The Mean Value Theorem

April 20, 2016

1 Introduction

The basic ideas and tools of differential calculus were developed well before 1800, but mathematicians were still struggling to build a rigorous foundation for these ideas. J. L. Lagrange was one of the leaders of the movement to create a solid theory of the derivative. He tried to create his theory around Taylor series expansions during the period 1790-1810, but he was not entirely successful. Cauchy was a pivotal character in building the theory of calculus. He was more successful with his work, mostly published in his 1823 *Calcul Infinitésimal* [C]. He gave a good, near-modern definition of limits, defined the derivative as

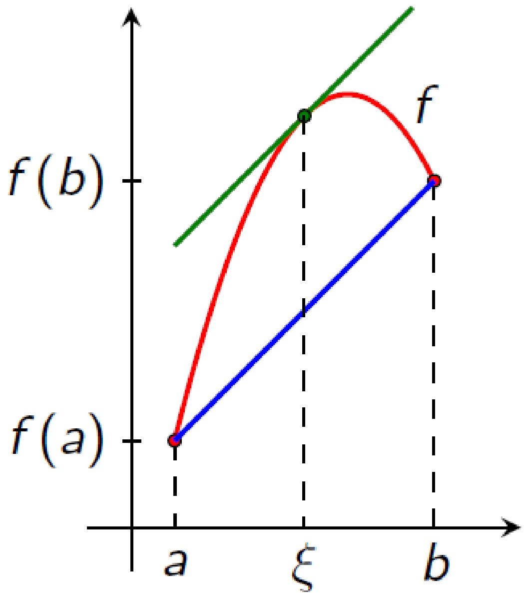
$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

and then used it to prove a number of basic derivative properties, in many places building on Lagrange's efforts.

The Mean Value Theorem (MVT) has come to be recognized as a fundamental result in a modern theory of the differential calculus. As you may recall from introductory calculus courses, under suitable conditions for a function $f : [a, b] \to \mathbb{R}$, we can find a value ξ so that

$$f'(\xi) = \frac{f(b) - f(a)}{b - a} \quad \text{or} \quad f(b) - f(a) = f'(\xi)(b - a) \tag{1}$$

This result is quite plausible from a geometric argument, as the diagram below indicates. The MVT provides a crucial link between change in function values and the derivative at a point. While this results looks pretty clear, it is not so simple to prove analytically, without recourse to a proof by picture. In this project, we will read in Section 2 about Cauchy's efforts to tackle this problem and his proposal for "suitable conditions" on f. Then in Section 3 we will explore a very different approach some forty years later by mathematicians Serret and Bonnet.



Exercise 1 *Use the diagram above to explain why (1) makes sense geometrically.*

The next exercise should give you some initial appreciation of the "suitable conditions" issue for the MVT to be valid.

Exercise 2 *Show that there is no such* ξ *value for (1) with* $f(x) = 2|x-3| + 1,$ *and* $a = 0, b = 4.$ *Interpret this result in terms of the graph of* f*.*

2 Cauchy's Mean Value Theorem

As mentioned in the Introduction, Cauchy defined the derivative as a limit of the difference quotient or ratio, as

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

Building on the work of predecessors such as Lagrange, he knew he needed to state and prove an acceptable version of the MVT in order to expand his rigorous theory of the calculus.

Before we read about Cauchy's MVT and his proof below, we need to know that he uses the term "average" in an unusual way. He calls an **average** of given quantities $b_1, b_2, \ldots, b_n$ a new quantity between the smallest and the largest of those under consideration, and he denotes an average by $M(a_1, a_2, \ldots, a_n)$. For example, any value between 1 and 17 is an average of the values $1, 3, 4, 17$. Cauchy proves several results about his averages, including the following theorem, which we will need to follow his proof below.

Theorem XII. Let $b_1, b_2, b_3, \ldots, b_n$ all have the same sign and $a_1, a_2, \ldots, a_n$ any n quantities. Then

$$\frac{a_1 + a_2 + \cdots + a_n}{b_1 + \cdots + b_n} \text{ is an average of } \frac{a_1}{b_1}, \frac{a_2}{b_2}, \ldots, \frac{a_n}{b_n}.$$

Equivalently,

$$\min_j \frac{a_j}{b_j} \le \frac{a_1 + a_2 + \cdots + a_n}{b_1 + \cdots + b_n} \le \max_j \frac{a_j}{b_j}$$

The following exercise outlines Cauchy's proof of his Theorem XII.

Exercise 3 *Let g be the largest $\frac{a_j}{b_j}$ value and let k denote the minimum $\frac{a_j}{b_j}$ value. Consider the case where $b_j \ge 0$ for all j. (The case where $b_j \le 0$ for all j is similar)*

(a) Show that $gb_j - a_j \ge 0$ and $a_j - kb_j \ge 0$ for all j.

(b) Show that

$$\frac{g\sum b_j - \sum a_j}{\sum b_j} = g - \frac{\sum a_j}{\sum b_j} \ge 0 \quad \textit{and} \quad \frac{\sum a_j - k\sum b_j}{\sum b_j} = \frac{\sum a_j}{\sum b_j} - k \ge 0$$

(c) Show that $k \le \frac{\sum a_j}{\sum b_j} \le g$, which is what we needed to show.

Before Cauchy can prove his MVT, he first proves a crucial inequality. Here is Cauchy, stating and proving what is generally called his Mean Value Inequality theorem.

We now make known a relationship worthy of remark which exists between the derivative $f'(x)$ of any function $f(x)$, and the ratio of the finite differences

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

If, in this ratio, we attribute to x a particular value x_0, and if we make, in addition, $x_0 + h = X$, it will take the form

$$\frac{f(X) - f(x_0)}{X - x_0}.$$

This granted, we will establish the following without difficulty.

THEOREM. If, the function $f(x)$ being continuous between the limits $x = x_0, x = X$, we denote by A the smallest, and by B the largest values that the derivative function receives in this interval, the ratio of the finite differences,

$$\frac{f(X) - f(x_0)}{X - x_0} \tag{2}$$

will necessarily be contained between A and B.

Proof. Denote by δ, ϵ two very small numbers, the first being selected so that, for the numerical values[1] of i less than δ, and for any value of x contained between the limits x_0, X, the ratio

$$\frac{f(x+i) - f(x)}{i}$$

[1] Cauchy uses the term "numerical value" to mean the absolute value, in modern terminology.

always remains greater than $f'(x)-\epsilon$ and less than $f'(x)+\epsilon$. If, between the limits x_0, X, we interpose $n-1$ new values of the variable x, namely,

$$x_1, x_2, \ldots x_{n-1}$$

in a manner to divide the difference $X - x_0$ into elements

$$x_1 - x_0, \quad x_2 - x_1, \quad \ldots, \quad X - x_{n-1},$$

which, all being of the same sign and having numerical values less than δ, the fractions

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0}, \quad \frac{f(x_2) - f(x_1)}{x_2 - x_1}, \quad \ldots, \quad \frac{f(X) - f(x_{n-1})}{X - x_{n-1}} \tag{3}$$

are found contained, the first between the limits $f'(x_0) - \epsilon$, $f'(x_0) + \epsilon$, the second between the limits $f'(x_1) - \epsilon$, $f'(x_1) + \epsilon$, $\ldots$, will all be greater than than the $A - \epsilon$ and less than the quantity $B + \epsilon$. Moreover, the fractions in (3) having denominators of the same sign, if we divide the sum of their numerators by the sum of their denominators, we will obtain an *average* fraction, that is to say, one contained between the smallest and the largest of those that we consider (see Theorem XII). Expression (2), with which this average coincides, will therefore, itself be enclosed between the limits $A - \epsilon, B + \epsilon$; and, since this conclusion remains valid however small the number ϵ, we can say that the expression (2) will be contained between A and B.

Exercise 4 *What assumption is Cauchy making about the derivative function f' in his theorem statement?*

We should celebrate this proof: historians credit it as the first time the symbols ϵ and δ appear in a published analysis proof!

Let's decipher the proof. The first key idea is to divide the interval $[x_0, X]$ sufficiently so that the following claim is valid:

Claim U. The fractions in (3) are found contained, the first between the limits $f'(x_0) - \epsilon$, $f'(x_0) + \epsilon$, the second between the limits $f'(x_1) - \epsilon$, $f'(x_1) + \epsilon$, $\ldots$"

Exercise 5 *Draw a diagram illustrating Cauchy's new values $x_1, x_2, \ldots x_{n-1}$ and Claim U.*

Exercise 6 *What properties of f and/or f' are needed to justify Cauchy's Claim U? Explain.*

Exercise 7 *Explain in your own words and with inequalities, why the fractions in (3) "will all be greater than the $A - \epsilon$ and less than the quantity $B + \epsilon$."*

Cauchy next argues that

$$A - \epsilon < \frac{f(X) - f(x_0)}{X - x_0} < B + \epsilon \tag{4}$$

using his "average" concept.

Exercise 8 *Rewrite Cauchy's proof of (4) in your own words, showing the algebraic details.*

As we saw in Exercise 6, we need one more hypothesis on f' to make Cauchy's proof solid. In 1884, 51 years after Cauchy's proof, the mathematician Giuseppe Peano (1858–1932) offered the following example to demonstrate the problem with Cauchy's proof without an additional hypothesis on f':

$$f(x) = x^2 \sin \frac{1}{x}, \quad x_0 = 0, \quad x_1 = \frac{1}{(2n+1)\pi}, \quad x_2 = \frac{1}{2n\pi}, \quad X = 1 \tag{5}$$

for $n \in \mathbb{N}$ sufficiently large.

Exercise 9 *Find a value for $f(0)$ so that Peano's function f is continuous at 0. Prove your assertion.*

(a) Find a value for $f'(0)$ so that f is differentiable at 0. Prove your assertion with the definition of derivative.

(b) Use standard calculus rules to find $f'(x)$ for $x \neq 0$.

(c) Show that f' is not continuous at $x = 0$.

Exercise 10 *Show that for Peano's f, x_1, x_2 defined in (5) and suitably large n,*

$$f'(x_1) - \epsilon \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'(x_1) + \epsilon$$

is false for a given small ϵ.

Exercise 11 *Use the previous exercises to explain why Cauchy's proof won"t work if f' is not assumed to be continuous on the interval.*

As we can see from this proof by Cauchy, mathematicians in his era had not fully grasped all the subtleties of the dependence of δ on x as well as ϵ in the definition of the derivative.

Exercise 12 *Write a corrected version of Cauchy's mean value inequality theorem with modern notation. Be sure to include the assumption that f' is continuous on the interval and final claim*

$$A \leq \frac{f(X) - f(x_0)}{X - x_0} \leq B\,.$$

Use the exercise results above to give a modern proof of this theorem.

Immediately after Cauchy gives his proof of the mean value inequality theorem, he gives the following result.

Corollary. If the derivative function $f'(x)$ is itself continuous between the limits $x = x_0, x = X$, by passing from one limit to the other, this function will vary in a manner to always remain contained between the two values A and B, and to successively take all the intermediate values. Therefore, any average quantity between A and B will then be a value of $f'(x)$ corresponding to a value of x included between the limits x_0 and $X = x_0 + h$, or to what amounts to the same thing, to a value of x of the form

$$x_0 + \theta h = x_0 + \theta (X - x_0),$$

θ denoting a number less than unity. By applying this remark to expression (2), we will conclude that there exists between the limits 0 and 1, a value of θ that works to satisfy the equation

$$\frac{f(X) - f(x_0)}{X - x_0} = f'(x_0 + \theta (X - x_0))$$

or what amounts to the same thing, the following

$$\frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0 + \theta h) \tag{6}$$

We can recognize this corollary as being a version of the Mean Value Theorem mentioned in the introduction! Recall that in his Mean Value Inequality theorem, Cauchy proved that

$$A \leq \frac{f(X) - f(x_0)}{X - x_0} \leq B \tag{7}$$

which we saw required the continuity of f'.

Exercise 13 *What theorem is Cauchy applying to assert the existence of his θ value? In what interval must the value $c = x_0 + \theta h$ lie?*

Exercise 14 *Write Cauchy's corollary (his mean value theorem) with modern notation.*

While Cauchy's *proof* requires the continuity of f' on the closed interval $[x_0, X]$, there are examples where the MVT holds true even without the continuity of f' on the closed interval $[x_0, X]$. In the next section, we will read about a very different approach to proving an improved version of the MVT, developed well after Cauchy, one which does not require the continuity of f'.

Exercise 15 *Define $f : [0,1] \to \mathbb{R}$ by $f(x) = \sqrt{1 - x^2}$.*

(a) Find a value θ for which equation (6) holds for f on $[0,1]$

(b) Explain why we can't apply Cauchy's version of the MVT to this example.

The following exercises are not needed for the flow of the project, but will sharpen your skills in working with Mean Value theorems, and show the power of Cauchy's version of the MVT.

Exercise 16 *Suppose f is differentiable on $I = [a,b]$ with $f'(x) = 0$ on I. Use Cauchy's MVT to show that f is a constant function on I.*

Sometimes properties of a function or functions are more easily shown by considering an "auxiliary function" defined in terms of the original function(s). We will see an example of this in the next section. The next exercise also demonstrates this technique.

Exercise 17 *Suppose that f and g are differentiable on $I = [a, b]$ with $f'(x) = g'(x)$ on I. Show there is a constant C for which $f(x) = g(x) + C$ on I. Hint: Apply the prior exercise to "auxiliary function" $f - g$.*

Exercise 18 *Explain why the previous exercise justifies the "+C" you tacked onto antiderivatives in your introductory calculus courses.*

3 Bonnet's Version of the Mean Value Theorem.

Cauchy's version of the MVT was the best available for another four decades after he proved it. A very different proof appeared in J. Serret's 1868 *Cours de calcul infinitésimal.* J. Serret (1819-1885) credits his colleague O. Bonnet (1819-1892) with the proof. Both men were accomplished French mathematicians. Here is the theorem and proof from Serret's book.

THEOREM I. *Suppose $f(x)$ is a function of x that is continuous for x values between given limits, and which, for these values, has a specific derivative $f'(x)$. If x_0 and X designate two values of x between these same limits, we have*

$$\frac{f(X) - f(x_0)}{X - x_0} = f'(x_1),$$

x_1 being a value between x_0 and X.

Indeed, the ratio

$$\frac{f(X) - f(x_0)}{X - x_0}$$

has, by hypothesis, a finite value, and if one names this value A, one has

$$[f(X) - AX] - [f(x_0) - Ax_0] = 0 \tag{8}$$

Designating by $\phi(x)$ the function of x defined by formula

$$\phi(x) = [f(x) - Ax] - [f(x_0) - Ax_0], \tag{9}$$

one has, due to equality (9),

$$\phi(x_0) = 0, \quad \phi(X) = 0, \tag{10}$$

so that $\phi(x)$ vanishes at x_0 and X.

Suppose, to fix the ideas, $X > x_0$ and as x increases from x_0 to X; the function $\phi(x)$ is initially zero. If we assume that it is not constantly zero, for the values of x between x_0 and X, it will have to begin to increase through positive values or to decrease by taking negative values, either from $x = x_0$ or from a value of x between x_0 and X. If the values in question are positive, as $\phi(x)$ is continuous and vanishes at X, obviously there will be a value x_1 between x_0 and X for which

$$\phi(x_1)$$

is greater than or at least equal to nearby values

$$\phi(x_1 - h), \quad \phi(x_1 + h),$$

h being an amount as small as we please. If the function $\phi(x)$, ceasing to be zero, takes negative values, the same reasoning proves that there is a value x_1 between x_0 and X for which

$$\phi(x_1)$$

is less than or at most equal to nearby values

$$\phi(x_1 - h), \quad \phi(x_1 + h).$$

So, in both cases, the value of x_1 will be such that the differences

$$\phi(x_1 - h) - \phi(x_1), \quad \phi(x_1 + h) - \phi(x_1),$$

will have the same sign, and, therefore, the ratios

$$\frac{\phi(x_1 - h) - \phi(x_1)}{-h}, \quad \frac{\phi(x_1 + h) - \phi(x_1)}{h}, \tag{11}$$

will have opposite signs.

Note that we do not exclude the hypothesis in which one of the previous ratios evaluates to zero, which requires the function $\phi(x)$ to have the same value for values of x in a finite interval. In particular, if the $\phi(x)$ function is always zero for x valuesbetween x_0 and X, the ratios (11) are both zero.

The ratios (11) tend to the same limit when h tends to zero, because we assume that the $f(x)$ function has a specific derivative, and the same thing occurs, consequently, with respect to $\phi(x)$. Anyway these ratios are of opposite signs, so the limit is zero. Thus we have

$$\lim \frac{\phi(x_1 + h) - \phi(x_1)}{h} = 0, \tag{12}$$

or, because of the equation (9),

$$\lim \left[\frac{f(x_1 + h) - f(x_1)}{h} - A \right] = 0,$$

that is to say

$$A = \lim \frac{f(x_1 + h) - f(x_1)}{h} = f'(x_1) \tag{13}$$

Therefore we have

$$\frac{f(X) - f(x_0)}{X - x_0} = f'(x_1)$$

or

$$f(X) - f(x_0) = (X - x_0) f'(x_1) \tag{14}$$

as was promised.

We assumed $X > x_0$, but as the above formula does not change by permuting the letters x_0, X, it [(14)] is obviously independent of this assumption.

If we set

$$X = x_0 + h,$$

the quantity x_1, between x_0 and $x_0 + h$, may be represented by $x_0 + \theta h$, θ being a quantity between 0 and 1; we can write

$$f(x_0 + h) - f(x_0) = hf'(x_0 + \theta h) \tag{15}$$

REMARK. The preceding demonstration is due to M. Ossian Bonnet. It should be noted that it in no way implies the continuity of the derivative $f'(x)$; it only requires that the derivative exists and has a specific value.

After studying Serret's proof, we can break it into three main parts. First, Serret shows the existence of a special point x_1 for his auxiliary function ϕ. Second, he proves the limit (12) for ϕ at this point x_1. Third, he rewrites this information in terms of the original function f and arrives at his goal, equation (14).

Exercise 19 *To get a sense of this auxiliary function ϕ, consider $f(x) = \sqrt{1 - x^2}$ on $[0, 1]$ from Exercise 15. Sketch the graph of ϕ for this example, and find the value of x_1. What are the properties of ϕ at x_1?*

Now look carefully at the first part of Serret's proof, where he shows the existence of x_1.

Exercise 20 *Show that the equations in (8) and (10) are valid.*

In his theorem statement, Serret states that f is continuous "between" two given limits and that x_1 is "between" x_0 and X, but it is not clear whether he means *strictly* between, or means to include the endpoints. He is similarly vague about where f' exits. Let's try to clarify these issues as we dissect the proof.

Exercise 21 *Using modern terminology and facts about continuous functions, how would you justify the existence of x_1 for ϕ? Does f need to be continuous on $[x_0, X]$ or just on (x_0, X)? Explain.*

In his existence proof of x_1, Serret mentions three cases for ϕ, but only fully explores the case where "the values in question are positive".

Exercise 22 *What are the other two cases for ϕ? Explain why we claim the existence of an x_1 in the open interval (x_0, X) for all three cases.*

Let's record our findings as a Lemma, for which you should be able to provide a modern proof.

Lemma 23 *Suppose $g : [a, b] \to \mathbb{R}$ is continuous on closed interval $[a, b]$. If $g(a) = g(b)$, then g has a local extremum at an interior point $c \in (a, b)$.*

Exercise 24 *Use the discussion above to prove Lemma 23 with a case argument.*

Now that we have the first part of Serret's proof fully clarified, let's turn to the second part, where he proves the limit (12) for ϕ at this point x_1. We recognize this limit as the derivative of ϕ at x_1. Recall that Serret assumes the existence of f' at x values "*between given limits*", which is a bit vague.

Exercise 25 *What derivative laws justify Serret's claim that ϕ has a derivative at x_1? Explain why we only need to require the hypothesis that f' exists on the open interval (x_0, X).*

Exercise 26 *Use Serret's ideas to give a modern proof of the following theorem.*

Theorem 27 *Suppose $g : [a, b] \to \mathbb{R}$. If g has a local extremum at an interior point $c \in (a, b)$ and $g'(c)$ exists, then $g'(c) = 0$.*

The third part of Serret's proof is perhaps the easiest, since he just has to repackage his statements about ϕ into a conclusion about f, which he restates in a few ways. Observe that (15) looks a lot like Cauchy's (6). This plus his final remark that his proof "only requires that the derivative exists" suggest that Serret was well aware of Cauchy's result and its deficiency.

Exercise 28 *For what θ values is (15) valid? Explain.*

Exercise 29 *State a modern version of the MVT based on Serret's work, adjusting his theorem statement as needed.*

Exercise 30 *Show that your modern MVT statement applies to the example $f(x) = \sqrt{1 - x^2}$ on $[0, 1]$ discussed in Exercise 15.*

Exercise 31 *Consider the function $f(x) = x^2 \sin(1/x)$ studied in Exercise 9. Show that the modern MVT can be applied to f on $[0, 1]$ while Cauchy's version cannot.*

Theorem 27 was considered pretty obvious, using geometrical arguments, to mathematicians of the 18th century. This result should be familiar from introductory calculus, where it is used to solve max/min problems. This result is often credited to French mathematician P. Fermat (1601-1665), who gave a version of the result for the special case of quadratic functions in terms of tangent lines in [F], before Newton and Leibniz even developed calculus! The modern version is often called Fermat's theorem, but be aware that Fermat has several famous theorems named after him.

We can combine Lemma 23 and Theorem 27 to get a proof of the following theorem.

Theorem 32 (Rolle) *Suppose $g : [a, b] \to \mathbb{R}$ is continuous on closed interval $[a, b]$ and is differentiable on open interval (a, b). If $g(a) = g(b)$, then there is at least one point c in (a, b) for which $g'(c) = 0$.*

Exercise 33 *Write out a proof of Theorem 32 using Lemma 23 and Theorem 27.*

The first known formal proof of this result was given by Michel Rolle (1652-1719). He proved a version of this theorem in 1690 for polynomials [R]. Interestingly, the theorem was not called "Rolle's" theorem until 1834, when mathematicians were growing more interested in foundations of the calculus.

The MVT is a great tool for proving many interesting results you may recall from introductory calculus, such as the following.

Theorem 34 *Suppose $f : I \to \mathbb{R}$ is differentiable on interval I. Then f is increasing on I if and only if $f'(x) \geq 0$ for all $x \in I$.*

Exercise 35 *Use the MVT to prove this theorem in one direction. Use the definition of derivative in the other direction.*

Exercise 36 *Suppose g is differentiable and g' is continuous on $[a, b]$. Show that there is a constant K for which $|g(x) - g(y)| \leq K|x - y|$ for all $x, y \in [a, b]$.*

4 Conclusion

Mathematicians learned how to define and apply the derivative well before they put the concept on a firm mathematical foundation. Important results such as Rolle's Theorem and the MVT turn out to be more difficult to prove properly than one might think. The Mean Value Theorem is fundamental, and can be used to prove many results from an introductory Calculus course.

References

[C] Cauchy, A.L. 1823. *Résumé des leçons données à l'École royale polytechnique sur le calcul infinitésimal.* Paris: De Bure.

[E] Euler, L. 1755. *Institutiones calculi differentialis* , St Petersburg.

[F] Fermat, P. 1891-1912, *Oeuvres de Fermat.* Gauthier-Villars, Paris.

[LH] L'Hôpital, G. 1696. *Analyse des infiniment petits pour l'intelligence des lignes courbes.* Paris : Montalant.

[N] Newton, I. 1704, *Tractatus de quadratura curvarum.* London.

[R] Rolle, Michel (1690). Traité d'Algebre. E. Michallet, Paris.

[S] J. Serret, 1868. *Cours de calcul infinitésimal,* Paris.