Chapter 5: The Derivative

Definition. Let I be an interval, $c \in I$, and $f : I \to \mathbb{R}$. If

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

exists, it is called the *derivative of f at c* and denoted $f'(c)$, and f is called *differentiable at c*. If $f'(c)$ exists for every $c \in I$, then f is *differentiable on I*.

Proposition. *If $f : I \to \mathbb{R}$ is differentiable at c in I, then f is continuous at c.*

This should actually be clear: Because the denominator in the definition of the derivative is going to 0, the only way the limit could exist is for the the numerator to be going to 0 as well; otherwise, the quotient is becoming unbounded. But here is a proof, anyway:

Proof. Let $\varepsilon > 0$ be given, and pick $\delta > 0$ s.t., first, if $x \in I$ and $|x - c| < \delta$, then

$$\left| \frac{f(x) - f(c)}{x - c} - f'(c) \right| < \frac{\varepsilon}{2} \ ,$$

and second, $\delta < \min(1, \varepsilon/2|f'(c)|)$ (unless $f'(c) = 0$, in which case we just need $\delta < 1$). Then for $x \in I$ with $|x - c| < \delta$, we have

$$|f(x) - f(c)| - |f'(c)||x - c| \le \left| \frac{f(x) - f(c)}{x - c} - f'(c) \right| |x - c| < \frac{\varepsilon}{2} \cdot 1 \ ,$$

so $|f(x) - f(c)| < (\varepsilon/2) + |f'(c)|(\varepsilon/2|f'(c)|) = \varepsilon$. $\square$

Immediately we admit that we don't want to use the δ-ε definition of the derivative any more than necessary, so we prove (or at least stipulate) the rules of differentiation of functions combined by the operations of arithmetic, that we learned so well in calc class:

Theorem. (Algebraic Derivative Theorems) *If $f, g : I \to \mathbb{R}$ are differentiable at $c \in I$, then so are $f + g$, af where a is a constant, fg and (unless $g(c) = 0$) f/g; and*

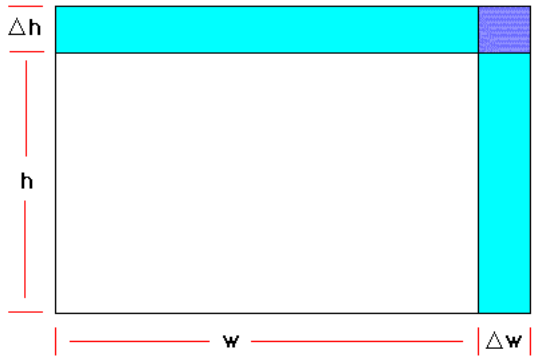
- $(f + g)'(c) = f'(c) + g'(c)$
- $(af)'(c) = af'(c)$
- $(fg)'(c) = f(c)g'(c) + f'(c)g(c)$
- $(f/g)'(c) = (g(c)f'(c) - f(c)g'(c))/(g(c))^2$

The proofs for addition and scalar multiplication are left to you, and the proof for multiplication can be seen from the following diagram:

$$A(t) = h(t)w(t)$$

$$\frac{dA}{dt} = \lim_{\Delta t \to 0} \frac{h(t+\Delta t)w(t+\Delta t) - h(t)w(t)}{\Delta t}$$

$$\Delta h = h(t+\Delta t) - h(t) \qquad \Delta w = w(t+\Delta t) - w(t)$$



$$\begin{aligned}
\frac{dA}{dt} &= \lim_{\Delta t \to 0} \frac{\Delta A}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{h\,\Delta w + \Delta h\, w + \Delta h\,\Delta w}{\Delta t} \\
&= \lim_{\Delta t \to 0} \left(h \underbrace{\frac{\Delta w}{\Delta t}}_{\to \frac{dw}{dt}} + \underbrace{\frac{\Delta h}{\Delta t}}_{\to \frac{dh}{dt}} w + \underbrace{\Delta h}_{\to 0} \underbrace{\frac{\Delta w}{\Delta t}}_{\to \frac{dw}{dt}} \right) \\
&= h \frac{dw}{dt} + \frac{dh}{dt} w
\end{aligned}$$

Here is the proof for the quotient rule:

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

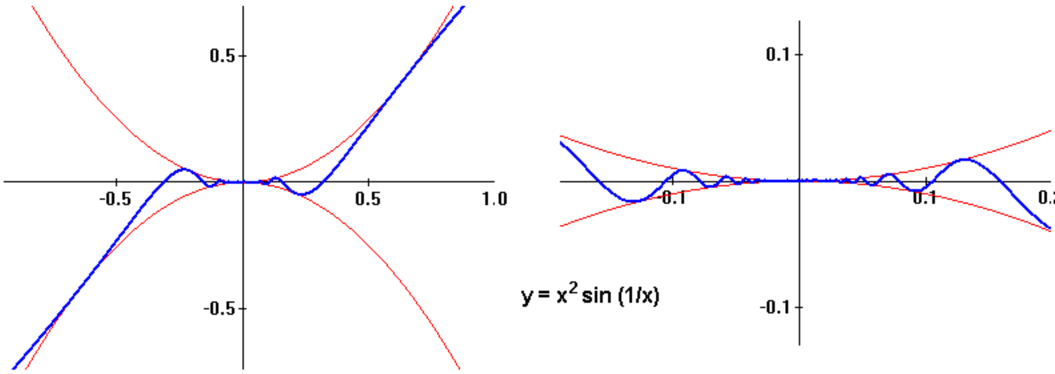
Recall that the Chain Rule is the formula for the derivative of the composition of two functions, say f that takes x-values to u-values followed by g that takes u-values to y-values. It is tempting to write:

$$\frac{g(f(x)) - g(f(c))}{x-c} = \frac{g(f(x)) - g(f(c))}{f(x) - f(c)} \cdot \frac{f(x)-f(c)}{x-c}\ , \qquad (*)$$

then to say that, as $x \to c$, the second difference quotient approaches $f'(c)$ and, because $f(x) \to f(c)$, the first difference quotient goes to $g'(f(c))$. The problem is that there may be a set of x-values with limit point c at which $f(x) = f(c)$, so that the first difference quotient is not defined in a whole interval around c, and hence it is not crystal clear that the limit of that difference quotient is $g'(f(c))$. (Here is an example of an f that takes the same value at $c = 0$ and at a set of points with limit 0:

$$f(x) = \begin{cases} x^2 \sin\frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

Here is the graph of f:



We will make another use of this function shortly.) But that is a nitpick, which we fix up by giving the difference quotient the value at $f(c)$ that it is approaching anyway as the denominator goes to 0 (namely $g'(f(c))$), so that it becomes continuous at $f(c)$:

Theorem. (Chain Rule) *Suppose I, J are intervals and $f : I \to J$ is differentiable at $c \in I$ and $g : J \to \mathbb{R}$ is differentiable at $f(c)$. Then $g \circ f : I \to \mathbb{R}$ is differentiable at c and $(g \circ f)'(c) = g'(f(c))f'(c)$.*

Proof. Define $h : J \to \mathbb{R}$ by

$$h(u) = \begin{cases} \frac{g(u)-g(f(c))}{u-f(c)} & \text{if } u \neq f(c) \\ g'(f(c)) & \text{if } u = f(c) \end{cases}$$

Then h is continuous at $f(c)$ and we have

$$\frac{g(f(x)) - g(f(c))}{x - c} = h(f(x)) \cdot \frac{f(x) - f(c)}{x - c}$$

whenever $x \neq c$. (Because $h(f(x))$ has a meaning even where $x \neq c$ but $f(x) = f(c)$, both sides of the equation are 0 at such points, whereas one side of $(*)$ didn't make sense at such points.) Taking the limit as $x \to c$ gives the desired result. $\square$

Next, we prove a very familiar theorem from calculus:

Theorem. (Fermat) *If $f : I \to \mathbb{R}$ attains an extremum at an <u>interior</u> point c of I and f is differentiable at c, then $f'(c) = 0$.*

Proof. We may assume c is a minimum on I, i.e., $f(c) \leq f(x)$ for all $x \in I$. Then $f(x) - f(c) \geq 0$ for all $x \in I$, so

- if $x < c$, we have $x - c < 0$, so $(f(x) - f(c))/(x - c) \leq 0$, so

$$f'(c) = \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} \leq 0 \; ;$$

  while

- if $x > c$, we have $x - c > 0$, so $(f(x) - f(c))/(x - c) \geq 0$, so

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

The only way both of these can be true is if $f'(c) = 0$. $\square$

The following theorem is also simple, but it is not usually proved in calc classes because it isn't used there.

Theorem. (Darboux) *If $f : I \to \mathbb{R}$ is differentiable on I, then f' has the "Intermediate Value Property" on I, i.e., if $a, b \in I$ and $f'(a) < r < f'(b)$ (or vice versa), then there is a c between a, b (and hence in I) for which $f'(c) = r$.*

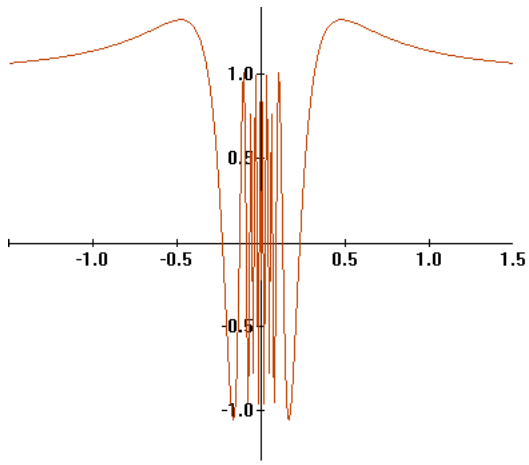
Before we give the proof, let us recall the function f from earlier:

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

The graph of the function shows how the x^2 factor squeezes the otherwise wildly oscillating function so that the derivative at the origin is 0. Here is the verification:

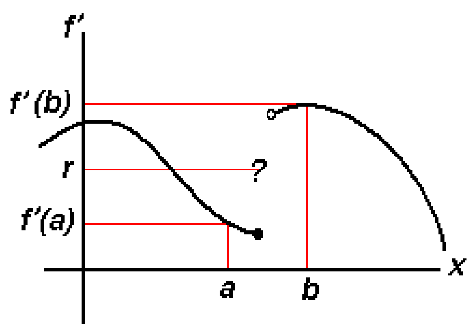
$$f'(0) = \lim_{x \to 0} \frac{x^2 \sin \frac{1}{x} - 0}{x - 0} = \lim_{x \to 0} x \sin \frac{1}{x} = 0 \ .$$

But we can compute the derivative at all other points by the usual formula, and we see that it oscillates wildly:

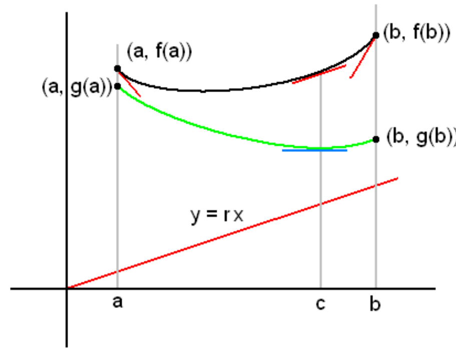


y = -cos(1/x) + 2x sin(1/x)

So f' is discontinuous at $x = 0$. Therefore, it is possible for a derivative to be discontinuous. But Darboux's theorem says that a derivative f' cannot have a jump discontinuity, because that would leave a gap in the f'-values:



The idea of the proof is, given f and an r between the slopes of f at its ends a, b, subtract the line rx from f to get g, so that 0 is between the slopes of g at a, b. Then g decreases from a and increases into b (in this picture — if $f(a) > r > f(b)$, then g increases from a and decreases into b), so g has an interior extremum, where g' is 0 and hence f' is r.



Proof. (of Darboux) Define $g(x) = f(x) - rx$. Then $g'(a) = f'(a) - r < 0$, so $\lim_{x \to a^+}(g(x) - g(a))/(x - a) < 0$, so there is an $x_1 > a$ for which $g(x_1) < g(a)$. (We have just shown that g goes <u>down</u> in value as it x leaves a.) Thus, g does not attain its minimum at a. Similarly, because $g'(b) > 0$, (g comes <u>up</u> in value as x comes into b, i.e.,) there is an $x_2 < b$ for which $g(x_2) < g(b)$, so g doesn't attain its minimum at b, either. But as a continuous function on the compact set $[a, b]$, it <u>does</u> attain its minimum somewhere, at the interior point c of $[a, b]$. By Fermat, $g'(c) = 0$, so $f'(c) = r$. $\square$

Here is another familiar result:

Theorem. (Mean Value Theorem) *Let f be continuous on $[a, b]$ and differentiable on (a, b). Then there is a c in (a, b) for which*

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

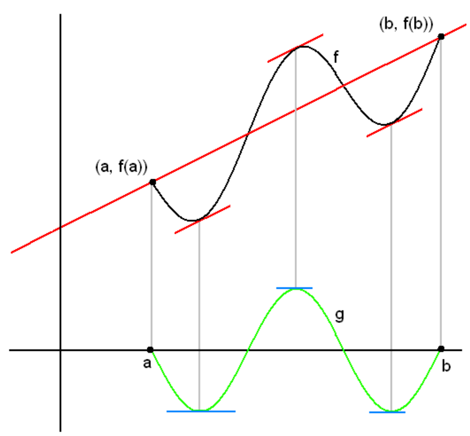
In words: Under these hypotheses, there is at least one point in the interval where the tangent line is parallel to the line joining the ends of the graph at the endpoints of the interval.

Proof. Define $g : [a, b] \to \mathbb{R}$ by

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

(So all we are doing is subtracting from f the linear function joining its endpoints.) Then g is continuous on $[a, b]$ and differentiable on (a, b), and $g(a) = 0 = g(b)$. Thus, one or more of the following hold:

- g is constantly 0 on $[a, b]$, in which case every point of $[a, b]$ is both a maximum and minimum of g, or
- $g(x) < 0$ for some x in $[a, b]$, in which case g attains its minimum at some point c of (a, b), or
- $g(x) > 0$ for some x in $[a, b]$, in which case g attains its maximum at some point c of (a, b).



By Fermat's theorem, $g'(c) = 0$ for some c in (a, b). Now $g'(x) = f'(x) - (f(b) - f(a))/(b - a)$, so the result follows. $\square$

It might have been simpler just to mimic the proof of Darboux and let

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

All we would lose is g being 0 at a, b; but all we need is $g(a) = g(b)$, which we still have.

Rolle's Theorem is a special case of the Mean Value Theorem, when the function agrees at the endpoints of the interval. Some books make it a lemma and prove the MVT by appealing to it; we didn't bother to separate it at the time, so for us it's a corollary.

Corollary. (Rolle) *If g is ctn on $[a, b]$ and diff on (a, b), and $g(a) = g(b)$, then $\exists\, c \in (a, b)$ s.t. $g'(c) = 0$.*

The following result will be handy when we talk about integrals, because they will involve antiderivatives.

Corollary. *If $f, g : I \to \mathbb{R}$ are differentiable on I and $f' = g'$ on I, then $\exists\, k \in \mathbb{R}$ s.t. $f(x) = g(x) + k$ $\forall\, x \in I$.*

Proof. Set $h = f - g$. Then $h' = 0$, and we want to show that h is a constant function. So assume BWOC that $\exists\, a, b \in I$ s.t. $h(a) \neq h(b)$. Then by the MVT $\exists\, c$ between a, b s.t. $h'(c) = (h(b) - h(a))/(b - a) \neq 0$, $\times$. □

We actually have use for a stronger version of the MVT:

Theorem. (Generalized MVT) *If f, g are continuous on $[a, b]$ and differentiable on (a, b), then $\exists\, c \in (a, b)$ s.t. $(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c)$. If $g'(x) \neq 0$ for all $x \in (a, b)$, then we can rewrite the above to*

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

Proof. Set $h(x) = (f(b) - f(a))g(x) - (g(b) - g(a))f(x)$. Then h is continuous on $[a, b]$ and differentiable on (a, b), and $h(a) = f(b)g(a) - g(b)f(a) = -f(a)g(b) + g(a)f(b) = h(b)$. So by Rolle's theorem, $\exists\, c \in (a, b)$ s.t. $h'(c) = 0$, and the first conclusion follows.

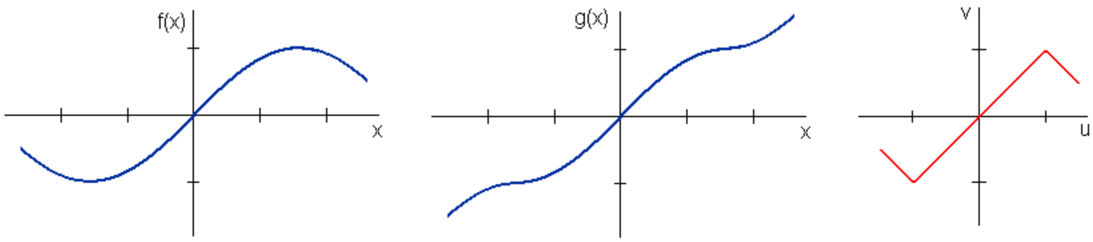
For the last sentence, we know $g'(c) \neq 0$, so we only need to show that $g(b) - g(a) \neq 0$. But by the MVT, $\exists\, d \in (a, b)$ s.t. $g(b) - g(a) = g'(d)(b - a)$; because neither of $g'(d)$ and $b - a$ is 0, neither is $g(b) - g(a)$. □

Exercise 5.3.4, in addition to asking for a proof of the GMVT, also asks for a geometric interpretation. Suppose we regard $f(x)$ and $g(x)$ as parametric equations of a curve in, say, the (u, v)-plane — i.e., the image of a function from the x-line into the (u, v)-plane: $u = g(x)$ and $v = f(x)$. From that point of view, $f'(c)/g'(c)$ is the slope of the tangent to the curve at the point $x = c$; so the GMVT asserts that the slope of the line segment joining the endpoints of the curve is equal to the slope of the tangent to the curve at some point. However, the hypothesis that g' never be 0 is necessary:

Example. Consider the functions on the interval $-5\pi/6 \leq x \leq 5\pi/6$:

$$f(x) = \sin x \ , \qquad g(x) = \begin{cases} -2 - \sin x & \text{if } -\frac{5\pi}{6} < x < -\frac{\pi}{2} \\ \sin x & \text{if } -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ 2 - \sin x & \text{if } \frac{\pi}{2} < x < \frac{\pi}{2} \end{cases}$$

The graphs of f and g are smooth curves, but the image of the parametric curve $u = g(x)$, $v = f(x)$ is not smooth, and sure enough, there is no point at which the tangent has the same slope, $1/3$, as the segment joining the endpoints $(-3/2, -1/2)$ and $(3/2, 1/2)$. At the "corners" of the image, both f' and g' are 0.



We need the Generalized MVT to get a complete proof of L'Hôpital's rule; but generally the idea is this: If $f(a) = g(a) = 0$, then

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

so as $x \to a$, the limit of the left side is the same as the limit of the right, which should be $f'(a)/g'(a)$.

Corollary. (L'Hôpital's Rule, $0/0$ case): *Suppose f, g are continuous on an interval containing a and differentiable on the interval except perhaps at a, and $f(a) = g(a) = 0$. If $\lim_{x\to a}(f'(x)/g'(x))$ exists, then so does $\lim_{x\to a}(f(x)/g(x))$, and they are equal.*

The proof of this is an exercise, but here is a sketch: Given $\varepsilon > 0$, pick $\delta > 0$ to make f'/g' ε-close to its limit L. Then if x is δ-close to a, use the Generalized MVT to pick an x_1 that is even closer and ...

Corollary. (L'Hôpital's Rule, ∞/∞ case): *Suppose f, g are differentiable on an open interval with endpoint a and $\lim_{x\to a} g(x) = \infty$. If $\lim_{x\to a}(f'(x)/g'(x))$ exists, then so does $\lim_{x\to a}(f(x)/g(x))$, and they are equal.*

The proof of this version is in the text; it is left for you to read.

In going rather directly through the big calculus results in this chapter, I've left some loose ends that I want to tie up before we leave it:

(I) To apply L'Hôpital's rule at $x = a$, $f(a)/g(a)$ <u>must</u> be indeterminate: The limits

$$\lim_{x\to 2} \frac{x^2 - 3x + 4}{x^2 + 1} \qquad \text{and} \qquad \lim_{x\to 2} \frac{2x - 3}{2x}$$

are $2/5$ and $1/4$ respectively; they aren't equal, even though the numerator and denominator of the second fraction are the derivatives of the numerator and denominator of the first.

(II) As an example for a homework problem, let's compute the derivative of the function a^x (for a a positive constant) in two ways:

(i) Assuming a^x is defined by $a^x = \exp(x \ln a)$ where $\exp(u) = \sum_{n=0}^{\infty} u^n/n!$ and $\ln a$ is the real number s.t. $\exp(\ln(a)) = a$: We have

$$\exp'(u) = \sum_{n=1}^{\infty} \frac{nu^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{u^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{u^n}{n!} = \exp(u) \ .$$

Thus, by the Chain Rule,

$$\frac{d}{dx}(a^x) = \exp'(x \ln a) \cdot \ln a = \exp(x \ln a) \cdot \ln a = a^x \ln a \ .$$

(ii) Assuming that a^x is defined as the limit of a^q as q moves through a sequence of rationals with limit x: The laws of exponents work for rational numbers, and they extend to the reals by continuity. So

$$\frac{d}{dx}(a^x) = \lim_{h \to 0} \frac{a^{x+h} - a^x}{h} = \lim_{h \to 0} a^x \frac{a^h - 1}{h} = a^x \lim_{h \to 0} \frac{a^h - 1}{h} \ .$$

Now the last limit is the slope of the tangent line to $y = a^x$ at $x = 0$; we define that slope to be $\ln a$, and the derivative formula follows.

(III) I think the true-or-false questions in Exercise 5.2.8 on page 137 are too tricky to assign, but they are interesting, so let's work through them. Here is the problem:

True or false? Provide an argument for those that are true and a counterexample for each one that is false.

(a) If a derivative function is not constant, then the derivative must take on some irrational values.

(b) If f' exists on an open interval, and there is some point c where $f'(c) > 0$, then there exists a δ-neighborhood $V_\delta(c)$ around c in which $f'(x) > 0$ for all $x \in V_\delta(c)$.

(c) If f is differentiable on an interval containing 0 and if $\lim_{x \to 0} f'(x) = L$, then it must be that $L = f'(0)$.

(d) Repeat conjecture (c) but drop the assumption that $f'(0)$ necessarily exists. If $f'(x)$ exists for all $x \neq 0$ and if $\lim_{x \to 0} f'(x) = L$, then $f'(0)$ exists and equals L.

Answers:

(a) False, if we are talking about <u>any</u> derivative; for example, the absolute value function has a derivative that takes on only the values 1 and -1. But that derivative is not defined on an interval; 0 divides the domain into two pieces, and the absolute value is not differentiable there. If we require that the function be differentiable on an interval, the statement is true: The Intermediate Value Property (Darboux' theorem) says that, if $f'(a) \neq f'(b)$, then f' must take on all the values between them, and some of those values are irrational.

(b) False: This would be true if f' were continuous, but we have an example, namely $f(x) = x^2 \sin(1/x)$ (and 0 at 0) for which the derivative oscillates wildly near 0. For this example, $f'(0) = 0$, but if we add a term (say, $0.001x$) to make the derivative positive at 0, it will still be negative at points inside any neighborhood of 0.

(c) True: Because we have the hypothesis that f is differentiable at 0, $f'(0)$ exists. If L were not equal to $f'(0)$, then we could pick a small ε-neighborhood around L to separate it from $f'(0)$. But then we would have a gap between the $f'(x)$-values with x near 0 and $f'(0)$, so the Intermediate Value Property would be violated. So we must have $L = f'(0)$.

(d) Again, the answer is conditional, this time on whether f is continuous at 0:

- If f is allowed to be discontinuous (and hence not differentiable) at 0, then the statement is false: A counterexample is the step function that is -1 if $x < 0$ and 1 if $x > 0$.
- If f is required to be continuous at 0 (and differentiable near 0), then here is a proof showing that $L = f'(0)$: Let $\varepsilon > 0$ be given, and pick $\delta > 0$ for which, if x is in the domain of f' and $0 < |x - 0| < \delta$, then $|f'(x) - L| < \varepsilon$. Now take an x in the domain of f for which $0 < |x - 0| < \delta$; then by the MVT there is an x_1 between x and 0 for which $(f(x) - f(0))/(x - 0) = f'(x_1)$. Then because x_1 is in the domain of f' and satisfies $0 < |x_1 - 0| < \delta$, we have

  $$\left| \frac{f(x) - f(0)}{x - 0} - L \right| = |f'(x_1) - L| < \varepsilon \,.$$

  Therefore, $f'(0) = \lim_{x \to 0}(f(x) - f(0))/(x - 0) = L$.//

(IV) Some examples, in all of which the value of the function at 0 is assumed to be 0. A function that is:

- discontinuous at 0 but has the IVP: $\sin(1/x)$
- continuous but not differentiable at 0: $x \sin(1/x)$
- differentiable, but the derivative is discontinuous at 0: $x^2 \sin(1/x)$
- twice differentiable but the second derivative is discontinuous at 0: $x^4 \sin(1/x)$

I think you get the drift: Multiply $\sin(1/x)$ by enough factors of x to get the level of "smoothness" you want at 0.

(V) Remark: Consider the very similar functions

$$g(x) = \begin{cases} -1 \text{ if } x < 0 \\ \ \ 1 \text{ if } x \geq 0 \end{cases}, \qquad h(x) = \begin{cases} -1 \text{ if } x < 0 \\ \ \ 1 \text{ if } x > 0 \end{cases}$$

Because g has a jump discontinuity at 0 but is defined in a neighborhood of it, g cannot be the derivative of anything. But h, which is not defined at 0, is the derivative of the absolute value function.

`[At this point students are ready to do the tenth problem set, §5.2 and §5.3.]`

(VI) I'd like to throw in here a result that isn't in the text and isn't in many calc books anymore, though it used to be in every calc student's repetoire, for classifying points where the derivative is zero into local maxima, local minima, or neither:

Proposition. (Second Derivative Test) *Let f a differentiable function from an interval I into $\mathbb{R}$, and suppose $f'(c) = 0$ and $f''(c)$ exists at an interior point c of I. If $f''(c) < 0$ [respectively > 0], then c is a local maximum [respectively minimum] of f.*

Note that the signs are reverse from what you might think: A negative second derivative means a local maximum. Note also that the test says nothing about the case where $f''(c)$ is zero or doesn't exist. The test has fallen out of favor, probably, because second derivatives can be messy to compute, and they tend to be 0 when the first derivative is 0, so that the test doesn't give any information.

Proof. Suppose $f''(c) < 0$. Then there is a $\delta > 0$ for which $(c-\delta, c+\delta) \subseteq I$ and, if $0 < |x-c| < \delta$, then

$$\frac{f'(x) - f'(c)}{x-c} = \frac{f'(x)}{x-c} < 0 \ .$$

Thus, if $c - \delta < x < c$, we have $x - c < 0$, so $f'(x) > 0$; so f is increasing on $(c-\delta, c)$. Similarly, f is decreasing on $(c, c+\delta)$. Thus, $f(c) \geq f(x)$ for $x \in (c-\delta, c+\delta)$, i.e., c is a local maximum for f. The proof for $f''(c) < 0$ is similar. $\square$