A MEAN VALUE INEQUALITY FOR ALL FUNCTIONS

PETE L. CLARK

1. Introduction

Over the past week I attended five short talks on this linchpin of calculus:

Theorem 1 (Mean Value Theorem, henceforth "MVT")**.** *Let $f : [a,b] \to \mathbb{R}$ be a function. Suppose that f is continuous and that its restriction to (a,b) is differentiable. Then there is $c \in (a,b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{1}$$

As one of the speakers pointed out, MVT is *not* just the equation (1): rather, (1) is its conclusion, which holds under the continuity and differentiability hypotheses.

Most freshman calculus students will struggle to find meaning in MVT unless it is placed in good context. The speakers did this very well. All of them gave the geometric interpretation involving a tangent line parallel to the secant line. Most also gave the physical interpretation: suppose $f(x)$ gives the position of a particle at time x. Then $f(b) - f(a)$ is the displacement of the particle between time a and time b, while $b - a$ is the elapsed time, so the quantity $\frac{f(b)-f(a)}{b-a}$ is the particle's *average* velocity on the time interval $[a,b]$. In turn $f'(c)$ can be interpreted as the *instantaneous* velocity at time c. Thus MVT asserts that the average velocity is also equal to the instantaneous velocity for at least one point in time.

Several of the speakers gave the following application: **MVT could be used by the highway patrol in order to award speeding tickets even when the speeding is not directly witnessed.** Namely, if the authorities know that at time a your position is $f(a)$ – perhaps you pass a camera that views your license plate – and also that at time $b > a$ your position is $f(b)$, they can apply MVT to conclude that at some point in between your instantaneous velocity must have been $\frac{f(b)-f(a)}{b-a}$. So if for the entire stretch of the highway between $f(a)$ and $f(b)$ the speed limit was at most $M < \frac{f(b)-f(a)}{b-a}$, they can write you a ticket.

Everything so far sounds quite familiar. But during the last lecture a new thought poked through: there is a snag in this argument. I didn't see it for years. Do you see it now?

Here it is: in order to apply MVT we need to know that f is continuous on $[a,b]$ and differentiable on (a,b). Is it clear that one must drive in a differentiable manner?

I think it is not! For instance, if you are driving at 30 mph and get rear ended, that will instantly bump up your speed. It seems natural to model this position function as having a corner point at the point of collision. But even one point of nondifferentiability can falsify the conclusion of MVT: e.g. the function

$$f : [-1,1] \to \mathbb{R}, \ x \mapsto |x|$$

is differentiable at every point other than $x = 0$ with derivative either 1 or -1: so $f'(c)$, when it exists, gets nowhere near the average velocity, which is 0.

If called to testify on the nondifferentiability problem in traffic court, I would say: while it's highly plausible that a driver's position must be a continuous function of time, differentiability feels like a simplifying assumption made so as to mathematically model a real world situation. There are also results suggesting that a "generic" continuous function is not differentiable.[1]

What is your reaction to this objection? Mine is to question our apparent definition of "exceeding the speed limit of M" as $f'(c) > M$ for at least one c in the interval (a, b). Isn't this more an *interpretation* of the derivative than a principled requirement that it exist?

Here is a definition of exceeding a speed limit M that applies to any function $f : [a, b] \to \mathbb{R}$ whatsoever: for $M > 0$, a function $f : [a, b] \to \mathbb{R}$ is an **M-speeder** if for all $\delta > 0$ there are $c, d \in [a, b]$ with $0 < d - c < \delta$ such that $\frac{|f(d)-f(c)|}{d-c} > M$. Thus you are an M-speeder if there are arbitrarily short subintervals of $[a, b]$ on which your average speed exceeds M. This definition seems physically appealing and even in accordance with the type of unsafe driving speed limits are designed to prevent: covering too much ground in too short a time for you and others to respond appropriately to changes in traffic conditions.

And here is the main result of this note:

Theorem 2. *Let $f : [a, b] \to \mathbb{R}$ be a function. Then:*

a) *The function f is an M-speeder for all $M < \frac{|f(b)-f(a)|}{b-a}$.*
b) *More precisely: there is $c \in [a, b]$ such that:*
   (i) *if $\frac{f(b)-f(a)}{b-a} > M$, then for all $\delta > 0$ there is $x \in [a, b]$ such that $0 < |c - x| < \delta$ and $\frac{f(x)-f(c)}{x-c} > M$; and*
   (ii) *if $\frac{f(b)-f(a)}{b-a} < -M$, then for all $\delta > 0$, there is $x \in [a, b]$ such that $0 < |c - x| < \delta$ and $\frac{f(x)-f(c)}{x-c} < -M$.*
c) *If f is continuous, the number c of part b) can be chosen to lie in (a, b).*

Thus the indirect procedure for awarding speeding tickets seems justified after all.

We give the proof of Theorem 2 in §2. In §3 we explore some complements, including a version of Theorem 2 for functions taking values in any metric space. In §4 we will discuss more conventional Mean Value Inequalities and see that they are implied by our results.

2. The Proof

For $f : [a, b] \to \mathbb{R}$ and $x \in (a, b)$, we put

$$\overline{D}_f(x) := \inf_{\epsilon>0} \sup \left\{ \frac{f(x+h) - f(x)}{h} \;\middle|\; 0 < |h| < \epsilon \right\} \in [-\infty, \infty]$$

$$\underline{D}_f(x) := \sup_{\epsilon>0} \inf \left\{ \frac{f(x+h) - f(x)}{h} \;\middle|\; 0 < |h| < \epsilon \right\} \in [-\infty, \infty].$$

We define $\overline{D}_f(a)$ and $\underline{D}_f(a)$ (resp. $\overline{D}_f(b)$ and $\underline{D}_f(b)$) similarly, but restricting to $0 < h < \epsilon$ (resp. restricting to $-\epsilon < h < 0$). Thus f is differentiable at $x \in [a, b]$ – in the one-sided sense at the endpoints – if and only if $\overline{D}_f(x)$ and $\underline{D}_f(x)$ are equal and finite, and if so this common value is $f'(x)$. However, both functions are defined (as extended real numbers) for arbitrary f.[2]

[1] A "generic" continuous function $f : [a, b] \to \mathbb{R}$ is differentiable *at no point* in the sense that the somewhere differentiable functions form a meager subset of the Baire space of all continuous functions [M, §49].

[2] These are two-sided versions of the **Dini derivatives**, though you need not know what that means.

We will make use of the following reframing of the completeness of $\mathbb{R}$ as an inductive principle.

Theorem 3 (Real Induction)**.** *Let $a < b$ be real numbers, and let $S \subseteq [a, b]$. We suppose:*

(RI1) *We have $a \in S$.*
(RI2) *If $x \in [a, b)$, then $x \in S \implies [x, y] \subseteq S$ for some $y > x$.*
(RI3) *If $x \in (a, b]$ and $[a, x) \subseteq S$, then $x \in S$.*

Then $S = [a, b]$.

Proof. This is [Cl19, Thm. 1]. $\square$

Theorem 4. *Let $f : [a, b] \to \mathbb{R}$, and let $A, B \in \mathbb{R}$.*

a) *If $\overline{D}_f(x) \leq A$ for all $x \in [a, b]$, then $\frac{f(b)-f(a)}{b-a} \leq A$.*
b) *If $\underline{D}_f(x) \geq B$ for all $x \in [a, b]$, then $\frac{f(b)-f(a)}{b-a} \geq B$.*

Proof. a) For $\epsilon > 0$, put

$$S_\epsilon := \{x \in [a, b] \mid f(x) - f(a) \leq (A + \epsilon)(x - a)\}.$$

It suffices to show:

$$\forall \epsilon > 0, \; S_\epsilon = [a, b]. \tag{2}$$

For if so, then for all $\epsilon > 0$, because $x \in S_\epsilon$ we have

$$f(b) - f(a) \leq (A + \epsilon)(b - a),$$

and it follows that $\frac{f(b)-f(a)}{b-a} \leq A$. We will show that $S_\epsilon = [a, b]$ by Real Induction.
(RI1): It is immediate that $a \in S_\epsilon$.
(RI2): Let $x \in [a, b) \cap S_\epsilon$. Since $\overline{D}_f(x) \leq A$, there is $\delta > 0$ such that for all $y \in [x, x + \delta]$ we have $\frac{f(y)-f(x)}{y-x} \leq A + \epsilon$, and thus we have

$$f(y) - f(a) - (f(y) - f(x)) + (f(x) - f(a)) \leq (A+\epsilon)(y - x) + (A+\epsilon)(x - a) = (y - a)(A+\epsilon),$$

so $[x, x + \delta] \subseteq S_\epsilon$.
(RI3): Let $x \in (a, b] \cap S_\epsilon$. Since $\overline{D}_f(x) \leq A$, there is $\delta > 0$ such that for all $y \in [x - \delta, x]$ we have $\frac{f(x)-f(y)}{x-y} \leq A + \epsilon$, and thus we have

$$f(x) - f(a) = (f(x) - f(y)) + (f(y) - f(a)) \leq (A+\epsilon)(x - y) + (A+\epsilon)(y - a) = (x - a)(A+\epsilon),$$

so $x \in S_\epsilon$.
b) This is very similar to part a). Or, since $\underline{D}_{-f}(x) = -\overline{D}_f(x)$, we can reduce to part a). $\square$

Part b) of Theorem 2 follows from Theorem 4:

- If $\frac{f(b)-f(a)}{b-a} > M$, then by Theorem 4a) we have $\overline{D}_f(c) > M$ for some $c \in [a, b]$, so for all $\delta > 0$ there is $x \in [a, b]$ with $0 < |c - x| < \delta$ such that $\frac{f(x)-f(c)}{x-c} > M$.
- If $\frac{f(b)-f(a)}{b-a} < -M$, then by Theorem 4b) we have $\underline{D}_f(c) < -M$ for some $c \in [a, b]$, so for all $\delta > 0$ there is $x \in [a, b]$ with $0 < |c - x| < \delta$ such that $\frac{f(x)-f(c)}{x-c} < -M$.

Clearly one of these holds if $|\frac{f(b)-f(a)}{b-a}| > M$, so f is an M-speeder, and part a) follows. As for part c), suppose that f is continuous. Then there are $a < a' < b' < b$ such that $M < |\frac{f(b')-f(a')}{b'-a'}|$. Applying part b) to the restriction of f to $[a', b']$ gives us a $c \in [a', b'] \subset (a, b)$.

Example 5. *Let* $f : [a, b] \to \mathbb{R}$ *by* $f(x) = \begin{cases} a & \text{if } x = a \\ b & \text{if } a < x \leq b \end{cases}$, *and let* $0 < M < 1 = \frac{f(b)-f(a)}{b-a}$. *The only* $c \in [a, b]$ *that satisfies the conclusion of Theorem 2b) is* $c = a$. *Thus in Theorem 2c) the hypothesis of continuity cannot be removed.*

3. Some Complements

3.1. **Infinite Speeders.** Let us say that a function $f : [a, b] \to \mathbb{R}$ is an **infinite-speeder** if it is an M-speeder for all $M > 0$.

Our next result has a clear moral: do not drive discontinuously!

Proposition 6. *Every discontinuous function* $f : [a, b] \to \mathbb{R}$ *is an infinite-speeder.*

Proof. Suppose that f is discontinuous at $c \in [a, b]$. Then there is $\epsilon > 0$ such that for all $\eta > 0$, there is $x \in [a, b]$ with $0 < |x - c| < \eta$ and $|f(x) - f(c)| \geq \epsilon$, so

$$\left| \frac{f(x) - f(c)}{x - c} \right| > \frac{\epsilon}{\eta}.$$

Let $M, \delta > 0$. Taking $\eta := \min\left(\delta, \frac{\epsilon}{M}\right)$ shows that f is an M-speeder. □

However, the converse of Proposition 6 fails: there are continuous infinite-speeders.

Example 7.

a) *The function* $f : [-1, 1] \to \mathbb{R}$ *by* $x \mapsto x^{\frac{1}{3}}$ *is an infinite-speeder. The vertical tangent line at* 0 *means* $\overline{D}_f(0) = \underline{D}_f(0) = \infty$, *and the result follows from this.*
b) *More generally, any function* $f : [a, b] \to \mathbb{R}$ *for which either* $\overline{D}_f$ *or* $\underline{D}_f$ *is unbounded is an infinite-speeder. The function* $g : [-1, 1] \to \mathbb{R}$ *by* $\begin{cases} x \mapsto x^2 \sin(1/x^2) & x \neq 0 \\ 0 & x = 0 \end{cases}$ *is everywhere differentiable but its derivative is unbounded on any interval containing* 0. *So even a differentiable function can be an infinite-speeder.*
c) *For* $C \geq 0$, *a function* $f : [a, b] \to \mathbb{R}$ *is* ***C-Lipschitz*** *if for all* $x, y \in [a, b]$, *we have* $|f(x) - f(y)| \leq C|x - y|$. *Evidently a* C*-Lipschitz function is not an* M*-speeder for any* $M \geq C$. *In particular, if* $f : [a, b] \to \mathbb{R}$ *is a* C^1*-function – i.e.,* f' *exists and is continuous (hence bounded, by the Extreme Value Theorem) – then* f *is* $\max_{x \in [a,b]} |f'(x)|$*-Lipschitz, by MVT, so is not an infinite-speeder.*

3.2. **Strong Speeders.** Say a function $f : [a, b] \to \mathbb{R}$ is a **strong M-speeder** if there is $c \in [a, b]$ such that for all $\delta > 0$ there is $x \in [a, b]$ with $0 < |x - c| < \delta$ such that $\left| \frac{f(x)-f(c)}{x-c} \right| > M$. By Theorem 2, every $f : [a, b] \to \mathbb{R}$ is a strong M-speeder for all $0 < M < \frac{|f(b)-f(a)|}{b-a}$. Above I went with my honest intuition for what should constitute a speeding violation. Should I have made the stronger definition instead? In fact it doesn't matter:

Theorem 8. *For* $f : [a, b] \to \mathbb{R}$ *and* $M > 0$, *the following are equivalent:*

(i) *There are* $a \leq c < d \leq b$ *such that* $\frac{|f(d)-f(c)|}{d-c} > M$.
(ii) *The function* f *is a strong* M*-speeder.*
(iii) *The function* f *is an* M*-speeder.*

Proof. If (i) holds, then applying Theorem 2 to $f|_{[c,d]}$ shows that $f|_{[c,d]}$ is a strong M-speeder, which implies that f itself is a strong M-speeder, which implies that f is an M-speeder, which certainly implies that there are $a \leq c < d \leq b$ such that $\frac{|f(d)-f(c)|}{d-c} > M$. □

Say $f : [a,b] \to \mathbb{R}$ is a **strong infinite-speeder** if there is $c \in [a,b]$ such that for all $M, \delta > 0$ there is $x \in [a,b]$ with $0 < |x-c| < \delta$ such that $\frac{|f(x)-f(c)|}{x-c} > M$. The proof of Proposition 6 shows that every discontinuous function is a strong infinite-speeder. A function f is a strong infinite-speeder if and only if $\overline{D}_f(c) = \infty$ or $\underline{D}_f(c) = -\infty$ for some $c \in [a,b]$, so there are continuous strong infinite-speeders (Example 7a)) but no differentiable ones.

3.3. **Speeding in Metric Spaces.** The notion of M-speeder can easily be adapted to functions taking values in a metric space: if (X,d) is a metric space and $M > 0$, we say a function $f : [a,b] \to X$ is an **M-speeder** if for all $\delta > 0$ there are $x, y \in [a,b]$ with $0 < |x-y| < \delta$ and $\frac{d(f(x),f(y))}{|x-y|} > M$. Our main result continues to hold in this context.

Theorem 9. *Let (X,d) be a metric space, and let $f : [a,b] \to X$ be a function. If*

$$0 < M < \frac{d(f(a),f(b))}{b-a},$$

then f is an M-speeder. Indeed there is $c \in [a,b]$ such that for all $\delta > 0$ there is $x \in [a,b]$ such that $0 < |c-x| < \delta$ and $\frac{d(f(c),f(x))}{|x-c|} > M$. If f is continuous, we may take $c \in (a,b)$.

Proof. Let $0 < M < \frac{d(f(a),f(b))}{b-a}$. Define $\mathcal{D} : [a,b] \to \mathbb{R}$ by $x \mapsto d(f(x),f(a))$. Since $\mathcal{D}(a) = 0$ and $\mathcal{D}(b) = d(f(a),f(b))$, we have $M < \left|\frac{\mathcal{D}(b)-\mathcal{D}(a)}{b-a}\right|$, so applying Theorem 2 to $\mathcal{D} : [a,b] \to \mathbb{R}$, we get $c \in [a,b]$ such that for all $\delta > 0$ there is $x \in [a,b]$ with $0 < |c-x| < \delta$ and $\left|\frac{\mathcal{D}(c)-\mathcal{D}(x)}{c-x}\right| > M$. By the Reverse Triangle Inequality [GT, Prop. 2.1] we have

$$\frac{d(f(c),f(x))}{|c-x|} \geq \left|\frac{\mathcal{D}(c)-\mathcal{D}(x)}{c-x}\right| > M.$$

The proof that we can take $c \in (a,b)$ if f is continuous is the same as in Theorem 2. $\square$

4. Debriefing on Mean Value Inequalities

To ticket differentiable speeders, one only needs the following weakening of MVT:

Theorem 10 (Mean Value Inequality, henceforth "MVI")**.** *Let $f : [a,b] \to \mathbb{R}$ be a continuous function whose restriction to (a,b) is differentiable. If $m \leq f'(x) \leq M$ for all $x \in (a,b)$, then*

$$m(y-x) \leq f(y) - f(x) \leq M(y-x) \tag{3}$$

for all $x, y \in [a,b]$ with $x \leq y$.

Our Theorem 2 is essentially Theorem 10 with the differentiability hypothesis removed: in particular, Theorem 2 implies Theorem 10. It is mentioned in [Cl19, Thm. 1.8] that one can use Real Induction to give a nice proof of MVI. The proof itself did not make it into the published version, but it is very close to the proof of Theorem 4 given here. So the main point of this note is that the proof of MVI actually works to show something stronger in which differentiability need not be assumed.[3]

MVI is also a well-known result. One can find in the literature a curious palaver about whether it is "more natural" than MVT and whether one should promote MVI to MVT's honorable position in the calculus curriculum. The suggestion that MVT should somehow be replaced by the weaker MVI goes back at least to 1967 works of Bers [Be67] and Cohen

[3]The use of Real Induction is *not* the point: it's just a proof technique I happen to like. One could certainly give a proof using least upper bounds instead – even essentially the same proof – and you may prefer to do so.

[Co67] and has also been made by Dieudonné [D] and Boas [Bo81], among others.

I'm afraid I have never been able to understand this controversy, but it may have something to do with the following point: we can differentiate vector-valued functions $f : [a,b] \to \mathbb{R}^N$, but when $N \geq 2$ it is hopeless to try to equate $\frac{f(b)-f(a)}{b-a}$ with $f'(c)$ at some $c \in [a,b]$, even after taking norms. For instance the function $f : [0,2\pi] \to \mathbb{R}^2$ given by $f(x) = (\cos x, \sin x)$ has $\frac{f(2\pi)-f(0)}{2\pi} = 0$ but $||f'(c)|| = 1$ for all c. In contrast MVI does extend to this setting: see e.g. [R, Thm. 9.19]. (I am still puzzled: that a result fails to generalize to vector-valued functions seems like a strange reason not to like or use it in the scalar-valued case!)

In fact MVI extends to functions $f : [a,b] \to W$, where $(W, ||\cdot||)$ is a normed linear space over $\mathbb{R}$. (We take precisely the usual definition of derivative: $f'(x) := \lim_{h \to 0} \frac{f(x+h)-f(x)}{h}$.)

Theorem 11. *Let $(W, ||\cdot||)$ be a normed linear space, and let $f : [a,b] \to W$ be a continuous function whose restriction to (a,b) is differentiable. If $||f'(x)|| \leq M$ for all $x \in (a,b)$, then*

$$||f(b) - f(a)|| \leq M(b-a).$$

Proof. See [C, p. 122]. □

Theorem 11 is a consequence of our Theorem 9. Unlike Theorem 10, Theorem 11 does not really use $f'(x)$ itself but only its norm $||f'(x)||$. This suggests a further generalization: for a metric space X and a function $f : [a,b] \to X$, one can define a **metric derivative**

$$|f'| : [a,b] \to \mathbb{R},\ |f'|(x) = \lim_{h \to 0} \frac{d(f(x+h), f(x))}{h},$$

provided this limit exists for all $x \in [a,b]$. This definition appears in the literature, e.g. [AGS, p. 24]. The following generalization of Theorem 11 is also a consequence of our Theorem 9.

Theorem 12. *Let (X,d) be a metric space, and let $f : [a,b] \to X$ be a continuous function whose restriction to (a,b) is metrically differentiable: $|f'|(x)$ exists for all $x \in (a,b)$. If $|f'|(x) \leq M$ for all $x \in (a,b)$, then*

$$d(f(b), f(a)) \leq M(b-a).$$

But let us end by reiterating our main point: actually we need not make any differentiability assumption at all, and it may be better not to do so.

References

[AGS] L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures.* Second edition. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008.

[Be67] L. Bers, *Classroom Notes: On Avoiding the Mean Value Theorem.* Amer. Math. Monthly 74 (1967), 583.

[Bo81] R.P. Boas, *Who Needs Those Mean-Value Theorems, Anyway?* The Two-Year College Mathematics Journal 12 (1981), 178–181.

[C] W. Cheney, *Analysis for applied mathematics.* Graduate Texts in Mathematics, 208. Springer-Verlag, New York, 2001.

[C-GT] P.L. Clark, *General topology.* `http://alpha.math.uga.edu/~pete/pointset.pdf`

[Cl19] P.L. Clark, *The instructor's guide to real induction.* Math. Mag. 92 (2019), 136–150.

[Co67] L.W. Cohen, *Classroom Notes: On Being Mean to the Mean Value Theorem.* Amer. Math. Monthly 74 (1967), 581–582.

[D] J. Dieudonné, *Foundations of Modern Analysis* 2nd edition, Academic Press, New York, 1969.

[M] J. Munkres, *Topology.* Second edition. Prentice Hall, Inc., Upper Saddle River, NJ, 2000.

[R] W. Rudin, *Principles of mathematical analysis.* Third edition. International Series in Pure and Applied Mathematics. McGraw-Hill, New York-Auckland-Düsseldorf, 1976.