Characterization of Riemann integrability – an elementary proof

L. Vesely, 2017

The aim of this note is to provide a very elementary proof, accessible to a first-year student, of the following important well-known theorem. (See below, for the definition of "almost everywhere".)

Theorem 0.1. *A function $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable if and only if f is bounded and almost everywhere continuous.*

For a standard proof, using measure theory, see e.g. the book [W. Rudin, *Principles of Mathematical Analysis*], §11.7. The proof that follows uses only the definition of a null set, and no other measure theoretical result. We shall prove separately the two implications. The proof of necessity is very elementary, while the clever proof of sufficiency was taken from the paper

[R.A. Gordon, *The use of tagged partitions in elementary real analysis*, Amer. Math. Monthly 105 (1998), no. 2, 107–117].

1. Preliminaries

Given a function $f\colon [a,b] \to \mathbb{R}$, the *oscillation* of f on a set $A \subset [a,b]$ is the quantity

$$\omega(f,A) := \sup f(A) - \inf f(A).$$

For $x \in (a,b)$, the *oscillation of f at the point x* is defined as

$$\omega(f,x) := \inf\{\omega(f,I) : \ I \subset (a,b) \text{ is an open interval, } x \in I\}.$$

Exercise 1.1. Le f be as above, and $x \in (a,b)$. Show that f is continuous at x if and only if $\omega(f,x) = 0$.

A *partition* of $[a,b]$ is a finite family of intervals $\{I_j : j = 1,\dots,n\}$, where $n \in \mathbb{N}$, $I_j = [x_{j-1}, x_j]$ $(1 \le j \le n)$, and $a = x_0 < x_1 < \cdots < x_n = b$. The length of an interval $I \subset \mathbb{R}$ is denoted by $|I|$, and the notation $f \in \mathcal{R}[a,b]$ means that f is Riemann integrable on $[a,b]$.

Exercise 1.2. Notice that the standard test of Riemann integrability can be formulated as follows: $f \in \mathcal{R}[a,b]$ *if and only if for every $\varepsilon > 0$ there exists a partition $\{I_j : j = 1,\dots,n\}$ of $[a,b]$ such that*

$$\sum_{j=1}^{n} \omega(f,I_j)|I_j| \le \varepsilon.$$

Definition 1.3. A set $A \subset \mathbb{R}$ is said to be a *null set* or to have *measure zero* if for every $\varepsilon > 0$ there exists a sequence $\{J_k\}_1^\infty$ of open (not necessarily nonempty) intervals such that $A \subset \bigcup_{k=1}^\infty J_k$ and $\sum_{k=1}^\infty |J_k| < \varepsilon$.
A property holds *almost everywhere* if it holds at each point outside a null set.

It is obvious that a subset of a null set is again a null set.

Exercise 1.4. Show that a finite or countable union of null sets is a null set, and that every at most countable set is a null set.

2. Proof of sufficiency

Theorem 2.1 (Sufficiency)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be bounded and almost everywhere continuous. Then $f \in \mathcal{R}[a,b]$.*

Our proof follows the elegant idea from the paper [R.A. Gordon, *The use of tagged partitions in elementary real analysis*, Amer. Math. Monthly 105 (1998), no. 2, 107–117]. It is based on the following notion of a "δ-fine tagged partition".

Definition 2.2. A *tagged partition* of $[a,b]$ is a set of the form

$$\{(c_j, I_j) : j = 1, \dots, n\}$$

where $\{I_j : j = 1, \dots, n\}$ is a partition of $[a,b]$, and $c_j \in I_j$ $(1 \le j \le n)$. So, a tagged partition is a partition for which we fix a distinguished point (a "tag", "etichetta" in Italian) in each of its intervals.

Given a function $\delta\colon [a,b] \to (0,\infty)$, the tagged partition is said to be *δ-fine* if

$$I_j \subset (c_j - \delta(c_j), c_j + \delta(c_j)) \quad \text{for each } 1 \le j \le n.$$

The following simple lemma could be easily proved also by using compactness of $[a,b]$ via open coverings. The proof we give here, based on completeness of $\mathbb{R}$ (that is, existence of supremum), is taken from Gordon's paper.

Lemma 2.3. *For every $\delta\colon [a,b] \to (0,\infty)$ there exists a δ-fine tagged partition of $[a,b]$.*

Proof. Let E be the set of all points $x \in (a,b]$ such that there exists a δ-fine tagged partition of $[a,x]$. Let us proceed in three easy steps.

(a) *E is nonempty.* Indeed, for every point $x \in (a, a+\delta(a))$ the singleton $\{(a,[a,x])\}$ is a δ-fine partition of $[a,x]$, and hence $x \in E$.

(b) *$s := \sup E$ is an element of E.* Assume this is not the case. Then there exists $u \in E$ such that $s - \delta(s) < u < s$. Now, if $\mathcal{P}$ is a δ-fine tagged partition of $[a,u]$, then $\mathcal{P}' := \mathcal{P} \cup \{(s,[u,s])\}$ is a δ-fine tagged partition of $[a,s]$. But this means that $s \in E$, which is a contradiction.

(c) $s = b$. Assume that $s < b$. Fix $v \in (s, b]$ such that $v < s + \delta(s)$, and a δ-fine tagged partition $\mathcal{P}$ of $[a, s]$. Then $\mathcal{P}' := \mathcal{P} \cup \{(s, [s, v])\}$ is a δ-fine tagged partition of $[a, v]$, contradicting the definition of s. We are done. □

Now we are ready to prove Theorem 2.1.

Proof of Theorem 2.1. Let $|f| \le M$ on $[a, b]$, and let D denote the set of discontinuity points of f in $[a, b]$. Given $\varepsilon > 0$, let $\{J_k\}_1^\infty$ be a sequence of open intervals such that $D \subset \bigcup_{k=1}^\infty J_k$ and $\sum_{k=1}^\infty |J_k| < \varepsilon$. For each $x \in [a, b]$, let us define $\delta(x) > 0$ as follows:

- if $x \notin D$, let $\delta(x) > 0$ be such that $|f(y) - f(x)| < \varepsilon$ whenever $|y - x| < \delta(x)$;
- if $x \in D$, there exists an index $k(x) \in \mathbb{N}$ such that $x \in J_{k(x)}$, and we fix some $\delta(x) > 0$ for which $(x - \delta(x), x + \delta(x)) \subset J_{k(x)}$.

Now, let $\{(c_j, I_j) : j = 1, \ldots, n\}$ be a δ-fine tagged partition of $[a, b]$ (Lemma 2.3). Put $\Delta := \{j : c_j \in D\}$. Then $I_j \subset J_{k(c_j)}$ whenever $j \in \Delta$, and hence

$$\sum_{j \in \Delta} |I_j| \le \sum_{k=1}^{\infty} |J_k| < \varepsilon.$$

Therefore,

$$\begin{aligned}\sum_{j=1}^{n} \omega(f, I_j)|I_j| &= \sum_{j\in\Delta} \omega(f, I_j)|I_j| + \sum_{j\notin\Delta} \omega(f, I_j)|I_j| \\ &\le 2M \sum_{j\in\Delta} |I_j| + 2\varepsilon \sum_{j\notin\Delta} |I_j| \\ &\le 2M\varepsilon + 2\varepsilon(b - a) = 2\varepsilon(M + b - a).\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, Exercise 1.2 implies that $f \in \mathcal{R}[a, b]$. □

Now we have the following easy corollary (cf. Exercise 1.4).

Corollary 2.4. *If $f : [a, b] \to \mathbb{R}$ is bounded and has at most countably many points of discontinuity, then $f \in \mathcal{R}[a, b]$.*

3. Proof of necessity

The proof of necessity is easier than that of the other implication. In the following proof, I° denotes the interior of an interval I.

Theorem 3.1 (Necessity). *If $f \in \mathcal{R}[a, b]$, then (it is bounded and) the set of its discontinuity points is a null set.*

Proof. For $\alpha > 0$, consider the set

$$D_\alpha := \{x \in (a, b) : \omega(f, x) > \alpha\}.$$

Let us show that each such set D_α is a null set. Fix an arbitrary $\varepsilon > 0$. By Exercise 1.4, there exists a partition $\{I_j : j = 1, \dots, n\}$ of $[a, b]$, such that $\sum_{j=1}^{n} \omega(f, I_j)|I_j| < \varepsilon\alpha$. Let E be the (finite) set of the extreme points of the intervals I_j $(1 \le j \le n)$, let and $\Delta := \{j : I_j^\circ \cap D_\alpha \neq \emptyset\}$. Then

$$D_\alpha = (D_\alpha \setminus E) \cup (D_\alpha \cap E) \subset \bigcup_{j\in\Delta} I_j^\circ \cup E. \tag{1}$$

Therefore we have

$$\varepsilon\alpha > \sum_{j\in\Delta} \omega(f, I_j)|I_j| > \alpha \sum_{j\in\Delta} |I_j|\,,$$

and hence $\sum_{j\in\Delta} |I_j| < \varepsilon$. Since E is finite, (1) easily implies that D_ε can be covered by finitely many open intervals sum of whose lengths is less than 2ε. Since $\varepsilon > 0$ was arbitrary, D_α is a null set.

To finish the proof, observe that the set D of points of discontinuity of f in $[a, b]$ is contained in the union $\bigcup_{n\in\mathbb{N}} D_{1/n} \cup \{a, b\}$ (see Exercise 1.1). By Exercise 1.4, D is a null set. □

4. An important corollary

Corollary 4.1. *If* $f, g \in \mathcal{R}[a, b]$*, then the functions*

$$|f|, \quad f \cdot g, \quad \max\{f, g\}, \quad \min\{f, g\} \quad \textit{(and similar)}$$

are Riemann integrable on $[a, b]$.

Proof. The sets of discontinuity points of these functions is clearly contained in the union of the sets of discontinuity points of f and g. Apply Theorem 0.1 together with Exercise 1.4. □

Exercise 4.2. Show that $|f| \in \mathcal{R}[a, b]$ does not imply that $f \in \mathcal{R}[a, b]$.
(*Hint:* consider a Dirichlet-like function with values ± 1.)