Rolle's Theorem & The Mean Value Theorem

C. Fortin & G. Indurskis

Champlain College St-Lambert

Dec. 3, 2018

Mathematical analysis is a modern area of research that evolved from Calculus and which now forms the theoretical foundation for it. You have already seen some statements from Analysis: The Extreme Value Theorem and the Intermediate Value Theorem are such statements. In this section, we will discuss two more important Analysis Theorems, which give the theoretical groundwork for some of the methods we have already (innocently) used in the past, and which also will enable us to prove new facts.

In everything that follows, we will often have to make the distinction between a *closed* and an *open* interval: A closed interval includes its endpoints where an open interval does not.

Rolle's Theorem

Theorem (Rolle's Theorem): Consider a closed interval $[a, b]$. If f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b) (it may not be differentiable at the endpoints), and $f(a) = f(b) = 0$ (a and b are roots of the function), then there is a point c in the open interval (a, b) where the tangent line is horizontal, that is $f'(c) = 0$.

This theorem was known to Indian mathematicians in the 12th century, but Michel Rolle, a French mathematician, provided the first known proof in the 17th century. Figure 1 illustrates the theorem whose conclusion is not too hard to believe, but not trivial to prove rigorously.[1]

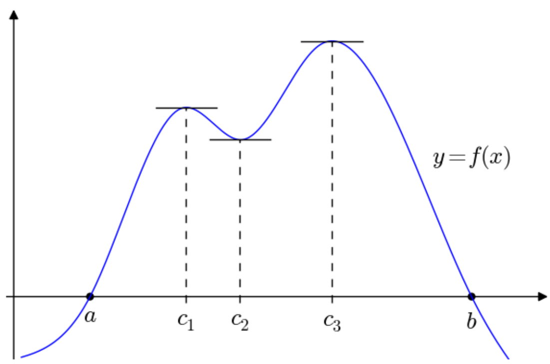


Figure 1: Rolle's Theorem guarantees the existence of some $c \in (a, b)$ where the derivative of f is 0. As shown on this figure, there may not be a unique c on the interval where the derivative is 0. In this example, there are three possible choices for c.

Example 1: Consider the function $f(x) = e^x \sin x - \sin 2x$. Show that there exists a value in the interval $(0, \pi)$ where the tangent line is horizontal.

Solution: To apply Rolle's Theorem, we first need to verify and justify the hypotheses of the theorem. Since f is the difference, the product, and the composition of continuous and differentiable functions on $\mathbb{R}$ (mainly the

[1]The proof of Rolle's Theorem relies on the Extreme Value Theorem and the fact that, at $x = c$, a local maximum or minimum of a differentiable function must satisfy $f'(c) = 0$ (this result is also called FERMAT'S THEOREM): If f is constant on the interval, c could be any value in the open interval (a, b), because $f'(c) = 0$ for all c when f is constant. If on the other hand f is not constant on the interval, there is a value within (a, b) where f is either positive or negative (the main point is that it is not zero). Using the Extreme Value Theorem, which holds since f is continuous on $[a, b]$, we have that one of the absolute extrema on $[a, b]$ is not at the endpoints (since f is zero there) and it must therefore be at $x = c$ for some $c \in (a, b)$. Because the function is differentiable on (a, b), by Fermat's Theorem, we can conclude that $f'(c) = 0$.

functions e^x, $\sin x$, and $2x$) then it is also continuous and differentiable on $\mathbb{R}$, and thus continuous and differentiable on $[0, \pi]$. We also have $f(0) = 1 \cdot 0 - 0 = 0$ and $f(\pi) = e^\pi \cdot 0 - 0 = 0$. Since the hypotheses of the theorem are verified, there exists $c \in (0, \pi)$ where $f'(c) = 0$. In other words, the tangent line at $x = c$ is horizontal. ♢

The previous example is a straightforward application of Rolle's Theorem and it is almost obvious from the question that it is the theorem that we need to use in the solution. Sometimes, it is not obvious that Rolle's Theorem is the right tool to answer a question.

Example 2: Show that the equation $x^5 + 3x^3 + x = 2$ has exactly one solution.

Solution: We write the equation as $x^5 + 3x^3 + x - 2 = 0$ and define the function $f(x) = x^5 + 3x^3 + x - 2$. The question can be rephrased in terms of f in the following way: show that f has exactly one root. First of all, we can ask ourselves if there is at least a root. Does the graph of f cross the x-axis? The right tool to answer this question is the Intermediate Value Theorem. Since f is a polynomial, it is continuous everywhere, and we can easily find two points where the function value has opposite signs, for example: $f(0) = -2 < 0$ and $f(1) = 3 > 0$. By the Intermediate Value Theorem f has a root in the interval $(0, 1)$.

But how do we know that this is the *only* root of f? This part is less obvious to show and this is where we use Rolle's Theorem. If f would have more than one root, there would be two distinct roots a and b, with $a < b$. In particular we would have $f(a) = f(b) = 0$. Because f is a polynomial, it is continuous and differentiable everywhere. Therefore, Rolle's Theorem would imply that $f'(c) = 0$ for some $c \in (a, b)$. Now $f'(c) = 5c^4 + 9c^2 + 1$, which is the sum of nonnegative terms (the first two) and a positive constant, and is therefore positive as well. It obviously cannot equal 0, and this means that our assumptions that two roots exist must be false in the first place. Hence the root of f must be unique. Notice that we did not use Rolle's Theorem directly here but its contrapositive: if f is continuous on $[a, b]$, differentiable on (a, b), and $f'(c)$ cannot equal 0 for any $c \in (a, b)$, then it is not possible to have $f(a) = f(b) = 0$. ♢

Activity 1: Consider the equation $6x^5 - 15x^4 + 20x^3 - 30x^2 + 30x = 30$.

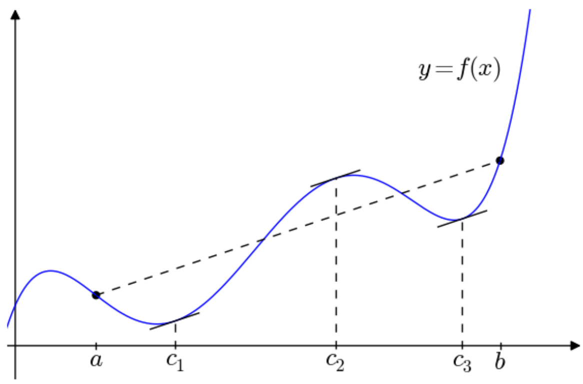
(a) Write the equation as $f(x) = 0$, where f is a polynomial with leading term $6x^5$.

(b) Find $f'(x)$ and verify that $f'(x) = 30(x^2 + 1)(x - 1)^2$.

(c) What is the maximum number of roots that f can have? What is the maximum number of solutions that the equation can have? Justify. ♦

The Mean Value Theorem

The main use of Rolle's Theorem is that we can use it to prove a generalized version of the theorem that is used more often in proofs: the Mean Value Theorem. This theorem can be thought of as a tilted version of Rolle's Theorem: In Rolle's theorem, a point $x = c$ is found where the function is locally as far as possible from the x-axis, where in the Mean Value Theorem we will consider points that are locally as far as possible from the secant line joining $(a, f(a))$ and $(b, f(b))$, which might be slanted and not necessarily the same as the x-axis. (See the exercises for a guided proof of this theorem.)



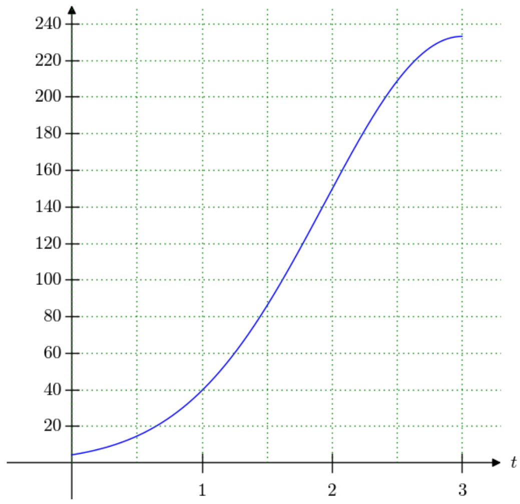
This figure illustrates the Mean Value Theorem. As in Rolle's Theorem the choice of c is not necessarily unique. In this example, there are three possible choices for c.

Theorem (Mean Value Theorem): Consider a closed interval $[a, b]$. If f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b) (it may not be differentiable at the end points), then there is a point $c \in (a, b)$ where the slope of the tangent line is equal to the slope of the secant line joining the points $(a, f(a))$ and $(b, f(b))$. This means that

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

Activity 2: The distance between Quebec City and Montreal is 233 km. A train going from Quebec City to Montreal has the following position function:

$$s(t) = 233e^{\frac{-(t-3)^2}{(1.5)^2}}$$



(a) Find the average speed of the train during the first three hours.

(b) Use a graphing device or the figure shown on the right to find approximately when the speed of the train is the same as the average speed during the first three hours.

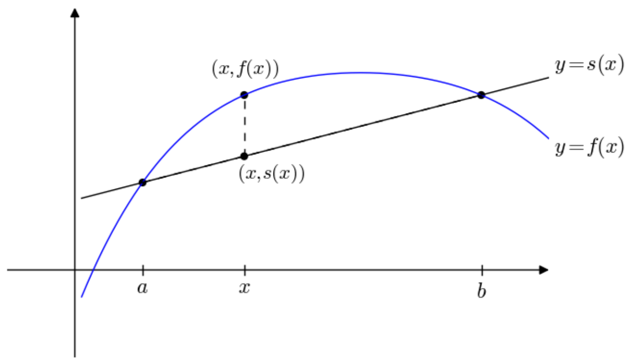
♦

Summary:

1. Rolle's Theorem says that a function which is continuous on an interval $[a, b]$ and differentiable on (a, b) which has roots at the endpoints of the interval must have a horizontal tangent line (a zero derivative) somewhere in (a, b). In can be used in particular to find the maximum number of solutions that the equation $f(x) = 0$ can have by looking at how many roots the derivative f' has.
2. The Mean Value Theorem is a generalization of Rolle's Theorem which says that a function which is continuous on an interval $[a, b]$ and differentiable on (a, b) must have a tangent line at $x = c \in (a, b)$ with a slope equal to the secant line going through $(a, f(a))$ and $(b, f(b))$. It can be used to relate the average rate of change of a function (the slope of the secant line) to an instantaneous rate of change (the derivative at $x = c$). For instance, we can relate the average velocity to the instantaneous velocity using this theorem.
3. The Mean Value Theorem is used as a tool for proving other results. In the exercises, you will do this to show several important results (which are worthwhile to remember):
   (i) A function which has a zero derivative everywhere is constant everywhere.
   (ii) Two functions that have the same derivatives everywhere must differ only by a constant. This result is always used when trying to find the *most general antiderivative* of a function, which is of great importance in Integral Calculus.
   (iii) The Mean Value Theorem is used to prove the FUNDAMENTAL THEOREM OF CALCULUS which (you guessed it) is seen in Integral Calculus.

Exercises:

1. Show that the function $f(x) = x^2e^x - 2 + 2x^2 - e^x$ has a horizontal tangent line in the interval $[-1, 1]$.
2. Show that the equation $\sin x + 2x = k$ has a unique solution for all possible values of k.
3. The Mean Value Theorem is often used to prove other results. In this exercise, we will use this theorem to prove a very important, yet obvious result: If f always has a horizontal tangent line on an open interval, then it is constant on that interval. More formally, if $f'(x) = 0$ for $x \in (a, b)$, then f is constant on (a, b). The idea of the proof is to show that for any two points x_1 and x_2 in the interval (a, b), the function values $f(x_1)$ and $f(x_2)$ are the same. Therefore all points in the interval (a, b) have the same function value.
   (i) Let x_1 and x_2 be two points in (a, b) with $a < x_1 < x_2 < b$. Why is the function differentiable on $[x_1, x_2]$?
   (ii) Why is the function continuous on $[x_1, x_2]$?
   (iii) Since f is differentiable on (x_1, x_2) and continuous on $[x_1, x_2]$, the Mean Value Theorem can be applied and there exists $c \in (x_1, x_2)$ such that
   $$f(x_2) - f(x_1) = f'(c)(x_2 - x_1).$$
   What can be said about $f'(c)$? Deduce that $f(x_1) = f(x_2)$.
4. Prove the following result: If two functions f and g have the same derivatives on an interval (a, b), then f and g only differ by a constant. More formally, if $f'(x) = g'(x)$ for all $x \in (a, b)$, then $f(x) = g(x) + c$ for $x \in (a, b)$, where c is a constant. To prove this, let $h(x) = f(x) - g(x)$, compute $h'(x)$, and apply the result of the previous exercise.
5. The previous result is very important in Integral calculus, but it can also be used to prove identities. For example, consider the identity
   $$2 \arcsin x = \arccos(1 - 2x^2),$$
   which holds for $0 \leq x \leq 1$. Let's use the previous result to show this identity. In order to do so, we shall define
   $$\begin{aligned} f(x) &= 2 \arcsin x, \\ g(x) &= \arccos(1 - 2x^2), \end{aligned}$$
   and show that $f(x) - g(x) = c$, where $c = 0$.
   (i) Compute $f'(x)$.
   (ii) Compute $g'(x)$ and simplify your answer as much as possible in order to show that $f'(x) = g'(x)$.
   (iii) Let $h(x) = f(x) - g(x)$ and apply the result of the previous exercise to show that $h(x) = c$ for $x \in (0, 1)$.
   (iv) In order to find the value of the constant c, find a value of $x \in (0, 1)$ where $h(x)$ is simple to evaluate. Replacing x by this value in h gives you exactly what c is since $h(x) = c$.
   (v) Evaluate $h(0)$ and $h(1)$ and conclude that $h(x) = c$ for all $x \in [0, 1]$.
   (vi) Deduce that for $0 \leq x \leq 1$
   $$2 \arcsin x = \arccos(1 - 2x^2).$$
6. In this exercise we will prove the Mean Value Theorem based on Rolle's Theorem. It might be worthwhile at this point to again read the comment relating both theorems which appeared just before the Mean Value Theorem was stated.
   (a) Recall that the equation of a line with slope m going through the point (x_1, y_1) can be written using the point-slope form
   $$y = m(x - x_1) + y_1.$$
   Find the slope of the secant line joining the points $(a, f(a))$ and $(b, f(b))$ and find the equation $y = s(x)$ of this secant line.
   (b) The distance, up to a sign, between a point $(x, f(x))$ on the graph of function f and a point $(x, s(x))$ on the secant line can be obtained by taking the difference between the y-coordinates of the points as shown below.



Create a function h which computes the distance, up to a sign, between the two points.

(c) Explain why h is continuous on $[a, b]$.

(d) Explain why h is differentiable on (a, b).

(e) Evaluate $h(a)$ and $h(b)$.

(f) Apply Rolle's Theorem to obtain the Mean Value Theorem.

7. In this exercise we will use the Mean Value Theorem to prove an important result in Integral Calculus, the FUNDAMENTAL THEOREM OF CALCULUS:

(a) Let x_i and x_{i+1} be two values on the x-axis, where $x_i < x_{i+1}$. Suppose that F is a function that satisfies the hypotheses of the Mean Value Theorem on the interval $[x_i, x_{i+1}]$ and that the derivative of F is f, that is $F' = f$. Show that the conclusion of the Mean Value Theorem is that there exists $x_i^* \in (x_i, x_{i+1})$ such that

$$F(x_{i+1}) - F(x_i) = f(x_i^*)(x_{i+1} - x_i). \tag{1}$$

(b) Consider the interval $[1, 9]$.

i Divide the interval in 4 equal parts, that is divide the interval in 4 subintervals of equal length.

ii Let the 4 subintervals be $[x_0, x_1]$, $[x_1, x_2]$, $[x_2, x_3]$, and $[x_3, x_4]$. Find the values of x_0, x_1, x_2, x_3, and x_4.

iii We can write the difference $F(9) - F(1) = F(x_4) - F(x_0)$ as

$$\begin{aligned}
&F(9) - F(1) = F(x_4) - F(x_0),\\
&= F(x_4) - F(x_3) + F(x_3) - F(x_2) + F(x_2) - F(x_1) + F(x_1) - F(x_0),\\
&= [F(x_4) - F(x_3)] + [F(x_3) - F(x_2)] + [F(x_2) - F(x_1)] + [F(x_1) - F(x_0)],\\
&= [F(x_1) - F(x_0)] + [F(x_2) - F(x_1)] + [F(x_3) - F(x_2)] + [F(x_4) - F(x_3)]
\end{aligned}$$

Since the subintervals have equal length, the difference between two consecutive values x_i is constant. Let the difference be Δx, that is $\Delta x = x_1 - x_0 = x_2 - x_1 = x_3 - x_4 = x_4 - x_3$.

Apply the Mean Value Theorem (see Equation 1) to each bracket in the last expression to justify the following equation

$$F(9) - F(1) = f(x_1^*)\Delta x + f(x_2^*)\Delta x + f(x_3^*)\Delta x + f(x_4^*)\Delta x,$$

where $x_i^* \in (x_{i-1}, x_i)$.

(c) (**Hard**) Consider the interval $[a, b]$ and divide the interval in n subintervals of equal length $\Delta x = (b - a)/n$ and let $x_0, x_1, \ldots, x_{n-1}, x_n$ be the endpoints of the subintervals. Show that

$$F(b) - F(a) = f(x_1^*)\Delta x + f(x_2^*)\Delta x + \ldots + f(x_n^*)\Delta x,$$

where $x_i^* \in (x_{i-1}, x_i)$. As a concluding remark, the limit

$$\lim_{n\to\infty} f(x_1^*)\Delta x + f(x_2^*)\Delta x + \ldots + f(x_n^*)\Delta x$$

is called a *definite integral* and it represents the net area between the function f and the x-axis for $a \leq x \leq b$. What this exercise has shown is that this difficult limit problem can be solved by simply evaluating the net change $F(b) - F(a)$. This result is called the *Fundamental Theorem of Calculus* and is one of the most important theorems in Integral Calculus (which you will be doing next semester), but also one of the most important theorems in mathematics in general.