Revisit Mean Value, Cauchy Mean Value and Lagrange Remainder Theorems

Wei-Chi Yang
e-mail: wyang@radford.edu
Department of Mathematics and Statistics
Radford University 24142 USA

Abstract

We demonstrate how evolving technological tools have led to advances in the teaching and learning of mathematics and in the production of mathematical research. In particular, the integration of dynamic geometry software (DGS) with a computer algebra system (CAS) has led to new methods for solving existing problems and has revealed the existence of new concepts waiting to be discovered. We demonstrate also how DGS software frequently provides the crucial insights and accessibility that motivate conjectures that can be proved analytically with the help of a CAS. There are two video clips which give some geometric insights of how we prove Mean Value Theorem and Cauchy Mean Value Theorem, they are located respectively at

http://mathandtech.org/eJMT_Yang_iss2_07/MVT/MVT.html

and

http://mathandtech.org/eJMT_Yang_iss2_07/CMV/CMV.html.

1 Introduction

We first give a geometric interpretation of how Mean the Value Theorem is proved and simulate the graph to which we normally apply the Rolle's Theorem. Next, we give a geometric description of how the Cauchy Mean-Value is stated and shed some light on how we can arrive at the function to which Rolle's Theorem is applied to yield the Cauchy Mean Value Theorem holds. We also show how to solve numerically for a number that satisfies the conclusion of the theorem.

Finally, we give an alternative interpretation of the Lagrange Remainder Theorem. This interpretation allows us to find and solve numerically for the number whose existence is guaranteed by the Theorem. It also allows us to approximate the remainder term for a given function.

2 Geometric Interpretation of Mean Value Theorem

The Mean Value Theorem can be stated as follows:

Theorem 1 *Suppose the function $f : [a, b] \to R$ is continuous on $[a, b]$ and differentiable on (a, b). Then there is a point x_0 in (a, b) at which*

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$

To prove this theorem, in many traditional text books, one introduces the function h defined at each number x by the following equation

$$h(x) = f(x) - f(a) - \left(\frac{f(b) - f(a)}{b - a}\right)(x - a). \qquad (1)$$

Then we use the fact that h satisfies the conditions for Rolle's theorem to deduce that there is a point c in (a, b) such that $h'(c) = 0$, and the Mean Value Theorem follows.

We give a geometric motivation on the graph of $y = h(x)$ and how the graph of $y = h(x)$ can be simulated with the help of a dynamic geometry system such as [ClassPad]. We consider a nice smooth function $f(x) = \cos(x)$ (i.e. satisfying the conditions of the Mean Value theorem) over the interval covering $(a, b) = (-\frac{\pi}{2}, 0.725)$ shown in Figure 1 below. We connect the line segment AB, where $A = (-\frac{\pi}{2}.0)$ and $B = (0.725, f(0.725))$ lying on $y = f(x)$ and ask the following question:

If we rotate line segment AB (while AB is attached to the graph of the function) *so that AB becomes a horizontal line segment, how would the graph of the original function appear?*

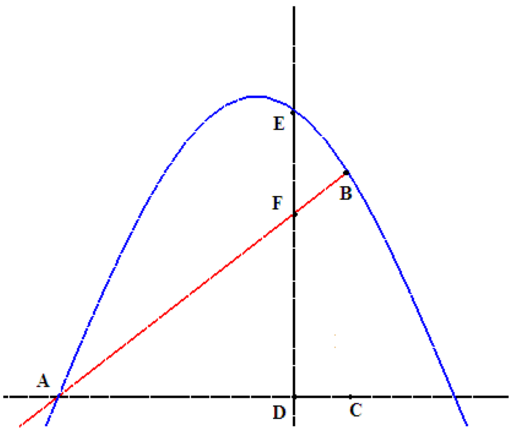


Figure 1. The graph of a function and a chord

We summarize the constructions below and we refer readers to a video clip to glimpse how dynamic geometry and computer algebra system play roles here. In this example, we select f and the line segment AB so that $f(x) > AB$, where $x \in [-\frac{\pi}{2}, 0.725]$ for demonstration purpose:

- Step 1. Construct a point D on AC and animate the point D along the line segment AC.
- Step 2. Construct the line passing through D and perpendicular to AC and intersect $y = f(x)$ and AB at E and F respectively.

- Step 3. We collect the (x, y) coordinate for the point D and the distance EF, We show the x-value of D in the first column and distance EF in the second column below:

-1.570796327	9.999833334E-3
-1.523943341	0.02143457875
-1.477090354	0.05278812417
-1.430237368	0.08395816613
-1.383384382	0.1148424706
-1.336531395	0.1453394306
-1.289678409	0.1753482893
-1.242825423	0.2047693614
-1.195972437	0.2335042513
-1.14911945	0.2614560697
-1.102266464	0.2885296459
-1.055413478	0.3146317368
-1.008560492	0.3396712314
-0.9617075054	0.363559351
-0.9148545191	0.386209844
-0.8680015329	0.4075391749
-0.8211485466	0.4274667082
-0.7742955604	0.4459148848
-0.7274425741	0.4628093926
-0.6805895878	0.4780793296
-0.6337366016	0.4916573592
-0.5868836153	0.5034798584
-0.5400306291	0.513487057
-0.4931776428	0.5216231693
-0.4463246565	0.5278365161
-0.3994716703	0.5320796383
-0.352618684	0.534309401
-0.3057656977	0.5344870885
-0.2589127115	0.5325784885
-0.2120597252	0.5285539682
-0.165206739	0.5223885385
-0.1183537527	0.5140619093
-0.07150076644	0.5035585338
-0.02464778018	0.4908676429
0.02220520608	0.4759832684
0.06905819234	0.4589042566
0.1159111786	0.4396342706
0.1627641649	0.418181782
0.2096171511	0.3945600528
0.2564701374	0.3687871059
0.3033231236	0.3408856858
0.3501761099	0.3108832085
0.3970290962	0.2788117015
0.4438820824	0.2447077333
0.4907350687	0.2086123333
0.537588055	0.1705709017
0.5844410412	0.1306331098
0.6312940275	0.08885279107
0.6781470137	0.04528782301
0.725	0

Table 1. Table for the scatter plot of $y = h(x)$.

- Step 4. We drag the x-value of D and distance EF back to the geometry strip (within ClassPad) to obtain the green graph in Figure 2. We observe that this is a scatter plot of $y = h(x)$.

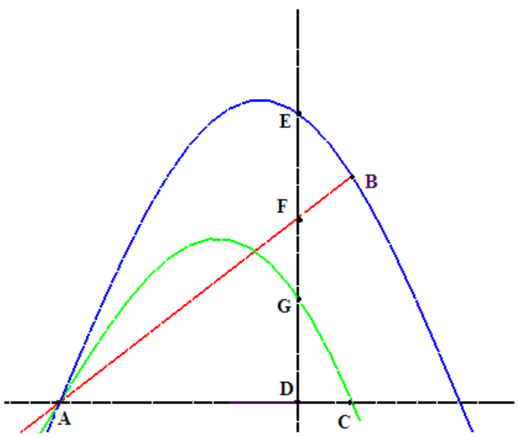


Figure 2. The graphs of two functions and a chord

First, we denote the intersection between the green curve and the vertical line to be G. Then since

$$EF = GD$$

and if the line segment AB is represented by $y = mx + b_1$, the green curve above can be represented by $y = f(x) - (mx + b_1)$ and if we denote this by $h(x)$, or

$$h(x) = f(x) - (mx + b_1)\,. \tag{2}$$

We see if $h'(c) = 0$ for some c in (a, b) then $f'(c) = m$. This is saying that the same c which makes the horizontal tangent for h will make the slope of the tangent line of f at $x = c$ to be the slope of the secant line AB. In other words, we are saying the motivation of equation (1) is given by equation (2). Finally, we find c, numerically, to be -0.3320693226.

3 Geometric Interpretation of Cauchy Mean Value Theorem

We use similar approach mentioned above to demonstrate how geometric interpretations of the Cauchy Mean Value Theorem can be explored with the help of DGS and CAS. The Cauchy Mean Value Theorem can be stated as follows:

Theorem 2 *Suppose the function $f : [a, b] \to R$ and $g : [a, b] \to R$ are continuous and that their restrictions to (a, b) are differentiable. Moreover, assume that $g'(t) \neq 0$ for all t in (a, b). Then there is a point t in (a, b) at which*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(t)}{g'(t).}$$

We see little geometric motivation of why we have two functions f and g and how the conclusion is obtained. This shall become clear later. We make the following observations.

1. Assume functions f and g satisfy the condition of the Cauchy Mean Value Theorem, the Theorem holds, can be interpreted as any number t for which the parametric curve P defined by the equation

   $$P(t) = [g(t), f(t)]$$

   for $a \leq t \leq b$ has slope equal to the slope of the secant that runs from the point $(g(a), f(a))$ to the point $(g(b), f(b))$.

2. Alternatively, if we apply the Mean Value Theorem to the graph of a polar equation $r = h(\theta)$, writing the polar equation in a parametric form

   $$[x(\theta), y(\theta)] = [h(\theta)\cos(\theta), h(\theta)\sin(\theta)] = [g(\theta), f(\theta)], \tag{3}$$

   then we will obtain the conclusion of Cauchy Mean Value Theorem (see [1]).

3. We use the following example to give motivations for the conclusion and the proof of Cauchy Mean Value Theorem. The technique used here can be applied to arbitrary case when the Theorem holds. A video clip which inspires the geometric interpretation of Cauchy Mean Value Theorem can be found here.

Example 3 *We consider the curve $P(t)$ of the form $[x(t), y(t)] = [t\cos t, t\sin t]$ in the interval $t \in [0, 2\pi]$, and let $A = (-\pi, 0)$, $C = (0, \frac{\pi}{2})$.(1) We want to find t on curve of $P(t)$ where the slope of the tangent line is same as the slope of AC when $t \in [\frac{\pi}{2}, \pi]$. (2) Describe how we can arrive at the function to which Rolle's Theorem is applied to yield the Cauchy Mean Value Theorem holds for $P(t)$ on AC when $t \in [\frac{\pi}{2}, \pi]$. In other words, we want to find the equation of the thick curve $Q(t)$ in Figure 3 below, when AC becomes a horizontal line segment.*

We sketch the graph of $[x(t), y(t)]$ and line segment AC below in Figure 3 and we write the equation of the thick curve $Q(t)$ as $[x(t), y_1(t)]$, which is what we would like to find and where we will apply Rolle's Theorem later.

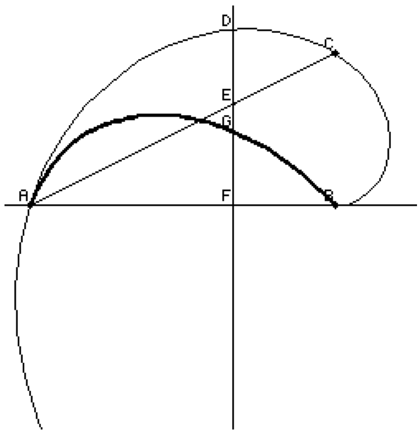


Figure 3. The graphs of $[t\cos t, t\sin t]$, a cord and another parametric curve

We briefly describe how we simulate the curve $Q(t)$ below:

Step 1. We construct the curve $[t\cos t, t\sin t]$.

Step 2. Drag and drop the points $A = (-\pi, 0), B = (0, 0)$ and $C = (0, \frac{\pi}{2})$ into the graph.

Step 3. Select the point D on $[t\cos t, t\sin t]$.

Step 4. Construct a perpendicular line passing through D and perpendicular to AB; and intersect the line segments AC, AB and thick graph respectively at E, F and G.

Step 5. Animate D along $[t\cos t, t\sin t]$ by properly selecting the range of the parameter for animation.

Step 6. Collect and drag the x-value of D and the distance DE into the curve of $P(t)$, we get the thick curve $Q(t)$.

To figure out the equation for the thick curve $Q(t)$, $[x(t), y_1(t)]$, the key here, similar to what we have done earlier for Mean Value Theorem, is

$$DE = GF. \tag{4}$$

Since $DE = f(t) - EF$, we can write

$$y_1(t) = GF = f(t) - EF. \tag{5}$$

First, we need to determine the equation for AC.

Since $A = (-\pi, 0)$,and $C = (0, \frac{\pi}{2})$, we get the slope of AC to be $m = \frac{1}{2}$. We note that $m = \dfrac{f(b) - f(a)}{g(b) - g(a)}$ and we get $y = \frac{1}{2}x + \frac{\pi}{2}$, the parametric equation for AC can be written as

$$\begin{aligned}[x(t), y(t)] &= [t\cos(t), m\cdot x(t)+b_1] \\ &= [t\cos(t), m\cdot g(t)+b_1] \\ &= [t\cos(t), \frac{1}{2}t\cos t+\frac{\pi}{2}].\end{aligned}$$

Thus we have

$$\begin{aligned}y_1(t) &= f(t)-y = f(t)-(m\cdot x(t)+b_1) \\ &= f(t)-(m\cdot g(t)+b_1) \\ &= f(t)-(m\cdot t\cos t+b_1) \\ &= t\sin(t)-\left(\frac{1}{2}t\cos t+\frac{\pi}{2}\right),\end{aligned} \tag{6}$$

and the parametric equation for Q (where we apply Rolle's Theorem) is

$$\begin{aligned}[x(t), y_1(t)] &= [t\cos(t), f(t)-(m\cdot t\cos t+b_1)] \\ &= [t\cos(t), t\sin t-\left(\frac{1}{2}t\cos t+\frac{\pi}{2}\right)].\end{aligned} \tag{7}$$

To find where Q has a horizontal tangent, we note $\frac{dy}{dx} = \frac{\frac{dy}{d\theta}}{\frac{dx}{d\theta}}$ and $\frac{dy}{dx} = 0$ implies that

$$\begin{aligned}\frac{dy_1}{dt} &= \frac{d}{dt}\left(f(t)-(m\cdot g(t)+b_1)\right) \\ &= \frac{d}{dt}\left(f(t)-(m\cdot t\cos t+b_1)\right)=0.\end{aligned} \tag{8}$$

Therefore we need to solve for t so that $f'(t) = mg'(t)$ or

$$\frac{f'(t)}{g'(t)} = m = \frac{f(b)-f(a)}{g(b)-g(a)}. \tag{9}$$

This is exactly what the Cauchy Mean Value theorem has stated.

Now it is a standard exercise to solve for t. We find t to be about 2.425497143 (when $m = \frac{1}{2}$, $a = -\pi$, and $b = 0$ respectively).

In summary, the equation of the straight line that runs from $A = (g(a), f(a))$ to $C = (g(b), f(b))$ is

$$y = f(a) + \frac{f(b)-f(a)}{g(b)-g(a)}\left(x-g(a)\right) \tag{10}$$

Furthermore, we conclude that the curve Q is defined to be

$$Q(t) = \left[g(t), f(t) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)}\left(g(t)-g(a)\right)\right] \tag{11}$$

and that this curve Q has a horizontal tangent when

$$f'(t) - 0 - \frac{f(b)-f(a)}{g(b)-g(a)}g'(t) = 0 \tag{12}$$

which says that

$$\frac{f'(t)}{g'(t)} = \frac{f(b) - f(a)}{g(b) - g(a)}. \tag{13}$$

Obviously, the Cauchy Mean Value is an extension of Mean Value Theorem. We naturally can replace the chord by a smooth curve connecting A and C in the previous example; and the technique we used (with the help of DGS and CAS) earlier for simulating the curve $Q(t)$ can be applied analogously.

4 Alternative Interpretations of Lagrange Remainder Theorem

Many traditional theorems guarantee us the existence of a solution. With the CAS, we can attempt to approximate where the solution is. For example, the Lagrange Remainder Theorem (see below) guarantees us the existence of the point c, we will show with the help of a CAS, we can find such c (if the equations are solvable by a CAS).

Theorem 4 *Let I be a neighborhood of the point x_0 and let n be a nonnegative integer. Suppose that the function $f : I \to R$ has $n+1$ derivatives. Then for each $x \neq x_0$ in I, there is a point c in (x, x_0) or (x_0, x) such that*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}.$$

4.1 Finding a solution numerically

We consider the function $f(x) = x^{\frac{1}{3}}$, and we pick $x_0 = 8$, we want to find the number c that is guaranteed by the Lagrange Remainder Theorem for the function f at x_0, $I = (7, 9)$ and $n = 2$. The 2nd degree Taylor polynomial for f at $x_0 = 8$ is

$$p_2(x) = \sqrt[3]{8} + \frac{(x-8)\sqrt[3]{8}}{24} - \frac{(x-8)^2\sqrt[3]{8}}{576}. \tag{14}$$

We write the remainder

$$R_2(x, 8) = R_2(x) = f(x) - p_2(x) = \sqrt[3]{x} - \left(\sqrt[3]{8} + \frac{(x-8)\sqrt[3]{8}}{24} - \frac{(x-8)^2\sqrt[3]{8}}{576}\right), \tag{15}$$

and we demonstrate for each $x \neq 8$ in $(7, 9)$, there is a point c in $(x, 8)$ or $(8, x)$ such that

$$R_2(x) = \frac{R_2^{(3)}(c)}{3!}(x - 8)^3,$$

where $R_2^{(3)}(c) = f^{(3)}(c) - p_2^{(3)}(c) = f^{(3)}(c)$. Assume we choose $x = 7.5$, we set $g(x) = \frac{R_2^{(3)}(x)}{3!}(7.5 - 8)^3$. We want to find c in $(7.5, 8)$ such that $R_2(7.5) = g(x)$. This is solvable by a CAS, we find $c = 7.872817270$.

Here is an alternative way to interpret the theorem. We define

$$F(x) = R_2^{(3)}(x) = f^{(3)}(x).$$

and

$$G(x) = \frac{R_2(x) \cdot 3!}{(x-8)^3}.$$

We graph $y = F(x)$ (in green) and $y = G(x)$ (in red) together in Figure 4 in the interval $[7, 8]$ below and make the following observations:

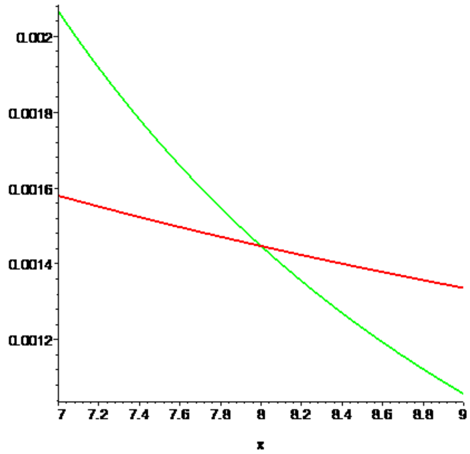


Figure 4. Graphs of $y = F(x)$ and $y = G(x)$

Remark 5

1. The Lagrange Remainder Theorem guarantees us that if we pick a point x in $(7, 9)$ except $x = 8$ for the function G, we can find a c so that $G(x) = F(c)$.

2. Therefore, if we pick any x in $(7, 8)$ on $y = G(x)$, we can find the corresponding point (by going horizontal direction) on $y = F(x)$ so that $F(c) = G(x)$. Similarly, if we pick $(8, 9)$ on $y = G(x)$, we can find the corresponding point on $y = F(x)$ so that $F(c) = G(x)$. For example, if we pick $x = 7.5$, we see $G(7.5) = 0.00150993$ and if we solve $F(x) = 0.00150993$,we get $x = 7.872810231$,which is consistent with the answer we obtained earlier.

3. In other words, any point x in $(7, 9)$ except $x = 8$, the complex expression $G(x) = \dfrac{R_2(x) \cdot 3!}{(x-8)^3} = \sqrt[3]{x} - \left(\sqrt[3]{8} + \dfrac{(x-8)\sqrt[3]{8}}{24} - \dfrac{(x-8)^2\sqrt[3]{8}}{576}\right)$, can be computed by a simpler expression $F(c) = R_2^{(3)}(c) = f^{(3)}(c)$ for a proper c.

A CAS can be used to help us to anticipate that such a number c should exist it can help us to appreciate the Lagrange Remainder Theorem by showing us how to estimate the number c. Thus, author proposes the following alternative interpretation of the Lagrange Remainder Theorem.

Theorem 6 *Let I be a neighborhood of the point x_0 and let n be a nonnegative integer. Suppose that the function $f : I \to R$ has $n+1$ derivatives. Then for each $x \neq x_0$ in I, we can find the point c in (x, x_0) or (x_0, x) such that*

$$G(x) = F(c),$$

where $G(x) = \dfrac{\left(f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k\right)(n+1)!}{(x-x_0)^{n+1.}}$ *and* $F(x) = f^{(n+1)}(x)$.

4.2 Approximating the Remainder Terms Graphically

Inspired by Theorem 6, we can apply Lagrange Remainder Theorem to approximate the remainder terms graphically. Assume the conditions of the Theorem 4 hold and we consider

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1.}. \tag{16}$$

We write

$$R_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k. \tag{17}$$

The Theorem 4 guarantees us for each $x \neq x_0$ in I, we can find the point c in (x, x_0) or (x_0, x) such that

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1.}. \tag{18}$$

Now we assume $c \in (x, x_0)$ and let $y \in [x, x_0]$. We define

$$F_n(y) = \frac{f^{(n+1)}(c)}{(n+1)!}(y-x_0)^{n+1.}. \tag{19}$$

Thus, we have $R_n(x) = F_n(x)$, $R_n(x_0) = F_n(x_0)$ and as we expected we have the following observations:

1. For each fixed non-negative integer n, $F_n(y)$ can be used to approximate $R_n(y)$ in the interval $[x, x_0]$.
2. We can interpret the generalized Mean Value Theorem as follows: Assume the conditions of Lagrange Theorem holds for the function f, for each $x \neq x_0$ in I, we can find the point c in (x, x_0) or (x_0, x) such that

$$\frac{R_n(x)}{\dfrac{(x-x_0)^{n+1.}}{(n+1)!}} = f^{(n+1)}(c). \tag{20}$$

3. We see that $R_n(y) -$ $F_n(y)$ converges to 0 uniformly in $[x, x_0]$ as $n \to \infty$.

We note that Taylor polynomials are to appropriate a function locally in a specified interval. We will use the following example to demonstrate how the remainder terms converge for a given function.

Example 7 *Consider* $f(x) = \cos x$ *and we pick* $x = 0$ *and* $x_0 = 0.725$.

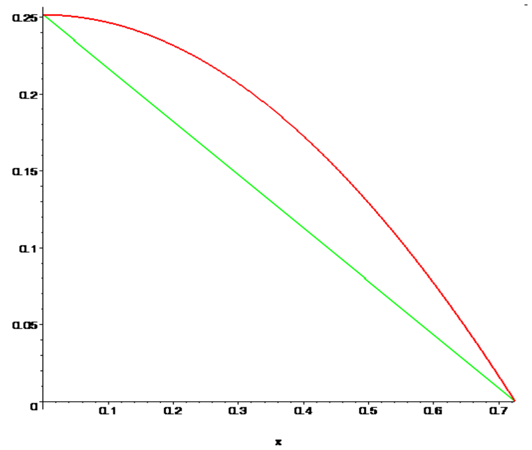
Case 1. For $n = 0$, it follows from Lagrange Theorem that there is a $c \in (x, x_0)$ such that

$$R_0(x) = f(x) - f(x_0) = \frac{f'(c)}{1!}(x - x_0). \tag{21}$$

In this case, we can solve c to be about 0.3542609993408. [We note that we are solving the c such that $f'(c) = \dfrac{f(x) - f(x_0)}{x - x_0}$, where c is guaranteed by the Mean Value Theorem.] Next we define

$$F_0(y) = \frac{f'(c)}{1!}(y - x_0) \tag{22}$$

for $y \in [x, x_0]$. We sketch $z = R_0(y)$ and $z = F_0(y)$ for $y \in [x, x_0]$ as follows ($z = R_0(y)$ in red and $z = F_0(y)$ in green):



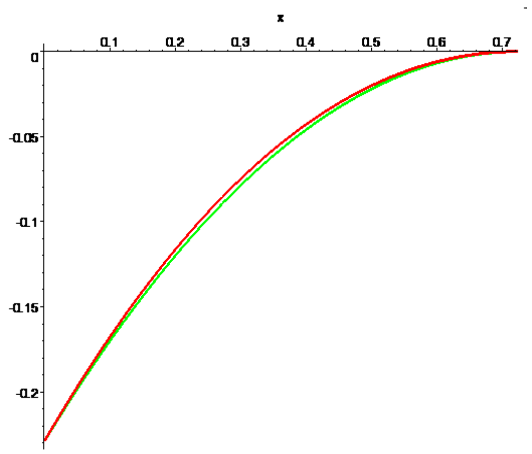
Case 2. For $n = 1$, it follows from Lagrange Theorem that there is a $c_1 \in (x, x_0)$ such that

$$R_1(x) = f(x) - f(x_0) - \frac{f'(x_0)}{1!}(x - x_0) = \frac{f''(c_1)}{2!}(x - x_0)^{2.}. \tag{23}$$

In this case, we can solve c_1 to be about 0.5107442329. Next we define

$$F_1(y) = \frac{f''(c_1)}{2!}(y - x_0)^2 \tag{24}$$

for $y \in [x, x_0]$. We sketch $z = R_1(y)$ and $z = F_1(y)$ for $y \in [x, x_0]$ as follows ($z = R_1(y)$ in red and $z = F_1(y)$ in green):



Remark 8 *We note that for* $f(x) = \cos x$ *in* $[x, x_0] = [0, 0.725]$*, the remainder* $R_1(y)$ *and* $F_1(y)$ *are already very close in the interval* $[0, 0.725]$*, which is not surprising since* $\cos x$ *can be approximated nicely locally with a degree two polynomial. Thus, by computing appropriate values* c, c_1 *and etc. we are able to approximate the remainder* $R_n(y)$ *in* $[x, x_0]$ *by using* $F_n(y)$*, which is a polynomial of degree* $n + 1$*.*

5 Conclusion

In this paper, we demonstrated geometrically how the proofs of Mean Value and Cauchy Mean Value Theorems are merely taking the difference between two quantities (see equations (2) and (5)). Furthermore, to understand how the Cauchy Mean Value Theorem is stated, we simply apply Mean Value Theorem on a parametric curve. The equation (11) can be made intuitive by equations (4) and (5) with animation and help of a Dynamic Geometry Software. We further use graphical and numerical capabilities within software packages (see [ClassPad] and [Maple]) to give alternative interpretations to Lagrange Remainder's Theorem. The evolving technological tools allow us to experiment abstract concepts and make mathematics more accessible to more students.

6 Acknowledgement

The author thanks Jonathan Lewin for many insightful discussions.

References

[1] Fulks, Watson, *Advanced Calculus*, 3rd edition, Wiley, 1978, ISBN 0-471-02195-4.

Software Packages

- [ClassPad] ClassPad Manager, a product of CASIO Computer Ltd., http://www.classpad.net or http://www.classpad.org/.
- [Maple] Maple 10.0.6, a product of Maplesoft, http://www.maplesoft.com/.

Supplemental Electronic Materials

- Yang, W.-C., *a video clip* which summarizes how ClassPad Manager is used in interpreting the Mean Value Theorem geometrically.
- Yang, W.-C., *a video clip* which summarizes how ClassPad Manager is used in interpreting the Cauchy Mean Value Theorem geometrically.
- Yang, W.-C., *ClassPad eActivity* interpreting the Mean Value Theorem geometrically at https://php.radford.edu/~ejmt/Content/Papers/v1n1p4/MVT.vcp.
- Yang, W.-C., *ClassPad eActivity* interpreting the Cauchy Mean Value Theorem geometrically at https://php.radford.edu/~ejmt/Content/Papers/v1n1p4/CMV.vcp.