Chapter 14

The First Derivative Test, Redux

> *"Geometry has always been considered as an exact science, and indeed as the source of the exactness which is widespread among other parts of mathematics. . . But it seems that this feature of exactness does not reign anymore in geometry since the new system of infinitely small quantities has been mixed to it. I do not see that this system has produced anything for the truth and it would seem to me that it often conceals mistakes."*
>
> – Michel Rolle (1652-1719)

Assuming that all of the properties of limits we talked about in Section 13.1 can be proved, we have seen that all of the differentiation rules we developed intuitively using differentials in Chapter 3 can be made rigorous using limits.

The question we need to address now is: Does Definition 15 also recover First Derivative Test? If not then our definition is inadequate and we need a better one. So this question really must be addressed.

14.1 Fermat's Theorem

> *"It is by logic that we prove, but by intuition that we discover."*
>
> – Henri Poincaré (1854-1912)

It is actually a little surprising how much effort it takes to prove the First Derivative Test. We will start by proving some preliminary results that will make it a little easier to follow the logic behind the proof of the First Derivative Test.

We'll begin with Fermat's Theorem (Theorem 1) which says that if f attains a maximum (or minimum) at $x = a$ then $f'(a) = 0$. More formally:

Theorem 34: **[Fermat's Theorem]**
If $f(a)$ is a local extremum (either a maximum or a minimum) of $f(x)$ at $x = a$, and $f(x)$ is differentiable at every point in an open interval containing $x = a$ then $f'(a) = 0$.

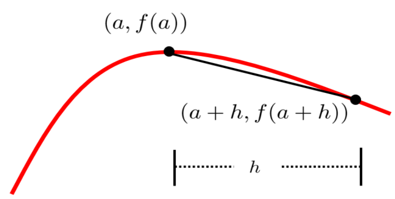
Recall that Fermat's Theorem does not say that if $f'(a) = 0$ then $f(a)$ is an extremum. In fact, we know that this is not true. Rather, it states the converse: If we know that $f(a)$ is an extremum and $f'(a)$ exists, then $f'(a) = 0$.

It is very rare that we can develop a proof of a theorem by directly writing down the logical steps in order. Usually the process takes a lot of trying, backtracking, trying again, and so on much as we described in our analogy in Chapter 1 about finding your way out of a forest.

Of course, in a textbook we can't follow all of the bad ideas we might have just to see that they are, in fact, bad ideas. So we will use the following **Scrapwork** construct when we are just "thinking about" a problem. The scrapwork is not the proof. The purpose of scrapwork is to engage our intuition and to begin organizing our intuitive understanding so that a rigorous proof will emerge. So not every statement we make inside a scrapwork construct will necessarily be fully rigorous. If you see a gap in the logic inside a scrapwork construct watch to see how it gets filled in the proof.

Scrapwork

Notice that in Definition 15 the quantity $\frac{f(a+h)-f(a)}{h}$ is the slope of a particular secant line, as in the sketch at the right.



If, as shown in the sketch, $(a, f(a))$ is a (local) maximum then the slope of the secant line in our diagram when $h > 0$ must be negative. Since the slope $\frac{f(a+h)-f(a)}{h}$ is *strictly* less than zero then

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

Drill #514:
Draw a similar diagram to convince yourself that $f'(a)$ must also be *greater than or equal to zero* when $h < 0$.

We can now rigorously prove Fermat's Theorem.

End Of Scrapwork

Proof of Fermat's Theorem: We will only prove the case when $f(a)$ is a local maximum. The case of a local minimum is very similar.

Since $f(a)$ is a local maximum there is an interval containing a such that for any h $f(a+h) \leq f(a)$. Thus $f(a+h) - f(a) \leq 0$ as seen in the sketch above. If $h < 0$ then $\frac{f(a+h)-f(a)}{h} \geq 0$ and so

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \geq 0. \tag{14.1}$$

However if $h > 0$ then

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \leq 0. \tag{14.2}$$

The only way that both Equations (14.1) and (14.2) can be true is if $f'(a) = 0$.

■

Problem #515: Use our proof of the maximum case as a guide to constructing a proof of Fermat's Theorem when $f(a)$ is a local minimum.

14.2 Rolle's Lemma and the Mean Value Theorem

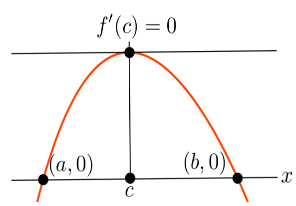
Michel Rolle
(1659-1719)

We will use Fermat's Theorem and Theorem 5 (the Extreme Value Theorem) to prove our next result, often called Rolle's Lemma. Michel Rolle was a contemporary of Newton, Leibniz, and Berkeley. Like Bishop Berkeley, Rolle was an early critic of Calculus. He once described it as a "collection of ingenious fallacies." Were he alive today he might be horrified to know that his lemma has become a fundamental part of the modern development of Calculus.

The distinction between a theorem and a lemma is very slight and (sometimes) rather arbitrary. Typically we call some statement[1] a theorem if it requires proof and it is important. We call a statement a lemma if it requires proof but its only purpose is to be used to help prove a theorem[2]. In the present instance we will use Rolle's Lemma to prove the Mean Value Theorem. Then we will use the Mean Value Theorem to prove the First Derivative Test.

Lemma 35: **[Rolle's Lemma]**

Suppose f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b). Suppose further that $f(a) = f(b)$. Then there is at least one number, c, in the interval (a, b) such that $f'(c) = 0$.



When Rolle's Lemma is visualized, as in the sketch at the right, it is clear what is going on. If $f(a) = f(b) = 0$ the between the points $(a, 0)$ and $(b, 0)$ the graph of $f(x)$ will either rise to a maximum or drop to a minimum (not shown) at some point c. In either case, by Fermat's Theorem, the derivative of f at c will be zero.

[1] Technically speaking, a statement is not a theorem until it has actually been proved. Until then it is a conjecture.

[2] This is not a hard-and-fast rule by any means. Sometimes we first prove a lemma as an aid to proving a theorem, only to find that the lemma is actually more important. However, having been originally dubbed a lemma the result is known ever after as a lemma. It is all very chaotic.

Rolle's Theorem is actually a special (simple) case of the Mean Value Theorem, so it is worth taking a few minutes to identify the essential features of our visualization. In particular, from our sketch it appears that the following statements are true.

- The line through $(a, f(a))$ and $(b, f(b))$ intersects the graph of $f(x)$ at $x = a$ and $x = b$.
- The slope of the line through $(a, f(a))$ and $(b, f(b))$ and the slope of the line tangent to the graph of $f(x)$ at $x = c$ are the same.

The second statement above actually is the statement of the Rolle's Theorem. Despite how clearly the truth of the statement is in our sketch, an analytical proof is still required because our sketch does not capture all of the possible ways that Rolle's Theorem can manifest. This is demonstrated in Drill 516 below.

Drill #516:
Sketch the graphs of two functions, different from each other and from our sketch above, which also satisfy the conditions of Rolle's Lemma.

Problem #517:
Sketch the graph of a function which violates the requirement that f be continuous on $[a, b]$, but which otherwise satisfies the conditions of Rolle's Lemma. Use your sketch to explain why continuity on the closed interval is required for Rolle's Lemma to be valid.
(**Hint:** If this problem looks completely obvious to you then you're probably missing the point. You want to show that it is necessary for f to be continuous on the *closed* interval, not the open interval.)

Proof of Rolle's Theorem: By The Extreme Value Theorem (5) there is are points α and β in the interval $[a, b]$, such that $f(\alpha)$ is a global maximum, and $f(\beta)$ is a global minimum.

There are two possibilities for α and β:

1. **Both α and β are endpoints of the interval**: In that case since $f(a) = f(b)$ the global maximum and the global minimum are equal. The only way that can happen is if the function is constant on the interval $[a, b]$, and if f is constant then $f'(x) = 0$ for every x in the interval (a, b). So we take c to be any point in (a, b).
2. **At least one of α or β is not an endpoint of the interval:** In that case by Fermat's Theorem, either $f'(\alpha) = 0$, or $f'(\beta) = 0$. So we take $c = \alpha$ or $c = \beta$ as appropriate.

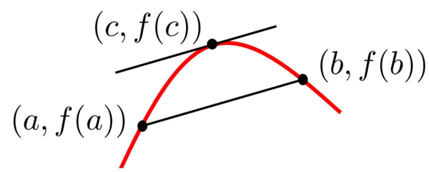
In either case the existence of c in the interval (a, b), with $f'(c) = 0$ is guaranteed. ■

Theorem 36: **[The Mean Value Theorem]**
Suppose $f(x)$ is continuous on some closed interval, $[a, b]$, and f is differentiable on (a, b). Then there is at least one number c in the open interval (a, b) such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{14.3}$$

Scrapwork

The Mean Value Theorem is visualized in the sketch below. Comparing this with our visualization of Rolle's Theorem we see that the two facts that we pointed out above still appear to be true.

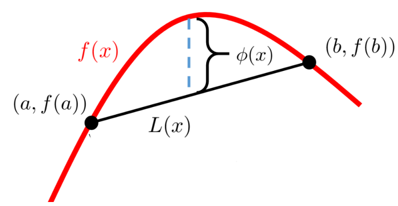


- The line through $(a, f(a))$ and $(b, f(b))$ intersects the graph of $f(x)$ at a and b.
- The slope of the line through $(a, f(a))$ and $(b, f(b))$ and the slope of the line tangent to the graph of $f(x)$ at $x = c$ are the same.

Once again, the second of these is actually a statement of the Mean Value Theorem.

If we're going to use Rolle's Theorem we need to create a function – we'll call it $\phi(x)$ – that satisfies all of the conditions of Rolle's Theorem. From the diagram at the right we have

$$\phi(x) = f(x) - L(x).$$



Do you see that $\phi(x)$ is the function we need?

Drill #518:

Show that $L(x) = \frac{f(b)-f(a)}{b-a}(x-a) + f(a)$.
(**Hint:** $L(x)$ is a straight line and we have the coordinates of two points, $(a, f(a))$ and $(b, f(b))$ on the line.)

Problem #519:

Show that $\phi(x) = f(x) - L(x)$ from the diagram above satisfies all of the conditions of Rolle's Theorem. That is show that,

(a) $\phi(x)$ is continuous on $[a, b]$.

(b) $\phi(x)$ is differentiable on (a, b).

(c) $\phi(a) = \phi(b) = 0$

End Of Scrapwork

Proof of The Mean Value Theorem:

Observe that

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

satisfies all of the conditions of Rolle's Theorem. Therefore, by Rolle's Lemma there is a point c, between a and b such that $\phi'(c) = 0$. Therefore

$$f'(c) - \frac{f(b) - f(a)}{b - a} = 0$$

and so

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

which completes the proof. ■

In French the Mean Value Theorem is known as the *théorème des accroissements finis* or literally the "theorem of the finite increments."

To see why this is an accurate description let $y = f(x)$. Then $f(b) - f(a) = \Delta y$ and $b - a = \Delta x$ so that we can re-express Equation (14.3) as $\Delta y = f'(c)\Delta x$. In this form it is clear that the Mean Value Theorem relates the finite ratio $\frac{\Delta y}{\Delta x}$ to the instantaneous rate of change $f'(c)$. You saw this in our physical interpretation with velocities, and in the first derivative test as we translated information about the derivative into information about changes in the function at the values $x = a$ and $x = b$.

14.3 The Proof of the First Derivative Test

A problem with Theorem 3 as we stated it in Chapter 8 is that it tells us whether a function is increasing or decreasing one point at a time. But what we really want are conditions on a function, $f(x)$ and an interval (α, β) which allow us to conclude that if $a, b \in (\alpha, \beta)$ and $a \geq b$ then $f(a) \geq f(b)$.

We restate the First Derivative Test to reflect this new, and deeper, understanding.

Theorem 37: [First Derivative Test]
Suppose $f(x)$ is continuous on the interval $[\alpha, \beta]$, differentiable on the interval (α, β). Suppose further that $b > a$ and both a and b are in the interval (α, β).

(a) If $f'(x) > 0$ on the interval (α, β) then $f(b) > f(a)$, and

(b) If $f'(x) < 0$ on the interval (α, β) then $f(b) < f(a)$.

Proof of Part (a): We want to use the Mean Value Theorem on the interval $[a, b]$ so we begin by verifying that the conditions of the Mean Value Theorem are satisfied on that interval. Observe that $[a, b]$ is a subinterval of $[\alpha, \beta]$ so $f(x)$ is continuous on $[a, b]$ and differentiable on (a, b).

By the Mean Value Theorem there is a number, c, in the interval (a, b) such that

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

Since both $b-a>0$ and $f'(c)>0$ are positive, $f(b)-f(a)$ must be positive as well. Therefore

$$f(b)-f(a)>0 \text{ or } f(b)>f(a).$$

■

Problem #520:

Prove part (b) of the First Derivative Test in two different ways:

(a) By modifying our proof of part (a) of the First Derivative Test as needed.

(b) Let $g(x)=-f(x)$ and apply part (a) of this problem.
(**Comment:** Don't forget to show first that $g(x)$ satisfies the conditions of the Mean Value Theorem.)

While we are in this frame of mind, we'll take a moment to notice that we can use the Mean Value Theorem to prove, rigorously, something that we have alluded to a few times but have never addressed directly. It is clear from our differentiation rules that if two functions differ by a constant, then they have the same derivative. We've mentioned that the converse is true, namely if two functions have the same derivative on an interval then they must differ by a constant. This can be proved in a manner similar to the proof above.

Problem #521:

(a) Suppose $f'(x)=0$ on the interval (α,β) and that a and b are two points in that interval. Use an argument similar to the proof of the first derivative test to show that $f(a)=f(b)$.

(b) Explain how the result of part (a) says that $f(x)$ must be constant on (α,β).

(c) Show that if $f'(x)=g'(x)$ on the interval (α,β), then $f(x)=g(x)+c$ for some constant c.
(**Hint:** Consider the function $F(x)=f(x)-g(x)$.)

(d) What can be said if $f'(x)=g'(x)$ for all x in some set S which is *not* an interval?