RESEARCH

Real Analysis Exchange
Vol. 40(2), 2014/2015, pp. 319–354

Rodrigo López Pouso,* Department of Mathematical Analysis, Institute of Mathematics, Faculty of Mathematics, University of Santiago de Compostela, 15782, Santiago de Compostela, Spain. email: `rodrigo.lopez@usc.es`

Adrián Rodríguez, Department of Mathematical Analysis, Faculty of Mathematics, University of Santiago de Compostela, 15782, Santiago de Compostela, Spain. email: `adrian.rguez.sanjurjo@gmail.com`

A NEW UNIFICATION OF CONTINUOUS, DISCRETE, AND IMPULSIVE CALCULUS THROUGH STIELTJES DERIVATIVES

Abstract

We study a simple notion of derivative with respect to a function which we assume to be nondecreasing and continuous from the left everywhere. Derivatives of this type were already considered by Young in 1917 and Daniell in 1918, in connection with the fundamental theorem of calculus for Stieltjes integrals. We show that our definition contains as a particular case the delta derivative in time scales, thus providing a new unification of the continuous and the discrete calculus. Moreover, we can consider differential equations in the new sense, and we show that not only dynamic equations on time scales, but also ordinary differential equations with impulses at fixed times are particular cases. We study almost everywhere differentiation of monotone functions and the fundamental theorems of calculus which connect our new derivative with Lebesgue–Stieltjes and Kurzweil–Stieltjes integrals. These fundamental theorems are the key for reducing differential equations with the new derivative to generalized integral equations, for which many theoretical results are already available thanks to Kurzweil, Schwabik and their followers.

Mathematical Reviews subject classification: Primary: 26A24, 26A36; Secondary: 26A45, 34A34

Key words: Differentiation, Lebesgue–Stieltjes integration, Kurzweil–Stieltjes integration, Fundamental theorem of calculus

Received by the editors September 4, 2014

Communicated by: Luisa Di Piazza

*Partially supported by Ministerio de Economía y Competitividad, Spain, and FEDER projects MTM2010-15314 and MTM2013-43014-P.

1 Introduction and definition

Slavík [25] and Federson *et al.* [10, 11] have recently shown that many classes of generalized differential equations (such as equations on time scales or equations with impulses) can be reduced to a single class of integral problems using generalized Stieltjes integration. This paper is the result of the authors' attempt to establish the differential counterpart of the main ideas in [10, 11, 25].

Here and henceforth we assume that $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a monotone nondecreasing function which is continuous from the left everywhere. Nothing else is needed for the main definition in this paper.

Definition 1.1. *The derivative with respect to g (or g–derivative) of a real–valued real function f at a point $x \in \mathbb{R}$ is denoted and defined as follows, provided that the corresponding limit exists:*

$$f'_g(x) = \lim_{y \to x} \frac{f(y) - f(x)}{g(y) - g(x)} \quad \textit{if } g \textit{ is continuous at } x, \textit{ or} \tag{1}$$

$$f'_g(x) = \lim_{y \to x^+} \frac{f(y) - f(x)}{g(y) - g(x)} \quad \textit{if } g \textit{ is discontinuous at } x. \tag{2}$$

We say that f is g–differentiable at x_0 if $f'_g(x_0)$ exists.

Definition 1.1 is a simplified version of a definition already used by Daniell [7] in 1918, see also Young [27] and [8]. Some more classical references for derivatives with respect to functions are due to Lebesgue [15], Saks [22, Chapter 9, part 5], [6, 12, 16, 20], and, more recently, Gradinaru [14] and Averna and Preiss [1]. Probably our results remain valid without left–continuity of g, but removing that assumption introduces more technicalities in the definition and in the proofs and does not improve the applicability of the results for the purposes considered in this paper. Moreover, our definition of absolutely continuous function with respect to g is particularly simple thanks to left–continuity; see Definition 5.1.

The limit in (1) can only be understood as a limit when y tends to x and $g(y) \neq g(x)$. In particular, (1) does not make sense at the points of the set

$$C_g = \{x \in \mathbb{R} : g \text{ is constant on } (x - \varepsilon, x + \varepsilon) \text{ for some } \varepsilon > 0\}, \tag{3}$$

and no matter how f is. This seeming drawback is not important because the set C_g is negligible in a sense to be made precise.

Another preliminary remark about (1), which also reenforces the interpretation of the g–derivative as a relative rate of change, is the following one: if

$g'(x)$ exists and it is positive, and if $f'(x)$ exists, then $f'_g(x)$ exists and

$$f'_g(x) = \frac{f'(x)}{g'(x)}. \tag{4}$$

Since this formula is not valid in general, we have to study the g–derivative directly. In fact, the set of points where g' exists can be very small in the sense needed in this paper, as we prove in Proposition 4.7.

It is obvious from (2) that a special role will be played by the set of discontinuity points of g, which we denote by

$$D_g = \{x \in \mathbb{R} : g(x^+) - g(x) > 0\}, \tag{5}$$

where, as usual, $g(x^+)$ stands for the right–hand limit of g at x. We shall show in Section 3 that useful derivators g have $C_g \neq \emptyset$ or $D_g \neq \emptyset$ or both.

Notice that D_g is countable because g is monotone, and that for each $x \in D_g$, the limit (2) exists if and only if $f(x^+)$ exists and, in that case,

$$f'_g(x) = \frac{f(x^+) - f(x)}{g(x^+) - g(x)}.$$

Choosing only the right–hand limit to define $f'_g(x)$ at $x \in D_g$ is necessary for the theory that we are going to set up and also agrees with the applications. A better explanation of this fact will be given in the next section in terms of the Lebesgue–Stieltjes measure induced by g.

The rest of this paper is organized as follows: Section 2 contains basic information on Definition 1.1 and its relation with the Lebesgue–Stieltjes measure generated by g. In Section 3 we show that Definition 1.1 is really meaningful by proving that it contains as a particular case the delta derivative in the theory of time scales, and it even allows us to express differential equations with impulses in the form of a unique differential equation with an adequate g–derivative. The remaining sections contain the main results in this paper: in Section 4 we prove a generalization of the Lebesgue Differentiation Theorem with usual derivatives replaced by g–derivatives, and in Section 5 we prove the fundamental theorems of calculus for the Lebesgue–Stieltjes integral in terms of the g–derivative. In Section 6 we prove similar results for the Kurzweil–Stieltjes integral.

2 Preliminary observations

Basic results on derivatives have their analogue for g–derivatives. For completeness we include the most important of them under generous assumptions. Proofs are easy, so we omit them for brevity. See [14] for related results.

Proposition 2.1. *Let $x \in \mathbb{R} \setminus D_g$ and assume $f'_g(x)$ exists. Then $x \notin C_g$ and f is continuous from the right (respectively, from the left) at x provided that $g(y) > g(x)$ for all $y > x$ (respectively, $g(y) < g(x)$ for all $y < x$). If $x \in D_g$, then $f'_g(x)$ exists if and only if $f(x^+)$ exists.*

Proposition 2.2. *Let f_1, f_2 be real–valued real functions defined on a neighborhood of a point $x \in \mathbb{R}$ such that $g(y) \neq g(x)$ for $y \neq x$. If f_i is g–differentiable at x for $i = 1, 2$, then the following results hold:*

1. *The function $c_1 f_1 + c_2 f_2$ if g–differentiable at x for any choice of $c_i \in \mathbb{R}$, $i = 1, 2$, and*
$$(c_1 f_1 + c_2 f_2)'_g(x) = c_1 (f_1)'_g(x) + c_2 (f_2)'_g(x).$$
2. *The product $f_1 f_2$ is g–differentiable at x and*
$$(f_1 f_2)'_g(x) = (f_1)'_g(x) f_2(x^+) + f_1(x^+)(f_2)'_g(x).$$
3. *If $f_2(x) f_2(x^+) \neq 0$, then f_1/f_2 is g–differentiable at x and*
$$(f_1/f_2)'_g(x) = \frac{(f_1)'_g(x) f_2(x) - f_1(x)(f_2)'_g(x)}{f_2(x) f_2(x^+)}.$$

We have two versions of the chain rule for the g–derivative of a composition. We only consider points $x \in \mathbb{R} \setminus D_g$ because similar formulas for $x \in D_g$ are more complicated and less useful. Notice that usual derivatives are involved.

Theorem 2.3. (Chain rule for g–derivatives) *Let f be a real–valued real function defined on a neighborhood of $x \in \mathbb{R} \setminus D_g$, and let h be another function defined in a neighborhood of $f(x)$. The following results hold for the g–derivative of the composition $h \circ f$ at x:*

1. *If $h'(f(x))$ and $f'_g(x)$ exist, then $(h \circ f)'_g(x)$ exists and*
$$(h \circ f)'_g(x) = h'(f(x)) f'_g(x)$$
exists.
2. *If $h'_g(f(x))$, $g'(f(x))$ and $f'_g(x)$ exist, then $(h \circ f)'_g(x)$ exists and*
$$(h \circ f)'_g(x) = h'_g(f(x)) g'(f(x)) f'_g(x)$$
exists.

Next we focus our attention on the basic aspects of the Lebesgue–Stieltjes measure induced by g. Since g is nondecreasing and left–continuous, it generates a unique Lebesgue–Stieltjes measure $\mu_g : \mathcal{M}_g \longrightarrow [0, +\infty]$, where $\mathcal{M}_g$ is a σ–algebra of subsets of the reals containing all Borel sets. The usual way of constructing μ_g starts with the fundamental formula

$$\mu_g([a, b)) = g(b) - g(a) \quad \text{for every } a, b \in \mathbb{R},\ a < b,$$

and then we define an outer measure as

$$\mu_g^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu_g([a_n, b_n)) \ : \ A \subset \bigcup_{n=1}^{\infty} [a_n, b_n) \right\} \quad \text{for every } A \subset \mathbb{R}.$$

The σ–algebra $\mathcal{M}_g$ is then defined as the family of all subsets of the reals which satisfy a Carathéodory–type identity, and μ_g is defined as the restriction of μ_g^* to $\mathcal{M}_g$. See [5, 21, 23] for more details.

We shall use standard denominations such as "set of g–measure zero" for any null set with respect to μ_g; a property "holds g–almost everywhere" if it holds outside a set of g–measure zero; a "g–measurable function" is any function $f : E \in \mathcal{M}_g \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ such that $f^{-1}(V) \in \mathcal{M}_g$ for any open subset $V \subset \overline{\mathbb{R}}$.

While μ_g shares many properties with the Lebesgue measure, the main difference between the two is the fact that μ_g may have atoms. Specifically, for any $x \in D_g$ we use a basic property of measures to compute

$$\begin{aligned} \mu_g(\{x\}) &= \mu_g \left(\bigcap_{n=1}^{\infty} [x, x + 1/n) \right) = \lim_{n \to \infty} \mu_g([x, x + 1/n)) \\ &= \lim_{n \to \infty} [g(x + 1/n) - g(x)] = g(x^+) - g(x) > 0. \qquad (6) \end{aligned}$$

We now proceed to explain that our choice of a right–hand side limit in (2) is necessary. One of the main results in this paper (which we prove in Section 5) concerns almost everywhere g–differentiability of indefinite Lebesgue–Stieltjes integrals.

Theorem 2.4. *Assume that $f : [a, b) \longrightarrow \overline{\mathbb{R}}$ is integrable on $[a, b)$ with respect to μ_g, and consider its indefinite Lebesgue–Stieltjes integral*

$$F(x) = \int_{[a,x)} f \, d\mu_g \quad \textit{for all } x \in [a, b].$$

Then there is a g–measurable set $N \subset [a, b]$ such that $\mu_g(N) = 0$ and

$$F_g'(x) = f(x) \quad \textit{for all } x \in [a, b] \setminus N.$$

Our next example shows that Theorem 2.4 is false if we replace the right–hand side limit in (2) by the left–hand side one.

Example 2.1. *Let $g(x) = x$ for $x \le 1$ and $g(x) = x + 1$ for $x > 1$, which is nondecreasing and left–continuous.*

Consider now a function $f(x) = 1$ for all $x \in [0, 1)$, $f(1) = c \in \mathbb{R}$, and $f(x) = 2$ for all $x \in (1, 2)$. We claim that f is Lebesgue–Stieltjes integrable on $[0, 2)$ with respect to μ_g for any choice of $c \in \mathbb{R}$. Indeed, f is a Borel–measurable function, hence $\mathcal{M}_g$–measurable, and now the additivity property of the integral with respect to measurable partitions of its domain gives

$$\int_{[0,2)} |f|\, d\mu_g = \int_{[0,1)} |f|\, d\mu_g + \int_{\{1\}} |f|\, d\mu_g + \int_{(1,2)} |f|\, d\mu_g = 1 + |c| + 2 < \infty.$$

With the notation of Theorem 2.4, we compute for $x \in [0, 1]$

$$F(x) = \int_{[0,x)} f d\mu_g = \mu_g([0, x)) = x,$$

and for $x \in (1, 2]$ we have

$$F(x) = \int_{[0,x)} f\, d\mu_g = \mu_g([0, 1)) + \int_{\{1\}} f\, d\mu_g + \int_{(1,x)} f\, d\mu_g = 2x + c - 1.$$

Notice that, in agreement with Theorem 2.4, we have for any $c \in \mathbb{R}$ that

$$F'_g(1) = \lim_{y \to 1^+} \frac{F(y) - F(1)}{g(y) - g(1)} = c = f(1),$$

and, on the other hand,

$$\lim_{y \to 1^-} \frac{F(y) - F(1)}{g(y) - g(1)} = 1,$$

which is not necessarily equal to $f(1)$. Note also that we cannot simply dump the point $x = 1$ in the exceptional set N because 1 is an atom for μ_g.

Attention is turned now to the set of points around which g is constant. Let us see that it is a g–null set.

Proposition 2.5. *Let C_g be as in (3). Then $\mu_g(C_g) = 0$.*

PROOF. The set C_g is open, hence a countable union of disjoint open intervals. By countable additivity of μ_g, it then suffices to prove that if (a,b) is one of those open intervals, then $\mu_g(a,b)=0$. To prove it, we use the fact that g is constant on (a,b):

$$\begin{aligned}\mu_g(a,b) = \mu_g\left(\bigcup_{n=1}^{\infty}[a+1/n,b)\right) &= \lim_{n\to\infty}\mu_g([a+1/n,b)) \\ &= \lim_{n\to\infty}[g(b)-g(a+1/n)] = 0.\end{aligned}$$

□

Notice that the definition of g–derivative only makes sense from one side at the endpoints of the connected components of C_g which do not belong to D_g. At some steps of later argumentations, it will be convenient to disregard those points which, fortunately, form a g–null set.

Proposition 2.6. *Let C_g and D_g be as in (3) and (5), respectively. If $C_g = \bigcup_{n=1}^{\infty}(a_n,b_n)$, where the intervals are pairwise disjoint, then the set*

$$N_g = \{a_n,\, b_n \,:\, n\in\mathbb{N}\}\setminus D_g \tag{7}$$

has zero g–measure.

PROOF. Notice that N_g is a countable union of singletons which have zero g–measure because those points do not belong to D_g. □

Remark 2.1. *As an immediate consequence of Proposition 2.6, we note that a property of x holds g–almost everywhere in some $E\in\mathcal{M}_g$ if and only if it holds g–almost everywhere in $E\setminus A$, where*

$$A = C_g\cup N_g. \tag{8}$$

Notice also that if $x\notin A$, then

$$g(y)\neq g(x) \quad \text{for every } y\neq x,$$

and therefore the limit in (1) must be studied from both sides (if f is defined on both sides of x).

The set C_g is a subset of $\{x\in\mathbb{R} : \exists\, g'(x)=0\}$ which can be very big with respect to the Lebesgue measure: see [26, Theorem 4.54] for examples of strictly increasing continuous functions g having zero derivative almost everywhere (in Lebesgue's sense). However, we shall prove that $\mu_g(\{x\in\mathbb{R} : \exists\, g'(x)=0\})=0$ in Proposition 4.6.

3 What is it good for?

Before carrying out deeper analysis, we illustrate the applicability of Definition 1.1. In this section only, we change notation and use t as independent variable, as it is customary in the settings that we are going to consider.

3.1 A new unification of discrete and continuous calculus

Let $\mathbb{T}$ be a time scale; i.e., a nonempty closed subset of the reals. Following Slavík [25], we introduce a function $g : \mathbb{R} \longrightarrow \mathbb{T}$ defined as

$$g(t) = \inf\{s \in \mathbb{T} \,:\, s \geq t\} \quad \text{for all } t \in \mathbb{R}, \tag{9}$$

and not to be confused with the usual forward jump operator

$$\sigma(t) = \inf\{s \in \mathbb{T} \,:\, s > t\} \quad \text{for all } t \in \mathbb{T} \text{ (by convention, } \inf \emptyset = \sup \mathbb{T}).$$

It is easy to prove that g is nondecreasing and left–continuous everywhere. Note also that $t_0 \in \mathbb{T}$ implies that $g(t_0) = t_0$.

Now for every function $f : \mathbb{T} \longrightarrow \mathbb{R}$ we define its Slavík extension to be the function $\bar{f} : \mathbb{R} \longrightarrow \mathbb{R}$ defined as

$$\bar{f}(t) = f(g(t)) \quad \text{for all } t \in \mathbb{R}. \tag{10}$$

We are now in a position to prove that delta derivatives in time scales can be seen as g–derivatives. See [3] for details on calculus on time scales.

Theorem 3.1. *Let $\mathbb{T}$ be a time scale and let $f : \mathbb{T} \longrightarrow \mathbb{R}$ be continuous from the left at every right–scattered point. Let g and $\bar{f}$ be as in (9) and (10), respectively. The function f is Δ–differentiable at $t_0 \in \mathbb{T}$ if and only if $\bar{f}$ is g–differentiable at t_0, and*

$$\bar{f}'_g(t_0) = f^{\Delta}(t_0). \tag{11}$$

Proof. We consider two cases separately.

Case I. We assume that t_0 is right–scattered. The function f is left–continuous at t_0 by our assumptions. Hence there exists

$$f^{\Delta}(t_0) = \frac{f(\sigma(t_0)) - f(t_0)}{\sigma(t_0) - t_0}.$$

On the other hand, for any $t \in (t_0, \sigma(t_0))$ we have $g(t) = \sigma(t_0)$, and g is discontinuous at t_0. Therefore, $\bar{f}'_g(t_0)$ exists and

$$\bar{f}'_g(t_0) = \lim_{t \to t_0^+} \frac{\bar{f}(t) - \bar{f}(t_0)}{g(t) - g(t_0)} = \frac{f(\sigma(t_0)) - f(t_0)}{\sigma(t_0) - t_0}.$$

Case II. We assume that t_0 is right–dense. In this case f is Δ–differentiable at t_0 if and only if there exists

$$f'(t_0) = \lim_{t \to t_0,\, t \in \mathbb{T}} \frac{f(t) - f(t_0)}{t - t_0}, \tag{12}$$

and in such a case we have $f^{\Delta}(t_0) = f'(t_0)$.

Since g is continuous at t_0 and $g(t_0) = t_0$, the function $\bar{f}$ is g–differentiable at t_0 if and only if there exists

$$\lim_{t \to t_0} \frac{f(g(t)) - f(t_0)}{g(t) - t_0}, \tag{13}$$

which, if exists, is equal to $\bar{f}'_g(t_0)$.

Now we assume that (12) exists and we prove that (13) exists and has the same value. We take a sequence $\{t_n\}_{n=1}^{\infty}$ which converges to t_0 and $g(t_n) \neq t_0$ for all $n \in \mathbb{N}$ (such sequences exist: since t_0 is right–dense we have $t_0 = g(t_0) < g(t)$ for all $t > t_0$). We have $g(t_n) \to t_0$ and $g(t_n) \in \mathbb{T} \setminus \{t_0\}$ for all $n \in \mathbb{N}$. Therefore the existence of (12) guarantees the existence of

$$\lim_{n \to \infty} \frac{f(g(t_n)) - f(t_0)}{g(t_n) - t_0} = \lim_{t \to t_0,\, t \in \mathbb{T}} \frac{f(t) - f(t_0)}{t - t_0}.$$

Since the sequence was arbitrary, we deduce that (13) exists and coincides with (12).

The proof of the converse is similar: Let $\{t_n\}_{n=1}^{\infty}$ be a sequence in $\mathbb{T} \setminus \{t_0\}$ which tends to t_0. If we assume that (13) exists, then there exists

$$\lim_{n \to \infty} \frac{f(t_n) - f(t_0)}{t_n - t_0} = \lim_{n \to \infty} \frac{f(g(t_n)) - f(t_0)}{g(t_n) - t_0} = \lim_{t \to t_0} \frac{f(g(t)) - f(t_0)}{g(t) - t_0}.$$

□

Theorem 3.1 only concerns points in $\mathbb{T}$. Notice that for the remaining points we have $\mathbb{R} \setminus \mathbb{T} = C_g$, where C_g is as in (3) for the function g defined in (9), and it has g–measure zero by virtue of Proposition 2.5.

3.2 Differential equations with impulses as g–differential equations

Essential ideas for this part come from [10, 24].

Consider the differential equation with impulses at fixed times

$$\begin{cases} x' = f(t,x), \quad t > t_0,\ t \neq t_k,\ k = 1, 2, \ldots, m, \\ \Delta x(t_k) = I_k(x(t_k)), \quad t = t_k, \end{cases} \tag{14}$$

where $\Delta x(t_k) = x(t_k^+) - x(t_k)$, $t_0 < t_1 < t_2 < \cdots < t_m$, and $I_k : \mathbb{R} \longrightarrow \mathbb{R}$ are given functions for each $k \in \{1, 2, \dots, m\}$. A classical solution of (14) is a bounded function $x : [t_0, t_m + \varepsilon] \longrightarrow \mathbb{R}$, $\varepsilon > 0$, which satisfies the differential equation everywhere in $[t_0, t_m + \varepsilon] \setminus \{t_1, t_2, \dots, t_m\}$, and for each $k \in \{1, 2, \dots, m\}$, $x(t_k^-)$ exists and $x(t_k^-) = x(t_k)$, and $x(t_k^+)$ exists and $\Delta x(t_k) = I_k(x(t_k))$.

Following [10] we define a function $g : \mathbb{R} \longrightarrow \mathbb{R}$ by means of

$$g(t) = t + \sum_{k=1}^{m} \chi_{(t_k, +\infty)}(t), \tag{15}$$

where χ_I is the characteristic function of the interval I. Obviously, g is monotone nondecreasing and continuous from the left everywhere.

Now it is easy to prove by means of the definitions that $x : [t_0, t_m + \varepsilon] \longrightarrow \mathbb{R}$, $\varepsilon > 0$, is a classical solution of (14) if and only if it is a g–differentiable solution of the g–differential equation

$$x'_g(t) = \tilde{f}(t, x(t)) \quad \text{for all } t \in [t_0, t_m + \varepsilon], \tag{16}$$

where for $(t, x) \in [t_0, t_m + \varepsilon] \times \mathbb{R}$ we define

$$\tilde{f}(t, x) = \begin{cases} f(t, x), & \text{if } t \neq t_k, \\ \\ I_k(x), & \text{if } t = t_k \text{ for some } k. \end{cases}$$

Summing up, equation (16) can be an ordinary differential equation, a dynamic equation on a time scale or a differential equation with impulses, depending on the derivator g that we use.

In this section, we have considered classical solutions for simplicity. In fact, the equivalence between problems (14) and (16) is also valid in the more general context of Carathéodory solutions. The results in the following sections are the key for studying (16) in the Carathéodory sense. Specifically, a Carathéodory solution of (16) is a g–absolutely continuous function satisfying (16) g–almost everywhere; see Definition 5.1.

4 A Lebesgue–Stieltjes differentiation theorem

This section is devoted to proving the following generalization of the celebrated Lebesgue Differentiation Theorem of monotone functions.

Theorem 4.1. *If $f : [a, b] \longrightarrow \mathbb{R}$ is monotone nondecreasing, then there exists $N \subset [a, b]$ such that $\mu_g(N) = 0$ and*

$$f'_g(x) \quad \text{exists for every } x \in [a, b) \setminus N.$$

In particular, if g is continuous at b, then $f'_g(x)$ exists for g–almost all $x \in [a, b]$.

A very clear elementary proof of Lebesgue's differentiation theorem is due to Botsko [4]. All of the following material leading to the proof of Theorem 4.1 is nothing but a slight modification of Botsko's arguments. We include every detail for self–containedness and for the convenience of readers. Moreover, our generalization of Botsko Covering Lemma (Lemma 4.3) will be used again in Section 6 in the proof of a fundamental theorem of calculus for the Kurzweil–Stieltjes integral.

Lemma 4.2. *Let $f : [a, b] \longrightarrow \mathbb{R}$, let $P = \{x_0, x_1, x_2, ..., x_n\}$ be a partition of $[a, b]$, let S be a nonempty subset of $\{1, 2, 3, ..., n\}$, and let $\alpha > 0$. If $f(a) \leq f(b)$ and*

$$\frac{f(x_k) - f(x_{k-1})}{g(x_k) - g(x_{k-1})} < -\alpha$$

for each $k \in S$, then

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| > |f(b) - f(a)| + \alpha\, L,$$

where $L = \sum_{k \in S} (g(x_k) - g(x_{k-1}))$. The same result is true if $f(a) \geq f(b)$ and

$$\frac{f(x_k) - f(x_{k-1})}{g(x_k) - g(x_{k-1})} > \alpha \quad \text{for each } k \in S.$$

PROOF. Since $f(a) \leq f(b)$, we get the result from the following computations:

$$\begin{aligned} |f(b) - f(a)| = f(b) - f(a) &= \sum_{k=1}^{n} (f(x_k) - f(x_{k-1})) \\ &= \sum_{k \in S} (f(x_k) - f(x_{k-1})) + \sum_{k \notin S} (f(x_k) - f(x_{k-1})) \\ &< -\alpha \sum_{k \in S} (g(x_k) - g(x_{k-1})) + \sum_{k \notin S} (f(x_k) - f(x_{k-1})) \\ &\leq -\alpha L + \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| . \end{aligned}$$

The other case follows from the previous one with f replaced by $-f$. □

Next we generalize the Botsko Covering Lemma, and we also include in the statement the analogue to Remark 2 in [4]. For brevity, we use a nonelementary shortcut: the inner regularity of the Lebesgue–Stieltjes measure; see [21, Theorem 2.18].

Lemma 4.3. *Let D_g be as in (5), let E be a subset of (a,b) which is not of g–measure zero; i.e., there is some $\varepsilon_0 > 0$ such that $\mu_g^*(E) = \varepsilon_0$. Then*

1. *If $\mathcal{I}$ is any collection of open subintervals of $[a,b]$ that covers E, then there exists a finite disjoint subcollection $\{I_1, I_2, \dots, I_N\}$ of $\mathcal{I}$ such that*

$$\sum_{k=1}^{N} \mu_g(I_k) > \frac{\varepsilon_0}{3}.$$

2. *If P is a finite subset of $[a,b] \setminus D_g$ and $\mathcal{I}$ is any collection of open subintervals of $[a,b]$ that covers $E \setminus P$, then there exists a finite disjoint subcollection $\{I_1, I_2, \dots, I_N\}$ of $\mathcal{I}$ such that*

$$\sum_{k=1}^{N} \mu_g(I_k) > \frac{\varepsilon_0}{4}.$$

Proof. Since $V = \cup_{I \in \mathcal{I}} I$ is open and $E \subset V$, we have

$$\mu_g(V) = \mu_g^*(V) \geq \mu_g^*(E) \geq \varepsilon_0.$$

By inner regularity of μ_g, we can find a compact set $K \subset V$ such that

$$\mu_g(K) \geq 2\varepsilon_0/3.$$

By compactness, there exists a finite number of intervals $J_1, J_2, \dots, J_p$ from $\mathcal{I}$ which cover K. We may assume by discarding some of the intervals, if necessary, that no interval in $\{J_k\}_{k=1}^p$ is a subset of the union of the remaining intervals in $\{J_k\}_{k=1}^p$. Thus each J_i contains a point x_i that does not belong to $\cup_{k \neq i} J_k$ and, by renumbering the J_k's if necessary, we may assume that $x_1 < x_2 < x_3 < \cdots < x_p$. Therefore, both $\{J_1, J_3, J_5, \dots\}$ and $\{J_2, J_4, J_6, \dots\}$ are finite disjoint subcollections of $\mathcal{I}$. Clearly either $\sum_k \mu_g(J_{2k-1}) \geq (\sum_{k=1}^p \mu_g(J_k))/2$ or $\sum_k \mu_g(J_{2k}) \geq (\sum_{k=1}^p \mu_g(J_k))/2$. Thus depending on which of the two previous inequalities holds, we have found a finite disjoint subcollection $\mathcal{J}$ of $\mathcal{I}$ such that

$$\sum_{I \in \mathcal{J}} \mu_g(I) \geq \left(\sum_{k=1}^{p} \mu_g(J_k) \right) /2 \geq \mu_g(K)/2 \geq \varepsilon_0/3.$$

The second claim follows from the first one as indicated in [4, Remark 2]: we apply the first part to a covering of E obtained by adding open intervals centered at the points of P such that the sum of their g–measures is sufficiently small (which can be achieved because P contains no point of D_g). □

We are ready for proving Theorem 4.1.

Proof of Theorem 4.1. First, since f is nondecreasing

$$f'_g(x) = \frac{f(x^+) - f(x)}{g(x^+) - g(x)} \quad \text{exists for every } x \in [a,b) \cap D_g.$$

Since either $a \in D_g$ or $\mu_g(\{a\}) = 0$, it suffices to prove that f'_g exists for g–almost all $x \in (a,b) \setminus D_g$. In fact, we can even restrict our attention to a smaller set: by Remark 2.1, it suffices to prove that f'_g exists for g–almost all $x \in (a,b) \setminus (D_g \cup A)$, where A is as in (8). Moreover, if $x \in (a,b) \setminus (D_g \cup A)$, then

$$g(y) \neq g(x) \quad \text{for all } y \neq x,$$

and we can then define the Dini upper and lower g–derivatives as follows:

$$\overline{f'_g}(x) = \limsup_{y \to x} \frac{f(y) - f(x)}{g(y) - g(x)}, \quad \underline{f'_g}(x) = \liminf_{y \to x} \frac{f(y) - f(x)}{g(y) - g(x)}.$$

As a final reduction, note that the set of all those $x \in (a,b) \setminus (D_g \cup A)$ such that f is discontinuous at x is countable and has zero g–measure. Therefore, it is sufficient to show that

$$F = \{x \in (a,b) \setminus (D_g \cup A) \,:\, f \text{ is continuous at } x \text{ and } \overline{f'_g}(x) > \underline{f'_g}(x)\}$$

has zero g–measure.

Clearly F is the countable union of the sets

$$E_{r,s} = \{x \in F \,:\, \overline{f'_g}(x) > r > s > \underline{f'_g}(x)\}$$

for rational numbers r and s with $r > s > 0$. Thus we need show only that each $E_{r,s}$ has g–measure zero.

Suppose on the contrary that for some choice of r and s the set $E = E_{r,s}$ does not have g–measure zero, and let $\varepsilon_0 > 0$ be as in Lemma 4.3. If $\alpha = (r-s)/2$, $\beta = (r+s)/2$, and $h(x) = f(x) - \beta g(x)$ for all $x \in [a,b]$, then clearly α and β are positive and

$$E = \{x \in F \,:\, h \text{ is continuous at } x,\ \overline{h'_g}(x) > \alpha, \text{ and } \underline{h'_g}(x) < -\alpha\}.$$

Since $\{\sum_P |h(x_k)-h(x_{k-1})| \,:\, P$ is a partition of $[a,b]$, $P\cap D_g \subset \{a,b\}\}$ is bounded above, we can let T be the least upper bound of this set. Because α and ε_0 are both positive, there exists a partition $P=\{x_0,x_1,\ldots,x_n\}$ of $[a,b]$ such that $x_k \notin D_g$ for any $k\in\{1,2,\ldots,n-1\}$ and

$$\sum_{k=1}^{n} |h(x_k)-h(x_{k-1})| > T - \frac{\alpha\varepsilon_0}{4}. \tag{17}$$

Now let x belong to $E\setminus P$, which means that $x\in E\cap(x_{k-1},x_k)$ for some k. Since $\underline{h'_g}(x) < -\alpha$, $\overline{h'_g}(x) > \alpha$, and g and h are continuous at x, we can choose $a_x, b_x \in (a,b)\setminus D_g$ such that $a_x < x < b_x$, $(a_x,b_x)\subset(x_{k-1},x_k)$ and

$$\frac{h(b_x)-h(a_x)}{g(b_x)-g(a_x)} < -\alpha \qquad \text{or} \qquad > \alpha,$$

according to whether $h(x_{k-1}) \le h(x_k)$ or $h(x_{k-1}) > h(x_k)$. Notice that $\mu_g(a_x,b_x) = g(b_x)-g(a_x)$ because g is continuous at a_x.

Thus $\mathcal{I} = \{(a_x,b_x) \,:\, x\in E\setminus\{x_1,x_2,\ldots,x_{n-1}\}\}$ is a collection of open subintervals of (a,b) that covers $E\setminus\{x_1,x_2,\ldots,x_{n-1}\}$, and $\{x_1,x_2,\ldots,x_{n-1}\}\cap D_g=\emptyset$. By the second part in Lemma 4.3, there exists a finite disjoint subcollection $\{I_1,I_2,\ldots,I_N\}$ of $\mathcal{I}$ such that

$$\sum_{k=1}^{N} \mu_g(I_k) > \frac{\varepsilon_0}{4}. \tag{18}$$

Now let $Q=\{y_0,y_1,\ldots,y_q\}$ be the partition of $[a,b]$ determined by the points of P and the endpoints of the intervals $I_1,I_2,\ldots,I_N$. For each $[x_{k-1},x_k]$ containing at least one of the intervals in $\{I_1,I_2,\ldots,I_N\}$, we infer from Lemma 4.2 that

$$\sum_{[y_{i-1},y_i]\subset[x_{k-1},x_k]} |h(y_i)-h(y_{i-1})| > |h(x_k)-h(x_{k-1})| + \alpha\, L_k, \tag{19}$$

where the summation is taken over the closed intervals determined by Q that are contained in $[x_{k-1},x_k]$ and L_k is the sum of the g–measures of those intervals $I_1,I_2,\ldots,I_N$ that are contained in $[x_{k-1},x_k]$. Summing inequality (19) over k and using (17) and (18), we get

$$\sum_{k=1}^{q} |h(y_k)-h(y_{k-1})| > \sum_{k=1}^{n} |h(x_k)-h(x_{k-1})| + \alpha\sum_{k=1}^{N}\mu_g(I_k) > T,$$

which contradicts the definition of T.

We have thus proven that for g–almost all $x \in [a,b) \setminus D_g$ we have

$$\overline{f'_g}(x) = \underline{f'_g}(x),$$

so it remains to prove that $\underline{f'_g}(x)$ is finite for g–almost all $x \in [a,b) \setminus D_g$. Once again, we replicate Botsko's proof of [4, Theorem 2] with all the necessary modifications.

Suppose that $E = \{x \in (a,b) \setminus (D_g \cup A) \, : \, \underline{f'_g}(x) = +\infty\}$ does not have g–measure zero. Let $\varepsilon_0 > 0$ be as in Lemma 4.3 and let $M > 3(f(b)-f(a))/\varepsilon_0$. If x lies in E, then $\underline{f'_g}(x) > M$ and there exist $a_x, b_x \in (a,b) \setminus D_g$ such that $a_x < x < b_x$ and

$$\frac{f(b_x) - f(a_x)}{g(b_x) - g(a_x)} > M.$$

Thus $\mathcal{I} = \{(a_x, b_x) \, : \, x \in E\}$ covers E and by Lemma 4.3 contains a finite disjoint subcollection $\{I_1, I_2, \dots, I_N\}$ such that

$$\sum_{k=1}^{N} \mu_g(I_k) > \frac{\varepsilon_0}{3}.$$

Let $I_k = (a_k, b_k)$ for each k and recall that $\mu_g(I_k) = g(b_k) - g(a_k)$ because g is continuous at a_k. Since f is nondecreasing, we have

$$f(b) - f(a) \geq \sum_{k=1}^{N} (f(b_k) - f(a_k)) > M \sum_{k=1}^{N} (g(b_k) - g(a_k)) > f(b) - f(a),$$

a contradiction, and so the proof of Theorem 4.1 is complete. □

One of the most important consequences of Theorem 4.1 is the Lebesgue–Stieltjes integrability of the g–derivative of monotone functions. From now on we denote by $\mathcal{L}^1_g([a,b))$ the set of all functions $f : [a,b) \longrightarrow \overline{\mathbb{R}}$ which are Lebesgue–Stieltjes integrable with respect to g on $[a,b)$.

Theorem 4.4. *If $f : [a,b] \longrightarrow \mathbb{R}$ is monotone, then $f'_g \in \mathcal{L}^1_g([a,b))$.*

Proof. We assume without loss of generality that f is nondecreasing, and we write $[a,b) \cap D_g = \{\xi_j \, : \, j \in J\}$, where $J \subset \mathbb{N}$ and $\xi_j \neq \xi_i$ if $j \neq i$.

For each $n \in \mathbb{N}$ we consider the partition $\{x_{n,0}, x_{n,1}, x_{n,2}, \dots, x_{n,2^n}\}$ which divides $[a,b]$ into 2^n subintervals of equal length, i.e.

$$x_{n,k} = a + k\frac{b-a}{2^n} \quad \text{for } k \in \{0, 1, 2, \dots, 2^n\}.$$

Now we construct a simple function $f_n : [a,b) \longrightarrow \mathbb{R}$ as

$$f_n(x) = \begin{cases} f'_g(x) = \dfrac{f(x^+) - f(x)}{g(x^+) - g(x)}, & \text{if } x \in \{\xi_1, \xi_2, \dots, \xi_n\}, \\[2ex] \dfrac{f(x_{n,k}) - f(x_{n,k-1})}{g(x_{n,k}) - g(x_{n,k-1})}, & \text{if } x \in [x_{n,k-1}, x_{n,k}) \setminus D_g \text{ and} \\ & \quad g(x_{n,k}) > g(x_{n,k-1}), \\[2ex] 0, & \text{otherwise.} \end{cases}$$

We are going to prove that

$$\lim_{n\to\infty} f_n(x) = f'_g(x) \quad \text{for } g\text{–almost all } x \in [a,b).$$

Clearly, $f_n(x) \to f'_g(x)$ for all $x \in [a,b) \cap D_g$, so we have to prove that the same is true g–almost everywhere in $(a,b) \setminus D_g$ or, equivalently, g–almost everywhere in $(a,b) \setminus (D_g \cup A)$, where A is as in Remark 2.1.

Note that the points $x_{n,k}$ which do not belong to D_g form a countable g–null set, so Theorem 4.1 ensures that for g–almost all $x \in (a,b) \setminus (D_g \cup A)$ we have that $f'_g(x)$ exists and $x \neq x_{n,k}$ for all n and k. Fix one of those $x \in (a,b) \setminus (D_g \cup A)$. For every $n \in \mathbb{N}$ there exists a unique $k = k(n,x) \in \{1, 2, \dots, 2^n\}$ such that $x_{n,k-1} < x < x_{n,k}$. Since $x \notin A$, we have $g(x_{n,k-1}) < g(x) < g(x_{n,k})$, and therefore

$$f_n(x) = \frac{f(x_{n,k}) - f(x_{n,k-1})}{g(x_{n,k}) - g(x_{n,k-1})}.$$

Hence

$$\begin{aligned} \left| f'_g(x) - f_n(x) \right| &= \left| f'_g(x) - \frac{f(x_{n,k}) - f(x) + f(x) - f(x_{n,k-1})}{g(x_{n,k}) - g(x_{n,k-1})} \right| \\ &= \left| f'_g(x) \frac{g(x_{n,k}) - g(x) + g(x) - g(x_{n,k-1})}{g(x_{n,k}) - g(x_{n,k-1})} \right. \\ &\quad - \frac{f(x_{n,k}) - f(x)}{g(x_{n,k}) - g(x)} \frac{g(x_{n,k}) - g(x)}{g(x_{n,k}) - g(x_{n,k-1})} \\ &\quad \left. - \frac{f(x) - f(x_{n,k-1})}{g(x) - g(x_{n,k-1})} \frac{g(x) - g(x_{n,k-1})}{g(x_{n,k}) - g(x_{n,k-1})} \right| \\ &\le \left| f'_g(x) - \frac{f(x_{n,k}) - f(x)}{g(x_{n,k}) - g(x)} \right| + \left| f'_g(x) - \frac{f(x) - f(x_{n,k-1})}{g(x) - g(x_{n,k-1})} \right|, \end{aligned}$$

which implies that $f_n(x) \to f'_g(x)$ as $n \to \infty$. Hence $f_n(x) \to f'_g(x)$ for g–almost every $x \in [a,b)$, and, in particular, f'_g is g–measurable because each f_n is Borel–measurable.

The Fatou Lemma for integrals with respect to abstract positive measures (see [21, Theorem 1.28]) yields

$$\int_{[a,b)} |f'_g|\, d\mu_g = \int_{[a,b)} f'_g\, d\mu_g \le \liminf \int_{[a,b)} f_n\, d\mu_g,$$

and the last term is finite because for every $n \in \mathbb{N}$ we have

$$\begin{aligned}
\int_{[a,b)} f_n\, d\mu_g &= \int_{[a,b)\cap D_g} f_n\, d\mu_g + \int_{\bigcup_{k=1}^{2^n}[x_{n,k-1},x_{n,k})\setminus D_g} f_n\, d\mu_g \\
&\le \sum_{j\in J} f'_g(\xi_j)(g(\xi_j^+) - g(\xi_j)) + \sum_{k=1}^{2^n} \int_{[x_{n,k-1},x_{n,k})\setminus D_g} f_n\, d\mu_g \\
&\le \sum_{j\in J}(f(\xi_j^+) - f(\xi_j)) + \sum_{k=1}^{2^n}(f(x_{n,k}) - f(x_{n,k-1})) \\
&\le 2(f(b) - f(a)).
\end{aligned}$$

□

Corollary 4.5. *If $f : [a,b] \longrightarrow \mathbb{R}$ has bounded variation on $[a,b]$, then f'_g exists g–almost everywhere on $[a,b)$ and $f'_g \in \mathcal{L}^1_g([a,b))$.*

Proof. By the Jordan Decomposition Theorem, $f = f_1 - f_2$ with f_i monotone nondecreasing ($i = 1,2$), so the result follows from Theorem 4.4 and the elementary identity $f'_g(x) = {f_1}'_g(x) - {f_2}'_g(x)$ for g–almost all $x \in [a,b)$. □

An interesting consequence of Lemma 4.3 concerns the usual derivative of the derivator g.

Proposition 4.6. $\mu_g(\{x \in \mathbb{R} : \exists\, g'(x) = 0\}) = 0.$

Proof. It suffices to prove that for any $(a,b) \subset \mathbb{R}$ the set

$$E = \{x \in (a,b) : \exists\, g'(x) = 0\}$$

is g–null. Assume for contradiction that $\mu_g^*(E) = \varepsilon_0 > 0$. For each $x \in E$ there exist $a_x, b_x \in (a,b) \setminus D_g$ such that $a_x < x < b_x$ and

$$\frac{g(b_x) - g(a_x)}{b_x - a_x} < \frac{\varepsilon_0}{3(b-a)}.$$

Since $\{(a_x, b_x)\}_{x \in E}$ is an open cover of E, we deduce by Lemma 4.3 that there exists a finite disjoint family of these intervals, which we denote by $\{(a_k, b_k)\}_{k=1}^N$, such that

$$\sum_{k=1}^N \mu_g((a_k, b_k)) \geq \varepsilon_0/3.$$

However

$$\sum_{k=1}^N \mu_g((a_k, b_k)) = \sum_{k=1}^N (g(b_k) - g(a_k)) < \frac{\varepsilon_0}{3(b-a)} \sum_{k=1}^N (b_k - a_k) \leq \varepsilon_0/3,$$

a contradiction. □

Remark 4.1. *An alternative concise proof of Proposition 4.6 leans on Theorem 4.1: Let $f(x) = x$ for all $x \in [a,b]$. The set*

$$E = \{x \in (a,b) \,:\, \exists\, g'(x) = 0\}$$

is a subset of the points of (a,b) such that f_g' does not exist, which is g–null by virtue of Theorem 4.1.

We emphasize that Proposition 4.6 does not imply that $g'(x) > 0$ for g–almost all $x \in \mathbb{R}$. As an instance, consider the case when g is constant on $\mathbb{R}$, whose associated Lebesgue–Stieltjes measure is constantly equal to zero. Our next proposition ensures the existence of more complicated examples: we can have nondecreasing continuous functions g which are not constant and yet g' exists only in a g–null set. Examples of this type show that, in general, the g–derivative cannot be reduced g–almost everywhere to usual derivatives by means of (4).

In the following proposition, we use the notation "$\exists\, g'(x)$" to mean that $\lim_{y \to x} (g(y) - g(x))/(y - x)$ is a real number.

Proposition 4.7. *There exist nonconstant functions $g : \mathbb{R} \longrightarrow \mathbb{R}$ which are continuous, nondecreasing, and*

$$\mu_g\left(\{x \in \mathbb{R} \,:\, \exists\, g'(x)\}\right) = 0. \tag{20}$$

Proof. As proven in [26, Theorem 4.54], there exists a strictly increasing continuous function $f : \mathbb{R} \longrightarrow \mathbb{R}$ such that $f'(y) = 0$ almost everywhere in Lebesgue's sense. Let $[a,b] \subset f(\mathbb{R})$, with $a < b$, and let $g : [a,b] \to \mathbb{R}$ be an inverse of f; i.e., $f(g(x)) = x$ for all $x \in [a,b]$. Finally, we extend the definition of g to assume the value $g(a)$ everywhere on the left of a and $g(b)$ on the right of b. This construction guarantees that g is nondecreasing.

To prove (20) we note that $g' = 0$ everywhere on $(-\infty, a) \cup (b, +\infty)$, which is a g–null set because g is constant in those intervals. Therefore, the problem is reduced to showing that the set

$$E = \{x \in (a,b) \,:\, \nexists\, g'(x)\}$$

is g–null. Since g is nondecreasing, we have

$$E = \{x \in (a,b) \,:\, \exists\, g'(x) = 0\} \cup \{x \in (a,b) \,:\, \exists\, g'(x) > 0\},$$

and the first set in this union is g–null by virtue of Proposition 4.6. Now for the second set. A classical result on differentiation of inverse functions yields

$$F = \{x \in (a,b) \,:\, \exists\, g'(x) > 0\} \subset \{x \in (a,b) \,:\, \exists\, f'(g(x)) > 0\}.$$

Since g is a homeomorphism from (a,b) onto $g(a,b) = (g(a), g(b))$, we have

$$F \subset g^{-1}\left(\{y \in g(a,b) \,:\, \exists\, f'(y) > 0\}\right).$$

Now let $\varepsilon > 0$ be fixed; since $\{y \in g(a,b) \,:\, \exists\, f'(y) > 0\}$ has zero Lebesgue measure, we can find an open set V such that $\{y \in g(a,b) \,:\, \exists\, f'(y) > 0\} \subset V \subset g(a,b)$ and $m(V) < \varepsilon$, where m stands for the Lebesgue measure. We can express $V = \cup_n (\alpha_n, \beta_n)$ with pairwise disjoint intervals $(\alpha_n, \beta_n) \subset g(a,b)$. Hence $W = \cup_n g^{-1}(\alpha_n, \beta_n) = \cup_n (a_n, b_n)$ is an open set which contains F and

$$\begin{aligned}\mu_g(W) &= \sum_{n=1}^{\infty} \mu_g((a_n, b_n)) = \sum_{n=1}^{\infty} (g(b_n) - g(a_n)) \\ &= \sum_{n=1}^{\infty} (\beta_n - \alpha_n) = m(V) < \varepsilon.\end{aligned}$$

We have proven that for any $\varepsilon > 0$ we can find an open set W which contains F and such that $\mu_g(W) < \varepsilon$; hence $\mu_g(F) = 0$. This completes the proof. $\square$

5 The Fundamental Theorems of Calculus

This section is devoted to proving the Fundamental Theorems of Calculus for the Lebesgue–Stieltjes integral, labeled as Theorem 2.4 and Theorem 5.4. Both results can probably be deduced from Theorem 2.9.8 in [9] or Theorem 2.12 (1) in [18], abstract measure theoretic results proven by means of coverings. Moreover, Theorem 5.4 was essentially proven by Daniell [7]. We are grateful to the anonymous referee for having brought to our attention the monograph [13], where Stieltjes derivatives are studied and a version of Theorem 5.4 is proven in its Section IV.18. Our contributions to Theorems 2.4 and 5.4 are new proofs in modern elementary terms and a new and simpler characterization of absolute continuity with respect to a function.

For the proof of Theorem 2.4, we use Stromberg's approach to the Lebesgue–Stieltjes integral in [26, p. 283–284]. First we need the following generalization of the Fubini Theorem on almost everywhere differentiation of series. We omit its proof because it is essentially the same given in [26] for the particular case the Lebesgue measure, but using Theorem 4.1 instead of the classical Lebesgue Differentiation Theorem.

Lemma 5.1. *Let $[a,b] \subset \mathbb{R}$ and let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real–valued nondecreasing functions on $[a,b]$. If the series*

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \quad \textit{converges for all } x \in [a,b],$$

then

$$f'_g(x) = \sum_{n=1}^{\infty} (f_n)'_g(x) \quad \textit{for } g\textit{–almost all } x \in [a,b).$$

Following Stromberg [26, p. 283-284 and Theorem 6.48] we now prove Theorem 2.4. Stromberg's approach avoids the use of technical results on coverings and yields a simple proof (compare it with the proof of our Theorem 6.5, a generalization of Theorem 2.4 in the context of Kurzweil–Stieltjes integrals). In the next proof we are going to use the set $M_0(g)$ of all step functions whose points of discontinuity are each points at which g is continuous.

Proof of Theorem 2.4. We follow [26, p. 320] and we consider several cases separately:

Case 1: If $f = \chi_{(\alpha,\beta)}$, where $(\alpha,\beta) \subset (a,b)$, $\alpha,\beta \notin D_g$, then

$$F(x) = \mu_g([a,x) \cap (\alpha,\beta)) = \begin{cases} 0, & \text{if } x \le \alpha, \\ g(x) - g(\alpha), & \text{if } \alpha < x \le \beta, \\ g(\beta) - g(\alpha), & \text{if } \beta < x < b. \end{cases}$$

We compute $F'_g(x) = \chi_{(\alpha,\beta)}(x)$ for all $x \in [a,b] \setminus (A \cup \{\alpha,\beta\})$, where A is as in Remark 2.1. Hence $F'_g(x) = f(x)$ for g–almost all $x \in [a,b)$.

Case 2: If $f \in M_0(g)$, we deduce that $F'_g = f$ g–a.e. from Case 1.

Case 3: There exists a nondecreasing sequence $\{\phi_n\}_{n=1}^{\infty}$ in $M_0(g)$ such that $\lim_{n\to\infty} \phi_n(x) = f(x)$ for g–almost all $x \in [a, b)$. We define

$$\Phi_n(x) = \int_{[a,x)} \phi_n \, d\mu_g,$$

and then

$$F(x) = \lim_{n\to\infty} \Phi_n(x) = \Phi_1(x) + \sum_{k=2}^{\infty} (\Phi_k(x) - \Phi_{k-1}(x))$$

for all $x \in [a, b]$. Since each summand is a nondecreasing step function of x, we can apply Case 2 and Fubini's Lemma 5.1 to deduce that for g–almost every $x \in [a, b)$ we have

$$\begin{aligned} F'_g(x) &= {\Phi_1}'_g(x) + \sum_{k=2}^{\infty} ({\Phi_k}'_g(x) - {\Phi_{k-1}}'_g(x)) \\ &= \phi_1(x) + \sum_{k=2}^{\infty} (\phi_k(x) - \phi_{k-1}(x)) = \lim_{n\to\infty} \phi_n(x) = f(x). \end{aligned}$$

General Case. For any $f \in \mathcal{L}_g^1([a, b))$ we have $f = f_1 - f_2$, where each of the f_i's is a limit of a nondecreasing sequence of step functions in the conditions of Case 3. □

We shall prove that functions F in the conditions of Theorem 2.4 are absolutely continuous with respect to g, according to the following definition.

Definition 5.1. *A function $F : [a, b] \longrightarrow \mathbb{R}$ is absolutely continuous with respect to g (or g–absolutely continuous) if for each $\varepsilon > 0$ there is some $\delta > 0$ such that for any familiy $\{(a_n, b_n)\}_{n=1}^{m}$ of pairwise disjoint open subintervals of $[a, b]$ the inequality*

$$\sum_{n=1}^{m} (g(b_n) - g(a_n)) < \delta$$

implies

$$\sum_{n=1}^{m} |F(b_n) - F(a_n)| < \varepsilon.$$

The notion of absolute continuity of a function with respect to another function has already been considered under different forms. See, for instance, Daniell [7], whose definition involves the total variation of the associated Borel measures. To the best of the authors' knowledge, the specific form of Definition 5.1 is new.

Exactly as in the classical case when g is the identity, g–absolutely continuous functions form a vector space and, moreover, we have the following result.

Proposition 5.2. *Let f and F be as in Theorem 2.4, then F is g–absolutely continuous on $[a,b]$.*

Proof. It suffices to consider the case $f(x) \geq 0$ g–almost everywhere, as the general case is just a difference of two such functions. Let $\varepsilon > 0$ be fixed. Since $f \in \mathcal{L}^1_g([a,b))$, there exists $\delta > 0$ such that if $E \in \mathcal{M}_g$ is such that $\mu_g(E) < \delta$, then

$$\int_E f \, d\mu_g < \varepsilon.$$

In particular, if $\{(a_n, b_n)\}_{n=1}^m$ are open intervals in the conditions of the definition of g–absolute continuity for the previous value of δ, and we call $E = \cup [a_n, b_n)$, then

$$\mu_g(E) = \sum_{n=1}^m \mu_g([a_n, b_n)) = \sum_{n=1}^m (g(b_n) - g(a_n)) < \delta. \tag{21}$$

Using the definition of F and (21) we obtain

$$\begin{aligned}\sum_{n=1}^m |F(b_n) - F(a_n)| &= \sum_{n=1}^m (F(b_n) - F(a_n)) = \sum_{n=1}^m \int_{[a_n, b_n)} f \, d\mu_g \\ &= \int_E f \, d\mu_g < \varepsilon.\end{aligned}$$

□

The following proposition is very important in our next proof of the second fundamental theorem of calculus and it is also of independent interest.

Proposition 5.3. *If F is g–absolutely continuous on $[a,b]$, then it has bounded variation and it is continuous from the left at every $x \in [a,b)$. Moreover, F is continuous in $[a,b] \setminus D_g$, where D_g is the set of discontinuity points of g, and if g is constant on some $(\alpha, \beta) \subset [a,b]$, then F is constant on (α, β) as well.*

PROOF. If g is constant in $[a,b]$, then so F must be, and there is nothing to prove. For the nontrivial case we use the following simple observation:

Claim. If there exists $c>0$ such that the total variation of F on $[\alpha,\beta]$ is bounded above by c for any subinterval $[\alpha,\beta]\subset(a,b)$, then F has bounded variation on $[a,b]$.

To prove the claim, note that for each $x\in(a,b)$ we have

$$|F(x)|\le|F(x)-F((a+b)/2)|+|F((a+b)/2)|\le c+|F((a+b/2))|,$$

and therefore $|F|$ is bounded on $[a,b]$. Let $K>0$ be one of its upper bounds. If $\{x_0,x_1,\dots,x_n\}$ is an arbitrary partition of $[a,b]$, then we have

$$\sum_{k=1}^{n}|F(x_k)-F(x_{k-1})|=|F(x_1)-F(a)|+|F(b)-F(x_{n-1})| \\ +\sum_{k=2}^{n-1}|F(x_k)-F(x_{k-1})|\le 4K+c,$$

and the claim is proven.

To prove that F has bounded variation on $[a,b]$, we take $\varepsilon=1$ in the definition of g–absolutely continuous function and we get some value $\delta>0$ such that for any familiy $\{(a_n,b_n)\}_{n=1}^m$ of pairwise disjoint open subintervals of $[a,b]$ the inequality

$$\sum_{n=1}^{m}(g(b_n)-g(a_n))<\delta$$

implies

$$\sum_{n=1}^{m}|F(b_n)-F(a_n)|<1.$$

We consider now a partition $\{y_0,y_1,\dots,y_n\}$ of $[g(a),g(b)]$ such that $0<y_k-y_{k-1}<\delta$ for any $k=1,2,\dots,n$, and we define

$$I_k=g^{-1}([y_{k-1},y_k)),\ k=1,2,\dots,n.$$

Since g is nondecreasing the I_k's are intervals, but not necessarily open or closed, and some of them could be empty. Anyway, $[a,b]=\cup I_k$, and it suffices to prove that F has bounded variation on the closure of each I_k. In the nontrivial case we have $\bar{I}_k=[a',b']$ with $a'<b'$. If $[\alpha,\beta]\subset(a',b')$ and $\{x_0,x_1,\dots,x_m\}$ is a partition of $[\alpha,\beta]$, then

$$\sum_{i=1}^{m}(g(x_i)-g(x_{i-1}))=g(\beta)-g(\alpha)\le y_k-y_{k-1}<\delta,$$

which implies that

$$\sum_{i=1}^{m} |F(x_i) - F(x_{i-1})| < 1.$$

We deduce from the claim that F has bounded variation on each $\bar{I}_k$, thus proving that F has bounded variation on $[a, b]$.

The fact that F is left–continuous is a consequence of the left–continuity of g. Indeed, let $x_0 \in (a, b]$ and $\varepsilon > 0$ be fixed, and let $\delta > 0$ be given by definition of g–absolute continuity of F. Since g is left–continuous at x_0, there exists $\delta' > 0$ such that if $x \in [a, x_0)$ and $0 < x_0 - x < \delta'$, then

$$g(x_0) - g(x) < \delta,$$

and therefore $|F(x) - F(x_0)| < \varepsilon$.

In a similar way we prove that F is right–continuous at those $x \in [a, b)$ such that g is right–continuous, and therefore F is continuous everywhere in the set $[a, b] \setminus D_g$.

Finally, if g is constant on some $(\alpha, \beta) \subset [a, b]$, then for every $x \in (\alpha, \beta)$ we have $|g(x) - g((\alpha + \beta)/2))| = 0$, which implies $|F(x) - F((\alpha + \beta)/2)| = 0$. Hence F is constant on (α, β). □

It follows from the previous proposition that if $F : [a, b] \longrightarrow \mathbb{R}$ is g–absolutely continuous, then there exist two monotone nondecreasing and left–continuous functions F_i $(i = 1, 2)$ such that $F = F_1 - F_2$. Let us denote by $\mathcal{B}([a, b])$ the σ–algebra of Borel subsets of $[a, b]$, and let us denote by $\mu_i : \mathcal{B}([a, b]) \longrightarrow [0, +\infty)$ the restriction to $\mathcal{B}([a, b])$ of the Lebesgue–Stieltjes measure generated by F_i $(i = 1, 2)$. It is worth recalling at this moment that the μ_i's are outer regular, which means that for every $E \in \mathcal{B}([a, b])$ we have

$$\mu_i(E) = \inf\{\mu_i(V) \,:\, E \subset V,\, V \text{ open}\}.$$

The (restriction to $\mathcal{B}([a, b])$ of the) Lebesgue–Stieltjes measure generated by F can now be defined by means of the formula

$$\mu_F(E) = \mu_1(E) - \mu_2(E) \quad (E \in \mathcal{B}([a, b])), \tag{22}$$

and we note that μ_F is a well–defined signed measure because the μ_i's are finite positive measures. An important consequence of the definition (22) is the fact that for any $(\alpha, \beta) \subset [a, b]$ we have

$$\mu_F((\alpha, \beta)) = F(\beta) - F(\alpha^+).$$

We are in a position to prove the main result in this section.

Theorem 5.4. (Fundamental Theorem of Calculus for the Lebesgue–Stieltjes integral) *A funcion $F : [a, b] \longrightarrow \mathbb{R}$ is g–absolutely continuous on $[a, b]$ if and only if the following three conditions are fulfilled:*

1. *$F'_g(x)$ exists for g–almost all $x \in [a, b]$;*
2. *$F'_g \in \mathcal{L}^1_g([a, b))$; and*
3. *For each $x \in [a, b]$ we have*

$$F(x) = F(a) + \int_{[a,x)} F'_g(x)\, d\mu_g. \tag{23}$$

PROOF. Proposition 5.2 ensures that the three conditions are sufficient for F to be g–absolutely continuous, so we only have to prove the converse. To do it, we use a lemma and the Radon–Nikodym Theorem. For better readability, we postpone the proof of the lemma.

Lemma 5.5. *The measure μ_F is absolutely continuous with respect to μ_g; i.e., if $E \subset [a, b]$ is a Borel set and $\mu_g(E) = 0$, then $\mu_F(E) = 0$.*

We deduce from Lemma 5.5 and the Radon–Nikodym Theorem (see [21, Theorem 6.10]) that there exists a unique Borel measurable function $h : [a, b) \longrightarrow \mathbb{R}$ which is μ_g–integrable and such that

$$\mu_F(E) = \int_E h\, d\mu_g \quad \text{for any Borel set } E \subset [a, b).$$

In particular, if $E = [a, x)$, $x \in [a, b]$, then

$$F(x) - F(a) = \mu_F([a, x)) = \int_{[a,x)} h\, d\mu_g.$$

It only remains to invoke Theorem 2.4, which ensures that $F'_g(x) = h(x)$ for g–almost all $x \in [a, b)$. □

Remark 5.1. *Notice that the proof implies that F'_g is the Radon–Nikodym derivative of the measure μ_F with respect to μ_g.*

While the Radon–Nikodym Theorem is the best shortcut for proving Theorem 5.4, it is not elementary at all. The ideas in [17] can probably be adapted to get an elementary but longer proof of Theorem 5.4.

Proof of Lemma 5.5. We start proving that to each $\varepsilon > 0$, there exists $\delta > 0$ such that if V is an open set and $\mu_g(V) < \delta$, then $|\mu_F(V)| < \varepsilon$.

Let $\varepsilon > 0$ be fixed and let $\delta > 0$ be given by definition of g–absolute continuity of F with ε replaced by $\varepsilon/2$. Now we fix an open set V such that $\mu_g(V) < \delta$. Without loss of generality we assume that $V \subset (a,b)$ and $V = \cup(a_n, b_n)$ for a pairwise disjoint family of open intervals $\{(a_n, b_n)\}_{n\in\mathbb{N}}$. For every $n \in \mathbb{N}$ we take $a_n' \in (a_n, b_n)$, and for each $m \in \mathbb{N}$ we have

$$\sum_{n=1}^{m}(g(b_n) - g(a_n')) = \mu_g\left(\bigcup_{n=1}^{m}[a_n', b_n)\right) \leq \mu_g(V) < \delta,$$

which implies that

$$\sum_{n=1}^{m}|F(b_n) - F(a_n')| < \varepsilon/2.$$

Letting every a_n' tend to a_n, we obtain

$$\sum_{n=1}^{m}|F(b_n) - F(a_n^+)| \leq \varepsilon/2 \quad \text{for each fixed } m \in \mathbb{N}.$$

Hence

$$|\mu_F(V)| = \left|\sum_{n=1}^{\infty}\mu_F(a_n, b_n)\right| \leq \sum_{n=1}^{\infty}|F(b_n) - F(a_n^+)| < \varepsilon,$$

which proves the claim we state at the beginning of this proof.

We are now in a position to finish quickly the proof of our lemma. Let $E \in \mathcal{B}([a,b])$ be such that $\mu_g(E) = 0$; by outer regularity, there exist open sets $V_n \subset [a,b]$, $n \in \mathbb{N}$, such that $E \subset V_n$ for all $n \in \mathbb{N}$ and

$$\lim_{n\to\infty}\mu_g(V_n) = \mu_g(E) \quad \text{and} \quad \lim_{n\to\infty}\mu_i(V_n) = \mu_i(E),$$

where μ_i, $i = 1, 2$, are as in (22). By the first part of the proof we know that $\lim_{n\to\infty}\mu_F(V_n) = 0$ because $\lim_{n\to\infty}\mu_g(V_n) = \mu_g(E) = 0$. Hence

$$\mu_F(E) = \mu_1(E) - \mu_2(E) = \lim_{n\to\infty}\mu_F(V_n) = 0.$$

□

6 Fundamental theorems for Kurzweil integrals

The main results proven in Section 5 will be generalized in the context of Kurzweil–Stieljes integration. Interested readers are referred to [19] for recent developments on Kurzweil–Stieltjes integration.

Definition 6.1. *We say that $f : [a,b] \longrightarrow \mathbb{R}$ is Kurzweil-Stieltjes integrable on $[a,b]$ with respect to g if there exists $L \in \mathbb{R}$ such that for all $\varepsilon > 0$ there is a gauge $\delta : [a,b] \longrightarrow (0,+\infty)$ such that for every δ–fine partition $P = \{([x_{j-1},x_j],\xi_j); j = 1,\dots,n\}$ of $[a,b]$ we have*

$$\left| L - \sum_{j=1}^{n} f(\xi_j)(g(x_j) - g(x_{j-1})) \right| < \varepsilon. \tag{24}$$

The number $L \in \mathbb{R}$ is the Kurzweil-Stieltjes integral of f on $[a,b]$ with respect to g and will be denoted by $\int_a^b f(s)\, dg(s)$ or simply $\int_a^b f\, dg$.

We denote by $K_g([a,b])$ the set of all functions which are Kurzweil-Stieltjes integrable with respect to g on $[a,b]$.

Our proof of the first fundamental theorem of calculus for the Kurzweil–Stieltjes integral leans on the following generalization of the Straddle Lemma (which does not straddle points in D_g).

Lemma 6.1. *If $F : [a,b] \longrightarrow \mathbb{R}$ is g–differentiable at some $x_0 \in [a,b]$, then the following assertions are true:*

1. *If $x_0 \in D_g$, then for each $\varepsilon > 0$ there exists $\delta_\varepsilon(x_0) > 0$ such that the relations $x_0 \le v < x_0 + \delta_\varepsilon(x_0)$ and $v \in [a,b]$ imply*

$$|F(v) - F(x_0) - F'_g(x_0)(g(v) - g(x_0))| \le \varepsilon(g(v) - g(x_0)).$$

2. *If $x_0 \notin D_g$ and F is constant on every subinterval where g is constant, then for each $\varepsilon > 0$ there exists $\delta_\varepsilon(x_0) > 0$ such that the relations $x_0 - \delta_\varepsilon(x_0) < u \le x_0 \le v < x_0 + \delta_\varepsilon(x_0)$ and $u, v \in [a,b]$ imply*

$$|F(v) - F(u) - F'_g(x_0)(g(v) - g(u))| \le \varepsilon(g(v) - g(u)).$$

Proof. If $x_0 \in D_g$, then $x_0 < b$ (for if $x_0 = b \in D_g$, then F could not be g–differentiable at x_0). Now the definition of g–derivative directly implies the result for the first part of the lemma.

Assume now that F is g–differentiable at some $x_0 \notin D_g$ and that F is constant on the subintervals where so is g. First, note that $x_0 \notin C_g$ because the definition of g–derivative makes no sense at the points of C_g. Hence $g(x) > g(x_0)$ for $x > x_0$, or $g(x) < g(x_0)$ for $x < x_0$, or both.

Assume that $g(x) > g(x_0)$ for $x > x_0$ and there is some $\rho > 0$ such that $g(x) = g(x_0)$ for all $x \in [x_0 - \rho, x_0]$ (the proof in the remaning cases is similar and we omit it). The assumptions ensure that F is constant on $[x_0 - \rho, x_0] \cap [a, b]$.

Since F is g–differentiable at x_0 and g is constant on the left of x_0, then it must be $x_0 < b$. We deduce from the definition of g–derivative that for each $\varepsilon > 0$ we can find $\delta_\varepsilon(x_0) \in (0, \rho)$ such that if $0 \le v - x_0 < \delta_\varepsilon(x_0)$ and $v \in [a, b]$, then

$$|F(v) - F(x_0) - F'_g(x_0)(g(v) - g(x_0))| \le \varepsilon(g(v) - g(x_0)).$$

Therefore if $x_0 - \delta_\varepsilon(x_0) < u \le x_0 \le v < x_0 + \delta_\varepsilon(x_0)$ and $u, v \in [a, b]$, then $u \in [x_0 - \rho, x_0] \cap [a, b]$ and

$$\begin{aligned} |F(v) - F(u) - F'_g(x_0)(g(v) - g(u))| &= |F(v) - F(x_0) - F'_g(x_0)(g(v) - g(x_0))| \\ &\le \varepsilon(g(v) - g(x_0)) = \varepsilon(g(v) - g(u)). \end{aligned}$$

□

Next we prove a fundamental theorem for the Kurzweil–Stieltjes integral.

Theorem 6.2. *Let $F : [a, b] \longrightarrow \mathbb{R}$ be g–differentiable everywhere in $[a, b] \setminus C_g$, where C_g is as in (3). Assume also that F is continuous from the left at the points of $(a, b] \cap D_g$ and that F is constant on every subinterval of $[a, b]$ where g is. If $h : [a, b] \longrightarrow \mathbb{R}$ coincides with F'_g in $[a, b] \setminus C_g$, then $h \in K_g([a, b])$ and*

$$F(x) = F(a) + \int_a^x h(t)\, dg(t) \quad \textit{for all } x \in [a, b]. \tag{25}$$

Proof. We are going to prove the result with x replaced by b in (25). The remaining cases of $x \in (a, b]$ can be treated in an analogous way, and (25) is trivial for $x = a$.

Let $\varepsilon > 0$ be fixed. We have to find a gauge $\delta : [a, b] \longrightarrow (0, +\infty)$ such that for any δ–fine tagged partition $P = ([x_{j-1}, x_j], \xi_j)_{j=1}^n$ of $[a, b]$ we have

$$\left| F(b) - F(a) - \sum_{j=1}^n h(\xi_j)(g(x_j) - g(x_{j-1})) \right| \le \varepsilon. \tag{26}$$

Let $\tilde{\varepsilon} = \varepsilon/(g(b) - g(a) + 1)$.

First, for each $x \in C_g$ there is some $\rho(x) > 0$ such that g is constant on $(x - \rho(x), x + \rho(x))$, and then so is F on $(x - \rho(x), x + \rho(x)) \cap [a, b]$ by our assumptions.

Second, for any $x \in [a,b]\backslash C_g$ we have $h(x) = F_g'(x)$. By Lemma 6.1, if $x \notin D_g$, then there exists $\delta_\varepsilon(x) > 0$ such that the relations $x - \delta_\varepsilon(x) < u \leq x \leq v < x + \delta_\varepsilon(x)$ and $u, v \in [a,b]$ imply

$$|F(v) - F(u) - h(x)(g(v) - g(u))| \leq \tilde{\varepsilon}(g(v) - g(u)). \tag{27}$$

If $x \in D_g$ there exists $\delta_\varepsilon(x) > 0$ such that the relations $x \leq v < x + \delta_\varepsilon(x)$ and $v \in [a,b]$ imply

$$|F(v) - F(x) - h(x)(g(v) - g(x))| \leq \tilde{\varepsilon}(g(v) - g(x)). \tag{28}$$

We write $[a,b] \cap D_g = \{\gamma_i \,:\, i \in J\}$, where $J \subset \mathbb{N}$ and $\gamma_i \neq \gamma_j$ if $i \neq j$. Since F and g are continuous from the left at each γ_i, we can find $\delta_i > 0$ such that for all $x \in (\gamma_i - \delta_i, \gamma_i]$ we have

$$|F(\gamma_i) - F(x)| < \frac{\tilde{\varepsilon}}{2^{i+1}} \tag{29}$$

and

$$|g(\gamma_i) - g(x)| < \frac{\tilde{\varepsilon}}{2^{i+1}(|h(\gamma_i)| + 1)}. \tag{30}$$

We are now in a position to define an adequate gauge $\delta : [a,b] \longrightarrow (0, +\infty)$ as follows:

$$\delta(x) = \begin{cases} \rho(x), & \text{if } x \in C_g, \\ \min\{\delta_i, \delta_\varepsilon(x)\}, & \text{if } x = \gamma_i \text{ for some } i \in J, \\ \delta_\varepsilon(x), & \text{if } x \in [a,b]\backslash(C_g \cup D_g). \end{cases}$$

Now let $P = ([x_{j-1}, x_j], \xi_j)_{j=1}^n$ be a δ–fine tagged partition of $[a,b]$ and let us prove that (26) is satisfied. Since $F(b) - F(a) = \sum_{j=1}^n (F(x_j) - F(x_{j-1}))$, it suffices to prove that

$$\sum_{j=1}^n |F(x_j) - F(x_{j-1}) - h(\xi_j)(g(x_j) - g(x_{j-1}))| \leq \varepsilon. \tag{31}$$

The simplest terms in the previous sum are those for which the tag $\xi_j \in C_g$. In that case the definition of the gauge implies that g and F are constant on $[x_{j-1}, x_j]$, and therefore the corresponding term is equal to zero (notice that the specific values of h in C_g play no role).

Assume now that $[x_{j-1}, x_j]$ has tag $\xi_j = \gamma_i$ for some $i \in J$. We use (28), (29) and (30) to obtain

$$
\begin{aligned}
&\Big|F(x_j) - F(x_{j-1}) - h(\gamma_i)(g(x_j) - g(x_{j-1}))\Big| \\
&\quad \leq \Big|F(x_j) - F(\gamma_i) - h(\gamma_i)(g(x_j) - g(\gamma_i))\Big| \\
&\qquad + \Big|F(\gamma_i) - F(x_{j-1}) - h(\gamma_i)(g(\gamma_i) - g(x_{j-1}))\Big| \\
&\quad \leq \tilde{\varepsilon}(g(x_j) - g(\gamma_i)) + |F(\gamma_i) - F(x_{j-1})| + |h(\gamma_i)(g(\gamma_i) - g(x_{j-1}))| \\
&\quad \leq \tilde{\varepsilon}(g(x_j) - g(\gamma_i)) + \frac{\tilde{\varepsilon}}{2^{i+1}} + \frac{\tilde{\varepsilon}}{2^{i+1}} \\
&\quad \leq \tilde{\varepsilon}(g(x_j) - g(x_{j-1})) + \frac{\tilde{\varepsilon}}{2^i}.
\end{aligned}
$$

Finally, if $\xi_j \in [a, b] \setminus (C_g \cup D_g)$, then the definition of the gauge and (27) imply that

$$
\Big|F(x_j) - F(x_{j-1}) - h(\xi_j)(g(x_j) - g(x_{j-1}))\Big| \leq \tilde{\varepsilon}(g(x_j) - g(x_{j-1})).
$$

Summing up, we have proven for the left-hand term in (31) that

$$
\begin{aligned}
&\sum_{j=1}^{n} |F(x_j) - F(x_{j-1}) - h(\xi_j)(g(x_j) - g(x_{j-1}))| \\
&\qquad \leq \tilde{\varepsilon} \sum_{j=1}^{n} (g(x_j) - g(x_{j-1})) + \tilde{\varepsilon} \sum_{i \in J} \frac{1}{2^i} < \tilde{\varepsilon}(g(b) - g(a) + 1) = \varepsilon.
\end{aligned}
$$

The proof is complete. □

Since Kurzweil–Stieltjes integration extends Lebesgue–Stieltjes integration (see [23, Theorem 24.36]), we immediately obtain the following result, which complements the information given in Theorem 5.4 and generalizes a well–known result when g is the identity (see [21, Theorem 7.21]).

Corollary 6.3. *Let $F : [a, b] \longrightarrow \mathbb{R}$ satisfy the conditions in Theorem 6.2. If $F_g' \in \mathcal{L}_g^1([a, b))$, then F is g–absolutely continuous on $[a, b]$ and (25) holds in the Lebesgue–Stieltjes sense with integrals on $[a, x)$.*

Our next corollary extends to g–derivatives the well–known Barrow's rule.

Corollary 6.4. *Let $F : [a, b] \longrightarrow \mathbb{R}$ satisfy the conditions in Theorem 6.2. If $h : [a, b] \longrightarrow \mathbb{R}$ is continuous and coincides with F_g' in $[a, b] \setminus C_g$, then (25) holds in the Riemann–Stieltjes sense.*

When $C_g = \emptyset = D_g$ (for instance, when g is the identity), the assumptions in Theorem 6.2 reduce to "F is g–differentiable everywhere in $[a,b]$." Therefore Theorem 6.2 contains as a particular case the usual fundamental theorem for the Kurzweil integral.

It is worthy of remark that the assumptions in Theorem 6.2 cannot be removed in general. Indeed, (25) fails for $x = b$ if $g(x) = 0$ for all $x \in \mathbb{R}$ and $F(x) = 0$ for all $x \in [a,b)$ and $F(b) = 1$ (in this case g is constant on $[0,1]$ while F is not). Similarly easy examples show that we cannot remove the left–continuity assumption of F at discontinuity points of g: (25) fails for $x = b$ if $g(x) = x$ for $x \le (a+b)/2$ and $g(x) = x+1$ for $x > (a+b)/2$, and $F(x) = 0$ for all $x \in [a,(a+b)/2)$, $F(x) = 1$ for all $x \in [(a+b)/2, b]$ (notice that $F'_g(x) = 0$ for all $x \in [a,b]$).

A second fundamental theorem for the Kurzweil–Stieltjes integral complements Theorem 2.4. Interestingly, we base the proof on the generalization of the Botsko Covering Lemma that we prove as Lemma 4.3. Anyway it is fair to mention that the basic arguments are borrowed from [2, Theorem 5.9], where the particular case of $g(x) = x$ is treated.

Theorem 6.5. *Let $f \in K_g([a,b])$ and define for each $x \in [a,b]$*

$$F(x) = \int_a^x f(y)\, dg(y).$$

Then F is regulated on $[a,b]$ and it is continuous from the right (or from the left) where g is. Moreover, there exists a g–null set $Z \subset [a,b]$ such that

$$F'_g(x) = f(x) \quad \text{for all } x \in [a,b)\backslash Z. \tag{32}$$

In particular, if g is continuous at b, then $F'_g(x) = f(x)$ for g–almost all $x \in [a,b]$.

PROOF. The continuity properties of F are proven in [23, Theorem 24.25]. We know from [24, Theorem 1.19] that if $x \in [a,b) \cap D_g$, then

$$f(x) = \frac{F(x^+) - F(x)}{g(x^+) - g(x)},$$

hence $F'_g(x) = f(x)$; see also [10, Theorem 2.2].

Let us prove now that $F'_{g+}(x) = f(x)$ for g–almost all $x \in (a,b)\backslash D_g$, where $F'_{g+}(x)$ denotes the right-hand side g–derivative. As in the proof of Theorem 4.1, it suffices to prove that $F'_{g+}(x) = f(x)$ for g–almost all $x \in (a,b)\backslash(D_g \cup A)$, where $A = C_g \cup N_g$ is as in (2.1). Remember that for every $x \in (a,b)\backslash(D_g \cup A)$ we have $g(x) \neq g(y)$ if $x \neq y$.

Let E denote the set of points of $(a,b)\backslash(D_g \cup A)$ at which either F'_{g+} does not exist or F'_{g+} is different from f. If $x \in E$, then we can find $\alpha(x) > 0$ such that for all $s > 0$ there is some $y_{x,s} \in [a,b]$ such that $x < y_{x,s} < x + s$ and

$$\left| \frac{F(y_{x,s}) - F(x)}{g(y_{x,s}) - g(x)} - f(x) \right| > \alpha(x).$$

Since F and g are continuous from the left at $y_{x,s}$ and x, there exist $a_{x,s} \in (x - s, x)$ and $b_{x,s} \in (x, y_{x,s})$ such that

$$\left| \frac{F(b_{x,s}) - F(a_{x,s})}{g(b_{x,s}) - g(a_{x,s})} - f(x) \right| > \alpha(x).$$

We can assume without loss of generality that $a_{x,s}, b_{x,s} \notin D_g$, and therefore

$$\begin{aligned} \Big| F(b_{x,s}) - F(a_{x,s}) - f(x)(g(b_{x,s}) - g(a_{x,s})) \Big| &> \alpha(x)(g(b_{x,s}) - g(a_{x,s})) \\ &= \alpha(x)\mu_g((a_{x,s}, b_{x,s})). \end{aligned} \tag{33}$$

For each $n \in \mathbb{N}$ we define $E_n = \{x \in E : \alpha(x) \geq 1/n\}$, and we note that $E = \cup E_n$. So it suffices to prove that $\mu_g(E_n) = 0$ for each $n \in \mathbb{N}$. Reasoning by contradiction, assume that for some fixed $n \in \mathbb{N}$ we have $\mu_g^*(E_n) = \varepsilon_0 > 0$.

Since $f \in K_g([a,b])$ we can find a gauge δ_0 in $[a,b]$ such that if $P = ([x_{j-1}, x_j], \xi_j)_{j=1}^m$ is a δ_0–fine partition of $[a,b]$, then

$$\left| \sum_{j=1}^m f(\xi_j)(g(x_j) - g(x_{j-1})) - \int_a^b f \, dg \right| \leq \frac{\varepsilon_0}{7n}. \tag{34}$$

The family of open intervals $\mathcal{F} = \{(a_{x,s}, b_{x,s}) : x \in E_n \ , s = \delta_0(x)\}$ covers E_n. We deduce from Lemma 4.3 that there are intervals from $\mathcal{F}$, $I_1 = (a_1, b_1)$, $I_2 = (a_2, b_2), \ldots, I_p = (a_p, b_p)$ which are pairwise disjoint and satisfy

$$\sum_{j=1}^p \mu_g(I_j) \geq \varepsilon_0/3. \tag{35}$$

For each I_j we denote by x_j the point $x \in E_n$ for which (33) is satisfied with $a_{x,s} = a_j$ and $b_{x,s} = a_j$, and $(a_j, b_j) \subset (x_j - \delta_0(x), x_j + \delta_0(x))$. Since

$([a_j, b_j], x_j)_{j=1}^p$ is a subpartition of a δ_0–fine partition of $[a, b]$ for which (34) holds, we deduce from the Saks–Henstock Lemma that

$$\sum_{j=1}^{p} \left| f(x_j)(g(b_j) - g(a_j)) - \int_{a_j}^{b_j} f \, dg \right| \le 2\frac{\varepsilon_0}{7n}; \tag{36}$$

see [23, Theorem 24.23 (ii)]. On the other hand, the definition of F and (33) ensure that

$$\begin{aligned} \sum_{j=1}^{p} \left| f(x_j)(g(b_j) - g(a_j)) - \int_{a_j}^{b_j} f \, dg \right| &\ge \sum_{j=1}^{p} \alpha(x_j)(g(b_j) - g(a_j)) \\ &\ge \frac{1}{n} \sum_{j=1}^{p} \mu_g(a_j, b_j). \end{aligned} \tag{37}$$

Now (37) and (36) yield

$$\sum_{j=1}^{p} \mu_g(I_j) \le 2\varepsilon_0/7 < \varepsilon_0/3,$$

which contradicts (35). We have thus proven that $\mu_g(E) = 0$, and therefore $F'_{g_+}(x) = f(x)$ for g–almost all $x \in [a, b)$.

A similar argument shows that an analogous result is true for F'_{g_-}, the left-hand side g–derivative of F, and then the proof is complete. □

7 Acknowledgements

The authors are very grateful to M. Tvrdý and J. Maly for reading a first draft of this paper and sending us many bibliographical references on differentiation with respect to functions. The authors also wish to express their gratitude to the anonymous referee for many fruitful remarks.

References

[1] V. Aversa and D. Preiss, *Lusin's theorem for derivatives with respect to a continuous function*, Proc. Amer. Math. Soc., **127(11)** (1999), 3229–3235.

[2] R. G. Bartle, *A Modern Theory of Integration*, Grad. Studies in Math., **32**, Amer. Math. Soc., 2001.

[3] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales: an Introduction with Applications*, Birkhäuser, Boston, 2001.

[4] M. W. Botsko, *An elementary proof of Lebesgue's differentiation theorem*, Amer. Math. Monthly, **110** (2003), 834–838.

[5] F. E. Burk, *A garden of integrals*, Math. Assoc. Amer., 2007.

[6] R. Cacciopoli, *Sul lemma fondamentale del calcolo integrale,* Atti. Mem. Accad. Sci. Padova, **50** (1934), 93–98.

[7] P. J. Daniell, *Differentiation with respect to a function of limited variation*, Trans. Amer. Math. Soc., **19(4)** (1918), 353–362.

[8] P. J. Daniell, *Stieltjes derivatives*, Bull. Amer. Math. Soc., **26(10)** (1920), 444–448.

[9] H. Federer, *Geometric Measure Theory*, Springer, New York, 1969.

[10] M. Federson, J. G. Mesquita and A. Slavík, *Measure functional differential equations and functional dynamic equations on time scales*, J. Diff. Eq., **252(6)** (2012), 3816–3847.

[11] M. Federson, J. G. Mesquita and A. Slavík, *Basic results for functional differential and dynamic equations involving impulses*, Math. Nachr., **286(2-3)** (2013), 181–204.

[12] W. Feller, *On differential operators and boundary conditions*, Comm. Pure Appl. Math., **8** (1955), 203–216.

[13] K. M. Garg, *Relativization of some aspects of the theory of functions of bounded variation*, Dissertationes Math. (Rozprawy Mat.), **320**, 1992.

[14] M. Gradinaru, *On the derivative with respect to a function with applications to Riemann–Stieltjes integral*, Collection: Seminar on Mathematical Analysis (Cluj-Napoca, 1989-1990), 21-28, "Babes-Bolyai" Univ., Cluj-Napoca, 1990.

[15] H. Lebesgue, *Leons sur l'intégration et la recherche des fonctions primitives*, Paris, 1928.

[16] J. Liberman, *Théorème de Denjoy sur la dérivée d'une fonction arbitraire par rapport á une fonction continue*, Rec. Math. [Mat. Sbornik] N. S. **9** (1941), 221–236.

[17] R. López Pouso, A simple proof of the fundamental theorem of calculus for the Lebesgue integral, Studia Univ. Babes-Bolyai Mathematica, **58(2)** (2013), 139–145.

[18] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability*, Cambridge Univ. Press, 1995.

[19] G. Monteiro and M. Tvrdý, *On Kurzweil-Stieltjes integral in a Banach space*, Math. Bohem. **137(4)** (2012), 365–381.

[20] J. Petrovski, *Sur l'unicité de la fonction primitive par rapport á une fonction continue arbitraire*, Rec. Math. Soc. Math. Moscou **41** (1934), 48–58.

[21] W. Rudin, *Real and Complex Analysis*, Third Edition, McGraw-Hill, 1987.

[22] S. Saks, *Theory of the Integral*, Second Edition, Hafner Publishing Company, 1937.

[23] E. Schechter, *Handbook of Analysis and its Foundations*, Academic Press, 1996.

[24] Š. Schwabik, *Generalized Ordinary Differential Equations*, Series in Real Analysis, **5**, World Scientific, 1992.

[25] A. Slavík,*Dynamic equations on time scales and generalized ordinary differential equations*, J. Math. Anal. Appl., **385(1)** (2012), 534–550.

[26] K. R. Stromberg, *Introduction to Classical Real Analysis*, Wadsworth, Inc., 1981.

[27] W. H. Young, *On integrals and derivatives with respect to a function*, Proc. London Math. Soc., (1917) s2-15 (1): 35-63. doi: 10.1112/plms/s2-15.1.35.