KURZWEIL-STIELJES INTEGRAL

AND ITS APPLICATIONS

Giselle Antunes Monteiro,
Antonín Slavík
and
Milan Tvrdý

„Il semble donc que les sommes de Riemann-Stieltjes aient encore un bel avenir devant elles en calcul intégral, et qu'elles pourront réserver encore, dans les mains d'habiles analystes, d'intéressantes surprises.“

Jean Mawhin

Giselle Antunes Monteiro (email: monteiro@mat.savba.sk)
Mathematical Institute, Slovak Academy of Sciences, Bratislava, Slovak Republic.

Supported by the SASPRO Programme, Marie Curie Actions (European Union's 7th Framework Programme under REA grant agreement No. 609427), the Slovak Academy of Sciences and by the Academic Human Resource Program of the Czech Academy of Sciences.

Antonín Slavík (email: slavik@karlin.mff.cuni.cz)
Charles University in Prague, Faculty of Mathematics and Physics, Czech Republic.

Milan Tvrdý (email: tvrdy@math.cas.cz)
Institute of Mathematics, Czech Academy of Sciences, Praha, Czech Republic.

Supported by by the grant No. 14-06958S of the Grant Agency of the Czech Republic and by RVO: 67985840.

This is a preliminary version of the monograph in preparation for World Scientific Publishing Co.

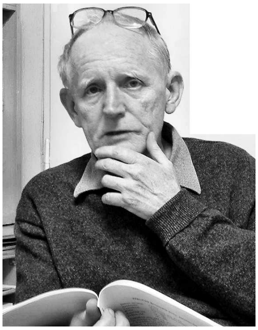
25.1. 2017

Preface

This text is a very preliminary version of the monograph we would like to complete in a near future. Actually, it is a continuation of Štefan Schwabik's monograph [125] written in Czech and dedicated to the theory of integral over one-dimensional intervals. In this textbook the author succeeded to explain not only the classic concepts of the Newton and Riemann integrals but also the McShane integral and above all, the Kurzweil integral, a non-absolutely convergent integral defined in a constructive way. This definition belongs to Jaroslav Kurzweil


Thomas Joannes Stieltjes


Jaroslav Kurzweil


Štefan Schwabik

and was introduced first in the seminal paper published in 1957 in Czechoslovak Mathematical Journal (see [78]). The new integral, which is today called Kurzweil integral in the world's mathematical literature, or rather Kurzweil-Henstock integral (Ralph Henstock, a specialist in the theory of the integral from the United Kingdom, published the definition of an analogous integral in 1960 independently of J. Kurzweil), has since then turned out to be very inspiring not only for the theory of integral (since it includes classic and well-known concepts of Riemann and Newton integrals including their improper modifications and, even, harder to manage integrals of Lebesgue and Perron) but also for the theory of differential and integral equations. From the methodical point of view the emphasis put on the Kurzweil integral made it possible to Š. Schwabik to focus on non-absolutely convergent integrals which were considered to be more difficult to explain in older methodology of integral theory. Kurzweil's concept of integral is in finite dimensional setting equivalent to Perron integral which is non-absolutely convergent. His definition seemingly and almost mechanically „copies" Riemann's definition which is with its illustrative nature and strong geometric interpretation the most reasonable one for a student. Nonetheless, just the comparison with Riemann's definition shows how ingenious is its unobtrusive but at the same time very effective Kurzweil modification. A great advantage of Kurzweil's integral is also the fact that it does not need generalization for improper integrals – that the Hake theorem holds true here (i.e. the theorem concerning the limiting process with respect to the upper and/or lower bounds of the integration domain).

In the integrals of Riemann, Newton, Lebesgue, Perron, Kurzweil a given function is integrated with respect to the identity function. Some physical problems have exacted the extension of concept of integral to integrals in which the given function is integrated with respect to a function which does not have to be the identity in general. The first time when such integral appeared was in a famous Stieltjes' treatise [136] from, dedicated to the connection between the convergence of chain fractions and the problem how to describe the distribution of matter on a solid line segment when all moments of this line segment of natural orders are known.

The integrals of this kind have been since then called *Stieltjes integrals* and integrals of a function f (*integrand*) with respect to function g (*integrator*) over interval $[a,b]$ are since then denoted by $\int_a^b f \,\mathrm{d}\, g$. To various modifications of the definition, which with time arose, the names of the authors of these modifications are then usually added. Soon there were integrals of: Riemann-Stieltjes, Perron-Stieltjes or Lebesgue-Stieltjes. Another significant impulse which turned attention to the Stieltjes integral was the fundamental Riesz's result from 1909 (see [112]) stating that every continuous linear functional in the space of continuous functions can be expressed using the Stieltjes integral. Soon, in 1910, H. Lebesgue (see [86]) proved that for a continuous function f and a function g of bounded variation Stieltjes integral can be, using suitable substitution, expressed as the Lebesgue integral in the form of

$$\int_a^{v(b)} f(w(t))\, h(t) \,\mathrm{d}\, t,$$

where $v(x)$ is the variation of the function g over the interval $[a,x]$, w is the generalized inverse function of v, $w(t) = \inf\{s \in [a,b] : v(s) = t\}$ for $t \in [a,b]$, and

$$h(t) = \frac{\mathrm{d}\, g(w(t))}{\mathrm{d}\, t} \quad \text{for a.e. } t \in [a, v(b)].$$

In this way, H. Lebesgue arrived at the concept of Lebesgue-Stieltjes integral of a function f with respect to g. Few years after Riesz's result, in 1912, Stieltjes integral appeared also in the monograph [109] by O. Perron. During the next roughly two decades the Stieltjes integral and its modifications were the subject of investigation of many outstanding persons of the theory of functions: W. H. Young ([157]), C. J. de la Vallée Poussin ([149]), E. B. Van Vleck ([150]), T. H. Hildebrandt ([52]), L. C. Young ([155] and [156]), A. J. Ward ([151]) and others. In 1933 S. Saks dedicated a whole chapter in his famous monograph [116] to the Lebesgue-Stieltjes integral and to functions of bounded variation. Up to now the integrals of Stieltjes kind have found wide use in many fields: e.g. in the theory of curve integrals, the theory of probability, the theory of hysteresis, the theory of functional-differential, generalized differential equations etc. A range of monographs are dedicated to the history of the theory of integral. However, we are not aware if any one of them is dedicated to the history of the Stieltjes integral more thoroughly. There exists a superb piece of work in Czech „*Malý průvodce historií integrálu*“ (in English: Little Guide Through the History of Integral) written by Š. Schwabik and P. Šarmanová. Unfortunately, there was not space enough in it to include also the Stieltjes integral. For the French speakers let us mention at least a brief historical essay by [94] J. Mawhin.

Given the limited assigned extent of his monograph, Štefan Schwabik could not include the natural generalization of Kurzweil's concept of integral to Stieltjes integrals there although at that time we already had „Kurzweil's theory“ of Stieltjes integral in our joint works (cf. e.g. [131]) dedicated to generalized differential equations processed and prepared to a considerable extent. It is our ambition to continue in his work and to complete his monograph with the theory of Stieltjes integral with emphasis on Kurzweil's definition and some of its applications.

Our book is divided into 8 chapters. In the introductory chapter there are briefly described two of many motivations to study the Stieltjes integral: the problems of moments and curve integrals. The next three chapters are preparatory and provide an overview of properties of the categories of functions with which this book works most often: functions of bounded variation, absolutely continuous functions and regulated functions. Chapter 5 provides a survey of the basic properties of the classical notions of the Riemann-Stieltjes integral. The core of the whole book is then Chapter 6 dedicated to the definition of the Stieltjes integral in Kurzweil's sense. There, the advantages of this definition are demonstrated: the width of the class of functions integrable in this sense, the broad range of properties of this integral, in particular, the validity of very general convergence theorems including Hake's theorem. The final two chapters describe some selected applications in functional analysis and in the theory of generalized differential equations.

From the bibliography dealing with related topics, we can recommend the monograph [55] by T. H. Hildebrandt and also the unobtrusive but modernly approached monograph [95] by R. M. McLeod from 1981 including even the Kurzweil-Stieltjes integral. Other stimuli can be found in monographs [29], [70], [117], [115] or lecture notes [92]. Two demanding monographs [83] and [84] by J. Kurzweil from 2000 and 2002 dedicated to topological problems related to integrating do not directly concern the Stieltjes integration. Integrals and generalized differential equations studied in Kurzweil's latest monograph [85] cover both the Kurzweil-Stieltjes integral and the linear generalized equations which we discuss in Chapters 6 and 8 of this book. An outstanding supplement of this publication will be, besides already mentioned Schwabik's monograph [125], also his other monograph [122] dedicated especially to generalized differential equations.

Contents

Conventions and notation

(i) $\mathbb{N}$ is the set of natural numbers (excluding zero), $\mathbb{Q}$ is the set of rational numbers, and $\mathbb{R}$ is the set of real numbers. $\mathbb{R}^m$ is the m-dimensional Euclidean space consisting of real m-vectors (m-tuples of real numbers). If $x \in \mathbb{R}^m$, then its i-th component is denoted by x_i. We write $x = \big(x_i\big)_{i=1}^m$ or, unless a misunderstanding may occur, $x = \big(x_i\big)$. For a Banach space X, the norm of its element x is denoted by $\|x\|_X$. If $X = \mathbb{R}^m$ for some $m \in \mathbb{N}$ and $x = \big(x_i\big) \in \mathbb{R}^m$, we write $|x|$ instead of $\|x\|_{\mathbb{R}^m}$ and define $|x| = \sum_{i=1}^m |x_i|$.

(ii) $\{x \in A : B(x)\}$ stands for the set of all elements x of the set A which satisfy the condition $B(x)$.

For given sets P, Q, the symbol $P \setminus Q$ represents the set

$$P \setminus Q = \{x \in P : x \notin Q\}.$$

As usual, $P \subset Q$ means that P is a subset of Q (every element of set P is also an element of set Q). Unless it may cause a misunderstanding, we write $\{x_n\}$ instead of $\{x_n \in \mathbb{R} : n \in \mathbb{N}\}$ or $\{x_n \in \mathbb{R} : n = 1, \dots, m\}$. A sequence $\{x_n\}$ is called *non-repeating* if $x_k \neq x_n$ whenever $k \neq n$.

(iii) For a given $a \in \mathbb{R}$, we set $a^+ = \max\{a, 0\}$ and $a^- = \max\{-a, 0\}$. (Let us recall that $a^+ + a^- = |a|$ and $a^+ - a^- = a$ for every $a \in \mathbb{R}$.) Furthermore,

$$\operatorname{sgn}(a) = \begin{cases} 1 & \text{if } a > 0, \\ -1 & \text{if } a < 0, \\ 0 & \text{if } a = 0. \end{cases}$$

(iv) If $-\infty < a < b < \infty$, then $[a, b]$ is the *closed* interval $\{t \in \mathbb{R} : a \le t \le b\}$ and (a, b) is the *open* interval $\{t \in \mathbb{R} : a < t < b\}$. The corresponding *half-closed*, or *half-open* intervals are denoted by $[a, b)$ and $(a, b]$. In all cases, the points a, b are called the endpoints of the interval. If $a = b \in \mathbb{R}$, we say that the interval $[a, b]$ *degenerates* to a one-point set, and we write $[a, b] = [a]$. If I is an interval (closed or open or half-open) with endpoints $a, b,$ then $|I| = |b - a|$ stands for its length. Of course, $|[a]| = 0$.

(v) A finite set of points $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \dots, \alpha_m\}$ of the interval $[a, b]$ is called a *division of the interval,* $[a, b]$ if $a = \alpha_0 < \alpha_1 < \cdots < \alpha_m = b$. The set of all divisions of the interval $[a, b]$ is denoted by $\mathscr{D}[a, b]$.

If $\boldsymbol{\alpha} \in \mathscr{D}[a, b]$, then, unless otherwise stated, its elements will be denoted by α_j, $|\boldsymbol{\alpha}|$ is the length of the largest subinterval $[\alpha_{j-1}, \alpha_j]$, and $\nu(\boldsymbol{\alpha})$ is the number of subintervals, i.e.,

$$\alpha_{\nu(\boldsymbol{\alpha})} = b \quad \text{and} \quad |\boldsymbol{\alpha}| = \max_{j=1,\dots,\nu(\boldsymbol{\alpha})} (\alpha_j - \alpha_{j-1}) \quad \text{for } \boldsymbol{\alpha} \in \mathscr{D}[a,b].$$

We will write also $\boldsymbol{\alpha} = \{\alpha_j\}$. If $\boldsymbol{\alpha}' \supset \boldsymbol{\alpha}$, then we say that $\boldsymbol{\alpha}'$ is a *refinement* of $\boldsymbol{\alpha}$.

(vi) For a given set $M \subset \mathbb{R}$ the symbol χ_M denotes the characteristic function of M, i.e.,

$$\chi_M(t) = \begin{cases} 1 & \text{if } t \in M, \\ 0 & \text{if } t \notin M. \end{cases}$$

(vii) The supremum (or the infimum) of a set $M \subset \mathbb{R}$ is denoted by $\sup M$ (or $\inf M$). If $m = \sup M$ is an element of M (or $m = \inf M$ is an element of M), i.e., if m is a maximum (or a minimum) of M, we write $m = \max M$ (or $m = \min M$). If M is the set of all values $F(x)$ of a function F over the domain B (i.e., if $M = \{F(x) : x \in B\}$), we write $\sup_{x \in B} F(x)$, and similarly for the infimum, maximum or minimum, respectively.

(viii) Let X be a Banach space. The notation $f : [a,b] \to X$ means that f is a function from the interval $[a,b]$ into X. In such a case we say that f is a *vector-valued function*. If $X = \mathbb{R}$, then f is said to be a real-valued function. For functions $f : [a,b] \to X$ and $g : [a,b] \to X$ and a real number λ, we define the functions $f + g$ and λf by

$$(f+g)(x) = f(x) + g(x) \quad \text{for } x \in [a,b]$$

and

$$(\lambda f)(x) = \lambda f(x) \quad \text{for } x \in [a,b].$$

For a given function $f : [a,b] \to X$, we set $\|f\|_\infty = \sup_{t \in [a,b]} |f(t)|$. (Of course, if f is unbounded on $[a,b]$, then $\|f\|_\infty = \infty$.)

(ix) If $\{x_n\}$ is an infinite sequence of real numbers which has a limit

$$\lim_{n \to \infty} x_n = x \in \mathbb{R} \cup \{-\infty\} \cup \{\infty\},$$

we will write also shortly $x_n \to x$. Given a sequence of functions $\{f_n\}$ defined on $[a,b]$, the symbol $f_n \to f$ stands for the pointwise convergence, i.e., it means that $f_n(x) \to f(x)$ for each $x \in [a,b]$. If $\{f_n\}$ converges uniformly to f on $[a,b]$, i.e. $\lim_{n \to \infty} \|f_n - f\|_\infty = 0$, we write also $f_n \rightrightarrows f$ on $[a,b]$.

(x) If $f:[a,b]\to\mathbb{R},\ t\in[a,b),\ s\in(a,b]$ and the one-sided limits

$$\lim_{\tau\to t+} f(\tau) \quad \text{and} \quad \lim_{\tau\to s-} f(\tau)$$

exist and are finite, we denote

$$f(t+)=\lim_{\tau\to t+} f(\tau), \quad f(s-)=\lim_{\tau\to s-} f(\tau),$$

$$\Delta^+ f(t)=f(t+)-f(t), \quad \Delta^- f(s)=f(s)-f(s-),$$

$$\Delta\, f(t)=f(t+)-f(t-) \quad \text{for } t\in(a,b).$$

Customarily, for function defined on the interval $[a,b]$, the following convention is used:

$$f(a-)=f(a), \quad f(b+)=f(b), \quad \Delta^- f(a)=\Delta^+ f(b)=0.$$

(xi) $\mathrm{C}([a,b])$ is the space of all continuous real functions on the interval $[a,b]$, with a norm defined by

$$\|f\|=\max_{x\in[a,b]} |f(x)| \quad \text{for } f\in\mathrm{C}([a,b]).$$

$\mathrm{L}^1([a,b])$ is the space of all real functions that are Lebesgue integrable on the interval $[a,b]$, with the convention

$$f=g\in\mathrm{L}^1([a,b]) \iff f(x)=g(x) \text{ for almost all } x\in[a,b],$$

and with the norm defined by

$$\|f\|_1=\int_a^b |f(x)|\,\mathrm{d}x \quad \text{for } f\in\mathrm{L}^1[a,b].$$

The space of all continuous vector-valued functions from $[a,b]$ to a Banach space X is denoted by $\mathrm{C}([a,b],X)$. Symbols such as $\mathrm{L}^1([a,b],X)$ corresponding to other function spaces have a similar meaning.

(xii) If M is a subset of a Banach space X, then $\mathrm{cl}(M)$ stands for the closure of M in X and $\mathrm{Lin}(M)$ denotes the linear span of M, i.e., the set of all elements $x\in M$ of the form

$$x=\sum_{j=1}^{m} c_j\, x_j,$$

where $m\in\mathbb{N},\ c_1,c_2,\dots,c_m\in\mathbb{R}$ and $x_1,x_2,\dots,x_m\in M.$

(xiii) The set of all continuous linear mappings from a Banach space X to a Banach space Y is denoted by $\mathscr{L}(X, Y)$. It is known that $\mathscr{L}(X,Y)$ is a Banach space when equipped with the operator norm

$$\|A\|_{\mathscr{L}(X,Y)} = \sup_{\substack{x\in X \\ |x\|_X\le 1}} \|A\,x\|_Y \quad \text{for } A\in\mathscr{L}(X,Y).$$

If $Y = X$ or $Y = \mathbb{R}$, we write simply $\mathscr{L}(X) = \mathscr{L}(X,X)$ or $X^* = \mathscr{L}(X,\mathbb{R})$, respectively. In particular, $\mathscr{L}(\mathbb{R}^m,\mathbb{R}^n)$ is the space of $m\times n$ real matrices and $\mathscr{L}(\mathbb{R}^m,\mathbb{R})$ is the space of column m-vectors, which coincides with the space $\mathbb{R}^m$. indexsymbSPACES!$\mathscr{L}(X, Y)$, $\mathscr{L}(\mathbb{R}^m,\mathbb{R}^n)$

(xiv) Given a matrix $A\in\mathscr{L}(\mathbb{R}^m,\mathbb{R}^n)$, the element in the i-th line and j-th column is denoted by $a_{i,j}$. We write

$$A = \big(a_{i,j}\big)_{\substack{i=1,\dots,m \\ j=1,\dots,n}}.$$

For every $n\in\mathbb{N}$, the symbol I stands for the unit $n\times n$-matrix, i.e.,

$$I = \big(e_{i,j}\big)_{\substack{i=1,\dots,n, \\ j=1,\dots,n}} \quad \text{where } e_{i,j} = \begin{cases} 1 & \text{if } i=j, \\ 0 & \text{if } i\ne j. \end{cases}$$

The norm in $\mathscr{L}(\mathbb{R}^m,\mathbb{R}^n)$ is defined by

$$|A| = \max_{j=1,\dots,m} \sum_{i=1}^{n} |a_{i,j}| \quad \text{for } A = \big(a_{i,j}\big)_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in\mathscr{L}(\mathbb{R}^m,\mathbb{R}^n).$$

In particular, for $x\in\mathbb{R}^m = \mathscr{L}(\mathbb{R}^m,\mathbb{R})$, we have $|x| = \sum_{i=1}^m |x_i|$, which agrees with the definition in (i). Furthermore, for $A\in\mathscr{L}(\mathbb{R}^m,\mathbb{R}^n)$, we have

$$|A| = \sup\big\{|A\,x| : |x|\le 1\big\},$$

i.e., the norm of a matrix coincides with the operator norm on $\mathscr{L}(\mathbb{R}^m,\mathbb{R}^n)$ (provided that the norm introduced in (i) is used on $\mathbb{R}^n$).

(xv) To a certain extent, standard logic symbols are used. For example,

$$\forall\,\varepsilon>0\ \exists\,\delta>0 : (A(\delta)\wedge B(\delta)) \implies C(\varepsilon)$$

means

“for every $\varepsilon>0$ there exists a $\delta>0$ such that if both $A(\delta)$ and $B(\delta)$ are true, then $C(\varepsilon)$ holds, as well.”

Chapter 1

Introduction

1.1 Areas of planar regions and moments

It is well known that the value of the Riemann integral

$$\int_a^b f(x)\,\mathrm{d}x$$

of a nonnegative continuous function $f:[a,b]\to\mathbb{R}$ equals the area of the plane region M bounded by the graph of f and by the lines $y=0$, $x=a$, and $x=b$. This conclusion is justified by the following consideration:

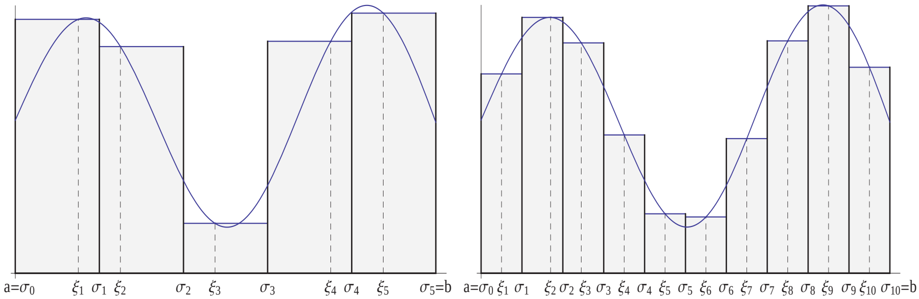
Let $\alpha_0, \alpha_1, \ldots, \alpha_m$ be points of the interval $[a,b]$ such that

$$a=\alpha_0<\alpha_1<\cdots<\alpha_m=b,$$

i.e., the set $\boldsymbol{\alpha}=\{\alpha_0,\alpha_1,\ldots,\alpha_m\}$ is a division of the interval $[a,b]$. Moreover, for every $j\in\{1,\ldots,m\}$ let us choose a point $\xi_j\in[\alpha_{j-1},\alpha_j]$, the so called tag of the interval $[\alpha_{j-1},\alpha_j]$, and denote by $\boldsymbol{\xi}=\{\xi_1,\ldots,\xi_m\}$ the corresponding set of tags. The area of the region M can be approximated by the sum of the areas of the rectangles created above the line segments $[\alpha_{j-1},\alpha_j]$ with the height $f(\xi_j)$, i.e., by

$$S(\boldsymbol{\alpha},\boldsymbol{\xi})=\sum_{j=1}^{m} f(\xi_j)\,(\alpha_j-\alpha_{j-1}). \tag{1.1.1}$$

As the following pictures indicate, finer divisions of the interval $[a,b]$ lead to



Riemann integral sums

higher accuracy of the approximation that we get. It can be expected that with a suitably defined limiting process based on refining the divisions of the interval $[a,b]$, the sums $S(\boldsymbol{\alpha},\boldsymbol{\xi})$ approach (independently of the choice of the tags $\boldsymbol{\xi}$) a certain real number $S(M)$, which is equal to the area of the region M. For a moment, let us settle with the intuitive idea about such a limiting process. We will describe it more accurately later. The result is the concept of the Riemann integral of a function f over an interval $[a,b]$ (that is, "from a to b"), which is denoted by $\int_a^b f(x)\,\mathrm{d}x$ and satisfies

$$S(M) = \int_a^b f(x)\,\mathrm{d}x.$$

A similar problem is to determine the *moment* (also known as the static moment) of a planar or spatial region. Let us restrict ourselves to a bounded line segment $[a,b]$ lying on the real axis $\mathbb{R}$. The moment of a point mass $x \in [a,b]$ of mass μ with respect to the origin is given by $|x|\mu$. If the mass of the line segment is concentrated at a finite number of points $x_i \in [a,b]$, $i = 1,\ldots,n$, while the mass of the point x_i is equal to μ_i, then the moment of the line segment $[a,b]$ with respect to the origin is equal to the sum $\sum_{i=1}^m |x_i|\,\mu_i$. In the general case when the mass of the line segment is not concentrated at a finite number of points, we can proceed as follows:
Let a division $\boldsymbol{\alpha} = \{\alpha_0,\alpha_1,\ldots,\alpha_m\}$ of the interval $[a,b]$ be given and, for $j \in \{1,\ldots,m\}$, let ξ_j be the tag of the interval $I_j = [\alpha_{j-1},\alpha_j]$. Further, for each $x \in [a,b]$, let $\mu(x)$ be the mass of the line segment $[a,x]$. Then, for every j, the mass of the subinterval I_j is given by $\mu(\alpha_j) - \mu(\alpha_{j-1})$. Assuming that the mass of each subinterval I_j is concentrated at the tag ξ_j, the moment of the line segment I_j with respect to the origin is approximately equal to

$$|\xi_j|\,(\mu(\alpha_j) - \mu(\alpha_{j-1})),$$

and the moment of the whole line segment $[a,b]$ can be approximated by the sum

$$S(\boldsymbol{\alpha},\boldsymbol{\xi}) = \sum_{j=1}^m |\xi_j|\,(\mu(\alpha_j) - \mu(\alpha_{j-1})). \tag{1.1.2}$$

Again, we can expect that $S(\boldsymbol{\alpha},\boldsymbol{\xi})$ will approach the actual value of the moment if the division $\boldsymbol{\alpha}$ becomes finer, i.e., if it contains more elements. Indeed, with a suitable definition of the limiting process, the sums (1.1.2) approach a certain number S, which equals the static moment of the line segment $[a,b]$ relative to the origin. We denote

$$S = \int_a^b |x|\,\mathrm{d}\mu(x).$$

The expression on the right-hand side is called the *Stieltjes integral* of the function $x \mapsto |x|,\ x \in [a,b]$, with respect to μ over the interval $[a,b]$. Of course, this very special function can be replaced by any "reasonable" function f defined on the interval $[a,b]$. In such a way, we can determine the moment of inertia of the line segment $[a,b]$ as $\int_a^b x^2\,\mathrm{d}\mu(x)$, and, in general, the moment of the k-th order as $\int_a^b |x|^k\,\mathrm{d}\mu(x)$.

1.2 Line integrals

LINE INTEGRAL OF THE FIRST KIND

A continuous mapping $\varphi:[a,b]\to\mathbb{R}^3$ is called a *path* in $\mathbb{R}^3$. The *length of the path* φ will be denoted by the symbol $\Lambda(\varphi,[a,b])$.

Let φ be a path in $\mathbb{R}^3$ having a finite length. Moreover, assume that the mapping φ is injective. Let us imagine that φ describes the shape of a wire and $f(x)\in\mathbb{R}$ is its density at the point x. The mass of the part of the wire corresponding to an interval $[c,d]\subset[a,b]$ is approximately given by $f(\varphi(\xi))\,\Lambda(\varphi,[c,d])$, where ξ is a point in the interval $[c,d]$.

Let $\boldsymbol{\alpha}=\{\alpha_0,\alpha_1,\dots,\alpha_m\}$ be a division of $[a,b]$ and $\boldsymbol{\xi}=\{\xi_1,\dots,\xi_m\}$ be the set of its tags, i.e. $\xi_j\in[\alpha_{j-1},\alpha_j]$ for $j\in\{1,\dots,m\}$.

Set $v(t)=\Lambda(\varphi,[a,t])$ for $t\in[a,b]$. Then the sum

$$\sum_{j=1}^{m} f(\varphi(\xi_j))\,(v(\alpha_j)-v(\alpha_{j-1}))$$

approximates the mass of the whole wire. Again, it is natural to expect that this approximation will be more precise when the division becomes finer. If such a limiting process leads to a uniquely determined value M, this value will be equal to the mass of the whole wire and we write

$$M=\int_{\varphi} f\,\mathrm{d}s \quad \text{or also} \quad M=\int_a^b f\circ\varphi\,\mathrm{d}v.$$

The former expression is called the *line integral of the first kind* of the function f along the path φ, while the latter expression represents the equivalent concept of the *Stieltjes integral* of a scalar function $f\circ\varphi$ with respect to the scalar function v.

LINE INTEGRAL OF THE SECOND KIND

Consider a point mass travelling along a path $\varphi:[a,b]\to\mathbb{R}^3$ under the influence of a force field $f:\mathbb{R}^3\to\mathbb{R}^3$. Then $f(\varphi(t))\in\mathbb{R}^3$ is the the force acting on the mass at time t.

Now, let $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \ldots, \alpha_m\}$ be a division of the interval $[a,b]$ and let $\boldsymbol{\xi} = \{\xi_1, \ldots, \xi_m\}$ be a corresponding set of tags. The scalar product

$$f(\varphi(\xi_j)) \cdot \big(\varphi(\alpha_j) - \varphi(\alpha_{j-1})\big) = \sum_{k=1}^{3} f_k(\varphi(\xi_j))\,\big(\varphi_k(\alpha_j) - \varphi_k(\alpha_{j-1})\big)$$

approximates the work done by the force f as the mass moves from the point $\varphi(\alpha_{j-1})$ to the point $\varphi(\alpha_j)$. Hence, the sum

$$\sum_{j=1}^{m} f(\varphi(\xi_j)) \cdot \big(\varphi(\alpha_j) - \varphi(\alpha_{j-1})\big) = \sum_{k=1}^{3}\sum_{j=1}^{m} f_k(\varphi(\xi_j))\,\big(\varphi_k(\alpha_j) - \varphi_k(\alpha_{j-1})\big)$$

approximates the total work done by the force field f during the whole motion of the given mass. If the values of these sums approach a uniquely determined value with respect to the refinements of the divisions of the interval $[a,b]$, this value will be equal to the total work done by the force field f during the motion of the given point mass along the path φ. It is usually denoted by

$$\int_{\varphi} f \qquad \text{or also} \qquad \int_a^b f \circ \varphi \,\mathrm{d}\varphi = \sum_{k=1}^{3} \int_a^b f_k(\varphi)\,\mathrm{d}\varphi_k.$$

The former expression is called the *line integral of the second kind* of the vector function f along the path φ, while the latter expression represents the equivalent concept of the *Stieltjes integral* of the (composite) vector function $f \circ \varphi \colon [a,b] \to \mathbb{R}^3$ with respect to the vector function $\varphi : [a,b] \to \mathbb{R}^3$.

Chapter 2

Functions of bounded variation

In this chapter, we define the variation of a real function defined on a compact interval and derive some basic properties of functions having bounded variation. Such functions are useful in a whole range of physical and technical problems, in probability theory, in the theory of Fourier series, in differential equations and other areas of mathematics.

2.1 Definitions and basic properties

Let $-\infty < a < b < \infty$. Recall that by a division of the interval $[a,b]$, we mean a finite set $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \dots, \alpha_m\}$ of points of the interval $[a,b]$ such that

$$a = \alpha_0 < \alpha_1 < \cdots < \alpha_m = b,$$

while $\mathscr{D}[a,b]$ stands for the set of all divisions of the interval $[a,b]$. Furthermore, the elements of the division $\boldsymbol{\alpha}$ of $[a,b]$ are usually denoted by α_j and

$$\nu(\boldsymbol{\alpha}) = m, \ \ \alpha_{\nu(\boldsymbol{\alpha})} = b \ \text{ and } \ |\boldsymbol{\alpha}| = \max_{j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}} (\alpha_j - \alpha_{j-1}) \ \text{ for } \boldsymbol{\alpha} \in \mathscr{D}[a,b].$$

If $\boldsymbol{\beta} \supset \boldsymbol{\alpha}$, we say that $\boldsymbol{\beta}$ is a *refinement* of $\boldsymbol{\alpha}$.

2.1.1 Definition. For a given function $f:[a,b] \to \mathbb{R}$ and a division $\boldsymbol{\alpha}$ of the interval $[a,b]$, we define

$$V(f,\boldsymbol{\alpha}) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} |f(\alpha_j) - f(\alpha_{j-1})| \quad \text{and} \quad \operatorname{var}_a^b f = \sup_{\boldsymbol{\alpha} \in \mathscr{D}[a,b]} V(f,\boldsymbol{\alpha}).$$

For $a = b$, we define $\operatorname{var}_a^b f = \operatorname{var}_a^a f = 0$. The quantity $\operatorname{var}_a^b f$ is called the *variation of the function f on the interval* $[a,b]$. If $\operatorname{var}_a^b f < \infty$, we say that the function f has *bounded variation* on $[a,b]$ (or *is of bounded variation on* $[a,b]$). The set of functions of bounded variation on $[a,b]$ is denoted by $\mathrm{BV}([a,b])$.

The concept of variation is closely related to the problem of rectifiability of curves. Let us recall the definition of the length of a parametric curve $c:[a,b] \to \mathbb{R}^n$ (where c is a continuous mapping). For each division $\boldsymbol{\alpha}$ of the interval $[a,b]$, the sum

$$\lambda(c,\boldsymbol{\alpha}) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} |c(\alpha_j) - c(\alpha_{j-1})|$$

equals the length of the polygonal path connecting the vertices $c(\alpha_0), \ldots, c(\alpha_{\nu(\boldsymbol{\alpha})})$. The length $\Lambda(c,[a,b])$ of the curve c is then defined as

$$\Lambda(c,[a,b]) = \sup_{\boldsymbol{\alpha}\in\mathscr{D}[a,b]} \lambda(c,\boldsymbol{\alpha}).$$

The next theorem provides a necessary and sufficient condition for the length of a curve to be finite. To distinguish it from the Jordan decomposition theorem, we call it JORDAN'S SECOND THEOREM.

2.1.2 Theorem (JORDAN'S SECOND THEOREM)**.** *Consider a parametric curve* $c:[a,b]\to\mathbb{R}^n$, *where* $c(t) = (c_1(t),\ldots,c_n(t))$ *for each* $t\in[a,b]$. *Then the length of* c *is finite if and only if each of the functions* $c_1,\ldots,c_n$ *has bounded variation on the interval* $[a,b]$.

Proof. If $x_1, \ldots, x_n$ are arbitrary real numbers, then

$$x_i^2 \le x_1^2 + \cdots + x_n^2 \le (|x_1| + \cdots + |x_n|)^2, \quad i\in\{1,\ldots,n\},$$

and therefore

$$|x_i| \le \sqrt{x_1^2 + \cdots + x_n^2} \le |x_1| + \cdots + |x_n|, \quad i\in\{1,\ldots,n\}. \tag{2.1.1}$$

Thus, for an arbitrary division $\boldsymbol{\alpha}$ of $[a,b]$ and each $i\in\{1,\ldots,n\}$, we obtain

$$\begin{aligned}\sum_{j=1}^{\nu(\boldsymbol{\alpha})} |c_i(\alpha_j) - c_i(\alpha_{j-1})| &\le \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \sqrt{\sum_{k=1}^{n} (c_k(\alpha_j) - c_k(\alpha_{j-1}))^2} \\ &\le \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \sum_{k=1}^{n} |c_k(\alpha_j) - c_k(\alpha_{j-1})| \\ &= \sum_{k=1}^{n} \sum_{j=1}^{\nu(\boldsymbol{\alpha})} |c_k(\alpha_j) - c_k(\alpha_{j-1})|.\end{aligned}$$

This means that

$$V(c_i,\boldsymbol{\alpha}) \le \lambda(c,\boldsymbol{\alpha}) \le \sum_{k=1}^{n} V(c_k,\boldsymbol{\alpha}).$$

Passing to the supremum, we get

$$\operatorname{var}_a^b c_i \le \Lambda(c,[a,b]) \le \sum_{k=1}^{n} \operatorname{var}_a^b c_k,$$

wherefrom the statement of our theorem follows immediately. □

In practice, one often deals with planar curves defined by an equation $y = f(x)$, where $f : [a,b] \to \mathbb{R}$ is a continuous function. The corresponding parametric curve $c : [a,b] \to \mathbb{R}^2$ is given by $c(t) = (t, f(t))$, $t \in [a,b]$, and Theorem 2.1.2 implies the following statement.

2.1.3 Corollary. *The graph of a function $f : [a,b] \to \mathbb{R}$ has finite length if and only if f has bounded variation on the interval $[a,b]$.*

2.1.4 Example. Let $f : [a,b] \to \mathbb{R}$ be continuous and such that $|f'(x)| \le M$ for all $x \in (a,b)$, where $M \in \mathbb{R}$ is independent of x. Then, by the mean value theorem, the inequality

$$|f(y) - f(x)| \le M\,|y - x|$$

holds for all $x, y \in [a,b]$. Hence, for each division $\boldsymbol{\alpha}$ of the interval $[a,b]$, we have

$$V(f, \boldsymbol{\alpha}) \le M \sum_{j=1}^{\nu(\boldsymbol{\alpha})} (\alpha_j - \alpha_{j-1}) = M\,(b-a).$$

We can see that *every continuous and real valued function on $[a,b]$ which has bounded derivative in its interior (a,b) is of bounded variation.*

If $|f'|$ is in addition Riemann integrable on $[a,b]$ (e.g., if f' is continuous on (a,b)), then the variation of f on $[a,b]$ can be calculated as follows.

2.1.5 Theorem. *If $f : [a,b] \to \mathbb{R}$ is such that $|f'|$ is Riemann integrable, then*

$$\operatorname{var}_a^b f = \int_a^b |f'(x)|\,\mathrm{d}x. \tag{2.1.2}$$

Proof. Let $\varepsilon > 0$ be given. The existence of the Riemann integral $\int_a^b |f'(x)|\,\mathrm{d}x$ means that there exists a $\delta > 0$ such that the inequality

$$\left| \sum_{j=1}^{\nu(\boldsymbol{\alpha})} |f'(\xi_j)|\,(\alpha_j - \alpha_{j-1}) - \int_a^b |f'(x)|\,\mathrm{d}x \right| < \frac{\varepsilon}{2} \tag{2.1.3}$$

holds for each division $\boldsymbol{\alpha}$ of $[a,b]$ such that $|\boldsymbol{\alpha}| < \delta$ and for every choice of points ξ_j such that

$$\xi_j \in [\alpha_{j-1}, \alpha_j] \quad \text{for } j = 1, \ldots, \nu(\boldsymbol{\alpha}). \tag{2.1.4}$$

On the other hand, by the definition of variation, there exists $\boldsymbol{\alpha} \in \mathscr{D}[a,b]$ such that $|\boldsymbol{\alpha}| < \delta$ and

$$\operatorname{var}_a^b f \geq V(f,\boldsymbol{\alpha}) > \operatorname{var}_a^b f - \frac{\varepsilon}{2}. \tag{2.1.5}$$

By the mean value theorem, there are points $\xi_j, j = 1, \ldots, \nu(\boldsymbol{\alpha})$, satisfying (2.1.4) and such that

$$V(f,\boldsymbol{\alpha}) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} |f'(\xi_j)| \, (\alpha_j - \alpha_{j-1}).$$

This, together with (2.1.3) and (2.1.5), implies that

$$\begin{aligned}
&\left|\operatorname{var}_a^b f - \int_a^b |f'(x)| \, \mathrm{d}x\right| \\
&\leq |\operatorname{var}_a^b f - V(f,\boldsymbol{\alpha})| + \left|\sum_{j=1}^{\nu(\boldsymbol{\alpha})} |f'(\xi_j)| \, (\alpha_j - \alpha_{j-1}) - \int_a^b |f'(x)| \, \mathrm{d}x\right| \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, it follows that (2.1.2) is true. □

2.1.6 Exercises. Determine $\operatorname{var}_a^b f$ and estimate the length of the graph of the following functions f:

a) $f(x) = \sin^2 x, \quad a = 0, \; b = \pi,$

b) $f(x) = x^3 - 3x + 4, \quad a = 0, \; b = 2,$

c) $f(x) = \cos x + x \sin x, \quad a = 0, \; b = 2\pi.$

2.1.7 Remark. By Definition 2.1.1, it is obvious that $\operatorname{var}_a^b f$ is nonnegative for every function $f : [a,b] \to \mathbb{R}$. Furthermore,

$$\operatorname{var}_a^b f = \sup_{\boldsymbol{\alpha} \supset \boldsymbol{\beta}} V(f,\boldsymbol{\alpha}) \tag{2.1.6}$$

holds for any division $\boldsymbol{\beta} \in \mathscr{D}[a,b]$. This follows from several elementary observations: First, because

$$\{V(f,\boldsymbol{\alpha}) : \boldsymbol{\alpha} \in \mathscr{D}[a,b] \;\text{ and }\; \boldsymbol{\alpha} \supset \boldsymbol{\beta}\} \subset \{V(f,\boldsymbol{\alpha}) : \boldsymbol{\alpha} \in \mathscr{D}[a,b]\},$$

it follows that

$$\sup_{\boldsymbol{\alpha} \supset \boldsymbol{\beta}} V(f,\boldsymbol{\alpha}) \leq \operatorname{var}_a^b f.$$

Moreover, by the triangle inequality, for arbitrary two divisions $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}'$ of $[a,b]$ such that $\boldsymbol{\alpha}' \supset \boldsymbol{\alpha}$ and for any function $f:[a,b]\to\mathbb{R}$ we have

$$V(f,\boldsymbol{\alpha}) \le V(f,\boldsymbol{\alpha}').$$

Finally, if $\boldsymbol{\alpha}\in\mathscr{D}[a,b]$ is given and $\boldsymbol{\alpha}'=\boldsymbol{\alpha}\cup\boldsymbol{\beta}$, then $\boldsymbol{\alpha}'\supset\boldsymbol{\beta}$ and thus

$$V(f,\boldsymbol{\alpha}) \le V(f,\boldsymbol{\alpha}').$$

This means that for every $d\in\{V(f,\boldsymbol{\alpha}):\boldsymbol{\alpha}\in\mathscr{D}[a,b]\}$ there exists

$$d'\in\{V(f,\boldsymbol{\alpha}):\boldsymbol{\alpha}\in\mathscr{D}[a,b] \text{ and } \boldsymbol{\alpha}\supset\boldsymbol{\beta}\}$$

such that $d\le d'$. Thus

$$\operatorname{var}_a^b f \le \sup_{\boldsymbol{\alpha}\supset\boldsymbol{\beta}} V(f,\boldsymbol{\alpha}),$$

which implies (2.1.6).

2.1.8 Exercises. Prove the following properties of the variation and of functions of bounded variation:

(i) If $[c,d]\subset[a,b]$, then

$$|f(d)-f(c)| \le \operatorname{var}_c^d f \le \operatorname{var}_a^b f$$

holds for every function $f:[a,b]\to\mathbb{R}$.

(ii) If $f\in\mathrm{BV}([a,b])$, then

$$|f(x)| \le |f(a)| + \operatorname{var}_a^b f < \infty \quad \text{for every } x\in[a,b]. \tag{2.1.7}$$

(iii) $\operatorname{var}_a^b f = d\in\mathbb{R}$ if and only if for each $\varepsilon>0$ there is $\boldsymbol{\alpha}_\varepsilon\in\mathscr{D}[a,b]$ such that

$$d-\varepsilon \le V(f,\boldsymbol{\alpha}) \le d$$

holds for all refinements $\boldsymbol{\alpha}$ of $\boldsymbol{\alpha}_\varepsilon$.

(iv) $\operatorname{var}_a^b f=\infty$ if and only if for each $K>0$ there is a division $\boldsymbol{\alpha}_K\in\mathscr{D}[a,b]$ such that $V(f,\boldsymbol{\alpha}_K)\ge K$.

(v) $\operatorname{var}_a^b f=\infty$ if and only if there exists a sequence $\{\boldsymbol{\alpha}^n\}$ of divisions of $[a,b]$ such that

$$\lim_{n\to\infty} V(f,\boldsymbol{\alpha}^n)=\infty.$$

(vi) If, for a given function $f:[a,b]\to\mathbb{R}$, there is an $L\in\mathbb{R}$ such that

$$|f(x)-f(y)| \le L\,|x-y| \quad \text{for all } x,y\in[a,b],$$

then $\operatorname{var}_a^b f\le L\,(b-a)$. (In such a case, we say that f satisfies the *Lipschitz condition* on $[a,b]$, or that f is *Lipschitz continuous* on $[a,b]$.)

2.1.9 Remark. The inequality (2.1.7) implies that every function with bounded variation on $[a,b]$ is necessarily bounded on $[a,b]$..

2.1.10 Example. Let

$$f(x)=\begin{cases} 0 & \text{if } x=0,\\ x\,\sin\left(\dfrac{\pi}{x}\right) & \text{if } x\in(0,2]. \end{cases}$$

Notice that $f(x)=0$ if and only if $x=0$ or $x=\dfrac{1}{k}$ for some $k\in\mathbb{N}$. Furthermore, for $x\in(0,2]$ we have

$$f(x)=\begin{cases} x & \text{if and only if } \; x=y_k=\dfrac{2}{4k+1},\, k\in\mathbb{N}\cup\{0\},\\ -x & \text{if and only if } \; x=z_k=\dfrac{2}{4k-1},\, k\in\mathbb{N}. \end{cases}$$

Thus, for a given $n\in\mathbb{N}$ and for $\boldsymbol{\alpha}^n=\{0,y_n,z_n,\dots,y_1,z_1,2\}$, we have

$$\begin{aligned} V(f,\boldsymbol{\alpha}^n) &= |f(0)-f(y_n)|+\sum_{k=1}^{n}|f(y_{k-1})-f(z_k)|+\sum_{k=1}^{n}|f(y_k)-f(z_k)|\\ &= y_n+\sum_{k=1}^{n}\big(y_{k-1}+z_k\big)+\sum_{k=1}^{n}\big(y_k+z_k\big)\\ &= y_0+2\sum_{k=1}^{n}\big(y_k+z_k\big)=2+4\sum_{k=1}^{n}\frac{8\,k}{16k^2-1}\ge 2\left(1+\sum_{k=1}^{n}\frac{1}{k}\right). \end{aligned}$$

It is known that $\sum_{k=1}^{\infty}\frac{1}{k}=\infty$. Hence, $\lim_{n\to\infty}V(f,\boldsymbol{\alpha}^n)=\infty$ and, consequently, $\text{var}_0^2 f=\infty$.

We can easily determine the variation of a monotone function.

2.1.11 Theorem. *If f is monotone on $[a,b]$, then*

$$\text{var}_a^b f=|f(b)-f(a)|.$$

Proof. If f is nonincreasing on $[a,b]$ and $\boldsymbol{\alpha}\in\mathscr{D}[a,b]$, then

$$\begin{aligned} V(f,\boldsymbol{\alpha}) &= \sum_{j=1}^{m}\big[f(\alpha_{j-1})-f(\alpha_j)\big]\\ &= \big[f(a)-f(\alpha_1)\big]+\big[f(\alpha_1)-f(\alpha_2)\big]+\cdots\\ &\qquad +\big[f(\alpha_{m-2})-f(\alpha_{m-1})\big]+\big[f(\alpha_{m-2})-f(b)\big]\\ &= f(a)-f(b), \end{aligned}$$

i.e., $\operatorname{var}_a^b f = f(a) - f(b) = \big|f(b) - f(a)\big|$.

In a similar way, one can show that if f is nondecreasing on $[a,b]$, then

$$\operatorname{var}_a^b f = f(b) - f(a) = \big|f(b) - f(a)\big|.$$

□

2.1.12 Exercises. (i) Prove that the function $f:[a,b]\to\mathbb{R}$ has a bounded variation if and only if there exists a nondecreasing function $h:[a,b]\to\mathbb{R}$ such that

$$|f(y) - f(x)| \le h(y) - h(x) \quad \text{for } x, y \in [a,b], x \le y.$$

(ii) Prove that the inequality

$$\sum_{j=1}^{\nu(\boldsymbol{\alpha})} \Big(|\Delta^+ f(\alpha_{j-1})| + |f(\alpha_j-) - f(\alpha_{j-1}+)| + |\Delta^- f(\alpha_j)|\Big) \le \operatorname{var}_a^b f$$

holds for each function $f:[a,b]\to\mathbb{R}$ and each division $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \dots, \alpha_{\nu(\boldsymbol{\alpha})}\}$ of $[a,b]$.

Hint: Consider expressions

$$\sum_{j=1}^{\nu(\boldsymbol{\alpha})} \Big(|f(\alpha_{j-1}+\delta) - f(\alpha_{j-1})| + |f(\alpha_j-\delta) - f(\alpha_{j-1}+\delta)| + |f(\alpha_j) - f(\alpha_j-\delta)|\Big)$$

and let $\delta \to 0$.

2.1.13 Examples. (i) A simple example of a bounded variation function which does not have bounded derivative (and hence the statement from Example 2.1.4 (i) is not applicable) is $f(x) = \sqrt{x}$, $x \in [0,1]$. Indeed, since f is increasing on $[0,1]$, by Theorem 2.1.11 we have $\operatorname{var}_0^1 f = 1$.

(ii) An example of a discontinuous function with a bounded variation is

$$f(x) = \begin{cases} 0 & \text{if } x = 0, \\ \dfrac{1}{k} & \text{if } x \in (0,1] \text{ and } x \in (\tfrac{1}{k+1}, \tfrac{1}{k}] \text{ for some } k \in \mathbb{N}. \end{cases}$$

This function is nondecreasing on the interval $[0,1]$, and therefore $\operatorname{var}_0^1 f = 1$ by Theorem 2.1.11.

2.1.14 Theorem. *For every* $c \in [a,b]$ *and every function* $f:[a,b]\to\mathbb{R}$, *we have*

$$\operatorname{var}_a^b f = \operatorname{var}_a^c f + \operatorname{var}_c^b f.$$

Proof. Let $f:[a,b]\to\mathbb{R}$ and $c\in[a,b]$ be given. If $c=a$ or $c=b$, the statement of the theorem is trivial. Thus, assume that $c\in(a,b)$.

Let $\widetilde{\boldsymbol{\alpha}}=\{a,c,b\}$ and let $\boldsymbol{\alpha}\in\mathscr{D}[a,b]$ be an arbitrary refinement of $\widetilde{\boldsymbol{\alpha}}$. Then necessarily $c\in\boldsymbol{\alpha}$ and we can split the division $\boldsymbol{\alpha}$ in two parts: the division $\boldsymbol{\alpha}'$ of the interval $[a,c]$ and the division $\boldsymbol{\alpha}''$ of the interval $[c,b]$, i.e., $\boldsymbol{\alpha}=\boldsymbol{\alpha}'\cup\boldsymbol{\alpha}''$, where $\boldsymbol{\alpha}'\in\mathscr{D}[a,c]$ and $\boldsymbol{\alpha}''\in\mathscr{D}[c,b]$. Obviously, we have

$$V(f,\boldsymbol{\alpha})=V(f,\boldsymbol{\alpha}')+V(f,\boldsymbol{\alpha}''), \tag{2.1.8}$$

wherefrom, by Remark 2.1.7, we deduce

$$\operatorname{var}_a^b f=\sup_{\boldsymbol{\alpha}\supset\widetilde{\boldsymbol{\alpha}}} V(f,\boldsymbol{\alpha})\le\operatorname{var}_a^c f+\operatorname{var}_c^b f. \tag{2.1.9}$$

On the other hand, for every two divisions $\boldsymbol{\alpha}'\in\mathscr{D}[a,c]$ and $\boldsymbol{\alpha}''\in\mathscr{D}[c,b]$, their union $\boldsymbol{\alpha}=\boldsymbol{\alpha}'\cup\boldsymbol{\alpha}''$ is a division of the interval $[a,b]$ and (2.1.8) holds again. This implies

$$\operatorname{var}_a^c f+\operatorname{var}_c^b f=\sup_{\boldsymbol{\alpha}'\in\mathscr{D}[a,c]} V(f,\boldsymbol{\alpha}')+\sup_{\boldsymbol{\alpha}''\in\mathscr{D}[c,b]} V(f,\boldsymbol{\alpha}'')\le\operatorname{var}_a^b f,$$

which completes the proof of (2.1.9). □

2.1.15 Example. Let $n\in\mathbb{N}$. Consider the function $f_n:[0,2]\to\mathbb{R}$ given by

$$f_n(x)=\begin{cases}0 & \text{if } 0\le x\le\frac{1}{n},\\ x\,\sin\left(\dfrac{\pi}{x}\right) & \text{if } \frac{1}{n}<x\le 2.\end{cases}$$

Its derivative

$$f_n'(x)=\begin{cases}0 & \text{if } 0\le x<\frac{1}{n},\\ \sin\left(\dfrac{\pi}{x}\right)-\dfrac{\pi}{x}\,\cos\left(\dfrac{\pi}{x}\right) & \text{if } \frac{1}{n}<x\le 2\end{cases}$$

is bounded on $(0,\frac{1}{n})$ and on $(\frac{1}{n},2)$. Evidently, $\operatorname{var}_0^{1/n} f_n=0$. By Example 2.1.4(i), $\operatorname{var}_{1/n}^2 f_n<\infty$. Thus, Theorem 2.1.14 yields $\operatorname{var}_0^1 f_n<\infty$ for every $n\in\mathbb{N}$.

The following evident statement implies that the set $\mathrm{BV}([a,b])$ is closed under pointwise addition of functions and multiplication by scalars (see Conventions and Notation, item (viii)).

2.1.16 Lemma. *For given functions $f,g:[a,b]\to\mathbb{R}$ and a real number c, the relations*

$$\operatorname{var}_a^b(f+g)\le\operatorname{var}_a^b f+\operatorname{var}_a^b f \tag{2.1.10}$$

and

$$\operatorname{var}_a^b(c\,f)=|c|\operatorname{var}_a^b f \tag{2.1.11}$$

are true. Furthermore, $\operatorname{var}_a^b f=0$ *if and only if f is constant on $[a,b]$.*

Proof. It suffices to notice that

$$V(f+g,\boldsymbol{\alpha}) \le V(f,\boldsymbol{\alpha}) + V(g,\boldsymbol{\alpha}) \quad \text{and} \quad V(c\,f,\boldsymbol{\alpha}) = |c|\,V(f,\boldsymbol{\alpha})$$

holds for every division $\boldsymbol{\alpha}$ of $[a,b]$, and that $\operatorname{var}_a^b f = 0$ implies $|f(x) - f(a)| = 0$ for each $x \in [a,b]$. □

2.1.17 Example. Show that the inequality

$$\operatorname{var}_a^b (f+g) \ge |\operatorname{var}_a^b f - \operatorname{var}_a^b g| \tag{2.1.12}$$

holds for any couple of functions $f, g \in \mathrm{BV}([a,b])$.
Hint: Notice that $f = (f+g) - g$ and $g = (f+g) - f$ and make use of (2.1.10).

2.1.18 Remark. A trivial example

$$f(x) = \cos\, x,\ g(x) = -\cos\, x, \ \text{ with } \operatorname{var}_0^{\pi/2} f = \operatorname{var}_0^{\pi/2} g = 1 \ \text{ and } \ f + g \equiv 0,$$

shows that in general, the inequality in (2.1.10) cannot be replaced by equality. On the other hand, it is possible to formulate conditions sufficient to ensure that (2.1.10) holds with equality. This is done by the following lemma.

2.1.19 Lemma. *Let* $f \in \mathrm{BV}([a,b])$ *and* $g \in \mathrm{BV}([a,b])$ *be such that for each* $\varepsilon > 0$ *there are* $n \in \mathbb{N}$ *and* $a_j, b_j \in [a,b]$, $j \in \{1, \dots, n\}$, *such that*

$$a \le a_1 \le b_1 \le \dots \le a_n \le b_n \le b, \tag{2.1.13}$$

$$\sum_{j=1}^{n} \operatorname{var}_{a_j}^{b_j} f > \operatorname{var}_a^b f - \varepsilon, \tag{2.1.14}$$

$$\sum_{j=1}^{n} \operatorname{var}_{a_j}^{b_j} g < \varepsilon. \tag{2.1.15}$$

Then

$$\operatorname{var}_a^b (f+g) = \operatorname{var}_a^b f + \operatorname{var}_a^b g. \tag{2.1.16}$$

Proof. Let $\varepsilon > 0$ be given and let $\{a_j, b_j\} \subset [a,b]$ with $j \in \{1, \dots, n\}$ be such that relations (2.1.13)-(2.1.15) are true.

Put $b_0 = a$ and $a_{n+1} = b$. Then

$$\operatorname{var}_a^b f = \sum_{j=1}^{n} \operatorname{var}_{a_j}^{b_j} f + \sum_{j=0}^{n} \operatorname{var}_{b_j}^{a_{j+1}} f$$

This together with (2.1.14) means that

$$\sum_{j=0}^{n} \operatorname{var}_{b_j}^{a_{j+1}} f < \varepsilon. \tag{2.1.17}$$

Similarly, by (2.1.15),

$$\sum_{j=0}^{n} \operatorname{var}_{b_j}^{a_{j+1}} g > \operatorname{var}_a^b g - \varepsilon. \tag{2.1.18}$$

Now, using (2.1.13), (2.1.14), (2.1.17), (2.1.18) and Example 2.1.17, we deduce that

$$\begin{aligned}\operatorname{var}_a^b (f+g) &= \sum_{j=1}^{n} \operatorname{var}_{a_j}^{b_j} (f+g) + \sum_{j=0}^{n} \operatorname{var}_{b_j}^{a_{j+1}} (f+g)\\ &\geq \sum_{j=1}^{n} (\operatorname{var}_{a_j}^{b_j} f - \operatorname{var}_{a_j}^{b_j} g) + \sum_{j=0}^{n} (\operatorname{var}_{b_j}^{a_{j+1}} g - \operatorname{var}_{b_j}^{a_{j+1}} f)\\ &> \operatorname{var}_a^b f - 2\,\varepsilon + \operatorname{var}_a^b g - 2\,\varepsilon = \operatorname{var}_a^b f + \operatorname{var}_a^b g - 4\,\varepsilon,\end{aligned}$$

i.e.

$$\operatorname{var}_a^b (f+g) > \operatorname{var}_a^b f + \operatorname{var}_a^b g - 4\,\varepsilon$$

holds for every $\varepsilon > 0$. Consequently,

$$\operatorname{var}_a^b (f+g) \geq \operatorname{var}_a^b f + \operatorname{var}_a^b g$$

wherefrom, by (2.1.10) the desired equality (2.1.16) follows. □

2.1.20 Remark. Some important examples of functions f, g satisfying the assumptions of Lemma 2.1.19 will be provided later, cf. Propositions 2.5.7 and 3.3.5.

2.1.21 Theorem. *A function* $f : [a,b] \to \mathbb{R}$ *has bounded variation on* $[a,b]$ *if and only if there exist nondecreasing functions* $f_1, f_2 : [a,b] \to \mathbb{R}$ *such that* $f = f_1 - f_2$.

Proof. If f_1 and f_2 are nondecreasing on $[a,b]$ and $f = f_1 - f_2$, then, by Theorem 2.1.11, both f_1 and f_2 have bounded variation on $[a,b]$. Hence, by (2.1.10), we have $\operatorname{var}_a^b f < \infty$.

Conversely, assume that $f \in \mathrm{BV}([a,b])$, and define

$$f_1(x) = \operatorname{var}_a^x f \quad \text{and} \quad f_2(x) = f_1(x) - f(x) \quad \text{for } x \in [a,b].$$

Let $x, y \in [a,b]$ and $y \geq x$. Then, by Theorem 2.1.14, $f_1(y) = f_1(x) + \operatorname{var}_x^y f$, and since the variation is always nonnegative, it follows that f_1 is nondecreasing on $[a,b]$. Furthermore, by Theorem 2.1.14 we have

$$f_2(y) = f_1(x) + \operatorname{var}_x^y f - f(y)$$

and

$$f_2(y) - f_2(x) = \operatorname{var}_x^y f - (f(y) - f(x)) \geq 0$$

(see Exercise 2.1.8 (i)). This means that the function f_2 is also nondecreasing and the proof is complete. □

2.1.22 Exercise. Let $f \in \mathrm{BV}([a,b])$. Prove that the functions

$$\mathfrak{p}\,(x) = \begin{cases} 0 & \text{if } x = a, \\ \sup\limits_{\boldsymbol{\alpha} \in \mathscr{D}[a,x]} \sum\limits_{j=1}^{\nu(\boldsymbol{\alpha})} \big(f(\alpha_j) - f(\alpha_{j-1})\big)^+ & \text{if } x \in (a,b] \end{cases}$$

and

$$\mathfrak{n}(x) = \begin{cases} 0 & \text{if } x = a, \\ \sup\limits_{\boldsymbol{\alpha} \in \mathscr{D}[a,x]} \sum\limits_{j=1}^{\nu(\boldsymbol{\alpha})} \big(f(\alpha_j) - f(\alpha_{j-1})\big)^- & \text{if } x \in (a,b] \end{cases}$$

are nondecreasing and nonnegative on $[a,b]$, and that the relations

$$f(x) = f(a) + \mathfrak{p}\,(x) - \mathfrak{n}(x) \quad \text{and} \quad \operatorname{var}_a^x f = \mathfrak{p}\,(x) + \mathfrak{n}(x)$$

hold for all $x \in [a,b]$.

2.1.23 Corollary. *For any function $f \in \mathrm{BV}([a,b])$ and all $t \in [a,b)$ and $s \in (a,b]$ there exist finite limits*

$$f(t+) = \lim_{\tau \to t+} f(\tau) \quad \textit{and} \quad f(s-) = \lim_{\tau \to s-} f(\tau).$$

Proof. By Theorem 2.1.21, we can assume that f is nondecreasing on $[a,b]$. Then $f(a) \leq f(x) \leq f(b)$ for every $x \in [a,b]$. Consequently,

$$f(a) \leq \sup_{x \in [a,s)} f(x) \leq f(b) \quad \text{for } s \in (a,b]$$

and

$$f(a) \leq \inf_{x \in (t,b]} f(x) \leq f(b) \quad \text{for } t \in [a,b).$$

Next, we will show that

$$f(t+) = \inf_{x \in (t,b]} f(x) \quad \text{if } t \in [a,b). \tag{2.1.19}$$

Let $d = \inf_{x \in (t,b]} f(x)$ and choose an arbitrary $\varepsilon > 0$. Then, by the definition of the infimum, there is a $t' \in (t,b]$ such that $d \leq f(t') < d + \varepsilon$. By the monotonicity of the function f, it follows that $d \leq f(x) < d + \varepsilon$ for every $x \in (t,t']$, wherefrom the relation (2.1.19) follows immediately.

In a similar way, one can show that

$$f(s-) = \sup_{x \in [a,s)} f(x) \quad \text{if } s \in (a,b]. \tag{2.1.20}$$

□

2.2 Space of functions of bounded variation

By Lemma 2.1.16, every linear combination of functions of bounded variation has a bounded variation, too. It follows that the set $\mathrm{BV}([a,b])$ is a linear space whose zero element is given by the identically zero function. We will show that, with suitably chosen norm, $\mathrm{BV}([a,b])$ becomes a normed linear space.

2.2.1 Theorem. $\mathrm{BV}([a,b])$ *is a normed linear space with respect to the norm defined by*

$$\|f\|_{\mathrm{BV}} = |f(a)| + \operatorname{var}_a^b f \quad \text{for } f \in \mathrm{BV}([a,b]). \tag{2.2.1}$$

Proof. $\mathrm{BV}([a,b])$ is a linear space by Lemma 2.1.16. By the same lemma, the relations

$$\|f+g\|_{\mathrm{BV}} \le \|f\|_{\mathrm{BV}} + \|g\|_{\mathrm{BV}} \quad \text{and} \quad \|c\,f\|_{\mathrm{BV}} = |c|\,\|f\|_{\mathrm{BV}} \tag{2.2.2}$$

hold for all $f, g \in \mathrm{BV}([a,b])$ and every $c \in \mathbb{R}$. Finally, if $\|f\|_{\mathrm{BV}} = 0$, then $f(a) = 0$ and $\operatorname{var}_a^b f = 0$. Hence, by Lemma 2.1.16, $f(x) - f(a) = 0$ on $[a,b]$, i.e., f is the zero element of $\mathrm{BV}([a,b])$. Consequently, the relation (2.2.1) defines a norm on $\mathrm{BV}([a,b])$. □

Next, we prove that $\mathrm{BV}([a,b])$ is a Banach space with respect to the norm given by (2.2.1). This fact will enable us to use the results of functional analysis in the study of the bounded variation functions.

2.2.2 Theorem. $\mathrm{BV}([a,b])$ *is a Banach space.*

Proof. It is sufficient to prove that $\mathrm{BV}([a,b])$ is complete, i.e., that every Cauchy sequence in $\mathrm{BV}([a,b])$ has a limit in $\mathrm{BV}([a,b])$. To this aim, let $\{f_n\} \subset BV([a,b])$ be a Cauchy sequence, i.e.,

$$\left.\begin{array}{l} \text{for each } \varepsilon > 0 \text{ there is } n_\varepsilon \in \mathbb{N} \text{ such that} \\ |f_n(x) - f_m(x)| \le \|f_n - f_m\|_{\mathrm{BV}} < \varepsilon \text{ for } x \in [a,b] \text{ and } n, m \ge n_\varepsilon. \end{array}\right\} \tag{2.2.3}$$

a) By (2.2.3), the sequence $\{f_n(x)\}$ is a Cauchy sequence of real numbers for every $x \in [a,b]$. Hence, for any $x \in [a,b]$ it has a finite limit

$$\lim_{n\to\infty} f_n(x) = f(x).$$

b) Let an arbitrary $\varepsilon > 0$ be given and let $n_\varepsilon \in \mathbb{N}$ be determined by (2.2.3). Then for every $x \in [a,b]$ we have

$$|f(x) - f_{n_\varepsilon}(x)| = \lim_{n\to\infty} |f_n(x) - f_{n_\varepsilon}(x)| \le \varepsilon,$$

and hence the inequalities

$$|f(x) - f_n(x)| \leq |f(x) - f_{n_\varepsilon}(x)| + |f_{n_\varepsilon}(x) - f_n(x)| < 2\,\varepsilon$$

hold for every $n \geq n_\varepsilon$ and every $x \in [a,b]$. This means that

$$\lim_{n\to\infty} \|f - f_n\|_\infty = 0.$$

In other words, the sequence $\{f_n\}$ tends to f uniformly on $[a,b]$.

c) By (2.2.2) and (2.2.3), there is $n_1 \in \mathbb{N}$ such that

$$\operatorname{var}_a^b f_n \leq \|f_n\|_{\mathrm{BV}} \leq \|f_{n_1}\|_{\mathrm{BV}} + 1 \quad \text{for } n \geq n_1.$$

Consequently, the sequence $\{\operatorname{var}_a^b f_n\}$ of real numbers is bounded by a certain $d \in [0,\infty)$. As a result, we have

$$V(f,\boldsymbol{\alpha}) = \lim_{n\to\infty} V(f_n,\boldsymbol{\alpha}) \leq d \quad \text{for all } \boldsymbol{\alpha} \in \mathscr{D}[a,b],$$

which implies

$$\operatorname{var}_a^b f = \sup_{\boldsymbol{\alpha}\in\mathscr{D}[a,b]} V(f,\boldsymbol{\alpha}) \leq d.$$

In particular, $f \in \mathrm{BV}([a,b])$.

d) It remains to show that

$$\lim_{n\to\infty} \|f - f_n\|_{\mathrm{BV}} = 0. \tag{2.2.4}$$

Let an arbitrary $\varepsilon > 0$ be given. By (2.2.3), there exists an $n_\varepsilon \in \mathbb{N}$ such that

$$V(f_n - f_m,\boldsymbol{\alpha}) \leq \operatorname{var}_a^b (f_n - f_m) < \varepsilon \text{ for all } n,m \geq n_\varepsilon \text{ and } \boldsymbol{\alpha} \in \mathscr{D}[a,b],$$

wherefrom, by letting $m \to \infty$, we deduce that

$$V(f_n - f,\boldsymbol{\alpha}) = \lim_{m\to\infty} V(f_n - f_m,\boldsymbol{\alpha}) \leq \varepsilon \text{ for all } n \geq n_\varepsilon \text{ and } \boldsymbol{\alpha} \in \mathscr{D}[a,b],$$

and consequently

$$\lim_{n\to\infty} \operatorname{var}_a^b (f_n - f) = 0.$$

This fact together with part a) of the proof implies that (2.2.4) is true. □

2.3 Bounded variation and continuity

2.3.1 Definition. Let a function $f:[a,b]\to\mathbb{R}$ be given. We say that $x\in(a,b)$ is its *point of discontinuity of the first kind* if both the one-sided limits

$$f(x-)=\lim_{t\to x-}f(t)\quad\text{and}\quad f(x+)=\lim_{t\to x+}f(t)$$

exist and are finite, while either $f(x-)\neq f(x)$ and/or $f(x+)\neq f(x)$. Analogously, a is the point of discontinuity of the first kind of f if the limit

$$f(a+)=\lim_{t\to a+}f(t)$$

is finite and $f(a+)\neq f(a)$, and b is the point of discontinuity of the first kind of f if the limit

$$f(b-)=\lim_{t\to b-}f(t)$$

is finite and $f(b-)\neq f(b)$.

By Corollary 2.1.23, functions of bounded variation can only have discontinuities of the first kind. Now, let us have a closer look on the properties of bounded variation functions related to continuity.

2.3.2 Theorem. *Every function $f\in\mathrm{BV}([a,b])$ has at most countably many discontinuities in the interval $[a,b]$.*

Proof. By virtue of Theorem 2.1.21, it is enough to prove the statement for the case when f is a nondecreasing function. Let D be the set of all discontinuity points of f. For each $x\in D$, choose an arbitrary rational $q(x)$ in the interval $(f(x-),f(x+))$. Since f is nondecreasing, it follows that $q(x)\neq q(y)$ whenever $x,y\in D$ and $x\neq y$. Hence, the mapping q provides a one-to-one correspondence between D and a subset of rational numbers. This proves that D is at most countable. □

Let $f\in\mathrm{BV}([a,b])$ and

$$v(x)=\operatorname{var}_a^x f\quad\text{for } x\in[a,b].\tag{2.3.1}$$

From the proof of Theorem 2.1.21, we know that the functions v and $v-f$ are nondecreasing on $[a,b]$. We now will show that the function v inherits the continuity properties of the function f.

2.3.3 Lemma. *Let* $f \in \mathrm{BV}([a,b])$ *and let* $v:[a,b] \to \mathbb{R}$ *be defined by* (2.3.1). *Then*

$$\Delta^{-} v(x) = |\Delta^{-} f(x)| \quad \textit{for} \;\; x \in (a,b] \tag{2.3.2}$$

and

$$\Delta^{+} v(x) = |\Delta^{+} f(x)| \quad \textit{for} \;\; x \in [a,b). \tag{2.3.3}$$

Proof. a) If $x \in (a,b]$, then

$$v(x) - v(s) = \operatorname{var}_s^x f \ge |f(x) - f(s)| \quad \text{holds for every} \;\; s \in [a,x].$$

For $s \to x-$, we get the inequality

$$\Delta^{-} v(x) \ge |\Delta^{-} f(x)| \quad \text{for} \;\; x \in (a,b]. \tag{2.3.4}$$

Let $\varepsilon > 0$ be given. Choose a $\delta > 0$ such that

$$|f(x-) - f(s)| < \frac{\varepsilon}{2} \quad \text{for} \;\; s \in (x-\delta, x). \tag{2.3.5}$$

Furthermore, choose a division $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \ldots, \alpha_m\}$ of $[a,x]$ such that

$$\alpha_{m-1} \ge x - \tfrac{\delta}{2} \quad \text{and} \quad v(x) - V(f, \boldsymbol{\alpha}) < \frac{\varepsilon}{2}. \tag{2.3.6}$$

Then, by (2.3.6) and (2.3.5), we have

$$\begin{aligned} v(x) - \sum_{j=1}^{m-1} |f(\alpha_j) - f(\alpha_{j-1})| &< V(f,\boldsymbol{\alpha}) + \frac{\varepsilon}{2} - \sum_{j=1}^{m-1} |f(\alpha_j) - f(\alpha_{j-1})| \\ &= |f(x) - f(\alpha_{m-1})| + \frac{\varepsilon}{2} \le |\Delta^{-} f(x)| + |f(x-) - f(\alpha_{m-1})| + \frac{\varepsilon}{2} \\ &\le |\Delta^{-} f(x)| + \varepsilon. \end{aligned}$$

Moreover, since

$$\sum_{j=1}^{m-1} |f(\alpha_j) - f(\alpha_{j-1})| \le v(\alpha_{m-1})$$

and $v(t) \ge v(\alpha_{m-1})$ for every $t \in (\alpha_{m-1}, x)$, it follows that

$$\begin{aligned} v(x) - v(t) \le v(x) - v(\alpha_{m-1}) \le v(x) - \sum_{j=1}^{m-1} |f(\alpha_j) - f(\alpha_{j-1})| \\ \le |\Delta^{-} f(x)| + \varepsilon \end{aligned}$$

holds for every $t \in (\alpha_{m-1}, x)$ and every $\varepsilon > 0$. Letting $t \to x-$ and recalling that $\varepsilon > 0$ can be arbitrary, we get the inequality $\Delta^- v(x) \le |\Delta^- f(x)|$, which together with (2.3.4) proves the equality (2.3.2).

b) The second relation (2.3.3) can be proved by a symmetry argument based on reflecting the graph of f about the vertical axis: Let $\widetilde{f} : [-b, -a] \to \mathbb{R}$ be given by $\widetilde{f}(x) = f(-x)$, $x \in [-b, -a]$. Observe that $\operatorname{var}_c^d f = \operatorname{var}_{-d}^{-c} \widetilde{f}$ whenever $[c, d] \subset [a, b]$, and $\Delta^+ f(x) = -\Delta^- \widetilde{f}(-x)$ for each $x \in [a, b)$. Therefore,

$$\begin{aligned} \Delta^+ v(x) &= \lim_{\delta \to 0+} \left(\operatorname{var}_a^{x+\delta} f - \operatorname{var}_a^x f \right) = \lim_{\delta \to 0+} \operatorname{var}_x^{x+\delta} f \\ &= \lim_{\delta \to 0+} \operatorname{var}_{-x-\delta}^{-x} \widetilde{f} = \lim_{\delta \to 0+} \left(\operatorname{var}_{-b}^{-x} \widetilde{f} - \operatorname{var}_{-b}^{-x-\delta} \widetilde{f} \right) \\ &= |\Delta^- \widetilde{f}(-x)| = |\Delta^+ f(x)|, \end{aligned}$$

where the first equality on the last line follows from part (a) applied to the function $\widetilde{f}$ at the point $-x$. □

The following statement is an immediate consequence of Lemma 2.3.3.

2.3.4 Corollary. *Let $f \in \mathrm{BV}([a, b])$ and let the function $v : [a, b] \to \mathbb{R}$ be defined by the relation* (2.3.1)*. Then f is right-continuous at a point $x \in [a, b)$ if and only if the function v is right-continuous at this point. Similarly, f is left-continuous at the point $x \in (a, b]$ if and only if the function v is left-continuous at this point.*

From the next theorem it will follow that the sum of absolute values of the jumps of a bounded variation function is always finite. For its proof, we need the following statement.

2.3.5 Lemma. *If $f : [a, b] \to \mathbb{R}$ has bounded variation and the functions $\mathfrak{p}$ and $\mathfrak{n}$ are defined as in Exercise* 2.1.22*, then*

$$\Delta^- \mathfrak{p}(x) = (\Delta^- f(x))^+, \ \Delta^- \mathfrak{n}(x) = (\Delta^- f(x))^- \quad \text{for } x \in (a, b], \tag{2.3.7}$$

and

$$\Delta^+ \mathfrak{p}(x) = (\Delta^+ f(x))^+, \ \Delta^+ \mathfrak{n}(x) = (\Delta^+ f(x))^- \quad \text{for } x \in [a, b). \tag{2.3.8}$$

The proof can be carried out analogously as the proof of Lemma 2.3.3; it suffices to work with

$$(f(\alpha_j) - f(\alpha_{j-1}))^+ \text{ or } (f(\alpha_j) - f(\alpha_{j-1}))^-$$

instead of $|f(\alpha_j) - f(\alpha_{j-1})|$. The detailed proof is left as an exercise for the reader.

2.3.6 Theorem. *Let* $f \in \mathrm{BV}([a,b])$ *and let* $D = \{s_k\}$ *be a non-repeating sequence (i.e.,* $s_k \neq s_\ell$ *whenever* $k \neq \ell$*) of points from the open interval* (a,b)*. Then*

$$|\Delta^+ f(a)| + \sum_{k=1}^{\infty} \Big(|\Delta^+ f(s_k)| + |\Delta^- f(s_k)| \Big) + |\Delta^- f(b)| \leq \operatorname{var}_a^b f. \qquad (2.3.9)$$

Proof. a) First, assume that f is nondecreasing. Then

$$\begin{aligned} &|\Delta^+ f(a)| + \sum_{k=1}^{\infty} \Big(|\Delta^+ f(s_k)| + |\Delta^- f(s_k)| \Big) + |\Delta^- f(b)| \\ &\qquad = \Delta^+ f(a) + \sum_{k=1}^{\infty} \Delta\, f(s_k) + \Delta^- f(b). \end{aligned}$$

Choose an arbitrary $n \in \mathbb{N}$. Denote $\alpha_0 = a$, $\alpha_k = s_k$ for $k \in \{1, \ldots, n\}$, and $\alpha_{n+1} = b$. Then

$$\begin{aligned} &\Delta^+ f(a) + \sum_{k=1}^{n} \Delta\, f(s_k) + \Delta^- f(b) \\ =&\Delta^+ f(\alpha_0) + \sum_{k=1}^{n} \Big(f(\alpha_k+) - f(\alpha_k-) \Big) + \Delta^- f(\alpha_{n+1}) \\ =& - f(\alpha_0) + \sum_{k=0}^{n} \Big(f(\alpha_k+) - f(\alpha_{k+1}-) \Big) + f(\alpha_{n+1}) \\ \leq& - f(\alpha_0) + f(\alpha_{n+1}) = f(b) - f(a) = \operatorname{var}_a^b f, \end{aligned}$$

where the last inequality follows from the fact that f is nondecreasing. Passing to the limit $n \to \infty$, we obtain (2.3.9).

b) Now, let f be an arbitrary function of bounded variation on $[a,b]$ and let the functions $\mathfrak{p}$ and $\mathfrak{n}$ be defined as in Exercise 2.1.22. We know that $f = f(a) + \mathfrak{p} - \mathfrak{n}$ on $[a,b]$. Obviously,

$$\Delta^+ f(t) = \Delta^+ \mathfrak{p}(t) - \Delta^+ \mathfrak{n}(t) \quad \text{and} \quad \Delta^- f(s) = \Delta^- \mathfrak{p}(s) - \Delta^- \mathfrak{n}(s)$$

for $t \in [a,b)$ and $s \in (a,b]$. Finally, using Lemma 2.3.5, we can easily deduce that the relations

$$|\Delta^+ f(t)| = \Delta^+ \mathfrak{p}(t) + \Delta^+ \mathfrak{n}(t) \quad \text{and} \quad |\Delta^- f(s)| = \Delta^- \mathfrak{p}(s) + \Delta^- \mathfrak{n}(s) \quad (2.3.10)$$

hold for $t \in [a,b)$, $s \in (a,b]$.

By the first part of the proof we have

$$\Delta^+ \mathfrak{p}(a) + \sum_{k=1}^{\infty} \Big(\Delta^+ \mathfrak{p}(s_k) + \Delta^- \mathfrak{p}(s_k) \Big) + \Delta^- \mathfrak{p}(b) \leq \mathfrak{p}(b)$$

and

$$\Delta^+\mathfrak{n}(a) + \sum_{k=1}^{\infty}\Big(\Delta^+\mathfrak{n}(s_k) + \Delta^-\mathfrak{n}(s_k)\Big) + \Delta^-\mathfrak{n}(b) \le \mathfrak{n}(b).$$

By adding these two inequalities, we obtain

$$\begin{aligned}|\Delta^+ f(a)| + \sum_{k=0}^{n}\Big(|\Delta^+ f(s_k)|+|\Delta^- f(s_k)|\Big) + |\Delta^- f(b)|\\ \le \mathfrak{p}(b) + \mathfrak{n}(b) = \operatorname{var}_a^b f.\end{aligned}$$

□

2.3.7 Remark. Let $f:[a,b]\to\mathbb{R}$ have a bounded variation and let the set D of its discontinuity points in (a,b) be infinite. By Theorem 2.3.2, the elements of D can be arranged into a sequence $\{s_k\}$. (Naturally, there is an infinite number of such sequences.) By Theorem 2.3.6, the series

$$\sum_{k=1}^{\infty}\Big(|\Delta^+ f(s_k)| + |\Delta^- f(s_k)|\Big)$$

is absolutely convergent and its sum does not depend on the ordering of points of D. Since for $x\in(a,b)$, the expression $|\Delta^+ f(x)| + |\Delta^- f(x)|$ is nonzero if and only if $x\in D$, it makes sense to define

$$\begin{aligned}\sum_{a<x<b}\Big(|\Delta^+ f(x)|+|\Delta^- f(x)|\Big) &= \sum_{x\in D}\Big(|\Delta^+ f(x)|+|\Delta^- f(x)|\Big)\\ &= \sum_{k=1}^{\infty}\Big(|\Delta^+ f(s_k)|+|\Delta^- f(s_k)|\Big),\end{aligned}$$

where $\{s_k\}$ is an arbitrary non-repeating sequence of points from (a,b) such that $D=\{s_k\}$. The symbols

$$\sum_{a\le x<b},\quad \sum_{a<x\le b},\quad \sum_{a\le x\le b},\quad \text{or}\quad \sum_{x\in[a,b)},\quad \sum_{x\in(a,b]},\quad \sum_{x\in[a,b]}\quad \text{etc.}$$

should be understood in an analogous way.

Theorem 2.3.6 implies the following result.

2.3.8 Corollary. *Each function $f\in\mathrm{BV}([a,b])$ satisfies the inequality*

$$\sum_{x\in[a,b)}|\Delta^+ f(x)| + \sum_{x\in(a,b]}|\Delta^- f(x)| \le \operatorname{var}_a^b f. \tag{2.3.11}$$

2.4 Derivatives of bounded variation functions

In this section we will consider the properties of functions of bounded variation related to differentiation. To this aim, let us first recall the concept of the Lebesgue outer measure.

For an arbitrary set $M \subset \mathbb{R}$, we define its *Lebesgue outer measure*

$$\mu^*(M) := \inf \sum_k (b_k - a_k),$$

where the infimum is taken over all at most countable collections $\{(a_k, b_k)\}$ of open intervals in $\mathbb{R}$ such that

$$M \subset \bigcup_k (a_k, b_k).$$

The Lebesgue outer measure is either a nonnegative real number, or ∞. Furthermore, $\mu^*(M_1) \le \mu^*(M_2)$ whenever $M_1 \subset M_2$.

Obviously, for any finite collection $\{I_k\}$ of disjoint open intervals in $\mathbb{R}$, the equality

$$\mu^*\left(\bigcup_k I_k\right) = \sum_k |I_k|$$

holds.

We say that a certain property holds *almost everywhere* (a.e.) on the interval $[a,b]$ if there exists a set $N \subset [a,b]$ whose Lebesgue outer measure is zero and such that the property holds for all $x \in [a,b] \setminus N$. Equivalently, we say that the property holds for almost all $x \in [a,b]$.

If not stated otherwise, in what follows by *outer measure* we always understand the Lebesgue outer measure.

2.4.1 Exercises. Prove the following assertions:

(i) Every countable set $S \subset \mathbb{R}$ has outer measure equal to zero.

(ii) If I is an interval, then $\mu^*(I)$ equals its length $|I|$.

(iii) If $\{M_k\}$ is a countable collection of subsets of $\mathbb{R}$, then

$$\mu^*\left(\bigcup_k M_k\right) \le \sum_k \mu^*(M_k).$$

(iv) The union of countably many sets whose outer measure is zero has outer measure equal to zero.

2.4.2 Theorem (LEBESGUE DIFFERENTIATION THEOREM)**.** *Every monotone function* $f:[a,b]\to\mathbb{R}$ *has a finite derivative* $f'(x)$ *for almost all* $x\in[a,b]$.

The proof of Theorem 2.4.2 can be found in many real analysis textbooks (see e.g. Appendix E in [11], Theorem 7.9 in [16], Theorem 4.9 in [43] or Theorem 6.2.1 in [111].

2.4.3 Remark. In particular, by Theorem 2.1.21, each function $f\in\mathrm{BV}([a,b])$ has a bounded derivative almost everywhere on the interval $[a,b]$. It is even known (see Theorem 3.2.1) that the derivatives of functions of bounded variation are Lebesgue integrable. However, the seemingly natural equation

$$f(x)-f(a)=\int_a^x f'(t)\,\mathrm{d}t \quad \text{for } x\in[a,b]$$

is not true for every function $f\in\mathrm{BV}([a,b])$. For example, there exist functions $f\in\mathrm{BV}([a,b])$ which are non-constant on $[a,b]$ and such that $f'=0$ a.e. on $[a,b]$.

2.4.4 Definition. A function $f\in\mathrm{BV}([a,b])$ is called *singular* if $f'(x)=0$ for almost all $x\in[a,b]$.

2.5 Step functions

The simplest example of non-constant singular functions are functions of the form $f=\chi_{[a,c]}$, where $c\in(a,b)$. Their natural generalizations are the *finite step functions*, sometimes called also *simple functions*.

2.5.1 Definition. A function $f:[a,b]\to\mathbb{R}$ is *a finite step function on* $[a,b]$ if there exists a division $\boldsymbol{\alpha}=\{\alpha_0,\alpha_1,\dots,\alpha_m\}$ of $[a,b]$ such that f is constant on each of the open intervals (α_{j-1},α_j), $j=1,\dots,m$. The set of finite step functions on the interval $[a,b]$ is denoted by $\mathrm{S}([a,b])$.

By definition, the function $f:[a,b]\to\mathbb{R}$ is a finite step function if and only if there exists an $m\in\mathbb{N}$, sets $\{\widetilde{c}_j : j=0,1,\dots,m\}\subset\mathbb{R}$, $\{\widetilde{d}_j : j=1,\dots,m\}\subset\mathbb{R}$ and a division $\{a=s_0<s_1<\dots<s_m=b\}$ of $[a,b]$ such that

$$f(s_j)=\widetilde{c}_j \quad \text{for } j=0,1,\dots,m,$$

and

$$f(x)=\widetilde{d}_j \quad \text{for } x\in(s_{j-1},s_j) \text{ and } j=1,\dots,m,$$

i.e.,

$$
\begin{aligned}
f(x) &= \sum_{j=0}^{m} \widetilde{c}_j\, \chi_{[s_j]}(x) + \sum_{j=1}^{m} \widetilde{d}_j\, \chi_{(s_{j-1},s_j)}(x) \\
&= \widetilde{c}_0\, (\chi_{[a,b]}(x) - \chi_{(a,b]}(x)) + \sum_{j=1}^{m-1} \widetilde{c}_j\, (\chi_{[s_j,b]} - \chi_{(s_j,b]}(x)) \\
&\qquad + \widetilde{c}_m\, \chi_{[b]}(x) + \sum_{j=1}^{m-1} \widetilde{d}_j\, (\chi_{(s_{j-1},b]}(x) - \chi_{[s_j,b]}(x)) \\
&\qquad + \widetilde{d}_m\, (\chi_{(s_{m-1},b]}(x) - \chi_{[b]}(x)) \\
&= \widetilde{c}_0 + \sum_{j=0}^{m-1} (\widetilde{d}_{j+1} - \widetilde{c}_j)\, \chi_{(s_j,b]}(x) \\
&\qquad + \sum_{j=1}^{m-1} (\widetilde{c}_j - \widetilde{d}_j)\, \chi_{[s_j,b]}(x) + (\widetilde{c}_m - \widetilde{d}_m)\, \chi_{[b]}(x).
\end{aligned}
$$

Equivalently,

$$
\left.
\begin{aligned}
f(x) = c + \sum_{j=0}^{m-1} c_j\, \chi_{(s_j,b]}(x) + \sum_{j=1}^{m-1} d_j\, \chi_{[s_j,b]}(x) + d\, \chi_{[b]}(x) \\
\text{for } x \in [a,b],
\end{aligned}
\right\} \tag{2.5.1}
$$

where

$$c = \widetilde{c}_0, \quad c_j = \widetilde{d}_{j+1} - \widetilde{c}_j \quad \text{for } j = 0, 1, \dots, m-1,$$

and

$$d_j = \widetilde{c}_j - \widetilde{d}_j \quad \text{for } j = 1, \dots, m-1, \quad d = \widetilde{c}_m - \widetilde{d}_m.$$

Obviously, $f(a) = c$,

$$f(x-) = f(x) \text{ for } x \in (a,b\,] \setminus \{s_k\}, \;\; f(x) = f(x+) \text{ for } x \in [a,b) \setminus \{s_k\}$$

and

$$
\begin{aligned}
\Delta^+ f(s_j) &= \widetilde{d}_{j+1} - \widetilde{c}_j = c_j \quad && \text{for } j = 0, 1, \dots, m-1, \\
\Delta^- f(s_j) &= \widetilde{c}_j - \widetilde{d}_j = d_j && \text{for } j = 1, \dots, m.
\end{aligned}
$$

A generalization of a finite step function is provided by *step functions*, sometimes called also *jump functions*.

2.5.2 Definition. A function $f : [a,b] \to \mathbb{R}$ is *a step function on* $[a,b]$ if either f is a finite step function or there exist $c, c_0, d \in \mathbb{R}$, a non-repeating sequence

$\{s_k\} \subset (a,b)$, and sequences $\{c_k\} \subset \mathbb{R}$ and $\{d_k\} \subset \mathbb{R}$ such that

$$\sum_{k=1}^{\infty} (|c_k| + |d_k|) < \infty \tag{2.5.2}$$

and

$$\left.\begin{aligned} f(x) = c + c_0\,\chi_{(a,b]}(x) + \sum_{k=1}^{\infty}\Big(c_k\,\chi_{(s_k,b]}(x) + d_k\,\chi_{[s_k,b]}(x)\Big) \\ + d\,\chi_{[b]}(x) \quad \text{for } x \in [a,b]. \end{aligned}\right\} \tag{2.5.3}$$

The set of all step functions on the interval $[a,b]$ is denoted by $\mathrm{B}([a,b])$.

If $f \in \mathrm{B}([a,b])$ is not a finite step function, then the sequence $\{s_k\}$ from Definition 2.5.2 is infinite and, in general, it is not possible to reorder it into an increasing sequence. However, thanks to condition (2.5.2) we have

$$\sum_{k=1}^{\infty} \big|c_k\,\chi_{(s_k,b]}(x) + d_k\,\chi_{[s_k,b]}(x)\big| \le \sum_{k=1}^{\infty} \big(|c_k| + |d_k|\big) < \infty. \tag{2.5.4}$$

This means that the series on the right-hand side of (2.5.3) is absolutely convergent for each $x \in [a,b]$. Hence, the values $f(x)$ do not depend on the particular ordering of the sequence $\{s_k\}$. Consequently, for each $x \in [a,b]$ the relation (2.5.3) can be equivalently rewritten as

$$f(x) = \begin{cases} c & \text{if } x = a, \\ c + c_0 + \displaystyle\sum_{a<s_k<x} c_k + \sum_{a<s_k\le x} d_k & \text{if } x \in (a,b), \\ c + c_0 + \displaystyle\sum_{a<s_k<b} c_k + \sum_{a<s_k<b} d_k + d & \text{if } x = b, \end{cases} \tag{2.5.5}$$

where the sum $\sum_{a<s_k<x} c_k$ runs over all indices $k \in \mathbb{N}$ such that $s_k \in (a,x)$, the sum $\sum_{a<s_k\le x} c_k$ runs over all indices $k \in \mathbb{N}$ such that $s_k \in (a,x]$, and analogously in the case of $\sum_{a<s_k<b} d_k$.

From (2.5.5) we see that the function $f \in \mathrm{B}([a,b])$ defined by (2.5.3) satisfies $f(a) = c$. From (2.5.5) we see that if there is a $\delta > 0$ such that $(a, a+\delta) \cap \{s_k\} = \emptyset$, then

$$f(a+) = c + c_0. \tag{2.5.6}$$

In the general case, one has to take into account that since

$$\lim_{t\to a+} \Bigg(\sum_{a<s_k<t} c_k + \sum_{a<s_k\le t} d_k\Bigg)$$

is in fact the limit of the remainder of an absolutely convergent series, it must be zero. Hence (2.5.6) holds also in the case when a is a limit point of the set $\{s_k\}$. Similar argument can be used to prove the following formulas:

$$\left.\begin{aligned} &f(x-) = c + c_0 + \sum_{a<s_k<x} c_k + \sum_{a<s_k<x} d_k \quad \text{if} \quad x \in (a,b],\\ &\text{and}\\ &f(x+) = c + c_0 + \sum_{a<s_k\le x} c_k + \sum_{a<s_k\le x} d_k \quad \text{if} \quad x \in [a,b). \end{aligned}\right\} \tag{2.5.7}$$

Subtracting (2.5.5) from (2.5.7) leads to

$$f(x-) = f(x) = f(x+) \quad \text{for } x \in (a,b) \setminus \{s_k\} \tag{2.5.8}$$

and

$$\left.\begin{aligned} &\Delta^+ f(s_k) = c_k \qquad \text{for } k \in \mathbb{N}, \quad \Delta^+ f(a) = c_0,\\ &\Delta^- f(s_k) = d_k \qquad \text{for } k \in \mathbb{N}, \quad \Delta^- f(b) = d. \end{aligned}\right\} \tag{2.5.9}$$

Thus, the relation (2.5.3) from Definition 2.5.2 can be reformulated in further two equivalent ways:

$$f(x) = \begin{cases} c & \text{if } x = a,\\ c + \Delta^+ f(a) + \sum\limits_{a<s_k<x} \Delta^+ f(s_k) + \sum\limits_{a<s_k\le x} \Delta^- f(s_k) & \\ & \text{if } x \in (a,b),\\ c + \Delta^+ f(a) + \sum\limits_{a<s_k<b} \Delta^+ f(s_k) + \sum\limits_{a<s_k<b} \Delta^- f(s_k) + \Delta^- f(b) & \\ & \text{if } x = b, \end{cases} \tag{2.5.10}$$

or

$$\left.\begin{aligned} f(x) = f(a) &+ \sum_{d\in D} \big[\Delta^+ f(d)\, \chi_{(d,b]}(x) + \Delta^- f(d)\, \chi_{[d,b]}(x)\big]\\ &+ \Delta^+ f(a)\, \chi_{(a,b]}(x) + \Delta^- f(b)\, \chi_{[b]}(x) \quad \text{for } x \in [a,b], \end{aligned}\right\} \tag{2.5.11}$$

where $D = \{s_k\}$. Recall that we assume that $D \subset (a,b)$. Hence, with respect to (2.5.9) we can see that D is the set of points of discontinuity of f in the open interval (a,b), while the set of points of discontinuity of f on the closed interval $[a,b]$ is contained in the set $D \cup \{a\} \cup \{b\}$. (The points a, b do not necessarily have to be discontinuity points of f.)

2.5.3 Theorem. $\mathrm{S}([a,b]) \subset \mathrm{B}([a,b]) \subset \mathrm{BV}([a,b])$ *and the inequality*

$$\operatorname{var}_a^b f = |\Delta^+ f(a)| + \sum_{x\in(a,b)} \Big(|\Delta^+ f(x)| + |\Delta^- f(x)|\Big) + |\Delta^- f(b)| < \infty \tag{2.5.12}$$

holds for each step function $f \in \mathrm{B}([a,b])$.

Proof. By Definition 2.5.1, we have

$$\mathrm{S}([a,b]) \subset \mathrm{B}([a,b]) \quad \text{and} \quad \mathrm{S}([a,b]) \subset \mathrm{BV}([a,b]).$$

Obviously, (2.5.12) holds for all finite step functions $f \in \mathrm{S}([a,b])$. Thus, we can restrict ourselves to the case that $f \in \mathrm{B}([a,b]) \setminus \mathrm{S}([a,b])$, i.e., we may assume that f is given by (2.5.5), and (2.5.2) is true. Using (2.5.9) we have

$$\left.\begin{aligned} &|\Delta^+ f(a)| + \sum_{a<x<b} \Big(|\Delta^+ f(x)| + |\Delta^- f(x)|\Big) + |\Delta^- f(b)| \\ &\qquad\qquad = |c_0| + \sum_{k=1}^{\infty} \Big(|c_k| + |d_k|\Big) + |d| < \infty. \end{aligned}\right\} \tag{2.5.13}$$

Notice that

$$|f(y) - f(x)| \le \sum_{x\le s_k<y} |c_k| + \sum_{x<s_k\le y} |d_k|$$

holds for arbitrary $x,\ y \in (a,b)$ such that $x<y$. Furthermore,

$$|f(y) - f(a)| \le |c_0| + \sum_{a<s_k<y} |c_k| + \sum_{a<s_k\le y} |d_k| \quad \text{if } a<y<b,$$

and

$$|f(b) - f(x)| \le \sum_{x\le s_k<b} |c_k| + \sum_{x<s_k<b} |d_k| + |d| \qquad \text{if } a<x<b.$$

Hence, for any division $\boldsymbol{\alpha}$ of $[a,b]$ with $\nu(\boldsymbol{\alpha}) \ge 3$ we can deduce

$$\begin{aligned} V(f,\boldsymbol{\alpha}) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} |f(\alpha_j) - f(\alpha_{j-1})| &\le |c_0| + \left(\sum_{a<s_k<\alpha_1} |c_k| + \sum_{a<s_k\le\alpha_1} |d_k|\right) \\ &+ \sum_{j=2}^{\nu(\boldsymbol{\alpha})-1} \left(\sum_{\alpha_{j-1}\le s_k<\alpha_j} |c_k| + \sum_{\alpha_{j-1}<s_k\le\alpha_j} |d_k|\right) \\ &+ \left(\sum_{\alpha_{\nu(\boldsymbol{\alpha})-1}\le s_k<b} |c_k| + \sum_{\alpha_{\nu(\boldsymbol{\alpha})-1}<s_k<b} |d_k|\right) + |d| \\ &\le |c_0| + \sum_{k=1}^{\infty} (|c_k| + |d_k|) + |d|. \end{aligned}$$

Consequently, by (2.5.2) we have

$$\operatorname{var}_a^b f \le |c_0| + \sum_{k=1}^{\infty} (|c_k| + |d_k|) + |d| < \infty, \tag{2.5.14}$$

i.e., $f \in \mathrm{BV}([a,b])$ and $\mathrm{B}([a,b]) \subset \mathrm{BV}([a,b])$. Finally, using Theorem 2.3.6 we get

$$|\Delta^+ f(a)| + \sum_{x\in(a,b)} \Big(|\Delta^+ f(x)| + |\Delta^- f(x)|\Big) + |\Delta^- f(b)| \le \operatorname{var}_a^b f. \tag{2.5.15}$$

Now, using (2.5.13), (2.5.14) and (2.5.15), we obtain the relation (2.5.12). □

Obviously, if f is a finite step function on $[a,b]$, then $f' = 0$ on $[a,b] \setminus M$ where $M \subset [a,b]$ is a finite (possibly empty) set. Finite step functions on $[a,b]$ are thus singular on $[a,b]$. We will show that every step function on $[a,b]$ is also singular on $[a,b]$. For this purpose, we need the following statement known as the *little Fubini theorem*.

2.5.4 Theorem ("LITTLE" FUBINI). *Let $\{f_k\}$ be a sequence of nondecreasing functions on $[a,b]$ such that the series $f(x) = \sum_{k=1}^{\infty} f_k(x)$ converges for every $x \in [a,b]$. Then $f'(x) = \sum_{k=1}^{\infty} f_k'(x) \in \mathbb{R}$ for almost all $x \in [a,b]$.*

Proof. a) Denote

$$g(x) = f(x) - f(a), \quad g_k(x) = f_k(x) - f_k(a) \quad \text{for } k \in \mathbb{N}, x \in [a,b],$$

and

$$s_n(x) = \sum_{k=1}^{n} g_k(x) \quad \text{for } n \in \mathbb{N}, x \in [a,b].$$

Then the functions $g, g_k,\ k \in \mathbb{N}$, are nonnegative and nondecreasing on $[a,b]$, while

$$g(x) = \sum_{k=1}^{\infty} g_k(x) = \lim_{n\to\infty} s_n(x) \quad \text{for } x \in [a,b].$$

By Theorem 2.4.2, for every $k \in \mathbb{N}$ there exists a set $D_k \subset ab$ of zero outer measure such that the function g_k has a finite derivative $g_k'(x)$ for every $x \in [a,b] \setminus D_k$. Similarly, there exists a finite derivative $g'(x)$ for every $x \in [a,b] \setminus D$ where $D \subset [a,b]$ has zero outer measure. Thus, if we let $\widetilde{D} = D \cup \bigcup_{k=1}^{\infty} D_k$, we can summarize that there exist finite derivatives $g'(x)$, $g_k'(x)$ for each each $k \in \mathbb{N}$ and each $x \in [a,b] \setminus \widetilde{D}$. By Exercise 2.4.1 (iv) the set $\widetilde{D}$ also has zero outer measure.

For any $x \in [a,b] \setminus \widetilde{D}$ and $\xi \in [a,b]$ such that $\xi \neq x$, we have

$$\sum_{k=1}^{\infty} \frac{g_k(\xi) - g_k(x)}{\xi - x} = \frac{g(\xi) - g(x)}{\xi - x}.$$

Since every term in the sum on the left-hand side is nonnegative (because g_k is nondecreasing), it follows that

$$\frac{s_n(\xi) - s_n(x)}{\xi - x} = \sum_{k=1}^{n} \frac{g_k(\xi) - g_k(x)}{\xi - x} \leq \frac{g(\xi) - g(x)}{\xi - x}$$

holds for any $x \in [a,b]$, $\xi \in [a,b] \setminus \{x\}$ and $n \in \mathbb{N}$. Letting $\xi \to x$ we get

$$s_n'(x) = \sum_{k=1}^{n} g_k'(x) \leq g'(x) \quad \text{for } x \in [a,b] \setminus \widetilde{D} \text{ and } n \in \mathbb{N}.$$

Since $g_k'(x) \geq 0$ for $x \in [a,b] \setminus \widetilde{D}$ and $k \in \mathbb{N}$, the sequence $\{s_n'(x)\}$ is bounded and nondecreasing for each $x \in [a,b] \setminus \widetilde{D}$. Thus, for every $x \in [a,b] \setminus \widetilde{D}$ there exists a finite limit

$$\lim_{n\to\infty} s_n'(x) = \sum_{k=1}^{\infty} g_k'(x) \leq g'(x), \tag{2.5.16}$$

i.e., the series $\sum_{k=1}^{\infty} g_k'(x)$ converges for almost all $x \in [a,b]$.

b) On the other hand, for every $\ell \in \mathbb{N}$ there exists n_ℓ such that

$$0 \leq g(b) - s_{n_\ell}(b) < \frac{1}{2^\ell}.$$

Since both g and s_{n_ℓ} are nondecreasing on $[a,b]$, it follows that

$$0 \leq g(x) - s_{n_\ell}(x) = \sum_{k=n_\ell+1}^{\infty} g_k(x) \leq \sum_{k=n_\ell+1}^{\infty} g_k(b) = g(b) - s_{n_\ell}(b) < \frac{1}{2^\ell}$$

and hence also

$$0 \leq \sum_{\ell=1}^{\infty} \big(g(x) - s_{n_\ell}(x)\big) \leq \sum_{\ell=1}^{\infty} \frac{1}{2^\ell} = 1 \quad \text{for } x \in [a,b].$$

Repeating the considerations from part a) with $\{g_k\}$ replaced by $\big\{g(x) - s_{n_\ell}(x)\big\}$ we deduce that the series $\sum_{\ell=1}^{\infty} \big(g'(x) - s_{n_\ell}'(x)\big)$ is convergent for almost all $x \in [a,b]$. In particular,

$$\lim_{\ell\to\infty} \big(g'(x) - s_{n_\ell}'(x))\big) = 0 \quad \text{for almost all } x \in [a,b].$$

Naturally, this could not be true if the inequality in (2.5.16) was strict. Consequently,

$$f'(x) = (f(x) - f(a))' = g'(x) = \sum_{k=1}^{\infty} g_k'(x)$$
$$= \sum_{k=1}^{\infty} \big(f_k(x) - f(a)\big)' = \sum_{k=1}^{\infty} f_k'(x) \quad \text{for almost all } x \in [a,b].$$

This completes the proof. □

2.5.5 Theorem. *Every step function on* $[a,b]$ *is singular on* $[a,b]$.

Proof. Let f be a step function on $[a,b]$. If $f \in \mathrm{S}([a,b])$, the statement of the theorem is obvious. Otherwise, f has the form (2.5.3) and (2.5.2) holds. Define

$$v_a(x) = \begin{cases} 0 & \text{if } x = a, \\ |c_0| & \text{if } x > a, \end{cases} \qquad v_b(x) = \begin{cases} 0 & \text{if } x < b, \\ |d| & \text{if } x = b, \end{cases}$$

and

$$v_k(x) = \left\{ \begin{array}{cl} 0 & \text{if } x < s_k, \\ |d_k| & \text{if } x = s_k, \\ |c_k| + |d_k| & \text{if } x > s_k \end{array} \right\} \qquad \text{and } k \in \mathbb{N}.$$

All these functions are nondecreasing on $[a,b]$,

$$v_a'(x) = 0 \text{ for } x \neq a, \quad v_b'(x) = 0 \text{ for } x \neq b,$$
$$v_k'(x) = 0 \text{ for } k \in \mathbb{N} \text{ and } x \neq s_k.$$

Moreover, by (2.5.2) we have

$$\sum_{k=1}^{\infty} |v_k(x)| \le \sum_{k=1}^{\infty} \big(|c_k| + |d_k|\big) < \infty \quad \text{for } x \in [a,b].$$

Thus, the series $\sum_{k=1}^{\infty} v_k(x)$ is absolutely convergent for each $x \in [a,b]$ and the function

$$v(x) = v_a(x) + \sum_{k=1}^{\infty} v_k(x) + v_b(x)$$

is well defined for each $x \in [a,b]$. In view of Theorem 2.5.4, we have

$$v'(x) = v_a'(x) + \sum_{k=1}^{\infty} v_k'(x) + v_b'(x) = 0 \quad \text{for almost all } x \in [a,b].$$

Now, since

$$\left|\frac{f(x)-f(y)}{x-y}\right| \le \left|\frac{v(x)-v(y)}{x-y}\right|$$

holds for all $x, y \in [a,b]$ such that $x \ne y$, it follows easily that $f'(x) = 0$ for almost all $x \in [a,b]$. □

2.5.6 Remark. A well-known example of a continuous, nondecreasing and singular function is the so-called Cantor function; see e.g. [43], pages 14–15.

Next two assertions are interesting in the context of Remark 2.1.18 and Lemma 2.1.19.

2.5.7 Proposition. *Let f be a step function on $[a,b]$ and let $g \in \mathrm{BV}([a,b])$ be continuous on $[a,b]$. Then* $\operatorname{var}_a^b(f+g) = \operatorname{var}_a^b f + \operatorname{var}_a^b g$.

Proof. We will verify that the assumptions of Lemma 2.1.19 are satisfied.

Let f be given as in Definition 2.5.2, where $K = \mathbb{N}$ and $D = \{s_k\}$ is the set of discontinuity points of f in (a,b). By Theorem 2.5.3, we have

$$|\Delta^+ f(a)| + \sum_{k=1}^{\infty} \big(|\Delta^- f(s_k)| + |\Delta^+ f(s_k)|\big) + |\Delta^- f(b)| = \operatorname{var}_a^b f < \infty.$$

Let $\varepsilon > 0$ be given and let $n \in \mathbb{N}$ be such that

$$|\Delta^+ f(a)| + \sum_{k=1}^{n} \big(|\Delta^- f(s_k)| + |\Delta^+ f(s_k)|\big) + |\Delta^- f(b)| > \operatorname{var}_a^b f - \frac{\varepsilon}{2}.$$

Let $x_0, x_1, \dots, x_n, x_{n+1}$ be such that

$$\{s_k\}_{k=1}^n = \{x_k\}_{k=1}^n \quad \text{and} \quad a = x_0 < x_1 < \dots < x_n < x_{n+1} = b.$$

Further, choose points a_k, b_k, $k \in \{0, 1, \dots, n, n+1\}$, in such a way that

$$a = x_0 = a_0 < b_0 < a_1 < x_1 < b_1 < \dots < a_n < x_n < b_n < a_{n+1} < x_{n+1} = b_{n+1} = b,$$

$$|f(b_k) - f(x_k)| > |\Delta^+ f(x_k)| - \frac{\varepsilon}{4(n+2)} \quad \text{for } k \in \{0, 1, \dots, n\}$$

and

$$|f(x_k) - f(a_k)| > |\Delta^- f(x_k)| - \frac{\varepsilon}{4(n+2)} \quad \text{for } k \in \{1, \dots, n, n+1\}.$$

Furthermore, as $v_g(x) = \operatorname{var}_a^x g$ is continuous on $[a,b]$ by Corollary 2.3.4, we can also assume that

$$v_g(b_k) - v_g(a_k) < \frac{\varepsilon}{n+2} \quad \text{for all } k \in \{0, \dots, n+1\}.$$

To summarize, we have

$$\sum_{k=0}^{n+1} \operatorname{var}_{a_k}^{b_k} g < \varepsilon$$

and

$$\sum_{k=0}^{n+1} \operatorname{var}_{a_k}^{b_k} f \geq \sum_{k=0}^{n+1} \left(|f(b_k) - f(x_k)| + |f(x_k) - f(a_k)|\right)$$
$$> \sum_{k=0}^{n+1} \left(|\Delta^+ f(x_k)| + |\Delta^- f(x_k)| - \frac{\varepsilon}{2(n+2)}\right) > \operatorname{var}_a^b f - \varepsilon.$$

Thus, the assumptions of Lemma 2.1.19 and the proof is complete. □

2.6 Jordan decomposition of a function of bounded variation

2.6.1 Theorem. *For each* $f \in \mathrm{BV}([a,b])$ *there are* $f_1 \in \mathrm{BV}([a,b]) \cap \mathrm{C}([a,b])$ *and* $f_2 \in \mathrm{B}([a,b])$ *such that* $f = f_1 + f_2$ *on* $[a,b]$.

If $f = \widetilde{f}_1 + \widetilde{f}_2$ *is another decomposition with* $\widetilde{f}_1 \in \mathrm{C}([a,b]) \cap \mathrm{BV}([a,b])$ *and* $\widetilde{f}_2 \in \mathrm{B}([a,b])$, *then the functions* $f_1 - \widetilde{f}_1$ *and* $f_2 - \widetilde{f}_2$ *are constant on* $[a,b]$.

Proof. a) Let D be the set of all discontinuity points of f in the open interval (a,b). The set D contains at most countably many points, i.e.,

$$D = \{s_k \in (a,b) \colon k \in K\}, \text{ where } K = \{1,\dots,m\} \text{ for some } m \in \mathbb{N} \text{ or } K = \mathbb{N}.$$

Define

$$\left.\begin{aligned} f_2(x) &= f(a) + \Delta^+ f(a)\, \chi_{(a,b]}(x) \\ &\quad + \sum_{k \in K} \left(\Delta^+ f(s_k)\, \chi_{(s_k,b]}(x) + \Delta^- f(s_k)\, \chi_{[s_k,b]}(x)\right) \\ &\quad + \Delta^- f(b)\, \chi_{[b]}(x) \quad \text{for } x \in [a,b]. \end{aligned}\right\} \tag{2.6.1}$$

By Corollary 2.3.8 we have

$$|\Delta^+ f(a)| + \sum_{k \in \mathbb{K}} \left(|\Delta^+ f(s_k)| + |\Delta^- f(s_k)|\right) + |\Delta^- f(b)| \leq \operatorname{var}_a^b f < \infty$$

and thus $f_2 \in \mathrm{B}([a,b])$. Further, using (2.5.9) we get

$$\left.\begin{aligned} &\Delta^+ f_2(t) = \Delta^+ f(t) \text{ and } \Delta^- f_2(s) = \Delta^- f(s) \\ &\qquad\qquad \text{for } t \in [a,b) \text{ and } s \in (a,b]. \end{aligned}\right\} \tag{2.6.2}$$

Hence

$$\big(f(t+)-f_2(t+)\big)-\big(f(t)-f_2(t)\big)=\Delta^+f(t)-\Delta^+f_2(t)=0 \text{ for } t\in[a,b)$$

and

$$\big(f(s)-f_2(s)\big)-\big(f(s-)-f_2(s-)\big)=\Delta^-f(s)-\Delta^-f_2(s)=0 \text{ for } s\in(a,b].$$

Thus, the function $f_1=f-f_2$ is continuous on $[a,b]$ and $f=f_1+f_2$ on $[a,b]$.

b) Let $f=\widetilde{f}_1+\widetilde{f}_2$ where $\widetilde{f}_1\in\mathrm{C}([a,b])\cap\mathrm{BV}([a,b])$ and $\widetilde{f}_2\in\mathrm{B}([a,b])$. Then the relations

$$0=\Delta^+\widetilde{f}_1(t)=\big(f(t+)-\widetilde{f}_2(t+)\big)-\big(f(t)-\widetilde{f}_2(t)\big)=\Delta^+f(t)-\Delta^+\widetilde{f}_2(t)$$

and

$$0=\Delta^-\widetilde{f}_1(s)=\big(f(s)-\widetilde{f}_2(s)\big)-\big(f(s-)-\widetilde{f}_2(s-)\big)=\Delta^-f(s)-\Delta^-\widetilde{f}_2(s)$$

hold for all $t\in[a,b)$ and $s\in(a,b]$. Using (2.6.2), we see that

$$\left.\begin{aligned}&\Delta^+\widetilde{f}_2(t)=\Delta^+f_2(t)=\Delta^+f(t) \qquad &&\text{for } t\in[a,b)\\ &\text{and}\\ &\Delta^-\widetilde{f}_2(s)=\Delta^-f_2(s)=\Delta^-f(s) \qquad &&\text{for } s\in(a,b].\end{aligned}\right\} \tag{2.6.3}$$

Since $\widetilde{f}_2$ is a step function whose discontinuities are contained in D, there exist real numbers $\widetilde{c}$, $\widetilde{c}_0$, $\widetilde{d}$ and sequences $\{\widetilde{c}_k\}$, $\{\widetilde{d}_k\}$ such that

$$\widetilde{f}_2(x)=\widetilde{c}+\widetilde{c}_0\,\chi_{(a,b]}(x)+\sum_{k\in K}\Big(\widetilde{c}_k\,\chi_{(s_k,b]}(x)+\widetilde{d}_k\,\chi_{[s_k,b]}(x)\Big)+\widetilde{d}\,\chi_{[b]}(x)$$

for all $x\in[a,b]$, where

$$\begin{aligned}&\sum_{k\in K}(|\widetilde{c}_k|+|\widetilde{d}_k|)<\infty,\\ &\Delta^+\widetilde{f}_2(s_k)=\widetilde{c}_k \ \text{ for } k\in\mathbb{N}, \quad \Delta^+\widetilde{f}_2(a)=\widetilde{c}_0,\\ &\Delta^-\widetilde{f}_2(s_k)=\widetilde{d}_k \ \text{ for } k\in\mathbb{N}, \quad \Delta^-\widetilde{f}_2(b)=d.\end{aligned}$$

Using (2.6.3), we have

$$\begin{aligned}\widetilde{f}_2(x)&=\widetilde{c}+\Delta^+f(a)\,\chi_{(a,b]}(x)+\Delta^-f(b)\,\chi_{[b]}(x)\\ &\qquad+\sum_{k\in\mathbb{K}}\Big(\Delta^+f(s_k)\,\chi_{(s_k,b]}(x)+\Delta^-f(s_k)\,\chi_{[s_k,b]}(x)\Big)\\ &=(\widetilde{c}-f(a))+f_2(x) \quad \text{for } x\in[a,b].\end{aligned}$$

It follows that on the whole interval $[a,b]$, $\widetilde{f}_2 - f_2$ is equal to the constant $\varkappa := \widetilde{c} - f(a)$. Hence also

$$\begin{aligned} f_1(x) - \widetilde{f}_1(x) &= (f(x) - f_2(x)) - (f(x) - \widetilde{f}_2(x)) \\ &= \widetilde{f}_2(x) - f_2(x) = \varkappa \quad \text{for } x \in [a,b]. \end{aligned}$$

□

2.6.2 Remark. By Theorem 2.6.1 every function of bounded variation can be decomposed into the sum of a continuous function and a step function. Such a decomposition is called the *Jordan decomposition* of a function of bounded variation.

2.6.3 Definition. Every function f_2 assigned to f by Theorem 2.6.1 is called the *jump part* of the function f. The difference $f - f_2$ is called the *continuous part* of the function f. The jump part and the continuous part of the function f are usually denoted by f^{B} and f^{C}, respectively.

2.6.4 Exercises. Let $\{f_n\}$ be a sequence of functions with bounded variations on $[a,b]$ and let $\{f_n^{\mathrm{C}}\}$ and $\{f_n^{\mathrm{B}}\}$ be the sequences of continuous and jump parts of $\{f_n\}$, respectively.

(i) Show that $\operatorname{var}_a^b f_n = \operatorname{var}_a^b f_n^{\mathrm{C}} + \operatorname{var}_a^b f_n^{\mathrm{B}}$ for each $n \in \mathbb{N}$.

(ii) Prove that $\lim\limits_{n\to\infty} \operatorname{var}_a^b f_n = 0$ if and only if

$$\lim_{n\to\infty} (\operatorname{var}_a^b f_n^{\mathrm{C}}) = \lim_{n\to\infty} (\operatorname{var}_a^b f_n^{\mathrm{B}}) = 0.$$

For dealing with step functions, it is useful to know that every step function may be approximated in the norm of the space $\mathrm{BV}([a,b])$ by finite step functions. This is the content of the following lemma, which will be particularly useful in Chapter 6.

2.6.5 Lemma. *For each step function* $f \in \mathrm{B}([a,b])$ *there is a sequence* $\{f_n\} \subset \mathrm{S}([a,b])$ *of finite step functions such that*

$$\lim_{n\to\infty} \|f - f_n\|_{\mathrm{BV}} = 0. \tag{2.6.4}$$

Proof. Let $f \in \mathrm{B}([a,b])$. If the set D of its discontinuity points in (a,b) is finite, then $f \in \mathrm{S}([a,b])$ and the assertion of the lemma is obvious. Therefore assume that $D = \{s_k\}$ is a non-repeating infinite sequence. By Theorem 2.3.6 an Corollary 2.3.8, the series

$$\sum_{k=1}^{\infty} \Big(\Delta^+ f(s_k)\, \chi_{(s_k,b]}(x) + \Delta^- f(s_k)\, \chi_{[s_k,b]}(x) \Big) \tag{2.6.5}$$

is absolutely convergent for $x \in [a,b]$. Therefore (cf. (2.5.3) and (2.5.9))

$$\left. \begin{aligned} f(x) = f(a) + \Delta^+ f(a)\, \chi_{(a,b]}(x) \\ + \sum_{k=1}^{\infty} \Big(\Delta^+ f(s_k)\, \chi_{(s_k,b]}(x) + \Delta^- f(s_k)\, \chi_{[s_k,b]}(x) \Big) \\ + \Delta^- f(b)\, \chi_{[b]}(x) \quad \text{for } x \in [a,b]. \end{aligned} \right\} \tag{2.6.6}$$

Define

$$\left. \begin{aligned} f_n(x) = f(a) + \Delta^+ f(a)\, \chi_{(a,b]}(x) \\ + \sum_{k=1}^{n} \Big(\Delta^+ f(s_k)\, \chi_{(s_k,b]}(x) + \Delta^- f(s_k)\, \chi_{[s_k,b]}(x) \Big) \\ + \Delta^- f(b)\, \chi_{[b]}(x) \quad \text{for } x \in [a,b] \ \text{ and } \ n \in \mathbb{N}. \end{aligned} \right\} \tag{2.6.7}$$

Then $f_n \in \mathrm{S}([a,b])$ for each $n \in \mathbb{N}$. Moreover, for each $n \in \mathbb{N}$ we have

$$f(a) = f_n(a), \tag{2.6.8}$$

and

$$f(x) - f_n(x) = \sum_{k=n+1}^{\infty} \Big(\Delta^+ f(s_k)\, \chi_{(s_k,b]}(x) + \Delta^- f(s_k)\, \chi_{[s_k,b]}(x) \Big) \text{ for } x \in [a,b].$$

Now, by Theorem 2.5.3 we have

$$\operatorname{var}_a^b(f - f_n) = \sum_{k=n+1}^{\infty} \big(|\Delta^- f(s_k)| + |\Delta^+ f(s_k)| \big). \tag{2.6.9}$$

Since the right hand side of (2.6.9) is the remainder of an absolutely convergent series, it converges to 0 as $n \to \infty$. This means that $\lim_{n\to\infty} \operatorname{var}_a^b(f - f_n) = 0$ and hence (2.6.4) holds due to (2.6.8). This completes the proof. □

2.7 Pointwise convergence

2.7.1 Example. Consider again the functions

$$f_n(x) = \begin{cases} 0 & \text{if } 0 \le x \le \frac{1}{n}, \\ x \sin\left(\dfrac{\pi}{x}\right) & \text{if } \frac{1}{n} \le x \le 2 \end{cases} \qquad \text{for } n \in \mathbb{N}$$

and

$$f(x) = \begin{cases} 0 & \text{if } x = 0, \\ x \sin\left(\dfrac{\pi}{x}\right) & \text{if } x > 0. \end{cases}$$

From Example 2.1.15 we know that $\operatorname{var}_0^2 f_n < \infty$ for each $n \in \mathbb{N}$. It can be easily verified that $\{f_n\}$ converges to f uniformly on $[0,2]$, while by Example 2.1.10, f has unbounded variation on $[0,2]$.

The previous example shows that the limit of bounded variation functions need not have bounded variation even if the convergence is uniform. On the other hand, the following theorem shows that uniform boundedness of variations of f_n together with the pointwise convergence $f_n \to f$ already guarantee that the limit function f has bounded variation. (Notice that, using the argument from Example 2.1.10, it is possible to show that the sequence $\{f_n\}$ from Examples 2.1.15 and 2.7.1 satisfies $\lim_{n\to\infty} \operatorname{var}_0^2 f_n = \infty$, and therefore $\sup_{n\in\mathbb{N}} \operatorname{var}_0^2 f_n = \infty$.)

2.7.2 Theorem. *Let $f : [a,b] \to \mathbb{R}$ be given and let $\{f_n\}$ be a sequence of functions such that*

$$\operatorname{var}_a^b f_n \le \varkappa < \infty \quad \textit{for} \quad n \in \mathbb{N}, \quad \textit{and} \quad \lim_{n\to\infty} f_n(x) = f(x) \quad \textit{for } x \in [a,b].$$

Then $\operatorname{var}_a^b f \le \varkappa$.

Proof. Given an arbitrary $\boldsymbol{\alpha} \in \mathscr{D}[a,b]$, we have

$$V(f,\boldsymbol{\alpha}) = \lim_{n\to\infty} V(f_n,\boldsymbol{\alpha}) \le \varkappa.$$

Consequently, $\operatorname{var}_a^b f \le \varkappa$. □

2.7.3 Exercise. Let

$$f(x) = \begin{cases} 2^{-k} & \text{if } x = \frac{1}{k+1} \text{ for some } k \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Prove that $f \in \mathrm{BV}([0,1])$.

We now formulate and prove Helly's Choice Theorem, which will be useful e.g. in the proof of compactness of certain operators defined on the space $\mathrm{BV}([a,b])$. The theorem states that every sequence of functions with uniformly bounded variations has a subsequence which is pointwise convergent to a function of bounded variation.

2.7.4 Theorem (HELLY'S CHOICE THEOREM)**.** *Let $\{f_n\} \subset \mathrm{BV}([a,b])$, $\varkappa \in \mathbb{R}$, $|f_n(a)| \le \varkappa$ and $\operatorname{var}_a^b f_n \le \varkappa$ for all $n \in \mathbb{N}$.*

Then there exist a function $f \in \mathrm{BV}([a,b])$ and a subsequence $\{n_k\}$ of $\mathbb{N}$ such that

$$|f(a)| \le \varkappa, \quad \operatorname{var}_a^b f \le \varkappa \quad \textit{and} \quad \lim_{k\to\infty} f_{n_k}(x) = f(x) \quad \textit{for} \quad x \in [a,b].$$

To prove Theorem 2.7.4 we need the following two assertions.

Assertion 1. *Let* $|f_n(x)| \le M < \infty$ *for all* $x \in [a,b]$ *and all* $n \in \mathbb{N}$. *Then for any countable subset* P *of* $[a,b]$, *there is a subsequence* $\{n_k\}$ *of* $\mathbb{N}$ *such that* $\lim_{k\to\infty} f_{n_k}(p)$ *exists and is finite for all* $p \in P$.

Proof. Let $P = \{p_k\}$. We have $|f_n(p_k)| \le M < \infty$ for all $n, k \in \mathbb{N}$. Hence, by the Bolzano-Weierstrass theorem, there is a sequence $\{n_{k,1} : k \in \mathbb{N}\}$ and a number $q_1 \in \mathbb{R}$ such that

$$\lim_{k\to\infty} f_{n_{k,1}}(p_1) = q_1.$$

Similarly, there are $\{f_{n_{k,2}} : k \in \mathbb{N}\} \subset \{f_{n_{k,1}} : k \in \mathbb{N}\}$ and $q_2 \in \mathbb{R}$ such that

$$\lim_{k\to\infty} f_{n_{k,2}}(p_2) = q_2 \in \mathbb{R}, \text{ and } \lim_{k\to\infty} f_{n_{k,2}}(p_1) = q_1 \in \mathbb{R}.$$

In this way, for each $j \in \{2, 3, \dots\}$ we can find a sequence

$$\{f_{n_{k,j}} : k \in \mathbb{N}\} \subset \{f_{n_{k,j-1}} : k \in \mathbb{N}\}$$

and a number $q_j \in \mathbb{R}$ such that

$$\lim_{k\to\infty} f_{n_{k,\ell}}(p_\ell) = q_\ell \in \mathbb{R} \quad \text{for all } \ell \in \{1, \dots, j\}.$$

Put $f_{n_k} = f_{n_{k,k}}$ for $k \in \mathbb{N}$. Then

$$\lim_{k\to\infty} f_{n_k}(p_j) = q_j \in \mathbb{R} \quad \text{for } j \in \mathbb{N}.$$

□

Assertion 2. *Let functions* f_n, $n \in \mathbb{N}$, *be nondecreasing on* $[a,b]$ *and let* $M \in [0,\infty)$ *be such that* $||f_n||_\infty \le M$ *for all* $n \in \mathbb{N}$. *Then there is a subsequence* $\{n_k\}$ *of* $\mathbb{N}$ *and a nondecreasing function* $f : [a,b] \to \mathbb{R}$ *such that*

$$\lim_{k\to\infty} f_{n_k}(x) = f(x) \quad \textit{for } x \in [a,b].$$

Proof. Let $P = \big(\mathbb{Q} \cap (a,b)\big) \cup \{a\} \cup \{b\}$ be the set of all rational numbers from the open interval (a,b) together with the points a, b. Then P is countable and $[a,b] \setminus P \subset (a,b)$. By Assertion 1 there is a subsequence $\{n_k\} \subset \mathbb{N}$ and a function $\varphi : P \to \mathbb{R}$ such that

$$\lim_{k\to\infty} f_{n_k}(p) = \varphi(p) \quad \text{for } p \in P.$$

Obviously, $\varphi(p') \le \varphi(p'')$ if $p', p'' \in P$ and $p' \le p''$. Furthermore, define

$$\varphi(x) = \sup_{p \in P \cap [a,x)} \varphi(p) \text{ for } x \in (a,b) \setminus P.$$

The function φ is nondecreasing on $[a,b]$ and

$$\varphi(x) = \lim_{\substack{p \to x- \\ p \in P}} \varphi(p) \quad \text{for } x \in (a,b) \setminus P.$$

We will show that

$$\lim_{k\to\infty} f_{n_k}(x_0) = \varphi(x_0) \tag{2.7.1}$$

whenever $x_0 \in (a,b)$ and φ is continuous at x_0. Indeed, assume that $x_0 \in (a,b)$ is a continuity point of φ and let an arbitrary $\varepsilon > 0$ be given. Then there is a $\delta_\varepsilon > 0$ such that

$$\varphi(x_0) - \varepsilon < \varphi(x) < \varphi(x_0) + \varepsilon \quad \text{for all } x \in (x_0 - \delta_\varepsilon, x_0 + \delta_\varepsilon).$$

Further, let us choose $r' \in P \cap (x_0 - \delta_\varepsilon, x_0)$ and $r'' \in P \cap (x_0, x_0 + \delta_\varepsilon)$ arbitrarily. Then

$$\varphi(x_0) - \varepsilon < \varphi(r') \le \varphi(x_0) \le \varphi(r'') < \varphi(x_0) + \varepsilon.$$

Moreover, there is a k_ε such that

$$\varphi(r') - \varepsilon < f_{n_k}(r') < \varphi(r') + \varepsilon$$

and

$$\varphi(r'') - \varepsilon < f_{n_k}(r'') < \varphi(r'') + \varepsilon$$

for all $k \ge k_\varepsilon$. Hence, for each $k \ge k_\varepsilon$ we have

$$\begin{aligned} \varphi(x_0) - 2\,\varepsilon < \varphi(r') - \varepsilon < f_{n_k}(r') \le f_{n_k}(x_0) \\ \le f_{n_k}(r'') < \varphi(r'') + \varepsilon < \varphi(x_0) + 2\,\varepsilon. \end{aligned}$$

Thus, (2.7.1) is true.

To summarize, we have proved that if D is the set of all discontinuity points of the function φ in (a,b), then

$$\lim_{k\to\infty} f_{n_k}(x) = \varphi(x) \quad \text{for } x \in [a,b] \setminus D.$$

By Theorem 2.3.2, the set D is countable. Thus, we can use once more Assertion 1 to prove that there is a subsequence

$$\{f_{n_{k_\ell}} : \ell \in \mathbb{N}\} \subset \{f_{n_k} : k \in \mathbb{N}\}$$

of $\{f_{n_k}\}$ which has a limit $\psi(x) \in \mathbb{R}$ for each $x \in D$. Now, define

$$f(x) = \begin{cases} \varphi(x), & \text{if } x \in [a,b] \setminus D, \\ \psi(x), & \text{if } x \in D. \end{cases}$$

Then

$$\lim_{\ell\to\infty} f_{n_{k_\ell}}(x) = f(x) \quad \text{for } x \in [a,b],$$

and f is nondecreasing on $[a,b]$ because it is the pointwise limit of a sequence of nondecreasing functions. The proof of Assertion 2 is complete. □

Proof of Theorem 2.7.4

For given $n \in \mathbb{N}$ and $x \in [a,b]$, let

$$g_n(x) = \operatorname{var}_a^x f_n \quad \text{and} \quad h_n(x) = g_n(x) - f_n(x).$$

For each $n \in \mathbb{N}$ we have $f_n = g_n - h_n$ and the functions g_n, h_n with $n \in \mathbb{N}$ are nondecreasing on $[a,b]$ (see Exercise 2.1.22). Furthermore,

$$\|g_n\|_\infty \le \operatorname{var}_a^b f_n \le \varkappa \quad \text{and} \quad \|h_n\|_\infty \le \|f_n\|_\infty + \|g_n\|_\infty \le 3\,\varkappa \quad \text{for } n \in \mathbb{N}.$$

By Assertion 2, there exist functions $g, h \in \mathrm{BV}([a,b])$ and a sequence $\{n_k\} \subset \mathbb{N}$ such that

$$\|g\|_\infty \le \varkappa,\ \ \|h\|_\infty \le 2\,\varkappa,\ \ \operatorname{var}_a^b g \le \varkappa,\ \ \operatorname{var}_a^b h \le 2\,\varkappa,$$

$$\lim_{k\to\infty} g_{n_k}(x) = g(x) \quad \text{and} \quad \lim_{k\to\infty} h_{n_k}(x) = h(x) \quad \text{for all } x \in [a,b].$$

Denote $f = g - h$. Then

$$\lim_{k\to\infty} f_{n_k}(x) = \lim_{k\to\infty} \big(g_{n_k}(x) - h_{n_k}(x)\big) = g(x) - h(x) = f(x)$$

for all $x \in [a,b]$. Obviously, $|f(a)| \le \varkappa$. Finally, Theorem 2.7.2 implies that $\operatorname{var}_a^b f \le \varkappa$. This completes the proof. □

2.8 Variation on elementary sets

First, motivated by Definition 6.1 from [44], we introduce the definition of the variation over arbitrary intervals.

2.8.1 Definition. Let J be a bounded interval in $\mathbb{R}$. We say that a finite set

$$\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \ldots, \alpha_{\nu(\boldsymbol{\alpha})}\} \subset J$$

is a *generalized division of* J if $\alpha_0 < \alpha_1 < \cdots < \alpha_{\nu(\boldsymbol{\alpha})}$.

The set of all generalized divisions of the interval J is denoted by $\mathscr{D}^*(J)$.

Let $f:[a,b]\to\mathbb{R}$ and let J be an arbitrary subinterval of $[a,b]$. Then we define the variation of f on J by

$$\operatorname{var}(f,J)=\sup\left\{\sum_{j=1}^{\nu(\boldsymbol{\alpha})}|f(\alpha_j)-f(\alpha_{j-1})|:\boldsymbol{\alpha}\in\mathscr{D}^*(J)\right\}.$$

We say that f is of bounded variation on J if $\operatorname{var}(f,J)<\infty$. In such a case, we write $f\in\mathrm{BV}(J)$. We also set $\operatorname{var}(f,\emptyset)=0$ and $\operatorname{var}(f,[c])=0$ for $c\in[a,b]$.

2.8.2 Remark. It is easy to see that Definition 2.8.1 coincides with the definition of the variation in the sense of Definition 2.1.1 if J is a compact interval; that is, for $f:[a,b]\to\mathbb{R}$ and $J=[c,d]\subset[a,b]$, we have $\operatorname{var}(f,J)=\operatorname{var}_c^d f$. For this reason, in the case of a compact interval J, we may always restrict ourselves to divisions containing the endpoints of J.

Moreover, it is easy to see that if J is a bounded interval and $f\in\mathrm{BV}(J)$, then f is bounded on J.

The next proposition follows immediately from Definition 2.8.1.

2.8.3 Proposition. *Let $f:[a,b]\to\mathbb{R}$ and let J_1 and J_2 be subintervals of $[a,b]$ such that $J_2\subset J_1$. Then* $\operatorname{var}(f,J_2)\le\operatorname{var}(f,J_1)$.

In particular, if J is a subinterval of $[a,b]$ and $f\in\mathrm{BV}(J)$, then $f\in\mathrm{BV}(I)$ for every interval $I\subset J$.

The next theorem presents formulas for the variation over half-open and open intervals.

2.8.4 Theorem. *Let $f:[a,b]\to\mathbb{R}$ and $c,\ d\in[a,b]$, with $c<d$.*

(i) *If $f\in\mathrm{BV}([c,d))$, then*

$$\operatorname{var}(f,[c,d))=\lim_{\delta\to0+}\operatorname{var}_c^{d-\delta}f=\sup_{t\in[c,d)}\operatorname{var}_c^t f.$$

(ii) *If $f\in\mathrm{BV}((c,d])$, then*

$$\operatorname{var}(f,(c,d])=\lim_{\delta\to0+}\operatorname{var}_{c+\delta}^{d}f=\sup_{t\in(c,d]}\operatorname{var}_t^d f.$$

(iii) *If $f\in\mathrm{BV}((c,d))$, then*

$$\operatorname{var}(f,(c,d))=\lim_{\delta\to0+}\operatorname{var}_{c+\delta}^{d-\delta}f.$$

Proof. We prove only the assertion (i); the other ones follow in a similar way.

For a fixed $\delta > 0$, consider a division $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \ldots, \alpha_{\nu(\boldsymbol{\alpha})}\}$ of $[c, d-\delta]$. Of course, $\boldsymbol{\alpha}$ is also a generalized division of $[c,d)$ and hence

$$\sum_{j=1}^{\nu(\boldsymbol{\alpha})} |f(\alpha_j) - f(\alpha_{j-1})| \le \text{var}\,(f, [c,d)).$$

Thus, taking the supremum over all divisions of $[c, d-\delta]$, we get

$$\text{var}_c^{d-\delta} f \le \text{var}\,(f, [c,d)).$$

Since this inequality holds for every $\delta > 0$, it follows that

$$M := \sup_{t \in [c,d)} \text{var}_c^t f = \lim_{\delta \to 0+} \text{var}_c^{d-\delta} f \le \text{var}\,(f, [c,d)).$$

Now, assume that $M < \text{var}\,(f, [c,d))$. Then, for $\varepsilon = \text{var}\,(f, [c,d)) - M$, there exists a generalized division $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \ldots, \alpha_{\nu(\boldsymbol{\alpha})}\}$ of $[c,d)$ such that

$$M = \text{var}\,(f, [c,d)) - \varepsilon < \sum_{j=1}^{\nu(\boldsymbol{\alpha})} |f(\alpha_j) - f(\alpha_{j-1})| \le \text{var}_c^{\alpha_{\nu(\boldsymbol{\alpha})}} f \le M,$$

a contradiction. This completes the proof of (i). □

Dealing with functions taking values in a metric space, Chistyakov presents in [21] an extensive study of the properties of the variation over subsets of the real line. Here, we call the reader's attention to a particular result (see [21], Corollary 4.7) connecting the variation over arbitrary intervals and the usual variation over a compact interval. This will be the content of Theorem 2.8.6, whose proof is included for the sake of completeness. To this aim, the next lemma will be useful.

2.8.5 Lemma. *Let* $f : [a,b] \to \mathbb{R}$, $a \le c < d \le b$ *and* $f \in \text{BV}((c,d))$. *Then both the limits* $f(c+)$ *and* $f(d-)$ *exist.*

Proof. Let $\varepsilon > 0$ and an increasing sequence $\{t_n\} \subset (c,d)$ tending to d be given. By Theorem 2.8.4 (iii) there is $\delta > 0$ such that

$$0 < \text{var}\,(f, (c,d)) - \text{var}_{c+\delta}^{d-\delta} f < \varepsilon.$$

Choose $n_0 \in \mathbb{N}$ in such a way that $t_n > d - \delta$ for every $n \ge n_0$. Therefore, for $n > m \ge n_0$, we have

$$\begin{aligned} |f(t_n) - f(t_m)| &\le \text{var}_{t_m}^{t_n} f = \text{var}_{c+\delta}^{t_n} f - \text{var}_{c+\delta}^{t_m} f \\ &\le \text{var}\,(f, (c,d)) - \text{var}_{c+\delta}^{d-\delta} f < \varepsilon, \end{aligned}$$

wherefrom the existence of the limit $f(d-)$ follows immediately.

The existence of the limit $f(c+)$ can be proved analogously. □

2.8.6 Theorem. *Let* $f:[a,b]\to\mathbb{R}$ *and* $a\le c<d\le b$.

(i) *If* $f\in \mathrm{BV}([c,d))$, *then* $f(d-)$ *exists and*

$$\mathrm{var}_c^d f=\mathrm{var}(f,[c,d))+|\Delta^- f(d)|.$$

(ii) *If* $f\in \mathrm{BV}((c,d])$, *then* $f(c+)$ *exists and*

$$\mathrm{var}_c^d f=\mathrm{var}(f,(c,d])+|\Delta^+ f(c)|.$$

(iii) *If* $f\in \mathrm{BV}((c,d))$, *then both the limits* $f(c+)$ *and* $f(d-)$ *exist and*

$$\mathrm{var}_c^d f=\mathrm{var}(f,(c,d))+|\Delta^+ f(c)|+|\Delta^- f(d)|.$$

Proof. The existence of all the limits follows from Lemma 2.8.5.

Assume that $f\in\mathrm{BV}([c,d))$. Let $\varepsilon>0$ and

$$\boldsymbol{\alpha}=\{\alpha_0,\alpha_1,\dots,\alpha_{m+1}\}\in\mathcal{D}[c,d]$$

be given. We can choose $\xi\in[c,d]$ in such a way that

$$\alpha_m<\xi<d \ \text{ and } \ |f(d-)-f(\xi)|<\varepsilon.$$

Consequently,

$$\begin{aligned}\sum_{j=1}^{m+1}&|f(\alpha_j)-f(\alpha_{j-1})|\\ &\le\sum_{j=1}^{m}|f(\alpha_j)-f(\alpha_{j-1})|+|f(\xi)-f(\alpha_m)|\\ &\quad+|f(d-)-f(\xi)|+|\Delta^- f(d)|\\ &<\mathrm{var}_c^\xi f+\varepsilon+|\Delta^- f(d)|\le\mathrm{var}(f,[c,d))+\varepsilon+|\Delta^- f(d)|.\end{aligned}$$

As $\boldsymbol{\alpha}\in\mathcal{D}[c,d]$ and $\varepsilon>0$ were chosen arbitrarily, we conclude that

$$\mathrm{var}_c^d f\le\mathrm{var}(f,[c,d))+|\Delta^- f(d)|. \tag{2.8.1}$$

On the other hand, for any $\delta>0$ we have

$$|f(d)-f(d-\delta)|\le\mathrm{var}_{d-\delta}^d f=\mathrm{var}_c^d f-\mathrm{var}_c^{d-\delta} f.$$

Hence, letting $\delta\to0+$ we get

$$|\Delta^- f(d)|\le\mathrm{var}_c^d f-\lim_{\delta\to0+}\mathrm{var}_c^{d-\delta} f=\mathrm{var}_c^d f-\mathrm{var}(f,[c,d)),$$

wherefrom we conclude that

$$\operatorname{var}_c^d f \geq \operatorname{var}(f,[c,d)) + |\Delta^- f(d)|.$$

This completes the proof of (i).

Similarly, we can prove the assertions (ii) and (iii). □

2.8.7 Corollary. *Let* $f:[a,b]\to\mathbb{R}$ *and* $c,\ d\in[a,b]$, *with* $c<d$. *Then the following assertions are equivalent:*

(i) $f\in \mathrm{BV}([c,d])$,

(ii) $f\in \mathrm{BV}((c,d])$,

(iii) $f\in \mathrm{BV}([c,d))$,

(iv) $f\in \mathrm{BV}((c,d))$.

2.8.8 Remark. In view of Theorem 2.8.6, we can also observe that for every $f:[a,b]\to\mathbb{R}$ and $c\in[a,b]$ we have

$$\lim_{\delta\to 0+} \operatorname{var}_{c-\delta}^{c+\delta} f = |\Delta^- f(c)| + |\Delta^+ f(c)|$$

provided the one-sided limits exist at the point c (see Proposition I.2.8 in [59]). Furthermore, by Theorem 2.8.6, if $f\in \mathrm{BV}([a,b])\cap \mathrm{C}([a,b])$, then

$$\operatorname{var}(f,[c,d)) = \operatorname{var}(f,(c,d)) = \operatorname{var}(f,(c,d]) = \operatorname{var}_c^d f$$

for $c,\ d\in[a,b]$ such that $c<d$.

We now extend the notion of the variation on intervals to elementary sets.

2.8.9 Definition. Let $E\subset\mathbb{R}$ be bounded. We say that E is an *elementary set* if it is a finite union of intervals.

A collection of intervals $\{J_k : k=1,\dots,N\}$ is called a *minimal decomposition* of E if $E=\bigcup_{k=1}^N J_k$ and the union $J_k\cup J_\ell$ is not an interval whenever $k\neq\ell$.

If $S\subset\mathbb{R}$, then $\mathcal{E}(S)$ stands for the set of all elementary subsets of S.

Note that the minimal decomposition of an elementary set is uniquely determined. Moreover, the intervals of such decomposition are pairwise disjoint. Having this in mind, we define the variation over elementary sets as follows.

2.8.10 Definition. 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 $\{J_k : k=1,\dots,N\}$ is the minimal decomposition of E.

It is worth highlighting that if $f \in \mathrm{BV}([a,b]) \cap \mathrm{C}([a,b])$, then $\operatorname{var}(f,\,\cdot\,)$ defines a finitely additive measure on $\mathcal{E}([a,b])$. More precisely, we have

$$\operatorname{var}(f,E) \le \operatorname{var}_a^b f \quad \text{for any} \quad E \in \mathcal{E}([a,b])$$

and

$$\operatorname{var}(f,E_1 \cup E_2) = \operatorname{var}(f,E_1) + \operatorname{var}(f,E_2)$$

whenever $E_1,\ E_2 \in \mathcal{E}([a,b])$ and $E_1 \cap E_2 = \emptyset$.

2.8.11 Remark. Let us note that e.g. in [43] Definition 2.8.1 is applied also to arbitrary subsets E of $[a,b]$. Unfortunately, such a definition is not convenient for our purposes, as the variation would lose the additivity property even for continuous functions. Indeed, let $a<c<d<b$ and $E=[a,c]\cup[d,b]$. Then, according to such a definition we would have

$$\operatorname{var}(f,E) \ge \operatorname{var}_a^c f + \operatorname{var}_d^b f + |f(d)-f(c)| > \operatorname{var}(f,[a,c]) + \operatorname{var}(f,[d,b])$$

whenever $f(d) \ne f(c)$. This is why for elementary subsets of $[a,b]$ we define the variation in a way different from Gordon's in [43].

Chapter 3

Absolutely continuous functions

A special case of functions of bounded variation are absolutely continuous functions, which are closely related to the Lebesgue integration theory and which are well-known from Carathéodory's theory of ordinary differential equations. The integrals contained in this chapter are the Lebesgue ones.

3.1 Definition and basic properties

3.1.1 Definition. A function $f:[a,b]\to\mathbb{R}$ is *absolutely continuous* on the interval $[a,b]$ if for every $\varepsilon>0$ there exists $\delta>0$ such that

$$\sum_{j=1}^{m}|f(\beta_j)-f(\alpha_j)|<\varepsilon \tag{3.1.1}$$

holds for every finite set of intervals $\{[\alpha_j,\beta_j]:j=1,2,,\dots,m\}$ satisfying

$$\left.\begin{aligned} &a\le\alpha_1<\beta_1\le\alpha_2<\beta_2\le\dots<\beta_{m-1}\le\alpha_m<\beta_m\le b\\ &\text{and}\\ &\sum_{j=1}^{m}(\beta_j-\alpha_j)<\delta.\end{aligned}\right\} \tag{3.1.2}$$

The set of functions which are absolutely continuous on $[a,b]$ is denoted by $\mathrm{AC}([a,b])$.

3.1.2 Exercise. Prove the statement:
Every Lipschitz function on the interval $[a,b]$ (see Exercise 2.1.8 (iv)) *is absolutely continuous on this interval. In particular, if the derivative* f' *of the function* f *is continuous on* $[a,b]$ [1]*, then* f *is absolutely continuous on* $[a,b]$.

3.1.3 Theorem. *If* f *is absolutely continuous on* $[a,b]$ *and* $[c,d]\subset[a,b]$*, then* f *is absolutely continuous on* $[c,d]$*, too.*

If $a<c<b$ *and* f *is absolutely continuous on* $[a,c]$ *and* $[c,b]$*, then* f *is absolutely continuous on* $[a,b]$.

[1] More precisely, f' is continuous on (a,b) and there are finite limits $f'(a+)=\lim\limits_{t\to a+}f'(t)$, $f'(b-)=\lim\limits_{t\to b-}f'(t)$ and $f'(a)=f'(a+)$, $f'(b)=f'(b-)$.

Proof. The first statement is evident.

So, assume $c \in (a,b)$, $f \in \mathrm{AC}[a,c]$ and $f \in \mathrm{AC}[c,b]$ and let $\varepsilon > 0$ be given. We can choose $\delta > 0$ such that

$$\sum_{j=1}^{m} |f(\beta_j) - f(\alpha_j)| < \frac{\varepsilon}{2}$$

holds for every system of intervals $\{[\alpha_j, \beta_j] : j = 1, 2, , \ldots, m\}$ such that

$$\left.\begin{aligned} & a \le \alpha_1 < \beta_1 \le \alpha_2 < \beta_2 \ldots < \beta_{m-1} \le \alpha_m < \beta_m \le c \\ \text{and} & \\ & \sum_{j=1}^{m} (\beta_j - \alpha_j) < \delta. \end{aligned}\right\} \tag{3.1.3}$$

Simultaneously

$$\sum_{j=1}^{p} |f(\delta_j) - f(\gamma_j)| < \frac{\varepsilon}{2}$$

holds for every system of intervals $\{[\gamma_j, \delta_j] : j = 1, \ldots, p\}$ such that

$$c \le \gamma_1 < \delta_1 \le \gamma_2 < \delta_2 \ldots < \delta_{p-1} \le \gamma_p < \delta_p \le b \quad \text{and} \quad \sum_{j=1}^{p} (\delta_j - \gamma_j) < \delta. \tag{3.1.4}$$

Now, consider a system of intervals $\{[a_j, d_j] : j = 1, 2, , \ldots, n\}$ such that

$$a \le a_1 < d_1 \le a_2 < d_2 \ldots < d_{n-1} \le a_n < d_n \le b \quad \text{and} \quad \sum_{j=1}^{n} (d_j - a_j) < \delta. \tag{3.1.5}$$

We may assume that c does not belong to any of the intervals $(a_j, d_j), j{=}1, \ldots, n$. (If we had $c \in (a_k, d_k)$ for some $k \in \{1, \ldots, n\}$, we would divide the interval $[a_k, d_k]$ into the union $[a_k, c] \cup [c, d_k]$ and the new system would again satisfy (3.1.5).) Therefore we can divide the given system $\{[a_j, d_j] : j = 1, 2, , \ldots, n\}$ into the systems

$$\{[\alpha_j, \beta_j] : j = 1, 2, , \ldots, m\} \quad \text{and} \quad \{[\gamma_j, \delta_j] : j = 1, 2, , \ldots, p\}$$

satisfying (3.1.3) and (3.1.4). Thus, the sum $\sum_{j=1}^{n} |f(d_j) - f(a_j)|$ can be divided into two sums which are both less than $\frac{\varepsilon}{2}$. As a result,

$$\sum_{j=1}^{n} |f(d_j) - f(a_j)| < \varepsilon.$$

This completes the proof. □

3.1.4 Example. By Exercise 3.1.2 every function which has a continuous derivative on (a,b) is absolutely continuous on $[a,b]$. A simple example of absolutely continuous function on $[a,b]$ which does not have a continuous derivative on (a,b) is the function

$$f(x) = \begin{cases} x-a & \text{for } x \in [a, \dfrac{a+b}{2}], \\ b-x & \text{for } x \in [\dfrac{a+b}{2}, b], \end{cases}$$

which is obviously absolutely continuous on the intervals $[a, \frac{a+b}{2}]$ and $[\frac{a+b}{2}, b]$. By Theorem 3.1.3 this means that f is absolutely continuous also on $[a,b]$.

3.1.5 Remark. If $f:[a,b]\to\mathbb{R}$, $\mathbb{K}\subset\mathbb{N}$ and if for every $\varepsilon>0$ there exists $\delta>0$ such that

$$\sum_{j\in\mathbb{K}} |f(\beta_j)-f(\alpha_j)| < \varepsilon \tag{3.1.6}$$

holds for every (not necessarily finite) system of intervals

$$\{[\alpha_j,\beta_j]\subset[a,b]: j\in\mathbb{K}\}$$

satisfying

$$(\alpha_j,\beta_j)\cap(\alpha_k,\beta_k)=\emptyset \quad \text{for } j\neq k \text{ and } \sum_{j\in\mathbb{K}}(\beta_j-\alpha_j)<\delta, \tag{3.1.7}$$

then the function $f:[a,b]\to\mathbb{R}$ is, of course, absolutely continuous on $[a,b]$.

The following lemma shows that also the converse implication holds.

3.1.6 Lemma. *If $f\in \mathrm{AC}([a,b])$, then for every $\varepsilon>0$ there exists $\delta>0$ such that the inequality* (3.1.6) *holds for any (possibly infinite) system*

$$\{[\alpha_j,\beta_j]\subset[a,b]: j\in\mathbb{K}\}$$

of subintervals of the interval $[a,b]$ satisfying (3.1.7).

Proof. Assume $f\in\mathrm{AC}([a,b])$. Obviously, it is sufficient to prove the statement of the lemma for the case when $\mathbb{K}=\mathbb{N}$. Let $\varepsilon>0$ be given and let $\delta>0$ be determined by Definition 3.1.1 for $\varepsilon/2$ instead of ε. Let $\{[\alpha_j,\beta_j]: j\in\mathbb{N}\}$ be a system of subintervals in $[a,b]$ satisfying (3.1.7). Then for every $m\in\mathbb{N}$ we have

$$\sum_{j=1}^{m}(\beta_j-\alpha_j)<\delta, \quad \text{and thus} \quad \sum_{j=1}^{m}|f(\beta_j)-f(\alpha_j)|<\frac{\varepsilon}{2}.$$

Hence

$$\sum_{j=1}^{\infty} |f(\beta_j) - f(\alpha_j)| = \lim_{m\to\infty} \sum_{j=1}^{m} |f(\beta_j) - f(\alpha_j)| \le \frac{\varepsilon}{2} < \varepsilon.$$

This completes the proof. □

3.1.7 Theorem. *Every function which is absolutely continuous on an interval* $[a,b]$ *has bounded variation on this interval.*

Proof. Let $f \in \mathrm{AC}([a,b])$. Choose $\delta > 0$ such that

$$\sum_{j=1}^{m} |f(d_j) - f(a_j)| < 1$$

holds for every finite system of intervals $\{[a_j,d_j] : j = 1,2,,\dots,m\}$ satisfying (3.1.2). Next, choose a division $\{x_0,x_1,\dots,x_k\}$ of $[a,b]$ such that

$$0 < x_i - x_{i-1} < \delta \qquad \text{for every } i = 1,\dots,k.$$

Then for every $i = 1,\dots,k$ and every division $\boldsymbol{\alpha}^i = \{\alpha_0^i, \alpha_1^i, \dots, \alpha_{m_i}^i\}$ of the interval $[x_{i-1},x_i]$, we have

$$\sum_{j=1}^{m_i} (\alpha_j^i - \alpha_{j-1}^i) = x_i - x_{i-1} < \delta,$$

and consequently (by Theorem 2.1.14)

$$\mathrm{var}_a^b\, f = \sum_{i=1}^{k} \mathrm{var}_{x_{i-1}}^{x_i}\, f = \sum_{i=1}^{k} \sup_{\boldsymbol{\alpha}^i \in \mathscr{D}[x_{i-1},x_i]} V(f,\boldsymbol{\alpha}^i) \le k < \infty.$$

□

3.1.8 Theorem. *If* $f,g \in \mathrm{AC}([a,b])$, *then*

$$|f|,\ f+g,\ f\,g,\ \max\{f,g\},\ \min\{f,g\} \in \mathrm{AC}([a,b]).$$

If, in addition, $|f(x)| > 0$ *on* $[a,b]$, *then* $\dfrac{1}{f} \in \mathrm{AC}([a,b])$.

Proof. Let $f,\ g \in \mathrm{AC}([a,b])$.

a) $|f(x)| \le |f(x) - f(y)| + |f(y)|$ holds for any $x,\ y \in [a,b]$. Hence

$$|f(x) - f(y)| \ge \big||f(x)| - |f(y)|\big| \quad \text{for all } x,y \in [a,b]$$

and consequently

$$\sum_{j=1}^{m} \big||f(\beta_j)| - |f(\alpha_j)|\big| \le \sum_{j=1}^{m} |f(\beta_j) - f(\alpha_j)|.$$

This shows that $|f| \in \mathrm{AC}([a,b])$.

b) The statements $f+g \in \mathrm{AC}([a,b])$ and $f\,g \in \mathrm{AC}([a,b])$ follow from the inequalities

$$|(f(x)+g(x)) - (f(y)+g(y))| \le |f(x)-f(y)| + |g(x)-g(y)|$$

and

$$|f(x)\,g(x) - f(y)\,g(y)| \le \|f\|\,|g(x)-g(y)| + \|g\|\,|f(x)-f(y)|.$$

c) For any $x \in [a,b]$ we have

$$\max\{f(x), g(x)\} = \frac{1}{2}\Big(f(x)+g(x)+|f(x)-g(x)|\Big)$$

and

$$\min\{f(x), g(x)\} = \frac{1}{2}\Big(f(x)+g(x)-|f(x)-g(x)|\Big).$$

Therefore

$$\max\{f,g\} \in \mathrm{AC}([a,b]) \text{ and } \min\{f,g\} \in \mathrm{AC}([a,b])$$

holds as a consequence of a) and b).

d) Finally, if, in addition, $|f| > 0$ on $[a,b]$, then there exists $\mu > 0$ such that $|f(x)| \ge \mu$ holds for $x \in [a,b]$. Hence

$$\Big|\frac{1}{f(x)} - \frac{1}{f(y)}\Big| \le \frac{|f(x)-f(y)|}{\mu^2}$$

is true for all $x, y \in [a,b]$. Now, it is easy to show that $\dfrac{1}{f} \in \mathrm{AC}([a,b])$. □

We will close this section by stating and proving two further interesting properties of absolutely continuous functions.

3.1.9 Lemma. *Let* $f \in \mathrm{AC}([a,b])$ *and* $v(x) = \mathrm{var}_a^x f$ *for* $x \in [a,b]$. *Then* v *is also absolutely continuous on* $[a,b]$.

Proof. Assume that $\varepsilon > 0$ is given and let $\delta > 0$ be such that

$$\sum_{j=1}^{m} |f(d_j) - f(a_j)| < \frac{\varepsilon}{2}$$

is true for every system of intervals $\{[a_j, d_j] : j = 1, 2, , \dots, m\}$ satisfying (3.1.2).

Let $[\alpha_j,\beta_j]$, $j=1,\dots,n$, be an arbitrary system of intervals satisfying (3.1.3) in which $m=n$. For any $j\in\{1,\dots,m\}$, let $\boldsymbol{\alpha}^j=\{\alpha_0^j,\alpha_1^j,\dots,\alpha_{n_j}^j\}$ be an arbitrary division of the interval $[\alpha_j,\beta_j]$. Then

$$\sum_{j=1}^{n}\sum_{i=1}^{n_j}\big(\alpha_i^j-\alpha_{i-1}^j\big)=\sum_{j=1}^{n}\big[\beta_j-\alpha_j\big]<\delta,$$

and hence

$$\sum_{j=1}^{n}V(f,\boldsymbol{\alpha}^j)=\sum_{j=1}^{n}\sum_{i=1}^{n_j}\big|f(\sigma_i^j)-f(\sigma_{i-1}^j)\big|<\frac{\varepsilon}{2}.$$

This implies that

$$\sum_{j=1}^{n}\big(v(\beta_j)-v(\alpha_j)\big)=\sum_{j=1}^{n}\operatorname{var}_{\alpha_j}^{\beta_j}f=\sum_{j=1}^{n}\Big(\sup_{\boldsymbol{\alpha}^j\in\mathscr{D}[\alpha_j,\beta_j]}V(f,\boldsymbol{\alpha}^j)\Big)\le\frac{\varepsilon}{2}<\varepsilon.$$

This completes the proof of the lemma. □

3.1.10 Corollary. *A function $f:[a,b]\to\mathbb{R}$ is absolutely continuous on the interval $[a,b]$ if and only if there exist functions f_1 and f_2 which are nondecreasing and absolutely continuous on $[a,b]$ and such that $f=f_1-f_2$ on the interval $[a,b]$.*

Proof. a) Let $f=f_1-f_2$ on $[a,b]$, where f_1, f_2 are absolutely continuous and nondecreasing on $[a,b]$. Then by Theorem 3.1.8 f is absolutely continuous on $[a,b]$, too.

b) Let $f\in\mathrm{AC}([a,b])$. By Theorems 3.1.7 and 2.1.21 there exist such functions f_1, f_2 nondecreasing on $[a,b]$ that $f=f_1-f_A$. By the proof of Theorem 2.1.21 we can set

$$f_1(x)=\operatorname{var}_a^x f\ \text{ and }\ f_2(x)=f_1(x)-f(x)\ \text{ for } x\in[a,b].$$

By Theorem 3.1.8, it is sufficient to prove that f_1 is absolutely continuous on $[a,b]$. But that follows from Lemma 3.1.9. □

3.2 Absolutely continuous functions and Lebesgue integral

Let us recall that by Theorem 2.4.2 every function of bounded variation on an interval $[a,b]$ has a bounded derivative $f'(x)$ for a.e. $x\in[a,b]$. By Theorem 3.1.7

every function which is absolutely continuous on $[a,b]$ thus has the same property. In the remaining part of this chapter, we will recall some other basic properties of the derivatives of absolutely continuous functions and the connection between absolute continuity and indefinite Lebesgue integral. In the cases when the proofs or their parts are based on the theory of measure in the extent beyond this text, the proofs (or their parts) are not included and we only refer to accessible literature. The integral in this section is supposed to be the Lebesgue one.

By the next theorem the derivatives of the functions of bounded variation (and thus all the more so of absolutely continuous functions) are Lebesgue integrable. Its proof substantially uses a range of knowledge from the theory of measure and Lebesgue integration which will not fit in this text. The full proof can be found in relevant literature (see e.g. Theorem 4.10 in [43] or Theorem 6.2.9 in [111]).

3.2.1 Theorem. *If a function $f:[a,b]\to\mathbb{R}$ has a bounded variation on $[a,b]$, then its derivative f' is Lebesgue integrable on $[a,b]$.*

If f is also nondecreasing on $[a,b]$, then

$$0\le\int_a^b f'(x)\,\mathrm{d}x\le f(b)-f(a). \tag{3.2.1}$$

The next statement concerns the differentiation of indefinite integrals of integrable functions. For the proof see, e.g. Theorem 4.12 in [43] or Theorem 6.3.1 in [111]. Let us recall (cf. Conventions and Notation (xi)) that $\mathrm{L}^1([a,b])$ stands for the space of all real functions that are Lebesgue integrable on $[a,b]$.

3.2.2 Theorem. *If $g\in\mathrm{L}^1([a,b])$ and*

$$f(x)=\int_a^x g(t)\,\mathrm{d}t \quad \textit{for } x\in[a,b],$$

then f is absolutely continuous on $[a,b]$ and $f'(x)=g(x)$ for a.e. $x\in[a,b]$.

Let a function $g\in\mathrm{L}^1([a,b])$ be given. By Theorem 3.2.2 its indefinite Lebesgue integral f is absolutely continuous on $[a,b]$ and $f'=g$ a.e. on $[a,b]$. We want to show that f is absolutely continuous on $[a,b]$ if and only if f is the indefinite integral of some Lebesgue integrable function. The following lemma known as Riesz's lemma is essential for the proof of such a statement. For the proof see e.g. Lemma 7.5 in [16].

3.2.3 Lemma (RIESZ RISING SUN LEMMA). *Let $f\in\mathrm{C}[a,b]$ and*

$$E=\{x\in(a,b): \textit{there is } \xi\in(x,b] \textit{ such that } f(\xi)>f(x)\}.$$

Then the set E is open and it is a union of at most countable system of disjoint open intervals (a_k,d_k) while $f(a_k)\le f(d_k)$ holds for any of them.

3.2.4 Lemma. *If* $f \in \mathrm{AC}([a,b])$ *is nondecreasing on* $[a,b]$ *and* $f'(x)=0$ *for a.e.* $x \in [a,b]$, *then* f *is constant on* $[a,b]$.

Proof. Due to its monotonicity, the function f maps the interval $[a,b]$ on the interval $[f(a),f(b)]$. We will prove that $f(a)=f(b)$.

Let $\varepsilon>0$ be given and let $\delta>0$ be as in Lemma 3.1.6. Let Z be the set of all $x \in [a,b]$ for which $f'(x)=0$ holds. Its complement $[a,b]\setminus Z$ has zero measure ($\mu([a,b]\setminus Z)=0$) by assumption. This means that there exists a finite or countable system $\{(\alpha_j,\beta_j): j\in\mathbb{K}\}$ satisfying (3.1.7) and

$$[a,b]\setminus Z \subset \bigcup_{j\in\mathbb{K}} (\alpha_j,\beta_j).$$

The image $f([a,b]\setminus Z)$ of the set $[a,b]\setminus Z$ is thus contained in the union of open intervals $\{(f(\alpha_j),f(\beta_j)): j\in\mathbb{K}\}$. Since (3.1.6) holds by Lemma 3.1.6, the set $f([a,b]\setminus Z)$ has zero measure, i.e.

$$\mu(f([a,b]\setminus Z))=0. \tag{3.2.2}$$

Now, let $x\in Z$. Then $f'(x)=0$ and thus there is $\Delta>0$ such that

$$\frac{f(t)-f(x)}{t-x} < \varepsilon \quad \text{for every } t \text{ such that } 0<|t-x|<\Delta.$$

This implies

$$\varepsilon\, x - f(x) < \varepsilon\, t - f(t) \quad \text{for every } t\in(x,x+\Delta).$$

By Riesz's lemma 3.2.3, which we apply to the function $\varepsilon\, x - f(x)$ instead of $f(x)$, the set Z is thus contained in the union of a finite or countable system of disjoint intervals $\{(a_k,d_k)\subset[a,b]: k\in\mathbb{K}\}$, while

$$\varepsilon\, a_k - f(a_k) \le \varepsilon\, d_k - f(d_k), \quad i.e. f(d_k)-f(a_k)\le \varepsilon\,(d_k-a_k)$$

holds for every $k\in\mathbb{K}$. Hence

$$\sum_{k\in\mathbb{K}} [f(d_k)-f(a_k)] \le \varepsilon \sum_{k\in\mathbb{K}} [d_k-a_k] \le \varepsilon\,(b-a).$$

Now we can already deduce that the set $f(Z)$ has also zero measure, i.e.

$$\mu(f(Z))=0. \tag{3.2.3}$$

By (3.2.2) and (3.2.3) the interval $[f(a),f(b)]=f(Z)\cup(f([a,b]\setminus Z))$ has zero length, i.e., thanks to the monotonicity of f, $f(a)=f(x)=f(b)$ for every $x\in(a,b)$. □

3.2.5 Theorem. *A function* $f:[a,b]\to\mathbb{R}$ *is absolutely continuous on* $[a,b]$ *if and only if*

$$f(x)-f(a)=\int_a^x g(t)\,\mathrm{d}t \qquad \textit{for } x\in[a,b] \tag{3.2.4}$$

for some function $g\in\mathrm{L}^1([a,b])$. *Then* $f'=g$ *a.e. on* $[a,b]$.

Proof. a) Let $g\in\mathrm{L}^1([a,b])$ and

$$f(x)=f(a)+\int_a^x g(t)\,\mathrm{d}t \quad \text{for } x\in[a,b].$$

Then f is absolutely continuous on $[a,b]$ and by Theorem 3.2.2, $f'=g$ a.e. on $[a,b]$.

b) First, assume the function $f\in\mathrm{AC}([a,b])$ is nondecreasing on $[a,b]$. By Theorems 3.1.7 and 3.2.1, $f'\in\mathrm{L}^1([a,b])$. Set

$$h(x)=\int_a^x f'(t)\,\mathrm{d}t \quad \text{and} \quad g(x)=f(x)-h(x) \quad \text{for } x\in[a,b].$$

We will show that the function g is nondecreasing on $[a,b]$, too. By Theorem 3.2.1 we have

$$\begin{aligned} g(y)-g(x)&=\big(f(y)-h(y)\big)-\big(f(x)-h(x)\big)\\ &=\big(f(y)-f(x)\big)-\int_x^y f'(t)\,\mathrm{d}t\ge 0 \end{aligned}$$

for all points $x,y\in[a,b]$ such that $x\le y$.

Moreover, by Theorem 3.2.2 the function h is absolutely continuous on $[a,b]$ and $h'=f'$ a.e. on $[a,b]$. Hence, $g'=(f-h)'=0$ a.e. on $[a,b]$. By Lemma 3.2.4 the function g is therefore constant on $[a,b]$. Thus we get

$$g(x)=f(x)-h(x)=f(a)-h(a)=f(a) \quad \text{for } x\in[a,b],$$

i.e.

$$f(x)=f(a)+h(x)=f(a)+\int_a^x f'(t)\,\mathrm{d}t \quad \text{for } x\in[a,b].$$

This means that (3.2.4) holds for every function $f\in\mathrm{AC}([a,b])$ nondecreasing on $[a,b]$.

In the general case of $f\in \mathrm{AC}([a,b])$, by Corollary 3.1.10, there exist functions f_1, f_2 absolutely continuous on $[a,b]$, nondecreasing on $[a,b]$ and such that $f=f_1-f_2$ on $[a,b]$. We thus have

$$\begin{aligned} f(x) = f_1(x)-f_2(x) &= \Big(f_1(a)+\int_a^x f_1'(t)\,\mathrm{d}t\Big)-\Big(f_2(a)+\int_a^x f_2'(t)\,\mathrm{d}t\Big)\\ &= f(a)+\int_a^x f'(t)\,\mathrm{d}t \quad \text{for } x\in[a,b]. \end{aligned}$$

The proof has been completed. □

Next result is an extension of Theorem 2.1.5

3.2.6 Theorem. *If* $f\in \mathrm{AC}([a,b])$, *then* $\mathrm{var}_a^b\, f = \int_a^b |f'(t)|\,\mathrm{d}t$.

Proof. Put $v(a)=0$ and $v(x)=\mathrm{var}_a^x\, f$ for $x\in[a,b]$. Then, using Lemma 3.1.9 and Theorem 3.2.5, we get

$$v\in \mathrm{AC}([a,b]) \text{ and } v(x)=\int_a^x v'(t)\,\mathrm{d}t \quad \text{for } x\in[a,b].$$

In particular,

$$\mathrm{var}_a^b\, f = v(b) = \int_a^b v'(t)\,\mathrm{d}t.$$

Furthermore, by the proof of Theorem 2.1.21, we know that $v-f$ is nondecreasing on $[a,b]$. Similarly, we can verify that $v+f$ is nondecreasing on $[a,b]$, as well. Since $v'-f'\geq 0$ and $v'+f'\geq 0$ almost everywhere in $[a,b]$, it follows that $|f'(t)|\leq v'(t)$ for almost all $t\in[a,b]$ and thus

$$\mathrm{var}_a^b\, f \geq \int_a^b |f'(t)|\,\mathrm{d}t. \tag{3.2.5}$$

On the other hand, for an arbitrary division $\boldsymbol{\alpha}$ of $[a,b]$ we have

$$V(f,\boldsymbol{\alpha}) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \Big|\int_{\alpha_{j-1}}^{\alpha_j} f'(t)\,\mathrm{d}t\Big| \leq \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \int_{\alpha_{j-1}}^{\alpha_j} |f'(t)|\,\mathrm{d}t = \int_a^b |f'(t)|\,\mathrm{d}t,$$

i.e.,

$$\mathrm{var}_a^b\, f \leq \int_a^b |f'(t)|\,\mathrm{d}t,$$

which together with (3.2.5) completes the proof. □

3.2.7 Exercise. Prove the following assertion:
Let $f\in \mathrm{AC}([a,b])$ *and let* v *be given as in the proof of Theorem 3.2.6. Then* $v'(t)=|f'(t)|$ *for almost all* $t\in[a,b]$.

3.3 Lebesgue decomposition of functions of bounded variation

We know (see Theorem 2.6.1 and Remark 2.6.2) that every function of bounded variation on $[a,b]$ can be decomposed into a sum of a continuous function and a step function or into a difference of two functions nondecreasing on $[a,b]$ (see Theorem 2.1.21). Another option of decomposition of functions of bounded variation is offered by the following theorem.

3.3.1 Theorem (LEBESGUE DECOMPOSITION THEOREM). *For every function* $f \in \mathrm{BV}([a,b])$, *there exist an absolutely continuous function* $f^{\,\mathrm{AC}}$, *a singular continuous function* $f^{\,\mathrm{SC}}$ *and a step function* $f^{\,\mathrm{B}}$ *such that*

$$f = f^{\,\mathrm{AC}} + f^{\,\mathrm{SC}} + f^{\,\mathrm{B}} \quad \textit{on } [a,b].$$

If $f = f_1 + f_2 + f_3$, *where the function* f_1 *is absolutely continuous on* $[a,b]$, *the function* f_2 *is singular and continuous on* $[a,b]$ *and the function* f_3 *is a step function on* $[a,b]$, *then the functions* $f^{\,\mathrm{AC}} - f_1$, $f^{\,\mathrm{SC}} - f_2$ *and* $f^{\,\mathrm{B}} - f_3$ *are constant on* $[a,b]$.

Proof. a) By Theorem 2.6.1 there exists a step function $f^{\,\mathrm{B}}$ such that the function $f^{\,\mathrm{C}} = f - f^{\,\mathrm{B}}$ is continuous on $[a,b]$. Furthermore, $f' \in \mathrm{L}^1([a,b])$ due to Theorem 3.2.1. Set

$$f^{\,\mathrm{AC}}(x) = \int_a^x f'(t)\,\mathrm{d}t \quad \text{and} \quad f^{\,\mathrm{SC}}(x) = f^{\,\mathrm{C}}(x) - f^{\,\mathrm{AC}}(x) \;\text{ for } x \in [a,b].$$

By Theorems 2.5.5 and 3.2.2 we have $(f^{\,\mathrm{B}})' = 0$ a.e. on $[a,b]$ and $(f^{\,\mathrm{AC}})' = f'$ a.e. on $[a,b]$, respectively. This means that

$$(f^{\,\mathrm{SC}})' = f' - (f^{\,\mathrm{AC}})' - (f^{\,\mathrm{B}})' = 0 \quad \text{a.e. on } [a,b].$$

b) Let $f = f_1 + f_2 + f_3$, where $f_1 \in \mathrm{AC}([a,b])$, f_2 is singular and continuous on $[a,b]$ and $f_3 \in \mathrm{B}[a,b]$. By Theorem 2.6.1 the differences

$$\big(f^{\,\mathrm{AC}} + f^{\,\mathrm{SC}}\big) - \big(f_1 + f_2\big) \;\text{ and }\; f^{\,\mathrm{B}} - f_3$$

are constant on $[a,b]$. Since

$$f^{\,\mathrm{AC}} + f^{\,\mathrm{SC}} + f^{\,\mathrm{B}} = f_1 + f_2 + f_3,$$

it means that there exists such $c \in \mathbb{R}$ that

$$\big(f^{\,\mathrm{AC}} + f^{\,\mathrm{SC}}\big) - \big(f_1 + f_2\big) = f_3 - f^{\,\mathrm{B}} = c.$$

Hence

$$\left(f^{\mathrm{AC}} - f_1\right) = c - \left(f^{\mathrm{SC}} - f_2\right) \quad \text{and} \quad \left(f^{\mathrm{AC}} - f_1\right)' = 0 \quad \text{a.e. on } [a,b].$$

As both the functions f^{AC} and f_1 are absolutely continuous on the interval $[a,b]$, it follows by Theorem 3.2.5 that the difference $f^{\mathrm{AC}} - f_1$ is constant on $[a,b]$. This completes the proof. □

3.3.2 Definition. If $f \in \mathrm{BV}([a,b])$, then the function f^{AC}, or f^{SC}, or f^{B} from Theorem 3.3.1 is called the *absolutely continuous part*, or the *continuous singular part*, or the *jump part* of the function f, respectively. In addition, the sum $f^{\mathrm{SC}} + f^{\mathrm{B}}$ is called the *singular part* of f and denoted by f^{SING}.

3.3.3 Exercise. Prove the following statement:

$$f^{\mathrm{AC}}(x) - f^{\mathrm{AC}}(a) = \int_a^x f'(t)\, \mathrm{d}t$$

holds for every function $f \in \mathrm{BV}([a,b])$ and every $x \in [a,b]$.

Next assertion is a useful addition to Theorem 3.3.1.

3.3.4 Theorem. *If $f \in \mathrm{BV}([a,b])$ is nondecreasing on $[a,b]$, then the functions f^{AC}, f^{SC}, and f^{B} from Theorem 3.3.1 are nondecreasing on $[a,b]$, too.*

Proof. Let $f \in \mathrm{BV}([a,b])$ be nondecreasing on $[a,b]$ and let the functions f^{AC}, f^{SC}, f^{B} be assigned to the function f by Theorem 3.3.1. Furthermore, let $\{s_k\}$ be the set of the points of discontinuity of the function f and x, y be any pair of points from $[a,b]$ such that $x \le y$.

Since f is nondecreasing on $[a,b]$, we have

$$\Delta^+ f(t) \ge 0 \;\text{ and }\; \Delta^- f(s) \ge 0 \text{ for } t \in [a,b),\; s \in (a,b],$$

and therefore

$$f^{\mathrm{B}}(y) - f^{\mathrm{B}}(x) = \sum_{x < s_k \le y} \Delta^- f(s_k) + \sum_{x \le s_k < y} \Delta^+ f(s_k) \ge 0.$$

The jump part f^{B} of the function f is thus nondecreasing on $[a,b]$.

Let g be the continuous part of f, i.e. $g = f - f^{\mathrm{B}}$. By Corollary 2.3.8 we have

$$f^{\mathrm{B}}(y) - f^{\mathrm{B}}(x) \le \operatorname{var}_x^y f = f(y) - f(x),$$

and hence

$$g(y) - g(x) = \big(f(y) - f(x)\big) - \big(f^{\mathrm{B}}(y) - f^{\mathrm{B}}(x)\big) \ge 0.$$

The continuous part of the function f is thus nondecreasing on $[a,b]$.

For a.e. $t \in [a,b]$ we have

$$f'(t) = \lim_{s \to t} \frac{f(s) - f(t)}{s - t} \in \mathbb{R}.$$

Since f is nondecreasing on $[a,b]$, the inequality $f'(t) \geq 0$ holds for a.e. $t \in [a,b]$. By the proof of Theorem 3.3.1 we thus get

$$f^{\text{AC}}(y) - f^{\text{AC}}(x) = \int_x^y f'(t)\, \mathrm{d}t \geq 0 \text{ whenever } x,\, y \in [a,b] \text{ and } x \leq y.$$

This means that f^{AC} is nondecreasing on $[a,b]$.

By Theorem 2.5.5, $(f^{\text{B}})' = 0$ a.e. on $[a,b]$, and hence

$$g' = f' - (f^{\text{B}})' = f' \quad \text{a.e. on } [a,b].$$

Using (3.2.1) and the proof of Theorem 3.3.1 we can deduce that

$$g(y) - g(x) \geq \int_x^y g'(t)\, \mathrm{d}t = \int_x^y f'(t)\, \mathrm{d}t = f^{\text{AC}}(y) - f^{\text{AC}}(x)$$

is true, i.e.

$$\begin{aligned} f^{\text{SC}}(y) - f^{\text{SC}}(x) &= \big(g(y) - f^{\text{AC}}(y)\big) - \big(g(x) - f^{\text{AC}}(x)\big) \\ &= \big(g(y) - g(x)\big) - \big(f^{\text{AC}}(y) - f^{\text{AC}}(x)\big) \geq 0. \end{aligned}$$

The continuous singular part f^{SC} of the function f is thus nondecreasing on $[a,b]$, too. This completes the proof. □

We can now state the following assertion, which is in some sense complementary to Proposition 2.5.7.

3.3.5 Proposition. *Let f be singular on $[a,b]$ and let g be absolutely continuous on $[a,b]$. Then* $\operatorname{var}_a^b (f+g) = \operatorname{var}_a^b f + \operatorname{var}_a^b g$.

For the proof we will need the Vitali Covering Theorem, which is based on the notion of a *Vitali cover*.

3.3.6 Definition. Let E be a subset of $\mathbb{R}$ and let $\mathcal{V}$ be a collection of nondegenerate [2] closed subintervals of $[a,b]$. We say that $\mathcal{V}$ is a Vitali cover of E, if for each $\varepsilon > 0$ and any $x \in E$, there is an interval $[\alpha,\beta] \in \mathcal{V}$ containing x and such that $\beta - \alpha < \varepsilon$.

[2] It means that no singletons (i.e. one-point sets) are allowed.

3.3.7 Theorem (Vitali Covering Theorem)**.** *Let E be a subset of $[a,b]$ and let $\mathcal{V}$ be a Vitali cover of E. Then, for every $\varepsilon > 0$, there is a finite collection $\{I_1, \dots, I_N\}$ of disjoint intervals from $\mathcal{V}$ such that*

$$\mu^*(E \setminus \bigcup_{j=1}^{N} I_j) < \varepsilon,$$

*where μ^* stands for the outer Lebesgue measure (*see section 2.4*).*

Proof of Proposition 3.3.5.

We will verify that the assumptions of Lemma 2.1.19 are satisfied.

Let $\varepsilon > 0$ be given. By Lemma 3.1.9 the function $v_g(x) = \operatorname{var}_a^x g$ is absolutely continuous. Hence, we can choose $\delta > 0$ in such a way that

$$\sum_{j=1}^{m} \operatorname{var}_{\alpha_j}^{\beta_j} g = \sum_{j=1}^{m} |v_g(\beta_j) - v_g(\alpha_j)| < \varepsilon \quad \text{whenever} \quad \sum_{j=1}^{m} (\beta_j - \alpha_j) < \delta. \quad (3.3.1)$$

Further, choose a division $\boldsymbol{\sigma}$ of $[a,b]$ in such a way that

$$V(f, \boldsymbol{\sigma}) > \operatorname{var}_a^b f - \frac{\varepsilon}{2}. \quad (3.3.2)$$

As f is singular, there is a set $N \subset [a,b]$ such that

$$f'(x) = 0 \quad \text{for all} \quad x \in [a,b] \setminus N \quad \text{and} \quad \mu(N) = 0.$$

Let $\mathcal{V}$ be the set of all nondegenerate intervals $[\xi, \eta] \subset [a,b] \setminus \boldsymbol{\sigma}$ such that the inequality

$$|f(\eta) - f(\xi)| \le \frac{\varepsilon}{2} \frac{\eta - \xi}{b - a} \quad (3.3.3)$$

holds. Obviously, $\mathcal{V}$ is a Vitali cover of the set $E := [a,b] \setminus (N \cup \boldsymbol{\sigma})$. Hence, by Theorem 3.3.7, there is a finite system $\big\{[\xi_j, \eta_j] : j \in \{1, \dots, r\}\big\}$ of disjoint intervals such that

$$a < \xi_1 < \eta_1 < \dots < \xi_r < \eta_r < b \quad \text{and} \quad \mu^*(E \setminus \bigcup_{j=1}^{r} [\xi_j, \eta_j]) < \delta.$$

Now, let $\boldsymbol{\alpha}$ be the division of $[a,b]$ consisting of all elements of the set $\{\xi_j, \eta_j : j \in \{1, \dots, r\}\} \cup \boldsymbol{\sigma}$. Let $\mathscr{K}$ be the set of all indices $k \in \{1, \dots, \nu(\boldsymbol{\alpha})\}$ for which the intersection $(\alpha_{k-1}, \alpha_k) \cap [\xi_j, \eta_j]$ is empty for each $j \in \{1, \dots, r\}$. Then

$$\bigcup_{k \in \mathscr{K}} (\alpha_{k-1}, \alpha_k) \subset E \setminus \bigcup_{j=1}^{r} [\xi_j, \eta_j].$$

Hence

$$\sum_{k\in\mathscr{K}}(\alpha_k-\alpha_{k-1})=\mu^*\Big(\bigcup_{k\in\mathscr{K}}(\alpha_{k-1},\alpha_k)\Big)\le\mu^*\Big(E\setminus\bigcup_{j=1}^{r}[\xi_j,\eta_j]\Big)<\delta.$$

Consequently, we can apply (3.3.1) to get

$$\sum_{k\in\mathscr{K}}\operatorname{var}_{\alpha_{k-1}}^{\alpha_k} g<\varepsilon. \tag{3.3.4}$$

On the other hand, by (3.3.2) and since $\boldsymbol{\alpha}$ is a refinement of $\boldsymbol{\sigma}$, we have

$$\operatorname{var}_a^b f-\frac{\varepsilon}{2}<V(f,\boldsymbol{\alpha})=\sum_{k\in\mathscr{K}}|f(\alpha_k)-f(\alpha_{k-1})|+\sum_{k\in\mathscr{K}'}|f(\alpha_k)-f(\alpha_{k-1})|, \tag{3.3.5}$$

where $\mathscr{K}'=\{1,\dots,\nu(\boldsymbol{\alpha})\}\setminus\mathscr{K}$. Of course,

$$\sum_{k\in\mathscr{K}'}|f(\alpha_k)-f(\alpha_{k-1})|=\sum_{j=1}^{r}|f(\eta_j)-f(\xi_j)|$$

and, due to (3.3.3),

$$\sum_{k\in\mathscr{K}'}|f(\alpha_k)-f(\alpha_{k-1})|<\frac{\varepsilon}{2\,(b-a)}\sum_{j=1}^{r}(\eta_j-\xi_j)<\frac{\varepsilon}{2}\,.$$

Moreover,

$$\sum_{k\in\mathscr{K}}|f(\alpha_k)-f(\alpha_{k-1})|\le\sum_{k\in\mathscr{K}}\operatorname{var}_{\alpha_{k-1}}^{\alpha_k} f.$$

To summarize, by (3.3.5) we have

$$\begin{aligned}\sum_{k\in\mathscr{K}}\operatorname{var}_{\alpha_{k-1}}^{\alpha_k} f&\ge\sum_{k\in\mathscr{K}}|f(\alpha_k)-f(\alpha_{k-1})|\\&=V(f,\boldsymbol{\alpha})-\sum_{k\in\mathscr{K}'}|f(\alpha_k)-f(\alpha_{k-1})|\\&>V(f,\boldsymbol{\alpha})-\frac{\varepsilon}{2}>\operatorname{var}_a^b f-\varepsilon,\end{aligned}$$

i.e.

$$\sum_{k\in\mathscr{K}}\operatorname{var}_{\alpha_{k-1}}^{\alpha_k} f>\operatorname{var}_a^b f-\varepsilon. \tag{3.3.6}$$

Now, if we relabel the points α_k in such a way that it will be

$$\big\{[\alpha_{k-1},\alpha_k]:k\in\mathscr{K}\big\}=\big\{[a_j,b_j]:j\in\{1,\dots,n\}\big\},$$

we can check that, thanks to (3.3.4) and (3.3.6), the conditions (2.1.13)-(2.1.15) of Lemma 2.1.19 are satisfied. This completes the proof. □

3.3.8 Exercises. Use Proposition 3.3.5 to prove the following statements:

- *Let* $\{f_n\}$ *be a sequence of functions with bounded variations on* $[a,b]$ *and let* $\{f_n^{\text{AC}}\}$ *and* $\{f_n^{\text{SING}}\}$ *be the sequences of absolutely continuous and singular parts of* $\{f_n\}$, *respectively. Then*

$$\lim_{n\to\infty}(\text{var}_a^b f_n) = 0 \text{ if and only if } \lim_{n\to\infty}(\text{var}_a^b f_n^{\text{AC}}) = \lim_{n\to\infty}(\text{var}_a^b f_n^{\text{SING}}) = 0.$$

- *If* $\in \text{BV}([a,b])$, *then* $\text{var}_a^b f \geq \int_a^b |f'(t)|\,\text{d}t$.

By Theorem 3.1.8, $\text{AC}([a,b])$ is a linear subspace of $\text{BV}([a,b])$. We will close this chapter by a further corollary of Proposition 3.3.5 which shows that, when equipped with the norm of $\text{BV}([a,b])$, the space $\text{AC}([a,b])$ becomes a Banach space.

3.3.9 Proposition. $\text{AC}([a,b])$ *is a Banach space when equipped with the norm*

$$\|f\|_{\text{AC}} := \|f\|_{\text{BV}} = |f(a)| + \text{var}_a^b f \quad \textit{for } f \in \text{AC}([a,b]).$$

Proof. We will show that $\text{AC}([a,b])$ is a closed subspace of $\text{BV}([a,b])$. To this aim assume that $\{f_n\}$ is a sequence of absolutely continuous functions which is convergent in the BV norm to a function $f:[a,b]\to\mathbb{R}$, i.e.,

$$\lim_{n\to\infty}\|f_n - f\|_{\text{BV}} = 0. \tag{3.3.7}$$

Clearly, $f \in \text{BV}([a,b])$ and $f = f^{\text{AC}} + f^{\text{SING}}$ on $[a,b]$, where f^{AC} is the absolutely continuous part of f and f^{SING} is the singular part of f. Without any loss of generality we may assume that $f_n(a) = f(a) = 0$ for all $n \in \mathbb{N}$. Then, thanks to Proposition 3.3.5, relation (3.3.7) can be rewritten as

$$\begin{aligned} 0 &= \lim_{n\to\infty}\|f - f_n\|_{\text{BV}} = \lim_{n\to\infty}\text{var}_a^b(f - f_n) \\ &= \lim_{n\to\infty}\text{var}_a^b(f^{\text{AC}} + f^{\text{SING}} - f_n) = \lim_{n\to\infty}\text{var}_a^b(f^{\text{AC}} - f_n) + \text{var}_a^b f^{\text{SING}}, \end{aligned}$$

which is possible only if $\text{var}_a^b f^{\text{SING}} = 0$, i.e., if $f^{\text{SING}} \equiv 0$ on $[a,b]$. In other words, $f = f^{\text{AC}} \in \text{AC}([a,b])$, wherefrom the proof immediately follows. □

More details about absolutely continuous functions can be found e.g. in monographs [5], [43], [70], or in the lecture notes [92].

Chapter 4

Regulated functions

The analysis of functions of bounded variation is one of the crucial keys for the development of Stieltjes integration theory. Of similar importance is the class of regulated functions, which represent a very natural generalization of both the continuous functions and the functions of bounded variation. This chapter is fully devoted to the study of regulated functions.

Throughout the chapter, we assume that $-\infty < a < b < \infty$. For a given function $f:[a,b] \to \mathbb{R}$, we set

$$\|f\|_\infty = \sup_{t\in[a,b]} |f(t)|.$$

4.1 Introduction

4.1.1 Definition. A function $f:[a,b]\to\mathbb{R}$ is said to be regulated on $[a,b]$ if the left limit $f(t+)$ exists and is finite for every $t\in[a,b)$, and the right limit $f(t-)$ exists and is finite for every $t\in(a,b]$. The set of all regulated functions on $[a,b]$ will be denoted by $\mathrm{G}([a,b])$. Recall that

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

4.1.2 Remark. Evidently, the following relations hold:

$$\mathrm{BV}([a,b]) \cup \mathrm{C}([a,b]) \subset \mathrm{G}([a,b]),$$
$$\mathrm{G}([a,b]) \setminus \mathrm{C}([a,b]) \neq \emptyset \text{ and } \mathrm{G}([a,b]) \setminus \mathrm{BV}([a,b]) \neq \emptyset$$

For an example of a regulated function which does not have bounded variation, see Example 2.1.10.

4.1.3 Theorem. *If a sequence* $\{f_n\}$ *of regulated functions converges uniformly on the interval* $[a,b]$ *to a function* $f:[a,b]\to\mathbb{R}$, *then this function is also regulated on* $[a,b]$.

Proof. Let $t\in[a,b)$ and let $\{t_k\}\subset(t,b]$ be an arbitrary decreasing sequence such that $t_k\to t$ for $k\to\infty$. Given an arbitrary $\varepsilon>0$, choose $n_0\in\mathbb{N}$ and $k_0\in\mathbb{N}$ such that

$$\|f-f_{n_0}\|_\infty < \frac{\varepsilon}{3} \text{ and } |f_{n_0}(t_k) - f_{n_0}(t_\ell)| < \frac{\varepsilon}{3} \quad \text{for all } k,\ell \geq k_0.$$

Then

$$\begin{aligned}|f(t_k)-f(t_\ell)| &\le |f(t_k)-f_{n_0}(t_k)|+|f_{n_0}(t_k)-f_{n_0}(t_\ell)| \\ &\qquad +|f(t_\ell)-f_{n_0}(t_\ell)| \\ &\le 2\,\|f-f_{n_0}\|_\infty+|f_{n_0}(t_k)-f_{n_0}(t_\ell)|<\varepsilon \quad \text{for } k,\ell\ge k_0.\end{aligned}$$

Consequently, there exists a limit

$$f(t+)=\lim_{k\to\infty} f(t_k)\in\mathbb{R}.$$

Similarly, we would show that for every $t\in(a,b]$ there exists a limit $f(t-)\in\mathbb{R}$. □

4.1.4 Exercises. (i) In the context of Theorem 4.1.3, prove that

$$f(t+)=\lim_{n\to\infty} f_n(t+) \ \text{ for every } t\in[a,b)$$

and

$$f(t-)=\lim_{n\to\infty} f_n(t-) \ \text{ for every } t\in(a,b].$$

This statement represents a special case of the Moore-Osgood theorem; an even stronger result will be obtained in Lemma 4.2.3.

(ii) Let $f(x)=x$ if $x=1/k$ for a certain $k\in\mathbb{N}$, and $f(x)=0$ otherwise. Show that f is regulated on $[0,1]$.

(iii) Let $f_D(x)=1$ if x is a rational number, and $f_D(x)=0$ otherwise (f_D is the *Dirichlet function*). Show that f_D is not regulated on $[0,1]$.

Let us now formulate the crucial result of this chapter.

4.1.5 Theorem (HÖNIG)**.** *The following three statements are equivalent:*

(i) $f\in\mathrm{G}([a,b])$.

(ii) *There exists a sequence* $\{f_n\}\subset\mathrm{S}([a,b])$ *which converges uniformly to* f *on* $[a,b]$.

(iii) *For every* $\varepsilon>0$ *there exists a division* $\boldsymbol{\alpha}$ *of* $[a,b]$ *such that*

$$|f(t)-f(s)|<\varepsilon$$

holds for every $j\in\{1,\dots,\nu(\boldsymbol{\alpha})\}$ *and each pair* $t,s\in(\alpha_{j-1},\alpha_j)$.

Proof. a) The implication (ii) $\Longrightarrow$ (i) is proved by Theorem 4.1.3.
b) Assume (i) holds and let an arbitrary $\varepsilon > 0$ be given. Denote by B the set of all points $\tau \in (a,b]$ with the following property:

$$\left.\begin{array}{l}\text{There is a division } \boldsymbol{\alpha} \text{ of } [a,\tau] \text{ such that } |f(t)-f(s)|<\varepsilon \\ \text{for each pair } t,s \in (\alpha_{j-1},\alpha_j), \text{ where } j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}.\end{array}\right\} \tag{4.1.1}$$

Our goal is to prove that $b \in B$. First, we show that B is nonempty. By Definition 4.1.1, there is a $\delta_a \in (0, b-a)$ such that

$$|f(t)-f(a+)| < \frac{\varepsilon}{2} \quad \text{holds for all } \ t \in (a, a+\delta_a).$$

Thus, for arbitrary $t,s \in (a,a+\delta_a)$, we get

$$|f(t)-f(s)| \le |f(t)-f(a+)| + |f(s)-f(a+)| < \varepsilon.$$

Denote $\tau = a+\delta_a$. Then $\{a,\tau\}$ is a division of $[a,\tau]$ satisfying (4.1.1). This means that the set B is nonempty and $\tau^* := \sup B \in (a,b]$.

Next, we will show that $\tau^* \in B$. Indeed, by Definition 4.1.1 we can choose a $\delta_1 \in (0,\tau^*-a)$ in such a way that

$$|f(t)-f(\tau^*-)| < \frac{\varepsilon}{2} \quad \text{holds for all } \ t \in (\tau^*-\delta_1,\tau^*).$$

Hence, for arbitrary $t,s \in (\tau^*-\delta_1,\tau^*)$, we have

$$|f(t)-f(s)| \le |f(t)-f(\tau^*-)| + |f(s)-f(\tau^*-)| < \varepsilon. \tag{4.1.2}$$

Furthermore, by the definition of the supremum, there is a $\tau \in B \cap (\tau^*-\delta_1,\tau^*)$. Let $\boldsymbol{\alpha}$ be a division of $[a,\tau]$ such that (4.1.1) is true and let $\widetilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha} \cup \{\tau^*\}$. Then $\widetilde{\boldsymbol{\alpha}} = \{\alpha_0,\alpha_1,\dots,\tau,\tau^*\}$ is a division of $[a,\tau^*]$ with $\nu(\widetilde{\boldsymbol{\alpha}}) = \nu(\boldsymbol{\alpha})+1$, whose division points are

$$\widetilde{\alpha}_j = \begin{cases} \alpha_j & \text{if} \quad j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}, \\ \tau^* & \text{if} \quad j = \nu(\widetilde{\boldsymbol{\alpha}}). \end{cases}$$

Using (4.1.1) and (4.1.2), we get $|f(t)-f(s)| < \varepsilon$ for all $t,s \in (\widetilde{\alpha}_{j-1},\widetilde{\alpha}_j)$, where $j \in \{1,\dots,\nu(\widetilde{\boldsymbol{\alpha}})\}$. This means that $\tau^* \in B$.

Finally, we prove that $\tau^* = b$. Assume, on the contrary, that $\tau^* < b$. By Definition 4.1.1, we can choose a $\delta_2 \in (0, b-\tau^*)$ in such a way that

$$|f(t)-f(\tau^*+)| < \frac{\varepsilon}{2} \quad \text{holds for all } \ t \in (\tau^*,\tau^*+\delta_2).$$

Similarly as before in this proof, we can deduce that the inequality

$$|f(t)-f(s)| \le |f(t)-f(\tau^*+)| + |f(s)-f(\tau^*+)| < \varepsilon \tag{4.1.3}$$

holds for arbitrary $t, s \in (\tau^*, \tau^* + \delta_2)$. Let $\boldsymbol{\alpha}$ be a division of the interval $[a, \tau^*]$ such that (4.1.1) holds. Set $\tau = \tau^* + \delta_2$ and $\widetilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha} \cup \{\tau\}$. Then

$$\widetilde{\boldsymbol{\alpha}} = \{\alpha_0, \alpha_1, \dots, \tau^*, \tau\}$$

is a division of $[a, \tau]$ with $\nu(\widetilde{\boldsymbol{\alpha}}) = \nu(\boldsymbol{\alpha}) + 1$, whose division points are

$$\widetilde{\alpha}_j = \begin{cases} \alpha_j & \text{if} \quad j \in \{1, \dots, \nu(\boldsymbol{\alpha})\}, \\ \tau & \text{if} \quad j = \nu(\widetilde{\boldsymbol{\alpha}}). \end{cases}$$

Using (4.1.1) and (4.1.3), we have

$$|f(t)-f(s)| < \varepsilon \quad \text{for all } t, s \in (\widetilde{\alpha}_{j-1}, \widetilde{\alpha}_j) \text{ and } j \in \{1, \dots, \nu(\widetilde{\boldsymbol{\alpha}}\}.$$

It follows that $\tau \in B$. However, since we have $\tau > \tau^*$, this contradicts the definition of $\tau^* = \sup B$. Hence, $\tau^* = b$, and the proof of the implication (i) $\Longrightarrow$ (iii) is complete.

c) Assume that (iii) holds. Let $n \in \mathbb{N}$ be given and let $\boldsymbol{\alpha}$ be a division of $[a, b]$ such that $|f(t)-f(s)| < \frac{1}{n}$ for all $t, s \in (\alpha_{j-1}, \alpha_j)$ and $j \in \{1, \dots, \nu(\boldsymbol{\alpha})\}$.

For every $j \in \{1, \dots, \nu(\boldsymbol{\alpha})\}$ choose an arbitrary $\tau_j \in (\alpha_{j-1}, \alpha_j)$ and put

$$f_n(t) = \begin{cases} f(t) & \text{if } t \in \boldsymbol{\alpha}, \\ f(\tau_j) & \text{if } t \in (\alpha_{j-1}, \alpha_j). \end{cases}$$

Obviously, $f_n \in \mathrm{S}([a, b])$ and $\|f - f_n\|_\infty < \frac{1}{n}$ for every $n \in \mathbb{N}$, i.e. $f_n \rightrightarrows f$ on $[a, b]$ when $n \to \infty$. This proves the implication (iii) $\Longrightarrow$ (ii). □

4.1.6 Corollary. *Every regulated function* $f : [a, b] \to \mathbb{R}$ *is bounded.*

Proof. By statement (iii) of Hönig's Theorem 4.1.5, there is a division $\boldsymbol{\alpha}$ of the interval $[a, b]$ such that

$$|f(t)-f(s)| \le 1 \quad \text{whenever } t, s \in (\alpha_{j-1}, \alpha_j) \text{ and } j \in \{1, \dots, \nu(\boldsymbol{\alpha})\}.$$

For every $j \in \{1, \dots, \nu(\boldsymbol{\alpha})\}$, choose an arbitrary $\tau_j \in (\alpha_{j-1}, \alpha_j)$. Then

$$|f(t)| \le |f(\tau_j)| + 1 \quad \text{for } t \in (\alpha_{j-1}, \alpha_j) \text{ and } j \in \{1, \dots, \nu(\boldsymbol{\alpha})\}.$$

Hence $|f(t)| \le M$ for all $t \in [a, b]$, where

$$M = \max\{|f(\alpha_0)|, \dots, |f(\alpha_{\nu(\boldsymbol{\alpha})})|, |f(\tau_1)| + 1, \dots, |f(\tau_{\nu(\boldsymbol{\alpha})})| + 1\} < \infty. \quad \square$$

4.1.7 Corollary. *For every regulated function* $f:[a,b]\to\mathbb{R}$ *and every* $\varepsilon>0$*, there are at most finitely many points* $t\in[a,b]$ *such that*

$$t\in[a,b)\ \textit{ and }\ |\Delta^+ f(t)|>\varepsilon\ \textit{ or }\ t\in(a,b]\ \textit{ and }\ |\Delta^- f(x)|>\varepsilon.$$

Proof. Let $\varepsilon>0$ be given. By the statement (iii) from Hönig's Theorem 4.1.5, we can find a division $\boldsymbol{\alpha}$ of $[a,b]$ such that

$$|f(t)-f(s)|<\varepsilon \quad \text{for } t,s\in(\alpha_{j-1},\alpha_j) \text{ and } j\in\{1,\dots,\nu(\boldsymbol{\alpha})\}.$$

This implies that

$$|\Delta^+ f(t)|\le\varepsilon \quad\text{and}\quad |\Delta^- f(t)|\le\varepsilon \quad \text{for } t\in[a,b]\setminus\boldsymbol{\alpha},$$

wherefrom the statement of the corollary follows immediately. □

4.1.8 Theorem. *Every regulated function* $f:[a,b]\to\mathbb{R}$ *has at most countably many discontinuities.*

Proof. For each $k\in\mathbb{N}$, denote

$$D_k^+=\{t\in[a,b):|\Delta^+ f(t)|>\tfrac{1}{k}\} \quad\text{and}\quad D_k^-=\{t\in(a,b]:|\Delta^- f(t)|>\tfrac{1}{k}\}.$$

Then

$$D^+=\bigcup_{k\in\mathbb{N}} D_k^+=\{t\in[a,b):|\Delta^+ f(t)|>0\}$$

is the set of all points where f is discontinuous from the right, and

$$D^-=\bigcup_{k\in\mathbb{N}} D_k^-=\{t\in(a,b]:|\Delta^- f(t)|>0\}$$

is the set of all points where the function f is discontinuous from the left. Obviously, $D=D^+\cup D^-$ is the set of all discontinuity points of f on $[a,b]$.

By Corollary 4.1.7 every set D_k^+, D_k^-, $k\in\mathbb{N}$, is finite. As a result, D is at most countable. □

4.1.9 Corollary. *Let* $f\in\mathrm{G}([a,b])$ *and*

$$\widetilde{f}(t)=\begin{cases} f(t+) & \textit{if } t\in[a,b),\\ f(b) & \textit{if } t=b,\end{cases} \tag{4.1.4}$$

$$\widehat{f}(t)=\begin{cases} f(a) & \textit{if } x=a,\\ f(t-) & \textit{if } t\in(a,b].\end{cases} \tag{4.1.5}$$

Then both $\widetilde{f}$ *and* $\widehat{f}$ *are regulated on* $[a,b]$ *and*

$$\widetilde{f}(t+) = f(t+) \text{ if } t \in [a,b), \quad \widetilde{f}(t-) = f(t-) \text{ if } t \in (a,b], \tag{4.1.6}$$

and

$$\widehat{f}(t+) = f(t+) \text{ if } t \in [a,b), \quad \widehat{f}(t-) = f(t-) \text{ if } t \in (a,b]. \tag{4.1.7}$$

Proof. a) Let $\varepsilon > 0$ be given. By Hönig's Theorem 4.1.5 (iii), there exists a division $\boldsymbol{\alpha}$ of the interval $[a,b]$ such that the inequality

$$|f(t) - f(s)| < \frac{\varepsilon}{2}$$

holds whenever $t,s \in (\alpha_{j-1},\alpha_j)$ for some $j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$. In particular,

$$|f(t+\delta) - f(s+\delta)| < \frac{\varepsilon}{2}$$

holds for every pair $t,s \in [\alpha_{j-1},\alpha_j)$ with $j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$ and every $\delta > 0$ such that $t+\delta, s+\delta \in (\alpha_{j-1},\alpha_j)$. Therefore

$$|f(t+) - f(s+)| = \lim_{\delta\to 0+} |f(t+\delta) - f(s+\delta)| \le \frac{\varepsilon}{2} < \varepsilon$$

holds for each $j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$ and each pair $t,s \in [\alpha_{j-1},\alpha_j)$, as well. In other words,

$$|\widetilde{f}(t) - \widetilde{f}(s)| < \varepsilon \text{ for every } j \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \text{ and } t,s \in [\alpha_{j-1},\alpha_j). \tag{4.1.8}$$

Similarly, it can be shown that

$$|\widehat{f}(t) - \widehat{f}(s)| < \varepsilon \text{ for every } j \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \text{ and } t,s \in (\alpha_{j-1},\alpha_j]. \tag{4.1.9}$$

By Hönig's Theorem 4.1.5, it follows that both $\widetilde{f}$ and $\widehat{f}$ are regulated on $[a,b]$.

b) Let $x \in [a,b)$ and $\varepsilon > 0$ be given, and let $\boldsymbol{\alpha}$ be a division of $[a,b]$ such that $|f(t) - f(s)| < \frac{\varepsilon}{2}$ for every pair $t,s \in (\alpha_{j-1},\alpha_j)$ and every $j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$. There is a unique index $i \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$ such that $x \in [\alpha_{i-1},\alpha_i)$. By (4.1.8), we have

$$|\widetilde{f}(t) - f(x+)| = |\widetilde{f}(t) - \widetilde{f}(x)| < \varepsilon \quad \text{for } t \in (x,\alpha_i).$$

In other words, $\widetilde{f}(x+) = f(x+)$. This proves the first statement from (4.1.6).

c) Analogously, let $x \in (a,b]$, $\varepsilon > 0$, and let $\boldsymbol{\alpha}$ be a division of $[a,b]$ such that $|f(t) - f(s)| < \frac{\varepsilon}{2}$ holds for every pair $t,s \in (\alpha_{j-1},\alpha_j)$ and every $j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$. There is a unique $i \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$ such that $x \in (\alpha_{i-1},\alpha_i]$. If $t \in (\alpha_{i-1},x)$ and $0 < \delta < \min\{x-t, x-\alpha_{i-1}\}$, then

$$\alpha_{i-1} < x-\delta < \alpha_i \quad \text{and} \quad \alpha_{i-1} < t+\delta < x.$$

Therefore, by the definition of the division $\boldsymbol{\alpha}$, we have

$$|f(x-)-\widetilde{f}(t)|=\lim_{\delta\to 0+}|f(x-\delta)-f(t+\delta)|\le\frac{\varepsilon}{2}<\varepsilon\quad\text{for } t\in(\alpha_{i-1},x).$$

In other words $\widetilde{f}(x-)=f(x-)$ and this completes the proof of the second statement from (4.1.6).

d) The relations (4.1.7) can be proved similarly to (4.1.6). □

4.2 The space of regulated functions and its subspaces

The set $\mathrm{G}([a,b])$ is a linear space equipped with the natural operations of pointwise addition and multiplication by scalars, i.e.,

$$\left.\begin{aligned}&(f+g)(t)=f(t)+g(t) &&\text{for } f,g\in\mathrm{G}([a,b]),\ t\in[a,b],\\ &(c\,f)(t)=c\,f(t) &&\text{for } c\in\mathbb{R},\ f\in\mathrm{G}([a,b]),\ t\in[a,b].\end{aligned}\right\}\quad(4.2.1)$$

It is also easy to verify that

$$\|f\|_{\mathrm{G}}:=\|f\|_\infty=\sup_{x\in[a,b]}|f(x)|\qquad(4.2.2)$$

defines a norm on $\mathrm{G}([a,b])$.

4.2.1 Theorem. $\mathrm{G}([a,b])$ *is a Banach space with respect to the operations* (4.2.1) *and the norm* (4.2.2).

Proof. It suffices to show that the space $\mathrm{G}([a,b])$ is complete with respect to the norm given by (4.2.2). Thus, assume that $\{f_n\}$ is a Cauchy sequence in $\mathrm{G}([a,b])$.

Using the completeness of the space $\mathbb{R}$, analogously to the parts a) and b) of the proof of Theorem 2.2.2, we can prove that there is a function $f\colon[a,b]\to\mathbb{R}$ such that $f_n\rightrightarrows f$. By Theorem 4.1.3 it follows that $f\in\mathrm{G}([a,b])$ and this completes the proof. □

4.2.2 Remark. (i) By Definition 2.5.1, $f\in\mathrm{S}([a,b])$ if and only if there exists a division $\boldsymbol{\alpha}$ of the interval $[a,b]$ such that f is constant on every subinterval (α_{j-1},α_j). Every function from $\mathrm{S}([a,b])$ is a finite linear combination of functions of the form $\chi_{(\alpha,\beta)}$ and $\chi_{[\tau]}$, where (α,β) is an arbitrary subinterval in $[a,b]$ and τ is any point in $[a,b]$. Note that

$$\chi_{(\alpha,\beta)}=\chi_{(\alpha,b]}-\chi_{[\beta,b]}\quad\text{for all } \alpha,\beta\in[a,b],\ \alpha<\beta$$

and

$$\chi_{[\tau]} = \chi_{[\tau,b]} - \chi_{(\tau,b]} \quad \text{for all } \tau \in [a,b).$$

Hence $f \in \mathrm{S}([a,b])$ if and only if f is a finite linear combination of functions of the forms $\chi_{[b]}$, $\chi_{[\tau,b]}$, $\chi_{(\tau,b]}$, where τ can be an arbitrary point in $[a,b)$, i.e.,

$$\mathrm{S}([a,b]) = \mathrm{Lin}\Big(\{\chi_{[\tau,b]},\ \chi_{(\tau,b]},\ \chi_{[b]} : \tau \in [a,b)\}\Big), \tag{4.2.3}$$

where Lin (M) denotes the linear span of the set M.

(ii) Similarly, we can show that also

$$\mathrm{S}([a,b]) = \mathrm{Lin}\Big(\{\chi_{[a,\tau]},\ \chi_{[a,\tau)},\ \chi_{[a]} : \tau \in (a,b]\}\Big). \tag{4.2.4}$$

(iii) By Hönig's Theorem 4.1.5, the set $\mathrm{S}([a,b])$ is dense in $\mathrm{G}([a,b])$, i.e., $\mathrm{cl}(\mathrm{S}([a,b])) = \mathrm{G}([a,b])$, where $\mathrm{cl}(M)$ stands for the closure of a set M.

4.2.3 Lemma. *Let* $\{f_n\} \subset \mathrm{G}([a,b])$ *and* $f_n \rightrightarrows f$ *on* $[a,b]$. *For* $n \in \mathbb{N}$, *set*

$$\widetilde{f}_n(t) = \begin{cases} f_n(t+), & \text{if } t \in [a,b),\\ f_n(b), & \text{if } t = b, \end{cases} \qquad \widehat{f}_n(t) = \begin{cases} f_n(a), & \text{if } t = a,\\ f_n(t-), & \text{if } t \in (a,b], \end{cases}$$

and

$$\widetilde{f}(t) = \begin{cases} f(t+), & \text{if } t \in [a,b),\\ f(b), & \text{if } t = b, \end{cases} \qquad \widehat{f}(t) = \begin{cases} f(a), & \text{if } t = a,\\ f(t-), & \text{if } t \in (a,b]. \end{cases}$$

Then $\widetilde{f}_n \rightrightarrows \widetilde{f}$ *and* $\widehat{f}_n \rightrightarrows \widehat{f}$ *on* $[a,b]$.

Proof. By Corollary 4.1.9, the functions $\widetilde{f}$, $\widetilde{f}_n$, $\widehat{f}$, $\widehat{f}_n$, $n \in \mathbb{N}$, are regulated on $[a,b]$. Let $\varepsilon > 0$ be given. Choose $n_\varepsilon \in \mathbb{N}$ such that $|f_n(s) - f(s)| < \frac{\varepsilon}{2}$ for every $n \geq n_\varepsilon$ and every $s \in [a,b]$. Letting $s \to t$ from the right we get that

$$|\widetilde{f}_n(t) - \widetilde{f}(t)| = \lim_{s \to t+} |f_n(s) - f(s)| \leq \tfrac{\varepsilon}{2} < \varepsilon$$

holds for every $t \in [a,b)$ and every $n \geq n_\varepsilon$. Consequently,

$$\lim_{n\to\infty} \|\widetilde{f}_n - \widetilde{f}\|_\infty = 0, \ \text{ i.e., } \ \widetilde{f}_n \rightrightarrows \widetilde{f} \quad \text{on } [a,b].$$

Similarly, we would show that $\widehat{f}_n \rightrightarrows \widehat{f}$ on $[a,b]$. □

In the remaining part of this chapter, we present several statements which will be useful later (in particular, in Chapters 6 and 7). Note that if the assumptions of Lemma 4.2.3 are satisfied, then it follows that

$$f(t+) = \lim_{n\to\infty} f_n(t+) \ \text{ for each } t \in [a,b),$$

and

$$f(t-) = \lim_{n\to\infty} f_n(t-) \ \text{ for each } t \in (a,b]\,;$$

this observation leads to the following corollary.

4.2.4 Corollary. *The sets*

$$\mathrm{G_L}([a,b]) = \{f \in \mathrm{G}([a,b]) : f(t-) = f(t) \textit{ for } t \in (a,b]\},$$

$$\widetilde{\mathrm{G}}_\mathrm{L}([a,b]) = \{f \in \mathrm{G}([a,b]) : f(t-) = f(t) \textit{ for } t \in (a,b)\},$$

$$\mathrm{G_R}([a,b]) = \{f \in \mathrm{G}([a,b]) : f(t+) = f(t) \textit{ for } t \in [a,b)\},$$

$$\widetilde{\mathrm{G}}_\mathrm{R}([a,b]) = \{f \in \mathrm{G}([a,b]) : f(t+) = f(t) \textit{ for } t \in (a,b)\},$$

$$\mathrm{G_{reg}}([a,b]) = \{f \in \mathrm{G}([a,b]) : f(t-) + f(t+) = 2\,f(t) \textit{ for } t \in (a,b),\ f(a+) = f(a),\ f(b-) = f(b)\},$$

$$\widetilde{\mathrm{G}}_\mathrm{reg}([a,b]) = \{f \in \mathrm{G}([a,b]) : f(t-) + f(t+) = 2\,f(t) \textit{ for } t \in (a,b)\}$$

are closed in $\mathrm{G}([a,b])$.

4.2.5 Remark. If a regulated function f satisfies $f(t-)+f(t+)=2f(t)$ for $t \in (a,b)$, we say that f is *regular* on (a,b). Functions from the space $\mathrm{G_{reg}}([a,b])$ are said to be regular on the closed interval $[a,b]$.

4.2.6 Lemma. *The following relations hold:*

$$\mathrm{cl}(\mathrm{G_L}([a,b]) \cap \mathrm{S}([a,b])) = \mathrm{G_L}([a,b]),$$

$$\mathrm{cl}(\widetilde{\mathrm{G}}_\mathrm{L}([a,b]) \cap \mathrm{S}([a,b])) = \widetilde{\mathrm{G}}_\mathrm{L}([a,b]),$$

$$\mathrm{cl}(\mathrm{G_R}([a,b]) \cap \mathrm{S}([a,b])) = \mathrm{G_R}([a,b]),$$

$$\mathrm{cl}(\widetilde{\mathrm{G}}_\mathrm{R}([a,b]) \cap \mathrm{S}([a,b])) = \widetilde{\mathrm{G}}_\mathrm{R}([a,b]),$$

$$\mathrm{cl}(\mathrm{G_{reg}}([a,b]) \cap \mathrm{S}([a,b])) = \mathrm{G_{reg}}([a,b]),$$

$$\mathrm{cl}(\widetilde{\mathrm{G}}_\mathrm{reg}([a,b]) \cap \mathrm{S}([a,b])) = \widetilde{\mathrm{G}}_\mathrm{reg}([a,b]).$$

Proof. We will prove only the next to last assertion, the other ones can be proved similarly.

Let arbitrary $f \in \mathrm{G_{reg}}([a,b])$ and $\varepsilon > 0$ be given. By Hönig's Theorem (Theorem 4.1.5), there is a $\varphi \in \mathrm{S}([a,b])$ such that

$$\|f - \varphi\|_\infty < \varepsilon. \tag{4.2.5}$$

It follows that

$$\left.\begin{aligned} &|f(t-)-\varphi(t-)| = \lim_{s\to t-} |f(s)-\varphi(s)| \le \varepsilon \quad \text{for } t\in(a,b], \\ \text{and} \\ &|f(t+)-\varphi(t+)| = \lim_{s\to t+} |f(s)-\varphi(s)| \le \varepsilon \quad \text{for } t\in[a,b). \end{aligned}\right\} \tag{4.2.6}$$

Define

$$\widetilde{\varphi}(t) = \begin{cases} \varphi(a+) & \text{if } t=a, \\ \frac{1}{2}\Big(\varphi(t+)+\varphi(t-)\Big) & \text{if } t\in(a,b), \\ \varphi(b-) & \text{if } t=b. \end{cases} \tag{4.2.7}$$

Then $\widetilde{\varphi}\in \mathrm{S}([a,b])\cap \mathrm{G}_{\mathrm{reg}}([a,b])$. Furthermore, by (4.2.5) and (4.2.7), we have

$$|f(a)-\widetilde{\varphi}(a)| = |f(a+)-\varphi(a+)| \le \varepsilon, \quad |f(b)-\widetilde{\varphi}(b)| = |f(b-)-\varphi(b-)| \le \varepsilon$$

and, by (4.2.6) and (4.2.7),

$$\begin{aligned} |f(t)-\widetilde{\varphi}(t)| &= \left|\tfrac{1}{2}\left(f(t+)+f(t-)\right)-\tfrac{1}{2}\left(\varphi(t+)+\varphi(t-)\right)\right| \\ &\le \tfrac{1}{2}\Big(\big|f(t+)-\varphi(t+)\big|+\big|f(t-)-\varphi(t-)\big|\Big) \le \varepsilon \end{aligned}$$

for $t\in(a,b)$. In other words, we have $\|f-\widetilde{\varphi}\|_\infty \le \varepsilon$, wherefrom the desired equality $\mathrm{cl}(\mathrm{G}_{\mathrm{reg}}([a,b])\cap \mathrm{S}([a,b])) = \mathrm{G}_{\mathrm{reg}}([a,b])$ follows. □

4.2.7 Exercise. Prove the remaining assertions of Lemma 4.2.6.

4.2.8 Lemma. *The following relations hold:*

$$\begin{aligned} \mathrm{G_L}([a,b])\cap \mathrm{S}([a,b]) &= \mathrm{Lin}\Big(\{\chi_{[a,\tau]} : \tau\in[a,b]\}\Big), \\ \widetilde{\mathrm{G}}_{\mathrm{L}}[a,b]\cap \mathrm{S}([a,b]) &= \mathrm{Lin}\Big(\{\chi_{[a,\tau]},\ \chi_{[b]} : \tau\in[a,b]\}\Big), \\ \mathrm{G_R}([a,b])\cap \mathrm{S}([a,b]) &= \mathrm{Lin}\Big(\{\chi_{[\tau,b]} : \tau\in[a,b]\}\Big), \\ \widetilde{\mathrm{G}}_{\mathrm{R}}[a,b]\cap \mathrm{S}([a,b]) &= \mathrm{Lin}\Big(\{\chi_{[a]},\chi_{[\tau,b]} : \tau\in[a,b]\}\Big), \\ \mathrm{G}_{\mathrm{reg}}([a,b])\cap \mathrm{S}([a,b]) &= \mathrm{Lin}\Big(\{\chi_{[a,b]}, \frac{1}{2}\chi_{[\tau]}+\chi_{(\tau,b]} : \tau\in(a,b)\}\Big), \\ \widetilde{\mathrm{G}}_{\mathrm{reg}}[a,b]\cap \mathrm{S}([a,b]) &= \mathrm{Lin}\Big(\{\chi_{[a,b]},\ \chi_{(a,b]},\ \frac{1}{2}\chi_{[\tau]}+\chi_{(\tau,b]},\ \chi_{[b]} : \tau\in(a,b)\}\Big). \end{aligned}$$

Proof. Notice that the first statement follows from Remark 4.2.2 (ii); from the set (4.2.4) of functions generating the whole set $\mathrm{S}([a,b])$, we have selected those which are left-continuous on $(a,b]$.

Next, we will show the proof of the fifth relation. To this aim, let $f \in \mathrm{S}([a,b]) \cap \mathrm{G_{reg}}([a,b])$ be given. Then there are $m \in \mathbb{N}$, $x_0, x_1, \ldots, x_{m+1} \in \mathbb{R}$ and a division $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \ldots, \alpha_m\}$ of $[a,b]$ such that

$$f(t) = \begin{cases} x_1, & \text{if } \ t \in [a, \alpha_1), \\ x_j, & \text{if } \ t \in (\alpha_{j-1}, \alpha_j) \text{ for a certain } j \in \{2, \ldots, m\}, \\ \frac{x_j + x_{j+1}}{2}, & \text{if } \ t = \alpha_j \text{ for a certain } j \in \{1, \ldots, m-1\}, \\ x_m, & \text{if } \ t \in (\alpha_{m-1}, b]. \end{cases}$$

Therefore,

$$\left. \begin{aligned} f(t) &= \chi_{[a,\alpha_1)}(t)\, x_1 + \sum_{j=2}^{m-1} \chi_{(\alpha_{j-1},\alpha_j)}(t)\, x_j + \chi_{(\alpha_{m-1},b]}(t)\, x_m \\ &\quad + \frac{1}{2} \Big(\sum_{j=1}^{m-1} \chi_{[\alpha_j]}(t) \, \big(x_j + x_{j+1}\big) \Big) \quad \text{for } \ t \in [a,b]. \end{aligned} \right\} \tag{4.2.8}$$

This relation can be rearranged as follows:

$$\begin{aligned} f(t) &= \chi_{[a,b]}(t)\, x_1 - \chi_{[\alpha_1]}(t)\, x_1 - \chi_{(\alpha_1,b]}(t)\, x_1 \\ &\quad + \sum_{j=2}^{m} \chi_{(\alpha_{j-1},b]}(t)\, x_j - \sum_{j=2}^{m-1} \chi_{[\alpha_j]}(t)\, x_j - \sum_{j=2}^{m-1} \chi_{(\alpha_j,b]}(t)\, x_j \\ &\quad + \frac{1}{2} \Big(\sum_{j=1}^{m-1} \chi_{[\alpha_j]}(t) \, \big(x_j + x_{j+1}\big) \Big) \\ &= \chi_{[a,b]}(t)\, x_1 + \sum_{j=1}^{m-1} \chi_{(\alpha_j,b]}(t)\, x_{j+1} - \sum_{j=1}^{m-1} \chi_{(\alpha_j,b]}(t)\, x_j \\ &\quad - \sum_{j=1}^{m-1} \chi_{[\alpha_j]}(t)\, x_j + \frac{1}{2} \Big(\sum_{j=1}^{m-1} \chi_{[\alpha_j]}(t) \, \big(x_j + x_{j+1}\big) \Big) \\ &= \chi_{[a,b]}(t)\, x_1 + \sum_{j=1}^{m-1} \big[\chi_{(\alpha_j,b]}(t) + \tfrac{1}{2}\, \chi_{[\alpha_j]}(t)\big] \ (x_{j+1} - x_j) \\ &= \chi_{[a,b]}(t)\, \widetilde{x}_1 + \sum_{j=2}^{m} \big[\chi_{(\alpha_{j-1},b]}(t) + \tfrac{1}{2}\, \chi_{[\alpha_{j-1}]}(t)\big]\, \widetilde{x}_j, \end{aligned}$$

where

$$\widetilde{x}_1 = x_1, \quad \text{and} \quad \widetilde{x}_j = x_j - x_{j-1} \quad \text{for } \ j \in \{2, \ldots, m\}. \tag{4.2.9}$$

This means that

$$f \in \mathrm{Lin}\Big(\{\chi_{[a,b]}, \frac{1}{2}\chi_{[\tau]}+\chi_{(\tau,b]} : \tau \in (a,b)\}\Big),$$

wherefrom the fifth statement of the lemma follows.

The other statements of the lemma may be proved in a similar way. □

4.3 Relatively compact subsets of $\mathrm{G}([a,b])$

Recall that a subset M of a Banach space X is *relatively compact* if any sequence of its elements contains a convergent subsequence. It is known (see e.g. [30], Theorem I.6.15 or [154], Theorem, p.13) that M is relatively compact if and only if it is *totally bounded*, i.e., if for each $\varepsilon > 0$ there is a finite set $D_\varepsilon \subset X$ such that for every $x \in M$ there exists a $d \in D_\varepsilon$ satisfying $\|x-d\|_X < \varepsilon$. Such a set D_ε is called an ε*-net* for M in X.

The following assertion is not surprising.

4.3.1 Lemma. *Each totally bounded set is bounded.*

Proof. Let $M \subset X$ be totally bounded and let

$$D = \{d_1, d_2, \dots, d_m\} \subset X$$

be such that for each $x \in M$, there is $\widetilde{d}_x \in D$ satisfying $\|x-\widetilde{d}_x\|_X < 1$. Thus, for an arbitrary $x \in M$ we have

$$\|x\|_X \le \|x-\widetilde{d}_x\|_X + \|\widetilde{d}_x\|_X \le K,$$

where $K = 1 + \max\{\|d_1\|_X, \dots, \|d_n\|_X\}$ does not depend on $x \in M$. □

In the space $\mathrm{C}([a,b])$ of continuous functions we have the following criterion for relative compactness known as the Arzelà-Ascoli theorem. Its proof can be found in many functional analysis textbooks, see e.g. Theorem 8.2.12 in [143].

4.3.2 Theorem (Arzelà-Ascoli). *A subset M of the space $\mathrm{C}([a,b])$ is relatively compact if and only if the following conditions are satisfied:*

(i) *There is a $c^* \in [0,\infty)$ such that $\|f\|_\infty \le c^*$ for each $f \in M$.*

(ii) *For each $\varepsilon > 0$ there is a $\delta > 0$ such that $|f(t)-f(s)| < \varepsilon$ holds for each $f \in M$ and each $t, s \in [a,b]$ satisfying $|t-s| < \delta$.*

If the condition (i) from the Arzelà-Ascoli theorem is satisfied, we say that the set M is *uniformly bounded*, while if the condition (ii) is satisfied, we say that the set M is *equicontinuous*. Thus, Theorem 4.3.2 can be reformulated as follows:

A subset of $\mathrm{C}([a,b])$ *is relatively compact if and only if it is uniformly bounded and equicontinuous.*

We will derive an analogous criterion for subsets of the space $\mathrm{G}([a,b])$ with the notion of equicontinuity replaced by the related notion of *equiregulatedness*. The definition of this notion resembles the definition of equicontinuity with ordinary limits replaced by the one-sided ones.

4.3.3 Definition. A subset M of $\mathrm{G}([a,b])$ is called *equiregulated* if the following conditions hold:

- For each $\varepsilon > 0$ and $\tau \in (a,b]$ there is a $\delta_1(\tau) \in (0, \tau - a)$ such that

$$|f(\tau-) - f(t)| < \varepsilon \quad \text{for all } t \in (\tau - \delta_1(\tau), \tau) \text{ and } f \in M.$$

- For each $\varepsilon > 0$ and $\tau \in [a,b)$ there is a $\delta_2(\tau) \in (0, b - \tau)$ such that

$$|f(\tau+) - f(t)| < \varepsilon \quad \text{for all } t \in (\tau, \tau + \delta_2(\tau)) \text{ and } f \in M.$$

The next characterization of equiregulated sets of functions will be helpful later.

4.3.4 Lemma. *The following statements are equivalent:*

(i) $M \subset G([a,b])$ *is equiregulated.*

(ii) *For every* $\varepsilon > 0$ *there exists a division* $\boldsymbol{\alpha}$ *of* $[a,b]$ *such that for every* $f \in M$, $j \in \{1, \ldots, \nu(\boldsymbol{\alpha})\}$ *and* $s, t \in (\alpha_{j-1}, \alpha_j)$, *we have* $|f(s) - f(t)| < \varepsilon$.

Proof. a) The proof of the implication (i) $\Longrightarrow$ (ii) is almost identical with the proof of the implication (i) $\Longrightarrow$ (iii) in Theorem 4.1.5; we leave it as an exercise for the reader.

b) Let an arbitrary $\varepsilon > 0$ be given, and let $\boldsymbol{\alpha}$ be the corresponding division from condition (ii).

Choose an arbitrary $\tau \in (a,b]$. There is a unique $j \in \{1, \ldots, \nu(\boldsymbol{\alpha})\}$ such that $\tau \in (\alpha_{j-1}, \alpha_j]$. For all $t, s \in (\alpha_{j-1}, \tau)$ and $f \in M$ we have $|f(s) - f(t)| < \varepsilon$. Letting $s \to \tau-$ we get

$$|f(\tau-) - f(t)| \le \varepsilon \quad \text{for all } t \in (\tau - \delta_1, \tau) \text{ and } f \in M,$$

where $\delta_1 = \tau - \alpha_{j-1}$.

Analogously, if $\tau \in [a,b)$, there exists a unique $j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$ such that $\tau \in [\alpha_{j-1},\alpha_j)$. Hence, $|f(s)-f(t)|<\varepsilon$ holds for all $t,s \in (\tau,\alpha_j)$ and $f \in M$. Letting $s \to \tau+$ yields

$$|f(\tau+)-f(t)| \le \varepsilon \quad \text{for all } t \in (\tau,\tau+\delta_2), \quad f \in M,$$

where $\delta_2 = \alpha_j - \tau$.

This shows that $M \subset G([a,b])$ is equiregulated. □

4.3.5 Exercise. Prove the implication (i) $\Longrightarrow$ (ii) from Lemma 4.3.4.

We now proceed to the analogue of the Arzelà-Ascoli theorem in the world of regulated functions, which reads as follows.

4.3.6 Theorem (FRAŇKOVÁ)**.** *A subset M of $\mathrm{G}([a,b])$ is relatively compact if and only if it is uniformly bounded and equiregulated.*

Proof. a) Let $M \subset \mathrm{G}([a,b])$ be relatively compact. We will show that M is uniformly bounded and equiregulated. The uniform boundedness of functions from M follows from Lemma 4.3.1. It remains to show that M is equiregulated.

Let $\varepsilon > 0$ and $\tau \in [a,b]$ be given and let $F = \{f_1, f_2, \dots, f_m\}$ be an $\varepsilon/3-$net for the set M in $\mathrm{G}([a,b])$. This means that

$$\text{for any } f \in M \text{ there is an } \widetilde{f} \in F \text{ such that } \|f-\widetilde{f}\|_\infty < \frac{\varepsilon}{3}. \tag{4.3.1}$$

Consequently, the inequalities

$$\left.\begin{aligned} &|f(t-)-\widetilde{f}(t-)| \le \tfrac{\varepsilon}{3} \quad \text{for } t \in (a,b],\\ &|f(t+)-\widetilde{f}(t+)| \le \tfrac{\varepsilon}{3} \quad \text{for } t \in [a,b) \end{aligned}\right\} \tag{4.3.2}$$

hold for any $f \in M$ and any $\widetilde{f}$ satisfying (4.3.1). All the functions $f_k \in F$ are regulated on $[a,b]$. Hence, for a given $\tau \in (a,b]$ and every $k \in \{1,2,\dots,m\}$, there is a $\delta_k^1 \in (0,\tau-a)$ such that

$$|f_k(t)-f_k(\tau-)| < \varepsilon \quad \text{for } t \in (\tau-\delta_k^1,\tau). \tag{4.3.3}$$

Similarly, for a given $\tau \in [a,b)$ and every $k \in \{1,2,\dots,m\}$, there is a $\delta_k^2 \in (0,b-\tau)$ such that

$$|f_k(t)-f_k(\tau+)| < \varepsilon \quad \text{for } t \in (\tau,\tau+\delta_k^2). \tag{4.3.4}$$

Set

$$\delta = \begin{cases} \min\{\delta_k^i : i=1,2;\ k=1,\dots,m\} & \text{if } \tau \in (a,b),\\ \min\{\delta_k^1 : k=1,\dots,m\} & \text{if } \tau = b,\\ \min\{\delta_k^2 : k=1,\dots,m\} & \text{if } \tau = a. \end{cases}$$

If $a \le \tau-\delta < t < \tau \le b$, then by (4.3.1)–(4.3.4) the inequalities

$$|f(t)-f(\tau-)| \le |f(t)-\widetilde{f}(t)| + |\widetilde{f}(t)-\widetilde{f}(\tau-)| + |\widetilde{f}(\tau-)-f(\tau-)| < \varepsilon$$

hold for any $f \in M$ and any $\widetilde{f}$ corresponding to f by (4.3.1). Similarly, we can prove that

$$|f(t)-f(\tau+)| < \varepsilon \quad \text{for all } f \in M$$

whenever $a \le \tau-\delta < t < \tau \le b$. Consequently, the set M is equiregulated.

b) Now, assume that M is uniformly bounded and equiregulated. We will show that M is relatively compact in $\mathrm{G}([a,b])$. It suffices to show that M is totally bounded, i.e., that for every $\varepsilon > 0$ the set M has a finite $\varepsilon-$net in $\mathrm{G}([a,b])$.

Let an arbitrary $\varepsilon > 0$ be given, and let $\boldsymbol{\alpha}$ be the corresponding division of $[a,b]$ from part (ii) of Lemma 4.3.4.

Since M is uniformly bounded, there is a $c^* > 0$ such that $\|f\|_\infty \le c^*$ for all $f \in M$. Let $\boldsymbol{z} = \{z_0, z_1, \dots, z_n\}$ be a division of $[-c^*, c^*]$ such that

$$|\boldsymbol{z}| = \max_{1\le j\le n}(z_j - z_{j-1}) < \frac{\varepsilon}{2}.$$

Let F be the set of all functions $\widetilde{f}:[a,b]\to\mathbb{R}$ which are constant on each of the intervals (α_{j-1},α_j), $j=1,\dots,\nu(\boldsymbol{\alpha})$, and whose values belong to the set $\boldsymbol{z}$. The number of elements of F is obviously finite.

We will show that F is an $\varepsilon-$net for M in $\mathrm{G}([a,b])$. To this aim, consider an arbitrary function $f \in M$. By the definition of $\boldsymbol{z}$, we know that

- for each $j \in \{0,1,\dots,\nu(\boldsymbol{\alpha})\}$ there is a $k_j \in \{0,1,\dots,n\}$ such that

$$\big|f(\alpha_j) - z_{k_j}\big| < \frac{\varepsilon}{2},$$

- for each $j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$ there is an $\ell_j \in \{0,1,\dots,n\}$ such that

$$\Big|f\left(\frac{\alpha_{j-1}+\alpha_j}{2}\right) - z_{\ell_j}\Big| < \frac{\varepsilon}{2}.$$

Furthermore, by the definition of $\boldsymbol{\alpha}$, we have

$$|f(t)-z_{\ell_j}| \le \Big|f(t) - f\left(\frac{\alpha_{j-1}+\alpha_j}{2}\right)\Big| + \Big|f\left(\frac{\alpha_{j-1}+\alpha_j}{2}\right) - z_{\ell_j}\Big| < \varepsilon$$

for all $j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$ and $t \in (\alpha_{j-1},\alpha_j)$, as well. Let us define

$$\widetilde{f}(t) = \begin{cases} z_{k_j} & \text{if} \quad t=\alpha_j \quad \text{for some } j \in \{0,1,\dots,m\}, \\ z_{\ell_j} & \text{if} \quad t \in (\alpha_{j-1},\alpha_j) \quad \text{for some } j \in \{1,\dots,m\}. \end{cases}$$

Obviously, $\widetilde{f} \in F$ and $\|f - \widetilde{f}\|_\infty < \varepsilon$. Thus F is an $\varepsilon-$net for M in $\mathrm{G}([a,b])$, and the proof is complete. □

The next assertion shows that the condition of uniform boundedness can be weakened.

4.3.7 Corollary. *A subset M of the space $\mathrm{G}([a,b])$ is relatively compact if and only if it is equiregulated and*

$$\text{the set } \{f(t) : f \in M\} \text{ is bounded for each } t \in [a,b]. \tag{4.3.5}$$

Proof. If $M \subset \mathrm{G}([a,b])$ is uniformly bounded, then it obviously satisfies condition (4.3.5). Hence, by Fraňková's Theorem 4.3.6, any relatively compact subset M of $\mathrm{G}([a,b])$ is equiregulated and satisfies (4.3.5). It remains to prove the reverse implication. To this aim, assume that M is equiregulated and satisfies condition (4.3.5). We will show that M is uniformly bounded. By Lemma 4.3.4, we can choose a division $\boldsymbol{\alpha}$ of $[a,b]$ such that

$$\left.\begin{aligned} &|f(t) - f(s)| < 1 \\ &\qquad \text{for all } t, s \in (\alpha_{j-1}, \alpha_j),\ j \in \{1, \ldots, m\} \ \text{ and } \ f \in M, \end{aligned}\right\} \tag{4.3.6}$$

where $m = \nu(\boldsymbol{\alpha})$. By our assumption (4.3.5), there exist constants

$$\gamma_j,\ j = 0, 1, \ldots, m, \ \text{ and } \ \widetilde{\gamma}_j,\ j = 1, \ldots, m,$$

such that the estimates

$$\left.\begin{aligned} &|f(\alpha_j)| \le \gamma_j && \text{for } j = 0, 1, \ldots, m, \\ &\left|f(\tfrac{1}{2}(\alpha_{j-1} + \alpha_j))\right| \le \widetilde{\gamma}_j && \text{for } j = 1, \ldots, m \end{aligned}\right\} \tag{4.3.7}$$

hold for all $f \in M$. This, together with (4.3.6), implies that the estimate

$$\left.\begin{aligned} &|f(t)| < \left|f(\tfrac{1}{2}(\alpha_{j-1} + \alpha_j))\right| + 1 \le \widetilde{\gamma}_j + 1 \\ &\qquad \text{if } t \in (\alpha_{j-1}, \alpha_j) \ \text{ and } \ j \in \{1, \ldots, m\} \end{aligned}\right\} \tag{4.3.8}$$

holds for each $f \in M$. According to (4.3.7) and (4.3.8) we have $\|f\|_\infty < c^*$ for any $f \in M$, where

$$\gamma^* = \max\{\gamma_j : j = 0, 1, \ldots, m\}, \quad \widetilde{\gamma}^* = \max\{\widetilde{\gamma}_j : j = 1, \ldots, m\},$$
$$c^* = \max\{\gamma^*, \widetilde{\gamma}^*\} + 1.$$

Hence, the set M is uniformly bounded and the proof is complete. □

We conclude this section by another useful criterion for the relative compactness in $\mathrm{G}([a,b])$.

4.3.8 Corollary. *Let* $M \subset \mathrm{G}([a,b])$. *Assume that the set* $\{f(a): f \in M\}$ *is bounded and there exists a nondecreasing function* $h:[a,b] \to \mathbb{R}$ *such that*

$$|f(t)-f(s)| \le |h(t)-h(s)| \text{ for all } t,s \in [a,b] \text{ and } f \in M. \tag{4.3.9}$$

Then M *is relatively compact in* $\mathrm{G}([a,b])$.

Proof. By assumption, there is $K \in [0,\infty)$ such that $|f(a)| \le K$ for any $f \in M$. Consequently,

$$|f(x)| \le |f(a)| + |f(x)-f(a)| \le K + h(b) - h(a)$$

for all $x \in [a,b]$ and $f \in M$. Thus, the set M is uniformly bounded.

Now, let an arbitrary $\varepsilon > 0$ be given. Obviously, $h \in \mathrm{G}([a,b])$. Hence, by Hönig's Theorem 4.1.5, there is a division $\boldsymbol{\alpha}$ of $[a,b]$ such that

$$|h(t)-h(s)| < \varepsilon \text{ for all } t,s \in (\alpha_{j-1},\alpha_j) \text{ and } j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}.$$

It follows that

$$|f(t)-f(s)| \le |h(t)-h(s)| < \varepsilon$$

for arbitrary $t,s \in (\alpha_{j-1},\alpha_j)$, $j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$, and $f \in M$. By Lemma 4.3.4, the set M is equiregulated. Finally, by Fraňková's Theorem 4.3.6, M is relatively compact in $\mathrm{G}([a,b])$. □

4.3.9 Remark. Corollary 4.3.8 provides a sufficient condition for the relative compactness of sets in $\mathrm{G}([a,b])$. Note that this condition is not necessary: If (4.3.9) holds, it is easy to verify that $\mathrm{var}_a^b f \le h(b)-h(a)$ for each $f \in M$. However, in general, regulated functions need not have bounded variation.

The next assertion shows that if $M \subset \mathrm{G}([a,b])$ is a pointwise convergent sequence of functions which satisfy (4.3.9), then necessarily this sequence converges uniformly.

4.3.10 Corollary. *Assume that* $\{f_n\} \subset \mathrm{G}([a,b])$ *is a sequence which is pointwise convergent to* $f:[a,b] \to \mathbb{R}$. *Moreover, suppose there exists a nondecreasing function* $h:[a,b] \to \mathbb{R}$ *such that*

$$|f_n(t)-f_n(s)| \le |h(t)-h(s)| \quad \text{for all } t,s \in [a,b] \text{ and } n \in \mathbb{N}.$$

Then $\{f_n\}$ *converges uniformly to* f *on* $[a,b]$.

Proof. Corollary 4.3.8 implies that each subsequence of $\{f_n\}$ has a subsequence which is uniformly convergent. Obviously, the uniform limit of this subsequence is necessarily f.

Now, suppose that $\{f_n\}$ is not uniformly convergent to f. Then there exists an $\varepsilon > 0$ such that for each $k \in \mathbb{N}$, there is an index $n_k \in \mathbb{N}$ with the property that $\|f_{n_k} - f\|_\infty \geq \varepsilon$. It follows that $\{f_{n_k}\}_k$ has no subsequence which is uniformly convergent to f, and this is a contradiction. □

More information about regulated functions can be found in Hönig's monograph [60] (see Section I.3 there). Other useful results (e.g., characterization of compact sets in $\mathrm{G}([a,b])$ or a generalization of Helly's Choice Theorem) are included in D. Fraňková's paper [39].

Chapter 5

Riemann-Stieltjes integral

The answer to some problems mentioned in the introductory chapter is provided by the Riemann-Stieltjes integral, which is a natural generalization of the well-known Riemann integral.

5.1 Definition and basic properties

Recall that a set $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \dots, \alpha_m\}$ of points from an interval $[a,b]$ is called a *division of the interval* $[a,b]$ if

$$a = \alpha_0 < \alpha_1 < \dots < \alpha_m = b.$$

The set of all divisions of the interval $[a,b]$ is denoted by $\mathscr{D}[a,b]$. The elements of a division $\boldsymbol{\alpha}$ of $[a,b]$ are usually denoted by α_j, $\nu(\boldsymbol{\alpha})$ is the index of the maximum element (i.e., $\alpha_{\nu(\boldsymbol{\alpha})} = b$) and

$$|\boldsymbol{\alpha}| = \max_{j=1,\dots,m} (\alpha_j - \alpha_{j-1}).$$

We say that a division $\boldsymbol{\alpha}'$ of $[a,b]$ is a *refinement* of $\boldsymbol{\alpha}$ if $\boldsymbol{\alpha}' \supset \boldsymbol{\alpha}$.

5.1.1 Definition. A pair $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of finite subsets of $[a,b]$ is called a *partition* (or also a *tagged division*) of the interval $[a,b]$ if $\boldsymbol{\alpha}$ is a division of $[a,b]$, $\boldsymbol{\xi} = \{\xi_1, \dots, \xi_{\nu(\boldsymbol{\alpha})}\}$, and

$$\alpha_{j-1} \le \xi_j \le \alpha_j \quad \text{for all } j = 1, \dots, \nu(\boldsymbol{\alpha}).$$

We say that ξ_j is the *tag* of the subinterval $[\alpha_{j-1}, \alpha_j]$ and $\boldsymbol{\xi}$ is the *set of tags* of the division $\boldsymbol{\alpha}$.

Sequences of divisions or partitions will be denoted by $\{\boldsymbol{\alpha}^n\}$ or $\{(\boldsymbol{\beta}^n, \boldsymbol{\eta}^n)\}$, respectively; we use upper indices to avoid confusion with the elements of the sets $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\eta}$, etc.

5.1.2 Definition. Given a pair of functions $f, g : [a,b] \to \mathbb{R}$ and a partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of the interval $[a,b]$, we define

$$S(f, \mathrm{d}g, P; [a,b]) := \sum_{j=1}^{\nu(\boldsymbol{\alpha})} f(\xi_j)\,[g(\alpha_j) - g(\alpha_{j-1})].$$

If $[a,b]$ and f, g are fixed and no misunderstanding can happen, we write $S(f, \mathrm{d}g, P)$, $S(\boldsymbol{\alpha}, \boldsymbol{\xi})$, or even $S(P)$ instead of $S(f, \mathrm{d}g, P; [a,b])$.

5.1.3 Definition. Let $f, g : [a,b] \to \mathbb{R}$.
(i) We say that the *Riemann-Stieltjes* (δ)*-integral* (shortly (δ)RS-integral) of f with respect to g

$$(\delta)\int_a^b f(x)\,\mathrm{d}g(x) \quad \Big(\text{we also write } (\delta)\int_a^b f\,\mathrm{d}g\Big)$$

exists and has a value $I \in \mathbb{R}$ if

$$\left.\begin{array}{l}\text{for every } \varepsilon > 0 \text{ there is a } \delta_\varepsilon > 0 \text{ such that}\\ \quad |S(P) - I| < \varepsilon\\ \text{for all partitions } P = (\boldsymbol{\alpha}, \boldsymbol{\xi}) \text{ of } [a,b] \text{ such that } |\boldsymbol{\alpha}| < \delta_\varepsilon.\end{array}\right\} \tag{5.1.1}$$

(ii) We say that the *Riemann-Stieltjes* (σ)*-integral* (shortly (σ)RS-integral) of f with respect to g

$$(\sigma)\int_a^b f(x)\,\mathrm{d}g(x) \quad \Big(\text{we also write } (\sigma)\int_a^b f\,\mathrm{d}g\Big)$$

exists and has a value $I \in \mathbb{R}$ if

$$\left.\begin{array}{l}\text{for every } \varepsilon > 0 \text{ there is a division } \boldsymbol{\alpha}_\varepsilon \text{ of } [a,b] \text{ such that}\\ \quad |S(P) - I| < \varepsilon\\ \text{for all partitions } P = (\boldsymbol{\alpha}, \boldsymbol{\xi}) \text{ of } [a,b] \text{ such that } \boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon.\end{array}\right\} \tag{5.1.2}$$

(iii) For any $c \in [a,b]$ we set

$$(\delta)\int_c^c f\,\mathrm{d}g = (\sigma)\int_c^c f\,\mathrm{d}g = 0.$$

If the integral $(\delta)\int_a^b f\,\mathrm{d}g$ or $(\sigma)\int_a^b f\,\mathrm{d}g$ exists, we define

$$(\delta)\int_b^a f\,\mathrm{d}g = -(\delta)\int_a^b f\,\mathrm{d}g \quad \text{or} \quad (\sigma)\int_b^a f\,\mathrm{d}g = -(\sigma)\int_a^b f\,\mathrm{d}g,$$

respectively.

5.1.4 Remark. Our (δ)RS-integral corresponds to the original Stieltjes' definition, while the (σ)RS-integral is also known as the *Moore-Pollard* integral.

The classical Riemann integral is a special case of the (δ)RS-integral for $g(x) \equiv x$ on $[a,b]$.

If we speak about the RS-integral without distinguishing between the (δ) or (σ) variant, we mean that the given statement holds for both integrals. In such and other cases when no misunderstanding can occur we do not include the symbol (δ) or (σ) before the integral sign.

The function f in the integral $\int_a^b f\,\mathrm{d}g$ is called the *integrand*, while the function g is called the *integrator*.

5.1.5 Exercise. Prove that Definition 5.1.3 is correct in the sense that the value of the integral is determined uniquely, i.e., if $I_1 \in \mathbb{R}$ and $I_2 \in \mathbb{R}$ satisfy (5.1.1) (with I replaced by I_1 or I_2), then $I_1 = I_2$ (and similarly for (5.1.2)).

5.1.6 Exercises. Prove the following assertions for both kinds of the RS-integral:

(i) If the function g is constant, then $\int_a^b f\,\mathrm{d}g = 0$ for any function $f : [a,b] \to \mathbb{R}$.

(ii) If the function f is constant, then $\int_a^b f\,\mathrm{d}g = f(a)\,(g(b) - g(a))$ for any function $g : [a,b] \to \mathbb{R}$.

From Definition 5.1.3 we can easily conclude that the (δ)RS-integral is a special case of the (σ)RS-integral in the following sense.

5.1.7 Theorem. *If $(\delta)\int_a^b f\,\mathrm{d}g$ exists, then $(\sigma)\int_a^b f\,\mathrm{d}g$ exits as well and has the same value.*

Proof. The statement follows immediately from the fact that the inequality $|\boldsymbol{\alpha}''| \le |\boldsymbol{\alpha}'|$ holds for all divisions $\boldsymbol{\alpha}', \boldsymbol{\alpha}''$ of $[a,b]$ such that $\boldsymbol{\alpha}'' \supset \boldsymbol{\alpha}'$. □

5.1.8 Remark. Let an arbitrary $\delta_0 > 0$ be given. Then, in Definition 5.1.1 (i), the condition (5.1.1) can be replaced by the following weaker condition:

$$\left.\begin{aligned} &\text{For every } \varepsilon > 0 \text{ there is a } \delta_\varepsilon \in (0,\delta_0) \text{ such that} \\ &\quad |S(P) - I| < \varepsilon \\ &\text{for all partitions } P = (\boldsymbol{\alpha},\boldsymbol{\xi}) \text{ of } [a,b] \text{ such that } |\boldsymbol{\alpha}| < \delta_\varepsilon. \end{aligned}\right\} \qquad (5.1.1')$$

Similarly, if a division $\boldsymbol{\alpha}_0$ of $[a,b]$ is given, then, in Definition 5.1.3 (ii), the condition (5.1.2) can be replaced by the following weaker condition:

$$\left.\begin{aligned} &\text{For every } \varepsilon > 0 \text{ there is an } \boldsymbol{\alpha}_\varepsilon \in \mathscr{D}[a,b] \text{ such that } \boldsymbol{\alpha}_\varepsilon \supset \boldsymbol{\alpha}_0 \text{ and} \\ &\quad |S(P) - I| < \varepsilon \\ &\text{for all partitions } P = (\boldsymbol{\alpha},\boldsymbol{\xi}) \text{ of } [a,b] \text{ such that } \boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon. \end{aligned}\right\} \qquad (5.1.2')$$

5.1.9 Exercise. Verify the statements mentioned in Remark 5.1.8.

5.1.10 Example. Let $a = -1,\ b = 1$ and

$$f(x) = \begin{cases} 0 & \text{if } x \le 0, \\ 1 & \text{if } x > 0, \end{cases} \quad \text{and} \quad g(x) = \begin{cases} 1 & \text{if } x < 0, \\ 0 & \text{if } x \ge 0. \end{cases}$$

Set $\boldsymbol{\alpha}_0 = \{-1, 0, 1\}$. Then, for every division $\boldsymbol{\alpha}$ of $[-1,1]$ which is a refinement of $\boldsymbol{\alpha}_0$ and for every partition $P = (\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$, there is a $k \in \mathbb{N}$ such that

$\alpha_k = 0$. Then

$$S(P) = f(\xi_k)\,(g(0) - g(\alpha_{k-1})) + f(\xi_{k+1})\,(g(\alpha_{k+1}) - g(0)) = 0,$$

because

$$f(\xi_k) = 0 \quad \text{and} \quad g(\alpha_{k+1}) - g(0) = 0.$$

By the second part of Remark 5.1.8, we see that $(\sigma)\int_{-1}^{1} f\,\mathrm{d}g = 0$.

On the other hand, for every partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of the interval $[-1,1]$ such that $0 \notin \boldsymbol{\alpha}$ there is a $k \in \mathbb{N}$ such that $\alpha_{k-1} < 0 < \alpha_k$, and consequently

$$S(P) = f(\xi_k)\,(g(\alpha_k) - g(\alpha_{k-1})) = -f(\xi_k) = -\begin{cases} 0 & \text{if } \ \xi_k \le 0, \\ 1 & \text{if } \ \xi_k > 0. \end{cases}$$

Now it is clear that the integral $(\delta)\int_{-1}^{1} f\,\mathrm{d}g$ does not exist.

The following two lemmas hold for both kinds of the RS-integral and are direct corollaries of Definition 5.1.3.

5.1.11 Lemma. (i) *If the integral* $\int_a^b f\,\mathrm{d}g$ *exists, then*

$$\left|\int_a^b f\,\mathrm{d}g\right| \le \|f\|_\infty \operatorname{var}_a^b g.$$

(ii) *If, in addition,* $g \in \mathrm{BV}([a,b])$ *and the integral* $\int_a^b f(x)\,\mathrm{d}(\operatorname{var}_a^x g)$ *exists, then*

$$\left|\int_a^b f\,\mathrm{d}g\right| \le \int_a^b |f(x)|\,\mathrm{d}(\operatorname{var}_a^x g) \le \|f\|_\infty \operatorname{var}_a^b g.$$

5.1.12 Remark. We will show later (cf. Corollary 5.3.10) that if f is bounded on $[a,b]$ then for both kinds of the RS-integral, the existence of the integral $\int_a^b f(x)\,\mathrm{d}(\operatorname{var}_a^x g)$ already follows from the existence of the integral $\int_a^b f\,\mathrm{d}g$.

5.1.13 Lemma. *Let* $f, f_1, f_2,\ g, g_1, g_2 : [a,b] \to \mathbb{R}$ *and let all the integrals*

$$\int_a^b f_1\,\mathrm{d}g, \ \int_a^b f_2\,\mathrm{d}g, \ \int_a^b f\,\mathrm{d}g_1 \ \textit{and} \ \int_a^b f\,\mathrm{d}g_2$$

exist. Then for any $c_1, c_2 \in \mathbb{R}$, *the following relations hold:*

$$\int_a^b (c_1 f_1 + c_2 f_2)\,\mathrm{d}g = c_1 \int_a^b f_1\,\mathrm{d}g + c_2 \int_a^b f_2\,\mathrm{d}g,$$
$$\int_a^b f\,\mathrm{d}(c_1 g_1 + c_2 g_2) = c_1 \int_a^b f\,\mathrm{d}g_1 + c_2 \int_a^b f\,\mathrm{d}g_2.$$

5.1.14 Exercises. (i) Prove Lemmas 5.1.11 and 5.1.13.

(ii) Prove the following statement for both kinds of the RS-integral:

If $f, g : [a, b] \to \mathbb{R}$ *are such that* g *is nondecreasing and the integral* $\int_a^b f\,\mathrm{d}g$ *exists, then*

$$\Big(\inf_{x\in[a,b]} f(x)\Big)\,(g(b)-g(a)) \le \int_a^b f\,\mathrm{d}g \le \Big(\sup_{x\in[a,b]} f(x)\Big)\,(g(b)-g(a)).$$

Both kinds of the RS-integral are, in a sense, generalized limits of integral sums with respect to partitions. It is thus not surprising that the following statement, which is an analogue of the classical Bolzano-Cauchy condition, holds.

5.1.15 Theorem (BOLZANO-CAUCHY CONDITION)**.**
Given a pair of functions $f, g : [a, b] \to \mathbb{R}$, *the integral* $(\delta)\int_a^b f\,\mathrm{d}g$ *exists if and only if*

$$\left.\begin{aligned} &\text{for every } \varepsilon>0 \text{ there is a } \delta_\varepsilon>0 \text{ such that}\\ &\quad |S(P)-S(Q)|<\varepsilon\\ &\text{for all partitions } P=(\boldsymbol{\alpha},\boldsymbol{\xi}),\ Q=(\boldsymbol{\beta},\boldsymbol{\eta}) \text{ of } [a,b]\\ &\text{such that } |\boldsymbol{\alpha}|<\delta_\varepsilon \text{ and } |\boldsymbol{\beta}|<\delta_\varepsilon. \end{aligned}\right\} \tag{5.1.3}$$

Similarly, the integral $(\sigma)\int_a^b f\,\mathrm{d}g$ *exists if and only if*

$$\left.\begin{aligned} &\text{for every } \varepsilon>0 \text{ there is a division } \boldsymbol{\alpha}_\varepsilon \text{ of } [a,b] \text{ such that}\\ &\quad |S(P)-S(Q)|<\varepsilon\\ &\text{for all partitions } P=(\boldsymbol{\alpha},\boldsymbol{\xi}), Q=(\boldsymbol{\beta},\boldsymbol{\eta}) \text{ of } [a,b]\\ &\text{with } \boldsymbol{\alpha}\supset\boldsymbol{\alpha}_\varepsilon \text{ and } \boldsymbol{\beta}\supset\boldsymbol{\alpha}_\varepsilon. \end{aligned}\right\} \tag{5.1.4}$$

Proof. The necessity of the conditions (5.1.3) and (5.1.4) for the existence of the corresponding integrals is obvious from Definition 5.1.3.

We will prove that condition (5.1.4) guarantees the existence of the integral $(\sigma)\int_a^b f\,\mathrm{d}g$. If (5.1.4) is satisfied, there is a sequence $\{P_k\}=\{(\boldsymbol{\alpha}^k,\boldsymbol{\xi}^k)\}$ of partitions of $[a, b]$ such that

$$\left.\begin{aligned} &|S(P)-S(P_k)|<\tfrac{1}{k}\\ &\qquad \text{for all partitions } P=(\boldsymbol{\alpha},\boldsymbol{\xi}) \text{ of } [a,b] \text{ such that } \boldsymbol{\alpha}\supset\boldsymbol{\alpha}^k,\\ &\text{and}\\ &\boldsymbol{\alpha}^k\subset\boldsymbol{\alpha}^\ell \quad \text{for } \ell\in\mathbb{N} \text{ and } \ell\ge k. \end{aligned}\right\} \tag{5.1.5}$$

In particular,

$$|S(P_k)-S(P_\ell)|<\tfrac{1}{k} \quad \text{whenever } k,\ell\in\mathbb{N} \text{ and } \ell\ge k. \tag{5.1.6}$$

The sequence $\{S(P_k)\}$ is a Cauchy sequence of real numbers. Hence there exists an $I \in \mathbb{R}$ such that $\lim_{k\to\infty} S(P_k) = I$. Now, let $\varepsilon > 0$ be given. Choose k_ε in such a way that

$$\frac{1}{k_\varepsilon} < \frac{\varepsilon}{2} \quad \text{and} \quad |S(P_{k_\varepsilon}) - I| < \frac{\varepsilon}{2}. \tag{5.1.7}$$

Then, combining (5.1.5) and (5.1.7), we deduce that

$$|S(P) - I| \le |S(P) - S(P_{k_\varepsilon})| + |S(P_{k_\varepsilon}) - I| < \varepsilon$$

for every partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a, b]$ such that $\boldsymbol{\alpha} \supset \boldsymbol{\alpha}^{k_\varepsilon}$. Therefore $I = (\sigma)\int_a^b f\,\mathrm{d}g$.

In a similar way we can prove that condition (5.1.3) implies the existence of the integral $(\delta)\int_a^b f\,\mathrm{d}g$. □

5.1.16 Exercises. (i) Prove the assertion of Theorem 5.1.15 for (δ)RS-integrals.

(ii) Prove that conditions (5.1.3) or (5.1.4) are respectively equivalent to the following ones:

For every $\varepsilon > 0$ there is a $\delta_\varepsilon > 0$ such that
$$|S(P) - S(Q)| < \varepsilon$$
for all partitions $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$, $Q = (\boldsymbol{\beta}, \boldsymbol{\eta})$ of $[a, b]$ such that $|\boldsymbol{\alpha}| < \delta_\varepsilon$, $\boldsymbol{\beta} \supset \boldsymbol{\alpha}$. (5.1.3')

For every $\varepsilon > 0$ there is a division $\boldsymbol{\alpha}_\varepsilon$ of $[a, b]$ such that
$$|S(P) - S(Q)| < \varepsilon$$
for all partitions $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$, $Q = (\boldsymbol{\beta}, \boldsymbol{\eta})$ of $[a, b]$ such that $\boldsymbol{\beta} \supset \boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon$. (5.1.4')

Hint: Let $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ be divisions of $[a, b]$ and $\boldsymbol{\alpha}' = \boldsymbol{\alpha} \cup \boldsymbol{\beta}$. Then $\boldsymbol{\alpha}'$ is also a division of $[a, b]$, $\boldsymbol{\alpha}' \supset \boldsymbol{\alpha}$, $\boldsymbol{\alpha}' \supset \boldsymbol{\beta}$ and

$$\big|S(\boldsymbol{\alpha}, \boldsymbol{\xi}) - S(\boldsymbol{\beta}, \boldsymbol{\eta})\big| \le \big|S(\boldsymbol{\alpha}, \boldsymbol{\xi}) - S(\boldsymbol{\alpha}', \boldsymbol{\xi}')\big| + \big|S(\boldsymbol{\alpha}', \boldsymbol{\xi}') - S(\boldsymbol{\beta}, \boldsymbol{\eta})\big|$$

for all partitions $(\boldsymbol{\alpha}, \boldsymbol{\xi})$, $(\boldsymbol{\beta}, \boldsymbol{\eta})$ and $(\boldsymbol{\alpha}', \boldsymbol{\xi}')$ of $[a, b]$.

The following theorem is a direct corollary of Theorem 5.1.15. It is valid for both kinds of the RS-integral.

5.1.17 Theorem. *If the integral* $\int_a^b f\,\mathrm{d}g$ *exists and* $[c, d] \subset [a, b]$, *then the integral* $\int_c^d f\,\mathrm{d}g$ *exists, as well.*

Proof. Assume that the integral $(\sigma)\int_a^b f\,\mathrm{d}g$ exists and let $\varepsilon > 0$ be given. By Theorem 5.1.15 there is a division $\boldsymbol{\alpha}_\varepsilon$ of $[a,b]$ such that

$$|S(P) - S(P')| < \varepsilon \tag{5.1.8}$$

holds for all partitions $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$, $P' = (\boldsymbol{\alpha}', \boldsymbol{\xi}')$ of $[a,b]$ such that $\boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon$ and $\boldsymbol{\alpha}' \supset \boldsymbol{\alpha}_\varepsilon$. By Remark 5.1.8, we can assume that $\{c,d\} \subset \boldsymbol{\alpha}_\varepsilon$ and we can thus decompose $\boldsymbol{\alpha}_\varepsilon$ in such a way that

$$\boldsymbol{\alpha}_\varepsilon = \boldsymbol{\beta}^- \cup \boldsymbol{\beta}_\varepsilon \cup \boldsymbol{\beta}^+,$$

where $\boldsymbol{\beta}^-$ is a division of $[a,c]$, $\boldsymbol{\beta}_\varepsilon$ is a division of $[c,d]$, and $\boldsymbol{\beta}^+$ is a division of $[d,b]$. Now, let $(\boldsymbol{\beta}, \boldsymbol{\eta})$ and $(\boldsymbol{\beta}', \boldsymbol{\eta}')$ be partitions of $[c,d]$ such that $\boldsymbol{\beta} \supset \boldsymbol{\beta}_\varepsilon$ and $\boldsymbol{\beta}' \supset \boldsymbol{\beta}_\varepsilon$. Define

$$\boldsymbol{\alpha} = \boldsymbol{\beta}^- \cup \boldsymbol{\beta} \cup \boldsymbol{\beta}^+,\ \boldsymbol{\eta} = (\boldsymbol{\eta}^-, \boldsymbol{\eta}, \boldsymbol{\eta}^+) \ \text{ and } \ \boldsymbol{\alpha}' = \boldsymbol{\beta}^- \cup \boldsymbol{\beta}' \cup \boldsymbol{\beta}^+,\ (\boldsymbol{\eta}^-, \boldsymbol{\eta}', \boldsymbol{\eta}^+),$$

where $(\boldsymbol{\beta}^-, \boldsymbol{\eta}^-)$ is a partition of $[a,c]$ and $(\boldsymbol{\beta}^+, \boldsymbol{\eta}^+)$ is a partition of $[d,b]$. Obviously, $(\boldsymbol{\alpha}, \boldsymbol{\xi})$ and $(\boldsymbol{\alpha}', \boldsymbol{\xi}')$ are partitions of $[a,b]$, $\boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon$, $\boldsymbol{\alpha}' \supset \boldsymbol{\alpha}_\varepsilon$,

$$S(\boldsymbol{\alpha}, \boldsymbol{\xi}) = S(\boldsymbol{\beta}^-, \boldsymbol{\eta}^-) + S(\boldsymbol{\beta}, \boldsymbol{\eta}) + S(\boldsymbol{\beta}^+, \boldsymbol{\eta}^+)$$

and

$$S(\boldsymbol{\alpha}', \boldsymbol{\xi}') = S(\boldsymbol{\beta}^-, \boldsymbol{\eta}^-) + S(\boldsymbol{\beta}', \boldsymbol{\eta}') + S(\boldsymbol{\beta}^+, \boldsymbol{\eta}^+).$$

Thus, by (5.1.8) we have

$$|S(\boldsymbol{\beta}, \boldsymbol{\eta}) - S(\boldsymbol{\beta}', \boldsymbol{\eta}')| = |S(\boldsymbol{\alpha}, \boldsymbol{\xi}) - S(\boldsymbol{\alpha}', \boldsymbol{\xi}')| < \varepsilon$$

and, by Theorem 5.1.15, this yields the existence of the integral $\int_c^d f\,\mathrm{d}g$. □

The statement of the theorem for the (δ)RS-integral can be proved analogously; we leave it as an exercise to the reader. □

5.1.18 Exercise. Prove Theorem 5.1.17 for the (δ)RS-integral.

The following statement holds for both kinds of the RS-integral.

5.1.19 Theorem. *If the integral $\int_a^b f\,\mathrm{d}g$ exists and $c \in [a,b]$, then also both the integrals $\int_a^c f\,\mathrm{d}g$ and $\int_c^b f\,\mathrm{d}g$ exist and satisfy*

$$\int_a^b f\,\mathrm{d}g = \int_a^c f\,\mathrm{d}g + \int_c^b f\,\mathrm{d}g.$$

Proof. If $c = a$ or $c = b$, the statement of the theorem is trivial. Thus, let $c \in (a,b)$ and let the integral $\int_a^b f\,\mathrm{d}g$ exist. Then the existence of the integrals $\int_a^c f\,\mathrm{d}g$ and $\int_c^b f\,\mathrm{d}g$ is guaranteed by Theorem 5.1.17.

Let $\varepsilon > 0$. Choose partitions $P' = (\boldsymbol{\alpha}', \boldsymbol{\xi}')$ of $[a,c]$ and $P'' = (\boldsymbol{\alpha}'', \boldsymbol{\xi}'')$ of $[c,b]$ in such a way that

$$\left.\begin{array}{l} \left|S(P') - \int_a^c f \, \mathrm{d}g\right| + \left|S(P'') - \int_c^b f \, \mathrm{d}g\right| + \left|S(P) - \int_a^b f \, \mathrm{d}g\right| < \varepsilon, \\ \text{where } \boldsymbol{\alpha} = \boldsymbol{\alpha}' \cup \boldsymbol{\alpha}'', \ \boldsymbol{\xi} = \boldsymbol{\xi}' \cup \boldsymbol{\xi}'' \text{ and } P = (\boldsymbol{\alpha}, \boldsymbol{\xi}). \end{array}\right\} \tag{5.1.9}$$

Obviously, $S(P) = S(P') + S(P'')$. Hence,

$$\begin{aligned} &\left|\int_a^b f \, \mathrm{d}g - \int_a^c f \, \mathrm{d}g - \int_c^b f \, \mathrm{d}g\right| \\ &\quad \le \left|\int_a^b f \, \mathrm{d}g - S(P)\right| + |S(P) - S(P') - S(P'')| \\ &\qquad + \left|S(P') - \int_a^c f \, \mathrm{d}g\right| + \left|S(P'') - \int_c^b f \, \mathrm{d}g\right| < \varepsilon. \end{aligned}$$

As $\varepsilon > 0$ was arbitrary, this completes the proof. □

5.1.20 Exercise. Why does the existence of the integrals

$$\int_a^b f \, \mathrm{d}g, \quad \int_a^c f \, \mathrm{d}g, \quad \int_c^b f \, \mathrm{d}g$$

imply the existence of partitions P' of $[a,c]$ and P'' of $[c,b]$ such that (5.1.9) holds?

The converse of Theorem 5.1.19 is easily shown to be valid for the (σ) RS-integral.

5.1.21 Theorem. *If $c \in [a,b]$ and if the integrals*

$$I_1 = (\sigma)\int_a^c f \, \mathrm{d}g \quad \textit{and} \quad I_2 = (\sigma)\int_c^b f \, \mathrm{d}g$$

exist, then also the integral $(\sigma)\int_a^b f \, \mathrm{d}g$ exists and equals $I_1 + I_2$.

Proof. Let $\varepsilon > 0$ be given. Choose divisions $\boldsymbol{\alpha}'_\varepsilon$ of $[a,c]$ and $\boldsymbol{\alpha}''_\varepsilon$ of $[c,b]$ such that

$$\begin{aligned} &\left|S(P') - I_1\right| < \varepsilon \text{ for all partitions } P' = (\boldsymbol{\alpha}', \boldsymbol{\xi}') \text{ of } [a,c] \text{ such that } \boldsymbol{\alpha}' \supset \boldsymbol{\alpha}'_\varepsilon, \\ &\left|S(P'') - I_2\right| < \varepsilon \text{ for all partitions } P'' = (\boldsymbol{\alpha}'', \boldsymbol{\xi}'') \text{ of } [c,b] \text{ such that } \boldsymbol{\alpha}'' \supset \boldsymbol{\alpha}''_\varepsilon. \end{aligned}$$

Now, let $\boldsymbol{\alpha}_\varepsilon = \boldsymbol{\alpha}'_\varepsilon \cup \boldsymbol{\alpha}''_\varepsilon$. Since $c \in \boldsymbol{\alpha}_\varepsilon$, every partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$ satisfying $\boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon$ can be decomposed to partitions $P' = (\boldsymbol{\alpha}', \boldsymbol{\xi}')$ of $[a,c]$ and $P'' = (\boldsymbol{\alpha}'', \boldsymbol{\xi}'')$ of $[c,b]$ in such a way that

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}' \cup \boldsymbol{\alpha}'' \quad \text{and} \quad \boldsymbol{\xi} = \boldsymbol{\xi}' \cup \boldsymbol{\xi}'', \quad \text{where } \boldsymbol{\alpha}' \supset \boldsymbol{\alpha}'_\varepsilon \text{ and } \boldsymbol{\alpha}'' \supset \boldsymbol{\alpha}''_\varepsilon.$$

Moreover, $S(P) = S(P') + S(P'')$, and the definitions of $\boldsymbol{\alpha}'_\varepsilon$ and $\boldsymbol{\alpha}''_\varepsilon$ imply

$$\big|S(P) - (I_1 + I_2)\big| \le \big|S(P') - I_1\big| + \big|S(P'') - I_2\big| < 2\,\varepsilon$$

for every partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$ such that $\boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon$. This completes the proof of the theorem. □

5.1.22 Remark. To get an analogous statement also for the (δ)RS-integral, we have to assume the pseudo-additivity of the functions f, g at the point c; see Definition 5.2.1 and Exercise 5.2.10.

For the existence of the (δ)RS-integral, we have the following necessary and sufficient condition.

5.1.23 Theorem. *For each pair* $f, g : [a,b] \to \mathbb{R}$, *the integral* $(\delta)\int_a^b f\,\mathrm{d}g$ *exists if and only if*

$$\left.\begin{aligned} &\lim_{n\to\infty} S(P_n) \in \mathbb{R} \text{ for each sequence } \{P_n\} = \{(\boldsymbol{\alpha}^n, \boldsymbol{\xi}^n)\}\\ &\qquad\qquad \text{of partitions of } [a,b] \text{ such that } \lim_{n\to\infty} |\boldsymbol{\alpha}^n| = 0. \end{aligned}\right\} \quad (5.1.10)$$

Proof. The necessity of the condition (5.1.10) for the existence of the integral $(\delta)\int_a^b f\,\mathrm{d}g$ is obvious; it remains to prove its sufficiency.

Thus, assume that (5.1.10) holds and let the sequences $\{P_n\} = \{(\boldsymbol{\alpha}^n, \boldsymbol{\xi}^n)\}$ and $\{\widetilde{P}_m\} = \{(\widetilde{\boldsymbol{\alpha}}^n, \widetilde{\boldsymbol{\xi}}^n)\}$ of partitions of $[a,b]$ be such that

$$\lim_{n\to\infty} |\boldsymbol{\alpha}^n| = \lim_{n\to\infty} |\widetilde{\boldsymbol{\alpha}}^n| = 0$$

and

$$\lim_{n\to\infty} S(\boldsymbol{\alpha}^n, \boldsymbol{\xi}^n) = I \in \mathbb{R} \quad \text{and} \quad \lim_{n\to\infty} S(\widetilde{\boldsymbol{\alpha}}^n, \widetilde{\boldsymbol{\xi}}^n) = \widetilde{I} \in \mathbb{R}.$$

Now, consider the sequence $\{Q_n\}$ of partitions of $[a,b]$ given by

$$Q_{2k-1} = P_k, \quad Q_{2k} = \widetilde{P}_k \quad \text{for } k \in \mathbb{N}.$$

By our assumption, the sequence $\big\{S(Q_n)\big\}$ has a finite limit $J \in \mathbb{R}$, and since it contains both the sequences $\{S(P_n)\}$ and $\{S(\widetilde{P}_n)\}$, we necessarily have $I = \widetilde{I} = J$. This means that the value of the limit

$$I = \lim_{n\to\infty} S(P_n)$$

does not depend on the choice of the sequence $\{P_n\}$ of partitions of $[a,b]$ for which $\lim_{n\to\infty} |\boldsymbol{\alpha}^n| = 0$.

Now, assume that $(\delta)\int_a^b f\,\mathrm{d}g \ne I$. Then there exists an $\widetilde{\varepsilon} > 0$ such that for every $k \in \mathbb{N}$ there is a partition $P_k = (\boldsymbol{\alpha}^k, \boldsymbol{\xi}^k)$ of $[a,b]$ such that

$$|\boldsymbol{\alpha}^k| < 1/k \quad \text{and} \quad |S(P_k) - I| > \widetilde{\varepsilon}.$$

In other words, we have a sequence $\{P_k\} = \{(\boldsymbol{\alpha}^k, \boldsymbol{\xi}^k)\}$ of partitions of $[a,b]$ such that

$$\lim_{k\to\infty} |\boldsymbol{\alpha}^k| = 0 \quad \text{and} \quad \lim_{k\to\infty} S(P_k) \neq I,$$

which is a contradiction. Therefore $(\delta)\int_a^b f\,\mathrm{d}g = I$. □

5.1.24 Remark. Let $x_0 \in (a,b)$, c, $d \in \mathbb{R}$ and

$$g(x) = \begin{cases} c' & \text{if } \ x \in [a, x_0), \\ c \in [c', c''] & \text{if } \ x = x_0, \\ c'' & \text{if } \ x \in (x_0, d], \end{cases}$$

and let an arbitrary $f \in \mathrm{G}([a,b])$ be given.
(i) Consider a sequence of divisions $\{\boldsymbol{\alpha}^n\}$ of $[a,b]$ such that $|\boldsymbol{\alpha}^n| \to 0$, while for every $n \in \mathbb{N}$ there is a k_n such that $\alpha^n_{k_n-1} < x_0 < \alpha^n_{k_n}$. Further, let $\boldsymbol{\xi}^n, \boldsymbol{\eta}^n, \boldsymbol{\zeta}^n$ be the sequences of sets of tags corresponding to $\boldsymbol{\alpha}^n$, $n \in \mathbb{N}$, and such that

$$\xi^n_{k_n} = x_0, \quad \alpha^n_{k_n-1} \le \eta^n_{k_n} < x_0 \quad \text{and} \quad x_0 < \zeta^n_{k_n} \le \alpha^n_k \qquad \text{for every } n \in \mathbb{N}.$$

Then

$$S(\boldsymbol{\alpha}^n, \boldsymbol{\xi}^n) = f(x_0)\,(c'' - c') \ = f(x_0)\,\Delta g(x_0),$$

$$S(\boldsymbol{\alpha}^n, \boldsymbol{\eta}^n) = f(\eta^n_{k_n})\,(c'' - c') = f(\eta^n_{k_n})\,\Delta g(x_0),$$

and

$$S(\boldsymbol{\alpha}^n, \boldsymbol{\zeta}^n) = f(\zeta^n_{k_n})\,(c'' - c') = f(\zeta^n_{k_n})\,\Delta g(x_0)$$

for every $n \in \mathbb{N}$. Thus, if there is an $I \in \mathbb{R}$ such that

$$\lim_{n\to\infty} S(\boldsymbol{\alpha}^n, \boldsymbol{\xi}^n) = I$$

for each sequence $(\boldsymbol{\alpha}^n, \boldsymbol{\xi}^n)$ of partitions of $[a,b]$ fulfilling $\lim_{n\to\infty} |\boldsymbol{\alpha}^n| = 0$, then either

$$g(x_0-) = c' = g(x_0) = c = g(x_0+) = c'' \quad \text{or} \quad f(x_0-) = f(x_0) = f(x_0+)$$

must hold. In view of Theorem 5.1.23, we can expect that if the integral $(\delta)\int_a^b f\,\mathrm{d}g$ exists, then the functions f and g have no common point of discontinuity.

(ii) Now, let $\boldsymbol{\alpha}_0$ be an arbitrary division of $[a,b]$ containing x_0. For each its refinement $\boldsymbol{\alpha}$, there is an index $k = k(\boldsymbol{\alpha})$ such that $x_0 = \alpha_k$. Hence for each partition $(\boldsymbol{\alpha}, \boldsymbol{\xi})$ such that $\boldsymbol{\alpha} \supset \boldsymbol{\alpha}_0$ we have

$$S(\boldsymbol{\alpha}, \boldsymbol{\xi}) = \begin{cases} f(\xi_{k-1})\, \Delta^- g(x_0) + f(\xi_k)\, \Delta^+ g(x_0) & \text{if } \xi_{k-1} < x_0 < \xi_k, \\ f(x_0)\, \Delta^- g(x_0) + f(\xi_k))\, \Delta^+ g(x_0) & \text{if } \xi_{k-1} = x_0 < \xi_k, \\ f(\xi_{k-1})\, \Delta^- g(x_0) + f(x_0))\, \Delta^+ g(x_0) & \text{if } \xi_{k-1} < x_0 = \xi_k, \\ f(x_0)\, \Delta^- g(x_0) + f(x_0)\, \Delta^+ g(x_0) & \text{if } \xi_{k-1} = x_0 = \xi_k. \end{cases}$$

and

$$\begin{aligned} &f(x_0-)\, \Delta^- g(x_0) + f(\xi_k)\, \Delta^+ g(x_0), \quad && f(x_0)\, \Delta^- g(x_0) + f(x_0+)\, \Delta^+ g(x_0), \\ &f(x_0-)\, \Delta^- g(x_0) + f(x_0))\, \Delta^+ g(x_0), && f(x_0)\, \Delta^- g(x_0) + f(x_0)\, \Delta^+ g(x_0) \end{aligned}$$

are the accumulation points of the set

$$\Sigma = \Big\{ S(\boldsymbol{\alpha}, \boldsymbol{\xi}) : (\boldsymbol{\alpha}, \boldsymbol{\xi}) \text{ is a partition of } [a,b] \text{ and } \boldsymbol{\alpha} \supset \boldsymbol{\alpha}_0 \Big\}$$

Of course, the integral $(\sigma) \int_a^b f\, \mathrm{d}g$ can exist only if the set Σ will have exactly one accumulation point. It is easy to see that this can happen only if

$$\Delta^+ f(x_0)\, \Delta^+ g(x_0) = \Delta^- f(x_0)\, \Delta^- g(x_0) = 0$$

5.2 Pseudo-additivity

The notion of *pseudo-additivity* enables us to better clarify the mutual relationship between the (δ) and (σ) integrals.

5.2.1 Definition. We say that functions $f, g : [a,b] \to \mathbb{R}$ satisfy the condition of pseudo-additivity at a point $x \in (a,b)$, if

$$\left.\begin{aligned} &\text{for every } \varepsilon > 0 \text{ there is a } \delta_\varepsilon > 0 \text{ such that} \\ &\quad \big| f(\xi)\,(g(x+\delta'') - g(x-\delta')) - f(\xi')\,(g(x) - g(x-\delta')) \\ &\qquad\qquad - f(\xi'')\,(g(x+\delta'') - g(x)) \big| < \varepsilon \\ &\text{holds whenever } \delta', \delta'' \in (0, \delta_\varepsilon) \text{ and} \\ &\quad \xi \in [x-\delta', x+\delta''],\ \xi' \in [x-\delta', x],\ \xi'' \in [x, x+\delta'']. \end{aligned}\right\} \tag{PA}$$

5.2.2 Remark. It is sometimes more convenient to reformulate the condition (PA) as follows:

$$\left.\begin{aligned}&\text{For every } \varepsilon>0 \text{ there is a } \delta_\varepsilon>0 \text{ such that}\\ &\quad \big|f(\xi)\,(g(x'')-g(x'))-f(\xi')\,(g(x)-g(x'))\\ &\qquad\qquad -f(\xi'')\,(g(x'')-g(x))\big|<\varepsilon\\ &\text{holds whenever } x'\in(x-\delta_\varepsilon,x),\ x''\in(x,x+\delta_\varepsilon) \text{ and}\\ &\quad \xi\in[x',x''],\ \xi'\in[x',x],\ \xi''\in[x,x''].\end{aligned}\right\} \qquad \text{(PA')}$$

5.2.3 Example. Let

$$f(x)=\begin{cases}0 & \text{if } x\le 0,\\ 1 & \text{if } x>0,\end{cases} \qquad g(x)=\begin{cases}1 & \text{if } x\le 0,\\ 0 & \text{if } x>0.\end{cases}$$

If $x'<0<x''$, $\xi\in[x',x'']$, $\xi'\in[x',0]$, $\xi''\in[0,x'']$, then

$$\begin{aligned}&\big|f(\xi)\,(g(x'')-g(x'))-f(\xi')\,(g(0)-g(x'))-f(\xi'')\,(g(x'')-g(0))\big|\\ &\quad=\big|-f(\xi)+f(\xi'')\big|=1\end{aligned}$$

whenever $\xi\le 0$ and $\xi''>0$. Thus, the functions f,g do not satisfy the condition (PA) at the point 0.

5.2.4 Lemma. *If $f,g:[a,b]\to\mathbb{R}$ satisfy the condition of pseudo-additivity at $x\in(a,b)$, then at least one of the functions f,g is continuous at x.*

On the other hand, if one of the functions f,g is continuous at $x\in(a,b)$ and the other one is bounded on a neigborhood of x, then the functions f,g satisfy the condition of pseudo-additivity at the point x.

Proof. a) Let $x\in(a,b)$ and f,g satisfy the condition (PA') of pseudo-additivity at x. If we substitute $\xi=\xi'$ into (PA'), we get that for each $\varepsilon>0$ there is a $\delta_\varepsilon>0$ such that

$$\big|f(\xi)-f(\xi'')\big|\,\big|g(x'')-g(x)\big|=\big|f(\xi')-f(\xi'')\big|\,\big|g(x'')-g(x)\big|<\varepsilon$$

holds whenever

$$x'\in(x-\delta_\varepsilon,x),\ x''\in(x,x+\delta_\varepsilon) \text{ and } \xi'=\xi\in[x',x],\ \xi''\in[x,x'']. \tag{5.2.1}$$

In particular, it has to be

$$|\Delta^+ f(x)|\,|\Delta^+ g(x)|=|\Delta f(x)|\,|\Delta^+ g(x)|=0.$$

Therefore, f has to be continuous at x whenever g is not continuous from the right at x. Similarly, by setting $\xi=\xi''$ in (PA'), we can deduce that if g is not continuous from the left at x, then f has to be continuous at x.

b) Let $x \in (a,b)$, $x' \in [a,x)$, $x'' \in (x,b]$, $\xi \in [x',x'']$, $\xi' \in [x',x]$, $\xi'' \in [x,x'']$. Then

$$\begin{aligned}
&\big|f(\xi)\,(g(x'') - g(x')) - f(\xi')\,(g(x) - g(x')) - f(\xi'')\,(g(x'') - g(x))\big| \\
&\quad = \big|\big(f(\xi) - f(\xi')\big)\,\big(g(x) - g(x')\big) - \big(f(\xi'') - f(\xi)\big)\,\big(g(x'') - g(x)\big)\big| \\
&\quad \le \big|f(\xi) - f(\xi')\big|\,\big|g(x) - g(x')\big| + \big|f(\xi'') - f(\xi)\big|\,\big|g(x'') - g(x)\big| \\
&\quad \le \big(|f(\xi) - f(x)| + |f(x) - f(\xi')|\big)\,\big|g(x) - g(x')\big| \\
&\qquad\quad + \big(|f(\xi'') - f(x)| + |f(x) - f(\xi)|\big)\,\big|g(x'') - g(x)\big|,
\end{aligned}$$

wherefrom the second statement of the lemma easily follows. □

5.2.5 Lemma. *If* $f, g : [a,b] \to \mathbb{R}$ *are such that the integral* $(\delta)\int_a^b f\,\mathrm{d}g$ *exists, then they satisfy the condition of pseudo-additivity at every point* $x \in (a,b)$.

Proof. Assume that the integral $(\delta)\int_a^b f\,\mathrm{d}g$ exists and at the same time, (PA') does not hold at some point $x \in (a,b)$. Then there exists an $\widetilde{\varepsilon} > 0$ such that for every $\delta > 0$ it is possible to find points $x', x'', \eta', \eta'', \eta$ such that

$$\left.\begin{aligned}
&x' \in (x-\delta, x),\ x'' \in (x, x+\delta), \\
&\eta \in [x', x''],\ \eta' \in [x', x] \ \text{ and } \ \eta'' \in [x, x''] \\
\text{and}\quad &\\
&\big|f(\eta)\,(g(x'') - g(x')) - f(\eta')\,(g(x) - g(x')) \\
&\qquad\qquad - f(\eta'')\,(g(x'') - g(x))\big| \ge \widetilde{\varepsilon}.
\end{aligned}\right\} \tag{5.2.2}$$

Now, let $\delta > 0$ and $x', x'', \eta', \eta'', \eta$ be given such that (5.2.2) hold and let $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ with be a partition of $[a,b]$ such that

$$|\boldsymbol{\alpha}| < \delta,\ \ \alpha_{k-1} = x' < x < x'' = \alpha_k \ \text{ and } \ \xi_k = \eta \ \text{ for some } k \in \{1, \dots, \nu(\boldsymbol{\alpha})\}.$$

Put $\widetilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha} \cup \{x\}$, $\widetilde{\boldsymbol{\xi}} = (\xi_1, \dots, \xi_{k-1}, \eta', \eta'', \xi_{k+1}, \dots, \xi_{\nu(\boldsymbol{\alpha})})$ and $\widetilde{P} = (\widetilde{\boldsymbol{\alpha}}, \widetilde{\boldsymbol{\xi}})$. Note that $|\widetilde{\boldsymbol{\alpha}}| < \delta$. We have

$$\begin{aligned}
|S(P) - S(\widetilde{P})| &= \big|f(\xi_k)\,[g(\alpha_k) - g(\alpha_{k-1})] - f(\eta')\,[g(x) - g(\alpha_{k-1})] \\
&\qquad - f(\eta'')\,[g(\alpha_k) - g(x)]\big| \\
&= \big|f(\eta)\,[g(x'') - g(x')] - f(\eta')\,[g(x) - g(x')] \\
&\qquad - f(\eta'')\,[g(x'') - g(x)]\big| \\
&\ge \varepsilon.
\end{aligned}$$

This means that the condition (5.1.3) is not satisfied and hence, in view of Theorem 5.1.15, the integral $(\delta)\int_a^b f\,\mathrm{d}g$ does not exist. □

The following statement is a corollary of Lemmas 5.2.4 and 5.2.5.

5.2.6 Theorem. *Let* $f,g:[a,b]\to\mathbb{R}$ *be such that the integral* $(\delta)\int_a^b f\,\mathrm{d}g$ *exists. Then for each* $x\in(a,b)$ *at least one of the functions* f,g *is continuous at* x.

We know that the (δ)RS-integral is a special case of the (σ)RS-integral (see Theorem 5.1.7). The following theorem shows that the concept of pseudo-additivity enables us to clarify the relationship between these integrals in the opposite direction, too.

5.2.7 Theorem. *Let* $f,g:[a,b]\to\mathbb{R}$. *Then the integral* $(\delta)\int_a^b f\,\mathrm{d}g$ *exists if and only if the integral* $(\sigma)\int_a^b f\,\mathrm{d}g$ *exists and the pair* f,g *satisfies the condition of pseudo-additivity at every point* $x\in(a,b)$.

Proof. First, assume that $(\delta)\int_a^b f\,\mathrm{d}g$ exists. Then, by Theorem 5.1.7, $(\sigma)\int_a^b f\,\mathrm{d}g$ exists as well and has the same value. Furthermore, by Lemma 5.2.5, the functions f,g satisfy the condition of pseudo-additivity at every point $x\in(a,b)$.

Conversely, assume that the integral $(\sigma)\int_a^b f\,\mathrm{d}g=I$ exists and the functions f,g satisfy the condition of pseudo-additivity at every point $x\in(a,b)$. Let $\varepsilon>0$ be given and let the division $\boldsymbol{\alpha}_\varepsilon=\{s_0,s_1,\dots,s_r\}$ of $[a,b]$ be such that $r\ge 2$ and

$$|S(Q)-I|<\varepsilon \text{ for all partitions } Q=(\boldsymbol{\beta},\boldsymbol{\eta}) \text{ of } [a,b] \text{ such that } \boldsymbol{\beta}\supset\boldsymbol{\alpha}_\varepsilon. \tag{5.2.3}$$

Set

$$\delta_*:=\min\big\{s_i-s_{i-1}: i\in\{1,\dots,r\}\big\}. \tag{5.2.4}$$

Since the functions f,g satisfy the condition of pseudo-additivity on (a,b), there exists a $\delta_\varepsilon\in(0,\delta_*)$ such that for every $i\in\{1,\dots,r-1\}$, the inequality

$$\left.\begin{aligned}&\big|f(\xi)\,(g(s_i'')-g(s_i'))\\&\qquad -f(\xi')\,(g(s_i)-g(s_i'))-f(\xi'')\,(g(s_i'')-g(s_i))\big|<\frac{\varepsilon}{r-1}\\&\text{holds whenever}\\&\qquad s_i'\in(s_i-\delta_\varepsilon,s_i),\ \ s_i''\in(s_i,s_i+\delta_\varepsilon),\\&\qquad \xi\in[s_i',s_i''],\ \ \xi'\in[s_i',s_i],\ \ \xi''\in[s_i,s_i''].\end{aligned}\right\} \tag{5.2.5}$$

Let $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ be a partition of $[a,b]$ and let $|\boldsymbol{\alpha}|<\delta_\varepsilon$.

By (5.2.4), for any $j\in\{1,\dots,m\}$ the set $(\alpha_{j-1},\alpha_j)\cap\boldsymbol{\alpha}_\varepsilon$ is either a singleton or empty. Let

$$U_1=\big\{j\in\{1,\dots,\nu(\boldsymbol{\alpha})\}:(\alpha_{j-1},\alpha_j)\cap\boldsymbol{\alpha}_\varepsilon=\emptyset\big\},$$
$$U_2=\{1,\dots,\nu(\boldsymbol{\alpha})\}\setminus U_1.$$

Then, for every $j\in U_2$ there is a unique $i(j)\in\{1,\dots,r-1\}$ such that $s_{i(j)}\in(\alpha_{j-1},\alpha_j)$. Thus, the cardinality of U_2 does not exceed $r-1$.

Now, set $\boldsymbol{\beta}=\boldsymbol{\alpha}\cup\boldsymbol{\alpha}_\varepsilon$. Then

$$|\boldsymbol{\beta}|<\delta_\varepsilon<\delta_* \tag{5.2.6}$$

and for every $j\in U_1$, there exists a unique $k(j)\in\{1,\dots,\nu(\boldsymbol{\beta})\}$ such that

$$[\beta_{k(j)-1},\beta_{k(j)}]=[\alpha_{j-1},\alpha_j]. \tag{5.2.7}$$

If $j\in U_2$, then there exists exactly one $\ell(j)\in\{1,\dots,\nu(\boldsymbol{\beta})-1\}$ such that

$$\beta_{\ell(j)-1}=\alpha_{j-1},\ \ \beta_{\ell(j)}=s_{i(j)},\ \ \beta_{\ell(j)+1}=\alpha_j. \tag{5.2.8}$$

Choose a partition $Q=(\boldsymbol{\beta},\boldsymbol{\eta})$ of $[a,b]$ such that

$$\eta_{k(j)}=\xi_j, \quad \text{if } j\in U_1. \tag{5.2.9}$$

Now, we compare the integral sums $S(P)$ and $S(Q)$. We have

$$S(P)=\sum_{j\in U_1} f(\xi_j)\,(g(\alpha_j)-g(\alpha_{j-1}))+\sum_{j\in U_2} f(\xi_j)\,(g(\alpha_j)-g(\alpha_{j-1})).$$

Let $V_1=\{k(j):j\in U_1\},\ V_2=\{1,\dots,\nu(Q)\}\setminus V_1$. Then by (5.2.7)–(5.2.9),

$$\begin{aligned}
S(Q)&=\sum_{k\in V_1} f(\eta_k)\,(g(\beta_k)-g(\beta_{k-1}))+\sum_{k\in V_2} f(\eta_k)\,(g(\beta_k)-g(\beta_{k-1}))\\
&=\sum_{j\in U_1} f(\eta_{k(j)})\,(g(\beta_{k(j)})-g(\beta_{k(j)-1}))+\sum_{k\in V_2} f(\eta_k)\,(g(\beta_k)-g(\beta_{k-1}))\\
&=\sum_{j\in U_1} f(\xi_j)\,(g(\alpha_j)-g(\alpha_{j-1}))\\
&\quad+\sum_{j\in U_2}\big(f(\eta_{\ell(j)})\,(g(\beta_{\ell(j)})-g(\beta_{\ell(j)-1}))+f(\eta_{\ell(j)+1})\,(g(\beta_{\ell(j)+1})-g(\beta_{\ell(j)}))\big)\\
&=\sum_{j\in U_1} f(\xi_j)\,(g(\alpha_j)-g(\alpha_{j-1}))\\
&\quad+\sum_{j\in U_2}\big(f(\eta_{\ell(j)})\,(g(s_{i(j)})-g(\alpha_{j-1}))+f(\eta_{\ell(j)+1})\,(g(\alpha_j)-g(s_{i(j)}))\big).
\end{aligned}$$

Hence

$$\begin{aligned}
S(P)-S(Q)=&\sum_{j\in U_2} f(\xi_j)\,(g(\alpha_j)-g(\alpha_{j-1}))\\
&-\sum_{j\in U_2}\big(f(\eta_{\ell(j)})\,(g(s_{i(j)})-g(\alpha_{j-1}))+f(\eta_{\ell(j)+1})\,(g(\alpha_j)-g(s_{i(j)}))\big),
\end{aligned}$$

i.e., $|S(P)-S(Q)| \le \sum_{j\in U_2} |W_j|$, where

$$W_j = f(\xi_j)\,(g(\alpha_j)-g(\alpha_{j-1})) \\ \qquad - f(\eta_{\ell(j)})\,(g(s_{i(j)})-g(\alpha_{j-1})) - f(\eta_{\ell(j)+1})\,(g(\alpha_j)-g(s_{i(j)})).$$

Let us recall that by (5.2.6) and (5.2.8) we have

$$[\alpha_{j-1},\alpha_j]\subset(s_{i(j)}-\delta_\varepsilon, s_{i(j)}+\delta_\varepsilon),\ \ \xi_j\in[\alpha_{j-1},\alpha_j],$$

$$\eta_{\ell(j)}\in[\alpha_{j-1},s_{i(j)}],\ \ \eta_{\ell(j)+1}\in[s_{i(j)},\alpha_j] \quad \text{for } j\in U_2.$$

By (5.2.5), $|W_j| < \dfrac{\varepsilon}{r-1}$ for every $j\in U_2$, and therefore (using the fact that the cardinality of U_2 does not exceed $r-1$) we have

$$\big|S(P)-S(Q)\big| \le \sum_{j\in U_2} \big|W_j\big| < \varepsilon.$$

Finally, by (5.2.3) and in view of the definition of $\boldsymbol{\beta}$, we get

$$|S(P)-I| \le |S(P)-S(Q)| + |S(Q)-I| < 2\,\varepsilon.$$

This means that $(\delta)\int_a^b f\,\mathrm{d}g = I$. □

5.2.8 Corollary. *Assume that* $(\sigma)\int_a^b f\,\mathrm{d}g$ *exists and for each* $x\in(a,b)$, *at least one of the functions* $f,g:[a,b]\to\mathbb{R}$ *is continuous at* x, *while the other one is bounded on a neighborhood of* x. *Then* $(\delta)\int_a^b f\,\mathrm{d}g$ *exists and equals* $(\sigma)\int_a^b f\,\mathrm{d}g$.

Proof. By Lemma 5.2.4, the pair f,g satisfies the condition of pseudo-additivity at every point $x\in(a,b)$. By Theorem 5.2.7, the integral $(\delta)\int_a^b f\,\mathrm{d}g$ exists as well. The equality $(\delta)\int_a^b f\,\mathrm{d}g = (\sigma)\int_a^b f\,\mathrm{d}g$ follows from Theorem 5.1.7. □

5.2.9 Remark. In particular, if $g(x)\equiv x$ and f is bounded on $[a,b]$, then the definitions of the integrals $(\delta)\int_a^b f(x)\,\mathrm{d}x$ and $(\sigma)\int_a^b f(x)\,\mathrm{d}x$ coincide.

5.2.10 Exercise. Prove the following statement:
If $c\in[a,b]$, *the integrals* $(\delta)\int_a^c f\,\mathrm{d}g$ *and* $(\delta)\int_c^b f\,\mathrm{d}g$ *exist, and* f,g *satisfy the condition of pseudo-additivity at* c, *then the integral* $(\delta)\int_a^b f\,\mathrm{d}g$ *exists and*

$$(\delta)\int_a^b f\,\mathrm{d}g = (\delta)\int_a^c f\,\mathrm{d}g + (\delta)\int_c^b f\,\mathrm{d}g.$$

Hint: Make use of Theorems 5.1.21 and 5.2.7.

5.3 Absolute integrability

We now introduce the following notation.

5.3.1 Definition. Let $-\infty < c < d < \infty$ and $f, g : [c,d] \to \mathbb{R}$. Then we define

$$\left.\begin{aligned} &\mathfrak{S}_{f\Delta g}[c,d] \\ &\quad = \big\{|S(f,\mathrm{d}g,P) - S(f,\mathrm{d}g,P')| : P, P' \text{ are partitions of } [c,d]\big\} \end{aligned}\right\} \tag{5.3.1}$$

and

$$\omega(S_{f\Delta g};[c,d]) = \sup \mathfrak{S}_{f\Delta g}[c,d]. \tag{5.3.2}$$

The following modifications of the Bolzano-Cauchy condition will be useful.

5.3.2 Theorem. *Let $f, g : [a,b] \to \mathbb{R}$. Then the following assertions hold:*

(i) *The integral* $(\delta)\displaystyle\int_a^b f\,\mathrm{d}g$ *exists if and only if*

$$\left.\begin{aligned} &\textit{for every } \varepsilon > 0 \textit{ there is a } \delta_\varepsilon > 0 \textit{ such that} \\ &\qquad \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \omega(S_{f\Delta g};[\alpha_{j-1},\alpha_j]) < \varepsilon \\ &\textit{holds for all divisions } \boldsymbol{\alpha} \textit{ of } [a,b] \textit{ such that } |\boldsymbol{\alpha}| < \delta_\varepsilon. \end{aligned}\right\} \tag{5.3.3}$$

(ii) *The integral* $(\sigma)\displaystyle\int_a^b f\,\mathrm{d}g$ *exists if and only if*

$$\left.\begin{aligned} &\textit{for every } \varepsilon > 0 \textit{ there is a division } \boldsymbol{\alpha}_\varepsilon \textit{ of } [a,b] \textit{ such that} \\ &\qquad \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \omega(S_{f\Delta g};[\alpha_{j-1},\alpha_j]) < \varepsilon \\ &\textit{holds for all divisions } \boldsymbol{\alpha} \textit{ of } [a,b] \textit{ such that } \boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon. \end{aligned}\right\} \tag{5.3.4}$$

Proof. We will show that the condition (5.3.3) is necessary and sufficient for the existence of the (δ)RS-integral.

a) Assume that (5.1.3) holds. Let $\widetilde{\varepsilon} > 0$, $\varepsilon = \widetilde{\varepsilon}/2$, and let δ_ε be defined by the condition (5.1.3). Let a division $\boldsymbol{\alpha}$ of $[a,b]$ be such that $|\boldsymbol{\alpha}| < \delta_\varepsilon$. Denote $m = \nu(\boldsymbol{\alpha})$ and, for every $j \in \{1,\dots,m\}$, choose partitions $P_j = (\boldsymbol{\alpha}^j, \boldsymbol{\xi}^j)$, $\widetilde{P}_j = (\widetilde{\boldsymbol{\alpha}}^j, \widetilde{\boldsymbol{\xi}}^j)$ of $[\alpha_{j-1},\alpha_j]$ such that

$$\omega(S_{f\Delta g},[\alpha_{j-1},\alpha_j]) < S(P_j) - S(\widetilde{P}_j) + \frac{\varepsilon}{m}. \tag{5.3.5}$$

Further, define

$$\boldsymbol{\beta} = \bigcup_{j=1}^{m} \boldsymbol{\alpha}^j, \;\; \boldsymbol{\eta} = \bigcup_{j=1}^{m} \boldsymbol{\xi}^j, \;\; \widetilde{\boldsymbol{\beta}} = \bigcup_{j=1}^{m} \widetilde{\boldsymbol{\alpha}}^j, \;\; \widetilde{\boldsymbol{\eta}} = \bigcup_{j=1}^{m} \widetilde{\boldsymbol{\xi}}^j.$$

Then $Q = (\boldsymbol{\beta}, \boldsymbol{\eta})$ and $\widetilde{Q} = (\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\eta}})$ are partitions of $[a, b]$, and

$$|\boldsymbol{\beta}| < \delta_\varepsilon \text{ and } |\widetilde{\boldsymbol{\beta}}| < \delta_\varepsilon.$$

Hence, using (5.1.3) and (5.3.5) we get

$$\sum_{j=1}^{m} \omega\left(S_{f\Delta g}; [\alpha_{j-1}, \alpha_j]\right) < \sum_{j=1}^{m} \Big(S(P_j) - S(\widetilde{P}_j) + \frac{\varepsilon}{m}\Big)$$
$$= S(Q) - S(\widetilde{Q}) + \varepsilon < 2\,\varepsilon = \widetilde{\varepsilon}.$$

Since $\widetilde{\varepsilon} > 0$ was arbitrary, it follows that condition (5.3.3) is satisfied.

b) For the proof of the reverse implication, assume that (5.3.3) holds. We will prove that condition (5.1.3') is satisfied.

Choose an arbitrary $\varepsilon > 0$. Let δ_ε be defined by (5.3.3) and let $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ and $\widetilde{P} = (\widetilde{\boldsymbol{\alpha}}, \widetilde{\boldsymbol{\xi}})$ be partitions of $[a, b]$ such that $|\boldsymbol{\alpha}| < \delta_\varepsilon$ and $\widetilde{\boldsymbol{\alpha}} \supset \boldsymbol{\alpha}$. Set $m = \nu(P)$. Then for each $j \in \{1, \ldots, m\}$ there is a partition $P_j = (\boldsymbol{\alpha}^j, \boldsymbol{\xi}^j)$ of $[\alpha_{j-1}, \alpha_j]$ such that

$$\widetilde{\boldsymbol{\alpha}} = \bigcup_{j=1}^{m} \boldsymbol{\alpha}^j \text{ and } \widetilde{\boldsymbol{\xi}} = \bigcup_{j=1}^{m} \boldsymbol{\xi}^j.$$

In view of the assumption (5.3.3), we get

$$|S(P) - S(\widetilde{P})| \le \sum_{j=1}^{m} \big|f(\xi_j)\,(g(\alpha_j) - g(\alpha_{j-1})) - S(P_j)\big|$$
$$\le \sum_{j=1}^{m} \omega(S_{f\Delta g}; [\alpha_{j-1}, \alpha_j]) < \varepsilon.$$

Therefore, (5.1.3') holds.

The equivalence of the condition (5.3.4) with the Bolzano-Cauchy condition for the existence of the (σ)RS-integral can be proved analogously; the detailed proof is left as an exercise for the reader. □

5.3.3 Exercise. Prove the statement of Theorem 5.3.2 for (σ)RS-integrals.

5.3.4 Lemma. *Let* $f, g : [a, b] \to \mathbb{R}$ *and* $[c, d] \subset [a, b]$. *Then*

$$\omega_{[c,d]}(f)\, |g(d) - g(c)| \leq \omega(S_{f\Delta g}; [c, d]) \leq \omega_{[c,d]}(f)\, \mathrm{var}_c^d\, g, \tag{5.3.6}$$

where

$$\omega_I(f) = \sup_{s,t \in I} |f(t) - f(s)|$$

denotes the modulus of oscillation of f *over interval* I.

Proof. a) Let $\boldsymbol{\alpha} = \boldsymbol{\beta} = \{c, d\}$, $\xi, \eta \in [c, d]$ and $\boldsymbol{\xi} = \{\xi\}$, $\boldsymbol{\eta} = \{\eta\}$. Then $(\boldsymbol{\alpha}, \boldsymbol{\xi})$ and $(\boldsymbol{\beta}, \boldsymbol{\eta})$ are partitions of $[c, d]$, and thus

$$|f(\xi) - f(\eta)|\, |g(d) - g(c)| \in \mathfrak{S}(f, \mathrm{d}g, [c, d])$$

and

$$\omega_{[c,d]}(f)\, |g(d) - g(c)| \leq \sup\, \mathfrak{S}_{f\Delta g}[c, d] = \omega\,(S_{f\Delta g}; [c, d]).$$

b) On the other hand, if $Q = (\boldsymbol{\beta}, \boldsymbol{\eta})$ and $\widetilde{Q} = (\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\eta}})$ are partitions of $[c, d]$ and $\boldsymbol{\alpha} = \boldsymbol{\beta} \cup \widetilde{\boldsymbol{\beta}}$, then $\boldsymbol{\alpha}$ is a division of $[c, d]$ and

$$\big|S(Q) - S(\widetilde{Q})\big| = \Big| \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \big(f(\eta_j') - f(\widetilde{\eta}_j')\big)\, \big(g(\alpha_j) - g(\alpha_{j-1})\big) \Big|,$$

where $\eta_j' = \eta_k$ if $[\alpha_{j-1}, \alpha_j] \subset [\beta_{k-1}, \beta_k]$ and $\widetilde{\eta}_j' = \widetilde{\eta}_k$ if $[\alpha_{j-1}, \alpha_j] \subset [\widetilde{\beta}_{k-1}, \widetilde{\beta}_k]$. Consequently,

$$\begin{aligned} \big|S(Q) - S(\widetilde{Q})\big| &\leq \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \big|f(\eta_j') - f(\widetilde{\eta}_j')\big|\, \big|g(\alpha_j) - g(\alpha_{j-1})\big| \\ &\leq \omega_{[c,d]}(f)\, V(g, \boldsymbol{\alpha}) \leq \omega_{[c,d]}(f)\, \mathrm{var}_c^d\, g, \end{aligned}$$

and finally

$$\omega(S_{f\Delta g}; [c, d]) = \sup\, \mathfrak{S}_{f\Delta g}[c, d] \leq \omega_{[c,d]}(f)\, \mathrm{var}_c^d\, g.$$

This proves that the inequalities (5.3.6) hold. □

5.3.5 Remark. If $\mathrm{var}_c^d\, g = \infty$, then the second inequality from (5.3.6) is trivial.

The following result indicates the role of bounded functions in the Riemann-Stieltjes integration theory. It implies that if $\int_a^b f\, \mathrm{d}g$ exists, then f is bounded on the complement of a finite collection of intervals where g is constant.

5.3.6 Theorem. *Let* $f, g : [a,b] \to \mathbb{R}$ *be such that* $\int_a^b f \, \mathrm{d}g$ *exists. Then there exists a division* $\boldsymbol{\alpha}$ *of* $[a,b]$ *such that for each* $j \in \{1, \ldots, \nu(\boldsymbol{\alpha})\}$, f *is bounded on* $[\alpha_{j-1}, \alpha_j]$, *or* g *is constant on* $[\alpha_{j-1}, \alpha_j]$.

Proof. Without loss of generality, assume that the integral $(\sigma)\int_a^b f \, \mathrm{d}g$ exists. Choose an arbitrary $\varepsilon > 0$. By Theorem 5.3.2, there is a division $\boldsymbol{\alpha}$ of $[a,b]$ such that

$$\sum_{j=1}^{\nu(\boldsymbol{\beta})} \omega(S_{f\Delta g}; [\beta_{j-1}, \beta_j]) < \varepsilon$$

holds for all divisions $\boldsymbol{\beta}$ of $[a,b]$ satisfying $\boldsymbol{\beta} \supset \boldsymbol{\alpha}$. In particular, we have

$$\omega(S_{f\Delta g}; [c,d]) < \varepsilon \quad \text{for each } [c,d] \subset [\alpha_{j-1}, \alpha_j] \text{ with } j \in \{1, \ldots, \nu(\boldsymbol{\alpha})\}.$$

For each $[c,d] \subset [\alpha_{j-1}, \alpha_j]$, we also have (cf. the proof of Lemma 5.3.4)

$$|f(\xi) - f(\eta)| \, |g(d) - g(c)| \le \omega(S_{f\Delta g}; [c,d]) < \varepsilon \quad \text{whenever } \xi, \eta \in [c,d].$$

Observe that if $\omega_{[c,d]}(f) = \infty$, then the previous inequality necessarily implies that $|g(d) - g(c)| = 0$, i.e., $g(c) = g(d)$.

Let us prove that for each $j \in \{1, \ldots, \nu(\boldsymbol{\alpha})\}$, f is bounded on $[\alpha_{j-1}, \alpha_j]$, or g is constant on $[\alpha_{j-1}, \alpha_j]$.

If f is unbounded on $[\alpha_{j-1}, \alpha_j]$ for a certain $j \in \{1, \ldots, \nu(\boldsymbol{\alpha})\}$, then

$$\omega_{[\alpha_{j-1}, \alpha_j]}(f) = \infty.$$

Our previous reasoning leads to the conclusion that

$$g(\alpha_j) = g(\alpha_{j-1}) = \gamma \quad \text{for a certain } \gamma \in \mathbb{R}.$$

It remains to show that $g(t) = \gamma$ for each $t \in (\alpha_{j-1}, \alpha_j)$. Note that we necessarily have

$$\omega_{[\alpha_{j-1}, t]}(f) = \infty \ \text{ or } \ \omega_{[t, \alpha_j]}(f) = \infty.$$

The former possibility implies that $g(t) = g(\alpha_{j-1}) = \gamma$, while the latter one implies
$g(t) = g(\alpha_j) = \gamma$. This completes the proof. □

5.3.7 Remark. Assume that $\int_a^b f \, \mathrm{d}g$ exists. Since the value of the integral does not change if we change arbitrarily the values of f on the intervals where g is constant, we see from Theorem 5.3.6 that it is always possible to find a bounded function $\widetilde{f} : [a,b] \to \mathbb{R}$ such that $\int_c^d f \, \mathrm{d}g = \int_c^d \widetilde{f} \, \mathrm{d}g$ whenever $[c,d] \subset [a,b]$.

The next statement provides other necessary and sufficient conditions for the existence of both kinds of RS-integrals.

5.3.8 Theorem. *Let* $f:[a,b]\to\mathbb{R}$, $g\in\mathrm{BV}([a,b])$ *and* $v(x)=\mathrm{var}_a^x\, g$ *for* $x\in[a,b]$. *Then:*
(i) *The integral* $(\sigma)\int_a^b f\,\mathrm{d}g$ *exists if and only if the integral* $(\sigma)\int_a^b f\,\mathrm{d}v$ *exists.*
(ii) *If* f *is bounded on* $[a,b]$, *then the integral* $(\delta)\int_a^b f\,\mathrm{d}g$ *exists if and only if the integral* $(\delta)\int_a^b f\,\mathrm{d}v$ *exists.*

Proof. a) For every interval $[c,d]\subset[a,b]$ we have $\mathrm{var}_c^d\, v=v(d)-v(c)$. Hence, by Lemma 5.3.4 (with g replaced by v),

$$\sum_{j=1}^{\nu(\boldsymbol{\alpha})}\omega_{[\alpha_{j-1},\alpha_j]}(f)\,\big(v(\alpha_j)-v(\alpha_{j-1})\big)=\sum_{j=1}^{\nu(\boldsymbol{\alpha})}\omega(S_{f\,\Delta\, v};[\alpha_{j-1},\alpha_j])$$

holds for any division $\boldsymbol{\alpha}$ of $[a,b]$. Consequently, using Lemma 5.3.4 for an arbitrary division $\boldsymbol{\alpha}$ of $[a,b]$ we deduce

$$\sum_{j=1}^{\nu(\boldsymbol{\alpha})}\omega(S_{f\,\Delta\, g};[\alpha_{j-1},\alpha_j])\le\sum_{j=1}^{\nu(\boldsymbol{\alpha})}\omega_{[\alpha_{j-1},\alpha_j]}(f)\ \mathrm{var}_{\alpha_{j-1}}^{\alpha_j}\, g$$
$$=\sum_{j=1}^{\nu(\boldsymbol{\alpha})}\omega_{[\alpha_{j-1},\alpha_j]}(f)\,\big(v(\alpha_j)-v(\alpha_{j-1})\big)=\sum_{j=1}^{\nu(\boldsymbol{\alpha})}\omega(S_{f\,\Delta\, v};[\alpha_{j-1},\alpha_j]).$$

Using Theorem 5.3.2, we can now easily prove that for both kinds of the RS-integral, the existence of $\int_a^b f\,\mathrm{d}v$ implies the existence of $\int_a^b f\,\mathrm{d}g$.
b) Assume that the integral $(\sigma)\int_a^b f\,\mathrm{d}g$ exists. We will prove that then the integral $(\sigma)\int_a^b f\,\mathrm{d}v$ exists as well.

Let $\varepsilon>0$ be given. By Theorem 5.3.2, there exists a division $\boldsymbol{\beta}$ of $[a,b]$ such that

$$\sum_{j=1}^{\nu(\boldsymbol{\alpha})}\omega\left(S_{f\Delta\, g};[\alpha_{j-1},\alpha_j]\right)<\varepsilon \tag{5.3.7}$$

holds for each its refinement $\boldsymbol{\alpha}\supset\boldsymbol{\beta}$. We can also assume that

$$0\le\mathrm{var}_a^b\, g-V(g,\boldsymbol{\alpha})<\varepsilon \tag{5.3.8}$$

holds for every division $\boldsymbol{\alpha}$ of $[a,b]$ such that $\boldsymbol{\alpha}\supset\boldsymbol{\beta}$. Finally, according to Theorem 5.3.6, we can suppose that for each $\boldsymbol{\alpha}\supset\boldsymbol{\beta}$ and $j\in\{1,\ldots,\nu(\boldsymbol{\alpha})\}$, f is bounded on $[\alpha_{j-1},\alpha_j]$, or g is constant on $[\alpha_{j-1},\alpha_j]$. For each such division, denote by $J_{\boldsymbol{\alpha}}$ the set of all $j\in\{1,\ldots,\nu(\boldsymbol{\alpha})\}$ such that f is bounded on $[\alpha_{j-1},\alpha_j]$.

Let $\boldsymbol{\alpha}$ be a division of $[a,b]$ such that $\boldsymbol{\alpha}\supset\boldsymbol{\beta}$. Using Lemma 5.3.4 we obtain

$$\begin{aligned}
\sum_{j=1}^{\nu(\boldsymbol{\alpha})}\omega(S_{f\,\Delta\,v};[\alpha_{j-1},\alpha_j]) &= \sum_{j\in J_{\boldsymbol{\alpha}}}\omega(S_{f\,\Delta\,v};[\alpha_{j-1},\alpha_j]) \\
&\le \sum_{j\in J_{\boldsymbol{\alpha}}}\omega_{[\alpha_{j-1},\alpha_j]}(f)\ \mathrm{var}_{\alpha_{j-1}}^{\alpha_j}\,g \\
&\le \sum_{j\in J_{\boldsymbol{\alpha}}}\omega_{[\alpha_{j-1},\alpha_j]}(f)\,|g(\alpha_j)-g(\alpha_{j-1})| \\
&\quad + \sum_{j\in J_{\boldsymbol{\alpha}}}\omega_{[\alpha_{j-1},\alpha_j]}(f)\,\big(\mathrm{var}_{\alpha_{j-1}}^{\alpha_j}\,g-|g(\alpha_j)-g(\alpha_{j-1})|\big) \\
&\le \sum_{j\in J_{\boldsymbol{\alpha}}}\omega(S_{f\,\Delta\,g};[\alpha_{j-1},\alpha_j]) \\
&\quad + \Big(\max_{j\in J_{\boldsymbol{\alpha}}}\omega_{[\alpha_{j-1},\alpha_j]}(f)\Big)\sum_{j\in J_{\boldsymbol{\alpha}}}\big(\mathrm{var}_{\alpha_{j-1}}^{\alpha_j}\,g-|g(\alpha_j)-g(\alpha_{j-1})|\big) \\
&< \varepsilon+\Big(\max_{j\in J_{\boldsymbol{\alpha}}}\omega_{[\alpha_{j-1},\alpha_j]}(f)\Big)\,\big(\mathrm{var}_a^b\,g-V(g,\boldsymbol{\alpha})\big) \\
&< \varepsilon+\Big(\max_{j\in J_{\boldsymbol{\beta}}}\omega_{[\beta j-1,\beta_j]}(f)\Big)\,\varepsilon.
\end{aligned}$$

By Theorem 5.3.2 we conclude that the integral $(\sigma)\int_a^b f\,\mathrm{d}v$ exists.

c) It remains to prove that if the function f is bounded on $[a,b]$, then the existence of the integral $(\delta)\int_a^b f\,\mathrm{d}g$ implies the existence of the integral $(\delta)\int_a^b f\,\mathrm{d}v$. In this situation, Theorems 5.1.7 and 5.2.6 imply that the integral $(\sigma)\int_a^b f\,\mathrm{d}g$ exists and the functions f,g have no common point of discontinuity in (a,b). Moreover, by Lemma 2.3.3, the functions f, v have no common point of discontinuity in (a,b), either. Finally, since the integral $(\sigma)\int_a^b f\,\mathrm{d}v$ exists by part b) of this proof, the existence of the integral $(\delta)\int_a^b f\,\mathrm{d}v$ follows from Corollary 5.2.8. (As g has a bounded variation on $[a,b]$, the functions g and v are bounded on $[a,b]$.) □

5.3.9 Theorem. *Assume that* $f:[a,b]\to\mathbb{R}$, $g\in\mathrm{BV}([a,b])$, *and the integral* $\int_a^b f\,\mathrm{d}g$ *exists. Then the integral* $\int_a^b |f|\,\mathrm{d}g$ *exists as well.*

Proof. By Theorem 2.1.21 and Lemma 5.1.13 we can restrict ourselves to the case when g is nondecreasing on $[a,b]$. Then $\mathrm{var}_c^d\,g=g(d)-g(c)$ for each $[c,d]\subset[a,b]$. Thus, due to Lemma 5.3.4, the relations

$$\sum_{j=1}^{\nu(\boldsymbol{\alpha})}\omega(S_{f\Delta\,g};[\alpha_{j-1},\alpha_j])=\sum_{j=1}^{\nu(\boldsymbol{\alpha})}\omega_{[\alpha_{j-1},\alpha_j]}(f)\,\big(g(\alpha_j)-g(\alpha_{j-1})\big)$$

and

$$\sum_{j=1}^{\nu(\boldsymbol{\alpha})} \omega(S_{|f|\Delta g};[\alpha_{j-1},\alpha_j]) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \omega_{[\alpha_{j-1},\alpha_j]}(|f|)\,\big(g(\alpha_j)-g(\alpha_{j-1})\big)$$

hold for all divisions $\boldsymbol{\alpha}$ of $[a,b]$. On the other hand,

$$\big||f(x)|-|f(y)|\big| \le |f(x)-f(y)| \quad \text{for all } x,y\in[a,b].$$

Thus we have $\omega_{[c,d]}(|f|) \le \omega_{[c,d]}(f)$ for any $[c,d]\subset[a,b]$ and, therefore,

$$\sum_{j=1}^{\nu(\boldsymbol{\alpha})} \omega(S_{|f|\Delta g};[\alpha_{j-1},\alpha_j]) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \omega_{[\alpha_{j-1},\alpha_j]}(|f|)\,\big(g(\alpha_j)-g(\alpha_{j-1})\big)$$
$$\le \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \omega_{[\alpha_{j-1},\alpha_j]}(f)\,\big(g(\alpha_j)-g(\alpha_{j-1})\big) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \omega\,(S_{f\Delta g};[\alpha_{j-1},\alpha_j]).$$

The statement of the theorem now follows from Theorem 5.3.2. □

The next assertion is a direct corollary of Lemma 5.1.11 and Theorems 5.3.8 and 5.3.9.

5.3.10 Corollary. *Let* $f:[a,b]\to\mathbb{R}$, $g\in\mathrm{BV}([a,b])$, *and* $v(x)=\mathrm{var}_a^x\, g$ *for* $x\in[a,b]$. *Then:*

(i) *If the integral* $(\sigma)\int_a^b f\,\mathrm{d}g$ *exists, then the integral* $(\sigma)\int_a^b |f|\,\mathrm{d}v$ *exists as well and*

$$\Big|(\sigma)\int_a^b f\,\mathrm{d}g\Big| \le (\sigma)\int_a^b |f|\,\mathrm{d}v \le \|f\|_\infty\,\mathrm{var}_a^b g.$$

(ii) *If the integral* $(\delta)\int_a^b f\,\mathrm{d}g$ *exists and the function* f *is bounded on* $[a,b]$, *then the integral* $(\delta)\int_a^b |f|\,\mathrm{d}v$ *exists as well and*

$$\Big|(\delta)\int_a^b f\,\mathrm{d}g\Big| \le (\delta)\int_a^b |f|\,\mathrm{d}v \le \|f\|_\infty\,\mathrm{var}_a^b g.$$

5.4 Substitution

All statements in this section hold in the same form for both kinds of the RS-integral. The next result is a fairly straightforward consequence of Definition 5.3.1.

5.4.1 Lemma. *If the integral* $\int_a^b f\,\mathrm{d}g$ *exists, then the inequality*

$$\Big|\int_a^b f\,\mathrm{d}g - S(P)\Big| \le \sum_{j=1}^{\nu(P)} \omega(S_{f\Delta g};[\alpha_{j-1},\alpha_j]) \tag{5.4.1}$$

holds for any partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ *of* $[a,b]$.

Proof. Let $\varepsilon > 0$ and a partition $P = (\widetilde{\boldsymbol{\alpha}}, \widetilde{\boldsymbol{\xi}})$ of $[a,b]$ be given. For both kinds of the RS-integral, we can choose a partition $\widetilde{P} = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$ such that $\widetilde{\boldsymbol{\alpha}} \supset \boldsymbol{\alpha}$ and

$$\left| \int_a^b f \,\mathrm{d}g - S(\widetilde{P}) \right| < \varepsilon.$$

Since $\widetilde{\boldsymbol{\alpha}}$ is a refinement of $\boldsymbol{\alpha}$, we can split it so that

$$\widetilde{\boldsymbol{\alpha}} = \bigcup_{j=1}^{m} \widetilde{\boldsymbol{\alpha}}^j, \quad \text{where } \widetilde{\boldsymbol{\alpha}}^j \text{ is a division of } [\alpha_{j-1}, \alpha_j] \text{ for } j \in \{1, \dots, m\}.$$

Similarly, $\widetilde{\boldsymbol{\xi}} = \{\widetilde{\boldsymbol{\xi}}^1, \widetilde{\boldsymbol{\xi}}^2, \dots, \widetilde{\boldsymbol{\xi}}^m\}$, where $\widetilde{\boldsymbol{\xi}}^j$ are sets such that

$$P_j = (\widetilde{\boldsymbol{\alpha}}^j, \widetilde{\boldsymbol{\xi}}_j) \text{ are partitions of } [\alpha_{j-1}, \alpha_j] \text{ for } j \in \{1, \dots, m\}.$$

We thus have

$$\begin{aligned}
\left| \int_a^b f \,\mathrm{d}g - S(P) \right| &\le \left| \int_a^b f \,\mathrm{d}g - S(\widetilde{P}) \right| + \left| S(\widetilde{P}) - S(P) \right| \\
&< \varepsilon + \sum_{j=1}^{m} \left| f(\xi_j)\,(g(\alpha_j) - g(\alpha_{j-1})) - S(\widetilde{P}_j) \right| \\
&\le \varepsilon + \sum_{j=1}^{m} \omega(S_{f\Delta g}; [\alpha_{j-1}, \alpha_j]).
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, it follows that (5.4.1) holds. □

5.4.2 Corollary. *If the integral $\int_a^b f \,\mathrm{d}g$ exists and $[c,d] \subset [a,b]$, then*

$$\left| \int_c^d f \,\mathrm{d}g - f(\xi)\,(g(d) - g(c)) \right| \le \omega(S_{f\Delta g}; [c,d])$$

holds for every $\xi \in [c,d]$.

The following substitution theorem is another corollary of Lemma 5.4.1.

5.4.3 Theorem (SUBSTITUTION THEOREM)**.** *Let $f,\ g,\ h : [a,b] \to \mathbb{R}$ be such that f is bounded and $\int_a^b g \,\mathrm{d}h$ exists. Then one of the integrals*

$$\int_a^b f(x) \,\mathrm{d}\left(\int_a^x g \,\mathrm{d}h \right) \quad \textit{and} \quad \int_a^b f g \,\mathrm{d}h$$

exists if and only if the other exists. In this case we have

$$\int_a^b f(x) \,\mathrm{d}\left(\int_a^x g \,\mathrm{d}h \right) = \int_a^b f g \,\mathrm{d}h. \tag{5.4.2}$$

Proof. Consider the function $w:[a,b]\to\mathbb{R}$ given by

$$w(x)=\int_a^x g\,\mathrm{d}h,\qquad x\in[a,b].$$

For any partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$, we have

$$\begin{aligned}
&|S(fg,\mathrm{d}h,P)-S(f,\mathrm{d}w,P)|\\
&\quad=\left|\sum_{j=1}^{\nu(P)} f(\xi_j)\,g(\xi_j)\,(h(\alpha_j)-h(\alpha_{j-1}))-\sum_{j=1}^{\nu(P)} f(\xi_j)\,(w(\alpha_j)-w(\alpha_{j-1}))\right|\\
&\quad\le\sum_{j=1}^{\nu(P)}|f(\xi_j)|\left|g(\xi_j)\,(h(\alpha_j)-h(\alpha_{j-1}))-\int_{\alpha_{j-1}}^{\alpha_j} g\,\mathrm{d}h\right|\\
&\quad\le\|f\|_\infty\left(\sum_{j=1}^{\nu(P)}\left|g(\xi_j)\,(h(\alpha_j)-h(\alpha_{j-1}))-\int_{\alpha_{j-1}}^{\alpha_j} g\,\mathrm{d}h\right|\right).
\end{aligned}$$

Now, Corollary 5.4.2 yields

$$\big|S(fg,\mathrm{d}h,P)-S(f,\mathrm{d}w,P)\big|\le\|f\|_\infty\sum_{j=1}^{\nu(P)}\omega(S_{g\Delta h};[\alpha_{j-1},\alpha_j]),$$

and the proof of (5.4.2) is completed by using Theorem 5.3.2. □

Setting $h(t)\equiv t$ in Theorem 5.4.3, we get the following statement.

5.4.4 Corollary. *If $f:[a,b]\to\mathbb{R}$ is bounded, $g:[a,b]\to\mathbb{R}$ is Riemann integrable, and $p(x)=\int_a^x g(t)\,\mathrm{d}t$, then one of the integrals*

$$\int_a^b f\,\mathrm{d}p\quad\textit{and}\quad\int_a^b f(x)\,g(x)\,\mathrm{d}x$$

exists if and only if the other exists. In this case, we have

$$\int_a^b f\,\mathrm{d}p=\int_a^b f(x)\,g(x)\,\mathrm{d}x.$$

5.4.5 Theorem (SECOND SUBSTITUTION THEOREM).

Assume that $\phi:[c,d]\to\mathbb{R}$ is continuous, strictly monotone and maps the interval $[c,d]$ onto $[a,b]$. Then for arbitrary functions $f,g:[a,b]\to\mathbb{R}$, the following statement holds:

$$\textit{If}\ \int_a^b f(x)\,\mathrm{d}g(x)\ \textit{exists, then}\ \int_c^d f(\phi(x))\,\mathrm{d}g(\phi(x))\ \textit{exists as well,}$$

and the relation

$$\pm\int_c^d f(\phi(x))\,\mathrm{d}g(\phi(x)) = \int_a^b f(x)\,\mathrm{d}g(x) \tag{5.4.3}$$

holds with the plus sign if ϕ is increasing, and with the minus sign if ϕ is decreasing.

Proof. Assume, for example, that ϕ is decreasing. Then $b=\phi(c)$ and $a=\phi(d)$. For a given partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of the interval $[c,d]$, set

$$\beta_{\nu(P)-j}=\phi(\alpha_j) \text{ and } \eta_{\nu(P)-j}=\phi(\xi_j) \quad \text{for } j\in\{1,\ldots,\nu(P)\}.$$

Then $Q=(\boldsymbol{\beta},\boldsymbol{\eta})$, where $\boldsymbol{\beta}=\{\beta_0,\beta_1,\ldots,\beta_{\nu(P)}\}$, $\boldsymbol{\eta}=\{\eta_1,\eta_2,\ldots,\eta_{\nu(P)}\}$, is a partition of $[a,b]$. We write $\boldsymbol{\beta}=\phi(\boldsymbol{\alpha})$, $\boldsymbol{\eta}=\phi(\boldsymbol{\xi})$, and $Q=\phi(P)$. Obviously, if $\boldsymbol{\alpha}\supset\boldsymbol{\alpha}'$, then also $\phi(\boldsymbol{\alpha})\supset\phi(\boldsymbol{\alpha}')$. Since ϕ is uniformly continuous on $[c,d]$, we have

$$\lim_{|\boldsymbol{\alpha}|\to 0}|\phi(\boldsymbol{\alpha})|=0.$$

Moreover,

$$\sum_{j=1}^{\nu(P)} f(\phi(\xi_j))\big(g(\phi(\alpha_j))-g(\phi(\alpha_{j-1}))\big) = -\sum_{i=1}^{\nu(Q)} f(\eta_j)\big(g(\beta_j)-g(\beta_{j-1})\big)$$

holds for every partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[c,d]$. This fact easily implies the statement of the theorem; we leave the details to the reader. The case when ϕ is increasing can be handled in a similar way. □

The following theorem is yet another variant of the substitution theorem. Its proof is left as an exercise for the reader.

5.4.6 Theorem. *Let $\phi:[a,b]\to[\phi(a),\phi(b)]$ be increasing and continuous, and let $\psi:[\phi(a),\phi(b)]\to[a,b]$ be the inverse of ϕ. Moreover, let an arbitrary function $f:[a,b]\to\mathbb{R}$ be given. Then, if one of the integrals*

$$\int_a^b f(x)\,\mathrm{d}x, \qquad \int_{\phi(a)}^{\phi(b)} f(\psi(x))\,\mathrm{d}\psi(x)$$

exists, the other exists as well, and

$$\int_a^b f(x)\,\mathrm{d}x = \int_{\phi(a)}^{\phi(b)} f(\psi(x))\,\mathrm{d}\psi(x).$$

5.4.7 Exercise. Prove Theorem 5.4.6 for both kinds of the RS-integral.

5.5 Integration by parts

The following statement is a generalization of the classical integration by parts formula for the Riemann integral (see Exercise 5.5.3).

5.5.1 Theorem (INTEGRATION BY PARTS). *If one of the integrals* $\int_a^b f\,\mathrm{d}g$ *and* $\int_a^b g\,\mathrm{d}f$ *exists, then the other exists as well, and we have*

$$\int_a^b f\,\mathrm{d}g + \int_a^b g\,\mathrm{d}f = f(b)g(b) - f(a)g(a). \tag{5.5.1}$$

Proof. a) Let an arbitrary partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$ be given. Set $m = \nu(P)$. Rearranging the terms in the sum $S(f, \mathrm{d}g, P)$, we get

$$\begin{aligned} S(f,\mathrm{d}g,P) &= f(\xi_1)\,(g(\alpha_1) - g(a)) + f(\xi_2)\,(g(\alpha_2) - g(\alpha_1)) \\ &\quad + \cdots + f(\xi_m)\,(g(b) - g(\alpha_{m-1})) \\ &= -f(a)\,g(a) - (f(\xi_1) - f(a))\,g(a) - (f(\xi_2) - f(\alpha_1))\,g(\alpha_1) \\ &\quad - (f(\alpha_1) - f(\xi_1))\,g(\alpha_1) - \cdots - (f(\xi_m) - f(\alpha_{m-1}))\,g(\alpha_{m-1}) \\ &\quad - (f(\alpha_{m-1}) - f(\xi_{m-1}))\,g(\alpha_{m-1}) - (f(b) - f(\xi_m))\,g(b) + f(b)g(b) \\ &= f(b)\,g(b) - f(a)\,g(a) - S(g,\mathrm{d}f,Q), \end{aligned}$$

where the partition $Q = (\boldsymbol{\beta}, \boldsymbol{\eta})$ of $[a,b]$ is such that

$$\begin{aligned} \boldsymbol{\beta} &= \{a, \xi_1, \alpha_1, \xi_2, \alpha_2, \ldots, \alpha_{m-1}, \xi_m, b\}, \\ \boldsymbol{\eta} &= \{a, \alpha_1, \alpha_1, \alpha_2, \alpha_2, \ldots, \alpha_{m-1}, \alpha_{m-1}, b\}. \end{aligned}$$

Clearly, $\boldsymbol{\beta}$ is a refinement of $\boldsymbol{\alpha}$. [1]

b) Assume that the integral $(\sigma)\int_a^b g\,\mathrm{d}f$ exists and let $\varepsilon > 0$ be given. Choose a division $\boldsymbol{\beta}_\varepsilon$ of $[a,b]$ such that

$$\left| S(g,\mathrm{d}f,Q) - (\sigma)\int_a^b g\,\mathrm{d}f \right| < \varepsilon$$

holds for all partitions $Q = (\boldsymbol{\beta}, \boldsymbol{\eta})$ of $[a,b]$ such that $\boldsymbol{\beta} \supset \boldsymbol{\beta}_\varepsilon$.

Let $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ be an arbitrary partition of $[a,b]$ such that $\boldsymbol{\alpha} \supset \boldsymbol{\beta}_\varepsilon$. By the first part of the proof there is a partition $Q = (\boldsymbol{\beta}, \boldsymbol{\eta})$ of $[a,b]$ such that

$$\begin{aligned} S(f,\mathrm{d}g,P) - f(b)\,g(b) + f(a)\,g(a) + (\sigma)\int_a^b g\,\mathrm{d}f \\ = (\sigma)\int_a^b g\,\mathrm{d}f - S(g,\mathrm{d}f,Q), \end{aligned}$$

[1] Of course, if $\xi_j = \alpha_{j-1}$ or $\xi_j = \alpha_j$ for some j, we have to leave out ξ_j from $\boldsymbol{\beta}$ and the corresponding tag from $\boldsymbol{\eta}$.

while $\boldsymbol{\beta} \supset \boldsymbol{\alpha}$ and hence also $\boldsymbol{\beta} \supset \boldsymbol{\beta}_\varepsilon$. Consequently,

$$\left| S(f, \mathrm{d}g, P) - f(b)g(b) + f(a)g(a) + (\sigma) \int_a^b g \, \mathrm{d}f \right| \\ = \left| (\sigma) \int_a^b g \, \mathrm{d}f - S(g, \mathrm{d}f, Q) \right| < \varepsilon.$$

This means that the integral $(\sigma) \int_a^b f \, \mathrm{d}g$ exists and (5.5.1) holds.

By interchanging the roles of f and g, we immediately see that if $(\sigma) \int_a^b f \, \mathrm{d}g$ exists, then $(\sigma) \int_a^b g \, \mathrm{d}f$ exists as well and (5.5.1) holds.

c) Also the statement of the theorem for (δ) RS-integrals follows easily from the relation (5.5.1); the details are left to the reader. □

5.5.2 Exercise. Prove Theorem 5.5.1 for the (δ) RS-integral.

5.5.3 Exercise. The classical integration by parts theorem for the Riemann integral reads as follows: *Assume that $f, g : [a,b] \to \mathbb{R}$ are Riemann integrable and let $F, G : [a,b] \to \mathbb{R}$ be given by*

$$F(x) = \int_a^x f, \quad G(x) = \int_a^x g \quad \textit{for } x \in [a,b].$$

If one of the integrals $\int_a^b f\, G$ and $\int_a^b F\, g$ exists, then the other exists as well, and we have

$$\int_a^b f\, G + \int_a^b F\, g = F(b)\, G(b) - F(a)\, G(a).$$

Show that this result is a consequence of Theorem 5.5.1.
Hint: Use Corollary 5.4.4.

5.6 Uniform convergence and existence of the integral

All statements in this section except Theorem 5.6.4 and Exercise 5.6.5 hold in the same form for both kinds of the RS-integral.

5.6.1 Theorem. *Assume that $g \in \mathrm{BV}([a,b])$, $f : [a,b] \to \mathbb{R}$ is bounded, $f_n : [a,b] \to \mathbb{R}$, $n \in \mathbb{N}$, are such that the integral $\int_a^b f_n \, \mathrm{d}g$ exists for every $n \in \mathbb{N}$, and*

$$\lim_{n \to \infty} \|f_n - f\|_\infty = 0. \tag{5.6.1}$$

Then the integral $\int_a^b f\, \mathrm{d}g$ *exists as well and*

$$\lim_{n\to\infty} \int_a^b f_n\, \mathrm{d}g = \int_a^b f\, \mathrm{d}g. \tag{5.6.2}$$

Proof. If $\mathrm{var}_a^b g = 0$, then, by Lemma 2.1.16, g is constant on $[a,b]$ and the statement of the theorem is obvious. Therefore, assume that $\mathrm{var}_a^b g > 0$.

Let $\varepsilon > 0$ be given. According to (5.6.1), we can choose $n_\varepsilon \in \mathbb{N}$ such that

$$\|f_n - f\|_\infty < \frac{\varepsilon}{\mathrm{var}_a^b g} \quad \text{and} \quad \|f_n\|_\infty < \|f\|_\infty + 1 \qquad \text{for all } n \ge n_\varepsilon. \tag{5.6.3}$$

Furthermore, Lemma 5.1.11 (i) implies that

$$\left| \int_a^b f_n\, \mathrm{d}g \right| \le \|f_n\|_\infty \mathrm{var}_a^b g \le (\|f\|_\infty + 1)\, \mathrm{var}_a^b g \qquad \text{for all } n \ge n_\varepsilon.$$

Hence, there are an increasing subsequence $\{n_k\}$ of $\mathbb{N}$ and $I \in \mathbb{R}$ such that

$$\lim_{k\to\infty} \int_a^b f_{n_k}\, \mathrm{d}g = I.$$

In particular, there exists a $k_\varepsilon \in \mathbb{N}$ such that

$$n_{k_\varepsilon} \ge n_\varepsilon \quad \text{and} \quad \left| \int_a^b f_{n_{k_\varepsilon}}\, \mathrm{d}g - I \right| < \varepsilon. \tag{5.6.4}$$

Now, let $\boldsymbol{\alpha}_\varepsilon$ be a division of $[a,b]$ such that

$$\left.\begin{array}{l} \left| S(f_{n_{k_\varepsilon}}, \mathrm{d}g, P) - \displaystyle\int_a^b f_{n_{k_\varepsilon}} \mathrm{d}g \right| < \varepsilon \\ \qquad \text{whenever } P = (\boldsymbol{\alpha}, \boldsymbol{\xi}) \text{ ia a partition of } [a,b] \text{ such that } \boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon. \end{array}\right\} \tag{5.6.5}$$

Since $n_{k_\varepsilon} \ge n_\varepsilon$, it follows from (5.6.3) that

$$\left| S(f, \mathrm{d}g, P) - S(f_{n_{k_\varepsilon}}, \mathrm{d}g, P) \right| \le \|f - f_{n_{k_\varepsilon}}\|_\infty \,\mathrm{var}_a^b g < \varepsilon$$

for every partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$. Further, using (5.6.4)–(5.6.5), we deduce that

$$\begin{aligned} \left| S(f, \mathrm{d}g, P) - I \right| &\le \left| S(f, \mathrm{d}g, P) - S(f_{n_{k_\varepsilon}}, \mathrm{d}g, P)) \right| \\ &\quad + \left| S(f_{n_{k_\varepsilon}}, \mathrm{d}g, P) - \int_a^b f_{n_{k_\varepsilon}}\, \mathrm{d}g \right| + \left| \int_a^b f_{n_{k_\varepsilon}}\, \mathrm{d}g - I \right| < 3\,\varepsilon \end{aligned}$$

for each partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$ such that $\boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon$. Thus,

$$\int_a^b f \,\mathrm{d}g = I.$$

Finally, since by Lemmas 5.1.11 and 5.1.13 we have

$$\left| \int_a^b f_n \,\mathrm{d}g - \int_a^b f \,\mathrm{d}g \right| \le \|f_n - f\|_\infty \operatorname{var}_a^b g,$$

equality (5.6.2) follows from our assumption (5.6.1). This proves the statement for the (σ)RS-integral. The proof for the (δ)RS-integral is analogous and is left to the reader. □

5.6.2 Exercise. Prove the statement of Theorem 5.6.1 for (δ)RS-integrals.

5.6.3 Theorem. *Let $f \in \mathrm{C}([a,b])$ and $g \in \mathrm{BV}([a,b])$. Then both the integrals $\int_a^b f \,\mathrm{d}g$ and $\int_a^b g \,\mathrm{d}f$ exist.*

Proof. By Theorems 2.1.21, 5.1.7, 5.5.1 and Lemma 5.1.13, it is sufficient to prove the existence of the integral $(\delta) \int_a^b f \,\mathrm{d}g$ in the case when g is nondecreasing on $[a,b]$.

Let $\varepsilon > 0$ be given. If $g(b) = g(a)$, then g is constant on $[a,b]$ and hence $(\delta) \int_a^b f \,\mathrm{d}g = 0$. Thus, we can assume that $g(b) - g(a) > 0$. Next, since every function continuous on a compact interval is also uniformly continuous on this interval, we can find a $\delta_\varepsilon > 0$ such that

$$|f(x) - f(y)| < \frac{\varepsilon}{g(b) - g(a)} \quad \text{for all } x, y \in [a,b] \text{ such that } |x-y| < \delta_\varepsilon. \tag{5.6.6}$$

Now, consider two partitions $P = (\boldsymbol{\alpha}, \boldsymbol{\xi}), Q = (\boldsymbol{\beta}, \boldsymbol{\eta})$ of $[a,b]$ such that $|\boldsymbol{\alpha}| < \delta_\varepsilon$ and $\boldsymbol{\beta} \supset \boldsymbol{\alpha}$. We will show that $|S(P) - S(Q)| < \varepsilon$. By Theorem 5.1.15 and Exercise 5.1.16 (ii), this will guarantee the existence of the integral $(\delta) \int_a^b f \,\mathrm{d}g$.

Denote $m = \nu(\boldsymbol{\alpha})$. Since $\boldsymbol{\beta} \supset \boldsymbol{\alpha}$, the elements of $\boldsymbol{\beta}$ can be for $j \in \{1, \dots, m\}$ and $i \in \{1, \dots n_j\}$ denoted by β_i^j, where $\alpha_{j-1} = \beta_0^j < \dots < \beta_{n_j}^j = \alpha_j$. The tag corresponding to $[\beta_{i-1}^j, \beta_i^j]$ will be denoted by η_i^j. Then

$$S(P) = \sum_{j=1}^m f(\xi_j)\,(g(\alpha_j) - g(\alpha_{j-1})) = \sum_{j=1}^m f(\xi_j) \sum_{i=1}^{n_j} (g(\beta_i^j) - g(\beta_{i-1}^j))$$

and

$$|S(P) - S(Q)| \le \sum_{j=1}^m \sum_{i=1}^{n_j} |f(\xi_j) - f(\eta_i^j)|\,(g(\beta_i^j) - g(\beta_{i-1}^j)).$$

Since $|\xi_j - \eta_i^j| < |\boldsymbol{\alpha}| < \delta_\varepsilon$ for all $j \in \{1,\dots,m\}$ and $i \in \{1,\dots,n_j\}$, it follows from (5.6.6) that

$$|S(P) - S(Q)| < \frac{\varepsilon}{g(b)-g(a)} \sum_{j=1}^{m}\sum_{i=1}^{n_j} (g(\beta_i^j) - g(\beta_{i-1}^j))$$
$$= \frac{\varepsilon}{g(b)-g(a)} (g(b)-g(a)) = \varepsilon.$$

This completes the proof. □

5.6.4 Theorem. (i) *If* $f \in \mathrm{BV}([a,b])$ *is left-continuous on* $(a,b]$*, then the integrals*

$$(\sigma)\int_a^b f\,\mathrm{d}g \quad \textit{and} \quad (\sigma)\int_a^b g\,\mathrm{d}f$$

exist for every $g \in \mathrm{G}([a,b])$ *which is right-continuous on* $[a,b)$.

(ii) *If* $f \in \mathrm{BV}([a,b])$ *is right-continuous on* $[a,b)$*, then the integrals*

$$(\sigma)\int_a^b f\,\mathrm{d}g \quad \textit{and} \quad (\sigma)\int_a^b g\,\mathrm{d}f.$$

exist for every $g \in \mathrm{G}([a,b])$ *which is left-continuous on* $(a,b]$.

Proof. In both cases, it suffices (thanks to Theorem 5.5.1) to prove the existence of the integral $(\sigma)\int_a^b g\,\mathrm{d}f$.

Let $g \in \mathrm{G}([a,b])$ be right-continuous on $[a,b)$, i.e., $g \in \mathrm{G_R}([a,b])$. By Lemmas 4.2.6 and 4.2.8 we have

$$\mathrm{G_R}([a,b]) = \mathrm{cl}(\mathrm{G_R}([a,b]) \cap \mathrm{S}([a,b])) = \mathrm{cl}(\mathrm{Lin}\Big(\{\chi_{[\tau,b]} : \tau \in [a,b]\}\Big)).$$

Thus, by Lemma 5.1.13 and Theorem 5.6.1, it is sufficient to prove that the integral $(\sigma)\int_a^b g\,\mathrm{d}f$ exists if $g = \chi_{[\tau,b]}$ for some $\tau \in [a,b]$.

If $g = \chi_{[a,b]}$, i.e., $\tau = a$ and $g = 1$ on $[a,b]$, then $(\sigma)\int_a^b g\,\mathrm{d}f = f(b) - f(a)$ (see Exercise 5.1.6 (ii)). Therefore, we may assume that $\tau \in (a,b]$ and $g = \chi_{[\tau,b]}$. We will show that

$$(\sigma)\int_a^b g\,\mathrm{d}f = f(b) - f(\tau). \tag{5.6.7}$$

By Remark 5.1.8 we can restrict ourselves to partitions $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$ such that $\tau \in \boldsymbol{\alpha}$. For every such partition P, let $k(P)$ denote the unique index $k \in \{1,\dots,\nu(P)\}$ such that $\tau = \alpha_k$. Then $\alpha_{k(P)-1} \le \xi_{k(P)} \le \alpha_{k(P)} = \tau$, and

$$S(P) = \begin{cases} f(b) - f(\tau) & \text{if } \xi_{k(P)} < \tau, \\ f(b) - f(\alpha_{k(P)-1}) & \text{if } \xi_{k(P)} = \tau. \end{cases}$$

Consequently,

$$\big|S(P)-\big(f(b)-f(\tau)\big)\big| = \begin{cases} 0 & \text{if } \xi_{k(P)} < \tau, \\ |f(\tau)-f(\alpha_{k(P)-1})| & \text{if } \xi_{k(P)} = \tau. \end{cases} \tag{5.6.8}$$

Thanks to the continuity of the function f at τ from the left, we can choose a division $\boldsymbol{\alpha}_\varepsilon$ of $[a,b]$ containing the point τ and such that

$$|f(\tau)-f(\alpha_{k(P)-1})| < \varepsilon$$

holds for any $\boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon$. By (5.6.8), we have

$$\big|S(P)-\big(f(b)-f(\tau)\big)\big| < \varepsilon$$

for all partitions $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$ such that $\boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon$, which implies (5.6.7) and the proof of the statement (i) is complete.

The second statement can be proved similarly. □

5.6.5 Exercises. (i) For both kinds of the RS-integral, prove the following assertion: *If* $f \in \mathrm{BV}([a,b])$ *is continuous, then the integral* $\int_a^b f\,\mathrm{d}g$ *exists for every* $g \in \mathrm{G}([a,b])$.

(ii) Give a detailed proof of the statement (ii) in Theorem 5.6.4.

5.6.6 Remark. Let us mention (without proof) another interesting existence result. It was proved in 1936 by L. C. Young, one of the pioneers of integration theory (see [156]): *Assume that* $f:[a,b]\to\mathbb{R}$ *and* $g:[a,b]\to\mathbb{R}$ *satisfy*

$$|f(x)-f(y)| \le K\,|x-y|^\alpha \quad \textit{and} \quad |g(x)-g(y)| \le L\,|x-y|^\beta \quad \textit{for } x,y \in [a,b],$$

where $K,L \in [0,\infty)$, $\alpha,\beta \in (0,\infty)$, $\alpha+\beta>1$. *Then* $(\delta)\int_a^b f\,\mathrm{d}g$ *exists.*

5.7 Pointwise convergence

In order to derive a convergence theorem for integrals $\int_a^b f_n\,\mathrm{d}g$ when the sequence $\{f_n\}$ is not uniformly convergent, we introduce the following concepts of the Darboux upper and lower integrals.

5.7.1 Definition. Let $g:[a,b]\to\mathbb{R}$ be nondecreasing. For a function $f:[a,b]\to\mathbb{R}$ and a division $\boldsymbol{\alpha}$ of the interval $[a,b]$, put

$$\overline{S}(f,\mathrm{d}g,\boldsymbol{\alpha}) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \Big(\sup_{x\in[\alpha_{j-1},\alpha_j]} f(x)\Big)\big(g(\alpha_j)-g(\alpha_{j-1})\big),$$

$$\underline{S}(f,\mathrm{d}g,\boldsymbol{\alpha}) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \Big(\inf_{x\in[\alpha_{j-1},\alpha_j]} f(x)\Big)\big(g(\alpha_j)-g(\alpha_{j-1})\big).$$

Furthermore, we define the *upper integral* of f with respect to g as

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

and the *lower integral* of f with respect to g as

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

If the functions f, g are fixed, we write simply $\overline{S}(\boldsymbol{\alpha})$ instead of $\overline{S}(f, \mathrm{d}g, \boldsymbol{\alpha})$ and $\underline{S}(\boldsymbol{\alpha})$ instead of $\underline{S}(f, \mathrm{d}g, \boldsymbol{\alpha})$.

5.7.2 Lemma. *Let* $f : [a,b] \to \mathbb{R}$ *and let* $g : [a,b] \to \mathbb{R}$ *be nondecreasing. Then*

$$\overline{\int_a^b} f \, \mathrm{d}g = \underline{\int_a^b} f \, \mathrm{d}g = I \in \mathbb{R} \tag{5.7.1}$$

if and only if $(\sigma) \int_a^b f \, \mathrm{d}g = I$.

Proof. a) Assume that (5.7.1) holds. Since g is nondecreasing, it follows directly from Definition 5.7.1 that

$$\underline{S}(\boldsymbol{\alpha}) \le S(\boldsymbol{\alpha}, \boldsymbol{\xi}) \le \overline{S}(\boldsymbol{\alpha}) \ \text{ for all partitions } \ (\boldsymbol{\alpha}, \boldsymbol{\xi}) \ \text{ of } \ [a,b],$$

and

$$\underline{S}(\widetilde{\boldsymbol{\alpha}}) \ge \underline{S}(\boldsymbol{\alpha}) \ \text{ and } \ \overline{S}(\widetilde{\boldsymbol{\alpha}}) \le \overline{S}(\boldsymbol{\alpha}) \ \text{ if } \ \widetilde{\boldsymbol{\alpha}} \supset \boldsymbol{\alpha}.$$

Using the first fact, it is not difficult to verify that for each $k \in \mathbb{N}$ there is a division $\boldsymbol{\alpha}^k$ of $[a,b]$ such that the inequalities

$$I - \frac{1}{k} < \underline{S}(\boldsymbol{\alpha}^k) \le S(\boldsymbol{\alpha}^k, \boldsymbol{\xi}^k) \le \overline{S}(\boldsymbol{\alpha}^k) < I + \frac{1}{k}$$

hold for whenever $(\boldsymbol{\alpha}^k, \boldsymbol{\xi}^k)$ is a partition of $[a,b]$. For a given $\varepsilon > 0$, choose $k_\varepsilon > \frac{1}{\varepsilon}$ and set $\boldsymbol{\alpha}_\varepsilon = \boldsymbol{\alpha}^{k_\varepsilon}$. Then

$$I - \varepsilon < \underline{S}(\boldsymbol{\alpha}^{k_\varepsilon}) \le \underline{S}(\boldsymbol{\alpha}) \le S(\boldsymbol{\alpha}, \boldsymbol{\xi}) \le \overline{S}(\boldsymbol{\alpha}) \le \overline{S}(\boldsymbol{\alpha}^{k_\varepsilon}) < I + \varepsilon$$

whenever $\boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon$ and $(\boldsymbol{\alpha}, \boldsymbol{\xi})$ is a partition of $[a,b]$. It follows that $(\sigma) \int_a^b f \, \mathrm{d}g = I$.

b) Assume that $(\sigma) \int_a^b f \mathrm{d}g$ exists. Let $\varepsilon > 0$ be given. By Theorem 5.1.15 there exists a division $\boldsymbol{\alpha}$ such that the inequality $\big| S(\boldsymbol{\alpha}, \boldsymbol{\xi}) - S(\boldsymbol{\alpha}, \boldsymbol{\eta}) \big| < \frac{\varepsilon}{2}$, or

$$\left| \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \big(f(\xi_j) - f(\eta_j) \big) \big(g(\alpha_j) - g(\alpha_{j-1}) \big) \right| < \frac{\varepsilon}{2},$$

holds for all sets of tags $\boldsymbol{\xi}, \boldsymbol{\eta}$ of the division $\boldsymbol{\alpha}$. Passing to the supremum and infimum on every interval $[\alpha_{j-1}, \alpha_j]$, we get

$$\begin{aligned}0 &\le \overline{S}(\boldsymbol{\alpha}) - \underline{S}(\boldsymbol{\alpha}) \\ &= \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \left(\sup_{x\in[\alpha_{j-1},\alpha_j]} f(x) - \inf_{x\in[\alpha_{j-1},\alpha_j]} f(x) \right) (g(\alpha_j) - g(\alpha_{j-1})) \le \frac{\varepsilon}{2} < \varepsilon.\end{aligned}$$

Consequently,

$$\overline{\int_a^b} f \, \mathrm{d}g \le \overline{S}(\boldsymbol{\alpha}) < \underline{S}(\boldsymbol{\alpha}) + \varepsilon \le \underline{\int_a^b} f \, \mathrm{d}g + \varepsilon$$

and finally also

$$0 \le \overline{\int_a^b} f \, \mathrm{d}g - \underline{\int_a^b} f \, \mathrm{d}g < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we conclude that

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

By the first part of the proof, both the integrals are equal to $(\sigma) \int_a^b f \mathrm{d}g$. □

5.7.3 Remark. If

$$\overline{\int_a^b} f \, \mathrm{d}g = \underline{\int_a^b} f \, \mathrm{d}g \in \mathbb{R},$$

then the common value of both the integrals is called the *Darboux-Stieltjes integral*. Lemma 5.7.2 implies that this integral is in fact equivalent to the (σ)RS-integral.

Our next goal is to prove two main results of this section, Osgood's bounded convergence theorem and Helly's convergence theorem. To this aim, we need the following statement known as Arzelà's lemma. Its proof is pretty long and can be found e.g. in [55], Lemma II.15.8.

5.7.4 Lemma (ARZELÀ'S LEMMA)**.** *For every $k \in \mathbb{N}$, let $\{J_{k,j} : j \in U_k\}$ be a finite collection of intervals in $[a,b]$. Assume there exists a $C > 0$ such that for every $k \in \mathbb{N}$, the length of the union $\bigcup_{j\in U_k} J_{k,j}$ is greater than C. Then there exist infinite sequences $\{k_\ell\}$ and $\{j_\ell\}$ such that $j_\ell \in U_{k_\ell}$ for every $\ell \in \mathbb{N}$ and $\bigcap_{\ell\in\mathbb{N}} J_{k_\ell, j_\ell} \ne \emptyset$.*

5.7.5 Theorem (OSGOOD'S CONVERGENCE THEOREM). *Assume that the function* $f:[a,b]\to\mathbb{R}$ *and the sequence* $\{f_n\}$ *of functions defined on* $[a,b]$ *satisfy*

$$\left.\begin{aligned} &\lim_{n\to\infty} f_n(x)=f(x) \quad \textit{for } x\in[a,b] \\ \textit{and}\quad & \\ &|f_n(x)|\le M<\infty \quad \textit{for } x\in[a,b] \textit{ and } n\in\mathbb{N}. \end{aligned}\right\} \tag{5.7.2}$$

Further, let $g\in \mathrm{BV}([a,b])$ *be such that the integrals* $\int_a^b f\,\mathrm{d}g$ *and* $\int_a^b f_n\,\mathrm{d}g$ *exist for every* $n\in\mathbb{N}$. *Then*

$$\lim_{n\to\infty}\int_a^b f_n\,\mathrm{d}g=\int_a^b f\,\mathrm{d}g. \tag{5.7.3}$$

Proof. By Corollary 5.3.10, the integral $\int_a^b |f_n(x)-f(x)|\,\mathrm{d}(\mathrm{var}_a^x\, g)$ exists for every $n\in\mathbb{N}$ and the inequality

$$\left|\int_a^b f_n(x)\,\mathrm{d}g(x)-\int_a^b f(x)\,\mathrm{d}g(x)\right|\le\int_a^b |f_n(x)-f(x)|\,\mathrm{d}(\mathrm{var}_a^x\, g) \tag{5.7.4}$$

holds. Therefore, it is sufficient to prove that the theorem holds if the functions f_n are non-negative, $f=0$ and g is nondecreasing. Under these assumptions, we need to prove that

$$\lim_{n\to\infty}\int_a^b f_n\,\mathrm{d}g=0. \tag{5.7.5}$$

Without loss of generality, assume that g is nonconstant (otherwise the statement is obvious).

Suppose that (5.7.5) does not hold. Then, due to Lemma 5.7.2, there is an $\varepsilon>0$ and a subsequence $\{f_{n_k}\}$ such that

$$\underline{\int_a^b} f_{n_k}\,\mathrm{d}g=\int_a^b f_{n_k}\,\mathrm{d}g>\varepsilon \quad \text{for all } k\in\mathbb{N}.$$

By Definition 5.7.1, this means that for every $k\in\mathbb{N}$ there is a division $\boldsymbol{\alpha}^k$ of $[a,b]$ such that

$$\underline{S}_k(\boldsymbol{\alpha}^k)>\varepsilon, \quad \text{where } \underline{S}_k(\boldsymbol{\alpha}^k)=\underline{S}(f_{n_k},\mathrm{d}g,\boldsymbol{\alpha}^k).$$

Further, put $m_k=\nu(\boldsymbol{\alpha}^k)$ and

$$\phi_{k,j}=\inf_{x\in[\alpha_{j-1}^k,\alpha_j^k]} f_{n_k}(x) \ \text{ for } k\in\mathbb{N} \ \text{ and } \ j\in\{1,\dots,m_k\}.$$

For a given $\eta > 0$, let U_k stand for the set of indices j such that $\phi_{k,j} > \eta$, while $V_k = \{1, \ldots, m_k\} \setminus U_k$. Obviously,

$$M \sum_{j \in U_k} (g(\alpha_j^k) - g(\alpha_{j-1}^k)) + \eta \sum_{j \in V_k} (g(\alpha_j^k) - g(\alpha_{j-1}^k)) > \varepsilon$$

or

$$M \sum_{j \in U_k} (g(\alpha_j^k) - g(\alpha_{j-1}^k)) > \varepsilon - \eta(g(b) - g(a)).$$

For $\eta = \dfrac{\varepsilon}{2(g(b) - g(a))}$ we get

$$\sum_{j \in U_k} (g(\alpha_j^k) - g(\alpha_{j-1}^k)) > \frac{\varepsilon}{2\,M} > 0.$$

Note that the intervals $J_{k,j} = [g(\alpha_{j-1}^k), g(\alpha_j^k)]$, $j \in U_k$, are nonoverlapping, and thus the length of their union is

$$\sum_{j \in U_k} |J_{k,j}| > \frac{\varepsilon}{2\,M} > 0.$$

Hence, by Arzelà's lemma, there exist a point y_0 and sequences $\{k_\ell\}$ and $\{j_\ell\}$ such that $j_\ell \in U_{k_\ell}$ for every $\ell \in \mathbb{N}$ and $y_0 \in \bigcap_{\ell \in \mathbb{N}} J_{k_\ell, j_\ell}$. This gives

$$y_0 \in [g(\alpha_{j_\ell - 1}^{k_\ell}), g(\alpha_{j_\ell}^{k_\ell})] \quad \text{for every } \ell \in \mathbb{N}.$$

Since g is nondecreasing on $[a, b]$, there exists a point $x_0 \in [a, b]$ such that

$$y_0 \in [g(x_0-), g(x_0+)],\ x_0 \in [\alpha_{j_\ell - 1}^{k_\ell}, \alpha_{j_\ell}^{k_\ell}] \ \text{ and } \ j_\ell \in U_{k_\ell} \text{ for every } \ell \in \mathbb{N}.$$

By the definition of the sets U_k it follows that $f_{n_{k_\ell}}(x_0) > \eta$ for every $\ell \in \mathbb{N}$. However, this contradicts our assumption that $\lim_{n \to \infty} f_n(x) = 0$. Thus, (5.7.5) is true. □

The next assertion is complementary to Osgood's theorem.

5.7.6 Theorem (HELLY'S CONVERGENCE THEOREM). *Let* $f:[a, b] \to \mathbb{R}$ *be continuous and let* $g : [a, b] \to \mathbb{R}$, $\{g_n\} \subset \mathrm{BV}([a, b])$ *and* $\gamma \in [0, \infty)$ *are such that*

$$\operatorname{var}_a^b g_n \le \gamma < \infty \ \textit{for all } n \in \mathbb{N} \quad \textit{and} \quad \lim_{n \to \infty} g_n(x) = g(x) \ \textit{for all } x \in [a, b].$$

Then $\operatorname{var}_a^b g \le \gamma$ *and*

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

Proof. By Theorem 2.7.2 we have $\operatorname{var}_a^b g \le \gamma$; by Theorem 5.6.3, all the integrals $\int_a^b f\,\mathrm{d}g_n$, $n \in \mathbb{N}$, and $\int_a^b f\,\mathrm{d}g$ exist. Let $\varepsilon > 0$ be given. The continuity of f implies that there is a division $\boldsymbol{\alpha}$ of $[a,b]$ such that

$$\left.\begin{aligned} &|f(x)-f(y)| < \frac{\varepsilon}{3\,\gamma} \\ &\qquad\text{for all } x,y \in [\alpha_{j-1},\alpha_j] \text{ and } j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}. \end{aligned}\right\} \tag{5.7.6}$$

Let $\boldsymbol{\xi} = (\alpha_1, \alpha_2, \dots, \alpha_m)$. Then for each $n \in \mathbb{N}$ we have

$$\begin{aligned} &\int_a^b f\,\mathrm{d}g_n - S(f,\mathrm{d}g_n,(\boldsymbol{\alpha},\boldsymbol{\xi})) \\ &\quad = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \left(\int_{\alpha_{j-1}}^{\alpha_j} f(x)\,\mathrm{d}g_n(x) - f(\alpha_j) \int_{\alpha_{j-1}}^{\alpha_j} \mathrm{d}g_n(x) \right) \\ &\quad = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \int_{\alpha_{j-1}}^{\alpha_j} \big(f(x) - f(\alpha_j)\big)\,\mathrm{d}g_n(x). \end{aligned}$$

Using (5.7.6) and Lemma 5.1.11, we get

$$\left| \int_a^b f(x)\,\mathrm{d}g_n(x) - S(f,\mathrm{d}g_n,(\boldsymbol{\alpha},\boldsymbol{\xi})) \right| \le \frac{\varepsilon}{3\,\gamma} \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \operatorname{var}_{\alpha_{j-1}}^{\alpha_j} g_n \le \frac{\varepsilon}{3\,\gamma}\,\gamma = \frac{\varepsilon}{3}.$$

Similarly, we can derive the following inequality with the function g_n replaced by g:

$$\left| \int_a^b f(x)\,\mathrm{d}g(x) - S(f,\mathrm{d}g,(\boldsymbol{\alpha},\boldsymbol{\xi})) \right| < \frac{\varepsilon}{3}.$$

Since $g_n(x) \to g(x)$ for every $x \in [a,b]$, it is clear that

$$\lim_{n\to\infty} \big| S(f,\mathrm{d}g_n,(\boldsymbol{\alpha},\boldsymbol{\xi})) - S(f,\mathrm{d}g,(\boldsymbol{\alpha},\boldsymbol{\xi})) \big| = 0.$$

Thus, there exists an $n_0 \in \mathbb{N}$ such that

$$\big| S(f,\mathrm{d}g_n,(\boldsymbol{\alpha},\boldsymbol{\xi})) - S(f,\mathrm{d}g,(\boldsymbol{\alpha},\boldsymbol{\xi})) \big| < \frac{\varepsilon}{3} \quad \text{for } n \ge n_0.$$

Using the last three inequalities we finally get

$$\begin{aligned} \left| \int_a^b f\,\mathrm{d}g_n - \int_a^b f\,\mathrm{d}g \right| &\le \left| \int_a^b f\,\mathrm{d}g_n - S(f,\mathrm{d}g_n,(\boldsymbol{\alpha},\boldsymbol{\xi})) \right| \\ &\quad + \big| S(f,\mathrm{d}g_n,(\boldsymbol{\alpha},\boldsymbol{\xi})) - S(f,\mathrm{d}g,(\boldsymbol{\alpha},\boldsymbol{\xi})) \big| \\ &\quad + \left| S(f,\mathrm{d}g,(\boldsymbol{\alpha},\boldsymbol{\xi})) - \int_a^b f\,\mathrm{d}g \right| < \varepsilon \end{aligned}$$

for all $n \ge n_0$, which implies the desired equality

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

5.8 Consequences of RS-integrability

In this section we investigate some consequences of the existence of the integral $\int_a^b f\,\mathrm{d}g$. To prove the first result, we need the following two auxiliary lemmas.

5.8.1 Lemma. *If* $a_n \geq 0$ *for all* $n \in \mathbb{N}$ *and* $\sum_{n=1}^{\infty} a_n = \infty$, *then there exists a sequence* $\{c_n\}$ *such that*

$$c_n > 0 \text{ for all } n \in \mathbb{N}, \quad \lim_{n\to\infty} c_n = 0 \text{ and } \sum_{n=1}^{\infty} c_n\, a_n = \infty. \tag{5.8.1}$$

Proof. The sequence $\{s_n\} = \left\{ \sum_{k=1}^{n} a_k \right\}$ is nondecreasing and

$$\lim_{n\to\infty} s_n = \infty. \tag{5.8.2}$$

In particular, for sufficiently large n $(n \geq n_0)$, all s_n are positive. Therefore, we can define

$$c_n = \begin{cases} 1 & \text{if } n < n_0, \\ \frac{1}{s_n} & \text{if } n \geq n_0. \end{cases}$$

Obviously, $c_n > 0$ for every $n \in \mathbb{N}$ and $\lim_{n\to\infty} c_n = 0$. On the other hand, we have

$$\sum_{k=n}^{m} c_k\, a_k = \sum_{k=n}^{m} \frac{a_k}{s_k} \geq \frac{1}{s_m} \sum_{k=n}^{m} a_k = \frac{1}{s_m} \left(\sum_{k=1}^{m} a_k - \sum_{k=1}^{n-1} a_k \right) = 1 - \frac{s_{n-1}}{s_m}$$

for all $m, n \in \mathbb{N}$ such that $m > n \geq n_0$. In view of (5.8.2), for each $n \in \mathbb{N}$ there is an $m_n > n$ such that $\frac{s_{n-1}}{s_{m_n}} < \frac{1}{2}$. Consequently,

$$\sum_{k=n}^{m_n} c_k\, a_k > \frac{1}{2}.$$

This means that (5.8.1) is true. □

5.8.2 Lemma. *Let* $g : [a,b] \to \mathbb{R}$ *be given.*

(i) *If*

$$x_0 \in (a,b] \text{ and } \operatorname{var}_x^{x_0} g = \infty \quad \text{for every } x \in [a, x_0), \tag{5.8.3}$$

then there exists an increasing sequence $\{x_k\}$ *of points in* $[a, x_0)$ *such that*

$$\lim_{k\to\infty} x_k = x_0 \quad \text{and} \quad \sum_{k=1}^{\infty} |g(x_{k+1}) - g(x_k)| = \infty. \tag{5.8.4}$$

(ii) *If*

$$x_0 \in [a,b) \text{ and } \operatorname{var}_{x_0}^x g = \infty \quad \text{for every } x \in (x_0, b], \tag{5.8.5}$$

then there exists a decreasing sequence $\{x_k\}$ *of points in* $(x_0, b]$ *such that* (5.8.4) *is true.*

Proof. i) Assume (5.8.3). First, we will prove that

$$\sup\{\operatorname{var}_y^x g : x \in (y, x_0)\} = \infty \quad \text{for each } y \in [a, x_0). \tag{5.8.6}$$

Assume the opposite. In particular, let there be $M \in [0, \infty)$ and $y \in [a, x_0)$ such that

$$\sup\{\operatorname{var}_y^x g : x \in (y, x_0)\} \le M. \tag{5.8.7}$$

Set $M^* = M + |g(x_0) - g(y)|$. Then, due to (5.8.3), we can choose a division $\{y_0, y_1, \dots, y_m\}$ of the interval $[y, x_0]$ such that

$$\sum_{j=1}^{m} |g(y_j) - g(y_{j-1})| > 3\, M^*.$$

Since

$$\begin{aligned} |g(y_m) - g(y_{m-1})| &= |g(x_0) - g(y_{m-1})| \\ &\le |g(x_0) - g(y)| + |g(y) - g(y_{m-1})| \le M^*, \end{aligned}$$

we have

$$\sum_{j=1}^{m-1} |g(y_j) - g(y_{j-1})| > 2\, M^*,$$

and thus $\operatorname{var}_y^{y_{m-1}} g > 2\, M^*$, which contradicts (5.8.7). Hence, (5.8.6) is true.

Now, let us construct the desired sequence. Set $u_1 = a$ and choose $u_2 \in (a, x_0)$ such that $u_2 > x_0 - 1$ and $\operatorname{var}_{u_1}^{u_2} g > 1$. If $u_1, u_2, \dots, u_\ell \in [a, x_0)$ are already defined and satisfy

$$u_\ell \in (u_{\ell-1}, x_0) \cap (x_0 - \frac{1}{\ell - 1}, x_0) \text{ and } \operatorname{var}_{u_{\ell-1}}^{u_\ell} g > 1,$$

find $u_{\ell+1}$ such that

$$u_{\ell+1} \in (u_\ell, x_0) \cap (x_0 - \frac{1}{\ell}, x_0) \text{ and } \operatorname{var}_{u_\ell}^{u_{\ell+1}} g > 1.$$

The sequence $\{u_\ell\}$ is increasing and

$$\lim_{\ell\to\infty} u_\ell = x_0. \tag{5.8.8}$$

By the definition of variation, for every $\ell \in \mathbb{N}$ there is a division

$$\boldsymbol{\alpha}^\ell = \{\alpha_0^\ell, \alpha_1^\ell, \ldots, \alpha_{m_\ell}^\ell\}$$

of the interval $[u_\ell, u_{\ell+1}]$ such that

$$\sum_{j=1}^{m_\ell} |g(\alpha_j^\ell) - g(\alpha_{j-1}^\ell)| > 1$$

holds. Then

$$\sum_{\ell=1}^{\infty}\left(\sum_{j=1}^{m_\ell} |g(\alpha_j^\ell) - g(\alpha_{j-1}^\ell)|\right) \geq \sum_{\ell=1}^{\infty} 1 = \infty. \tag{5.8.9}$$

Let us reorder the elements of the sets $\boldsymbol{\alpha}^\ell$, $\ell \in \mathbb{N}$, to the sequence $\{x_k\}$ in such a way that

$$x_1 = a = \alpha_0^1,$$
$$x_{k+1} = \alpha_{j+1}^\ell \quad \text{if } x_k = \alpha_j^\ell \text{ and } j < m_\ell - 1,$$

and

$$x_{k+1} = \alpha_0^{\ell+1} \quad \text{if } x_k = \alpha_{m_\ell-1}^\ell.$$

By (5.8.8) and (5.8.9), the sequence $\{x_k\}$ has the required properties. This completes the proof of the assertion (i).

ii) The proof of the second assertion is quite analogous and we can leave the details to readers. □

5.8.3 Theorem. *If the integral* $\int_a^b f\,\mathrm{d}g$ *exists for every continuous function* $f:[a,b]\to\mathbb{R}$, *then* $g:[a,b]\to\mathbb{R}$ *has bounded variation.*

Proof. By Theorem 5.1.7, we can restrict ourselves to the (σ)RS-integral.

Notice that by Heine-Borel Theorem and thanks to the additivity of variation as a function of intervals (cf. Theorem 2.1.14), it follows that a given function $\varphi:[a,b]\to\mathbb{R}$ has bounded variation if and only if the following conditions are satisfied:

$$\left.\begin{array}{l}
\bullet \text{ For each } x \in (a,b], \text{ there is a } \delta_1 \in (0, x-a) \text{ such that} \\
\qquad \operatorname{var}_{x-\delta_1}^{x} \varphi < \infty. \\
\bullet \text{ For each } x \in [a,b), \text{ there is a } \delta_2 \in (0, b-x) \text{ such that} \\
\qquad \operatorname{var}_{x}^{x+\delta_2} \varphi < \infty.
\end{array}\right\} \tag{5.8.10}$$

Consequently, it can happen that $\operatorname{var}_a^b g = \infty$ only if there exists an $x_0 \in [a,b]$ such that at lest one of the conditions (5.8.10) is not satisfied for $x = x_0$. First, assume that $x_0 \in (a,b]$ is such that $\operatorname{var}_x^{x_0} g = \infty$ for each $x \in [a, x_0)$. Then, by the first part of Lemma 5.8.2 there is an increasing sequence $\{x_k\}$ of points in (a, x_0) such that

$$\lim_{k\to\infty} x_k = x_0 \quad \text{and} \quad \sum_{k=1}^{\infty} |g(x_{k+1}) - g(x_k)| = \infty.$$

By Lemma 5.8.1, there is a sequence $\{c_k\}$ of positive numbers such that

$$\lim_{k\to\infty} c_k = 0 \quad \text{and} \quad \sum_{k=1}^{\infty} c_k\,|g(x_{k+1}) - g(x_k)| = \infty.$$

Now, let $\xi_k = \frac{x_{k+1}+x_k}{2}$ for each $k \in \mathbb{N},$ define

$$f(x) = \begin{cases} 0 & \text{if } x < x_1, \text{ or } x \ge x_0, \text{ or } x \in \{x_k\}, \\ c_k \operatorname{sgn}(g(x_{k+1}) - g(x_k)) & \text{if } x = \xi_k, \end{cases}$$

and extend the function f linearly to $[a,b]$. This implies that f will be continuous on $[a,b]$. We have

$$\sum_{k=1}^{\infty} f(\xi_k)\,(g(x_{k+1}) - g(x_k)) = \infty.$$

In particular, for each $M > 0$ there is an $N_M \in \mathbb{N}$ such that

$$\sum_{k=1}^{N_M} f(\xi_k)\,(g(x_{k+1}) - g(x_k)) > M.$$

For a given $M > 0,$ set

$$\boldsymbol{\alpha}_M = \{a, x_1, x_2, \ldots, x_{N_M}, x_{N_M+1}, b\}, \quad \boldsymbol{\xi}_M = (a, \xi_1, \xi_2, \ldots, \xi_{N_M}, b).$$

Then $P_M = (\boldsymbol{\alpha}_M, \boldsymbol{\xi}_M)$ is a partition of $[a,b]$ and

$$S(P_M) = \sum_{k=1}^{N_M} f(\xi_k)\,(g(x_{k+1}) - g(x_k)) > M.$$

(Recall that $f(a) = f(b) = 0.$) But this means that the integral $(\sigma)\int_a^b f\,\mathrm{d}g$ cannot have a finite value.

If the latter condition from (5.8.10) is not satisfied, i.e., there is an $x_0 \in [a,b)$ such that $\operatorname{var}_{x_0}^x g = \infty$ for every $x \in (x_0, b]$, then the proof is similar, just the second part of Lemma 5.8.2 should be used instead of the first one. □

5.8.4 Exercise. Formulate and prove an analogue of Lemma 5.8.2 which is necessary to complete the proof of Theorem 5.8.3 if there is an $x_0 \in [a,b)$ such that $\operatorname{var}_{x_0}^{x} g = \infty$ for every $x \in (x_0, b]$.

5.8.5 Theorem. *Let* $f : [a,b] \to \mathbb{R}$ *be such that the integral* $\int_a^b f \, \mathrm{d}g$ *exists for every step function* g. *Then* f *is continuous.*

Proof. Clearly, we can restrict ourselves to the (σ)RS-integral. Let $x_0 \in (a,b)$, ç, $c' c'' \in \mathbb{R}$, $(c'-c)\,(c''-c)\,(c''-c') \neq 0$ and let the function $g : [a,b] \to \mathbb{R}$ be defined as in Remark 5.1.24, i.e.

$$g(x) = c'\,\chi_{[a,x_0)}(x) + c\,\chi_{[x_0]}(x) + c''\,\chi_{(x_0,b]}(x) \quad \text{for } x \in [a,b].$$

Arguing like in Remark 5.1.24, we can see that the integral $(\sigma)\int_a^b f \, \mathrm{d}g$ can exist only if

$$f(x_0-) = f(x_0) = f(x_0+).$$

Right-continuity of f at a and left-continuity at b can be proved similarly. □

5.9 Mean value theorems

The results presented in this section apply to both kinds of the RS-integral.

5.9.1 Theorem (MEAN VALUE THEOREM)**.** *If* f *is continuous on* $[a,b]$ *and* g *is nondecreasing on* $[a,b]$, *then there exists an* $x_0 \in [a,b]$ *such that*

$$\int_a^b f \, \mathrm{d}g = f(x_0)\,(g(b) - g(a)). \tag{5.9.1}$$

Proof. Theorem 5.6.3 guarantees the existence of the integral $\int_a^b f \, \mathrm{d}g$. Since g is nondecreasing on $[a,b]$, we have

$$m\,(g(b) - g(a)) \le S(P) \le M\,(g(b) - g(a))$$

for every partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$, where

$$m = \min\{f(x) : x \in [a,b]\} \text{ and } M = \max\{f(x) : x \in [a,b]\}.$$

It follows that

$$m(g(b) - g(a)) \le \int_a^b f \, \mathrm{d}g \le M(g(b) - g(a)).$$

Since f is continuous, it takes on all values from the interval $[m, M]$. In particular, there is an $x_0 \in [a,b]$ such that (5.9.1) holds. □

5.9.2 Theorem (SECOND MEAN VALUE THEOREM). *If f is continuous on $[a,b]$ and g is nondecreasing on $[a,b]$, then there is an $x_0 \in [a,b]$ such that*

$$\int_a^b f(x)\, g(x)\, \mathrm{d}x = g(a) \int_a^{x_0} f(x)\, \mathrm{d}x + g(b) \int_{x_0}^b f(x)\, \mathrm{d}x. \tag{5.9.2}$$

Proof. The function f is Riemann integrable on $[a,b]$. Set

$$h(x) = \int_a^x f(t)\, \mathrm{d}t \quad \text{for } x \in [a,b].$$

By virtue of Corollary 5.4.4 (substitution theorem), Theorem 5.5.1 (integration by parts) and Theorem 5.9.1, there is an $x_0 \in [a,b]$ such that

$$\begin{aligned}\int_a^b f(x)\, g(x)\, \mathrm{d}x &= \int_a^b g\, \mathrm{d}h = h(b)\, g(b) - \int_a^b h\, \mathrm{d}g \\ &= \Big(\int_a^b f\, \mathrm{d}x\Big)\, g(b) - \Big(\int_a^{x_0} f\, \mathrm{d}x\Big)\, (g(b) - g(a)) \\ &= g(a) \int_a^{x_0} f(x)\, \mathrm{d}x + g(b) \int_{x_0}^b f(x)\, \mathrm{d}x.\end{aligned}$$

□

5.10 Other integrals of Stieltjes type

Let functions $f, g : [a,b] \to \mathbb{R}$ and a division $\boldsymbol{\alpha}$ of $[a,b]$ be given. Set

$$S_M(\boldsymbol{\alpha}) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \frac{f(\alpha_j) + f(\alpha_{j-1})}{2}\, (g(\alpha_j) - g(\alpha_{j-1})),$$

$$S_{CL}(\boldsymbol{\alpha}) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} f(\alpha_{j-1})\, (g(\alpha_j) - g(\alpha_{j-1})),$$

$$S_{CR}(\boldsymbol{\alpha}) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} f(\alpha_j)\, (g(\alpha_j) - g(\alpha_{j-1})).$$

If $S(\boldsymbol{\alpha}, \boldsymbol{\xi})$ in Definition 5.1.3 is replaced by $S_M(\boldsymbol{\alpha})$, $S_{CL}(\boldsymbol{\alpha})$, or $S_{CR}(\boldsymbol{\alpha})$, we get the definitions of the *main integral*, *left Cauchy integral*, and *right Cauchy integral*, respectively. Again, we distinguish between their (δ) and (σ) variants according to the choice of the limiting process. For each of these integrals, the class of integrable functions includes all RS-integrable functions. However, not all properties of RS-integrals are maintained. For example, an analogue of Theorem 5.4.3 (substitution theorem) does not hold for the central integral. More details can be found in Section II.19 of T. H. Hildebrandt's monograph [55].

5.11 Exercises

(i) In the following examples, investigate the existence and value of the Stieltjes integral $\int_a^b f\,\mathrm{d}g$ of each kind introduced in this chapter:

(a) $[a,b]=[0,\pi],\ f(x)=x$ and $g(x)=\sin\,x$ for $x\in[a,b]$,

(b) $[a,b]=[-1,1],\ f(x)=x$ and $g(x)=\exp(|x|)$ for $x\in[a,b]$,

(c) $[a,b]=[0,1],\ f:[a,b]\to\mathbb{R}$ and $g(x)=\begin{cases}0 & \text{if } x\in[0,\frac{1}{2}),\\ c & \text{if } x=\frac{1}{2},\\ d & \text{if } x\in(\frac{1}{2},1].\end{cases}$

(ii) Let

$$f(x)=\begin{cases}0 & \text{if } x<0,\\ 1 & \text{if } x\ge 0,\end{cases}\qquad g(x)=\begin{cases}0 & \text{if } x\le 0,\\ 1 & \text{if } x>0.\end{cases}$$

Investigate the existence and the value of the following integrals (consider both (δ) and (σ) RS-integrals):

$$\int_{-1}^{1} g\,\mathrm{d}f,\ \int_{-1}^{0} g\,\mathrm{d}f,\ \int_{0}^{1} g\,\mathrm{d}f,\ \int_{-1}^{1} g\,\mathrm{d}g,\ \int_{-1}^{0} g\,\mathrm{d}g,\ \int_{0}^{1} g\,\mathrm{d}g.$$

(iii) Determine the value of the integral $(\delta)\int_0^1 x^2\,\mathrm{d}g(x)$, where

$$g(x)=\begin{cases}x & \text{if } x\in[0,\frac{1}{2}],\\ \frac{1}{x} & \text{if } x\in(\frac{1}{2},1].\end{cases}$$

(iv) Provide an exact definition of the line integral of the first kind mentioned in Section 1.2, and formulate some of its basic properties that follow from the results obtained in the present chapter.

Chapter 6

Kurzweil-Stieltjes integral

Riemann-Stieltjes integral is widely used everywhere where it is possible to limit ourselves to the cases when the integrand and the integrator have no common points of discontinuity (or, in the case of (σ) RS-integral, there are no points at which both functions have discontinuity on the same side). For some applications (e.g. in the theory of hysteresis and related variational inequalities, see [17], [76] and [77]), the Stieltjes integral which has no requirements on the continuity of the integrated and integrating functions is needed. It appears that the most suitable integral from this point of view is the integral which we will call the Kurzweil-Stieltjes integral. However, its generality is not its only asset. In particular, let us mention also the simplicity of its definition and a relatively easy way of the proofs. Unfortunately, monographic literature has not devoted sufficient attention to this concept. As far as we know, a brief treatise of this integral can be found in chapter 24 of Schechter's monograph [117] from 1997 (however, it is called the Henstock-Stieltjes integral there). Furthermore, McLeod's monograph [95] from 1980, where it is called *gauge integral*, and several extensive sections (2.6 and 2.7 and partially also 2.8) of the monograph [29] by Dudley and Norvaiša deal with this integral (called there the Henstock-Kurzweil integral) in more detail. Let us notice that this integral is a special case of the generalized nonlinear integral which has been introduced in Kurzweil's seminal work [78] from 1957 as a tool for explaining some of the convergence effects occurring in the theory of non-linear differential equations. A year later, in [79], Kurzweil explicitly used this special Stieltjes form of its integral to consider generalized differential equations covering e.g. the equations whose right hand sides contain terms with Dirac distributions. During the 70's of the last century, the term Kurzweil-Stieltjes integral (or Perron-Stieltjes integral by Kurzweil's definition) was already commonly used in the works dealing with the generalized nonlinear differential equations (see e.g. [119] or [131] and the papers cited there).

The aim of this chapter is to present the theory of the Kurzweil-Stieltjes integral as comprehensively as possible.

6.1 Definition and basic properties

Let us recall that the finite ordered subset $\{\alpha_0, \alpha_1, \dots, \alpha_m\}$ of an interval $[a,b]$ is called a *division* of the interval $[a,b]$ if

$$a = \alpha_0 < \alpha_1 < \dots < \alpha_m = b.$$

The set of all divisions of the interval $[a,b]$ is denoted by $\mathscr{D}[a,b]$. The number of subintervals forming a division is usually denoted by $\nu(\boldsymbol{\alpha})$, i.e. $\alpha_{\nu(\boldsymbol{\alpha})}=b$. Finally, we also set

$$|\boldsymbol{\alpha}| = \max_{j\in\{1,\dots,\nu(\boldsymbol{\alpha})\}} (\alpha_j - \alpha_{j-1}) \quad \text{for } \boldsymbol{\alpha}\in\mathscr{D}[a,b].$$

We say that $\boldsymbol{\alpha}'\in\mathscr{D}[a,b]$ is a *refinement of* $\boldsymbol{\alpha}\in\mathscr{D}[a,b]$, if $\boldsymbol{\alpha}'\supset\boldsymbol{\alpha}$.

The pair $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of finite ordered sets

$$\boldsymbol{\alpha}=\{\alpha_0,\alpha_1,\dots,\alpha_{\nu(\boldsymbol{\alpha})}\} \quad \text{and} \quad \boldsymbol{\xi}=\{\xi_1,\dots,\xi_{\nu(\boldsymbol{\alpha})}\}$$

is called a *partition* of the interval $[a,b]$ if $\boldsymbol{\alpha}$ is a division of the interval $[a,b]$ and

$$\alpha_{j-1}\le\xi_j\le\alpha_j \quad \text{for } j\in\{1,\dots,\nu(\boldsymbol{\alpha})\}.$$

We say that ξ_j is the *tag* of the subinterval $[\alpha_{j-1},\alpha_j]$ and $\boldsymbol{\xi}$ is the *set of the tags*. For partitions $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ we will also write $\nu(P)$ instead of $\nu(\boldsymbol{\alpha})$. (We can also say that $\nu(P)$ is the number of tags contained in partition P.)

6.1.1 Definition. Every positive function $\delta:[a,b]\to(0,\infty)$ is called a *gauge* on the interval $[a,b]$. A set of gauges on $[a,b]$ is denoted by $\mathscr{G}[a,b]$.

If δ is a gauge on $[a,b]$, we say that the partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of the interval $[a,b]$ is δ*-fine* if

$$[\alpha_{j-1},\alpha_j]\subset(\xi_j-\delta(\xi_j),\xi_j+\delta(\xi_j)) \quad \text{for all } j=1,\dots,\nu(\boldsymbol{\alpha}). \tag{6.1.1}$$

Consider functions $f,g:[a,b]\to\mathbb{R}$ and a partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$. Analogously to RS-integrals, we define

$$S(f,\mathrm{d}g,P;[a,b])\Big(=S(f,\mathrm{d}g,(\boldsymbol{\alpha},\boldsymbol{\xi});[a,b])\Big)=\sum_{j=1}^{\nu(P)} f(\xi_j)\,[g(\alpha_j)-g(\alpha_{j-1})]$$
$$\Big(=\sum_{j=1}^{\nu(\boldsymbol{\alpha})} f(\xi_j)\,[g(\alpha_j)-g(\alpha_{j-1})]\Big).$$

According to the concrete situation, we will usually write simply e.g. $S(P)$ or $S(\boldsymbol{\alpha},\boldsymbol{\xi})$ or $S(f,\mathrm{d}g,P)$ or $S(f,\mathrm{d}g,(\boldsymbol{\alpha},\boldsymbol{\xi}))$ instead of $S(f,\mathrm{d}g,P;[a,b])$ or $S(f,\mathrm{d}g,(\boldsymbol{\alpha},\boldsymbol{\xi});[a,b])$.

6.1.2 Definition. Let $f,g:[a,b]\to\mathbb{R}$ and $I\in\mathbb{R}$. We say that the *Kurzweil-Stieltjes integral* (KS-integral)

$$\int_a^b f(x)\,\mathrm{d}[g(x)]$$

exists and has the value of $I \in \mathbb{R}$ if

$$\left.\begin{array}{l} \text{for every } \varepsilon > 0 \text{ there exists a gauge } \delta_\varepsilon \text{ on } [a,b] \text{ such that} \\ \qquad |I - S(P)| < \varepsilon \\ \text{holds for all } \delta_\varepsilon-\text{fine partitions of } [a,b]. \end{array}\right\} \tag{6.1.2}$$

As for RS-integrals, we will also use the shortened notation

$$\int_a^b f\,\mathrm{d}g = \int_a^b f(x)\,\mathrm{d}[g(x)].$$

and define

$$\int_b^a f\,\mathrm{d}g = -\int_a^b f\,\mathrm{d}g \text{ and } \int_a^a f\,\mathrm{d}g = 0.$$

If $g(x) \equiv x$, then we speak about KH-integral (Kurzweil-Henstock integral) instead of KS-integral and we denote it $\int_a^b f(x)\,\mathrm{d}x$, or in a shorter way $\int_a^b f\,\mathrm{d}x$.

This definition is correct due to the following two lemmas.

6.1.3 Lemma (COUSIN). *The set of all δ-fine partitions of the interval $[a,b]$ is nonempty for every gauge δ on $[a,b]$.*

Proof. Consider a gauge δ on $[a,b]$. For a given $c \in (a,b]$, let us denote by $\mathscr{A}(\delta;[a,c])$ the set of all δ-fine partitions of the interval $[a,c]$ and let M be the set of all $c \in (a,b]$ for which $\mathscr{A}(\delta;[a,c])$ is nonempty.

Let
$$c = \min\{a+\delta(a), b\}, \ \boldsymbol{\alpha} = \{a,c\} \text{ and } \boldsymbol{\xi} = \{a\}.$$

Since $\delta(a) > 0$, we have $c \in (a,b]$ and $(\boldsymbol{\alpha},\boldsymbol{\xi}) \in \mathscr{A}(\delta;[a,c])$, i.e. $c \in M$. The set M is thus nonempty and therefore $d = \sup M > -\infty$.

Next, we will show that d does belong to the set M. As $\delta(d) > 0$, and by the definition of d, there is a $c \in (d-\delta(d), d] \cap M$. Hence, there exists also a δ-fine partition $(\boldsymbol{\alpha}',\boldsymbol{\xi}')$ of $[a,c]$. Let $c < d$. (In the opposite case, trivially $d = c \in M$.) Set $\boldsymbol{\alpha} = \boldsymbol{\alpha}' \cup \{d\}$ and $\boldsymbol{\xi} = \boldsymbol{\xi}' \cup \{d\}$. Then $(\boldsymbol{\alpha},\boldsymbol{\xi})$ is a partition of the interval $[a,d]$, and as $[c,d] \subset (d-\delta(d), d+\delta(d))$, it is δ-fine. Therefore $d \in M$.

If $d = b$, the proof is finished. Thus, assume $d < b$ and choose an arbitrary tagged division $(\boldsymbol{\alpha}'',\boldsymbol{\xi}'')$ of $[a,d]$ and $\gamma \in (d, d+\delta(d)) \cap (d,b)$. (Such a γ exists because $\delta(d) > 0$.) Thus, we have

$$[d,\gamma] \subset (d-\delta(d), d+\delta(d)).$$

Therefore $\big(\boldsymbol{\alpha}'' \cup \{\gamma\}, \boldsymbol{\xi}'' \cup \{d\}\big)$ is a δ-fine partition of the interval $[a,\gamma]$, i.e. $\gamma \in M$. Since $\gamma > d$, we get a contradiction with the definition $d = \sup M$. We thus have $d = \sup M = b$ and the proof is completed. □

6.1.4 Lemma. *The value of the integral* $\int_a^b f\,\mathrm{d}g$ *is defined uniquely by condition* (6.1.2).

Proof. Assume there exist $I_1,\,I_2\in\mathbb{R},\ I_1\neq I_2,$ such that (6.1.2) holds and if we substitute $I=I_i,\ i=1,2.$ Set $\widetilde{\varepsilon}=\frac{1}{2}\,|I_1-I_2|.$ Then there exist gauges δ_1 and δ_2 such that

$$|S(P)-I_1|<\widetilde{\varepsilon}\quad\text{for every } \delta_1-\text{fine partition } P \text{ of } [a,b],$$

and

$$|S(P)-I_2|<\widetilde{\varepsilon}\quad\text{for every } \delta_2-\text{fine partition } P \text{ of } [a,b],$$

Define a gauge δ by

$$\delta(x)=\min\{\delta_1(x),\delta_2(x)\}\quad\text{for } x\in[a,b].$$

Then every δ-fine partition of $[a,b]$ is simultaneously both δ_1-fine and δ_2-fine. Consequently, for every δ-fine partition P of $[a,b]$ we have

$$\begin{aligned}2\,\widetilde{\varepsilon}&=|I_1-I_2|=|I_1-S(P)+S(P)-I_2|\\&\leq|I_1-S(P)|+|S(P)-I_2|<2\,\widetilde{\varepsilon}.\end{aligned}$$

This being impossible, it has to be $I_1=I_2$. □

If not stated otherwise, in the following text the symbol of the integral will always stand for the KS-integral.

6.1.5 Remark. If gauges δ,δ_0 are such that $\delta_0\leq\delta$ on $[a,b]$, then every δ_0-fine partition of $[a,b]$ is also δ-fine. Therefore, if some condition is satisfied for all δ_0-fine partitions of $[a,b]$, a fortiori it is satisfied also for all δ-fine partitions. Hence, if the gauge δ_0 is given, we can limit ourselves in Definition 6.1.2 to the gauges δ_ε, for which $\delta_\varepsilon\leq\delta_0$ on $[a,b]$.

Moreover, Definition 6.1.2 will not change if (6.1.1) is replaced by the condition

$$[\alpha_{j-1},\alpha_j\,]\subset[\xi_j-\delta(\xi_j),\xi_j+\delta(\xi_j)]\quad\text{for } j=1,\dots,\nu(\boldsymbol{\alpha}).$$

The next result provides a Bolzano-Cauchy condition for the existence of the KS-integral.

6.1.6 Theorem (Bolzano–Cauchy condition)**.** *Let* $f,g:[a,b]\to\mathbb{R}.$ *Then the integral* $\int_a^b f\,\mathrm{d}g$ *exists if and only if*

$$\left.\begin{array}{l}\textit{for each } \varepsilon>0 \textit{ there is a gauge } \delta_\varepsilon \textit{ on } [a,b] \textit{ such that}\\ \big|S(P)-S(Q)\big|<\varepsilon \textit{ holds for all } \delta_\varepsilon\textit{-fine partitions } P,\,Q.\end{array}\right\}\quad(6.1.3)$$

Proof. a) If the integral $\int_a^b f\,\mathrm{d}g = I \in \mathbb{R}$ exists, then by Definition 6.1.2 for every $\varepsilon > 0$ there is a gauge δ_ε on $[a,b]$ such that

$$|S(P) - I| < \frac{\varepsilon}{2} \quad \text{for all } \delta_\varepsilon - \text{fine partitions } of [a,b].$$

Hence, for every pair P, Q of δ_ε-fine partitions we have

$$|S(P) - S(Q)| \le |S(P) - I| + |S(Q) - I| < \varepsilon,$$

which means that (8.3.6) is true.

b) Now, assume that the Bolzano–Cauchy condition (8.3.6) holds and for a given $\varepsilon > 0$, let δ_ε stand for the gauge given by (8.3.6). Set

$$I(\varepsilon) = \{S(P) : P \text{ is a } \delta_\varepsilon\text{-fine partition of } [a,b]\}.$$

By Cousin's Lemma 6.1.3, the $I(\varepsilon)$ is nonempty for every $\varepsilon > 0$ and by (8.3.6) we have

$$\operatorname{diam} I(\varepsilon) = \sup\{|S(P) - S(Q)| : P, Q \text{ are } \delta_\varepsilon\text{-fine partitions of } [a,b]\} < \varepsilon. \quad (6.1.4)$$

Furthermore,

$$I(\varepsilon_1) \subset I(\varepsilon_2) \text{ whenever } \varepsilon_1 < \varepsilon_2.$$

Thus, by Cantor's intersection theorem for complete metric spaces (see e.g. Theorem 6.52 in [51] or Theorem 5.1.17 in [143]), the intersection $\bigcap_{\varepsilon>0} \mathrm{cl}\, I(\varepsilon)$ is a one point set $\{I\}$ with $I \in \mathbb{R}$. As a result of (6.1.4), it follows that

$$|S(f, \mathrm{d}g, P) - I| \le \varepsilon$$

holds for every δ-fine partition P of $[a,b]$. In other words, $\int_a^b f\,\mathrm{d}g = I$ and this completes the proof. □

6.1.7 Remark. As in the case of RS-integrals (see Exercise 5.1.16), if a division $\boldsymbol{\alpha}$ of $[a,b]$ is given, we can weaken the condition (8.3.6) in the following way:

$$\left.\begin{array}{l} \text{for all } \varepsilon > 0 \text{ there is a gauge } \delta_\varepsilon \text{ on } [a,b] \text{ such that} \\ \qquad \Big|S(P) - S(Q)\Big| < \varepsilon \\ \text{for all } \delta_\varepsilon\text{-fine partitions } P = (\boldsymbol{\alpha}, \boldsymbol{\xi}),\ Q = (\boldsymbol{\beta}, \boldsymbol{\eta}) \text{ of } [a,b] \\ \text{such that } \boldsymbol{\beta} \supset \boldsymbol{\alpha}. \end{array}\right\} \quad (6.1.5)$$

The KS-integral has the usual linear properties.

6.1.8 Theorem. *Let* $f,\ f_1,\ f_2,\ g,\ g_1,\ g_2 : [a,b] \to \mathbb{R}$ *and let the integrals*

$$\int_a^b f_1\,\mathrm{d}g, \quad \int_a^b f_2\,\mathrm{d}g, \quad \int_a^b f\,\mathrm{d}g_1 \ \text{ and } \int_a^b f\,\mathrm{d}g_2$$

exist. Then for any $c_1, c_2 \in \mathbb{R}$,

$$\int_a^b (c_1\,f_1 + c_2\,f_2)\,\mathrm{d}g = c_1 \int_a^b f_1\,\mathrm{d}g + c_2 \int_a^b f_2\,\mathrm{d}g$$

and

$$\int_a^b f\,\mathrm{d}[\,c_1\,g_1 + c_2\,g_2] = c_1 \int_a^b f\,\mathrm{d}g_1 + c_2 \int_a^b f\,\mathrm{d}g_2$$

hold.

Proof. Let us show for example the proof of the first statement.

Let $\varepsilon > 0$ be given. By our assumption there are gauges δ_1 and δ_2 on $[a,b]$ such that

$$\left| S(f_1, \mathrm{d}g, P) - \int_a^b f_1\,\mathrm{d}g \right| < \varepsilon \quad \text{for all } \delta_1\text{-fine partitions } P \text{ on } [a,b]$$

and

$$\left| S(f_2, \mathrm{d}g, P) - \int_a^b f_2\,\mathrm{d}g \right| < \varepsilon \quad \text{for all } \delta_2\text{-fine partitions } P \text{ on } [a,b]$$

Set $\delta_\varepsilon(x) = \min\{\delta_1(x), \delta_2(x)\}$ for $x \in [a,b]$ and $h = c_1\,f_1 + c_2\,f_2$. Since for a given partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ on ab we have

$$\begin{aligned} S(h, \mathrm{d}g, P) &= \sum_{j=1}^{\nu(P)} (c_1\,f_1(\xi_j) + c_2\,f_2(\xi_j))\,[g(\alpha_j) - g(\alpha_{j-1})] \\ &= c_1\,S(f_1, \mathrm{d}g, P) + c_2\,S(f_2, \mathrm{d}g, P), \end{aligned}$$

we get

$$\begin{aligned} &\left| S(h, \mathrm{d}g, P) - c_1 \int_a^b f_1\,\mathrm{d}g - c_2 \int_a^b f_2\,\mathrm{d}g \right| \\ &\qquad \le |c_1| \left| S(f_1, \mathrm{d}g, P) - \int_a^b f_1\,\mathrm{d}g \right| + |c_2| \left| S(f_2, \mathrm{d}g, P) - \int_a^b f_2\,\mathrm{d}g \right| \\ &\qquad < (|c_1| + |c_2|)\,\varepsilon, \end{aligned}$$

wherefrom our statement immediately follows.

The second statement of the theorem would be proved similarly and can be left as an exercise for the reader. □

6.1.9 Exercise. Prove the second statement of Theorem 6.1.8.

6.1.10 Theorem. *If the integral* $\int_a^b f\,\mathrm{d}g$ *exists and if* $[c,d] \subset [a,b]$, *then the integral* $\int_c^d f\,\mathrm{d}g$ *exists, too.*

Proof. Assume that the integral $\int_a^b f\,\mathrm{d}g$ exists and $a<c<d<b$. By Theorem 6.1.6 there exists a gauge δ_ε on $[a,b]$ such that

$$|S(P)-S(P')|<\varepsilon \quad \text{for all } \delta_\varepsilon\text{-fine partitions } P, P' \text{ of } [a,b]. \tag{6.1.6}$$

Let $Q=(\boldsymbol{\beta},\boldsymbol{\eta})$ and $Q'=(\boldsymbol{\beta}',\boldsymbol{\eta}')$ be arbitrary δ_ε-fine partitions of $[c,d]$. Further, let us fix arbitrarily a δ_ε-fine partition $Q^-=(\boldsymbol{\beta}^-,\boldsymbol{\eta}^-)$ of $[a,c]$ and a δ_ε-fine partition
$Q^+=(\boldsymbol{\beta}^+,\boldsymbol{\eta}^+)$ of $[c,d]$ and set $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ and $P'=(\boldsymbol{\alpha}',\boldsymbol{\xi}')$, where

$$\boldsymbol{\alpha}=\boldsymbol{\beta}^-\cup\boldsymbol{\beta}\cup\boldsymbol{\beta}^+,\ \boldsymbol{\xi}=\boldsymbol{\eta}^-\cup\boldsymbol{\eta}\cup\boldsymbol{\eta}^+$$

and

$$\boldsymbol{\alpha}'=\boldsymbol{\beta}^-\cup\boldsymbol{\beta}'\cup\boldsymbol{\beta}^+,\ \boldsymbol{\xi}'=\boldsymbol{\eta}^-\cup\boldsymbol{\eta}'\cup\boldsymbol{\eta}^+.$$

Then P and P' are δ_ε-fine partitions of $[a,b]$ and

$$S(P)=S(Q^-)+S(Q)+S(Q^+) \text{ and } S(P')=S(Q^-)+S(Q')+S(Q^+).$$

Thus, in view of (6.1.6) we have

$$|S(Q)-S(Q')|=|S(P)-S(P')|<\varepsilon,$$

wherefrom by Theorem 6.1.6 the existence of the integral $\int_c^d f\,\mathrm{d}g$ follows. □

6.1.11 Theorem. *Let* $f,g:[a,b]\to\mathbb{R}$ *and* $c\in[a,b]$. *Then the integral* $\int_a^b f\,\mathrm{d}g$ *exists if and only if both the integrals* $\int_a^c f\,\mathrm{d}g$ *and* $\int_c^b f\,\mathrm{d}g$ *exist. In such case, the equality*

$$\int_a^b f\,\mathrm{d}g=\int_a^c f\,\mathrm{d}g+\int_c^b f\,\mathrm{d}g$$

holds.

Proof. If $c=a$ or $c=b$, the statement of the theorem is trivial. Thus, let $c\in(a,b)$.

a) If the integral $\int_a^b f\,\mathrm{d}g$ exists, then by Theorem 6.1.10 both the integrals $\int_a^c f\,\mathrm{d}g$ and $\int_c^b f\,\mathrm{d}g$ exist, too.

b) Let

$$\int_a^c f\,\mathrm{d}g = I_1 \text{ and } \int_c^b f\,\mathrm{d}g = I_2.$$

Let $\varepsilon > 0$ be given. Choose gauges δ'_ε on $[a,c]$ and δ''_ε on $[c,b]$ such that

$$\left.\begin{aligned} &|S(P') - I_1| < \frac{\varepsilon}{2} \quad \text{for all } \delta'_\varepsilon\text{-fine partitions } P' \text{ on [a,c]} \\ &\text{and} \\ &|S(P'') - I_2| < \frac{\varepsilon}{2} \quad \text{for all } \delta''_\varepsilon\text{-fine partitions } P'' \text{ on [c,b].} \end{aligned}\right\} \tag{6.1.7}$$

Now, we define the gauge δ_ε on $[a,b]$ by

$$\delta_\varepsilon(x) = \begin{cases} \min\left\{\delta'_\varepsilon(x), \frac{1}{4}(c-x)\right\} & \text{if } x \in [a,c), \\ \min\{\delta'_\varepsilon(c), \delta''_\varepsilon(c)\} & \text{if } x = c, \\ \min\left\{\delta''_\varepsilon(x), \frac{1}{4}(x-c)\right\} & \text{if } x \in (c,b]. \end{cases}$$

Then

$$x + \delta_\varepsilon(x) \le x + \frac{1}{4}(c-x) < c \quad \text{if } x < c,$$

and

$$x - \delta_\varepsilon(x) \ge x - \frac{1}{4}(x-c) > c \quad \text{if } x > c.$$

Therefore $c \in [x-\delta_\varepsilon(x), x+\delta_\varepsilon(x)]$ for no $x \ne c$. Hence, for every δ_ε-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$, there exists a $k \in \{1, \ldots, \nu(\boldsymbol{\alpha})\}$ such that $\xi_k = c$. Thus, we can assume that

$$\alpha_{k-1} < \alpha_k = \xi_k = c = \xi_{k+1} < \alpha_{k+1}.$$

If

$$\alpha_{k-1} < c = \xi_k < \alpha_k,$$

we would adjust the corresponding term in the sum $S(P)$ as follows:

$$f(c)\,[g(\alpha_k) - g(\alpha_{k-1})] = f(c)\,[g(\alpha_k) - g(c)] + f(c)\,[g(c) - g(\alpha_{k-1})].$$

Thus, there are δ_ε-fine partitions $P' = (\boldsymbol{\alpha}', \boldsymbol{\xi}')$ of $[a,c]$ and $P' = (\boldsymbol{\alpha}'', \boldsymbol{\xi}'')$ of $[c,b]$ such that

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}' \cup \boldsymbol{\alpha}'', \; \boldsymbol{\xi} = \boldsymbol{\xi}' \cup \boldsymbol{\xi}'' \text{ and } S(P) = S(P') + S(P'').$$

If we take into consideration the relation (6.1.7), we can see that

$$\begin{aligned}|S(P)-(I_1+I_2)| &= |S(P')+S(P'')-(I_1+I_2)|\\ &\le |S(P')-I_1|+|S(P'')-I_2|<\varepsilon\end{aligned}$$

holds for every δ_ε-fine partition P of $[a,b]$, i.e. $\int_a^b f\,\mathrm{d}g = I_1+I_2$. □

The next lemma provides a crucial characterization of the KS-integration. Its proof is based on a slight modification of the choice of a suitable gauge used in the proof of the previous statement.

6.1.12 Lemma. *For any finite set D of points of the interval $[a,b]$, there exists a gauge δ on $[a,b]$ such that $D\subset\boldsymbol{\xi}$ for every δ-fine partition $(\boldsymbol{\alpha},\boldsymbol{\xi})$ of the interval $[a,b]$.*

Proof. Let $D=\{s_1,\dots,s_k\}$ and $a\le s_1<\dots<s_k\le b$. Set

$$\delta(x)=\begin{cases}\frac14\min\{|x-s_j|:j\in\{1,\dots,k\}\} & \text{if } x\notin D,\\ 1, & \text{if } x\in D.\end{cases}$$

For given $j\in\{1,\dots,k\}$ we have

$$\xi+\delta(\xi)<\xi+\frac14(s_j-\xi)<s_j \qquad \text{if } \xi\in(s_{j-1},s_j)$$

and

$$\xi-\delta(\xi)>\xi-\frac14(s_j-\xi)>s_j \qquad \text{if } \xi\in(s_j,s_{j+1}).$$

That is,

$$s_j\in[\xi-\delta(\xi),\xi+\delta(\xi)] \quad \text{if and only if } \xi=s_j.$$

Hence, $s_j\in\boldsymbol{\xi}$ for each $j\in\{1,\dots,k\}$ and each δ-fine partition $(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$. In other words, $D\subset\boldsymbol{\xi}$ for each δ_ε-fine partition $(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$. □

6.2 Relationship to Riemann-Stieltjes and Perron-Stieltjes integrals

If the integral $(\delta)\int_a^b f\,\mathrm{d}g$ exists, then the KS-integral $\int_a^b f\,\mathrm{d}g$ exists, too, and has the same value. Indeed, if

$$(\delta)\int_a^b f\,\mathrm{d}g = I\in\mathbb{R},$$

then for every $\varepsilon > 0$ there is $\Delta_\varepsilon > 0$ such that

$$|S(\boldsymbol{\alpha},\boldsymbol{\xi}) - I| < \varepsilon \quad \text{for all partitions } (\boldsymbol{\alpha},\boldsymbol{\xi}) \text{ of } [a,b] \text{ such that } |\boldsymbol{\alpha}| < \Delta_\varepsilon.$$

Then $\delta_\varepsilon(x) \equiv \Delta_\varepsilon/2$ is a gauge with the properties guaranteeing the relation

$$\int_a^b f \,\mathrm{d}g = I = (\delta)\int_a^b f \,\mathrm{d}g.$$

On the other hand, if $\int_a^b f \,\mathrm{d}g = I \in \mathbb{R}$ and if for every $\varepsilon > 0$ we can find a gauge δ_ε on $[a,b]$ such that

$$\inf\{\delta_\varepsilon(x) : x \in [a,b]\} > 0 \ \text{ and } \ |S(\boldsymbol{\alpha},\boldsymbol{\xi}) - I| < \varepsilon \text{ for each } \delta\text{-fine partition } (\boldsymbol{\alpha},\boldsymbol{\xi}),$$

then

$$(\delta)\int_a^b f \,\mathrm{d}g = I$$

will be true, as well. Indeed, setting $\Delta_\varepsilon = \inf\{\delta_\varepsilon(x) : x \in [a,b]\}$, we get that

$$|S(\boldsymbol{\alpha},\boldsymbol{\xi}) - I| < \varepsilon \quad \text{for every partition } (\boldsymbol{\alpha},\boldsymbol{\xi}) \text{ of } [a,b] \text{ such that } |\boldsymbol{\alpha}| < \Delta_\varepsilon$$

holds.

The relationship between the (σ) RS-integral and the KS-integral is not that evident.

6.2.1 Theorem. *If the integral* $(\sigma)\displaystyle\int_a^b f \,\mathrm{d}g$ *exists, then the* KS-*integral* $\int_a^b f \,\mathrm{d}g$ *exists as well and*

$$\int_a^b f \,\mathrm{d}g = (\sigma)\int_a^b f \,\mathrm{d}g.$$

Proof. Assume that $(\sigma)\displaystyle\int_a^b f \,\mathrm{d}g = I \in \mathbb{R}$. Let $\varepsilon > 0$ be given and let $\boldsymbol{\alpha}_\varepsilon$ be a division of the interval $[a,b]$ such that

$$\big|S(\boldsymbol{\alpha},\boldsymbol{\xi}) - I\big| < \varepsilon \quad \text{for every partition } (\boldsymbol{\alpha},\boldsymbol{\xi}) \text{ of } [a,b] \text{ such that } \boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon.$$

By Lemma 6.1.12 there exists a gauge δ_ε such that

$$\boldsymbol{\alpha}_\varepsilon \subset \boldsymbol{\xi} \quad \text{for each } \delta_\varepsilon - \text{fine partition } (\boldsymbol{\alpha},\boldsymbol{\xi}) \text{ of } [a,b]. \tag{6.2.1}$$

Now, let $(\boldsymbol{\alpha},\boldsymbol{\xi})$ be an arbitrary δ_ε-fine partition of $[a,b]$. Then

$$\left.\begin{aligned} S(\boldsymbol{\alpha},\boldsymbol{\xi}) &= \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \Big[f(\xi_j)\,[g(\alpha_j) - g(\xi_j)] + f(\xi_j)\,[g(\xi_j) - g(\alpha_{j-1})]\Big] \\ &= S(\widetilde{\boldsymbol{\alpha}},\widetilde{\boldsymbol{\xi}}), \end{aligned}\right\} \tag{6.2.2}$$

where

$$\widetilde{\boldsymbol{\alpha}} = \{\alpha_0, \xi_1, \alpha_1, \xi_2, \dots, \xi_{\nu(\boldsymbol{\alpha})}, \alpha_{\nu(\boldsymbol{\alpha})}\} \text{ and } \widetilde{\boldsymbol{\xi}} = \{\xi_1, \xi_1, \xi_2, \xi_2, \dots, \xi_{\nu(\boldsymbol{\alpha})}, \xi_{\nu(\boldsymbol{\alpha})}\}.$$

(Of course, if $\alpha_{k-1} = \xi_k$ or $\xi_k = \alpha_k$ for some k, then we would remove the intervals $[\alpha_{k-1}, \xi_k]$ and $[\xi_k, \alpha_k]$ and the corresponding tags from $(\widetilde{\boldsymbol{\alpha}}, \widetilde{\boldsymbol{\xi}})$.) By (6.2.1) we have $\boldsymbol{\alpha}_\varepsilon \subset \boldsymbol{\xi} \subset \widetilde{\boldsymbol{\alpha}}$. Finally, in view of (6.2.2) and due to the definition of $\boldsymbol{\alpha}_\varepsilon$, we get

$$|S(\boldsymbol{\alpha}, \boldsymbol{\xi}) - I| = |S(\widetilde{\boldsymbol{\alpha}}, \widetilde{\boldsymbol{\xi}}) - I| < \varepsilon,$$

which means that $\int_a^b f \, \mathrm{d}g = I$. This completes the proof. □

Notice that the proof of the previous theorem also contains the proof of the following statement.

6.2.2 Lemma. *Let $f, g : [a,b] \to \mathbb{R}$. Then for every partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$, there exists a partition $\widetilde{P} = (\widetilde{\boldsymbol{\alpha}}, \widetilde{\boldsymbol{\xi}})$ of $[a,b]$ such that*

$$\boldsymbol{\xi} \subset \widetilde{\boldsymbol{\alpha}} \cap \widetilde{\boldsymbol{\xi}} \quad \textit{and} \quad S(\widetilde{P}) = S(P).$$

6.2.3 Examples. Let us consider for a while the special case when the integrator is the identity function, i.e. when the KS-integral reduces to the KH-integral.

(i) The KH-integral is obviously a generalization of the classical Riemann integral.

(ii) Let $f(x) = 0$ for $x \in [a,b] \setminus D$ where D is a subset of $[a,b]$ of zero measure. Let an arbitrary $\varepsilon > 0$ be given and let M be the set of those $t \in [a,b]$ for which $f(x) \neq 0$. By assumption, $\mu(M) = 0$ holds for the Lebesgue measure $\mu(M)$ of M. Set

$$M_n = \{x \in [a,b] : n - 1 \le f(x) < n\} \quad \text{for } n \in \mathbb{N}.$$

Obviously,

$$M = \bigcup_{n \in \mathbb{N}} \quad \text{and} \quad \mu(M_n) = 0 \quad \text{for every } n \in \mathbb{N}.$$

In particular, for each $n \in \mathbb{N}$ there is an open subset G_n of $[a,b]$ such that

$$M_n \subset G_n \text{ and } \mu(G_n) < \frac{\varepsilon}{n\, 2^n}.$$

Now, define a gauge δ_ε in such a way that

$$\delta_\varepsilon(x) = 1 \text{ if } x \notin D \text{ and } (x - \delta_\varepsilon(x), x + \delta_\varepsilon(x)) \subset G_n \text{ if } x \in M_n \text{ for some } n \in \mathbb{N}.$$

Let $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ be a δ_ε-fine partition of $[a,b]$ and let $m = \nu(P)$. Then

$$\begin{aligned}
|S(P)| &= \left|\sum_{j=1}^{m} f(\xi_j)\,[\alpha_j - \alpha_{j-1}]\right| = \left|\sum_{n=1}^{\infty} \sum_{\substack{j=1\\ \xi_j\in M_n}}^{m} f(\xi_j)\,[\alpha_j - \alpha_{j-1}]\right| \\
&\le \sum_{n=1}^{\infty} \left|\sum_{\substack{j=1\\ \xi_j\in M_n}}^{m} f(\xi_j)\,[\alpha_j - \alpha_{j-1}]\right| \le \sum_{n=1}^{\infty} n \left(\sum_{\substack{j=1\\ \xi_j\in M_n}}^{m} (\alpha_j - \alpha_{j-1})\right) \\
&< \varepsilon \left(\sum_{n=1}^{\infty} \frac{1}{2^n}\right) = \varepsilon.
\end{aligned}$$

According to Definition 6.1.2 it follows that $\int_a^b f(x)\,\mathrm{d}x = 0$.

(iii) Let the Newton integral (N) $\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$ exist where the function F is continuous on $[a,b]$ and

$$F'(x) = f(x) \text{ for each } x \in (a,b),\ F'(a+) = f(a),\ F'(b-) = f(b). \tag{6.2.3}$$

We will show that then the KH-integral $\int_a^b f(x)\,\mathrm{d}x$ also exists and its value equals $F(b) - F(a)$.

Let $\varepsilon > 0$ be given. Due to (6.2.3), for each $\xi \in [a,b]$ there exists a $\delta_\varepsilon(\xi) > 0$ such that

$$|F(x) - F(\xi) - f(\xi)\,(x - \xi)| < \frac{\varepsilon}{b-a}\,|x - \xi|$$

is true for all $x \in [a,b] \cap (\xi - \delta_\varepsilon(\xi), \xi + \delta_\varepsilon(\xi))$. Now, let $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ be an arbitrary
δ_ε-fine partition of $[a,b]$ and put $m = \nu(P)$. Then

$$\begin{aligned}
&\big|F(\alpha_j) - F(\alpha_{j-1}) - f(\xi_j)\,[\alpha_j - \alpha_{j-1}]\big| \\
&\quad\le \big|F(\alpha_j) - F(\xi_j) - f(\xi_j)\,[\alpha_j - \xi_j]\big| \\
&\qquad + \big|F(\xi_j) - F(\alpha_{j-1}) - f(\xi_j)\,[\xi_j - \alpha_{j-1}]\big| \\
&\quad< \frac{\varepsilon}{b-a}\,\big(|\alpha_j - \xi_j| + |\xi_j - \alpha_{j-1}|\big) = \frac{\varepsilon}{b-a}\,[\alpha_j - \alpha_{j-1}]
\end{aligned}$$

for every $j \in \{1, \dots, m\}$ and hence

$$\big|[F(b) - F(a)] - S(P)\big| = \left|\sum_{j=1}^{m} \Big(F(\alpha_j) - F(\alpha_{j-1}) - f(\xi_j)\,[\alpha_j - \alpha_{j-1}]\Big)\right|$$

$$\leq \sum_{j=1}^{m} \left| F(\alpha_j) - F(\alpha_{j-1}) - f(\xi_j)\,[\alpha_j - \alpha_{j-1}] \right|$$

$$< \frac{\varepsilon}{b-a} \sum_{j=1}^{m} [\alpha_j - \alpha_{j-1}] = \varepsilon,$$

i.e. $\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$.

6.2.4 Remark. Consider the Dirichlet function

$$f_D(x) = \begin{cases} 1 & \text{if } x \text{ is rational,} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$

(cf. Exercise 4.1.4 (iii)) and an arbitrary bounded interval $[a,b]$. By Example 6.2.3 (ii), we have

$$\int_a^b f_D(x)\,\mathrm{d}x = 0.$$

On the other hand, the Riemann integral (R) $\int_a^b f_D\,\mathrm{d}x$ does not exist whenever the interval $[a,b]$ is nondegenerate. Indeed, let $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \ldots, \alpha_m\}$ be an arbitrary division of $[a,b]$. Choosing all the tags $\xi_j \in [\alpha_{j-1}, \alpha_j]$ rational, we get $S(\boldsymbol{\alpha}, \boldsymbol{\xi}) = b - a \neq 0$, while, if all the tags are irrational, we have $S(\boldsymbol{\alpha}, \boldsymbol{\xi}) = 0$, wherefrom the assertion immediately follows.

The Kurzweil-Stieltjes integral is closely connected with the integral known as the *Perron-Stieltjes* integral even though its definition actually belongs to A. J. Ward [151]. Ward's definition is also described in Section VI.8 of the monograph by S. Saks [116] and is based on the terms *majorant* and *minorant*.

6.2.5 Definition. Let $f, g : [a,b] \to \mathbb{R}$. We say that $M : [a,b] \to \mathbb{R}$ is a *majorant* for f with respect to g if there exists a gauge δ on $[a,b]$ such that

$$(t-\tau)\,\big[M(t) - M(\tau)\big] \geq (t-\tau)\, f(\tau)\,\big[g(t) - g(\tau)\big]$$

holds for every $\tau \in [a,b]$ and every $t \in [a,b] \cap (\tau - \delta(\tau), \tau + \delta(\tau))$.

Similarly, $m : [a,b] \to \mathbb{R}$ is the *minorant* for f with respect to g if there exist a gauge δ on $[a,b]$ such that

$$(t-\tau)\,\big[m(t) - m(\tau)\big] \leq (t-\tau)\, f(\tau)\,\big[g(t) - g(\tau)\big]$$

holds for every $\tau \in [a,b]$ and every $t \in [a,b] \cap (\tau - \delta(\tau), \tau + \delta(\tau))$.

The set of all majorants for f with respect to g is denoted by $\mathfrak{M}(f \Delta g)$ whereas $\mathfrak{m}(f \Delta g)$ stands for the set of all minorants for f with respect to g.

6.2.6 Definition. Let $f, g : [a,b] \to \mathbb{R}$ and

$$\mathfrak{M}(f\Delta g) \neq \emptyset \neq \mathfrak{m}(f\Delta g). \tag{6.2.4}$$

Then we define

$$(\mathrm{PS})\overline{\int_a^b} f\,\mathrm{d}g = \inf\left\{M(b) - M(a) : M \in \mathfrak{M}[a,b]\right\}$$

and

$$(\mathrm{PS})\underline{\int_a^b} f\,\mathrm{d}g = \sup\left\{m(b) - m(a) : m \in \mathfrak{m}[a,b]\right\}.$$

The quantity $(\mathrm{PS})\overline{\int_a^b} f\,\mathrm{d}g$ is called the *upper Perron-Stieltjes integral* of the function f with respect to g (from a to b) while $(\mathrm{PS})\underline{\int_a^b} f\,\mathrm{d}g$ is the *lower Perron-Stieltjes integral* of the function f with respect to g (from a to b).

In the cases when (6.2.4) does not hold, we set

$$(\mathrm{PS})\overline{\int_a^b} f\,\mathrm{d}g = \infty \qquad \text{if} \quad \mathfrak{M}(f\Delta g) = \emptyset,$$

and

$$(\mathrm{PS})\underline{\int_a^b} f\,\mathrm{d}g = -\infty \qquad \text{if} \quad \mathfrak{m}(f\Delta g) = \emptyset.$$

It is not surprising that the following statement holds.

6.2.7 Lemma. *For any functions $f, g : [a,b] \to \mathbb{R}$ the inequality*

$$(\mathrm{PS})\underline{\int_a^b} f\,\mathrm{d}g \le (\mathrm{PS})\overline{\int_a^b} f\,\mathrm{d}g \tag{6.2.5}$$

holds.

Proof. If at least one of the sets $\mathfrak{M}(f\Delta g)$ or $\mathfrak{m}(f\Delta g)$ is empty, then the inequality (6.2.5) is trivially satisfied. Therefore, assume that (6.2.4) holds.

Choose arbitrary majorant $M \in \mathfrak{M}(f\Delta g)$ and minorant $m \in \mathfrak{m}(f\Delta g)$. By the definition there are gauges δ_1 and δ_2 on $[a,b]$ such that

$$(t-\tau)\left[M(t) - M(\tau)\right] \ge (t-\tau)\, f(\tau)\left[g(t) - g(\tau)\right]$$
$$\text{for } \tau \in [a,b] \text{ and } t \in [a,b] \cap (\tau - \delta_1(\tau), \tau + \delta_1(\tau)),$$

and

$$(t-\tau)\,\big[m(t)-m(\tau)\big] \le (t-\tau)\,f(\tau)\,\big[g(t)-g(\tau)\big]$$
$$\text{for } \tau \in [a,b] \text{ and } t \in [a,b] \cap (\tau-\delta_2(\tau), \tau+\delta_2(\tau))$$

hold. Set $\delta(x) = \min\{\delta_1(x), \delta_2(x)\}$ for $x \in [a,b]$. Then

$$m(t)-m(\tau) \le f(\tau)\,\big[g(t)-g(\tau)\big] \le M(t)-M(\tau)$$
$$\text{for } \tau \in [a,b] \text{ and } t \in [a,b] \cap [\tau, \tau+\delta(\tau)),$$

and

$$m(\tau)-m(t) \le f(\tau)\,\big[g(\tau)-g(t)\big] \le M(\tau)-M(t)$$
$$\text{for } \tau \in [a,b] \text{ and } t \in [a,b] \cap [\tau-\delta(\tau), \tau].$$

Thus, for any δ-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$ and every $j \in \{1,\dots,\nu(P)\}$ we have

$$m(\alpha_j)-m(\xi_j) \le f(\xi_j)\,\big[g(\alpha_j)-g(\xi_j)\big] \le M(\alpha_j)-M(\xi_j)$$

and

$$m(\xi_j)-m(\alpha_{j-1}) \le f(\xi_j)\,\big[g(\xi_j)-g(\alpha_{j-1})\big] \le M(\xi_j)-M(\alpha_{j-1}).$$

Summing these inequalities for $j = 1,\dots,\nu(P)$, we get

$$m(b)-m(a) \le S(P) \le M(b)-M(a). \tag{6.2.6}$$

This implies that the inequality

$$m(b)-m(a) \le M(b)-M(a)$$

holds for every $M \in \mathfrak{M}(f\Delta g)$ and every $m \in \mathfrak{m}(f\Delta g)$. Hence,

$$\text{(PS)}\underline{\int_a^b} f\,\mathrm{d}g = \sup\big\{m(b)-m(a) : m \in \mathfrak{m}[a,b]\big\}$$
$$\le \inf\big\{M(b)-M(a) : M \in \mathfrak{m}[a,b]\big\} = \text{(PS)}\overline{\int_a^b} f\,\mathrm{d}g.$$

This completes the proof of (6.2.5). □

6.2.8 Remark. Notice that the proof of the previous lemma contains also the proof of the following assertion:
For given functions $f, g : [a,b] \to \mathbb{R}$, $M \in \mathfrak{M}(f\Delta g)$ and $m \in \mathfrak{m}(f\Delta g)$, there is a gauge δ on $[a,b]$ such that the inequalities (6.2.6) hold for each δ-fine partition P of $[a,b]$.

6.2.9 Definition. If

$$(\mathrm{PS})\underline{\int_a^b} f\,\mathrm{d}g = (\mathrm{PS})\overline{\int_a^b} f\,\mathrm{d}g \in \mathbb{R},$$

we say that the *Perron-Stieltjes integral* (in short, PS-integral)

$$(\mathrm{PS})\int_a^b f\,\mathrm{d}g$$

of the function f with respect to g from a to b exists. Its value is determined by the common value of the upper and lower integral, i.e.

$$(\mathrm{PS})\int_a^b f\,\mathrm{d}g = (\mathrm{PS})\underline{\int_a^b} f\,\mathrm{d}g = (\mathrm{PS})\overline{\int_a^b} f\,\mathrm{d}g.$$

The relationship between the PS-integral and the KS-integral is described by the following theorem.

6.2.10 Theorem. *The integral* $(\mathrm{PS})\displaystyle\int_a^b f\,\mathrm{d}g$ *exists if and only if the* KS-*integral* $\int_a^b f\,\mathrm{d}g$ *exists and in that case, both integrals have the same value*

$$\int_a^b f\,\mathrm{d}g = (\mathrm{PS})\int_a^b f\,\mathrm{d}g. \tag{6.2.7}$$

Proof. a) Assume that $(\mathrm{PS})\displaystyle\int_a^b f\,\mathrm{d}g = I \in \mathbb{R}$ and let an arbitrary $\varepsilon > 0$ be given. By definition, there are a majorant $M \in \mathfrak{M}(f\Delta g)$ and a minorant $m \in \mathfrak{m}(f\Delta g)$ such that

$$M(b) - M(a) - \frac{\varepsilon}{2} < I < m(b) - m(a) + \frac{\varepsilon}{2},$$

or equivalently

$$I - \frac{\varepsilon}{2} \le m(b) - m(a) \le M(b) - M(a) < I + \frac{\varepsilon}{2}. \tag{6.2.8}$$

According to Remark 6.2.8 there is a gauge δ on $[a,b]$ such that (6.2.6) holds for each δ-fine partition P of $[a,b]$. Combining the inequalities (6.2.6) and (6.2.8), we show that the inequalities

$$I - \frac{\varepsilon}{2} \le m(b) - m(a) \le S(P) \le M(B) - M(a) < I + \frac{\varepsilon}{2}$$

hold for every δ-fine partition P of the interval $[a,b]$. This means that the KS–integral $\int_a^b f\,\mathrm{d}g$ exists and relation (6.2.7) is true.

b) Assume that $\int_a^b f\,\mathrm{d}g = I \in \mathbb{R}$ and an arbitrary $\varepsilon > 0$ is given. By Definition 6.1.2 there is a gauge δ_ε such that

$$I - \frac{\varepsilon}{2} < S(P) < I + \frac{\varepsilon}{2} \quad \text{for all } \delta_\varepsilon\text{-fine partitions } P \text{ of } [a,b]. \tag{6.2.9}$$

Define $M(a) = m(a) = 0$,

$$M(x) = \sup\big\{S(Q) : Q \text{is a } \delta_\varepsilon\text{-fine partition of } [a,x]\big\} \quad \text{for } x \in (a,b]$$

and

$$m(x) = \inf\big\{S(Q) : Q \text{is a } \delta_\varepsilon\text{-fine partition of } [a,x]\big\} \quad \text{for } x \in (a,b].$$

By (6.2.9) we have

$$I - \varepsilon < I - \frac{\varepsilon}{2} \le m(b) - m(a) \le M(b) - M(a) \le I + \frac{\varepsilon}{2} < I + \varepsilon. \tag{6.2.10}$$

Let $x \in [a,b)$ and $t \in [x, x+\delta_\varepsilon) \cap [a,b]$. Further, let $Q = (\boldsymbol{\beta}, \boldsymbol{\eta})$ be an arbitrary δ_ε-fine partition of $[a,x]$ and $\widetilde{Q} = (\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\eta}})$, where $\widetilde{\boldsymbol{\beta}} = \boldsymbol{\beta} \cup \{t\}$ and $\widetilde{\boldsymbol{\eta}} = \boldsymbol{\eta} \cup \{x\}$. Then $\widetilde{Q}$ is a δ_ε fine partition of $[a,t]$ and

$$S(Q) + f(x)\,[g(t) - g(x)] = S(\widetilde{Q}). \tag{6.2.11}$$

Passing to the supremum on both sides of equality (6.2.11) we obtain the inequality

$$M(t) \ge M(x) + f(x)\,[g(t) - g(x)] \text{ for } x \in [a,b) \text{ and } t \in [x, x+\delta_\varepsilon) \cap [a,b].$$

Analogously,

$$M(x) \ge M(t) + f(x)\,[g(x) - g(t)] \text{ for } x \in (a,b] \text{ and } t \in (x-\delta_\varepsilon, x] \cap [a,b].$$

Similarly, we can also prove the following inequalities

$$m(t) \le m(x) + f(x)\,[g(t) - g(x)] \text{ for } x \in [a,b] \text{ and } t \in [x, x+\delta_\varepsilon) \cap [a,b]$$

and

$$m(x) \le m(t) + f(x)\,[g(x) - g(t)] \text{ for } x \in [a,b] \text{ and } t \in (x-\delta_\varepsilon, x] \cap [a,b].$$

It follows that M and m are, respectively, the majorant and minorant for f with respect to g. Consequently, having in mind Definition 6.2.6, Lemma 6.2.7 and (6.2.10), we can see that the inequalities

$$I - \varepsilon < m(b) - m(a) \le (\mathrm{PS})\underline{\int_a^b} f\,\mathrm{d}g \le (\mathrm{PS})\overline{\int_a^b} f\,\mathrm{d}g \le M(b) - M(a) < I + \varepsilon$$

are true. As $\varepsilon > 0$ was arbitrary, this yields

$$(\mathrm{PS})\underline{\int_a^b} f\,\mathrm{d}g = (\mathrm{PS})\overline{\int_a^b} f\,\mathrm{d}g = (\mathrm{PS})\int_a^b f\,\mathrm{d}g = I,$$

which completes the proof. $\square$

6.2.11 Remark. Notice that if $M \in \mathfrak{M}(f \Delta g)$ and the integral $\int_a^b f\,\mathrm{d}g$ exists, then by Theorem 6.2.10 and Definitions 6.2.6 and 6.2.9 the estimate

$$\int_a^b f\,\mathrm{d}g \le M(b) - M(a)$$

is true. Analogously, if $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ are such that the integral $\int_a^b f\,\mathrm{d}g$ exists, and if there are a gauge δ on $[a,b]$ and a function $u : [a,b] \to \mathbb{R}$ nondecreasing on $[a,b]$ such that

$$\left.\begin{aligned} &|t-\tau|\,|f(\tau)|\,|g(t)-g(\tau)| \le (t-\tau)\big(u(t)-u(\tau)\big) \\ &\qquad \text{holds whenever } \tau \in [a,b] \text{ and } t \in (\tau-\delta(\tau), \tau+\delta(\tau)) \cap [a,b], \end{aligned}\right\} \tag{6.2.12}$$

then

$$\begin{aligned} \big|S(P)\big| &\le \sum_{j=1}^{\nu(P)} |f(\xi_j)|\,\big(|g(\alpha_j)-g(\xi_j)| + |g(\xi_j)-g(\alpha_{j-1})|\big) \\ &\le \sum_{j=1}^{\nu(P)} \big(u(\alpha_j)-u(\alpha_{j-1})\big) = u(b)-u(a) \end{aligned}$$

for every δ-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$, wherefrom the estimate

$$\left|\int_a^b f\,\mathrm{d}g\right| \le u(b)-u(a) \tag{6.2.13}$$

instantaneously follows.

6.3 Existence of integral

In Examples 6.2.3, we determined the values of some simple KH-integrals directly from the definition. Now, we want to show that in some simple examples it is also possible to determine directly the values of KS-integrals.

6.3.1 Examples. (i) If $f = f(a)$ on $[a,b]$, then clearly

$$\int_a^b f\,\mathrm{d}g = f(a)\,[g(b)-g(a)] \quad\text{and}\quad \int_a^b g\,\mathrm{d}f = 0$$

for every function $g\colon[a,b]\to\mathbb{R}$.

(ii) For any function $f\colon[a,b]\to\mathbb{R}$, the following relations are true:

$$\int_a^b f\,\mathrm{d}\chi_{(\tau,b]} = f(\tau) \qquad \text{if } \tau\in[a,b), \tag{6.3.1}$$

$$\int_a^b f\,\mathrm{d}\chi_{[\tau,b]} = f(\tau) \qquad \text{if } \tau\in(a,b], \tag{6.3.2}$$

$$\int_a^b f\,\mathrm{d}\chi_{[a,\tau]} = -f(\tau) \quad \text{if } \tau\in[a,b), \tag{6.3.3}$$

$$\int_a^b f\,\mathrm{d}\chi_{[a,\tau)} = -f(\tau) \quad \text{if } \tau\in(a,b] \tag{6.3.4}$$

and

$$\int_a^b f\,\mathrm{d}\chi_{[\tau]} = \begin{cases} -f(a) & \text{if } \tau=a,\\ 0 & \text{if } \tau\in(a,b),\\ f(b) & \text{if } \tau=b. \end{cases} \tag{6.3.5}$$

Let us show the proofs of (6.3.1) and (6.3.2). The other ones then follow by Theorem 6.1.11.

a) Let $\tau\in[a,b)$ and $g(x)=\chi_{(\tau,b]}(x)$ on $[a,b]$. Then $g\equiv 0$ on $[a,\tau]$ and by the example (i) we have

$$\int_a^\tau f\,\mathrm{d}g = 0.$$

By Lemma 6.1.12, there is a gauge δ on $[\tau,b]$ such that $\tau=\alpha_0=\xi_1$ holds for each δ-fine partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[\tau,b]$. Moreover, we have

$$g(\alpha_j)-g(\alpha_{j-1})=0 \quad \text{for } j\in\{2,3,\dots,\nu(P)\}.$$

Thus,

$$S(P)=f(\tau)\,[g(\alpha_1)-g(\tau)]=f(\tau) \ \text{ for each } \delta\text{-fine partition } P=(\boldsymbol{\alpha},\boldsymbol{\xi}) \text{ of } [\tau,b],$$

which implies that

$$\int_\tau^b f\,\mathrm{d}g = f(\tau).$$

Relation (6.3.1) now follows by Theorem 6.1.11.

b) Relationship (6.3.2) will be proved analogously. We have

$$\tau \in (a,b], \quad g(x) = \chi_{[\tau,b]}(x) \quad \text{and} \quad \int_\tau^b f \, \mathrm{d}g = 0.$$

By Lemma 6.1.12, there is a gauge δ on $[a,\tau]$ such that $\alpha_{\nu(P)} = \xi_{\nu(P)} = \tau$ for each δ-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,\tau]$. Thus

$$S(P) = f(\tau)\,[g(\tau) - g(\alpha_{\nu(P)-1})] = f(\tau)$$

for every δ-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,\tau]$. Therefore,

$$\int_a^\tau f \, \mathrm{d}g = f(\tau).$$

Relation (6.3.2) follows again by using Theorem 6.1.11.

(iii) For any function g regulated on $[a,b]$, the following relations are true:

$$\int_a^b \chi_{(\tau,b]} \, \mathrm{d}g = g(b) - g(\tau+) \quad \text{if } \tau \in [a,b), \tag{6.3.6}$$

$$\int_a^b \chi_{[\tau,b]} \, \mathrm{d}g = g(b) - g(\tau-) \quad \text{if } \tau \in (a,b], \tag{6.3.7}$$

$$\int_a^b \chi_{[a,\tau]} \, \mathrm{d}g = g(\tau+) - g(a) \quad \text{if } \tau \in [a,b), \tag{6.3.8}$$

$$\int_a^b \chi_{[a,\tau)} \, \mathrm{d}g = g(\tau-) - g(a) \quad \text{if } \tau \in (a,b] \tag{6.3.9}$$

and

$$\int_a^b \chi_{[\tau]} \, \mathrm{d}g = \begin{cases} \Delta^+ g(a) & \text{if } \tau = a, \\ \Delta\, g(\tau) & \text{if } \tau \in (a,b), \\ \Delta^- g(b) & \text{if } \tau = b. \end{cases} \tag{6.3.10}$$

Again, we limit ourselves to the proof of the first two relations.

a) First, let $\tau \in [a,b)$ and $f(x) = \chi_{(\tau,b]}(x)$. We have

$$\int_a^\tau f \, \mathrm{d}g = 0.$$

Let $\varepsilon > 0$ be given and let $\eta_\varepsilon > 0$ be such that

$$|g(\tau+) - g(x)| < \varepsilon \quad \text{for } x \in (\tau, \tau + \eta_\varepsilon).$$

By Lemma 6.1.12 we can choose a gauge δ on $[\tau, b]$ such that $\tau = \alpha_0 = \xi_1$ for each δ-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[\tau, b]$. Set

$$\delta_\varepsilon(x) = \begin{cases} \min\{\eta_\varepsilon, \delta(\tau)\} & \text{if } x = \tau, \\ \delta(x), & \text{if } x \in (\tau, b]. \end{cases}$$

Let an arbitrary δ_ε-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[\tau, b]$ be given. Then

$$\tau = \alpha_0 = \xi_1 \text{ and } \alpha_1 \in (\tau, \tau + \eta_\varepsilon)$$

and consequently

$$\begin{aligned} &\Big| S(P) - [g(b) - g(\tau+)] \Big| \\ &\quad = \Big| [g(b) - g(\alpha_{\nu(P)-1})] + [g(\alpha_{\nu(P)-1}) - g(\alpha_{\nu(P)-2}] \\ &\qquad + \dots + [g(\alpha_2) - g(\alpha_1)] - [g(b) - g(\tau+)] \Big| \\ &\quad = |g(\tau+) - g(\alpha_1)| < \varepsilon. \end{aligned}$$

Thus,

$$\int_\tau^b f \, \mathrm{d}g = g(b) - g(\tau+),$$

and, due to Theorem 6.1.11,

$$\int_a^b f \, \mathrm{d}g = \int_a^\tau f \, \mathrm{d}g + \int_\tau^b f \, \mathrm{d}g = g(b) - g(\tau+),$$

i.e. (6.3.6) is true.

b) Let $\tau \in (a, b]$ and $f(x) = \chi_{[\tau,b]}(x)$ for $x \in [a, b]$. Then

$$\int_\tau^b f \, \mathrm{d}g = g(b) - g(\tau).$$

Let an arbitrary $\varepsilon > 0$ be given. Choose $\eta_\varepsilon > 0$ such that $|g(\tau-) - g(x)| < \varepsilon$ for every $x \in (\tau - \eta_\varepsilon, \tau)$. Further, using Lemma 6.1.12 we can choose a gauge δ on $[a, \tau]$ such that

$$\tau = \alpha_{\nu(P)} = \xi_{\nu(P)} \quad \text{for each } \delta\text{-fine partition } P = (\boldsymbol{\alpha}, \boldsymbol{\xi}) \text{ of } [a, \tau].$$

Now, set

$$\delta_\varepsilon(x) = \begin{cases} \delta(x), & \text{if } \ x \in [a,\tau), \\ \min\{\eta_\varepsilon, \delta(\tau)\} & \text{if } \ x = \tau. \end{cases}$$

Let an arbitrary δ-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,\tau]$ be given. Then

$$\tau = \alpha_{\nu(P)} = \xi_{\nu(P)}, \ \ \alpha_{\nu(P)-1} \in (\tau - \eta_\varepsilon, \tau) \ \text{ and } \ S(P) = g(\tau) - g(\alpha_{\nu(P)-1})$$

and consequently

$$\begin{aligned} \Big|S(P) - [g(\tau) - g(\tau-)]\Big| &= \Big|[g(\tau) - g(\alpha_{\nu(P)-1})] - [g(\tau) - g(\tau-)]\Big| \\ &= \Big|g(\tau-) - g(\alpha_{\nu(P)-1})\Big| < \varepsilon, \end{aligned}$$

wherefrom

$$\int_a^\tau f \, \mathrm{d}g = g(\tau) - g(\tau-).$$

Finally, using Theorem 6.1.11, we get

$$\int_a^b f \, \mathrm{d}g = \int_a^\tau f \, \mathrm{d}g + \int_\tau^b f \, \mathrm{d}g = g(b) - g(\tau-).$$

As far as the existence of the integral is concerned, since every finite step function is a finite linear combination of functions of the form $\chi_{(\tau,b]}$, $\chi_{[\tau,b]}$, $\chi_{[b]}$ where τ can be an arbitrary point from $[a,b)$ (see (2.5.1)), we can summarize the above examples into the following statement.

6.3.2 Corollary. *If* $g \in \mathrm{G}([a,b])$ *and* $f \in \mathrm{S}[a,b]$, *then both integrals*

$$\int_a^b f \, \mathrm{d}g \quad \textit{and} \quad \int_a^b g \, \mathrm{d}f$$

exist.

6.3.3 Exercise. Prove the following statement:
Let $h : [a,b] \to \mathbb{R}$, $c \in \mathbb{R}$, $D = \{d_1, \ldots, d_n\} \subset [a,b]$ *and* $h(x) = c$ *for* $x \in [a,b] \setminus D$. *Then*

$$\int_a^b f \, \mathrm{d}h = f(b)\, h(b) - f(a)\, h(a) - \big(f(b) - f(a)\big)\, c$$

holds for every $f : [a,b] \to \mathbb{R}$.
Hint: Write the function h in the form $h(x) = c + \sum_{k=1}^n \big[h(d_k) - c\big]\, \chi_{[d_k]}(x)$ and use the results of Examples 6.3.1.

The next three theorems give us the basic estimates for KS-integrals under the assumption that these integrals exist. The first two require no other assumptions about the functions f and g. Nevertheless, it is obvious that these theorems have a practical meaning only under the assumption that f is bounded on $[a,b]$ and g has a bounded variation on $[a,b]$.

6.3.4 Theorem. *Let* $f, g : [a,b] \to \mathbb{R}$. *Then*

$$|S(P)| \le \|f\| \operatorname{var}_a^b g \quad \textit{holds for each partition } P \textit{ of } [a,b]. \tag{6.3.11}$$

If the integral $\int_a^b f\,\mathrm{d}g$ *exists, then*

$$\Big|\int_a^b f\,\mathrm{d}g\Big| \le \|f\| \operatorname{var}_a^b g. \tag{6.3.12}$$

If, in addition, the integral $\int_a^b |f(x)|\,\mathrm{d}[\operatorname{var}_a^x g]$ *exists, then we have also*

$$\Big|\int_a^b f\,\mathrm{d}g\Big| \le \int_a^b |f(x)|\,\mathrm{d}[\operatorname{var}_a^x g] \le \|f\| \operatorname{var}_a^b g. \tag{6.3.13}$$

Proof. For every partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$ we have

$$\begin{aligned} |S(P)| &\le \sum_{j=1}^{\nu(P)} |f(\xi_j)|\,|g(\alpha_j) - g(\alpha_{j-1})| \\ &\le \sum_{j=1}^{\nu(P)} |f(\xi_j)| \left(\operatorname{var}_{\alpha_{j-1}}^{\alpha_j} g\right) \le \|f\| \operatorname{var}_a^b g, \end{aligned}$$

wherefrom the assertion of the theorem immediately follows. □

6.3.5 Theorem. *Let* $f, g : [a,b] \to \mathbb{R}$. *Then*

$$\left.\begin{aligned} &|S(P)| \le (|f(a)| + |f(b)| + \operatorname{var}_a^b f)\,\|g\| \\ &\qquad\qquad \textit{for each partition } P = (\boldsymbol{\alpha}, \boldsymbol{\xi}) \textit{ of } [a,b]. \end{aligned}\right\} \tag{6.3.14}$$

Furthermore,

$$\Big|\int_a^b f\,\mathrm{d}g\Big| \le \left(|f(a)| + |f(b)| + \operatorname{var}_a^b f\right) \|g\| \tag{6.3.15}$$

holds whenever the integral $\int_a^b f\,\mathrm{d}g$ *exists.*

Proof. For an arbitrary partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$ we have

$$\begin{aligned}S(P)&=f(\xi_1)\,[g(\alpha_1)-g(a)]+f(\xi_2)\,[g(\alpha_2)-g(\alpha_1)]\\&\qquad\qquad+\cdots+f(\xi_m)\,[g(b)-g(\alpha_{m-1})]\\&=f(b)\,g(b)-f(a)\,g(a)\\&\qquad-[f(\xi_1)-f(a)]\,g(a)-[f(\xi_2)-f(\xi_1)]\,g(\alpha_1)\\&\qquad\qquad-\cdots-[f(b)-f(\xi_m)]\,g(b)\\&=f(b)\,g(b)-f(a)\,g(a)-\sum_{j=0}^{\nu(P)}[f(\xi_{j+1})-f(\xi_j)]\,g(\alpha_j),\end{aligned}$$

where $\xi_0=a$ and $\xi_{m+1}=b$. Inequalities (6.3.14) and (6.3.15) now immediately follow. □

6.3.6 Remark. Usually, instead of (6.3.14) or (6.3.15), slightly less sharp estimates

$$|S(P)|\le 2\,\|f\|_{\mathrm{BV}}\,\|g\|\quad\text{for every partition } P=(\boldsymbol{\alpha},\boldsymbol{\xi}) \text{ of } [a,b] \tag{6.3.16}$$

and

$$\left|\int_a^b f\,\mathrm{d}g\right|\le 2\,\|f\|_{\mathrm{BV}}\,\|g\| \tag{6.3.17}$$

will be sufficient for our aims.

Theorem 6.3.4 enables us to prove the simplest convergence theorem.

6.3.7 Theorem. *Let* $f:[a,b]\to\mathbb{R}$, $g\in\mathrm{BV}([a,b])$ *and let the sequence* $\{f_n\}$ *of functions defined on the interval* $[a,b]$ *be such that*

$$\lim_{n\to\infty}\|f_n-f\|=0, \tag{6.3.18}$$

and all the integrals $\int_a^b f_n\,\mathrm{d}g$, $n\in\mathbb{N}$, *exist. Then the integral* $\int_a^b f\,\mathrm{d}g$ *exists, too, and*

$$\lim_{n\to\infty}\int_a^b f_n\,\mathrm{d}g=\int_a^b f\,\mathrm{d}g. \tag{6.3.19}$$

Proof. By assumption (6.3.18), for a given $\varepsilon>0$ there is $n_1\in\mathbb{N}$ such that

$$\|f_n-f_m\|<\frac{\varepsilon}{1+\operatorname{var}_a^b g}\quad\text{for all } m,n\ge n_1.$$

Thus, by Theorems 6.1.8 and 6.3.4, we have

$$\Big|\int_a^b f_n\,\mathrm{d}g-\int_a^b f_m\,\mathrm{d}g\Big|\le\|f_n-f_m\|\operatorname{var}_a^b g<\varepsilon\quad\text{for all } n\ge n_1.$$

Hence, the sequence $\Big\{\int_a^b f_n\,\mathrm{d}g\Big\}\subset\mathbb{R}$ is a Cauchy one and consequently there exists a number $I\in\mathbb{R}$ such that

$$\lim_{n\to\infty}\int_a^b f_n\,\mathrm{d}g=I.\tag{6.3.20}$$

To show that $\int_a^b f\,\mathrm{d}g=I$, let $\varepsilon>0$ be given. Then we can choose $n_0\in\mathbb{N}$ such that

$$\Big|\int_a^b f_{n_0}\,\mathrm{d}g-I\Big|<\varepsilon\quad\text{and}\quad\|f_{n_0}-f\|<\varepsilon.\tag{6.3.21}$$

Moreover, let δ_0 be a gauge on $[a,b]$ such that

$$\Big|S_{n_0}(P)-\int_a^b f_{n_0}\,\mathrm{d}g\Big|<\varepsilon\tag{6.3.22}$$

holds for all δ_0-fine partitions P of $[a,b]$, where $S_{n_0}(P)=S(f_{n_0},\mathrm{d}g,P)$. Now, let a δ_0-fine partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$ be given. Then by (6.3.21) we have

$$\begin{aligned}\big|S(P)-S_{n_0}(P)\big|&=\Big|\sum_{j=1}^{\nu(P)}\big(f(\xi_j)-f_{n_0}(\xi_j)\big)\,\big[g(\alpha_j)-g(\alpha_{j-1})\big]\Big|\\&\le\|f_{n_0}-f\|\operatorname{var}_a^b g<\varepsilon\operatorname{var}_a^b g,\end{aligned}$$

and, furthermore, using (6.3.21) and (6.3.22) we get

$$\begin{aligned}|S(P)-I|\le|S(P)-S_{n_0}(P)|&+\Big|S_{n_0}(P)-\int_a^b f_{n_0}\,\mathrm{d}g\Big|\\&+\Big|\int_a^b f_{n_0}\,\mathrm{d}g-I\Big|\\&<\varepsilon\,(\operatorname{var}_a^b g+2).\end{aligned}$$

Therefore

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

Now, we can formulate the first existence result.

6.3.8 Theorem. *Let* $f \in \mathrm{G}([a,b])$ *and* $g \in \mathrm{BV}([a,b])$. *Then both the integrals*

$$\int_a^b f \,\mathrm{d}g \quad \textit{and} \quad \int_a^b |f(x)| \,\mathrm{d}\big[\mathrm{var}_a^x\, g\big]$$

exist and estimates (6.3.12) *and* (6.3.13) *hold.*

Proof. By Theorem 4.1.5, there exists a sequence $\{f_n\}$ of finite step functions which converges uniformly on $[a,b]$ to the function f. Further, by Corollary 6.3.2 all the integrals

$$\int_a^b f_n \,\mathrm{d}g, \qquad n \in \mathbb{N},$$

exist. By Theorem 6.3.7, this means that the integral $\int_a^b f \,\mathrm{d}g$ exists as well and (6.3.19) holds.

Obviously, $|f| \in \mathrm{G}([a,b])$ and by the previous part of the proof, also the integral

$$\int_a^b |f(x)| \,\mathrm{d}\big[\mathrm{var}_a^x\, g\,\big]$$

exists. Thus, according to Theorem 6.3.4, both the relations (6.3.12) and (6.3.13) are true. □

The following convergent result is kind of symmetric to Theorem 6.3.7.

6.3.9 Theorem. *Let* $f : [a,b] \to \mathbb{R}$ *be bounded on* $[a,b]$, $g \in \mathrm{BV}([a,b])$ *and let the sequence* $\{g_n\}$ *of functions defined on the interval* $[a,b]$ *be such that*

$$\lim_{n\to\infty} \mathrm{var}_a^b(g_n - g) = 0,$$

and all the integrals $\int_a^b f \,\mathrm{d}g_n$, $n \in \mathbb{N}$, *exist. Then the integral* $\int_a^b f \,\mathrm{d}g$ *exists as well and*

$$\lim_{n\to\infty} \int_a^b f \,\mathrm{d}g_n = \int_a^b f \,\mathrm{d}g \tag{6.3.23}$$

holds.

Proof. The proof is onward formally similar to the proof of Theorem 6.3.7. For a given $\varepsilon > 0$ there is $n_1 \in \mathbb{N}$ such that

$$\mathrm{var}_a^b\,(g - g_n) = \mathrm{var}\,(g_n - g) < \frac{\varepsilon}{2} \quad \text{for all } n \in \mathbb{N},$$

and consequently

$$\operatorname{var}_a^b(g_n - g_m) < \operatorname{var}_a^b(g_n - g) + \operatorname{var}(g - g_m) < \varepsilon \quad \text{for all } m, n \in \mathbb{N}.$$

By Theorems 6.1.8 and 6.3.5 (see also Remark 6.3.6)

$$\Big|\int_a^b f\,\mathrm{d}g_n - \int_a^b f\,\mathrm{d}g_m\Big| \le \|f\| \operatorname{var}_a^b(g_n - g_m) \quad \text{for all } m, n \in \mathbb{N}.$$

Hence, the sequence $\Big\{\displaystyle\int_a^b f\,\mathrm{d}g_n\Big\}$ is a Cauchy one and there exists $I \in \mathbb{R}$ such that

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

Given $\varepsilon > 0,$ choose $n_0 \in \mathbb{N}$ and a gauge δ_0 on $[a,b]$ such that

$$\Big|\int_a^b f\,\mathrm{d}g_{n_0} - I\Big| < \varepsilon, \quad \operatorname{var}_a^b(g_{n_0} - g) < \varepsilon$$

and

$$\Big|S_{n_0}(P) - \int_a^b f\,\mathrm{d}g_{n_0}\Big| < \varepsilon \ \text{ for all } \delta_0\text{-fine partitions } P \text{ of } [a,b],$$

where $S_{n_0}(P) = S(f, \mathrm{d}g_{n_0}, P).$ Then for every δ_0-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$ we have

$$\begin{aligned}\big|S(P) - S_{n_0}(P)\big| &= \Big|\sum_{j=1}^{\nu(P)} f(\xi_j)\,\big[g(\alpha_j) - g(\alpha_{j-1}) - g_{n_0}(\alpha_j) + g_{n_0}(\alpha_{j-1})\big]\Big| \\ &\le \|f\|\; V(g_{n_0} - g, \boldsymbol{\alpha}) \le \|f\| \operatorname{var}_a^b(g_{n_0} - g) \le \varepsilon\,\|f\|,\end{aligned}$$

and therefore

$$\begin{aligned}|S(P) - I| &\le \big|S(P) - S_{n_0}(P)\big| + \Big|S_{n_0}(P) - \int_a^b f\,\mathrm{d}g_{n_0}\Big| + \Big|\int_a^b f\,\mathrm{d}g_{n_0} - I\Big| \\ &< \varepsilon\,(\|f\| + 2).\end{aligned}$$

This gives the equality

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

which concludes the proof. □

We know that if the function f is regulated on $[a,b]$ and g has a bounded variation on $[a,b]$, then by Theorem 6.3.8 the integral $\int_a^b f\,\mathrm{d}g$ exists. However, in applications, we often need to work with the KS-integral also in the reverse situation when $f\in\mathrm{BV}([a,b])$ and $g\in\mathrm{G}([a,b])$. The following convergence theorem will serve us well when proving the existence of an integral in such a situation.

6.3.10 Theorem. *Let* $f\in\mathrm{BV}([a,b])$, $g:[a,b]\to\mathbb{R}$, *and let the sequence* $\{g_n\}$ *of functions defined on the interval* $[a,b]$ *be such that*

$$\lim_{n\to\infty}\|g_n-g\|=0, \tag{6.3.24}$$

and the integrals $\int_a^b f\,\mathrm{d}g_n$ *exist for* $n\in\mathbb{N}$.

Then the integral $\int_a^b f\,\mathrm{d}g$ *exists and*

$$\lim_{n\to\infty}\int_a^b f\,\mathrm{d}g_n=\int_a^b f\,\mathrm{d}g. \tag{6.3.25}$$

Proof. a) By assumption (6.3.24), for a given $\varepsilon>0$ there is an $n_1\in\mathbb{N}$ such that

$$\|g_n-g_m\|<\frac{\varepsilon}{2\,\|f\|_{\mathrm{BV}}}\quad\text{for all } m,n\ge n_1. \tag{6.3.26}$$

Our assumptions ensure that all the integrals

$$I_n:=\int_a^b f\,\mathrm{d}g_n,\quad n\in\mathbb{N},$$

are defined. Further, by (6.3.26) and Theorem 6.3.5 (see also Remark 6.3.6) the relations

$$\left|\int_a^b f\,\mathrm{d}[g_n-g_m]\right|\le 2\,\|f\|_{\mathrm{BV}}\,\|g_n-g_m\|<\varepsilon\quad\text{for all } m,n\ge n_1$$

hold. The sequence $\left\{\int_a^b f\,\mathrm{d}g_n\right\}\subset\mathbb{R}$ is a Cauchy one and thus has a finite limit, i.e. there exists a number $I\in\mathbb{R}$ such that

$$\lim_{n\to\infty}I_n=\lim_{n\to\infty}\int_a^b f\,\mathrm{d}g_n=I. \tag{6.3.27}$$

b) We will prove that $\int_a^b f\,\mathrm{d}g=I$. To this aim, let $\varepsilon>0$ be arbitrary and choose an $n_0\in\mathbb{N}$ such that simultaneously

$$|I_{n_0}-I|<\varepsilon\quad\text{and}\quad\|g_{n_0}-g\|<\frac{\varepsilon}{2\,\|f\|_{\mathrm{BV}}}. \tag{6.3.28}$$

Further, for an arbitrary partition P of $[a,b]$ put

$$S_{n_0}(P) := S(f, \mathrm{d}g_{n_0}, P) = \sum_{j=1}^{\nu(P)} f(\xi_j)\,[g_{n_0}(\alpha_j) - g_{n_0}(\alpha_{j-1})]$$

and choose a gauge δ_0 on $[a,b]$ such that

$$|S_{n_0}(P) - I_{n_0}| < \varepsilon \quad \text{for each } \delta_0\text{-fine partition } P \text{ of } [a,b]. \tag{6.3.29}$$

By Remark 6.3.6 and due to inequalities (6.3.28), (6.3.29) we deduce that

$$\begin{aligned}|S(P) - I| &\le |S(P) - S_{n_0}(P)| + |S_{n_0}(P) - I_{n_0}| + |I_{n_0} - I| \\ &= |S(f, \mathrm{d}[g - g_{n_0}], P)| + |S_{n_0}(P) - I_{n_0}| + |I_{n_0} - I| \\ &\le 2\,\|f\|_{\mathrm{BV}}\,\|g - g_{n_0}\| + |S_{n_0}(P) - I_{n_0}| + |I_{n_0} - I| < 3\,\varepsilon\end{aligned}$$

is true for each δ_0-fine partition P of $[a,b]$, that is

$$|S(P) - I| < 3\,\varepsilon \quad \text{for all } \delta_0\text{-fine partitions } P \text{ of } [a,b].$$

But this means that $\int_a^b f\,\mathrm{d}g = I$. The proof has been completed. □

Finally, we are able to prove the following important existence result which is somehow reverse to Theorem 6.3.8.

6.3.11 Theorem. *If* $f \in \mathrm{BV}([a,b])$ *and* $g \in \mathrm{G}([a,b])$, *then the integral* $\int_a^b f\,\mathrm{d}g$ *exists and the estimate* (6.3.15) *holds.*

Proof. Choose a sequence $\{g_n\}$ of finite step functions which converges uniformly on $[a,b]$ to g (see Theorem 4.1.5). By Corollary 6.3.2 the integrals $\int_a^b f\,\mathrm{d}g_n$ exist for all $n \in \mathbb{N}$.. This means that, by Theorem 6.3.5, the integral $\int_a^b f\,\mathrm{d}g$ exists as well and the estimate (6.3.15) is true. □

We know that the uniform limit of regulated functions is a regulated function (see Theorem 4.1.3). The following convergent statement is thus a direct corollary of Theorems 6.3.10 and 6.3.11.

6.3.12 Corollary. *If* $g, g_n \in \mathrm{G}([a,b])$ *for* $n \in \mathbb{N}$ *and* (6.3.24) *holds, then*

$$\lim_{n\to\infty} \int_a^b f\,\mathrm{d}g_n = \int_a^b f\,\mathrm{d}g \tag{6.3.30}$$

for every $f \in \mathrm{BV}([a,b])$.

Another variation of the convergent statement not covered by the aforementioned theorems will follow. However, its proof is based on the same principle as the proof of Theorem 6.3.10.

6.3.13 Theorem. *Let the function* g *be bounded on* $[a,b]$ *and let the sequence* $\{f_n\} \subset \mathrm{BV}([a,b])$ *be such that*

$$\int_a^b f_n\,\mathrm{d}g \quad \textit{exists for every } n \in \mathbb{N} \textit{ and } \lim_{n\to\infty} \|f_n - f\|_{\mathrm{BV}} = 0.$$

Then $f \in \mathrm{BV}([a,b])$, *the integral* $\int_a^b f\,\mathrm{d}g$ *exists and*

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

Proof. By Theorems 6.1.8 and 6.3.5 (see also Remark 6.3.6)

$$\left| \int_a^b f_n\,\mathrm{d}g - \int_a^b f_m\,\mathrm{d}g \right| \le 2\,\|g\|\,\|f_n - f_m\|_{\mathrm{BV}} \quad \text{for all } m,n \in \mathbb{N}.$$

Hence, the sequence $\left\{ \int_a^b f_n\,\mathrm{d}g \right\}$ is a Cauchy one and consequently there is $I \in \mathbb{R}$ such that

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

We will show that $\int_a^b f\,\mathrm{d}g = I$. To this aim, let $\varepsilon > 0$ be given and let $n_0 \in \mathbb{N}$ be such that

$$\left| \int_a^b f_{n_0}\,\mathrm{d}g - I \right| < \varepsilon \quad \text{and} \quad \|f_{n_0} - f\|_{\mathrm{BV}} < \varepsilon.$$

Furthermore, choose a gauge δ_ε on $[a,b]$ in such a way that

$$\left| S_{n_0}(P) - \int_a^b f_{n_0}\,\mathrm{d}g \right| < \varepsilon$$

holds for all δ_ε-fine partitions P, where $S_{n_0}(P) = S(f_{n_0}, \mathrm{d}g, P)$. By (6.3.14), for any partition P of $[a,b]$ we have

$$\begin{aligned} \big|S(P) - S_{n_0}(P)\big| &\le \Big(|f(a) - f_{n_0}(a)| + |f(b) - f_{n_0}(b)| + \mathrm{var}_a^b\,(f - f_{n_0}) \Big)\,\|g\| \\ &\le 2\,\|f - f_{n_0}\|_{\mathrm{BV}}\,\|g\|. \end{aligned}$$

Altogether, for every δ_ε-fine partition of $[a,b]$ we obtain

$$\begin{aligned}|S(P)-I| &\le \big|S(P)-S_{n_0}(P)\big| + \Big|S_{n_0}(P)-\int_a^b f_{n_0}\,\mathrm{d}g\Big| \\ &\quad + \Big|\int_a^b f_{n_0}\,\mathrm{d}g - I\Big| \\ &< 2\,\|f-f_{n_0}\|_{\mathrm{BV}}\,\|g\| + 2\,\varepsilon < \varepsilon\,2\,(\|g\|+1)\end{aligned}$$

and thus

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

To complete this section, we will show how to determine the value of the integral $\int_a^b f\,\mathrm{d}g$ when the value of $\int_a^b f^{\mathrm{C}}\,\mathrm{d}g$ is known for the continuous part f^{C} of f. The sum symbol

$$\sum_{d\in D}\big[\Delta^- f(d)\,(g(b)-g(d-)) + \Delta^+ f(d)\,(g(b)-g(d+))\big] \tag{6.3.31}$$

will appear in next corollary where D is the set of the points of discontinuity of the function $f\in\mathrm{BV}([a,b])$ in the open interval (a,b). The set D has at most countably many elements. If it is finite, then the meaning of the symbol (6.3.31) is evident. If D is infinite, then there exists a one-to-one mapping $k\in\mathbb{N}\to d_k\in D$ such that $D=\{d_k\}$. In general, this mapping is not uniquely determined. However, as the series

$$\sum_{k=1}^{\infty}\big[\Delta^- f(d_k)\,(g(b)-g(d_k-)) + \Delta^+ f(d_k)\,(g(b)-g(d_k+))\big]$$

is absolutely convergent, the concrete ordering of the set D does not matter. Therefore, we can use the notation (6.3.31) or

$$\sum_{a<x<b}\big[\Delta^- f(x)\,(g(b)-g(x-)) + \Delta^+ f(x)\,(g(b)-g(x+))\big] \tag{6.3.32}$$

as in Remark 2.3.7.

6.3.14 Corollary. *If $f\in\mathrm{BV}([a,b])$, $g\in\mathrm{G}([a,b])$, D is the set of the points of discontinuity of the function f in (a,b) and f^{C} is the continuous part of f, $f^{\mathrm{C}}(a)=f(a)$, then*

$$\left.\begin{aligned}\int_a^b f\,\mathrm{d}g &= \int_a^b f^{\mathrm{C}}\,\mathrm{d}g + \Delta^+ f(a)\,(g(b)-g(a+)) + \Delta^- f(b)\,\Delta^- g(b) \\ &\quad + \sum_{d\in D}\Big[\Delta^- f(d)\,(g(b)-g(d-)) + \Delta^+ f(d)\,(g(b)-g(d+))\Big].\end{aligned}\right\} \tag{6.3.33}$$

Proof. If the set D is finite, the validity of the relationship (6.3.33) is obvious. Assume D is infinite, i.e. $D = \{d_k\}$. By Theorem 2.3.6

$$\left.\begin{aligned} &\sum_{k=1}^{\infty} \left|\Delta^- f(d_k)\,\chi_{[d_k,b]}(x) + \Delta^+ f(d_k)\,\chi_{(d_k,b]}(x)\right| \\ &\qquad \le \sum_{k=1}^{\infty} \left[|\Delta^- f(d_k)| + |\Delta^+ f(d_k)|\right] < \infty \end{aligned}\right\} \tag{6.3.34}$$

holds for every $x \in [a,b]$. The series on the left hand side of the inequality (6.3.34) is thus absolutely convergent for every $x \in [a,b]$ and hence we may define

$$\begin{aligned} f^{\mathrm{B}}(x) &= \Delta^+ f(a)\,\chi_{(a,b]}(x) + \Delta^- f(b)\,\chi_{[b]}(x) \\ &\quad + \sum_{k=1}^{\infty} \left[\Delta^- f(d_k)\,\chi_{[d_k,b]}(x) + \Delta^+ f(d_k)\,\chi_{(d_k,b]}(x)\right] \text{ for } x \in [a,b] \end{aligned}$$

and

$$\begin{aligned} f_n^{\mathrm{B}}(x) &= \Delta^+ f(a)\,\chi_{(a,b]}(x) + \Delta^- f(b)\,\chi_{[b]}(x) \\ &\quad + \sum_{k=1}^{n} \left[\Delta^- f(d_k)\,\chi_{[d_k,b]}(x) + \Delta^+ f(d_k)\,\chi_{(d_k,b]}(x)\right] \end{aligned}$$

for $x \in [a,b]$ and $n \in \mathbb{N}$. By Theorem 2.6.1, f^{B} is the jump part of the function f, while $f^{\mathrm{C}} = f - f^{\mathrm{B}}$ is its continuous part. Moreover, $f^{\mathrm{B}}(a) = 0$, $f^{\mathrm{C}}(a) = f(a)$ and

$$f^{\mathrm{B}}(x) - f_n^{\mathrm{B}}(x) = \sum_{k=n+1}^{\infty} \left[\Delta^- f(d_k)\,\chi_{[d_k,b]}(x) + \Delta^+ f(d_k)\,\chi_{(d_k,b]}(x)\right]$$

for $x \in [a,b]$ and $n \in \mathbb{N}$. By Definition 2.5.2, $f^{\mathrm{B}} - f_n^{\mathrm{B}}$ is a step function and by Theorem 2.5.3

$$\operatorname{var}_a^b(f^{\mathrm{B}} - f_n^{\mathrm{B}}) \le \sum_{k=n+1}^{\infty} \left[|\Delta^- f(d_k)| + |\Delta^+ f(d_k)|\right]. \tag{6.3.35}$$

As the right hand side of (6.3.35) is the remainder of an absolutely convergent series, we have

$$\lim_{n\to\infty} \operatorname{var}_a^b(f^{\mathrm{B}} - f_n^{\mathrm{B}}) \le \lim_{n\to\infty} \sum_{k=n+1}^{\infty} \left[|\Delta^- f(d_k)| + |\Delta^+ f(d_k)|\right] = 0. \tag{6.3.36}$$

By Theorem 6.3.13, this implies that

$$\int_a^b f^{\mathrm{B}}\,\mathrm{d}g = \lim_{n\to\infty} \int_a^b f_n^{\mathrm{B}}\,\mathrm{d}g \in \mathbb{R}. \tag{6.3.37}$$

On the other hand, by (6.3.6), (6.3.7) and Theorem 6.1.8 we have

$$\left.\begin{aligned}\int_a^b f_n^{\mathrm{B}}\,\mathrm{d}g &= \Delta^+ f(a)\,(g(b)-g(a+)) + \Delta^- f(b)\,\Delta^- g(b)\\ &+\sum_{k=1}^{n}\Big(\Delta^- f(d_k)\,[g(b)-g(d_k-)] + \Delta^+ f(d_k)\,[g(b)-g(d_k+)]\Big)\end{aligned}\right\}\quad(6.3.38)$$

for all $n\in\mathbb{N}$. Thus, due to (6.3.37) and (6.3.38), we get

$$\left.\begin{aligned}\int_a^b f^{\mathrm{B}}\,\mathrm{d}g &= \Delta^+ f(a)\,(g(b)-g(a+)) + \Delta^- f(b)\,\Delta^- g(b)\\ &+\sum_{k=1}^{\infty}\Big(\Delta^- f(d_k)\,(g(b)-g(d_k-)) + \Delta^+ f(d_k)\,(g(b)-g(d_k+))\Big).\end{aligned}\right\}\quad(6.3.39)$$

Finally, since

$$\begin{aligned}&\sum_{k=1}^{\infty}\big|\Delta^- f(d_k)\,(g(b)-g(d_k-)) + \Delta^+ f(d_k)\,(g(b)-g(d_k+))\big|\\ &\qquad\le 2\,\|g\|\sum_{k=1}^{\infty}\Big(|\Delta^- f(d_k)|+|\Delta^+ f(d_k)|\Big)\le 2\,\|g\|\,(\mathrm{var}_a^b f)<\infty\end{aligned}$$

due to Corollary 2.3.8, we can see that the series on the right hand side of (6.3.39) converges absolutely. Thus, we can rewrite it in the form

$$\begin{aligned}&\sum_{k=1}^{\infty}\Big(\Delta^- f(d_k)\,(g(b)-g(d_k-)) + \Delta^+ f(d_k)\,(g(b)-g(d_k+))\Big)\\ &\qquad=\sum_{d\in D}\Big(\Delta^- f(d)\,(g(b)-g(d)) + \Delta^+ f(d)\,(g(b)-g(d))\Big)\end{aligned}$$

or, equivalently,

$$\begin{aligned}&\sum_{k=1}^{\infty}\Big(\Delta^- f(d_k)\,(g(b)-g(d_k-)) + \Delta^+ f(d_k)\,(g(b)-g(d_k+))\Big)\\ &\qquad=\sum_{a<x<b}\Big(\Delta^- f(x)\,(g(b)-g(x)) + \Delta^+ f(x)\,(g(b)-g(x))\Big).\end{aligned}$$

If $f=\widetilde{f}^{\mathrm{C}}+\widetilde{f}^{\mathrm{B}}$ is another decomposition of the function f into continuous and jump part, then by Theorem 2.6.1 there is a constant $c\in\mathbb{R}$ such that

$$\widetilde{f}(x)-f^{\mathrm{C}}(x)=f^{\mathrm{B}}(x)-\widetilde{f}^{\mathrm{B}}(x)=c\quad\text{for all }\ x\in[a,b].$$

Then naturally

$$\begin{aligned}
&\int_a^b \widetilde{f}^{\,\mathrm{C}}\,\mathrm{d}g + \int_a^b \widetilde{f}^{\,\mathrm{B}}\,\mathrm{d}g \\
&\quad = \int_a^b f^{\,\mathrm{C}}\,\mathrm{d}g + c\,[g(b)-g(a)] + \int_a^b f^{\,\mathrm{B}}\,\mathrm{d}g - c\,[g(b)-g(a)] \\
&\quad = \int_a^b f^{\,\mathrm{C}}\,\mathrm{d}g + \int_a^b f^{\,\mathrm{B}}\,\mathrm{d}g.
\end{aligned}$$

Hence (6.3.33) holds, while the value of the integral does not depend on the choice of the decomposition of the function f into its continuous and jump parts. □

In the situation symmetrical to Corollary 6.3.14, we have

6.3.15 Lemma. *Let* $f:[a,b]\to\mathbb{R}$ *be bounded, let* $g\in\mathrm{BV}([a,b])$ *be a step function and let* D *be the set of the points of discontinuity of* g *in* (a,b). *Then*

$$\int_a^b f\,\mathrm{d}g = f(a)\,\Delta^+g(a) + \sum_{d\in D} f(d)\,\Delta g(d) + f(b)\,\Delta^-g(b). \tag{6.3.40}$$

Proof. Let $D=\{s_k\}$, where $\{s_k\}$ is an infinite non-repeating sequence of points of (a,b). Then as in the proof of Lemma 2.6.5, we can write

$$\begin{aligned}
g(x) &= g(a) + \Delta^+g(a)\,\chi_{(a,b]}(x) + \Delta^-g(b)\,\chi_{[b]}(x) \\
&\quad + \sum_{k=1}^{\infty}\Big(\Delta^+g(s_k)\,\chi_{(s_k,b]}(x) + \Delta^-g(s_k)\,\chi_{[s_k,b]}(x)\Big) \quad \text{for } x\in[a,b],
\end{aligned}$$

where the series on the right hand side is absolutely and uniformly convergent on $[a,b]$. For $x\in[a,b]$ and $n\in\mathbb{N}$ define

$$\begin{aligned}
g_n(x) &= g(a) + \Delta^+g(a)\,\chi_{(a,b]}(x) + \Delta^-g(b)\,\chi_{[b]}(x) \\
&\quad + \sum_{k=1}^{n}\Big(\Delta^+g(s_k)\,\chi_{(s_k,b]}(x) + \Delta^-g(s_k)\,\chi_{[s_k,b]}(x)\Big).
\end{aligned}$$

By Lemma 2.6.5, we have

$$\lim_{n\to\infty}\|g_n - g\|_{\mathrm{BV}} = 0.$$

By Theorem 6.3.9, this implies that the integral $\int_a^b f\,\mathrm{d}g$ exists and

$$\lim_{n\to\infty}\int_a^b f\,\mathrm{d}g_n = \int_a^b f\,\mathrm{d}g. \tag{6.3.41}$$

On the other hand, using (6.3.1), (6.3.2) and Theorem 6.1.8, we can determine the integrals on the left-hand side of (6.3.41). In particular, for each $n \in \mathbb{N}$ we have

$$\int_a^b f \, \mathrm{d}g_n = f(a)\,\Delta^+ g(a) + \sum_{k=1}^{n} f(s_k)\,\Delta g(s_k) + f(b)\,\Delta^- g(b). \tag{6.3.42}$$

Furthermore, as f is bounded, it is easy to see that the series

$$\sum_{k=1}^{\infty} f(s_k)\,\Delta g(s_k)$$

is absolutely convergent. Consequently,

$$\begin{aligned}\lim_{n\to\infty} &\Big(f(a)\,\Delta^+ g(a) + \sum_{k=1}^{n} f(s_k)\,\Delta g(s_k) + f(b)\,\Delta^- g(b)\Big)\\ &= \Big(f(a)\,\Delta^+ g(a) + \sum_{k=1}^{\infty} f(s_k)\,\Delta g(s_k) + f(b)\,\Delta^- g(b)\Big).\end{aligned}$$

Summarizing, we conclude

$$\begin{aligned}\int_a^b f \, \mathrm{d}g &= \lim_{n\to\infty} \int_a^b f \, \mathrm{d}g_n\\ &= \Big(f(a)\,\Delta^+ g(a) + \sum_{k=1}^{\infty} f(s_k)\,\Delta g(s_k) + f(b)\,\Delta^- g(b)\Big)\\ &= \Big(f(a)\,\Delta^+ g(a) + \sum_{d\in D} f(d)\,\Delta g(d) + f(b)\,\Delta^- g(b)\Big).\end{aligned}$$

This completes the proof. □

6.3.16 Corollary. *If $f \in \mathrm{G}([a,b])$, $g \in \mathrm{BV}([a,b])$, D is the set of the discontinuity points of g in (a,b) and g^{C} is the continuous part of g, then*

$$\int_a^b f \, \mathrm{d}g = \int_a^b f \, \mathrm{d}g^{\mathrm{C}} + f(a)\,\Delta^+ g(a) + \sum_{d\in D} f(d)\,\Delta g(d) + f(b)\,\Delta^- g(b). \tag{6.3.43}$$

P r o o f follows immediately from the Jordan decomposition of g (cf. Theorem 2.6.1) and Lemma 6.3.15. □

6.3.17 Lemma. *Let $h \in \mathrm{G}([a,b])$, $c \in \mathbb{R}$ and*

$$h(x) = c \quad \textit{for } x \in [a,b] \setminus D, \tag{6.3.44}$$

where the set $D \subset (a,b)$ is at most countable. Then

$$h(x-) = h(x+) = h(a+) = h(b-) = c \quad \textit{for all } x \in (a,b). \tag{6.3.45}$$

Proof. The equalities $h(t-)=c$ and $h(s+)=c$ evidently hold for all $t\in(a,b]$ and $s\in[a,b)$ which are not points of density of D. On the other hand, if $x\in(a,b]$ is density point of D, we can choose an increasing sequence $\{x_k\}\subset[a,x)\setminus D$ which tends to x. Clearly, $\lim_{k\to\infty} h(x_k)=c$. Since $h\in\mathrm{G}([a,b])$, this, due to the uniqueness of limits, means that $h(x-)=c$, as well. Similarly, the equality $h(x+)=c$ can be proved also for every density point $x\in[a,b)$ of D. □

We remark that a function h defined by the equality in (6.3.44) need not be regulated. A simple example of this fact is given by the function

$$h(t)=\begin{cases}1, & \text{if } t=a+\frac{1}{k},\ k\in\mathbb{N},\\ 0, & \text{otherwise.}\end{cases}$$

More generally, it suffices to have $D=\{d_k\}$ with

$$\lim_{k\to\infty} d_k=p\notin D \quad\text{and}\quad \lim_{k\to\infty} f(d_k)\neq c.$$

6.3.18 Lemma. *Let $h\in\mathrm{BV}([a,b])$, $c\in\mathbb{R}$ and let $D\subset(a,b)$ be an at most countable set such that* (6.3.44) *holds. Then*

$$\left.\begin{aligned}\int_a^b h\,\mathrm{d}g=c\,[g(b)-g(a)\,]+(h(a)-c)\,\Delta^+g(a)\\ +\sum_{d\in D}(h(d)-c)\,\Delta g(d)+(h(b)-c\,)\,\Delta^-g(b)\end{aligned}\right\}\tag{6.3.46}$$

holds for each function $g\in\mathrm{G}([a,b])$.

Proof. Since $h\in\mathrm{G}([a,b])$, by Lemma 6.3.17 we have (6.3.45). Hence, the function $h^{\mathrm{C}}(x)\equiv h(a)$ is the continuous part of h, $h^{\mathrm{B}}=h-h^{\mathrm{C}}$,

$$\Delta^+h(a)=c-h(a),\ \ \Delta^-h(b)=h(b)-c$$

and

$$\Delta^-h(x)=h(x)-c=-\Delta^+h(x)\quad\text{for all }\ x\in(a,b).$$

Now, for an arbitrary $g\in\mathrm{G}([a,b])$, using Corollary 6.3.14 where we replace f by h, we get

$$\begin{aligned}\int_a^b h\,\mathrm{d}g=h(a)\,[g(b)-g(a)\,]+(c-h(a))\,[g(b)-g(a+)\,]\\ +\sum_{d\in D}(h(d)-c)\,[g(b)-g(d-)-g(b)+g(d+)\,]\\ +\ (h(b)-c\,)\,\Delta^-g(b)\end{aligned}$$

$$
\begin{aligned}
&= h(a)\,[g(b)-g(a)\,] + (h(a)-c)\,[g(b)-g(a)\,] \\
&\qquad -\ (h(a)-c)\,[g(a+)-g(a)\,] \\
&\qquad + \sum_{d\in D}(h(d)-c)\,\Delta g(d) + (h(b)-c\,)\,\Delta^- g(b) \\
&= c\,[g(b)-g(a)\,] + (h(a)-c)\,\Delta^+ g(a) \\
&\qquad + \sum_{d\in D}(h(d)-c)\,\Delta g(d) + (h(b)-c\,)\,\Delta^- g(b),
\end{aligned}
$$

i.e. (6.3.46) holds. □

6.3.19 Lemma. *Let* $h\in \mathrm{G}([a,b])$, $c\in\mathbb{R}$ *and an at most countable set* $D\subset(a,b)$ *be such that* (6.3.44) *holds. Then* (6.3.46) *and*

$$
\int_a^b g\,\mathrm{d}h = g(b)\,h(b) - g(a)\,h(a) - c\,\big(g(b)-g(a)\big) \tag{6.3.47}
$$

hold for each $g\in\mathrm{BV}([a,b])$.

Proof. Assume the set D is infinite, i.e. $D=\{d_k\}$.

a) By Lemma 6.3.17 we have

$$
h(x-)=h(x+)=h(a+)=h(b-)=c \quad \text{for each } x\in(a,b)
$$

and

$$
\Delta^- h(x) = h(x)-c = -\Delta^+ h(x) \quad \text{for each } x\in(a,b).
$$

The function $h:[a,b]\to\mathbb{R}$ satisfies (6.3.44) if and only if

$$
h(x) = c + \begin{cases} h(x)-c & \text{if } x\in D', \\ 0, & \text{if } x\notin D', \end{cases}
$$

where $D'=D\cup\{a\}\cup\{b\}$. For $n\in\mathbb{N}$, set $D'_n=\{d_k\}_{k=1}^n\cup\{a\}\cup\{b\}$ and

$$
h_n(x) = c + \begin{cases} h(x)-c & \text{if } x\in D'_n, \\ 0, & \text{if } x\notin D'_n. \end{cases}
$$

Then for every $n\in\mathbb{N}$ we get $h_n(a)=h(a)$, $h_n(b)=h(b)$,

$$
\left.
\begin{aligned}
h_n(x) = c + (h(a)-c)\,\chi_{[a]}(x) + \sum_{k=1}^n (h(d_k)-c)\,\chi_{[d_k]}(x) \\
+ (h(b)-c)\,\chi_{[b]}(x) \quad \text{for } x\in[a,b]
\end{aligned}
\right\} \tag{6.3.48}
$$

and

$$|h(x) - h_n(x)| = \begin{cases} |h(x) - c|, & \text{if } \ x \in D \setminus D'_n, \\ 0, & \text{if } \ x \notin D \setminus D'_n. \end{cases}$$

Let $\varepsilon > 0$ be given. Then by Corollary 4.1.7 the set of those $k \in \mathbb{N}$ for which

$$|\Delta^- h(x)| = |\Delta^+ h(x)| = |h(d_k) - c| \geq \varepsilon$$

can have only an at most finite number of elements. Hence, there is an $n_\varepsilon \in \mathbb{N}$ such that $\widetilde{D}_\varepsilon \subset D'_n$ for $n \geq n_\varepsilon$. Hence,

$$|h(x) - h_n(x)| = 0 \quad \text{for } \ x \in \widetilde{D}_\varepsilon.$$

Obviously, we have also

$$|h(x) - h_n(x)| < \varepsilon \quad \text{if } \ x \in [a, b] \setminus \widetilde{D}_\varepsilon.$$

Thus, $\|h - h_n\| < \varepsilon$ whenever $n \geq n_\varepsilon$. In other words,

$$\lim_{n\to\infty} \|h - h_n\| = 0. \tag{6.3.49}$$

b) Now, by (6.3.48), (6.3.5) and Theorem 6.1.8 (see also Exercise 6.3.3) we check that the equalities

$$\begin{aligned} \int_a^b g \,\mathrm{d}h_n &= \sum_{k=1}^n (h(d_k) - c) \Big(\int_a^b g \,\mathrm{d}\chi_{[d_k]} \Big) \\ &= g(b)\,[h(b) - c\,] - g(a)\,[h(a) - c\,] \\ &= g(b)\,h(b) - g(a)\,h(a) - c\,[g(b) - g(a)\,] \end{aligned}$$

hold for every $n \in \mathbb{N}$. Thus, by (6.3.49) and by Corollary 6.3.12 we have

$$\int_a^b g \,\mathrm{d}h = \lim_{n\to\infty} \int_a^b g \,\mathrm{d}h_n = g(b)\,h(b) - g(a)\,h(a) - c\,[g(b) - g(a)\,],$$

i.e. (6.3.47) holds.

c) Similarly, by (6.3.48), (6.3.7)–(6.3.9) and Theorem 6.1.8 we get

$$\begin{aligned} \int_a^b h_n \,\mathrm{d}g = c\,(g(b) - g(a)) + (h(a) - c) \int_a^b \chi_{[a]} \,\mathrm{d}g \\ + \sum_{k=1}^n (h(d_k) - c) \int_a^b \chi_{[d_k]} \,\mathrm{d}g + (h(b) - c) \int_a^b \chi_{[b]} \,\mathrm{d}g \\ = c\,(g(b) - g(a)) + (h(a) - c)\,\Delta^+ g(a) \\ + \sum_{k=1}^n (h(d_k) - c)\,\Delta\, g(d_k) + (h(b) - c)\,\Delta^- g(b) \end{aligned}$$

for every $n \in \mathbb{N}$. However, by Corollary 4.1.7, by (6.3.49) and Theorem 6.3.7, we have also

$$\begin{aligned}\int_a^b h\,\mathrm{d}g &= \lim_{n\to\infty}\int_a^b h_n\,\mathrm{d}g\\ &= c\,(g(b)-g(a)) + (h(a)-c)\,\Delta^+ g(a)\\ &\qquad + \lim_{n\to\infty}\sum_{k=1}^{n}(h(d_k)-c)\,\Delta\, g(d_k) + (h(b)-c)\,\Delta^- g(b)\\ &= c\,(g(b)-g(a)) + (h(a)-c)\,\Delta^+ g(a)\\ &\qquad + \sum_{k=1}^{\infty}(h(d_k)-c)\,\Delta\, g(d_k) + (h(b)-c)\,\Delta^- g(b),\end{aligned}$$

i.e. (6.3.46) holds.

The proof in the case when D is finite is obvious. □

6.3.20 Exercise. Show that the following assertion is true.
Let $h \in \mathrm{BV}([a,b])$, $c \in \mathbb{R}$ *and an at most countable set* $D \subset (a,b)$ *be such that* (6.3.44) *holds. Then* (6.3.47) *holds for every* $g \in \mathrm{G}([a,b])$.
Hint: Use Lemma 6.3.19 and Theorem 6.3.7.

6.4 Integration by parts

The aim of this section is to prove the integration by parts theorem for KS-integrals. Before that, let us recall the convention (x) from Conventions and Notation according to which we assume

$$f(a-) = f(a) \quad \text{and} \quad f(b+) = f(b),$$

i.e.

$$\Delta^- f(a) = \Delta^+ f(b) = 0,\ \ \Delta f(a) = \Delta^+ f(a),\ \ \Delta f(b) = \Delta^- f(b) \tag{6.4.1}$$

for any $f \in \mathrm{G}([a,b])$.

First step of the proof will be the following lemma dealing with the case when one of the considered functions is a finite step function.

6.4.1 Lemma. *Let* $f : [a,b] \to \mathbb{R}$ *be a finite step function and let* $g \in \mathrm{G}([a,b])$. *Then both the integrals*

$$\int_a^b f\,\mathrm{d}g \quad \textit{and} \quad \int_a^b g\,\mathrm{d}f$$

exist and

$$\left.\begin{aligned}\int_a^b f\,\mathrm{d}g + \int_a^b g\,\mathrm{d}f = f(b)\,g(b) - f(a)\,g(a)\\ + \sum_{a\le x\le b}\Big(\Delta^- f(x)\,\Delta^- g(x) - \Delta^+ f(x)\,\Delta^+ g(x)\Big)\end{aligned}\right\}\quad(6.4.2)$$

Proof. Since g is a finite step function, there are a division $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \ldots, \alpha_m\}$ of $[a,b]$ and real numbers $c_1, \ldots, c_m \in \mathbb{R}$ such that $g(t) = c_j$ for $t \in (\alpha_{j-1}, \alpha_j)$. Using the results of Examples 6.3.1 we obtain

$$\begin{aligned}\int_a^b f\,\mathrm{d}g =& f(a)\,\Delta^+ g(a) + f(b)\,\Delta^- g(b)\\ &+ \sum_{j=1}^m c_j\,[g(\alpha_j-) - g(\alpha_{j-1}+)] + \sum_{j=1}^{m-1} f(\alpha_j)\,[g(\alpha_j+) - g(\alpha_j-)],\end{aligned}$$

and also

$$\int_a^b g\,\mathrm{d}f = \sum_{j=1}^m \big[c_j\,(g(\alpha_{j-1}) - g(\alpha_j)) + f(\alpha_j)\,g(\alpha_j) - f(\alpha_{j-1})\,g(\alpha_{j-1})\big].$$

Therefore

$$\begin{aligned}\int_a^b f\,\mathrm{d}g + \int_a^b g\,\mathrm{d}f =& f(b)\,g(b) - f(a)\,g(a) - \sum_{j=1}^m c_j\,[\Delta^+ g(\alpha_{j-1}) + \Delta^- g(\alpha_j)]\\ &+ \sum_{j=1}^m [f(\alpha_{j-1})\,\Delta^+ g(\alpha_{j-1}) + f(\alpha_j)\,\Delta^- g(\alpha_j)].\end{aligned}$$

Noting that

$$\Delta^+ f(\alpha_{j-1}) = c_j - f(\alpha_{j-1}) \quad \text{and} \quad \Delta^- f(\alpha_j) = f(\alpha_j) - c_j \quad \text{for } j \in \{1, \ldots, m\},$$

the equality (6.4.2) follows. □

6.4.2 Theorem (INTEGRATION BY PARTS)**.** *Let* $f \in \mathrm{G}([a,b])$ *and* $g \in \mathrm{BV}([a,b])$. *Then both the integrals*

$$\int_a^b f\,\mathrm{d}g \quad \textit{and} \quad \int_a^b g\,\mathrm{d}f$$

exist and (6.4.2) *holds.*

Proof. Let $\{f_n\}$ be a sequence of finite step functions which tends uniformly to f on $[a,b]$. Then by Lemma 6.4.1 we have

$$\left.\begin{aligned}&\int_a^b f_n\,\mathrm{d}g+\int_a^b g\,\mathrm{d}f_n-f_n(b)\,g(b)+f_n(a)\,g(a)\\&\qquad=\sum_{a\le x\le b}\Big(\Delta^- f_n(x)\,\Delta^- g(x)-\Delta^+ f_n(x)\,\Delta^+ g(x)\Big)\end{aligned}\right\}\tag{6.4.3}$$

for any $n\in\mathbb{N}$. By Theorems 6.3.7 and 6.3.10, the relation

$$\lim_{n\to\infty}\Big(\int_a^b f_n\,\mathrm{d}g+\int_a^b g\,\mathrm{d}f_n-f_n(b)\,g(b)+f_n(a)\,g(a)\Big)$$
$$=\int_a^b f\,\mathrm{d}g+\int_a^b g\,\mathrm{d}f-f(b)\,g(b)+f(a)\,g(a)$$

holds. Furthermore, taking into account that

$$|\Delta^+ f(t)|\le 2\,\|f\|,\qquad |\Delta^- f(t)|\le 2\,\|f\|$$

and

$$|\Delta^+(f_n-f)(t)|\le 2\,\|f_n-f\|\ \text{ and }\ |\Delta^-(f_n-f)(t)|\le 2\,\|f_n-f\|$$

for $t\in[a,b]$, we obtain the following estimates

$$\sum_{a\le t\le b}\big|\Delta^+ f(t)\,\Delta^+ g(t)-\Delta^- f(t)\,\Delta^- g(t)\big|$$
$$\le 2\,\|f\|\sum_{a\le t\le b}\Big(|\Delta^+ g(t)|+|\Delta^- g(t)|\Big)\le 2\,\|f\|\,\mathrm{var}_a^b g$$

and

$$\sum_{a\le t\le b}\big|\Delta^+(f_n-f)(t)\,\Delta^+ g(t)-\Delta^-(f_n-f)(t)\,\Delta^- g(t)\big|$$
$$\le 2\,\|f_n-f\|\sum_{a\le t\le b}\Big(|\Delta^+ g(t)|+|\Delta^- g(t)|\Big)\le 2\,\|f_n-f\|\,\mathrm{var}_a^b g.$$

Consequently, the sum

$$\sum_{a\le t\le b}\big(\Delta^+ f(t)\,\Delta^+ g(t)-\Delta^- f(t)\,\Delta^- g(t)\big)$$

is absolutely convergent and

$$\lim_{n\to\infty}\sum_{a\le x\le b}\Big(\Delta^- f_n(x)\,\Delta^- g(x)-\Delta^+ f_n(x)\,\Delta^+ g(x)\Big)$$
$$=\sum_{a\le x\le b}\Big(\Delta^- f(x)\,\Delta^- g(x)-\Delta^+ f(x)\,\Delta^+ g(x)\Big).$$

Summarizing, letting $n\to\infty$ in (6.4.3) we obtain (6.4.2). □

6.4.3 Remark. We could see that the integration by parts theorem does not hold for the KS-integral in the form we know for RS-integrals. The reason is that the domain of functions which are KS-integrable is significantly wider than that of RS-integrable functions.

6.5 The indefinite integral

The Saks-Henstock lemma is an indispensable tool in the study of deeper properties of the Kurzweil-Stieltjes integral.

6.5.1 Lemma (SAKS-HENSTOCK)**.** *Let* $f, g: [a,b] \to \mathbb{R}$ *be such that the integral* $\int_a^b f \, \mathrm{d}g$ *exists. Let* $\varepsilon > 0$ *be given and let* δ *be a gauge on* $[a,b]$ *such that*

$$\left| S(P) - \int_a^b f \, \mathrm{d}g \right| < \varepsilon \quad \textit{for all } \delta\textit{-fine partitions } P \textit{ of } [a,b].$$

If $\{([s_j, t_j], \theta_j) : j = 1, 2, \ldots, n\}$ *is an arbitrary system satisfying*

$$\left.\begin{array}{l} a \le s_1 \le \theta_1 \le t_1 \le s_2 \le \cdots \le s_n \le \theta_n \le t_n \le b, \\ [s_j, t_j] \subset (\theta_j - \delta(\theta_j), \theta_j + \delta(\theta_j)) \quad \textit{for } j = 1, \ldots, n, \end{array}\right\} \tag{6.5.1}$$

then

$$\left| \sum_{j=1}^{n} \left(f(\theta_j)\,(g(t_j) - g(s_j)) - \int_{s_j}^{t_j} f \, \mathrm{d}g \right) \right| \le \varepsilon. \tag{6.5.2}$$

Proof. Assume that the system $\Big\{([s_j, t_j], \theta_j) : j \in \{1, 2, \ldots, n\}\Big\}$ satisfies conditions (6.5.1). Set $t_0 = a$ and $s_{n+1} = b$.

Now, let $\eta > 0$ and $j \in \{0, 1, \ldots, n\}$ be given. Assume that $t_j < s_{j+1}$. Then by Remark 6.1.5, there are a gauge δ_j on $[t_j, s_{j+1}]$ and a δ_j-fine partition $P_j = (\boldsymbol{\alpha}^j, \boldsymbol{\xi}^j)$ of $[t_j, s_{j+1}]$ such that $\delta_j(x) \le \delta(x)$ for $x \in [t_j, s_{j+1}]$ and

$$\left| S(P_j) - \int_{t_j}^{s_{j+1}} f \, \mathrm{d}g \right| < \frac{\eta}{n+1}. \tag{6.5.3}$$

Now, form a δ-fine partition $Q = (\boldsymbol{\beta}, \boldsymbol{\eta})$ of the interval $[a,b]$ such that

$$S(Q) = \sum_{j=1}^{n} f(\theta_j)\,(g(t_j) - g(s_j)) + \sum_{j=0}^{n} S(P_j).$$

(If $t_j = s_{j+1}$, we set $S(P_j) = 0$.) Hence,

$$\left| \sum_{j=1}^{n} \left(f(\theta_j)\,(g(t_j) - g(s_j)) - \int_{s_j}^{t_j} f\,\mathrm{d}g \right) + \sum_{j=0}^{n} \left(S(P_j) - \int_{t_j}^{s_{j+1}} f\,\mathrm{d}g \right) \right|$$
$$= \left| S(Q) - \int_a^b f\,\mathrm{d}g \right| < \varepsilon.$$

This together with (6.5.3) yields

$$\left| \sum_{j=1}^{n} f(\theta_j)\,(g(t_j) - g(s_j)) - \int_{s_j}^{t_j} f\,\mathrm{d}g \right|$$
$$\leq \left| S(Q) - \int_a^b f\,\mathrm{d}g \right| + \left| \sum_{j=0}^{n} \left(S(P_j) - \int_{t_j}^{s_{j+1}} f\,\mathrm{d}g \right) \right| < \varepsilon + \eta.$$

Since $\eta > 0$ was arbitrary, (6.5.2) follows. □

Sometimes it is useful to have an estimate similar to (6.5.2), but with the absolute value inside the sum. Such estimate is easily obtained directly from the Saks-Henstock lemma.

6.5.2 Corollary. *Let* $f, g : [a, b] \to \mathbb{R}$ *be such that the integral* $\int_a^b f\,\mathrm{d}g$ *exists. Let* $\varepsilon > 0$ *be given and let* δ *be a gauge on* $[a, b]$ *such that*

$$\left| S(P) - \int_a^b f\,\mathrm{d}g \right| < \varepsilon \quad \text{for all } \delta\text{-fine partitions of } [a, b].$$

If $\{([s_j, t_j], \theta_j) : j = 1, 2, \ldots, n\}$ *is an arbitrary system satisfying*

$$a \leq s_1 \leq \theta_1 \leq t_1 \leq s_2 \leq \cdots \leq s_n \leq \theta_n \leq t_n \leq b,$$
$$[s_j, t_j] \subset (\theta_j - \delta(\theta_j), \theta_j + \delta(\theta_j)) \quad \text{for } j \in \{1, \ldots, n\},$$

then

$$\sum_{j=1}^{n} \left| f(\theta_j)\,(g(t_j) - g(s_j)) - \int_{s_j}^{t_j} f\,\mathrm{d}g \right| \leq 2\,\varepsilon. \tag{6.5.4}$$

Proof. Consider the sets

$$J^+ = \{j \in \{1, \ldots, n\} : f(\theta_j)\,(g(t_j) - g(s_j)) - \int_{s_j}^{t_j} f\,\mathrm{d}g > 0\},$$
$$J^- = \{j \in \{1, \ldots, n\} : f(\theta_j)\,(g(t_j) - g(s_j)) - \int_{s_j}^{t_j} f\,\mathrm{d}g < 0\}.$$

According to Lemma 6.5.1, we have

$$0 \le \sum_{j \in J^+} \left(f(\theta_j)\,(g(t_j) - g(s_j)) - \int_{s_j}^{t_j} f \, \mathrm{d}g \right) \le \varepsilon,$$

$$0 \le \sum_{j \in J^-} \left(\int_{s_j}^{t_j} f \, \mathrm{d}g - f(\theta_j)\,(g(t_j) - g(s_j)) \right) \le \varepsilon.$$

Adding these inequalities gives (6.5.4). □

6.5.3 Theorem. *Let* $\int_a^b f \, \mathrm{d}g$ *exist and let* $c \in [a,b]$. *Then*

$$\lim_{\substack{x \to c \\ x \in [a,b]}} \left(\int_a^x f \, \mathrm{d}g + f(c)\,(g(c) - g(x)) \right) = \int_a^c f \, \mathrm{d}g. \tag{6.5.5}$$

Proof. Let $\varepsilon > 0$ be given and let δ_ε be a gauge on $[a,b]$ such that

$$\left| S(P) - \int_a^b f \, \mathrm{d}g \right| < \varepsilon \quad \text{for all } \ \delta_\varepsilon - \text{fine partitions } of [a,b].$$

For each $x \in (c, c + \delta_\varepsilon(c)) \cap [a,b]$, the system

$$\{([s_1, t_1], \theta_1)\}, \ \text{ where } \ t_1 = x \ \text{ and } \ s_1 = \theta_1 = c,$$

satisfies conditions (6.5.1). Therefore, by Lemma 6.5.1, we get

$$\left| f(c)\,(g(x) - g(c)) - \int_c^x f \, \mathrm{d}g \right| \le \varepsilon. \tag{6.5.6}$$

Similarly, if $x \in (c - \delta_\varepsilon(c), c) \cap [a,b]$, then, applying Lemma 6.5.1 to the system $\{[x,c], c\}$, we get

$$\left| f(c)\,(g(c) - g(x)) - \int_x^c f \, \mathrm{d}g \right| \le \varepsilon.$$

So, inequality (6.5.6) holds for each $x \in (c - \delta_\varepsilon(c), c + \delta_\varepsilon(c)) \cap [a,b]$. Hence

$$\left| \int_a^c f \, \mathrm{d}g - \int_a^x f \, \mathrm{d}g - f(c)\,[g(c) - g(x)] \right| = \left| \int_c^x f \, \mathrm{d}g - f(c)\,[g(x) - g(c)] \right| \le \varepsilon,$$

i.e., (6.5.5) holds. □

6.5.4 Corollary. *Suppose that* $\int_a^b f \, \mathrm{d}g$ *exists,* g *is regulated on* $[a,b]$, *and let*

$$h(x) = \int_a^x f \, \mathrm{d}g \quad \textit{for} \ \ x \in [a,b].$$

Then the following statements hold:

(i) *h is regulated and satisfies*

$$h(t+) = h(t) + f(t)\,\Delta^+ g(t) \quad \textit{for } t \in [a,b),$$
$$h(t-) = h(t) - f(t)\,\Delta^- g(t) \quad \textit{for } t \in (a,b].$$

(ii) *If f is bounded and g has bounded variation, then h has bounded variation, too.*

Proof. The first assertion follows immediately from Theorem 6.5.3. To prove the second assertion, let us consider an arbitrary division $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \ldots, \alpha_m\}$ of the interval $[a,b]$. By Theorems 6.3.4 and 2.1.14 we have

$$V(h,\boldsymbol{\alpha}) = \sum_{j=1}^{m} |h(\alpha_j) - h(\alpha_{j-1})| = \sum_{j=1}^{m} \left| \int_{\alpha_{j-1}}^{\alpha_j} f\,\mathrm{d}g \right|$$
$$\leq \|f\| \sum_{j=1}^{m} \operatorname{var}_{\alpha_{j-1}}^{\alpha_j} g = \|f\| \operatorname{var}_a^b g < \infty,$$

and therefore $\operatorname{var}_a^b h < \infty$. □

6.5.5 Theorem (HAKE). (i) *Assume that $\int_a^x f\,\mathrm{d}g$ exist for each $x \in [a,b)$ and*

$$\lim_{x \to b-} \left(\int_a^x f\,\mathrm{d}g + f(b)\,[g(b) - g(x)] \right) = I \in \mathbb{R}.$$

Then $\int_a^b f\,\mathrm{d}g = I$.

(ii) *Assume that $\int_x^b f\,\mathrm{d}g$ exist for each $x \in (a,b]$ and*

$$\lim_{x \to a+} \left(\int_x^b f\,\mathrm{d}g + f(a)\,[g(x) - g(a)] \right) = I \in \mathbb{R}.$$

Then $\int_a^b f\,\mathrm{d}g = I$.

Proof. (i) a) Let $\varepsilon > 0$ be given. Choose a $\Delta > 0$ in such a way that

$$\left| \int_a^x f\,\mathrm{d}g + f(b)\,[g(b) - g(x)] - I \right| < \varepsilon \quad \text{for each } x \in [b - \Delta, b). \tag{6.5.7}$$

Set $x_k = b - \dfrac{b-a}{k+1}$ for $k \in \mathbb{N} \cup \{0\}$. Then the sequence $\{x_k\}$ is increasing, $\lim_{k\to\infty} x_k = b$ and

$$\left.\begin{array}{l} \text{for a given } k \in \mathbb{N}, \ \text{there is a gauge } \delta_k \text{ on } [a, x_k] \text{ such that} \\[2mm] \qquad \left| S(P) - \displaystyle\int_a^{x_k} f\,\mathrm{d}g \right| < \dfrac{\varepsilon}{2^k} \\[2mm] \text{for all } \delta_k\text{-fine partitions } P \text{ of } [a, x_k]. \end{array}\right\} \tag{6.5.8}$$

b) Let δ_0 be a gauge on $[a,b)$ such that

$$\delta_0(s) \le \delta_k(s) \quad \text{and} \quad [s-\delta_0(s), s+\delta_0(s)] \subset [a, x_k]$$

for every $k \in \mathbb{N}$ and $s \in [x_{k-1}, x_k)$. Furthermore, for a given $s \in [a,b)$, let $\kappa(s)$ stand for the uniquely determined natural number k such that $s \in [x_{k-1}, x_k)$.

c) We will prove that

$$\left. \begin{array}{l} \left| S(T) - \displaystyle\int_a^x f \, \mathrm{d}g \right| < \varepsilon \\ \qquad\qquad \text{for all } x \in [a,b) \text{ and all } \delta_0\text{-fine partitions } T \text{ of } [a,x]. \end{array} \right\} \tag{6.5.9}$$

To this aim, assume that $x \in [a,b)$ is given and $p = \kappa(x)$ (i.e. $x \in [x_{p-1}, x_p)$). Moreover, let $T = (\boldsymbol{\tau}, \boldsymbol{\theta})$ be an arbitrary δ_0-fine partition of $[a,x]$. Set $\nu(T) = r$. For every $k \in \mathbb{N} \cap [1,p]$ and every $j \in \mathbb{N} \cap [1,r]$ such that $\kappa(\theta_j) = k$, we have

$$\theta_j - \delta_k(\theta_j) \le \theta_j - \delta_0(\theta_j) \le \tau_{j-1} < \tau_j \le \theta_j + \delta_0(\theta_j) \le \theta_j + \delta_k(\theta_j).$$

Thanks to (6.5.8), we see that for every $k \in \{1, \dots, p\}$, the assumptions (6.5.1) of Lemma 6.5.1 are satisfied if the system $\{([s_j, t_j], \theta_j) : j = 1, \dots, n\}$ is replaced by $\{([\tau_{j-1}, \tau_j], \theta_j) : j = 1, \dots, r, \ \kappa(\theta_j) = k\}$. Therefore,

$$\left| \sum_{\kappa(\theta_j)=k} f(\theta_j) \, [g(\tau_j) - g(\tau_{j-1})] - \int_{\tau_{j-1}}^{\tau_j} f \, \mathrm{d}g \right| \le \frac{\varepsilon}{2^k} \quad \text{for each } k \in \{1, \dots, p\}.$$

Finally,

$$\begin{aligned} &\left| S(T) - \int_a^x f \, \mathrm{d}g \right| \\ &\quad = \left| \sum_{k=1}^p \sum_{\kappa(\theta_j)=k} \left(f(\theta_j) \, [g(\tau_j) - g(\tau_{j-1})] - \int_{\tau_{j-1}}^{\tau_j} f \, \mathrm{d}g \right) \right| \\ &\quad \le \sum_{k=1}^p \left| \sum_{\kappa(\theta_j)=k} \left(f(\theta_j) \, [g(\tau_j) - g(\tau_{j-1})] - \int_{\tau_{j-1}}^{\tau_j} f \, \mathrm{d}g \right) \right| < \sum_{k=1}^\infty \frac{\varepsilon}{2^k} = \varepsilon, \end{aligned}$$

i.e. (6.5.9) is true.

d) Set $\delta^*(x) = \min\{b - x, \ \delta_0(x)\}$ for $x \in [a,b)$, $\delta^*(x) = \Delta$ for $x = b$ and let $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ be an arbitrary δ^*-fine partition of $[a,b]$. Put $m = \nu(P)$. Then $\xi_m = \alpha_m = b$,

$\alpha_{m-1} \in (b-\Delta, b)$ and

$$\begin{aligned}\Big|S(P)-I\Big| &= \Big|\sum_{j=1}^{m-1} f(\xi_j)\,[g(\alpha_j)-g(\alpha_{j-1})] + f(b)\,[g(b)-g(\alpha_{m-1})] - I\Big| \\ &\le \Big|\sum_{j=1}^{m-1} f(\xi_j)\,[g(\alpha_j)-g(\alpha_{j-1})] - \int_a^{\alpha_{m-1}} f\,\mathrm{d}g\Big| \\ &\qquad + \Big|\int_a^{\alpha_{m-1}} f\,\mathrm{d}g + f(b)\,[g(b)-g(\alpha_{m-1})] - I\Big|.\end{aligned}$$

Finally, using (6.5.9) and (6.5.7) (where we set $x=\alpha_{m-1}$), we obtain

$$\Big|S(P)-I\Big| < 2\,\varepsilon, \quad \text{i.e.} \quad \int_a^b f\,\mathrm{d}g = I.$$

The proof of the second statement can be done analogously and is left as an exercise for the reader. □

6.5.6 Exercise. Prove the statement (ii) of Theorem 6.5.5 and its variant: *Assume the integral* $\int_a^b f\,\mathrm{d}g$ *exists. Let* $x \in [a,b)$ *be given and let*

$$\lim_{t\to x+}\Big(\int_a^t f\,\mathrm{d}g - f(x)\,[g(t)-g(x)]\Big) = I \in \mathbb{R}.$$

Then $\int_a^x f\,\mathrm{d}g = I$.

6.5.7 Examples. Using Hake's theorem, we can easily and in an universal way derive the formulas obtained in Examples 6.3.1 directly from the definition by using suitable choices of the gauge. E.g. the formula

$$\int_a^b f\,\mathrm{d}\chi_{[\tau,b]} = f(\tau),$$

where $\tau \in (a,b]$ and f is arbitrary, can be derived in the following way:

$$\begin{aligned}\int_a^b f\,\mathrm{d}\chi_{[\tau,b]} &= \int_a^\tau f\,\mathrm{d}\chi_{[\tau,b]} \\ &= \lim_{t\to\tau-}\Big(\int_a^t f\,\mathrm{d}\chi_{[\tau,b]} + f(\tau)\,[\chi_{[\tau,b]}(\tau)-\chi_{[\tau,b]}(t)]\Big) = f(\tau).\end{aligned}$$

Similarly, for $\tau \in [a,b)$ and $g \in \mathrm{G}([a,b])$, using Hake's theorem we get

$$\begin{aligned}\int_a^b \chi_{[a,\tau]}\,\mathrm{d}g &= \int_a^\tau 1\,\mathrm{d}g + \int_\tau^b \chi_{[a,\tau]}\,\mathrm{d}g \\ &= g(\tau)-g(a) + \lim_{t\to\tau+}\Big(\int_t^b \chi_{[a,\tau]}\,\mathrm{d}g + 1\,[g(t)-g(\tau)]\Big) \\ &= g(\tau+)-g(a),\end{aligned}$$

i.e. (6.3.8) holds.

6.5.8 Exercise. Using Hake's theorem, prove all the other formulas from Examples 6.3.1.

6.6 Substitution

The next theorem on substitution in KS-integrals will be substantially helpful in the next chapter.

6.6.1 Theorem (SUBSTITUTION THEOREM). *Let the function $g:[a,b]\to\mathbb{R}$ be given and assume that the function $h:[a,b]\to\mathbb{R}$ is bounded and such that the integral $\int_a^b f\,\mathrm{d}g$ exists. Then whenever one of the integrals*

$$\int_a^b h(x)\,\mathrm{d}\Big[\int_a^x f\,\mathrm{d}g\Big], \qquad \int_a^b h(x)\,f(x)\,\mathrm{d}[\,g(x)],$$

exists, the other one exists as well, in which case the equality

$$\int_a^b h(x)\,\mathrm{d}\Big[\int_a^x f\,\mathrm{d}g\Big] = \int_a^b h(x)\,f(x)\,\mathrm{d}[\,g(x)]$$

holds.

Proof. First, notice that by Theorem 6.1.10 the function $w(x)=\int_a^x f\,\mathrm{d}g$ is defined for every $x\in[a,b]$.

a) Assume the integral $\int_a^b hf\,\mathrm{d}g$ exists. Let $\varepsilon>0$ be given and let δ_ε be a gauge on $[a,b]$ such that the inequalities

$$\sum_{j=1}^{\nu(P)}\Big|h(\xi_j)\,f(\xi_j)\,[g(\alpha_j)-g(\alpha_{j-1})]-\int_{\alpha_{j-1}}^{\alpha_j} hf\,\mathrm{d}g\Big|<\varepsilon$$

and

$$\sum_{j=1}^{\nu(P)}\Big|f(\xi_j)\,[g(\alpha_j)-g(\alpha_{j-1})]-\int_{\alpha_{j-1}}^{\alpha_j} f\,\mathrm{d}g\Big|<\varepsilon$$

hold for all δ_ε-fine partitions $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$. (Such a gauge exists by Corollary 6.5.2.)

Let a δ_ε-fine partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ be given and $m=\nu(P)$. Then

$$\begin{aligned}
&\Big|\sum_{j=1}^m h(\xi_j)\,[w(\alpha_j)-w(\alpha_{j-1})]-\int_a^b hf\,\mathrm{d}g\Big|\\
&\quad\le\sum_{j=1}^m\Big|h(\xi_j)\int_{\alpha_{j-1}}^{\alpha_j} f\,\mathrm{d}g-h(\xi_j)\,f(\xi_j)\,[g(\alpha_j)-g(\alpha_{j-1})]\Big|\\
&\qquad+\sum_{j=1}^m\Big|h(\xi_j)\,f(\xi_j)\,[g(\alpha_j)-g(\alpha_{j-1})]-\int_{\alpha_{j-1}}^{\alpha_j} hf\,\mathrm{d}g\Big|\\
&\quad\le\|h\|\sum_{j=1}^m\Big|\int_{\alpha_{j-1}}^{\alpha_j} f\,\mathrm{d}g-f(\xi_j)\,[g(\alpha_j)-g(\alpha_{j-1})]\Big|\\
&\qquad+\sum_{j=1}^m\Big|h(\xi_j)\,f(\xi_j)\,[g(\alpha_j)-g(\alpha_{j-1})]-\int_{\alpha_{j-1}}^{\alpha_j} hf\,\mathrm{d}g\Big|\\
&\quad\le(\|h\|+1)\,\varepsilon,
\end{aligned}$$

i.e. the integral $\int_a^b h\,\mathrm{d}w$ exists and $\int_a^b h\,\mathrm{d}w=\int_a^b hf\,\mathrm{d}g$ holds.

b) The converse implication would be proved similarly, again with a substantial use of Corollary 6.5.2. □

Of course, statements analogous to substitution theorems presented for the RS-integrals in Section 5.4, hold for the KS-integral, as well. We will mention at least one of them.

6.6.2 Theorem. *Assume that the function $\phi:[c,d]\to\mathbb{R}$ is increasing and maps the interval $[c,d]$ onto the interval $[a,b]$. Furthermore, let $f:[a,b]\to\mathbb{R}$. Then, if any one of the integrals*

$$\int_a^b f(x)\,\mathrm{d}[\,g(x)],\qquad \int_c^d f(\phi(x))\,\mathrm{d}[\,g(\phi(x))]$$

exists, the other one exists as well and the equality

$$\int_c^d f(\phi(x))\,\mathrm{d}[\,g(\phi(x))]=\int_a^b f(x)\,\mathrm{d}[\,g(x)] \tag{6.6.1}$$

holds.

Proof. Notice that as ϕ is increasing and maps the interval $[c,d]$ onto the interval $[a,b]$, both ϕ and its inversion ϕ^{-1} have to be continuous.

For a given partition $Q=(\boldsymbol{\beta},\boldsymbol{\eta})$ of $[c,d]$ and $j\in\{1,\dots,\nu(Q)\}$, set

$$\alpha_j=\phi(\beta_j),\quad \xi_j=\phi(\eta_j)$$

and

$$\boldsymbol{\alpha}=\{\alpha_0,\alpha_1,\dots,\alpha_{\nu(Q)}\},\quad \boldsymbol{\xi}=\{\xi_1,\dots,\xi_{\nu(Q)}\}.$$

Then $P:=(\boldsymbol{\alpha},\boldsymbol{\xi})$ is partition of $[a,b]$ and $\nu(P)=\nu(Q)$. We write

$$P=\phi(Q) \text{ and } Q=\phi^{-1}(P).$$

Obviously, $\phi^{-1}(P)$ is a partition of $[c,d]$ for every partition P of $[a,b]$.

Further, for a given gauge $\widetilde{\delta}$ on $[c,d]$, define $\delta:[a,b]\to(0,\infty)$ in such a way that

$$\left.\begin{aligned}&\phi^{-1}(\tau+\delta(\tau))<\phi^{-1}(\tau)+\widetilde{\delta}(\phi^{-1}(\tau)) && \text{if } \tau\in[a,b)\\ \text{and}\\ &\phi^{-1}(\tau-\delta(\tau))>\phi^{-1}(\tau)-\widetilde{\delta}(\phi^{-1}(\tau)) && \text{if } \tau\in(a,b].\end{aligned}\right\}\tag{6.6.2}$$

Now, for any δ-fine partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$ we get by (6.6.2)

$$\beta_j=\phi^{-1}(\alpha_j)\le\phi^{-1}(\xi_j+\delta(\xi_j))<\phi^{-1}(\xi_j)+\widetilde{\delta}(\phi^{-1}(\xi_j))=\eta_j+\widetilde{\delta}(\eta_j)$$

and

$$\beta_{j-1}=\phi^{-1}(\alpha_{j-1})\ge\phi^{-1}(\xi_j-\delta(\xi_j))>\phi^{-1}(\xi_j)-\widetilde{\delta}(\phi^{-1}(\xi_j))=\eta_j-\widetilde{\delta}(\eta_j)$$

for all $j\in\{1,\dots,\nu(P)\}$. In other words, $\phi^{-1}(P)$ is $\widetilde{\delta}$-fine whenever P is δ-fine. Similarly, for every gauge δ on $[a,b]$ we can find a gauge $\widetilde{\delta}$ on $[c,d]$ such that $\phi(Q)$ is δ-fine as soon as Q is $\widetilde{\delta}$-fine.

Now, since the equality

$$\sum_{j=1}^{\nu(P)} f(\xi_j)\,[g(\alpha_j)-g(\alpha_{j-1})]=\sum_{j=1}^{\nu(Q)} f(\phi(\eta_j))\,[g(\phi(\beta_j))-g(\phi(\beta_{j-1}))]$$

holds for every partition P of ab and $Q=\phi^{-1}(P)$, the proof of the theorem follows. □

6.6.3 Exercises. (i) Formulate and prove an analogue of Theorem 6.6.2 for the case when ϕ is decreasing.
(ii) Formulate and prove an analogue of Theorem 5.4.6 for the KS-integral.

6.6.4 Remark. Theorem 6.6.2 can be generalized in several ways. For example, the following version of substitution theorem found its use in the application of the theory of hysteresis in economics (see [74]):

Assume that the function $f:[a,b]\to\mathbb{R}$ *is bounded on* $[a,b]$ *and regulated on* $[\alpha,b]$ *for every* $\alpha\in(a,b)$. *Moreover, let the function* $\phi:[a,b]\to\mathbb{R}$ *be nondecreasing on* $[a,b]$ *and let* $\phi(a)=c,\ \phi(b)=d$. *Finally, let the function* $g\in\mathrm{BV}([c,d])$ *be continuous from the right on* $[c,d)$. *Set*

$$\psi(s)=\inf\{t\in[a,b]:\ s\le\phi(t)\}\quad for s\in[c,d].$$

Then the relation

$$\int_\alpha^b f(t)\,\mathrm{d}[g(\phi(t))]=\int_{\phi(\alpha)}^{\phi(b)} f(\psi(s))\,\mathrm{d}[g(s)]$$

holds for every $\alpha\in[a,b]$.

6.7 Absolute integrability

The Kurzweil-Stieltjes integral is a nonabsolutely convergent integral – the existence of $\int_a^b f\,\mathrm{d}g$ does not necessarily imply the existence of $\int_a^b |f|\,\mathrm{d}g$. In this section, we collect some sufficient and necessary conditions for the existence of the latter integral.

We restrict our attention to the case when g is nondecreasing. In this situation, if both $\int_a^b f\,\mathrm{d}g$ and $\int_a^b |f|\,\mathrm{d}g$ exist, we have the inequality

$$\left|\int_a^b f\,\mathrm{d}g\right|\le\int_a^b |f|\,\mathrm{d}g.$$

This fact follows immediately from the definition of the integral, since

$$|S(f,\mathrm{d}g,P)|\le S(|f|,\mathrm{d}g,P)\quad\text{for each partition } P \text{ of } [a,b].$$

6.7.1 Theorem. *Assume that* $g:[a,b]\to\mathbb{R}$ *is nondecreasing and* $\int_a^b f\,\mathrm{d}g$ *exists. Let*

$$F(x)=\int_a^x f\,\mathrm{d}g\quad\textit{for}\ x\in[a,b].$$

Then $\int_a^b |f|\,\mathrm{d}g$ *exists if and only if* F *has bounded variation on* $[a,b]$. *In this case, we have*

$$\int_a^b |f|\,\mathrm{d}g=\mathrm{var}_a^b F.$$

Proof. Suppose that $\int_a^b |f|\,\mathrm{d}g$ exists. If $\boldsymbol{\alpha}$ is an arbitrary division of $[a,b]$, then

$$V(F,\boldsymbol{\alpha}) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} |F(\alpha_j) - F(\alpha_{j-1})| = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \left| \int_{\alpha_{j-1}}^{\alpha_j} f\,\mathrm{d}g \right|$$
$$\le \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \int_{\alpha_{j-1}}^{\alpha_j} |f|\,\mathrm{d}g = \int_a^b |f|\,\mathrm{d}g,$$

and therefore $\operatorname{var}_a^b F$ is finite.

Conversely, suppose that F has bounded variation on $[a,b]$. Consider an arbitrary $\varepsilon > 0$ and let $\boldsymbol{\alpha}$ be a division of $[a,b]$ satisfying

$$\operatorname{var}_a^b F - \varepsilon \le V(F,\boldsymbol{\alpha}) \le \operatorname{var}_a^b F.$$

Observe that if $\boldsymbol{\beta}$ is a refinement of $\boldsymbol{\alpha}$, then

$$\operatorname{var}_a^b F - \varepsilon \le V(F,\boldsymbol{\alpha}) \le V(F,\boldsymbol{\beta}) \le \operatorname{var}_a^b F. \tag{6.7.1}$$

Let δ be a gauge on $[a,b]$ with the following properties:

- If P is a δ-fine partition of $[a,b]$, then $\left|\int_a^b f\,\mathrm{d}g - S(f,\mathrm{d}g,P)\right| < \varepsilon$.
- If $i \in \{1,\ldots,\nu(\boldsymbol{\alpha})\}$ and $t \in (\alpha_{i-1},\alpha_i)$, then
$$(t-\delta(t), t+\delta(t)) \subset (\alpha_{i-1},\alpha_i).$$
- If $i \in \{0,\ldots,\nu(\boldsymbol{\alpha})\}$, then
$$(\alpha_i - \delta(\alpha_i), \alpha_i + \delta(\alpha_i)) \subset (\alpha_{i-1},\alpha_{i+1})$$
with the convention that $\alpha_{-1} = -\infty$ and $\alpha_{\nu(\boldsymbol{\alpha})+1} = \infty$.

Let $P = (\boldsymbol{\beta},\boldsymbol{\xi})$ be an arbitrary δ-fine partition of $[a,b]$. The last two properties of δ ensure that each interval $[\beta_{i-1},\beta_i]$ is either contained in a single interval $[\alpha_{j-1},\alpha_j]$, or $[\beta_{i-1},\beta_i] \subset [\alpha_{j-1},\alpha_{j+1}]$ and $\xi_i = \alpha_j$. By splitting all intervals of the second type in two subintervals $[\beta_{i-1},\xi_i]$ and $[\xi_i,\beta_i]$ that share the same tag ξ_i, we can obtain a new partition P' such that $S(|f|,\mathrm{d}g,P) = S(|f|,\mathrm{d}g,P')$. Thus, without loss of generality, we can assume that $\boldsymbol{\beta}$ is a refinement of $\boldsymbol{\alpha}$.

We now use the fact that g is nondecreasing, the reverse triangle inequality $||x|-|y|| \le |x-y|$, and finally Corollary 6.5.2 to obtain

$$\left|S(|f|,\mathrm{d}g,P) - V(F,\boldsymbol{\beta})\right| = \left| \sum_{i=1}^{\nu(\boldsymbol{\beta})} \left(|f(\xi_i)\,(g(\beta_i)-g(\beta_{i-1}))| - \left| \int_{\beta_{i-1}}^{\beta_i} f\,\mathrm{d}g \right| \right) \right|$$
$$\le \sum_{i=1}^{\nu(\boldsymbol{\beta})} \left| f(\xi_i)\,(g(\beta_i)-g(\beta_{i-1})) - \int_{\beta_{i-1}}^{\beta_i} f\,\mathrm{d}g \right| \le 2\varepsilon.$$

By combining this estimate with (6.7.1), we conclude that

$$\left|S(|f|, \mathrm{d}g, P) - \operatorname{var}_a^b F\right| \le |S(|f|, \mathrm{d}g, P) - V(F, \boldsymbol{\beta})| + \left|V(F, \boldsymbol{\beta}) - \operatorname{var}_a^b F\right| \le 3\varepsilon,$$

which means that $\int_a^b |f| \, \mathrm{d}g$ exists and equals $\operatorname{var}_a^b F$. □

6.7.2 Example. Let $\{t_n\}$ be an increasing sequence of points from the open interval $(0,1)$ tending to $\tau \in (0,1]$. Consider the function $g:[0,1] \to \mathbb{R}$ given by

$$g(t) = \sum_{n=1}^{\infty} \chi_{[t_n,1]}(t) \, \frac{1}{n^2}. \tag{6.7.2}$$

It is not hard to see that g is nondecreasing and right-continuous on $[0,1)$. Let $f:[0,1] \to \mathbb{R}$ be defined as

$$f(t) = \begin{cases} (-1)^{n+1} \, n & \text{if } \ t = t_n \ \text{ for some } \ n \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Claim 1. The integral $\int_0^1 f \, \mathrm{d}g$ exists.

Note that, for each $t \in [0,\tau)$, the integral $\int_0^t f \, \mathrm{d}g$ exists (due to Corollary 6.3.2 and using the fact that the restriction of g to the interval $[0,t]$ defines a finite step function). Moreover, by Example 6.3.1, we have

$$\int_0^{t_1} f \, \mathrm{d}g = f(t_1) \, g(t_1) = 1,$$

and

$$\begin{aligned} \int_{t_n}^{t_{n+1}} f \, \mathrm{d}g &= \int_{t_n}^{t_{n+1}} f \, \mathrm{d}\big[g(t_n)\chi_{[t_n,t_{n+1})} + g(t_{n+1})\chi_{[t_{n+1}]}\big] \\ &= f(t_{n+1}) \, [g(t_{n+1}) - g(t_n)] = \frac{(-1)^{n+2}}{n+1}, \end{aligned}$$

for each $n \in \mathbb{N}$. Since

$$\lim_{t \to \tau-} \int_0^t f \, \mathrm{d}g + f(\tau)\Delta^+ g(\tau) = \lim_{n\to\infty} \int_0^{t_n} f \, \mathrm{d}g = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \ln \, 2,$$

applying Hake's Theorem (Theorem 6.5.5), we conclude that the integral $\int_0^\tau f \, \mathrm{d}g$ exists and equals $\ln 2$. If $\tau = 1$, the claim is proved. If $\tau < 1$, then $\int_\tau^1 f \, \mathrm{d}g = 0$ (because g is constant on $[\tau,1]$), wherefrom the proof of the claim follows.

Claim 2. $|f|$ is not integrable with respect to g on $[0,1]$.

Let $F(t) = \int_0^t f \, \mathrm{d}g$ for $t \in [0,1]$. We will show that $\mathrm{var}_0^1 F = \infty$. For every $n \in \mathbb{N}$ consider the division $D_n = \{0, t_1, t_2, \dots, t_n, 1\}$ of $[0,1]$. We have

$$V(F, D_n) = \sum_{k=1}^{n-1} \left| \int_{t_k}^{t_{k+1}} f \, \mathrm{d}g \right| + \left| \int_0^{t_1} f \, \mathrm{d}g \right| + \left| \int_{t_n}^1 f \, \mathrm{d}g \right|$$
$$= \sum_{k=1}^{n} \frac{1}{k} + \left| \int_{t_n}^1 f \, \mathrm{d}g \right|.$$

Since $V(F, D_n) \leq \mathrm{var}_0^1 F$ for every $n \in \mathbb{N}$, $\mathrm{var}_0^1 F$ cannot be finite. Thus the claim is a consequence of Theorem 6.7.1.

6.7.3 Exercises. Let $\{t_n\}$ be an increasing sequence of points from the open interval (a,b) tending to $\tau \in (a,b]$. Let $\{c_n\}$ be a sequence of nonnegative numbers such that the series $\sum c_n$ converges. Define a function $g : [a,b] \to \mathbb{R}$ by

$$g(t) = \sum_{n : t_n \leq t} c_n.$$

For a given $f : [a,b] \to \mathbb{R}$ prove that:

(i) The integral $\int_a^b f \, \mathrm{d}g$ exists if and only if the series $\sum c_n f(t_n)$ converges and in such a case it is given by

$$\int_a^b f \, \mathrm{d}g = \sum_{n=1}^{\infty} c_n f(t_n).$$

(ii) The integral $\int_a^b |f| \, \mathrm{d}g$ exists if and only if the series $\sum c_n f(t_n)$ is absolutely convergent.

6.7.4 Theorem. *Assume that* $g : [a,b] \to \mathbb{R}$ *is nondecreasing,* $f, h : [a,b] \to \mathbb{R}$ *are such that* $\int_a^b f \, \mathrm{d}g$ *and* $\int_a^b h \, \mathrm{d}g$ *exist and* $|f(t)| \leq h(t)$ *for each* $t \in [a,b]$. *Then* $\int_a^b |f| \, \mathrm{d}g$ *exists.*

Proof. If $[c,d] \subset [a,b]$ and P is a partition of $[c,d]$, then

$$|S(f, \mathrm{d}g, P)| \leq S(h, \mathrm{d}g, P).$$

Thus, it follows from the definition of the integral that

$$\left| \int_c^d f \, \mathrm{d}g \right| \leq \int_c^d h \, \mathrm{d}g \quad \text{for each } [c,d] \subset [a,b].$$

Let

$$F(x) = \int_a^x f\,\mathrm{d}g \quad \text{for } x \in [a,b].$$

If $\boldsymbol{\alpha}$ is an arbitrary division of $[a,b]$, we obtain the estimate

$$V(F,\alpha) = \sum_{i=1}^{\nu(\boldsymbol{\alpha})} |F(\alpha_i) - F(\alpha_{i-1})| = \sum_{i=1}^{\nu(\boldsymbol{\alpha})} \left| \int_{\alpha_{i-1}}^{\alpha_i} f\,\mathrm{d}g \right| \le \sum_{i=1}^{\nu(\boldsymbol{\alpha})} \int_{\alpha_{i-1}}^{\alpha_i} h\,\mathrm{d}g = \int_a^b h\,\mathrm{d}g,$$

which shows that F has bounded variation on $[a,b]$. Thus, the existence of the integral $\int_a^b |f|\,\mathrm{d}g$ follows from Theorem 6.7.1. □

6.7.5 Theorem. *If $g:[a,b]\to\mathbb{R}$ is nondecreasing and $f_1, f_2:[a,b]\to\mathbb{R}$ are such that the integrals $\int_a^b f_1\,\mathrm{d}g$, $\int_a^b f_2\,\mathrm{d}g$, $\int_a^b |f_1|\,\mathrm{d}g$ and $\int_a^b |f_2|\,\mathrm{d}g$ exist, then $\int_a^b \max(f_1,f_2)\,\mathrm{d}g$ and $\int_a^b \min(f_1,f_2)\,\mathrm{d}g$ exist as well.*

Proof. By Theorem 6.7.4, the integral $\int_a^b |f_1 - f_2|\,\mathrm{d}g$ exists, because

$$|f_1 - f_2| \le |f_1| + |f_2|$$

and the integrals $\int_a^b (f_1 - f_2)\,\mathrm{d}g$, $\int_a^b |f_1|\,\mathrm{d}g$ and $\int_a^b |f_2|\,\mathrm{d}g$ exist. Since

$$\max(f_1,f_2) = \frac{f_1 + f_2 + |f_1 - f_2|}{2} \qquad \text{and} \qquad \min(f_1,f_2) = \frac{f_1 + f_2 - |f_1 - f_2|}{2},$$

the existence of $\int_a^b \max(f_1,f_2)\,\mathrm{d}g$ and $\int_a^b \min(f_1,f_2)\,\mathrm{d}g$ follows from the linearity of the integral (i.e., from Theorem 6.1.8). □

6.8 Convergence theorems

In this section, we present several convergence theorems for the Kurzweil-Stieltjes integral that do not require uniform convergence.

We start by introducing the concept of uniform integrability (also known as equiintegrability).

6.8.1 Definition. Consider a sequence of functions $f_n:[a,b]\to\mathbb{R}$, $n\in\mathbb{N}$, and a function $g:[a,b]\to\mathbb{R}$. Then $\{f_n\}$ is called uniformly integrable with respect to g, if the following two conditions are satisfied:

- The integral $\displaystyle\int_a^b f_n\,\mathrm{d}g$ exists for each $n\in\mathbb{N}$.

- For every $\varepsilon > 0$, there exists a gauge δ on $[a,b]$ such that the inequality
$$\left|\int_a^b f_n \,\mathrm{d}g - S(f_n, \mathrm{d}g, P)\right| < \varepsilon$$
holds for each δ-fine partition P of $[a,b]$ and for every $n \in \mathbb{N}$.

The next result is a basic convergence theorem for the Kurzweil-Stieltjes integral. Although the assumption of uniform integrability might be difficult to verify, the result will play a key role in deriving the other convergence theorems given later in this section.

6.8.2 Theorem. *If $\{f_n\}$ is uniformly integrable with respect to g and*
$$\lim_{n\to\infty} f_n(t) = f(t) \quad \textit{for all } t \in [a,b],$$
then both the integral $\int_a^b f \,\mathrm{d}g$ and the limit $\lim_{n\to\infty}\int_a^b f_n \,\mathrm{d}g$ exist, and the equality
$$\int_a^b f \,\mathrm{d}g = \lim_{n\to\infty}\int_a^b f_n \,\mathrm{d}g$$
is true.

Moreover,
$$\lim_{n\to\infty}\left(\sup_{t\in[a,b]}\left|\int_a^t f_n \,\mathrm{d}g - \int_a^t f \,\mathrm{d}g\right|\right) = 0 \tag{6.8.1}$$
holds whenever g is bounded on $[a,b]$.

Proof. a) Consider an arbitrary $\varepsilon > 0$ and let δ be the corresponding gauge from Definition 6.8.1. Choose an arbitrary δ-fine partition P of $[a,b]$. Since
$$\lim_{n\to\infty} S(f_n, \mathrm{d}g, P) = S(f, \mathrm{d}g, P),$$
there exists an $n_0 \in \mathbb{N}$ such that
$$|S(f_m, \mathrm{d}g, P) - S(f_n, \mathrm{d}g, P)| < \varepsilon$$
holds for all $m, n \ge n_0$. Using this fact together with the definition of uniform integrability, we get
$$\begin{aligned}&\left|\int_a^b f_m \,\mathrm{d}g - \int_a^b f_n \,\mathrm{d}g\right| \\ &\quad\le \left|\int_a^b f_m \,\mathrm{d}g - S(f_m, \mathrm{d}g, P)\right| + |S(f_m, \mathrm{d}g, P) - S(f_n, \mathrm{d}g, P)| \\ &\qquad + \left|S(f_n, \mathrm{d}g, P) - \int_a^b f_n \,\mathrm{d}g\right| < 3\,\varepsilon\end{aligned}$$

for all $m, n \geq n_0$. In particular, $\{\int_a^b f_n \, dg\}$ is a Cauchy sequence and thus it has a finite limit:

$$\lim_{n\to\infty} \int_a^b f_n \, dg = L \in \mathbb{R}.$$

Now, let P be an arbitrary δ-fine partition of $[a,b]$. Choose an $n_1 \in \mathbb{N}$ such that

$$|S(f_{n_1}, dg, P) - S(f, dg, P)| < \varepsilon \quad \text{and} \quad \left| \int_a^b f_{n_1} \, dg - L \right| < \varepsilon.$$

Then

$$\begin{aligned} |S(f, dg, P) - L| &\leq |S(f, dg, P) - S(f_{n_1}, dg, P)| \\ &\quad + \left| S(f_{n_1}, dg, P) - \int_a^b f_{n_1} \, dg \right| + \left| \int_a^b f_{n_1} \, dg - L \right| < 3\varepsilon. \end{aligned}$$

It follows that $\int_a^b f \, dg$ exists and equals L.

b) To prove (6.8.1), let $h_n(t) = f_n(t) - f(t)$ for $n \in \mathbb{N}$ and $t \in [a,b]$, and assume that g is bounded on $[a,b]$. Note that the sequence $\{h_n\}$ tends pointwise on $[a,b]$ to 0 and it is uniformly integrable with respect to g. Choose an arbitrary $\varepsilon > 0$ and find a gauge δ on $[a,b]$ such that

$$\left| \int_a^b h_n \, dg - S(h_n, dg, P) \right| < \varepsilon \tag{6.8.2}$$

for each δ-fine partition P of $[a,b]$. Let $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ be such a partition. Since $h_n(t) \to 0$ for $t \in [a,b]$ and g is bounded, there exists an $n_0 \in \mathbb{N}$ such that

$$|h_n(\xi_i)| \, \|g\| < \frac{\varepsilon}{2\,\nu(\boldsymbol{\alpha})} \quad \text{for all } n \geq n_0 \text{ and } i \in \{1, \dots, \nu(\boldsymbol{\alpha})\}. \tag{6.8.3}$$

Let arbitrary $t \in [a,b]$ and $n \in \mathbb{N} \cap [n_0, \infty)$ be given and let $j \in \{1, \dots, \nu(\boldsymbol{\alpha})\}$ be such that $t \in [\alpha_{j-1}, \alpha_j]$. Then,

$$\begin{aligned} \left| \int_a^t h_n \, dg \right| &= \left| \sum_{i=1}^{j-1} \int_{\alpha_{i-1}}^{\alpha_i} h_n \, dg + \int_{\alpha_{j-1}}^t h_n \, dg \right| \\ &\leq \left| \sum_{i=1}^{j-1} \left(\int_{\alpha_{i-1}}^{\alpha_i} h_n \, dg - h_n(\xi_i) \, (g(\alpha_i) - g(\alpha_{i-1})) \right) \right. \\ &\quad \left. + \int_{\alpha_{j-1}}^t h_n \, dg - h_n(\xi_j) \, (g(t) - g(\alpha_{j-1})) \right| \\ &\quad + \sum_{i=1}^{j-1} |h_n(\xi_i)| \, |g(\alpha_i) - g(\alpha_{i-1})| + |h_n(\xi_j)| \, |g(t) - g(\alpha_{j-1})|. \end{aligned}$$

Due to (6.8.3) we have

$$\sum_{i=1}^{j-1} |h_n(\xi_i)|\,|g(\alpha_i)-g(\alpha_{i-1})| + |h_n(\xi_j)|\,|g(t)-g(\alpha_{j-1})| \\ \le \sum_{i=1}^{j} 2\,|h_n(\xi_i)|\,\|g\| < \varepsilon.$$

Furthermore, if $t \ge \xi_j$, then, having in mind (6.8.2), we may apply the Saks-Henstock lemma to the system

$$\{([\alpha_{i-1},\alpha_i],\xi_i) : i=1,\dots,j-1\} \cup \{([\alpha_{j-1},t],\xi_j)\}$$

and deduce that

$$\bigg| \sum_{i=1}^{j-1} \left(\int_{\alpha_{i-1}}^{\alpha_i} h_n\,\mathrm{d}g - h_n(\xi_i)\,(g(\alpha_i)-g(\alpha_{i-1})) \right) \\ + \int_{\alpha_{j-1}}^{t} h_n\,\mathrm{d}g - h_n(\xi_j)\,(g(t)-g(\alpha_{j-1})) \bigg| \le \varepsilon.$$

On the other hand, if $t < \xi_j$, then applying the Saks-Henstock lemma to the systems

$$\big\{([\alpha_{i-1},\alpha_i],\xi_i) : i=1,\dots,j\big\} \ \text{ and } \ \big\{([t,\alpha_j],\xi_j)\big\},$$

we get

$$\bigg| \sum_{i=1}^{j-1} \left(\int_{\alpha_{i-1}}^{\alpha_i} h_n\,\mathrm{d}g - h_n(\xi_i)\,(g(\alpha_i)-g(\alpha_{i-1})) \right) \\ + \int_{\alpha_{j-1}}^{t} h_n\,\mathrm{d}g - h_n(\xi_j)\,(g(t)-g(\alpha_{j-1})) \bigg| \\ = \bigg| \sum_{i=1}^{j} \left(\int_{\alpha_{i-1}}^{\alpha_i} h_n\,\mathrm{d}g - h_n(\xi_i)\,(g(\alpha_i)-g(\alpha_{i-1})) \right) \\ - \left(\int_{t}^{\alpha_j} h_n\,\mathrm{d}g - h_n(\xi_j)\,(g(\alpha_j)-g(t)) \right) \bigg| \le 2\,\varepsilon.$$

To summarize, we have shown that

$$\left| \int_a^t h_n\,\mathrm{d}g \right| \le 3\,\varepsilon \quad \text{for all } n \ge n_0 \ \text{ and } \ t \in [a,b],$$

which completes the proof of (6.8.1). □

The next example shows that the boundedness of the integrator g is essential to ensure the uniform convergence of the indefinite integrals in (6.8.1).

6.8.3 Example. Let $f_n(t)=1/n$ for $t\in[a,b]$ and let $g:[a,b]\to\mathbb{R}$ be arbitrary with $g(a)=0$. Then $\{f_n\}$ tends pointwise on $[a,b]$ to the zero function. Furthermore, as $S(f_n,\mathrm{d}g,P)=(1/n)\,g(b)$ for each $n\in\mathbb{N}$ and each partition P of $[a,b]$, we see that $\int_a^b f_n\,\mathrm{d}g=(1/n)\,g(b)$ for each $n\in\mathbb{N}$ and the sequence $\{f_n\}$ is uniformly integrable with respect to g. However, we claim that the indefinite integrals $F_n(t)=\int_a^t f_n\,\mathrm{d}g$ tend uniformly to the zero function if and only if g is bounded. Indeed, we have $F_n(t)=(1/n)\,g(t)$ for $t\in[a,b]$. Clearly, if g is bounded on $[a,b]$, then $\{F_n\}$ tends uniformly on $[a,b]$ to the zero function. On the other hand, if g is unbounded and an arbitrary $\varepsilon>0$ is given, then for each $n\in\mathbb{N}$, there is a $t\in[a,b]$ such that

$$|F_n(t)|=|(1/n)\,g(t)|>\varepsilon,$$

i.e., F_n does not converge uniformly to the zero function.

The remaining convergence theorems in this section provide more transparent conditions that imply uniform integrability. To proceed, we need the following auxiliary lemma.

6.8.4 Lemma. *For each $\ell\in\mathbb{N}$, let $\mathcal{S}_\ell$ be the set of all couples $(\boldsymbol{\sigma},\boldsymbol{s})$, where $\boldsymbol{\sigma}$ is a division of $[a,b]$ and $\boldsymbol{s}=\{s_1,\dots,s_{\nu(\boldsymbol{\sigma})}\}$ is a finite sequence of integers greater than or equal to ℓ.*

Let $g:[a,b]\to\mathbb{R}$ and let $f_n:[a,b]\to\mathbb{R}$, $n\in\mathbb{N}$, be a sequence of functions such that the integral $\int_a^b f_n\,\mathrm{d}g$ exists for each $n\in\mathbb{N}$. Moreover, assume that there are $B,C\in\mathbb{R}$ such that if $\boldsymbol{\sigma}$ is a division of $[a,b]$ and $s_1,\dots,s_{\nu(\boldsymbol{\sigma})}\in\mathbb{N}$, then

$$B\le\sum_{j=1}^{\nu(\boldsymbol{\sigma})}\int_{\sigma_{j-1}}^{\sigma_j} f_{s_j}\,\mathrm{d}g\le C. \tag{6.8.4}$$

Then for each $\varepsilon>0$ and $\ell\in\mathbb{N}$ there exist $(\boldsymbol{\pi}^\ell,\boldsymbol{p}^\ell)$, $(\boldsymbol{\rho}^\ell,\boldsymbol{r}^\ell)\in\mathcal{S}_\ell$ such that the following statements are true:

(i) *The inequalities*

$$\sum_{j=1}^{\nu(\boldsymbol{\pi}^\ell)}\int_{\pi_{j-1}^\ell}^{\pi_j^\ell} f_{p_j^\ell}\,\mathrm{d}g-\frac{\varepsilon}{2^\ell}<\sum_{j=1}^{\nu(\boldsymbol{\sigma})}\int_{\sigma_{j-1}}^{\sigma_j} f_{s_j}\,\mathrm{d}g<\sum_{j=1}^{\nu(\boldsymbol{\rho}^\ell)}\int_{\rho_{j-1}^\ell}^{\rho_j^\ell} f_{r_j^\ell}\,\mathrm{d}g+\frac{\varepsilon}{2^\ell} \tag{6.8.5}$$

hold for each $(\boldsymbol{\sigma},\boldsymbol{s})\in\mathcal{S}_\ell$.

(ii) *Assume that* $\mathcal{I} = \bigcup_{i \in I} [c_i, d_i]$ *is a finite union of nonoverlapping intervals in* $[a,b]$ *such that for each* $i \in I$ *there is a* $j(i) \in \{1, \dots, \nu(\boldsymbol{\pi}^\ell)\}$ *satisfying* $[c_i, d_i] \subset [\pi^\ell_{j(i)-1}, \pi^\ell_{j(i)}]$. *Then the inequality*

$$\sum_{i \in I} \int_{c_i}^{d_i} f_{p^\ell_{j(i)}} \, \mathrm{d}g - \frac{\varepsilon}{2^\ell} < \sum_{i \in I} \int_{c_i}^{d_i} f_n \, \mathrm{d}g \tag{6.8.6}$$

holds for each $n \geq \ell$.

(iii) *Assume that* $\mathcal{I} = \bigcup_{i \in I} [c_i, d_i]$ *is a finite union of nonoverlapping intervals in* $[a,b]$ *such that for each* $i \in I$ *there exists a* $j(i) \in \{1, \dots, \nu(\boldsymbol{\rho}^\ell)\}$ *satisfying* $[c_i, d_i] \subset [\rho^\ell_{j(i)-1}, \rho^\ell_{j(i)}]$. *Then the inequality*

$$\sum_{i \in I} \int_{c_i}^{d_i} f_n \, \mathrm{d}g < \sum_{i \in I} \int_{c_i}^{d_i} f_{r^\ell_{j(i)}} \, \mathrm{d}g + \frac{\varepsilon}{2^\ell} \tag{6.8.7}$$

holds for each $n \geq \ell$.

Proof. According to (6.8.4), the values of all sums

$$\sum_{j=1}^{\nu(\boldsymbol{\sigma})} \int_{\sigma_{j-1}}^{\sigma_j} f_{s_j} \, \mathrm{d}g, \quad \text{where } (\boldsymbol{\sigma}, \boldsymbol{s}) \in \mathcal{S}_\ell,$$

are contained in a bounded subset of $\mathbb{R}$. Thus, the existence of the couples $(\boldsymbol{\pi}^\ell, \boldsymbol{p}^\ell)$ and $(\boldsymbol{\rho}^\ell, \boldsymbol{r}^\ell) \in \mathcal{S}_\ell$ having the properties from statement (i) follows from the definitions of infimum and supremum.

To prove statement (ii), consider the division $\boldsymbol{\sigma} = \boldsymbol{\pi}^\ell \cup \{c_i, d_i : i \in I\}$ of $[a,b]$. For each $j \in \{1, \dots, \nu(\boldsymbol{\sigma})\}$, there exists a $k(j) \in \{1, \dots, \nu(\boldsymbol{\pi}^\ell)\}$ such that $[\sigma_{j-1}, \sigma_j] \subset [\pi^\ell_{k(j)-1}, \pi^\ell_{k(j)}]$. Define $\boldsymbol{s}$ as follows: If $j \in \{1, \dots, \nu(\boldsymbol{\sigma})\}$ is such that $[\sigma_{j-1}, \sigma_j] \subset \mathcal{I}$, let $s_j = n$. Otherwise, let $s_j = p^\ell_{k(j)}$. Then $(\boldsymbol{\sigma}, \boldsymbol{s}) \in \mathcal{S}_\ell$, and relation (6.8.5) holds. Obviously,

$$\begin{aligned}
\sum_{j=1}^{\nu(\boldsymbol{\sigma})} \int_{\sigma_{j-1}}^{\sigma_j} f_{s_j} \, \mathrm{d}g &= \sum_{\substack{j \in \{1, \dots, \nu(\boldsymbol{\sigma})\} \\ [\sigma_{j-1}, \sigma_j] \subset \mathcal{I}}} \int_{\sigma_{j-1}}^{\sigma_j} f_n \, \mathrm{d}g + \sum_{\substack{j \in \{1, \dots, \nu(\boldsymbol{\sigma})\} \\ [\sigma_{j-1}, \sigma_j] \not\subset \mathcal{I}}} \int_{\sigma_{j-1}}^{\sigma_j} f_{p^\ell_{k(j)}} \, \mathrm{d}g \\
&= \sum_{i \in I} \int_{c_i}^{d_i} f_n \, \mathrm{d}g + \sum_{\substack{j \in \{1, \dots, \nu(\boldsymbol{\sigma})\} \\ [\sigma_{j-1}, \sigma_j] \not\subset \mathcal{I}}} \int_{\sigma_{j-1}}^{\sigma_j} f_{p^\ell_{k(j)}} \, \mathrm{d}g
\end{aligned}$$

Furthermore, for each $m \in \{1, \dots, \nu(\boldsymbol{\pi}^\ell)\}$, we have

$$[\pi^\ell_{m-1}, \pi^\ell_m] = \bigcup_{\substack{j \in \{1, \dots, \nu(\boldsymbol{\sigma})\} \\ k(j) = m}} [\sigma_{j-1}, \sigma_j],$$

and therefore

$$\begin{aligned}\sum_{j=1}^{\nu(\boldsymbol{\pi}^\ell)}\int_{\pi^\ell_{j-1}}^{\pi^\ell_j} f_{p^\ell_j}\,\mathrm{d}g &= \sum_{m=1}^{\nu(\boldsymbol{\pi}^\ell)}\sum_{\substack{j\in\{1,\dots,\nu(\boldsymbol{\sigma})\}\\ k(j)=m}}\int_{\sigma_{j-1}}^{\sigma_j} f_{p^\ell_m}\,\mathrm{d}g = \sum_{j=1}^{\nu(\sigma)}\int_{\sigma_{j-1}}^{\sigma_j} f_{p^\ell_{k(j)}}\,\mathrm{d}g\\
&= \sum_{\substack{j\in\{1,\dots,\nu(\boldsymbol{\sigma})\}\\ [\sigma_{j-1},\sigma_j]\subset\mathcal{I}}}\int_{\sigma_{j-1}}^{\sigma_j} f_{p^\ell_{k(j)}}\,\mathrm{d}g + \sum_{\substack{j\in\{1,\dots,\nu(\boldsymbol{\sigma})\}\\ [\sigma_{j-1},\sigma_j]\not\subset\mathcal{I}}}\int_{\sigma_{j-1}}^{\sigma_j} f_{p^\ell_{k(j)}}\,\mathrm{d}g\\
&= \sum_{i\in I}\int_{c_i}^{d_i} f_{p^\ell_{j(i)}}\,\mathrm{d}g + \sum_{\substack{j\in\{1,\dots,\nu(\boldsymbol{\sigma})\}\\ [\sigma_{j-1},\sigma_j]\not\subset\mathcal{I}}}\int_{\sigma_{j-1}}^{\sigma_j} f_{p^\ell_{k(j)}}\,\mathrm{d}g.\end{aligned}$$

Now, by subtracting

$$\sum_{\substack{j\in\{1,\dots,\nu(\boldsymbol{\sigma})\}\\ [\sigma_{j-1},\sigma_j]\not\subset\mathcal{I}}}\int_{\sigma_{j-1}}^{\sigma_j} f_{p^\ell_{k(j)}}\,\mathrm{d}g$$

from the first inequality in (6.8.5), we get (6.8.6).

Statement (iii) can be proved in a similar way. □

The following convergence theorem is due to D. Preiss and Š. Schwabik. It is a special case of Theorem 5.5 from J. Kurzweil's book [81], which is concerned with a more general type of integral; see also Theorem 1.28 and Remark 1.30 in [122].

6.8.5 Theorem. *Let $g\in\mathrm{BV}([a,b])$ and let $\{f_n\}$ be a sequence of real valued functions defined on $[a,b]$ and satisfying the following conditions:*

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

(ii) $\lim\limits_{n\to\infty} f_n(t)=f(t)$ *for $t\in[a,b]$.*

(iii) *There are $B,C\in\mathbb{R}$ such that the inequalities*

$$B\le\sum_{j=1}^{\nu(\boldsymbol{\sigma})}\int_{\sigma_{j-1}}^{\sigma_j} f_{s_j}\,\mathrm{d}g\le C$$

hold for all divisions $\boldsymbol{\sigma}$ of $[a,b]$ and all $s_1,\dots,s_{\nu(\boldsymbol{\sigma})}\in\mathbb{N}$.

Then $\{f_n\}$ is uniformly integrable with respect to g, the integral $\int_a^b f\,\mathrm{d}g$ exists, and

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

Moreover, we have

$$\lim_{n\to\infty}\left(\sup_{t\in[a,b]}\left|\int_a^t f_n\,\mathrm{d}g-\int_a^t f\,\mathrm{d}g\right|\right)=0.$$

Proof. According to Theorem 6.8.2, it suffices to prove that $\{f_n\}$ is uniformly integrable with respect to g. Choose an arbitrary $\varepsilon>0$. For each $\ell\in\mathbb{N}$, let $(\boldsymbol{\pi}^\ell,\boldsymbol{p}^\ell)$ and $(\boldsymbol{\rho}^\ell,\boldsymbol{r}^\ell)$ be as in Lemma 6.8.4. Also, for each $\ell\in\mathbb{N}$, there is a gauge δ_ℓ on $[a,b]$ such that for each δ_ℓ-fine partition P of $[a,b]$, we have

$$\left|S(f_\ell,\mathrm{d}g,P)-\int_a^b f_\ell\,\mathrm{d}g\right|<\varepsilon, \tag{6.8.8}$$

$$\left|S(f_{p_i^\ell},\mathrm{d}g,P)-\int_a^b f_{p_i^\ell}\,\mathrm{d}g\right|<\frac{\varepsilon}{2^\ell\nu(\boldsymbol{\pi}^\ell)}\quad\text{for } i\in\{1,\ldots,\nu(\boldsymbol{\pi}^\ell)\}, \tag{6.8.9}$$

$$\left|S(f_{r_i^\ell},\mathrm{d}g,P)-\int_a^b f_{r_i^\ell}\,\mathrm{d}g\right|<\frac{\varepsilon}{2^\ell\nu(\boldsymbol{\rho}^\ell)}\quad\text{for } i\in\{1,\ldots,\nu(\boldsymbol{\rho}^\ell)\}. \tag{6.8.10}$$

Moreover, assume that $\delta_\ell(t)<\operatorname{dist}\left(t,(\boldsymbol{\pi}^\ell\cup\boldsymbol{\rho}^\ell)\setminus\{t\}\right)$ holds for each $t\in[a,b]$.

Due to assumption (ii), for each $t\in[a,b]$ we can choose an $N(t)\in\mathbb{N}$ such that

$$|f(t)-f_n(t)|<\varepsilon\quad\text{for all }\ n\ge N(t).$$

Let us put

$$\delta(t)=\min\{\delta_1(t),\ldots,\delta_{N(t)}(t)\}\quad\text{for }\ t\in[a,b].$$

Let $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ be an arbitrary δ-fine partition of $[a,b]$ and $n\in\mathbb{N}$. Our goal is to obtain estimates for the terms appearing in $S(f_n,\mathrm{d}g,P)$. If $i\in\{1,\ldots,\nu(\boldsymbol{\alpha})\}$ is such that $N(\xi_i)\ge n$, then $\delta(\xi_i)\le\delta_n(\xi_i)$. Thus, assumption (6.8.8) together with the Saks-Henstock lemma imply

$$\sum_{\substack{i\in\{1,\ldots,\nu(\boldsymbol{\alpha})\}\\ N(\xi_i)\ge n}} f_n(\xi_i)\,(g(\alpha_i)-g(\alpha_{i-1}))\le\sum_{\substack{i\in\{1,\ldots,\nu(\boldsymbol{\alpha})\}\\ N(\xi_i)\ge n}}\int_{\alpha_{i-1}}^{\alpha_i} f_n\,\mathrm{d}g+\varepsilon. \tag{6.8.11}$$

It remains to estimate those terms in $S(f_n,\mathrm{d}g,P)$ for which $N(\xi_i)<n$. To this aim, consider a fixed $\ell\in\{1,\ldots,n-1\}$ together with all indices $i\in\{1,\ldots,\nu(\boldsymbol{\alpha})\}$ fulfilling $N(\xi_i)=\ell$. Due to the construction of the gauge δ_ℓ, each corresponding interval $[\alpha_{i-1},\alpha_i]$ contains at most one of the division points of $\boldsymbol{\pi}^\ell$ and, in that case, the tag ξ_i must coincide with this division point. If we split the interval $[\alpha_{i-1},\alpha_i]$ into a pair of intervals $[\alpha_{i-1},\xi_i]$ and $[\xi_i,\alpha_i]$ that share the same tag

ξ_i, the value of $S(f_n, \mathrm{d}g, P)$ remains unchanged. Thus, without loss of generality, we can assume that for every $i \in \{1, \dots, \nu(\boldsymbol{\alpha})\}$ satisfying $N(\xi_i) = \ell$, there is a $j(i) \in \{1, \dots, \nu(\boldsymbol{\pi}^\ell)\}$ such that $[\alpha_{i-1}, \alpha_i] \subset [\pi^\ell_{j(i)-1}, \pi^\ell_{j(i)}]$. Then

$$\begin{aligned}
&\sum_{\substack{i \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \\ N(\xi_i)=\ell}} f_n(\xi_i)\,(g(\alpha_i) - g(\alpha_{i-1})) \\
&\quad = \sum_{\substack{i \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \\ N(\xi_i)=\ell}} (f_n(\xi_i) - f_{p^\ell_{j(i)}}(\xi_i))\,(g(\alpha_i) - g(\alpha_{i-1})) \\
&\qquad + \sum_{\substack{i \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \\ N(\xi_i)=\ell}} \left(f_{p^\ell_{j(i)}}(\xi_i)\,(g(\alpha_i) - g(\alpha_{i-1})) - \int_{\alpha_{i-1}}^{\alpha_i} f_{p^\ell_{j(i)}}\,\mathrm{d}g \right) \\
&\qquad + \sum_{\substack{i \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \\ N(\xi_i)=\ell}} \int_{\alpha_{i-1}}^{\alpha_i} f_{p^\ell_{j(i)}}\,\mathrm{d}g.
\end{aligned}$$

Since $(\boldsymbol{\pi}^\ell, \boldsymbol{p}^\ell) \in \mathcal{S}_\ell$, we also have $p^\ell_{j(i)} \geq \ell$. Hence, if $N(\xi_i) = \ell < n$, the definition of $N(\xi_i)$ gives

$$|f_n(\xi_i) - f_{p^\ell_{j(i)}}(\xi_i)| \leq |f_n(\xi_i) - f(\xi_i)| + |f(\xi_i) - f_{p^\ell_{j(i)}}(\xi_i)| < 2\,\varepsilon,$$

and consequently

$$\Big| (f_n(\xi_i) - f_{p^\ell_{j(i)}}(\xi_i))\,(g(\alpha_i) - g(\alpha_{i-1})) \Big| < 2\,\varepsilon \operatorname{var}_{\alpha_{i-1}}^{\alpha_i} g.$$

Next, we notice that

$$\begin{aligned}
&\left| \sum_{\substack{i \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \\ N(\xi_i)=\ell}} \left(f_{p^\ell_{j(i)}}(\xi_i)\,(g(\alpha_i) - g(\alpha_{i-1})) - \int_{\alpha_{i-1}}^{\alpha_i} f_{p^\ell_{j(i)}}\,\mathrm{d}g \right) \right| \\
&\quad \leq \sum_{k=1}^{\nu(\boldsymbol{\pi}^\ell)} \left| \sum_{\substack{i \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \\ N(\xi_i)=\ell,\ j(i)=k}} \left(f_{p^\ell_k}(\xi_i)\,(g(\alpha_i) - g(\alpha_{i-1})) - \int_{\alpha_{i-1}}^{\alpha_i} f_{p^\ell_k}\,\mathrm{d}g \right) \right| \\
&\quad \leq \sum_{k=1}^{\nu(\boldsymbol{\pi}^\ell)} \frac{\varepsilon}{2^\ell \nu(\boldsymbol{\pi}^\ell)} = \frac{\varepsilon}{2^\ell},
\end{aligned}$$

where the last inequality is a consequence of the assumption (6.8.9) and the Saks-Henstock lemma.

To summarize, for each $\ell \in \{1, \dots, n-1\}$ the previous estimates imply

$$\begin{aligned}
&\sum_{\substack{i \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \\ N(\xi_i)=\ell}} f_n(\xi_i)\,(g(\alpha_i)-g(\alpha_{i-1})) \\
&\quad < \sum_{\substack{i \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \\ N(\xi_i)=\ell}} \left(\int_{\alpha_{i-1}}^{\alpha_i} f_{p^\ell_{j(i)}}\,\mathrm{d}g + 2\,\varepsilon\,\mathrm{var}_{\alpha_{i-1}}^{\alpha_i} g \right) + \frac{\varepsilon}{2^\ell} \\
&\quad < \sum_{\substack{i \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \\ N(\xi_i)=\ell}} \left(\int_{\alpha_{i-1}}^{\alpha_i} f_n\,\mathrm{d}g + 2\,\varepsilon\,\mathrm{var}_{\alpha_{i-1}}^{\alpha_i} g \right) + 2\,\frac{\varepsilon}{2^\ell},
\end{aligned}$$

where the last inequality follows from (6.8.6). Summation over $\ell \in \{1, \dots, n-1\}$ yields

$$\sum_{\substack{i \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \\ N(\xi_i)<n}} f_n(\xi_i)\,(g(\alpha_i)-g(\alpha_{i-1})) < \sum_{\substack{i \in \{1,\dots,\nu(\boldsymbol{\alpha})\} \\ N(\xi_i)<n}} \int_{\alpha_{i-1}}^{\alpha_i} f_n\,\mathrm{d}g + 2\,\varepsilon\,\mathrm{var}_a^b g + 2\,\varepsilon.$$

Adding the last inequality and (6.8.11), we conclude that the estimate

$$S(f_n, \mathrm{d}g, P) = \sum_{i=1}^{\nu(\boldsymbol{\alpha})} f_n(\xi_i)\,(g(\alpha_i)-g(\alpha_{i-1})) < \int_a^b f_n\,\mathrm{d}g + 2\,\varepsilon\,\mathrm{var}_a^b g + 3\,\varepsilon$$

holds for all $n \in \mathbb{N}$. Proceeding in a similar way (using (6.8.7) and (6.8.10)), we can show that also

$$S(f_n, \mathrm{d}g, P) > \int_a^b f_n\,\mathrm{d}g - 2\,\varepsilon\,\mathrm{var}_a^b g - 3\,\varepsilon$$

holds for all $n \in \mathbb{N}$. As a consequence,

$$\left| S(f_n, \mathrm{d}g, P) - \int_a^b f_n\,\mathrm{d}g \right| < 2\,\varepsilon\,\mathrm{var}_a^b g + 3\,\varepsilon$$

for each δ-fine partition P of $[a,b]$ and each $n \in \mathbb{N}$, i.e., $\{f_n\}$ is uniformly integrable with respect to g. □

A straightforward consequence of Theorem 6.8.5 is the dominated convergence theorem.

6.8.6 Theorem (DOMINATED CONVERGENCE THEOREM)**.**
Let $g:[a,b] \to \mathbb{R}$ be nondecreasing and let $\{f_n\}$ be a sequence of real valued functions defined on $[a,b]$ and satisfying the following conditions:

- *The integral* $\int_a^b f_n \,\mathrm{d}g$ *exists for each* $n \in \mathbb{N}$.
- $\lim_{n\to\infty} f_n(t) = f(t)$ *for* $t \in [a,b]$.
- *There exist functions* $h_1, h_2 : [a,b] \to \mathbb{R}$ *such that the integrals* $\int_a^b h_1 \,\mathrm{d}g$, $\int_a^b h_2 \,\mathrm{d}g$ *exist and* $h_1 \le f_n \le h_2$ *on* $[a,b]$ *for each* $n \in \mathbb{N}$.

Then $\{f_n\}$ *is uniformly integrable with respect to* g, *the integral* $\int_a^b f \,\mathrm{d}g$ *exists and*

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

Moreover, we have

$$\lim_{n\to\infty} \left(\sup_{t\in[a,b]} \left| \int_a^t f_n \,\mathrm{d}g - \int_a^t f \,\mathrm{d}g \right| \right) = 0.$$

Proof. The statement is a consequence of Theorem 6.8.5. To see this, let

$$B = \int_a^b h_1 \,\mathrm{d}g, \qquad C = \int_a^b h_2 \,\mathrm{d}g.$$

If $\boldsymbol{\sigma}$ is a division of $[a,b]$ and $s_1, \ldots, s_{\nu(\boldsymbol{\sigma})} \in \mathbb{N}$, then

$$B = \sum_{j=1}^{\nu(\boldsymbol{\sigma})} \int_{\sigma_{j-1}}^{\sigma_j} h_1 \,\mathrm{d}g \le \sum_{j=1}^{\nu(\boldsymbol{\sigma})} \int_{\sigma_{j-1}}^{\sigma_j} f_{s_j} \,\mathrm{d}g \le \sum_{j=1}^{\nu(\boldsymbol{\sigma})} \int_{\sigma_{j-1}}^{\sigma_j} h_2 \,\mathrm{d}g = C.$$

This shows that the assumptions of Theorem 6.8.5 are satisfied, and the proof is complete. □

6.8.7 Remark. In the previous theorem, the condition $h_1 \le f_n \le h_2$ on $[a,b]$ can be weakened. It is enough to assume that for each interval $[c,d] \subset [a,b]$ and each $n \in \mathbb{N}$, we have

$$\int_c^d h_1 \,\mathrm{d}g \le \int_c^d f_n \,\mathrm{d}g \le \int_c^d h_2 \,\mathrm{d}g.$$

In a similar way, we can derive the bounded convergence theorem. In comparison with the dominated convergence theorem, the dominating hypothesis on $\{f_n\}$ is stronger. On the other hand, g is no longer assumed to be nondecreasing, but merely of bounded variation.

6.8.8 Theorem (BOUNDED CONVERGENCE THEOREM).
Let $g \in \mathrm{BV}([a,b])$ and let $\{f_n\}$ be a sequence of real valued functions defined on $[a,b]$ and satisfying the following conditions:

- *The integral $\int_a^b f_n \, \mathrm{d}g$ exists for each $n \in \mathbb{N}$.*
- $\lim\limits_{n\to\infty} f_n = f$ *on* $[a,b]$.
- *There exists a constant $M \ge 0$ such that $|f_n(t)| \le M$ for all $n \in \mathbb{N}$ and $t \in [a,b]$.*

Then $\{f_n\}$ is uniformly integrable with respect to g, the integral $\int_a^b f \, \mathrm{d}g$ exists and

$$\lim_{n\to\infty} \left(\sup_{t\in[a,b]} \left| \int_a^t f_n \, \mathrm{d}g - \int_a^t f \, \mathrm{d}g \right| \right) = 0.$$

Proof. If $\boldsymbol{\sigma}$ is a division of $[a,b]$ and $s_1, \ldots, s_{\nu(\boldsymbol{\sigma})} \in \mathbb{N}$, then

$$\left| \sum_{j=1}^{\nu(\boldsymbol{\sigma})} \int_{\sigma_{j-1}}^{\sigma_j} f_{s_j} \, \mathrm{d}g \right| \le \sum_{j=1}^{\nu(\boldsymbol{\sigma})} \left| \int_{\sigma_{j-1}}^{\sigma_j} f_{s_j} \, \mathrm{d}g \right| \le \sum_{j=1}^{\nu(\boldsymbol{\sigma})} M \operatorname{var}_{\sigma_{j-1}}^{\sigma_j} g = M \operatorname{var}_a^b g.$$

Hence, the assumptions of Theorem 6.8.5 are satisfied with

$$B = -M \operatorname{var}_a^b g, \qquad C = M \operatorname{var}_a^b g,$$

and the proof is complete. □

Another important consequence of Theorem 6.8.5 is the monotone convergence theorem.

6.8.9 Theorem (MONOTONE CONVERGENCE THEOREM).
Let $g : [a,b] \to \mathbb{R}$ be nondecreasing and $\{f_n\}$ be a sequence of functions such that the integral $\int_a^b f_n \, \mathrm{d}g$ exists for each $n \in \mathbb{N}$, and $\lim\limits_{n\to\infty} f_n = f$ on $[a,b]$. Suppose, in addition, that one of the following conditions holds:

- *$\{f_n\}$ is a nondecreasing sequence and* $\lim\limits_{n\to\infty} \int_a^b f_n \, \mathrm{d}g < \infty$.
- *$\{f_n\}$ is a nonincreasing sequence and* $\lim\limits_{n\to\infty} \int_a^b f_n \, \mathrm{d}g > -\infty$.

Then $\{f_n\}$ *is uniformly integrable with respect to* g, *the integral* $\int_a^b f\,\mathrm{d}g$ *exists and*

$$\int_a^b f\,\mathrm{d}g = \lim_{n\to\infty}\int_a^b f_n\,\mathrm{d}g. \tag{6.8.12}$$

Moreover, we have

$$\lim_{n\to\infty}\left(\sup_{t\in[a,b]}\left|\int_a^t f_n\,\mathrm{d}g - \int_a^t f\,\mathrm{d}g\right|\right) = 0. \tag{6.8.13}$$

Proof. Suppose that the first condition holds. Let

$$B = \int_a^b f_1\,\mathrm{d}g, \qquad C = \lim_{n\to\infty}\int_a^b f_n\,\mathrm{d}g.$$

If $\boldsymbol{\sigma}$ is a division of $[a,b]$ and $s_1,\dots,s_{\nu(\boldsymbol{\sigma})}\in\mathbb{N}$, then

$$B = \sum_{j=1}^{\nu(\boldsymbol{\sigma})}\int_{\sigma_{j-1}}^{\sigma_j} f_1\,\mathrm{d}g \le \sum_{j=1}^{\nu(\boldsymbol{\sigma})}\int_{\sigma_{j-1}}^{\sigma_j} f_{s_j}\,\mathrm{d}g \le \sum_{j=1}^{\nu(\boldsymbol{\sigma})}\int_{\sigma_{j-1}}^{\sigma_j}\big(f_{\max(s_1,\dots,s_{\nu(\boldsymbol{\sigma})})}\big)\,\mathrm{d}g \le C.$$

Thus, the assumptions of Theorem 6.8.5 are satisfied, and (6.8.12), (6.8.13) hold. If the second condition holds, it is enough to consider the sequence $\{-f_n\}$, which obviously satisfies the first condition. □

The monotone convergence theorem has a number of useful corollaries, which will be needed later. The first result is an analogue of Levi's theorem for term-by-term integration of infinite series.

6.8.10 Theorem. *Suppose that* $g:[a,b]\to\mathbb{R}$ *is nondecreasing,* $f_k:[a,b]\to\mathbb{R}$ *is a nonnegative function for each* $k\in\mathbb{N}$, *and* $f=\sum_{k=1}^\infty f_k$ *on* $[a,b]$. *If the integral* $\int_a^b f_k\,\mathrm{d}g$ *exists for each* $k\in\mathbb{N}$ *and if the sum* $\sum_{k=1}^\infty\int_a^b f_k\,\mathrm{d}g$ *is finite, then the sequence* $\{s_n\}$ *given by* $s_n=\sum_{k=1}^n f_k$ *on* $[a,b]$ *is uniformly integrable with respect to* g *and*

$$\int_a^b f\,\mathrm{d}g = \sum_{k=1}^\infty\int_a^b f_k\,\mathrm{d}g.$$

Moreover, we have

$$\lim_{n\to\infty}\left(\sup_{t\in[a,b]}\left|\int_a^t s_n\,\mathrm{d}g - \int_a^t f\,\mathrm{d}g\right|\right) = 0.$$

Proof. The sequence $\{s_n\}$ is nondecreasing, the integral $\int_a^b s_n \,\mathrm{d}g$ exists for each $n \in \mathbb{N}$, and $\lim_{n\to\infty} s_n = f$ on $[a,b]$. Moreover, we have

$$\begin{aligned}\lim_{n\to\infty} \int_a^b s_n \,\mathrm{d}g &= \lim_{n\to\infty} \int_a^b \left(\sum_{k=1}^n f_k\right) \mathrm{d}g \\ &= \lim_{n\to\infty} \sum_{k=1}^n \int_a^b f_k \,\mathrm{d}g = \sum_{k=1}^\infty \int_a^b f_k \,\mathrm{d}g < \infty.\end{aligned}$$

Thus, the statement of the present theorem follows from the monotone convergence theorem. □

6.8.11 Remark. Suppose that $g:[a,b] \to \mathbb{R}$ is nondecreasing and $f = \sum_{k=1}^\infty f_k$ on $[a,b]$, where each $f_k:[a,b]\to\mathbb{R}$ is a nonnegative function such that $\int_a^b f_k \,\mathrm{d}g$ exists. If we know that $\int_a^b f \,\mathrm{d}g$ exists, then

$$\sum_{k=1}^n \int_a^b f_k \,\mathrm{d}g = \int_a^b \left(\sum_{k=1}^n f_k\right) \mathrm{d}g \le \int_a^b f \,\mathrm{d}g \quad \text{for each } n \in \mathbb{N}.$$

Hence, $\sum_{k=1}^\infty \int_a^b f_k \,\mathrm{d}g$ is finite and equals $\int_a^b f \,\mathrm{d}g$ by Theorem 6.8.10. In other words, we have established the following result:

$\int_a^b \sum_{k=1}^\infty f_k \,\mathrm{d}g$ exists if and only if $\sum_{k=1}^\infty \int_a^b f_k \,\mathrm{d}g$ is finite; in this case, both expressions have the same value.

The monotone convergence theorem also leads to the following result.

6.8.12 Lemma. *If $g:[a,b]\to\mathbb{R}$ is nondecreasing and $\{f_k\}$ is a sequence of nonnegative functions such that the integral $\int_a^b f_k \,\mathrm{d}g$ exists for each $k \in \mathbb{N}$, then $\int_a^b (\inf_{k\in\mathbb{N}} f_k) \,\mathrm{d}g$ exists as well.*

Proof. For each $n \in \mathbb{N}$, let $h_n = \min\{f_1, \dots, f_n\}$, and note that Theorem 6.7.5 guarantees the existence of the integral $\int_a^b h_n \,\mathrm{d}g \ge 0$. The sequence $\{h_n\}$ is nonincreasing and pointwise convergent to $\inf_{k\in\mathbb{N}} f_k$. Hence, the monotone convergence theorem implies the existence of the integral $\int_a^b (\inf_{k\in\mathbb{N}} f_k) \,\mathrm{d}g$. □

Another useful corollary of the monotone convergence theorem is Fatou's lemma.

6.8.13 Lemma (FATOU'S LEMMA)**.** *Let $g:[a,b]\to\mathbb{R}$ be nondecreasing and let $\{f_n\}$ be a sequence of functions such that the integral $\int_a^b f_n \,\mathrm{d}g$ exists for each $n \in \mathbb{N}$. Then the following statements hold:*

(i) *Assume there is a function* $\varphi:[a,b]\to\mathbb{R}$ *such that* $\int_a^b \varphi\,\mathrm{d}g$ *exists and* $f_n \ge \varphi$ *on* $[a,b]$ *for each* $n\in\mathbb{N}$. *If* $\liminf_{n\to\infty} f_n(x) < \infty$ *for each* $x\in[a,b]$ *and* $\liminf_{n\to\infty}\int_a^b f_n\,\mathrm{d}g < \infty$, *then* $\int_a^b (\liminf_{n\to\infty} f_n)\,\mathrm{d}g$ *exists, and we have*

$$\int_a^b (\liminf_{n\to\infty} f_n)\,\mathrm{d}g \le \liminf_{n\to\infty}\int_a^b f_n\,\mathrm{d}g.$$

(ii) *Assume there is a function* $\psi:[a,b]\to\mathbb{R}$ *such that* $\int_a^b \psi\,\mathrm{d}g$ *exists and* $f_n \le \psi$ *for each* $n\in\mathbb{N}$. *If* $\limsup_{n\to\infty} f_n(x) > -\infty$ *for each* $x\in[a,b]$ *and* $\limsup_{n\to\infty}\int_a^b f_n\,\mathrm{d}g > -\infty$, *then* $\int_a^b (\limsup_{n\to\infty} f_n)\,\mathrm{d}g$ *exists, and we have*

$$\int_a^b (\limsup_{n\to\infty} f_n)\,\mathrm{d}g \ge \limsup_{n\to\infty}\int_a^b f_n\,\mathrm{d}g.$$

Proof. Let us prove the first statement. Without loss of generality, we can assume that $\varphi = 0$ (otherwise, it is enough to consider the sequence $\{f_n - \varphi\}$). For each $n\in\mathbb{N}$, let $h_n = \inf_{k\ge n} f_k$. The sequence $\{h_n\}$ consists of nonnegative functions, it is nondecreasing, and pointwise convergent to $\liminf_{n\to\infty} f_n$. By Lemma 6.8.12, the integral $\int_a^b h_n\,\mathrm{d}g$ exists for each $n\in\mathbb{N}$. Obviously, we have

$$\int_a^b h_n\,\mathrm{d}g \le \int_a^b f_n\,\mathrm{d}g,$$

and therefore

$$\lim_{n\to\infty}\int_a^b h_n\,\mathrm{d}g = \liminf_{n\to\infty}\int_a^b h_n\,\mathrm{d}g \le \liminf_{n\to\infty}\int_a^b f_n\,\mathrm{d}g < \infty.$$

The monotone convergence theorem implies the existence of the integral $\int_a^b (\lim_{n\to\infty} h_n)\,\mathrm{d}g$, and we get

$$\int_a^b (\liminf_{n\to\infty} f_n)\,\mathrm{d}g = \int_a^b (\lim_{n\to\infty} h_n)\,\mathrm{d}g = \lim_{n\to\infty}\int_a^b h_n\,\mathrm{d}g \le \liminf_{n\to\infty}\int_a^b f_n\,\mathrm{d}g.$$

To prove that the second statement holds, it is enough to consider the sequence $\{-f_n\}$ and apply the first statement. □

6.9 Integration over elementary sets

Up to now we have been discussing the integration over a fixed given interval, more precisely, integration from some lower bound to some upper bound. In the theory of the Lebesgue integral one can meet also the possibility of integration over more general sets. The generality of the Kurzweil-Stieltjes integral (in the sense of a large domain of integrable functions) seems to be the source of troubles with defining the Kurzweil-Stieltjes integral over general sets, in particular in the cases when we do not want to restrict ourselves to continuous integrators. A reasonable compromise seems to be to restrict the considerations to elementary sets (cf. Definition 2.8.10). In this case the following definition turned out to be useful for our purposes.

6.9.1 Definition. Let $f:[a,b]\to\mathbb{R},\ g:[a,b]\to\mathbb{R}$ and an elementary subset E of $[a,b]$ be given. The Kurzweil-Stieltjes integral of f with respect to g over E is given by

$$\int_E f\,\mathrm{d}g = \int_a^b (f\,\chi_E)\,\mathrm{d}g$$

provided the integral on the right-hand side exists.

According to Definition 6.9.1, the existence of the integral $\int_E f\,\mathrm{d}g$ means that there is an $I\in\mathbb{R}$ such that for every $\varepsilon>0$ there is a gauge δ on $[a,b]$ such that

$$\big|S(f\chi_E,\mathrm{d}g,P)-I\big|<\varepsilon \quad \text{for all } \delta\text{-fine partitions } P \text{ on } [a,b].$$

As we will see later, the value of the integral $\int_E f\,\mathrm{d}g$ can depend on the choice of the interval $[a,b]$ which contains E. Hence, throughout this section the interval $[a,b]$ is assumed to be fixed. Further, we use extensively the following convention mentioned in the point (x) of Conventions and notation: Functions g defined on the interval $[a,b]$ are supposed to be extended to some open interval containing $[a,b]$ in such a way that

$$g(a-)=g(a),\quad g(b+)=g(b),\quad \text{i.e. } \Delta^- g(a)=\Delta^+ g(b)=0.$$

The basic properties of the Kurzweil-Stieltjes integral over elementary sets given below are immediate consequences of the corresponding properties of the Kurzweil-Stieltjes integral presented in the previous sections.

6.9.2 Theorem. *Let E be an elementary subset of $[a,b]$. Assume that the functions $f,f_1,f_2,g,g_1,g_2:[a,b]\to\mathbb{R}$ are such that the integrals*

$$\int_E f_1\,\mathrm{d}g,\quad \int_E f_2\,\mathrm{d}g,\quad \int_E f\,\mathrm{d}g_1 \ \text{ and } \int_E f\,\mathrm{d}g_2$$

exist. Then for any $c_1,c_2\in\mathbb{R}$,

$$\int_E (c_1f_1+c_2f_2)\,\mathrm{d}g = c_1\int_E f_1\,\mathrm{d}g + c_2\int_E f_2\,\mathrm{d}g$$

and

$$\int_E f\,\mathrm{d}[c_1\,g_1 + c_2\,g_2] = c_1 \int_E f\,\mathrm{d}[g_1] + c_2 \int_E f\,\mathrm{d}[g_2]$$

hold.

Using Theorem 6.9.2 the following additivity result of the integral with respect to elementary sets can be justified.

6.9.3 Theorem. *Let E_1 and E_2 be disjoint elementary subsets of $[a,b]$. Assume that $f, g : [a,b] \to \mathbb{R}$ are such that both the integrals*

$$\int_{E_1} f\,\mathrm{d}g \quad \textit{and} \quad \int_{E_2} f\,\mathrm{d}g$$

exist. Then the integral $\int_{E_1 \cup E_2} f\,\mathrm{d}g$ exists as well and

$$\int_{E_1 \cup E_2} f\,\mathrm{d}g = \int_{E_1} f\,\mathrm{d}g + \int_{E_2} f\,\mathrm{d}g.$$

6.9.4 Remark. Note that if E is an elementary subset of $[a,b]$ and $h = 0$ on E, then obviously $\int_E h\,\mathrm{d}g = 0$ for each $g : [a,b] \to \mathbb{R}$. Therefore, if the integral $\int_E f_1\,\mathrm{d}g$ exists, then the equality

$$\int_E f_1\,\mathrm{d}g = \int_E f_2\,\mathrm{d}g$$

holds for each function f_2 such that $f_2 = f_1$ on E.

Later in this section we will provide conditions ensuring the existence of the integral over elementary sets (see Corollary 6.9.10). For now, let us simply highlight the following result which is a consequence of Theorem 6.3.8.

6.9.5 Theorem. *Let $f \in \mathrm{G}([a,b])$ and $g \in \mathrm{BV}([a,b])$. If E is an elementary subset of $[a,b]$, then the integral $\int_E f\,\mathrm{d}g$ exists.*

Proof. It is enough to observe that $f\,\chi_E : [a,b] \to \mathbb{R}$ is a regulated function whenever $f \in \mathrm{G}([a,b])$ and E is an elementary subset of $[a,b]$. □

Now we are going to consider the properties of integrals over arbitrary bounded intervals. For degenerate intervals, i.e., singleton sets, we have the following result.

6.9.6 Theorem. *Let $\tau \in [a,b]$ and let $g : [a,b] \to \mathbb{R}$ have finite one-sided limits $g(\tau-)$ and $g(\tau+)$. Then the integral $\int_{[\tau]} f\,\mathrm{d}g$ exists and*

$$\int_{[\tau]} f\,\mathrm{d}g = f(\tau)\,\Delta g(\tau). \tag{6.9.1}$$

Proof. Note that $\big(f\chi_{[\tau]}\big)(t) = f(\tau)\chi_{[\tau]}(t)$ for all $t \in [a,b]$. Hence,

$$\int_{[\tau]} f \, \mathrm{d}g = \int_a^b f\chi_{[\tau]} \, \mathrm{d}g = f(\tau) \int_a^b \chi_{[\tau]} \, \mathrm{d}g$$

and the result follows from Example 6.3.1 (iii). □

Integration over subintervals of $[a,b]$ is described by the following assertion.

6.9.7 Theorem. *Let* $f : [a,b] \to \mathbb{R}$, $g \in \mathrm{G}([a,b])$ *and* $a \le c < d \le b$. *Then if one of the integrals*

$$\int_{[c,d]} f \, \mathrm{d}g, \quad \int_{(c,d)} f \, \mathrm{d}g, \quad \int_{[c,d)} f \, \mathrm{d}g, \quad \int_{(c,d]} f \, \mathrm{d}g, \quad \int_c^d f \, \mathrm{d}g \tag{6.9.2}$$

exists, all the others exist as well. In this case, we have the following equalities:

$$\int_{[c,d]} f \, \mathrm{d}g = \quad f(c)\,\Delta^- g(c) + \int_c^d f \, \mathrm{d}g + f(d)\,\Delta^+ g(d), \tag{6.9.3}$$

$$\int_{(c,d)} f \, \mathrm{d}g = -f(c)\,\Delta^+ g(c) + \int_c^d f \, \mathrm{d}g - f(d)\,\Delta^- g(d), \tag{6.9.4}$$

$$\int_{[c,d)} f \, \mathrm{d}g = \quad f(c)\,\Delta^- g(c) + \int_c^d f \, \mathrm{d}g - f(d)\,\Delta^- g(d), \tag{6.9.5}$$

$$\int_{(c,d]} f \, \mathrm{d}g = -f(c)\,\Delta^+ g(c) + \int_c^d f \, \mathrm{d}g + f(d)\,\Delta^+ g(d). \tag{6.9.6}$$

Proof. Note that

$$\int_{[c,d]} f \, \mathrm{d}g = \int_a^c (f\chi_{[c,d]}) \, \mathrm{d}g + \int_c^d f \, \mathrm{d}g + \int_d^b (f\chi_{[c,d]}) \, \mathrm{d}g.$$

Clearly, the first and third integral on the right-hand side exist by Examples 6.3.1 (iii). Thus, the integral over $[c,d]$ exists if and only if the integral $\int_c^d f \, \mathrm{d}g$ exists. In this case,

$$\begin{aligned}\int_{[c,d]} f \, \mathrm{d}g &= f(c) \int_a^c \chi_{[c]} \, \mathrm{d}g + \int_c^d f \, \mathrm{d}g + f(d) \int_d^b \chi_{[d]} \, \mathrm{d}g \\ &= f(c)\,\Delta^- g(c) + \int_c^d f \, \mathrm{d}g + f(d)\,\Delta^+ g(d).\end{aligned}$$

Since by Theorem 6.9.6, both integrals $\int_{[c]} f \, \mathrm{d}g$ and $\int_{[d]} f \, \mathrm{d}g$ exist, the equivalence between the existence of the integral over $[c,d]$ and the existence of integral over the open or half-open corresponding intervals can be easily derived from the relations

$$\chi_{[c,d]} = \chi_{[c]} + \chi_{(c,d)} + \chi_{[d]} = \chi_{[c]} + \chi_{(c,d]} = \chi_{[c,d)} + \chi_{[d]}.$$

The equalities (6.9.4), (6.9.5) and (6.9.6) follow from (6.9.3) using Theorem 6.9.6. The detailed proof is left as an exercise for the reader. □

6.9.8 Remark. According to Theorem 6.9.7, given $f:[a,b]\to\mathbb{R},\ g\in \mathrm{C}([a,b])$ and $a\le c\le d\le b$, such that one of the integrals in (6.9.2) exists, then

$$\int_{[c,d)} f\,\mathrm{d}g = \int_{(c,d]} f\,\mathrm{d}g = \int_{(c,d)} f\,\mathrm{d}g = \int_c^d f\,\mathrm{d}g = \int_{[c,d]} f\,\mathrm{d}g.$$

We are now ready to evaluate integrals over elementary sets. To this end, we will make use of the notion of minimal decomposition of an elementary set (see Definition 2.8.9).

6.9.9 Theorem. *Let $f:[a,b]\to\mathbb{R},\ g\in \mathrm{G}([a,b])$ and let E be an elementary subset of $[a,b]$ with the minimal decomposition $\{J_k:\ k=1,\dots,N\}$. If the integral $\int_E f\,\mathrm{d}g$ exists, then also the integrals $\int_{J_k} f\,\mathrm{d}g$ exist for all $k\in\{1,\dots,N\}$ and*

$$\int_E f\,\mathrm{d}g = \sum_{k=1}^{N}\int_{J_k} f\,\mathrm{d}g. \tag{6.9.7}$$

Proof. For $k=1,\dots,N$, let $c_k=\inf J_k$ and $d_k=\sup J_k$. By the hypothesis, the integral

$$\int_E f\,\mathrm{d}g = \int_a^b (f\chi_E)\,\mathrm{d}g$$

exists, and hence, by Theorem 6.1.10, so do all the integrals

$$\int_{c_k}^{d_k} (f\chi_E)\,\mathrm{d}g \quad \text{for } k\in\{1,\dots,N\}.$$

Note that

$$\int_{c_k}^{d_k} f(\chi_E-\chi_{J_k})\,\mathrm{d}g = 0, \qquad \text{for } k\in\{1,\dots,N\}$$

(due to the fact that $f(\chi_E-\chi_{J_k})$ vanishes on $[c_k,d_k]$) and, similarly, the integrals

$$\int_a^{c_k} (f\chi_{J_k})\,\mathrm{d}g \quad \text{and} \quad \int_{d_k}^{b} (f\chi_{J_k})\,\mathrm{d}g$$

are either zero or integrals over singleton sets (depending on whether c_k and/or d_k belong to J_k). Since all of these integrals exist, we conclude that

$$\int_{J_k} f\,\mathrm{d}g = \int_a^b (f\,\chi_{J_k})\,\mathrm{d}g$$

exist for all $k \in \{1,\dots,N\}$. Having in mind that the intervals of the minimal decomposition are pairwise disjoint, (6.9.7) follows from Theorem 6.9.3. □

Theorem 6.9.9 means that once the integral over an elementary set exists, so do the integrals over subintervals of the minimal decomposition. This implies immediately the following assertions.

6.9.10 Corollary. *Let E be an elementary subset of $[a,b]$ and let $f:[a,b]\to\mathbb{R}$ and $g\in\mathrm{G}([a,b])$ be such that the integral $\int_E f\,\mathrm{d}g$ exists. Then the integral $\int_T f\,\mathrm{d}g$ exists for every elementary subset T of E.*

In particular, if the integral $\int_a^b f\,\mathrm{d}g$ exists, then $\int_T f\,\mathrm{d}g$ exists for every elementary subset T of $[a,b]$.

Proof. Let T be an elementary subset of E and let $\{I_\ell : \ell=1,\dots,p\}$ be its minimal decomposition. Assume that $\{J_k : k=1,\dots,N\}$ is the minimal decomposition of E. Fixed an arbitrary $\ell\in\{1,\dots,p\}$, there exists $k_\ell\in\{1,\dots,N\}$ such that $I_\ell\subset J_{k_\ell}$. By Theorem 6.9.9, we know that the integral $\int_a^b (f\,\chi_{J_{k_\ell}})\,\mathrm{d}g$ exists; therefore $\int_{a_\ell}^{b_\ell}(f\,\chi_{J_{k_\ell}})\,\mathrm{d}g$ also exists, where $a_\ell=\inf I_\ell$ and $b_\ell=\sup I_\ell$. Applying Theorem 6.9.7 we conclude that the integral $\int_{I_\ell}(f\,\chi_{J_{k_\ell}})\,\mathrm{d}g$ exists, while $\chi_{J_{k_\ell}}\chi_{I_\ell}=\chi_{I_\ell}$ implies that

$$\int_{I_\ell} f\,\mathrm{d}g = \int_{I_\ell} (f\,\chi_{J_{k_\ell}})\,\mathrm{d}g$$

Since $\ell\in\{1,\dots,p\}$ is arbitrary and the intervals of the minimal decomposition are pairwise disjoint, using Theorem 6.9.3 the existence of the integral over T has been proved. □

6.9.11 Corollary. *Let E_1 and E_2 be elementary subsets of $[a,b]$ and let $f:[a,b]\to\mathbb{R}$ and $g\in\mathrm{G}([a,b])$ be such that both the integrals*

$$\int_{E_1} f\,\mathrm{d}g \quad \textit{and} \int_{E_2} f\,\mathrm{d}g$$

exist.

Then the integral $\int_{E_1\cup E_2} f\,\mathrm{d}g$ exists as well and

$$\int_{E_1\cup E_2} f\,\mathrm{d}g = \int_{E_1} f\,\mathrm{d}g + \int_{E_2} f\,\mathrm{d}g - \int_{E_1\cap E_2} f\,\mathrm{d}g. \tag{6.9.8}$$

Similarly, if the integral $\int_{E_1 \cup E_2} f \, \mathrm{d}g$ *exists, then both the integrals*

$$\int_{E_1} f \, \mathrm{d}g \quad \textit{and} \quad \int_{E_2} f \, \mathrm{d}g$$

exist and the equality (6.9.8) *holds.*

The following estimates are, in a sense, the analogues of the estimates from Section 6.3.

6.9.12 Theorem. *Let* J *be a subinterval of* $[a,b]$ *and let* $c = \inf J$ *and* $d = \sup J$ *be such that* $c < d$*. Assume that* $f : [a,b] \to \mathbb{R}$ *and* $g \in \mathrm{BV}([a,b])$ *are such that the integral* $\int_J f \, \mathrm{d}g$ *exists. Then the following assertions are true:*

(i) *If* $J = [c,d]$*, then*

$$\left| \int_J f \, \mathrm{d}g \right| \le \Big(\sup_{t \in J} |f(t)| \Big) \operatorname{var}(g,J) + |f(c)| |\Delta^- g(c)| + |f(d)| |\Delta^+ g(d)|.$$

(ii) *If* $J = [c,d)$*, then*

$$\left| \int_J f \, \mathrm{d}g \right| \le \Big(\sup_{t \in J} |f(t)| \Big) \operatorname{var}(g,J) + |f(c)| \, |\Delta^- g(c)|.$$

(iii) *If* $J = (c,d]$*, then*

$$\left| \int_J f \, \mathrm{d}g \right| \le \Big(\sup_{t \in J} |f(t)| \Big) \operatorname{var}(g,J) + |f(d)| \, |\Delta^+ g(d)|.$$

(iv) *If* $J = (c,d)$*, then*

$$\left| \int_J f \, \mathrm{d}g \right| \le \Big(\sup_{t \in J} |f(t)| \Big) \operatorname{var}(g,J).$$

Proof. (i) In the case $J = [c,d]$, the inequality is an obvious consequence of (6.9.3) and Theorem 6.3.4.

(ii) By (6.9.5) and Theorem 6.5.3 we get

$$\begin{aligned} \int_{[c,d)} f \, \mathrm{d}g &= f(c) \, \Delta^- g(c) + \int_c^d f \, \mathrm{d}g - f(d) \, \Delta^- g(d) \\ &= f(c) \, \Delta^- g(c) + \lim_{s \to d-} \int_c^s f \, \mathrm{d}g. \end{aligned}$$

This together with Theorem 2.8.4 implies

$$\Big|\int_{[c,d)} f\,\mathrm{d}g\Big| \le |f(c)|\,|\Delta^- g(c)| + \lim_{s\to d-}\Big(\sup_{t\in[c,s]} |f(t)|\,\mathrm{var}_c^s g\Big)$$
$$= |f(c)|\,|\Delta^- g(c)| + \Big(\sup_{t\in[c,d)} |f(t)|\Big)\,\mathrm{var}(g,[c,d)).$$

(iii) To obtain the corresponding estimate in the case when $J=(c,d]$, we use (6.9.6) and follow the same arguments used in the proof of (ii).

(iv) For an arbitrary but fixed $\tau\in(c,d)$, by (6.9.4) we have

$$\int_{(c,d)} f\,\mathrm{d}g = \int_c^d f\,\mathrm{d}g - f(c)\,\Delta^+ g(c) - f(d)\,\Delta^- g(d)$$
$$= \int_c^\tau f\,\mathrm{d}g - f(c)\,\Delta^+ g(c) + \int_\tau^d f\,\mathrm{d}g - f(d)\,\Delta^- g(d).$$

Thus, applying Theorem 6.5.3 we obtain

$$\int_{(c,d)} f\,\mathrm{d}g = \lim_{s\to c+}\int_s^\tau f\,\mathrm{d}g + \lim_{s\to d-}\int_\tau^s f\,\mathrm{d}g,$$

and consequently

$$\Big|\int_{(c,d)} f\,\mathrm{d}g\Big| \le \lim_{s\to c+}\Big(\sup_{t\in[s,\tau]}|f(t)|\,\mathrm{var}_s^\tau g\Big) + \lim_{s\to d-}\Big(\sup_{t\in[\tau,s]}|f(t)|\,\mathrm{var}_\tau^s g\Big)$$
$$= \Big(\sup_{t\in(c,d)}|f(t)|\Big)\Big(\lim_{s\to c+}\mathrm{var}_s^\tau g + \lim_{s\to d-}\mathrm{var}_\tau^s g\Big).$$

The estimate in (iv) then follows from Theorem 2.8.4. □

As a consequence of Definition 2.8.10, Theorem 6.9.12 and Theorem 6.9.9 we have the following result.

6.9.13 Corollary. *Let E be an elementary subset of $[a,b]$ and let $f:[a,b]\to\mathbb{R}$ and $g\in\mathrm{BV}([a,b])\cap\mathrm{C}([a,b])$ be such that the integral $\int_E f\,\mathrm{d}g$ exists. Then*

$$\Big|\int_E f\,\mathrm{d}g\Big| \le \Big(\sup_{t\in E}|f(t)|\Big)\,\mathrm{var}(g,E).$$

6.10 Integrals of vector, matrix and complex functions

Let us recall that, by part (xiii) of Conventions and Notation, the norm of a matrix $A \in \mathscr{L}(\mathbb{R}^m, \mathbb{R}^n)$ is denoted by the symbol $|A|$ and defined by

$$|A| = \max_{j=1,\dots,m} \sum_{i=1}^{n} |a_{i,j}|.$$

Vectors from $\mathbb{R}^n$ are identified with $n \times 1$ matrices (i.e., they are treated as column vectors). In other words, we identify the spaces $\mathbb{R}^n$ and $\mathscr{L}(\mathbb{R}^n, \mathbb{R})$. Consequently, the norm of a vector $x \in \mathbb{R}^n$ is

$$|x| = \sum_{i=1}^{n} |x_i|.$$

Clearly, we have $|A\,x| \le |A|\,|x|$ for all $A \in \mathscr{L}(\mathbb{R}^m, \mathbb{R}^n),\ x \in \mathbb{R}^n$. It is also known that

$$|A| = \sup\big\{|A\,x| : x \in \mathbb{R}^n, |x| \le 1\big\}.$$

Let $F : [a,b] \to \mathscr{L}(\mathbb{R}^m, \mathbb{R}^p)$ and $G : [a,b] \to \mathscr{L}(\mathbb{R}^p, \mathbb{R}^n)$ be matrix-valued functions with components

$$f_{i,k}, \quad i \in \{1,\dots,m\}, \quad k \in \{1,\dots,p\},$$

and

$$g_{k,j}, \quad k \in \{1,\dots,p\}, \quad j \in \{1,\dots,n\},$$

respectively. Then, if all the integrals

$$\int_a^b f_{i,k}\,\mathrm{d}g_{k,j} \quad i \in \{1,\dots,m\}, \quad k \in \{1,\dots,p\}, \quad j \in \{1,\dots,n\}$$

exist, the symbols

$$\int_a^b F(t)\,\mathrm{d}G(t) \quad \text{or} \quad \int_a^b F\,\mathrm{d}G$$

stand for the $m \times n$ matrix $M \in \mathscr{L}(\mathbb{R}^m, \mathbb{R}^n)$ with elements

$$m_{i,j} = \sum_{k=1}^{p} \int_a^b f_{i,k}\,\mathrm{d}g_{k,j}, \quad i = 1,\dots,m, \quad j = 1,\dots,n.$$

6.10.1 Exercise. Verify that if $\int_a^b F(t)\,\mathrm{d}G(t)$ exists and equals $M \in \mathscr{L}(\mathbb{R}^m, \mathbb{R}^n)$, then for each $\varepsilon > 0$, there is a gauge δ on $[a,b]$ such that the inequality

$$\Big|\sum_{\ell=1}^{\nu(P)} F(\xi_\ell)\,(G(\alpha_\ell) - G(\alpha_{\ell-1})) - M\Big| < \varepsilon$$

holds for each δ-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$.

Analogously, we can define the integrals $\int_a^b \mathrm{d}F\,G$, or $\int_a^b F\,\mathrm{d}G\,H$, where the values of F, G and H are matrices of appropriate dimensions.

The variation of a matrix-valued function $F : [a,b] \to \mathscr{L}(\mathbb{R}^m, \mathbb{R}^n)$ is defined by the same formula as in the scalar case, i.e.,

$$\operatorname{var}_a^b F = \sup_{\boldsymbol{\alpha} \in \mathscr{D}[a,b]} \sum_{j=1}^{\nu(\boldsymbol{\alpha})} |F(\alpha_j) - F(\alpha_{j-1})|.$$

One can easily verify the inequalities

$$\max_{\substack{i=1,\dots,m\\ j=1,\dots,n}} \big(\operatorname{var}_a^b f_{i,j}\big) \le \operatorname{var}_a^b F \le \sum_{i=1}^m \sum_{j=1}^n \operatorname{var}_a^b f_{i,j}.$$

This means that a matrix-valued function $F : [a,b] \to \mathscr{L}(\mathbb{R}^m, \mathbb{R}^n)$ has bounded variation if and only if all its components have bounded variation. Similarly, F is continuous or regulated if and only if all its components have the same property.

The results obtained in this and in the preceding chapter are easily generalized for matrix or vector functions. One only needs to keep in mind that matrix multiplication is in general not commutative, i.e., it is not allowed to change the order of matrix-valued functions involved in various products. For example, the integration by parts formula (Theorem 6.4.2) has to be formulated as follows:

If $F : [a,b] \to \mathscr{L}(\mathbb{R}^m, \mathbb{R}^p)$ is regulated and $G : [a,b] \to \mathscr{L}(\mathbb{R}^p, \mathbb{R}^n)$ has bounded variation, then both the integrals

$$\int_a^b F\,\mathrm{d}G \quad \textit{and} \quad \int_a^b \mathrm{d}F\,G$$

exist, and we have

$$\int_a^b F\,\mathrm{d}G + \int_a^b \mathrm{d}F\,G = F(b)\,G(b) - F(a)\,G(a) \\ + \sum_{x \in [a,b]} \Big(\Delta^- F(x)\,\Delta^- G(x) - \Delta^+ F(x)\,\Delta^+ G(x)\Big).$$

Similarly, the substitution theorem (Theorem 6.6.1) for matrix functions reads as follows:

Let $H:[a,b]\to\mathscr{L}(\mathbb{R}^m,\mathbb{R}^p)$ *be bounded and let* $F:[a,b]\to\mathscr{L}(\mathbb{R}^p,\mathbb{R}^q)$ *and* $G:[a,b]\to\mathscr{L}(\mathbb{R}^q,\mathbb{R}^n)$ *be such that the integral* $\int_a^b F\,\mathrm{d}G$ *exists. If one of the integrals*

$$\int_a^b H(x)\,\mathrm{d}\Big[\int_a^x F\,\mathrm{d}G\Big], \qquad \int_a^b (H\,F)\,\mathrm{d}G,$$

exists, then the other exists as well, and we have

$$\int_a^b H(x)\,\mathrm{d}\Big[\int_a^x F\,\mathrm{d}G\Big] = \int_a^b (H\,F)\,\mathrm{d}G.$$

Finally, let us consider Kurzweil-Stieltjes integrals of complex-valued functions. Given a pair of functions $f,g:[a,b]\to\mathbb{C}$ with real parts f_1,g_1 and imaginary parts f_2,g_2, we define

$$\begin{aligned}\int_a^b f\,\mathrm{d}g &= \int_a^b (f_1+\mathrm{i}f_2)\,\mathrm{d}(g_1+\mathrm{i}g_2)\\ &= \int_a^b f_1\,\mathrm{d}g_1 - \int_a^b f_2\,\mathrm{d}g_2 + \mathrm{i}\left(\int_a^b f_1\,\mathrm{d}g_2 + \int_a^b f_2\,\mathrm{d}g_1\right)\end{aligned}$$

whenever all four integrals on the right-hand side exist.

Again, most results obtained in this chapter, such as the integration by parts formula or substitution theorem, are still valid for complex-valued functions. We leave the verification of this fact up to the reader; the proofs are straightforward and based on the decomposition of complex functions into the real and imaginary parts.

6.10.2 Exercise. Given a pair of functions $f,g:[a,b]\to\mathbb{C}$, verify that if $\int_a^b f\,\mathrm{d}g$ exists and equals $I\in\mathbb{C}$, then for each $\varepsilon>0$, there is a gauge δ on $[a,b]$ such that the inequality

$$\left|\sum_{j=1}^{\nu(P)} f(\xi_j)\,(g(\alpha_j)-g(\alpha_{j-1})) - I\right| < \varepsilon$$

holds for each δ-fine partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$. In this context, the symbol $|z|$ denotes the absolute value of a complex number z.

6.10.3 Exercise. Let $f,g:[a,b]\to\mathbb{C}$ be such that the integral $\int_a^b f\,\mathrm{d}g$ exists. Show that

$$\overline{\int_a^b f\,\mathrm{d}g} = \int_a^b \bar{f}\,\mathrm{d}\bar{g},$$

where the symbol $\bar{z}$ denotes the complex conjugate of a complex number z.

6.11 Relation between Lebesgue-Stieltjes and Kurzweil-Stieltjes integrals

The goal of this section is to clarify the relationship between Kurzweil-Stieltjes integrals and Lebesgue-Stieltjes integrals. We assume some basic familiarity with measure theory and Lebesgue integration. Since the rest of the book makes no use of the results from this section, readers who are not interested in Lebesgue-Stieltjes integrals can skip this part.

First, let us recall some facts about Lebesgue-Stieltjes measures. More details can be found e.g. in Section 22 of [153].

Throughout this section, we always assume that $g:\mathbb{R}\to\mathbb{R}$ is nondecreasing. Then the outer Lebesgue-Stieltjes measure of an arbitrary set $E\subset\mathbb{R}$ is given by

$$\mu_g^*(E) := \inf\left\{\sum_{n=1}^{\infty}(g(b_n+)-g(a_n+)) : E\subset\bigcup_{n=1}^{\infty}(a_n,b_n]\right\}.$$

The outer measure is either a nonnegative real number, or ∞. The function μ_g^* is defined on the collection of all subsets of $\mathbb{R}$, but it need not be σ-additive. By restricting μ_g^* to a certain σ-algebra of sets that are called μ_g-measurable, we get the Lebesgue-Stieltjes measure μ_g, which is σ-additive. This measure has the following properties:

- Each Borel set (in particular, each open or closed set) is μ_g-measurable.
- The measures of various types of intervals are calculated as follows:

$$\begin{aligned}\mu_g([a,b]) &= g(b+)-g(a-),\\ \mu_g((a,b]) &= g(b+)-g(a+),\\ \mu_g([a,b)) &= g(b-)-g(a-),\\ \mu_g((a,b)) &= g(b-)-g(a+).\end{aligned}$$

  In particular, the measure of a singleton $\{a\}$ is $\mu_g(\{a\}) = g(a+)-g(a-)$.

- If $E\subset\mathbb{R}$ is μ_g-measurable and $\varepsilon>0$, there exists an open set G such that $E\subset G$ and $\mu_g(G\setminus E)<\varepsilon$.
- If $E\subset\mathbb{R}$ satisfies $\mu_g^*(E)=0$, then E is μ_g-measurable and $\mu_g(E)=0$.

In Lebesgue's integration theory, it is common to deal with functions whose values can be not only real numbers, but also $\pm\infty$. For this reason, we set $\mathbb{R}^* = \mathbb{R}\cup\{\pm\infty\}$.

A function $f:\mathbb{R}\to\mathbb{R}^*$ is called μ_g-measurable if the set $\{x\in\mathbb{R}: f(x)>r\}$ is μ_g-measurable for each $r\in\mathbb{R}$. A function $s:\mathbb{R}\to\mathbb{R}$ is called simple if it can be expressed in the form

$$s=\sum_{j=1}^{n} c_j\chi_{A_j},$$

where $c_1,\dots,c_n$ are distinct real numbers, and $A_1,\dots,A_n$ are disjoint subsets of $\mathbb{R}$. Obviously, s is μ_g-measurable if and only if the sets $A_1,\dots,A_n$ are μ_g-measurable. If s is μ_g-measurable and nonnegative (i.e., if $c_1,\dots,c_n\geq 0$), we define its Lebesgue-Stieltjes integral by the formula

$$\int_{\mathbb{R}} s\,\mathrm{d}\mu_g=\sum_{j=1}^{n} c_j\mu_g(A_j).$$

If $c_j=0$ and $\mu_g(A_j)=\infty$ for a certain $j\in\{1,\dots,n\}$, we use the convention that $0\cdot\infty=0$. The value of the integral is either a nonnegative real number, or ∞.

For each nonnegative μ_g-measurable function $f:\mathbb{R}\to\mathbb{R}^*$, there exists a nondecreasing sequence of simple nonnegative functions which is convergent to f. We define

$$\int_{\mathbb{R}} f\,\mathrm{d}\mu_g=\sup\left\{\int_{\mathbb{R}} s\,\mathrm{d}\mu_g\colon s \text{ is a simple } \mu_g\text{-measurable function with } 0\leq s\leq f\right\}.$$

Finally, for an arbitrary μ_g-measurable function $f:\mathbb{R}\to\mathbb{R}^*$, we define

$$\int_{\mathbb{R}} f\,\mathrm{d}\mu_g=\int_{\mathbb{R}} f^+\,\mathrm{d}\mu_g-\int_{\mathbb{R}} f^-\,\mathrm{d}\mu_g$$

whenever the difference on the right-hand side makes sense (i.e., it is not of the form $\infty-\infty$).

If $E\subset\mathbb{R}$ is a μ_g-measurable set, we let

$$\int_{E} f\,\mathrm{d}\mu_g=\int_{\mathbb{R}} f\chi_E\,\mathrm{d}\mu_g$$

whenever the last integral exists. Since the values of f outside E are unimportant, we can assume that f is defined only on E (and extend it to $\mathbb{R}$ in an arbitrary way).

If $g(x)=x$ for all $x\in\mathbb{R}$, then μ_g is simply the Lebesgue measure, and the Lebesgue-Stieltjes integral reduces to the ordinary Lebesgue integral.

Note that if $\int_{\mathbb{R}} f\,\mathrm{d}\mu_g$ exists and is finite, then $\int_{\mathbb{R}} |f|\,\mathrm{d}\mu_g$ also exists and is finite, because $|f|$ is μ_g-measurable, $|f|=f^++f^-$, and the integrals of f^+ and

f^- are finite. For this reason, we say that the Lebesgue-Stieltjes integral is *absolutely convergent*.

Our goal is to prove that Lebesgue-Stieltjes integrability implies Kurzweil-Stieltjes integrability. This fact will be a fairly straightforward consequence of the next lemma.

6.11.1 Lemma. *If $E \subset [a,b]$ is a μ_g-measurable set, then the Kurzweil-Stieltjes integral $\int_a^b \chi_E \, \mathrm{d}g$ exists, and we have*

$$\int_a^b \chi_E \, \mathrm{d}g = \int_{(a,b)} \chi_E \, \mathrm{d}\mu_g + \chi_E(a)\, \Delta^+ g(a) + \chi_E(b)\, \Delta^- g(b).$$

Proof. Let $\widetilde{E} = E \cap (a,b)$ and observe that

$$\int_{(a,b)} \chi_E \, \mathrm{d}\mu_g = \int_{\mathbb{R}} \chi_{\widetilde{E}} \, \mathrm{d}\mu_g = \mu_g(\widetilde{E}).$$

Since $E \cap \{a\}$ and $E \cap \{b\}$ are either empty or singleton sets, we have

$$\int_a^b \chi_{E\cap\{a\}} \, \mathrm{d}g = \chi_E(a)\, \Delta^+ g(a), \qquad \int_a^b \chi_{E\cap\{b\}} \, \mathrm{d}g = \chi_E(b)\, \Delta^+ g(b).$$

Thus, it suffices to show that

$$\int_a^b \chi_{\widetilde{E}} \, \mathrm{d}g = \mu_g(\widetilde{E}). \tag{6.11.1}$$

Let $H = (a,b) \setminus \widetilde{E}$. Consider an arbitrary $\varepsilon > 0$. It is well known (cf. e.g. Lemma 22.10 in [153]) that there exist open sets $G_1, G_2 \subset \mathbb{R}$ such that

$$\widetilde{E} \subset G_1, \quad \mu_g(G_1) < \mu_g(\widetilde{E}) + \varepsilon, \quad H \subset G_2, \quad \mu_g(G_2) < \mu_g(H) + \frac{\varepsilon}{2}.$$

Let δ be a gauge on $[a,b]$ with the following properties:

- $\delta(x) < \operatorname{dist}(x, \mathbb{R} \setminus G_1)$ for every $x \in \widetilde{E}$, and $\delta(x) < \operatorname{dist}(x, \mathbb{R} \setminus G_2)$ for every $x \in H$.
- $\delta(x) < x - a$ for each $x \in (a,b]$, and $\delta(x) < b - x$ for each $x \in [a,b)$.
- If $x \in (a, a+\delta(a))$, then $g(x) - g(a+) < \varepsilon/4$; if $x \in (b - \delta(b), b)$, then $g(b-) - g(x) < \varepsilon/4$.

Let $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ be a δ-fine partition of $[a,b]$. The first property from the definition of δ ensures that if $\xi_i\in\widetilde{E}$, then $[\alpha_{i-1},\alpha_i]\subset G_1$, and if $\xi_i\in H$, then $[\alpha_{i-1},\alpha_i]\subset G_2$. The second property guarantees that $\xi_1=a$ and $\xi_{\nu(P)}=b$. We can write

$$\bigcup_{\xi_i\in\widetilde{E}}[\alpha_{j-1},\alpha_j]=\bigcup_{j\in J}[\beta_{j-1},\beta_j],$$

where the right-hand side is a finite union of disjoint intervals contained in G_1 (note that the intervals on the left-hand side are nonoverlapping, but in general not disjoint). Consequently, we get

$$\begin{aligned}S(\chi_{\widetilde{E}},P)&=\sum_{\xi_i\in\widetilde{E}}(g(\alpha_i)-g(\alpha_{i-1}))=\sum_{j\in J}(g(\beta_j)-g(\beta_{j-1}))\\&\le\sum_{j\in J}(g(\beta_j+)-g(\beta_{j-1}-))=\sum_{j\in J}\mu_g([\beta_{j-1},\beta_j])\\&\le\mu_g(G_1)<\mu_g(\widetilde{E})+\varepsilon.\end{aligned}$$

To obtain a lower bound for $S(\chi_{\widetilde{E}},P)$, we write

$$\bigcup_{\xi_i\in H}[\alpha_{i-1},\alpha_i]=\bigcup_{k\in K}[\gamma_{k-1},\gamma_k],$$

where the right-hand side is a finite union of disjoint intervals contained in G_2. Recalling that $\xi_1=a\notin\widetilde{E}$ and $\xi_{\nu(P)}=b\widetilde{E}$, we get

$$\begin{aligned}S(\chi_{\widetilde{E}},P)&=g(b)-g(a)-\sum_{\xi_i\in[a,b]\setminus\widetilde{E}}(g(\alpha_i)-g(\alpha_{i-1}))\\&=g(b)-g(a)\\&\quad-\left(g(\alpha_1)-g(\alpha_0)+\sum_{\xi_i\in H}\big(g(\alpha_i)-g(\alpha_{i-1})\big)+g(\alpha_{\nu(P)})-g(\alpha_{\nu(P)-1})\right)\\&=g(\alpha_{\nu(P)-1})-g(\alpha_1)-\sum_{\xi_i\in H}(g(\alpha_i)-g(\alpha_{i-1}))\\&=g(\alpha_{\nu(P)-1})-g(\alpha_1)-\sum_{k\in K}(g(\gamma_k)-g(\gamma_{k-1}))\\&>g(b-)-\frac{\varepsilon}{4}-g(a+)-\frac{\varepsilon}{4}-\sum_{k\in K}(g(\gamma_k+)-g(\gamma_{k-1}-))\\&\ge g(b-)-g(a+)-\frac{\varepsilon}{2}-\sum_{k\in K}\mu_g([\gamma_{k-1},\gamma_k])\\&\ge\mu_g((a,b))-\frac{\varepsilon}{2}-\mu_g(G_2)\ge\mu_g((a,b))-\mu_g(H)-\varepsilon=\mu_g(\widetilde{E})-\varepsilon.\end{aligned}$$

We have proved that

$$|S(\chi_{\widetilde{E}}, P) - \mu_g(\widetilde{E})| < \varepsilon$$

for each δ-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$, which implies that the identity (6.11.1) holds and the proof is complete. □

Using the definition of a simple function and its Lebesgue-Stieltjes integral, we get the next result.

6.11.2 Corollary. *If $s : [a, b] \to \mathbb{R}$ is a μ_g-measurable simple function, then the Kurzweil-Stieltjes integral $\int_a^b s \,\mathrm{d}g$ exists, and we have*

$$\int_a^b s \,\mathrm{d}g = \int_{(a,b)} s \,\mathrm{d}\mu_g + s(a)\,\Delta^+ g(a) + s(b)\,\Delta^- g(b). \tag{6.11.2}$$

6.11.3 Theorem. *If $f : [a, b] \to \mathbb{R}$ and the Lebesgue-Stieltjes integral $\int_{(a,b)} f \,\mathrm{d}\mu_g$ has a finite value, then the Kurzweil-Stieltjes integral $\int_a^b f \,\mathrm{d}g$ exists, as well, and*

$$\int_a^b f \,\mathrm{d}g = \int_{(a,b)} f \,\mathrm{d}\mu_g + f(a)\,\Delta^+ g(a) + f(b)\,\Delta^- g(b). \tag{6.11.3}$$

If $g(a+) = g(a)$ and $g(b+) = g(b)$, then

$$\int_a^b f \,\mathrm{d}g = \int_{(a,b]} f \,\mathrm{d}\mu_g. \tag{6.11.4}$$

If $g(a-) = g(a)$ and $g(b-) = g(b)$, then

$$\int_a^b f \,\mathrm{d}g = \int_{[a,b)} f \,\mathrm{d}\mu_g. \tag{6.11.5}$$

Proof. If the Lebesgue-Stieltjes integral exists, then f (which is considered to be zero outside $[a, b]$) is necessarily μ_g-measurable. Each of the functions f^+, f^- is nonnegative and μ_g-measurable, and therefore it is the limit of a nondecreasing sequence of nonnegative μ_g-measurable simple functions. The Kurzweil-Stieltjes and Lebesgue-Stieltjes integrals of these simple functions exist and satisfy the relation (6.11.2). Using the monotone convergence theorems for the Kurzweil-Stieltjes and Lebesgue-Stieltjes integrals, we get

$$\int_a^b f^+ \,\mathrm{d}g = \int_{(a,b)} f^+ \,\mathrm{d}\mu_g + f^+(a)\,\Delta^+ g(a) + f^+(b)\,\Delta^- g(b),$$
$$\int_a^b f^- \,\mathrm{d}g = \int_{(a,b)} f^- \,\mathrm{d}\mu_g + f^-(a)\,\Delta^+ g(a) + f^-(b)\,\Delta^- g(b),$$

which immediately implies (6.11.3).

If $g(a+)=g(a)$ and $g(b+)=g(b)$, observe that $\Delta^+g(a)=0$ and

$$f(b)\,\Delta^-g(b)=f(b)\,(g(b+)-g(b-))=f(b)\,\mu_g(\{b\})=\int_{\{b\}} f\,\mathrm{d}\mu_g.$$

Similarly, if $g(a-)=g(a)$ and $g(b-)=g(b)$, then $\Delta^-g(b)=0$ and

$$f(a)\,\Delta^+g(a)=f(a)\,(g(a+)-g(a-))=f(a)\,\mu_g(\{a\})=\int_{\{a\}} f\,\mathrm{d}\mu_g.$$

These facts together with (6.11.3) imply (6.11.4) and (6.11.5), respectively. □

6.11.4 Exercise. In Section 6.9 we have introduced Kurzweil-Stieltjes integrals of the form $\int_E f\,\mathrm{d}g$, where $f,g:[a,b]\to\mathbb{R}$ and E is an elementary subset of $[a,b]$. Suppose that we extend g to $\mathbb{R}$ in such a way that $g(a-)=g(a)$ and $g(b+)=g(b)$. Show that if $I\subset[a,b]$ is an interval of an arbitrary type and the Lebesgue-Stieltjes integral $\int_I f\,\mathrm{d}\mu_g$ exists, then the Kurzweil-Stieltjes integral $\int_I f\,\mathrm{d}g$ exists as well and has the same value. Conclude that the same statement holds if I is replaced by an elementary set $E\subset[a,b]$. Let $f:[a,b]\to\mathbb{R}$, $g:\mathbb{R}\to\mathbb{R}$ and a subinterval I of $[a,b]$ be given and let $g(a-)=g(a)$ and $g(b+)=g(b)$. Assume that the Lebesgue-Stieltjes integral $\int_I f\,\mathrm{d}g$ exists. Then the Kurzweil-Stieltjes integral $\int_I f\,\mathrm{d}g$ exists as well, and both integrals have the same value.

Our next goal is to prove a partial converse to Theorem 6.11.3 and show that for nonnegative functions f, Kurzweil-Stieltjes integrability implies Lebesgue-Stieltjes integrability. We begin by establishing two auxiliary results.

6.11.5 Lemma. *Let $f:[a,b]\to\mathbb{R}$ be a nonnegative function such that the Kurzweil-Stieltjes integral $\int_a^b f\,\mathrm{d}g$ exists, and $E=\{x\in[a,b]: f(x)\geq 1\}$. Then for each $\varepsilon>0$, there exists a μ_g-measurable set $G\subset[a,b]$ containing E and such that $\int_a^b \chi_G\,\mathrm{d}g\leq\varepsilon+\int_a^b f\,\mathrm{d}g$.*

Proof. We have $f=f\chi_{\{a\}}+f\chi_{(a,b]}$ and $E=E_1\cup E_2$, where $E_1=E\cap\{a\}$ and $E_2=E\cap(a,b]$. Note that both $\int_a^b f\chi_{\{a\}}\,\mathrm{d}g$ and $\int_a^b f\chi_{(a,b]}\,\mathrm{d}g$ exist. We will construct measurable sets $G_1,G_2\subset[a,b]$ such that $E_1\subset G_1$, $E_2\subset G_2$,

$$\int_a^b \chi_{G_1}\,\mathrm{d}g\leq\frac{\varepsilon}{2}+\int_a^b f\chi_{\{a\}}\,\mathrm{d}g,$$
$$\int_a^b \chi_{G_2}\,\mathrm{d}g\leq\frac{\varepsilon}{2}+\int_a^b f\chi_{(a,b]}\,\mathrm{d}g.$$

Then $G=G_1\cup G_2$ contains E, and

$$\int_a^b \chi_G\,\mathrm{d}g\leq\int_a^b(\chi_{G_1}+\chi_{G_2})\,\mathrm{d}g\leq\varepsilon+\int_a^b f\,\mathrm{d}g.$$

If $f(a)<1$, it suffices to take $G_1=\emptyset$. Otherwise, if $f(a)\geq 1$, find a $c\in(a,b]$ such that $g(c)-g(a+)\leq\varepsilon/2$, and let $G_1=[a,c)$. Then $E_1\subset\{a\}\subset G_1$ and

$$\begin{aligned}\int_a^b \chi_{G_1}\,\mathrm{d}g &= g(c-)-g(a)=g(c-)-g(a+)+g(a+)-g(a)\\ &\leq g(c)-g(a+)+\frac{f(a)\,\Delta^+g(a)}{f(a)}\\ &= g(c)-g(a+)+\frac{1}{f(a)}\int_a^b f\chi_{\{a\}}\,\mathrm{d}g\\ &\leq \frac{\varepsilon}{2}+\int_a^b f\chi_{\{a\}}\,\mathrm{d}g.\end{aligned}$$

Note that g is right-continuous in $[a,b)$ with at most countably many exceptions. Hence, there exists a sequence of divisions $\{\boldsymbol{\alpha}^n\}$ of $[a,b]$ such that for each $n\in\mathbb{N}$, $\boldsymbol{\alpha}^{n+1}$ is a refinement of $\boldsymbol{\alpha}^n$, $|\boldsymbol{\alpha}^n|\leq(b-a)/2^n$, and g is right-continuous at the division points $\alpha_1^n,\ldots,\alpha_{\nu(\boldsymbol{\alpha})-1}^n$.

Let δ be a gauge on $[a,b]$ such that

$$|S(f\chi_{(a,b]},\mathrm{d}g,P)-\int_a^b f\chi_{(a,b]}\,\mathrm{d}g|<\varepsilon/2$$

for each δ-fine partition P of $[a,b]$. We now construct a collection $\mathcal{I}$ of interval-point pairs of the form $((u,v],\tau)$. In the beginning, let $\mathcal{I}=\emptyset$. For each $n\in\mathbb{N}$, perform the following step: Find all intervals $[\alpha_{j-1}^n,\alpha_j^n]$ which are not contained in any interval in $\mathcal{I}$ and such that there exists a point $\tau\in[\alpha_{j-1}^n,\alpha_j^n]\cap E_2$ satisfying $[\alpha_{j-1}^n,\alpha_j^n]\subset(\tau-\delta(\tau),\tau+\delta(\tau))$; then add $((\alpha_{j-1}^n,\alpha_j^n],\tau)$ to $\mathcal{I}$. If there are several possible choices of τ for a given $[\alpha_{j-1}^n,\alpha_j^n]$, take only one of these interval-point pairs.

This procedure leads to a collection $\mathcal{I}=\{((u_k,v_k],\tau_k): k\in K\}$, where $K\subset N$ is either finite or countable. All intervals in $\mathcal{I}$ are pairwise disjoint, and all points τ_k satisfy $f(\tau_k)\geq 1$ (because $\tau_k\in E_2$). Moreover, we have

$$E_2\subset\bigcup_{k\in K}(u_k,v_k].$$

Indeed, if $x\in E_2$, take $n\in\mathbb{N}$ such that $(b-a)/2^n<\delta(x)$. Since x belongs to $(\alpha_{j-1}^n,\alpha_j^n]$ for a certain $j\in\{1,\ldots,\nu(\boldsymbol{\alpha}^n)\}$, and (recall that $|\boldsymbol{\alpha}^n|\leq(b-a)/2^n$)

$$x-\frac{b-a}{2^n}\leq\alpha_j^n-\frac{b-a}{2^n}\leq\alpha_{j-1}^n\leq\alpha_j^n\leq\alpha_{j-1}^n+\frac{b-a}{2^n}\leq x+\frac{b-a}{2^n},$$

we see that that $[\alpha_{j-1}^n,\alpha_j^n]\subset(x-\delta(x),x+\delta(x))$. Thus, either $(\alpha_{j-1}^n,\alpha_j^n]$ was added to $\mathcal{I}$ in the n-th stage of its construction, or it is contained in an interval added earlier. In any case, $x\in\bigcup_{k\in K}(u_k,v_k]$.

If L is an arbitrary finite subset of K, then the collection $\{([u_k,v_k],\tau_k):k\in L\}$ can be extended to a δ-fine tagged partition P of $[a,b]$, and therefore

$$\sum_{k\in L}(g(v_k)-g(u_k))\leq\sum_{k\in L}f(\tau_k)\,(g(v_k)-g(u_k))$$
$$\leq S(f\chi_{(a,b]},\mathrm{d}g,P)<\frac{\varepsilon}{2}+\int_a^b f\chi_{(a,b]}\,\mathrm{d}g.$$

It follows that

$$\sum_{k\in K}(g(v_k)-g(u_k))\leq\frac{\varepsilon}{2}+\int_a^b f\chi_{(a,b]}\,\mathrm{d}g.$$

Let $G_2=\bigcup_{k\in K}(u_k,v_k]$. Since G_2 is μ_g-measurable, the integral $\int_a^b\chi_{G_2}\mathrm{d}g$ exists by Lemma 6.11.1, and we have

$$\int_a^b\chi_{G_2}\,\mathrm{d}g\leq\sum_{k\in K}\int_a^b\chi_{(u_k,v_k]}\,\mathrm{d}g.$$

If $v_k=b$, then $\int_a^b\chi_{(u_k,v_k]}\,\mathrm{d}g=g(v_k)-g(u_k+)$. If $v_k<b$, then g is right-continuous at v_k, and

$$\int_a^b\chi_{(u_k,v_k]}\,\mathrm{d}g=g(v_k+)-g(u_k+)=g(v_k)-g(u_k+).$$

Thus, we get

$$\int_a^b\chi_{G_2}\,\mathrm{d}g\leq\sum_{k\in K}\big(g(v_k)-g(u_k+)\big)\leq\sum_{k\in K}\big(g(v_k)-g(u_k)\big)\leq\frac{\varepsilon}{2}+\int_a^b f\chi_{(a,b]}\,\mathrm{d}g,$$

which completes the proof. □

6.11.6 Lemma. *Let $f:[a,b]\to\mathbb{R}$ be a nonnegative function such that the Kurzweil-Stieltjes integral $\int_a^b f\,\mathrm{d}g$ exists. Then for each $\varepsilon>0$, there exists a nonnegative μ_g-measurable function $\varphi:[a,b]\to\mathbb{R}$ satisfying $f(x)\leq\varphi(x)+\varepsilon$ for all $x\in[a,b]$, and $\int_a^b\varphi\,\mathrm{d}g\leq\varepsilon+\int_a^b f\,\mathrm{d}g$.*

Proof. Let $\varepsilon>0$ be given. Suppose first that $f(a)=f(b)=0$. For each $n\in\mathbb{N}\cup\{0\}$, let

$$f_n(x)=\begin{cases}0 & \text{if } f(x)\leq n\,\varepsilon,\\ f(x)-n\,\varepsilon & \text{if } f(x)\in[n\,\varepsilon,(n+1)\,\varepsilon],\\ \varepsilon & \text{if } f(x)\geq(n+1)\,\varepsilon,\end{cases}$$

and observe that

$$f(x)=\sum_{n=0}^{\infty} f_n(x) \quad \text{for } x\in[a,b]. \tag{6.11.6}$$

We claim that $\int_a^b f_n\,\mathrm{d}g$ exists for each $n\in\mathbb{N}\cup\{0\}$. To see this, note that

$$f_n(x)=\min\{\max\{f(x)-n\,\varepsilon,0\},\varepsilon\} \quad \text{for } x\in[a,b].$$

Since $|f(x)-n\,\varepsilon|\le f(x)+n\,\varepsilon$ and both $\int_a^b(f-n\,\varepsilon)\,\mathrm{d}g$ and $\int_a^b(f+n\,\varepsilon)\,\mathrm{d}g$ exist, Theorem 6.7.4 implies that $\int_a^b|f-n\,\varepsilon|\,\mathrm{d}g$ exists as well. According to Theorem 6.7.5, this means that $\int_a^b\max\{f-n\,\varepsilon,0\}\,\mathrm{d}g$ exists and consequently

$$\int_a^b \min\{\max\{f-n\,\varepsilon,0\},\varepsilon\}\,\mathrm{d}g$$

exists, too; this proves the claim.

Using (6.11.6) and Levi's theorem (Theorem 6.8.10 and Remark 6.8.11), we get

$$\int_a^b f\,\mathrm{d}g=\sum_{n=0}^{\infty}\int_a^b f_n\,\mathrm{d}g. \tag{6.11.7}$$

For each $n\in\mathbb{N}\cup\{0\}$, Lemma 6.11.5 implies the existence of a μ_g-measurable set $G_n\subset[a,b]$ such that

$$\{x\in[a,b]:f_n(x)=\varepsilon\}=\{x\in[a,b]:\frac{1}{\varepsilon}f_n(x)\ge 1\}\subset G_n,$$

and

$$\int_a^b \chi_{G_n}\,\mathrm{d}g\le\frac{1}{2^{n+1}}+\frac{1}{\varepsilon}\int_a^b f_n\,\mathrm{d}g. \tag{6.11.8}$$

Since $f(a)=f(b)=0$, we can assume that all the sets G_n are contained in the open interval (a,b). Denote

$$\psi(x)=\varepsilon\sum_{n=0}^{\infty}\chi_{G_n}(x) \quad \text{for } x\in[a,b].$$

Since the last sum need not be convergent, the function ψ takes values in $[0,\infty]$. Using Levi's theorem for the Lebesgue-Stieltjes integral, Theorem 6.11.3 and (6.11.8), we get

$$\int_{[a,b]}\psi\,\mathrm{d}\mu_g=\varepsilon\sum_{n=0}^{\infty}\int_{[a,b]}\chi_{G_n}\,\mathrm{d}\mu_g=\varepsilon\sum_{n=0}^{\infty}\int_a^b\chi_{G_n}\,\mathrm{d}g\le\varepsilon+\int_a^b f\,\mathrm{d}g.$$

The finiteness of the integral implies that ψ must be finite μ_g-almost everywhere in $[a,b]$, i.e., if $N=\{x\in[a,b]:\psi(x)=\infty\}$, then $\mu_g(N)=0$. It follows that the function $\chi_N\, f$ is μ_g-measurable. Consequently, the function $\varphi:[a,b]\to\mathbb{R}$ given by

$$\varphi=\chi_{[a,b]\setminus N}\,\psi+\chi_N\, f=\varepsilon\sum_{n=0}^{\infty}\chi_{G_n\setminus N}+\chi_N\, f$$

is μ_g-measurable, nonnegative, and everywhere finite. Let

$$H_n=\{x\in[a,b]: n\varepsilon<f(x)<(n+1)\varepsilon\}=\{x\in[a,b]: 0<f_n(x)<\varepsilon\}\text{ for }n\in\mathbb{N}\cup\{0\}.$$

Note that all the sets H_n are pairwise disjoint, and

$$f_n(x)\le\varepsilon\,(\chi_{G_n}(x)+\chi_{H_n}(x))\quad\text{for }x\in[a,b].$$

Thus, if $x\in[a,b]\setminus N$, we obtain

$$f(x)=\sum_{n=0}^{\infty}f_n(x)\le\varepsilon\sum_{n=0}^{\infty}\chi_{G_n}(x)+\varepsilon\sum_{n=0}^{\infty}\chi_{H_n}(x)\le\varphi(x)+\varepsilon.$$

On the other hand, if $x\in N$, then $f(x)=\varphi(x)$.

Since $\mu_g(N)=0$, we have $\int_{(a,b)}\chi_N f\,\mathrm{d}\mu_g=0$, and thus (by Theorem 6.11.3)

$$\int_a^b\chi_N f\,\mathrm{d}g=0.$$

Levi's theorem for the Kurzweil-Stieltjes integral and (6.11.8) imply

$$\int_a^b\varphi\,\mathrm{d}g=\varepsilon\sum_{n=0}^{\infty}\int_a^b\chi_{G_n\setminus N}\,\mathrm{d}g+\int_a^b\chi_N\, f\,\mathrm{d}g\le\varepsilon\sum_{n=0}^{\infty}\int_a^b\chi_{G_n}\,\mathrm{d}g\le\varepsilon+\int_a^b f\,\mathrm{d}g.$$

Therefore, we have shown that φ has all the required properties.

In a general case when $f(a)$ and $f(b)$ are arbitrary, consider the function $f\chi_{(a,b)}$. By the previous part of the proof, there is a μ_g-measurable function $\widetilde{\varphi}$ such that

$$f\chi_{(a,b)}\le\widetilde{\varphi}+\varepsilon\ \text{ and }\ \int_a^b\widetilde{\varphi}\,\mathrm{d}g\le\varepsilon+\int_a^b f\chi_{(a,b)}\,\mathrm{d}g.$$

Taking $\varphi=\widetilde{\varphi}+f\chi_{\{a,b\}}$, we see that

$$f=f\chi_{(a,b)}+f\chi_{\{a,b\}}\le\widetilde{\varphi}+\varepsilon+f\chi_{\{a,b\}}=\varphi+\varepsilon,$$

and $\int_a^b\varphi\,\mathrm{d}g\le\varepsilon+\int_a^b f\,\mathrm{d}g$. □

6.11.7 Theorem. *Let* $g:[a,b]\to\mathbb{R}$ *be given and let* $f:[a,b]\to\mathbb{R}$ *be a nonnegative function such that the Kurzweil-Stieltjes integral* $\int_a^b f\,\mathrm{d}g$ *exists. Then the Lebesgue-Stieltjes integral* $\int_{(a,b)} f\,\mathrm{d}\mu_g$ *also exists and is finite.*

Proof. First, assume that $\int_a^b f\,\mathrm{d}g=0$. Define

$$E_k=\{x\in(a,b):f(x)\geq 1/k\}\quad\text{for }\ k\in\mathbb{N}.$$

For each $\varepsilon>0$, Lemma 6.11.5 implies the existence of a μ_g-measurable subset G of $[a,b]$ containing the set $\{x\in[a,b]:k\,f(x)\geq 1\}$ and satisfying $\int_a^b \chi_G\,\mathrm{d}g\leq\varepsilon$. Observing that $E_k\subset G\cap(a,b)$ and using Lemma 6.11.1, we get

$$\mu_g^*(E_k)\leq\mu_g(G\cap(a,b))=\int_{(a,b)}\chi_{G\cap(a,b)}\,\mathrm{d}\mu_g$$
$$=\int_a^b \chi_{G\cap(a,b)}\,\mathrm{d}g\leq\int_a^b\chi_G\,\mathrm{d}g\leq\varepsilon.$$

Since $\varepsilon>0$ was arbitrary, we have $\mu_g^*(E_k)=0$. Hence, E_k is μ_g-measurable and $\mu_g(E_k)=0$. It follows that

$$0=\mu_g\left(\bigcup_{k=1}^{\infty}E_k\right)=\mu_g(\{x\in(a,b):f(x)>0\}).$$

Consequently, $f\chi_{(a,b)}$ is μ_g-measurable and $\int_{(a,b)} f\,\mathrm{d}\mu_g=0$.

Next, assume that $\int_a^b f\,\mathrm{d}g>0$. Using Lemma 6.11.6 with $\varepsilon=1/k$, $k\in\mathbb{N}$, we get a sequence of nonnegative μ_g-measurable functions $\{\varphi_k\}$ satisfying $f\leq\varphi_k+1/k$ and

$$\int_a^b\varphi_k\,\mathrm{d}g\leq\frac{1}{k}+\int_a^b f\,\mathrm{d}g,\qquad k\in\mathbb{N}.$$

Hence, $h=\liminf_{k\to\infty}\varphi_k$ is μ_g-measurable, $f\leq h$, and Fatou's lemma implies

$$\int_a^b h\,\mathrm{d}g\leq\liminf_{k\to\infty}\int_a^b\varphi_k\,\mathrm{d}g\leq\liminf_{k\to\infty}\left(\frac{1}{k}+\int_a^b f\,\mathrm{d}g\right)=\int_a^b f\,\mathrm{d}g\leq\int_a^b h\,\mathrm{d}g,$$

i.e., $\int_a^b(h-f)\,\mathrm{d}g=0$.

By the first part of the proof, the Lebesgue-Stieltjes integral $\int_{(a,b)}(h-f)\,\mathrm{d}\mu_g$ exists and equals zero. For each $k\in\mathbb{N}$, Theorem 6.11.3 implies

$$\int_{(a,b)}\varphi_k\,\mathrm{d}\mu_g\leq\int_a^b\varphi_k\,\mathrm{d}g.$$

Using Fatou's lemma for the Lebesgue-Stieltjes integral, we get

$$\int_{(a,b)} h \,\mathrm{d}\mu_g \le \liminf_{k\to\infty} \left(\int_{(a,b)} \varphi_k \,\mathrm{d}\mu_g \right) \le \liminf_{k\to\infty} \left(\int_a^b \varphi_k \,\mathrm{d}g \right) = \int_a^b h \,\mathrm{d}g,$$

which shows that $\int_{(a,b)} h \,\mathrm{d}\mu_g$ is finite. Consequently,

$$\int_{(a,b)} f \,\mathrm{d}\mu_g = \int_{(a,b)} h \,\mathrm{d}\mu_g - \int_{(a,b)} (h-f) \,\mathrm{d}\mu_g$$

exists and is finite. □

Theorems 6.11.3 and 6.11.7 show that for nonnegative functions f, the Kurzweil-Stieltjes integral $\int_a^b f \,\mathrm{d}g$ exists if and only if the Lebesgue-Stieltjes integral $\int_{(a,b)} f \,\mathrm{d}\mu_g$ exists and is finite; the relation between their values is given by formula (6.11.3).

6.12 Relation of the Kurzweil-Stieltjes integral to other Stieltjes-type integrals

In Section 6.2 we have already clarified the relationships between the Kurzweil-Stieltjes (KS) integral on one side and the Riemann-Stieltjes (RS) integrals (both (δ) and (σ)), or the Perron-Stieltjes integral on the other side. We have touched also the relation with the Newton integral. The relationship with the Lebesgue-Stieltjes integral was discussed in the previous section. Now, in addition, we will briefly outline the relationship with some of the other known integrals of Stieltjes type.

YOUNG INTEGRAL

Let $f:[a,b]\to\mathbb{R}$ and $g\in \mathrm{G}([a,b])$. Define

$$S_Y(P) = \sum_{j=1}^{\nu(P)} \Big(f(\alpha_{j-1})\,\Delta^+ g(\alpha_{j-1}) + f(\xi_j)\,[g(\alpha_j-)-g(\alpha_{j-1}+)] + f(\alpha_j)\Delta^- g(\alpha_j) \Big)$$

for every tagged partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$. We say that the (σ) *Young-integral* $(\sigma\mathrm{Y})\int_a^b f \,\mathrm{d}g$ exists and has a value $I\in\mathbb{R}$ if

$$\left.\begin{array}{l} \text{for every } \varepsilon>0 \text{ there is a division } \boldsymbol{\alpha}_\varepsilon \text{ of } [a,b] \text{ such that} \\ \quad |S_Y(P)-I|<\varepsilon \\ \text{holds for all partitions } P=(\boldsymbol{\alpha},\boldsymbol{\xi}) \text{ of } [a,b] \text{ such that } \boldsymbol{\alpha}\supset\boldsymbol{\alpha}_\varepsilon \text{ and} \\ \quad \alpha_{j-1}<\xi_j<\alpha_j \quad \text{for all } j\in\{1,\dots,\nu(\boldsymbol{\alpha})\}. \end{array}\right\}$$

Notice that the expression $S_Y(P)$ can be equivalently rewritten as follows:

$$\left.\begin{aligned} S_Y(P) = f(a)\,\Delta^+ g(a) + \sum_{j=1}^{\nu(P)-1} f(\alpha_j)\,\Delta\, g(\alpha_j) + f(b)\Delta^- g(b)\\ + \sum_{j=1}^{\nu(P)} f(\xi_j)\,[g(\alpha_j-) - g(\alpha_{j-1}+)] \end{aligned}\right\} \tag{6.12.1}$$

Integral sums of the form (6.12.1) were introduced by W. H. Young in [157], where related (δ)-type integrals were discussed, as well. To our knowledge, a systematic study of the (σ) Young integral was initiated by T. H. Hildebrandt in [53]. More details are available in Section II.19 of his monograph [55]. The (σ) Young integral is more general than the corresponding RS-integrals. If the function f is regulated on $[a,b]$ and g has a bounded variation on $[a,b]$, then it is known that the integral $(\sigma\text{Y})\int_a^b f\,\mathrm{d}g$ exists and coincides with the KS-integral (KS) $\int_a^b f\,\mathrm{d}g$ (cf. Schwabik [120] and [121]). However, proceeding similarly as in the proofs of Theorems 6.3.8 and 6.3.11, it is possible to extend this assertion as follows.

6.12.1 Proposition. *Suppose f and g are regulated on $[a,b]$ and at least one of them has a bounded variation on $[a,b]$. Then both integrals*

$$(\text{KS})\int_a^b f\,\mathrm{d}g \quad \textit{and} \quad (\sigma\text{Y})\int_a^b f\,\mathrm{d}g$$

exist and have the same value.

The proof essentially follows the ideas of Section 6.3 and we leave it to the reader as the following (somewhat more advanced) exercise. □

6.12.2 Exercise. Prove Proposition 6.12.1.
Hint:

- Verify that the formulas (6.3.1)–(6.3.10) from Examples 6.3.1 hold also for the (σ) Young integral.

- Using Exercise 2.1.12 (ii) show that the estimate

  $$|S_Y(P)| \le \|f\|_\infty \operatorname{var}_a^b g$$

  holds for all partitions P of $[a,b]$ and all $f, g : [a,b] \to \mathbb{R}$.

- For $f : [a,b] \to \mathbb{R}$, $g \in \mathrm{G}([a,b])$ and $\alpha, \xi, \beta \in [a,b]$ such that $a \le \alpha \le \xi \le \beta \le b$, verify the equality

  $$\begin{aligned} f(\alpha)\,[g(\alpha+) - g(\alpha)] + f(\xi)\,[g(\beta-) - g(\alpha+)] + f(\beta)\,[g(\beta) - g(\beta-)]\\ = [f(\alpha) - f(\xi)]\,g(\alpha+) + [f(\xi) - f(\beta)]\,g(\beta-) + f(\beta)\,g(\beta) - f(\alpha)\,g(\alpha). \end{aligned}$$

Having this in mind, it is easy to see that the estimate

$$|S_Y(P)| \le (|f(a)| + |f(b)| + \text{var}_a^b f)\, \|g\|_\infty$$

is valid for all partitions P of $[a,b]$ and all $f, g : [a,b] \to \mathbb{R}$.

- Notice that due to the previous two steps, the estimates

$$\Big|(\sigma Y)\int_a^b f\, \mathrm{d}g\Big| \le \|f\|_\infty\, \text{var}_a^b g$$

and

$$\Big|(\sigma Y)\int_a^b f\, \mathrm{d}g\Big| \le 2\, \|f\|_{\text{BV}}\, \|g\|_\infty$$

hold whenever the corresponding integrals exist.

- Modify the proofs of Theorems 6.3.7 and 6.3.10 to show that their analogues are true also for the (σ) Young integral.

- Finally, complete the proof of Proposition 6.12.1 by proceeding as in the proofs of Theorems 6.3.8 and 6.3.11.

6.12.3 Remark. (i) If g is regulated on $[a,b]$ and

$$g(a) = g(t-) = g(s+) = g(b) \quad \text{for } t \in (a,b],\ s \in [a,b), \tag{6.12.2}$$

then, by (6.12.1), $S_Y(P) = 0$ for every function $f : [a,b] \to \mathbb{R}$ and every partition P of $[a,b]$, i.e., $(\sigma Y)\int_a^b f\, \mathrm{d}g = 0$. In general, this is no longer true for the KS-integral, as shown by the following example taken from [120] (see Example 2.1 there): Let

$$t_k = \frac{1}{k+1} \quad \text{for } k \in \mathbb{N} \quad \text{and} \quad g(t) = \begin{cases} 2^{-k} & \text{if } t = t_k \text{ for some } k \in \mathbb{N}, \\ 0 & \text{for } t \in [0,1] \setminus \{t_k : k \in \mathbb{N}\}, \end{cases}$$

and

$$f(t) = \begin{cases} 2^{k} & \text{if } t = t_k \text{ for some } k \in \mathbb{N}, \\ 0 & \text{for } t \in [0,1] \setminus \{t_k : k \in \mathbb{N}\}, \end{cases}$$

Evidently, g has a bounded variation on $[0,1]$ and $g(a) = g(t-) = g(s+) = g(b)$ for $t \in (a,b]$ and $s \in [a,b)$. (Notice that f is not regulated on $[0,1]$ as $f(0+)$ does not exist.)

Consider an arbitrary gauge δ and let $\ell \in \mathbb{N}$ be such that $t_\ell \in (0, \delta(0))$. We can choose a δ-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[0,1]$ in such a way that

$$\alpha_0 = \xi_1 = 0,\ \ \alpha_1 = \xi_2 = t_\ell \ \ \text{and} \ \ g(\alpha_j) = 0 \ \text{for } j \in \{2, \ldots, \nu(\boldsymbol{\alpha})\}.$$

Then

$$S(P)=f(\xi_2)\,[g(\alpha_2)-g(\alpha_1)]=-f(t_\ell)\,g(t_\ell)=-1$$

On the other hand, we can choose a δ-fine partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ in such a way that $\boldsymbol{\alpha}\cap\{t_k\}=\emptyset$. As $S(P)=0$ for such partitions, it follows that the KS-integral $\int_a^b f\,\mathrm{d}g$ does not exist.

(ii) Of course, if $g(x)=x$, then the Dirichlet function f_D (cf. Remark 6.2.4) is not (σ) Young integrable with respect to g. Thus, the existence of the KS-integral does not in general imply the existence of the (σ) Young integral.

SCHWABIK'S MODIFIED KS-INTEGRAL

To avoid the trouble illustrated by the first part of the above remark, Š. Schwabik introduced in [121] the notion of the modified KS-integral in the following way:

We say $I\in\mathbb{R}$ is the modified KS-integral of f with respect to g if for each $\varepsilon>0$ we can find a gauge δ such that $|S(P)-I|<\varepsilon$ holds for each δ-fine partition P of $[a,b]$ such that

$$a\le\xi_1<\alpha_1,$$
$$\alpha_{j-1}<\xi_j<\alpha_j \quad \text{for all } j\in\{2,\dots,\nu(\boldsymbol{\alpha})-1\}$$

and

$$\alpha_{\nu(\boldsymbol{\alpha})-1}<\xi_{\nu(\boldsymbol{\alpha})}\le b.$$

6.12.4 Exercises. (i) Show that in the example described in the first part of Remark 6.12.3, the modified KS-integral of f with respect to g equals 0.

(ii) Let $g(a)=g(t+)=g(s-)=g(b)$ for all $t\in(a,b]$ and all $s\in[a,b)$. Prove that the modified KS-integral of f with respect to g equals 0 for each $f:[a,b]\to\mathbb{R}$.

So far, there exists no systematic treatment of the modified KS-integral. However, it is evident that the modified KS-integral exists whenever the KS-integral exists and in such a case they have the same value. The applications presented in this book show that the KS-integral is sufficient for many purposes.

KREJČÍ'S KN-INTEGRAL

Recently (see [71]), P. Krejčí modified the notion of the KS-integral so that his integral, called the KN-integral, fully covers not only the (σ) Young integral but also Schwabik's modified KS-integral. His definition is based on a skillful reduction of the set of permissible partitions. We can formulate it as follows:

Let $f,g:[a,b]\to\mathbb{R}$ and $I\in\mathbb{R}$. We write

$$\text{(KN)}\int_a^b f\,\mathrm{d}g=I$$

if for every $\varepsilon > 0$ there are a gauge δ on $[a,b]$ and a countable set $A \subset [a,b]$ such that the inequality

$$|S(P) - I| < \varepsilon$$

holds for every δ-fine partition P such that none of its tags belongs to the set A.

DUSHNIK (INTERIOR) INTEGRAL

We say that the (σ) *Dushnik-integral* $(\sigma\mathrm{D}) \int_a^b f\, \mathrm{d}g$ exists and equals $I \in \mathbb{R}$ if

$$\left.\begin{array}{l} \text{for every } \varepsilon > 0 \text{ there is a division } \boldsymbol{\alpha}_\varepsilon \text{ of } [a,b] \text{ such that} \\ \quad |S(P) - I| < \varepsilon \\ \text{holds for all partitions } P = (\boldsymbol{\alpha}, \boldsymbol{\xi}) \text{ of } [a,b] \text{ such that } \boldsymbol{\alpha} \supset \boldsymbol{\alpha}_\varepsilon \text{ and} \\ \quad \alpha_{j-1} < \xi_j < \alpha_j \quad \text{for all } j \in \{1, \ldots, \nu(\boldsymbol{\alpha})\}. \end{array}\right\}$$

The Dushnik integral is also known as the *interior* integral, cf. Dushnik [31].

6.12.5 Exercise. Assume that f is the Dirichlet function and $g(x) = x$ on $[a,b]$. Show that the (σ) Dushnik integral $(\sigma\mathrm{D}) \int_a^b f\, \mathrm{d}g$ does not exist.

Thus, unlike the modified KS-integral and KN-integral, the (σ) Dushnik integral does not fully cover the KS-integral. Still, this concept of integral is sufficiently general for many purposes, because $(\sigma D) \int_a^b f\, \mathrm{d}g$ exists if f, g are regulated and one of them has a bounded variation. However, the value of the integral is in general different from the KS-integral. This is evident from the next exercise, which should be compared with Examples 6.3.1.

6.12.6 Exercises. (i) For any function $f \in \mathrm{G}([a,b])$, show that the following relations are true:

$$(\sigma\mathrm{D}) \int_a^b f\, \mathrm{d}\chi_{(\tau,b]} = f(\tau+) \qquad \text{if } \tau \in [a,b),$$

$$(\sigma\mathrm{D}) \int_a^b f\, \mathrm{d}\chi_{[\tau,b]} = f(\tau-) \qquad \text{if } \tau \in (a,b],$$

$$(\sigma\mathrm{D}) \int_a^b f\, \mathrm{d}\chi_{[a,\tau]} = -f(\tau+) \quad \text{if } \tau \in [a,b),$$

$$(\sigma\mathrm{D}) \int_a^b f\, \mathrm{d}\chi_{[a,\tau)} = -f(\tau-) \quad \text{if } \tau \in (a,b],$$

and

$$(\sigma\mathrm{D}) \int_a^b f\, \mathrm{d}\chi_{[\tau]} = \begin{cases} -f(a+) & \text{if } \tau = a, \\ -\Delta f(\tau) & \text{if } \tau \in (a,b), \\ f(b-) & \text{if } \tau = b. \end{cases}$$

(ii) For any function $g:[a,b]\to\mathbb{R}$, show that the following relations are true:

$$(\sigma\mathrm{D})\int_a^b \chi_{(\tau,b]}\,\mathrm{d}g = g(b)-g(\tau) \quad \text{if } \tau\in[a,b),$$

$$(\sigma\mathrm{D})\int_a^b \chi_{[\tau,b]}\,\mathrm{d}g = g(b)-g(\tau) \quad \text{if } \tau\in(a,b],$$

$$(\sigma\mathrm{D})\int_a^b \chi_{[a,\tau]}\,\mathrm{d}g = g(\tau)-g(a) \quad \text{if } \tau\in[a,b),$$

$$(\sigma\mathrm{D})\int_a^b \chi_{[a,\tau)}\,\mathrm{d}g = g(\tau)-g(a) \quad \text{if } \tau\in(a,b],$$

and

$$(\sigma\mathrm{D})\int_a^b \chi_{[\tau]}\,\mathrm{d}g = 0 \qquad \text{for } \tau\in[a,b].$$

The next result provides more information on the relation between (σ) Dushnik and KS-integrals.

6.12.7 Proposition. *Assume that f and g are regulated on $[a,b]$ and at least one of them has a bounded variation on $[a,b]$. Then both integrals*

$$(\sigma\mathrm{D})\int_a^b f\,\mathrm{d}g \quad \textit{and} \quad (\mathrm{KS})\int_a^b g\,\mathrm{d}f$$

exist and the equality

$$(\sigma\mathrm{D})\int_a^b f\,\mathrm{d}g + (\mathrm{KS})\int_a^b g\,\mathrm{d}f = f(b)\,g(a) - f(a)\,g(a) \tag{6.12.3}$$

holds.

Similarly to Proposition 6.12.1, the proof is left as an exercise to the reader.

6.12.8 Exercise. Prove Proposition 6.12.7.
Hint: Follow the steps of Exercise 6.12.2 with the (σ) Young integral replaced by the (σ) Dushnik integral. Take into account the results of Exercises 6.12.6 and observe that (6.12.3) holds if one of the functions f,g is regulated and the other is a finite step function.

Notice that, combining relation (6.12.3) with the integration-by-parts formula (6.4.2), we obtain the equality

$$(\sigma\mathrm{D})\int_a^b f\,\mathrm{d}g = (\mathrm{KS})\int_a^b f\,\mathrm{d}g - \sum_{a\le x\le b}\Big(\Delta^- f(x)\,\Delta^- g(x) - \Delta^+ f(x)\,\Delta^+ g(x)\Big)$$

valid whenever f and g are regulated on $[a,b]$ and at least one of them has a bounded variation on $[a,b]$.

We point out that a relation between the (σ) Young and (σ) Dushnik integrals analogous to Proposition 6.12.7 has been already known for a long time, cf. Theorem B in MacNerney [96] or Theorem 4.7 in Hönig [61].

A noteworthy study of the Dushnik integral is contained in the monograph [60] by Ch.S. Hönig, who extended its definition to functions with values in Banach spaces and subsequently used it to develop the theory of abstract Volterra-Stieltjes integral equations.

INTEGRATION IN ABSTRACT SPACES

The extension of integration to vector and matrix functions was shown in Section 6.10. One can act analogously even in the case of abstract functions, i.e. functions with values in Banach spaces. If X is a Banach space and $\mathscr{L}(X)$ is the corresponding Banach space of continuous linear operators on X and

$$F:[a,b]\to\mathscr{L}(X),\ G:[a,b]\to\mathscr{L}(X),\ g:[a,b]\to X,$$

then we can define KS-integrals

$$\int_a^b \mathrm{d}F\,g,\quad \int_a^b F\,\mathrm{d}g,\quad \int_a^b \mathrm{d}F\,G,\quad \int_a^b F\,\mathrm{d}G.$$

For example, $\int_a^b \mathrm{d}F\,g=I\in X$ if for every $\varepsilon>0$ there is a gauge δ on $[a,b]$ such that

$$\Big\|\sum_{j=1}^{\nu(P)} F(\xi_j)\,\big[g(\alpha_j)-g(\alpha_{j-1})\big]-I\,\Big\|_X<\varepsilon$$

holds for every δ-fine partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$. The notion of the variation can be easily transferred to abstract functions, too. For a function $f:[a,b]\to X$ and a division $\boldsymbol{\alpha}$ of the interval $[a,b]$, we define

$$V(f,\boldsymbol{\alpha})=\sum_{j=1}^{\nu(\boldsymbol{\alpha})}\|f(\alpha_j)-f(\alpha_{j-1})\|_X$$

and

$$\operatorname{var}_a^b f=\sup\{V(f,\boldsymbol{\alpha}):\boldsymbol{\alpha}\in\mathscr{D}[a,b]\}.$$

It is also obvious how to define the space $\mathrm{G}([a,b],X)$ of regulated functions with values in X. Then e.g. both integrals

$$\int_a^b \mathrm{d}F\,G \quad\text{and}\quad \int_a^b F\,\mathrm{d}G$$

exist if $F \in \mathrm{BV}([a,b],\mathscr{L}(X))$ and $G \in \mathrm{G}([a,b],\mathscr{L}(X))$ and most of the statements known for the integration of scalar functions (see [124], [128] and [102]) hold. There are some exceptions though: e.g. Corollary 6.5.2 holds only if the space X has finite dimension. This means that, inter alia, there are certain problems with the transfer of e.g. the Substitution Theorem to abstract integrals. In this brief information, it is worth mentioning that if the space X does not have a finite dimension, it makes sense to consider, instead of variation, then in general weaker notion of *semivariation* which is defined as follows:

For a given function $F:[a,b] \to \mathscr{L}(X)$ and a division $\boldsymbol{\alpha}$ of $[a,b]$, set first

$$V_a^b(F,\boldsymbol{\alpha}) = \sup\Big\{\Big\| \sum_{j=1}^{\nu(\boldsymbol{\alpha})} [F(\alpha_j) - F(\alpha_{j-1})]\, x_j \Big\|_X\Big\}$$

where the supremum is taken over all choices of elements $x_j \in X, j \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$ such that $\|x_j\|_X \le 1.$ Then the number

$$(\mathcal{B})\,\mathrm{var}_a^b F = \sup\{V_a^b(F,\boldsymbol{\alpha}) : \boldsymbol{\alpha} \in \mathscr{D}[a,b]\}$$

is called the *semivariation* of the function F on $[a,b]$ (see e.g. [60]). The assumptions concerning bounded variation can be usually (not always) weakened to bounded semivariation. It is known that if X has finite dimension, then the notions of variation and semivariation coincide.

Finally, let us note that the integration of functions with values in Hilbert, or reflexive Banach spaces, is useful e.g. in the theory of hysteresis (see e.g. [76] or [77]).

Chapter 7

Generalized linear differential equations

7.1 Introduction

All integrals in this chapter are KS-integrals whose definition is extended to matrix valued functions (i.e. the functions mapping the interval $[a,b]$ into the space of matrices) in the sense of Section 6.10. As we already explained in Section 6.10, all properties of KS-integral as well as of both kinds of RS-integral, which we have proved so far for scalar functions, hold for the vector and matrix valued functions, too, if the original order of matrices is kept. Therefore, in the proofs, for any needed properties of functions and integrals, we will refer to the corresponding statements proved in the previous chapters for scalar functions.

The following definition introduces the spaces of vector and matrix valued functions that will be used throughout this chapter.

7.1.1 Definition. (i) $\mathrm{G}([a,b],\mathbb{R}^n)$ is a Banach space of the functions $f\colon[a,b]\to\mathbb{R}^n$, which are regulated on $[a,b]$. The norm on $\mathrm{G}([a,b],\mathbb{R}^n)$ is defined by

$$\|f\| = \sup_{t\in[a,b]} |f(t)| \quad \text{for } f\in\mathrm{G}([a,b],\mathbb{R}^n)$$

where $|f(t)|$ is the norm of the vector $f(t)$ in $\mathbb{R}^n$.
(ii) $\mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$ is a Banach space of the functions $F\colon[a,b]\to\mathscr{L}(\mathbb{R}^n)$ which have bounded variation on $[a,b]$. The norm on $\mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$ is defined by

$$\|F\|_{\mathrm{BV}} = |F(a)| + \operatorname{var}_a^b F \quad \text{for } F\in\mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n)),$$

where $\operatorname{var}_a^b F$ is defined as in Section 6.10 and $|F(a)|$ is the norm of the matrix $F(a)$ in $\mathscr{L}(\mathbb{R}^n)$.

The spaces $\mathrm{BV}([a,b],\mathbb{R}^n)$, $\mathrm{C}([a,b],\mathscr{L}(\mathbb{R}^n))$ and $\mathrm{C}([a,b],\mathbb{R}^n)$ and their norms are defined similarly. A set of functions $f\colon[a,b]\to\mathbb{R}^n$ with a derivative that is continuous on the interval $[a,b]$ is denoted by $\mathrm{C}^1([a,b],\mathbb{R}^n)$. As usual, we define

$$f'(a) = f'(a+) \quad\text{and}\quad f'(b) = f'(b-) \quad \text{for } f\in\mathrm{C}^1([a,b],\mathbb{R}^n).$$

The topic of this chapter are the equations of form

$$x(t) - x(s) - \int_s^t \mathrm{d}A\, x = f(t) - f(s) \tag{7.1.1}$$

where $t,\ s\in[a,b],\ A$ is an $n\times n$-matrix valued function, f is an n-vector valued function and we look for n-vector valued function x satisfying the following definition.

7.1.2 Definition. A function $x:[a,b]\to\mathbb{R}^n$ is a solution of equation (7.1.1) on $[a,b]$ if the integral $\int_a^b \mathrm{d}A\,x$ exists and equation (7.1.1) is satisfied for all $t,s\in[a,b]$.

The equation (7.1.1) is called *generalized linear differential equation*.

7.1.3 Remark. Let $t_0\in[a,b]$ be given and let x satisfy the equation

$$x(t)-x(t_0)-\int_{t_0}^{t}\mathrm{d}A\,x=f(t)-f(t_0) \tag{7.1.2}$$

for $t\in[a,b]$. Then for any $s\in[a,b]$, we have

$$x(s)=x(t_0)+\int_{t_0}^{s}\mathrm{d}A\,x+f(s)-f(t_0).$$

If we subtract this equation from (7.1.2), we will find out that (7.1.1) holds for all $t,s\in[a,b]$, i.e. x is a solution of equation (7.1.1). Thus, the function $x:[a,b]\to\mathbb{R}^n$ is a solution of equation (7.1.1) on $[a,b]$ if and only if for some fixed $t_0\in[a,b]$ it satisfies (7.1.2) on $[a,b]$.

7.2 Differential equations with impulses

The motivation for studying generalized differential equations are among others the problems with impulses. A range of practical problems actually involve perturbations that have negligible persistence time compared to the time of the whole process which however significantly affect the studied process. The suitable model for describing such processes is usually *differential equations with impulses*, i.e. differential equations whose solutions does not have to be neither smooth nor continuous.

The source of the models with impulses is mainly physics (e.g. the description of clock mechanisms, oscillations of electromechanical systems, radiation of electric or magnetic waves in the environment with rapidly changing parameters, stabilization of Kapitza's pendulum, optimal regulation by bang-bang method) but also medical science (distribution of medicinal substances in a body, strategy of impulse vaccination in epidemiological models, investigation of the effect of mass measles vaccination), population dynamics (models with rapid changes in the amount of some populations) or economics (trade models which admit rapid changes of prices).

The simplest idealization of impulse processes are the processes described by the linear differential equations on which linear impulses act in finite amount of firmly given points.

Assume

$$\left.\begin{aligned} & r \in \mathbb{N}, \quad a < \tau_1 < \cdots < \tau_r < b, \\ & P \in \mathrm{C}([a,b], \mathscr{L}(\mathbb{R}^n)), \quad q \in \mathrm{C}([a,b], \mathbb{R}^n), \\ & B_k \in \mathscr{L}(\mathbb{R}^n), \quad d_k \in \mathbb{R}^n \ \text{ for } k = 1, \dots, r. \end{aligned}\right\} \tag{7.2.1}$$

(In this chapter, the symbols like B_k, or d_k, stand also for matrices, or vectors.)

Denote

$$D = \{\tau_1, \tau_2, \dots, \tau_r\}, \ \tau_0 = a, \ \tau_{r+1} = b$$

and, for a given regulated function $x: [a,b] \to \mathbb{R}^n$, define

$$\left.\begin{aligned} & x_{[1]}(t) = x(t) \quad \text{for } \ t \in [a, \tau_1] \\ & \text{and} \\ & x_{[k]}(t) = \begin{cases} x(\tau_{k-1}+) & \text{if } t = \tau_{k-1}, \\ x(t) & \text{if } t \in (\tau_{k-1}, \tau_k] \end{cases} \\ & \text{for } k \in \{2, 3, \dots, r{+}1\}. \end{aligned}\right\} \tag{7.2.2}$$

The linear impulse problem then consists of the linear differential equation

$$x' = P(t)\,x + q(t) \tag{7.2.3}$$

and linear impulse conditions

$$\Delta^+ x(\tau_k) = B_k\, x(\tau_k) + d_k, \quad k = 1, \dots, r \tag{7.2.4}$$

while the solution is defined by the following definition.

7.2.1 Definition. We say that a function $x: [a,b] \to \mathbb{R}^n$ is a solution to the impulse problem (7.2.3), (7.2.4) if

$$x_{[k]} \in \mathrm{C}^1([\tau_{k-1}, \tau_k]) \quad \text{for all } k \in \{1, \dots, r{+}1\}, \tag{7.2.5}$$

$$x'(t) = P(t)\,x(t) + q(t) \quad \text{for all } t \in [a,b] \setminus D \tag{7.2.6}$$

and x satisfies the impulse conditions (7.2.4).

7.2.2 Remark. Notice that a solution to problem (7.2.3), (7.2.4) always belongs to the space $\mathrm{G}([a,b], \mathbb{R}^n)$.

Now we will show that problem (7.2.3), (7.2.4) can be equivalently reformulated as a generalized linear differential equation of the form (7.1.2).

First, assume that $r=1$, and let $x:[a,b]\to\mathbb{R}^n$ be a solution of the impulse problem (7.2.3), (7.2.4). Integrating equation (7.2.6), we get

$$x(t)= x(a) + \int_a^t P(s)\,x(s)\,\mathrm{d}s + \int_a^t q(s)\,\mathrm{d}s \quad \text{for } t\in[a,\tau_1]$$

and

$$x(t)=x(\tau_1+) + \int_{\tau_1}^t P(s)\,x(s)\,\mathrm{d}s + \int_{\tau_1}^t q(s)\,\mathrm{d}s \quad \text{for } t\in(\tau_1,b].$$

Substituting (7.2.4) (where $k=r=1$) into the latter relation above, we get

$$\begin{aligned} x(t) &= x(\tau_1) + B_1\,x(\tau_1) + d_1 + \int_{\tau_1}^t P(s)\,x(s)\,\mathrm{d}s + \int_{\tau_1}^t q(s)\,\mathrm{d}s \\ &= x(a) + \int_a^t P(s)\,x(s)\,\mathrm{d}s + B_1\,x(\tau_1) + \int_a^t q(s)\,\mathrm{d}s + d_1, \end{aligned}$$

for $t\in(\tau_1,b]$ and therefore

$$\left.\begin{aligned} x(t) = x(a) + \int_a^t P(s)\,x(s)\,\mathrm{d}s + \chi_{(\tau_1,b]}(t)\,B_1\,x(\tau_1) \\ + \int_a^t q(s)\,\mathrm{d}s + \chi_{(\tau_1,b]}(t)\,d_1 \end{aligned}\right\} \text{for } t\in[a,b]. \tag{7.2.7}$$

For $t\in[a,b]$, set

$$A(t) = \int_a^t P(s)\,\mathrm{d}s + \chi_{(\tau_1,b]}(t)\,B_1 \quad\text{and}\quad f(t) = \int_a^t q(s)\,\mathrm{d}s + \chi_{(\tau_1,b]}(t)\,d_1$$

Then $A\in\mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$, $f\in\mathrm{BV}([a,b],\mathbb{R}^n)$ and

$$A(t-)=A(t) \quad\text{and}\quad f(t-)=f(t) \text{ for } t\in(a,b].$$

Moreover,

$$A(t+) = \int_a^t P(s)\,\mathrm{d}s + \chi_{[\tau_1,b]}(t)\,B_1 \quad\text{and}\quad f(t+) = \int_a^t q(s)\,\mathrm{d}s + \chi_{[\tau_1,b]}(t)\,d_1,$$

that is

$$\Delta^+ A(t) = \chi_{[\tau_1]}(t)\,B_1 \quad\text{and}\quad \Delta^+ f(t) = \chi_{[\tau_1]}(t)\,d_1 \quad \text{for } t\in[a,b).$$

By the Substitution Theorem 6.6.1 and formula (6.3.1) from Examples 6.3.1 (ii) (see also Examples 6.5.7), the equalities

$$\int_a^t \mathrm{d}A\,x = \int_a^t P(s)\,x(s)\,\mathrm{d}s + \chi_{(\tau_1,b]}(t)\,B_1\,x(\tau_1)$$

and

$$f(t) - f(a) = \int_a^t q(s)\,\mathrm{d}s + \chi_{(\tau_1,b]}(t)\,d_1$$

hold for $t \in [a,b]$ and $x \in \mathrm{G}([a,b],\mathbb{R}^n)$. Substituting to (7.2.7), we find out that x satisfies on $[a,b]$ (7.1.2), where $t_0 = a$.

Reversely, if $x \in \mathrm{G}([a,b],\mathbb{R}^n)$ verifies (7.1.2) on $[a,b]$, then (7.2.7) surely holds. Thus, if we define the functions $x_{[k]}$ as in (7.2.2), then (7.2.5) and (7.2.6) will be true. Furthermore, by Hake's Theorem (see also Exercise 6.5.6) $x(t-) = x(t)$ for each $t \in (a,b]$ and

$$\begin{aligned} x(t+) &= x(a) + \lim_{s\to t+} \int_a^s \mathrm{d}A\,x + f(t+) - f(a) \\ &= x(a) + \int_a^t \mathrm{d}A\,x + f(t) - f(a) + \Delta^+ A(t)\,x(t) + \Delta^+ f(t) \\ &= x(t) + \chi_{[\tau_1]}(t)\,\big(B_1\,x(t) + d_1\big) \quad \text{for every } t \in [a,b]. \end{aligned}$$

In particular, putting $t = \tau_1$, we find out that x meets the impulse condition (7.2.4) where $k = r = 1$.

Hence, by Remark 7.1.3, the problem (7.2.3), (7.2.4) is equivalent to generalized differential equation (7.1.1) if $r = 1$.

In the general case of $r \in \mathbb{N}$, we define

$$\left.\begin{aligned} A(t) &= \int_a^t P(s)\,\mathrm{d}s + \sum_{k=1}^r \chi_{(\tau_k,b]}(t)\,B_k \quad \text{for } t \in [a,b], \\ f(t) &= \int_a^t q(s)\,\mathrm{d}s + \sum_{k=1}^r \chi_{(\tau_k,b]}(t)\,d_k \quad \text{for } t \in [a,b]. \end{aligned}\right\} \tag{7.2.8}$$

By induction, we verify the following statement easily.

7.2.3 Theorem. *Assume* (7.2.1) *and* (7.2.8). *Then the impulse problem* (7.2.3), (7.2.4) *is equivalent to the generalized differential equation* (7.1.2), *i.e.* $x : [a,b] \to \mathbb{R}^n$ *is a solution of problem* (7.2.3), (7.2.4) *on* $[a,b]$ *if and only if it is a solution of equation* (7.1.1) *on* $[a,b]$.

□

7.3 Linear operators

Now, let us briefly recall some basic notions and results from functional analysis which we will need later. More detailed information can be found in the majority of the textbooks on functional analysis (see e.g. [70] or [115]). The basic overview is also included in the introduction part of [131].

Let X, Y be Banach spaces. The mapping $T: X \to Y$ is a *continuous operator* if

$$\lim_{n\to\infty} \|x_n - x\|_X = 0 \Longrightarrow \lim_{n\to\infty} \|T(x_n) - T(x)\|_Y = 0,$$

where $\|\cdot\|_X$ is the norm on X and $\|\cdot\|_Y$ is the norm on Y. The operator $L: X \to Y$ is called *linear* if

$$L(c_1\, x_2 + c_2\, x_2) = c_1\, L(x_1) + c_2\, L(x_2) \text{ holds for } x_1, x_2 \in X \text{ and } c_1, c_2 \in \mathbb{R}.$$

Moreover, we say that the linear operator L is *bounded* if there is a number $K \in [0, \infty)$ such that

$$\|L(x)\|_{\mathbb{Y}} \le K\, \|x\|_X \quad \text{for all } x \in X.$$

If L is a linear operator, then, as usual, we write $L\,x$ instead of $L(x)$. It is known that the linear operator $L: X \to Y$ is continuous if and only if it is bounded.

The set of linear bounded mappings of the space X into Y is denoted by $\mathscr{L}(X, Y)$. If $X = Y$, we write $\mathscr{L}(X)$ instead of $\mathscr{L}(X, X)$. On $\mathscr{L}(X, Y)$, the operations of adding the operators and multiplying the operators by a real number are established in an obvious way and $\mathscr{L}(X, Y)$ is then a Banach space with respect to the norm

$$L \in \mathscr{L}(X,Y) \to \|L\|_{\mathscr{L}(X,Y)} = \sup\big\{\|L\,x\|_Y : x \in X \text{ and } \|x\|_X \le 1\big\}.$$

It is known that the space $\mathscr{L}(\mathbb{R}^n)$ is equivalent with the space of matrices of form $n \times n$.

Finally, we say that $L \in \mathscr{L}(X, Y)$ is *compact* if it maps every set bounded in X onto a set which is relatively compact in Y, i.e. if for every sequence $\{x_n\}$ bounded in X, its value $\{L\,x_n\} \subset Y$ contains a subsequence that is convergent in Y. It is known that every compact linear operator is simultaneously continuous.

We will use the following two statements in the proofs of the main results of this chapter. The former one is a generalization of one of Fredholm's theorems known from the theory of integral equations. Its proof is included e.g. in the monographs by N. Dunford and J. T. Schwartz [30] or by M. Schechter [118].

7.3.1 Theorem (FREDHOLM ALTERNATIVE THEOREM)**.** *Let X be a Banach space and let the operator $L \in \mathscr{L}(X)$ be compact. Then the operator equation*

$$x - L\,x = g \tag{7.3.1}$$

has exactly one solution $x \in X$ *for every* $g \in X$ *if and only if the corresponding homogeneous equation*

$$x - L\,x = 0 \tag{7.3.2}$$

has only the trivial solution $x = 0 \in X$.

The second statement is known also from the elementary theory of matrices. Let us recall its general form borrowed from the monograph [140] (see Lemma 4.1-C).

7.3.2 Lemma. *Let* X *be a Banach space,* $L \in \mathscr{L}(X)$ *and* $\|L\|_{\mathscr{L}(X)} < 1$. *Then the operator* $[I - L]$ *has a bounded inverse* $[I - L]^{-1}$ *and the inequality*

$$\left\| [I - L]^{-1} \right\|_{\mathscr{L}(X)} \leq \frac{1}{1 - \|L\|_{\mathscr{L}(X)}}$$

is true.

7.4 Existence of solutions

Let us start our consideration of generalized linear differential equations by a simple observation based on known properties of the KS-integral.

7.4.1 Theorem. *Let* $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$ *and* $f \in \mathrm{G}([a,b], \mathbb{R}^n)$. *Then every solution* x *of equation* (7.1.1) *on* $[a,b]$ *is regulated on* $[a,b]$ *and satisfies the relations*

$$\left.\begin{aligned} &\Delta^- x(t) = \Delta^- A(t)\; x(t) + \Delta^- f(t) \quad \textit{for } t \in (a,b], \\ &\Delta^+ x(s) = \Delta^+ A(s)\, x(s) + \Delta^+ f(s) \quad \textit{for } s \in [a,b). \end{aligned}\right\} \tag{7.4.1}$$

P r o o f follows from Corollary 6.5.4 of Lemma 6.5.1 (Saks-Henstock Lemma). □

Thus, by virtue of Theorem 7.4.1, it is appropriate to look for solutions of generalized linear differential equations in the class $\mathrm{G}([a,b], \mathbb{R}^n)$.

The problem

$$x(t) - \widetilde{x} - \int_{t_0}^{t} \mathrm{d}A\, x = f(t) - f(t_0), \tag{7.4.2}$$

where the point $t_0 \in [a,b]$ and the vector $\widetilde{x} \in \mathbb{R}^n$ are given beforehand, is an analogue of the initial problems for linear ordinary differential equations.

7.4.2 Definition. A function $x\colon [a,b] \to \mathbb{R}^n$ is said to be a solution to the initial problem (7.4.2) on $[a,b]$ if equation (7.4.2) is satisfied for every $t \in [a,b]$.

7.4.3 Remark. Due to Remark 7.1.3, it is obvious that the function x is a solution to the initial problem (7.4.2) on $[a,b]$ if and only if it is a solution to equation (7.1.1) on $[a,b]$ and $x(t_0) = \widetilde{x}$.

For a given function $x \in \mathrm{G}([a,b],\mathbb{R}^n)$ and a point $t_0 \in [a,b]$, let the function $\mathscr{A}_{t_0} x$ be given by

$$(\mathscr{A}_{t_0} x)(t) = \int_{t_0}^{t} \mathrm{d}A\, x \quad \text{for } t \in [a,b]. \tag{7.4.3}$$

By Corollary 6.5.4 all the functions $\mathscr{A}_{t_0} x$ are regulated on $[a,b]$. The mapping

$$\mathscr{A}_{t_0} : x \in \mathrm{G}([a,b],\mathbb{R}^n) \to \mathscr{A}_{t_0} x \in \mathrm{G}([a,b],\mathbb{R}^n)$$

is obviously linear. Moreover, by Theorem 6.3.4 we have

$$\|\mathscr{A}_{t_0} x\| \le \left(\operatorname{var}_a^b A\right) \|x\| \;\text{ for all } x \in \mathrm{G}([a,b],\mathbb{R}^n).$$

Thus, for every $t_0 \in [a,b]$, $\mathscr{A}_{t_0}$ is a continuous linear operator on the space $\mathrm{G}([a,b],\mathbb{R}^n)$, i.e.

$$\mathscr{A}_{t_0} \in \mathscr{L}(\mathrm{G}([a,b],\mathbb{R}^n)).$$

Next, we will prove that (7.4.3) defines simultaneously a linear continuous operator mapping $\mathrm{G}([a,b],\mathbb{R}^n)$ into $\mathrm{BV}([a,b],\mathbb{R}^n)$.

7.4.4 Lemma. *Let* $A \in \mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$, $t_0 \in [a,b]$ *and let the function* $\mathscr{A}_{t_0} x$ *be defined for* $x \in \mathrm{G}([a,b],\mathbb{R}^n)$ *by* (7.4.3).

Then $\mathscr{A}_{t_0} x \in \mathrm{BV}([a,b],\mathbb{R}^n)$ *for every* $x \in \mathrm{G}([a,b],\mathbb{R}^n)$ *and the operator*

$$x \in \mathrm{G}([a,b],\mathbb{R}^n) \to \mathscr{A}_{t_0} x \in \mathrm{BV}([a,b],\mathbb{R}^n)$$

is bounded.

Proof. Let $\boldsymbol{\alpha}$ be an arbitrary division of the interval $[a,b]$. By Theorem 6.3.4

$$\begin{aligned}\sum_{j=1}^{\nu(\boldsymbol{\alpha})} \left|(\mathscr{A}_{t_0} x)(\alpha_j) - (\mathscr{A}_{t_0} x)(\alpha_{j-1})\right| &= \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \left| \int_{\alpha_{j-1}}^{\alpha_j} \mathrm{d}A\, x \right| \\ &\le \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \left(\operatorname{var}_{\alpha_{j-1}}^{\alpha_j} A\right) \|x\| = \left(\operatorname{var}_a^b A\right) \|x\|\end{aligned}$$

and

$$\left|(\mathscr{A}_{t_0}x)(a)\right| = \left|\int_{t_0}^{a} \mathrm{d}A\,x\right| \leq \left(\operatorname{var}_a^b A\right)\|x\|.$$

hold for every $x \in \mathrm{G}([a,b],\mathbb{R}^n)$. Hence $\mathscr{A}_{t_0}x \in \mathrm{BV}([a,b],\mathbb{R}^n)$ and

$$\|\mathscr{A}_{t_0}x\|_{\mathrm{BV}} = |(\mathscr{A}_{t_0}x)(a)| + \operatorname{var}_a^b(\mathscr{A}_{t_0}x) \leq 2\left(\operatorname{var}_a^b A\right)\|x\|$$

for every $x \in \mathrm{G}([a,b],\mathbb{R}^n)$. □

Using the operator $\mathscr{A}_{t_0}$ from (7.4.3), we can rewrite the initial problem (7.4.2) as the operator equation

$$x - \mathscr{A}_{t_0}x = g, \quad \text{where } g = \widetilde{x} + f - f(t_0).$$

Unfortunately, we do not have tools that would enable us to prove the compactness of the operator $\mathscr{A} \in \mathscr{L}(\mathrm{G}([a,b],\mathbb{R}^n))$. Therefore, we cannot apply the Fredholm Theorem (Theorem 7.3.1) directly and we have to proceed by a kind of indirect route. In the following theorem, using the Helly Choice Theorem (Theorem 2.7.4) and the Bounded Convergence Theorem (Theorem **??**), we will show that the operator $\mathscr{A}_{t_0}$ generates compact mapping of the space $\mathrm{BV}([a,b],\mathbb{R}^n)$ into itself.

7.4.5 Theorem. *Let* $t_0 \in [a,b]$, $A \in \mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$ *and* $L\,x = \mathscr{A}_{t_0}x$ *for* $x \in \mathrm{BV}([a,b],\mathbb{R}^n)$. *Then* L *is a compact linear operator on* $\mathrm{BV}([a,b],\mathbb{R}^n)$.

Proof. Since $\|x\| \leq \|x\|_{\mathrm{BV}}$ for each $x \in \mathrm{BV}([a,b],\mathbb{R}^n)$, it follows by Lemma 7.4.4 that $L \in \mathscr{L}(\mathrm{BV}([a,b],\mathbb{R}^n))$.

Next, we will prove that for an arbitrary sequence $\{x_n\}$ bounded in $\mathrm{BV}([a,b],\mathbb{R}^n)$, the set of its values $\{L\,x_n\} \subset \mathrm{BV}([a,b],\mathbb{R}^n)$ contains a subsequence that is convergent in $\mathrm{BV}([a,b],\mathbb{R}^n)$.

Thus, let the sequence $\{x_n\} \subset \mathrm{BV}([a,b])$ and the number $\varkappa \in [0,\infty)$ be such that

$$\|x_n\|_{\mathrm{BV}} \leq \varkappa < \infty \quad \text{for every } n \in \mathbb{N}.$$

By the Helly Choice Theorem (Theorem 2.7.4) there are a function $x \in \mathrm{BV}([a,b],\mathbb{R}^n)$ and an increasing subsequence $\{n_k\} \subset \mathbb{N}$ such that

$$\|x\|_{\mathrm{BV}} \leq 2\,\varkappa \text{ and } \lim_{k\to\infty} x_{n_k}(t) = x(t) \quad \text{for every } t \in [a,b].$$

Set $z_k(t) = x_{n_k}(t) - x(t)$ for $k \in \mathbb{N}$ and $t \in [a,b]$. Then

$$|z_k(t)| \leq 4\,\varkappa \text{ and } \lim_{k\to\infty} z_k(t) = 0 \quad \text{for } k \in \mathbb{N} \text{ and } t \in [a,b].$$

Since the integrals $\int_c^d \mathrm{d}A\, z_k$ and $\int_c^d \mathrm{d}[\mathrm{var}_a^s A]\, |z_k(s)|$ exist for all $c, d \in [a,b]$ and $k \in \mathbb{N}$, Theorem 6.3.4 guarantees that the inequalities

$$\sum_{j=1}^{\nu(\boldsymbol{\alpha})} |(L\, z_k)(\alpha_j) - (L\, z_k)(\alpha_{j-1})| = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \Big| \int_{\alpha_{j-1}}^{\alpha_j} \mathrm{d}A\, z_k \Big|$$
$$\leq \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \int_{\alpha_{j-1}}^{\alpha_j} \mathrm{d}\big[\mathrm{var}_a^s A\big]\, |z_k(s)| \leq \int_a^b \mathrm{d}\big[\mathrm{var}_a^s A\big]\, |z_k(s)|$$

hold for every division $\boldsymbol{\alpha}$ of $[a,b]$ and every $k \in \mathbb{N}$. We thus have

$$\mathrm{var}_a^b (L\, z_k) \leq \int_a^b \mathrm{d}\big[\mathrm{var}_a^s A\big]\, |z_k(s)| \quad \text{for } k \in \mathbb{N}.$$

By the Bounded Convergence Theorem **??**,

$$\lim_{k\to\infty} \int_a^b \mathrm{d}\big[\mathrm{var}_a^s A\big]\, |z_k(s)| = 0,$$

and hence

$$\lim_{k\to\infty} \mathrm{var}_a^b \big(L\, x_{n_k} - L\, x\big) = \lim_{k\to\infty} \mathrm{var}_a^b (L\, z_k) = 0.$$

Similarly,

$$\lim_{k\to\infty} \big|\big(L\, x_{n_k}(a) - L\, x(a)\big)\big| = \lim_{k\to\infty} |(L\, z_k)(a)| = \lim_{k\to\infty} \Big| \int_{t_0}^a \mathrm{d}A\, z_k \Big|$$
$$\leq \lim_{k\to\infty} \int_{t_0}^a \big[\mathrm{var}_a^s A\big]\, |z_k(s)| = 0.$$

This completes the proof of the theorem. □

The following statement is a corollary of Theorems 7.3.1 and 7.4.5.

7.4.6 Theorem. *Let* $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$ *and* $t_0 \in [a,b]$. *Then the initial value problem*

$$x(t) - \int_{t_0}^t \mathrm{d}A\, x = g(t) \tag{7.4.4}$$

has exactly one solution on $[a,b]$ *for every* $g \in \mathrm{BV}([a,b], \mathbb{R}^n)$ *if and only if the corresponding homogeneous problem*

$$x(t) - \int_{t_0}^t \mathrm{d}A\, x = 0 \tag{7.4.5}$$

has only trivial solution $x \equiv 0$ *on* $[a,b]$.

Proof. Equation (7.4.4) is equivalent with the operator equation

$$x - L\,x = g$$

where $L\,x = \mathscr{A}_{t_0}\,x$ for $x \in \mathrm{BV}([a,b],\mathbb{R}^n)$, i.e.

$$(L\,x)(t) = \int_{t_0}^{t} \mathrm{d}A\,x \quad \text{for } x \in \mathrm{BV}([a,b],\mathbb{R}^n) \text{ and } t \in [a,b].$$

By Theorem 7.4.5 L is linear compact operator on $\mathrm{BV}([a,b],\mathbb{R}^n)$. Hence, the proof can be completed by using Theorem 7.3.1. □

Now, assume that $\tau \in (t_0, b\,]$ and that the function $x \in \mathrm{G}([a,b],\mathbb{R}^n)$ satisfies equation (7.4.2) on $[t_0,\tau)$. Clearly, $x(t_0) = \widetilde{x}$ and, using Hake's theorem (Theorem 6.5.5, see also Examples 6.5.7 and Exercise 6.5.8), we easily verify the following relations

$$\begin{aligned} x(\tau-) &= \widetilde{x} + \lim_{s\to\tau-} \int_{t_0}^{s} \mathrm{d}A\,x + \big(f(\tau-) - f(t_0)\big) \\ &= \widetilde{x} + \int_{t_0}^{\tau} \mathrm{d}A\,x + f(\tau) - f(t_0) - \lim_{s\to\tau-} \int_{s}^{\tau} \mathrm{d}A\,x - \Delta^- f(\tau) \\ &= \widetilde{x} + \int_{t_0}^{\tau} \mathrm{d}A\,x + f(\tau) - f(t_0) - \Delta^- A(\tau)\,x(\tau) - \Delta^- f(\tau). \end{aligned}$$

Thus, if the function x should satisfy (7.4.2) also in τ, the value $x(\tau)$ has to be such that the equality

$$\big[I - \Delta^- A(\tau)\big]\,x(\tau) = x(\tau-) + \Delta^- f(\tau) \tag{7.4.6}$$

is true, where I stands for the $n \times n$-unit matrix (see Conventions and Notations (xiv)). From this, it is obvious that the solution to the initial problem (7.4.2) on $[t_0,\tau)$ can be extended to the point τ in an unique way if and only if

$$\det\big[I - \Delta^- A(\tau)\big] \neq 0. \tag{7.4.7}$$

Similarly, we can conclude that a function $x \in \mathrm{G}([a,b],\mathbb{R}^n)$ satisfying (7.4.2) on $(\tau, t_0]$, where $\tau \in [a, t_0)$ can be extended to the point τ if and only if

$$\big[I + \Delta^+ A(\tau)\big]\,x(\tau) = x(\tau+) - \Delta^+ f(\tau), \tag{7.4.8}$$

which will be true just if

$$\det\big[I + \Delta^+ A(\tau)\big] \neq 0. \tag{7.4.9}$$

Thus, we can expect that the conditions (7.4.7) and (7.4.9) should be essential for the existence of the solution to the problem (7.4.2).

7.4.7 Lemma. *Let* $A \in \mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$ *and* $t_0 \in [a,b]$. *Then problem* (7.4.4) *has exactly one solution for every function* $g \in \mathrm{BV}([a,b],\mathbb{R}^n)$ *if and only if*

$$\det\bigl(I - \Delta^- A(t)\bigr) \neq 0 \quad \textit{for all } t \in (t_0, b] \tag{7.4.10}$$

and

$$\det\bigl(I + \Delta^+ A(s)\bigr) \neq 0 \quad \textit{for all } s \in [a, t_0). \tag{7.4.11}$$

(*Here* $(t_0,b] = \emptyset$ *if* $t_0 = b$, *and* $[a,t_0) = \emptyset$ *if* $t_0 = a$.)

Proof. a) Assume that $t_0 \in [a,b)$, A satisfies (7.4.10) and (7.4.11) and x satisfies (7.4.5) on $[a,b]$. By Remark 7.1.3, x is a solution of (7.1.1) on $[a,b]$ while $x(t_0) = 0$. By Theorem 7.4.1, x is regulated on $[a,b]$ and the second equation in (7.4.1) yields

$$\Delta^+ x(t_0) = \Delta^+ A(t_0)\, x(t_0) = 0,$$

i.e. $x(t_0+) = 0$.

Set $\alpha(t) = \mathrm{var}_{t_0}^{\,t} A$ for $t \in [t_0, b\,]$. Then the function α is nondecreasing on the interval $[t_0, b\,]$. Thus, there is a finite limit $\alpha(t_0+)$ and we can choose a $\delta \in (0, b - t_0)$ such that $0 \le \alpha(t_0+\delta) - \alpha(t_0+) < 1/2$. From this and using Theorems 6.3.4 and 6.5.5, we derive for $t \in [t_0, t_0+\delta]$ the inequalities

$$\begin{aligned}
|x(t)| &\le \int_{t_0}^{t} |x|\,\mathrm{d}\alpha = \Delta^+\alpha(t_0)\,|x(t_0)| + \lim_{\tau \to t_0+} \int_{\tau}^{t} |x|\,\mathrm{d}\alpha \\
&= \lim_{\tau \to t_0+} \int_{\tau}^{t} |x|\,\mathrm{d}\alpha \le \bigl[\alpha(t_0+\delta) - \alpha(t_0+)\bigr] \Bigl(\sup_{t \in [t_0, t_0+\delta]} |x(t)|\Bigr) \\
&\le \frac{1}{2} \Bigl(\sup_{t \in [t_0, t_0+\delta]} |x(t)|\Bigr).
\end{aligned}$$

Hence

$$\Bigl(\sup_{t \in [t_0, t_0+\delta]} |x(t)|\Bigr) \le \frac{1}{2} \Bigl(\sup_{t \in [t_0, t_0+\delta]} |x(t)|\Bigr),$$

which is possible if and only if $x = 0$ on $[t_0, t_0+\delta]$.

Now, set $t^* = \sup\{\tau \in (t_0, b\,] : x = 0 \text{ on } [t_0, \tau]\}$. Obviously, $x = 0$ on $[t_0, t^*)$ and hence $x(t^*-) = 0$. Moreover, by (7.4.1) we have $0 = \bigl[I - \Delta^- A(t^*)\bigr]\, x(t^*)$. However, thanks to assumption (7.4.10), that is possible if and only if $x(t^*) = 0$.

Finally, assume that $t^* < b$. Using the same arguments as those we used to prove that there is a $\delta \in (0, b - t_0]$ such that x is zero on $[t_0, t_0+\delta]$, we would now show that there is an $\eta \in (0, b - t^*)$ such that x vanishes on $[t^*, t^*+\eta]$. This being in contradiction with the definition of t^*, it must be $t^* = b$. Thus, we proved that every solution of the problem (7.4.5) vanishes on $[t_0, b\,]$.

Similarly, using the assumption (7.4.11), we would prove that if $t_0 \in (a,b]$, then the solution x of the problem (7.4.5) vanishes also on $[a,t_0]$.

To summarize, we have proved that whenever (7.4.10) and (7.4.11) are satisfied, problem (7.4.5) will have only trivial solution on $[a,b]$. Consequently, by Theorem 7.4.6, problem (7.4.4) has exactly one solution on $[a,b]$ for each $g \in \mathrm{BV}([a,b],\mathbb{R}^n)$.

b) Assume that e.g. (7.4.10) does not hold. By Lemma 7.3.2

$$\det\left[I - \Delta^- A(t)\right] \neq 0 \ \text{ if } |\Delta^- A(t)| \leq 1/2.$$

On the other hand, by Corollary 4.1.7, the reverse inequality $|\Delta^- A(t)| > 1/2$ holds for at most finite number of points $t \in (t_0, b]$. Thus, the matrix $I - \Delta^- A(t)$ is not regular for at most finite number of points $t \in (t_0, b]$ and therefore we can choose a $t^* \in (t_0, b]$ such that

$$\det\left[I - \Delta^- A(t)\right] \neq 0 \text{ for } t \in (t_0, t^*) \ \text{ and } \ \det\left[I - \Delta^- A(t^*)\right] = 0.$$

Now, it is well known from linear algebra that in such a case there exists $d \in \mathbb{R}^n$ such that

$$\left[I - \Delta^- A(t^*)\right] c \neq d \quad \text{for every } \ c \in \mathbb{R}^n. \tag{7.4.12}$$

Define

$$g(t) = \begin{cases} 0 & \text{when } \ t \neq t^*, \\ d & \text{when } \ t = t^*. \end{cases}$$

Then $g \in \mathrm{BV}([a,b],\mathbb{R}^n)$ and $\Delta^- g(t^*) = d$. Assume that equation (7.4.4) has a solution x on $[a,b]$. Then by the first part of the proof $x = 0$ on $[a,t^*)$, and thus also $x(t^*-) = 0$. By Theorem 7.4.1 $\left[I - \Delta^- A(t^*)\right] x(t^*) = d$ has to hold. This is however in contradiction with the statement (7.4.12) and hence problem (7.4.4) cannot have a solution.

If (7.4.11) does not hold, then we will analogously find a point $t^* \in [a, t_0)$ and a function g such that

$$\left[I + \Delta^+ A(t^*)\right] c \neq \Delta^+ g(t^*) \quad \text{for every } \ c \in \mathbb{R}^n,$$

which again leads to the contradiction with Theorem 7.4.1. □

7.4.8 Theorem. *Let $A \in \mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$ and $t_0 \in [a,b]$. Then the initial problem (7.4.2) has exactly one solution for every function $f \in \mathrm{BV}([a,b],\mathbb{R}^n)$ and every vector $\widetilde{x} \in \mathbb{R}^n$ if and only if (7.4.10) and (7.4.11) hold.*

Proof. The theorem is a corollary of Lemma 7.4.7 if we set

$$g(t) = \widetilde{x} + f(t) - f(t_0) \quad \text{for } \ t \in [a,b].$$

□

7.5 A priori estimates of solutions

An important role in the theory of ordinary differential equations (e.g. for the proof the uniqueness of a solution to the initial problem or for the proof of the continuous dependence of solutions on some parameters) is played by the assertion called Gronwall lemma. Below we recall its statement. Its proof can be found in the majority of textbooks on ordinary differential equations, see. e.g. Auxiliary Theorem 4.3.1 in [82].

7.5.1 Lemma (GRONWALL). *Let the functions u and p be continuous and nonnegative on $[a,b]$, let $K \ge 0$ and let*

$$u(t) \le K + \int_a^t \big(p(s)\,u(s)\big)\,\mathrm{d}s \quad \textit{for } t \in [a,b].$$

Then

$$u(t) \le K\,\exp\Big(\int_a^t p(s)\,\mathrm{d}s\Big) \quad \textit{for } t \in [a,b].$$

For our purposes, the generalization of Gronwall lemma to the Stieltjes setting will be likewise important. To deduce it, we need the following auxiliary result.

7.5.2 Lemma. *If $h : [a,b] \to [0,\infty)$ is nondecreasing and left-continuous, then*

$$\int_a^b h^k\,\mathrm{d}h \le \frac{h^{k+1}(b) - h^{k+1}(a)}{k+1} \quad \textit{for every } k \in \mathbb{N} \cup \{0\}.$$

Proof. The existence of the integral $\int_a^b h^k\,\mathrm{d}h$ follows from Theorem 6.3.8. Consider an arbitrary $\varepsilon > 0$. There exists a gauge $\delta : [a,b] \to (0,\infty)$ such that for each δ-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$, we have

$$\Big|\int_a^b h^k\,\mathrm{d}h - S(P)\Big| < \varepsilon.$$

Moreover, using the left-continuity of h, we can assume that δ is chosen in such a way that

$$h^i(\tau) - h^i(t) < \varepsilon \quad \text{for } i \in \{0,\dots,k\},\ \tau \in (a,b],\ t \in (\tau - \delta(\tau), \tau]. \tag{7.5.1}$$

Now, let $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ be an arbitrary δ-fine partition of $[a,b]$. Then

$$\left.\begin{aligned} \int_a^b h^k\,\mathrm{d}h &= \int_a^b h^k\,\mathrm{d}h - S(P) + S(P) < \varepsilon + S(P) \\ &= \varepsilon + \sum_{j=1}^{\nu(P)} \big(h^k(\xi_j)\,(h(\alpha_j) - h(\xi_j)) + h^k(\xi_j)\,(h(\xi_j) - h(\alpha_{j-1}))\big). \end{aligned}\right\} \tag{7.5.2}$$

Note that for each $j \in \{1, \dots, \nu(P)\}$, we have

$$h^k(\xi_j) \le \frac{1}{k+1} \sum_{i=0}^{k} h^{k-i}(\alpha_j) h^i(\xi_j)$$

(since h is nondecreasing, the right-hand side corresponds to the average of $k+1$ terms, which are all greater than or equal to $h^k(\xi_j)$). Consequently,

$$\begin{aligned} h^k(\xi_j)\,(h(\alpha_j) - h(\xi_j)) &\le \frac{1}{k+1} \sum_{i=0}^{k} h^{k-i}(\alpha_j)\, h^i(\xi_j)\,(h(\alpha_j) - h(\xi_j)) \\ &= \frac{1}{k+1}(h^{k+1}(\alpha_j) - h^{k+1}(\xi_j)). \end{aligned}$$

For each $j \in \{1, \dots, \nu(P)\}$, the inequality (7.5.1) implies that

$$h^k(\xi_j) \le \frac{1}{k+1} \sum_{i=0}^{k} h^{k-i}(\xi_j) \left(h^i(\alpha_{j-1}) + \varepsilon \right).$$

Furthermore, since h is nondecreasing, we have $h^{k-i}(\xi_j) \le M$, where

$$M = \max\{h^i(b),\, i = 0, \dots, k\}.$$

Consequently,

$$\begin{aligned} &h^k(\xi_j)\,(h(\xi_j) - h(\alpha_{j-1})) \\ &\quad \le \frac{1}{k+1} \sum_{i=0}^{k} h^{k-i}(\xi_j) \left(h^i(\alpha_{j-1}) + \varepsilon \right) \, (h(\xi_j) - h(\alpha_{j-1})) \\ &\quad = \frac{1}{k+1} \sum_{i=0}^{k} h^{k-i}(\xi_j)\, h^i(\alpha_{j-1})\,(h(\xi_j) - h(\alpha_{j-1})) \\ &\qquad + \frac{\varepsilon}{k+1} \left(\sum_{i=0}^{k} h^{k-i}(\xi_j) \right) (h(\xi_j) - h(\alpha_{j-1})) \\ &\quad \le \frac{1}{k+1}(h^{k+1}(\xi_j) - h^{k+1}(\alpha_{j-1})) + \varepsilon\, M\, (h(\xi_j) - h(\alpha_{j-1})). \end{aligned}$$

By substituting the previous inequalities into (7.5.2), we get

$$\begin{aligned} \int_a^b h^k \,\mathrm{d}h &\le \varepsilon + \frac{1}{k+1} \sum_{j=1}^{\nu(P)} \left(h^{k+1}(\alpha_j) - h^{k+1}(\alpha_{j-1}) + \varepsilon\, M\, (h(\xi_j) - h(\alpha_{j-1})) \right) \\ &\le \varepsilon + \frac{h^{k+1}(b) - h^{k+1}(a)}{k+1} + \varepsilon\, M\, (h(b) - h(a)), \end{aligned}$$

which completes the proof. □

7.5.3 Exercise. Prove the following complementary statement to Lemma 7.5.2:

If $h:[a,b]\to[0,\infty)$ *is nonincreasing and right-continuous, then*

$$\int_a^b h^k\,\mathrm{d}h \ge \frac{h^{k+1}(b)-h^{k+1}(a)}{k+1} \quad \textit{for every } k\in\mathbb{N}\cup\{0\}.$$

7.5.4 Theorem (GENERALIZED GRONWALL LEMMA)**.** *Assume that* $u:[a,b]\to[0,\infty)$ *is bounded,* $h:[a,b]\to[0,\infty)$ *is nondecreasing and left-continuous,* $K\ge 0,\ L\ge 0,$ *and*

$$u(t)\le K+L\int_a^t u\,\mathrm{d}h \quad \textit{for}\ \ t\in[a,b]. \tag{7.5.3}$$

Then

$$u(t)\le K\,\exp\big(L\,[h(t)-h(a)]\big) \quad \textit{for}\ \ t\in[a,b]. \tag{7.5.4}$$

Proof. Let $\kappa\ge 0$ and $w_\kappa(t)=\kappa\,\exp\big(L\,[h(t)-h(a)]\big)$ for $t\in[a,b]$. Then

$$\begin{aligned}\int_a^t w_\kappa\,\mathrm{d}h &= \kappa\int_a^t \exp\big(L\,[h(s)-h(a)]\big)\,\mathrm{d}h(s)\\ &= \kappa\int_a^t\Big(\sum_{k=0}^\infty \frac{L^k}{k!}\,\big[h(s)-h(a)\big]^k\Big)\,\mathrm{d}h(s) \quad \text{for}\ \ t\in[a,b].\end{aligned}$$

Since, as is known, the series $\sum_{k=0}^\infty \frac{L^k}{k!}\,[h(t)-h(a)]^k$ converges uniformly on $[a,b]$, we can change the order of the operations of integrating and adding. If we now use Theorem 7.5.2, where we replace the function h by the difference $h-h(a)$, we get

$$\begin{aligned}\int_a^t w_\kappa\,\mathrm{d}h &= \kappa\sum_{k=0}^\infty\Big(\frac{L^k}{k!}\int_a^t\big[h(s)-h(a)\big]^k\Big)\,\mathrm{d}[h(s)]\\ &\le \kappa\sum_{k=0}^\infty\Big(\frac{L^k\,[h(t)-h(a)]^{k+1}}{(k+1)!}\Big)=\frac{\kappa}{L}\Big(\exp(L\,[h(t)-h(a)])-1\Big)\\ &= \frac{w_\kappa(t)-\kappa}{L}\end{aligned}$$

for $t\in[a,b]$. This means that the function w_κ satisfies the inequality

$$w_\kappa(t)\ge\kappa+L\int_a^t w_\kappa\,\mathrm{d}h \tag{7.5.5}$$

for every $\kappa\ge 0$ and $t\in[a,b]$. Let $\varepsilon>0$ be given and let $\kappa=K+\varepsilon$ and $v_\varepsilon=u-w_\kappa$. By subtracting the inequalities (7.5.3) and (7.5.5), we find out

$$v_\varepsilon(t)\le-\varepsilon+L\int_a^t v_\varepsilon\,\mathrm{d}h \quad \text{holds for}\ \ t\in[a,b]. \tag{7.5.6}$$

Specially, $v_\varepsilon(a) \le -\varepsilon < 0$. The remaining part of the proof will resemble the method used in the proof of Lemma 7.4.7. The functions u and w_κ are evidently bounded on $[a,b]$ for every $\kappa \ge 0$. Hence, the function v_ε is bounded on $[a,b]$, too. By Hake's theorem 6.5.5 (ii) we have for $t \in (a,b]$

$$\int_a^t v_\varepsilon \,\mathrm{d}h = v_\varepsilon(a)\,\Delta^+ h(a) + \lim_{\delta\to 0+}\int_{a+\delta}^t v_\varepsilon \,\mathrm{d}h$$
$$\le -\varepsilon\,\Delta^+ h(a) + \|v_\varepsilon\|\,[h(t) - h(a+)] \le \|v_\varepsilon\|\,[h(t) - h(a+)],$$

and therefore

$$v_\varepsilon(t) \le -\varepsilon + L\int_a^t v_\varepsilon \,\mathrm{d}h \le -\varepsilon + L\,\|v_\varepsilon\|\,[h(t) - h(a+)] \quad \text{for } t \in (a,b].$$

Choose an $\eta > 0$ in such a way that

$$L\,\|v_\varepsilon\|\,[h(t) - h(a+)] < \varepsilon/2 \quad \text{holds for } t \in [a, a+\eta].$$

Then $v_\varepsilon < 0$ on $[a, a+\eta]$. Set $t^* = \sup\{\tau \in [a,b] : v_\varepsilon < 0 \text{ on } [a,\tau]\}$.

We see that $t^* > a$ and $v_\varepsilon < 0$ on $[a, t^*)$. By repeated using Hake's theorem 6.5.5 (i), we get

$$v_\varepsilon(t^*) \le -\varepsilon + L\int_a^{t^*} v_\varepsilon \,\mathrm{d}h$$
$$= -\varepsilon + L\Big(v_\varepsilon(t^*)\Delta^- h(t^*) + \lim_{\delta\to 0+}\int_a^{t^*-\delta} v_\varepsilon \,\mathrm{d}h\Big) \le -\varepsilon < 0,$$

from (7.5.6) as $\Delta^- h(t^*) = 0$ and $\int_a^{t^*-\delta} v_\varepsilon \,\mathrm{d}h \le 0$ for every $\delta > 0$.

If $t^* < b$, we would repeat the previous method and show that there exists $\theta \in (0, b - t^*)$ such that $v_\varepsilon < 0$ on the interval $[a, t^* + \theta]$ which is in contradiction with Definition t^*. Hence $t^* = b$, $v_\varepsilon < 0$ on the whole $[a,b]$ and

$$u(t) < w_\kappa(t) = K\,\exp\big(L\,(h(t) - h(a))\big) + \varepsilon\,\exp\big(L\,(h(t) - h(a))\big) \text{ for } t \in [a,b].$$

Since $\varepsilon > 0$ was arbitrary, it means that (7.5.4) holds. □

7.5.5 Exercise. Prove the following variant of Theorem 7.5.4:

Let $u : [a,b] \to [0,\infty)$ be bounded on $[a,b]$, $h : [a,b] \to [0,\infty)$ be nondecreasing and continuous from right on $(a,b]$, $K \ge 0$, $L \ge 0$ and

$$u(t) \le K + L\int_t^b u \,\mathrm{d}h \quad \textit{for } t \in [a,b]. \tag{7.5.7}$$

Then

$$u(t) \le K\,\exp\big(L\,[h(b) - h(t)]\big) \quad \text{for } t \in [a,b]. \tag{7.5.8}$$

7.5.6 Remark. More general versions of generalized Gronwall lemma are included in the monographs by Š. Schwabik [122] (see Theorem 1.40) and J. Kurzweil [85] (see chapter 22).

In the following theorem, we will use generalized Gronwall lemma for deriving an important estimate for the solution of the problem (7.4.2).

7.5.7 Theorem. *Let* $t_0 \in [a,b]$ *and* $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$ *meet the conditions* (7.4.10) *and* (7.4.11), $f \in \mathrm{G}([a,b], \mathbb{R}^n)$, $\widetilde{x} \in \mathbb{R}^n$ *and let* x *be a solution of the initial problem* (7.4.2) *on* $[a,b]$. *Then*

$$\operatorname{var}_a^b (x - f) \le (\operatorname{var}_a^b A)\, \|x\| < \infty, \tag{7.5.9}$$

$$\left.\begin{aligned} c_{(A,t_0)} := \max\Big\{1, \sup_{t\in(t_0,b]} \big|[I - \Delta^- A(t)]^{-1}\big|,\\ \sup_{t\in[a,t_0)} \big|[I + \Delta^+ A(t)]^{-1}\big| \Big\} < \infty, \end{aligned}\right\} \tag{7.5.10}$$

$$\left.\begin{aligned} |x(t)| \le c_{(A,t_0)} \big(|\widetilde{x}| + 2\,\|f\|\big) \exp\big(2\,c_{(A,t_0)} \operatorname{var}_{t_0}^t A\big) \text{ for } t \in [t_0, b],\\ |x(t)| \le c_{(A,t_0)} \big(|\widetilde{x}| + 2\,\|f\|\big) \exp\big(2\,c_{(A,t_0)} \operatorname{var}_t^{t_0} A\big) \text{ for } t \in [a, t_0]. \end{aligned}\right\} \tag{7.5.11}$$

Proof. a) For any division $\boldsymbol{\alpha}$ of the interval $[a,b]$, we have

$$\sum_{j=1}^{\nu(\boldsymbol{\alpha})} \Big| x(\alpha_j) - f(\alpha_j) - x(\alpha_{j-1}) + f(\alpha_{j-1}) \Big|$$
$$= \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \Big| \int_{\alpha_{j-1}}^{\alpha_j} \mathrm{d}[A]\, x \Big| \le \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \big[(\operatorname{var}_{\alpha_{j-1}}^{\alpha_j} A)\, \|x\| \big] = (\operatorname{var}_a^b A)\, \|x\| < \infty.$$

This gives immediately (7.5.9).

b) Let $t \in (t_0, b]$ be such that $|\Delta^- A(t)| \le \frac{1}{2}$. Then, using Lemma 7.3.2, we get

$$\big|[I - \Delta^- A(t)]^{-1}\big| \le \frac{1}{1 - |\Delta^- A(t)|} < 2.$$

Since the set $\big\{t \in [a,b] \colon |\Delta^- A(t)| > \frac{1}{2}\big\}$ has at most finitely many elements, Corollary 4.1.7 implies that

$$\sup_{t\in(t_0,b]} \big|[I - \Delta^- A(t)]^{-1}\big| < \infty.$$

Similarly, we would argue to prove that

$$\sup_{t\in[a,t_0)} \big|[I + \Delta^+ A(t)]^{-1}\big| < \infty.$$

To summarize, (7.5.10) is true.

c) Let x satisfy (7.4.2). Set

$$B(t) = \begin{cases} A(t) & \text{if } t \in [a, t_0], \\ A(t-) & \text{if } t \in (t_0, b]. \end{cases}$$

Obviously, $A(t) - B(t) = \Delta^- A(t)$ and

$$\operatorname{var}_{t_0}^t (B - A) = \sum_{s \in (t_0, t]} |\Delta^- A(s)| \le \operatorname{var}_{t_0}^t A$$

for $t \in (t_0, b]$ (see Corollary 2.3.8). Hence

$$A - B \in \mathrm{BV}([a, b], \mathbb{R}^n) \quad \text{and} \quad \operatorname{var}_{t_0}^t B \le 2 \operatorname{var}_{t_0}^t A.$$

Furthermore, $\Delta^+ B(t_0) = \Delta^+ A(t_0)$ and hence, using Corollary 6.3.16, we get

$$\int_{t_0}^t \mathrm{d}[A - B]\, x = \Delta^- A(t)\, x(t) \quad \text{for } t \in (t_0, b].$$

The equation (7.4.2) is thus reduced to

$$[I - \Delta^- A(t)]\, x(t) = \widetilde{x} + \int_{t_0}^t \mathrm{d}[B]\, x + f(t) - f(t_0) \quad \text{for } t \in [t_0, b].$$

From here and having in mind that $c_{(A,t_0)} \ge 1$ we easily derive the inequality

$$|x(t)| \le K + L \int_{t_0}^t |x|\, \mathrm{d}h \quad \text{for } t \in [t_0, b],$$

where

$$K = c_{(A,t_0)} \left(|\widetilde{x}| + 2\, \|f\| \right), \quad L = c_{(A,t_0)} \text{ and } h(t) = \operatorname{var}_{t_0}^t B.$$

The function h is nondecreasing on $[t_0, b]$. Moreover, since B is continuous from the left on $(t_0, b]$, the function h is by Lemma 2.3.3 also continuous from the left on $(t_0, b]$. Now, applying the Generalized Gronwall Lemma 7.5.4 we get finally the former inequality in (7.5.11).

For the proof of the latter inequality in (7.5.11) we can argue in a similar way, while using the variation of the generalized Gronwall inequality from Exercise 7.5.5. □

7.5.8 Exercise. Under the assumptions of Theorem 7.5.7, prove the inequalities

$$0 < \sup_{t \in [a, t_0)} \left| [I + \Delta^+ A(t)]^{-1} \right| < \infty$$

and

$$|x(t)| \le c_{(A,t_0)} \left(|\widetilde{x}| + 2\, \|f\| \right) \exp\left(2\, c_{(A,t_0)} \operatorname{var}_t^{t_0} A \right) \quad \text{for } t \in [a, t_0]$$

in details.

7.6 Continuous dependence of solutions on parameters

Let $t_0 \in [a,b]$ and $A \in \mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$ meet the conditions (7.4.10) and (7.4.11), $\widetilde{x} \in \mathbb{R}^n$, $f \in \mathrm{G}([a,b],\mathbb{R}^n)$ and let x be a solution of problem (7.4.2) on $[a,b]$. Also, let y be a solution of

$$y(t) - \widetilde{y} - \int_{t_0}^{t} \mathrm{d}A\, y = g(t) - g(t_0) \tag{7.6.1}$$

on $[a,b]$, where $g \in \mathrm{G}([a,b],\mathbb{R}^n)$ and $\widetilde{y} \in \mathbb{R}^n$. Then

$$\big(x(t) - y(t)\big) = \big(\widetilde{x} - \widetilde{y}\big) + \int_{t_0}^{t} \mathrm{d}A\,\big(x - y\big) + \big(f(t) - g(t)\big) - \big(f(t_0) - g(t_0)\big)$$

for $t \in [a,b]$. Thus, by Theorem 7.5.7 we have

$$\|x - y\| \le c_{(A,t_0)}\,\big(\big|\widetilde{x} - \widetilde{y}\big| + 2\,\big\|f - g\big\|\big)\ \exp\big(2\,c_{(A,t_0)}\,\mathrm{var}_a^b A\big),$$

where $c_{(A,t_0)} \in (0,\infty)$ is defined in (7.5.10). We see that the „distance“ between the solutions of initial problems (7.4.2) and (7.6.1) is proportional to the „distance“ between the input data (i.e. the initial values of $\widetilde{x}$, $\widetilde{y}$ and the right hand side f, g) of these equations. This phenomenon is described in more details by the following theorem.

7.6.1 Theorem. *Let $t_0 \in [a,b]$ and $A \in \mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$ satisfy (7.4.10) and (7.4.11). Further, let f, $f_k \in \mathrm{G}([a,b],\mathbb{R}^n)$ and $\widetilde{x}$, $\widetilde{x}_k \in \mathbb{R}^n$ for $k \in \mathbb{N}$ are such that*

$$\lim_{k\to\infty} \|f_k - f\| = 0 \tag{7.6.2}$$

and

$$\lim_{k\to\infty} \widetilde{x}_k = \widetilde{x}. \tag{7.6.3}$$

Finally, let for each $k \in \mathbb{N}$ the initial problem

$$x_k(t) - \widetilde{x}_k - \int_{t_0}^{t} \mathrm{d}A\, x_k = f_k(t) - f_k(t_0) \tag{7.6.4}$$

have a solution x_k on $[a,b]$. Then problem (7.4.2) has a solution x on $[a,b]$ and

$$\lim_{k\to\infty} \big\|x_k - x\big\| = 0. \tag{7.6.5}$$

Proof. a) As a consequence of (7.6.2) and (7.6.3), there exists an k_0 such that

$$\|f_k\| \le \|f\| + 1 \text{ and } |\widetilde{x}_k| \le |\widetilde{x}| + 1 \quad \text{for } k \ge k_0.$$

Thus, by Theorem 7.5.7, we have

$$\|x_k\| \le \varkappa_0 < \infty \quad \text{for } k \ge k_0, \tag{7.6.6}$$

where

$$\varkappa_0 = c_{(A,t_0)} \big(|\widetilde{x}| + 2\,\|f\| + 3\big) \exp \big(2\, c_{(A,t_0)} \operatorname{var}_a^b A\big)$$

does not depend on k. Furthermore, by the same theorem, we also have

$$\operatorname{var}_a^b(x_k - f_k) \le \varkappa_0 \operatorname{var}_a^b A < \infty \quad \text{for } k \ge k_0.$$

Now, by Helly's Selection Theorem (Theorem 2.7.4) there are $y \in \mathrm{BV}([a,b],\mathbb{R}^n)$ and an increasing sequence $\{k_\ell\} \subset \mathbb{N}$ such that $k_1 \ge k_0$,

$$\|y\|_{\mathrm{BV}} \le 2 \max\{\varkappa_0 \operatorname{var}_a^b A, \varkappa_0 + \|f\| + 1\}$$

and

$$\lim_{\ell\to\infty} \big(x_{k_\ell}(t) - f_{k_\ell}(t)\big) = y(t) \quad \text{for } t \in [a,b].$$

Having in mind (7.6.2), we see that the limit

$$x(t) := \lim_{\ell\to\infty} x_{k_\ell}(t)$$

exists for every $t \in [a,b]$. By (7.6.6) the sequence $\{x_{k_\ell}\}$ is uniformly bounded. Hence, using the Bounded Convergence Theorem (Theorem **??**), we get

$$\lim_{\ell\to\infty} \int_{t_0}^t \mathrm{d}A\, x_{k_\ell} = \int_{t_0}^t \mathrm{d}A\, x \quad \text{for each } t \in [a,b].$$

Moreover, letting $\ell \to \infty$ in (7.6.4) (and therefore using assumptions (7.6.2) and (7.6.3)), we find out that x is a solution to problem (7.4.2) on $[a,b]$.

b) If for every $k \in \mathbb{N}$ we repeat the arguments from the introduction to this section with y replaced by x_k, g replaced by f_k and $\widetilde{y}$ replaced by $\widetilde{x} - k$, we will find out that

$$\|x - x_k\| \le K \big(\big|\widetilde{x} - \widetilde{x}_k\big| + 2\,\big\|f - f_k\|\big)$$

holds for every $k \in \mathbb{N}$, where

$$K = c_{(A,t_0)} \exp \big(2\, c_{(A,t_0)} \operatorname{var}_a^b A\big) < \infty$$

does not depend on k. Therefore, (7.6.5) holds, too. □

Now we are ready to extend Theorem 7.4.8 to the general case of $f \in \mathrm{G}([a,b],\mathbb{R}^n)$. The existence results, which we have up to now at our disposal, applies only to the cases when the right hand side f has a bounded variation on $[a,b]$.

7.6.2 Theorem. *Let* $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$, $t_0 \in [a,b]$ *and let* (7.4.10) *and* (7.4.11) *hold.*

Then the initial problem (7.4.2) *has exactly one solution on* $[a,b]$ *for every function* $f \in \mathrm{G}([a,b], \mathbb{R}^n)$ *and every vector* $\widetilde{x} \in \mathbb{R}^n$.

Proof. a) If we have two solutions x, y of the problem (7.4.2) on the interval $[a,b]$, their difference on $[a,b]$ will be a solution of homogeneous problem (7.4.5) which however has only trivial solution by Lemma 7.4.7. Therefore, $x \equiv y$ has to hold on $[a,b]$.

b) Set $\widetilde{x}_k = \widetilde{x}$ for $k \in \mathbb{N}$. By Theorem 4.1.5 there exists a sequence $\{f_k\}$ of step functions (therefore of functions from $\mathrm{BV}([a,b], \mathbb{R}^n)$) which converges uniformly on $[a,b]$ to f. By Theorem 7.4.8 there exists exactly one solution x_k of the problem (7.6.4) for every $k \in \mathbb{N}$ and by Theorem 7.6.1 the sequence $\{x_k\}$ converges uniformly to the solution of problem (7.4.2). □

In the remaining part of the section, we will investigate the initial problem

$$x(t) - \widetilde{x} - \int_{t_0}^{t} \mathrm{d}A\, x = f(t) - f(t_0) \tag{7.6.7}$$

as the limit of the problems

$$x_k(t) - \widetilde{x}_k - \int_{t_0}^{t} \mathrm{d}A_k\, x_k = f_k(t) - f_k(t_0), \tag{7.6.8}$$

where the kernels A_k depend on the parameter $k \in \mathbb{N}$. This case is slightly more complicated than the one we dealt with in Theorem 7.6.1. First, we will prove convergence theorem for KS-integrals for the situation which is not covered by the theorems from the Chapter 6.

7.6.3 Theorem. *Let* $f,\ f_k \in \mathrm{G}([a,b], \mathbb{R}^n)$, $A, A_k \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$ *for* $k \in \mathbb{N}$. *Assume that conditions* (7.6.2),

$$\lim_{k\to\infty} \|A_k - A\| = 0 \tag{7.6.9}$$

and

$$\alpha^* := \sup_{k \in \mathbb{N}} \operatorname{var}_a^b A_k < \infty \tag{7.6.10}$$

are satisfied. Then

$$\lim_{k\to\infty} \left(\sup_{t\in[a,b]} \left| \int_a^t \mathrm{d}A_k\, f_k - \int_a^t \mathrm{d}A\, f \right| \right) = 0.$$

Proof. Let $\varepsilon > 0$ be given. By Theorem 4.1.5 we can choose a function φ: $[a,b] \to \mathbb{R}^n$ such that its every component is a step function on $[a,b]$ and at the same time, $\|f - \varphi\| < \varepsilon$. Moreover, by (7.6.2) and (7.6.9) we can choose a $k_0 \in \mathbb{N}$ such that

$$\|f_k - f\| < \varepsilon \quad \text{and} \quad \|A_k - A\| < \varepsilon \quad \text{for } k \ge k_0.$$

For given $t \in [a,b]$ and $k \ge k_0$, using Theorems 6.3.4 and 6.3.5, we get

$$\begin{aligned}
&\Big|\int_a^t \mathrm{d}A_k\, f_k - \int_a^t \mathrm{d}A\, f\Big| \\
&\quad \le \Big|\int_a^t \mathrm{d}A_k\,(f_k - \varphi)\Big| + \Big|\int_a^t \mathrm{d}\big[A_k - A\big]\,\varphi\Big| + \Big|\int_a^t \mathrm{d}A\,(\varphi - f)\Big| \\
&\quad \le (\mathrm{var}_a^b A_k)\,\|f_k - \varphi\| + 2\,\|A_k - A\|\,\|\varphi\|_{\mathrm{BV}} + (\mathrm{var}_a^b A)\,\|\varphi - x\| \\
&\quad \le \alpha^*\,\big(\|f_k - f\| + \|f - \varphi\|\big) + 2\,\|A_k - A\|\,\|\varphi\|_{\mathrm{BV}} + (\mathrm{var}_a^b A)\,\|\varphi - f\| \\
&\quad \le \big(2\,\alpha^* + 2\,\|\varphi\|_{\mathrm{BV}} + \mathrm{var}_a^b A\big)\,\varepsilon = K\,\varepsilon,
\end{aligned}$$

where

$$K = \big(2\,\alpha^* + 2\,\|\varphi\|_{\mathrm{BV}} + \mathrm{var}_a^b A\big) \in (0,\infty)$$

does not depend neither on k nor on t. This completes the proof. □

The following auxiliary statement will be useful, too.

7.6.4 Lemma. *If $A,\ A_k \in \mathrm{G}([a,b],\mathbb{R}^n)$ for $k \in \mathbb{N}$ are such that $A_k \rightrightarrows A$ on $[a,b]$, then the following statements hold:*

1. *If $I - \Delta^- A(t)$ is invertible for each $t \in (t_0,b]$, then there exists a $k_0 \in \mathbb{N}$ such that $I - \Delta^- A_k(t)$ is invertible for all $k \ge k_0$, $t \in (t_0,b]$. Moreover,*

$$\sup_{t\in(t_0,b]} \big|(I - \Delta^- A_k(t))^{-1}\big| < 2 \sup_{t\in(t_0,b]} \big|(I - \Delta^- A(t))^{-1}\big| \tag{7.6.11}$$

   for each $k \ge k_0$.

2. *If $I + \Delta^+ A(s)$ is invertible for each $s \in [a,t_0)$, then there exists a $k_0 \in \mathbb{N}$ such that $I - \Delta^+ A_k(s)$ is invertible for all $k \ge k_0$, $t \in [a,t_0)$. Moreover,*

$$\sup_{t\in[a,t_0)} \big|(I + \Delta^+ A_k(s))^{-1}\big| < 2 \sup_{t\in[a,t_0)} \big|(I + \Delta^+ A(s))^{-1}\big| \tag{7.6.12}$$

   for each $k \ge k_0$.

Proof. We prove the first statement; the proof of the second one is similar and is left to the reader.

According to the proof of Theorem 7.5.7, the quantity

$$c = \sup_{t\in(t_0,b]} \left|(I - \Delta^- A(t))^{-1}\right|$$

is finite. By Lemma 4.2.3, we have $\Delta^- A_k \rightrightarrows \Delta^- A$ on $[a,b]$. Thus, there exists a $k_0 \in \mathbb{N}$ such that

$$|\Delta^- A_k(t) - \Delta^- A(t)| < \frac{1}{4(c+1)} \tag{7.6.13}$$

for all $t\in[a,b]$, $k \geq k_0$. For $t \in (t_0, b]$, we can write

$$\begin{aligned} I - \Delta^- A_k(t) &= (I - \Delta^- A(t)) - (\Delta^- A_k(t) - \Delta^- A(t)) \\ &= \big(I - \Delta^- A(t)\big)\big(I - T_k(t)\big), \end{aligned}$$

where

$$T_k(t) = \big(I - \Delta^- A(t)\big)^{-1}\big(\Delta^- A_k(t) - \Delta^- A(t)\big).$$

To prove that $I - \Delta^- A_k(t)$ is invertible, it suffices to show that $I - T_k(t)$ is invertible. By (7.6.13) we have

$$|T_k(t)| \leq \left|\big(I - \Delta^- A(t)\big)^{-1}\right| \cdot \left|\Delta^- A_k(t) - \Delta^- A(t)\right| < \frac{1}{4}$$

for all $t \in (t_0, b]$, $k \geq k_0$. Thus, Lemma 7.3.2 guarantees that $I - T_k(t)$ and consequently also $I - \Delta^- A_k(t)$ are invertible; moreover, $\left|\big(I - T_k(t)\big)^{-1}\right| < 2$. Hence, it follows that

$$\left|\big(I - \Delta^- A_k(t)\big)^{-1}\right| \leq \left|\big(I - T_k(t)\big)^{-1}\right| \cdot \left|\big(I - \Delta^- A(t)\big)^{-1}\right| < 2\left|(I - \Delta^- A(t))^{-1}\right|,$$

which proves the estimate (7.6.11). □

7.6.5 Exercise. Prove the second part of Lemma 7.6.4.

Now we are ready to formulate and prove the main result of this section.

7.6.6 Theorem. *Let* $A, A_k \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$, f, $f_k \in \mathrm{G}([a,b], \mathbb{R}^n)$, $\widetilde{x}, \widetilde{x}_k \in \mathbb{R}^n$ *for all* $k \in \mathbb{N}$, *where* $A_k \rightrightarrows A$, $f_k \rightrightarrows f$, $\widetilde{x}_k \to \widetilde{x}$ *for* $k \to \infty$. *Furthermore, assume that* $\sup_{k\in\mathbb{N}} \operatorname{var}_a^b A_k < \infty$ *and conditions* (7.4.10), (7.4.11) *hold.*

Then there exists a $k_0 \in \mathbb{N}$ *such that for every* $k \geq k_0$, *the equation*

$$x_k(t) = \widetilde{x}_k + \int_a^t \mathrm{d}A_k\, x_k + f_k(t) - f_k(a), \quad t \in [a,b], \tag{7.6.14}$$

has a unique solution $x_k : [a,b] \to \mathbb{R}^n$. *Moreover,* $x_k \rightrightarrows x$, *where* $x : [a,b] \to \mathbb{R}^n$ *is the unique solution of the equation*

$$x(t) = \widetilde{x} + \int_a^t \mathrm{d}A\, x + f(t) - f(a), \quad t \in [a,b]. \tag{7.6.15}$$

Proof. By Lemma 7.6.4, there is a $k_0 \in \mathbb{N}$ such that if $k \ge k_0$, $I - \Delta^- A_k(t)$ is invertible for all $t \in (t_0, b]$, and $I + \Delta^+ A_k(t)$ is invertible for all $t \in [a, t_0)$. Hence, Theorem 7.6.2 implies that equation (7.6.14) has a unique solution $x_k : [a,b] \to \mathbb{R}^n$ for every $k \ge k_0$, and equation (7.6.15) has a unique solution $x : [a,b] \to \mathbb{R}^n$. Let

$$w_k = (x_k - f_k) - (x - f), \quad k \in \mathbb{N}.$$

Then $x_k - x = w_k + \big(f_k - f\big)$, and the theorem will be proved if we show that $w_k \rightrightarrows 0$. Observe that

$$w_k(t) = \widetilde{w}_k + \int_a^t \mathrm{d}A_k\, w_k + h_k(t) - h_k(a) \quad \text{for } k \in \mathbb{N} \text{ and } t \in [a,b],$$

where $\widetilde{w}_k = (\widetilde{x}_k - f_k(a)) - (\widetilde{x} - f(a))$, and

$$h_k(t) = \int_a^t \mathrm{d}[A_k - A]\,(x - f) + \Big(\int_a^t \mathrm{d}A_k\, f_k - \int_a^t \mathrm{d}A\, f \Big). \tag{7.6.16}$$

If we denote $\alpha^* = \sup_{k\in\mathbb{N}} \operatorname{var}_a^b A_k$, then it follows from Theorem 7.5.7 that

$$|w_k(t)| \le c_k \left(|\widetilde{w}_k| + 2\,\|h_k\| \right) \exp\left(2\, c_k \alpha^*\right), \quad t \in [a,b], \quad k \in \mathbb{N}$$

where

$$c_k = \max\Big\{1, \sup_{t\in(t_0,b]} \big|(I - \Delta^- A_k(t))^{-1}\big|, \sup_{t\in[a,t_0)} \big|(I + \Delta^+ A_k(t))^{-1}\big| \Big\}.$$

The sequence $\{c_k\}$ is bounded, because Lemma 7.6.4 implies

$$c_k \le \max\Big\{1, 2 \sup_{t\in(t_0,b]} \big|(I - \Delta^- A(t))^{-1}\big|, 2 \sup_{t\in[a,t_0)} \big|(I + \Delta^+ A(t))^{-1}\big| \Big\}.$$

Next, notice that $\widetilde{w}_k = \widetilde{x}_k - \widetilde{x} + f(a) - f_k(a) \to 0$. Hence, to show that $w_k \rightrightarrows 0$, it is enough to prove that $h_k \rightrightarrows 0$.

By Theorem 7.6.3, we have

$$\lim_{k\to\infty} \sup_{t\in[a,b]} \Big| \int_a^t \mathrm{d}A_k\, f_k - \int_a^t \mathrm{d}A\, f \Big| = 0. \tag{7.6.17}$$

From Theorem 6.3.5, we have the estimate

$$\Big|\int_a^t \mathrm{d}[A_k - A]\,(x-f)\Big| \le 2\,\|A_k - A\|\,\|x-f\|_{\mathrm{BV}} \quad \text{for } t\in[a,b].$$

Since $(x-f)\in \mathrm{BV}([a,b],\mathbb{R}^n)$ by Theorem 7.5.7, it follows that

$$\lim_{k\to\infty}\sup_{t\in[a,b]}\Big|\int_a^t \mathrm{d}[\,A_k - A]\,(x-f)\Big| = 0. \tag{7.6.18}$$

Taking into account (7.6.17)–(7.6.18), we see from (7.6.16) that $h_k \rightrightarrows 0$, and the proof is complete. □

7.7 Fundamental matrices

The equation

$$x(t) - x(s) - \int_s^t \mathrm{d}A\,x = 0 \tag{7.7.1}$$

is a generalization of a homogeneous system of linear ordinary differential equations. Assume that $A\in\mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$, $t_0\in[a,b]$, and conditions (7.4.10) and (7.4.11) are satisfied. For every $\widetilde{x}\in\mathbb{R}^n$, Theorem 7.4.8 (with $f=0$ on $[a,b]$) implies that equation (7.7.1) has a unique solution $x\colon[a,b]\to\mathbb{R}^n$ satisfying $x(t_0)=\widetilde{x}$. By the first part of Corollary 6.5.4, the solution x is a regulated function. Thus, the second part of the same corollary implies that x has bounded variation.

Clearly, the relation between solutions x of (7.7.1) and their values at the point t_0 is a one-to-one correspondence. It is easy to verify that if x, y are solutions of (7.7.1) on $[a,b]$ and $c_1, c_2\in\mathbb{R}$, then $c_1\,x + c_2\,y$ is also a solution of (7.7.1) on $[a,b]$. These observations are summarized in the following statement.

7.7.1 Theorem. *Let $A\in\mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$ and let (7.4.10) and (7.4.11) hold. Then the set of all solutions of equation (7.7.1) on $[a,b]$ is a linear subspace of $\mathrm{BV}([a,b],\mathbb{R}^n)$ having dimension n.*

We now introduce an analogue of the classical notion of a fundamental matrix.

7.7.2 Definition. A matrix-valued function $X\colon[a,b]\to\mathscr{L}(\mathbb{R}^n)$ is called a fundamental matrix of equation (7.7.1) on the interval $[a,b]$ if

$$X(t) = X(s) + \int_s^t \mathrm{d}A\,X \quad \text{for all } t,s\in[a,b] \tag{7.7.2}$$

and $\det X(t)\neq 0$ for at least one $t\in[a,b]$.

7.7.3 Remark. If a matrix-valued function X satisfies the relation (7.7.2), then it is easy to verify that for any $c \in \mathbb{R}^n$, the function $x(t) = X(t)c$ is a solution to (7.7.1).

7.7.4 Lemma. *Assume that* $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$, $t_0 \in [a,b]$, *and conditions* (7.4.10) *and* (7.4.11) *are satisfied.*

Then for every matrix $\widetilde{X} \in \mathscr{L}(\mathbb{R}^n)$, *there exists a unique matrix-valued function* $X_{t_0} \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$ *such that*

$$X_{t_0}(t) = \widetilde{X} + \int_{t_0}^{t} \mathrm{d}A\, X_{t_0} \quad \textit{for all } t \in [a,b]. \tag{7.7.3}$$

Proof. For each $k \in \{1, \ldots, n\}$, let $\widetilde{x}_k$ denote the k-th column of the matrix $\widetilde{X}$. Thus, $\widetilde{x}_k \in \mathbb{R}^n$ for $k = 1, \ldots, n$, and $\widetilde{X} = \big(\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_n\big)$. For each $k \in \{1, \ldots, n\}$, Theorem 7.4.8 implies the existence of a unique function $x_k : [a,b] \to \mathbb{R}^n$ satisfying the equation

$$x_k(t) - \widetilde{x}_k - \int_{t_0}^{t} \mathrm{d}A\, x_k = 0 \quad \text{for } t \in [a,b].$$

By Corollary 6.5.4, x_k has bounded variation on $[a,b]$. The function

$$X_{t_0}(t) = \big(x_1(t), x_2(t), \ldots, x_n(t)\big)$$

(i.e., the matrix-valued function with the columns x_k, $k = 1, \ldots, n$) is therefore a unique solution of (7.7.3) and has bounded variation on $[a,b]$. □

7.7.5 Remark. If $t_0 \in [a,b]$, $\widetilde{X} \in \mathscr{L}(\mathbb{R}^n)$ and X_{t_0} is the function defined by Lemma 7.7.4, then the function $X = X_{t_0}$ obviously satisfies (7.7.2). Therefore, X is a fundamental matrix of equation (7.7.1) whenever $\det \widetilde{X} \neq 0$.

For simplicity, we will now assume that

$$\left.\begin{aligned} &\det\big(I - \Delta^- A(t)\big) \neq 0 \qquad \text{for every } t \in (a,b], \\ &\det\big(I + \Delta^+ A(s)\big) \neq 0 \qquad \text{for every } s \in [a,b). \end{aligned}\right\} \tag{7.7.4}$$

7.7.6 Lemma. *If* $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$, *condition* (7.7.4) *holds, and* X *is a fundamental matrix of* (7.7.1) *on* $[a,b]$, *then*

$$\det X(t) \neq 0 \quad \textit{for each } t \in [a,b]. \tag{7.7.5}$$

Proof. If X is a fundamental matrix of (7.7.1) and (7.7.5) does not hold, then there exist points $\tau_0, \tau_1 \in [a,b]$ such that

$$\det X(\tau_0) \neq 0 \qquad \text{and} \qquad \det X(\tau_1) = 0.$$

The latter equality implies that the columns $x_1(\tau_1), x_2(\tau_1), \ldots, x_n(\tau_1)$ of the matrix $X(\tau_1)$ are linearly dependent. Hence, there are coefficients $c_1, c_2, \ldots, c_n \in \mathbb{R}$, not all zero, such that

$$\sum_{k=1}^{n} c_k x_k(\tau_1) = 0.$$

Set $x(t) = \sum_{k=1}^{n} c_k\, x_k(t)$ for each $t \in [a,b]$. Then x is a solution to (7.7.1) (see Remark 7.7.3) with $x(\tau_1) = 0$, i.e., x is a solution of the initial-value problem

$$x(t) = \int_{\tau_1}^{t} \mathrm{d}A\, x \quad \text{for } t \in [a,b].$$

However, the same equation also has the trivial solution. Since Theorem 7.4.8 guarantees uniqueness of solutions (note that (7.7.4) implies that conditions (7.4.10) and (7.4.11) are satisfied for $t_0 = \tau_1$), we necessarily have $x = 0$ on $[a,b]$. In particular,

$$x(\tau_0) = \sum_{k=1}^{n} c_k\, x_k(\tau_0) = 0,$$

which contradicts the assumption that $\det X(\tau_0) \neq 0$. □

For functions of two variables, we use the following notation.

7.7.7 Notation. Consider a function $U : [a,b] \times [a,b] \to \mathscr{L}(\mathbb{R}^n)$. For each fixed $\tau \in [a,b]$, the symbol $U(\tau,\cdot)$ stands for the function $s \mapsto U(\tau,s)$ of a single variable $s \in [a,b]$. Similarly, the symbol $U(\cdot,\tau)$ denotes the function $t \mapsto U(t,\tau)$ of a single variable $t \in [a,b]$. Finally, we let

$$U(\tau,s+) = \lim_{\delta\to 0+} U(\tau,s+\delta), \quad U(\tau,s-) = \lim_{\delta\to 0+} U(\tau,s-\delta),$$
$$U(t+,\tau) = \lim_{\delta\to 0+} U(t+\delta,\tau), \quad U(t-,\tau) = \lim_{\delta\to 0+} U(t-\delta,\tau)$$

whenever the limits exist.

The following assertion follows from Lemmas 7.7.4 and 7.7.6.

7.7.8 Theorem. *If $A \in \mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$ satisfies (7.7.4), then there exists a unique matrix-valued function $U : [a,b] \times [a,b] \to \mathscr{L}(\mathbb{R}^n)$ such that*

$$U(t,s) = I + \int_{s}^{t} \mathrm{d}[A(\tau)]\, U(\tau,s) \quad \textit{for all } t,s \in [a,b] \times [a,b]. \tag{7.7.6}$$

For each $t_0 \in [a,b]$, the function $U(\cdot,t_0)$ is a fundamental matrix of (7.7.1) on $[a,b]$. Furthermore, U has the following properties:

(i) $U(\cdot, s) \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$ *for every* $s \in [a,b]$.

(ii) $U(t,t) = I$ *for every* $t \in [a,b]$.

(iii) $\det U(t,s) \neq 0$ *for all* $t, s \in [a,b]$.

Proof. For each $t, s \in [a,b]$, let $U(t,s) = X_s(t)$, where $X_s \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$ is the unique matrix-valued function satisfying

$$X_s(t) = I + \int_s^t \mathrm{d}A\, X_s \quad \text{for } t \in [a,b].$$

(see Lemma 7.7.4). Then U satisfies equation (7.7.6) and $U(t,t) = I$ for $t \in [a,b]$. By Remark 7.7.5, this means that for every $t_0 \in [a,b]$, the function $U(\cdot, t_0)$ is a fundamental matrix of (7.7.1) on $[a,b]$. Finally, the fact that $\det U(t,s) \neq 0$ for all $t, s \in [a,b]$ follows from Lemma 7.7.6. □

7.7.9 Theorem. *Assume that* $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$, $t_0 \in [a,b]$, $\widetilde{x} \in \mathbb{R}^n$, (7.7.4) *holds and* U *is the matrix-valued function defined by Theorem* 7.7.8. *Then* $x : [a,b] \to \mathbb{R}^n$ *is a solution of the initial-value problem*

$$x(t) - \widetilde{x} - \int_{t_0}^t \mathrm{d}A\, x = 0 \tag{7.7.7}$$

on $[a,b]$ *if and only if*

$$x(t) = U(t,t_0)\,\widetilde{x} \ \textit{ for } t \in [a,b]. \tag{7.7.8}$$

Proof. The function x given by the relation (7.7.8) is a solution of (7.7.7), because (7.7.6) implies

$$\int_{t_0}^t \mathrm{d}A\, x = \int_{t_0}^t \mathrm{d}[A(\tau)]\, U(\tau,s)\,\widetilde{x} = \big(U(t,t_0) - I\big)\,\widetilde{x} = x(t) - \widetilde{x} \quad \text{for } t \in [a,b].$$

By Theorem 7.4.8, this solution of (7.7.7) is necessarily unique. □

7.7.10 Definition. If $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$ and conditions (7.7.4) hold, then the matrix-valued function U defined by Theorem 7.7.8 is called the *Cauchy matrix* of equation (7.7.1) on $[a,b]$.

7.7.11 Corollary. *Let* $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$, $t_0 \in [a,b]$ *and* $\widetilde{X} \in \mathscr{L}(\mathbb{R}^n)$. *Further, assume that conditions* (7.7.4) *hold and* U *is the Cauchy matrix of* (7.7.1) *on* $[a,b]$. *Then a matrix-valued function* $X : [a,b] \to \mathscr{L}(\mathbb{R}^n)$ *satisfies the equation*

$$X(t) = \widetilde{X} + \int_{t_0}^t \mathrm{d}A\, X$$

if and only if $X(t) = U(t,t_0)\widetilde{X}$ *for all* $t \in [a,b]$.

Proof. For each $k = 1, \dots, n$, if x_k is the k-th column of the matrix-valued function X, we know from Theorem 7.7.9 that

$$x_k(t) = U(t, t_0)\widetilde{x}_k \quad \text{for all } t \in [a, b],$$

where $\widetilde{x}_k$ is the k-th column of the matrix $\widetilde{X}$. This completes the proof. □

7.7.12 Theorem. *If $A \in \mathrm{BV}([a, b], \mathscr{L}(\mathbb{R}^n))$, the conditions (7.7.4) are satisfied and U is the Cauchy matrix of (7.7.1) on $[a, b]$, then the relations*

$$U(t, r)\, U(r, s) = U(t, s), \tag{7.7.9}$$

$$(U(t, r))^{-1} = U(r, t) \tag{7.7.10}$$

hold for any triplet of points $t, s, r \in [a, b]$.

Proof. Let $r, s \in [a, b]$ be given. Using the definition of U, we have

$$\begin{aligned} U(t, s) &= I + \int_s^t \mathrm{d}[A(\tau)]\, U(\tau, s) \\ &= I + \int_s^r \mathrm{d}[A(\tau)]\, U(\tau, s) + \int_r^t \mathrm{d}[A(\tau)]\, U(\tau, s) \\ &= U(r, s) + \int_r^t \mathrm{d}[A(\tau)]\, U(\tau, s) \end{aligned}$$

for all $t \in [a, b]$. Hence, the relation (7.7.9) follows by Corollary 7.7.11.

Inserting $s = t$, we get

$$U(t, r)\, U(r, t) = U(t, t) = I,$$

which implies relation (7.7.10). □

7.7.13 Remark. If U is the Cauchy matrix for equation (7.7.1) on $[a, b]$, then, by Theorem 7.7.12, we have

$$U(t, s) = U(t, a)\, U(a, s) = U(t, a)U(s, a)^{-1} \quad \text{for all } t, s \in [a, b].$$

Thus, if we set $X(t) = U(t, a)$ for $t \in [a, b]$, then

$$U(t, s) = X(t)X(s)^{-1}$$

for all $t, s \in [a, b]$. Obviously, X is a fundamental matrix for (7.7.1) on $[a, b]$.

In the rest of this section, we will describe some additional properties of the Cauchy matrix of (7.7.1) on $[a, b]$.

7.7.14 Theorem. *Let* $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$, *let* (7.7.4) *hold and let* U *be the Cauchy matrix of* (7.7.1) *on* $[a,b]$. *Then there exists an* $M \in (0,\infty)$ *such that*

$$|U(t,s)| + \operatorname{var}_a^b U(\cdot,s) + \operatorname{var}_a^b U(t,\cdot\,) \le M \quad \textit{for all } t,s \in [a,b]. \tag{7.7.11}$$

Proof. a) For $k \in \{1,\dots,n\}$, let e_k stand for the kth column of the unit matrix I. Then $|e_k| = 1$ for every $k \in \{1,\dots,n\}$. By Theorem 7.5.7, for $t,s \in [a,b]$ and $k \in \{1,\dots,n\}$, we have

$$|u_k(t,s)| \le M_1 := c_A \, \exp\big(2\,c_A \operatorname{var}_a^b A\big) < \infty, \tag{7.7.12}$$

where, thanks to assumption (7.7.4), we can set

$$c_A := \max\Big\{1, \sup_{t\in(a,b]} \big|[I - \Delta^- A(t)]^{-1}\big|, \sup_{t\in[a,b)} \big|[I + \Delta^+ A(t)]^{-1}\big|\Big\}.$$

Thus,

$$|U(t,s)| = \max_{k=1,\dots,n} |u_k(t,s)| \le M_1 \quad \text{for } t,s \in [a,b]. \tag{7.7.13}$$

b) Let $t_1, t_2, s \in [a,b]$ and $t_1 \le t_2$. Then

$$|u_k(t_2,s) - u_k(t_1,s)| = \Big|\int_{t_1}^{t_2} \mathrm{d}[A(\tau)]\, u_k(\tau,s)\Big| \le M_1 \operatorname{var}_{t_1}^{t_2} A$$

for each $k \in \{1,\dots,n\}$. Hence, for all $s \in [a,b]$, $k \in \{1,\dots,n\}$ and all divisions $\boldsymbol{\alpha}$ of $[a,b]$, we have

$$V(u_k(\cdot,s),\boldsymbol{\alpha}) \le M_1 \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \operatorname{var}_{\alpha_{j-1}}^{\alpha_j} A = M_1 \operatorname{var}_a^b A =: M_2 < \infty.$$

Therefore

$$\operatorname{var}_a^b U(\cdot,s) \le \max_{k=1,\dots,n} \operatorname{var}_a^b u_k(\cdot,s) \le M_2 \quad \text{for all } s \in [a,b]. \tag{7.7.14}$$

c) Let $s_1, s_2 \in [a,b]$ and $s_1 \le s_2$. Then, for every $t \in [a,b]$ and any $k \in \{1,\dots,n\}$, we have

$$\begin{aligned}
&u_k(t,s_2) - u_k(t,s_1)\\
&\quad= \int_{s_2}^{t} \mathrm{d}[A(\tau)]\, u_k(\tau,s_2) - \int_{s_2}^{t} \mathrm{d}[A(\tau)]\, u_k(\tau,s_1) - \int_{s_1}^{s_2} \mathrm{d}[A(\tau)]\, u_k(\tau,s_1)\\
&\quad= -\int_{s_1}^{s_2} \mathrm{d}[A(\tau)] u_k(\tau,s_1) + \int_{s_2}^{t} \mathrm{d}[A(\tau)]\big(u_k(\tau,s_2) - u_k(\tau,s_1)\big).
\end{aligned}$$

Hence, the function $x(t) = u_k(t,s_2) - u_k(t,s_1)$ is a solution of the initial problem

$$x(t) = -\int_{s_1}^{s_2} \mathrm{d}[A(\tau)]u_k(\tau,s_1) + \int_{s_2}^{t} \mathrm{d}A\,x$$

on $[a,b]$. Consequently, by Theorem 7.7.9 we have

$$u_k(t,s_2) - u_k(t,s_1) = -U(t,s_2)\Big(\int_{s_1}^{s_2} \mathrm{d}[A(\tau)]u_k(\tau,s_1)\Big)$$

for $t \in [a,b]$ and $k \in \{1,\dots,n\}$ and, thus,

$$|u_k(t,s_2) - u_k(t,s_1)| \le M_1^2 \operatorname{var}_{s_1}^{s_2} A \quad \text{for } k \in \{1,\dots,n\} \text{ and } t \in [a,b]. \tag{7.7.15}$$

As a result, for all $t \in [a,b]$, $k = 1,\dots,n$ and all divisions $\boldsymbol{\alpha}$ of $[a,b]$, we get

$$V(u_k(t,\cdot),\boldsymbol{\alpha}) \le M_1^2 \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \operatorname{var}_{\alpha_{j-1}}^{\alpha_j} A = M_1^2 \operatorname{var}_a^b A < \infty.$$

Consequently

$$\operatorname{var}_a^b U(t,\cdot) \le \max_{k=1,\dots,n} \operatorname{var}_a^b u_k(t,\cdot) \le M_1^2 \operatorname{var}_a^b A < \infty \quad \text{for } t \in [a,b]$$

and hence

$$\operatorname{var}_a^b U(t,\cdot) \le M_1^2 \operatorname{var}_a^b A =: M_3 < \infty \quad \text{for } t \in [a,b]. \tag{7.7.16}$$

d) By virtue of (7.7.11)–(7.7.14), the statement of the theorem holds with

$$M = M_1 + M_2 + M_3. \qquad \square$$

7.7.15 Remark. Considering an arbitrary subinterval $[s_1,s_2]$ of $[a,b]$ in place of $[a,b]$, one can see that also the following estimates are true

$$\left.\begin{aligned} &\operatorname{var}_{s_1}^{s_2} U(t,\cdot) \le M_1^2 \operatorname{var}_{s_1}^{s_2} A \\ &\quad \text{for all } s_1,s_2 \in [a,b] \text{ such that } s_1 \le s_2 \text{ and all } t \in [a,b]. \end{aligned}\right\} \tag{7.7.17}$$

7.7.16 Theorem. *Let* $A \in \mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$, *let conditions* (7.7.4) *be satisfied and let* U *be the Cauchy matrix of equation* (7.7.1) *on* $[a,b]$. *Then*

$$\left.\begin{aligned} U(t+,s) &= \big[I + \Delta^+ A(t)\big]\,U(t,s) && \textit{for } t \in [a,b),\ s \in [a,b],\\ U(t-,s) &= \big[I - \Delta^- A(t)\big]\,U(t,s) && \textit{for } t \in (a,b],\ s \in [a,b],\\ U(t,s+) &= U(t,s)\,\big[I + \Delta^+ A(s)\big]^{-1} && \textit{for } t \in [a,b],\ s \in [a,b),\\ U(t,s-) &= U(t,s)\,\big[I - \Delta^+ A(s)\big]^{-1} && \textit{for } t \in [a,b],\ s \in (a,b]. \end{aligned}\right\} \tag{7.7.18}$$

Proof. First, notice that, by the previous theorem, the functions $U(\cdot,s)$ and $U(t,\cdot)$ have bounded variations on $[a,b]$ for all $t,s\in[a,b]$. Thus, all one-sided limits appearing in the relationships (7.7.18) are well justified.

a) The first two relations in (7.7.3) can be derived if we substitute successively the columns of the function U for x into relations (7.4.1) from Lemma 7.4.7.

b) Let $s\in[a,b)$ and $\delta\in(0,b-s)$ be given. Then from (7.7.6) we deduce that

$$\begin{aligned} U(t,s+\delta)-U(t,s) &= \int_{s+\delta}^{t} \mathrm{d}[A(\tau)]\,U(\tau,s+\delta) - \int_{s}^{t} \mathrm{d}[A(\tau)]\,U(\tau,s) \\ &= \int_{s+\delta}^{t} \mathrm{d}[A(\tau)]\,\big(U(\tau,s+\delta)-U(\tau,s)\big) - \int_{s}^{s+\delta} \mathrm{d}[A(\tau)]\,U(\tau,s) \end{aligned}$$

hold for each $t\in[a,b]$. Thus, the function $Y(t)=U(t,s+\delta)-U(t,s)$ satisfies the equation

$$Y(t)=\widetilde{Y}+\int_{s+\delta}^{t} \mathrm{d}[A(\tau)]\,Y(\tau) \quad \text{for } t\in[a,b],$$

where

$$\widetilde{Y}=-\int_{s}^{s+\delta} \mathrm{d}[A(\tau)]\,U(\tau,s).$$

Then, Corollary 7.7.11 yields

$$\begin{aligned} U(t,s+\delta)-U(t,s) &= U(t,s+\delta)\,\widetilde{Y} \\ &= -U(t,s+\delta)\int_{s}^{s+\delta} \mathrm{d}[A(\tau)]\,U(\tau,s) \quad \text{for } t\in[a,b]. \end{aligned}$$

Letting $\delta\to 0+$ and using Hake's Theorem 6.5.5, we get for $t\in[a,b]$

$$U(t,s+)-U(t,s)=-U(t,s+)\,\Delta^+A(s)\,U(s,s)=-U(t,s+)\,\Delta^+A(s),$$

that is

$$U(t,s)=U(t,s+)\,\big[I+\Delta^+A(s)\big].$$

Therefore the third relationship from (7.7.18) is true, as well. The remaining one would be proven analogously. □

7.8 Variation of constants formula

Let us now go back to the nonhomogeneous initial value problem

$$x(t)-\widetilde{x}-\int_{t_0}^{t} \mathrm{d}A\,x = f(t)-f(t_0) \tag{7.8.1}$$

(cf. (7.4.2)). We will assume that $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$ satisfies conditions (7.7.4). By Theorem 7.6.2, for any $\widetilde{x} \in \mathbb{R}^n$ and $f \in \mathrm{G}([a,b], \mathbb{R}^n)$, there is exactly one solution x of the initial value problem (7.8.1) and this solution is regulated on $[a,b]$.

The aim of this section is to show that this solution can be expressed in a form resembling the *variation of constants* formula known from the theory of ordinary differential equations. To establish this result, we need several auxiliary lemmas.

7.8.1 Lemma. *Consider a function* $K : [a,b] \times [a,b] \to \mathscr{L}(\mathbb{R}^n)$ *such that* $K(\cdot, s)$ *is regulated for each* $s \in [a,b]$*. Moreover, suppose there exists a nondecreasing function*
$h : [a,b] \to \mathbb{R}$ *such that*

$$|K(t,s_2) - K(t,s_1)| \le h(s_2) - h(s_1) \quad \textit{for } t \in [a,b] \textit{ and } [s_1,s_2] \subset [a,b]. \tag{7.8.2}$$

Then, for each $g \in \mathrm{G}([a,b])$ *the function*

$$\psi(t) = \int_a^b \mathrm{d}_s[K(t,s)]\, g(s) \qquad \textit{for } t \in [a,b], \tag{7.8.3}$$

is regulated, and

$$\psi(t-) = \int_a^b \mathrm{d}_s[K(t-,s)]\, g(s) \quad \textit{for } t \in (a,b],$$

$$\psi(t+) = \int_a^b \mathrm{d}_s[K(t+,s)]\, g(s) \quad \textit{for } t \in [a,b).$$

Proof. Let $g \in \mathrm{G}([a,b], \mathbb{R}^n)$ and let ψ be given by (7.8.3). Obviously, ψ is then well defined on $[a,b]$, because (7.8.2) implies that $K(t,\cdot)$ has a bounded variation for each $t \in [a,b]$.

a) First, let an arbitrary $\tau \in (a,b]$ be given and let $\{t_k\}$ be an arbitrary sequence of points from $[a,\tau)$ such that $\lim_{k\to\infty} t_k = \tau$. Having in mind (7.8.2), it is easy to verify (cf. also Exercise 2.1.12) that

$$\operatorname{var}_a^b K(t_k,\cdot) \le h(b) - h(a) \quad \text{for every } k \in \mathbb{N}, \tag{7.8.4}$$

$$\lim_{k\to\infty} K(t_k,s) = K(\tau-,s) \qquad \text{for every } s \in [a,b] \tag{7.8.5}$$

and, by Theorem 2.7.2, also

$$\operatorname{var}_a^b K(\tau-,\cdot) \le h(b) - h(a),$$

i.e. $K(\tau-,\cdot) \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$. Furthermore, thanks to (7.8.2) the convergence in (7.8.5) is by Corollary 4.3.10 uniform. Therefore, Theorem 7.6.3, where we put

$$A_k = K(t_k, \cdot),\ A = K(\tau-, \cdot) \quad \text{and} \quad f_k = g \quad \text{for} \quad k \in \mathbb{N},$$

ensures that the relations

$$\lim_{k\to\infty} \psi(t_k) = \lim_{k\to\infty} \int_a^b \mathrm{d}_s[K(t_k, s)]\, g(s) = \int_a^b \mathrm{d}_s[K(\tau-, s)]\, g(s)$$

are true. In particular,

$$\psi(\tau-) = \int_a^b \mathrm{d}_s[K(\tau-, s)]\, g(s) \quad \text{for all } \tau \in (a, b].$$

b) The proof of the relation

$$\psi(\tau+) = \int_a^b \mathrm{d}_s[K(\tau+, s)]\, g(s) \quad \text{for all } \tau \in [a, b)$$

is analogous and may be left to the reader.

All this together leads us to conclude that ψ is regulated. □

Next result is a corollary of the previous lemma.

7.8.2 Corollary. *Let* $t_0 \in [a, b],\ A \in \mathrm{BV}([a, b], \mathscr{L}(\mathbb{R}^n))$ *satisfy conditions* (7.7.4) *and let* U *be the Cauchy matrix of equation* (7.7.1) *on* $[a, b]$*. Then, for each* $g \in \mathrm{G}([a, b], \mathbb{R}^n)$ *the function*

$$\psi(t) := \int_{t_0}^t \mathrm{d}_s[U(t, s)]\, g(s) \quad \textit{for } t \in [a, b]$$

is regulated on $[a, b]$.

Proof. The function ψ is obviously well defined on $[a, b]$. Furthermore,

$$\psi(t) = \begin{cases} \displaystyle\int_a^{t_0} \mathrm{d}_s[K_a(t, s)]\, g(s) & \text{if } t \in [a, t_0], \\ \displaystyle\int_{t_0}^b \mathrm{d}_s[K_b(t, s)]\, g(s) & \text{if } t \in [t_0, b], \end{cases}$$

where

$$K_a(t, s) = \begin{cases} -U(t, t) & \text{if } a \le s \le t \le t_0, \\ -U(t, s) & \text{if } a \le t \le s \le t_0 \end{cases}$$

and

$$K_b(t,s) = \begin{cases} U(t,s) & \text{if } t_0 \le s < t \le b, \\ U(t,t) & \text{if } t_0 \le t \le s \le b. \end{cases}$$

By (7.7.17), there is a constant $M_1 \in (0,\infty)$ such that for every $t \in [a,t_0]$ and every subinterval $[s_1,s_2]$ of $[a,t_0]$ we have

$$|K_a(t,s_2) - K_a(t,s_1)| \le \operatorname{var}_{s_1}^{s_2} K_a(t,\cdot) \le \operatorname{var}_{s_1}^{s_2} U(t,\cdot) \le M_1^2 \operatorname{var}_{s_1}^{s_2} A,$$

i.e.,

$$|K_a(t,s_2) - K_a(t,s_1)| \le h_a(s_2) - h_a(s_1) \quad \text{for } t \in [a,t_0] \text{ and } [s_1,s_2] \subset [a,t_0],$$

where $h_a(s) = M_1^2 \operatorname{var}_a^s A$ for $s \in [a,t_0]$. Similarly,

$$|K_b(t,s_2) - K_b(t,s_1)| \le h_b(s_2) - h_b(s_1) \quad \text{for } t \in [t_0,b] \text{ and } [s_1,s_2] \subset [t_0,b],$$

where $h_b(s) = M_1^2 \operatorname{var}_{t_0}^s A$ for $s \in [t_0,b]$.

Now, taking either $\{[a,t_0], K_a, h_a\}$ or $\{[t_0,b], K_b, h_b\}$ in place of $\{[a,b], K, h\}$, we can apply Lemma 7.8.1 to justify that ψ is regulated on $[a,t_0]$ or $[t_0,b]$, respectively. □

Next auxiliary assertion deals with adjusting an iterated integral which will be useful for the proof of the main result of this section.

7.8.3 Lemma. *Let* $t_0 \in [a,b]$, $A \in \mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$ *satisfy conditions* (7.7.4) *and let* U *be the Cauchy matrix of equation* (7.7.1) *on* $[a,b]$*. Then, for every* $g \in \mathrm{G}([a,b],\mathbb{R}^n)$ *and* $t \in [a,b]$ *we have*

$$\left.\begin{aligned} &\int_{t_0}^t \mathrm{d}[A(r)]\Big(\int_{t_0}^r \mathrm{d}_s\,[U(r,s)]\,g(s)\Big) \\ &\qquad = \int_{t_0}^t \mathrm{d}_s[U(t,s)]\,g(s) + \int_{t_0}^t \mathrm{d}[A(s)]\,g(s). \end{aligned}\right\} \tag{7.8.6}$$

Proof. The fact that U has bounded variation in both variables together with Corollary 7.8.2 imply that all integrals in (7.8.6) have a sense.

a) Given an arbitrary $t > t_0$, we will consider the interval $[t_0,t]$ fixed along the proof (the case $t \le t_0$ can be treated in similar way). Recalling that regulated functions can be uniformly approximated by finite step functions, let us first prove (7.8.6) for functions g of the form

$$g = \chi_{(\tau,t]}\,\xi,\ g = \chi_{[\tau,t]}\,\xi,\ g = \chi_{[t]}\,\xi, \quad \text{where } \tau \in [t_0,t) \text{ and } \xi \in \mathbb{R}^n \text{ can be arbitrary.} \tag{7.8.7}$$

So, let $\tau \in [t_0,t)$ and $\xi \in \mathbb{R}^n$ be given and let $g = \chi_{(\tau,t]}\,\xi$. First, having in mind Examples 6.3.1 (cf. (6.3.6)), we can see that the relations

$$\int_{t_0}^t \mathrm{d}[A(s)]\,g(s) = \int_{\tau}^t \mathrm{d}[A(s)]\,g(s) = [A(t) - A(\tau+)]\,\xi \tag{7.8.8}$$

hold. Like in Corollary 7.8.2 put

$$\psi(r) := \int_{t_0}^{r} \mathrm{d}_s[U(r,s)]\, g(s) \quad \text{for } r \in [t_0, t].$$

Using again Examples 6.3.1 and having in mind that $U(r,r) = I$ for all r, it is not difficult to verify that

$$\psi(r) = \chi_{(\tau,t]}(r)\, [I - U(r,\tau+)]\, \xi \quad \text{for } r \in [t_0, t]. \tag{7.8.9}$$

Further, applying Theorem 6.5.3, we get

$$\begin{aligned}\int_{t_0}^{t} \mathrm{d}[A(r)]\, \psi(r) &= \int_{\tau}^{t} \mathrm{d}[A(r)]\, \psi(r) = \lim_{\sigma\to\tau+} \int_{\sigma}^{t} \mathrm{d}[A(r)]\, \psi(r)\\ &= \lim_{\sigma\to\tau+} \int_{\sigma}^{t} \mathrm{d}[A(r)]\, [I - U(r,\tau+)]\, \xi,\end{aligned}$$

i.e.,

$$\int_{t_0}^{t} \mathrm{d}[A(r)]\psi(r) = (A(t) - A(\tau+))\, \xi - \lim_{\sigma\to\tau+} \int_{\sigma}^{t} \mathrm{d}[A(r)]\, U(r,\tau+)\, \xi. \tag{7.8.10}$$

Next, applying Theorem 6.5.3 again, and then Theorem 7.7.16 and relation (7.7.6), we deduce

$$\begin{aligned}\lim_{\sigma\to\tau+} \int_{\sigma}^{t} \mathrm{d}[A(r)]\, U(r,\tau+)\, \xi &= \Big(\int_{\tau}^{t} \mathrm{d}[A(r)]\, U(r,\tau+) - \Delta^+ A(\tau)\, U(\tau,\tau+) \Big)\, \xi\\ &= \Big(\int_{\tau}^{t} \mathrm{d}[A(r)]\, U(r,\tau)[I + \Delta^+ A(\tau)]^{-1} - \Delta^+ A(\tau)\, U(\tau,\tau)\, [I + \Delta^+ A(\tau)]^{-1} \Big) \xi\\ &= \Big((U(t,\tau) - I)\, [I + \Delta^+ A(\tau)]^{-1} - \Delta^+ A(\tau)\, [I + \Delta^+ A(\tau)]^{-1} \Big)\, \xi\\ &= \Big(U(t,\tau)\, [I + \Delta^+ A(\tau)]^{-1} - [I + \Delta^+ A(\tau)]\, [I + \Delta^+ A(\tau)]^{-1} \Big)\, \xi\\ &= \Big(U(t,\tau+) - I \Big)\, \xi = -\psi(t),\end{aligned}$$

wherefrom we conclude using (7.8.8) – (7.8.10) that the relations

$$\int_{t_0}^{t} \mathrm{d}[A(r)]\, \psi(r) = (A(t) - A(\tau+))\, \xi + \psi(t) = \psi(t) + \int_{t_0}^{t} \mathrm{d}[A(s)]\, g(s)$$

are true, that is, (7.8.6) holds.

Analogously, we can show that relation (7.8.6) holds also for $t \le t_0$ and for all other functions g from the set (7.8.7).

b) Assume that g is an arbitrary regulated function on $[a,b]$ and let $\{g_n\}$ be a sequence of finite step functions converging uniformly to g on $[t_0,t]$. By the first part of the proof we have

$$\int_{t_0}^{t} \mathrm{d}[A(r)]\Big(\int_{t_0}^{r} \mathrm{d}_s\,[U(r,s)\,]\,g_n(s)\Big) = \int_{t_0}^{t} \mathrm{d}_s[U(t,s)]\,g_n(s) + \int_{t_0}^{t} \mathrm{d}[A(s)]\,g_n(s)$$

for each $n \in \mathbb{N}$. Furthermore, by Theorem 6.3.7

$$\lim_{n\to\infty} \int_{t_0}^{t} \mathrm{d}[A(s)]\,g_n(s) = \int_{t_0}^{t} \mathrm{d}[A(s)]\,g(s)$$

and

$$\lim_{n\to\infty} \int_{t_0}^{t} \mathrm{d}_s[U(t,s)]\,g_n(s) = \int_{t_0}^{t} \mathrm{d}_s[U(t,s)]\,g(s). \tag{7.8.11}$$

It remains to show that

$$\lim_{n\to\infty} \int_{t_0}^{t} \mathrm{d}[A(s)]\,\psi_n(s) = \int_{t_0}^{t} \mathrm{d}[A(s)]\,\psi(s), \tag{7.8.12}$$

where

$$\psi(r) = \int_{t_0}^{r} \mathrm{d}_s\,[U(r,s)\,]\,g(s) \quad\text{and}\quad \psi_n(r) = \int_{t_0}^{r} \mathrm{d}_s\,[U(r,s)\,]\,g_n(s),$$

for $n \in \mathbb{N}$ and $r \in [t_0,t]$. Note that, by Theorem 6.3.4

$$|\psi_n(r)| \le \operatorname{var}_{t_0}^{r} U(r,\cdot)\,\|g_n\| \quad \text{for } n \in \mathbb{N} \text{ and } r \in [t_0,t].$$

The fact that the sequence $\{g_n\}$ is uniformly bounded, together with Theorem 7.7.14 implies that the sequence $\{\psi_n\}$ is also uniformly bounded. Since (7.8.11) guarantees that $\lim_{n\to\infty} \psi_n(r) = \psi(r)$ for each $r \in [t_0,t]$, the convergence in (7.8.12) is a consequence of Theorem **??** and this completes the proof of the assertion. □

Now we can state and prove the promised analogy of the classical *variation of constants formula* for solutions of nonhomogeneous linear generalized differential equations.

7.8.4 Theorem. *Let $t_0 \in [a,b]$, $A \in \mathrm{BV}([a,b],\mathscr{L}(\mathbb{R}^n))$ satisfy conditions (7.7.4) and let U be the Cauchy matrix of equation (7.7.1) on $[a,b]$. Then problem (7.8.1) has for every $\widetilde{x} \in \mathbb{R}^n$ and every $f \in \mathrm{G}([a,b],\mathbb{R}^n)$ exactly one solution x on $[a,b]$. This solution is given by the formula*

$$x(t) = U(t,t_0)\,\widetilde{x} + f(t) - f(t_0) - \int_{t_0}^{t} \mathrm{d}_s[U(t,s)]\,\big(f(s) - f(t_0)\big) \quad \textit{for } t \in [a,b]. \tag{7.8.13}$$

Proof. Let an arbitrary $t \in [a,b]$ be given.

By Corollary 7.8.2, the function

$$x(s) = U(s,t_0)\,\widetilde{x} + f(s) - f(t_0) - \int_{t_0}^{s} \mathrm{d}_r[U(s,r)]\,(f(r) - f(t_0))$$

is regulated on $[a,b]$. Inserting it into the integral $\int_{t_0}^{t} \mathrm{d}[A(s)]\,x(s)$, we get

$$\begin{aligned}\int_{t_0}^{t} \mathrm{d}[A(s)]\,x(s) = \int_{t_0}^{t} \mathrm{d}[A(s)]\,U(s,t_0)\,\widetilde{x} + \int_{t_0}^{t} \mathrm{d}[A(s)]\,(f(s) - f(t_0))\\ - \int_{t_0}^{t} \mathrm{d}[A(s)]\left(\int_{t_0}^{s} \mathrm{d}_r\,[U(s,r)]\,(f(r) - f(t_0))\right),\end{aligned}$$

where both sides have a sense. Further, by Theorem 7.7.9 we have

$$\int_{t_0}^{t} \mathrm{d}[A(s)]\,U(s,t_0)\,\widetilde{x} = U(t,t_0)\,\widetilde{x} - \widetilde{x},$$

while Lemma 7.8.3 with $g(s) = f(s) - f(t_0)$ yields

$$\begin{aligned}&\int_{t_0}^{t} \mathrm{d}[A(r)]\left(\int_{t_0}^{r} \mathrm{d}_s\,[U(r,s)]\,(f(s) - f(t_0))\right)\\ &\qquad = \int_{t_0}^{t} \mathrm{d}_s[U(t,s)]\,(f(s) - f(t_0)) + \int_{t_0}^{t} \mathrm{d}[A(s)]\,(f(s) - f(t_0)).\end{aligned}$$

Therefore,

$$\begin{aligned}\int_{t_0}^{t} \mathrm{d}[A(s)]\,x(s) &= U(t,t_0)\,\widetilde{x} - \widetilde{x} - \int_{t_0}^{t} \mathrm{d}_s[U(t,s)]\,(f(s) - f(t_0))\\ &= x(t) - \widetilde{x} - (f(t) - f(t_0)),\end{aligned}$$

which proves the result. □

In the case that A is left-continuous on $(a,b]$ and $t_0 = a$, formula (7.8.13) can be somewhat simplified, if we define $X(t) = U(t,a)$ for $t \in [a,b]$ and

$$Y(s) = \begin{cases} U(a,s+), & \text{if } \; a \le s < b,\\ U(a,b), & \text{if } \; s = b.\end{cases} \tag{7.8.14}$$

7.8.5 Corollary. *Let* $t_0 = a$ *and let* $A \in \mathrm{BV}([a,b], \mathscr{L}(\mathbb{R}^n))$ *be left-continuous on* $(a,b]$. *Furthermore, let*

$$\det[I + \Delta^+ A(t)] \ne 0 \quad \textit{for } t \in [a,b)$$

and let U *be the Cauchy matrix for* (7.7.1),

$$X(t)=U(t,a) \quad \textit{pro} \;\; t\in[a,b]$$

and let Y *be given by* (7.8.14).

Then equation (7.6.7) *have for each* $\widetilde{x}\in\mathbb{R}^n$ *and each* $f\in G([a,b],\mathbb{R}^n)$ *left-continuous on* $(a,b]$ *a unique solution* x *on* $[a,b]$. *This solution is given by*

$$x(t)=X(t)\,\widetilde{x}+X(t)\left(\int_a^t Y\,\mathrm{d}\,f\right) \quad \textit{for} \;\; t\in[a,b]. \tag{7.8.15}$$

Proof. Let $\widetilde{x}\in\mathbb{R}^n$ and let $f\in G([a,b],\mathbb{R}^n)$ be continuous from the left on $(a,b]$. By Theorem 7.8.4, equation (7.6.7) has a unique solution x on $[a,b]$ and formula (7.8.13) can be rewritten as

$$x(t)=X(t)\,\widetilde{x}+\big(f(t)-f(a)\big)-X(t)\left(\int_a^t \mathrm{d}[X^{-1}(s)]\,\big(f(s)-f(a)\big)\right),$$

where $X^{-1}(s)=U(a,s)$ for $s\in[a,b]$. Due to (7.8.14), we have

$$X^{-1}(s)=Y(s)-\Delta^+X^{-1}(s) \quad \text{for} \;\; s\in[a,b).$$

By Lemma 6.3.19, the relation

$$\begin{aligned}\int_a^t &\mathrm{d}[X^{-1}(s)]\,\big(f(s)-f(a)\big)\\ &=\int_a^t \mathrm{d}[Y(s)]\,\big(f(s)-f(a)\big)-\Delta^+X^{-1}(t)\,\big(f(t)-f(a)\big)\end{aligned}$$

holds for every $t\in[a,b]$. Since f is continuous from the left on $(a,b]$ and Y is continuous from the right on $[a,b)$, using Integration by parts Theorem 6.4.2, we get

$$\begin{aligned}x(t)&=X(t)\,\widetilde{x}+\big(f(t)-f(a)\big)-X(t)\left(\int_a^t \mathrm{d}[X^{-1}(s)]\,\big(f(s)-f(a)\big)\right)\\ &=X(t)\,\widetilde{x}-\big(f(t)-f(a)\big)+X(t)\int_a^t Y\,\mathrm{d}\,f\\ &\qquad +X(t)\,\Delta^+X^{-1}(t)\,\big(f(t)-f(a)\big)-X(t)\,Y(t)\,\big(f(t)-f(a)\big)\end{aligned}$$

for every $t\in[a,b]$. Finally, as

$$\begin{aligned}X(t)\,&\Delta^+X^{-1}(t)\,\big(f(t)-f(a)\big)-X(t)\,Y(t)\,\big(f(t)-f(a)\big)\\ &=X(t)\,\big(X^{-1}(t+)-X^{-1}(t)\,(f(t)-f(a))\big)\\ &\qquad\qquad -X(t)\,X^{-1}(t+)\,\big(f(t)-f(a)\big)\\ &=-\big(f(t)-f(a)\big)\end{aligned}$$

for each $t\in[a,b]$, we get (7.8.15). □

7.8.6 Remark. Thanks to Theorem 7.8.4, or rather to its Corollary 7.8.5, it is already possible to successfully investigate for example the boundary value problems in which one looks for the function satisfying equation (7.7.7) on the interval $[a,b]$ and, simultaneously, some other conditions, for instance two-point conditions

$$M\,x(a)+N\,x(b)=0$$

where $M,N\in\mathscr{L}(\mathbb{R}^n)$.

7.9 Generalized elementary functions

We now show that the theory of generalized linear differential equations can be used to extend the definitions of the exponential, hyperbolic and trigonometric functions.

One possible way of introducing the classical exponential function is to define it as the unique solution of the initial-value problem

$$z'(t)=z(t), \quad z(0)=1.$$

More generally, for every continuous function p defined on the real line, the initial-value problem $z'(t)=p(t)\,z(t)$, $z(t_0)=1$, which can be written in the equivalent integral form

$$z(t)=1+\int_{t_0}^{t} p(s)z(s)\,\mathrm{d}s, \tag{7.9.1}$$

has the unique solution $z(t)=e^{\int_{t_0}^{t} p(s)\,\mathrm{d}s}$. Using Substitution Theorem 6.6.1, we can rewrite equation (7.9.1) as the generalized linear differential equation

$$z(t)=1+\int_{t_0}^{t} z(s)\,\mathrm{d}P(s) \tag{7.9.2}$$

with $P(s)=\int_{s_0}^{s} p$. In this section, we study equation (7.9.2) for an arbitrary function P with bounded variation (not necessarily differentiable or continuous). The solution of this equation will be called the generalized exponential function and denoted by $e_{\mathrm{d}P}$. If P is a real function, then $e_{\mathrm{d}P}$ is simply a special case of the Cauchy matrix U introduced in Definition 7.7.10 with $n=1$. In this scalar case, we will be able to obtain a much more detailed information about solutions of equation (7.9.2) than in the n-dimensional case studied in Section 7.7. Moreover, we will focus on the more general case when P is a complex-valued function. To this end, we need an existence and uniqueness theorem for generalized linear differential equations with complex coefficients.

7.9.1 Theorem. *Consider a function* $P:[a,b]\to\mathbb{C}$, *which has bounded variation on* $[a,b]$. *Let* $t_0\in[a,b]$ *and assume that* $1+\Delta^+P(t)\neq 0$ *for every* $t\in[a,t_0)$, *and* $1-\Delta^-P(t)\neq 0$ *for every* $t\in(t_0,b]$. *Then, for every* $\widetilde{z}\in\mathbb{C}$, *there exists a unique function* $z:[a,b]\to\mathbb{C}$ *such that*

$$z(t)=\widetilde{z}+\int_{t_0}^{t} z(s)\,\mathrm{d}P(s),\qquad t\in[a,b]. \tag{7.9.3}$$

The function z *has bounded variation on* $[a,b]$. *If* P *and* $\widetilde{z}$ *are real, then* z *is real as well.*

Proof. We decompose all complex quantities into real and imaginary parts as follows: $P=P_1+\mathrm{i}P_2$, $z=z_1+\mathrm{i}z_2$, and $\widetilde{z}=\widetilde{z}_1+\mathrm{i}\,\widetilde{z}_2$. Now, we see that equation (7.9.3) is equivalent to the following system of two equations with real coefficients:

$$\begin{aligned}
z_1(t)&=\widetilde{z}_1+\int_{t_0}^{t} z_1(s)\,\mathrm{d}P_1(s)-\int_{t_0}^{t} z_2(s)\,\mathrm{d}P_2(s)\\
z_2(t)&=\widetilde{z}_2+\int_{t_0}^{t} z_1(s)\,\mathrm{d}P_2(s)+\int_{t_0}^{t} z_2(s)\,\mathrm{d}P_1(s)
\end{aligned}$$

The system can be also written in the vector form

$$u(t)=\widetilde{u}+\int_{t_0}^{t}\mathrm{d}[A(s)]\,u(s),\qquad t\in[a,b] \tag{7.9.4}$$

with

$$\widetilde{u}=\begin{pmatrix}\widetilde{z}_1\\ \widetilde{z}_2\end{pmatrix},\quad u(t)=\begin{pmatrix}z_1(t)\\ z_2(t)\end{pmatrix},\quad A(t)=\begin{pmatrix}P_1(t) & -P_2(t)\\ P_2(t) & P_1(t)\end{pmatrix},\quad t\in[a,b].$$

Since P has bounded variation on $[a,b]$, it is clear that A has the same property. The condition $1+\Delta^+P(t)\neq 0$ implies

$$1+\Delta^+P_1(t)\neq 0 \ \text{ or } \ \Delta^+P_2(t)\neq 0\quad \text{for } t\in[a,t_0),$$

and similarly $1-\Delta^-P(t)\neq 0$ implies

$$1-\Delta^-P_1(t)\neq 0 \ \text{ or } \ \Delta^-P_2(t)\neq 0\quad \text{for } t\in(t_0,b].$$

In view of this, we have

$$\begin{aligned}
\det(I+\Delta^+A(t))&=(1+\Delta^+P_1(t))^2+(\Delta^+P_2(t))^2\neq 0\quad \text{for } t\in[a,t_0),\\
\det(I-\Delta^-A(t))&=(1-\Delta^-P_1(t))^2+(\Delta^-P_2(t))^2\neq 0\quad \text{for } t\in(t_0,b].
\end{aligned}$$

Hence, existence and uniqueness of solution to equation (7.9.3) follows from Theorem 7.4.8. By Corollary 6.5.4, the solution has bounded variation on $[a,b]$.

If P and $\widetilde{z}$ are real, the equation for z_2 simplifies to $z_2(t) = \int_{t_0}^t z_2 \,\mathrm{d}P_1$, whose solution is identically zero and therefore z is real. □

The previous existence and uniqueness theorem guarantees that the definition of the generalized exponential function is meaningful.

7.9.2 Definition. Consider a function $P\colon [a,b] \to \mathbb{C}$, which has bounded variation on $[a,b]$. Let $t_0 \in [a,b]$ and assume that $1 + \Delta^+ P(t) \neq 0$ for every $t \in [a,t_0)$, and $1 - \Delta^- P(t) \neq 0$ for every $t \in (t_0,b]$. Then we define the generalized exponential function $t \mapsto e_{\mathrm{d}P}(t,t_0)$, $t \in [a,b]$, as the unique solution $z\colon [a,b] \to \mathbb{C}$ of the generalized linear differential equation

$$z(t) = 1 + \int_{t_0}^t z(s)\,\mathrm{d}P(s).$$

To explore the properties of the generalized exponential function, we need the following auxiliary lemma, which can generalizes the formula

$$\int_a^b h^k(t)h'(t)\,\mathrm{d}t = \frac{h^{k+1}(b) - h^{k+1}(a)}{k+1}$$

to the case when h is continuous but not necessarily differentiable. (Recall that Lemma 7.5.2 and Exercise 7.5.3 deal with the situation when h is discontinuous.)

7.9.3 Lemma. *If $h\colon [a,b] \to \mathbb{C}$ is a continuous function with bounded variation, then*

$$\int_a^b h^k \,\mathrm{d}h = \frac{h^{k+1}(b) - h^{k+1}(a)}{k+1} \quad \textit{for every } k \in \mathbb{N} \cup \{0\}.$$

Proof. Since h has bounded variation and h^k is continuous for each $k \in \mathbb{N} \cup \{0\}$, the integral $\int_a^b h^k \,\mathrm{d}h$ exists as a Riemann-Stieltjes integral (see Theorem 5.6.3).

The statement of the lemma obviously holds for $k = 0$. Let us assume that it holds for a certain $k \in \mathbb{N} \cup \{0\}$, and show its validity for $k+1$. Using first the substitution theorem (Theorem 5.4.3) and then the integration by parts formula (Theorem 5.5.1), we get

$$\begin{aligned}
\int_a^b h^{k+1}(t)\,\mathrm{d}h(t) &= \int_a^b h(t)\cdot h^k(t)\,\mathrm{d}h(t) = \int_a^b h(t)\,\mathrm{d}\left(\int_a^t h^k \,\mathrm{d}h\right)\\
&= \int_a^b h(t)\,\mathrm{d}\left(\frac{h^{k+1}(t) - h^{k+1}(a)}{k+1}\right) = \frac{1}{k+1}\int_a^b h(t)\,\mathrm{d}h^{k+1}(t)\\
&= \frac{1}{k+1}\left(h^{k+2}(b) - h^{k+2}(a) - \int_a^b h^{k+1}(t)\,\mathrm{d}h(t)\right).
\end{aligned}$$

By equating the first and the last term in the previous chain of equalities and solving for $\int_a^b h^{k+1}\,\mathrm{d}h$, we obtain

$$\int_a^b h^{k+1}\,\mathrm{d}h = \frac{h^{k+2}(b) - h^{k+2}(a)}{k+2},$$

which completes the proof by induction. □

The next theorem summarizes the basic properties of the generalized exponential functions. Some of them were already established in Section 7.7 in the context of Cauchy matrices, but we repeat them here for reader's convenience. In each of the eight statements, we assume that the function P is such that all exponentials appearing in the given identity are defined. For example, in statement (v), it is necessary to assume that

$$1 + \Delta^+ P(t) \neq 0 \quad \text{for every } t \in [a, \max\{s, r\}),$$
$$1 - \Delta^- P(t) \neq 0 \quad \text{for every } t \in (\min\{s, r\}, b].$$

7.9.4 Theorem. *Let $P : [a, b] \to \mathbb{C}$ be a function with bounded variation. The generalized exponential function has the following properties:*

(i) *If P is constant, then $e_{\mathrm{d}P}(t, t_0) = 1$ for every $t \in [a, b]$.*

(ii) $e_{\mathrm{d}P}(t, t) = 1$ *for every $t \in [a, b]$.*

(iii) *The function $t \mapsto e_{\mathrm{d}P}(t, t_0)$ is regulated on $[a, b]$ and satisfies*

$$\begin{aligned}
\Delta^+ e_{\mathrm{d}P}(t, t_0) &= \Delta^+ P(t)\, e_{\mathrm{d}P}(t, t_0) && \text{for } t \in [a, b),\\
\Delta^- e_{\mathrm{d}P}(t, t_0) &= \Delta^- P(t)\, e_{\mathrm{d}P}(t, t_0) && \text{for } t \in (a, b],\\
e_{\mathrm{d}P}(t+, t_0) &= (1 + \Delta^+ P(t))\, e_{\mathrm{d}P}(t, t_0) && \text{for } t \in [a, b),\\
e_{\mathrm{d}P}(t-, t_0) &= (1 - \Delta^- P(t))\, e_{\mathrm{d}P}(t, t_0) && \text{for } t \in (a, b].
\end{aligned}$$

(iv) *The function $t \mapsto e_{\mathrm{d}P}(t, t_0)$ has bounded variation on $[a, b]$.*

(v) $e_{\mathrm{d}P}(t, s)\, e_{\mathrm{d}P}(s, r) = e_{\mathrm{d}P}(t, r)$ *for every $t, s, r \in [a, b]$.*

(vi) $e_{\mathrm{d}P}(t, s) = e_{\mathrm{d}P}(s, t)^{-1}$ *for every $t,\ s \in [a, b]$.*

(vii) $\overline{e_{\mathrm{d}P}(t, t_0)} = e_{\mathrm{d}\overline{P}}(t, t_0)$ *for every $t \in [a, b]$, where $\overline{z}$ denotes the complex conjugate of $z \in \mathbb{C}$.*

(viii) *If P is continuous, then $e_{\mathrm{d}P}(t, t_0) = e^{P(t) - P(t_0)}$ for every $t \in [a, b]$.*

Proof. The first two statements are obvious. Statement (iii) is a consequence of Corollary 6.5.4 (extended to complex-valued case), and statement (iv) follows from Theorem 7.9.1.

To prove statement (v), note that, given arbitrary $r, s \in [a,b]$, we have

$$\begin{aligned} e_{\mathrm{d}P}(t,r) &= 1 + \int_r^t e_{\mathrm{d}P}(\tau,r)\,\mathrm{d}P(\tau) \\ &= 1 + \int_r^s e_{\mathrm{d}P}(\tau,r)\,\mathrm{d}P(\tau) + \int_s^t e_{\mathrm{d}P}(\tau,r)\,\mathrm{d}P(\tau) \\ &= e_{\mathrm{d}P}(s,r) + \int_s^t e_{\mathrm{d}P}(\tau,r)\,\mathrm{d}P(\tau) \end{aligned}$$

for every $t \in [a,b]$. Hence, the function $y(t) = e_{\mathrm{d}P}(t,r)$ is a solution on $[a,b]$ of the generalized linear differential equation

$$x(t) = \widetilde{x} + \int_s^t x(s)\,\mathrm{d}P(s), \quad \text{where } \widetilde{x} = e_{\mathrm{d}P}(s,r).$$

On the other hand, it is not hard to see that $z(t) = e_{\mathrm{d}P}(t,s)\,\widetilde{x}$ for $t \in [a,b]$ is also a solution of the same equation. By the uniqueness of solutions, we have $y(t) = z(t)$ for all $t \in [a,b]$, which proves statement (v).

Statement (vi) is a direct consequence of statement (v). Indeed, for $t, s \in [a,b]$, we obtain

$$e_{\mathrm{d}P}(t,s)\,e_{\mathrm{d}P}(s,t) = e_{\mathrm{d}P}(t,t) = 1.$$

By the definition of the exponential function, we have

$$e_{\mathrm{d}P}(t,t_0) = 1 + \int_{t_0}^t e_{\mathrm{d}P}(s,t_0)\,\mathrm{d}P(s).$$

Taking the complex conjugate of both sides, we get

$$\overline{e_{\mathrm{d}P}(t,t_0)} = 1 + \int_{t_0}^t \overline{e_{\mathrm{d}P}(s,t_0)}\,\mathrm{d}\overline{P}(s),$$

which proves statement (vii).

To prove statement (viii), assume that P is a continuous function with bounded variation. Let $\tilde{P}(t) = P(t) - P(t_0)$ and $z(t) = e^{\tilde{P}(t)}$ for all $t \in [a,b]$. Using the uniform convergence theorem (Theorem 5.6.1) and Lemma 7.9.3, we get

$$1 + \int_{t_0}^t z(s)\,\mathrm{d}P(s) = 1 + \int_{t_0}^t z(s)\,\mathrm{d}\tilde{P}(s) = 1 + \int_{t_0}^t \left(\sum_{k=0}^{\infty} \frac{\tilde{P}(s)^k}{k!} \right) \mathrm{d}\tilde{P}(s)$$

$$
\begin{aligned}
&= 1 + \sum_{k=0}^{\infty} \left(\frac{1}{k!} \int_{t_0}^{t} \tilde{P}(s)^k \, \mathrm{d}\tilde{P}(s) \right) = 1 + \sum_{k=0}^{\infty} \frac{1}{k!} \frac{\tilde{P}(t)^{k+1} - \tilde{P}(t_0)^{k+1}}{k+1} \\
&= 1 + \sum_{k=0}^{\infty} \frac{\tilde{P}(t)^{k+1}}{(k+1)!} = z(t)
\end{aligned}
$$

for all $t \in [a,b]$. It follows that $e_{\mathrm{d}P}(t,t_0) = z(t) = e^{P(t)-P(t_0)}$. □

As an immediate consequence of parts (iii) and (vi) of Theorem 7.9.4, we see that the generalized exponential function is regulated with respect to both arguments; this fact will be used in the proof of the next result.

We now derive an explicit formula for the value of $e_{\mathrm{d}P}(t,t_0)$. It is sufficient to focus on the case when $t > t_0$; the formula for the case $t < t_0$ then follows easily from the identity $e_{\mathrm{d}P}(t,t_0) = e_{\mathrm{d}P}(t_0,t)^{-1}$.

The next theorem involves infinite products of (possibly complex) numbers. Recall that if $\sum_{k=1}^{\infty} a_k$ is an absolutely convergent series of complex numbers, then the infinite product $\prod_{k=1}^{\infty}(1+a_k)$ is also absolutely convergent; in particular, the product converges to the same value after an arbitrary rearrangement of the sequence $\{a_k\}$. Now, suppose that $P: [a,b] \to \mathbb{R}$ is a function of bounded variation with infinitely many discontinuity points in (a,b), which are arranged in a sequence $\{s_k\}$. By Theorem 2.3.6, the sum $\sum_{k=1}^{\infty}(|\Delta^+ P(s_k)| + |\Delta^- P(s_k)|)$ is finite. Hence, the products $\prod_{k=1}^{\infty}(1+\Delta^+ P(s_k))$ and $\prod_{k=1}^{\infty}(1-\Delta^- P(s_k))$ are absolutely convergent. Since the order of the factors is unimportant, we introduce a similar convention as in Remark 2.3.7, and write

$$
\begin{aligned}
\prod_{x \in (a,b)} (1+\Delta^+ P(x)) &= \prod_{k=1}^{\infty} (1+\Delta^+ P(s_k)), \\
\prod_{x \in (a,b)} (1-\Delta^- P(x)) &= \prod_{k=1}^{\infty} (1-\Delta^- P(s_k)).
\end{aligned}
$$

The symbols

$$
\prod_{x \in [a,b)}, \quad \prod_{x \in (a,b]}, \quad \text{and} \quad \prod_{x \in [a,b]}
$$

should be understood in an analogous way.

7.9.5 Theorem. *If* $t > t_0$, *then*

$$
e_{\mathrm{d}P}(t,t_0) = \frac{e^{P(t-)-P(t_0+)}}{e^{\sum_{\tau \in (t_0,t)} \Delta P(\tau)}} \frac{\prod_{\tau \in [t_0,t)} (1+\Delta^+ P(\tau))}{\prod_{\tau \in (t_0,t]} (1-\Delta^- P(\tau))}. \tag{7.9.5}
$$

Proof. We begin by verifying the formula in the case when P has only finitely many discontinuities in $[t_0,t]$. Let $\boldsymbol{\alpha}$ be a division of $[t_0,t]$ containing all discontinuity points of P. Then P is continuous on each interval (α_{k-1},α_k), where $k\in\{1,\dots,\nu(\boldsymbol{\alpha})\}$. Parts (ii), (iii), (v), (vi), (viii) of Theorem 7.9.4 imply

$$\begin{aligned}
e_{\mathrm{d}P}(t,t_0) &= \prod_{k=1}^{\nu(\boldsymbol{\alpha})} e_{\mathrm{d}P}(\alpha_k,\alpha_{k-1})\\
&= \prod_{k=1}^{\nu(\boldsymbol{\alpha})} e_{\mathrm{d}P}(\alpha_k,\alpha_k-)e_{\mathrm{d}P}(\alpha_k-,\alpha_{k-1}+)e_{\mathrm{d}P}(\alpha_{k-1}+,\alpha_{k-1})\\
&= \prod_{k=1}^{\nu(\boldsymbol{\alpha})} (1-\Delta^- P(\alpha_k))^{-1} e^{P(\alpha_k-)-P(\alpha_{k-1}+)}(1+\Delta^+P(\alpha_{k-1}))\\
&= \frac{e^{P(t-)-P(t_0+)}}{e^{\sum_{k=1}^{\nu(\boldsymbol{\alpha})-1}\Delta P(\alpha_k)}}\,\frac{\prod_{k=1}^{\nu(\boldsymbol{\alpha})}(1+\Delta^+P(\alpha_{k-1}))}{\prod_{k=1}^{\nu(\boldsymbol{\alpha})}(1-\Delta^-P(\alpha_k))},
\end{aligned}$$

which agrees with (7.9.5).

Now, assume that P has infinitely many discontinuities in $[t_0,t]$, and let $D=\{s_k\}$ be a sequence of all discontinuity points contained in (t_0,t). According to Theorem 2.6.1 and its proof (which remains valid for complex functions), we have the Jordan decomposition $P=P^{\mathrm{C}}+P^{\mathrm{B}}$, where

$$\begin{aligned}
P^{\mathrm{B}}(\tau) &= P(t_0)+\Delta^+P(t_0)\,\chi_{(t_0,t]}(\tau)+\Delta^-P(t)\,\chi_{[t]}(\tau)\\
&\quad+\sum_{k=1}^{\infty}\left(\Delta^+P(s_k)\,\chi_{(s_k,t]}(\tau)+\Delta^-P(s_k)\,\chi_{[s_k,t]}(\tau)\right), \qquad \tau\in[t_0,t],
\end{aligned}$$

is the jump part of P, and $P^{\mathrm{C}}=P-P^{\mathrm{B}}$ is the continuous part of P. For each $n\in\mathbb{N}$, let $P_n=P^{\mathrm{C}}+P_n^{\mathrm{B}}$, where

$$\begin{aligned}
P_n^{\mathrm{B}}(\tau) &= P(t_0)+\Delta^+P(t_0)\,\chi_{(t_0,t]}(\tau)+\Delta^-P(t)\,\chi_{[t]}(\tau)\\
&\quad+\sum_{k=1}^{n}\left(\Delta^+P(s_k)\,\chi_{(s_k,t]}(\tau)+\Delta^-P(s_k)\,\chi_{[s_k,t]}(\tau)\right), \qquad \tau\in[t_0,t].
\end{aligned}$$

For each $n\in\mathbb{N}$, we have $\Delta^+P_n(t_0)=\Delta^+P(t_0)$ and $\Delta^-P_n(t)=\Delta^-P(t)$; moreover, $\Delta^+P_n(s_k)=\Delta^+P(s_k)$ and $\Delta^-P_n(s_k)=\Delta^-P(s_k)$ for all $k\le n$. Since the discontinuities of P_n are contained in the finite set $\{t_0,s_1,\dots,s_n,t\}$, we obtain

$$\begin{aligned}
e_{\mathrm{d}P_n}(t,t_0) &= \frac{e^{P_n(t-)-P_n(t_0+)}}{e^{\sum_{\tau\in(t_0,t)}\Delta P_n(\tau)}}\,\frac{1+\Delta^+P_n(t_0)}{1-\Delta^-P_n(t)}\,\frac{\prod_{\tau\in(t_0,t)}(1+\Delta^+P_n(\tau))}{\prod_{\tau\in(t_0,t)}(1-\Delta^-P_n(\tau))}\\
&= \frac{e^{P_n(t-)-P_n(t_0+)}}{e^{\sum_{k=1}^{n}\Delta P(s_k)}}\,\frac{1+\Delta^+P(t_0)}{1-\Delta^-P(t)}\,\frac{\prod_{k=1}^{n}(1+\Delta^+P(s_k))}{\prod_{k=1}^{n}(1-\Delta^-P(s_k))}.
\end{aligned}$$

According to the proof of Lemma 2.6.5, the sequence $\{P_n\}$ is convergent to P in the BV norm. It follows that the variations of P_n, $n \in \mathbb{N}$, are uniformly bounded, and $\{P_n\}$ is uniformly convergent to P. Hence, Theorem 7.6.6 (continuous dependence of solutions to generalized linear differential equations with respect to the right-hand side) implies

$$\begin{aligned} e_{\mathrm{d}P}(t,t_0) &= \lim_{n\to\infty} e_{\mathrm{d}P_n}(t,t_0) \\ &= \frac{e^{P(t-)-P(t_0+)}}{e^{\sum_{k=1}^{\infty} \Delta P(s_k)}} \frac{1+\Delta^+ P(t_0)}{1-\Delta^- P(t)} \frac{\prod_{k=1}^{\infty}(1+\Delta^+ P(s_k))}{\prod_{k=1}^{\infty}(1-\Delta^- P(s_k))}. \end{aligned}$$

The series $\sum_{k=1}^{\infty} \Delta P(s_k)$ and the two infinite products on the right-hand side are absolutely convergent, which proves that formula (7.9.5) holds. □

To obtain some additional properties of the generalized exponential function, we need the following auxiliary result, which is a fairly straightforward consequence of Lemma 6.3.18.

7.9.6 Lemma. *If* $f:[a,b]\to\mathbb{R}$ *has bounded variation and* $g:[a,b]\to\mathbb{R}$ *is regulated, then*

$$\int_a^b \Delta^+ f \,\mathrm{d}g = \sum_{x\in(a,b)} \Delta^+ f(x)\,\Delta g(x) + \Delta^+ f(a)\,\Delta^+ g(a), \tag{7.9.6}$$

$$\int_a^b \Delta^- f \,\mathrm{d}g = \sum_{x\in(a,b)} \Delta^- f(x)\,\Delta g(x) + \Delta^- f(b)\,\Delta^- g(b), \tag{7.9.7}$$

with the convention that $\Delta^- f(a) = 0,\ \Delta^+ f(b) = 0.$

Proof. Since f has bounded variation, it has only finitely or countably many discontinuities. Moreover, by Corollary 2.3.8, the sums

$$\sum_{x\in[a,b]} |\Delta^+ f(x)| \quad \text{and} \quad \sum_{x\in[a,b]} |\Delta^- f(x)|$$

are finite. This means that $\Delta^+ f$ and $\Delta^- f$ are step functions with bounded variation. The formulas (7.9.6) and (7.9.7) are now an immediate consequence of Lemma 6.3.18 with $c = 0$ and $h = \Delta^+ f$ or $h = \Delta^- f$, respectively. □

In the next theorem, we prove that the product of two exponentials $e_{\mathrm{d}P}$ and $e_{\mathrm{d}Q}$ equals the exponential of a certain function denoted by $P \oplus Q$. We have $P\oplus Q = P+Q$ if P, Q are continuous, but the general definition of $P\oplus Q$ is more complicated and takes into account the jumps of P and Q. We make the following agreement: If $c > d$, then a sum of the form $\sum_{s\in[c,d)} h(s)$ should be interpreted as $-\sum_{s\in(d,c]} h(s)$, and the sum $\sum_{s\in(c,d]} h(s)$ should be understood as $\sum_{s\in[d,c)} h(s)$.

7.9.7 Theorem. *Assume that* $P, Q : [a,b] \to \mathbb{C}$ *have bounded variation and*

$$(1+\Delta^+P(t))\,(1+\Delta^+Q(t)) \neq 0 \quad \textit{for every} \;\; t \in [a,t_0)$$
$$(1-\Delta^-P(t))\,(1-\Delta^-Q(t)) \neq 0 \quad \textit{for every} \;\; t \in (t_0,b].$$

Then

$$e_{\mathrm{d}P}(t,t_0)\, e_{\mathrm{d}Q}(t,t_0) = e_{\mathrm{d}(P\oplus Q)}(t,t_0), \quad t \in [a,b],$$

where

$$(P\oplus Q)(t) = P(t) + Q(t) + \int_{t_0}^{t} \Delta^+Q(s)\,\mathrm{d}P(s) - \int_{t_0}^{t} \Delta^-P(s)\,\mathrm{d}Q(s), \tag{7.9.8}$$

with the convention that $\Delta^+Q(t) = 0, \;\Delta^-P(t_0) = 0.$ *Equivalently, we have*

$$(P\oplus Q)(t) = P(t) + Q(t) + \sum_{s\in[t_0,t)} \Delta^+Q(s)\Delta^+P(s) - \sum_{s\in(t_0,t]} \Delta^-Q(s)\Delta^-P(s).$$

Proof. By Lemma 7.9.6, we have

$$\begin{aligned}
&\int_{t_0}^{t} \Delta^+Q(s)\,\mathrm{d}P(s) - \int_{t_0}^{t} \Delta^-P(s)\,\mathrm{d}Q(s)\\
&\quad = \sum_{s\in(t_0,t)} \big(\Delta^+Q(s)\,\Delta P(s) - \Delta^-P(s)\,\Delta Q(s)\big)\\
&\qquad + \Delta^+Q(t_0)\,\Delta^+P(t_0) - \Delta^-P(t)\,\Delta^-Q(t)\\
&\quad = \sum_{s\in(t_0,t)} \big(\Delta^+Q(s)\,(\Delta^+P(s)+\Delta^-P(s)) - \Delta^-P(s)\,(\Delta^+Q(s)+\Delta^-Q(s))\big)\\
&\qquad + \Delta^+Q(t_0)\,\Delta^+P(t_0) - \Delta^-P(t)\,\Delta^-Q(t)\\
&\quad = \sum_{s\in[t_0,t)} \Delta^+Q(s)\,\Delta^+P(s) - \sum_{s\in(t_0,t]} \Delta^-Q(s)\,\Delta^-P(s).
\end{aligned}$$

Hence, the two formulas for $P\oplus Q$ are equivalent. For $t\in[a,b]$, let

$$R(t) = \sum_{s\in(t_0,t]} \Delta^-Q(s)\Delta^-P(s) \quad \text{and} \quad T(t) = \sum_{s\in[t_0,t)} \Delta^+Q(s)\Delta^+P(s).$$

Using the definition of variation together with Corollary 2.3.8, it is not difficult to check that R, T have bounded variation. Furthermore, the definitions of R, T imply

$$\Delta^-R(t) = \Delta^-Q(t)\,\Delta^-P(t), \quad \Delta^+R(t) = 0,$$
$$\Delta^-T(t) = 0, \quad \Delta^+T(t) = \Delta^+Q(t)\,\Delta^+P(t).$$

In view of this, it is clear that $P \oplus Q$ has bounded variation on $[a,b]$. Moreover,

$$\begin{aligned}1-\Delta^-(P\oplus Q)(t) &= 1-\Delta^-P(t)-\Delta^-Q(t)-\Delta^-T(t)+\Delta^-R(t)\\ &= 1-\Delta^-P(t)-\Delta^-Q(t)+\Delta^-Q(t)\,\Delta^-P(t)\\ &= \big(1-\Delta^-P(t)\big)\big(1-\Delta^-Q(t)\big)\neq 0 \quad \text{for } t\in(t_0,b].\end{aligned}$$

Proceeding in a similar way, one can show that

$$1+\Delta^+(P\oplus Q)(t) = \big(1+\Delta^+P(t)\big)\big(1+\Delta^+Q(t)\big)\neq 0 \quad \text{for } t\in[a,t_0).$$

Therefore, the exponential function $t\mapsto e_{\mathrm{d}(P\oplus Q)}(t,t_0)$ is defined.

For $t\in[a,b]$, integration by parts gives

$$\begin{aligned}&e_{\mathrm{d}P}(t,t_0)\,e_{\mathrm{d}Q}(t,t_0) = e_{\mathrm{d}P}(t_0,t_0)\,e_{\mathrm{d}Q}(t_0,t_0)\\ &\quad+\int_{t_0}^t e_{\mathrm{d}P}(s,t_0)\,\mathrm{d}[e_{\mathrm{d}Q}(s,t_0)]+\int_{t_0}^t e_{\mathrm{d}Q}(s,t_0)\,\mathrm{d}[e_{\mathrm{d}P}(s,t_0)]\\ &\quad+\sum_{s\in[t_0,t)}\Delta^+e_{\mathrm{d}P}(s,t_0)\,\Delta^+e_{\mathrm{d}Q}(s,t_0)-\sum_{s\in(t_0,t]}\Delta^-e_{\mathrm{d}P}(s,t_0)\,\Delta^-e_{\mathrm{d}Q}(s,t_0).\end{aligned}$$

Let us examine the terms on the right-hand side. Obviously,

$$e_{\mathrm{d}P}(t_0,t_0)e_{\mathrm{d}Q}(t_0,t_0)=1.$$

Using the substitution theorem, we have

$$\begin{aligned}\int_{t_0}^t e_{\mathrm{d}P}(s,t_0)\,\mathrm{d}[e_{\mathrm{d}Q}(s,t_0)] &= \int_{t_0}^t e_{\mathrm{d}P}(s,t_0)\,\mathrm{d}\left[1+\int_{t_0}^s e_{\mathrm{d}Q}(u,t_0)\,\mathrm{d}Q(u)\right]\\ &= \int_{t_0}^t e_{\mathrm{d}P}(s,t_0)e_{\mathrm{d}Q}(s,t_0)\,\mathrm{d}Q(s),\end{aligned}$$

and

$$\begin{aligned}\int_{t_0}^t e_{\mathrm{d}Q}(s,t_0)\,\mathrm{d}[e_{\mathrm{d}P}(s,t_0)] &= \int_{t_0}^t e_{\mathrm{d}Q}(s,t_0)\,\mathrm{d}\left[1+\int_{t_0}^s e_{\mathrm{d}P}(u,t_0)\,\mathrm{d}P(u)\right]\\ &= \int_{t_0}^t e_{\mathrm{d}Q}(s,t_0)\,e_{\mathrm{d}P}(s,t_0)\,\mathrm{d}P(s).\end{aligned}$$

Finally, by performing an algebraic manipulation and using subsequently Lemma 7.9.6 part (3) of Theorem 7.9.4 and the substitution theorem, we get

$$\begin{aligned}&\sum_{s\in[t_0,t)}\Delta^+e_{\mathrm{d}P}(s,t_0)\,\Delta^+e_{\mathrm{d}Q}(s,t_0)-\sum_{s\in(t_0,t]}\Delta^-e_{\mathrm{d}P}(s,t_0)\,\Delta^-e_{\mathrm{d}Q}(s,t_0)\\ &=\Delta^+e_{\mathrm{d}P}(t_0,t_0)\,\Delta^+e_{\mathrm{d}Q}(t_0,t_0)+\sum_{s\in(t_0,t)}[\Delta^+e_{\mathrm{d}P}(s,t_0)+\Delta^-e_{\mathrm{d}P}(s,t_0)]\,\Delta^+e_{\mathrm{d}Q}(s,t_0)\\ &\quad-\sum_{s\in(t_0,t)}\Delta^-e_{\mathrm{d}P}(s,t_0)\,[\Delta^-e_{\mathrm{d}Q}(s,t_0)+\Delta^+e_{\mathrm{d}Q}(s,t_0)]-\Delta^-e_{\mathrm{d}P}(t,t_0)\,\Delta^-e_{\mathrm{d}Q}(t,t_0))\end{aligned}$$

$$
\begin{aligned}
&= \Delta^+ e_{\mathrm{d}P}(t_0,t_0)\,\Delta^+ e_{\mathrm{d}Q}(t_0,t_0) + \sum_{s\in(t_0,t)} \Delta e_{\mathrm{d}P}(s,t_0)\,\Delta^+ e_{\mathrm{d}Q}(s,t_0)\\
&\quad - \sum_{s\in(t_0,t)} \Delta^- e_{\mathrm{d}P}(s,t_0)\,\Delta e_{\mathrm{d}Q}(s,t_0) - \Delta^- e_{\mathrm{d}P}(t,t_0)\,\Delta^- e_{\mathrm{d}Q}(t,t_0))\\
&= \int_{t_0}^{t} \Delta^+ e_{\mathrm{d}Q}(s,t_0)\,\mathrm{d}[e_{\mathrm{d}P}(s,t_0)] - \int_{t_0}^{t} \Delta^- e_{\mathrm{d}P}(s,t_0)\,\mathrm{d}[e_{\mathrm{d}Q}(s,t_0)]\\
&= \int_{t_0}^{t} \Delta^+ Q(s) e_{\mathrm{d}Q}(s,t_0)\, e_{\mathrm{d}P}(s,t_0)\,\mathrm{d}P(s) - \int_{t_0}^{t} \Delta^- P(s) e_{\mathrm{d}P}(s,t_0)\, e_{\mathrm{d}Q}(s,t_0)\,\mathrm{d}Q(s)\\
&= \int_{t_0}^{t} e_{\mathrm{d}Q}(s,t_0)\, e_{\mathrm{d}P}(s,t_0)\,\mathrm{d}\left[\int_{t_0}^{s} \Delta^+ Q(u)\,\mathrm{d}P(u) - \int_{t_0}^{s} \Delta^- P(u)\,\mathrm{d}Q(u)\right].
\end{aligned}
$$

By combining the previous results, we obtain

$$
e_{\mathrm{d}P}(t,t_0)\, e_{\mathrm{d}Q}(t,t_0) = 1 + \int_{t_0}^{t} e_{\mathrm{d}P}(s,t_0)\, e_{\mathrm{d}Q}(s,t_0)\,\mathrm{d}\left[(P\oplus Q)(s)\right],
$$

with $P\oplus Q$ given by (7.9.8). □

7.9.8 Exercise. Obviously, the binary operation $\oplus$ introduced in the previous theorem is commutative. Verify that

$$
\begin{aligned}
((P\oplus Q)\oplus R)(t) &= (P\oplus(Q\oplus R))(t) = P(t)+Q(t)+R(t)\\
&\quad + \sum_{s\in[t_0,t)} \big(\Delta^+Q(s)\Delta^+P(s) + \Delta^+R(s)\Delta^+P(s) + \Delta^+Q(s)\Delta^+R(s)\big)\\
&\quad + \sum_{s\in[t_0,t)} \Delta^+P(s)\,\Delta^+Q(s)\,\Delta^+R(s)\\
&\quad - \sum_{s\in(t_0,t]} \big(\Delta^-Q(s)\,\Delta^-P(s) + \Delta^-R(s)\,\Delta^-P(s) + \Delta^-Q(s)\,\Delta^-R(s)\big)\\
&\quad + \sum_{s\in(t_0,t]} \Delta^-P(s)\,\Delta^-Q(s)\,\Delta^-R(s),
\end{aligned}
$$

i.e., the operation $\oplus$ is also associative.

The next result shows that the reciprocal value of an exponential function is again an exponential function.

7.9.9 Theorem. *Assume that* $P:[a,b]\to\mathbb{C}$ *has bounded variation and*

$$
\begin{aligned}
&1+\Delta^+P(t)\neq 0 \quad \textit{for every } t\in[a,b),\\
&1-\Delta^-P(t)\neq 0 \quad \textit{for every } t\in(a,b].
\end{aligned}
$$

Then

$$(e_{\mathrm{d}P}(t,t_0))^{-1} = e_{\mathrm{d}(\ominus P)}(t,t_0) \quad \textit{for} \ \ t \in [a,b], \tag{7.9.9}$$

where

$$(\ominus P)(t) = -P(t) + \sum_{s\in[t_0,t)} \frac{(\Delta^+ P(s))^2}{1+\Delta^+ P(s)} - \sum_{s\in(t_0,t]} \frac{(\Delta^- P(s))^2}{1-\Delta^- P(s)}.$$

Proof. For $t \in [a,b]$, we have $(\ominus P)(t) = -P(t) + R_1(t) - R_2(t)$, where

$$R_1(t) = \sum_{s\in[t_0,t)} \frac{(\Delta^+ P(s))^2}{1+\Delta^+ P(s)}, \quad R_2(t) = \sum_{s\in(t_0,t]} \frac{(\Delta^- P(s))^2}{1-\Delta^- P(s)}.$$

These functions have bounded variation on $[a,b]$ and satisfy

$$\begin{aligned} \Delta^- R_1(t) &= 0, \qquad \Delta^+ R_1(t) = \frac{(\Delta^+ P(t))^2}{1+\Delta^+ P(t)}, \\ \Delta^+ R_2(t) &= 0, \qquad \Delta^- R_2(t) = \frac{(\Delta^- P(t))^2}{1-\Delta^- P(t)}. \end{aligned} \tag{7.9.10}$$

Thus, $\ominus P$ has bounded variation on $[a,b]$ and

$$1+\Delta^+(\ominus P)(t) = 1-\Delta^+ P(t) + \frac{(\Delta^+ P(t))^2}{1+\Delta^+ P(t)} = \frac{1}{1+\Delta^+ P(t)} \neq 0 \quad \text{for } t \in [a,t_0),$$

$$1-\Delta^-(\ominus P)(t) = 1+\Delta^- P(t) + \frac{(\Delta^- P(t))^2}{1-\Delta^- P(t)} = \frac{1}{1-\Delta^- P(t)} \neq 0 \quad \text{for } t \in (t_0,b],$$

which implies that the exponential function $e_{\mathrm{d}(\ominus P)}$ is defined.

Using the relations (7.9.10) together with the definition of $\oplus$ given in Theorem 7.9.7, we obtain

$$\begin{aligned} (P \oplus (\ominus P))(t) &= P(t) - P(t) + R_1(t) - R_2(t) \\ &\quad + \sum_{s\in[t_0,t)} \Delta^+(-P+R_1-R_2)(s)\, \Delta^+ P(s) \\ &\quad - \sum_{s\in(t_0,t]} \Delta^-(-P+R_1-R_2)(s)\, \Delta^- P(s) = R_1(t) - R_2(t) \\ &\quad + \sum_{s\in[t_0,t)} \left(-(\Delta^+ P(s))^2 + \frac{(\Delta^+ P(s))^3}{1+\Delta^+ P(s)} \right) - \sum_{s\in(t_0,t]} \left(-(\Delta^- P(s))^2 - \frac{(\Delta^- P(s))^3}{1-\Delta^- P(s)} \right) \\ &= R_1(t) - R_2(t) - \sum_{s\in[t_0,t)} \frac{(\Delta^+ P(s))^2}{1+\Delta^+ P(s)} + \sum_{s\in(t_0,t]} \frac{(\Delta^- P(s))^2}{1-\Delta^- P(s)} = 0. \end{aligned}$$

Now, it follows from Theorem 7.9.7 and part (1) of Theorem 7.9.4 that

$$e_{\mathrm{d}P}(t,t_0)\, e_{\mathrm{d}(\ominus P)}(t,t_0) = e_{\mathrm{d}(P\oplus(\ominus P))}(t,t_0) = 1,$$

which proves the relation (7.9.9). □

Our next goal is to investigate the sign of the generalized exponential function corresponding to a real function P.

7.9.10 Theorem. *Consider a function* $P\colon [a,b] \to \mathbb{R}$, *which has bounded variation and*

$$1 + \Delta^+ P(t) \neq 0 \quad \textit{for every } \ t \in [a,b),$$
$$1 - \Delta^- P(t) \neq 0 \quad \textit{for every } \ t \in (a,b].$$

Then, for every $t_0 \in [a,b]$, *the following statements hold:*

(i) $e_{\mathrm{d}P}(t,t_0) \neq 0$ *for all* $t \in [a,b]$.

(ii) *If* $t \in [a,b)$ *and* $1 + \Delta^+ P(t) < 0$, *then* $e_{\mathrm{d}P}(t,t_0)\, e_{\mathrm{d}P}(t+,t_0) < 0$.

(iii) *If* $t \in (a,b]$ *and* $1 - \Delta^- P(t) < 0$, *then* $e_{\mathrm{d}P}(t,t_0)\, e_{\mathrm{d}P}(t-,t_0) < 0$.

(iv) *If* $t \in (a,b)$, $1 + \Delta^+ P(t) > 0$ *and* $1 - \Delta^- P(t) > 0$, *then* $t \mapsto e_{\mathrm{d}P}(t,t_0)$ *does not change sign in the neighborhood of* t.

Proof. If $e_{\mathrm{d}P}(t,t_0) = 0$ for a certain $t \in [a,b]$, we can use Theorem 7.9.4 to obtain

$$1 = e_{\mathrm{d}P}(t_0,t_0) = e_{\mathrm{d}P}(t_0,t) e_{\mathrm{d}P}(t,t_0) = 0,$$

which is a contradiction. This proves statement (i). Statements (ii) and (iii) follow immediately from part (iii) of Theorem 7.9.4. Finally, if $1 + \Delta^+ P(t) > 0$ and $1 - \Delta^- P(t) > 0$, then $e_{\mathrm{d}P}(t+,t_0)$ and $e_{\mathrm{d}P}(t-,t_0)$ have the same sign as $e_{\mathrm{d}P}(t,t_0)$, which proves (iv). □

According to the previous theorem, the exponential function changes sign at all points t such that $1 + \Delta^+ P(t) < 0$ or $1 - \Delta^- P(t) < 0$. Since P has bounded variation, we conclude that the interval $[a,b]$ can contain only finitely many points where the exponential function changes its sign.

The next theorem describes the class of all real functions P for which the generalized exponential function remains positive.

7.9.11 Theorem. *Let* $\mathcal{P}_+$ *be the class consisting of all functions* $P\colon [a,b] \to \mathbb{R}$ *that have bounded variation and satisfy* $1 + \Delta^+ P(t) > 0$ *for every* $t \in [a,b)$, *and* $1 - \Delta^- P(t) > 0$ *for every* $t \in (a,b]$. *Then the following statements hold:*

(i) $P \in \mathcal{P}_+$ *if and only if* $1+\Delta^+P(t) \neq 0$ *for every* $t \in [a,b)$, $1-\Delta^-P(t) \neq 0$ *for every* $t \in (a,b]$, *and the inequality* $e_{\mathrm{d}P}(t,t_0) > 0$ *holds for all* $t, t_0 \in [a,b]$.

(ii) *If* $P, Q \in \mathcal{P}_+$, *then* $P \oplus Q \in \mathcal{P}_+$.

(iii) *If* $P \in \mathcal{P}_+$, *then* $\ominus P \in \mathcal{P}_+$.

Proof. The first statement follows from Theorem 7.9.10 and the fact that $e_{\mathrm{d}P}(t_0,t_0)$ is positive. Statement (ii) is a consequence of the formulas

$$1-\Delta^-(P\oplus Q)(t) = \big(1-\Delta^-P(t)\big)\,\big(1-\Delta^-Q(t)\big),$$
$$1+\Delta^+(P\oplus Q)(t) = \big(1+\Delta^+P(t)\big)\,\big(1+\Delta^+Q(t)\big),$$

which were obtained in the proof of Theorem 7.9.7. Similarly, the last statement is a consequence of the formulas

$$1+\Delta^+(\ominus P)(t) = \frac{1}{1+\Delta^+P(t)}, \qquad 1-\Delta^-(\ominus P)(t) = \frac{1}{1-\Delta^-P(t)},$$

which were obtained in the proof of Theorem 7.9.9. □

Using the exponential function, we can now introduce the generalized hyperbolic functions.

7.9.12 Definition. Consider a function $P:[a,b]\to\mathbb{C}$, which has bounded variation on $[a,b]$. Let $t_0 \in [a,b]$ and assume that

$$1-(\Delta^+P(t))^2 \neq 0 \text{ for every } t\in[a,t_0) \text{ and } 1-(\Delta^-P(t))^2 \neq 0 \text{ for every } t\in(t_0,b]$$

Then we define the generalized hyperbolic functions $\cosh_{\mathrm{d}P}$ and $\sinh_{\mathrm{d}P}$ by the formulas

$$\cosh_{\mathrm{d}P}(t,t_0) = \frac{e_{\mathrm{d}P}(t,t_0)+e_{\mathrm{d}(-P)}(t,t_0)}{2} \quad \text{for } t\in[a,b],$$
$$\sinh_{\mathrm{d}P}(t,t_0) = \frac{e_{\mathrm{d}P}(t,t_0)-e_{\mathrm{d}(-P)}(t,t_0)}{2} \quad \text{for } t\in[a,b].$$

Note that the condition $1-(\Delta^+P(t))^2 \neq 0$ is equivalent to

$$(1+\Delta^+P(t))(1+\Delta^+(-P)(t)) \neq 0,$$

and $1-(\Delta^-P(t))^2 \neq 0$ is equivalent to

$$(1-\Delta^-P(t))(1-\Delta^-(-P)(t)) \neq 0.$$

Therefore, $e_{\mathrm{d}P}$ and $e_{\mathrm{d}(-P)}$ are well defined.

Obviously, the two hyperbolic functions are real if P is real, and for $P(s)=s$, we obtain the classical hyperbolic functions:

$$\cosh_{\mathrm{d}P}(t,t_0)=\cosh(t-t_0), \quad \sinh_{\mathrm{d}P}(t,t_0)=\sinh(t-t_0).$$

More generally, if P is continuous, then

$$\cosh_{\mathrm{d}P}(t,t_0)=\cosh(P(t)-P(t_0)) \text{ and } \sinh_{\mathrm{d}P}(t,t_0)=\sinh(P(t)-P(t_0)).$$

In the next theorem, we obtain the analogues of the well-known formulas

$$(\cosh z)'=\sinh z, \quad (\sinh z)'=\cosh z \text{ and } \cosh^2 z-\sinh^2=1.$$

7.9.13 Theorem. *Consider a function* $P:[a,b]\to\mathbb{C}$, *which has bounded variation and*

$$1-(\Delta^+P(t))^2\neq 0 \quad \textit{for every } t\in[a,t_0),$$
$$1-(\Delta^-P(t))^2\neq 0 \quad \textit{for every } t\in(t_0,b].$$

The generalized hyperbolic functions have the following properties:

(i) $\cosh_{\mathrm{d}P}(t_0,t_0)=1,\ \sinh_{\mathrm{d}P}(t_0,t_0)=0.$

(ii) $\cosh_{\mathrm{d}P}(t,t_0)=1+\int_{t_0}^t \sinh_{\mathrm{d}P}(s,t_0)\,\mathrm{d}P(s)$ *for* $t\in[a,b]$.

(iii) $\sinh_{\mathrm{d}P}(t,t_0)=\int_{t_0}^t \cosh_{\mathrm{d}P}(s,t_0)\,\mathrm{d}P(s)$ *for* $t\in[a,b]$.

(iv) $\cosh^2_{\mathrm{d}P}(t,t_0)-\sinh^2_{\mathrm{d}P}(t,t_0)=e_{\mathrm{d}Q}(t,t_0)$ *for* $t\in[a,b]$, *where*

$$Q(t)=(P\oplus(-P))(t)=\int_{t_0}^t (\Delta^-P(s)-\Delta^+P(s))\,\mathrm{d}P(s)$$
$$=\sum_{s\in(t_0,t]}(\Delta^-P(s))^2-\sum_{s\in[t_0,t)}(\Delta^+P(s))^2$$

with the convention that $\Delta^+P(t)=0,\ \Delta^-P(t_0)=0.$

Proof. The first statement is obvious. Using the definition of the generalized exponential function, we obtain

$$\cosh_{\mathrm{d}P}(t,t_0)=\frac{1}{2}\left(1+\int_{t_0}^t e_{\mathrm{d}P}(s,t_0)\,\mathrm{d}P(s)+1+\int_{t_0}^t e_{\mathrm{d}(-P)}(s,t_0)\,\mathrm{d}(-P)(s)\right)$$
$$=1+\frac{1}{2}\int_{t_0}^t (e_{\mathrm{d}P}(s,t_0)-e_{\mathrm{d}(-P)}(s,t_0))\,\mathrm{d}P(s)=1+\int_{t_0}^t \sinh_{\mathrm{d}P}(s,t_0)\,\mathrm{d}P(s),$$

which proves statement (ii). Similarly we ca prove statement (iii). To verify the last one, observe that

$$
\begin{aligned}
&\cosh^2_{\mathrm{d}P}(t,t_0)-\sinh^2_{\mathrm{d}P}(t,t_0)\\
&\quad=\left(\frac{e_{\mathrm{d}P}(t,t_0)+e_{\mathrm{d}(-P)}(t,t_0)}{2}\right)^2-\left(\frac{e_{\mathrm{d}P}(t,t_0)-e_{\mathrm{d}(-P)}(t,t_0)}{2}\right)^2\\
&\quad=e_{\mathrm{d}P}(t,t_0)e_{\mathrm{d}(-P)}(t,t_0)=e_{\mathrm{d}(P\oplus(-P))}(t,t_0).
\end{aligned}
$$

From Theorem 7.9.7, we have

$$
\begin{aligned}
(P\oplus(-P))(t)&=-\int_{t_0}^{t}\Delta^+P(s)\,\mathrm{d}P(s)+\int_{t_0}^{t}\Delta^-P(s)\,\mathrm{d}P(s)\\
&=-\sum_{s\in[t_0,t)}(\Delta^+P(s))^2+\sum_{s\in(t_0,t]}(\Delta^-P(s))^2,
\end{aligned}
$$

which completes the proof. □

Finally, we introduce the generalized trigonometric functions.

7.9.14 Definition. Consider a function $P:[a,b]\to\mathbb{C}$, which has bounded variation on $[a,b]$. Let $t_0\in[a,b]$ and assume that

$$1+(\Delta^+P(t))^2\neq 0 \text{ for every } t\in[a,t_0) \text{ and } 1+(\Delta^-P(t))^2\neq 0 \text{ for every } t\in(t_0,b].$$

Then we define the generalized trigonometric functions $\cos_{\mathrm{d}P}$ and $\sin_{\mathrm{d}P}$ by the formulas

$$
\begin{aligned}
\cos_{\mathrm{d}P}(t,t_0)&=\frac{e_{\mathrm{d}(\mathrm{i}P)}(t,t_0)+e_{\mathrm{d}(-\mathrm{i}P)}(t,t_0)}{2}=\cosh_{\mathrm{d}(iP)}(t,t_0) &&\text{for } t\in[a,b],\\
\sin_{\mathrm{d}P}(t,t_0)&=\frac{e_{\mathrm{d}(\mathrm{i}P)}(t,t_0)-e_{\mathrm{d}(-\mathrm{i}P)}(t,t_0)}{2\mathrm{i}}=-i\sinh_{\mathrm{d}(iP)}(t,t_0) &&\text{for } t\in[a,b].
\end{aligned}
$$

Note that the condition $1+(\Delta^+P(t))^2\neq 0$ is equivalent to

$$(1+\Delta^+(\mathrm{i}P)(t))(1+\Delta^+(-\mathrm{i}P)(t))\neq 0,$$

and $1+(\Delta^-P(t))^2\neq 0$ is equivalent to

$$(1-\Delta^-(\mathrm{i}P)(t))(1-\Delta^-(-\mathrm{i}P)(t))\neq 0.$$

Therefore, $e_{\mathrm{d}(\mathrm{i}P)}$ and $e_{\mathrm{d}(-\mathrm{i}P)}$ are well defined. If P is a real function, both conditions are always satisfied.

Again, it is easy to see that for $P(s)=s$, our definitions coincide with the classical trigonometric functions:

$$\cos_{\mathrm{d}P}(t,t_0)=\cos(t-t_0),\quad \sin_{\mathrm{d}P}(t,t_0)=\sin(t-t_0).$$

Also, if P is continuous, then

$$\cos_{\mathrm{d}P}(t,t_0)=\cos(P(t)-P(t_0)) \text{ and } \sin_{\mathrm{d}P}(t,t_0)=\sin(P(t)-P(t_0)).$$

If P is real, the trigonometric functions are real as well: By part (7) of Theorem 7.9.4,

$$e_{\mathrm{d}(\mathrm{i}P)}+e_{\mathrm{d}(-\mathrm{i}P)}=e_{\mathrm{d}(\mathrm{i}P)}+\overline{e_{\mathrm{d}(\mathrm{i}P)}},$$

which is purely real. Similarly,

$$e_{\mathrm{d}(\mathrm{i}P)}-e_{\mathrm{d}(-\mathrm{i}P)}=e_{\mathrm{d}(\mathrm{i}P)}-\overline{e_{\mathrm{d}(\mathrm{i}P)}},$$

which is purely imaginary.

We now derive the analogues of the well-known formulas

$$(\cos z)'=-\sin z,\quad (\sin z)'=\cos z \text{ and } \cos^2 z+\sin^2=1.$$

7.9.15 Theorem. *Consider a function $P:[a,b]\to\mathbb{C}$, which has bounded variation and*

$$1+(\Delta^+P(t))^2\neq 0 \quad \textit{for every } t\in[a,t_0),$$
$$1+(\Delta^-P(t))^2\neq 0 \quad \textit{for every } t\in(t_0,b].$$

The generalized trigonometric functions have the following properties:

(i) $\cos_{\mathrm{d}P}(t_0,t_0)=1,\ \sin_{\mathrm{d}P}(t_0,t_0)=0.$

(ii) $\cos_{\mathrm{d}P}(t,t_0)=1-\int_{t_0}^t \sin_{\mathrm{d}P}(s,t_0)\,\mathrm{d}P(s)$ *for* $t\in[a,b]$.

(iii) $\sin_{\mathrm{d}P}(t,t_0)=\int_{t_0}^t \cos_{\mathrm{d}P}(s,t_0)\,\mathrm{d}P(s)$ *for* $t\in[a,b]$.

(iv) $\cos^2_{\mathrm{d}P}(t,t_0)+\sin^2_{\mathrm{d}P}(t,t_0)=e_{\mathrm{d}Q}(t,t_0)$ *for* $t\in[a,b]$, *where*

$$\begin{aligned}Q(t)=(\mathrm{i}P\oplus(-\mathrm{i}P))(t)&=\int_{t_0}^t(\Delta^+P(s)-\Delta^-P(s))\,\mathrm{d}P(s)\\&=\sum_{s\in[t_0,t)}(\Delta^+P(s))^2-\sum_{s\in(t_0,t]}(\Delta^-P(s))^2\end{aligned}$$

with the convention that $\Delta^+P(t)=0,\ \Delta^-P(t_0)=0.$

Proof. The first statement is obvious. Using the definition of the generalized exponential function, we get

$$\begin{aligned}\cos_{\mathrm{d}P}(t,t_0) &= \frac{1}{2}\left(1+\int_{t_0}^{t} e_{\mathrm{d}(\mathrm{i}P)}(s,t_0)\,\mathrm{d}[\mathrm{i}P(s)]+1+\int_{t_0}^{t} e_{\mathrm{d}(-\mathrm{i}P)}(s,t_0)\,\mathrm{d}[-\mathrm{i}P(s)]\right)\\ &= 1+\frac{\mathrm{i}}{2}\int_{t_0}^{t}\left(e_{\mathrm{d}(\mathrm{i}P)}(s,t_0)-e_{\mathrm{d}(-\mathrm{i}P)}(s,t_0)\right)\mathrm{d}P(s)\\ &= 1+\frac{\mathrm{i}}{2}\int_{t_0}^{t}\left(2\mathrm{i}\,\sin_{\mathrm{d}P}(s,t_0)\right)\mathrm{d}P(s) = 1-\int_{t_0}^{t}\sin_{\mathrm{d}P}(s,t_0)\,\mathrm{d}P(s),\end{aligned}$$

which proves statement (ii). The proof of statement (iii) is similar. To prove the last one, we observe that

$$\begin{aligned}&\cos^2_{\mathrm{d}P}(t,t_0)+\sin^2_{\mathrm{d}P}(t,t_0)\\ &= \left(\frac{e_{\mathrm{d}(\mathrm{i}P)}(t,t_0)+e_{\mathrm{d}(-\mathrm{i}P)}(t,t_0)}{2}\right)^2+\left(\frac{e_{\mathrm{d}(\mathrm{i}P)}(t,t_0)-e_{\mathrm{d}(-\mathrm{i}P)}(t,t_0)}{2\mathrm{i}}\right)^2\\ &= e_{\mathrm{d}(\mathrm{i}P)}(t,t_0)\,e_{\mathrm{d}(-\mathrm{i}P)}(t,t_0) = e_{\mathrm{d}(\mathrm{i}P\oplus(-\mathrm{i}P))}(t,t_0).\end{aligned}$$

From Theorem 7.9.7, we have

$$\begin{aligned}(\mathrm{i}P\oplus(-\mathrm{i}P))(t) &= \int_{t_0}^{t}\Delta^+P(s)\,\mathrm{d}P(s)-\int_{t_0}^{t}\Delta^-P(s)\,\mathrm{d}P(s)\\ &= \sum_{s\in[t_0,t)}(\Delta^+P(s))^2-\sum_{s\in(t_0,t]}(\Delta^-P(s))^2,\end{aligned}$$

which completes the proof. □

As far as an additional literature to this chapter is concerned, we can recommend for example the monographs [60], [85], [122], [147] or the articles [1], [46], [103], [126], [127].

Chapter 8

Miscellaneous additional topics

In this chapter we present selected applications of the Kurzweil-Stieltjes integral.

The following two sections are concerned with topics in functional analysis, namely, the general form of continuous linear functionals on $\mathrm{C}([a,b])$ and $\mathrm{G}([a,b])$, respectively.

8.1 Continuous linear functionals on the space of continuous functions

One of the most important tasks of functional analysis is to find explicit representations of continuous linear functionals on function spaces.

Recall that *linear functionals* on a space X are linear mappings of X into $\mathbb{R}$. The set of all linear functionals on X is a linear space when equipped with the usual operations of addition and scalar multiplication (defined pointwise). Further, if X is a Banach space equipped with a norm $\|\cdot\|_X$, then it is well known that a linear functional Φ on X is continuous if and only if it is bounded, i.e., if there is a number $K\in[0,\infty)$ such that $|\Phi(x)|\le K\,\|x\|_X$ holds for all $x\in X$. The space of continuous linear functionals on the Banach space X is denoted by X^* and is called the *dual* (or *adjoint*) *space* to X. Furthermore, X^* is a Banach space with respect to the norm given by

$$\|\Phi\|_{X^*}=\sup\{|\Phi(x)|: x\in X, \|x\|_X\le 1\}\quad \text{for } \Phi\in X^*.$$

It is known that continuous linear functionals on the space $\mathrm{C}([a,b])$ are well described by means of the Riesz representation formula involving the classical Riemann-Stieltjes (δ)-integral. In [55], we can find a proof of the Riesz theorem based on the Bernstein polynomials approximation of continuous functions. Herein, following the standards of many books in functional analysis, we rely on the Hahn-Banach Theorem to obtain the general form of continuous linear functionals on $\mathrm{C}([a,b])$.

8.1.1 Theorem (HAHN-BANACH). *Let X be a Banach space, and let $Y\subset X$ be its subspace. If $\Phi\in Y^*$ is an arbitrary continuous linear functional on Y, then there exists a continuous linear functional $\widetilde{\Phi}$ on X such that*

$$\widetilde{\Phi}(y)=\Phi(y)\quad \textit{for } y\in Y \quad \textit{and}\quad \|\widetilde{\Phi}\|_{X^*}=\|\Phi\|_{Y^*}. \tag{8.1.1}$$

Note that our formulation of the Hahn-Banach Theorem is not the most general one (see e.g. [30]). However, it is quite sufficient for our purposes.

Now we present a general form of continuous linear functionals on the space of continuous functions.

8.1.2 Theorem (RIESZ). *Φ is a continuous linear functional on $\mathrm{C}([a,b])$ if and only if there is a function $p \in \mathrm{BV}([a,b])$ such that $p(a)=0$ and*

$$\Phi(x) = (\delta)\int_a^b x\,\mathrm{d}p \quad \textit{for any function } x \in \mathrm{C}([a,b]). \tag{8.1.2}$$

In such a case, $\|\Phi\|_{(\mathrm{C}([a,b]))^*} = \operatorname{var}_a^b p$.

Proof. a) Let $x \in \mathrm{C}([a,b])$ and $p \in \mathrm{BV}([a,b])$ be given. Then by Theorem 5.6.3 the integral $(\delta)\int_a^b x\,\mathrm{d}p$ exists and, by Lemma 5.1.11, the inequality

$$\left|(\delta)\int_a^b x\,\mathrm{d}p\right| \le \left(\operatorname{var}_a^b p\right)\|x\|$$

is true. Hence, the mapping $\Phi_p : \mathrm{C}([a,b]) \to \mathbb{R}$ given by

$$\Phi_p(x) = (\delta)\int_a^b x\,\mathrm{d}p$$

is a continuous linear functional on $\mathrm{C}([a,b])$, and

$$\|\Phi_p\|_{(\mathrm{C}([a,b]))^*} \le \operatorname{var}_a^b p. \tag{8.1.3}$$

b) Let an arbitrary $\Phi \in (\mathrm{C}([a,b]))^*$ be given. Denote by X the set of all bounded functions on $[a,b]$. Obviously, X is a Banach space with respect to the supremum norm and $\mathrm{C}([a,b])$ is a closed subspace of X. For the rest of the proof, put $Y = \mathrm{C}([a,b])$.

By Theorem 8.1.1, there is a functional $\widetilde{\Phi} \in X^*$ such that $\|\widetilde{\Phi}\|_{X^*} = \|\Phi\|_{Y^*}$ and $\widetilde{\Phi}(y) = \Phi(y)$ for all $y \in Y$. Put

$$p(a) = 0 \quad \text{and} \quad p(t) = \widetilde{\Phi}(\chi_{[a,t]}) \quad \text{for } t \in (a,b]. \tag{8.1.4}$$

We will prove that $p \in \mathrm{BV}([a,b])$. To this aim, let $\boldsymbol{\alpha}$ be an arbitrary division of $[a,b]$. Denote $m = \nu(\boldsymbol{\alpha})$. Then

$$V(p,\boldsymbol{\alpha}) = \sum_{j=1}^m |p(\alpha_j) - p(\alpha_{j-1})| = \sum_{j=1}^m c_j\,[p(\alpha_j) - p(\alpha_{j-1})],$$

where $c_j = \text{sgn}[p(\alpha_j) - p(\alpha_{j-1})]$ for $j \in \{1,\dots,m\}$. Hence, with respect to the definition (8.1.4), we obtain

$$\begin{aligned} V(p,\boldsymbol{\alpha}) &= c_1\,\widetilde{\Phi}(\chi_{[a,\alpha_1]}) + \sum_{j=2}^{m} c_j\,\big[\widetilde{\Phi}(\chi_{[a,\alpha_j]}) - \widetilde{\Phi}(\chi_{[a,\alpha_{j-1}]})\big] \\ &= c_1\,\widetilde{\Phi}(\chi_{[a,\alpha_1]}) + \sum_{j=2}^{m} c_j\,\widetilde{\Phi}(\chi_{(\alpha_{j-1},\alpha_j]}) \\ &= \widetilde{\Phi}\Big(c_1\,\chi_{[a,\alpha_1]} + \sum_{j=2}^{m} c_j\,\chi_{(\alpha_{j-1},\alpha_j]}\Big) = \widetilde{\Phi}(h), \end{aligned}$$

where

$$h(t) = c_1\,\chi_{[a,\alpha_1]}(t) + \sum_{j=2}^{m} c_j\,\chi_{(\alpha_{j-1},\alpha_j]}(t) \quad \text{for } t \in [a,b].$$

Obviously $\|h\|_X = \|h\| = 1$ and hence $V(p,\boldsymbol{\alpha}) \le \|\widetilde{\Phi}\|_{X^*} = \|\Phi\|_{Y^*}$ for any division $\boldsymbol{\alpha}$ of $[a,b]$. This means that

$$\operatorname{var}_a^b p \le \|\Phi\|_{Y^*}. \tag{8.1.5}$$

It remains to show that (8.1.2) is true or, in other words, $\Phi = \Phi_p$. Let $x \in Y$ and $\varepsilon > 0$ be given. Since the function x is uniformly continuous on $[a,b]$, there is a $\delta > 0$ such that

$$|x(t) - x(s)| < \varepsilon \quad \text{whenever } t,s \in [a,b] \text{ and } |t-s| < \delta.$$

Without loss of generality we can assume that $\delta > 0$ is such that

$$\Big|S(x,\mathrm{d}p,P) - (\delta)\int_a^b x\,\mathrm{d}p\Big| < \varepsilon$$

for all partitions $P = (\boldsymbol{\beta},\boldsymbol{\eta})$ of $[a,b]$ with $|\boldsymbol{\beta}| < \delta$. Let $\boldsymbol{\alpha}$ be an arbitrary division of $[a,b]$ such that $|\boldsymbol{\alpha}| < \delta$, and consider the function

$$x_{\boldsymbol{\alpha}}(t) = \begin{cases} x(\alpha_1) & \text{if } t \in [a,\alpha_1], \\ x(\alpha_j) & \text{if } t \in (\alpha_{j-1},\alpha_j] \text{ and } j \in \{2,3,\dots,\nu(\boldsymbol{\alpha})\}. \end{cases}$$

It is easy to see that $\|x - x_{\boldsymbol{\alpha}}\|_X \le \varepsilon$ and

$$x_{\boldsymbol{\alpha}}(t) = x(\alpha_1)\,\chi_{[a,\alpha_1]}(t) + \sum_{j=2}^{\nu(\boldsymbol{\alpha})} x(\alpha_j)\,\chi_{(\alpha_{j-1},\alpha_j]}(t) \quad \text{for } t \in [a,b].$$

This together with (8.1.4) implies that

$$\begin{aligned}
\widetilde{\Phi}(x_{\boldsymbol{\alpha}}) &= x(\alpha_1)\,\widetilde{\Phi}(\chi_{[a,\alpha_1]}) + \sum_{j=2}^{\nu(\boldsymbol{\alpha})} x(\alpha_j)\,\big[\widetilde{\Phi}(\chi_{[a,\alpha_j]}) - \widetilde{\Phi}(\chi_{[a,\alpha_{j-1}]})\big] \\
&= x(\alpha_1)\,\big[p\,(\alpha_1) - p\,(a)\big] + \sum_{j=2}^{\nu(\boldsymbol{\alpha})} x(\alpha_j)\,\big[p\,(\alpha_j) - p\,(\alpha_{j-1})\big] \\
&= \sum_{j=1}^{\nu(\boldsymbol{\alpha})} x(\alpha_j)\,\big[p\,(\alpha_j) - p\,(\alpha_{j-1})\big] = S(x, \mathrm{d}p, (\boldsymbol{\alpha}, \boldsymbol{\xi})),
\end{aligned}$$

where $\boldsymbol{\xi} = \{\alpha_1, \ldots, \alpha_{\nu(\boldsymbol{\alpha})}\}$. Therefore

$$\begin{aligned}
\Big|\Phi(x) - (\delta)\int_a^b x\,\mathrm{d}p\Big| &= \Big|\widetilde{\Phi}(x) - (\delta)\int_a^b x\,\mathrm{d}p\Big| \\
&\le \Big|\widetilde{\Phi}(x) - \widetilde{\Phi}(x_{\boldsymbol{\alpha}})\Big| + \Big|\widetilde{\Phi}(x_{\boldsymbol{\alpha}}) - (\delta)\int_a^b x\,\mathrm{d}p\Big| \\
&< \|\widetilde{\Phi}\|_{X^*}\,\|x - x_{\boldsymbol{\alpha}}\|_X + \Big|S(x, \mathrm{d}p, (\boldsymbol{\alpha}, \boldsymbol{\xi})) - (\delta)\int_a^b x\,\mathrm{d}p\Big| \\
&< \|\widetilde{\Phi}\|_{X^*}\varepsilon + \varepsilon = \big(\|\Phi\|_{Y^*} + 1\big)\,\varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ can be arbitrarily small, we conclude that

$$\Phi(x) = \Phi_p(x) = (\delta)\int_a^b x\,\mathrm{d}p \quad \text{for } x \in \mathrm{C}([a,b]).$$

Finally, by (8.1.3) and (8.1.5), we have $\|\Phi\|_{(\mathrm{C}([a,b]))^*} = \|\Phi\|_{Y^*} = \operatorname{var}_a^b p.$ □

The correspondence

$$\Phi \in (\mathrm{C}([a,b]))^* \mapsto p \in \mathrm{BV}([a,b])$$

is not uniquely determined by the relation (8.1.2). Indeed, if p_1 and p_2 are functions such that $p_1 = p_2$ except for a countable set and (8.1.2) holds for p_1, then the relation is also satisfied for p_2. This is a consequence of the following lemma.

8.1.3 Lemma. *Let* $g \in \mathrm{BV}([a,b])$. *Then*

$$(\delta)\int_a^b f\,\mathrm{d}g = 0 \quad \textit{holds for any function } f \in \mathrm{C}([a,b]) \tag{8.1.6}$$

if and only if there is an at most countable set $D \subset (a,b)$ *such that*

$$g(t) = g(a) \quad \textit{for } t \in [a,b] \setminus D. \tag{8.1.7}$$

Proof. Without any loss of generality we may assume that $g(a)=0$.

a) Assume (8.1.7). Let f be an arbitrary continuous function on $[a,b]$. We will show that

$$(\delta)\int_a^b f\,\mathrm{d}g=0. \tag{8.1.8}$$

Clearly, if $D=\emptyset$, the integral is zero. Let $D=\{d\}$, where $d\in(a,b)$, and let $\varepsilon>0$ be given. Choose $\delta>0$ in such a way that

$$|f(t)-f(s)|<\varepsilon \quad \text{whenever } t,s\in[a,b] \text{ and } |t-s|<2\,\delta.$$

Consider an arbitrary partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$ such that $|\boldsymbol{\alpha}|<\delta$. Clearly, $S(f,\mathrm{d}g,P)=0$ if $d\notin\boldsymbol{\alpha}$. Otherwise, $d=\alpha_j$ for some $j\in\{1,\dots,\nu(\boldsymbol{\alpha})-1\}$ and we have

$$|S(f,\mathrm{d}g,P)|=\big|f(\xi_j)-f(\xi_{j+1})\big|\,|g(d)|<\varepsilon\,\|g\|,$$

which proves that (8.1.8) holds in the case when D is a singleton set. From the linearity of the integral with respect to the integrator, it follows that (8.1.8) holds if the set D is finite.

Now, assume that D is countable, i.e., $D=\{d_k\}$. For each $n\in\mathbb{N}$, put

$$g_n(t)=\begin{cases} g(t) & \text{if } t\in\{d_1,d_2,\dots,d_n\},\\ 0 & \text{if } t\in[a,b]\setminus\{d_1,d_2,\dots,d_n\}\end{cases}$$

Clearly, the sequence $\{g_n\}$ is pointwise convergent to g and $\operatorname{var}_a^b g_n\le\operatorname{var}_a^b g$ for all $n\in\mathbb{N}$. Furthermore, by the previous part of the proof we have $(\delta)\int_a^b f\,\mathrm{d}g_n=0$ for every $n\in\mathbb{N}$. Equality (8.1.8) is then a consequence of the Helly's convergence theorem (Theorem 5.7.6).

b) Let (8.1.6) hold. Inserting $f(t)\equiv 1$ into (8.1.6) we find that $g(b)=g(a)=0$. Put

$$f(t)=(\delta)\int_a^t g(s)\,\mathrm{d}s \quad \text{for } t\in[a,b].$$

Then $f\in\mathrm{C}([a,b])$ and, by Integration by parts Theorem (Theorem 5.5.1) and Substitution Theorem (Theorem 5.4.3), we obtain

$$\begin{aligned}(\delta)\int_a^b f\,\mathrm{d}g&=f(b)\,g(b)-f(a)\,g(a)-(\delta)\int_a^b g\,\mathrm{d}f\\ &=-(\delta)\int_a^b g^2(t)\,\mathrm{d}t.\end{aligned}$$

Let $t_0 \in (a,b)$ be such that g is continuous at t_0. Then, if it were $g(t_0) \neq 0$, we could find an $\eta > 0$ such that

$$g^2(t) > 0 \quad \text{for } t \in (t_0 - \eta, t_0 + \eta).$$

This leads to

$$(\delta)\int_a^b f\,\mathrm{d}g = -(\delta)\int_a^b g^2(t)\,\mathrm{d}t \leq -(\delta)\int_{t_0-\eta}^{t_0-\eta} g^2(t)\,\mathrm{d}t < 0,$$

which contradicts the assumption (8.1.6). Therefore, $g(t)$ can be nonzero only if t is a discontinuity point of g. By Theorem 2.3.2 there are at most countably many such points. This shows that (8.1.7) is true. □

8.1.4 Remark. From Theorem 8.1.2 and Lemma 8.1.3 we deduce that for any continuous linear functional Φ on the space $\mathrm{C}([a,b])$, there exists a unique function $p \in \mathrm{BV}([a,b])$ such that

$$\left.\begin{aligned} & p(a) = 0, \; p(t+) = p(t) \;\text{ for } t \in (a,b) \\ \text{and} \\ & \Phi(x) = (\delta)\int_a^b x\,\mathrm{d}p \quad \text{for any } x \in \mathrm{C}([a,b]). \end{aligned}\right\} \tag{8.1.9}$$

Functions $p \in \mathrm{BV}([a,b])$ that are continuous from the right on (a,b) and such that $p(a) = 0$ are called *normalized function of bounded variation*. The set of all such functions will be denoted by the symbol $\mathrm{NBV}([a,b])$. Obviously, $\mathrm{NBV}([a,b])$ is a closed subset of $\mathrm{BV}([a,b])$. Moreover, by Theorem 8.1.2 and Lemma 8.1.3, the spaces $(\mathrm{C}([a,b]))^*$ and $\mathrm{NBV}([a,b])$ are isomorphic, i.e., the mapping

$$\Phi \in (\mathrm{C}([a,b]))^* \mapsto p \in \mathrm{NBV}([a,b]) \tag{8.1.10}$$

is one-to-one. Note that the same statement holds if $\mathrm{NBV}([a,b])$ is replaced by the space of functions of bounded variation on $[a,b]$ that are left-continuous on (a,b) and vanish in some fixed point $c \in [a,b]$.

Let $\Phi \in (\mathrm{C}([a,b]))^*$ be given and let $p \in \mathrm{NBV}([a,b])$ be determined by (8.1.10). The following theorem shows that then the equality $\|\Phi\|_{(\mathrm{C}([a,b]))^*} = \|p\|_{\mathrm{BV}}$ is true, i.e., the spaces $(\mathrm{C}([a,b]))^*$ and $\mathrm{NBV}([a,b])$ are isometrically isomorphic.

8.1.5 Theorem. *If* $p \in \mathrm{NBV}([a,b])$ *and* $\Phi(x) = (\delta)\int_a^b x\,\mathrm{d}p$ *for* $x \in \mathrm{C}([a,b])$, *then*

$$\|\Phi\|_{(\mathrm{C}([a,b]))^*} = \|p\|_{\mathrm{BV}} = \operatorname{var}_a^b p.$$

More precisely, the mapping $\Phi \in (\mathrm{C}([a,b]))^* \mapsto p \in \mathrm{NBV}([a,b])$, *where* p *is determined by relation* (8.1.9), *is an isometry.*

Proof. By Lemma 5.1.11 we have $\left|(\delta)\displaystyle\int_a^b x\,\mathrm{d}p\right| \le \left(\operatorname{var}_a^b p\right)\|x\|$. This means that the inequality

$$\|\Phi\|_{(\mathrm{C}([a,b]))^*} \le \|p\|_{\mathrm{BV}} \tag{8.1.11}$$

is true. We will prove that for every $\varepsilon > 0$ there exists a function $\widetilde{x} \in \mathrm{C}([a,b])$ such that

$$\|\widetilde{x}\| = 1 \quad \text{and} \quad |\Phi(\widetilde{x})| > \operatorname{var}_a^b p - \varepsilon. \tag{8.1.12}$$

Let an arbitrary $\varepsilon > 0$ be given. Choose a division $\boldsymbol{\alpha}$ of $[a,b]$ in such a way that $\nu(\boldsymbol{\alpha}) \ge 2$ and

$$V(p,\boldsymbol{\beta}) > \operatorname{var}_a^b p - \frac{\varepsilon}{3} \quad \text{for any its refinement } \boldsymbol{\beta}. \tag{8.1.13}$$

Set $m = \nu(\boldsymbol{\alpha})$. Recall that the right-continuity of p on (a,b) implies that the function $v(t) = \operatorname{var}_a^t p$ is right-continuous on (a,b) (Corollary 2.3.4). Therefore, for any $j \in \{1,\dots,m-1\}$ there is a point $t_j \in (\alpha_j, \alpha_{j+1})$ such that

$$\operatorname{var}_{\alpha_j}^{t_j} p = v(t_j) - v(\alpha_j) < \frac{\varepsilon}{3(m-1)}. \tag{8.1.14}$$

Put

$$\widetilde{x}(t) = \begin{cases} \operatorname{sgn}\big(p(\alpha_1) - p(a)\big) & \text{if } t \in [a,\alpha_1] \\ \operatorname{sgn}\big(p(\alpha_{j+1}) - p(t_j)\big) & \text{if } t \in [t_j,\alpha_{j+1}],\ j \in \{1,\dots,m-1\}, \end{cases}$$

and extend the function $\widetilde{x}$ to a continuous function on $[a,b]$ in such a way that it will be linear on the intervals $[\alpha_j, t_j]$, $j = 1,\dots,m-1$. Obviously, $\|\widetilde{x}\| = 1$. Moreover,

$$(\delta)\int_a^b \widetilde{x}\,\mathrm{d}p = \big|p(\alpha_1) - p(a)\big| + \sum_{j=1}^{m-1}\big|p(\alpha_{j+1}) - p(t_j)\big| + \sum_{j=1}^{m-1}(\delta)\int_{\alpha_j}^{t_j}\widetilde{x}\,\mathrm{d}p.$$

Since $|\widetilde{x}(t)| \le 1$ for $t \in [a,b]$, it follows by (8.1.14) that

$$\begin{aligned}
&\left|(\delta)\int_a^b \widetilde{x}\,\mathrm{d}p\right| \\
&\quad \ge \big|p(\alpha_1) - p(a)\big| + \sum_{j=1}^{m-1}\big|p(\alpha_{j+1}) - p(t_j)\big| - \sum_{j=1}^{m-1}\operatorname{var}_{\alpha_j}^{t_j} p \\
&\quad \ge V(p,\widetilde{\boldsymbol{\alpha}}) - 2\sum_{j=1}^{m-1}\operatorname{var}_{\alpha_j}^{t_j} p > V(p,\widetilde{\boldsymbol{\alpha}}) - \frac{2\,\varepsilon}{3},
\end{aligned}$$

where $\widetilde{\boldsymbol{\alpha}} = \{a, \alpha_1, t_1, \alpha_2 \ldots, \alpha_{m-1}, t_{m-1}, b\}$. By (8.1.13) we get

$$\left|(\delta)\int_a^b \widetilde{x}\,\mathrm{d}p\right| > V(p, \widetilde{\boldsymbol{\alpha}}) - \frac{2\,\varepsilon}{3} > \operatorname{var}_a^b p - \varepsilon,$$

that is, (8.1.12) holds. Hence $\sup\limits_{\|x\|\le 1} |\Phi(x)| \ge \operatorname{var}_a^b p$, wherefrom by (8.1.11) the assertion of the theorem follows. □

In view of Theorem 8.1.5, the space $\mathrm{NBV}([a,b])$ can be identified with the space $(\mathrm{C}([a,b]))^*$.

8.1.6 Exercise. Prove the following assertion :
For a given continuous linear functional Φ on $\mathrm{C}([a,b])$, there exists a unique function $p \in \mathrm{BV}([a,b])$ such that

$$p\,(b) = 0, \;\; p\,(t-) = p\,(t) \;\; \textit{for} \;\; t \in (a,b)$$

and

$$\Phi(x) = (\delta)\int_a^b x\,\mathrm{d}p \quad \textit{for any} \;\; x \in \mathrm{C}([a,b]).$$

Furthermore, in Theorem 8.1.5, *the space* $\mathrm{NBV}([a,b])$ *can be replaced by the space of functions left-continuous on* (a,b) *and such that* $p\,(b) = 0$.

8.2 Continuous linear functionals on spaces of regulated functions

For continuous linear functionals on the space of regulated functions, an analogue of the Riesz representation (Theorem 8.2.1) requires a more general notion than the Riemann-Stieltjes integral. Results of this type are available in [60], [67], where the Dushnik integral was used, and in [17], where the Young integral was used. Herein, we show that continuous linear functionals on $\mathrm{G}([a,b])$ are well described by means of the Kurzweil-Stieltjes integral. To this end, we introduce the following notation:

8.2.1 Definition. We will say that $f : [a,b] \to \mathbb{R}$ is a summable function if f vanishes except for a countable set and $\sum\limits_{a \le t \le b} |f(t)| < \infty$. For simplicity, we denote

$$s[f] = \sum_{a \le t \le b} |f(t)|.$$

It is not difficult to see that a summable function $f : [a,b] \to \mathbb{R}$ has bounded variation. The following lemmas concerning summable functions will be useful later.

8.2.2 Lemma. *If* $r:[a,b]\to\mathbb{R}$ *is a summable function and* $x:[a,b]\to\mathbb{R}$ *is bounded, then the sum*

$$\Psi_r(x)=\sum_{a\le t\le b} r(t)\,x(t) \tag{8.2.1}$$

converges and $|\Psi_r(x)|\le s[r]\,\|x\|$.

Proof. By the definition of a summable function, the sum $\sum_{a\le t\le b}|r(t)|$ is finite. Therefore

$$|\Psi_r(x)|\le\sum_{a\le t\le b}|r(t)\,x(t)|\le\|x\|\sum_{a\le t\le b}|r(t)|,$$

is also finite and the result follows. □

8.2.3 Lemma. *If* Φ *is a continuous linear functional on* $\mathrm{G}([a,b])$, *then the function* $r:[a,b]\to\mathbb{R}$, *given by* $r(t)=\Phi(\chi_{[t]})$ *for* $t\in[a,b]$, *is summable.*

Proof. For each $n\in\mathbb{N}$, let $M_n=\{t\in[a,b]:r(t)\ge 1/n\}$. Assume that there exists $N\in\mathbb{N}$ such that M_N is infinite. Let $T\subset M_N$ be a finite set with m elements. Then $\|\chi_T\|=1$ and

$$\Phi(\chi_T)=\sum_{t\in T}\Phi(\chi_{[t]})=\sum_{t\in T}r(t)\ge\frac{m}{N}.$$

Note that $m\in\mathbb{N}$ can be arbitrarily large (by taking a sufficiently large set T), which contradicts the fact that Φ is bounded. Hence, M_n is finite for every $n\in\mathbb{N}$ and, consequently, the set $\{t\in[a,b]:r(t)>0\}=\bigcup_{n=1}^{\infty}M_n$ is countable. In a similar way, we can show that $\{t\in[a,b]:r(t)<0\}$ is also countable.

Assume $\{t_k\}$ is a non-repeating sequence of points in $[a,b]$ such that $r(t)\ne 0$ if and only if $t=t_k$ for some $k\in\mathbb{N}$. If $\{t_k\}$ is finite, then r is clearly summable. Otherwise, for each $n\in\mathbb{N}$ we get

$$\sum_{k=1}^{n}|r(t_k)|=\Big|\sum_{k=1}^{n}r(t_k)\,\lambda_k\Big|=\Big|\Phi\Big(\sum_{k=1}^{n}\lambda_k\,\chi_{[t_k]}\Big)\Big|\le\|\Phi\|_{(\mathrm{G}([a,b]))^*},$$

where $\lambda_k=\mathrm{sgn}(r(t_k))$ for $k\in\{1,\dots,n\}$. Thus the series $\sum_{k=1}^{\infty}|r(t_k)|$ converges, and we conclude that r is summable. □

Let $p \in \mathrm{BV}([a,b])$, $q \in \mathbb{R}$, and let $r:[a,b] \to \mathbb{R}$ be a summable function. For each $x \in \mathrm{G}([a,b])$, define

$$\left.\begin{aligned} \Phi^+_{p,q,r}(x) &= q\,x(a) + \int_a^b p\,\mathrm{d}x + \sum_{a \le t < b} r(t)\,\Delta^+ x(t), \\ \Phi^-_{p,q,r}(x) &= q\,x(a) + \int_a^b p\,\mathrm{d}x + \sum_{a < t \le b} r(t)\,\Delta^- x(t). \end{aligned}\right\} \qquad (8.2.2)$$

Lemma 8.2.2 together with the results of Chapter 6 ensures that both $\Phi^+_{p,q,r}$ and $\Phi^-_{p,q,r}$ are well-defined and linear. Moreover,

$$|\Phi^+_{p,q,r}(x)| \le \big(|q| + |p(a)| + |p(b)| + \mathrm{var}_a^b p + 2\,s[r]\big)\,\|x\| \quad \text{for } x \in \mathrm{G}([a,b]),$$

and similarly for $\Phi^-_{p,q,r}$. In summary, for each triple (p,q,r), the identities (8.2.2) define continuous linear functionals on $\mathrm{G}([a,b])$.

The following theorem shows that any continuous linear functionals on the space of regulated functions can be described by an identity of the form (8.2.2).

8.2.4 Theorem. *If Φ is a continuous linear functional on $\mathrm{G}([a,b])$, then there exist $p, \widetilde{p} \in \mathrm{BV}([a,b])$, $q \in \mathbb{R}$ and a summable function $r:[a,b] \to \mathbb{R}$ such that*

$$\Phi(x) = q\,x(a) + \int_a^b p\,\mathrm{d}x - \sum_{a \le t < b} r(t)\,\Delta^+ x(t) \quad \textit{for } x \in \mathrm{G}([a,b]) \qquad (8.2.3)$$

and

$$\Phi(x) = q\,x(a) + \int_a^b \widetilde{p}\,\mathrm{d}x + \sum_{a < t \le b} r(t)\,\Delta^- x(t) \quad \textit{for } x \in \mathrm{G}([a,b]). \qquad (8.2.4)$$

Proof. Let $p, r:[a,b] \to \mathbb{R}$ be given by

$$p(t) = \Phi(\chi_{[t,b]}) \quad \text{and} \quad r(t) = \Phi(\chi_{[t]}) \quad \text{for } t \in [a,b].$$

By Lemma 8.2.3, the function $r:[a,b] \to \mathbb{R}$ is summable. We will prove that $p \in \mathrm{BV}([a,b])$. To this end, consider an arbitrary division $\boldsymbol{\alpha}$ of $[a,b]$. Taking

$$c_j = \mathrm{sgn}(p\,(\alpha_{j-1}) - p\,(\alpha_j)) \quad \text{for } j \in \{1, \dots, \nu(\boldsymbol{\alpha})\},$$

we get

$$V(p,\boldsymbol{\alpha}) = \Big|\sum_{j=1}^{\nu(\boldsymbol{\alpha})} c_j\,[p\,(\alpha_{j-1}) - p\,(\alpha_j)]\Big| = \Big|\sum_{j=1}^{\nu(\boldsymbol{\alpha})} c_j\,\Phi(\chi_{[\alpha_{j-1},\alpha_j)})\Big| = |\Phi(h_{\boldsymbol{\alpha}})|,$$

where $h_{\boldsymbol{\alpha}} = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} c_j\,\chi_{[\alpha_{j-1},\alpha_j)}$. Since $\|h_{\boldsymbol{\alpha}}\| \le 1$, it follows that $\mathrm{var}_a^b\,p \le \|\Phi\|_{(\mathrm{G}([a,b]))^*}$.

Further, put $q = \Phi(\chi_{[a,b]})$ and

$$\Phi_{p,q,r}(x) = q\,x(a) + \int_a^b p\,\mathrm{d}x - \sum_{a \le t < b} r(t)\,\Delta^+ x(t) \quad \text{for } x \in \mathrm{G}([a,b]). \tag{8.2.5}$$

To show that $\Phi = \Phi_{p,q,r}$, it is enough to verify that

$$\Phi(x) = \Phi_{p,q,r}(x) \quad \text{for every } x \in \mathrm{S}([a,b]) \tag{8.2.6}$$

(the result then follows from the continuity of both functionals Φ and $\Phi_{p,q,r}$, together with the fact that $\mathrm{S}([a,b])$ is dense in $\mathrm{G}([a,b])$). Let $x \in \mathrm{S}([a,b])$ be given by

$$x = \sum_{j=0}^{m} c_j\,\chi_{[t_j]} + \sum_{j=1}^{m} d_j\,\chi_{(t_{j-1},t_j)},$$

or equivalently,

$$x = \sum_{j=0}^{m-1} (c_j - d_{j+1})\,\chi_{[t_j]} + c_m\,\chi_{[b]} + \sum_{j=1}^{m} d_j\,\chi_{[t_{j-1},t_j)},$$

where $\{t_0, t_1, \ldots, t_m\}$ is a division of $[a,b]$ and $c_j,\ d_j \in \mathbb{R}$ for all j. Then

$$\Phi(x) = \sum_{j=0}^{m-1} (c_j - d_{j+1})\,r(t_j) + c_m\,r(b) + \sum_{j=1}^{m} d_j\,[p(t_{j-1}) - p(t_j)].$$

On the other hand, by Examples 6.3.1 we have

$$\int_a^b p\,\mathrm{d}x = c_m\,p(b) - c_0\,p(a) + \sum_{j=1}^{m} d_j\,[p(t_{j-1}) - p(t_j)].$$

Noting that $p(b) = r(b)$ and $\Delta^+ x(t_j) = d_{j+1} - c_j$ for $j \in \{0, \ldots, m-1\}$, from the expressions above we obtain

$$\Phi(x) = c_0\,q + \int_a^b p\,\mathrm{d}x - \sum_{j=0}^{m-1} \Delta^+ x(t_j)\,r(t_j),$$

therefore (8.2.6) holds.

In order to prove (8.2.4), put $\widetilde{p}(t) = p(t) - r(t)$ for $t \in [a,b]$. Noticing that $\widetilde{p} \in \mathrm{BV}([a,b])$ and applying Lemma 6.3.18 we get

$$\int_a^b r\,\mathrm{d}x = r(a)\,\Delta^+ x(a) + \sum_{a<t<b} r(t)\,\Delta\,x(t) + r(b)\,\Delta^- x(b) \quad \text{for } x \in \mathrm{G}([a,b]).$$

This combined with (8.2.3) yields

$$\Phi(x) = q\,x(a) + \int_a^b \widetilde{p}\,\mathrm{d}x + \int_a^b r\,\mathrm{d}x - \sum_{a \le t < b} r(t)\,\Delta^+ x(t)$$
$$= q\,x(a) + \int_a^b \widetilde{p}\,\mathrm{d}x + \sum_{a < t \le b} r(t)\,\Delta^- x(t),$$

that is, (8.2.4) holds. □

The identities in Theorem 8.2.4 depend on the value of the regulated function at the initial point a. With a dependence on the end point b, two other representation formulas for continuous linear functionals on $\mathrm{G}([a,b])$ can be derived.

8.2.5 Theorem. *If* Φ *is a continuous linear functional on* $\mathrm{G}([a,b])$, *then there exist* $p, \widetilde{p} \in \mathrm{BV}([a,b])$, $q \in \mathbb{R}$ *and a summable function* $r : [a,b] \to \mathbb{R}$ *such that*

$$\Phi(x) = q\,x(b) - \int_a^b p\,\mathrm{d}x - \sum_{a \le t < b} r(t)\,\Delta^+ x(t) \quad \textit{for} \;\; x \in \mathrm{G}([a,b]). \qquad (8.2.7)$$

and

$$\Phi(x) = q\,x(b) - \int_a^b \widetilde{p}\,\mathrm{d}x + \sum_{a < t \le b} r(t)\,\Delta^- x(t) \quad \textit{for} \;\; x \in \mathrm{G}([a,b]). \qquad (8.2.8)$$

8.2.6 Exercise. Prove Theorem 8.2.5.
Hint: Consider $p, \widetilde{p} : [a,b] \to \mathbb{R}$ given respectively by

$$p(t) = \begin{cases} \Phi(\chi_{[a,t)}) & \text{if} \;\; t \in (a,b], \\ 0 & \text{if} \;\; t = a, \end{cases}$$

and $\widetilde{p}(t) = p(t) + r(t)$ for $t \in [a,b]$.

In the case when the function $x \in \mathrm{G}([a,b])$ is left-continuous on $(a,b]$, the identity (8.2.4) reduces to

$$\Phi(x) = q\,x(a) + \int_a^b p\,\mathrm{d}x,$$

while (8.2.8) yields

$$\Phi(x) = q\,x(b) - \int_a^b p\,\mathrm{d}x.$$

Analogous expressions hold if we consider functions $x \in \mathrm{G}([a,b])$ right-continuous on $[a,b)$, and use identities (8.2.3) and (8.2.7). Therefore, from Theorems 8.2.4 and 8.2.5 we obtain the following representation formulas for continuous linear functionals defined on certain subspaces of $\mathrm{G}([a,b])$.

8.2.7 Theorem. (i) Φ *is a continuous linear functional on* $\mathrm{G_R}([a,b])$ *if and only if there exist* $p, \tilde{p} \in \mathrm{BV}([a,b])$ *and* $q \in \mathbb{R}$ *such that*

$$\left.\begin{aligned} \Phi(x) &= q\,x(a) + \int_a^b p\,\mathrm{d}x, \\ \Phi(x) &= q\,x(b) - \int_a^b \widetilde{p}\,\mathrm{d}x, \end{aligned}\right\} \tag{8.2.9}$$

for every $x \in \mathrm{G_R}([a,b])$.

(ii) Φ *is a continuous linear functional on* $\mathrm{G_L}([a,b])$ *if and only if there exist* $p, \widetilde{p} \in \mathrm{BV}([a,b])$ *and* $q \in \mathbb{R}$ *such that*

$$\left.\begin{aligned} \Phi(x) &= q\,x(a) + \int_a^b p\,\mathrm{d}x, \\ \Phi(x) &= q\,x(b) - \int_a^b \widetilde{p}\,\mathrm{d}x, \end{aligned}\right\} \tag{8.2.10}$$

for every $x \in \mathrm{G_L}([a,b])$.

Proof. (i) Given $p, \tilde{p} \in \mathrm{BV}([a,b])$ and $q \in \mathbb{R}$, it is not difficult to see that each identity in (8.2.9) defines a continuous linear functional Φ on $\mathrm{G_R}([a,b])$.

Now, consider an arbitrary functional $\Phi \in (\mathrm{G_R}([a,b]))^*$ and, for $x \in \mathrm{G}([a,b])$, define

$$\widetilde{x}(t) = \begin{cases} x(t+) & \text{if } t \in [a,b), \\ x(b) & \text{if } t = b. \end{cases}$$

and $\widetilde{\Phi}(x) = \Phi(\widetilde{x})$. By Corollary 4.1.9, $\widetilde{x} \in \mathrm{G_R}([a,b])$, and hence the mapping $\widetilde{\Phi} : \mathrm{G}([a,b]) \to \mathbb{R}$ is well defined and linear. Furthermore,

$$|\widetilde{\Phi}(x)| \le \|\Phi\|_{(\mathrm{G_R}([a,b]))^*} \|\widetilde{x}\| \le \|\Phi\|_{(\mathrm{G_R}([a,b]))^*} \|x\| \quad \text{for } x \in \mathrm{G}([a,b]),$$

that is, $\widetilde{\Phi}$ is a continuous linear functional on $\mathrm{G}([a,b])$. Applying Theorems 8.2.4 and 8.2.5, the equalities in (8.2.9) follow from the fact that $\widetilde{\Phi} = \Phi$ on $\mathrm{G_R}([a,b])$.

The proof of (ii) is analogous. □

Next we show that Theorem 8.2.7(i) infers an isomorphism between $\mathrm{BV}([a,b])$ and the dual space of $\mathrm{G_R}([a,b])$.

8.2.8 Theorem. *For* $p \in \mathrm{BV}([a,b])$, *let*

$$\Phi_p(x) = p(a)\,x(a) + \int_a^b p\,\mathrm{d}x \quad \textit{for } x \in \mathrm{G_R}([a,b]).$$

Then, the mapping

$$p \in \mathrm{BV}([a,b]) \mapsto \Phi_p \in (\widetilde{\mathrm{G}}_{\mathrm{R}}([a,b]))^* \tag{8.2.11}$$

is an isomorphism.

Proof. Clearly, (8.2.11) defines a linear mapping and, by Theorem **??**, it is also surjective. Now, consider $p \in \mathrm{BV}([a,b])$ such that $\Phi_p \equiv 0$. Since by Examples 6.3.1 we have

$$\Phi_p(\chi_{[a,b]}) = p(a),$$

$$\Phi_p(\chi_{[\tau,b]}) = p\,(\tau) \;\text{ if }\; \tau \in (a,b],$$

we conclude that $p \equiv 0$, showing that the mapping is one-to-one.

Finally, from Theorem 6.3.5 it follows that

$$\|\Phi_p\|_{(\mathrm{G_L}([a,b]))^*} \le 2\,|p\,(a)| + |p\,(b)| + \operatorname{var}_a^b p \le 3\,\|p\|_{\mathrm{BV}},$$

which implies that (8.2.11) is continuous. □

Similarly, one can show that $\mathrm{BV}([a,b])$ is isomorphic to the dual space of $\mathrm{G_L}([a,b])$.

8.2.9 Theorem. *For* $p \in \mathrm{BV}([a,b])$, *let*

$$\Phi_p(x) = p(b)\,x(b) - \int_a^b p \,\mathrm{d}x, \quad x \in \widetilde{\mathrm{G}}_{\mathrm{L}}([a,b]).$$

Then, the mapping

$$p \in \mathrm{BV}([a,b]) \mapsto \Phi_p \in (\mathrm{G_L}([a,b]))^*$$

is an isomorphism.

8.2.10 Exercise. Prove Theorem 8.2.9.

8.2.11 Remark. It is worth highlighting that the representations given by (8.2.10) differ from the one presented in [60] not only in the integral used but also in its form. According to [60], a functional $\Phi \in (\mathrm{G_L}([a,b]))^*$ can be described by the equality

$$\Phi(x) = (\sigma\mathrm{D})\int_a^b x \,\mathrm{d}p \quad \text{for } x \in \mathrm{G_L}([a,b]),$$

where $p \in \mathrm{BV}([a,b])$, $p(a) = 0$, and the integral is understood as the (σ) Dushnik integral, cf. Section 6.12. On the other hand, the representation by means of the Kurzweil-Stieltjes integral, besides adding an extra term, has regulated functions x in the role of integrators.

Regarding continuous linear functionals on $\mathrm{G_{reg}}([a,b])$, $\widetilde{\mathrm{G}}_{\mathrm{reg}}([a,b])$, $\widetilde{\mathrm{G}}_{\mathrm{L}}([a,b])$ and $\widetilde{\mathrm{G}}_{\mathrm{R}}([a,b])$, their general representations can be obtained by following the same arguments as those used in the proof of Theorem 8.2.7. For example, for $\widetilde{\mathrm{G}}_{\mathrm{reg}}([a,b])$ we have the following statement.

8.2.12 Theorem. Φ *is a continuous linear functional on* $\widetilde{\mathrm{G}}_{\mathrm{reg}}([a,b])$ *if and only if there exist* $p\in\mathrm{BV}([a,b])$ *and* $q\in\mathbb{R}$ *such that* $\Phi=\Phi_{p,q}$, *where*

$$\Phi_{p,q}(x)=q\,x(a)+\int_a^b p\,\mathrm{d}x \quad \textit{for}\;\; x\in\widetilde{\mathrm{G}}_{\mathrm{reg}}([a,b]).$$

Moreover, the mapping

$$(p,q)\in\mathrm{BV}([a,b])\times\mathbb{R}\mapsto\Phi_{p,q}\in(\widetilde{\mathrm{G}}_{\mathrm{reg}}([a,b]))^*$$

is an isomorphism.

8.2.13 Exercise. Prove Theorem 8.2.12.
Hint: Consider the function $p:[a,b]\to\mathbb{R}$ given by

$$p\,(t)=\begin{cases}\Phi(\chi_{(a,b]}) & \text{if}\;\; t=a,\\ \Phi(\frac{1}{2}\chi_{[t]}+\chi_{(t,b]}) & \text{if}\;\; t\in(a,b),\\ \Phi(\chi_{[b]}) & \text{if}\;\; t=b.\end{cases}$$

8.3 Adjoint classes of KS-integrable functions

In mathematical analysis, to understand classes of functions which are integrable (in some sense) is fairly crucial for applications to differential equations. When the integration process is of the Stieltjes type though, there are two possible ways of addressing the question of integrability. First, for a fixed integrator g, we can ask for which functions f the integral $\int f\,\mathrm{d}g$ exists. Second, for a fixed integrand f, we can ask for which class of functions g the integral $\int f\,\mathrm{d}g$ exists. To put it another way, in Stieltjes-type integration a class of functions $\mathcal{B}$ determines a class $\mathcal{A}$ so that the integral $\int f\,\mathrm{d}g$ exists provided $f\in\mathcal{A}$ and $g\in\mathcal{B}$. Related notion is that of adjoint classes defined as follows (see also [18]).

8.3.1 Definition. Let $\mathcal{A}$ and $\mathcal{B}$ be two classes of functions defined on $[a,b]$, and consider an integration process T. We say that $\mathcal{A}$ is *adjoint with* $\mathcal{B}$ *regarding the integral* T if the following three conditions are satisfied:

(i) The integral $\displaystyle {}^{(T)}\!\!\int_a^b f\,\mathrm{d}g$ exists for every $f\in\mathcal{A}$ and every $g\in\mathcal{B}$.

(ii) If the integral ${}^{(T)}\!\!\int_a^b f\,\mathrm{d}g$ exists for every $g\in\mathcal{B}$, then $f\in\mathcal{A}$.

(iii) If the integral ${}^{(T)}\!\!\int_a^b f\,\mathrm{d}g$ exists for every $f\in\mathcal{A}$, then $g\in\mathcal{B}$.

Notice that in Definition 8.3.1 the order of sets $\mathcal{A}$ and $\mathcal{B}$ is important as it refers to the classes of integrands and integrators, respectively. Further, notice also that if $\mathcal{A}$ and $\mathcal{B}$ are adjoint with respect to a given integral T, then neither $\mathcal{A}$ nor $\mathcal{B}$ can be enlarged so that the integrability condition (i) is preserved. Moreover, conditions (ii) and (iii) mean that the integrability with respect to a given integral can be regarded as a tool to characterize some particular properties of functions.

In light of the definition above, we can see that Theorems 5.6.3, 5.8.3 and 5.8.5 imply that $\mathrm{C}([a,b])$ and $\mathrm{BV}([a,b])$ are adjoint classes of Riemann-Stieltjes integrable functions. From this, we know that the existence of the Riemann-Stieltjes integral with respect to every function of bounded variation ensures continuity. Obviously, we cannot expect such a property to hold in the theory of Kurzweil-Stieltjes integration. Indeed, the following result, based on a simple application of the bounded convergence theorem, indicates a whole class of discontinuous functions for which the Kurzweil-Stieltjes integral with respect to functions of bounded variation always exists.

8.3.2 Proposition. *Let $D=\{d_k\}\subset[a,b]$, $c\in\mathbb{R}$ and let $f:[a,b]\to\mathbb{R}$ be bounded and such that*

$$f(t)=c \quad \textit{for } t\in[a,b]\setminus D.$$

Then the integral $\int_a^b f\,\mathrm{d}g$ exists for all $g\in\mathrm{BV}([a,b])$.

Proof. If D is finite, the assertion of the corollary is obvious. So, let D be infinite. For each $n\in\mathbb{N}$, put $D_n=\{d_1,d_2,\dots,d_n\}$ and define

$$f_n(t)=\begin{cases} c & \text{if } t\in[a,b]\setminus D_n,\\ f(t) & \text{if } t\in D_n.\end{cases}$$

Since f_n is a finite step function, the integral $\int_a^b f_n\,\mathrm{d}g$ exists for every $g\in\mathrm{BV}([a,b])$ (cf. Corollary 6.3.2). It is easy to see that the sequence $\{f_n\}$ is uniformly bounded by $K=\max\{\|f\|,|c|\}$ and $f_n(t)$ tends to $f(t)$ for every $t\in[a,b]$. Hence the result follows from Theorem 6.8.8 (bounded convergence theorem). □

Adjoint classes of integrable functions have been studied in connection to a variety of integration theories in a handful of papers. Among those dealing with some generalizations of the Riemann-Stieltjes integral we can mention, for instance, [19], [20] and [138]. In what follows we will address the question of adjoint classes regarding the KS-integral. Having in mind Theorems 6.3.8 and 6.3.11, we investigate adjoint classes of KS-integrable functions in two directions. First, we show that the class of regulated functions, $\mathrm{G}([a,b])$, together with functions of bounded variation, $\mathrm{BV}([a,b])$, cannot be called adjoint with respect to the KS-integral. Alternatively, the reverse order, that is, $\mathrm{BV}([a,b])$ and $\mathrm{G}([a,b])$, yields a pair of adjoint classes of KS-integrable functions.

Next example shows that condition (ii) of Definition 8.3.1 fails to be true when $\mathcal{A}=\mathrm{G}([a,b])$ and $\mathcal{B}=\mathrm{BV}([a,b])$.

8.3.3 Example. Let $c\in\mathbb{R}$ and consider a function $f:[0,1]\to\mathbb{R}$ given by

$$f(t)=c \quad \text{for } t\in[a,b]\setminus D,$$

where $D=\{d_k\}\subset[a,b]$ is infinite and such that $\lim_{k\to\infty} d_k=d\notin D$ while $f(d_k)$ does not converge to c. Thus, f is not regulated, but by Proposition 8.3.2 the integral $\int_a^b f\,\mathrm{d}g$ exists for all $g\in\mathrm{BV}([a,b])$.

The example above leads us to conclude that $\mathrm{G}([a,b])$ is not adjoint with $\mathrm{BV}([a,b])$ regarding the KS-integral. Yet, we might wonder whether condition (iii) of Definition 8.3.1 holds for these classes of functions. As we will see in the following example, the answer is again negative.

8.3.4 Example. Let $g:[0,1]\to\mathbb{R}$ be given by

$$g(t)=\begin{cases}\frac{1}{k} & \text{if } t=\frac{1}{k} \text{ for some } k\in\mathbb{N} \text{ such that } k\geq 2,\\ 0 & \text{otherwise.}\end{cases} \tag{8.3.1}$$

Claim 1. $\operatorname{var}_a^b g=\infty$.

Note that $\Delta^- g(\frac{1}{k})=-\Delta^+ g(\frac{1}{k})=\dfrac{1}{k}$ for $k\in\mathbb{N}$ and hence

$$\sum_{0\leq t<1}|\Delta^+ g(t)|+\sum_{0<t\leq 1}|\Delta^- g(t)|=2\sum_{k=1}^{\infty}\frac{1}{k}=\infty.$$

On the other hand, if we had $g\in\mathrm{BV}([0,1])$, then by Corollary 2.3.8 we would have

$$\sum_{0\leq t<1}|\Delta^+ g(t)|+\sum_{0<t\leq 1}|\Delta^- g(t)|\leq \operatorname{var}_0^1 g.$$

Therefore $\operatorname{var}_0^1 g=\infty$.

Claim 2. $\int_0^1 f\,\mathrm{d}g=0$ *for every* $f:[0,1]\to\mathbb{R}$.

Let $f:[0,1]\to\mathbb{R}$ and $\varepsilon>0$ be given. Put $D=\{0\}\cup\{\frac{1}{k}:k\in\mathbb{N},\ k\geq 2\}$ and define

$$\delta(t)=\begin{cases}\operatorname{dist}(t,D) & \text{if } t\notin D,\\ \operatorname{dist}(t,D\setminus\{t\}) & \text{if } t=\frac{1}{k} \text{ for some } k\in\mathbb{N} \text{ such that } k\geq 2,\\ \eta, & \text{if } t=0,\end{cases}$$

where $\eta>0$ is such that

$$|f(0)|\,|g(s)|<\varepsilon \ \text{ for } \ s\in(0,\eta) \tag{8.3.2}$$

(such a number exits as $g(0+)=g(0)=0$). Let $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ be a δ-fine partition of $[0,1]$. Due to the definition of δ we have $\xi_1=0$ and for each $j\in\{2,\dots,\nu(\boldsymbol{\alpha})\}$ the subinterval $[\alpha_{j-1},\alpha_j]$ contains at most one point of D. Moreover, it is not difficult to see that if $\alpha_j=\xi_j=\frac{1}{k}$ for some $j\in\{2,\dots,\nu(\boldsymbol{\alpha})-1\}$ and $k\geq 2$, then $\xi_{j+1}=\xi_j=\frac{1}{k}$. Hence, in such case,

$$\begin{aligned}&f(\xi_j)\,(g(\alpha_j)-g(\alpha_{j-1}))+f(\xi_{j+1})\,(g(\alpha_{j+1})-g(\alpha_j))\\&\qquad=f(\tfrac{1}{k})\,(g(\alpha_{j+1})-g(\alpha_{j-1}))=0,\end{aligned}$$

since none of the points α_{j-1} and α_{j+1} belong to $D\setminus\{0\}$. Let $\widetilde{P}$ be the partition obtained from P by combining two adjacent subintervals whenever they have a common tag. All subintervals determined by $\widetilde{P}$, except the first one, contribute nothing to $S(f,\mathrm{d}g,\widetilde{P})$. Therefore

$$\big|S(f,\mathrm{d}g,P)\big|=\big|S(f,\mathrm{d}g,\widetilde{P})\big|=|f(0)\,g(\alpha_1)|<\varepsilon,$$

where the last inequality is due to (8.3.2). This proves Claim 2.

Concerning the function g in (8.3.1), it is worth highlighting that, unlike what we observed for the integration in the Kurzweil sense, the RS-integral fails to exist even for all continuous functions. This is a direct consequence of the fact that $\mathrm{C}([a,b])$ and $\mathrm{BV}([a,b])$ are adjoint with respect to the RS-integral. (See Theorem 5.8.3 for details.)

The example above not only invalidates condition (iii) for the classes $\mathrm{G}([a,b])$ and $\mathrm{BV}([a,b])$, but also shows that there is no class of functions $\mathcal{A}$ adjoint with $\mathrm{BV}([a,b])$ regarding the KS-integral. Nevertheless, as will see in the following theorem, integrability with respect to functions of bounded variation ensures boundedness of the integrator.

8.3.5 Theorem. *If $f:[a,b]\to\mathbb{R}$ is a function such that the integral $\int_a^b f\,\mathrm{d}g$ exists for every $g\in\mathrm{BV}([a,b])$, then f is bounded.*

Proof. For contradiction, assume that there exists an unbounded function f satisfying the hypothesis. Without loss of generality, assume that f is unbounded from above (otherwise, consider the function $-f$). Thus, there exist $c \in [a,b]$ and a sequence $\{t_n\}$ in $[a,b]$ such that

$$\lim_{n\to\infty} t_n = c \quad \text{and} \quad \lim_{n\to\infty} f(t_n) = +\infty.$$

Note that the intersection of $\{t_n : n \in \mathbb{N}\}$ with at least one of the intervals $[a,c)$, $(c,b]$ is an infinite set. Thus, we can assume that the sequence $\{t_n\}$ is monotone and $\{f(t_n)\}$ is increasing, with $f(t_n) > 0$ for all $n \in \mathbb{N}$. Denote

$$y_0 = 0, \qquad y_n = \frac{1}{f(t_n)} - \frac{1}{f(t_{n+1})}, \qquad s_n = \sum_{k=0}^{n} y_k, \quad \text{for } n \in \mathbb{N}.$$

If $\{t_n\}$ increases to c, define $g\colon [a,b] \to \mathbb{R}$ by

$$g(t) = \begin{cases} 0, & \text{if } t \in [a,t_1], \\ s_{n-1}, & \text{if } t \in [t_n, t_{n+1}),\ n \in \mathbb{N}, \\ \lim_{n\to\infty} s_n, & \text{if } t \in [c,b]. \end{cases}$$

Note that g is continuous at c and has bounded variation, because it is monotone. By hypothesis, the integral $\int_a^b f\,\mathrm{d}g$ exists; thus, using Hake's Theorem 6.5.5 we can write

$$\int_a^c f\,\mathrm{d}g = \lim_{n\to\infty}\left(\int_a^{t_n} f\,\mathrm{d}g + f(c)[g(c) - g(t_n)]\right) = \lim_{n\to\infty}\int_a^{t_n} f\,\mathrm{d}g. \quad (8.3.3)$$

For each $k \in \mathbb{N}$, calculating the integral we obtain

$$\int_{t_k}^{t_{k+1}} f\,\mathrm{d}g = \int_{t_k}^{t_{k+1}} f\,\mathrm{d}\big(s_{k-1}\chi_{[t_k,t_{k+1})} + s_k\chi_{[t_{k+1}]}\big) = f(t_{k+1})(s_k - s_{k-1}) = f(t_{k+1})\,y_k$$

(see Examples 6.3.1), and consequently

$$\int_a^{t_n} f\,\mathrm{d}g = \int_a^{t_1} f\,\mathrm{d}g + \sum_{k=1}^{n-1}\int_{t_k}^{t_{k+1}} f\,\mathrm{d}g = \sum_{k=1}^{n-1} f(t_{k+1})\,y_k. \quad (8.3.4)$$

We claim that $\sum_{n=1}^{\infty} f(t_{n+1})\,y_n$ diverges. Indeed, since the sequence $\{f(t_n)\}$ is increasing and unbounded, for each $n \in \mathbb{N}$ we can choose $m_n \in \mathbb{N}$, such that

$$m_n > n \quad \text{and} \quad \frac{f(t_n)}{f(t_{m_n+1})} < \frac{1}{2}.$$

Therefore,

$$\sum_{k=n}^{m_n} f(t_{k+1})\, y_k = \sum_{k=n}^{m_n} \frac{f(t_{k+1}) - f(t_k)}{f(t_k)} \geq \frac{1}{f(t_{m_n+1})} \sum_{k=n}^{m_n} [f(t_{k+1}) - f(t_k)] =$$
$$= 1 - \frac{f(t_n)}{f(t_{m_n+1})} > \frac{1}{2}.$$

This means that $\sum_{n=1}^{\infty} f(t_{n+1})\, y_n$ does not satisfy the Cauchy condition of convergence, and consequently the series diverges. Having this in mind, the equality (8.3.4) together with (8.3.3) contradicts the existence of the integral $\int_a^c f \,\mathrm{d}g$.

If, on the other hand, $\{t_n\}$ decreases to c, we redefine g accordingly and using similar argument we get that $\int_c^b f \,\mathrm{d}g$ equals a divergent series; again a contradiction. In summary, we conclude that a function satisfying the hypothesis must be bounded. □

Now we turn our attention to the following question: Is $\mathrm{BV}([a,b])$ adjoint with $\mathrm{G}([a,b])$ regarding KS-integral? We know from Theorem 6.3.11 that condition (i) of Definition 8.3.1 is satisfied for these classes of functions. The verification of the remaining ones is contained in Propositions 8.3.6 and 8.3.8.

8.3.6 Proposition. *If $f:[a,b] \to \mathbb{R}$ is a function such that the integral $\int_a^b f \,\mathrm{d}g$ exists for every $g \in \mathrm{G}([a,b])$, then $f \in \mathrm{BV}([a,b])$.*

Proof. For contradiction, assume that there exists a function f such that $\mathrm{var}_a^b f = \infty$, while the integral $\int_a^b f \,\mathrm{d}g$ exists for each $g \in \mathrm{G}([a,b])$. Recall that a function belongs to $\mathrm{BV}([a,b])$ if and only if each point of $[a,b]$ has a neighborhood on which the variation is finite. Hence, as $f \notin \mathrm{BV}([a,b])$, there must exist a point c which satisfies either $c \in (a,b]$ and $\mathrm{var}_t^c f = \infty$ for every $t \in [a,c)$, or $c \in [a,b)$ and $\mathrm{var}_c^t f = \infty$ for every $t \in (c,b]$. By Lemma 5.8.2 (i), in the former case there is an increasing sequence $\{t_k\}$ in (a,c) such that

$$\lim_{k\to\infty} t_k = c \quad \text{and} \quad \sum_{k=1}^{\infty} |f(t_{k+1}) - f(t_k)| = \infty$$

and further, due to Lemma 5.8.1, we can choose a sequence $\{y_k\}$ of positive numbers in such a way that

$$\lim_{k\to\infty} y_k = 0 \quad \text{and} \quad \sum_{k=1}^{\infty} y_k |f(t_{k+1}) - f(t_k)| = \infty. \tag{8.3.5}$$

Let $\lambda_k = \mathrm{sgn}\,\big(f(t_k) - f(t_{k+1})\big)$ for $k \in \mathbb{N}$. Define

$$g(t) = \begin{cases} y_k\, \lambda_k & \text{if } t \in (t_k, t_{k+1}) \text{ and } k \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, $g \in \mathrm{G}([a,b])$ and $g(c-) = g(c) = 0$. Since $\int_a^b f\,\mathrm{d}g$ exists, the integral also exists in every subinterval of $[a,b]$. In particular, $\int_a^{t_1} f\,\mathrm{d}g = 0$ and for each $k \in \mathbb{N}$, by Examples 6.3.1 (ii), we obtain

$$\int_{t_k}^{t_{k+1}} f\,\mathrm{d}g = y_k\,\lambda_k \int_{t_k}^{t_{k+1}} f\,\mathrm{d}\big[\chi_{(t_k,t_{k+1})}\big] = y_k\,\lambda_k\,\big(f(t_k) - f(t_{k+1})\big),$$

that is,

$$\int_{t_k}^{t_{k+1}} f\,\mathrm{d}g = y_k\,|f(t_{k+1}) - f(t_k)|.$$

Due to the convergence of the sequence $\{t_n\}$ and Hake's Theorem 6.5.5, we can write

$$\int_a^c f\,\mathrm{d}g = \lim_{n\to\infty}\left(\int_a^{t_n} f\,\mathrm{d}g + f(c)\,[g(c) - g(t_n)]\right)$$
$$= \int_a^{t_1} f\,\mathrm{d}g + \lim_{n\to\infty}\sum_{k=1}^{n-1}\int_{t_k}^{t_{k+1}} f\,\mathrm{d}g = \lim_{n\to\infty}\sum_{k=1}^{n-1} y_k|f(t_{k+1}) - f(t_k)|.$$

Hence, in view of (8.3.5), $\int_a^c f\,\mathrm{d}g$ diverges, which is a contradiction. Therefore we conclude that a function satisfying the hypothesis must have a bounded variation.

In the latter case, i.e., when there is $c \in [a,b)$ such that $\mathrm{var}_c^t f = \infty$ for every $t \in (c,b]$, the proof is similar, but relies on part (ii) of Lemma 5.8.2. □

We remark that, in the proposition above, $\mathrm{G}([a,b])$ cannot be replaced by the class $\mathrm{C}([a,b])$. In other words, the existence of the Kurzweil–Stieltjes integral with respect to continuous integrators does not ensure bounded variation. To illustrate this fact we can again make use of the function given in (8.3.1).

8.3.7 Example. Let $f\colon [0,1] \to \mathbb{R}$ be given by

$$f(t) = \begin{cases} \frac{1}{k} & \text{if } t = \frac{1}{k} \text{ for some } k \in \mathbb{N} \text{ such that } k \geq 2, \\ 0 & \text{otherwise.} \end{cases}$$

We know that $f \notin \mathrm{BV}([0,1])$ (see Example 8.3.1). Let us prove that $\int_0^1 f\,\mathrm{d}g = 0$ for every $g \in C[0,1]$. Consider an arbitrary $\varepsilon > 0$. Since g is uniformly continuous, for each $k \in \mathbb{N}$ there exists $\rho_k > 0$ such that

$$\omega_{Q_k}(g) < \frac{\varepsilon}{k}, \quad Q_k = (\tfrac{1}{k} - \rho_k, \tfrac{1}{k} + \rho_k) \cap [0,1],$$

where $\omega_{Q_k}(g)$ denotes the modulus of oscillation of g on Q_k. Put

$$D = \{0\} \cup \{\frac{1}{k} : k \in \mathbb{N},\ k \ge 2\}$$

and define a gauge δ on $[0,1]$ by

$$\delta(t) = \begin{cases} \text{dist}\,(t, D) & \text{if } t \notin D, \\ \min\left\{\rho_k\,,\ \text{dist}\,(t, D \setminus \{t\})\right\} & \text{if } t = \frac{1}{k},\ k \in \mathbb{N},\ k \ge 2, \\ 1 & \text{if } t = 0. \end{cases}$$

Let $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ be a δ-fine partition of $[0,1]$. Then, for each $j \in \{2, \ldots, \nu(\boldsymbol{\alpha})\}$, the subinterval $[\alpha_{j-1}, \alpha_j]$ contains at most one point of D. Let Λ be the set of all indices $j \in \{1, \ldots, \nu(\boldsymbol{\alpha})\}$ such that $\xi_j = \frac{1}{k_j}$ for some $k_j \in \mathbb{N}$. Clearly, $[\alpha_{j-1}, \alpha_j] \subset Q_{k_j}$ for every $j \in \Lambda$. Therefore,

$$|S(f, \mathrm{d}g, P)| = \Big|\sum_{j \in \Lambda} f(\xi_j)\,(g(\alpha_j) - g(\alpha_{j-1}))\Big| \le \sum_{j \in \Lambda} \frac{1}{k_j}\,\omega_{Q_{k_j}}(g) < \varepsilon \sum_{n=1}^{\infty} \frac{1}{n^2}$$

wherefrom it follows that $\int_0^1 f\,\mathrm{d}g = 0$.

To conclude that $\mathrm{BV}([a,b])$ and $\mathrm{G}([a,b])$ are adjoint classes it remains to verify also condition (iii). Such a characterization of a regulated function via the integrability in the KS-sense has been already investigated in Proposition 2.60 of [29]. Indeed, the result in [29] shows that we need not test the integrability of the whole space $\mathrm{BV}([a,b])$ but simply the finite step functions of the form χ_J where J is an arbitrary interval. Herein we present a slightly different proof for the result in [29].

8.3.8 Proposition. *Let $g : [a,b] \to \mathbb{R}$. If, for every interval J, the integral $\int_a^b \chi_J\,\mathrm{d}g$ exists, then $g \in \mathrm{G}([a,b])$.*

Proof. Given $c \in [a,b)$ we will prove that $g(c+)$ exists. To this aim, let $f = \chi_{(c,b]}$ and let $\varepsilon > 0$ be given. Thus, we can choose a gauge δ on $[a,b]$ such that

$$\Big|S(f, \mathrm{d}g, P) - \int_a^b f\,\mathrm{d}g\Big| < \frac{\varepsilon}{2} \quad \text{for every } \delta\text{-fine partition } P \text{ of } [a,b]. \quad (8.3.6)$$

In view of Lemma 6.1.12 we can assume that the gauge δ is such that every δ-fine partition contains c as the tag of some subinterval.

Fix an arbitrary δ-fine partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]$ and let $\ell \in \{1, \ldots, \nu(\boldsymbol{\alpha})\}$ be such that $\xi_\ell = c$. For each $t \in (c, \alpha_\ell)$, define the partition

$$\widetilde{P}_t = P_t^- \cup P_t \cup Q_t \cup P_t^+,$$

where $P_t^- = (\{\alpha_0,\dots,\alpha_{\ell-1}\},\{\xi_1,\dots,\xi_{\ell-1}\})$, $P_t = (\{\alpha_{\ell-1},t\},\{c\})$, Q_k is an arbitrary δ-fine partition of $[t,\alpha_\ell]$ and $P_t^+ = (\{\alpha_\ell,\dots,\alpha_{\nu(\boldsymbol{\alpha})}\},\{\xi_{\ell+1},\dots,\xi_{\nu(\boldsymbol{\alpha})}\})$. Thus, $\widetilde{P}_t$ is a δ-fine partition of $[a,b]$ and a simple calculation shows that

$$S(f,\mathrm{d}g,\widetilde{P}_t) = S(f,\mathrm{d}g,Q_t) + S(f,\mathrm{d}g,P_t^+) = g(b) - g(t).$$

This together with (8.3.6) implies that for each pair $u,v\in(c,\alpha_\ell)$ we get

$$|g(u)-g(v)| = |S(f,\mathrm{d}g,\widetilde{P}_u) - S(f,\mathrm{d}g,\widetilde{P}_v)| < \varepsilon,$$

showing that the Cauchy condition for the existence of $g(c+)$ is satisfied. Analogously, we can show the existence of $g(c-)$ for every $c\in(a,b]$. □

Propositions 8.3.6 and 8.3.8 lead to the following corollary.

8.3.9 Corollary. $\mathrm{BV}([a,b])$ *is adjoint with* $\mathrm{G}([a,b])$ *regarding the KS-integral.*

8.4 Distributions

In this section we outline some applications of the Kurzweil-Stieltjes integral in the theory of distributions, which are understood in the sense of L. Schwartz [**?**]. Let us recall some of the basic notions and definitions.

8.4.1 Definition. The symbol $\mathfrak{D}[a,b]$ stands for the set of functions $\varphi:\mathbb{R}\to\mathbb{R}$, which are infinitely differentiable and such that $\varphi^{(k)}(t)=0$ for all $t\in\mathbb{R}\setminus(a,b)$ and $k\in\mathbb{N}\cup\{0\}$. Functions from $\mathfrak{D}[a,b]$ are called *test functions* on $[a,b]$.

The set $\mathfrak{D}[a,b]$ is a linear space when equipped with the usual operations of addition and scalar multiplication. We say that a sequence φ_n converges to φ in $\mathfrak{D}[a,b]$ if and only if

$$\lim_{n\to\infty}\|\varphi_n^{(k)} - \varphi^{(k)}\| = 0 \quad \text{for any } k\in\mathbb{N}\cup\{0\}.$$

With the topology induced by the notion of convergence above, $\mathfrak{D}[a,b]$ is a topological vector space.

Typical examples of test functions on $[a,b]$ are given by:

$$\varphi_{c,d}(t) = \begin{cases} \exp\left(\frac{1}{c-t}+\frac{1}{t-d}\right) & \text{for } t\in(c,d),\\ 0 & \text{for } t\in\mathbb{R}\setminus(c,d), \end{cases} \tag{8.4.1}$$

where $[c,d]$ is an arbitrary closed subinterval of $[a,b]$.

8.4.2 Definition. Continuous linear functionals on the topological vector space $\mathfrak{D}[a,b]$ are called *distributions* on $[a,b]$, and the set of all distributions on $[a,b]$ is denoted by the symbol $\mathfrak{D}^*[a,b]$.

In other words, the set $\mathfrak{D}^*[a,b]$ is the dual space to $\mathfrak{D}[a,b]$.

For a given distribution $f\in\mathfrak{D}^*[a,b]$ and a test function $\varphi\in\mathfrak{D}[a,b]$, the value $f(\varphi)$ is traditionally denoted by $\langle f,\varphi\rangle$.

8.4.3 Remark. Let $f\in \mathrm{L}^1([a,b])$ be given and let

$$\langle f,\varphi\rangle=\int_a^b f(t)\,\varphi(t)\,\mathrm{d}t \quad \text{for } \varphi\in\mathfrak{D}[a,b]$$

(where the integral is understood as the Lebesgue integral). The mapping $\varphi\mapsto\langle f,\varphi\rangle$ defines a distribution on $[a,b]$, which will be also denoted by the symbol f. We say that the distribution f is determined by the function f.

The null element of the space $\mathfrak{D}^*[a,b]$ is the distribution that maps each test function to zero. Notice that this distribution is determined by an arbitrary measurable function f which vanishes almost everywhere in $[a,b]$. In particular, if $f\in \mathrm{G}([a,b])$, then $f=0\in\mathfrak{D}^*[a,b]$ if and only if $f(t-)=f(s+)=0$ for $t\in(a,b]$, $s\in[a,b)$. Likewise, if $f\in \mathrm{G}_{\text{reg}}([a,b])$, then $f=0\in\mathfrak{D}^*[a,b]$ if and only if $f(t)=0$ for all $t\in[a,b]$. Consequently, if $g\in \mathrm{L}^1([a,b])$, then there is at most one function $f\in \mathrm{G}_{\text{reg}}([a,b])$ such that $f=g$ a.e. on $[a,b]$. Furthermore, for $f,g\in \mathrm{L}^1([a,b])$, the equality $f=g$ holds in the sense of $\mathfrak{D}^*[a,b]$ if and only if $f=g$ a.e. on $[a,b]$.

8.4.4 Definition. For a given distribution $f\in\mathfrak{D}^*[a,b]$, its *distributional derivative* f' is defined by $\langle f',\varphi\rangle=-\langle f,\varphi'\rangle$ for $\varphi\in\mathfrak{D}[a,b]$.

Similarly, for each $k\in\mathbb{N}$, we define

$$\langle f^{(k)},\varphi\rangle=(-1)^k\langle f,\varphi^{(k)}\rangle \quad \text{for } \varphi\in\mathfrak{D}[a,b].$$

Note that distributional derivatives of absolutely continuous functions are determined by their classical derivatives.

8.4.5 Example. Given an arbitrary $\tau\in(a,b)$, by δ_τ we denote the *Dirac δ-distribution* (concentrated in τ) defined by

$$\langle\delta_\tau,\varphi\rangle=\varphi(\tau) \quad \text{for every } \varphi\in\mathfrak{D}[a,b].$$

One can show that δ_τ corresponds to the distributional derivative of the *Heaviside function* given by $h_\tau(t)=H(t-\tau)$, $t\in[a,b]$, where

$$H(t)=\begin{cases}0 & \text{for } t<0,\\ \frac{1}{2} & \text{for } t=0,\\ 1 & \text{for } t>0.\end{cases}$$

Indeed, by Theorems 6.4.2 and 6.6.1, as well as relations (6.3.1) and (6.3.5), we get

$$\begin{aligned}\langle h_\tau', \varphi\rangle &= -\langle h_\tau, \varphi'\rangle = -\int_a^b h_\tau(t)\,\varphi'(t)\,\mathrm{d}t\\ &= -\int_a^b h_\tau(t)\,\mathrm{d}\Big[\int_a^t \varphi'(s)\mathrm{d}s\Big] = -\int_a^b h_\tau\,\mathrm{d}\varphi = \int_a^b \varphi\,\mathrm{d}h_\tau = \varphi(\tau),\end{aligned}$$

which shows that $\delta_\tau = h_\tau'$ in the sense of distributions.

8.4.6 Theorem. *Let $f \in \mathrm{L}^1([a,b])$. Then its distributional derivative f' is the zero distribution if and only if there is a constant $c \in \mathbb{R}$ such that $f(t) = c$ for almost all $t \in [a,b]$.*

Proof. Let $f(t) = c$ for almost all $t \in [a,b]$ and $\varphi \in \mathfrak{D}[a,b]$, then

$$\langle f', \varphi\rangle = -\langle c, \varphi'\rangle = -c\int_a^b \varphi'(s)\,\mathrm{d}s = -c\,(\varphi(b) - \varphi(a)) = 0.$$

Conversely, assume that distributional derivative f' is the zero distribution. [1] Given an arbitrary test function $\varphi \in \mathfrak{D}[a,b]$, let

$$\rho(t) = \begin{cases} \displaystyle\int_a^t \big(\varphi(s) - a_0\,\Theta(s)\big)\,\mathrm{d}s & \text{for } t \in [a,b],\\ 0 & \text{for } t \in \mathbb{R}\setminus[a,b],\end{cases}$$

where

$$a_0 = \int_a^b \varphi(s)\,\mathrm{d}s, \qquad \Theta(t) = \frac{\varphi_{a,b}(t)}{\displaystyle\int_a^b \varphi_{a,b}(s)\,\mathrm{d}s},$$

and $\varphi_{a,b}$ is the function given by (8.4.1). Then

$$\int_a^b \Theta(s)\,\mathrm{d}s = 1$$

wherefrom it follows easily that $\rho(a) = \rho(b) = 0$, i.e., $\rho \in \mathfrak{D}[a,b]$. Furthermore,

$$\rho'(t) = \varphi(t) - a_0\,\Theta(t) \quad \text{for } t \in [a,b].$$

Hence $0 = \langle f, \rho'\rangle = \langle f, \varphi\rangle - \Big(\int_a^b \varphi(s)\,\mathrm{d}s\Big)\langle f, \Theta\rangle$. Therefore, by letting $c = \langle f, \Theta\rangle$, we get

$$\langle f, \varphi\rangle = \Big(\int_a^b \varphi(s)\,\mathrm{d}s\Big)\langle f, \Theta\rangle = \int_a^b c\,\varphi(s)\,\mathrm{d}s$$

for any $\varphi \in \mathfrak{D}[a,b]$. Thus, $f = c$ in the sense of distributions. □

[1] Originally we had the following: "On the other hand, assume that $\langle f, \varphi'\rangle = 0$ for any $\varphi \in \mathfrak{D}[a,b]$."

8.4.7 Exercise. For $f \in \mathrm{L}^1([a,b])$ and $k \in \mathbb{N} \cup \{0\}$ show that $f^{(k)} = 0 \in \mathfrak{D}^*[a,b]$ if and only if there are $c_0, c_1, \dots, c_{k-1} \in \mathbb{R}$ such that

$$f(t) = c_0 + c_1\,t + \cdots + c_{k-1}\,t^{k-1} \quad \text{for almost all } t \in [a,b].$$

An important problem of the theory of distributions is a proper definition of the product of two distributions. The following classical definitions apply only to very special kinds of distributions.

8.4.8 Definition. (i) If $f, g \in \mathrm{L}^1([a,b])$ are such that $f\,g \in \mathrm{L}^1([a,b])$, then

$$\langle f\,g, \varphi\rangle = \int_a^b f(t)\,g(t)\,\varphi(t)\,\mathrm{d}t \quad \text{for } \varphi \in \mathfrak{D}[a,b].$$

(ii) If $f \in \mathfrak{D}^*[a,b]$ and $g : [a,b] \to \mathbb{R}$ is infinitely differentiable on $[a,b]$, then $\langle fg, \varphi\rangle = \langle f, g\varphi\rangle$.

To deal with differential equations with distributional coefficients, it is useful to have a reasonable definition of the distributions $f\,g'$ and $f'\,g$, where $f \in \mathrm{G}([a,b])$ and $g \in \mathrm{BV}([a,b])$. Obviously, Definition 8.4.8 does not cover such cases. To formulate proper definitions for such couples, the Kurzweil-Stieltjes integral turns to be very helpful.

8.4.9 Definition. If $f \in \mathrm{G}([a,b])$ and $g \in \mathrm{BV}([a,b])$, then we define

$$\langle f'\,g, \varphi\rangle = \int_a^b \varphi\,g\;\mathrm{d}f \qquad \text{and} \qquad \langle f\,g', \varphi\rangle = \int_a^b \varphi f\,\mathrm{d}g \quad \text{for } \varphi \in \mathfrak{D}[a,b].$$

8.4.10 Theorem. Let $f \in \mathrm{G}([a,b])$ and $g \in \mathrm{BV}([a,b])$ be such that

$$\Delta^+ f(t)\,\Delta^+ g(t) = \Delta^- f(t)\,\Delta^- g(t) \quad \text{for every } t \in (a,b). \tag{8.4.2}$$

Then

$$(f\,g)' = f\,g' + f'\,g.$$

Proof. Using Definition 8.4.4 together with Integration by parts Theorem (Theorem 6.4.2) and Substitution Theorem (Theorem 6.6.1), for $\varphi \in \mathfrak{D}[a,b]$ we obtain

$$\begin{aligned}
\langle (f\,g)', \varphi\rangle &= -\langle f\,g, \varphi'\rangle \\
&= -\int_a^b f(t)\,g(t)\,\varphi'(t)\mathrm{d}t = -\int_a^b f(t)\,g(t)\,\mathrm{d}\varphi(t) \\
&= \int_a^b \varphi(t)\,\mathrm{d}\big[f(t)\,g(t) - f(a)\,g(a)\big] \\
&= \int_a^b \varphi(t)\,\mathrm{d}\Big[\int_a^t g\,\mathrm{d}f + \int_a^t f\,\mathrm{d}g\Big] \\
&= \int_a^b \varphi\,g\,\mathrm{d}f + \int_a^b \varphi\,f\,\mathrm{d}g = \langle f'\,g, \varphi\rangle + \langle f\,g', \varphi\rangle.
\end{aligned}$$

□

8.4.11 Remark. Let $f \in \mathrm{G}([a,b])$ and $g \in \mathrm{BV}([a,b])$. Then the condition (8.4.2) is obviously satisfied e.g. if

- both functions are regular (see Remark 4.2.5),
- at least one of them is continuous on (a,b),
- one of them is left-continuous on (a,b) and the other is right-continuous on (a,b).

8.4.12 Exercises.

- Let $\tau \in (a,b)$, $\varkappa \in \mathbb{R}$ and

$$g(t) = \begin{cases} 0 & \text{if} \quad t < \tau, \\ \varkappa & \text{if} \quad t = \tau, \\ 1 & \text{if} \quad t > \tau. \end{cases}$$

  Prove that $\int_a^b \varphi\, g\, \mathrm{d}g = \varphi(\tau)\, \varkappa$ for any $\varphi \in \mathrm{BV}([a,b])$.

  Hint: Using Exercise 6.3.3, determine the values of the integrals $\int_a^\tau \varphi\, g\, \mathrm{d}g$ and $\int_\tau^b \varphi\, g\, \mathrm{d}g$.

- If h_τ and δ_τ are the Heaviside and Dirac distributions concentrated at a point $\tau \in (a,b)$, show that $h_\tau\, \delta_\tau = \delta_\tau/2$.

8.5 Integration on time scales

The time scale calculus, which originated in the work of S. Hilger [57], is a popular tool that provides a unification of the continuous and discrete calculus. It is concerned with functions $f : \mathbb{T} \to \mathbb{R}$, where $\mathbb{T}$ is a *time scale* – an arbitrary nonempty closed set $\mathbb{T} \subset \mathbb{R}$. As we will see, the choice $\mathbb{T} = \mathbb{R}$ leads to the classical continuous calculus, while $\mathbb{T} = \mathbb{Z}$ corresponds to the discrete calculus. Another frequently studied time scale is $\mathbb{T} = q^{\mathbb{Z}} = \{q^n : n \in \mathbb{Z}\}$, where $q > 1$; this leads to the quantum calculus. The basic operations of the time scale calculus are the Δ-derivative, ∇-derivative, Δ-integral, and ∇-integral. The main goal of this section is show that both types of integrals on time scales are special cases of the Kurzweil-Stieltjes integral.

Let us start by introducing some basic notation. If $t \in \mathbb{T}$ and $t < \sup \mathbb{T}$, we denote

$$\sigma(t) = \inf\{s \in \mathbb{T} : s > t\}, \qquad \mu(t) = \sigma(t) - t.$$

Moreover, if $t = \sup \mathbb{T} < \infty$, we define $\sigma(t) = t$, $\mu(t) = 0$. The functions $\sigma : \mathbb{T} \to \mathbb{T}$ and $\mu : \mathbb{T} \to [0, \infty)$ are referred to as the *forward jump operator* and *forward graininess*, respectively. If $t \in \mathbb{T}$ satisfies $\sigma(t) > t$, we say that t is *right-scattered*; otherwise, if $\sigma(t) = t$ and $t < \sup \mathbb{T}$, then t is called *right-dense*.

Similarly, if $t \in \mathbb{T}$ and $t > \inf \mathbb{T}$, let

$$\rho(t) = \sup\{s \in \mathbb{T} : s < t\}, \qquad \nu(t) = t - \rho(t).$$

Moreover, if $t = \inf \mathbb{T} > -\infty$, we define $\rho(t) = t$, $\nu(t) = 0$. The functions $\rho : \mathbb{T} \to \mathbb{T}$ and $\nu : \mathbb{T} \to [0, \infty)$ are called the *backward jump operator* and *backward graininess*, respectively. If $t \in \mathbb{T}$ satisfies $\rho(t) < t$, we say that t is *left-scattered*; otherwise, if $\rho(t) = t$ and $t > \inf \mathbb{T}$, then t is called *left-dense*.

For an arbitrary pair of real numbers $a < b$, we use the following notation:

$$[a,b]_{\mathbb{T}} = [a,b] \cap \mathbb{T}, \ [a,b)_{\mathbb{T}} = [a,b) \cap \mathbb{T}, \ (a,b]_{\mathbb{T}} = (a,b] \cap \mathbb{T}, \ (a,b)_{\mathbb{T}} = (a,b) \cap \mathbb{T}.$$

These sets are referred to as the time scale intervals, and the subscript $\mathbb{T}$ helps to distinguish them from ordinary intervals.

A function $f : \mathbb{T} \to \mathbb{R}$ is called *rd-continuous*, if it is continuous at all right-dense points and regulated on $\mathbb{T}$. Similarly, f is called *ld-continuous*, if it is continuous at all left-dense points and regulated on $\mathbb{T}$.

Given a function $f : \mathbb{T} \to \mathbb{R}$, we can introduce the Δ-derivative and the ∇-derivative of f at a point $t \in \mathbb{T}$. Although our main interest lies in integration theory, we include the definitions of both derivatives and some of their properties (for more details, see [14, 15]).

8.5.1 Definition. Consider a function $f : \mathbb{T} \to \mathbb{R}$ and a point $t \in \mathbb{T}$.

- Suppose that $t < \sup \mathbb{T}$, or $t = \sup \mathbb{T}$ and $\rho(t) = t$. We say that the Δ-derivative $f^{\Delta}(t)$ exists and equals $D \in \mathbb{R}$, if for every $\varepsilon > 0$, there is a $\delta > 0$ such that

  $$|f(\sigma(t)) - f(s) - D(\sigma(t) - s)| < \varepsilon |\sigma(t) - s|$$

  for all $s \in (t - \delta, t + \delta)_{\mathbb{T}}$.

- Suppose that $t > \inf \mathbb{T}$, or $t = \inf \mathbb{T}$ and $\sigma(t) = t$. We say that the ∇-derivative $f^{\nabla}(t)$ exists and equals $D \in \mathbb{R}$, if for every $\varepsilon > 0$, there is a $\delta > 0$ such that

  $$|f(\rho(t)) - f(s) - D(\rho(t) - s)| < \varepsilon |\rho(t) - s|$$

  for all $s \in (t - \delta, t + \delta)_{\mathbb{T}}$.

The following remarks should help to clarify the meaning of both derivatives, as well as the difference between them:

- If $\mathbb{T}=\mathbb{R}$, then $f^{\Delta}(t)=f^{\nabla}(t)=f'(t)$, i.e., both derivatives coincide with the classical derivative.
- If $\mathbb{T}=\mathbb{Z}$, then $f^{\Delta}(t)=f(t+1)-f(t)$ and $f^{\nabla}(t)=f(t)-f(t-1)$, i.e., the Δ- and ∇-derivative reduce to the forward difference and backward difference, respectively.
- More generally, if $t\in\mathbb{T}$ satisfies $\sigma(t)>t$, then

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

  Similarly, if $t\in\mathbb{T}$ satisfies $\rho(t)<t$, then

  $$f^{\nabla}(t)=\frac{f(t)-f(\rho(t))}{t-\rho(t)}=\frac{f(t)-f(\rho(t))}{\nu(t)}.$$

We now proceed to Δ- and ∇-integrals of a function $f:[a,b]_{\mathbb{T}}\to\mathbb{R}$, which are in a certain sense inverse operations to the Δ- and ∇-derivatives. As in the classical calculus, there exist definitions in the spirit of Newton, Riemann, Lebesgue, and Kurzweil. Moreover, these integrals also have their Stieltjes-type counterparts. We are primarily interested in Kurzweil integrals, but it is instructive to begin with Riemann integrals.

In the rest of this section, we always assume that $a,\ b\in\mathbb{T}$. A partition of $[a,b]_{\mathbb{T}}$ is a partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]$ such that both $\boldsymbol{\alpha}$ and $\boldsymbol{\xi}$ are subsets of $\mathbb{T}$. We keep the notation used earlier in this book and write

$$S(P)=\sum_{j=1}^{\nu(\boldsymbol{\alpha})} f(\xi_j)(\alpha_j-\alpha_{j-1}).$$

(Throughout this section, the symbol ν has two different meanings: $\nu(\boldsymbol{\alpha})$ denotes the number of subintervals in a division $\boldsymbol{\alpha}$, while $\nu(t)$ is the backward graininess at a point $t\in\mathbb{T}$. The meaning of ν will be always clear from the context.)

The definition of the classical Riemann integral relies on partitions of $[a,b]$ such that the distance of each two successive division points does not exceed a certain $\delta>0$. However, if we replace $[a,b]$ by the time scale interval $[a,b]_{\mathbb{T}}$, such a partition need not exist. For this reason, the Riemann Δ- and ∇-integrals involve a modified type of partitions that are described in the next definition.

8.5.2 Definition. Given a $\delta>0$, the symbol $\mathcal{P}_{\delta}([a,b]_{\mathbb{T}})$ will denote all partitions $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]_{\mathbb{T}}$ such that for each $i\in\{1,\ldots,\nu(\boldsymbol{\alpha})\}$ we have either $\alpha_i-\alpha_{i-1}\le\delta$, or $\sigma(\alpha_{i-1})=\alpha_i$.

We are now able to define the Riemann Δ- and ∇-integrals.

8.5.3 Definition. Consider a function $f:[a,b]_{\mathbb{T}}\to\mathbb{R}$.

- We say that the Riemann Δ-integral $\int_a^b f(t)\Delta t$ exists and equals $I\in\mathbb{R}$, if for every $\varepsilon>0$, there is a $\delta>0$ such that $|S(P)-I|<\varepsilon$ for all partitions $P\in\mathcal{P}_\delta([a,b]_{\mathbb{T}})$ satisfying $\xi_i\in[\alpha_{i-1},\rho(\alpha_i)]_{\mathbb{T}}$ for each $i\in\{1,\dots,\nu(\boldsymbol{\alpha})\}$.

- We say that the Riemann ∇-integral $\int_a^b f(t)\nabla t$ exists and equals $I\in\mathbb{R}$, if for every $\varepsilon>0$, there is a $\delta>0$ such that $|S(P)-I|<\varepsilon$ for all partitions $P\in\mathcal{P}_\delta([a,b]_{\mathbb{T}})$ satisfying $\xi_i\in[\sigma(\alpha_{i-1}),\alpha_i]_{\mathbb{T}}$ for each $i\in\{1,\dots,\nu(\boldsymbol{\alpha})\}$.

A complete theory of Riemann Δ- and ∇-integrals can be found in [15]; here we limit ourselves to several remarks:

- If $\mathbb{T}=\mathbb{R}$, then $\int_a^b f(t)\Delta t=\int_a^b f(t)\nabla t=\int_a^b f(t)\,\mathrm{d}t$, i.e., both integrals coincide with the classical Riemann integral.

- The requirement $\xi_i\in[\alpha_{i-1},\rho(\alpha_i)]_{\mathbb{T}}$ in the definition of the Δ-integral means that if α_i is left-scattered, then it cannot serve as a tag for $[\alpha_{i-1},\alpha_i]_{\mathbb{T}}$. Thus, if $t\in\mathbb{T}$ satisfies $\sigma(t)>t$, then the only possible partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ that is taken into account in the definition of $\int_t^{\sigma(t)} f(t)\Delta t$ is $t=\alpha_0=\xi_1<\alpha_1=\sigma(t)$. Therefore, $\int_t^{\sigma(t)} f(t)\Delta t=f(t)(\sigma(t)-t)=f(t)\mu(t)$.

  Similarly, the requirement $\xi_i\in[\sigma(\alpha_{i-1}),\alpha_i]_{\mathbb{T}}$ in the definition of the ∇-integral ensures that if α_{i-1} is right-scattered, then it cannot serve as a tag for $[\alpha_{i-1},\alpha_i]_{\mathbb{T}}$. Hence, if $t\in\mathbb{T}$ satisfies $\rho(t)<t$, the only possible partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ appearing in the definition of $\int_{\rho(t)}^t f(t)\nabla t$ is $\rho(t)=\alpha_0<\xi_1=\alpha_1=t$. Therefore, $\int_{\rho(t)}^t f(t)\nabla t=f(t)(t-\rho(t))=f(t)\nu(t)$.

- If $\mathbb{T}=\mathbb{Z}$, then $\int_a^b f(t)\Delta t=\sum_{t=a}^{b-1} f(t)$ and $\int_a^b f(t)\nabla t=\sum_{t=a+1}^{b} f(t)$ (see the previous remark; for $\delta\le 1$, there is only one possible choice for the partition of $[a,b]_{\mathbb{T}}$ appearing in Definition 8.5.3).

- If $f:[a,b]_{\mathbb{T}}\to\mathbb{R}$ is regulated, then it is Riemann Δ- and ∇-integrable, and the indefinite integrals

$$F_1(t)=\int_a^t f(s)\Delta s\quad\text{and}\quad F_2(t)=\int_a^t f(s)\nabla s,\quad t\in[a,b]_{\mathbb{T}},$$

  are continuous. Moreover, if f is rd-continuous, then $F_1^\Delta(t)=f(t)$ for all $t\in[a,b)_{\mathbb{T}}$; if f is ld-continuous, then $F_2^\nabla(t)=f(t)$ for all $t\in(a,b]_{\mathbb{T}}$.

- If F is continuous, $F^\Delta=f$ on $[a,b)_{\mathbb{T}}$ and f is Δ-integrable, then $\int_a^b f(t)\Delta t=F(b)-F(a)$. Similarly, if F is continuous, $F^\nabla=f$ on $(a,b]_{\mathbb{T}}$ and f is ∇-integrable, then $\int_a^b f(t)\nabla t=F(b)-F(a)$.

We now turn our attention to Kurzweil Δ- and ∇-integrals. As in the classical case, they have several advantages over the Riemann (or Lebesgue) integrals: the class of integrable functions is larger, and the assumptions of the fundamental theorem of calculus are less restrictive (each function which is a Δ- or ∇-derivative is Kurzweil Δ- or ∇-integrable). In time scale calculus, the concepts of a gauge on $[a,b]_{\mathbb{T}}$ and a δ-fine partition of $[a,b]_{\mathbb{T}}$ take the following form:

8.5.4 Definition. Consider a pair of functions $\delta_L, \delta_R : [a,b]_{\mathbb{T}} \to (0,\infty)$. Then $\delta = (\delta_L, \delta_R)$ is called a Δ-gauge on $[a,b]_{\mathbb{T}}$ if $\delta_R(t) \ge \mu(t)$ for all $t \in [a,b)_{\mathbb{T}}$, and a ∇-gauge on $[a,b]_{\mathbb{T}}$ if $\delta_L(t) \ge \nu(t)$ for all $t \in (a,b]_{\mathbb{T}}$. If $\delta = (\delta_L, \delta_R)$ is either a Δ-gauge or a ∇-gauge on $[a,b]_{\mathbb{T}}$, a partition $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ of $[a,b]_{\mathbb{T}}$ is called δ-fine if

$$[\alpha_{j-1}, \alpha_j] \subset [\xi_j - \delta_L(\xi_j), \xi_j + \delta_R(\xi_j)] \quad \text{for all } j = 1, \dots, \nu(\boldsymbol{\alpha}). \tag{8.5.1}$$

8.5.5 Remark. We emphasize that the interval on the right-hand side of (8.5.1) is closed and cannot be replaced by an open interval. This is in contrast to Definition 6.1.1, where it does not matter whether we choose an open or closed interval (see Remark 6.1.5) – both choices lead to the same definition of the Kurzweil-Stieltjes integral.

Instead of introducing the Kurzweil Δ- and ∇-integrals, we choose a more general approach in the spirit of this book, and define the Kurzweil-Stieltjes Δ- and ∇-integrals of a function $f : [a,b]_{\mathbb{T}} \to \mathbb{R}$ with respect to a function $g : [a,b]_{\mathbb{T}} \to \mathbb{R}$. As in the rest of the book, if $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ is a tagged partition of $[a,b]_{\mathbb{T}}$, we write

$$S(f, \mathrm{d}g, P) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} f(\xi_j)(g(\alpha_j) - g(\alpha_{j-1})).$$

8.5.6 Definition. Consider a pair of functions $f, g : [a,b]_{\mathbb{T}} \to \mathbb{R}$.

- We say that the Kurzweil-Stieltjes Δ-integral $\int_a^b f(t)\Delta g(t)$ exists and equals $I \in \mathbb{R}$, if for every $\varepsilon > 0$, there is a Δ-gauge δ on $[a,b]_{\mathbb{T}}$ such that

  $$|S(f, \mathrm{d}g, P) - I| < \varepsilon$$

  for all δ-fine partitions P of $[a,b]_{\mathbb{T}}$.

- We say that the Kurzweil-Stieltjes ∇-integral $\int_a^b f(t)\nabla g(t)$ exists and equals $I \in \mathbb{R}$, if for every $\varepsilon > 0$, there is a ∇-gauge δ on $[a,b]_{\mathbb{T}}$ such that

  $$|S(f, \mathrm{d}g, P) - I| < \varepsilon$$

  for all δ-fine partitions P of $[a,b]_{\mathbb{T}}$.

If $g(t)=t$ for all $t\in[a,b]_{\mathbb{T}}$, the two integrals are referred to as the Kurzweil Δ-integral and ∇-integral, and they are denoted by $\int_a^b f(t)\Delta t$ and $\int_a^b f(t)\nabla t$, respectively. Let us show that Riemann integrability implies Kurzweil integrability.

8.5.7 Theorem. *Consider a function* $f:[a,b]_{\mathbb{T}}\to\mathbb{R}$.

- *If the Riemann* Δ*-integral* $\int_a^b f(t)\Delta t$ *exists, then the Kurzweil* Δ*-integral* $\int_a^b f(t)\Delta t$ *also exists and has the same value.*
- *If the Riemann* ∇*-integral* $\int_a^b f(t)\nabla t$ *exists, then the Kurzweil* ∇*-integral* $\int_a^b f(t)\nabla t$ *also exists and has the same value.*

Proof. We prove the first statement, and leave the second up to the reader. Suppose that the Riemann Δ-integral $I=\int_a^b f(t)\Delta t$ exists, and choose an arbitrary $\varepsilon>0$. By definition, there is a $\delta>0$ such that $|S(P)-I|<\varepsilon$ for all partitions $P\in\mathcal{P}_\delta([a,b]_{\mathbb{T}})$ satisfying $\xi_i\in[\alpha_{i-1},\rho(\alpha_i)]_{\mathbb{T}}$ for each $i\in\{1,\dots,\nu(\boldsymbol{\alpha})\}$. We define a Δ-gauge $\tilde{\delta}=(\tilde{\delta}_L,\tilde{\delta}_R)$ on $[a,b]_{\mathbb{T}}$ as follows:

$$\tilde{\delta}_L(t)=\begin{cases}\nu(t)/2 & \text{if } t\in(a,b]_{\mathbb{T}} \text{ and } \rho(t)<t,\\ \delta & \text{otherwise,}\end{cases}$$

$$\tilde{\delta}_R(t)=\begin{cases}\mu(t) & \text{if } t\in[a,b)_{\mathbb{T}} \text{ and } \sigma(t)>t,\\ \delta & \text{otherwise.}\end{cases}$$

Now, consider an arbitrary $\tilde{\delta}$-fine partition $P=(\boldsymbol{\alpha},\boldsymbol{\xi})$ of $[a,b]_{\mathbb{T}}$. By splitting each interval-point pair $([\alpha_{i-1},\alpha_i],\xi_i)$ into $([\alpha_{i-1},\xi_i],\xi_i)$ and $([\xi_i,\alpha_i],\xi_i)$, we get a partition $P'=(\boldsymbol{\alpha}',\boldsymbol{\xi}')$ which is still $\tilde{\delta}$-fine and satisfies $S(P)=S(P')$. By the definition a $\tilde{\delta}$-fine partition, we have

$$[\alpha'_{j-1},\alpha'_j]\subset[\xi'_j-\tilde{\delta}_L(\xi'_j),\xi'_j+\tilde{\delta}_R(\xi'_j)]\quad\text{for all } j=1,\dots,\nu(\boldsymbol{\alpha}').$$

It follows from the construction of P' that for each $j=1,\dots,\nu(\boldsymbol{\alpha}')$, either $\xi'_j=\alpha'_{j-1}$, or $\xi'_j=\alpha'_j$. In the latter case, we have $\alpha'_{j-1}\in[\xi'_j-\tilde{\delta}_L(\xi'_j),\xi'_j)_{\mathbb{T}}$. Observe that if $\rho(\xi'_j)<\xi'_j$, then the interval $[\xi'_j-\tilde{\delta}_L(\xi'_j),\xi'_j)_{\mathbb{T}}$ is empty (by the definition of $\tilde{\delta}_L$). This shows that the tag ξ'_j cannot coincide with the right endpoint α'_j if it is left-scattered, i.e., we always have $\xi'_j\in[\alpha_{j-1},\rho(\alpha_j)]_{\mathbb{T}}$.

Finally, observe that the inequality $\alpha'_j-\alpha'_{j-1}>\delta$ can be true only if $\xi'_j=\alpha'_{j-1}$ and $\alpha'_j=\sigma(\alpha'_{j-1})$; this shows that $P'\in\mathcal{P}_\delta([a,b]_{\mathbb{T}})$. Consequently,

$$|S(P)-I|=|S(P')-I|<\varepsilon,$$

i.e., the Kurzweil Δ-integral exists and equals the Riemann Δ-integral. □

Instead of developing a theory of Kurzweil-Stieltjes Δ- and ∇-integrals, we show that they are in fact special cases of the Kurzweil-Stieltjes integral from Definition 6.1.2. This means that all basic properties of the Kurzweil-Stieltjes Δ- and ∇-integrals can be obtained as corollaries of the results from Chapter 6.

First of all, we describe two possible ways of extending a function defined on the time scale interval $[a,b]_{\mathbb{T}}$ to the full interval $[a,b]$. For each $t \in [a,b]$, let

$$t^* = \inf\{s \in [a,b]_{\mathbb{T}} : s \geq t\},$$
$$t_* = \sup\{s \in [a,b]_{\mathbb{T}} : s \leq t\}.$$

Note that if $t \in [a,b]_{\mathbb{T}}$, then $t^* = t_* = t$; otherwise, t_* and t^* are elements of $[a,b]_{\mathbb{T}}$ satisfying $t_* < t < t^*$. Now, for an arbitrary function $g : [a,b]_{\mathbb{T}} \to \mathbb{R}$, we define the extension $g^* : [a,b] \to \mathbb{R}$ by

$$g^*(t) = g(t^*) \quad \text{for all } t \in [a,b], \tag{8.5.2}$$

and the extension $g_* : [a,b] \to \mathbb{R}$ by

$$g_*(t) = g(t_*) \quad \text{for all } t \in [a,b]. \tag{8.5.3}$$

These two extensions appear in the statement of the next result.

8.5.8 Theorem. *Consider a pair of functions $f, g : [a,b]_{\mathbb{T}} \to \mathbb{R}$. Let $\tilde{f} : [a,b] \to \mathbb{R}$ be an arbitrary function such that $\tilde{f}(t) = f(t)$ for all $t \in [a,b]_{\mathbb{T}}$. Then the following statements hold:*

1. *The integral $\int_a^b f(t)\Delta g(t)$ exists if and only if the integral $\int_a^b \tilde{f}(t)\mathrm{d}g^*(t)$ exists; in this case, both integrals have the same value.*

2. *The integral $\int_a^b f(t)\nabla g(t)$ exists if and only if the integral $\int_a^b \tilde{f}(t)\mathrm{d}g_*(t)$ exists; in this case, both integrals have the same value.*

Proof. We prove only the first statement; the proof of the second one is similar, and is left to the reader. We will repeatedly use the fact that if $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ is a partition of $[a,b]_{\mathbb{T}}$, then $\boldsymbol{\alpha}, \boldsymbol{\xi} \subset \mathbb{T}$, and therefore

$$\begin{aligned} S(\tilde{f}, \mathrm{d}g^*, P) &= \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \tilde{f}(\xi_j)(g^*(\alpha_j) - g^*(\alpha_{j-1})) \\ &= \sum_{j=1}^{\nu(\boldsymbol{\alpha})} f(\xi_j)(g(\alpha_j) - g(\alpha_{j-1})) = S(f, \mathrm{d}g, P). \end{aligned}$$

Suppose first that $\int_a^b \tilde{f}(t)\mathrm{d}g^*(t)$ exists. Given an arbitrary $\varepsilon > 0$, there is a gauge $\tilde{\delta} : [a,b] \to (0,\infty)$ such that $|S(\tilde{f}, \mathrm{d}g^*, P) - \int_a^b \tilde{f}(t)\mathrm{d}g^*(t)| < \varepsilon$ for each

$\tilde{\delta}$-fine partition P of $[a,b]$. Now, let $\delta = (\delta_L, \delta_R)$ be a Δ-gauge on $[a,b]_{\mathbb{T}}$ given by $\delta_L(t) = \tilde{\delta}(t)/2$ and $\delta_R(t) = \max(\tilde{\delta}(t)/2, \mu(t))$ for all $t \in [a,b]_{\mathbb{T}}$. Let $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ be an arbitrary δ-fine partition of $[a,b]_{\mathbb{T}}$. Note that P need not be $\tilde{\delta}$-fine: there might exist $i \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$ such that $\alpha_i \geq \xi_i + \tilde{\delta}(\xi_i)$. In this case, we necessarily have $\delta_R(\xi_i) = \mu(\xi_i) > 0$, $\alpha_i = \sigma(\xi_i)$, and therefore $g^*(t) = \alpha_i$ for all $t \in (\xi_i, \alpha_i]$. After replacing each such interval-point pair $([\alpha_{i-1}, \alpha_i], \xi_i)$ by the interval-point pair $([\alpha_{i-1}, \xi_i + \tilde{\delta}(\xi_i)/2], \xi_i)$ and an arbitrary $\tilde{\delta}$-fine partition of $[\xi_i + \tilde{\delta}(\xi_i)/2, \alpha_i]$, we get a partition P' of $[a,b]$, which is $\tilde{\delta}$-fine and satisfies $S(\tilde{f}, \mathrm{d}g^*, P') = S(\tilde{f}, \mathrm{d}g^*, P)$. Hence, we have

$$\left| S(f, \mathrm{d}g, P) - \int_a^b \tilde{f}(t)\mathrm{d}g^*(t) \right| = \left| S(\tilde{f}, \mathrm{d}g^*, P') - \int_a^b \tilde{f}(t)\mathrm{d}g^*(t) \right| < \varepsilon,$$

which proves that $\int_a^b f(t)\Delta g(t)$ exists and equals $\int_a^b \tilde{f}(t)\mathrm{d}g^*(t)$.

Conversely, assume that $\int_a^b f(t)\Delta g(t)$ exists. Given an arbitrary $\varepsilon > 0$, there is a Δ-gauge $\delta = (\delta_L, \delta_R)$ on $[a,b]_{\mathbb{T}}$ such that $|S(f, \mathrm{d}g, P) - \int_a^b f(t)\Delta g(t)| < \varepsilon$ for each δ-fine partition P of $[a,b]_{\mathbb{T}}$. Now, let $\tilde{\delta}: [a,b] \to (0,\infty)$ be a gauge given by

$$\tilde{\delta}(t) = \begin{cases} \min(\delta_L(t), \sup[t, t+\delta_R(t)]_{\mathbb{T}} - t) & \text{if } t \in (a,b)_{\mathbb{T}}, \\ \sup[a, a+\delta_R(a)]_{\mathbb{T}} - a & \text{if } t = a, \\ \delta_L(b) & \text{if } t = b, \\ \operatorname{dist}(t, [a,b]_{\mathbb{T}}) & \text{if } t \in [a,b] \setminus \mathbb{T}. \end{cases}$$

Let $P = (\boldsymbol{\alpha}, \boldsymbol{\xi})$ be an arbitrary $\tilde{\delta}$-fine partition of $[a,b]$. Note that P need not be a partition of $[a,b]_{\mathbb{T}}$ (i.e., the division points and tags need not be elements of $\mathbb{T}$). However, if we show that there exists a δ-fine partition P' of $[a,b]_{\mathbb{T}}$ such that $S(f, \mathrm{d}g, P') = S(\tilde{f}, \mathrm{d}g^*, P)$, then

$$\left| S(\tilde{f}, \mathrm{d}g^*, P) - \int_a^b f(t)\Delta g(t) \right| = \left| S(f, \mathrm{d}g, P) - \int_a^b f(t)\Delta g(t) \right| < \varepsilon,$$

and the proof will be complete.

Thus, suppose that P is not a partition of $[a,b]_{\mathbb{T}}$ (otherwise take $P' = P$). Since $\alpha_0 = a \in \mathbb{T}$, there exists the smallest $i \in \{1,\dots,\nu(\boldsymbol{\alpha})\}$ such that $\alpha_{i-1} \in \mathbb{T}$, and at least one of α_i, ξ_i is not an element of $\mathbb{T}$. However, since the definition of $\tilde{\delta}$ guarantees that we have either $\xi_i \in \mathbb{T}$, or $\xi_i \notin \mathbb{T}$ and $[\alpha_{i-1}, \alpha_i] \cap \mathbb{T} = \emptyset$, we necessarily have $\xi_i \in [a,b)_{\mathbb{T}}$ and $\alpha_i \notin \mathbb{T}$. We now modify the partition P as follows:

- Replace the interval-point pair $([\alpha_{i-1}, \alpha_i], \xi_i)$ by $([\alpha_{i-1}, \alpha_i^*], \xi_i)$.

- Remove all interval-point pairs $([\alpha_{j-1},\alpha_j],\xi_j)$ such that $[\alpha_{j-1},\alpha_j]\subset[\alpha_i,\alpha_i^*]$.
- If there is an interval-point pair $([\alpha_{j-1},\alpha_j],\xi_j)$ such that $\alpha_i\le\alpha_{j-1}<\alpha_i^*<\alpha_j$, replace it by $([\alpha_i^*,\alpha_j],\xi_j)$.

Denote the partition obtained in this way by $\tilde{P}$. Because g^* is constant on $[\alpha_i,\alpha_i^*]$, it is clear that $S(\tilde{f},\mathrm{d}g^*,P')=S(\tilde{f},\mathrm{d}g^*,P)$. Since

$$\alpha_i<\xi_i+\tilde{\delta}(\xi_i)\le\sup[\xi_i,\xi_i+\delta_R(\xi_i)]_{\mathbb{T}},$$

it follows from the definition of α_i^* that $\alpha_i^*\le\sup[\xi_i,\xi_i+\delta_R(\xi_i)]_{\mathbb{T}}$, and consequently

$$[\alpha_{i-1},\alpha_i^*]\subset[\xi_i-\delta_L(\xi_i),\xi_i+\delta_R(\xi_i)].$$

If necessary, we can repeat the procedure we have just described with P replaced by $\tilde{P}$. The procedure does not increase the total number of division points, and decreases the number of division points that are not elements of $\mathbb{T}$. Thus, after finitely many steps, we obtain the desired δ-fine partition P' of $[a,b]_{\mathbb{T}}$ such that $S(f,\mathrm{d}g,P')=S(\tilde{f},\mathrm{d}g^*,P')=S(\tilde{f},\mathrm{d}g^*,P)$, and the proof is complete. □

8.5.9 Corollary. *Consider a function $f:[a,b]_{\mathbb{T}}\to\mathbb{R}$. Let $\tilde{f}:[a,b]\to\mathbb{R}$ be an arbitrary function such that $\tilde{f}(t)=f(t)$ for all $t\in[a,b]_{\mathbb{T}}$. Also, let $g(t)=t^*$ and $h(t)=t_*$ for all $t\in[a,b]$. Then the following statements hold:*

1. *The integral $\int_a^b f(t)\Delta t$ exists if and only if the integral $\int_a^b \tilde{f}(t)\mathrm{d}g(t)$ exists; in this case, both integrals have the same value.*
2. *The integral $\int_a^b f(t)\nabla t$ exists if and only if the integral $\int_a^b \tilde{f}(t)\mathrm{d}h(t)$ exists; in this case, both integrals have the same value.*

Bibliographic remarks. More information about the time scale calculus can be found e.g. in [14, 57]. The basic reference on Riemann Δ- and ∇-integrals (as well as Lebesgue integrals on time scales) is [15]. The definition presented there is slighly different from our Definition 8.5.3: The conditions $\xi_i\in[\alpha_{i-1},\rho(\alpha_i)]_{\mathbb{T}}$ and $\xi_i\in[\sigma(\alpha_{i-1}),\alpha_i]_{\mathbb{T}}$ are replaced by $\xi_i\in[\alpha_{i-1},\alpha_i)_{\mathbb{T}}$ and $\xi_i\in(\alpha_{i-1},\alpha_i]_{\mathbb{T}}$, respectively. Nevertheless, the two definitions are equivalent; see [15, Remark 5.17].

Kurzweil Δ- and ∇-integrals were introduced in [108], and subsequently discussed in [141]. In [108], the definition of a Δ-gauge and a ∇-gauge is slightly less restrictive as it allows the possibility that $\delta_L(a)=0$ and $\delta_R(b)=0$. However, one can easily see that this modification has no influence on the definition of a δ-fine partition of $[a,b]_{\mathbb{T}}$.

Riemann-Stieltjes Δ- and ∇-integrals are treated in [105]. As far as we are aware, there is no literature dealing with the Kurzweil-Stieltjes Δ- and ∇-integrals. Special cases of Corollary 8.5.9 with $\tilde{f}(t)=f^*(t)$ were obtained in [38, 134].

8.6 Dynamic equations on time scales

The time scale calculus introduced in the previous section makes it possible to unify the theories of differential and difference equation by considering the so-called dynamic equations, where classical derivatives or differences are replaced by Δ- or ∇-derivatives. The general form of a Δ-dynamic equation is

$$x^{\Delta}(t) = f(x(t), t), \tag{8.6.1}$$

while ∇-dynamic equations have the form

$$x^{\nabla}(t) = f(x(t), t), \tag{8.6.2}$$

where in both cases $x : \mathbb{T} \to \mathbb{R}^n$ and $f : \mathbb{R}^n \times \mathbb{T} \to \mathbb{R}^n$. Integration leads to the equations

$$x(t) = x(t_0) + \int_{t_0}^{t} f(x(s), s)\Delta s, \tag{8.6.3}$$

$$x(t) = x(t_0) + \int_{t_0}^{t} f(x(s), s)\nabla s, \tag{8.6.4}$$

where $t_0, t \in \mathbb{T}$, and the integrals on the right-hand sides are the Kurzweil Δ- and ∇-integrals introduced in the previous section.

Equations (8.6.3) and (8.6.4) are the main subject of this section. Since the indefinite Kurzweil integrals are always continuous, each solution x of (8.6.3) or (8.6.4) necessarily has the same property. Moreover, if $s \mapsto f(x(s), s)$ is rd-continuous, then the integral on the right-hand side of (8.6.3) exists as a Riemann integral, it is Δ-differentiable, and the equation reduces back to (8.6.1); similarly, if $s \mapsto f(x(s), s)$ is ld-continuous, then the integral on the right-hand side of (8.6.4) is ∇-differentiable, and we get (8.6.2). Nevertheless, we focus on the more general integral equations (8.6.3) and (8.6.4) without imposing any continuity conditions on f.

Since we know that Kurzweil Δ- and ∇-integrals can be rewritten as Kurzweil-Stieltjes integrals, it comes as no surprise that equations (8.6.3) and (8.6.4) are equivalent to certain Kurzweil-Stieltjes integral equations. This relation is described in the next theorem, where we use the notation (8.5.2) and (8.5.3) from the previous section.

8.6.1 Theorem. *Suppose that $a, b, t_0 \in \mathbb{T}$, $a \le t_0 \le b$, and consider a function $f : B \times [a,b]_{\mathbb{T}} \to \mathbb{R}^n$, where $B \subset \mathbb{R}^n$. Let $\tilde{f} : B \times [a,b] \to \mathbb{R}^n$ be an arbitrary function such that $\tilde{f}(x,t) = f(x,t)$ for all $x \in B$, $t \in [a,b]_{\mathbb{T}}$. Also, let $g(t) = t^*$ and $h(t) = t_*$ for all $t \in [a,b]$. Then the following statements hold:*

1. *If a function* $x:[a,b]_{\mathbb{T}}\to B$ *satisfies*

$$x(t)=x(t_0)+\int_{t_0}^{t} f(x(s),s)\Delta s, \quad t\in[a,b]_{\mathbb{T}}, \tag{8.6.5}$$

then the function $y:[a,b]\to B$ *given by* $y=x^*$ *satisfies*

$$y(t)=y(t_0)+\int_{t_0}^{t} \tilde{f}(y(s),s)\,\mathrm{d}g(s), \quad t\in[a,b]. \tag{8.6.6}$$

Conversely, each function $y:[a,b]\to B$ *satisfying* (8.6.6) *has the form* $y=x^*$*, where* $x:[a,b]_{\mathbb{T}}\to B$ *satisfies* (8.6.5).

2. *If a function* $x:[a,b]_{\mathbb{T}}\to B$ *satisfies*

$$x(t)=x(t_0)+\int_{t_0}^{t} f(x(s),s)\nabla s, \quad t\in[a,b]_{\mathbb{T}}, \tag{8.6.7}$$

then the function $y:[a,b]\to B$ *given by* $y=x_*$ *satisfies*

$$y(t)=y(t_0)+\int_{t_0}^{t} \tilde{f}(y(s),s)\,\mathrm{d}h(s), \quad t\in[a,b]. \tag{8.6.8}$$

Conversely, each function $y:[a,b]\to B$ *satisfying* (8.6.8) *has the form* $y=x_*$*, where* $x:[a,b]_{\mathbb{T}}\to B$ *satisfies* (8.6.7).

Proof. Suppose that $x:[a,b]_{\mathbb{T}}\to B$ satisfies (8.6.5) and $y=x^*$. For each $t\in[a,b]$, Corollary 8.5.9 implies

$$\begin{aligned} y(t)=x(t^*)&=x(t_0)+\int_{t_0}^{t^*} f(x(s),s)\Delta s \\ &=x^*(t_0)+\int_{t_0}^{t^*} \tilde{f}(x^*(s),s)\,\mathrm{d}g(s) \\ &=y(t_0)+\int_{t_0}^{t^*} \tilde{f}(y(s),s)\,\mathrm{d}g(s). \end{aligned}$$

Since the function g is constant on $[t,t^*]$, we have $\int_t^{t^*} \tilde{f}(y(s),s)\,\mathrm{d}g(s)=0$, and therefore (8.6.6) holds.

Conversely, assume that $y:[a,b]\to B$ satisfies (8.6.6). Since the function g is constant on each interval $[t,t^*]$ with $t\in[a,b]$, it follows that y has the same property. Hence, $y=x^*$, where $x:[a,b]_{\mathbb{T}}\to B$ is the restriction of y to $[a,b]_{\mathbb{T}}$.

For each $t \in [a,b]_{\mathbb{T}}$, Corollary 8.5.9 implies

$$\begin{aligned} x(t) = y(t) = y(t_0) + \int_{t_0}^{t} \tilde{f}(y(s),s)\,\mathrm{d}g(s) \\ = x^*(t_0) + \int_{t_0}^{t} \tilde{f}(x^*(s),s)\,\mathrm{d}g(s) \\ = x(t_0) + \int_{t_0}^{t} f(x(s),s)\Delta s. \end{aligned}$$

The proof of the second statement is similar and is left to the reader. □

A consequence of the previous theorem is that all results on generalized differential equations are directly applicable to dynamic equations on time scales. We illustrate this fact in the context of linear equations, and show how the theory developed in Chapter 7 leads to results on dynamic equations.

Let $a : [\alpha,\beta]_{\mathbb{T}} \to \mathscr{L}(\mathbb{R}^n)$, $b : [\alpha,\beta]_{\mathbb{T}} \to \mathbb{R}^n$ be arbitrary functions. If we let $f(x,t) = a(t)x + b(t)$ for all $x \in \mathbb{R}^n$ and $t \in [\alpha,\beta]_{\mathbb{T}}$, equations (8.6.5) and (8.6.7) become

$$x(t) = x(t_0) + \int_{t_0}^{t} (a(s)x(s) + b(s))\Delta s, \quad t \in [\alpha,\beta]_{\mathbb{T}} \tag{8.6.9}$$

$$x(t) = x(t_0) + \int_{t_0}^{t} (a(s)x(s) + b(s))\nabla s, \quad t \in [\alpha,\beta]_{\mathbb{T}}, \tag{8.6.10}$$

Thus, if $\tilde{a} : [\alpha,\beta] \to \mathscr{L}(\mathbb{R}^n)$ and $\tilde{b} : [\alpha,\beta] \to \mathbb{R}^n$ are arbitrary functions such that $\tilde{a}(t) = a(t)$ and $\tilde{b}(t) = b(t)$ for all $t \in [\alpha,\beta]_{\mathbb{T}}$, Theorem 8.6.1 implies that equations (8.6.9), (8.6.10) are equivalent to the Kurzweil-Stieltjes integral equations

$$y(t) = y(t_0) + \int_{t_0}^{t} (\tilde{a}(s)y(s) + \tilde{b}(s))\mathrm{d}g(s), \quad t \in [\alpha,\beta], \tag{8.6.11}$$

$$y(t) = y(t_0) + \int_{t_0}^{t} (\tilde{a}(s)y(s) + \tilde{b}(s))\mathrm{d}h(s), \quad t \in [\alpha,\beta], \tag{8.6.12}$$

with $g(t) = t^*$ and $h(t) = t_*$ for all $t \in [\alpha,\beta]$. Using Theorem 6.6.1 (substitution theorem), we see that the last two equations can be rewritten as generalized linear differential equations

$$y(t) = y(t_0) + \int_{t_0}^{t} \mathrm{d}[A_1(s)]y(s) + B_1(t) - B_1(t_0), \quad t \in [\alpha,\beta], \tag{8.6.13}$$

$$y(t) = y(t_0) + \int_{t_0}^{t} \mathrm{d}[A_2(s)]y(s) + B_2(t) - B_2(t_0), \quad t \in [\alpha,\beta], \tag{8.6.14}$$

where $A_1, A_2 : [\alpha, \beta] \to \mathscr{L}(\mathbb{R}^n)$ and $B_1, B_2 : [\alpha, \beta] \to \mathbb{R}^n$ are given by

$$A_1(t) = \int_{t_0}^{t} \tilde{a}(s)\,\mathrm{d}g(s), \quad B_1(t) = \int_{t_0}^{t} \tilde{b}(s)\,\mathrm{d}g(s), \tag{8.6.15}$$

$$A_2(t) = \int_{t_0}^{t} \tilde{a}(s)\,\mathrm{d}h(s), \quad B_2(t) = \int_{t_0}^{t} \tilde{b}(s)\,\mathrm{d}h(s) \tag{8.6.16}$$

for all $t \in [\alpha, \beta]$.

These considerations lead to the following existence and uniqueness theorem for linear dynamic equations.

8.6.2 Theorem. *The following statements hold:*

1. *Suppose that* $a : [\alpha, \beta]_{\mathbb{T}} \to \mathscr{L}(\mathbb{R}^n)$, $b : [\alpha, \beta]_{\mathbb{T}} \to \mathbb{R}^n$ *are Kurzweil* Δ*-integrable,* $I + a(t)\mu(t)$ *is invertible for each* $t \in [\alpha, t_0)_{\mathbb{T}}$*, and there exists a Kurzweil* Δ*-integrable function* $m : [\alpha, \beta]_{\mathbb{T}} \to \mathbb{R}$ *such that*

$$\left| \int_u^v a(s)\Delta s \right| \leq \int_u^v m(s)\Delta s \quad \textit{whenever } u, v \in [\alpha, \beta]_{\mathbb{T}},\ u \leq v. \tag{8.6.17}$$

Then equation (8.6.9) *has a unique solution* $x : [\alpha, \beta]_{\mathbb{T}} \to \mathbb{R}^n$.

2. *Suppose that* $a : [\alpha, \beta]_{\mathbb{T}} \to \mathscr{L}(\mathbb{R}^n)$, $b : [\alpha, \beta]_{\mathbb{T}} \to \mathbb{R}^n$ *are Kurzweil* ∇*-integrable,* $I - a(t)\nu(t)$ *is invertible for each* $t \in (t_0, \beta]_{\mathbb{T}}$*, and there exists a Kurzweil* ∇*-integrable function* $m : [\alpha, \beta]_{\mathbb{T}} \to \mathbb{R}$ *such that*

$$\left| \int_u^v a(s)\nabla s \right| \leq \int_u^v m(s)\nabla s \quad \textit{whenever } u, v \in [\alpha, \beta]_{\mathbb{T}},\ u \leq v. \tag{8.6.18}$$

Then equation (8.6.10) *has a unique solution* $x : [\alpha, \beta]_{\mathbb{T}} \to \mathbb{R}^n$.

Proof. We prove the first statement; the proof of the second one is similar and is left to the reader.

Let $A_1 : [\alpha, \beta] \to \mathscr{L}(\mathbb{R}^n)$, $B_1 : [\alpha, \beta] \to \mathbb{R}^n$ be given by (8.6.15). Note that the function $g(t) = t^*$ is regulated (because it is nondecreasing) and left-continuous. According to Corollary 6.5.4, A_1 and B_1 are regulated on $[\alpha, \beta]$. For all $t \in (t_0, \beta]$, the matrix $I - \Delta^- A_1(t) = I - \tilde{a}(t)\Delta^- g(t) = I$ is invertible. If $t \in [\alpha, t_0)_{\mathbb{T}}$, then $I + \Delta^+ A_1(t) = I + \tilde{a}(t)\Delta^+ g(t) = I + a(t)\mu(t)$. On the other hand, if $t \in [\alpha, t_0) \setminus \mathbb{T}$, then $I + \Delta^+ A_1(t) = I + \tilde{a}(t)\Delta^+ g(t) = I$. Thus, $I + \Delta^+ A_1(t)$ is invertible for all $t \in [\alpha, t_0)$.

Note that $A_1(t^*) = A_1(t)$ for each $t \in [\alpha, \beta]$, since g is constant on $[t, t^*]$. Hence, if $\boldsymbol{\alpha}$ is an arbitrary division of $[\alpha, \beta]$, we get

$$V(A_1, \alpha) = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} |A_1(\alpha_j) - A_1(\alpha_{j-1})| = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} |A_1(\alpha_j^*) - A_1(\alpha_{j-1}^*)|$$
$$= \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \left| \int_{\alpha_{j-1}^*}^{\alpha_j^*} \tilde{a}(s)\, \mathrm{d}g(s) \right| = \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \left| \int_{\alpha_{j-1}^*}^{\alpha_j^*} a(s)\Delta s \right|$$
$$\leq \sum_{j=1}^{\nu(\boldsymbol{\alpha})} \int_{\alpha_{j-1}^*}^{\alpha_j^*} m(s)\Delta s = \int_{\alpha}^{\beta} m(s)\Delta s,$$

which shows that A_1 has bounded variation on $[\alpha, \beta]$.

Thus, all assumptions of Theorem 7.6.2 are satisfied, and equation (8.6.13) has a unique solution $y: [\alpha, \beta] \to \mathbb{R}^n$. By Theorem 8.6.1, equation (8.6.9) has a unique solution $x: [\alpha, \beta]_{\mathbb{T}} \to \mathbb{R}^n$, which is obtained as the restriction of y to $[\alpha, \beta]_{\mathbb{T}}$. □

8.6.3 Remark. In the theory of dynamic equations, the requirement that $I + a(t)\mu(t)$ is invertible for each $t \in [\alpha, t_0)_{\mathbb{T}}$ is known as μ-regressivity (or simply regressivity), while the condition that $I - a(t)\nu(t)$ is invertible for each $t \in (t_0, \beta]_{\mathbb{T}}$ is called ν-regressivity. For Δ-dynamic equations, it is usually supposed (see [14, Chapter 5]) that a and b are rd-continuous functions, while the common assumption for ∇-equations is ld-continuity of a and b (see [3]). Under these assumptions, both (8.6.17) and (8.6.18) hold, because a is bounded. Moreover, equations (8.6.9) and (8.6.10) reduce to $x^\Delta(t) = a(t)x(t) + b(t)$ and $x^\nabla(t) = a(t)x(t) + b(t)$, respectively. Our Theorem 8.6.2 is more general – no continuity or boundedness of a or b is assumed.

The next result is a continuous dependence theorem for solutions of linear Δ-dynamic equations.

8.6.4 Theorem. *Let $a: [\alpha, \beta]_{\mathbb{T}} \to \mathscr{L}(\mathbb{R}^n)$, $b: [\alpha, \beta]_{\mathbb{T}} \to \mathbb{R}^n$, $a_k: [\alpha, \beta]_{\mathbb{T}} \to \mathscr{L}(\mathbb{R}^n)$, $b_k: [\alpha, \beta]_{\mathbb{T}} \to \mathbb{R}^n$, $k \in \mathbb{N}$, be Kurzweil Δ-integrable. Suppose that $I + a(t)\mu(t)$ is invertible for each $t \in [a, t_0)_{\mathbb{T}}$, and there exists a Kurzweil Δ-integrable function $m: [\alpha, \beta]_{\mathbb{T}} \to \mathbb{R}$ such that*

$$\left| \int_u^v a_k(s)\Delta s \right| \leq \int_u^v m(s)\Delta s \quad \textit{for } u, v \in [\alpha, \beta]_{\mathbb{T}},\ u \leq v,\ k \in \mathbb{N}. \tag{8.6.19}$$

Furthermore, assume that

$$\lim_{k\to\infty}\sup_{t\in[\alpha,\beta]_{\mathbb{T}}}\left|\int_{\alpha}^{t}a_k(s)\Delta s-\int_{\alpha}^{t}a(s)\Delta s\right|=0, \tag{8.6.20}$$

$$\lim_{k\to\infty}\sup_{t\in[\alpha,\beta]_{\mathbb{T}}}\left|\int_{\alpha}^{t}b_k(s)\Delta s-\int_{\alpha}^{t}b(s)\Delta s\right|=0. \tag{8.6.21}$$

Finally, let $x_k^0\in\mathbb{R}^n$, $k\in\mathbb{N}$, *be a sequence satisfying*

$$\lim_{k\to\infty}x_k^0=x^0.$$

Then there exists a $k_0\in\mathbb{N}$ *such that the equations*

$$x_k(t)=x_k^0+\int_{t_0}^{t}(a_k(s)x_k(s)+b_k(s))\Delta s,\quad t\in[\alpha,\beta]_{\mathbb{T}},\quad k\geq k_0, \tag{8.6.22}$$

have unique solutions $x_k:[\alpha,\beta]_{\mathbb{T}}\to\mathbb{R}^n$. *Moreover,* $x_k\rightrightarrows x$, *where* $x:[\alpha,\beta]_{\mathbb{T}}\to\mathbb{R}^n$ *is the unique solution of the equation*

$$x(t)=x^0+\int_{\alpha}^{t}(a(s)x(s)+b(s))\Delta s,\quad t\in[\alpha,\beta]_{\mathbb{T}}. \tag{8.6.23}$$

Proof. Consider the functions

$$A(t)=\int_{t_0}^{t}\tilde{a}(s)\,\mathrm{d}g(s),\quad B(t)=\int_{t_0}^{t}\tilde{b}(s)\,\mathrm{d}g(s),$$

$$A_k(t)=\int_{t_0}^{t}\tilde{a}_k(s)\,\mathrm{d}g(s),\quad B_k(t)=\int_{t_0}^{t}\tilde{b}_k(s)\,\mathrm{d}g(s),\quad k\in\mathbb{N},$$

where $\tilde{a}:[\alpha,\beta]\to\mathscr{L}(\mathbb{R}^n)$, $\tilde{b}:[\alpha,\beta]\to\mathscr{L}(\mathbb{R}^n)$, $\tilde{a}_k:[\alpha,\beta]\to\mathscr{L}(\mathbb{R}^n)$, $\tilde{b}_k:[\alpha,\beta]\to\mathbb{R}^n$, $k\in\mathbb{N}$, are arbitrary functions satisfying $\tilde{a}(t)=a(t)$, $\tilde{b}(t)=b(t)$, $\tilde{a}_k(t)=a_k(t)$, $\tilde{b}_k(t)=b_k(t)$ for all $t\in[\alpha,\beta]_{\mathbb{T}}$. As in the proof of Theorem 8.6.2, one can show that A, A_k, B, B_k are regulated, left-continuous, $I+\Delta^+A(t)$ is invertible for each $t\in[a,t_0)_{\mathbb{T}}$, and

$$\operatorname{var}_{\alpha}^{\beta}A_k\leq\int_{\alpha}^{\beta}m(s)\Delta s,\quad k\in\mathbb{N}.$$

Corollary 8.5.9 and the assumptions (8.6.20) and (8.6.21) imply that $A_k\rightrightarrows A$ and $B_k\rightrightarrows B$ on $[\alpha,\beta]_{\mathbb{T}}$. Since A, A_k, B, B_k are constant on each interval $[t,t^*]$ with $t\in[\alpha,\beta]$, we conclude that $A_k\rightrightarrows A$ and $B_k\rightrightarrows B$ on $[\alpha,\beta]$.

Thus, by Theorem 7.6.6, there exists a $k_0\in\mathbb{N}$ such that for every $k\geq k_0$, the equation

$$y_k(t)=x_k^0+\int_{t_0}^{t}\mathrm{d}[A_k(s)]y_k(s)+B_k(t)-B_k(t_0),\quad t\in[\alpha,\beta], \tag{8.6.24}$$

has a unique solution $y_k : [a,b] \to \mathbb{R}^n$. Moreover, $y_k \rightrightarrows y$, where $y : [\alpha,\beta] \to \mathbb{R}^n$ is the unique solution of the equation

$$y(t) = x^0 + \int_{t_0}^{t} \mathrm{d}[A(s)]y(s) + B(t) - B(t_0), \quad t \in [\alpha,\beta].$$

By Theorem 8.6.1, the restrictions of y and y_k to $[\alpha,\beta]_{\mathbb{T}}$ are the unique solutions of the equations (8.6.23) and (8.6.22), respectively; this completes the proof. □

8.6.5 Exercise. Formulate and verify the counterpart of Theorem 8.6.4 for linear ∇-dynamic equations.

Bibliographic remarks. A special case of the relation between dynamic equations on time scales and Kurzweil-Stieltjes integral equations was described in [134]. Kurzweil-Stieltjes integral equations having the form

$$y(t) = y(t_0) + \int_{t_0}^{t} f(y(s),s)\,\mathrm{d}g(s),$$

where g is a nondecreasing functions, are sometimes called *measure differential equations*. They include not only dynamic equations on time scales, but also differential equations with impulses; see e.g. [38, 101]. Basic results concerning the existence, uniqueness and continuous dependence of solutions to nonlinear measure differential equations are available in [101, 135]. The applicability of the theory of linear generalized differential equations in the context of linear dynamic equations was demonstrated in [103], which contains a special case of Theorem 8.6.4. The concept of *Stieltjes differential equations* is closely related to measure differential equations, and is based on Stieltjes derivatives instead of Stieltjes integrals; more details can be found in [40].

References

[1] S. AFONSO, E. M. BONOTTO, M. FEDERSON AND Š. SCHWABIK. Discontinuous local semiflows for Kurzweil equations leading to LaSalle's invariance principle for non-autonomous systems with impulses. *J. Differential Equations* **250** (2011) 2969–3001.

[2] F. ALBIAC AND N. J. KALTON. *Topics in Banach space theory*. Springer-Verlag, Berlin-Heidelberg-New York, 2006.

[3] D. ANDERSON, J. BULLOCK, L. ERBE, A. PETERSON AND H. TRAN. Nabla Dynamic Equations. In M. Bohner, A. Peterson (eds.), *Advances in Dynamic Equations on Time Scales*, Birkhäuser, Boston, 2003, 47–83.

[4] T. M. APOSTOL. *Mathematical Analysis* (2nd ed.). Addison-Wesley, Reading, MA, 1974.

[5] J. APPELL, J. BANÁS AND N. MERENTES. *Bounded variation and around*. De Gruyter, Berlin, 2014.

[6] M. ASHORDIA. On the correctness of linear boundary value problems for systems of generalized ordinary differential equations. *Proc. Georgian Acad. Sci. Math.* **1** (1993), 385–394.

[7] M. ASHORDIA. Criteria of correctness of linear boundary value problems for systems of generalized ordinary differential equations. *Czechoslovak Math. J.* **46 (121)** (1996), 385–404.

[8] M. ASHORDIA. On the Opial type criterion for the well-posedness of the Cauchy problem for linear systems of generalized ordinary differential equations. *Math. Bohem.* **141** (2016), 183–215.

[9] L. BARBANTI. Simply regulated functions and semivariation in uniformly convex spaces. *Real Anal. Exchange* **24** (1) (1998/9), 405–410.

[10] R. G. BARTLE. A convergence theorem for generalized Riemann integrals. *Real Anal. Exchange* **20** (1994-95), 119–124.

[11] R. G. BARTLE. *A modern theory of integration*. Graduate Studies in Mathematics 32, AMS, 2001.

[12] R. G. BARTLE AND D. R. SHERBERT. *Introduction to Real Analysis*. Wiley, New York, 2000.

[13] O. Blasco, P. Gregori and J. M. Calabuig. Finite semivariation and regulated functions by means of bilinear maps. *Real Anal. Exchange* **26** (2) (2000/01), 603–608.

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

[15] M. Bohner and G. Guseinov. Riemann and Lebesgue Integration. In M. Bohner, A. Peterson (eds.), *Advances in Dynamic Equations on Time Scales*, Birkhäuser, Boston, 2003, 117–163.

[16] D. M. Bressoud. *A radical approach to Lebesgue's theory of integration.* Cambridge University Press, 2008.

[17] M. Brokate and P. Krejčí. Duality in the space of regulated functions and the play operator. *Mathematische Zeitschrift* **245** (2003) 667–688.

[18] H. Chen. A pair of adjoint classes of Riemann-Stieltjes integrable functions. *Real Anal. Exchange* **23** (1) (1998/99), 235–240.

[19] H. Chen. Adjoint classes of generalized-Stieltjes integrable functions. *Real Anal. Exchange* **24** (1) (1998/99), 139–148.

[20] H. Chen. Adjoint classes of Lebesgue-Stieltjes integrable functions. *Real Anal. Exchange* **26** (1) (2000/01), 421–428.

[21] V. V. Chistyakov. On mappings of bounded variation. *J. Dynam. Control Systems* **3** (1997), 261–289.

[22] V. V. Chistyakov. On the theory of set-valued maps of bounded variation of one real variable. *Sbornik: Mathematics* **189** (1998), 153–176.

[23] V. V. Chistyakov and O. E. Galkin. Mappings of bounded Φ-variation with arbitrary function Φ. *J. Dynam. Control Systems* **4** (1998), 217–247.

[24] C. Debiéve. Measures with finite semi-variation. *Bull. Belg. Math. Soc.* **4** (1997), 293–298.

[25] J. Diestel. *Sequences and Series in Banach Spaces*. Springer-Verlag, Berlin-Heidelberg-New York, 1984.

[26] J. Diestel and J. J. Uhl. *Vector Measures*. Math. Surveys, Vol. 15, Amer. Math. Soc., Providence, Rhode Island, 1977.

[27] M. Duchoň. A Helly type theorem in the setting of Banach spaces. *Int. J. Pure Appl. Math.* **66** (3) (2011), 263–273.

[28] M. DUCHOŇ AND P. VADOVIC. Functions with values in locally convex spaces with weakly compact semivariation. *Tatra Mt. Math. Publ.* **52** (2012), 133–139.

[29] R. M. DUDLEY AND R. NORVAIŠA. *Concrete Functional Calculus*. Springer-Verlag, Berlin-Heidelberg-New York, 2011.

[30] N. DUNFORD AND J. T. SCHWARTZ. *Linear Operators* I, II. Interscience Publishers, New York & London, 1958 and 1963.

[31] B. DUSHNIK. *On the Stieltjes integral*. University of Michigan, dissertation, 1931.

[32] A. DVORETZKY AND C. A. ROGERS. Absolute and unconditional convergence in normed linear spaces. *Proc. Nat. Acad. Sci. U. S. A.* **36** (1950), 192–197.

[33] W. F. EBERLEIN. Notes on integration I: the underlying convergence theorem. *Comm. Pure Appl. Math.* **10** (1957) 357–360.

[34] J. R. EDWARDS AND S. G. WAYMENT. An integral representation for linear operators on a space of continuous vector-valued functions. *Proc. Amer. Math. Soc.* **30** (1971) 260–262.

[35] M. FEDERSON. Substitution formulas for the Kurzweil and Henstock vector integrals. *Math. Bohem.* **127** (2002), 15–26.

[36] M. FEDERSON AND R. BIANCONI. A Fredholm-type theorem for linear integral equations of Stieltjes type. *Integration: Mathematical Theory and Applications* **1** (2), (2008), 135–168.

[37] M. FEDERSON, J. G. MESQUITA AND A. SLAVÍK. Measure functional differential and functional dynamic equations on time scales. *J. Differerential Equations* **252** (2012), 3816–3847.

[38] M. FEDERSON, J. G. MESQUITA AND A. SLAVÍK. Basic results for functional differential and dynamic equations involving impulses. *Math. Nachr.* **286** (2013), 181–204.

[39] D. FRAŇKOVÁ. Regulated functions. *Mathematica Bohemica* **116** (1991) 20–59.

[40] M. FRIGON AND R. L. POUSO. Theory and applications of first-order systems of Stieltjes differential equations. *Adv. Nonlinear Anal.* **6** (2017) 13–36.

[41] K.M. GARG. *Relativization of some aspects of the theory of functions of bounded variation*. Dissertationes Math. (Rozprawy Mat.) **320** (1992), 123 pp.

[42] C. GOFFMAN, G. MORAN AND D. WATERMAN. The structure of regulated functions. *Proc. Amer. Math. Soc.* **57** (1976), 61–65.

[43] R. A. GORDON. *The integrals of Lebesgue, Denjoy, Perron, and Henstock.* American Mathematical Society, 1994.

[44] M. GOWURIN. Über die Stieltjes Integration abstrakter Funktionen. *Fund. Math.* **27** (1936), 255–268.

[45] J. GROH. A nonlinear Volterra-Stieltjes integral equation and a Gronwal inequality in one dimension. *Illinois J. Math.* **24** (1980), 244–263.

[46] Z. HALAS, G. A. MONTEIRO AND M. TVRDÝ. Emphatic convergence and sequential solutions of generalized linear differential equations. *Mem. Differential Equations Math. Phys.* **54** (2011), 27–49.

[47] I. HALPERIN. *Introduction to the Theory of Distributions*. University of Toronto Press, Toronto, 1952.

[48] U. M. HANUNG, G. A. MONTEIRO AND M. TVRDÝ. Bounded convergence theorem for abstract Kurzweil–Stieltjes integral. *Monatshefte für Mathematik* **180** (2016) 409–434.

[49] R. HENSTOCK. *Lectures on the Theory of Integration*. World Scientific, Singapore, 1988.

[50] R. HENSTOCK. *The General Theory of Integration*. Clarendon Press, Oxford, 1991.

[51] E. HEWITT AND K. STROMBERG. *Real and Abstract Analysis: A modern treatment of the theory of functions of a real variable*. Springer-Verlag, Berlin-Heidelberg-New York, 1975 (Third printing).

[52] T. H. HILDEBRANDT. On integrals related to and extensions of the Lebesgue integrals. *Bull. Amer. Math. Soc.* (2) **24** (1917) 113–144; (1918) 177–202.

[53] T. H. HILDEBRANDT. Definitions of Stieltjes Integrals of the Riemann Type. *Amer. Math. Monthly* **45** (1938) 265–278.

[54] T. H. HILDEBRANDT. On systems of differentio-Stieltjes-integral equations. *Illinois J. Math.* **3** (1959) 352–373.

[55] T. H. HILDEBRANDT. *Theory of Integration*. Academic Press, New York & London, 1963.

[56] T. H. HILDEBRANDT. Linear Continuous Functionals on the Space (BV) with Weak Topologies. *Proc. Amer. Math. Soc.* **17** (1966) 658–664.

[57] S. HILGER. Analysis on measure chains - a unified approach to continuous and discrete calculus. *Results Math.* bf 18 (1990), 18–56.

[58] HILLE, E. AND PHILIPS, R. S.. *Functional Analysis and Semi-Groups*. American Mathematical Society, Providence, Rhode Island, 1974.

[59] CH. S. HÖNIG. *The abstract Riemann-Stieltjes integral and its applications to linear differential equations with generalized boundary conditions*. Notas do Instituto de Matemática e Estatistica da Universidade de S. Paulo, Série de Matemática, no. 1, 1973.

[60] CH. S. HÖNIG. *Volterra Stieltjes-Integral Equations*. North Holland & American Elsevier, Mathematics Studies 16, Amsterdam & New York, 1975.

[61] CH. S. HÖNIG. Volterra-Stieltjes integral equations. In: *Functional Differential Equations and Bifurcation, Proceedings of the Saõ Carlos Conference 1979* (Lecture Notes in Mathematics 799, Springer-Verlag, Berlin-Heidelberg-New York, 1980), pp. 173–216.

[62] CH. S. HÖNIG. There is no natural Banach space norm on the space of Kurzweil-Henstock-Denjoy-Perron integrable functions. *Seminário Brasileiro de Análise* **30** (1989), 387–397.

[63] J. JARNÍK. On some assumptions of the theorem on the continuous dependence on a parameter. *Čas. Pěst. Mat.* **86** (1961), 404–414.

[64] J. JARNÍK. On a certain modification of the theorem on the continuous dependence on a parameter. *Čas. Pěst. Mat.* **86** (1961), 415–424.

[65] J. JARNÍK. Dependence of solutions of a class of differential equations of the second order on a parameter. *Czechoslovak Math. J.* **15 (90)** (1965), 124–160.

[66] M. I. KADETS AND V. M. KADETS. *Series in Banach Spaces. Conditional and Unconditional Convergence*. Birkhäuser Verlag, Basel, 1997.

[67] H. S. KALTENBORN. Linear functional operations on functions having discontinuities of the first kind. *Bull. Amer. Math. Soc.* **40** (1934) 702–708.

[68] P. L. KAPITZA. Pendulum with vibrating suspension (in Russian). *Uspechi fiz. nauk* **XLIV**, 1 (1951), 7–20.

[69] P. L. KAPITZA. Dynamic stability of a pendulum with an oscilatting point of suspension (in Russian). *Zhurnal eksperimental'noi i teoreticheskoi fiziki*. **21** 5 (1951), 588–597. English transl. in:*Collected Papers by P. L. Kapitza*, vol. 2, 714–726, Pergamon Press, London, 1965.

[70] A. N. KOLMOGOROV AND S. V. FOMIN. *Elements of the Theory of Functions and Functional Analysis* [Two Volumes in One]. Martino Fine Books, United States, 2012).

[71] P. KREJČÍ. The Kurzweil integral with exclusion of negligible sets. *Mathematica Bohemica* **128** (2003) 277-292.

[72] P. KREJČÍ. The Kurzweil integral and hysteresis. In: *Proceedings of the International Workshop on Multi-Rate Processes and Hysteresis* (Cork, 3.4.2006–8.4.2006, eds: M. Mortell, R. O'Malley, A. Pokrovskii, V. Sobolev). *Journal of Physics: Conference Series* **55**, (2006) 144–154.

[73] P. KREJČÍ, J KURZWEIL. A nonexistence result for the Kurzweil integral. *Mathematica Bohemica* **127** (2002) 571–580.

[74] P. KREJČÍ, H. LAMBA, S. MELNIK AND D. RACHINSKII. Kurzweil integral representation of interacting Prandtl-Ishlinskii operators. *Discrete and Continuous Dynamical Systems B* **20** (2015), 2949–2965.

[75] P. KREJČÍ, H. LAMBA, G. A. MONTEIRO AND D. RACHINSKII The Kurzweil integral in financial market modeling. *Mathematica Bohemica* **141** 2016, 261–286.

[76] P. KREJČÍ, PH. LAURENÇOT. Generalized variational inequalities. *Journal of Convex Analysis* **9** (2002), 159-–183.

[77] P. KREJČÍ, M. LIERO. Rate independent Kurzweil processes. *Applications of Mathematics* **54** (2009) 117–145.

[78] J. KURZWEIL. Generalized ordinary differential equation and continuous dependence on a parameter. *Czechoslovak Mathematical Journal* **7(82)** (1957), 418–449.

[79] J. KURZWEIL. Generalized ordinary differential equations. *Czechoslovak Mathematical Journal* **8(83)** (1958) 360–387.

[80] J. KURZWEIL. *On integration by parts.* Czechoslovak Math. J., 8 (82) (1958), 356–359.

[81] J. KURZWEIL. *Nichtabsolut konvergente Integrale.* Teubner-Verlag, Leipzig, 1980.

[82] J. KURZWEIL. *Ordinary differential equations. (Introduction to the theory of ordinary differential equations in the real domain.)* Studies in Applied Mechanics, 13. Elsevier Scientific Publishing Co., Amsterdam, 1986.

[83] J. KURZWEIL. *Henstock-Kurzweil Integration:Its Relation to Topological Vector Spaces.* World Scientific, Singapore, 2000.

[84] J. KURZWEIL. *Integration Between the Lebesgue Integral and the Henstock-Kurzweil Integral: Its Relation to Local Convex Vector Spaces.* World Scientific, Singapore, 2002.

[85] J. KURZWEIL. Generalized ordinary differential equations (Not Absolutely Continuous Solutions). Series in Real Analysis–Vol. 11, World Scientific, Singapore, 2012.

[86] H. LEBESGUE. Sur l'intégrale de Stieltjes et sur les opérations functionnelles linéaires. *Comptes Rendus* **150** (1910), 86–88.

[87] P. Y LEE. *Lanzhou Lectures on Henstock Integration.* World Scientific, Singapore, 1989.

[88] P. Y LEE. AND R. VÝBORNÝ. *The Integral:An Easy Approach after Kurzweil and Henstock.* Cambridge Univ. Press, Cambridge, 2000.

[89] R. LEŚNIEWICZ AND W. ORLICZ. On generalized variations II. *Studia Math.* **45.1** (1973), 71–109
[http://matwbn.icm.edu.pl/ksiazki/sm/sm45/sm4518.pdf].

[90] J. W. LEWIN. A Truly Elementary Approach to the Bounded Convergence Theorem. *Amer. Math. Monthly* **93** (1986), 395–397.

[91] S. ŁOJASIEWICZ. Sur un effet asymptotique dans les équations différentielles dont les seconds membres contiennent des termes périodiques de pulsation et d'amplitude tendant a l'infini. *Ann. Pol. Math.* **1** (1955), 388–413.

[92] J. LUKEŠ, J. MALÝ. *Measure and Integral*. matfyzpress, Praha, 1995
[http://www.mff.cuni.cz/to.en/fakulta/mfp/download/books/lukes-maly_-_measure_and_integral.pdf].

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

[94] J. MAWHIN. L'eternal retour des sommes de Riemann-Stieltjes dans l'evolution du calcul intégral. *Bulletin de la Soc. Royale des Sciences de Liège* **70** (2001) (4-5-6) 345–364.

[95] R. M. MCLEOD. *The generalized Riemann integral*. Carus Monograph, No.2, Mathematical Association of America, Washington, 1980.

[96] J. S. MACNERNEY. An integration by parts formula. *Bull. Amer. Math. Soc.* **69** (1963) 803–805.

[97] G. A. MONTEIRO. On functions of bounded semivariation. *Real Anal. Exchange* **40** (2), 2015, 233-276.

[98] G. A. MONTEIRO. Adjoint classes of Kurzweil-Stieltjes integrable functions. *Monatshefte für Mathematik* (2017) [https://doi.org/10.1007/s00605-017-1071-9].

[99] G. A. MONTEIRO AND B. SATCO. Distributional, differential and integral problems: equivalence and existence results. EJQTDE, 2017 (7) 1-–26; [http://www.math.u-szeged.hu/ejqtde/]

[100] G. A. MONTEIRO AND A. SLAVÍK. Linear measure functional differential equations with infinite delay. *Math. Nachr.* **287** (2014), 1363–1382.

[101] G. A. MONTEIRO AND A. SLAVÍK. Extremal solutions of measure differential equations. *J. Math. Anal. Appl.* **444** (2016), 568–597.

[102] G. A. MONTEIRO AND M. TVRDÝ. On Kurzweil-Stieltjes integral in Banach space, *Mathematica Bohemica* **137** (2012), 365–381.

[103] G. A. MONTEIRO AND M. TVRDÝ. Generalized linear differential equations in a Banach space: Continuous dependence on a parameter. *Discrete Contin. Dyn. Syst.* **33** (2013) 283–303.

[104] G. A. MONTEIRO AND M. TVRDÝ. Continuous dependence of solutions of abstract generalized linear differential equations with potential converging uniformly with a weight. *Boundary Value Problems* **2014** (71), 2014, 1–18.
[http://www.boundaryvalueproblems.com/content/2014/1/71].

[105] D. MOZYRSKA, E. PAWLUSZEWICZ, D. F. M. TORRES. The Riemann-Stieltjes integral on time scales. *Aust. J. Math. Anal. Appl.* **7** (2010), 1–14.

[106] J. MUSIELAK AND W. ORLICZ. On generalized variations I. *Studia Math.* **18** (1959), 11-41,
[http://matwbn.icm.edu.pl/ksiazki/sm/sm18/sm1812.pdf].

[107] S. PERLMAN. *Functions of generalized variation*. *Fundamenta Mathematicae* **105** (1980), 199–211,
[http://matwbn.icm.edu.pl/ksiazki/fm/fm105/fm105117.pdf].

[108] A. PETERSON AND B. THOMPSON. Henstock-Kurzweil delta and nabla integrals. *J. Math. Anal. Appl.* **323** (2006), 162–178.

[109] O. PERRON. *Die Lehre von den Kettenbrüchen*. Teubner, Leipzig, 1913.

[110] W. F. PFEFFER. *The Riemann Approach to Integration:Local geometric theory.* Cambridge University Press, 1993.

[111] I. K. RANA, *An introduction to measure and integration* (second edition). American Mathematical Society, 2002.

[112] F. RIESZ. Sur les opérations functionelles linéaires. *Comptes Rendus Acad. Sci. Paris* **149** (1909), 974–977.

[113] H. L. ROYDEN. *Real Analysis*. McMillan, New York - London, 1963.

[114] W. RUDIN. *Real and Complex Analysis*. McGraw-Hill, New York, 1966.

[115] W. RUDIN. *Functional Analysis*. McGraw-Hill Series in Higher Mathematics, New York, 1973.

[116] S. SAKS. *Théorie de l'Intégrale*. Monografie Matematyczne, Warszawa, Lwòw, 1933.
English translation: *Theory of the Integral*. Monografie Matematyczne, Warszawa, Lwów, 1937.

[117] E. SCHECHTER. *Handbook of Analysis and its Foundations*. Academic Press, San Diego, 1997.

[118] M. SCHECHTER. *Principles of Functional Analysis*. Academic Press, New York & London, 1973.

[119] Š. SCHWABIK. Verallgemeinerte lineare Differentialgleichungssysteme. *Čas. pěst. mat.* **96** (1971) 183–211.

[120] Š. SCHWABIK. On the relation between Young's and Kurzweil's concept of Stieltjes integral. *Čas. pěst. mat.* **98** (1973) 237–251.

[121] Š. SCHWABIK. On a modified sum integral of Stieltjes type. *Čas. pěst. mat.* **98** (1973) 274–277.

[122] Š. SCHWABIK. *Generalized Ordinary Differential Equations*. Series in Real Analysis–Vol. 5, World Scientific, Singapore, 1992.

[123] Š. SCHWABIK. Linear operator in the space of regulated functions. *Mathematica Bohemica* **117** (1992), 79–92.

[124] Š. SCHWABIK. Abstract Perron-Stieltjes integral. *Mathematica Bohemica* **121** (1996), 425–447.

[125] Š. SCHWABIK. *Integrace v $\mathbb{R}$ (Kurzweilova teorie)*. Karolinum, Univerzita Karlova v Praze, 1999.

[126] Š. SCHWABIK. Linear Stieltjes integral equations in Banach spaces. *Mathematica Bohemica* **124** (1999), 433–457.

[127] Š. SCHWABIK. Linear Stieltjes integral equations in Banach spaces II: operator valued solutions. *Mathematica Bohemica* **125** (2000), 431–454.

[128] Š. SCHWABIK. A note on integration by parts for abstract Perron-Stieltjes integrals. *Mathematica Bohemica* **126** (2001), 613–629.

[129] Š. SCHWABIK. Operator-valued functions of bounded semivariation and convolutions. *Mathematica Bohemica* **126** (2001), 745–777.

[130] Š. SCHWABIK AND P. ŠARMANOVÁ. *Malý průvodce historií integral*. Prometheus, Praha, 1996 [http://dml.cz/handle/10338.dmlcz/400862].

[131] Š. SCHWABIK, M. TVRDÝ AND O. VEJVODA. *Differential and Integral Equations: Boundary Value Problems and Adjoint.* Academia and Reidel. Praha and Dordrecht, 1979 [http://dml.cz/handle/10338.dmlcz/400391].

[132] Š. SCHWABIK AND G. YE. *Topics in Banach space Integration.* World Scientific, Singapore, 2005.

[133] W. SIERPIŃSKI. Sur une popriété des fonctions qui n'ont que des disconinuités de premiére espéce. *Bulletin de la Section Scientifique de L'Académie Roumaine* **16** (1933), 1–4.

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

[135] A. SLAVÍK. Well-posedness results for abstract generalized differential equations and measure functional differential equations. *J. Differential Equations* **259** (2015), 666–707.

[136] T. J. STIELTJES. *Recherches sur les fractions continues*. Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys. **8** (1894), J1–122, **9** (1895) A1–47. Reprinted in *Oeuvres II* (P. Noordhoff, Groningen, 1918), 402–566.
English translation: *Investigations on continued fractions*. In: T. J. Stieltjes,

Collected Papers Vol. II. (Springer-Verlag, Berlin-Heidelberg-New York, 1993), 609–745.

[137] C. SWARTZ. *Introduction to Gauge Integrals*. World Scientific, Singapore, 2001.

[138] P. SWOROWSKI. Adjoint classes of functions in the H_1 sense. *Czechoslovak Math. J.* **57(132)** (2007), 505–522.

[139] T. TAO. *An Introduction to Measure Theory.* Graduate Studies in Mathematics 126, AMS, 2011.

[140] A. E. TAYLOR. *Introduction to functional analysis.* John Wiley,New York, 1967.

[141] B. S. THOMSON. Henstock-Kurzweil integrals on time scales. *PanAmer. Math. J.* **18** (2008), 1–19.

[142] B. L. D. THORP. On the Equivalence of Certain Notions of Bounded Variation. *J. London Math. Soc.* **43** (1968), 247–252.

[143] W. F. TRENCH. *Introduction to real analysis.* Pearson Education, 2003.

[144] D. H. TUCKER. A representation theorem for a continuous linear transformation on a space of continuous functions. *Proc. Amer. Math. Soc.* **16** (1965) 946–953.

[145] M. TVRDÝ. Regulated functions and the Perron–Stieltjes integral. *Čas. Pěst.Mat.* **114** (1989) 187—209.

[146] M. TVRDÝ. Generalized differential equations in the space of regulated functions (Boundary value problems and controllability). *Math. Bohem.* **116** (3) (1991) 225–244.

[147] M. TVRDÝ. *Differential and Integral Equations in the Space of Regulated Functions.* Memoirs on Differential Equations and Mathematical Physics vol.25 (2002), pp. 1–104.

[148] M. TVRDÝ. *Stieltjesův integrál. Kurzweilova teorie.* Univerzita Palackého v Olomouci, Olomouc, 2012.

[149] C. J. DE LA VALLÉE POUSSIN. *Intégrales de Lebesgue, Fonctions d'Ensemble. Classes de Baire.* Paris, 1916.

[150] E. B. VAN VLECK. Haskin's momental theorem and its connection with Stieltjes's problem of moments. *Trans. Amer. Math. Soc* **18** (1917), 326–330.

[151] A. J. WARD. The Perron-Stieltjes integral. *Mathematische Zeitschrift* **41** (1936), 578–604.

[152] D. WATERMAN. On convergence of Fourier series of functions of generalized bounded variation. *Studia Math.* **44** (1972), 107-117, [http://matwbn.icm.edu.pl/ksiazki/sm/sm44/sm44111.pdf]

[153] J. YEH. *Real Analysis: Theory of Measure and Integration* (3rd edition). World Scientific Publishing, 2014.

[154] K. YOSIDA. *Functional Analysis*. Springer-Verlag, Berlin-Heidelberg-New York, 1978 (Fifth Edition).

[155] L. C. YOUNG. *The Theory of Integration*. Cambridge, 1927.

[156] L. C. YOUNG. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Mathematica* **67** (1936), 251–282.

[157] W. H. YOUNG. Integration with respect to a function of bounded variation. *Proc. London Math. Soc.* (2) **13** (1914), 109–150.