Functions Having Finite Derivatives, Bounded Variation, Absolute Continuity, the Banach Zarecki Theorem and de La Vallée Poussin's Theorem

by Ng Tze Beng

We shall begin by examining the properties of the image under a function f of a set in which f has finite derivatives that are bounded by a constant**.** The first property we examine is the relation between the measure of such a set and the measure of its image. We state this property in the next theorem.

This result appears in Saks monograph on the theory of the integral and there are a number of proofs of the result. But I shall present a proof using some finiteness argument, a consequence of compactness and the triangle inequality.

Theorem 1. Suppose $f : [a, b] \to \mathbf{R}$ is a function. Suppose E is a subset of $[a, b]$ such that at each point x of E, f is differentiable and $|f'(x)| \le K$ for some constant $K \ge 0$. Then if m denotes the Lebesgue outer measure,

$$m(f(E)) \le K\, m(E) . \text{ ---------------------- } \textbf{(A)}$$

Proof . Now $E = \{ x \in [a, b] : |f'(x)| \le K \} \subseteq [a, b]$ and so E has finite outer measure. If E is finite or denumerable, then the set $f(E)$ is at most denumerable and so both $m(f(E))$ and $m(E)$ are zero and we have nothing to prove since both sides of the inequality are zero. We shall now assume that E is uncountably infinite. We may assume that neither a nor b is in E since adding or subtracting any finite number of points to E will not alter the inequality (A). Since the set of isolated points of E is countable, we may remove the set of isolated points from E without affecting the conclusion of the theorem, since the measure of a countable set is zero. We now assume that E has no isolated points.

For any $\varepsilon > 0$, by the definition of outer measure, there exists an open set U *in* $[a, b]$ such that $U \supseteq E$ and $m(U) \le m(E) + \varepsilon$.

Since for each e in E, $|f'(x)| \le K$, for $\varepsilon >$ o there exists a $\delta_e > 0$ such that

$$0 < |x-e| < \delta_e \Rightarrow \left| \left| \frac{f(x)-f(e)}{x-e} \right| - |f'(e)| \right| < \varepsilon.$$

and so

$$0 < |x-e| < \delta_e \Rightarrow \left| \frac{f(x)-f(e)}{x-e} \right| < |f'(e)| + \varepsilon \le K+\varepsilon.$$

Thus, we have,

$$|x-e| < \delta_e \Rightarrow |f(x)-f(e)| \le (K+\varepsilon)|x-e| . \text{ -------------- } (1)$$

Since U is open, we may choose $\delta_e > 0$ such that the open interval $(e - \delta_e , e + \delta_e) \subseteq U$. Denote $(e - \delta_e , e + \delta_e)$ by I_e. Then inequality (1) says that

$$x \in I_e \Rightarrow |f(x)-f(e)| \le (K+\varepsilon)|x-e| . \text{ ----------------- } (2)$$

Then the collection $\mathcal{C} = \{ I_e : e \in E\}$ covers E and the union $W = \cup \{V : V \in \mathcal{C}\} = \cup \{ I_e : e \in E\} \subseteq U$. In particular, the union W is open and so is a disjoint union of countable number of open intervals, i.e.,

$$W = \sqcup \{U_i : i \in B\},$$

where B the index set is a subset of the set $\mathbf{N}$ of natural numbers and each U_i is an open interval. We shall show next that for each i in B,

$$m(f(U_i \cap E)) \le (K+\varepsilon)\, m(U_i).\text{-------------------------- } (3)$$

Note that $U_i = \cup\{ I_e : e \in U_i \cap E \}$. Observe that each U_i is a path component of W.

Plainly for $e \in U_i \cap E$, $I_e \cap U_i \neq \varnothing$ and since $I_e \subseteq W$ and U_i is a path component of W, $I_e \subseteq U_i$. It follows that $\cup\{ I_e : e \in U_i \cap E \} \subseteq U_i$. For any x in U_i, $x \in I_e$ for some e in E, since $W = \cup \{ I_e : e \in E\}$ and so $I_e \cap U_i \neq \varnothing$. It follows, as in the above argument, that $I_e \subseteq U_i$ and so $e \in U_i \cap E$. Thus, $x \in I_e$ for some $e \in U_i \cap E$, that is, $x \in \cup\{ I_e : e \in U_i \cap E \}$ and so $U_i \subseteq \cup\{ I_e : e \in U_i \cap E \}$. This proves that $U_i = \cup\{ I_e : e \in U_i \cap E \}$.
Now take any $x < y$ in U_i. Since U_i is an open interval, the closed and bounded interval $[x, y]$ is contained in U_i. Now plainly the collection $\mathcal{B} = \{ I_e : e \in U_i \cap E \}$ is an open cover for $[x, y]$. Since $[x, y]$ is compact, there exists a finite subcover say

$$I_1, I_2, \ldots, I_n,$$

where $I_i = (e_i - \delta(e_i),\ e_i + \delta(e_i))$, for some e_i in E and $\delta(e_i)$ is as given in (1). *We* assume that the e_i's are ordered in an increasing order. Hence

$$[x, y] \subseteq I_1 \cup I_2 \cup \ldots \cup I_n$$

and $e_1 < e_2 < \ldots < e_n$.
We may assume that $x \in I_1$. This is seen as follows. If $x \notin I_1$, x must belong to I_j for some $1 < j \leq n$ and $x \notin I_i$ for for $1 \leq i < j$. Then $[x, y] \cap I_i = \varnothing$ for $1 \leq i < j$. It follows that $[x, y] \subseteq I_j \cup I_{j+1} \cup \ldots \cup I_n$ and so we can rename if need be I_j to be I_1, I_{j+1} to be I_2 and so on. By a similar argument we may assume that $y \in I_n$.

We may assume that $I_i \not\subset I_j$ for $j \neq i$. If $I_i \subseteq I_j$, then the collection of the I_k's without I_i still covers $[x, y]$ and so we can discard I_i and rename the I_j's.

We may also assume that $I_i \cap I_{i+1} \neq \varnothing$ for $1 \leq i \leq n-1$. We explain this as follows. Starting with I_1, suppose $I_1 \cap I_2 = \varnothing$. It follows that $e_1 + \delta(e_1) < e_2 - \delta(e_2)$. Then since $[x, y]$ is path connected, $I_1 \cap \cup\{ I_j : 1 < j \leq n\} \neq \varnothing$ implies for some $2 < j \leq n$, $I_1 \cap I_j \neq \varnothing$. Then $e_j - \delta(e_j) < e_1 + \delta(e_1) < e_2 - \delta(e_2)$. Hence, $\delta(e_j) > e_j - e_2 + \delta(e_2) > \delta(e_2)$ and so $e_j + \delta(e_j) > e_2 + \delta(e_2)$. Hence, $I_2 \subseteq I_j$. This contradicts that $I_2 \not\subset I_j$.
Therefore, $I_1 \cap I_2 \neq \varnothing$. This means $e_1 + \delta(e_1) > e_2 - \delta(e_2)$. We can repeat the same argument to show that $I_i \cap I_{i+1} \neq \varnothing$ for $i > 1$. For instance, take next the interval, $\left[\frac{e_1+\delta(e_1)+e_2+\delta(e_2)}{2}, y\right]$. Since $I_2 \not\subset I_1$, $e_2 + \delta(e_2) > e_1 + \delta(e_1)$ and so $\frac{e_1+\delta(e_1)+e_2+\delta(e_2)}{2} < e_2 + \delta(e_2)$. As $I_1 \cap I_2 \neq \varnothing$, $e_2 - \delta(e_2) < e_1 + \delta(e_1)$ so that $e_2 - \delta(e_2) < e_1 + \delta(e_1) < \frac{e_1+\delta(e_1)+e_2+\delta(e_2)}{2}$. Hence $\frac{e_1+\delta(e_1)+e_2+\delta(e_2)}{2} \in I_2$ but not in I_1.
Then $\{ I_j : 2 \leq j \leq n\}$ covers $\left[\frac{e_1+\delta(e_1)+e_2+\delta(e_2)}{2}, y\right]$. If it does not cover $\left[\frac{e_1+\delta(e_1)+e_2+\delta(e_2)}{2}, y\right]$, then there exists k such that $\frac{e_1+\delta(e_1)+e_2+\delta(e_2)}{2} < k < y$ and $k \notin \cup\{ I_j : 2 \leq j \leq n\}$ and so $k \in I_1$. This means $k < e_1 + \delta(e_1)$. But since $I_2 \not\subset I_1$, $e_2 + \delta(e_2) > e_1 + \delta(e_1)$ so that we have $k > \frac{e_1+\delta(e_1)+e_2+\delta(e_2)}{2} > e_1 + \delta(e_1)$ contradicting $k < e_1 + \delta(e_1)$. Hence, $\{ I_j : 2 \leq j \leq n\}$ covers $\left[\frac{e_1+\delta(e_1)+e_2+\delta(e_2)}{2}, y\right]$. By repeating the above argument on the interval, $\left[\frac{e_1+\delta(e_1)+e_2+\delta(e_2)}{2}, y\right]$ instead of $[x, y]$ and the cover $\{ I_j : 2 \leq j \leq n\}$, we can show that $I_2 \cap I_3 \neq \varnothing$. Thus, continuing the argument in this way, we have that $I_i \cap I_{i+1} \neq \varnothing$ for $1 \leq i \leq n-1$.
Therefore, we may assume that we have a sequence of points $x_1, x_2, \ldots, x_{n-1}$ such that

$$e_1 < x_1 < e_2 < x_2 < \ldots < e_{n-1} < x_{n-1} < e_n$$

and $x_i \in I_i \cap I_{i+1}$ for $1 \leq i \leq n-1$. Therefore, by (2) and the triangle inequality.

$$| f(x) - f(y)| \leq | f(x) - f(e_1)| + | f(e_1) - f(x_1)| + | f(x_1) - f(e_2)| + | f(e_2) - f(x_2)|$$

$+\cdots+|f(x_{n-2})-f(e_{n-1})|+|f(e_{n-1})-f(x_{n-1})|+|f(x_{n-1})-f(e_n)|+|f(e_n)-f(y)|$
$\le (K+\varepsilon)\{|x-e_1|+|e_1-x_1|+|x_1-e_2|+|e_2-x_2|+\cdots$
$+|x_{n-2}-e_{n-}|+|e_{n-1}-x_{n-1}|+|x_{n-1}-e_n|+|e_n-y|\}$
$\le (K+\varepsilon)\{|x-e_1|+|e_1-e_n|+|e_n-y|\}$
$\le (K+\varepsilon)\, m(I_1 \cup I_2 \cup \ldots \cup I_n) \le (K+\varepsilon)\, m(U_i)$.

Hence, the diameter of $f(U_i) \le (K+\varepsilon)\, m(U_i)$. It follows that $m(f(U_i \cap E)) \le (K+\varepsilon)$ $m(U_i)$. This proves (3).
Then using (3), we see that

$$\begin{aligned} m(f(E)) &= m(\cup\{f(U_i \cap E) : i \in B\}) \le \sum_{i\in B} m(f(U_i \cap E)) \\ &\le \sum_{i\in B}(K+\varepsilon)m(U_i) = (K+\varepsilon)m(W) \le (K+\varepsilon)m(U) \\ &\le (K+\varepsilon)(m(E)+\varepsilon). \end{aligned}$$

Since ε is arbitrary, we conclude that $m(f(E)) \le Km(E)$.

Theorem 2. Suppose $f:[a,b] \to \mathbf{R}$ is a measurable function. Suppose E is any measurable subset such that $f'(x)$ exists finitely for every x in E. Then

$$m(f(E)) \le \int_E |f'|,$$

where m is the Lebesgue outer measure.
Proof. Since f is measurable and finite on $[a, b]$, its Dini derivatives are measurable. (Banach Theorem). Consequently, f' is measurable on E and so $|f'|$ is measurable on E. Suppose now $g = |f'|$ is bounded on E, by a positive integer K, i.e., $|f'(x)| < K$ for each x in E. For any positive integer n and integer $i = 1, 2, \ldots, 2^n K$, let $E_{n,i} = g^{-1}\left(\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)\right) \cap E$. Define $g_n = \sum_{i=1}^{2^n K} \frac{i-1}{2^n}\chi_{E_{n,i}}$ for each positive integer n. Then (g_n) is a sequence of simple functions converging pointwise to g on E. In particular,

$$\int_E g_n \to \int_E g\ .$$

By Theorem 1, $m(f(E_{n,i})) \le \frac{i}{2^n} m(E_{n,i})$ for integer $i = 1, 2, \ldots, 2^n K$, Thus,

$$\begin{aligned} m(f(E)) = m(f(\bigcup_{i=1}^{2^n K} E_{n,i}) &\le \sum_{i=1}^{2^n K} \frac{i}{2^n} m(E_{n,i}) \\ &= \sum_{i=1}^{2^n K} \frac{i-1}{2^n} m(E_{n,i}) + \frac{1}{2^n}\sum_{i=1}^{2^n K} m(E_{n,i}) \\ &= \int_E g_n + \frac{1}{2^n} m(E). \quad \text{-----------------------} \quad (1) \end{aligned}$$

Therefore, $m(f(E)) \le \lim_{n\to\infty} (\int_E g_n + \frac{1}{2^n} m(E))$. Since, $\int_E g_n \to \int_E g$ and $\frac{1}{2^n} m(E) \to 0$, we conclude that

$$m(f(E)) \le \int_E g.$$

We now consider the case when g is unbounded. For each integer $k > 1$, let

$$E_k = g^{-1}([k-1, k)) \cap E.$$

Then it is obvious that E is a disjoint union of the E_k's. Note that on E_k, g is bounded by k. Hence, by what we have just shown, for each integer $k > 0$, $m(f(E_k)) \le \int_{E_k} g$. Therefore,

$$m(f(E)) \le \sum_{k=1}^{\infty} m(f(E_k) \le \sum_{k=1}^{\infty} \int_{E_k} g = \int_E g = \int_E |f'|.$$

This completes the proof of Theorem 2.

We have some easy consequences of the above theorems.

Theorem 3. Suppose f is defined and finite on $[a, b]$. Suppose $E = \{x \in [a, b]\colon f$ is differentiable at x and $f'(x) = 0\}$. Then $m(f(E)) = 0$.
Proof. By Theorem 1, $m(f(E)) \le 1/n\, m(E)$ for any positive integer n. Therefore, $m(f(E)) = 0$.

Recall a set is called a *null set* if its measure is zero.

Theorem 4. Suppose $f : [a, b] \to \mathbf{R}$ has a finite derivative at every point of $[a, b]$. Then f maps null sets onto null sets.
Proof. Suppose E is a null set in $[a, b]$. Then by Theorem 2,

$$m(f(E)) \le \int_E |f'| = 0 .$$

Hence $m(f(E)) = 0$. This proves the theorem.

Theorem 5. Suppose $f : [a, b] \to \mathbf{R}$ has a finite derivative at every point of $[a, b]$ and f' is Lebesgue integrable on $[a, b]$. Then for every closed and bounded interval $[c, d]$ in $[a, b]$,

$$\int_c^d |f'| \ge |f(d) - f(c)|.$$

Proof. Since f is continuous on $[a, b]$, $|f(d) - f(c)| \le m(f([c, d]))$. Since f is differentiable at every point of $[c, d]$, by Theorem 2,

$$m(f([c,d])) \le \int_{[c,d]} |f'| = \int_c^d |f'|.$$

It follows that $|f(d) - f(c)| \le \int_c^d |f'|$.

We can apply Theorem 5 to the next result.

Theorem 6. Suppose $f : [a, b] \to \mathbf{R}$ has a finite derivative at every point of $[a, b]$ and f' is Lebesgue integrable on $[a, b]$. Then f is absolutely continuous.
Proof. Since f' is Lebesgue integrable, $|f'|$ is also Lebesgue integrable on $[a, b]$. For each positive integer n, let $g_n = \min(|f'|, n)$. Then each g_n is Lebesgue integrable on $[a, b]$ and the sequence (g_n) converges pointwise to $|f'|$. In particular, for each n, $|g_n| = g_n \le |f'|$ and so by the Lebesgue Dominated Convergence Theorem,

$$\int_a^b g_n \to \int_a^b |f'|.$$

Hence, given any $\varepsilon > 0$, there exists a positive integer N such that

$$n \ge N \Rightarrow \left| \int_a^b |f'| - \int_a^b g_n \right| < \frac{\varepsilon}{2}.$$

It follows that

$$n \ge N \Rightarrow 0 \le \int_a^b (|f'| - g_n) < \frac{\varepsilon}{2}. \quad \text{-------------------------- (1)}$$

Now take $\delta = \frac{\varepsilon}{2N}$. Suppose $[a_i, b_i]$, $i = 1, 2, \ldots, k$ are non-overlapping intervals in $[a, b]$. If $\sum_{i=1}^{k} |b_i - a_i| < \delta$, then

$$\begin{aligned}\sum_{i=1}^{k} |f(b_i) - f(a_i)| &\le \sum_{i=1}^{k} \int_{a_i}^{b_i} |f'| \quad \text{by Theorem 5,}\\ &= \sum_{i=1}^{k} \int_{a_i}^{b_i} (|f'| - g_N) + \sum_{i=1}^{k} \int_{a_i}^{b_i} g_N \\ &\le \int_a^b (|f'| - g_N) + \sum_{i=1}^{k} \int_{a_i}^{b_i} N\end{aligned}$$

$$= \int_a^b (|f'| - g_N) + N \sum_{i=1}^{k} |b_i - a_i|$$
$$< \frac{\varepsilon}{2} + N\delta = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \text{, by (1).}$$

This shows that f is absolutely continuous on $[a, b]$.

Remark. Theorem 6 is Theorem 8.21 in Rudin's *Real and Complex Analysis* in an equivalent formulation.

More generally we may relax the requirement of everywhere differentiability but we need to impose the requirement that f maps null sets to null sets. This is a necessary condition for absolute continuity.

Theorem 7. Suppose $f : [a, b] \to \mathbf{R}$ is continuous and f' exists almost everywhere and is Lebesgue integrable on $[a, b]$. Suppose f maps null sets to null sets. Then f is absolutely continuous.

Proof. Let $E \subseteq [a, b]$ be the subset where f' exists at each point so that the measure of $[a, b] - E$ is zero. Then $|f'| = g$ almost everywhere, where $g(x) = |f'(x)|$ for x in E and $g(x) = 0$ for x outside E. Then there exists an increasing sequence of simple functions (g_n) converging pointwise to g almost everywhere and

$$\int_a^b g_n \to \int_a^b |f'| = \int_a^b g.$$

Thus, given any $\varepsilon > 0$, there exists a positive integer N such that

$$n \geq N \Rightarrow \left| \int_a^b g - \int_a^b g_n \right| = \int_a^b (g - g_n) < \frac{\varepsilon}{2}. \qquad \text{---------------- (1)}$$

Suppose $[a_i , b_i]$, $i = 1, 2, \ldots, k$ are non-overlapping intervals in $[a, b]$. Let $E_i = \{x \in [a_i , b_i]: f'(x) \text{ exists.}\}$. Then since f maps null sets to null sets and $m([a_i , b_i]-E_i) = 0$, $m(f([a_i , b_i]) = m(f(E_i))$. By Theorem 2, $m(f(E_i)) \leq \int_{E_i} |f'|$ and so for each i,

$$m(f([a_i, b_i])) \leq \int_{E_i} |f'| . \qquad \text{------------------------ (2)}$$

Since f is continuous, f is also continuous on $[a_i , b_i]$ and so by continuity,

$$|f(b_i) - f(a_i)| \leq m(f([a_i , b_i])) \text{ for each } i = 1, 2, \ldots, k.$$

Therefore, by (2) we have,

$$\sum_{i=1}^{k} |f(b_i) - f(a_i)| \leq \sum_{i=1}^{k} \int_{E_i} |f'| = \sum_{i=1}^{k} \int_{E_i} g$$
$$= \sum_{i=1}^{k} \int_{a_i}^{b_i} g, \text{ since } m([a_i , b_i]-E_i) = 0,$$
$$= \sum_{i=1}^{k} \int_{a_i}^{b_i} (g - g_N) + \sum_{i=1}^{k} \int_{a_i}^{b_i} g_N$$
$$\leq \int_a^b (g - g_N) + \sum_{i=1}^{k} \int_{a_i}^{b_i} K,$$

where $K > 0$ is an upper bound for g_N ,

$$< \frac{\varepsilon}{2} + K \sum_{i=1}^{k} |b_i - a_i| .$$

------------------------------- (3).

Take $\delta = \frac{\varepsilon}{2K}$, It follows from (3) that if $\sum_{i=1}^{k} |b_i - a_i| < \delta$, then

$$\sum_{i=1}^{k} |f(b_i) - f(a_i)| < \frac{\varepsilon}{2} + K\frac{\varepsilon}{2K} = \varepsilon.$$

This shows that f is absolutely continuous.

As a corollary we have the Banach Zarecki Theorem.

Theorem 8 (Banach Zarecki) . Suppose $f : [a, b] \to \mathbf{R}$ is continuous and is a function of bounded variation. Suppose f maps null sets to null sets. Then f is absolutely continuous.
Proof. Since f is of bounded variation, f is differentiable almost everywhere and f' is Lebesgue integrable. Therefore, by Theorem 7, f is absolutely continuous.

Remark. It is easy to see that if f is absolutely continuous on $[a, b]$, then f is continuous and of bounded variation on $[a, b]$. Any function of bounded variation on $[a, b]$ is the difference of two increasing functions (see for instance Theorem 13 of "*Monotone functions, function of bounded variation , fundamental theorem of Calculus*"). Since any increasing function on $[a, b]$ is differentiable almost everywhere on $[a, b]$ and its derived function is Lebesgue integrable on $[a, b]$, any function of bounded variation is therefore, differentiable almost everywhere on $[a, b]$ and its derivative is Lebesgue integrable on $[a, b]$. So if f is absolutely continuous on $[a, b]$, then f is differentiable almost everywhere on $[a, b]$ and f' is Lebesgue integrable on $[a, b]$. If f is absolutely continuous on $[a, b]$, then f maps null sets in $[a, b]$ to null sets (see for instance Proposition 9 of my article "*Change of variable or substitution in Riemann and Lebesgue Integration*"). Thus the converse of Theorem 7 and Theorem 8 are also true.

With a little thought we shall derive the following theorem.

Theorem 9. Suppose $f : [a, b] \to \mathbf{R}$ is absolutely continuous and $f'(x) = 0$ almost everywhere on $[a, b]$. Then f is a constant function.
Proof. It is enough to show that the range of f has measure zero. Let $E = \{x \in [a, b] : f'(x) = 0\}$. Then $m([a, b] - E) = 0$. By Theorem 3, $m(f(E)) = 0$. Since f is absolutely continuous, it maps null sets to null sets (see Proposition 9 of my article "*Change of variable or substitution in Riemann and Lebesgue Integration*"). It follows that $m(f([a, b] - E)) = 0$. Therefore, $m(f([a, b])) \leq m(f(E)) + m(f([a, b] - E)) = 0$. It follows that $m(f([a, b])) = 0$. Since f is continuous and $[a, b]$ is compact and connected, $f([a, b]))$ is compact and connected and so is either a non-trivial closed and bounded interval or a single point. Since a non-trivial closed and bounded interval has non-zero measure, $f([a, b]))$ must be a single point, consequently f is a constant function.

The next result is a consequence of a function having the property of being a continuous N function. In particular the result applies to an absolutely continuous function on $[a, b]$.

Theorem 10. Suppose $f : [a, b] \to \mathbf{R}$ is continuous and maps null sets to null sets, i.e., f is a continuous N function. Then f maps measurable sets to measurable sets.

Proof. Since the Lebesgue measure is a regular measure, for any measurable set E there is a subset, a F_σ set, K in $[a, b]$ such that $K \subseteq E$ and $m(E - K) = 0$. By a F_σ set K, we mean K is a countable union of closed sets in $[a, b]$. Thus

$$K = \bigcup_{n=1} K_n ,$$

where each K_n is a closed subset in $[a, b]$.
Each K_n is closed and bounded and so by the Heine Borel Theorem, is compact. Since f is continuous, each $f(K_n)$ is compact and so is closed and bounded by the Heine Borel Theorem. Since $f(K_n)$ is closed, it is measurable.
Therefore,

$$f(K) = \bigcup_{n=1} f(K_n),$$

being a countable union of measurable sets, is measurable.
Since f maps null sets to null sets, $m(f(E - K)) = 0$. It then follows that $f(E - K)$ is measurable. Hence,

$$f(E) = f(K) \cup f(E - K))$$

is a union of two measurable sets and so is measurable.

Corollary 11. Suppose $f : [a, b] \to \mathbf{R}$ is absolutely continuous. Then f maps measurable sets to measurable sets.
Proof. Since f is absolutely continuous on $[a, b]$, f maps null sets in $[a, b]$ to null sets (see for instance Proposition 9 of my article "*Change of variable or substitution in Riemann and Lebesgue Integration*"). Thus, f is a continuous N function and so by Theorem 10, f maps measurable sets to measurable sets.

For functions that are strictly increasing (or strictly decreasing) we have the following useful result for absolute continuity.

Theorem 12 (Zarecki). Suppose $f : [a, b] \to \mathbf{R}$ is strictly increasing and continuous.
(a) f is absolutely continuous if and only if $m(f(\{x \in [a, b]: f'(x) = \infty\})) = 0$.
(b) The inverse function, f^{-1}, is absolutely continuous, if and only if, $m(\{x \in [a, b]: f'(x) = 0\}) = 0$.
Proof.
(a) By Theorem 8, f is absolutely continuous, if and only if f, maps null sets to null sets. Since f is increasing, f is differentiable (finitely) almost everywhere on $[a, b]$. Hence $m(\{x \in [a, b]: f'(x) = \infty\}) = 0$. If f maps null sets to null sets, then $m(f(\{x \in [a, b]: f'(x) = \infty\})) = 0$.
Conversely, suppose $m(f(\{x \in [a, b]: f'(x) = \infty\})) = 0$. Let E be a set of measure 0 in $[a, b]$. Let $A = \{x \in [a, b]: f'(x) = \infty\}$, $B = \{x \in [a, b]: f'(x)$ does not exists and $f'(x) \neq \infty\}$. By the Theorem of De La Vallee Poussin, $m(f(B)) = 0$. Write $E = (E \cap A) \cup (E \cap B) \cup (E - (A \cup B))$. Then $m(E) = 0$ implies that $m(E - (A \cup B)) = 0$. By the Theorem of De La Vallee Poussin, we may assume that $f'(x)$ exists finitely on $E - (A \cup B)$. Therefore, by Theorem 2,

$$m(f(E - (A \cup B))) \leq \int_{E-(A \cup B)} |f'| = 0.$$

Hence $m(f(E - (A \cup B))) = 0$. Since $f(E \cap B) \subseteq f(B)$ and $m(f(B)) = 0$, $m(f(E \cap B)) = 0$. Since $E \cap A \subseteq A$ and $m(f(A)) = 0$, $m(f(E \cap A)) = 0$. Thus,

$$m(f(E)) \leq m(f(E - (A \cup B))) + m(f(E \cap A)) + m(f(E \cap B) = 0.$$

It follows that $m(f(E) = 0$. This means f maps null sets to null sets and it follows that f is absolutely continuous.

(b) Suppose f^{-1} is absolutely continuous. Let $C = \{x \in [a, b]: f'(x) = 0\}$. Then by Theorem 3, $m(f(C)) = 0$. Then since f^{-1} is absolutely continuous,

$$m(C) = m(f^{-1}(f(C))) = 0.$$

As in part (a), note that f^{-1} is absolutely continuous if and only if f^{-1} maps null sets to null sets.

Now assume that $m(C) = 0$.

Let E be a subset of $f([a, b]) = [c, d]$ of measure 0. Then $E = f(A)$, where $A = f^{-1}(E)$. We shall show that $m(A) = 0$.

By Theorem 15 of "*Functions of Bounded Variation and Johnson's Indicatrix*", $f' = 0$ almost everywhere on A.

Write $A = (A\cap C) \cup (A - C)$. Since $f' = 0$ almost everywhere on A, $m(A - C) = 0$. But $A\cap C \subseteq C$ and $m(C) = 0$ and so $m(A\cap C) = 0$. Hence $m(A) = m(f^{-1}(E)) = 0$. This completes the proof.

The proof of Theorem 12 (a) word for word with minor modification changing "increasing" to "of bounded variation" and "∞" to "$\pm\infty$" gives the following theorem.

Theorem 13. Suppose $f : [a, b] \to \mathbf{R}$ is continuous and of bounded variation. Then f is absolutely continuous, if and only if, $m(f(\{x \in [a, b]: f'(x) = \pm\infty\})) = 0$.

We shall now give a proof of the Theorem of De La Vallée Poussin.

Theorem 14 (De La Vallée Poussin). Suppose $f : [a, b] \to \mathbf{R}$ is a function of bounded variation. Then there is a subset N of $[a, b]$ such that

$$m(v_f(N)) = m(f(N)) = m(N) = 0,$$

where v_f is the total variation function of f, and for each x in $[a, b] - N$, $f'(x)$ and $v_f'(x)$ exist (finite or infinite) and that

$$v_f'(x) = |f'(x)|.$$

The following elementary proof is due to F. S. Cater.

The following technical lemma is the key to the proof.

Lemma 15. Suppose $f : [a, b] \to \mathbf{R}$ is a function of bounded variation. Let h and k be positive numbers such that $h < k$. Suppose $E = \{ x \in [a, b]$: there is a derived number of v_f at x greater than k and a derived number of f at x, whose absolute value is less than $h.\}$. Suppose $S = \{ x \in [a, b]$: there is a positive derived number and a negative derived number of f at $x\}$.
Then

$$m(v_f(E \cup S)) = m(f(E \cup S)) = m(E \cup S) = 0.$$

Proof. We assume that $E \cup S$ is non-denumerable, otherwise trivially all three sets have measure zero.

The first step is to choose some anchor partition for $[a, b]$ to approximate the total variation of f. Recall the definition of the total variation of a function of bounded variation,

$$v_f(b) = \sup\{ \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| : (P : a = x_0 < x_1 < \cdots < x_n = b) \text{ is a partition for } [a, b]\}.$$

Then given any $\varepsilon > 0$, there exists a partition $P : a = u_0 < u_1 < \cdots < u_n = b$ such that

$$v_f(b) - \varepsilon < \sum_{i=1}^{n} |f(u_i) - f(u_{i-1})| \le v_f(b), \text{ i.e.,}$$

$$v_f(b) < \sum_{i=1}^{n} |f(u_i) - f(u_{i-1})| + \varepsilon. \text{ ------------------------------ (A)}$$

Then for any partition, $Q: a = z_0 < z_1 < \cdots < z_t = b$, containing all the points of the partition P,

$$v_f(b) = \sum_{i=1}^{t} (v_f(z_i) - v_f(z_{i-1}) < \sum_{i=1}^{t} |f(z_i) - f(z_{i-1})| + \varepsilon \text{ ----------- (1)}$$

Let P denote also the set of points of the partition.$P: a = u_0 < u_1 < \cdots < u_n = b$.
We may assume also that f is continuous at every point of $E \cup S$. Then v_f is also continuous at every point of $E \cup S$. Since f is of bounded variation, the set of discontinuity of f is denumerable and so we may remove these points of discontinuity from $E \cup S$ without affecting the conclusion of the lemma.
Let U be an open set containing the image $v_f(E)$ such that $m(U) < m(v_f(E)) + \varepsilon$. Since U is open and v_f is continuous at e for each e in E, there exists an $\zeta > 0$ so that $(v_f(e) - \zeta, v_f(e) + \zeta) \subseteq U$ and corresponding to this $\zeta > 0$ there exists $\delta > 0$ such that

$$x \in (e - \delta, e + \delta) \Rightarrow v_f(x) \in (v_f(e) - \zeta, v_f(e) + \zeta).$$

Thus we can find arbitrary small non trivial intervals $[x, y]$ with $x \le e \le y$ such that $v_f(e) \in [v_f(x), v_f(y)] \subseteq (v_f(e) - \zeta, v_f(e) + \zeta)$. In particular, since v_f has a positive derived number $> k$ at e we can find arbitrary such small intervals $[x, y]$ such that

$$\frac{v_f(y) - v_f(x)}{y - x} > k.$$

(Note that since v_f has a positive derived number at e, the interval $[v_f(x), v_f(y)]$ is never degenerate.) Thus we can cover $v_f(E)$ by arbitrary such small closed intervals. Therefore, by the Vitali Covering Theorem, we can cover $v_f(E)$ almost everywhere by countable mutually disjoint closed interval,

$$\{[v_f(a_i), v_f(b_i)]\},$$

such that $[v_f(a_i), v_f(b_i)] \subseteq U$ and $v_f(b_i) - v_f(a_i) > k(b_i - a_i)$ for each i.
Therefore, the intervals, $\{[a_i, b_i]\}$, are also mutually disjoint and

$$m(v_f(E)) + \varepsilon > m(U) \ge \sum_i (v_f(b_i) - v_f(a_i)) > k\sum_i (b_i - a_i)$$

and so

$$\sum_i (b_i - a_i) < (m(v_f(E)) + \varepsilon)/k. \text{ ----------------------------- (2)}$$

Without loss of generality we may assume that the set of points of the partition . $P: a = u_0 < u_1 < \cdots < u_n = b$ does not contain any points of E. If P contains a point in E we may just remove this point from E. We may thus remove all the points in P that are in E from E without affecting the conclusion of the lemma as only a finite number of points is removed from E. We may take $\varepsilon = 1/N$ then by passing to the limit as N tends to infinity only at most a denumerable number of points are removed from E. Consequently as the measure of a set of denumerable number of points and its image under f or v_f is of measure zero, the conclusion of the lemma remains valid.

Since at each point e of $E - \{a_i, b_i : i = 1,2, \ldots\}$, there is a derived number of f whose absolute value is less than h, we may pick arbitrary small interval $[c, d]$ such that e is either one of the end points of the interval,

$$\frac{|f(d) - f(c)|}{d - c} < h,$$

$[c, d] \subseteq \bigcup_i [a_i, b_i]$ and that $P \cap [c, d] = \varnothing$. Note that $\{a_i, b_i : i = 1,2, \ldots\}$ is countable and so its image under v_f is also countable and so is of measure zero. Hence, again by the Vitali Covering Theorem, we can cover $v_f(E)$ almost everywhere with countable mutually disjoint closed intervals $\{[v_f(c_i), v_f(d_i)]\}$ such that $P \cap [c_i, d_i] = \varnothing$, $[c_i, d_i] \subseteq \bigcup_i [a_i, b_i]$,

$$| f(d_i) - f(c_i)| < h(d_i - c_i)$$

for each i and

$$m(v_f(E)) \le \sum_i (v_f(d_i) - v_f(c_i)). \quad \text{------------------- (3)}$$

But using (1) and the fact that for any $x < y$, $|f(y) - f(x)| \le v_f(y) - v_f(x)$, we can show that

$$\sum_i (v_f(d_i) - v_f(c_i)) \le \sum_i |f(d_i) - f(c_i)| + \varepsilon. \quad \text{---------------- (4)}$$

(Show this for finite number of the intervals $[c_i, d_i]$and pass to the limit.)
Since $\bigcup_i [c_i, d_i] \subseteq \bigcup_i [a_i, b_i]$,

$$\sum_i (d_i - c_i) \le \sum_i (b_i - a_i). \quad \text{------------------------- (5)}$$

Then from (3) and (4), we arrive at

$$m(v_f(E)) \le \sum_i (v_f(d_i) - v_f(c_i)) \le \sum_i |f(d_i) - f(c_i)| + \varepsilon$$

$$\le h \sum_i (d_i - c_i) + \varepsilon \le h \sum_i (b_i - a_i) + \varepsilon .$$

Thus, $$(m(v_f(E)) - \varepsilon)/h \le \sum_i (b_i - a_i)$$

and using (2) we get

$$(m(v_f(E)) - \varepsilon)/h < (m(v_f(E)) + \varepsilon)/k.$$

Since $\varepsilon = 1/N$ is arbitrary, by passing N to infinity we deduce that, $m(v_f(E))/h \le m(v_f(E))/k$. But since $h < k$, this is only possible if $m(v_f(E)) = 0$.

Now we proceed to show that $m(v_f(S)) = 0$. Using the fact that at each e in S, there is a positive derived number of f, and as in the above argument, we may pick arbitrary small interval $[r, s]$ such that e is either one of the end points of the interval,

$$\frac{f(s) - f(r)}{s - r} > 0$$

and $P \cap [r, s] = \varnothing$. Hence, we may cover $v_f(e)$ by arbitrary small intervals $[v_f(r), v_f(s)]$. Therefore, by the Vitali Covering Theorem, we may cover $v_f(S)$ almost everywhere by countable mutually disjoint closed intervals $\{[v_f(r_i), v_f(s_i)]\}$ such that $P \cap [r_i, s_i] = \varnothing$, $f(s_i) - f(r_i) > 0$ for each i and

$$m(v_f(S)) \le \sum_i (v_f(s_i) - v_f(r_i)) \le \sum_i f(s_i) - f(r_i) + \varepsilon, \quad \text{----------------- (6)}$$

where the last inequality is deduced using (1).
Similarly, as before using the negative derived number of f at each of the point e of S, we may cover $v_f(E)$ almost every where with countable mutually disjoint closed intervals $\{[v_f(p_i), v_f(q_i)]\}$ such that $P \cap [p_i, q_i] = \varnothing$, $[p_i, q_i] \subseteq \bigcup_i [r_i, s_i]$, $f(p_i) > f(q_i)$ for each i and

$$m(v_f(S)) \le \sum_i (v_f(q_i) - v_f(p_i)) \le \sum_i f(p_i) - f(q_i) + \varepsilon . \quad \text{---------------- (7)}$$

Since , $\bigcup_i [p_i, q_i] \subseteq \bigcup_i [r_i, s_i]$,

$$\sum_i f(p_i) - f(q_i) \le \sum_i (N_f(s_i) - N_f(r_i)),$$

$$\sum_i f(s_i) - f(r_i) \le \sum_i (P_f(s_i) - P_f(r_i)),$$

where N_f and P_f are the negative and positive variations of f. Therefore, because $v_f = N_f + P_f$,

$$\sum_i f(p_i) - f(q_i) + \sum_i f(s_i) - f(r_i) \leq \sum_i (v_f(s_i) - v_f(r_i)).$$

This inequality together with (6) and (7) yields,

$$\sum_i (v_f(s_i) - v_f(r_i)) - \varepsilon + \sum_i (v_f(q_i) - v_f(p_i)) - \varepsilon \leq \sum_i (v_f(s_i) - v_f(r_i))$$

and so

$$\sum_i (v_f(q_i) - v_f(p_i)) \leq 2\varepsilon.$$

Hence, $m(v_f(S)) \leq 2\varepsilon$. Since $\varepsilon = 1/N$ by passing to the limit as N tends to infinity, $m(v_f(S)) = 0$.

Therefore,

$$m(v_f(S \cup E)) \leq m(v_f(E)) + m(v_f(S)) = 0$$

and so $m(v_f(S \cup E)) = 0$.

Now for any $\varepsilon > 0$, take an open set U such that $v_f(E \cup S) \subseteq U$ and $m(U) \leq \varepsilon$. Since U is open, U is a countable union of mutually disjoint non-trivial intervals I_i. Then the collection $\{v_f^{-1}(I_i)\}$ covers $E \cup S$. Therefore,

$$m(f(S \cup E)) \leq m(f(v_f^{-1}(\bigcup I_i))) = \sum_i m(f(v_f^{-1}(I_i))) \leq \sum_i m(I_i) = m(U) \leq \varepsilon.$$

We have used the fact that $m(f(v_f^{-1}(I_i))) \leq m(I_i)$ for each i. We deduce this as follows. For any point x, y in $v_f^{-1}(I_i)$, $|f(x) - f(y)| \leq |v_f(x) - v_f(y)| \leq diameter(I_i)$. Therefore, the diameter of $f(v_f^{-1}(I_i)) \leq$ diameter of I_i = length of I_i = m(I_i). That means $m(f(v_f^{-1}(I_i))) \leq m(I_i)$. Since ε was arbitrary, $m(f(E \cup S)) = 0$.

It remains now to show that $m(E \cup S) = 0$.

Since f is of bounded variation, f is differentiable almost everywhere. So we may assume that f has finite derivative at every point of $E \cup S$. f is obviously not differentiable at every point of S since each point of S has a positive and negative derived numbers. Note that, since $|f'| = v_f'$ almost every where, we may look only at points x in E, where the derived number for f at x has the same absolute value as the only derived number of v_f at x. So since points in E do not have this property, E must have measure 0. It follows that $m(E \cup S) = 0$. We may alternatively prove directly that $m(E \cup S) = 0$ by using a Vitali covering argument.

16. Proof of de La Vallée Poussin Theorem (Theorem 14)

Let $E_{h,k} = \{ x \in [a, b]$: there is a derived number of v_f at x greater than k and a derived number of f at x, whose absolute value is less than h, $h < k.\}$.

Let $E = \bigcup\{E_{h,k} : h, k \text{ rational and } h < k\}$. Let $N = E \cup S$. We have already shown in the proof of Lemma 15 that $m(S) = m(f(S)) = m(v_f(S)) = 0$.

By Lemma 15, $m(E_{h,k}) = 0$ for each pair (h, k), $h < k$. Thus E is a countable union of sets of measure zero and so $m(N) = m(E \cup S) = 0$. Note that

$$m(f(E)) \leq \sum_{0 < h < k,\ h \text{ and } k \text{ rational}} m(f(E_{h,k})) = 0$$

since the set $f(E) = \bigcup\{f(E_{h,k}) : h, k \text{ rational and } h < k.\}$ is a countable union of sets, $f(E_{h,k})$, each of measure zero by Lemma 15. Thus $m(f(E)) = 0$. It follows that $m(f(N)) = 0$. Similarly, we show that $m(v_f(N)) = 0$.

We now prove the property of N as stated in the theorem. Take any x in $[a, b] - N$. Then x is not in S and not in any $E_{h,k}$. Hence f does not have a positive and a negative derived numbers at x. Moreover for any finite derived number DV of v_f at x,

$$DV \le |Df| \text{ for any derived number } Df \text{ of } f \text{ at } x.$$

Therefore, for any derived number DV of v_f at x, we have,

$$DV \le \inf\{ |Df| : Df \text{ is a derived number of } f \text{ at } x\}.$$

Note that if DV is a derived number of v_f at x, then there is a sequence (h_n) such that $h_n \neq 0$, $h_n \to 0$ and

$$DV = \lim_{n \to \infty} \frac{v_f(x+h_n) - v_f(x)}{h_n} .$$

Therefore, the sequence $\left(\frac{v_f(x+h_n) - v_f(x)}{h_n}\right)$ is bounded. Since we have for each n, $\left|\frac{f(x+h_n) - f(x)}{h_n}\right| \le \left|\frac{v_f(x+h_n) - v_f(x)}{h_n}\right|$, the sequence $\left(\frac{f(x+h_n) - f(x)}{h_n}\right)$ is also bounded. Hence, by the Bolzano Weierstrass Theorem, $\left(\frac{f(x+h_n) - f(x)}{h_n}\right)$ has a convergent subsequence, $\left(\frac{f(x+h_{n_k}) - f(x)}{h_{n_k}}\right)$ and

$$Df_1 = \lim_{k \to \infty} \frac{f(x+h_{n_k}) - f(x)}{h_{n_k}}$$

is a derived number of f at x. Moreover the subsequence $\left(\frac{v_f(x+h_{n_k}) - v_f(x)}{h_{n_k}}\right)$ converges to the same value DV and so we have

$$|Df_1| \le DV .$$

But $DV \le |Df_1|$ and so $DV = |Df_1|$. It follows that any derived number of v_f at x is equal to $\inf\{|Df| : Df \text{ is a derived number of } f \text{ at } x.\}$. Consequently there can be only one derived number of v_f at x and so v_f is differentiable at x. It follows that for any derived number Df of f at x,

$$|Df| \le v_f'(x)$$

and $v_f'(x) \le |Df|$ because $v_f'(x)$ is the infimum of all absolute values of the derived numbers of f at x. Thus, $|Df| = v_f'(x)$ for any derived number Df of f at x. Therefore, any derived number of f has one unique absolute value. Since f has no derived number of opposite sign at x, it can have only one unique derived number at x. That is to say, f is differentiable at x.

Suppose now that v_f has an infinite derived number at x, then since x is in $[a, b] - N$, any derived number Df of f at x must have $|Df| = \infty$. Consequently there is only one derived number of v_f at x, namely $+\infty$. Since f does not have derived number of opposite signs at x, it can have only one derived number at x either $+\infty$ or $-\infty$.

We have thus proved that f is differentiable (finite or infinite) at every point of $[a, b] - N$.

Revised Jan 2017

Ng Tze Beng
Email: tbengng@gmail.com