2017

The Mean Value Theorem

Dave Ruch
ruch@msudenver.edu

The Mean Value Theorem

David Ruch*

November 22, 2021

1 Introduction

The basic ideas and tools of differential calculus were developed well before 1800, but mathematicians of that time were still struggling to build a rigorous foundation for these ideas.[1] Joseph-Louis Lagrange (1736–1813) was one of the leaders of the movement to create a solid theory of the derivative. He tried to create his theory of the derivative around Taylor series expansions during the period 1790–1810, but he was not entirely successful. Augustin-Louis Cauchy (1789–1857), a pivotal character in building the theory of calculus, also built a theory of the derivative early in the 1800s. He was more successful than Lagrange, and published most of this work in his 1823 textbook on differential and integral calculus [Cauchy, 1823].[2] Cauchy gave a good, near-modern definition of limits, defining 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 efforts by Lagrange, Sylvestre François LaCroix (1765–1843) and others.

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

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

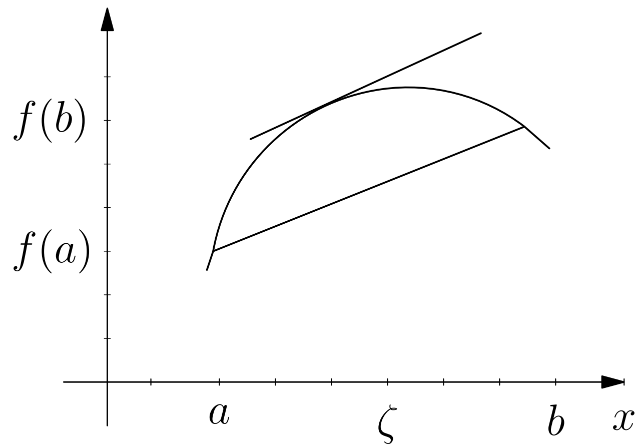
This result is quite plausible from a geometric argument, as the diagram below indicates. The MVT provides a crucial link between change in function values and the derivative at a point. While this result looks pretty clear, it is not so simple to prove analytically, without recourse to a proof by picture. 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 Joseph Serret (1819–1885) and Pierre Ossian Bonnet (1819–1892).

---

*Department of Mathematical and Computer Sciences, Metropolitan State University of Denver, Denver, CO; emailruch@msudenver.edu.

[1] See the project "Why be so Critical" [Barnett, 2016] and the article [Lützen, 2003] for more on this.

[2] All Cauchy source excerpts in this project are from the Seventh Lecture in this textbook, which is entitled *Résumé des leçons données à l'École royale polytechnique sur le calcul infinitésimal* (*Summary of lessons giving on the calculus at the École Royale Polytechnique)*, using the English translation in [Cates, 2012].



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

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

Task 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, which Cauchy himself referred to as the "ratio of the finite differences:"

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

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 Cauchy proved his MVT, he first proposed a crucial inequality. Here is Cauchy, stating 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[3] 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.

It turns out that Cauchy made some strong (perhaps hidden) assumptions in his proof of this theorem.

Task 3 Try applying Cauchy's Mean Value Inequality theorem to the function $f(x) = \sqrt{x}$ on the interval with endpoints $x_0 = 0$, $X = 4$. What are A, B for this example? Identify any ambiguities in Cauchy's theorem statement.

The first part of Cauchy's proof is given below. Read it carefully, thinking about the assumptions that he seems to have been making.

Proof. Denote by δ, ϵ two very small numbers, the first being selected so that, for the numerical values[4] of i less than δ, and for any value of x contained between the limits x_0, X, the ratio

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

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 $f'(x_1) - \epsilon$, $f'(x_1) + \epsilon$, &c ..., will all be greater than the quantity $A - \epsilon$ and less than the quantity $B + \epsilon$.

[3] Cauchy meant A and B are "output" function values of the derivative.

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

We should celebrate this proof: historians credit it as the first time the symbols ϵ and δ appear in a published analysis proof! However, there are some subtle problems with Cauchy's work from a modern viewpoint.

Task 4 Based on Task 3 and your reading of the proof, what assumption(s) was Cauchy making about the derivative function f' in his theorem statement and proof?

Let's decipher the proof. The very first sentence is especially important and subtle.

Task 5 How do you think Cauchy would have justified the very first sentence of the proof? What does the value of δ seem to depend on?

Task 6 Rewrite the very first sentence of Cauchy's proof using modern quantifiers to clarify the dependencies between ϵ, δ, i and x.

As we saw in Task 3, we need stronger hypotheses on f' than Cauchy explicitly gave in his Theorem statement to make Cauchy's proof solid by modern standards. In Task 5, we saw that Cauchy also set up certain dependencies between variables in his proof that go beyond the hypotheses in his Theorem statement.

In 1884, 61 years after Cauchy's proof was published, the mathematician Giuseppe Peano (1858–1932) offered the following example [Peano, 1884][5], which demonstrates a problem in Cauchy's first sentence without the additional hypothesis implied by the first sentence of Cauchy's proof :

$$f(x) = x^2 \sin \frac{1}{x}. \tag{4}$$

Tasks 7, 8, 9 below will reveal Peano's concern. We first record the modern version of Cauchy's first sentence in his proof, which asserts a sort of *uniform differentiability* property for f on the interval. We will call it Property UD (using "UD" to indicate "**u**niform **d**ifferentiability"):

> **Property UD.** For all $\epsilon > 0$, there exists $\delta > 0$ such that if $x, x+i \in [x_0, X]$ and $|i| < \delta$ then
> $$\left| \frac{f(x+i) - f(x)}{i} - f'(x) \right| < \epsilon.$$

Task 7 This task analyzes the behavior of Peano's function f at $x = 0$. Sketching a graph of f may give you some insights.

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

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

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

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

[5] The mathematician Gaston Darboux (1843-1917) used this function example in some 1875 private correspondence, but it is unlikely that Peano knew of Darboux's work since that correspondence wasn't published until the 1980s. It is thus generally believed that Peano devised his example independently.

Task 8 We are now ready for Peano's key observation. Suppose $n \in \mathbb{N}$. Choose interval endpoints $x_0 = 0, X = 1$. Consider points

$$x_1 = \frac{1}{(2n+1)\,\pi}, \quad x_2 = \frac{1}{2n\pi}.$$

Show that for Peano's f defined in (4) and Task 7, and arbitrary ϵ with $0 < \epsilon < 1$, the following statement is false:

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

Task 9 Use the previous tasks to explain why Peano's example function does not have Property UD even though its derivative exists and is bounded on the interval $[x_0, X] = [0, 1]$.

Hint: For any ϵ with $0 < \epsilon < 1$ and any $\delta > 0$, first show you can choose n so that $|x_1 - x_2| < \delta$.

What about the problematic Property UD in Cauchy's proof? This uniform differentiability property for f turns out to be valid once we add the crucial hypothesis that f' is *uniformly continuous* on the interval. However, the proof of Property UD for a function with uniformly continuous derivative is surprisingly difficult without using the MVT, so we omit it.

The next key idea in Cauchy's proof was to divide the interval $[x_0, X]$ finely enough so that he could apply Property UD. After this, Cauchy used some elaborate algebra and bounding[6] to show that

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

which completes the proof of his Mean Value Inequality theorem — with the added hypothesis that f' is *uniformly continuous* on the interval. As we can see from this proof by Cauchy, all the subtleties of the dependence of δ on x as well as ϵ in the definition of the derivative had not yet emerged at the time he wrote his proof. More on the history of these issues can be found in [Grabiner, 2010].[7]

Task 10 Write a corrected version of Cauchy's Mean Value Inequality theorem with modern notation. Be sure to include both (i) the assumption that f' is uniformly continuous on the interval, and (ii) his conclusion

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

Then add an explanation for why A and B exist as finite values under the hypothesis f' is uniformly continuous on the interval.

Immediately after Cauchy gave his proof of the Mean Value Inequality theorem, he stated the following result.

[6] This argument is omitted from the project and relies on Cauchy's Note II from his *Cours d'analyse*.

[7] The student project "Rigorous Debates over Debatable Rigor: Monster Functions in Introductory Analysis" [Barnett, 2017] also explores the emergence of these subtleties through excerpts the somewhat later work in analysis done by from Gaston Darboux (1842–1917), independently of Peano's work.

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 number[8] 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

$$\frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0 + \theta h). \tag{5}$$

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

$$A \leq \frac{f(X) - f(x_0)}{X - x_0} \leq B. \tag{6}$$

Task 11 What theorem was Cauchy applying to assert the existence of his θ value? In what interval must the value $c = x_0 + \theta h$ lie?

Task 12 Write Cauchy's corollary (his mean value theorem) with modern notation.

While Cauchy's *proof* required the uniform 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 uniform continuity of f'.

Task 13 Define $f : [0, 1] \to \mathbb{R}$ by $f(x) = \sqrt{1 - x^2}$.

(a) Find a value θ for which equation (5) holds for f on $[0, 1]$.

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

[8] Cauchy actually used the term "average quantity" in place of "number" for purposes of his proof.

The following tasks 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.

Task 14 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 task also demonstrates this technique.

Task 15 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 task to "auxiliary function" $f - g$.

Task 16 Explain why the previous task justifies the "$+C$" you tacked onto antiderivatives in your introductory calculus courses.

Task 17 Define $f(x) = \sec^2 x$ and $g(x) = \tan^2 x$.

(a) Define an appropriate interval I and use Task 15 to show there is a constant C for which $f(x) = g(x) + C$ on I.

(b) Use some trigonometry to find the value of C for this example. For what x values is $f(x) = g(x) + C$ known to be true for this example?

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 different version with weaker hypotheses and a very different proof appeared in Joseph Serret's 1868 *Cours de calcul différentiel et intégral* (*Course on differential and integral calculus*), [Serret, 1868]. Serret (1819–1885) credited his colleague Pierre Ossian Bonnet (1819–1892) with the proof. Both men were accomplished French mathematicians. Here is the theorem and proof from Serret's book. As you read this proof, try to identify the three separate "subclaims" that Serret proved to complete the proof.[9]

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.

[9] All translations of Serret excerpts in this project were prepared by the project author, 2017.

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

$$[f(X) - AX] - [f(x_0) - Ax_0] = 0. \tag{7}$$

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

$$\phi(x) = [f(x) - Ax] - [f(x_0) - Ax_0], \tag{8}$$

one has, due to equality (8),

$$\phi(x_0) = 0, \quad \phi(X) = 0, \tag{9}$$

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.[10] 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

$$\frac{\phi(x_1 - h) - \phi(x_1)}{-h}, \quad \frac{\phi(x_1 + h) - \phi(x_1)}{h}, \tag{10}$$

will have opposite signs.

Note that we do not exclude the hypothesis in which one of the previous ratios evaluates

[10] Serret assumed that h is positive throughout his proof.

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 values between x_0 and X, the ratios (10) are both zero.

The ratios (10) 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

$$\lim \frac{\phi(x_1+h)-\phi(x_1)}{h}=0, \tag{11}$$

or, because of the equation (8),

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

that is to say

$$A=\lim \frac{f(x_1+h)-f(x_1)}{h}=f'(x_1) \tag{12}$$

Therefore we have

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

or

$$f(X)-f(x_0)=(X-x_0)f'(x_1) \tag{13}$$

as was promised.

We assumed $X>x_0$, but as the above formula does not change by permuting the letters x_0, X, it [(13)] 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

$$f(x_0+h)-f(x_0)=hf'(x_0+\theta h) \tag{14}$$

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.

Task 18 Compare Serret's MVT statement with Cauchy's, Task 12. In what way is Serret's theorem superior to Cauchy's? Hint: Peek at Task 13.

Task 19 To begin analyzing Serret's proof, first outline the proof's main features. Try to identify three distinct major steps.

Task 20 To get a sense of Serret's "auxiliary" function ϕ, consider $f(x)=\sqrt{1-x^2}$ on $[0,1]$ from Task 13. 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 showed the existence of x_1.

Task 21 Show that the equations in (7) and (9) are valid.

In his theorem statement, Serret stated 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 meant *strictly* between, or meant to include the endpoints. He was similarly vague about where f' exists. Let's try to clarify these issues as we dissect the proof.

Task 22 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 for x_1, Serret mentioned three cases for ϕ, one being the case where "the values in question are positive".

Task 23 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 1. *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* $c \in (a, b)$.

Task 24 Use Serret's ideas and the discussion above to give a modern proof of Lemma 1 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 proved the limit (11) for ϕ at this point x_1. We recognize this limit as the derivative of ϕ at x_1. Recall that Serret assumed the existence of f' at x values "*between given limits*,"which is a bit vague.

Task 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).

Task 26 Use Serret's ideas to give a modern proof of the following Theorem 2:

Theorem 2. *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 had to repackage his statements about ϕ into a conclusion about f, which he restated in a few ways. Observe that (14) looks a lot like Cauchy's (5). This plus Serret's final remark that his own proof "only requires that the derivative exists" suggests that Serret was aware of Cauchy's result and its deficiency.

Task 27 For what θ values is (14) valid? Explain.

Task 28 State a modern version of the MVT based on Serret's work, adjusting his theorem statement as needed.

Task 29 Show that your modern MVT statement applies to the example $f(x) = \sqrt{1-x^2}$ on $[0,1]$ discussed in Task 13.

Task 30 Consider the function $f(x) = \begin{cases} x^2 \sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases}$ studied in Task 7. Show that the hypotheses of the modern MVT are met for f on $[0,1]$ while the hypotheses for Cauchy's version are not met for f on $[0,1]$.

Theorem 2 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. It is often credited to French mathematician Pierre Fermat (1601–1665), who first gave a version of the result in [Fermat, 1636], 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 1 and Theorem 2 to get a proof of the following theorem.

Theorem 3 (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 $g'(c) = 0$.*

Task 31 Write out a proof of Theorem 3 using Lemma 1 and Theorem 2.

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 [Rolle, 1690]. Interestingly, the theorem was not called "Rolle's Theorem" until the 1830s, when mathematicians were growing more interested in the foundations of the calculus [Cajori, 1919].

Task 32 Sketch a diagram demonstrating Rolle's theorem, labeling the point c. Next sketch a diagram with the graph of the ϕ function from Serret's proof with $a = x_0$ and $b = X$. Label the point x_1 from Serret's proof. Compare and contrast these diagrams with the diagram for the MVT given in the project introduction.

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

Theorem 4. *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 $x \in I$.*

Task 33 Use the MVT to prove Theorem 4 in one direction. Use the definition of derivative and Serret's ideas with (10) in the other direction.

Recall from Section 2 that Cauchy made the very strong assumption we called Property UD in the first sentence of his proof of the Mean Value Inequality, where Property UD necessarily holds provided f' is uniformly continuous on the interval. While this is difficult to prove directly, it can be proved more easily using the Mean Value Theorem. The next task confirms this.

Task 34 Suppose that $f : [a,b] \to \mathbb{R}$ and f' is continuous on $[a,b]$. Prove that for all $\epsilon > 0$ there exists a $\delta > 0$ for which

$$\text{if } x, u \in [a,b] \text{ and } 0 < |x-u| < \delta, \text{ then } \quad \left| f'(x) - \frac{f(x) - f(u)}{x-u} \right| < \epsilon. \tag{15}$$

Task 35 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 $x, y \in [a,b]$.

Task 36 Many difficult equations can be solved approximately by putting them into the form $g(x) = x$ and building a sequence of approximate solutions (x_n) defined by $x_n = g(x_{n-1})$. It can proved that this sequence (x_n) will converge to a solution of the equation $g(x) = x$ if (x_n) is **contractive**, which means that there is a constant C, $0 < C < 1$, for which

$$|x_{n+1} - x_n| \leq C|x_n - x_{n-1}| \quad \text{for all } n \in \mathbb{N}.$$

Suppose that the function $g : [a,b] \to [a,b]$ is differentiable and $|g'(x)| \leq 0.99$ on $[a,b]$. Let $x_1 \in [a,b]$ and define sequence (x_n) by $x_n = g(x_{n-1})$. Use the MVT to prove that (x_n) is contractive.

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 analytically than one might think. The Mean Value Theorem is fundamental, and can be used to prove many results from an introductory Calculus course.

As mathematicians delved into the ramifications of rigorous definitions of continuity and derivative during the 1800s, they found strange behavior in "monster function" examples such as the one proposed by Peano that we examined in Tasks 7–9, and again in 30. This function illustrates ways in which our intuition can go awry, as seen in the next task.

Task 37 In Serret's proof, he argued that the continuous function ϕ will have to begin to increase through positive values or decrease through negative values from its zero value at x_0. While this seems intuitively appealing, consider Peano's function f defined for $x \neq 0$ by $f(x) = x^2 \sin \dfrac{1}{x}$ and setting $f(0) = 0$ so f is continuous and differentiable at 0.

(a) Do you think Peano's function f begins to increase through positive values, or begins to decrease through negative values from its zero value at 0? Explain your answer.

(b) In our proof, we avoided appealing to this line of thinking from Serret's proof. How did we manage this? Hint: Review your task work around Lemma 1.

The historical story of the Mean Value Theorem is long and complicated, as we have seen. If you are interested in the history of mathematics, you might enjoy the following task. Talk to your instructor in advance about the appropriate answer length and desired depth of detail.

Task 38 Discuss the historical evolution of the Mean Value Theorem and the key mathematical ideas involved. You should discuss both precise mathematical statements of concepts and theorems, as well as proof methods and rigor.

References

J. H. Barnett. Why be so Critical? Nineteenth Century Mathematics and the Origins of Analysis. 2016. Available at `https://digitalcommons.ursinus.edu/triumphs\_analysis/1/`.

J. H. Barnett. Rigorous Debates over Debatable Rigor: Monster Functions in Introductory Analysis. 2017. Available at `https://digitalcommons.ursinus.edu/triumphs\_analysis/10/`.

U. Bottazzini. *The Higher Calculus: A History of Real and Complex Analysis from Euler to Weierstrass*. Springer-Verlag, New York, 1986.

F. Cajori. *A History of Mathematics*. The Macmillan Company, New York, 1919.

D. Cates. *Cauchy's Calcul Infinitésimal – A Complete English Translation*. Fairview Academic Press, Walnut Creek CA, 2012.

A.-L. Cauchy. *Résumé des leçons données à L'Ecole Royale Polytechnique sur le calcul infinitésimal (Summary of lessons given on the infinitesimal calculus at the Royal Polytechnic School)*. De Bure, Paris, 1823. English translation by D. Cates in [Cates, 2012].

P. Fermat. Methodus ad disquirendam maximam et minimam (Method for the study of maxima and minima). In C. Henry and P. Tannery, editors, *Oeuvres de Fermat*, volume 1, pages 133–136. Gauthier-Villars, Paris, 1636. First presentation of Fermat's method for finding maxima and minima, written as a letter to René Descartes (1596–1615) that was published posthumously (in French translation) in 1891 in the cited volume of Fermat's *Oeuvres* (or *Collected Works*).

J. Grabiner. *A Historian Looks Back: The Calculus as Algebra and Selected Writings*. Spectrum. Mathematical Association of America, Washington DC, 2010.

H. N. Jahnke, editor. *History of Analysis*. American Mathematical Society, Providence, RI, 2003.

J. Lützen. The Foundation of Analysis in the 19th Century. In H. N. Jahnke, editor, *A History of Analysis*, pages 155–196. American Mathematical Society, Providence, RI, 2003.

G. Peano. Correspondence, Tome 3. *Nouvelles Annales de Mathématiques*, pages 45–47, 1884.

J. Plante. A proof of Bonnet's Version of the Mean Value Theorem by Methods of Cauchy. *American Mathematical Monthly*, 124:269–273, 2017.

M. Rolle. *Traité d'Algebre*. E. Michallet, Paris, 1690.

J. Serret. *Cours de calcul différentiel et intégral (Course on differential and integral calculus)*, volume 1. Gauthier-Villars, Paris, 1868.

Notes to Instructor

PSP Content: Topics and Goals

This Primary Source Project (PSP) is designed for a course in Real Analysis. The first part of the project is over Cauchy's version of the Mean Value Theorem. Cauchy's proof was very clever but had some hidden assumptions.

The second part of the project focuses on the standard modern version seen in Introductory Analysis texts. This standard version of this proof is apparently due to Bonnet, and recorded in his student Serret's 1868 *Cours de calcul infinitésimal*.

The specific content goals of this project are to:

1. Study Cauchy's proof of the Mean Value Theorem.
2. Analyze the assumptions in Cauchy's proof and their necessity for his proof method. In particular, recognize and appreciate the relevance of *uniform* continuity and differentiability.
3. Analyze Serret's proof of the Mean Value Theorem.
4. Use Serret's proof to prove Fermat's Interior Extremum Theorem and Rolle's Theorem.
5. Use the Mean Value Theorem for applications.

Student Prerequisites

The project is written for a course in Real Analysis with the assumption that students have done a rigorous study of limits, continuity and the derivative. The concept of uniform continuity appears in a crucial way in Cauchy's proof. Instructors could use this as an opportunity to motivate and define uniform continuity if students have not seen it before.

PSP Design and Task Commentary

This project consists of four main sections, followed by a brief conclusion.

Section 1 Introduction

Some historical perspective for the MVT is given. A task designed to create some appreciation of suitable conditions under which the MVT conclusion holds is included.

Section 2 Cauchy

Task 5 is very important. It is a great chance to revisit the dependence of δ on x as well as ϵ in the definition of the derivative. The term "uniform continuity" is then explicitly mentioned later. The concept of *uniform* continuity needs to be defined or recalled here. The notion of uniform differentiability is addressed in some, but not all, modern texts. There is some debate among mathematical historians about Cauchy's treatment and understanding of global versus local behavior. See articles in [Grabiner, 2010] and [Jahnke, 2003] for more on this.

Section 3 Bonnet

Most standard treatments of the MVT first prove Fermat's theorem (if you have a local extremum, the derivative is zero there). This result is in turn used to prove Rolle's Theorem. Finally, Rolle's theorem is applied to a clever auxiliary function ϕ to obtain the MVT. Serret's proof of the MVT is a bit surprising. He immediately defined the crucial auxiliary function ϕ. Serret then proved Fermat's theorem for ϕ. Next he proved Rolle's theorem for ϕ. Finally he showed this gives us the MVT for the original function. Having students read this proof has some pedagogical advantages, since it assumes very little and students won't be able to track it word for word with a modern textbook or online search. Serret was a bit vague with his inequalities (strict vs. inclusive), which will force some careful reading. But once students unpack everything, they will have the standard big three results.

Serret didn't explicitly state that f has an extremum at x_1, nor was he explicit about x_1 being *strictly* between x_0 and X. The tasks are designed to clarify these points.

In Serret's proof, he argued that the function ϕ will " have to begin to increase through positive values."This is intuitively appealing, but Peano's function illustrates the flaws in this thinking. The PSP sidesteps this issue until the Conclusion, in order to keep Serret's long proof on track.

The last tasks can be used as an introduction to Lipschitz functions and convergence of fixed point iterations.

Suggestions for Classroom Implementation

This is roughly a two-week project under the following methodology (basically David Pengelley's "A, B, C" method (described on his website `https://web.nmsu.edu/~davidp/`):

1. Students do some advanced reading and light preparatory tasks before each class. This should be counted as part of the project grade to ensure students take it seriously. Be careful not to get carried away with the tasks or your grading load will get out of hand! Some instructor have students write questions or summaries based on the reading.

2. Class time is largely dedicated to students working in groups on the project - reading the material and working tasks. As they work through the project, the instructor circulates through the groups asking questions and giving hints or explanations as needed. Occasional student presentations may be appropriate. Occasional full class guided discussions may be appropriate, particularly for the beginning and end of class, and for difficult sections of the project. I have found that a "participation" grade suffices for this component of the student work. Some instructors collect the work. If a student misses class, I have them write up solutions to the tasks they missed. This is usually a good incentive not to miss class!

3. Some tasks are assigned for students to do and write up outside of class. Careful grading of these tasks is very useful, both to students and faculty. The time spent grading can replace time an instructor might otherwise spend preparing for a lecture.

If time does not permit a full implementation with this methodology, instructors can use more class time for guided discussion and less group work for difficult parts of the project.

LaTeX code of this entire PSP is available from the author by request to facilitate preparation of 'in-class task sheets' based on tasks included in the project. The PSP itself can also be modified by instructors as desired to better suit their goals for the course.

Sample Implementation Schedule (based on a 50-minute class period)

Full implementation of the project can be accomplished in 6 class days, as outlined below.

Students read through the introductory material and do Tasks 1, 2 before the first class. After discussing their results at the beginning of Class 1, students read the first Cauchy excerpt, do Task 3, read the second Cauchy excerpt and work on Tasks 4, 5.

As preparation for Class 2, students do Task 6. After discussing their results at the beginning of Class 2, students read about Peano's example and uniform differentiability, then work on and discuss Tasks 7–9.

As preparation for Class 3, students do Task 10. After discussing their results at the beginning of Class 3, students read the third "Corollary" Cauchy excerpt and work on Tasks 11–14. Tasks 15–17 can be assigned for homework.

As preparation for Class 4, students read Serret's MVT theorem statement and proof, and do Tasks 18, 19. After discussing their results at the beginning of Class 4, students work on Tasks 20-22.

As preparation for Class 5, students do Tasks 23, 24. After discussing their results at the beginning of Class 5, students work on Tasks 25–28. Tasks 29, 30 can be assigned for homework.

As preparation for Class 6, students do Tasks 31, 32. After discussing their results at the beginning of Class 6, students work on Tasks 33–36 as time permits. The remainder can be assigned for homework. Students should read the PSP conclusion and discuss ideas for Tasks 37, 38, which can be assigned for homework.

Connections to other Primary Source Projects

The following additional projects based on primary sources are also freely available for use in an introductory real analysis course; the PSP author name for each is listed parenthetically, along with the project topic if this is not evident from the PSP title. Shorter PSPs that can be completed in at most 2 class periods are designated with an asterisk (*). Classroom-ready versions of the last two projects listed can be downloaded from `https://digitalcommons.ursinus.edu/triumphs\_topology`; all other listed projects are available at `https://digitalcommons.ursinus.edu/triumphs\_analysis`.

- *Why be so Critical? 19th Century Mathematics and the Origins of Analysis** (Janet Heine Barnett)
- *Investigations into Bolzano's Bounded Set Theorem* (David Ruch)
- *Stitching Dedekind Cuts to Construct the Real Numbers* (Michael Saclolo)
  Also suitable for use in an Introduction to Proofs course.
- *Investigations Into d'Alembert's Definition of Limit** (David Ruch)
  A second version of this project suitable for use in a Calculus 2 course is also available.
- *Bolzano on Continuity and the Intermediate Value Theorem* (David Ruch)
- *An Introduction to a Rigorous Definition of Derivative* (David Ruch)
- *Rigorous Debates over Debatable Rigor: Monster Functions in Real* (Janet Heine Barnett; properties of derivatives, Intermediate Value Property)
- *The Definite Integrals of Cauchy and Riemann* (David Ruch)
- *Henri Lebesgue and the Development of the Integral Concept** (Janet Heine Barnett)
- *Euler's Rediscovery of e** (David Ruch; sequence convergence, series & sequence expressions for e)

- *Abel and Cauchy on a Rigorous Approach to Infinite Series* (David Ruch)
- *The Cantor set before Cantor** (Nicholas A. Scoville)
  Also suitable for use in a course on topology.
- *Topology from Analysis** (Nicholas A. Scoville)
  Also suitable for use in a course on topology.

Recommendations for Further Reading

The articles in [Jahnke, 2003], [Grabiner, 2010] and [Bottazzini, 1986] give some perspective on other 19th century works in analysis.

A recent article [Plante, 2017] gives a different proof of Bonnet's Mean Value Theorem using methods available to Cauchy, which might be of interest. The PSP author used ideas of Plante's to find a rather long and technical proof that Property UD holds for a function with uniformly continuous derivative, without using the Mean Value Theorem (see reference below). In some sense this "saves" Cauchy's Mean Value Theorem proof, but the details are long and not really fitting for the PSP. Interested readers can find these details in the following paper:

- Ruch, D. (2018). *Using Primary Source Projects to Learn Real Analysis, with Investigations from Cauchy and Abel.* Proceedings of the Conference on History of Mathematics and Teaching of Mathematics, University of Miskolc, Miskolc, Hungary. Available at:
  `http://www.uni-miskolc.hu/~matsefi/HMTM_2018_CD/`

Another good source for further reading is Boyer's classic:

- Boyer, C. (1959). *The History of The Calculus and Its Conceptual Development.* New York: Dover Publications. First published under the title *The Concepts of Calculus, A Critical and Historical Discussion of the Derivative and the Integral*, Hafner Publishing Company (1949).

Acknowledgments

The development of this student project has been partially supported by the TRansforming Instruction in Undergraduate Mathematics via Primary Historical Sources (TRIUMPHS) Program with funding from the National Science Foundation's Improving Undergraduate STEM Education Program under Grant Nos. 1523494, 1523561, 1523747, 1523753, 1523898, 1524065, and 1524098. Any opinions, findings, and conclusions or recommendations expressed in this project are those of the author and do not necessarily represent the views of the National Science Foundation.

For more information about TRIUMPHS, visit `https://blogs.ursinus.edu/triumphs/`.