INSTITUTE OF MATHEMATICS

THE CZECH ACADEMY OF SCIENCES

A new proof of the bounded convergence theorem for the Riemann–Stieltjes integral

Giselle A. Monteiro

Preprint No. 48-2018

PRAHA 2018

A new proof of the bounded convergence theorem for the Riemann–Stieltjes integral

Giselle A. Monteiro*

April 27, 2018

The bounded convergence theorem, whose first mention dates back to Arzelá's paper [1], is concerned with the convergence of integrals provided the sequence of Riemann integrable functions is uniformly bounded and has an integrable pointwise limit. One could say that this result is the Lebesgue dominated convergence theorem counterpart for Riemann integrals. Indeed, this is how many textbooks approach the bounded convergence theorem: as a trivial consequence of Lebesgue integration theory. The story is not much different when the Stieltjes–type integral is considered. In [5], Luxemburg briefly remarks that the result for Riemann–Stieltjes integral is strongly connected to the properties of the Stieltjes measure for intervals. In [3], the proof of the bounded convergence theorem for the Riemann–Stieltjes integral happens to be long and intricate. Reasoning by contradiction, the proof in [3] relies on a result known as Arzela's lemma, a result whose proof is rather technical unless a background on measure theory is assumed. Therefore, disregarding the knowledge of measure theory, the bounded convergence theorem makes its reputation as being a difficult problem. A number of authors have addressed the question of obtaining a proof independent of the theory of Lebesgue measure for such convergence result, see [5] and the references therein. An interesting elementary proof of the bounded convergence theorem for Riemann integrals was given in [4]; later, similar ideas have been applied to abstract Stieltjes-type integrals in [6]. Inspired by these two papers, our goal is to present a constructive proof of the following theorem:

Theorem A. *Let g be a function of bounded variation on $[a,b]$ and let $\{f_n\}$ be a sequence of real-valued functions defined on $[a,b]$ with the pointwise limit $f\colon [a,b]\to\mathbb{R}$. Assume that there exists a constant $M\geq 0$ such that $|f_n(t)|\leq M$ for all $n\in\mathbb{N}$ and $t\in[a,b]$. If the integrals $\int_a^b f\,\mathrm{d}g$ and $\int_a^b f_n\,\mathrm{d}g$ exist for each $n\in\mathbb{N}$, then*

$$\lim_{n\to\infty}\int_a^b f_n\,\mathrm{d}g=\int_a^b f\,\mathrm{d}g.$$

*Institute of Mathematics, Czech Academy of Sciences, Žitná 25, 115 67 Praha 1, Czech Republic (email: gam@math.cas.cz). Supported by RVO: 67985840.

An important step in early proofs of this theorem is the understanding of the concepts of figures and 'length' of a figure, see [2]. Closely related to the notion of figures is the idea of elementary sets that we consider in this work. Having this in mind, our approach relies on an extended notion of variation which has been briefly discussed in [6]. As we will see, the variation over elementary sets has some properties of a measure. The main analytic tool for proving the bounded convergence theorem is then an analogue of the result presented in [4], with the variation over elementary sets in the place of the Lebesgue measure (see also [6]).

We would like to highlight that the results contained in this notes are the outcome of the research carried out during the preparation of the monograph [7] which shall appear in 2018.

1 Preliminaries

To define the Riemann-Stieltjes we recall that a set $D = \{\alpha_0, \alpha_1, \dots, \alpha_{\nu(D)}\}$, with $\nu(D) \in \mathbb{N}$, is said to be a division of $[a, b]$ if $a = \alpha_0 < \alpha_1 < \cdots < \alpha_{\nu(D)} = b$. A partition of $[a, b]$ is a tagged division, that is, $P = (D, \xi)$ where D is a division of $[a, b]$ and $\xi = (\xi_1, \dots, \xi_{\nu(D)})$ with $\xi_j \in [\alpha_{j-1}, \alpha_j]$, $j = 1, \dots, \nu(D)$. Given a pair of functions $f, g : [a, b] \to \mathbb{R}$, the *Riemann–Stieltjes integral* $\int_a^b f \,\mathrm{d}g$ is defined and equals I if for every $\varepsilon > 0$, there exists a division D_ε of $[a, b]$ satisfying

$$\Big| \sum_{j=1}^{\nu(D)} f(\xi_j)\big(g(\alpha_j) - g(\alpha_{j-1})\big) - I \Big| < \varepsilon$$

for all partitions $P = (D, \xi)$ of $[a, b]$ such that $D \supset D_\varepsilon$.

The *(Jordan) variation of f on* $[a, b]$ is given by

$$\mathrm{var}_a^b f = \sup \sum_{j=1}^{\nu(D)} |f(\alpha_j) - f(\alpha_{j-1})|$$

where the supremum is taken over all divisions $D = \{\alpha_0, \alpha_1, \dots, \alpha_{\nu(D)}\}$ of $[a, b]$. If $\mathrm{var}_a^b f < \infty$ we say that f *is of bounded variation on* $[a, b]$. We remark that a function of bounded variation has at most countably many points of discontinuity. Moreover, the one–sided limits $f(t-)$ and $f(t+)$ exist at every point $t \in [a, b]$ (with the convention $f(a-) = f(a)$ and $f(b+) = f(b)$), and the following inequality holds

$$\sum_{t \in [a,b]} \Big(|\Delta^+ f(t)| + |\Delta^- f(t)| \Big) \le \mathrm{var}_a^b f, \tag{1.1}$$

where $\Delta^- f(t) = f(t) - f(t-)$ and $\Delta^+ f(t) = f(t+) - f(t)$ for $t \in [a, b]$.

The notion of Jordan variation can be easily extended to arbitrary intervals by rethinking the meaning of divisions. More precisely, for $f : [a, b] \to \mathbb{R}$ and a subinterval $J \subset [a, b]$, the variation of f over J is defined by

$$\operatorname{var}(f, J) = \sup \sum_{j=1}^{\nu(D)} |f(\alpha_j) - f(\alpha_{j-1})|,$$

where the supremum is taken over all finite subsets $D = \{\alpha_0, \alpha_1, \dots, \alpha_{\nu(D)}\}$ of J with $\alpha_0 < \alpha_1 < \cdots < \alpha_{\nu(D)}$. We use the convention that $\operatorname{var}(f, \emptyset) = 0$ and $\operatorname{var}(f, [c]) = 0$, where $[c]$ denotes the degenerate interval consisting of a single real number $c \in [a, b]$. Clearly, for $J = [c, d] \subset [a, b]$ we have $\operatorname{var}(f, J) = \operatorname{var}_c^d f$. Moreover, it follows from the definition that $\operatorname{var}(f, J_2) \le \operatorname{var}(f, J_1)$ for $J_2 \subset J_1$ subintervals of $[a, b]$.

Among the properties of variation over arbitrary intervals which have been discussed in [6], we highlight the following: *If f is a function of bounded variation on $[a, b]$, and $a \le c < d \le b$, then*

$$\operatorname{var}_c^d f = \operatorname{var}_{[c,d)} f + |\Delta^- f(d)|, \tag{1.2}$$

$$\operatorname{var}_c^d f = \operatorname{var}_{(c,d]} f + |\Delta^+ f(c)|, \tag{1.3}$$

$$\operatorname{var}_c^d f = \operatorname{var}_{(c,d)} f + |\Delta^+ f(c)| + |\Delta^- f(d)|. \tag{1.4}$$

Naturally, if f is a function of bounded variation which is also continuous, then

$$\operatorname{var}_c^d f = \operatorname{var}_{[c,d)} f = \operatorname{var}_{(c,d]} f = \operatorname{var}_{(c,d)} f \quad \text{for} \quad a \le c < d \le b.$$

It is well-known that a function f of bounded variation on $[a, b]$ can be decomposed into continuous and jump parts, namely, $f = f^{\mathrm{C}} + f^{\mathrm{B}}$ where f^{C} is continuous and f^{B} is given by

$$f^{\mathrm{B}}(t) = f(a) + \sum_{k=1}^{\infty} \left(\Delta^+ f(s_k)\, \chi_{(s_k,b]}(t) + \Delta^- f(s_k)\, \chi_{[s_k,b]}(t) \right) \quad \text{for} \quad t \in [a, b],$$

where $\{s_k\}$ is the set of discontinuity points of f in $[a, b]$, and χ_A stands for the characteristic function of the set $A \subset \mathbb{R}$. A simple computation shows that

$$\Delta^{\pm} f^{\mathrm{B}}(t) = \Delta^{\pm} f(t),\ t \in [a, b], \qquad \operatorname{var}_a^b f^{\mathrm{B}} = \sum_{t \in [a,b]} \Big(|\Delta^+ f(t)| + |\Delta^- f(t)| \Big). \tag{1.5}$$

(Note that (1.1) ensures that the sum above is finite). Therefore, using the relations (1.2)–(1.4), we can estimate the variation of the jump part, f^{B}, over arbitrary intervals. Indeed, assuming $J = [c, d)$ for some $c < d$ in $[a, b]$, it follows from (1.2) and (1.5) that

$$\operatorname{var}(f^{\mathrm{B}}, J) = \operatorname{var}_c^d f^{\mathrm{B}} - |\Delta^- f^{\mathrm{B}}(d)| = |\Delta^+ f(c)| + \sum_{s \in (c,d)} \Big(|\Delta^+ f(s)| + |\Delta^- f(s)| \Big).$$

Similar expressions can be derived in the cases $J = (c, d]$ or $J = (c, d)$, yielding the following lemma.

Lemma 1.1. *Let f be a function of bounded variation on $[a, b]$ and let D denote the set of discontinuity points of f in $[a, b]$. Then for every subinterval J of $[a, b]$ we have*

$$\operatorname{var}(f^{\mathrm{B}}, J) \le \sum_{s \in J \cap D} \Big(|\Delta^+ f(s)| + |\Delta^- f(s)| \Big).$$

We now consider the notion of variation over elementary sets introduced in [6]. By an *elementary set* we mean a finite union of bounded intervals. A collection of intervals $\{J_k \colon k = 1, \dots, N\}$ is called a *minimal decomposition* of an elementary set E if $E = \bigcup_{k=1}^N J_k$ and the union $J_k \cup J_\ell$ is not an interval whenever $k \neq \ell$. Note that the minimal decomposition of an elementary set is uniquely determined. Moreover, the intervals of such a decomposition are pairwise disjoint. Having this in mind, we define:

Definition 1.2. Given a function $f : [a, b] \to \mathbb{R}$ and an elementary subset E of $[a, b]$, the variation of f over E is

$$\operatorname{var}(f, E) = \sum_{k=1}^{N} \operatorname{var}(f, J_k),$$

where $\{J_k \colon k = 1, \dots, N\}$ is the minimal decomposition of E.

Out of curiosity, we list here some properties of the variation over elementary sets:

(i) For $E_2 \subset E_1$ elementary subsets of $[a, b]$, we have $\operatorname{var}(f, E_2) \le \operatorname{var}(f, E_1)$.

(ii) If f is of bounded variation on $[a, b]$, then $\operatorname{var}(f, E)$ is finite for every elementary subset E of $[a, b]$.

(iii) If f is of bounded variation on $[a, b]$, and E is an elementary subset of $[a, b]$, then for every $\varepsilon > 0$, there exists an elementary subset H_ε of E such that H_ε is closed in $[a, b]$ and
$$\operatorname{var}(f, E) - \varepsilon < \operatorname{var}(f, H_\varepsilon).$$

(iv) If f is of bounded variation and continuous on $[a, b]$, then for every E_1 and E_2 elementary subsets of $[a, b]$ we have
$$\operatorname{var}(f, E_1 \cup E_2) \le \operatorname{var}(f, E_1) + \operatorname{var}(f, E_2).$$
In the particular case when E_1 and E_2 are nonoverlapping, the equality holds.

The properties above indicate that the variation over elementary sets behaves somehow like a measure. Next, we present an assertion which resembles the property of continuity from above for measures. This property has been already established in [6] in the case when f is a continuous function with bounded variation. For the reader's convenience, we repeat its statement here. To this end, we fix the following notation: For each set $A \subset [a, b]$, let

$$v(A) := \sup\{\, \mathrm{var}(f, E) : E \text{ elementary subset of } A\}.$$

Lemma 1.3. *Let f be a function of bounded variation on $[a, b]$ which is also continuous. Assume that $\{A_n\}$ is a sequence of subsets of $[a, b]$ such that $A_{n+1} \subset A_n$ for $n \in \mathbb{N}$, and $\bigcap_n A_n \neq \emptyset$. Then $\lim_{n\to\infty} v(A_n) = 0$.*

The following theorem is a generalization of the lemma above and is an essential tool in our approach to a constructive proof of the bounded convergence theorem.

Theorem 1.4. *Let f be a function of bounded variation on $[a, b]$. Assume that $\{A_n\}$ is a sequence of subsets of $[a, b]$ such that $A_{n+1} \subset A_n$ for $n \in \mathbb{N}$, and $\bigcap_n A_n \neq \emptyset$. Then $\lim_{n\to\infty} v(A_n) = 0$.*

Proof. First, let us write $f = f^{\mathrm{C}} + f^{\mathrm{B}}$, where f^{C} and f^{B} are the continuous part and the jump part of the function f, respectively. For $A \subset [a, b]$ let

$$v^{\mathrm{C}}(A) = \sup\{\, \mathrm{var}(f^{\mathrm{C}}, E) : E \text{ elementary subset of } A\},$$
$$v^{\mathrm{B}}(A) = \sup\{\, \mathrm{var}(f^{\mathrm{B}}, E) : E \text{ elementary subset of } A\}.$$

Clearly $v(A) \leq v^{\mathrm{C}}(A) + v^{\mathrm{B}}(A)$ for every $A \subset [a, b]$. In view of Lemma 1.3 it suffices to prove that $\lim_{n\to\infty} v^{\mathrm{B}}(A_n) = 0$.

Let $\varepsilon > 0$ be given and denote by $\{s_k\}$ the set of discontinuity points of f in $[a, b]$. The absolute convergence of the series in (1.1) guarantees that there exists $n_0 \in \mathbb{N}$ satisfying

$$\sum_{k=n_0+1}^{\infty} \Big(|\Delta^+ f(s_k)| + |\Delta^- f(s_k)|\Big) < \frac{\varepsilon}{2}, \tag{1.6}$$

For each $n \in \mathbb{N}$, by the definition of $v^{\mathrm{B}}(A_n)$, we can choose an elementary subset E_n of A_n such that

$$v^{\mathrm{B}}(A_n) - \frac{\varepsilon}{2} < \mathrm{var}(f^{\mathrm{B}}, E_n). \tag{1.7}$$

Note that by Lemma 1.1 we have

$$\mathrm{var}(f^{\mathrm{B}}, E_n) \leq \sum_{s_k \in E_n} \Big(|\Delta^+ f(s_k)| + |\Delta^- f(s_k)|\Big) \leq \sum_{s_k \in A_n} \Big(|\Delta^+ f(s_k)| + |\Delta^- f(s_k)|\Big).$$

Since $\bigcap_n A_n = \emptyset$, there exists a number $n_1 \in \mathbb{N}$ such that

$$\{s_1, \dots, s_{n_0}\} \cap A_n = \emptyset \quad \text{for all} \quad n \geq n_1,$$

and as a consequence of (1.6) we obtain

$$\operatorname{var}(f^{\mathrm{B}}, E_n) \leq \sum_{s_k \in A_n} \Big(|\Delta^+ f(s_k)| + |\Delta^- f(s_k)| \Big) < \frac{\varepsilon}{2} \quad \text{for} \quad n \geq n_1.$$

This together with (1.7) implies that $v^{\mathrm{B}}(A_n) < \varepsilon$ for every $n \geq n_1$, showing that the equality $\lim_{n\to\infty} v^{\mathrm{B}}(A_n) = 0$ holds. □

2 Bounded convergence theorem

Let g be a function of bounded variation on $[a, b]$, and assume that $\{f_n\}$ and f satisfy the assumptions of the bounded convergence theorem (Theorem A). In this case, $\{|f_n - f|\}$ defines a sequence of nonnegative functions which is uniformly bounded and converges pointwise to zero. Moreover, the existence of the integrals $\int_a^b f \,\mathrm{d}g$ and $\int_a^b f_n \,\mathrm{d}g$ implies that the integral $\int_a^b |f_n(x) - f(x)| \,\mathrm{d}(\operatorname{var}_a^x g)$ exists for every $n \in \mathbb{N}$, and

$$\Big| \int_a^b f_n(x) \,\mathrm{d}g(x) - \int_a^b f(x) \,\mathrm{d}g(x) \Big| \leq \int_a^b |f_n(x) - f(x)| \,\mathrm{d}(\operatorname{var}_a^x g),$$

(see Theorem II.14.4 in [3] for details). Consequently, in order to obtain Theorem A it is enough to prove the following assertion:

Theorem B. *Let g be a nondecreasing function defined on $[a, b]$. Assume that $\{f_n\}$ is a sequence of functions defined on $[a, b]$ which satisfies the following conditions:*

(i) $\lim_{n\to\infty} f_n(t) = 0$ *for* $t \in [a, b]$*;*

(ii) There exists $M \geq 0$ such that $0 \leq f_n(t) \leq M$ for all $t \in [a, b]$ and $n \in \mathbb{N}$;

(iii) The integral $\int_a^b f_n \,\mathrm{d}g$ exist for each $n \in \mathbb{N}$.

Then

$$\lim_{n\to\infty} \int_a^b f_n \,\mathrm{d}g = 0. \tag{2.1}$$

In order to prove Theorem B, we can benefit from the relation between the Riemann–Stietjes integral and the Darboux–Stieltjes integral. The latter approach to Stieltjes integration is based on lower and upper sums. More precisely, for a pair of functions $f, g\colon [a,b] \to \mathbb{R}$, with f bounded and g nondecreasing, given a division $D = \{\alpha_0, \alpha_1, \ldots, \alpha_{\nu(D)}\}$ of $[a,b]$, we define

$$L(f, \mathrm{d}g, D) = \sum_{j=1}^{\nu(D)} m_j(g(\alpha_j) - g(\alpha_{j-1})), \quad U(f, \mathrm{d}g, D) = \sum_{j=1}^{\nu(D)} M_j(g(\alpha_j) - g(\alpha_{j-1})),$$

where $m_j = \inf_{t\in[\alpha_{j-1},\alpha_j]} f(t)$ and $M_j = \sup_{t\in[\alpha_{j-1},\alpha_j]} f(t)$, $j = 1, \ldots, \nu(D)$. The *lower integral* of f with respect to g is given by

$$\underline{\int_a^b} f\, \mathrm{d}g = \sup\Big\{L(f, \mathrm{d}g, D) : D \text{ is a division of } [a,b]\Big\},$$

while the *upper integral* of f with respect to g corresponds to

$$\overline{\int_a^b} f\, \mathrm{d}g = \inf\Big\{U(f, \mathrm{d}g, D) : D \text{ is a division of } [a,b]\Big\}.$$

In the case when $\overline{\int_a^b} f\, \mathrm{d}g = \underline{\int_a^b} f\, \mathrm{d}g$ is a finite number, the common value of both integrals is called the *Darboux–Stieltjes integral* of f with respect to g.

As observed in Section II.13 in [3], for g nondecreasing and f bounded, the existence of the Riemann–Stieltjes integral is equivalent to the existence of the Darboux–Stieltjes integral, and

$$\int_a^b f\, \mathrm{d}g = \overline{\int_a^b} f\, \mathrm{d}g = \underline{\int_a^b} f\, \mathrm{d}g. \tag{2.2}$$

We are now ready to prove Theorem B.

Proof of Theorem B. Without loss of generality, assume that g is nonconstant (otherwise the statement is obvious), and denote $\gamma = g(b) - g(a)$. Given $\varepsilon > 0$, for each $n \in \mathbb{N}$ define

$$A_n = \Big\{t \in [a,b] : f_m(t) \geq \frac{\varepsilon}{3\,\gamma} \text{ for some } \ m \geq n\Big\}.$$

It is not difficult to see that $A_{n+1} \subset A_n$ for $n \in \mathbb{N}$ and $\bigcap_n A_n = \emptyset$. Let

$$v(A_n) = \sup\{\,\mathrm{var}(g, E) : E \text{ elementary subset of } A_n\},\ n \in \mathbb{N},$$

(with $v(A_n) = 0$ in case $A_n = \emptyset$). From Theorem 1.4 we know that $\lim_{n\to\infty} v(A_n) = 0$, thus there exists $n_\varepsilon \in \mathbb{N}$ such that $v(A_n) < \frac{\varepsilon}{3\,M}$ for $n \geq n_\varepsilon$, and consequently

$$\mathrm{var}(g, E) < \frac{\varepsilon}{3\,M} \quad \text{for } \ n \geq n_\varepsilon \text{ and } E \text{ elementary subset of } A_n. \tag{2.3}$$

Now, fix an arbitrary $n \geq n_\varepsilon$. By (2.2) we have $\int_a^b f_n \, \mathrm{d}g = \underline{\int_a^b} f_n \, \mathrm{d}g$. Hence, we can choose a division $D = \{\alpha_0, \alpha_1, \dots, \alpha_{\nu(D)}\}$ of $[a, b]$ such that

$$\int_a^b f_n \, \mathrm{d}g - \frac{\varepsilon}{3} < L(f_n, \mathrm{d}g, D) = \sum_{j=1}^{\nu(D)} m_j \, (g(\alpha_j) - g(\alpha_{j-1})), \tag{2.4}$$

where $m_j = \inf_{t \in [\alpha_{j-1}, \alpha_j]} f_n(t)$. Consider the sets

$$\Lambda = \big\{ j \in \{1, \dots, \nu(D)\} : m_j \geq \frac{\varepsilon}{3\,\gamma} \big\}, \quad E = \bigcup_{j \in \Lambda} [\alpha_{j-1}, \alpha_j].$$

Clearly E is an elementary set (possibly empty) with $E \subset A_n$; and (2.3) yields

$$\sum_{j \in \Lambda} (g(\alpha_j) - g(\alpha_{j-1})) = \mathrm{var}(g, E) < \frac{\varepsilon}{3\,M}. \tag{2.5}$$

Noting that $m_j < \frac{\varepsilon}{3\gamma}$ for $j \notin \Lambda$, and using (2.4)-(2.5) we obtain

$$\begin{aligned}
\int_a^b f_n \, \mathrm{d}g &< \frac{\varepsilon}{3} + \sum_{j \in \Lambda} m_j \, (g(\alpha_j) - g(\alpha_{j-1})) + \sum_{j \notin \Lambda} m_j \, (g(\alpha_j) - g(\alpha_{j-1})) \\
&< \frac{\varepsilon}{3} + M \sum_{j \in \Lambda} (g(\alpha_j) - g(\alpha_{j-1})) + \frac{\varepsilon}{3\,\gamma} \sum_{j \notin \Lambda} (g(\alpha_j) - g(\alpha_{j-1})) \\
&\leq \frac{2\,\varepsilon}{3} + \frac{\varepsilon}{3\,\gamma} \, (g(b) - g(a)) < \varepsilon.
\end{aligned}$$

In summary,

$$\int_a^b f_n \, \mathrm{d}g < \varepsilon \quad \text{for every} \;\; n \geq n_\varepsilon,$$

which shows that (2.1) holds and concludes the proof. □

References

[1] C. Arzelà, Sulla integrazione per serie. I, II. *Rom. Acc. L. Rend.* (4) I (1885), 532–537, 566–569.

[2] R. Gordon, A convergence theorem for the Riemann integral. *Math. Mag.* **73** (2000) 141–147.

[3] T. H. Hildebrandt, *Introduction to theory of integration.* Academic Press, New York & London, 1963.

[4] J. W. Lewin, A truly elementary approach to the bounded convergence theorem. *Amer. Math. Monthly* **93** (1986), 395–397.

[5] W. A. J. Luxemburg, Arzela's dominated convergence theorem for the Riemann integral. *Amer. Math. Monthly* **78** (1971) 970–979.

[6] G.A. Monteiro, U.M. Hanung and M. Tvrdý, Bounded convergence theorem for abstract Kurzweil–Stieltjes integral. *Monatshefte für Mathematik* 180 (2016), 409–434.

[7] G. A. Monteiro, A. Slavík, M. Tvrdý, *Kurzweil–Stieltjes integral: theory and applications*. Series in Real Analysis - vol 15, World Scientific, Singapore, 2018.