ALMOST EVERYWHERE BOUNDS AND EVERYWHERE BOUNDS ARE EQUIVALENT FOR DERIVATIVES

WILLIE WY WONG

The main purpose of these notes is to record an alternative proof of the following theorem, which I learned from Pietro Majer[1]:

Theorem 1 (Main Theorem)**.** *Let* $[a,b]$ *be a closed interval in* $\mathbf{R}$, *and* $f : [a,b] \to \mathbf{R}$ *be a continuous function that is differentiable at every* $x \in (a,b)$. *Then* $\operatorname{ess\,sup}_{(a,b)} f' = \sup_{(a,b)} f'$.

For the benefit of the readers, we will recall the definition of the essential supremum. This however requires some preliminary discussions, which I will defer after explaining the motivation of these notes.

The proof that Majer give uses two concepts that are of the level of a "second course" or "third course" in real analysis: one of which is the *Vitali Covering Theorem* which are usually only covered in a course on measure theory of in harmonic analysis; the second is the concept of the Lebesgue measure of a set. It turns out that the theorem above can be explained using only the notion of an *outer measure*, which is definitely approachable for students in a first course in real analysis (viz. Lebesgue's characterization of Riemann integrable functions, which talks about discontinuity points being a "null set"). Given that the theorem can be understood by such students, it is natural to ask whether a *proof* can be written using more elementary language and techniques. It turns out that as Majer's proof does not use the full strength of the Vitali Theorem, only relying on a "finite approximation" version, we can easily replace it using the fairly elementary notion of *compactness*.

Next let us provide enough definitions for a first year analysis student to understand the theorem statement.

First, let $O \subseteq \mathbf{R}$ be an open set. If $x \in O$, we can define the *connected component* of x as follows: let $a = \inf\{y \in \mathbf{R} \mid [y,x] \subseteq O\}$ and let $b = \sup\{z \in \mathbf{R} \mid [x,z] \subseteq O\}$. Then we see that $(a,b) \subseteq O$ and $a, b \notin O$, and hence (a,b) is the largest open interval of O that contains x. That the rationals (a countable subset) is dense in $\mathbf{R}$ implies that every connected component of O contains at least one rational, and hence we have the well-known lemma

Lemma 2. *If* $O \subseteq \mathbf{R}$ *is open, then* O *is the disjoint union of a countable (finite or infinite) collection of open intervals.*

This means that we have a well-defined notion of total length for open sets.

Date: May 31, 2024.

WWY Wong is supported by a Simons Collaboration Grant, and wishes to also acknowledge the hospitality of the Mathematics Division of the National Center for Theoretical Sciences in Taipei, Taiwan, where these notes are written.

[1]See `https://mathoverflow.net/a/471863/3948`.

Definition 3.

(1) Given $-\infty < a < b < \infty$, the length of (a,b) is defined to be $|(a,b)| = b - a$.
(2) The length of an unbounded interval is defined to be $+\infty$; the length of the empty set is defined to be 0.
(3) Given an arbitrary open set O, we define its length by summing up the lengths of its connected components. More precisely, denote by $\mathscr{O}$ the countable set of disjoint open intervals such that $\cup\mathscr{O} = O$, we define

$$|O| = \sum_{I\in\mathscr{O}} |I|$$

with the sum allowed to take values in $[0,\infty]$.

We can now describe sets that are "arbitrarily small" in the sense of length.

Definition 4 (Null sets)**.** A subset $N \subseteq \mathbf{R}$ is said to be a *null set* if, for every $\epsilon > 0$, there exists an open set $O \supseteq N$ with $|O| < \epsilon$.

And finally we can define the notion of the essential supremum.

Definition 5. Given $f : [a,b] \to \mathbf{R}$, its *essential supremum* is defined to be

$$\operatorname{ess\,sup} f := \inf\Big\{\lambda \in \mathbf{R} \;\Big|\; f^{-1}((\lambda,\infty)) \text{ is a null set }\Big\}.$$

1. Initial Reductions Toward A Proof

From Definition 5, it is clear that $\operatorname{ess\,sup} g \leq \sup g$ for any function g. So the Main Theorem is proved provided that we can show, under the given hypotheses, that $\sup f'$ is not strictly larger than $\operatorname{ess\,sup} f'$. Now, set $m = \operatorname{ess\,sup} f'$, and let us consider the function $h(x) = f(x) - mx$. Then we have $h'(x) = f'(x) - m$. By definition $\operatorname{ess\,sup} h' = \operatorname{ess\,sup} f' - m = 0$, Our Theorem would follow if we can show that $0 = \operatorname{ess\,sup} h' \geq \sup h'$.

At this point, we will apply the standard result

Proposition 6. *If $f : [a,b] \to \mathbf{R}$ is continuous and differentiable in the interior of the interval, then the following are equivalent:*

(1) $f'(x) \leq 0$ for every $x \in (a,b)$;
(2) f is a decreasing function.

Proof. To show (1) $\implies$ (2), observe that if f is non-decreasing, there exists $c < d$ in its domain such that $f(c) < f(d)$. But the mean value theorem for differentiable functions implies that for some $w \in (c,d)$ we have $f'(w) = \frac{f(d)-f(c)}{d-c} > 0$.

To show (2) $\implies$ (1), use that $f'(x) = \lim_{t\to 0} \frac{f(x+t)-f(x)}{t}$ by definition, and when f is decreasing the quotient is always non-positive, so the limit is also non-positive. □

Hence our Main Theorem will follow from

Theorem 7 (Reduced Theorem)**.** *If $f : [a,b] \to \mathbf{R}$ is continuous and differentiable in the interior of the interval, and* $\operatorname{ess\,sup} f' \leq 0$, *then f is a decreasing function.*

We will prove Theorem 7 by contradiction, that is, we will show that the hypotheses of the theorem together with the assumption that f is non-decreasing will lead to a contradiction. Again noting that f is non-decreasing can be captured by the fact that there is some $[c,d] \subseteq [a,b]$ such that $f(c) < f(d)$. For convenience denote temporarily by $m = \frac{f(d)-f(c)}{d-c} > 0$ the slope of f through c and d. Taking $k : [c,d] \to \mathbf{R}$ to be defined by $k(x) = f(x) - \frac{1}{2}mx$. We have first that

$$k'(x) = f'(x) - \frac{1}{2}m \implies \text{ess sup}\, k' \leq -\frac{1}{2}m < 0.$$

Additionally, we still have

$$k(d) - k(c) = f(d) - f(c) - \frac{1}{2}m(d-c) = \frac{1}{2}m(d-c) > 0.$$

And hence, Theorem 7 will follow from the following:

Theorem 8. *There does not exist a function $f : [a,b] \to \mathbf{R}$ that satisfies simultaneously*

- *f is continuous and differentiable on (a,b), with* ess sup $f' < 0$,
- *and $f(a) < f(b)$.*

2. Preliminary Lemmas

In this section we recall some basic facts about differentiable functions.

Proposition 9. *If f is differentiable at x_0 with $f'(x_0) < 0$, then there exists an $\epsilon > 0$ such that*

- *every $y \in (x_0, x_0 + \epsilon)$ satisfies $f(y) < f(x)$;*
- *every $y \in (x_0 - \epsilon, x_0)$ satisfies $f(y) > f(x)$.*

Proof. Let $\delta < \frac{1}{2}|f'(x_0)|$, then the definition of differentiability means that there exists $\epsilon > 0$ such that on $(x_0 - \epsilon, x_0 + \epsilon)$ we have $|f(x) - f(x_0) - f'(x_0)(x - x_0)| < \delta|x - x_0|$. By the triangle inequality this gives

$$f'(x_0)(x - x_0) - \delta|x - x_0| \leq f(x) - f(x_0) \leq f'(x_0)(x - x_0) + \delta|x - x_0|.$$

When $x > x_0$, the latter inequality yields

$$f(x) - f(x_0) \leq \frac{1}{2}f'(x_0)(x - x_0) < 0.$$

When $x < x_0$, the former inequality gives

$$0 < \frac{1}{2}f'(x_0)(x - x_0) \leq f(x) - f(x_0).$$

And the proposition follows. □

Remark 10. Note that after replacing ϵ by $\epsilon/2$, we can also require that $f(x_0 + \epsilon) < f(x) < f(x_0 - \epsilon)$.

Proposition 11. *Let $f : [a,b] \to \mathbf{R}$ be differentiable at $x_0 \in (a,b)$. Then there exists some $\epsilon > 0$ and $M > 0$ such that for every c,d satisfying $x_0 - \epsilon < c < x_0 < d < x_0 + \epsilon$ we have $\left|\frac{f(d)-f(c)}{d-c}\right| < M$.*

Proof. Let $M = |2f'(x_0)|$. The definition of differentiability means that there exists an $\epsilon > 0$ such that on $(x_0 - \epsilon, x_0 + \epsilon)$ we have $|f(x) - f(x_0) - f'(x_0)(x - x_0)| < |f'(x_0)||x - x_0|$. This implies that $|f(x) - f(x_0)| < M|x - x_0|$. Since we have $c < x_0 < d$ we have by triangle inequality

$$|f(d) - f(c)| \leq |f(d) - f(x_0)| + |f(x_0) - f(c)| < M\Big(|d - x_0| + |x_0 - c|\Big) = M(d - c).$$

The claim follows. □

3. Main Proof

The proof of Theorem 8 has two main steps.

3.1. **Increase in Slope.**

Lemma 12. *Let $\lambda > 1$ be fixed. Suppose $f : [a,b] \to \mathbf{R}$ is continuous and differentiable on (a,b), with* $\operatorname{ess\,sup} f' < 0$, *and suppose $f(a) < f(b)$. Then there exists a subinterval $[c,d] \subseteq [a,b]$ such that*

$$\frac{f(d)-f(c)}{d-c} \geq \lambda \frac{f(b)-f(a)}{b-a}.$$

Proof. By hypothesis, there exists an open set O of $\mathbf{R}$ such that $|O| < \lambda^{-1}(b-a)$ and

$$O \supseteq \{a,b\} \cup (f')^{-1}((\operatorname{ess\,sup} f', \infty)).$$

Since O is open, we have $K = [a,b] \setminus O$ is closed and bounded, hence compact. Additionally, we know that $x \in K$ implies f is differentiable at x with $f'(x) \leq \operatorname{ess\,sup} f' < 0$. So by Proposition 9 and Remark 10, for each $x \in K$ there exists $\epsilon_x > 0$ such that every $y \in (x, x+\epsilon_x]$ satisfies $f(y) < f(x)$ and every $y \in [x-\epsilon_x, x)$ satisfies $f(y) > f(x)$. By shrinking ϵ_x if necessary we can ensure $x \pm \epsilon_x \in [a,b]$. The intervals $\{(x-\epsilon_x, x+\epsilon_x)\}$ form an open cover of K; compactness implies the existence of a finite subcover given by

$$\{(x_i - \epsilon_i, x_i + \epsilon_i)\}_{i \in \{1,\dots,N\}}.$$

We may assume that this subcover is minimal, in the sense that removing any of the intervals will leave the remaining set no longer a covering of K. Thus we may assume that whenever $i < j$, we have

$$x_i < x_j, \quad x_i - \epsilon_i < x_j - \epsilon_j, \quad x_i + \epsilon_i < x_j + \epsilon_j.$$

We next construct points u_i, v_i using the following rules:

- $u_1 = x_1 - \epsilon_1$, $v_N = x_N + \epsilon_N$.
- If $x_i + \epsilon_i < x_{i+1} - \epsilon_{i+1}$, we set $v_i = x_i + \epsilon_i$ and $u_{i+1} = x_{i+1} - \epsilon_{i+1}$.
- Else, choose some $z \in [x_{i+1} - \epsilon_{i+1}, x_i + \epsilon_i]$, and set $u_{i+1} = v_i = z$.

At the end of this construction, we have the points

$$a \leq u_1 < x_1 < v_1 \leq u_2 < x_2 < v_2 \leq u_3 \ldots < x_i < v_i \leq u_{i+1} < x_{i+1} < \ldots < x_N < v_N \leq b.$$

with the property that

- $\cup_{i=1}^N [u_i, v_i]$ covers K;
- $f(u_i) > f(v_i)$ for every $i \in \{1, \dots, N\}$.

Observe that $\tilde{O} := (a,b) \setminus \cup_{i=1}^N [u_i, v_i]$ is an open subset of O by definition, and we can write

$$\tilde{O} = (a, u_1) \cup (v_1, u_2) \cup \cdots \cup (v_i, u_{i+1}) \cup \cdots \cup (v_{N-1}, u_N) \cup (v_N, b)$$

as a finite union of (possibly empty) open intervals, which we label $I_0, \dots, I_N$. Necessarily we have

$$|\tilde{O}| \leq |O| < \lambda^{-1}(b-a).$$

Next we consider the telescoping sum

$$f(b) - f(a) = f(b) - f(v_N) + \Big(\underbrace{\sum_{i=1}^N f(v_i) - f(u_i)}_{<0} + \sum_{i=1}^{N+1} f(u_{i+1}) - f(v_i) \Big) + f(u_1) - f(a).$$

So, in terms of the intervals $I_0, \dots, I_N$, we can write (technically we need to cull any empty intervals from the list; those corresponds to situations where $u_{i+1} = v_i$ or $a = u_1$ or $b = v_N$, in which case the corresponding contribution to the sum is null anyway)

$$f(b) - f(a) < \sum_{i=0}^{N} f(\sup I_i) - f(\inf I_i).$$

Divide now by $b - a$ we can write

$$\frac{f(b) - f(a)}{b - a} < \frac{|\tilde{O}|}{b - a} \cdot \sum_{i=0}^{N} \frac{f(\sup I_i) - f(\inf I_i)}{|I_i|} \cdot \frac{|I_i|}{|\tilde{O}|}.$$

As $\sum \frac{|I_i|}{|\tilde{O}|} = 1$, we can interpret the sum as a weighted average. Since we have a lower bound for the weighted average, this is also a lower bound for at least one of the terms. Hence we conclude that there is a (non-degenerate) interval $I_j = [c, d]$ such that

$$\frac{f(b) - f(a)}{b - a} < \frac{|\tilde{O}|}{b - a} \frac{f(d) - f(c)}{d - c} < \lambda^{-1} \cdot \frac{f(d) - f(c)}{d - c}$$

as desired. □

3.2. **Driving the Contradiction.** Suppose a function satisfying the conditions of Theorem 8 exists on the interval $[a, b] = [a_0, b_0]$. Denote by $m = \frac{f(b)-f(a)}{b-a} > 0$. Notice that if $[c, d]$ is the interval provided by Lemma 12, then $f|_{[c,d]}$ is another function that satisfies the hypotheses of the Lemma itself. So we can apply the Lemma repeatedly and generated a sequence of nested intervals

$$[a_0, b_0] \supseteq [a_1, b_1] \supseteq [a_2, b_2] \dots$$

such that

$$\frac{f(b_i) - f(a_i)}{b_i - a_i} \geq \lambda^i m. \tag{13}$$

As a_i and b_i are bounded and monotonic, the two sequences converge to a_* and b_* respectively. We claim that $a_* = b_*$: suppose not, then the continuity of f implies that

$$\lim \frac{f(b_i) - f(a_0)}{b_i - a_i} = \frac{f(b_*) - f(a_*)}{b_* - a_*} < \infty.$$

This contradicts (13).

On the other hand, if $a_* = b_* = x_0$, by assumption we have f is differentiable at x_0. From Proposition 11 we extract some M, ϵ. As a_i and b_i converges to x_0, they are both eventually in $(x_0 - \epsilon, x_0 + \epsilon)$. But this implies that

$$\limsup \frac{f(b_i) - f(a_i)}{b_i - a_i} < M$$

which again contradicts (13).

Department of Mathematics, Michigan State University, East Lansing, MI, USA
Email address: wongwwy@math.msu.edu