Lebesgue's characterization of Riemann integrable functions

M. Müger

June 20, 2006

The aim of these notes is to give an elementary proof (i.e. without Lebesgue theory) of the following theorem:

1 THEOREM *A function $f : [a, b] \to \mathbb{R}$ is Riemann integrable iff it is bounded and the set*

$$S(f) = \{x \in [a, b] \mid f \text{ is not continuous at } x\}$$

has measure zero.

The proof will follow the strategy outlined in [3, Exercise 6.1.3 (b)-(d)]. For an alternative elementary (but more involved) proof cf. [1]. A short proof using the basics of Lebesgue theory is given in [2].

In order to fix the terminology, we briefly recall the relevant definitions:

2 DEFINITION *A partition P of $[a, b]$ is a finite sequence $a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$. We write $\Delta_i = [x_{i-1}, x_i]$ and $\Delta x_i = x_{i-1} - x_i$. The mesh of a partition P is defined by*

$$\lambda(P) = \max_{i=1,\ldots,n} \Delta x_i.$$

3 DEFINITION *A function $f : [a, b] \to \mathbb{R}$ is Riemann integrable (on $[a, b]$) if there exists $A \in \mathbb{R}$ (easily seen to be unique) such that for every $\varepsilon > 0$ there is $\lambda > 0$ such that the Riemann sum*

$$\sigma(f; P, \xi) = \sum f(\xi_i)\Delta x_i$$

satisfies $|\sigma(f; P, \xi) - A| < \varepsilon$ whenever P is a partition with $\lambda(P) < \delta$ and $\xi_i \in \Delta_i$ for all $i = 1, \ldots, n$. In this case we write $\int_a^b f(x)dx = A$.

4 DEFINITION *A subset $X \subset \mathbb{R}$ has measure zero if for every $\varepsilon > 0$ there is a sequence $(a_1, b_1), (a_2, b_2), \ldots$ of (open, bounded) intervals such that*

$$X \subset \bigcup_{i=1}^{\infty}(a_i, b_i) \quad \text{and} \quad \sum_{i=1}^{\infty} b_i - a_i < \varepsilon.$$

5 DEFINITION *Let $E \subset \mathbb{R}$ and $f : E \to \mathbb{R}$. Then the oscillation of f on E is defined as*

$$\omega(f, E) = \sup_{x,x' \in E} |f(x) - f(x')|.$$

6 DEFINITION *For $f : [a, b] \to \mathbb{R}$ and $x \in [a, b]$ we define*

$$\omega(f, x) = \inf_{\varepsilon>0} \omega(f, [a, b] \cap (x - \varepsilon, x + \varepsilon)),$$

$$S_\varepsilon(f) = \{x \in [a, b] \mid \omega(f, x) > \varepsilon\}.$$

7 REMARK Note that $S_0(f) = S(f)$, as defined in Theorem 1. □

We will use the following criterion:

8 Proposition *[3, p. 339] A function $f : [a, b] \to \mathbb{R}$ is Riemann integrable iff*

$$\lim_{\lambda(P)\to 0} \sum_{i=1}^{n} \omega(f, \Delta_i)\Delta x_i \ = 0. \tag{1}$$

Since a Riemann integrable function is bounded, cf. [3, p. 333], in the sequel we will always suppose f to be bounded. We define a function $\theta : \mathbb{R} \to \mathbb{R}$ by $\theta(x) = 0$ for $x \le 0$ and $\theta(x) = 1$ for $x > 0$

9 Proposition *[3, Exercise 6.1.3 (b)] A bounded function $f : [a, b] \to \mathbb{R}$ satisfies (1) iff for any $\varepsilon, \delta > 0$ there is a partition P of $[a, b]$ such that*

$$\sum_{i=1}^{n} \theta(\omega(f, \Delta_i) - \varepsilon)\Delta x_i < \delta. \tag{2}$$

In words: The sum of the lengths of the intervals of the partition on which the oscillation of the function is larger than ε is smaller than δ.

Proof. Assume that f satisfies (2) for some $\varepsilon, \delta > 0$. Splitting $\sum_{i=1}^{n} \omega(f, \Delta_i)\Delta x_i$ into two parts, we have

$$\sum_{i=1}^{n} \omega(f, \Delta_i)\Delta x_i = \sum_{i=1}^{n} \theta(\varepsilon - \omega(f, \Delta_i))\omega(f, \Delta_i)\Delta x_i + \sum_{i=1}^{n} \theta(\omega(f, \Delta_i) - \varepsilon)\omega(f, \Delta_i)\Delta x_i. \tag{3}$$

The first term is bounded from above by $\sum_{i=1}^{n} \varepsilon\, \Delta x_i = \varepsilon \cdot (b - a)$. On the other hand, the obvious inequalities $\omega(f, \Delta_i) \le \omega(f, [a, b]) < \infty$ (of which the second is the boundedness of f) implies that the second term in (3) is majorized by

$$\sum_{i=1}^{n} \theta(\omega(f, \Delta_i) - \varepsilon)\omega(f, [a, b])\Delta x_i,$$

which in view of (2) is smaller than $\delta \cdot \omega(f, [a, b])$. We have thus proven

$$\sum_{i=1}^{n} \omega(f, \Delta_i)\Delta x_i \ < \ \varepsilon \cdot (b - a) + \delta \cdot \omega(f, [a, b]). \tag{4}$$

Assume now that for every $\varepsilon, \delta > 0$ there is a partition P such that (2) holds. Then (4) implies that we can make $\sum_i \omega(f, \Delta_i)\Delta x_i$ arbitrarily small, thus (1) holds.

As to the opposite implication, let P be a partition of $[a, b]$ and let $\varepsilon > 0$. Considering only the contribution to $\sum_{i=1}^{n} \omega(f, \Delta_i)\Delta x_i$ of the intervals Δ_i on which the oscillation of f is larger than ε, we obtain

$$\varepsilon \sum_{i=1}^{n} \theta(\omega(f, \Delta_i) - \varepsilon)\Delta x_1 \ \le \ \sum_{i=1}^{n} \omega(f, \Delta_i)\Delta x_i.$$

Thus, if (1) holds, we can make $\varepsilon \sum_{i=1}^{n} \theta(\omega(f, \Delta_i) - \varepsilon)\Delta x_i$ smaller than any given positive number, thus also smaller than $\delta\varepsilon$ (where $\varepsilon, \delta > 0$). Thus, for any $\varepsilon, \delta > 0$ there is a partition P such that (2) holds. ■

10 Proposition *[3, Exercise 6.1.3 (c)] Let $f : [a, b] \to \mathbb{R}$ be bounded. Then the following are equivalent:*

(i) For every $\varepsilon, \delta > 0$ there is a partition P such that (2) holds.

(ii) For every $\varepsilon, \delta > 0$ there is a finite sequence of intervals $(a_1, b_1), \ldots, (a_m, a_m)$ such that

$$S_\varepsilon(f) \subset \bigcup_{i=1}^{m} (a_i, b_i) \quad \textit{and} \quad \sum_{i=1}^{m} b_i - a_i < \delta. \tag{5}$$

Proof. (i)⇒(ii). Let $\varepsilon, \delta > 0$, and pick a partition P satisfying (2). Let $I = \{i \in \{1, \ldots, n\} \mid \omega(f, \Delta_i) > \varepsilon\}$. Then $\{(x_{i-1}, x_i) \mid i \in I\}$ is a finite family of open intervals, and by (2) its total length is smaller than δ. If $i \notin I$ then $\omega(f, \Delta_i) \le \varepsilon$, thus $\omega(f, x) \le \varepsilon$ for every $x \in [x_{i-1}, x_i]$. Thus $S_\varepsilon(f) \subset \bigcup_{i\in I}(x_{i-1}, x_i)$ and (ii) holds.

(ii)⇒(i) Let $\varepsilon, \delta > 0$ and let $(a_1, b_1), \ldots, (a_m, b_m)$ be such that (5) holds. We may and will assume that these intervals are mutually disjoint. Since replacing the adjacent intervals (a, b) and (b, c) by (a, c) leaves the total length unchanged, we may also assume that no two intervals have a common boundary point. We write

$J = [a,b] - \bigcup_i (a_i, b_i)$. If $x \in J$ then $\omega(f,x) \le \varepsilon$. By definition of $\omega(f,x)$, this means that there is an open interval $U_x = (x-\alpha, x+\alpha)$ such that $\omega(f, U_x \cap [a,b]) \le 2\varepsilon$. (Of course, α depends on x.) Since J is compact, it can be covered by finitely many such sets: $J \subset U_{x_1} \cup \cdots \cup U_{x_n}$. We conclude that there exists a partition $P = \{a = x_0 < x_1 < \cdots < x_n = b\}$ such that each of the above (a_i, b_i) appears as (x_{j-1}, x_j) for some j, and such that $\omega(f, [x_{j-1}, x_j]) \le \varepsilon$ for the remaining intervals of the partition. By construction, $\omega(f, \Delta_i) \le 2\varepsilon$ for those intervals Δ_i that are not of the form $[a_i, b_i]$. Thus these Δ_i do not contribute to $\sum_{i=1}^n \theta(\omega(f,\Delta_i) - 2\varepsilon)\Delta x_i$, whereas (since we assume (ii)) the total length of the intervals $[a_i, b_i]$ is bounded by δ. Since $\varepsilon, \delta > 0$ were arbitrary, we have proven (i). ■

11 Proposition *[3, Exercise 6.1.3 (d)] Let $f : [a,b] \to \mathbb{R}$ be bounded. Then the following are equivalent:*

(i) For every $\varepsilon, \delta > 0$ there is a finite sequence of intervals $(a_1, b_1), \ldots, (a_m, b_m)$ such that

$$S_\varepsilon(f) \subset \bigcup_{i=1}^m (a_i, b_i) \quad \textit{and} \quad \sum_{i=1}^m b_i - a_i < \delta.$$

(This is the criterion of du Bois-Reymond.)

(ii) For every $\delta > 0$ there is a sequence of intervals $(a_1, b_1), (a_2, a_2), \ldots$ such that

$$S(f) \subset \bigcup_{i=1}^\infty (a_i, b_i) \quad \textit{and} \quad \sum_{i=1}^\infty b_i - a_i < \delta. \tag{6}$$

Equivalently, $S(f)$ has measure zero, which is Lebesgue's criterion.

Proof. (i)⇒(ii). Let $\delta > 0$. For every $n \in \mathbb{N}$ there are intervals $(a_{n,1}, b_{n,1}), \ldots, (a_{n,m(n)}, b_{n,m(n)})$ such that we have, for every $n \in \mathbb{N}$,

$$S_{2^{-n}}(f) \subset \bigcup_{i=1}^{m(n)} (a_{n,i}, b_{n,i}) \quad \text{and} \quad \sum_{i=1}^{m(n)} b_{n,i} - a_{n,i} < 2^{-n}\delta.$$

Now the family $\{(a_{n,i}, b_{n,i}),\ n \in \mathbb{N},\ i = 1, \ldots, m(n)\}$, contains $\bigcup_{n\in\mathbb{N}} S_{2^{-n}}(f) = S(f)$ and has total length bounded by $\sum_{n=1}^\infty 2^{-n}\delta = \delta$. Thus $S(f)$ has measure zero.

(ii)⇒(i). Let $\delta > 0$ and choose a sequence of intervals $(a_1, b_1), (a_2, b_2), \ldots$ satisfying (6). For any $\varepsilon > 0$, we have $S_\varepsilon(f) \subset S_0(f) = S(f)$, so we are done one once we can show that $S_\varepsilon(f)$ is closed. For then it is compact and thus covered by a finite subfamily of $\{(a_i, b_i),\ i \in \mathbb{N}\}$, implying (i).

Let $x \in [a,b]$ be a limit point of $S_\varepsilon(f)$. Thus any neighborhood of x contains a point x' with $\omega(f, x') > \varepsilon$. By definition, the latter means that every neighborhood of x' contains an x'' such that $|f(x') - f(x'')| > \varepsilon$. This implies that every neighborhood of x contains points x', x'' such that $|f(x') - f(x'')| > \varepsilon$. Therefore $\omega(f,x) > \varepsilon$, thus $x \in S_\varepsilon(f)$ and $S_\varepsilon(f)$ is closed. ■

Theorem 1 now follows immediately by combining Propositions 8,9,10,11.

References

[1] S. K. Berberian: *A first course in real analysis*. Springer, 1994.

[2] W. Rudin: *Principles of mathematical analysis*. 3rd edition. McGraw-Hill, 1976.

[3] V. A. Zorich: *Mathematical Analysis I*. Springer, 2004.