Measure zero and the characterization of Riemann integrable functions

Jeffrey Schenker

July 25, 2007

Let us define the length of an interval $I$ (open or closed) with endpoints $a < b$ to be
\[ \ell(I) = b - a. \] (1)

The extension of the notion of length to sets other than intervals, and to more general measures of length, is known as measure theory and is a basic part of a graduate course on real analysis. For the purpose of this discussion we need only the following notion from measure theory:

**Definition.** A set $S \subset \mathbb{R}$ is said to have measure zero if for every $\epsilon > 0$ there is a countable or finite collection of open intervals $I_j$, $j = 1, \ldots$, such that
\[ S \subset \bigcup_j I_j \quad \text{and} \quad \sum_j \ell(I_j) < \epsilon. \]

**Remark:** We could require the intervals $I_j$ to be disjoint, as was done in class, but nothing is gained by this.

Measure zero sets provide a characterization of Riemann integrable functions.

**Theorem 1.** A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if
\[ \{ x : f \text{ is not continuous at } x \} \text{ has measure zero.} \]

A proof of Theorem 1 can be found below.

Measure zero sets are “small,” at least insofar as integration is concerned. Because of this one defines

**Definition.** A proposition $A(x)$ which depends on a real number $x$ is said to be true almost everywhere if \( \{ x : A(x) \text{ is false} \} \) has measure zero.

Thus Theorem 1 states that a bounded function $f$ is Riemann integrable if and only if it is continuous almost everywhere.

The terminology “almost everywhere” is partially justified by the following

**Theorem 2.** If $f$ and $g$ are Riemann integrable on $[a, b]$ and $f(x) = g(x)$ almost everywhere, that is \( \{ x : f(x) \neq g(x) \} \) has measure zero, then
\[ \int_t^s f(x)dx = \int_t^s g(x)dx \]
for any $t, s \in [a, b]$. 

It is essential that we assume both $f$ and $g$ are Riemann integrable. Indeed, if $f$ is Riemann integrable and $f(x) = g(x)$ almost everywhere it may nonetheless happen that $g$ is not Riemann integrable. For example if $f(x) = 0$ and

$$g(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

then $f$ is Riemann integrable and $g$ is not, but $f(x) = g(x)$ almost everywhere since

**Lemma.** Any countable set has measure zero.

**Proof.** Exercise. □

Since Theorem 2 is really beyond the scope of this class we will not prove it here.

**Proof of Theorem 1**

($\Rightarrow$) First suppose $f$ is Riemann integrable and consider, for each $t > 0$, the set

$$S_t = \{ x \in [a,b] \mid \forall \delta > 0 \exists y \in [a,b] \text{ s.t. } |x - y| < \delta \text{ and } |f(x) - f(y)| > t \}.$$  

Then $S = \{ x : f \text{ is not continuous at } x \}$ satisfies

$$S = \bigcup_{t > 0} S_t.$$  

Because $S_t \subset S_s$ for $s < t$, we can replace the uncountable union $\bigcup_{t > 0} S_t$ by the countable union

$$S = \bigcup_{n=1}^{\infty} S_{1/n}.$$  

Thus, it suffices to show that each $S_t$ has measure zero, because of the following

**Lemma.** Let $S_j$, $j = 1, \ldots$, be a finite or countable collection of sets such that each $S_j$ has measure zero. Then $\bigcup_{j} S_j$ has measure zero.

**Proof of Lemma.** Let $\epsilon > 0$. Then for each $j = 1, \ldots$ there is a finite or countable collection of open intervals $I_k$, $k = 1, \ldots$, with $S_j \subset \bigcup_k I_k^j$ and $\sum_k \ell(I_k^j) < \epsilon/2^j$. Thus $\bigcup_j S_j \subset \bigcup_j \bigcup_k I_k^j$ and

$$\sum_j \sum_k \ell(I_k^j) < \epsilon \sum_{j=1}^{\infty} 2^{-j} = \epsilon.$$  

(Recall that a countable union of countable sets is countable. Explicitly, we enumerate $I_k^j$ as follows

$I_1 = I_1^1$, $I_2 = I_1^2$, $I_3 = I_2^1$, $I_4 = I_1^3$, $I_5 = I_2^2$, $I_6 = I_3^1$, \ldots.

That is, first we list indices with $j + k = 2$ then with $j + k = 3$, then with $j + k = 4$, etc.) □

2
Returning to the proof that $S_t$ has measure zero, let $\epsilon > 0$. Since $f$ is Riemann integrable, there is a partition $P$ of $[a, b]$ with $\text{Osc}(f, P) < \epsilon$. Let $a = x_0 < x_1 < x_2 < \cdots < x_n = b$ be the points of the partition $P$ and define an index set

$$J = \{j : S_t \cup [x_{j-1}, x_j] \neq \emptyset\}.$$

Thus $S_t \subset \bigcup_{j \in J} [x_{j-1}, x_j]$, however the intervals $[x_{j-1}, x_j]$ are closed. To obtain a covering by open intervals, let us enlarge the closed intervals $[x_{j-1}, x_j]$ a bit:

$$I_j = \left( x_{j-1} - \frac{x_j - x_{j-1}}{2}, x_j + \frac{x_j - x_{j-1}}{2} \right).$$

Thus $\ell(I_j) = 2(x_j - x_{j-1})$ and $S_t \subset \bigcup_{j \in J} I_j$.

It remains to estimate the sum

$$\sum_{j \in J} \ell(I_j) = 2 \sum_{j \in J} x_j - x_{j-1}. $$

On each interval $[x_{j-1}, x_j]$ with $j \in J$ we have

$$\sup_{x,y \in [x_{j-1}, x_j]} |f(x) - f(y)| > t$$

since there is a point $x \in S_t \cap [x_{j-1}, x_j]$. Thus

$$t \sum_{j \in J} (x_j - x_{j-1}) < \sum_{j \in J} \sup_{x,y \in [x_{j-1}, x_j]} |f(x) - f(y)| \cdot (x_j - x_{j-1}) \leq \text{Osc}(f, P),$$

since the oscillation involves the sum over all $j = 1, \ldots, n$ in place of just $j \in J$. Therefore

$$\sum_{j \in J} \ell(I_j) = 2 \sum_{j \in J} (x_j - x_{j-1}) < \frac{2}{t} \text{Osc}(f, P) < \frac{2}{t} \epsilon. $$

Since $\epsilon$ is arbitrary, we see that $S_t$ has measure zero. Thus $S = \bigcup_n S_t/n$ has measure zero by the Lemma.

$(\Leftarrow)$ Now suppose that $S = \{x : f$ is discontinuous at $x\}$ has measure zero. Let $\epsilon > 0$. Then there is a finite or countable collection $I_j, j = 1, \ldots$ of open intervals such that $S \subset \bigcup_j I_j$ and $\sum_j \ell(I_j) < \epsilon$. Consider the set

$$K = [a, b] \setminus \bigcup_j I_j = \{x : x \in [a, b] \text{ and } \forall_j x \notin I_j\}.$$

So $K$ is a closed subset of $[a, b]$ and thus is compact. Furthermore $f$ is continuous at each point $x \in K$ (since $S \cup K = \emptyset$). Thus, for each $x \in K$ there is $\delta_x > 0$ such that $|y - x| \leq \delta_x$ implies $|f(y) - f(x)| \leq \epsilon$. Now the open intervals $J_x = \{y : |y - x| < \delta_x\}$ cover $K$, that is $K \subset \bigcup_{x \in K} J_x$, since $x \in J_x$ for each $x$. By the Heine-Borel theorem there is a finite subcover, namely there are points $x_1, \ldots, x_m \in K$ such that

$$K \subset \bigcup_{j=1}^m \bar{J}_{x_j},$$

Let us form a partition out of the endpoints of the intervals of this finite subcover:

$$P = \{a, b\} \cup \{x_j - \delta_{x_j}, x_j + \delta_{x_j} : j = 1, \ldots, m\}.$$
The intervals \( \tilde{I}_{x_j} \) may overlap one another: we may have for example \( x_j - \delta x_j < x_k - \delta x_k < x_j + \delta_j \) for some pair \( j, k \). Thus we may have to rearrange labels to write the points in order. However \( \mathcal{P} \) is a finite set, so we may write \( \mathcal{P} = \{y_0 = a < y_1, \ldots < y_n = b\} \). Each point \( y_k, k = 1, \ldots, n - 1 \), is equal to \( x_j + \delta x_j \) for some \( j \). I claim that we have the following dichotomy for each interval \([y_{k-1}, y_k] \):

1. There is a \( j \in \{1, \ldots, m\} \) such that

\[
[y_{k-1}, y_k] \subset \overline{I}_{x_j},
\]

where \( \overline{I}_{x_j} = \{y : |y - x_j| \leq \delta_j\} \) is the closure of \( I_{x_j} \), or

2. For all \( j \in \{1, \ldots, m\} \), \([y_{k-1}, y_k] \) is disjoint from the open interval \( \overline{I}_{x_j} \).

To see this note that no endpoint of any interval \( \overline{I}_{x_j} \) can fall in the interior \((y_{k-1}, y_k)\), since we have listed the points \( y^* \) in increasing order. Thus, if \( \overline{I}_{x_j} \cap [y_{k-1}, y_k] \neq \emptyset \) then \((y_{k-1}, y_k) \subset \overline{I}_{x_j} \).

We break the oscillation of \( f \) in the partition \( \mathcal{P} \) into two pieces,

\[
\text{Osc}(f, \mathcal{P}) = \sum_{k \in \text{case 1}} \sup_{x, y \in [y_{k-1}, y_k]} |f(x) - f(y)| (y_k - y_{k-1}) + \sum_{k \in \text{case 2}} \sup_{x, y \in [y_{k-1}, y_k]} |f(x) - f(y)| (y_k - y_{k-1}).
\]

For \( k \) in case (1), we have for \( x, y \in [y_{k-1}, y_k] \),

\[
|f(x) - f(y)| \leq |f(x) - f(x_j)| + |f(x_j) - f(y)| \leq 2\epsilon,
\]

by the choice of \( \delta x_j \). For \( k \) in case (2), we have little control over \( |f(x) - f(y)| \), however \( |f(x) - f(y)| \leq 2M \) with \( M = \sup_x |f(x)| \). Thus

\[
\text{Osc}(f, \mathcal{P}) \leq 2\epsilon \sum_{k \in \text{case 1}} (y_k - y_{k-1}) + 2M \sum_{k \in \text{case 2}} (y_k - y_{k-1}).
\]

Now each interval \([y_{k-1}, y_k] \) for \( k \) in case 2 is disjoint for the compact set \( K \) and thus contained in the union of open intervals \( \cup_j I_j \) covering the measure zero set \( S \). Thus the total length of all intervals contributing to case 2 is bounded by \( \sum_j \ell(I_j) < \epsilon \). (Why? This is true and, perhaps, “intuitively obvious.” The proof is left as an exercise.) Thus

\[
\text{Osc}(f, \mathcal{P}) \leq (b - a + 2M)\epsilon,
\]

since the total length of all intervals contributing to case 1 is certainly less than the total of all intervals. As \( \epsilon \) is arbitrary, we see that \( f \) is Riemann integrable. \( \square \)