Mean, Meaner and the Meanest Mean Value Theorem

J. J. Koliha

> *. . . the real nature of the Mean Value Theorem*
> *is exhibited by writing it as an inequality.*
>
> J. Dieudonné

1 Introduction

The Mean Value Theorem of the elementary calculus keeps attracting the attention of mathematicians who ponder how to make its proof simple and elegant [1], how to generalize it [10], how to use it in proofs of other theorems [8], and, perversely, how to avoid it [2]. This MONTHLY has carried dozens of articles in which the Mean Value Theorem was the hero or the villain (for instance [5]).

As Jean Dieudonné pointed out in his *Foundations of Modern Analysis* [7], the Mean Value Theorem written as an equality $f(b) - f(a) = f'(c)(b - a)$ works only for real valued functions, and in general nothing is known about c except that it lies between a and b. In most applications we need an inequality or even an inclusion rather than an equality [9]. In this note we prove three versions of the Mean Value Theorem in the form of an inequality, and demonstrate their usefulness in applications. All three versions will be proved using the so-called full covers of $[a, b]$.

If I is a bounded interval with end points a, b, where $a < b$, and f is a complex valued function whose domain contains I, we write $f(I) = f(b) - f(a)$. The length of I is denoted by $|I|$. In this notation we can express the classical Cauchy's Mean Value Theorem as follows: *If* $f, g\colon [a, b] \to \mathbb{R}$ *are continuous on* $I = [a, b]$ *and differentiable on* (a, b) *with* $g'(x) \neq 0$ *for all* $x \in (a, b)$*, then there exists* $c \in (a, b)$ *such that*

$$\frac{f(I)}{g(I)} = \frac{f'(c)}{g'(c)}.$$

If we know that $|f'(x)| \le g'(x)$ (or $f'(x) \le g'(x)$) for all $x \in (a, b)$, we may conclude that $|f(I)| \le g(I)$ (or $f(I) \le g(I)$). This removes the dependence on the point c, which in general is not known, and motivates Theorem 1.

2 Full covers and Cousin's lemma

Definition. Let $[a, b]$ be a closed bounded interval and $\delta\colon [a, b] \to \mathbb{R}$ a strictly positive function. A family $\mathcal{C}$ of closed subintervals of $[a, b]$ is a *full* δ*-cover of* $[a, b]$ if, for any subinterval $[u, v]$ of $[a, b]$,

$$x \in [u, v] \text{ and } v - u < \delta(x) \implies [u, v] \in \mathcal{C}.$$

A *full cover* of $[a, b]$ is a full δ-cover of $[a, b]$ for some positive function $\delta\colon [a, b[\to \mathbb{R}$.

The following lemma is a result obtained in 1895 by Pierre Cousin in [6] presented in a slightly different formulation due to Thomson [11]. This very useful result is indispensable in theory of non-absolute integrals, and leads to the Riemann-like definition of the Lebesgue integral. A proof by contradiction based on a successive bisection of the interval $[a, b]$ was given by Botsko [4].

Lemma 1 (Cousin's lemma)**.** *Let $\mathcal{C}$ be a full cover for $[a, b]$. Then each closed interval contained in $[a, b]$ has a partition whose subintervals lie in $\mathcal{C}$.*

Full covers provide an extremely useful tool which can be used to prove many of the classical theorems of calculus and elementary real analysis (see [3, 4, 11]).

3 Mean Value Theorem

In our terminology, the 'meaner' the theorem, the stronger it is. Our meanest Mean Value Theorem tackles issues usually associated with the Lebesgue integral. For our three versions of the Mean Value Theorem we need a few concepts.

(i) A set $A \subset \mathbb{R}$ is *null* if for each $\varepsilon > 0$ there are open intervals I_n such that $\sum_{n=1}^{\infty} |I_n| < \varepsilon$ and $A \subset \bigcup_{n=1}^{\infty} I_n$. For instance, any countable set is null as is the Cantor ternary set obtained by successive removals of the open middle thirds from the interval $[0, 1]$.

(ii) We say that a property $\mathcal{P}$ holds *nearly everywhere* in a set $A \subset \mathbb{R}$ if it holds in $A \setminus D$, where $D \subset A$ is countable, and that it holds *almost everywhere* in A if it holds in $A \setminus E$, where $E \subset A$ is a null set.

(iii) A function $f\colon I \to \mathbb{C}$ is *absolutely continuous* on an interval I (bounded or unbounded) if for every $\varepsilon > 0$ there exists $\eta > 0$ such that for any finite family $\{I_k : k = 1, \ldots, n\}$ of non-overlapping subintervals of I,

$$\sum_{k=1}^{n} |I_k| < \eta \implies \sum_{k=1}^{n} |f(I_k)| < \varepsilon. \tag{1}$$

(iv) We say that a function $g\colon I \to \mathbb{R}$ is *increasing* (respectively *strictly increasing*) in I if $g(x_1) \leq g(x_2)$ (respectively $g(x_1) < g(x_2)$) when $x_1, x_2 \in I$ and $x_1 < x_2$. Analogously we define *decreasing* and *strictly decreasing.*

In Theorem 1 we assume that g is increasing. This condition may seem more restrictive than the assumption $g' \neq 0$ in I used in the classical Cauchy's Mean Value Theorem. In fact, it is *less* restrictive, since we can show that $g' \neq 0$ everywhere in I implies that g is strictly monotonic in I.

Theorem 1 (mean, meaner and the meanest Mean Value Theorem)**.** *Let* $f\colon [a,b] \to \mathbb{C}$*, let* $g\colon [a,b] \to \mathbb{R}$ *be increasing, and let the inequality* $|f'| \le g'$ *hold*

(A) *everywhere in* (a,b) *while the finite one-sided limits* $f(a+)$ *and* $f(b-)$ *exist, or*

(B) *nearly everywhere in* $[a,b]$ *while* f *is continuous in* $[a,b]$*, or*

(C) *almost everywhere in* $[a,b]$ *while* f *is absolutely continuous in* $[a,b]$*.*

Then $|f(J)| \le g(J)$ *whenever* $J \subset [a,b]$ *is a closed interval.*

Proof. We observe that (A) follows from (B) when the countable set $D \subset [a,b]$ involved in the definition of 'nearly everywhere' consists of the end points of $[a,b]$. To prove (B), it is enough to consider $J = [a,b]$. Let $|f'(t)| \le g'(t)$ for all $t \in J \setminus Q$, where $Q = \{s_1, s_2, s_3, \dots\}$ is a countable subset of J. Let $\varepsilon > 0$ be given, and let $\mathcal{C} = \bigcup_{n=0}^{\infty} \mathcal{C}_n$, where $\mathcal{C}_0$ (respectively $\mathcal{C}_n$ for $n > 0$) consist of all closed subintervals K of J such that $|f(K)| \le g(K) + \varepsilon|K|$ (respectively $s_n \in K$ and $|f(K)| \le (\frac{1}{2})^{n+1}\varepsilon\}$). We show that the family $\mathcal{C}$ is a full cover for J.

For each $x \in J \setminus Q$ there exists $\delta(x) > 0$ such that for any interval $K = [u,v] \subset J$ satisfying $x \in [u,v]$ and $v - u < \delta(x)$ we have

$$\left|\frac{f(v) - f(u)}{v - u} - f'(x)\right| \le \tfrac{1}{2}\varepsilon \quad \text{and} \quad \left|\frac{g(v) - g(u)}{v - u} - g'(x)\right| \le \tfrac{1}{2}\varepsilon. \tag{2}$$

For each $x = s_n \in Q$ there exists $\delta(x) > 0$ such that for any $K = [u,v] \subset J$ satisfying x and $v - u < \delta(x)$, we have

$$|f(v) - f(u)| \le (\tfrac{1}{2})^{n+1}\varepsilon. \tag{3}$$

If (2) holds, then

$$\begin{aligned} |f(v) - f(u)| &\le |f'(x)|(v - u) + \tfrac{1}{2}\varepsilon(v - u) \\ &\le g'(x)(v - u) + \tfrac{1}{2}\varepsilon(v - u) \le g(v) - g(u) + \varepsilon(v - u). \end{aligned}$$

Hence $|f(K)| \le g(K) + \varepsilon|K|$, that is, $K \in \mathcal{C}_0$. If $s_n \in K$ and $|K| < \delta(s_n)$, then $|f(K)| \le (\frac{1}{2})^{n+1}\varepsilon$, that is, $K \in \mathcal{C}_n$. This proves that $\mathcal{C}$ is a full δ-cover for J.

According to Cousin's lemma there exists a partition $\{t_0, t_1, \dots, t_p\}$ of $J = [a,b]$ such that $K_j = [t_{j-1}, t_j] \in \mathcal{C}$ for $j = 1, \dots, p$. We observe that each $\mathcal{C}_n$ with $n > 0$ contains at most two of the K_j. Then

$$\begin{aligned} |f(J)| = \Big|\sum_{j=1}^{p} f(K_j)\Big| &\le \sum_{j=1}^{p} |f(K_j)| = \sum_{K_j \in \mathcal{C}_0} |f(K_j)| + \sum_{n=1}^{\infty} \sum_{K_j \in \mathcal{C}_n} |f(K_j)| \\ &\le \sum_{K_j \in \mathcal{C}_0} (g(K_j) + \varepsilon|K_j|) + 2\sum_{n=1}^{\infty} (\tfrac{1}{2})^{n+1}\varepsilon \le \sum_{j=1}^{p} g(K_j) + \varepsilon \sum_{j=1}^{p} |K_j| + \varepsilon \\ &\le g(J) + (|J| + 1)\varepsilon. \end{aligned} \tag{4}$$

Since ε was arbitrary, $|f(J)| \le g(J)$. The assumption that g is increasing is needed in the second line of the preceding argument where we have to know that $g(K_j)$ is non-negative for each j.

(C) To prove the meanest Mean Value Theorem we again assume that $J = [a, b]$, and that $|f'(t)| \le g'(t)$ for all $t \in J \setminus E$, where E is a null subset of J. Let $\varepsilon > 0$ be given, and let $\eta > 0$ be the number corresponding to ε in the definition (1) of the absolute continuity of f on J. By the definition of a null set there exists a sequence of open intervals I_n such that $E \subset \bigcup_{n=1}^{\infty} I_n$ and $\sum_{n=1}^{\infty} |I_n| < \eta$. Let $\mathcal{C} = \bigcup_{n=0}^{\infty} \mathcal{C}_n$, where $\mathcal{C}_0$ (respectively $\mathcal{C}_n$ for $n > 0$) consists of all closed subintervals K of J such that $|f(K)| \le g(K) + \varepsilon|K|$ (respectively $K \subset I_n$). We show that the family $\mathcal{C}$ is a full cover for J.

Let $x \in J \setminus E$. Then there exists $\delta(x) > 0$ such that (2) holds whenever $x \in K = [u, v] \subset J$ and $v - u < \delta(x)$. The argument from the preceding part of the proof shows that $|f(K)| \le g(K) + \varepsilon|K|$. If $x \in E$, then there exists $\delta(x) > 0$ such that for any $K = [u, v] \subset J$ satisfying $x \in [u, v]$ and $v - u < \delta(x)$ we have $K \subset I_n$ for some n. Thus $\mathcal{C}$ is a full δ-cover for J. By Cousin's lemma, there exists a partition $\{t_0, t_1, \ldots, t_p\}$ of the interval $J = [a, b]$ for which each subinterval $K_j = [t_{j-1}, t_j]$ belongs to $\mathcal{C}$. Observe that all the K_j contained in I_n have the total length not exceeding $|I_n|$. Thus $\sum_{n=1}^{\infty} \sum_{K_j \subset I_n} |f(K_j)| < \varepsilon$ as the intervals K_j involved form a finite family of subintervals of J whose total length does not exceed η. Then

$$\begin{aligned}
|f(J)| = \Big|\sum_{j=1}^{p} f(K_j)\Big| &\le \sum_{j=1}^{p} |f(K_j)| = \sum_{K_j \in \mathcal{C}_0} |f(K_j)| + \sum_{n=1}^{\infty} \sum_{K_j \in \mathcal{C}_n} |f(K_j)| \\
&\le \sum_{K_j \in \mathcal{C}_0} (g(K_j) + \varepsilon|K_j|) + \sum_{n=1}^{\infty} \sum_{K_j \subset I_n} |f(K_j)| \\
&\le \sum_{j=1}^{p} g(K_j) + \varepsilon \sum_{j=1}^{p} |K_j| + \varepsilon \le g(J) + (|J| + 1)\varepsilon \quad (\text{as } g(K_j) \ge 0 \text{ for all } j).
\end{aligned}$$

Since ε was arbitrary, $|f(J)| \le g(J)$. □

In the proof of the preceding theorem we essentially followed Botsko [4]. Part (B) of the preceding theorem is proved in Dieudonné [7] for a Banach space valued function f. It is not difficult to see that our proof works also in this setting.

A modification of the preceding proof yields Theorem 2 in which f is real valued. Check the inequalities carefully. For instance, the string of inequalities (4) will start with

$$f(J) = \sum_{j=1}^{p} f(K_j) \le \sum_{K_j \in \mathcal{C}_0} f(K_j) + \sum_{n=1}^{\infty} \sum_{K_j \in \mathcal{C}_n} |f(K_j)|.$$

Theorem 2. *Let $f\colon [a, b] \to \mathbb{R}$, let $g\colon [a, b] \to \mathbb{R}$ be increasing, and let the inequality $f' \le g'$ hold subject to one of the conditions* (A), (B) *or* (C) *of Theorem* 1. *Then $f(J) \le g(J)$ whenever $J \subset [a, b]$ is a closed bounded interval.*

4 Applications

Now is the time to show that our three versions of the Mean Value Theorem yield the nice results which are known to follow from the classical form of the theorem, and that the inequality form of the theorem simplifies the proofs.

Proposition 1. *If* $f\colon [a,b] \to \mathbb{C}$ *satisfies* $f' = 0$ *subject to one of the conditions* (A), (B), (C), *then* f *is constant on* $[a,b]$.

Proof. An important special case of Theorem 1 is obtained when we set $g(t) = Mt$ for some constant $M \geq 0$. Under any of the three assumptions of Theorem 1 we get $|f(J)| \leq M|J|$ for any closed bounded interval $J \subset [a,b]$. If $M = 0$, we conclude that f is constant on $[a,b]$. □

This is a very powerful result indeed! If f is real valued and $f' = 0$ everywhere in (a,b), we obtain the classical result of calculus. If $f' = 0$ nearly everywhere and f is continuous, we reach the same conclusion. Finally, if $f' = 0$ almost everywhere, we need to assume that f is absolutely continuous to ensure the constancy of f.

Proposition 2. *If* $f\colon [a,b] \to \mathbb{R}$ *satisfies* $f' \leq 0$ (*respectively* $f' \geq 0$) *subject to one of the conditions* (A), (B), (C), *then* f *is decreasing* (*respectively increasing*) *in* $[a,b]$.

Proof. Assume first that $f' \leq 0$ subject to one of the conditions (A), (B) or (C). The constant function $g = 0$ is increasing on $[a,b]$, and f and g satisfy the hypotheses of Theorem 2. Hence for any interval $[u,v] \subset [a,b]$, $f(v) - f(u) = f([u,v]) \leq g([u,v]) = 0$, that is, f is decreasing. The conclusion for $f' \geq 0$ follows by applying the preceding result to $-f$. □

Proposition 3 (l'Hôpital's rule)**.** *Let* f *and* g *be real valued. If for some* $a \in \mathbb{R}$, $\lim_{x\to a} f(x) = 0 = \lim_{x\to a} g(x)$ *or* $\lim_{x\to a} f(x) = \infty = \lim_{x\to a} g(x)$, *then*

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = w \implies \lim_{x\to a} \frac{f(x)}{g(x)} = w \quad \text{where } w \in \mathbb{R}. \tag{5}$$

Proof. It is implicit from the existence of the limit on the left that there exists an open interval J containing a such that the derivatives f' and g' exist and $g' \neq 0$ in $J \setminus \{a\}$.

We prove the result first for the one sided limit from the right. Let $\varepsilon > 0$. Then there exists $\delta > 0$ such that $I := (a, a+\delta) \subset J$, and

$$\left| \frac{f'(x)}{g'(x)} - w \right| \leq \varepsilon \text{ if } a < x < a + \delta.$$

As mentioned earlier, g' does not change sign in I. By considering $-f$ and $-g$ if necessary, we may assume that $g' > 0$ in I; g is increasing in I by Proposition 2. Then

$$(w-\varepsilon)g'(x) \leq f'(x) \leq (w+\varepsilon)g'(x) \quad \text{for } x \in I.$$

Assume that $f(x)$ and $g(x)$ tend to ∞ as $x \to a$. Select $u \in (a, a+\delta)$ and keep it fixed. If $a < x < u$, then

$$(w-\varepsilon)(g(u)-g(x)) \leq f(u)-f(x) \leq (w+\varepsilon)(g(u)-g(x))$$

by Theorem 2 (A) applied to the interval $[x, u]$, and so

$$\left| \frac{f(x)-f(u)}{g(x)-g(u)} - w \right| \leq \varepsilon.$$

Let $h(x) = (1 - f(u)/f(x))/(1 - g(u)/g(x))$; then $\lim_{x \to a+} h(x) = 1$. There exists $\delta_1 \in (0, u-a)$ such that $h(x) \geq \frac{1}{2}$ and $|1 - h(x)| < \varepsilon$ whenever $a < x < a + \delta_1$. Let $a < x < a + \delta_1$ $(< u)$. Then

$$\begin{aligned} \left| \frac{f(x)}{g(x)} - w \right| &\leq 2 \left| \frac{f(x)}{g(x)} h(x) - wh(x) \right| \leq 2 \left| \frac{f(x)-f(u)}{g(x)-g(u)} - w \right| + 2|w - wh(x)| \\ &\leq 2\varepsilon + 2|w|\varepsilon = 2(1+|w|)\varepsilon. \end{aligned}$$

The limit from the left is treated similarly.

The case $\lim_{x \to a} f(x) = 0 = \lim_{x \to a} g(x)$ is left to the reader. □

References

[1] L. C. Barrett, Methods of proving mean valued theorem, this Monthly **69** (1962), 50–52.

[2] R. P. Boas, Jr., Classroom Notes: l'Hospital's Rule Without Mean-Value Theorems, this Monthly **76** (1969), 1051–1053.

[3] M. W. Botsko, A unified treatment of various theorems in elementary analysis, this Monthly **94** (1987), 450–452.

[4] M. W. Botsko, The use of full covers in real analysis, this Monthly **96** (1989), 328–333.

[5] L. W. Cohen, On being mean to the mean value theorem, this Monthly **74** (1967), 581–582.

[6] P. Cousin, Sur les fonctions de n variables complexes, *Acta Mathematica* **19** (1895), 1–62.

[7] J. Dieudonné, *Foundations of Modern Analysis*, 2nd and enlarged ed., Academic Press, New York, 1969.

[8] M. P. Drazin, Some inequalities arising from a generalized mean value theorem, this Monthly **62** (1955), 226–232.

[9] M. Furi and M. Martelli, On the mean value theorem, inequality, and inclusion, this Monthly **98** (1991), 840–846.

[10] S. Reich, On mean value theorems, this Monthly **76** (1969), 70–73.

[11] B. S. Thomson, On full covering properties, *Real Analysis Exchange* **6** (1980–81), 77–93.