# 9 Applications of the derivative

In this section, we discuss some applications of the derivative. All of these related, loosely, to getting information about the "shape" of a function (or more correctly about the shape of the graph of a function) from information about the derivative; but as we will see in examples, the applications go well beyond this limited scope.

## 9.1 Maximum and minimum points

At a very high level, this is what our intuition suggests: if a function $f$ is differentiable at $a$, then, since we interpret the derivative of $f$ at $a$ as being the slope of the tangent line to the graph of $f$ at the point $(a, f(a))$, we should have:

- if $f'(a) = 0$, the tangent line is horizontal, and at $a$ $f$ should have either a "local maximum" or a "local minimum";
- if $f'(a) > 0$, the tangent line has positive slope, and $f$ should be "locally increasing" near $a$; and
- if $f'(a) < 0$, the tangent line has negative slope, and $f$ should be "locally decreasing" near $a$.

This intuition is, unfortunately, *wrong*: for example, the function $f(x) = x^3$ has $f'(0) = 0$, but $f$ does not have a local maximum at $a$; in fact, it is increasing as we pass across $a = 0$.

More correctly, the intuition is only partly correct. What we do now is formalize some of the vague terms presented in quotes in the intuition above, and salvage it somewhat by presenting Fermat's principle.

Let $f$ be a function, and let $A$ be a subset of the domain of $f$.

**Definition of maximum point** Say that $x$ is a *maximum point* for $f$ on $A$ if

- $x \in A$ and
- $f(x) \geq f(y)$ for all $y \in A$.

In this case, say that $f(x)$ is *the*[121] *maximum value* of $f$ on $A$.

**Definition of minimum point** Say that $x$ is a *minimum point* for $f$ on $A$ if

- $x \in A$ and
- $f(x) \leq f(y)$ for all $y \in A$.

In this case, say that $f(x)$ is *the minimum value* of $f$ on $A$.

[121] "the": the maximum value is easily checked to be unique.

Maximum/minimum points are not certain to exist: consider $f(x) = x$, with $A = (0, 1)$; $f$ has neither a maximum point nor a minimum point on $A$. And if they exist, they are not certain to be unique: consider $f(x) = \sin x$ on $[0, 2\pi]$, which has maximum value 1 achieved at two maximum points, namely $\pi/2$ and $3\pi/2$, and minimum value 1 achieved at three minimum points, namely 0, $\pi$ and $2\pi$.

While having derivative equal to 0 doesn't ensure being at a max point, something is true in the converse direction: under certain conditions, being a maximum or minimum point, and being differentiable at the point, ensures that the derivative is 0. The specific conditions are that the function is defined on an *open interval*.

**Claim 9.1.** *(Fermat principle, part 1) Let $f : (a, b) \to \mathbb{R}$. If*

- *$x$ is a maximum point for $f$ on $(a, b)$, or a minimum point, and*
- *$f$ differentiable at $x$*

*then $f'(x) = 0$.*

Before giving the proof, some remarks are in order:

- As observed earlier via the example $f(x) = x^3$ at 0, the converse to Fermat principle is not valid: a function $f$ may be differentiable at a point, with zero derivative, but not have a maximum or minimum at that point.
- The claim becomes false if the function $f$ is considered on a *closed* interval $[a, b]$. For example, the function $f(x) = x$ on $[0, 1]$ has a maximum at 1 and a minimum at 0, is differentiable at both points[122], but at neither point in the derivative zero.[123]
- Fermat principle makes no assumptions about the function $f$ — it's not assumed to be differentiable everywhere, or even continuous. It's just a function.

**Proof**: Suppose $x$ is a maximum point, and that $f$ is differentiable at $x$. Consider the derivative of $f$ from below at $x$. We have, for $h < 0$,

$$\frac{f(x+h) - f(x)}{h} \geq 0$$

since $f(x+h) \leq f(x)$ ($x$ is a maximum point), so the ratio has non-positive numerator and negative denominator, so is positive. It follows that

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

[122] As differentiable as it can be ... differentiable from above at 0 and from below at 1.

[123] The state of Connecticut provides a real-world example: the highest point in the state is on a slope up to the summit of Mt. Frissell, whose peak is in Massachusetts.

Now consider the derivative of $f$ from above at $x$. We have, for $h > 0$,

$$\frac{f(x+h) - f(x)}{h} \leq 0$$

since $f(x+h) \leq f(x)$ still, and so the ratio has non-positive numerator and positive denominator, so is negative. It follows that

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

Since $f$ is differentiable at $x$ (by hypothesis), we have $f'(x) = f'_+(x) = f'_-(x)$, so $f'(x) \leq 0 \leq f'(x)$, making $f'(x) = 0$.

An almost identical argument works if $x$ is a minimum point. □

Fermat principle extends to "local" maxima and minima — points where a function has a maximum point or a minimum point, if the domain on which the function is viewed is made sufficient small around the point. Again let $f$ be a function, and let $A$ be a subset of the domain of $f$.

**Definition of local maximum point** Say that $x$ is a *local maximum point* for $f$ on $A$ if

- $x \in A$ and
- there is a $\delta > 0$ such that $f(x) \geq f(y)$ for all $y \in (x - \delta, x + \delta) \cap A$.

In this case, say that $f(x)$ is *a*[124] *local maximum value* of $f$ on $A$.

**Definition of local minimum point** Say that $x$ is a *local minimum point* for $f$ on $A$ if

- $x \in A$ and
- there is a $\delta > 0$ such that $f(x) \leq f(y)$ for all $y \in (x - \delta, x + \delta) \cap A$.

In this case, say that $f(x)$ is *a local minimum value* of $f$ on $A$.

Just like maximum/minimum points, local maximum/minimum points are not certain to exist: consider $f(x) = x$, with $A = (0, 1)$; $f$ has neither a local maximum point nor a local minimum point on $A$. And if they exist, they are not certain to be unique: consider $f(x) = 2x^2 - x^4$ defined on $[-2, 3]$. A look at the graph of this function shows that it has local maxima at both $-1$ and $1$ (both taking value $1$, although of course it isn't necessarily the case that multiple local maxima have to share the same value, in general[125]). It also has a local minima at $-2$, $0$ and $3$, with values $-8$, $0$ and $-63$. Notice that there are local minima at the endpoints of the interval, even though if the interval was extended slightly neither would be a local minimum. This is because the definition of $x$ being a local minimum

[124] "a": a local maximum value is clearly not necessarily unique; see examples below.

[125] Physically, a local maximum is the summit of a mountain, and of course different mountains in general have different heights.

of a set $A$ specifies that we should compare the function at $x$ to the function at all $y$ nearby to $x$ *that are also in* $A$.

There is an analog of the Fermat principle for local maxima and minima.

**Claim 9.2.** *(Fermat principle, part 2) Let* $f:(a,b)\to\mathbb{R}$*. If*

- $x$ *is a local maximum point for* $f$ *on* $(a,b)$*, or a local minimum point, and*
- $f$ *differentiable at* $x$

*then* $f'(x)=0$.

We do not present the proof here; it is in fact just a corollary of Claim 9.1. Indeed, if $x$ is a local maximum point for $f$ on $(a,b)$, then from the definition of local maximum and from the fact that the interval $(a,b)$ is open at both ends, it follows that there is some $\delta>0$ small enough that $(x-\delta,x+\delta)$ is completely contained in $(a,b)$, *and* that $x$ is a *maximum point* (as opposed to local maximum point) for $f$ on $(x-\delta,x+\delta)$; then if $f$ is differentiable at $x$ with derivative zero, Claim 9.1 shows that $f'(x)=0$.[126]

As with Claim 9.1, Claim 9.2 fails if $f$ is defined on a *closed* interval $[a,b]$, as the example $f(x)=2x^2-x^4$ on $[-2,3]$ discussed above shows.[127]

Fermat principle leads to an important definition.

**Definition of a critical point** $x$ is a *critical point* for a function $f$ if $f$ is differentiable at $x$, and if $f'(x)=0$.[128] The value $f(x)$ is then said to be a *critical value* of $f$.

Here's the point of critical points. Consider $f:[a,b]\to\mathbb{R}$. Where could a maximum point or a minimum point of $f$ be? Well, maybe at $a$ or $b$. If not at $a$ or $b$, then somewhere in $(a,b)$. And by Fermat principle, the only possibilities for a maximum point or a minimum point in $(a,b)$ are those points where $f$ is not differentiable or (and this is where Fermat principle comes in) where $f$ is differential and has derivative equal to 0; i.e., the critical point of $f$.

The last paragraph gives a proof of the following.

**Theorem 9.3.** *Suppose* $f:[a,b]\to\mathbb{R}$*. If a maximum point, or a minimum point, of* $f$ *exists (on* $[a,b]$*), then* $x$ *must be one of*

- $a$ *or* $b$
- *a critical point in* $(a,b)$ *or*
- *a point of non-differentiability in* $(a,b)$*.*

[126] I wrote "We do not present the proof here"; but then it seems I went and gave the proof.

[127] And as does Connecticut: the south slope of Mt. Frissell, crossing into Massachusetts, is a local maximum high point of Connecticut, but not a point with derivative zero. The highest *peak* in Connecticut, the highest point with derivative zero, is the summit of Mt. Bear, a little south of Mt. Frissell.

[128] Many authors also say that $x$ is a critical point for $f$ if $f$ is not differentiable at $x$.

*In particular, if $f$ is continuous on $[a,b]$ (and so, a maximum point and a minimum point exists, by the Extreme Value Theorem), then to locate a maximum point and/or a minimum point of $f$ it suffices to consider the values of $f$ at $a, b$, the critical points of $f$ in $(a,b)$ and the points of non-differentiability of $f$ in $(a,b)$.*

Often this theorem reduces the task of finding the maximum or minimum value of a function on an interval (*a priori* a task that involves checking infinitely many values) to that of find the maximum or minimum of a finite set of values. We give three examples:

- $f : [-1,1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1/3 & \text{if } x = -1 \text{ or } x = 1 \\ 1/2 & \text{if } x = 0 \\ |x| & \text{otherwise.} \end{cases}$$

  Here the endpoints of the closed interval on which the function is defined are $-1$ and 1. We have $f(-1) = f(1) = 1/3$. There are no critical points in $(-1,1)$, because where the function is differentiable (on $(-1,0)$ and $(0,1)$) the derivative is never 0. There is one point of non-differentiability in $(-1,1)$, namely the point 0, and $f(0) = 1/2$. It might seem that the theorem tells us that the maximum value of $f$ on $[-1,1]$ is 1/2 and the minimum value is 1/3. But this is clearly wrong, on both sides: $f(3/4) = 3/4 > 1/2$, for example, and $f(-1/4) = 1/4 < 1/3$. The issue is that the function $f$ *has no maximum* on $[-1,1]$ (it's not hard to check that $\sup\{f(x) : x \in [-1,1]\} = 1$ and $\inf\{f(x) : x \in [-1,1]\} = 0$, but that there are no $x$'s in $[-1,1]$ with $f(x) = 1$ or with $f(x) = -1$), and so the hypotheses of the theorem are not satisfied.

- $f : [0,4] \to \mathbb{R}$ defined by

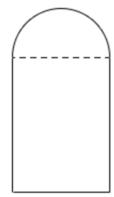
$$f(x) = \frac{1}{x^2 - 4x + 3}.$$

  Here the endpoints of the closed interval on which the function is defined are 0 and 4. We have $f(0) = f(4) = 1/3$. To find the critical points in $(0,4)$, we differentiate $f$ and set the derivative equal to 0:

$$f'(x) = \frac{-(2x-4)}{(x^2 - 4x + 3)^2} = 0 \text{ when } 2x - 4 = 0, \text{ or } x = 2.$$

  So there is one critical point (at 2), and $f(2) = -1$. It might seem that the theorem tells us that the maximum value of $f$ on $[0,4]$ is 1/3 and the minimum value is $-1$. But a quick look at the graph of the functions shows that this is quite wrong; the function takes arbitrarily large and arbitrarily small values on $[1,4]$, in particular near to 1 and near to 2. What went wrong was that, as with the last example, we did not verify the hypotheses of the theorem. The function $f$ may be written as $f(x) = 1/((x-1)((x-3)))$, and so is not defined at either 1 or 3, rendering the starting statement "$f : [0,4] \to \mathbb{R}$" meaningless. not satisfied.

- (A genuine example, a canonical example of a "calculus optimization problem") A piece of wire of length $L$ is to be bent into the shape of a Roman window — a rectangle below with a semicircle on top (see the figure below).



What is the maximum achievable area that can be enclosed by the wire with this shape?

We start by introducing names for the various variables of the problem. There are two reasonable variables: $x$, the base of the rectangle, and $y$, the height (these two values determine the entire shape). There is a relationship between these two numbers, namely $x + 2y + \pi(x/2) = L$ (the window has a base that is a straight line of length $x$, two vertical straight line sides of length $y$ each, and a semicircular cap of radius $x/2$, so length $\pi(x/2)$). The total area enclosed may be expressed as $A = xy + \pi x^2/8$ (the area of the rectangular base, plus the area of the semicircular cap). We use $x(1 + \pi/2) + 2y = L$ to express $y$ in terms of $x$: $y = (L - (1 + \pi/2)x)/2$, so that the area $A$ becomes a function $A(x)$ of $x$, namely

$$A(x) = (x/2)(L - (1 + \pi/2)x) + \pi x^2/8.$$

Clearly the smallest value of $x$ that we need consider is 0. The largest value is the one corresponding to $y = 0$, so $x = L/(1 + \pi/2)$. Therefore we are considering the problem of finding the maximum value of a continuous function $A$ on the closed interval $[0, L/(1 + \pi/2)]$. Because $A$ is continuous, we know that the maximum value exists. Because $A$ is everywhere differentiable, the theorem tells us that we need only consider $A$ at 0, $L/(1 + \pi/2)$, and any point between the two where $A'(x) = 0$. There is one such point, at $L/(2 + \pi/2)$.

We have

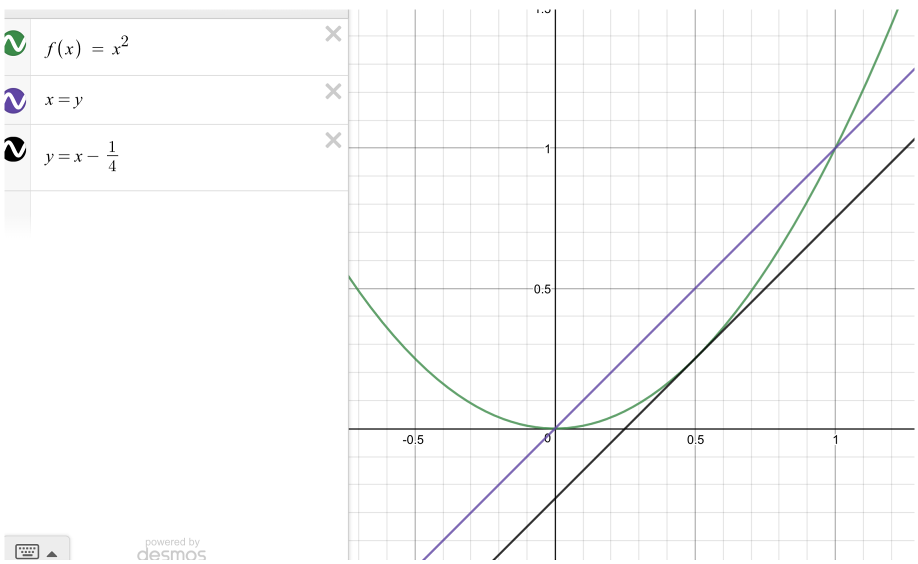
  - $A(0) = 0$,
  - $A(L/(2 + \pi/2)) = \frac{L^2}{2(2+\pi/2)^2} + \frac{\pi L^2}{8(2+\pi/2)^2}$ and
  - $A(L/(1 + \pi/2)) = \frac{\pi L^2}{8(1+\pi/2)^2}$.

The second of these is the largest, so is the largest achievable area.

## 9.2 The mean value theorem

To go any further with the study of the derivative, we need a tool that is to differentiation as the Intermediate and Extreme Value Theorems are to continuity. That tool is called the Mean Value Theorem (MVT). The MVT says, mathematically, that if a function is differentiable on an interval, then at some point between the start and end point of the interval, the slope of the tangent line to the function should equal the average slope over the whole interval, that is, the slope of the secant line joining the initial point of the interval to the terminal point. Informally, it says that if you travel from South Bend to Chicago averaging 60 miles per hour, then at some point on the journey you must have been traveling at exactly 60 miles per hour.

By drawing a graph of a generic differentiable function, it is fairly evident that the MVT *must* be true. The picture below shows the graph of $f(x) = x^2$. Between 0 and 1, the secant line is $y = x$, with slope 1, and indeed there is a number between 0 and 1 at which the slope of the tangent line to $f$ is 1, i.e., at which the derivative is 1, namely at $1/2$.




However, we need to be careful. If we choose to do our mathematics only in the world of rational numbers, then the notions of limits, continuity and differentiability make perfect sense; and just as it was possible to come up with examples of continuous functions in this "$\mathbb{Q}$-world" that satisfy the hypotheses of IVT and EVT, but do not satisfy their conclusions, it is also possible to come up with an example of a function on a closed interval that is differentiable in the $\mathbb{Q}$-world, but for which there is no point in the interval where the derivative is equal the slope of the secant line connecting the endpoints of the interval.[129] This says that to prove the MVT, the completeness axiom will be needed. But in fact we'll bypass completeness, and prove MVT using EVT (which itself required completeness).

**Theorem 9.4.** *(Mean value theorem) Suppose* $f : [a, b] \to \mathbb{R}$ *is*

- *continuous on* $[a, b]$, *and*

[129] Find one! (It will be on the homework $\cdots$).

- *differentiable on* $(a, b)$.

*Then there is* $c \in (a, b)$ *with*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

**Proof**: We begin with the special case $f(a) = f(b)$. In this case, we require $c \in (a, b)$ with $f'(c) = 0$.[130]

By the Extreme Value Theorem, $f$ has a maximum point and a minimum point in $[a, b]$. If there is a maximum point $c \in (a, b)$, then by Fermat principle, $f'(c) = 0$. If there is a minimum point $c \in (a, b)$, then by Fermat principle, $f'(c) = 0$. If neither of these things happen, then the maximum point and the minimum point must both occur at one (or both) of $a$ and $b$. In this case, both the maximum and the minimum of $f$ on $[a, b]$ are 0, so $f$ is constant on $[a, b]$, and so $f'(c) = 0$ for *all* $c \in (a, b)$.

We now reduce the general remaining case, $f(a) \neq f(b)$, to the case just considered. Set

$$L(x) = f(a) + (x - a)\frac{f(b) - f(a)}{b - a};$$

notice that the graph of this function is the line that passes through $(a, f(a))$ and $(b, f(b))$. Now let $h(x)$ be the (vertical) distance from the point $(x, f(x))$ to the point $(x, L(x))$, so

$$h(x) = f(x) - f(a) - \left(\frac{f(b) - f(a)}{b - a}\right)(x - a).$$

We have $h(a) = h(b) = 0$, and $h$ is continuous on $[a, b]$, and differentiable on $(a, b)$. So by the previous case, there is $c \in (a, b)$ with $h'(c) = 0$. But

$$h'(x) = f'(x) - \left(\frac{f(b) - f(a)}{b - a}\right),$$

so $f'(c) = \frac{f(b)-f(a)}{b-a}$. □

Note that both Rolle's theorem and the MVT fail if $f$ is not assumed to be differentiable on the whole of the interval $(a, b)$: consider the function $f(x) = |x|$ on $[-1, 1]$.

In the proof of the MVT, we used the fact that if $f : (a, b) \to \mathbb{R}$ is constant, then it is differentiable at all points, with derivative 0. What about the converse of this? If $f : (a, b) \to \mathbb{R}$ is differentiable at all points, with derivative 0, can we conclude that $f$ is constant? This seems a "fact" so obvious that it barely requires a proof: physically, it is asserting that if a particle has 0 velocity at all times, then it must always be located in the same position.

[130] This special case is often referred to as *Rolle's theorem*. It is traditional to make fun of Rolle's theorem; see e.g. this XKCD cartoon: https://xkcd.com/2042/. Before dismissing Rolle's theorem as a triviality, though, remember this: in $\mathbb{Q}$-world, it is false, and so its proof requires the high-level machinery of the completeness axiom.

But of course, it is not[131] be obvious. Indeed, if true, it must be a corollary of the completeness axiom, because in $\mathbb{Q}$-world, the function $f:[0,2]\to\mathbb{Q}$ given by

$$f(x)=\begin{cases} 0 & \text{if } x^2<2 \\ 1 & \text{if } x^2>2 \end{cases} cc$$

is continuous on $[0,2]$, differentiable on $(0,2)$, has derivative 0 everywhere, but certainly is not constant.

We will establish this converse, not directly from completeness, but from MVT.

**Claim 9.5.** *If $f:(a,b)\to\mathbb{R}$ is differentiable at all points, with derivative $0$, then $f$ is constant.*

**Proof**: Suppose that $f$ satisfies the hypotheses of the claim, but is not constant. Then there are $a<x_0<x_1<b$ with $f(x_0)\neq f(x_1)$. But then, applying MVT on the interval $[x_0,x_1]$, we find $c\in(x_0,x_1)\subseteq(a,b)$ with

$$f'(c)=\frac{f(x_1)-f(x_0)}{x_1-x_0}\neq 0,$$

a contradiction. We conclude that $f$ is constant on $(a,b)$. □

**Corollary 9.6.** *If $f,g:(a,b)\to\mathbb{R}$ are both differentiable at all points, with $f'=g'$ on all of $(a,b)$, then there is a constant such that $f$ and $g$ differ by that (same) constant at every point in $(a,b)$ (i.e., there's $c$ with $f(x)=g(x)+c$ for all $x\in(a,b)$).*

**Proof**: Apply Claim 9.5 on the function $f-g$. □

Our next application of MVT concerns the notions of a function increasing/decreasing on an interval. Throughout this definition, $I$ is some interval (maybe an open interval, like $(a,b)$ or $(a,\infty)$, or $(-\infty,b)$ or $(-\infty,\infty)$, or maybe a closed interval, like $[a,b]$, or maybe a mixture, like $(a,b]$ or $[a,b)$ or $(-\infty,b]$ or $[a,\infty)$).

**Definition of a function increasing** Say that $f$ is *increasing* on $I$, or *strictly increasing*[132], if whenever $a<b$ in $I$, $f(a)<f(b)$. Say that $f$ is *weakly increasing* on $I$ if whenever $a<b$ in $I$, $f(a)\leq f(b)$.

[131] at least, should not

[132] There is a truly annoying notational issue here. To some people, "increasing" means just what it has been defined to mean here, namely that as the input to the function increases, the output of the function genuinely increases, too. In this interpretation, the constant function is *not* increasing (it's *weakly* increasing). To other people, "increasing" means that as the input to the function increases, the output of the function either increases or stays the same. In this interpretation, the constant function *is* increasing. There is no resolution to this ambiguity, as both usages are firmly established in mathematics. So you have to be *very* careful, when someone talks about increasing/decreasing, that you know which interpretation they mean.

**Definition of a function decreasing** Say that $f$ is *decreasing* on $I$, or *strictly decreasing*, if whenever $a < b$ in $I$, $f(a) > f(b)$. Say that $f$ is *weakly decreasing* on $I$ if whenever $a < b$ in $I$, $f(a) \geq f(b)$.

**Definition of a function being monotone** Say that $f$ is *monotone* on $I$, or *strictly monotone*, if it is either increasing on $I$ or decreasing on $I$. Say that $f$ is *weakly monotone* on $I$ if it is either weakly increasing on $I$ or weakly decreasing on $I$.

**Claim 9.7.** *If $f'(x) > 0$ for all $x$ in some interval $I$, then $f$ is strictly increasing on $I$. If $f'(x) < 0$ for all $x \in I$, then $f$ is strictly decreasing on $I$.*

*If $f'(x) \geq 0$ for all $x$ in some interval $I$, then $f$ is weakly increasing on $I$. If $f'(x) \leq 0$ for all $x \in I$, then $f$ is weakly decreasing on $I$.*

**Proof**: Suppose $f'(x) > 0$ for all $x \in I$. Fix $a < b$ in $I$. By the MVT, there's $c \in (a, b)$ with

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

By hypothesis, $f'(c) > 0$, so $f(b) > f(a)$, proving that $f$ is strictly increasing on $I$.

All the other parts of the claim are proved similarly. □

The converse of this claim is not (entirely) true: if $f$ is strictly increasing on an interval, and differentiable on the whole interval, then it is not necessarily the case that $f'(x) > 0$ on the interval. The standard example here is $f(x) = x^3$, defined on the whole of the real line; it's strictly increasing, differentiable everywhere, but $f'(0) = 0$. On the other hand, we do have the following converse, which doesn't require MVT; it just comes from the definition of the derivative (similar to the proof of the Fermat principle). The proof is left as an exercise.

**Claim 9.8.** *If $f$ is weakly increasing on an interval, and differentiable on the whole interval, then $f'(x) \geq 0$ on the interval.*[133]

Now that we have established a way of identifying intervals on which a function is increasing and/or decreasing, we can develop some effective tools for identifying where functions have local minima/local maxima. The first of these gives a partial converse to the Fermat principle. Recall that Fermat principle says that if $f$ is defined on $(a, b)$, with $f$ differentiable at some $x \in (a, b)$, then if $f'(x) = 0$ $x$ *might* be a local minimum or local maximum; while if $f'(x) \neq 0$, $f$ cannot possibly be a local minimum or local maximum. This next claim gives some conditions under which we can say that $x$ *is* a local minimum or local maximum, when its derivative is 0.

**Claim 9.9.** *(First derivative test) Suppose $f$ is defined on $(a, b)$, and that $f$ is differentiable at $x \in (a, b)$, with $f'(x) = 0$. Suppose further that $f$ is differentiable near $x$*[134]*. If $f'(y) < 0$*

[133] And, since strictly increasing implies weakly increasing, it follows that if $f$ is strictly increasing on an interval, and differentiable on the whole interval, then $f'(x) \geq 0$ on the interval.

[134] Recall that "near $x$" means: in some interval $(x - \delta, x + \delta)$, $\delta > 0$.

*for all $y$ in some small interval to the left of $x$, and $f'(y) > 0$ in some small interval to right of $x$, then $x$ is a local minimum for $f$ on $(a,b)$; in fact, $x$ is a* strict *local minimum, meaning $f(x) < f(y)$ for all $y$ close to $x$. If, on the other hand, $f'(y) > 0$ for all $y$ in some small interval to the left of $x$, and $f'(y) < 0$ in some small interval to right of $x$, then $x$ is a strict local maximum for $f$ on $(a,b)$.*

**Proof**: We consider only the case where $x$ is claimed to be a strict local minimum (the other is very similar). We have that on some small interval $(x-\delta, x]$, $f$ has non-positive derivative (positive on $(x-\delta, x)$ and 0 at $x$), so, by Claim 9.7, $f$ is weakly decreasing on this interval. By the same token, $f$ is weakly increasing on $[x, x+\delta)$. This immediately says that $x$ is a local minimum point for $f$ on $(a,b)$.

To get the strictness: $f$ is strictly decreasing on $(x-\delta, x))$. For any $y$ in this interval, pick any $y'$ with $y < y' < x$. We have $f(y) > f(y')$ (because $f$ is strictly decreasing between $y$ and $y'$), and $f(y') \geq f(x)$ (because $f$ is weakly decreasing between $y'$ and $x$), so $f(y) > f(x)$; and by the same token $f(x) < f(y)$ for all $y$ in a small interval to the right of $x$. □

**Claim 9.10.** *(Second derivative test) Suppose $f$ is defined on $(a,b)$, and that $f$ is* twice *differentiable at $x \in (a,b)$, with $f'(x) = 0$.*

- *If $f''(x) > 0$, then $a$ is a (strict) local minimum for $f$ on $(a,b)$.*
- *If $f''(x) < 0$, then $a$ is a (strict) local maximum for $f$ on $(a,b)$.*
- *If $f''(x) = 0$ then anything can happen.*

**Proof**: We first consider the case where $f''(x) > 0$. We have

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

The denominator in the fraction at the end is negative. For the limit to be positive, the numerator must be negative for all sufficiently small (close to 0 and negative) $h$; in other words, $f'(y)$ must be negative on some small interval to the left of $x$. By a similar argument, $f'(y)$ must be positive on some small interval to the right of $x$. By Claim 9.9, $f$ has a strict local minimum on $(a,b)$ at $x$.

The case $f''(x) < 0$ is similar. To show that no conclusion can be reached when $f''(x) = 0$, consider the functions $f(x) = x^3$, $g(x) = x^4$ and $h(x) = -x^4$ at $x = 0$. In all three cases the functions have derivative 0 at 0, and second derivative 0 at 0. For $f$, 0 is neither a local maximum nor a local minimum point. For $g$, 0 is a local minimum. For $h$, 0 is a local maximum. □

## 9.3 Curve sketching

How do you get a good idea of the general appearance of the graph of a "reasonable" function (one which is continuous and differentiable at "most" points)? An obvious strategy is to use a graphing tool (such as `Desmos.com` or `WolframAlpha.com`). Here we'll describe a "by-hand" approach, that mostly utilizes information gleaned from the derivative. With powerful graphing tools available, this might seem pointless; but it's not. Here are two reasons why we might want to study curve sketching from first principles.

- It's a good exercise in reviewing the properties of the derivative, before applying them in situations where graphing tools may not be as helpful, and
- sometimes, graphing tools get things *very* wrong[135], and it's helpful to be able to do things by hand yourself, so that you can trouble-shoot when this happens.

The basic strategy that is often employed to sketch graphs of "reasonable" functions is as follows.

**Step 1** Identify the domain of the function. Express it as a union of intervals.

**Step 2** Identify the limiting behavior of the function at any *open* endpoints of intervals in the domain; this will usually involve one sided limits and/or limits at infinity, as well as possible infinite limits).

**Step 3** Find the derivative of the function, and identify critical points (where the derivative is 0), intervals where the derivative is positive (and so the function is increasing), and intervals where the derivative is negative (and so the function is decreasing).

**Step 4** Use the first derivative test to identify local maxima and minima.

**Step 5** Plot some obvious points (such as intercepts of axes, local minima and maxima, and points where the derivative does not exist).

**Step 6** Interpolate the graph between all these plotted points, in a manner consistent with the information obtained from the first four points.

There is also a zeroth step: check if the function is even, or is odd. This typically halves the work involved in curve sketching: if the function is even, then the graph is symmetric around the $y$-axis, and if it is odd, then the portion of the graph corresponding to negative $x$ is obtained from the portion corresponding to positive $x$ by reflection through the origin.

Our first example is $f(x) = x^3 + 3x^2 - 9x + 12$, which is neither even nor odd.

**Step 1** The domain of $f$ is all reals, or $(-\infty, \infty)$.

**Step 2** $\lim_{x\to\infty} x^3 + 3x^2 - 9x + 12 = \infty$ and $\lim_{x\to-\infty} x^3 + 3x^2 - 9x + 12 = -\infty$.

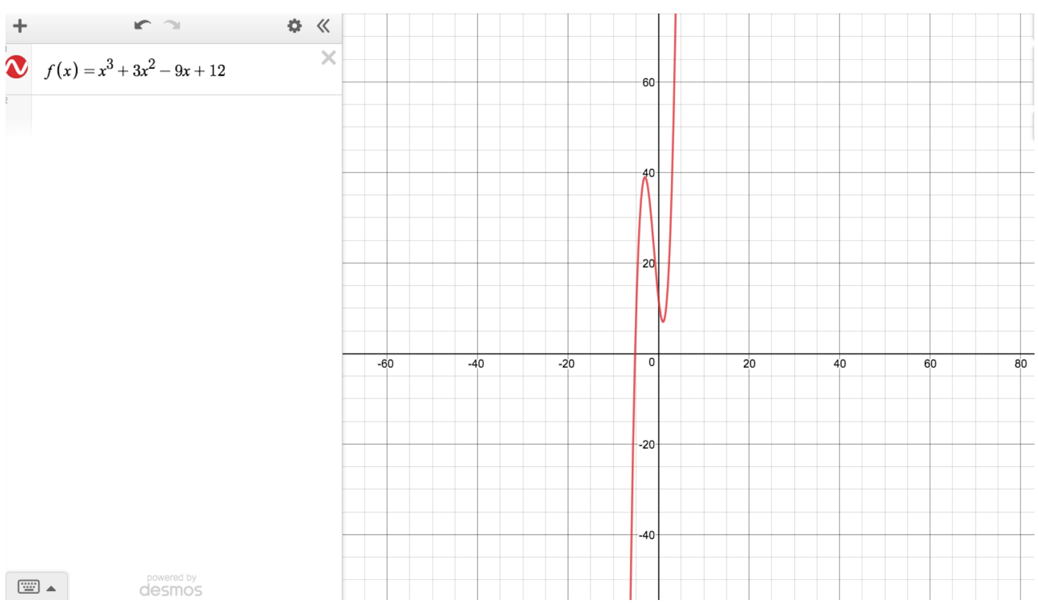
[135] Ask Desmos to graph the function $f(x) = [x \cdot (1/x)]$.

**Step 3** $f'(x) = 3x^2 + 6x - 9$. This is defined, and continuous, on all of $\mathbb{R}$, so to find intervals where it is positive or negative, it is enough to find where it is 0 — $3x^2 + 6x - 9 = 0$ is the same as $x^2 + 2x - 3 = 0$ or $x = (-2 \pm \sqrt{4 + 12})/2 = 1$ or $-3$. Removing these two numbers from $\mathbb{R}$ leaves intervals $(-\infty, -3)$, $(-3, 1)$ and $(1, \infty)$. By the IVT, on each of these intervals $f'$ must be either always positive or always negative (if $f'$ is both positive and negative on any of the intervals then by continuity of $f'$, $f'$ must be 0 somewhere on that interval, but it can't be since we have removed to points where $f'$ is 0). So we need to just test *one* point in each of $(-\infty, -3)$, $(-3, 1)$ and $(1, \infty)$, to determine the sign of $f'$ on the entire interval. Since $f'(-100) > 0$, $f'(0) < 0$ and $f'(100) > 0$, we find that $f$ is increasing on $(-\infty, -3)$, decreasing on $(-3, 1)$, and increasing on $(1, \infty)$.

**Step 4** By the first derivative test, there is a local maximum at $x = -3$ (to the left of $-3$ the derivative is positive, to the right it is negative, at $-3$ it is 0), a local minimum at $x = 1$, and no other local extrema.

**Step 5** At $x = 0$, $f(x) = 12$, so $(0, 12)$ is on the graph. The local maximum at $x = -3$ is the point $(-3, 39)$, and the local minimum at $x = 1$ is the point $(1, 7)$. The equation $f(x) = 0$ isn't obviously easy to solve, so we don't try to calculate any point at which the graph crosses the $x$-axis.

**Step 6** We are required to plot a curve that's defined on all reals. As we move from $-\infty$ in the positive direction, the curve increases from $-\infty$ until it reaches a local maximum at $(-3, 39)$. Then it drops to a local minimum at $(1, 7)$, passing through $(0, 12)$ along the way. From the local minimum at $(1, 7)$ it increases to $+\infty$ at $+\infty$. This is a verbal description of the graph; here's what is looks like visually, according to Desmos:




With what we know so far, we couldn't have sketched such an accurate graph; we know, for example, that $f$ decreases from $-3$ to 1, but how do we know that it decreases in the manner that it does (notice how it "bulges": between $-3$ and 1, for a while the graph is lying to the right of the straight line joining $(-3, 39)$ to $(1, 7)$, and then it moves to being on the left)? To get this kind of fine detail, we need to study the *second* derivative, and specifically the topic of *convexity*; that will come in a later section.

As a second example, consider $f(x) = x^2/(1 - x^2)$. This is an even function — $f(-x) = f(x)$ for all $x$ — so we only consider it on the interval $[0, \infty)$.

**Step 1** The domain of the function (with our attention restricted to $[0, \infty)$) is all non-negative numbers except $x = 1$, that is, $[0, 1) \cup (1, \infty)$.

**Step 2** We have

$$\lim_{x \to 0^-} \frac{x^2}{1 - x^2} = +\infty,$$

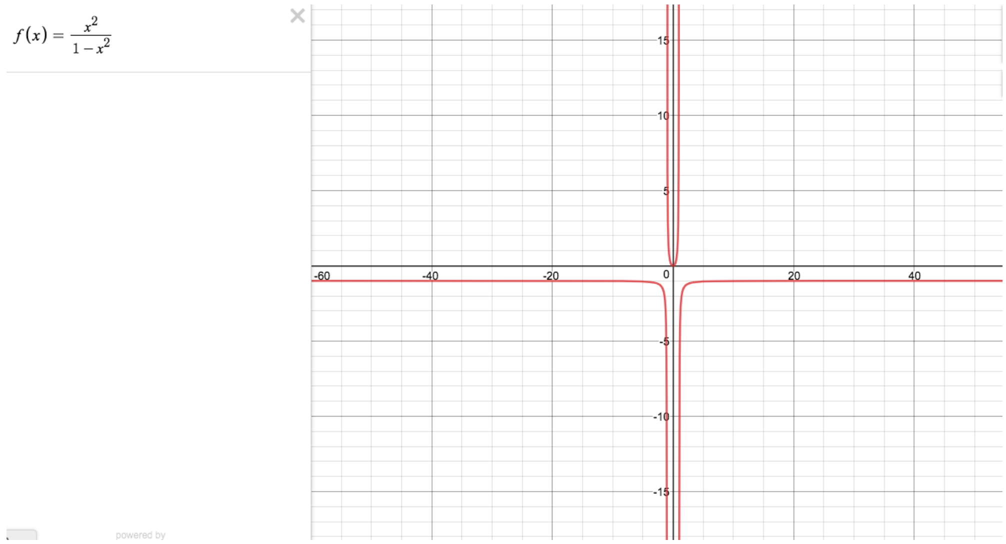
$$\lim_{x \to 0^+} \frac{x^2}{1 - x^2} = -\infty$$

and

$$\lim_{x \to \infty} \frac{x^2}{1 - x^2} = -1.$$

**Steps 3, 4, 5** We have $f'(x) = 2x/(1 - x^2)^2$, and the domain of $f'$ is the same at that of $f$: $[0, 1) \cup (1, \infty)$. The derivative is only equal to 0 at 0; at all other points it is positive. We conclude that $f$ is strictly increasing on $(0, 1)$ and on $(1, \infty)$, and it is weakly increasing on $[0, 1)$. The graph passes through the point $(0, 0)$, and it does not seem like there are any other obviously easy-to-identify points.

**Step 6** Moving from 0 to infinity: the graph starts at $(0, 0)$, and increases to infinity as $x$ approaches 1 (the line $x = 1$ is referred to as a *vertical asymptote* of the graph). To the right of 1, it (strictly) increases from $-\infty$ to $-1$ as $x$ moves from (just to the right of) 1 to ("just to the left of") $\infty$. (The line $y = -1$, that the graph approaches near infinity but doesn't reach, is referred to as a *horizontal asymptote* of the graph). To the left of the origin, the graph is the mirror image (the mirror being the $y$-axis) of what we have just described. Here is Desmos' rendering (for clarity, the aspect ratio has been changed from 1 : 1):




## 9.4 L'Hôpital's rule

What is $\lim_{x\to 1} \frac{x^2-1}{x^3-1}$? The function $f(x) = (x^2-1)/(x^3-1)$ is not continuous at 1 (it is not even defined at 1) so we cannot asses the limit by a direct evaluation. We can figure out the limit, via a little bit of algebraic manipulation, however: away from 1

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

Using our usual theorems about limits, we easily have $\lim_{x\to 1} \frac{x+1}{x^2+x+1} = 2/3$ (the function $g(x) = (x+1)/(x^2+x+1)$ is continuous at 1, with $g(1) = 2/3$, and $g$ agrees with $f$ at all reals other than 1).

We have calculated many such awkward limits using this kind of algebraic trickery. A common feature to many of these limits, is that the expression we are working with is a ratio, where both the numerator and denominator approach 0 near the input being approached in the limit calculation; this leads to the meaningless expression "0/0" when we attempt a "direct evaluation" of the limit as 0/0[136]. Using the derivative, there is a systematic way of approaching all limits of this kind, called *L'Hôpital's rule*.

Suppose that we want to calculate $\lim_{x\to a} f(x)/g(x)$, but a direct evaluation is impossible because $f(a) = g(a) = 0$. We can approximate both the numerator and the denominator of the expression, using the linearization. The linearization of $f$ near $a$ is $L_f(x) = f(a) +$

[136] A meaningless expression, that can take on any possible value, or no value. Consider the following examples:

- $\lim_{x\to 0} \frac{cx}{x} = c$, $c$ any real number;
- $\lim_{x\to 0} \frac{\pm x^2}{x} = \pm\infty$; and
- $\lim_{x\to 0} \frac{x\sin(1/x)}{x}$, which does not exist.

$f'(a)(x-a) = f'(a)(x-a)$, and the linearization of $g$ near $a$ is $L_g(x) = g(a) + g'(a)(x-a) = g'(a)(x-a)$.[137] Assuming that the linearization is a good approximation to the function it's linearizing, especially near the point of interest $a$, we get that near (but not at) $a$,

$$\frac{f(x)}{g(x)} \approx \frac{L_f(a)}{L_g(a)} = \frac{f'(a)(x-a)}{g'(a)(x-a)} = \frac{f'(a)}{g'(a)}$$ [138]

This strongly suggests that if $f, g$ are both differentiable at $a$, with $g'(a) \neq 0$ (and with $f(a) = g(a) = 0$), then

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

For example, with $f(x) = x^2 - 1$, $g(x) = x^3 - 1$, $a = 1$, so $f(a) = g(a) = 0$, $f'(x) = 2x$, $g'(x) = 3x^2$, so $f'(a) = 2$ and $g'(a) = 3$, we have

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

Before doing some examples, we try to formalize the linearization proof described above; along the way we keep track of all the various hypotheses we need to make on $f$ and $g$.

So, suppose $f(a) = g(a) = 0$. We have, if all the various limits exist,

$$\begin{aligned}
\lim_{x\to a} \frac{f(x)}{g(x)} &= \lim_{x\to a} \frac{f(x)-f(a)}{g(x)-g(a)} \quad (f(a) = g(a) = 0) \\
&= \lim_{x\to a} \frac{\frac{f(x)-f(a)}{x-a}}{\frac{g(x)-g(a)}{x-a}} \\
&= \frac{\lim_{x\to a} \frac{f(x)-f(a)}{x-a}}{\lim_{x\to a} \frac{g(x)-g(a)}{x-a}} \quad \text{(adding assumption here: bottom limit is non-zero)} \\
&= \frac{f'(a)}{g'(a)}.
\end{aligned}$$

Going backwards through this chain of equalities yields a proof of the following result, what turns out to be a fairly weak form of what we will ultimately call L'Hôpital's rule.

**Claim 9.11.** *Suppose that $f$ and $g$ are both differentiable at $a$ (so, in particular, defined in some small neighborhood around $a$, and also continuous at $a$), and that $g'(a) \neq 0$. If $f(a) = g(a) = 0$, then*

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

Here are a few examples.

[137] We're making the assumption here that $f, g$ are both differentiable at $a$.
[138] We're making another assumption here — that $g'(a) \neq 0$.

$\lim_{x\to 0} \frac{\sin x}{x}$ Here $f(x) = \sin x$, $g(x) = x$, all hypotheses of the claim are clearly satisfied, and

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

as we already knew.[139]

$\lim_{x\to 0} \frac{x}{\tan x}$ Recall(?) that $\tan x = \frac{\sin x}{\cos x}$, so by the quotient rule,

$$\tan' x = \frac{(\sin' x)(\cos x) - (\sin x)(\cos' x)}{(\cos x)^2} = \frac{(\cos x)^2 + (\sin x)^2}{(\cos x)^2} = \frac{1}{(\cos x)^2}.$$

It follows that all hypotheses of the claim are satisfied, and so

$$\lim_{x\to 0} \frac{x}{\tan x} = \frac{1}{1/(\cos 0)^2} = 1.$$

Alternately we could write $x/\tan x = (x\cos x)/(\sin x)$, and, since the derivative of $x \cos x$ is $-x\sin x + \cos x$, obtain

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

What we have so far is a very weak form of L'Hôpital's rule. It is not capable, for example, of dealing with

$$\lim_{x\to 1} \frac{x^3 - x^2 - x + 1}{x^3 - 3x + 2},$$

because although $f$ and $g$ are both 0 at 1, and both differentiable at 1, the derivative of $g$ at 1 is 0. We can, however, deal with this kind of expression using simple algebraic manipulation: away from 1

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

so

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

The issue L'Hôpital's rule is running into here is that what's causing $g$ to be zero at 1 is somehow "order 2"; one pass of differentiating only half deals with the problem.

There is a much more powerful version of L'Hôpital's rule that gets around this issue by making *far* fewer assumptions on $f$ and $g$: differentiability of $f$ and $g$ at $a$ is dropped (and so, continuity, and even existence), and replaced with the hypothesis that near $a$, the limit of $f'(x)/g'(x)$ exists (and so, at least, we are demanding that $f$ and $g$ be differentiable and continuous *near* $a$). Here is the strongest statement of L'Hôpital's rule.[140]

[139] But note that this is more of a reality check than an example. We used this particular limit to discover that the derivative of sin is cos, so using L'Hôpital (which requires knowing the derivative of sin) to calculate the limit, is somewhat circular!

[140] The proof is quite messy, and will only appear in these notes, not in class.

**Theorem 9.12.** *(L'Hôpital's rule) Let $f$ and $g$ be functions defined and differentiable near $a$*[141]*. Suppose that*

- $\lim_{x\to a} f(x) = 0$,
- $\lim_{x\to a} g(x) = 0$, *and*
- $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ *exists.*[142]

*Then* $\lim_{x\to a} \frac{f(x)}{g(x)}$ *exists, equals* $\lim_{x\to a} \frac{f'(x)}{g'(x)}$.

This version of L'Hôpital's rule is ideal for iterated applications. Consider, for example,

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

Does this exist? By L'Hôpital's rule, it does if

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

exists (and if so, the two limits have the same value). Does this second limit exist? Again by L'Hôpital's rule, it does if

$$\lim_{x\to 1} \frac{6x - 2}{6x}$$

exists (and if so, all three limits have the same value). But this last limit clearly exists and equals 2/3, so we conclude

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

In practice, we would be more likely to present the argument much more compactly as follows:

$$\begin{aligned} \text{``}\lim_{x\to 1} \frac{x^3 - x^2 - x + 1}{x^3 - 3x + 2} &= \lim_{x\to 1} \frac{3x^2 - 2x - 1}{3x^2 - 3} \quad \text{(by L'Hôpital's rule)} \\ &= \lim_{x\to 1} \frac{6x - 2}{6x} \quad \text{(by L'Hôpital's rule)} \\ &= \frac{2}{3}\text{''}, \end{aligned}$$

where all limits are seen to exist, and all applications of L'Hôpital's rule are seen to be valid, by considering the chain of equalities from bottom to top.

The proof of L'Hôpital's rule relies on a generalization of the Mean Value Theorem, known as the *Cauchy Mean Value Theorem*, that considers slopes of parameterized curve.

**Definition of a parameterized curve** A *parameterized curve* is a set of points of the form $(f(t), g(t))$, where $f$ and $g$ are functions; specifically it is $\{(f(t), g(t)) : t \in [a, b]\}$ where $[a, b]$ is (some subset of) the domain(s) of $f$ and of $g$.

[141]But not necessarily even defined *at* $a$.

[142]Note that we don't require $g'(a) \neq 0$: $g'(a)$ might not even exist!

Think of a particle moving in 2-dimensional space, with $f(t)$ denoting the $x$-coordinate of the point at time $t$, and $g(t)$ denoting the $y$-coordinate. Then the parameterized curve traces out the location of the particle as time goes from $a$ to $b$.

The graph of the function $f : [a, b] \to \mathbb{R}$ can be viewed as a parameterized curve — for example, it is $\{(t, f(t)) : t \in [a, b]\}$[143] On the other hand, not every parameterized curve is the graph of a function. For example, the curve $\{(\cos t, \sin t) : t \in [0, 2\pi]\}$ is a circle (the unit radius circle centered at $(0, 0)$), but is not the graph of a function.

We can talk about the *slope* of a parameterized curve at time $t$: using the same argument we made to motivate the derivative being the slope of the graph of a function, it makes sense to say that the slope of the curve $\{(f(t), g(t)) : t \in [a, b]\}$ at some time $t \in (a, b)$ is

$$\lim_{h\to 0} \frac{f(t+h) - f(t)}{g(t+h) - g(t)} = \lim_{h\to 0} \frac{(f(t+h) - f(t))/h}{(g(t+h) - g(t))/h} = \frac{\lim_{h\to 0}(f(t+h) - f(t))/h}{\lim_{h\to 0}(g(t+h) - g(t))/h} = \frac{f'(t)}{g'(t)},$$

assuming $f'(t)$, $g'(t)$ exist and $g'(t) \neq 0$.

We can also talk about the *average* slope of the curve, across the time interval $[a, b]$; it's

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

assuming $g(a) \neq g(b)$. The Cauchy Mean Value Theorem says that if the parameterized curve is suitably smooth, there is some point along the curve where the slope is equal to the average slope.

**Theorem 9.13.** *(Cauchy Mean Value Theorem) Suppose that $f, g : [a, b] \to \mathbb{R}$ are both continuous on $[a, b]$ and differentiable on $(a, b)$. There is $t \in (a, b)$ with*

$$(f(b) - f(a))g'(t) = (g(b) - g(a))f'(t).$$

Before turning to the (short) proof, some remarks are in order.

- If $g(b) \neq g(a)$ then the theorem says that

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(t)}{g'(t)}$$

  for some $t \in (a, b)$; that is, there is a point of the parameterized curve $\{(f(t), g(t)) : t \in [a, b]\}$ where the slope equal the average slope (as promised).

- If $g$ is the identity ($g(x) = x$) then the Cauchy Mean Value theorem says that if $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on $(a, b)$, then there is $t \in (a, b)$ with

$$f(b) - f(a) = (b - a)f'(t);$$

  this is *exactly* the Mean Value Theorem.

[143] But this representation is not unique. For example, $\{(t, f(t)\} : t \in [0, 1]$ and $\{(t^2, f(t^2)) : t \in [0, 1]\}$ both trace out the same graph, that of the squaring function on domain $[0, 1]$; but they are different parameterized curves, since the particles are moving at different speeds in each case.

- Using the Mean Value Theorem, we can quickly get something that looks a little like the Cauchy MVT: there's $t_1 \in (a,b)$ with

$$\frac{f(b)-f(a)}{b-a} = f'(t_1)$$

and $t_2 \in (a,b)$ with

$$\frac{g(b)-g(a)}{b-a} = g'(t_2),$$

from which it follows that

$$(f(b)-f(a))g'(t_2) = (g(b)-g(a))f'(t_1).$$

The power of the Cauchy MVT is that it is possible to take $t_1 = t_2$, and this can't be obviously deduced from the Mean Value Theorem.

**Proof** (of Cauchy Mean Value Theorem): Define

$$h(x) = (g(b)-g(a))f(x) - (f(b)-f(a))g(x).$$

This is continuous on $[a,b]$ and differentiable on $(a,b)$. Also,

$$\begin{aligned} h(a) &= (g(b)-g(a))f(a) - (f(b)-f(a))g(a) \\ &= g(b)f(a) - f(b)g(a) \\ &= (g(b)-g(a))f(b) - (f(b)-f(a))g(b) \\ &= h(b). \end{aligned}$$

By Rolle's theorem (or MVT) there is $t \in (a,b)$ with $h'(t) = 0$. But

$$h'(t) = (g(b)-g(a))f'(t) - (f(b)-f(a))g'(t)$$

so $h'(t) = 0$ says $(f(b)-f(a))g'(t) = (g(b)-g(a))f'(t)$, as required. □

**Proof** (of L'Hôpital's rule)[144]: To begin the proof of L'Hôpital's rule, note that a number of facts about $f$ and $g$ are implicit from the facts that $\lim_{x\to a} f(x) = 0$, $\lim_{x\to a} g(x) = 0$, and $\lim_{x\to a} f'(x)/g'(x)$ exists:

[144]Here's a sketch of the argument. $f$ and $g$ are both continuous on some interval $(a, a+\Delta)$ (because they are differentiable near $a$). Since $f, g \to 0$ near $a$, we can declare $g(a) = f(a) = 0$ to make the functions both continuous on $[a, \Delta$ (this may change the value of $f, g$ at $a$, but won't change any of the limits involved in L'Hôpital's rule). Now for each $b < \Delta$ we have (since $f, g$ are continuous on $[a,b]$ and differentiable on $(a,b)$) that $f(b)/g(b) = (f(b)-f(a))/(g(b)-g(a)) = f'(c)/g'(c)$ for some $c \in (a,b)$; this is Cauchy MVT. As $b$ approaches $a$ from above, the $c$ that comes out of CMVT approaches $a$, so near $a$ (from above) $f(b)/g(b)$ approaches $\lim_{c\to a^+} f'(c)/g'(c)$. A very similar argument gives the limit from below. Because $f, g$ are not known to be differentiable at $a$, CMVT can't be applied in any interval that has $a$ in its interior, which is why the argument gets split up into a "from above" and "from below" part.

- both $f$ and $g$ are differentiable and hence continuous in some open interval around $a$, except possibly at $a$ itself (neither $f$ nor $g$ are necessarily even defined at $a$) and
- there is some open interval around $a$ on which the derivative of $g$ is never 0 (again, we rule out considering the derivative of $g$ at $a$ here, as this quantity may not exist).

Combining these observations, we see that there exists a number $\delta > 0$ such that on $(a-\delta, a+\delta) \setminus \{a\}$ both $f$ and $g$ are continuous and differentiable and $g'$ is never 0.

Redefine $f$ and $g$ by declaring $f(a) = g(a) = 0$ (this may entailing increasing the domains of $f$ and/or $g$, or changing values at one point). Notice that after $f$ and $g$ have been re-defined, the hypotheses of L'Hôpital's rule remain satisfied, and if we can show the conclusion for the re-defined functions, then we trivially have the conclusion for the original functions (all this because in considering limits approaching $a$, we never consider values at $a$). Notice also that $f$ and $g$ are now both continuous at $a$, so are in fact continuous on the whole interval $(a-\delta, a+\delta)$.

In particular, this means that we can apply both the Mean Value Theorem and Cauchy's Mean Value Theorem on any interval of the form $[a, b]$ for $b < a+\delta$ or $[b, a]$ for $b > a-\delta$ (we have to split the argument into a consideration of two intervals, one to the right of $a$ and one to the left, because we do not know whether $f$ and/or $g$ are differentiable at $a$).

Given any $b$, $a < b < a+\delta$, we claim that $g(b) \neq 0$. Indeed, if $g(b) = 0$ then applying the Mean Value Theorem to $g$ on the interval $[a, b]$ we find that there is $c$, $a < c < b$, with $g'(c) = (g(b) - g(a))/(b-a) = 0$, but we know that $g'$ is never 0 on $(a, a+\delta)$. Similarly we find that $g(b) \neq 0$ for any $b$, $a-\delta < b < a$.

Fix an $x$, $a < x < a+\delta$. Applying Cauchy's Mean Value Theorem on the interval $[a, x]$ we find that there is an $\alpha_x$, $a < \alpha_x < x$, such that

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\alpha_x)}{g'(\alpha_x)}.$$

(Here we use $g(a) = f(a) = 0$ and the fact that $g(x) \neq 0$).

Since $\alpha_x \to a^+$ as $x \to a^+$, and since $\lim_{x \to a^+} f'(x)/g'(x)$ exists, it seems clear that

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = \lim_{x \to a^+} \frac{f'(x)}{g'(x)}, \tag{6}$$

and by similar reasoning on the interval $(a-\delta, a)$ we should have

$$\lim_{x \to a^-} \frac{f(x)}{g(x)} = \lim_{x \to a^-} \frac{f'(x)}{g'(x)}. \tag{7}$$

L'Hôpital's rule follows from a combination of (6) and (7).

Thus to complete the proof of L'Hôpital's rule we need to verify (6). Fix $\varepsilon > 0$. There is a $\delta' > 0$ such that $a < x < a+\delta'$ implies $|f'(x)/g'(x) - L| < \varepsilon$, where $L = \lim_{x \to a^+} f'(x)/g'(x)$. We may certainly assume that $\delta' < \delta$. But then $a < x < a+\delta$, and so we have that

$f(x)/g(x) = f'(\alpha_x)/g'(\alpha_x)$ where $a < \alpha_x < x < a + \delta'$. Since $\alpha_x$ is close enough to $a$ we have $|f'(\alpha_x)/g'(\alpha_x) - L| < \varepsilon$ and so $|f(x)/g(x) - L| < \varepsilon$. We have shown that $a < x < a + \delta'$ implies $|f(x)/g(x) - L| < \varepsilon$, which is the statement that $L = \lim_{x \to a^+} f(x)/g(x)$. This completes the verification of (6). □

The expressions that L'Hôpital's rule helps calculate the limits of, are often referred to as "indeterminates of the form 0/0" (for an obvious reason). There is a more general form of L'Hôpital's rule, that can deal with more "indeterminate" forms. In what follows, we use "lim" to stand for any of the limits

- $\lim_{x \to a}$,
- $\lim_{x \to a^-}$,
- $\lim_{x \to a^+}$,
- $\lim_{x \to \infty}$, or
- $\lim_{x \to -\infty}$,

and in interpreting the following claim, we understand that whichever version of "lim" we are thinking of for the first limit ($\lim f$), we are thinking of the *same* version for all the others ($\lim g$, $\lim f'/g'$ and $\lim f/g$).

**Claim 9.14.** *(General form of L'Hôpital's rule)*[145] *Suppose that* $\lim f(x)$ *and* $\lim g(x)$ *are either both* 0 *or are both* $\pm\infty$. *If*

$$\lim \frac{f'}{g'}$$

*has a finite value, or if the limit is* $\pm\infty$ *then*

$$\lim \frac{f}{g} = \frac{f'}{g'}.$$

We won't give a prove of this version of L'Hôpital's rule, but here's a sketch of how one of the variants goes. Suppose $\lim_{x\to\infty} f(x) = \infty$, $\lim_{x\to\infty} g(x) = \infty$, and $\lim_{x\to\infty} f'(x)/g'(x) = L$. Then we claim that $\lim_{x\to\infty} f(x)/g(x) = L$.

To show this we first have to argue a number of properties of $f$ and $g$, most of which are implicit in, or can be read out of, the statement that $\lim_{x\to\infty} f'(x)/g'(x)$ exists and is finite; verifying them all formally may be considered a good exercise in working with the definitions.

[145] As with the earlier version of L'Hôpital's rule, indeterminates of the form $\infty/\infty$ can have any limit, finite or infinite, or no limit. Consider, for example,

- $\lim_{x\to\infty} \frac{cx}{x} = c$, where $c$ can be any real;
- $\lim_{x\to\infty} \frac{\pm x^2}{x} = \pm\infty$; and
- $\lim_{x\to\infty} \frac{x(2+\sin x)}{x}$, which does not exist.

- At all sufficiently large number, $f$ is continuous;
- the same for $g$;
- for all sufficiently large $x$, $g'(x) \neq 0$; and
- if $x > N$ and $N$ is sufficiently large, then $g(x) - g(N) \neq 0$ (this follows from Rolle's theorem: if $N$ is large enough that $g'(c) \neq 0$ for all $c > N$, then if $g(x) = g(N)$ Rolle's theorem would imply that $g'(y) = 0$ for some $c \in (N, x)$, a contradiction).

Now write

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(N)}{g(x) - g(N)} \cdot \frac{f(x)}{f(x) - f(N)} \cdot \frac{g(x) - g(N)}{g(x)}. \quad (\star)$$

For each fixed $N$, the fact that $\lim_{x\to\infty} f(x) = \infty$ says that eventually (for all sufficiently large $x$) $f(x) - f(N) \neq 0$, so it makes sense to talk about $\lim_{x\to\infty} f(x)/(f(x) - f(N))$; and (again since $\lim_{x\to\infty} f(x) = \infty$) we have $\lim_{x\to\infty} f(x)/(f(x) - f(N)) = 1$. Similarly (since $\lim_{x\to\infty} g(x) = \infty$) we have $\lim_{x\to\infty}(g(x) - g(N))/g(x) = 1$. In both limits calculated here, we are using that $N$ is *fixed*, so that $f(N), g(N)$ are just fixed numbers.

Now for any $N$ that is large enough that $f$ and $g$ are both continuous on $[N, \infty)$ and differentiable on $(N, \infty)$, with $g'(x) \neq 0$ for any $x > N$ and $g(x) - g(N) \neq 0$ for any $x > N$ (such an $N$ exists, by our previous observations), the Cauchy Mean Value Theorem tells us that there is $c \in (N, x)$ with

$$\frac{f(x) - f(N)}{g(x) - g(N)} = \frac{f'(c)}{g'(c)}.$$

Because $\lim_{x\to\infty} f'(x)/g'(x) = L$, we can make the first term in $(\star)$ be as close as we want to $L$; and by then choosing $x$ sufficiently large, we can make the second and third terms in $(\star)$ be arbitrarily close to 1. In this way, the product of the three terms can be made arbitrarily close to $L$.

Good examples of the use of this more general form of L'Hôpital's rule are not so easy to come by at the moment; the rule really shows its strength when we deal with the exponential, logarithm and power functions, which we won't see until later. If you know about these functions, then the following example will make sense; if not, just ignore it.

Consider $f(x) = (\log x)/x$[146] What does this look like for large $x$? It's an indeterminate of the form $\infty/\infty$, so by L'Hôpital's rule the limit $\lim_{x\to\infty}(\log x)/x$ equals $\lim_{x\to\infty}(\log' x)/x' = \lim_{x\to\infty}(1/x)/1 = \lim_{x\to\infty} 1/x$, as long as this limit exists. Since this limit exists and equals

[146]Here $\log : (0, \infty) \to \mathbb{R}$ is the *natural logarithm* function, which has the property that if $\log x = y$ then $x = e^y$, where $e$ is a particular real number, approximately 2.71828, called the *base of the natural logarithm*. We'll see why such an odd looking function is "natural" next semester. The properties of log that we'll use in this example are that $\lim_{x\to\infty} \log x = \infty$, and that $\log'(x) = 1/x$.

0, it follows that

$$\lim_{x\to\infty} \frac{\log x}{x} = 0.$$[147]

Going to Desmos and looking at the graphs of $f(x) = \log x$ (entered as `ln x`) and $g(x) = x$, it seems pretty clear that for even quite small $x$, $\log x$ is dwarfed by $x$, so it is not surprising that the limit is 0. On the other hand, looking at the graphs of $f(x) = (\log x)^2$ and $g(x) = \sqrt{x}$, it's less clear what the limit

$$\lim_{x\to\infty} \frac{(\log x)^2}{\sqrt{x}}$$

might be. Looking at $x$ up to about, say, 180, it seems that $f(x)$ is growing faster than $\sqrt{x}$, but for larger values of $x$ the trend reverses, and at about $x = 5,500$, $g(x)$ has caught up with $f(x)$, and from there on seems to outpace it. This suggests that the limit might be 0. We can verify this using L'Hôpital's rule. With the usual caveat that the equalities are valid as long as the limits actually exist (which they will all be seen to do, by applications of L'Hôpital's rule, working from the back) we have

$$\begin{aligned}
\lim_{x\to\infty} \frac{(\log x)^2}{\sqrt{x}} &= \lim_{x\to\infty} \frac{2(\log x)(1/x)}{1/(2\sqrt{x})} \\
&= \lim_{x\to\infty} \frac{4\log x}{\sqrt{x}} \\
&= \lim_{x\to\infty} \frac{4/x}{1/(2\sqrt{x})} \\
&= \lim_{x\to\infty} \frac{8}{\sqrt{x}} \\
&= 0.
\end{aligned}$$

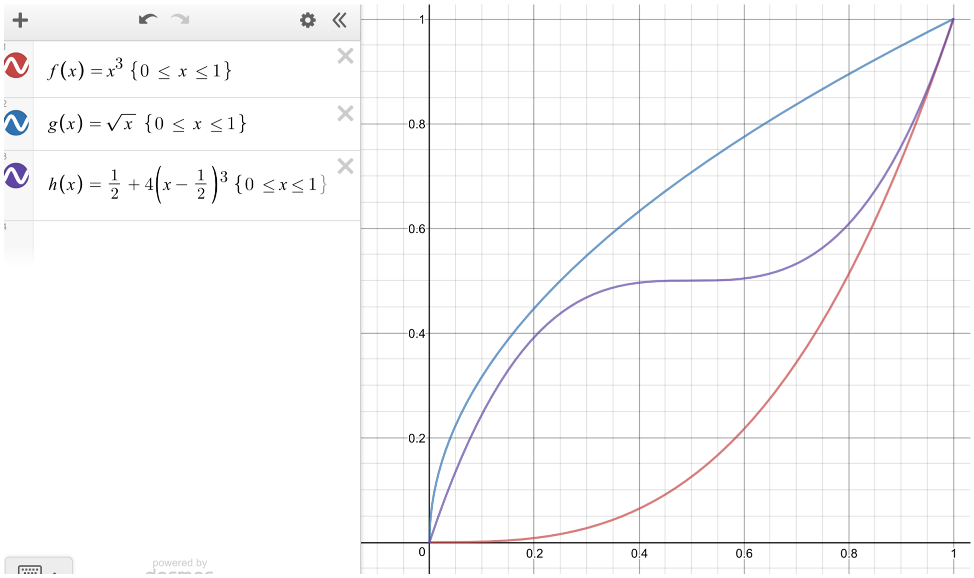
## 9.5 Convexity and concavity

Knowing that $f'(x) \geq 0$ for all $x \in [0,1]$ tells us that $f$ is (weakly) increasing on $[0,1]$, but that doesn't tell the whole story. Below there is illustrated the graphs of three functions,

- $f(x) = x^3$
- $g(x) = \sqrt{x}$ and
- $h(x) = \frac{1}{2} + 4\left(x - \frac{1}{2}\right)^3$,

all of which are increasing on $[0,1]$, but that otherwise look very different from each other.

[147] What about $\lim_{x\to\infty} x^{1/x}$? Write $x^{1/x} = e^{\log(x^{1/x})} = e^{(\log x)/x}$. Since $(\log x)/x$ approaches 0 as $x$ gets larger, it seems reasonable that $e^{(\log x)/x}$ approaches $e^0 = 1$; so $\lim_{x\to\infty} x^{1/x} = 1$. This is a very typical application of L'Hôpital's rule: we have two parts of a function that are competing with each other (in this case the $x$ in the base, causing $x^{1/x}$ to grow larger as $x$ grows, and the $1/x$ in the exponent, causing $x^{1/x}$ to grow smaller as $x$ grows), and L'Hôpital's rule (often) allows for a determination of which of the two "wins" in the limit.




The fine-tuning of the graph of a function "bulges"' is captured by the second derivative. Before delving into that, we formalize what we mean by the graph "bulging".

Let $f$ be a function whose domain includes the interval $I$.

**Definition of a function being convex** Say that $f$ is *strictly convex*, or just *convex*[148], on $I$ if for all $a, b \in I$, $a < b$, and for all $t$, $0 < t < 1$,

$$f((1-t)a + tb) < (1-t)f(a) + tf(b).$$

If instead $f((1-t)a + tb) \le (1-t)f(a) + tf(b)$ for all $a, b$ and $t$, say that $f$ is *weakly convex* on the interval.

**Definition of a function being concave** Say that $f$ is *strictly concave*, or just *concave*[149], on $I$ if for all $a, b \in I$, $a < b$, and for all $t$, $0 < t < 1$,

$$f((1-t)a + tb) > (1-t)f(a) + tf(b).$$

If instead $f((1-t)a + tb) \ge (1-t)f(a) + tf(b)$ for all $a, b$ and $t$, we say that $f$ is *weakly concave* on the interval.

Notice that as $t$ varies from 0 to 1, the value of $(1-t)a + tb$ varies from $a$ to $b$. The point

$$((1-t)a + tb, f((1-t)a + tb))$$

[148] Just as with "increasing", there is no universal convention about the meaning of the word "convex", without a qualifying adjective. By the word "convex", some people (including Spivak and me) mean what in the definition above is called "strictly convex", and others mean what above is called "weakly convex". It's a slight ambiguity that you have to learn to live with.

[149] Some authors, especially of 1000-page books called "Calculus and early transcendental functions, 45th edition", use "concave up" for what we are calling "convex", and "concave down" for what we are calling "concave". These phrases (the "up"-"down" ones) are almost never used in discussions among contemporary mathematicians.

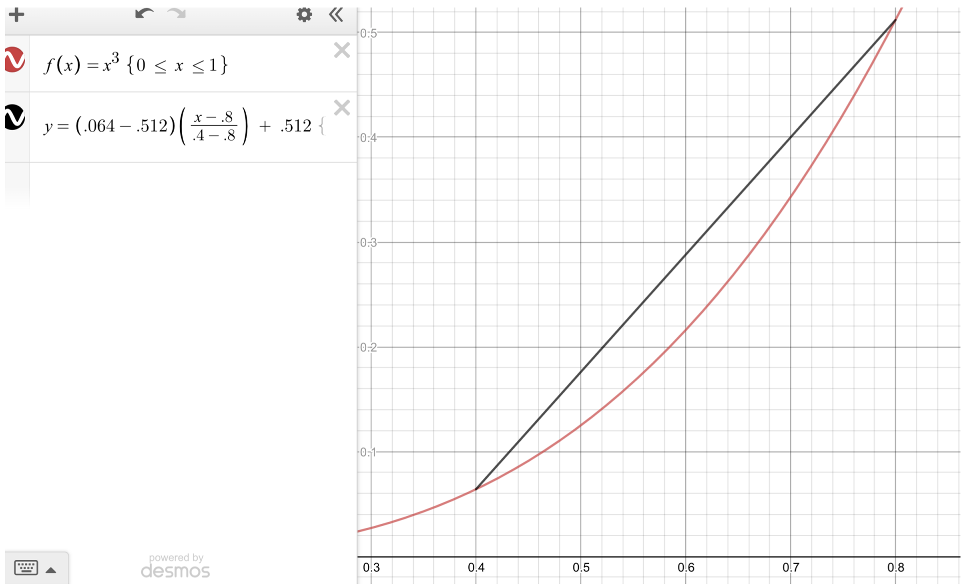
is a point on the graph of the function $f$, while the point

$$((1-t)a + tb, (1-t)f(a) + tf(b))$$

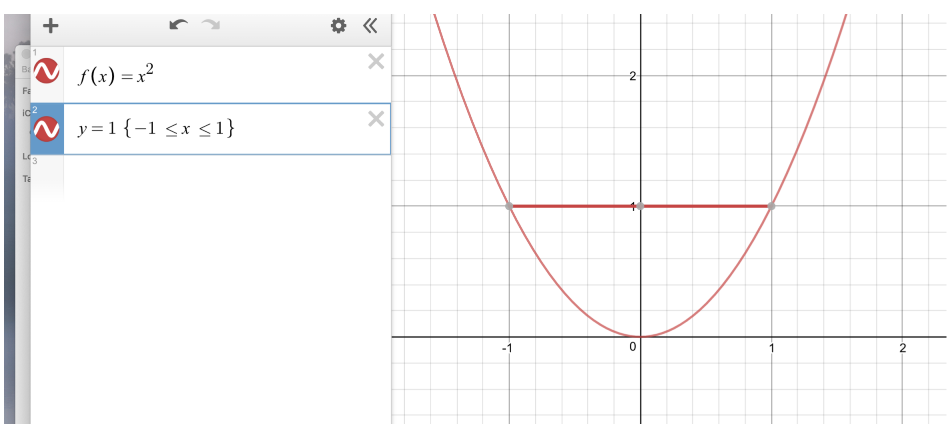
is a point on the secant line to the graph of the function $f$ between the points $(a, f(a))$ and $(b, f(b))$. So the graphical sense of convexity is that

> $f$ is convex on $I$ if the graph of $f$ lies below the graphs of all its secant lines on $I$.

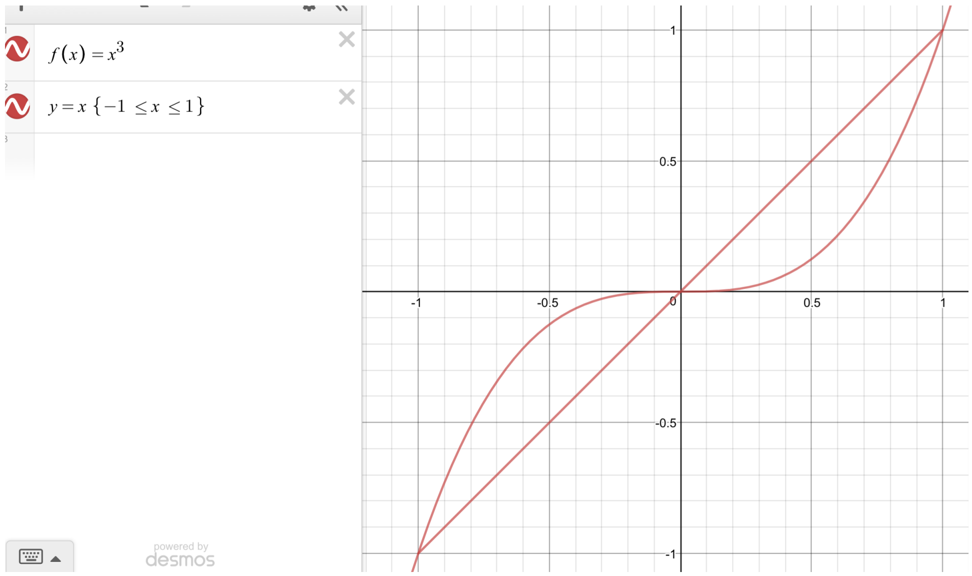
Illustrated below is the graph of $f(x) = x^3$, which lies below all its secant lines between 0 and 1, and so fairly clearly is convex on that interval. The picture below shows one secant line, from $(0.4, 0.64)$ to $(0.8, 0.512)$.




It is worth noting that convexity/concavity has nothing to do with $f$ increasing or decreasing. It should be fairly clear from the graph of $s(x) = x^2$ that this function is convex on the entire interval $(-\infty, \infty)$, even though it is sometimes decreasing and sometimes increasing. (The picture below shows a secant line lying above the graph of $s$, that straddles the local minimum).



On the other hand, it's clear that $f(x) = x^3$ is concave on $(-\infty, 0]$ and convex on $[0, \infty)$, though it is increasing throughout. (The picture below shows one secant line, on the negative side, lying *below* the graph of $f$, and another, on the positive side, lying above).



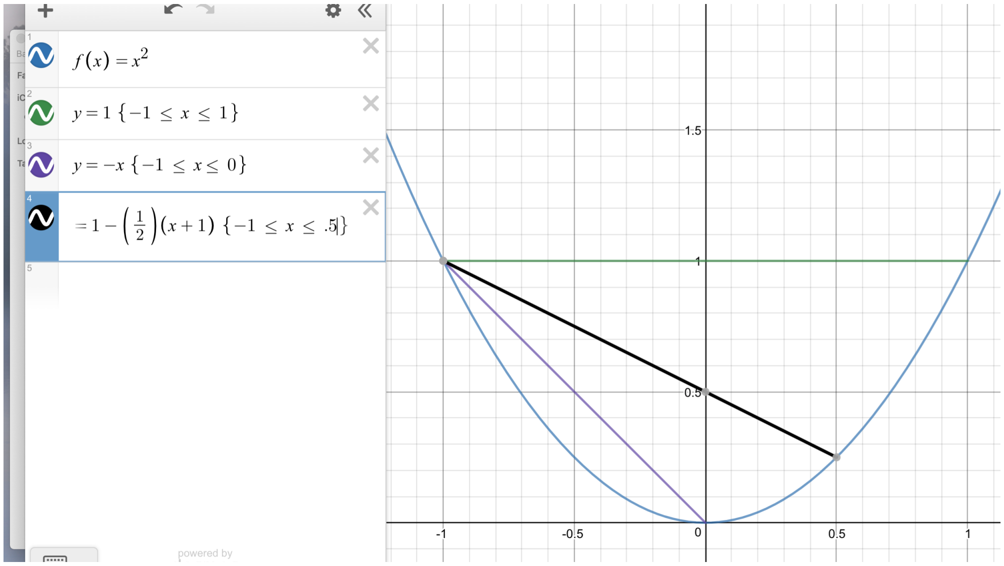

As was implicitly assumed in the last example discussed, just as convexity graphically means that secant lines lie above the graph, we have a graphical interpretation of concavity:

$f$ is concave on $I$ if the graph of $f$ lies *above* the graphs of all its secant lines on $I$.

In terms of proving properties about convexity and concavity, there is an easier way to think about concavity. The proof of the following very easy claim is left to the reader; it is evident from thinking about graphs.

**Claim 9.15.** *$f$ is concave of an interval $I$ if and only if $-f$ is convex on $I$.*

There is an alternate algebraic characterization of convexity and concavity, that will be very useful when proving things. If $f$ is concave on $I$, and $a, b \in I$ with $a < b$, then it seems clear from a graph that as $x$ runs from $a$ to $b$, the slope of the secant line from the point $(a, f(a))$ to the point $(x, f(x))$ is *increasing*. The picture below illustrates this with the square function, with $a = -1$ and $b = 1$. The slope of the secant line from $(-1, 1)$ to $(0, 0)$ is $-1$; from $(-1, 1)$ to $(1/2, 1/4)$ is $-1/2$; and from $(-1, 1)$ to $(1, 1)$ is $0$.




We capture this observation in the following claim, which merely says that as $x$ runs from $a$ to $b$, the slopes of all the secant lines are *smaller* than the slope of the secant line from $(a, f(a))$ to $(b, f(b))$.

**Claim 9.16.** • *$f$ is convex on $I$ if and only if for all $a, b \in I$ with $a < b$, and for all $x \in (a, b)$ we have*

$$\frac{f(x)-f(a)}{x-a} < \frac{f(b)-f(a)}{b-a}. \quad (\star)$$

*Also*[150]*, $f$ is convex on $I$ if and only if for all $a, b \in I$ with $a < b$, and for all $x \in (a, b)$ we have*

$$\frac{f(b)-f(a)}{b-a} < \frac{f(b)-f(x)}{b-x}.$$

• *$f$ is concave on $I$ if and only if for all $a, b \in I$ with $a < b$, and for all $x \in (a, b)$ we have*

$$\frac{f(x)-f(a)}{x-a} > \frac{f(b)-f(a)}{b-a} > \frac{f(b)-f(x)}{b-x}.$$

**Proof**: The key point is that any $x \in (a, b)$ has a unique representation as $x = (1-t)a + tb$ with $0 < t < 1$, specifically with

$$t = \frac{x-a}{b-a}$$

(it is an easy check that this particular $t$ works; that it is the unique $t$ that works follows from the fact that for $t \neq t'$, $(1-t)a + tb \neq (1-t')a + t'b$). So, $f$ being convex on $I$ says *precisely* that for $a < x < b \in I$,

$$f(x) < \left(1 - \frac{x-a}{b-a}\right) f(a) + \left(\frac{x-a}{b-a}\right) f(b).$$

[150] This next clause of the claim says that convexity also means that as $x$ runs from $a$ to $b$, the slopes of the secant lines from $(x, f(x))$ to $(b, f(b))$ *increase*. This can also easily be motivated by a picture.

Subtracting $f(a)$ from both sides, and dividing across by $x-a$, this is seen to be equivalent to

$$\frac{f(x)-f(a)}{x-a} < \frac{f(b)-f(a)}{b-a},$$

as claimed.

But now also note that

$$\left(1-\frac{x-a}{b-a}\right)f(a)+\left(\frac{x-a}{b-a}\right)f(b) = \left(\frac{b-x}{b-a}\right)f(a)+\left(1-\frac{b-x}{b-a}\right)f(b),$$

so $f$ being convex on $I$ also says *precisely* that for $a<x<b\in I$,

$$f(x) < \left(\frac{b-x}{b-a}\right)f(a)+\left(1-\frac{b-x}{b-a}\right)f(b),$$

which after similar algebra to before is equivalent to

$$\frac{f(b)-f(a)}{b-a} < \frac{f(b)-f(x)}{b-x},$$

also as claimed.

We now move on to the concavity statements. $f$ being concave means that $-f$ is convex, which (by what we have just proven) is equivalent to

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

for $a<x<b\in I$, and (multiplying both sides by $-1$) this is equivalent to

$$\frac{f(b)-f(a)}{b-a} > \frac{f(b)-f(x)}{b-x};$$

and the other claimed inequality for concavity is proved similarly. □

This alternate characterization of convexity and concavity allows us to understand the relationship between convexity and the derivative.

**Theorem 9.17.** *Suppose that $f$ is convex on an interval. If $f$ is differentiable at $a$ and $b$ in the interval, with $a<b$, then $f'(a)<f'(b)$ (and so, if $f$ is differentiable everywhere on the interval, then $f'$ is increasing on the interval).*

**Proof**: We will use our alternate characterization for convexity to show that

$$f'(a) < \frac{f(b)-f(a)}{b-a} < f'(b).$$

Pick any $b'\in(a,b)$. Applying our alternate characterization on concavity on the interval $[a,b']$, we have that for any $x\in(a,b')$,

$$\frac{f(x)-f(a)}{x-a} < \frac{f(b')-f(a)}{b'-a}.$$

Because $f$ is differentiable at $a$, we have[151] that

$$f'(a) = f'_+(a) = \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} \leq \frac{f(b') - f(a)}{b' - a}.$$

But now, also applying our alternate characterization of convexity on the interval $[a, b]$, and noting that $a < b' < b$, we have

$$\frac{f(b') - f(a)}{b' - a} < \frac{f(b) - f(a)}{b - a}.$$

It follows that

$$f'(a) < \frac{f(b) - f(a)}{b - a}.$$

So far, we have only used the first part of the alternate characterization of convexity (the part marked $(\star)$ above). Using the second part, an almost identical argument (which is left as an exercise) yields

$$\frac{f(b) - f(a)}{b - a} < f'(b),$$

and we are done. □

There is of course a similar theorem relating concavity and the derivative, which can be proven by using the fact that $f$ is concave iff $-f$ is convex (it is left as an exercise).

**Theorem 9.18.** *Suppose that $f$ is concave on an interval. If $f$ is differentiable at $a$ and $b$ in the interval, with $a < b$, then $f'(a) > f'(b)$ (and so, if $f$ is differentiable everywhere on the interval, then $f'$ is decreasing on the interval).*

There is a converse to these last two theorems.

**Theorem 9.19.** *Suppose that $f$ is differentiable on an interval. If $f'$ is increasing on the interval, then $f$ is convex, which if $f'$ is decreasing, then $f$ is concave.*

Before proving this, we make some comments.

- We are now in a position to use the first derivative to pin down intervals where a function is convex/concave — the intervals of convexity are precisely the intervals where $f'$ is increasing, and the intervals of concavity are those where $f'$ is decreasing. Of course, the easiest way to pin down intervals where $f'$ is increasing/decreasing is to look at the derivative of $f'$ (if it exists). That leads to the following corollary.

  **Corollary 9.20.** *If $f$ is twice differentiable, then the intervals where $f''$ is positive (so $f'$ is increasing) are the intervals of convexity, and the intervals where $f''$ is negative (so $f'$ is decreasing) are the intervals of concavity.*

[151] In the next line, we use a fact that we may not have formally proved, but is easy to prove (and very useful): suppose that $f(x) < M$ (for some constant $M$) for all $x > a$, and that $\lim_{x \to a^+} f(x)$ exists. Then $\lim_{x \to a^+} f(x) \leq M$.

The places where $f$ transitions from being concave to convex or vice-versa (usually, but not always, where $f''$ is zero), are called *points of inflection*.

- As an example, consider $f(x) = x/(1+x^2)$. Its domain is all reals. It goes to 0 as $x$ goes to both $+\infty$ and to $-\infty$. We have

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

which is

  - negative for $x < -1$ (so $x$ is decreasing on $(-\infty, -1)$),
  - positive for $-1 < x < 1$ (so $x$ is increasing on $(-1, 1)$), and
  - negative for $x > 1$ (so $x$ is decreasing on $(1, \infty)$).

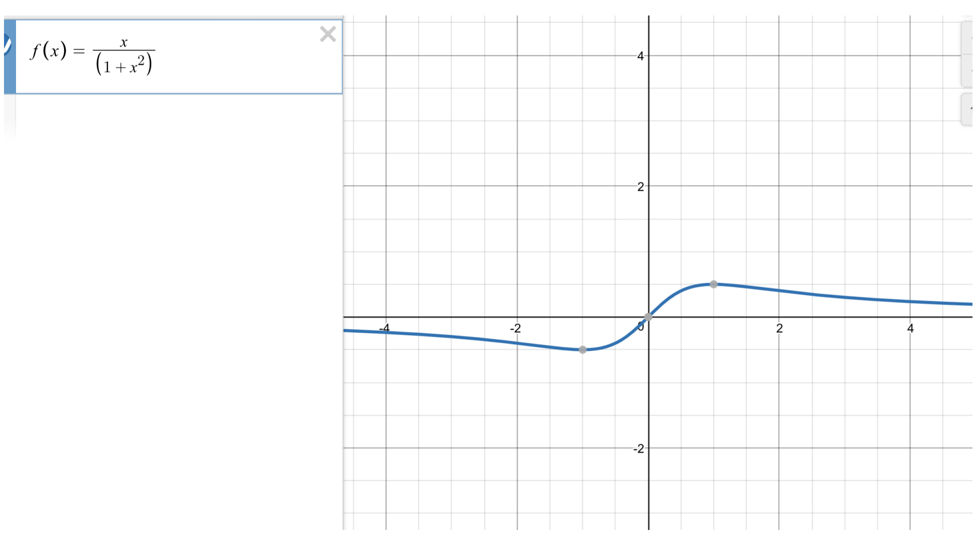
It follows that there is a local minimum at $(-1, -1/2)$ and a local maximum at $(1, 1/2)$. We also have

$$f''(x) = \frac{2x(x^2-3)}{(1+x^2)^3},$$

which is

  - negative for $x < -\sqrt{3}$ (so $x$ is concave on $(-\infty, -\sqrt{3})$),
  - positive for $-\sqrt{3} < x < 0$ (so $x$ is convex on $(-\sqrt{3}, 0)$),
  - negative for $0 < x < \sqrt{3}$ (so $x$ is concave on $(0, \sqrt{3})$), and
  - positive for $\sqrt{3} < x < \infty$ (so $x$ is convex on $(\sqrt{3}, \infty)$).

It follows that there is are points of inflection at $(-\sqrt{3}, -\sqrt{3}/4)$ and at $(\sqrt{3}, \sqrt{3}/4)$. Based on all of this information, it is not surprising to see that Desmos renders the graph of the function as follows.



Before proving Theorem 9.19, we need a preliminary lemma, the motivation for which which is that if $f$ is convex between $a$ and $b$, and $f(a) = f(b)$, then we expect the graph of $f$ to always be below the line joining $(a, f(a))$ to $(b, f(b))$.

**Lemma 9.21.** *Suppose $f$ is differentiable on an interval, with $f'$ increasing on the interval. For $a < b$ in the interval, if $f(a) = f(b)$ then for all $x \in (a, b)$, $f(x) < f(a)$ and $f(x) < f(b)$.*

**Proof**: Suppose there is an $x \in (a, b)$ with $f(x) \geq f(a)$ (and so also $f(x) \geq f(b)$). Then there is a maximum point of $f$ on $[a, b]$ at some particular $x \in (a, b)$. This is, of course, also a maximum point of $f$ on $(a, b)$. Since $f$ is differentiable everywhere, by Fermat principle $f'(x) = 0$. By the Mean Value Theorem applied to $[a, x]$, there is $x' \in (a, x)$ with

$$f'(x') = \frac{f(x) - f(a)}{x - a}. \quad (\star)$$

Now $f'$ is increasing on $[a, b]$ (by hypothesis), so $f'(x') < f'(x) = 0$. But $f(x) \geq f(a)$ (since $x$ is a maximum point for $f$ on $[a, b]$), so $(f(x) - f(a))/(x - a) \geq 0$. This contradicts the equality in $(\star)$ above.[152] □

**Proof** (of Theorem 9.19): Recall that we wish to show that if $f$ is differentiable on an interval, and if $f'$ is increasing on the interval, then $f$ is convex (the concavity part is left as an exercise; it follows as usual from the observation that $f$ is concave iff $-f$ is convex).

Given $a < b$ in the interval, set

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

We have $g(a) = g(b)$ (both are equal to $f(a)$). We also have

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

which is increasing on the interval, since $f'$ is. It follows from the preliminary lemma that for all $x \in (a, b)$, we have $g(x) < g(a)$ and $g(x) < g(b)$. The first of these says

$$f(x) - \frac{f(b) - f(a)}{b - a}(x - a) < f(a),$$

or

$$\frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(a)}{b - a},$$

which is the characterization $(\star)$ of convexity from earlier; so $f$ is convex on the interval.[153] □

[152]Note that we didn't actually use that $f(a) = f(b)$ in this proof, so what we actually showed was that if $f$ is differentiable on an interval, with $f'$ increasing on the interval, then for any $a < b$ in the interval, for all $x \in (a, b)$ we have $f(x) < \max\{f(a), f(b)\}$. The weaker result stated is, however, easier to comprehend visually, and it is the case that we will use in a moment.

[153]The second inequality, $g(x) < g(b)$, similarly reduces to the other characterization of convexity, but that isn't needed here.