Do we need Mean Value Theorem to prove $f'(x) = 0$ on (a, b) implies that f = constant on (a, b)?

By Ng Tze Beng

If we use the *Mean Value Theorem* here, then it is an immediate consequence of it. What does that mean? Basically that means the *Mean Value Theorem* does all the work for us. So how is the *Mean Value Theorem* proved? One proof involves the use of the *Extreme Value Theorem*. How is that proved? It involves the use of the completeness property of the real numbers. So we can ask the question: If we can define the notion of differentiability for a function from a non complete ordered field such as the rational numbers into itself, then does the *Mean Value Theorem* hold? We can obviously find examples of function from the rational numbers to the rational numbers where the *Mean Value Theorem* or *Rolle's Theorem* does not hold. An easy example would be a cubic polynomial function whose derived function is a quadratic with real non-rational roots, for instance $f(x) = x^3 - 6x + 1$. Is there a function from the rational numbers or an appropriate subset of it to the rational numbers whose derived function is zero but f is non-constant? An appropriate subset would be an intersection of a non-empty open interval with the rational numbers. Think of the holes that the rational numbers have. An easy example would be a function f defined by $f(x) = 1$ for any rational number $x > \sqrt{2}$ and $f(x) = 2$ for any rational number $x < \sqrt{2}$. f is not a constant function. Then the function $f: \mathbf{Q} \to \mathbf{Q}$ is differentiable and $f'(x) = 0$ for any rational number x. A more sophisticated example will be provided by g: $(-\sqrt{2}, \sqrt{2}) \cap \mathbf{Q} \to \mathbf{Q}$ where $g(x) = 1/2^{2n+2}$ for $x \in (\sqrt{2}/2^{n+1}, \sqrt{2}/2^n) \cap \mathbf{Q}$ or $x \in (-\sqrt{2}/2^n, -\sqrt{2}/2^{n+1}) \cap \mathbf{Q}$, n an integer ≥ 0 and $g(0) = 0$. Then g is differentiable and $g'(x) = 0$ for all x in $(-\sqrt{2}, \sqrt{2}) \cap \mathbf{Q}$ and g is not a constant function.

Theorem 1. $f'(x) = 0$ on (a, b) implies that f = constant on (a, b).

Now we prove the above using only the completeness property of the real numbers. We assume $b > a$. The proof is by contradiction. Suppoes that f is not constant. Then there exist u, v in (a, b), $u < v$ such that $f(u) \neq f(v)$. This means $f(v) - f(u) \neq 0$. Then we shall make use of the difference quotient $\frac{f(v) - f(u)}{v - u} = C \neq 0$ to deduce a contradiction. Suppose now that $C > 0$.

For now let us suppose that (not assuming anything on C)

$$f(v) - f(u) = C(v - u). \quad \text{-------------------------} \quad (*)$$

We are going to bisect the inteval $[u, v]$, pick the next interval from this bisection and continue bisecting in like manner.

Take the mid point $w = \frac{u + v}{2}$ of $[u, v]$. Then either

$$f(v) - f(w) \geq C(v - w) \quad \text{-----------------------------} \quad (1)$$

or

$$f(w) - f(u) \geq C(w - u). \quad \text{------------------------------} (2)$$

This is because if both (1) and (2) do not hold, then we would have

$$f(v) - f(w) < C(v - w) \text{ and } f(w) - f(u) < C(w - u)$$

which would imply that $f(v) - f(u) < C(v - u)$ contradicting (*).

If (1) holds, then we name $u_1 = w$ and $v_1 = v$. If (1) does not hold we name $u_1 = u$ and $v_1 = w$.

Let $k = (v - u)$. Then $|v_1 - u_1| = k/2$ and

$$f(v_1) - f(u_1) \geq C(v_1 - u_1). \quad \text{--------------------------} \quad (*1)$$

Obviously, $[u_1, v_1] \subset [u, v]$, $u \le u_1 < v_1 \le v$, $|u_1 - u| \le |v - u|/2 = k/2$ and $|v - v_1| \le |v - u|/2 = k/2$. We next take the mid point $w_1 = \dfrac{u_1 + v_1}{2}$ of $[u_1, v_1]$. Then we shall have either

$$f(v_1) - f(w_1) \ge C(v_1 - w_1) \quad \text{--------------------- --------} \; (3)$$

or

$$f(w_1) - f(u_1) \ge C(w_1 - u_1). \quad \text{------------------------------} (4)$$

Again this is because if both (3) and (4) do not hold then we would have $f(v_1) - f(w_1) < C(v_1 - w_1)$ and $f(w_1) - f(u_1) < C(w_1 - u_1)$ implying $f(v_1) - f(u_1) < C(v_1 - u_1)$ thus contradicting (*1).

If (3) holds, then we name $u_2 = w_1$ and $v_2 = v_1$. If (3) does not hold we name $u_2 = u_1$ and $v_2 = w_1$. Then $|v_2 - u_2| = k/2^2$,

$$f(v_2) - f(u_2) \ge C(v_2 - u_2). \quad \text{--------------------------} \quad (*2)$$

Obviously, $[u_2, v_2] \subset [u_1, v_1]$, $u_1 \le u_2 < v_2 \le v_1$, $|u_2 - u_1| \le |v_1 - u_1|/2 = k/2^2$ and $|v_1 - v_2| \le |v_1 - u_1|/2 = k/2^2$.

In this way, we obtained a nested sequence,

$$\subset [u_n, v_n] \subset \ldots \ldots \subset [u_2, v_2] \subset [u_1, v_1] \subset [u, v],$$

with the length of the interval $[u_n, v_n] = \dfrac{v-u}{2^n}$ approaches 0 as n tends to infinity,

an increasing sequence (not necessarily strictly increasing),

$$u_1 \le u_2 \le u_3 \le \ldots \ldots \le u_n \le \ldots,$$

satisfying, for all n, $u_n < v_n \le v$, $|u_n - u_{n-1}| \le k/2^n$ ------------------ (5)

and a decreasing sequence (not necessarily strictly decreasing),

$$v_1 \ge v_2 \ge v_3 \ge \ldots \ldots \ge v_n \ge \ldots,$$

satisfying, for all n, $u \le u_n < v_n$, $|v_n - v_{n-1}| \le k/2^n$ ------------------ (6)

and

$$f(v_n) - f(u_n) \ge C(v_n - u_n). \quad \text{-------------------------} \quad (*n)$$

Now we have a choice to proceed. We can use the Weierstrass characterization of completeness to conclude that the nested sequence $\{[u_n, v_n]\}_n$ must have a unique intersection. i.e., there is exactly one point x that belongs to $[u_n, v_n]$ for all n.

We can also note that the sequence or set $\{u_n\}$ is bounded above by v by (5). Therefore, by the completeness property of the real numbers, $\{u_n\}$ has a least upper bouned or supremum in **R** also denoted by x, *i.e.* $x = sup\ \{u_n\}$. Also by the completeness property of the real numbers since the sequence $\{v_n\}$ is bounded below by u by (6) it has a greatest lower bound or infimum in **R** denoted by y, that is, $y = inf\{v_n\}$.

We claim that $x = y$. From (5) any v_n is an upper bound for $\{u_n\}$. Hence $x = sup\ \{u_n\} \le v_n$ for each n. Therefore, x is a lower bound for $\{v_n\}$ and so $x \le y = inf\{v_n\}$. Can x be bigger than y? Suppose $x > y$. Then since $x = sup\ \{u_n\}$, there exists a u_j such that $y < u_j$. But since $y = inf\{v_n\}$ and $u_j < v_n$ for all n, $u_j \le y = inf\{v_n\}$. This contradicts $y < u_j$. Hence $x = y$. (It is obvious that x cannot be strictly less than y. Observe this as follows. For all n, $u_n \le x \le y \le v_n$. If $x < y$, then there exists an integer n such that $(v_n - u_n) = (v - u) / 2^n < (y - x)$. This is not possible since $(y - x) \le (v_n - u_n)$.)

In particular, we have $u_n \le x \le v_n$ for all n. That is the same as saying $x \in [u_n, v_n]$ for all n.

Next we shall show that $f'(x) \ge C$. That is $\lim_{y \to x} \dfrac{f(y) - f(x)}{y - x} \ge C$.

If on the contrary $\lim_{y \to x} \dfrac{f(y) - f(x)}{y - x} < C$, then there exists a $\delta > 0$ such that for all y with $0 < |y - x| < \delta$ we have

$$\frac{f(y) - f(x)}{y - x} < C \quad \text{--------------------------} \; (A).$$

If we can show that for any $\delta > 0$, we can find a x_δ such that $0 < |x_\delta - x| < \delta$ but $\frac{f(x_\delta) - f(x)}{x_\delta - x} \geq C$, then no $\delta > 0$ can exist so that (A) holds and so we can conclude that $f'(x) \geq C$. We shall now proceed to do just that.
For any $\delta > 0$, $x - \delta < x = sup\ \{u_n\}$ and so there exists integer N such that $x - \delta < u_N \leq x$. Likewise using the fact that $x = inf\{v_n\}$, there exists an integer M such that $x \leq v_M < x + \delta$. *Let* $K = \max(N, M)$. Then we have $x - \delta < u_N \leq u_K \leq x \leq v_K \leq v_M < x + \delta$ and $u_K < v_K$. *T*his means that both u_K and v_K lie in the interval $(x - \delta, x + \delta)$. If $x = u_K$, then let $x_\delta = v_K$. If $x = v_K$, then let $x_\delta = u_K$. In either case $\frac{f(x_\delta) - f(x)}{x_\delta - x} = \frac{f(v_K) - f(u_K)}{v_K - u_K} \geq C$ by (*K). If $u_K < x < v_K$, then as in the beginning of the proof one of the following must be true:

$$f(v_K) - f(x) \geq C(v_K - x) \text{ -------------- } (7)$$
$$f(x) - f(u_K) \geq C(x - u_K) \text{ ---------------- } (8).$$

This is because if both (7) and (8) do not hold, we would then get $f(v_K) - f(x) < C(v_K - x)$ and $f(x) - f(u_K) < C(x - u_K)$ implying that $f(v_K) - f(u_K) < C(v_K - u_K)$ *contradicting* (*K). If (7) holds, then we let $x_\delta = v_K$ and if (8) holds we let $x_\delta = u_K$. We then have

$$\frac{f(x_\delta) - f(x)}{x_\delta - x} \geq C \text{ ----------------- } (9).$$

Hence we conclude that if $C > 0$ this would give us $f'(x) \geq C > 0$ thus contradicting $f'(x) = 0$. Thus $C \leq 0$.
Suppose $C < 0$. We can either apply the above argument with the inequality "≥" replaced by "≤" throughout or we can consider using the function $g = -f$. We can rewrite (*) as

$$-f(v) - (-f(u)) = -C(v - u).$$

That is

$$g(v) - g(u) = (-C)(v - u). \text{ -------------------------- } (**)$$

Now $-C > 0$ and so (**) is similar to (*) and so we can concude that we can find an x in $[u, v] \subseteq (a, b)$ such that $g'(x) = -f'(x) \geq -C$, that is $f'(x) \leq C < 0$ thus contradicting $f'(x) = 0$. Therefore, $C = 0$. and so f must be a constant function.

Note that we have actually proved the following result:

Theorem 2: If $f:[a, b] \to \mathbf{R}$ is differentiable, then for any u, v in $[a, b]$ with $u < v$, there exists a point x in $[u, v]$ such that $f'(x) \geq \frac{f(v) - f(u)}{v - u}$.

Reversing the inequality "≥" by "≤" throughout, starting with (1) and (2) we would obtain the following:

Theorem 2': If $f:[a, b] \to \mathbf{R}$ is differentiable, for any u, v in $[a, b]$ with $u < v$ there exists a point x in $[u, v]$ such that $f'(x) \leq \frac{f(v) - f(u)}{v - u}$.

Theorem 3: If $f'(x) < 0$ on (a, b), then f is decreasing on (a, b).

Proof.
Take any u, v in (a, b) with $u < v$, then by Theorem 2, there exists a point x in $[u, v]$ such that $\frac{f(v)-f(u)}{v-u} \leq f'(x) < 0$. Hence $f(v) - f(u) < 0$ and so $f(v) < f(u)$. That means f is decreasing on (a, b).

Theorm 4 (Weak Mean Value Theorem). If $m \leq f'(x) \leq M$ on $[a, b]$, then for any u, v in $[a, b]$ with $u < v$,

$$m(v - u) \leq f(v) - f(u) \leq M(v - u).$$

Proof. By Theorem 2, $f(v) - f(u) \leq f'(y)(v - u)$ for some y in $[u, v]$ and so $f(v) - f(u) \leq M(v - u)$. By Theorem 2', there is a point y in $[u, v]$ such that $f(v) - f(u) \geq f'(y)(v - u) \geq m(v - u)$. Therefore, $m(v - u) \leq f(v) - f(u) \leq M(v - u)$.