_m)\big) - f(4^n x)}{h_m}$$

Prove the following

i. $|k_n| \le 4^n$ with equality when $n = m$.

ii. $n > m \implies k_n = 0$.

(*Hint: $|f(y) - f(z)| \le |y - z|$: when is this an equality?*)

(c) Use part (b) to prove that

$$\left|\frac{g(x + h_m) - g(x)}{h_m}\right| \ge \frac{1}{2}(3^m + 1)$$

Hence conclude that g is *nowhere differentiable.*

3.29 The Mean Value Theorem

A key result in elementary calculus, this should be very familiar from your previous studies.

Theorem 3.13 (Mean Value Theorem/MVT). *Let f be continuous on $[a,b]$ and differentiable on (a,b). Then there exists $\xi \in (a,b)$ such that $f'(\xi) = \frac{f(b)-f(a)}{b-a}$.*

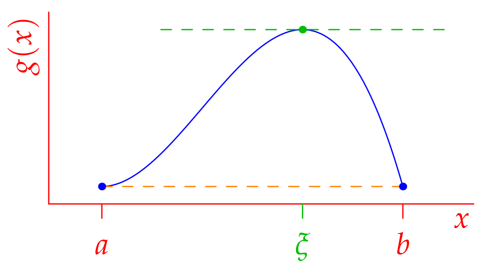
This follows easily from two lemmas.

Lemma 3.14.
1. *(Critical Points) Suppose g is bounded on (a,b) and attains its maximum or minimum at $\xi \in (a,b)$. If g is differentiable at ξ then $g'(\xi) = 0$.*
2. *(Rolle's Theorem) Suppose g is continuous on $[a,b]$, differentiable on (a,b), and $g(a) = g(b)$. Then there exists $\xi \in (a,b)$ such that $g'(\xi) = 0$.*

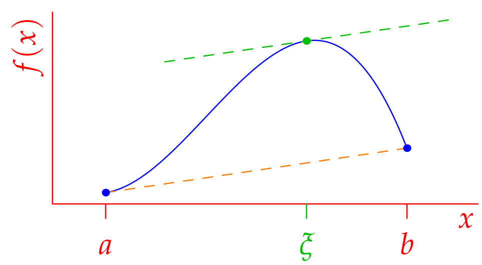
The main result is obtained by subtracting a straight line and applying Rolle's theorem to

$$g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$$

and observing that $g(a) = f(a) = g(b)$ and $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$.



Critical Points/Rolle's Theorem



Mean Value Theorem

In the pictures, the orange and green lines are *parallel*: the average slope over the interval $[a,b]$ equals the gradient/derivative $f'(\xi)$.

Proof of Lemma. 1. Suppose $\xi \in (a,b)$ is a maximum: that is, $g(x) \le g(\xi)$ for all $x \neq \xi$. Then

$$\frac{g(x)-g(\xi)}{x-\xi} \begin{cases} \le 0 & \text{whenever } x > \xi \\ \ge 0 & \text{whenever } x < \xi \end{cases}$$

Now take the one-sided limits: since g is differentiable at ξ, we see that

$$0 \le \lim_{x \to \xi^+} \frac{g(x)-g(\xi)}{x-\xi} = g'(\xi) = \lim_{x \to \xi^-} \frac{g(x)-g(\xi)}{x-\xi} \le 0$$

Otherwise said $g'(\xi) = 0$. The case when ξ is a minimum is similar.

2. By the Extreme Value Theorem (1.11), g is bounded and attains its bounds. If the extrema *both* occur at the endpoints a, b, then g is constant: any $\xi \in (a,b)$ satisfies the result. Otherwise, at least one extreme occurs at some $\xi \in (a,b)$: part 1 says that $g'(\xi) = 0$. ∎

Examples 3.15. 1. Let $f(x) = (x-1)^2(4-x) + x$ on $[a,b] = [1,4]$: this is roughly the above picture illustrating the mean value theorem. Compute the average slope and the derivative,

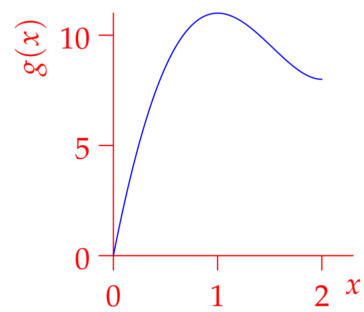
$$\frac{f(b)-f(a)}{b-a} = 1, \qquad f'(x) = 2(x-1)(4-x) - (x-1)^2 + 1 = -3x^2 + 12x - 8$$

and observe that

$$f'(\xi) = \frac{f(b)-f(a)}{b-a} \iff 3\xi^2 - 12\xi + 9 = 0 \iff \xi = 1 \text{ or } 3$$

Since only 3 lies in the interval $(1,4)$, this is the value ξ satisfying the mean value theorem.

2. We find the maximum and minimum values of $g(x) = x^4 - 14x^2 + 24x$ on the interval $[0,2]$. The function is differentiable, with



$$g'(x) = 4x^3 - 28x + 24 = 4(x-2)(x-1)(x+3)$$

By the Lemma, the locations of the extrema are either the end-points $x = 0, 2$ or locations with zero derivative ($x = 1$). Since

$$f(0) = 0, \quad f(1) = 11, \quad f(2) = 8$$

we conclude that $\max(f) = f(1) = 11$ and $\min(f) = f(0) = 0$.

Consequences of the Mean Value Theorem Several simple corollaries relate to monotonicity.

Definition 3.16. Suppose $f : I \to \mathbb{R}$ is defined on an interval I. We say that f is:

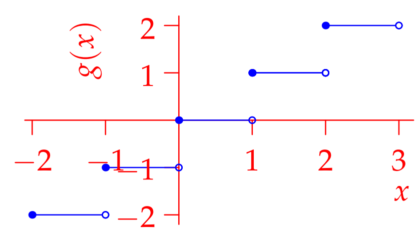
Increasing (monotone-up) on I if $x < y \implies f(x) \le f(y)$

Decreasing (monotone-down) on I if $x < y \implies f(x) \ge f(y)$

We say *strictly* increasing/decreasing if the inequalities are strict.

Examples 3.17. 1. $f : x \mapsto x^2$ is strictly increasing on $[0,\infty)$ and strictly decreasing on $(-\infty, 0]$.

2. The floor function $f : x \mapsto \lfloor x \rfloor$ (the greatest integer less than or equal to x) is increasing, but not strictly, on $\mathbb{R}$.



Corollary 3.18. *Suppose f is differentiable on an interval I. Then*

1. *$f' \ge 0$ on I $\iff$ f is increasing on I*
2. *$f' \le 0$ on I $\iff$ f is decreasing on I*
3. *$f' = 0$ on I $\iff$ f is constant on I*

Proof. (Part 1, $\Rightarrow$) Let $x < y$ where $x, y \in I$. By the mean value theorem, $\exists \xi \in (x, y)$ such that

$$\frac{f(y) - f(x)}{y - x} = f'(\xi) \quad \text{whence} \quad f'(\xi) \geq 0 \implies f(y) \geq f(x)$$

($\Leftarrow$) For the converse, use the definition of derivative: $f'(\xi) = \lim_{x \to \xi} \frac{f(x) - f(\xi)}{x - \xi}$. If f is increasing, then

$$x > \xi \implies f(x) \geq f(\xi) \implies f'(\xi) \geq 0$$

Parts 2 and 3 are similar. ∎

More care is required when relating $f' > 0$ to f being *strictly* increasing (see Exercise 5). The corollary also yields a couple of (hopefully familiar) flashbacks to elementary calculus.

Corollary 3.19. *Let I be an open interval.*

1. *(Anti-derivatives on an interval) If $f'(x) = g'(x)$ on I, then $\exists c$ such that $g(x) = f(x) + c$ on I.*
2. *(First derivative test) Suppose f is continuous on I and differentiable except perhaps at ξ. If*

$$\begin{cases} f'(x) < 0 & \textit{whenever } x < \xi, \textit{ and} \\ f'(x) > 0 & \textit{whenever } x > \xi \end{cases} \quad \textit{then } f \textit{ has its minimum value at } x = \xi$$

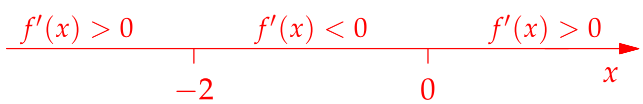
The statement for a maximum is similar.

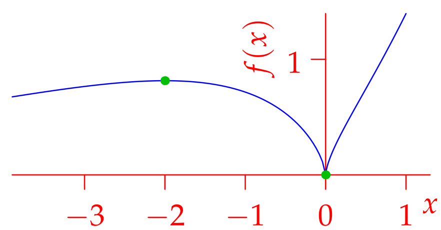
Examples 3.20. 1. Since $\frac{\mathrm{d}}{\mathrm{d}x} \sin(3x^2 + x) = (6x + 1)\cos(3x^2 + x)$ on (the interval) $\mathbb{R}$, whence all anti-derivatives of $f(x) = (6x + 1)\cos(3x^2 + x)$ are given by

$$\int f(x)\,\mathrm{d}x = \int (6x + 1)\cos(3x^2 + x)\,\mathrm{d}x = \sin(3x^2 + x) + c$$

As is typical in calculus, we use the *indefinite integral* notation $\int f(x)\,\mathrm{d}x$ for anti-derivatives.

2. If $f(x) = x^{2/3} e^{x/3}$, then $f'(x) = \frac{1}{3} x^{-1/3} (2 + x) e^{x/3}$.
By Lemma 3.14, the only possible critical points are at $x = 0$ or -2. The sign of the derivative is also clear:





By the 1st derivative test, f has a maximum at $x = -2$ and a minimum at $x = 0$.

We finish this section by tying together the mean and intermediate value theorems.

Theorem 3.21 (IVT for Derivatives). *Suppose f is differentiable on an interval I containing $a < b$, and that L lies between $f'(a)$ and $f'(b)$. Then $\exists \xi \in (a, b)$ such that $f'(\xi) = L$.*

If $f'(x)$ is *continuous*, this is just the intermediate value theorem applied to f'; surprisingly, continuity of f' is *not* required. A full proof is in Exercise 7.

Exercises 3.29. *Key concepts: Differentiability, Basic rules: linearity, power, product, chain, quotient*

1. Determine whether the conclusion of the mean value theorem holds for each function on the given interval. If so, find a suitable point ξ. If not, state which hypothesis fails.

   (a) x^2 on $[-1,2]$ (b) $\sin x$ on $[0,\pi]$ (c) $|x|$ on $[-1,2]$
   (d) $1/x$ on $[-1,1]$ (e) $1/x$ on $[1,3]$

2. Suppose f and g are differentiable on an interval I containing $a<b$ and that $f(a)=f(b)=0$. By considering $h(x)=f(x)e^{g(x)}$, prove that $f'(\xi)+f(\xi)g'(\xi)=0$ for some $\xi\in(a,b)$.

3. (a) Use the Mean Value Theorem to prove that $x<\tan x$ for all $x\in(0,\frac{\pi}{2})$.
   (b) Prove that $\frac{x}{\sin x}$ is *strictly* increasing on $(0,\frac{\pi}{2})$.
   (c) Prove that $x\le\frac{\pi}{2}\sin x$ for all $x\in[0,\frac{\pi}{2}]$.

4. Suppose that $|f(x)-f(y)|\le(x-y)^2$ for all $x,y\in\mathbb{R}$. Prove that f is a constant function.

5. (a) Prove that $f'>0$ on an interval $I\implies f$ is *strictly* increasing on I.
   (b) Show that the converse of part (a) is *false.*
   (c) Carefully prove the first derivative test (Corollary 3.19).

6. If f is differentiable on an interval I such that $f'(x)\neq 0$ for all $x\in I$, use the intermediate value theorem for derivatives to prove that f is either strictly increasing or strictly decreasing.

7. (Intermediate value theorem for derivatives) Let f,a,b and L be as in Theorem 3.21, define $g:I\to\mathbb{R}$ by $g(x)=f(x)-Lx$, and let $\xi\in[a,b]$ be such that
   $$g(\xi)=\min\{g(x):x\in[a,b]\}$$
   (a) Why can we be sure that ξ exists? If $\xi\in(a,b)$, explain why $f'(\xi)=L$.
   (b) Assume WLOG that $f'(a)<f'(b)$. Prove that $g'(a)<0<g'(b)$. By considering $\lim\limits_{x\to a^+}\frac{g(x)-g(a)}{x-a}$, show that $\exists x>a$ for which $g(x)<g(a)$. Hence complete the proof.

8. Suppose f' exists on (a,b), and is continuous except for a discontinuity at $c\in(a,b)$.

   (a) Suppose $\lim_{x\to c^+}f'(x)=L<f'(c)$. By taking $\epsilon=\frac{f'(c)-L}{2}$ in the definition of this limit and applying IVT for derivatives, obtain a contradiction.
   Hence argue that c cannot be a *removable* or a *jump* discontinuity.
   (b) Similarly, show that f' cannot have an *infinite* discontinuity by considering $\lim\limits_{x\to c^+}f'(x)=\infty$.
   (c) By parts (a) and (b), It remains to see that f' can have an essential discontinuity. Recall (Exercise 3.28.9) that
   $$f:\mathbb{R}\to\mathbb{R}:x\mapsto\begin{cases}x^2\sin\frac{1}{x} & x\neq 0\\ 0 & x=0\end{cases}$$
   is differentiable on $\mathbb{R}$, but has discontinuous derivative at $x=0$.
      i. Use $x_n=\frac{1}{2n\pi}$ and $y_n=\frac{1}{(2n+1)\pi}$ to show that f' has an essential discontinuity at $x=0$.
      ii. Prove that if $\lim s_n=0$ and $\lim f'(s_n)=M$, then $M\in[-1,1]$.
      iii. Prove that for any $L\in[-1,1]$, there is a sequence (t_n) for which $\lim f'(t_n)=L$.
      (*Hint: Use IVT for derivatives*)

3.30 L'Hôpital's Rule

We are often required to consider *indeterminate forms*: limits which do not yield easily to the standard limits laws. For instance, while it is tempting to write

$$\lim_{x\to 0} \frac{\sin 2x}{e^{3x}-1} = \frac{\lim \sin 2x}{\lim e^{3x}-1} = \frac{0}{0} \tag{$*$}$$

this is an incorrect application of the limit laws since the resulting quotient has no meaning.

Definition 3.22. An *indeterminate form* is any limit where a naïve application of the limit laws results in a meaningless expression: the primary types are $\frac{0}{0}$, $\frac{\infty}{\infty}$, $\infty-\infty$, $0\cdot\infty$, 0^0, 0^∞, and 1^∞.

Examples 3.23. 1. $\lim\limits_{x\to 7^+}(x-7)^{\frac{1}{x-7}}$ is an indeterminate form of type 0^∞.

2. Our motivating example $(*)$ may correctly be evaluated using the definition of the derivative:

$$\lim_{x\to 0} \frac{\sin 2x}{e^{3x}-1} = \lim_{x\to 0} \frac{\sin 2x - 0}{x-0}\,\frac{x-0}{e^{3x}-1} = \left(\frac{\mathrm{d}}{\mathrm{d}x}\Big|_{x=0} \sin 2x\right)\left(\frac{\mathrm{d}}{\mathrm{d}x}\Big|_{x=0} e^{3x}\right)^{-1} = \frac{2}{3}$$

By considering $\lim\limits_{x\to 0}\frac{3a\sin 2x}{2(e^{3x}-1)}$, we see that an indeterminate form of type $\frac{0}{0}$ can take *any value* a!

The approach generalizes, if non-rigorously: if f, g are differentiable at a and $f(a) = 0 = g(a)$, then

$$\boxed{\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f(x)-f(a)}{x-a}\cdot\frac{x-a}{g(x)-g(a)} = \frac{f'(a)}{g'(a)}}$$

Our goal is to fully justify this result and extend to several situations:

- One-sided limits, including when $a = \pm\infty$.
- When $\lim f(x) = 0$ exists, but $f(a)$ does not ($g(x), g(a)$ similarly).
- Indeterminate forms of type $\frac{\infty}{\infty}$ ($\lim f(x) = \infty$, etc.).
- When the RHS cannot be cleanly evaluated: for instance $g'(a) = 0$ or if the original limit is $\pm\infty$.

Here is the full result.

Theorem 3.24 (L'Hôpital's Rule). *Let $a \in \mathbb{R}\cup\{\pm\infty\}$ and suppose functions f and g satisfy:*

1. $\lim\limits_{x\to a}\frac{f'(x)}{g'(x)} = L$ *for some* $L \in \mathbb{R}\cup\{\pm\infty\}$*, and,*

2. *(a)* $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$, *or (b)* $\lim\limits_{x\to a} g(x) = \infty$ *(no condition on f)*

Then $\lim\limits_{x\to a}\frac{f(x)}{g(x)} = L$*. The same result holds for one-sided limits.*

The full proof is a behemoth—we postpone this until after several examples. In part because of this, and because examples can often be evaluated more instructively using elementary methods (as in the above example), l'Hôpital's rule is often discouraged in elementary calculus.

Examples 3.25. 1. If $f(x) = e^{4x}$ and $g(x) = 21x - 17$, then $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)}$ has type $\frac{\infty}{\infty}$. By l'Hôpital's rule,

$$\lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to\infty} \frac{4e^{4x}}{21} = \infty \implies \lim_{x\to\infty} \frac{e^{4x}}{21x - 17} = \infty$$

2. For an example of type $\frac{0}{0}$, consider $f(x) = x^2 - 9$ and $g(x) = \ln(4 - x)$:

$$\lim_{x\to 3^-} \frac{f'(x)}{g'(x)} = \lim_{x\to 3^-} \frac{2x}{-1/(4-x)} = \lim_{x\to 3^-} 2x(x-4) = -6 \implies \lim_{x\to 3^-} \frac{x^2 - 9}{\ln(4-x)} = -6$$

3. One can apply the rule repeatedly: for example

$$\lim_{x\to 0} \frac{e^{4x} - 1 - 4x}{x^2} = \lim_{x\to 0} \frac{4e^{4x} - 4}{2x} = \lim_{x\to 0} \frac{16e^{4x}}{2} = 8$$

This is a generally accepted abuse of protocol: one shouldn't really state the first limit until one knows the last limit exists! As long as everything works, you are fine. However...

4. It is crucially important that the limit $\lim \frac{f'}{g'}$ exists *before* applying l'Hôpital's rule! Consider $f(x) = x + \cos x$ and $g(x) = x$: certainly $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)}$ has type $\frac{\infty}{\infty}$, however

$$\lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to\infty} 1 - \sin x$$

does not exist! In this case the rule is unnecessary: appealing to the squeeze theorem,

$$\frac{f(x)}{g(x)} = 1 + \frac{\cos x}{x} \xrightarrow[x\to\infty]{} 1$$

5. For another reason for why l'Hôpital's rule is often prohibited in Freshman calculus, consider

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

This appears legitimate. However, recall (Exercise 3.28.12) that this limit is used to demonstrate $\frac{\mathrm{d}}{\mathrm{d}x} \sin x = \cos x$; to use this to calculate the limit on which it depends is circular logic!

The remaining indeterminate forms (Definition 3.22) may be modified so that l'Hôpital's rule applies.

Examples 3.26. 1. An indeterminate form of type $\infty - \infty$ may be transformed to one of type $\frac{0}{0}$ before applying the rule (twice):

$$\begin{aligned}
\lim_{x\to 0^+} \frac{1}{e^x - 1} - \frac{1}{x} &= \lim_{x\to 0^+} \frac{x + 1 - e^x}{x(e^x - 1)} && (\text{type } \tfrac{0}{0}) \\
&= \lim_{x\to 0^+} \frac{1 - e^x}{e^x - 1 + xe^x} && (\text{still type } \tfrac{0}{0}) \\
&= \lim_{x\to 0^+} \frac{-e^x}{2e^x + xe^x} = -\frac{1}{2}
\end{aligned}$$

2. For an indeterminate form of type 1^∞, we use the log laws & continuity of the exponential:

$$\begin{aligned}\lim_{x\to 0^+}(1+\sin x)^{1/x} &= \exp\left(\lim_{x\to 0^+}\frac{1}{x}\ln(1+\sin x)\right) && (\text{type } \tfrac{0}{0})\\ &= \exp\left(\lim_{x\to 0^+}\frac{\cos x}{1+\sin x}\right) = e^1 = e\end{aligned}$$

Proving l'Hôpital's Rule

The complete argument is very lengthy. It starts with an extension of the Mean Value Theorem.

Lemma 3.27 (Extended Mean Value Theorem). *Fix $a<b$, suppose f,g are continuous on $[a,b]$ and differentiable on (a,b). Then there exists $\xi\in(a,b)$ such that*

$$\big(f(b)-f(a)\big)g'(\xi) = \big(g(b)-g(a)\big)f'(\xi)$$

Proof. Apply the standard mean value theorem (really Rolle's theorem) to

$$h(t) = \big(f(b)-f(a)\big)g(t) - \big(g(b)-g(a)\big)f(t)$$

which satisfies $h(a)=h(b)$. ∎

Now for the main event. If you do nothing else, read the following proof of the simplest case. Everything else is a modification.

Proof (Case (a)/type $\frac{0}{0}$, with right limits). Suppose we have a form of type $\frac{0}{0} = \lim_{x\to a^+}\frac{f(x)}{g(x)}$ taking right-limits at a finite location a, and that the resulting limit L is finite.

First observe that condition 1 forces the existence of an interval (a,b) on which f,g are differentiable and $g'(x)\neq 0$. Everything follows from the definition the limit in condition 1, and Lemma 3.27:

$$\text{Given } \epsilon>0,\ \exists\delta\in(0,b-a) \text{ such that } a<\xi<a+\delta \implies \left|\frac{f'(\xi)}{g'(\xi)}-L\right|<\frac{\epsilon}{2} \qquad (*)$$

$$a<y<x<a+\delta \implies \exists\xi\in(y,x) \text{ such that } \frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(\xi)}{g'(\xi)} \qquad (\dagger)$$

Since $g'\neq 0$, the usual mean value theorem says

$$\exists c\in(y,x) \text{ such that } g(x)-g(y) = g'(c)(x-y)\neq 0$$

whence we never divide by zero in $(\dagger)$. Combining $(*)$ and $(\dagger)$, observe that

$$a<x<a+\delta \implies \left|\frac{f(x)}{g(x)}-L\right| \overset{2(a)}{=} \lim_{y\to a^+}\left|\frac{f(x)-f(y)}{g(x)-g(y)}-L\right| \overset{(\dagger)}{=} \lim_{y\to a^+}\left|\frac{f'(\xi)}{g'(\xi)}-L\right| \overset{(*)}{\leq} \frac{\epsilon}{2}<\epsilon$$

Note that $a<y<\xi(x,y)<x$ is a function of x,y here! Since $\epsilon>0$ is arbitrary, this is the required result. ∎

A complete proof for all indeterminate forms of type $\frac{0}{0}$ follows from some simple modifications.

If $a = -\infty$: Replace the blue part of $(*)$ as follows:

$$\text{Given } \epsilon > 0,\ \exists m \le b \text{ such that } \xi < m \implies \left| \frac{f'(\xi)}{g'(\xi)} - L \right| < \frac{\epsilon}{2}$$

The rest of the proof goes through after replacing a with $-\infty$ and $a + \delta$ with m.

If $L = \infty$: Replace the green parts of $(*)$ with Given $M > 0$ and $\frac{f'(\xi)}{g'(\xi)} > 2M$. Fixing the rest of the proof is again straightforward.

If $L = -\infty$: Replace the green parts of $(*)$ with Given $M > 0$ and $\frac{f'(\xi)}{g'(\xi)} < -2M$.

Left-limits: If f, g are differentiable on (c, a), then the blue part may be replaced with either:

- (a finite) $\exists \delta \in (0, a - c)$ such that $a - \delta < \xi < a$
- ($a = \infty$) $\exists m \ge c$ such that $\xi > m$

The blue and green parts of $(*)$ may be replaced independently.

Proof (Case (b), $\lim g(x) = \infty$*).* This requires a little more care.[15] Since $g' \neq 0$, and $\lim_{x \to a^+} g(x) = \infty$, Exercise 3.29.6 says that g is *strictly decreasing* on (a, b). By replacing b by some $\tilde{b} \in (a, b)$, if necessary, we may assume that

$$a < y < x < b \implies 0 < g(x) < g(y) \tag{‡}$$

Assume a and L are finite and obtain $(*)$ and $(\dagger)$ as before. Let $x \in (a, a + \delta)$ be fixed and multiply $(\dagger)$ by $\frac{g(y) - g(x)}{g(y)}$ (this is *positive* by (‡)): a little algebra and the triangle inequality tell us that

$$\begin{aligned} a < y < x &\implies \frac{f(y)}{g(y)} = \frac{f'(\xi)}{g'(\xi)} + \frac{f(x)}{g(y)} - \frac{g(x)}{g(y)} \cdot \frac{f'(\xi)}{g'(\xi)} \\ &\implies \left| \frac{f(y)}{g(y)} - L \right| \le \left| \frac{f'(\xi)}{g'(\xi)} - L \right| + \frac{1}{g(y)} \left(|f(x)| + |g(x)| \left(L + \frac{\epsilon}{2} \right) \right) \end{aligned}$$

Since $\lim_{y \to a^+} g(y) = \infty$ and x is fixed, we see that there exists $\eta \le x - a < \delta$ such that

$$y \in (a, a + \eta) \implies \frac{1}{g(y)} \left(|f(x)| + |g(x)| \left(L + \frac{\epsilon}{2} \right) \right) < \frac{\epsilon}{2}$$

Finally combine with $(*)$: given $\epsilon > 0, \exists \eta > 0$ such that $y \in (a, a + \eta) \implies \left| \frac{f(y)}{g(y)} - L \right| < \epsilon$.
The same modifications listed above complete the proof. ∎

[15] *Forms of type* $\frac{\infty}{\infty}$*?* Instead of assumption 2. (b), why not simply assume $\lim f = \lim g = \infty$ and write $\frac{f}{g} = \frac{1/g}{1/f}$ to obtain a form of type $\frac{0}{0}$? The problem is that the derivative of the 'new' denominator $\frac{\mathrm{d}}{\mathrm{d}x} \frac{1}{f} = \frac{-f'}{f^2}$ need not be non-zero on any interval (a, b) and so condition 1. need not hold. We could modify this, but it would make for a weaker theorem. Example 3.25.4 illustrates the issue: $f'(x) = 1 + \sin x$ has zeros on any unbounded interval.

After the 2. (b) case is proved and we know that $\lim \frac{f}{g} = L$, it is then clear that $\lim f$ must also be infinite (unless $L = 0$ in which case $\lim f$ could be anything and need not exist). This situation therefore really does deal with forms of type $\frac{\infty}{\infty}$.

Exercises 3.30. *Key concepts: Types of indeterminate forms, Formal statement of l'Hôpital's rule*

1. Evaluate the limits, if they exist:

   (a) $\lim_{x \to 0} \frac{x^3}{\sin x - x}$

   (b) $\lim_{x \to \frac{\pi}{2}^-} \tan x - \frac{2}{\pi - 2x}$

   (c) $\lim_{x \to 0} (\cos x)^{1/x^2}$

   (d) $\lim_{x \to 0} (1 + 2x)^{1/x}$

   (e) $\lim_{x \to \infty} (e^x + x)^{1/x}$

2. Suppose f is differentiable on (c, ∞) and that $\lim_{x \to \infty} [f(x) + f'(x)] = L$ is finite.

   (a) Prove that $\lim_{x \to \infty} f(x) = L$ and that $\lim_{x \to \infty} f'(x) = 0$.
   (*Hint: write* $f(x) = \frac{f(x)e^x}{e^x}$)

   (b) Does anything change if L exists and is *infinite*?

3. If $p_n(x)$ is a polynomial of degree n, use induction to prove that $\lim_{x \to \infty} p_n(x)e^{-x} = 0$

4. Let $f(x) = x + \sin x \cos x$, $g(x) = e^{\sin x} f(x)$ and $h(x) = \frac{2 \cos x}{e^{\sin x}(f(x) + 2 \cos x)}$

   (a) Prove that $\lim_{x \to \infty} f(x) = \infty = \lim_{x \to \infty} g(x)$ but that $\lim_{x \to \infty} \frac{f(x)}{g(x)}$ does not exist.

   (b) If $\cos x \neq 0$, and x is large, show that $\frac{f'(x)}{g'(x)} = h(x)$.

   (c) Prove that $\lim_{x \to \infty} h(x) = 0$. Explain why this does not contradict part (a)!

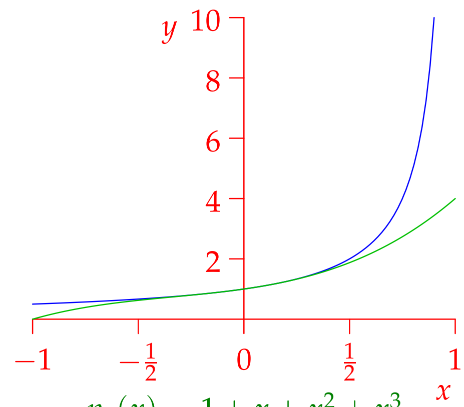
3.31 Taylor's Theorem

A primary goal of power series is the approximation of functions. With this in mind, there are two natural questions to ask of a function f:

1. Given $c \in \text{dom}(f)$, is there a series $\sum a_n(x-c)^n$ which equals $f(x)$ on an interval containing c?
2. If we take the first n terms of such a series, how accurate is this polynomial approximation?

Example 3.28. Recall the geometric series

$$f(x) = \frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \text{ whenever } -1 < x < 1$$



The polynomial approximation

$$p_n(x) = \sum_{k=0}^{n} x^k = 1 + x + \cdots + x^n = \frac{1-x^{n+1}}{1-x}$$

has error

$$R_n(x) = f(x) - p_n(x) = \frac{x^{n+1}}{1-x}$$

If x is close to 0, this is likely very small; for instance if $x \in \left[-\frac{1}{2}, \frac{1}{2}\right]$, then

$$|R_n(x)| \le \frac{1}{1-\frac{1}{2}}\left(\frac{1}{2}\right)^{n+1} = 2^{-n}$$

However, when x is close to 1 the error is unbounded!

The above behavior occurs in general: the truncated polynomials provide better approximations nearer the center of the series. To see this, we first need to consider higher-order derivatives.

Definition 3.29. We write f'' for the *second derivative* of f, namely the derivative of its derivative

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

The existence of $f''(a)$ presupposes that f' exists on an (open) interval containing a. We can similarly consider third, fourth, and higher-order derivatives. As a function, the n^{th} derivative is written

$$f^{(n)}(x) = \frac{\mathrm{d}^n f}{\mathrm{d}x^n}$$

By convention, the *zeroth derivative* is the function itself $f^{(0)}(x) = f(x)$. We say that f is *n times differentiable* at a if $f^{(n)}(a)$ exists, and *infinitely differentiable* (or *smooth*) if derivatives of all orders exist.

Example 3.30. $f(x) = x^2|x|$ is twice differentiable, with $f''(x) = 6|x|$. It is smooth everywhere except at $x = 0$, where third (and higher-order) derivatives do not exist.

Definition 3.31. Suppose f is n times differentiable at $x = c$. The n^{th} *Taylor polynomial* p_n of f centered at c is

$$p_n(x) := \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k = f(c) + f'(c)(x-c) + \frac{f''(c)}{2}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n$$

The *remainder* $R_n(x)$ is the error in the polynomial approximation

$$R_n(x) = f(x) - p_n(x) = f(x) - \sum_{j=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k$$

If f is infinitely differentiable at $x = c$, then its *Taylor series* centered at $x = c$ is the power series

$$\mathrm{T}_c f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!}(x-c)^n$$

When $c = 0$ this is known as a *Maclaurin series.*[16]

For simplicity we'll mostly work with Maclaurin series, with general situation hopefully being clear.

Examples 3.32. 1. If $f(x) = e^{3x}$, then $f^{(n)}(x) = 3^n e^x$, from which the Maclaurin series is

$$\mathrm{T}_0 f(x) = \sum_{n=0}^{\infty} \frac{3^n}{n!} x^n$$

2. If $g(x) = \sin 7x$, then the sequence of derivatives is

$$7\cos 7x,\quad -7^2 \sin 7x,\quad -7^3 \cos 7x,\quad 7^4 \sin 7x,\quad 7^5 \cos 7x,\quad -7^6 \sin 7x,\ \ldots$$

At $x = 0$, every even derivative is zero whereas the odd derivatives alternate in sign. The Maclaurin series is easily seen to be

$$\mathrm{T}_0 g(x) = \sum_{n=0}^{\infty} \frac{(-1)^n 7^{2n+1}}{(2n+1)!} x^{2n+1}$$

3. If $h(x) = \sqrt{x}$, then $h'(x) = \frac{1}{2}x^{-1/2}$, $h''(x) = \frac{-1}{2^2}x^{-3/2}$, and $h'''(x) = \frac{3}{2^3}x^{-5/2}$, from which the third Taylor polynomial centered at $c = 1$ is

$$\begin{aligned} p_2(x) &= h(1) + h'(1)(x-1) + \frac{h''(1)}{2}(x-1)^2 + \frac{h'''(1)}{6}(x-1)^3 \\ &= 1 + \frac{1}{2}(x-1) - \frac{1}{8}(x-1)^2 + \frac{1}{16}(x-1)^3 \end{aligned}$$

Rather than computing further examples, we first develop a little theory that makes verifying Taylor series much easier.

[16]Named for Englishman Brook Taylor (1685–1731) and Scotsman Colin Maclaurin (1698–1746). Taylor's general method expanded on examples discovered by James Gregory and Issac Newton in the mid-to-late 1600s.

Differentiation of Taylor Polynomials and Series

Suppose $P(x) = \sum a_j x^j$ is a power series with radius of convergence $R > 0$. As we saw previously (Theorem 2.31), $P(x)$ is differentiable term-by-term on $(-R, R)$. Indeed,

$$
\begin{aligned}
P'(x) &= \sum_{j=1}^{\infty} a_j j x^{j-1} \implies P'(0) = a_1 \\
P''(x) &= \sum_{j=2}^{\infty} a_j j(j-1) x^{j-2} \implies P''(0) = 2a_2 \\
P'''(x) &= \sum_{j=3}^{\infty} a_j j(j-1)(j-2) x^{j-3} \implies P'''(0) = 3!a_3 \\
&\vdots \\
P^{(k)}(x) &= \sum_{j=k}^{\infty} a_j j(j-1)\cdots(j-k+1) x^{j-k} = \sum_{j=k}^{\infty} \frac{j!a_j}{(j-k)!} x^{j-k} \implies P^{(k)}(0) = k!a_k
\end{aligned}
$$

Otherwise said, P is its own Maclaurin series! The same discussion holds for polynomials. Indeed if $P(x) = a_0 + a_1 x + \cdots + a_n x^n$ is a polynomial and f a function, then

$$P^{(k)}(0) = f^{(k)}(0) \iff a_k = \frac{f^{(k)}(0)}{k!}$$

If this holds for *all* $k \leq n$, then $P = p_n$ is the n^{th} Taylor polynomial of f! With a little modification, we've proved the following:

Theorem 3.33. 1. *If $f(x) = \sum a_n (x-c)^n$ is a power series defined on a neighborhood of c, then $T_c f(x) = f(x)$: the function is its own Taylor series!*

2. *The n^{th} Taylor polynomial of f centered at $x = c$ is the unique polynomial p_n of degree $\leq n$ whose value and first n derivatives agree with those of f at $x = c$: that is*

$$\forall k \leq n, \ \ p_n^{(k)}(c) = f^{(k)}(c)$$

This answers our first motivating question: a function can equal at most one power series with a given center. The second question requires a careful study of the *remainder*: we'll do this shortly.

Examples 3.34 (Common Maclaurin Series). These should be familiar from elementary calculus. Each function equals the given series form our previous discussions of power series: by the Theorem, each series is immediately the Maclaurin series of the given function.

$$
\begin{aligned}
e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} & x &\in \mathbb{R} & \frac{1}{1-x} &= \sum_{n=0}^{\infty} x^n & x &\in (-1,1) \\
\sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} & x &\in \mathbb{R} & \ln(1+x) &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n & x &\in (-1,1] \\
\cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} & x &\in \mathbb{R} & \tan^{-1} x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} & x &\in [-1,1]
\end{aligned}
$$

Examples 3.35 (Modifying Maclaurin Series). By substituting for x in a common series, we quickly obtain new series.

1. Substitute $x \mapsto 7x$ in the Maclaurin series for $\sin x$, to recover our earlier example
$$\sin 7x = \sum_{n=0}^{\infty} \frac{(-1)^n 7^{2n+1}}{(2n+1)!} x^{2n+1}, \quad x \in \mathbb{R}$$
Note how this requires almost no calculation: since the function equals a series, the Theorem says we have the Maclaurin series for $\sin 7x$!

2. Substitute $x \mapsto x^2$ in the Maclaurin series for e^x to obtain
$$e^{x^2} = \exp(x^2) = \sum_{n=0}^{\infty} \frac{1}{n!} x^{2n}, \quad x \in \mathbb{R}$$
This would be disgusting to verify directly, given the difficulty of repeatedly differentiating e^{x^2}.

3. We find the Taylor series for $f(x) = \frac{1}{5-x}$ centered at $x = 2$:
$$f(x) = \frac{1}{3+2-x} = \frac{1}{3(1-\frac{2-x}{3})} = \frac{1}{3} \sum_{n=0}^{\infty} \left(\frac{2-x}{3} \right)^n$$
which is valid whenever $-1 < \frac{2-x}{3} < 1 \iff -1 < x < 5$.

4. Fix $c \in \mathbb{R}$ and observe that, for all $x \in \mathbb{R}$,
$$e^x = e^{c+x-c} = e^c e^{x-c} = \sum_{n=0}^{\infty} \frac{e^c}{n!} (x-c)^n$$
We conclude that the series is the Taylor series of e^x centered at $x = c$. Of course this is easily verified using the definition, since $\frac{d^n}{dx^n}\Big|_{x=c} e^x = e^c$.

5. Combining the Theorem with the multiple-angle formula, we obtain the Taylor series for $\sin x$ centered at $x = c$:
$$\begin{aligned} \sin x &= \sin(c + x - c) = \sin c \cos(x-c) + \cos x \sin(x-c) \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n \sin c}{(2n)!} (x-c)^{2n} + \sum_{n=0}^{\infty} \frac{(-1)^n \cos c}{(2n+1)!} (x-c)^{2n+1} \end{aligned}$$

Definition 3.36. A function f is *analytic* on a its domain if every $c \in \operatorname{dom} f$ has a neighborhood on which $f(x)$ equals its Taylor series centered at c.

All the examples we've thus far seen are analytic on their domains; indeed the last two of Examples 3.35 *prove* this for the exponential and sine functions. Every analytic function is automatically smooth (infinitely differentiable), however the converse is *false* (Exercise 10). Analyticity is of greater importance in complex analysis where (amazingly!) it is equivalent to complex-differentiability.

Accuracy of Taylor Approximations

Our final goal is to estimate the accuracy of a Taylor polynomial as an approximation to its generating function. Otherwise said, we want to estimate the size of the remainder $R_n(x) = f(x) - p_n(x)$.

Theorem 3.37 (Taylor's Theorem: Lagrange's form). *Suppose f is $n+1$ times differentiable on an open interval I containing c and let $x \in I \setminus \{c\}$. Then there exists ξ between c and x for which the remainder centered at c satisfies*

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1}$$

Proof. For simplicity let $c = 0$. Fix $x \neq 0$, define a constant M_x and a function $g : I \to \mathbb{R}$ by

$$R_n(x) = \frac{M_x}{(n+1)!}x^{n+1} \quad \text{and} \quad g(t) = \frac{M_x}{(n+1)!}t^{n+1} + p_n(t) - f(t) = \frac{M_x}{(n+1)!}t^{n+1} - R_n(t)$$

Observe that

$$\begin{aligned} k \leq n+1 \implies g^{(k)}(x) &= \frac{M_x}{(n+1-k)!}t^{n+1-k} + p_n^{(k)}(t) - f^{(k)}(t) \qquad (*) \\ \implies g^{(k)}(0) &= p_n^{(k)}(0) - f^{(k)}(0) = 0 \quad \text{if } k \leq n \end{aligned}$$

where we invoked Theorem 3.33.
Now apply Rolle's Theorem repeatedly (WLOG assume $x > 0$):

- $\exists \xi_1$ between 0 and x such that $g'(\xi_1) = 0$.
- $\exists \xi_2$ between 0 and ξ_1 such that $g''(\xi_2) = 0$, etc.
- Iterate to obtain a sequence (ξ_k) such that

$$0 < \xi_{n+1} < \xi_n < \cdots < \xi_1 < x \quad \text{and} \quad g^{(k)}(\xi_k) = 0$$

Take $\xi = \xi_{n+1}$ and consider $(*)$: since $\deg p_n \leq n$, we see that

$$0 = g^{(n+1)}(\xi) = M_x - f^{(n+1)}(\xi) \implies R_n(x) = f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}$$

■

Corollary 3.38. *Suppose f is smooth on an open interval I containing c and that all derivatives $f^{(n)}$ of all orders are bounded on I. Then f equals its Taylor series (centered at c) on I.*

Proof. For simplicity, let $c = 0$. Suppose $\left|f^{(n+1)}(\xi)\right| \leq K$ for all $\xi \in I$. Choose any $N > |x|$ and observe that

$$n > N \implies |R_n(x)| \leq \frac{K\,|x|^{n+1}}{(n+1)!} = \frac{K\,|x|^{n+1}}{N!(N+1)\cdots(n+1)} \leq \frac{K\,|x|^N}{N!}\left(\frac{|x|}{N}\right)^{n+1-N} \xrightarrow[n\to\infty]{} 0$$

■

Examples 3.39. 1. The functions sine and cosine have derivatives bounded by 1 on $\mathbb{R}$, and thus both functions equal their Maclaurin series on $\mathbb{R}$. This removes the need to have previously justified these facts using the theory of differential equations.

2. The exponential function does not have bounded derivatives, however we can still apply Taylor's Theorem. For any fixed x, $\exists\xi$ between 0 and x such that

$$|R_n(x)| = \left|\frac{e^{\xi}}{(n+1)!}x^{n+1}\right| \xrightarrow[n\to\infty]{} 0$$

by the same argument in the Corollary. Thus e^x equals its Maclaurin series on the real line (we knew this already from Exercise 3.28.14).

3. Extending Example 3.32.3, we see that $h(x) = \sqrt{x}$ has linear approximation (1$^{\text{st}}$ Taylor polynomial) centered at $c = 9$

$$p_1(x) = h(9) + h'(9)(x-9) = 3 + \frac{1}{6}(x-9)$$

This yields the simple approximation

$$\sqrt{10} \approx p_1(10) = 3 + \frac{1}{6} = \frac{19}{6}$$

Taylor's Theorem can be used to estimate its accuracy (remember to shift the center to 9!):

$$R_1(10) = \frac{h''(\xi)}{2!}(10-9)^2 = -\frac{1}{2^2\cdot 2!}\xi^{-3/2} = -\frac{1}{8\xi^{3/2}} \quad \text{for some} \quad \xi \in (9,10)$$

Certainly $\xi^{-3/2} < 9^{-3/2} = \frac{1}{27}$, whence

$$-\frac{1}{216} < R_1(10) < 0 \implies \frac{19}{6} - \frac{1}{216} = \frac{683}{216} < \sqrt{10} < \frac{684}{216} = \frac{19}{6}$$

$\frac{19}{6}$ is therefore an overestimate for $\sqrt{10}$, but is accurate to within $\frac{1}{216} < 0.005$.

Alternative Versions of Taylor's Theorem

There are two further common expressions for the remainder in Taylor's Theorem. These are typically less easy to use than Lagrange's form but can sometimes provide sharper estimates for the remainder, particularly when x is far from the center of the series.

Corollary 3.40. *Suppose $f^{(n+1)}$ is continuous on an open interval I containing c, let $x \in I \setminus \{c\}$, and let $R_n(x) = f(x) - p_n(x)$ be the remainder for the Taylor polynomial centered at c. Then:*

1. *(Integral Remainder)* $R_n(x) = \displaystyle\int_c^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t$

2. *(Cauchy's Form)* $\exists\xi$ *between c and x such that* $R_n(x) = \dfrac{(x-\xi)^n}{n!}(x-c)f^{(n+1)}(\xi)$

Using these expressions it is possible to explicitly prove Newton's binomial series formula.

Corollary 3.41. *If* $\alpha \in \mathbb{R}$ *and* $|x| < 1$, *then*

$$\begin{aligned}(1+x)^\alpha &= 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n \\ &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \frac{\alpha(\alpha-1)(\alpha-2)(\alpha-3)}{4!}x^4 + \cdots\end{aligned}$$

If $\alpha \in \mathbb{N}_0$, this is the usual binomial theorem. Otherwise it is more interesting: for instance,

$$\begin{aligned}\sqrt{1+x} &= (1+x)^{1/2} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 + \cdots \\ \frac{1}{(1+x)^3} &= 1 - 3x + 6x^2 - 10x^3 + 15x^4 - \cdots\end{aligned}$$

Of course this last could easily be obtained from $\frac{1}{1+x} = \sum(-1)^n x^n$ by differentiating twice!

Exercises 3.31. *Key concepts: Taylor Series/Polynomials, Lagrange's form for Remainder*

1. Compute the Maclaurin series for $\cos x$ directly from the definition and use Taylor's Theorem to indicate why it converges to $\cos x$ for all $x \in \mathbb{R}$.
2. Repeat the previous exercise for $\sinh x = \frac{1}{2}(e^x - e^{-x})$ and $\cosh x = \frac{1}{2}(e^x + e^{-x})$.
3. Find the Maclaurin series for the function $\sin(3x^2)$. How do you know you are correct?
4. Find the Taylor series of $f(x) = x^4 - 3x^2 + 2x - 5$ at $x = 2$ and show that $T_2 f(x) = f(x)$.
5. Find a rational approximation to $\sqrt[3]{9}$ using the first Taylor polynomial for $f(x) = \sqrt[3]{x}$. Now use Taylor's Theorem to estimate its accuracy.
6. If $c \neq 1$, use the fact that $1 - x = (1-c)\left(1 - \frac{x-c}{1-c}\right)$ to obtain the Taylor series of $\frac{1}{1-x}$ centered at c. Hence conclude that $\frac{1}{1-x}$ is analytic on its domain $\mathbb{R} \setminus \{1\}$.
7. We prove that the Maclaurin series $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n$ converges to $\ln(1+x)$ whenever $0 < x \leq 1$.
   (a) Explicitly compute $\frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}} \ln(1+x)$.
   (b) Suppose $0 < x \leq 1$. Using Taylor's Theorem, prove that $\lim_{n\to\infty} R_n(x) = 0$.

   (*If* $-1 < x < 0$, *the argument is tougher, being similar to Exercise 11*)
8. Why can't we use Taylor's Theorem to approximate the error in $\frac{1}{1-x} = 1 + x + R_1(x)$ when $x \geq 1$? Try it when $x = 2$, what happens? What about when $x = -2$?
9. Prove Taylor's Theorem with integral remainder when $c = 0$ by using the following as an induction step: for each $n \in \mathbb{N}$, define
   $$A_n(x) = \int_0^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t$$
   and use integration by parts to prove that $A_{n+1} = A_n - \frac{x^{n+1}}{(n+1)!} f^{(n+1)}(0)$.

   (*The Cauchy form follows from the intermediate value theorem for integrals which we'll see later*)

10. Consider the function

$$f(x) = \begin{cases} e^{-1/x} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

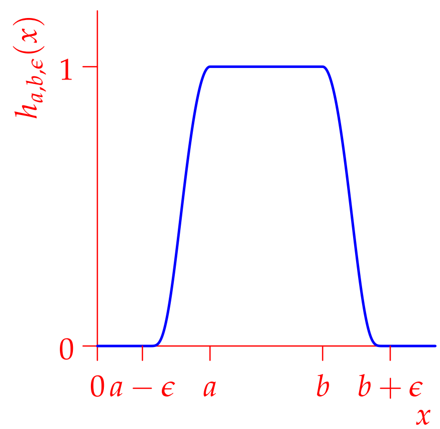
(a) Prove by induction that there exists a degree $2n$ polynomial q_n for which

$$f^{(n)}(x) = q_n\left(\tfrac{1}{x}\right) e^{-1/x} \text{ whenever } x > 0$$

(b) Prove that f is infinitely differentiable at $x = 0$ with $f^{(n)}(0) = 0$ (*use Exercise 3.30.3*).

The Maclaurin series of f is identically zero! Moreover, f is smooth (infinitely differentiable) on $\mathbb{R}$ but non-analytic at zero since it does not equal its Taylor series on any open interval containing zero.
A modification allows us to create bump functions, *which find wide use in analysis. If $a < b$, define*

$$g_{a,b} : x \mapsto f(x-a)f(b-x)$$



This is smooth on $\mathbb{R}$ but non-zero only on the interval (a, b). A further modification involving two such functions $g_{a,b}$ creates a smooth function on $\mathbb{R}$ which satisfies

$$h_{a,b,\epsilon}(x) = \begin{cases} 0 & \textit{if } x \le a-\epsilon \textit{ or } x \ge b+\epsilon \\ 1 & \textit{if } a \le x \le b \end{cases}$$

This 'switches on' rapidly from 0 to 1 near a and switches off similarly near b. By letting ϵ be small, we smoothly (but not uniformly) approximate the indicator function on $[a, b]$.

11. (Hard) We prove the binomial series formula (Corollary 3.41).

Let $f(x) = (1+x)^\alpha$ and $g(x) = 1 + \sum_{n=1}^{\infty} a_n x^n$ where $a_n = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}$. Our goal is to prove that $f = g$ on the interval $(-1, 1)$.

(a) Check that $f^{(n)}(0) = n!a_n$ so that g really is the Maclaurin series of f.

(b) i. Prove that the radius of convergence of g is 1.
ii. Prove that $\lim_{n\to\infty} na_n x^n = 0$ whenever $|x| < 1$.
iii. If $|x| < 1$ and ξ lies between 0 and x, prove that $\left|\frac{x-\xi}{1+\xi}\right| \le |x|$.
(*Hint: write $\xi = tx$ for some $t \in (0, 1) \ldots$*)

(c) Use Taylor's Theorem with Cauchy remainder to prove that

$$|R_n(x)| < (n+1)\,|a_{n+1}|\,|x|^{n+1}\,(1+\xi)^{\alpha-1}$$

Hence conclude that $g = f$ whenever $|x| < 1$.

(d) Here is an alternative argument for the full result:
i. Show that $(n+1)a_{n+1} + na_n = \alpha a_n$.
ii. Differentiate term-by-term to prove directly that g satisfies the differential equation $(1+x)g'(x) = \alpha g(x)$. Solve this to show that $g = f$ whenever $|x| < 1$.