A NOTE ON INTEGRATION BY PARTS FOR ABSTRACT PERRON-STIELTJES INTEGRALS

Štefan Schwabik, Praha

(Received July 28, 1999)

Abstract. Integration by parts results concerning Stieltjes integrals for functions with values in Banach spaces are presented. The background of the theory is the Kurzweil approach to integration based on Riemann type integral sums, which leads to the Perron integral.

Keywords: integration by parts, Kurzweil-Stieltjes integral, Perron-Stieltjes integral

MSC 2000: 26A39

Bilinear triples

Assume that X, Y and Z are Banach spaces and that there is a bilinear mapping $B\colon\ X\times Y\to Z$. We use the short notation $xy=B(x,y)$ for the value of the bilinear form B for $x\in X$, $y\in Y$ and assume that

$$\|B(x,y)\|_Z=\|xy\|_Z\leqslant\|x\|_X\|y\|_Y.$$

By $\|\cdot\|_X$ the norm in the Banach space X is denoted (and similarly for the other ones).

Triples of Banach spaces X, Y, Z with these properties are called *bilinear triples* and are denoted by $\mathcal{B}=(X,Y,Z)$ or shortly $\mathcal{B}$.

This work was supported by the grant 201/97/0218 of the Grant Agency of the Czech Republic

Assume that $[a,b]\subset\mathbb{R}$ is a bounded interval and that X is a Banach space. Given $x\colon [a,b]\to X$, the function x is of *bounded variation on* $[a,b]$ if

$$\operatorname{var}_a^b(x)=\sup\Big\{\sum_{j=1}^{k}\|x(\alpha_j)-x(\alpha_{j-1})\|_X\Big\}<\infty,$$

where the supremum is taken over all finite partitions

$$D\colon\ a=\alpha_0<\alpha_1<\ldots<\alpha_{k-1}<\alpha_k=b$$

of the interval $[a,b]$. The set of all functions $x\colon [a,b]\to X$ with $\operatorname{var}_a^b(x)<\infty$ will be denoted by $BV([a,b],X)$ or shortly $BV([a,b])$ if it is clear which Banach space X we have in mind.

Assume now that $\mathcal{B}=(X,Y,Z)$ is a bilinear triple of Banach spaces.

For $x\colon [a,b]\to X$ and a partition D of the interval $[a,b]$ define

$$V_a^b(x,D)=\sup\Big\{\Big\|\sum_{j=1}^{k}[x(\alpha_j)-x(\alpha_{j-1})]y_j\Big\|_Z\Big\},$$

where the supremum is taken over all possible choices of $y_j\in Y$, $j=1,\ldots,k$ with $\|y_j\|\leqslant 1$, and set

$$(\mathcal{B})\operatorname{var}_a^b(x)=\sup V_a^b(x,D),$$

where the supremum is taken over all finite partitions

$$D\colon\ a=\alpha_0<\alpha_1<\ldots<\alpha_{k-1}<\alpha_k=b$$

of the interval $[a,b]$.

A function $x\colon [a,b]\to X$ with $(\mathcal{B})\operatorname{var}_a^b(x)<\infty$ is called a *function with bounded $\mathcal{B}$-variation on* $[a,b]$ (sometimes also *a function of bounded semi-variation* [2], [3]).

The set of all functions $x\colon [a,b]\to X$ with $(\mathcal{B})\operatorname{var}_a^b(x)<\infty$ will be denoted by $(\mathcal{B})BV([a,b],X)$ or shortly by $(\mathcal{B})BV([a,b])$ if it is clear which bilinear triple (X,Y,Z) we have in mind.

1. Proposition. *If $\mathcal{B}=(X,Y,Z)$ is a bilinear triple then*

$$BV([a,b],X)\subset(\mathcal{B})BV([a,b],X)\tag{1}$$

and if $x \in BV([a,b],X)$, *then*

$$(\mathcal{B})\mathrm{var}_a^b(x) \leqslant \mathrm{var}_a^b(x).$$

P r o o f. For a given function $x\colon [a,b] \to X$ with $x \in BV([a,b],X)$, a partition D of $[a,b]$ and arbitrary $y_j \in Y$, $j = 1,\dots,k$ with $\|y_j\| \leqslant 1$ we have

$$\begin{aligned}\Big\|\sum_{j=1}^{k}(x(\alpha_j) - x(\alpha_{j-1}))y_j\Big\|_Z &\leqslant \sum_{j=1}^{k}\|x(\alpha_j) - x(\alpha_{j-1})\|_X\|y_j\|_Y \\ &\leqslant \sum_{j=1}^{k}\|x(\alpha_j) - x(\alpha_{j-1})\|_X \leqslant \mathrm{var}_a^b(x).\end{aligned}$$

Passing to the suprema corresponding to the definition of $(\mathcal{B})\mathrm{var}_a^b(x)$ in this inequality we immediately obtain the inclusion as well as the inequality claimed in the statement. □

R e m a r k. It is easy to show that if $x\colon [a,b] \to \mathbb{R}$ and $\mathcal{B} = (\mathbb{R},\mathbb{R},\mathbb{R})$ with the multiplication of reals as the corresponding bilinear form, then $x \in (\mathcal{B})BV([a,b])$ if and only if $x \in BV([a,b])$.

Indeed, in this case we have

$$V_a^b(x,D) = \sup\Big\{\Big|\sum_{j=1}^{k}[x(\alpha_j) - x(\alpha_{j-1})]y_j\Big|\Big\} = \sum_{j=1}^{k}|x(\alpha_j) - x(\alpha_{j-1})|$$

because we can take $y_j = 1$ if $x(\alpha_j) - x(\alpha_{j-1}) \geqslant 0$ and $y_j = -1$ if $x(\alpha_j) - x(\alpha_{j-1}) < 0$.

The same is true also if $x\colon [a,b] \to X$ and $\mathcal{B} = (X,\mathbb{R},X)$, where the Banach space X is finite-dimensional.

This shows that the concept of $\mathcal{B}$-variation of a function $x\colon [a,b] \to X$ is relevant only for infinite-dimensional Banach spaces X.

Regulated functions and step functions with values in a Banach space

Assume that $[a,b] \subset \mathbb{R}$ is a bounded interval and that X is a Banach space. Given $x\colon [a,b] \to X$, the function x is called *regulated on* $[a,b]$ if it has one-sided limits at every point of $[a,b]$, i.e. if for every $s \in [a,b)$ there is a value $x(s+) \in X$ such that

$$\lim_{t\to s+}\|x(t) - x(s+)\|_X = 0$$

and if for every $s \in (a,b]$ there is a value $x(s-) \in X$ such that

$$\lim_{t\to s-} \|x(t) - x(s-)\|_X = 0.$$

The set of all regulated functions $x\colon [a,b] \to X$ will be denoted by $G([a,b],X)$ or shortly $G([a,b])$ if it is clear which Banach space X we have in mind.

If $C([a,b],X)$ is the set of continuous functions $x\colon [a,b] \to X$ then evidently

$$C([a,b],X) \subset G([a,b],X). \tag{2}$$

Assume now that $\mathcal{B} = (X,Y,Z)$ is a bilinear triple of Banach spaces.

A function $x\colon [a,b] \to X$ is called *$\mathcal{B}$-regulated on* $[a,b]$ if for every $y \in Y$, $\|y\|_Y \leqslant 1$ the function $xy\colon [a,b] \to Z$ given by $t \mapsto x(t)y \in Z$ for $t \in [a,b]$ is regulated, i.e. $xy \in G([a,b],Z)$ for every $y \in Y$, $\|y\|_Y \leqslant 1$.

Similarly $y\colon [a,b] \to Y$ is called *$\mathcal{B}$-regulated on* $[a,b]$ if for every $x \in X$, $\|x\|_X \leqslant 1$ the function $xy\colon [a,b] \to Z$ given by $t \mapsto xy(t) \in Z$ for $t \in [a,b]$ is regulated, i.e. $xy \in G([a,b],Z)$ for every $x \in X$, $\|x\|_X \leqslant 1$.

For a given bilinear triple $\mathcal{B} = (X,Y,Z)$ the set of all $\mathcal{B}$-regulated functions $x\colon [a,b] \to X$ will be denoted by $(\mathcal{B})G([a,b],X)$ or shortly by $(\mathcal{B})G([a,b])$ if it is clear which bilinear triple (X,Y,Z) we have in mind.

A function $x\colon [a,b] \to X$ is called a *(finite) step function on* $[a,b]$ if there exists a finite partition

$$D\colon\ a = \alpha_0 < \alpha_1 < \ldots < \alpha_{k-1} < \alpha_k = b$$

of the interval $[a,b]$ such that x has a constant value on (α_{j-1},α_j) for every $j = 1,\ldots,k$.

The following results are well known for regulated functions.

2. Proposition. *$x \in G([a,b],X)$ if and only if x is the uniform limit of step functions. (See e.g.* [2, Theorem 3.1, p. 16]*.)*

If $x \in G([a,b],X)$ then:

a) *x is bounded, i.e. there exists $K > 0$ such that $\|x(s)\|_X \leqslant K$ for every $s \in [a,b]$,*

b) *for every $\varepsilon > 0$ the sets*

$$\{s \in [a,b); \|x(s+) - x(s)\| \geqslant \varepsilon\}, \quad \{s \in (a,b]; \|x(s) - x(s-)\| \geqslant \varepsilon\}$$

are finite,

c) *the set*

$$S = \{s \in [a,b]; x(s) \neq x(s+) \ \text{ or } \ x(s) \neq x(s-)\}$$

is at most countable (see e.g. [2, Corollary 3.2, p. 17]*),*

d) *$G([a,b],X)$ equipped with the norm $\|x\|_{G([a,b],X)} = \sup_{s\in[a,b]} \|x(s)\|_X$ for $x \in G([a,b],X)$ is a Banach space.*

3. Proposition. *If $\mathcal{B} = (X,Y,Z)$ is a bilinear triple and $x \in G([a,b],X)$ then $x \in (\mathcal{B})G([a,b],X)$, i.e. $G([a,b],X) \subset (\mathcal{B})G([a,b],X)$.*

P r o o f. For any $y \in Y$ with $\|y\|_Y \leqslant 1$ and $s,t \in [a,b]$ we have

$$\|x(t)y - x(s)y\|_Z \leqslant \|x(t) - x(s)\|_X \|y\|_Y \leqslant \|x(t) - x(s)\|_X$$

and this implies the statement (e.g. by the Bolzano-Cauchy condition for the existence of onesided limits of the function x). □

In addition to this we also have

4. Proposition. *If $x \in BV([a,b],X)$ then $x \in G([a,b],X)$, i.e.*

$$BV([a,b],X) \subset G([a,b],X) \subset (\mathcal{B})G([a,b],X). \tag{3}$$

P r o o f. For $s,t \in [a,b]$, $s \leqslant t$ we have

$$\|x(t) - x(s)\|_X \leqslant \operatorname{var}_{[s,t]}(x) = \operatorname{var}_{[a,t]}(x) - \operatorname{var}_{[a,s]}(x)$$

and this implies (e.g. by the Bolzano-Cauchy condition for the existence of onesided limits of the nondecreasing bounded real function $\operatorname{var}_{[a,t]}(x)$) that the onesided limits of the function $x\colon [a,b] \to X$ exist at any point of $[a,b]$, i.e. that $A \in G([a,b],X)$. □

R e m a r k. If the Banach space X is finite dimensional, then it is easy to check that a function $x\colon [a,b] \to X$ is $\mathcal{B}$-regulated if and only if it is regulated.

Stieltjes integration of vector valued functions

A finite system of points

$$\{\alpha_0, \tau_1, \alpha_1, \tau_2, \ldots, \alpha_{k-1}, \tau_k, \alpha_k\}$$

such that

$$a = \alpha_0 < \alpha_1 < \ldots < \alpha_{k-1} < \alpha_k = b$$

and

$$\tau_j \in [\alpha_{j-1}, \alpha_j] \quad \text{for } j = 1, \ldots, k$$

is called a *P-partition* of the interval $[a,b]$.

A function $\delta\colon [a,b] \to (0,\infty)$ is called a *gauge on* $[a,b]$.

For a given gauge δ on $[a,b]$ a P-partition $\{\alpha_0, \tau_1, \alpha_1, \tau_2, \dots, \alpha_{k-1}, \tau_k, \alpha_k\}$ of $[a,b]$ is called *δ-fine* if

$$[\alpha_{j-1}, \alpha_j] \subset (\tau_j - \delta(\tau_j), \tau_j + \delta(\tau_j)) \quad \text{ for } j = 1, \dots, k.$$

5. Cousin's Lemma. *Given an arbitrary gauge δ on $[a,b]$ there is a δ-fine P-partition of $[a,b]$.*

(See e.g. [4] and many other books on Henstock-Kurzweil integration.)

6. Definition. Assume that $\mathcal{B} = (X,Y,Z)$ is a bilinear triple and that functions $x\colon [a,b] \to X$ and $y\colon [a,b] \to Y$ are given.

We say that the Stieltjes integral $\int_a^b \mathrm{d}[x(s)]y(s)$ exists if there is an element $I \in Z$ such that for every $\varepsilon > 0$ there is a gauge δ on $[a,b]$ such that for

$$S(\mathrm{d}x, y, D) = \sum_{j=1}^{k} \bigl[x(\alpha_j) - x(\alpha_{j-1})\bigr] y(\tau_j)$$

we have

$$\bigl\|S(\mathrm{d}x, y, D) - I\bigr\|_Z < \varepsilon$$

provided D is a δ-fine P-partition of $[a,b]$. We denote $I = \int_a^b \mathrm{d}[x(s)]y(s)$. For the case $a = b$ it is convenient to set $\int_a^b \mathrm{d}[x(s)]y(s) = 0$ and if $b < a$, then $\int_a^b \mathrm{d}[x(s)]y(s) = -\int_b^a \mathrm{d}[x(s)]y(s)$.

Similarly we can define the Stieltjes integral $\int_a^b x(s)\mathrm{d}[y(s)]$ using Stieltjes integral sums of the form

$$S(x, \mathrm{d}y, D) = \sum_{j=1}^{k} x(\tau_j)\bigl[y(\alpha_j) - y(\alpha_{j-1})\bigr].$$

R e m a r k . Note that Cousin's Lemma 5 is essential for this definition. The Stieltjes integral introduced in this way is determined uniquely and has all the necessary elementary properties, see [5].

7. Proposition. *Assume that $\mathcal{B} = (X,Y,Z)$ is a bilinear triple, that*

$$x \in (\mathcal{B})G([a,b], X) \cap (\mathcal{B})BV([a,b], X)$$

and $y \in G([a,b], Y)$.

Then the integral $\int_a^b \mathrm{d}[x(s)]y(s)$ exists.

Symmetrically, if $x \in G([a,b],X)$ and

$$y \in (\mathcal{B})G([a,b],Y) \cap (\mathcal{B})BV([a,b],Y)$$

then the integral $\int_a^b x(s)\mathrm{d}[y(s)]$ exists.

See [5, Proposition 15].

Taking into account Proposition 3 we obtain the following

8. Corollary. *If $\mathcal{B} = (X,Y,Z)$ is a bilinear triple such that*

$$x \in G([a,b],X) \cap (\mathcal{B})BV([a,b],X) \quad \text{and} \quad y \in G([a,b],Y) \cap (\mathcal{B})BV([a,b],Y)$$

then both integrals

$$\int_a^b \mathrm{d}[x(s)]y(s) \quad \text{and} \quad \int_a^b \mathrm{d}[x(s)]y(s)$$

exist.

Integration by parts

Assume that $\mathcal{B} = (X,Y,Z)$ is a bilinear triple and that $x\colon [a,b] \to X$, $y\colon [a,b] \to Y$. For a P-partition $D = \{\alpha_0, \tau_1, \alpha_1, \tau_2, \ldots, \alpha_{k-1}, \tau_k, \alpha_k\}$ of the interval $[a,b]$ define

$$\Delta(x,y,D) = \sum_{j=1}^{k} [(x(\alpha_j) - x(\tau_j))(y(\alpha_j) - y(\tau_j)) - (x(\alpha_{j-1}) - x(\tau_j))(y(\alpha_{j-1}) - y(\tau_j))].$$

9. Definition. We say that $\Delta_a^b(x,y)$ exists if there is an element $J \in Z$ such that for every $\varepsilon > 0$ there is a gauge δ on $[a,b]$ such that

$$\|\Delta(x,y,D) - J\|_Z < \varepsilon$$

if D is a δ-fine P-partition of $[a,b]$. We then denote $J = \Delta_a^b(x,y)$.

Remark. The definition of the quantity $\Delta_a^b(x,y)$ is an integral-like definition when compared with the Definition 6.

The basic result is the following.

10. Theorem (Integration by parts)**.** *Assume that $\mathcal{B} = (X,Y,Z)$ is a bilinear triple and that $x\colon [a,b] \to X$, $y\colon [a,b] \to Y$.*

If two of the quantities

$$\int_a^b \mathrm{d}[x(s)]y(s), \quad \int_a^b x(s)\mathrm{d}[y(s)], \quad \Delta_a^b(x,y)$$

exist then the third exists as well and the equality

$$\int_a^b \mathrm{d}[x(s)]y(s) + \int_a^b x(s)\mathrm{d}[y(s)] = x(b)y(b) - x(a)y(a) - \Delta_a^b(x,y) \tag{4}$$

holds.

P r o o f. First of all let us show that for every P-partition

$$D = \{\alpha_0, \tau_1, \alpha_1, \tau_2, \ldots, \alpha_{k-1}, \tau_k, \alpha_k\}$$

of the interval $[a,b]$ we have

$$S(\mathrm{d}x, y, D) + S(x, \mathrm{d}y, D) + \Delta(x,y,D) = x(b)y(b) - x(a)y(a). \tag{5}$$

Indeed, by a simple algebraic manipulation we have

$$\begin{aligned}
&\sum_{j=1}^k [x(\alpha_j) - x(\alpha_{j-1})]y(\tau_j) + \sum_{j=1}^k x(\tau_j)[y(\alpha_j) - y(\alpha_{j-1})] \\
&\quad + \sum_{j=1}^k [(x(\alpha_j) - x(\tau_j))(y(\alpha_j) - y(\tau_j)) - (x(\alpha_{j-1}) - x(\tau_j))(y(\alpha_{j-1}) - y(\tau_j))] \\
&= \sum_{j=1}^k [x(\alpha_j)y(\tau_j) - x(\alpha_{j-1})y(\tau_j) + x(\tau_j)y(\alpha_j) - x(\tau_j)y(\alpha_{j-1}) + x(\alpha_j)y(\alpha_j) \\
&\quad - x(\tau_j)y(\alpha_j) - x(\alpha_j)y(\tau_j) + x(\tau_j)y(\tau_j) - x(\alpha_{j-1})y(\alpha_{j-1}) + x(\tau_j)y(\alpha_{j-1}) \\
&\quad + x(\alpha_{j-1})y(\tau_j) - x(\tau_j)y(\tau_j)] \\
&= \sum_{j=1}^k [x(\alpha_j)y(\alpha_j) - x(\alpha_{j-1})y(\alpha_{j-1})] \\
&= x(\alpha_k)y(\alpha_k) - x(\alpha_0)y(\alpha_0) = x(b)y(b) - x(a)y(a)
\end{aligned}$$

because $\alpha_k = b$ and $\alpha_0 = a$ for the P-partition D.

Suppose e.g. that the integrals $\int_a^b \mathrm{d}[x(s)]y(s)$, $\int_a^b x(s)\mathrm{d}[y(s)]$ exist. Then by their definition for every $\varepsilon > 0$ there is a gauge δ on $[a,b]$ such that for any δ-fine P-partition D of $[a,b]$ we have

$$\left\| \int_a^b \mathrm{d}[x(s)]y(s) - S(\mathrm{d}x, y, D) \right\|_Z < \frac{\varepsilon}{2} \tag{6}$$

and

$$\left\| \int_a^b x(s)\mathrm{d}[y(s)] - S(x, \mathrm{d}y, D) \right\|_Z < \frac{\varepsilon}{2}. \tag{7}$$

Then for any δ-fine P-partition D of $[a, b]$ we have by (5), (6) and (7)

$$\begin{aligned}
&\left\| \Delta_a^b(x, y) - x(b)y(b) + x(a)y(a) + \int_a^b \mathrm{d}[x(s)]y(s) + \int_a^b x(s)\mathrm{d}[y(s)] \right\|_Z \\
&\leqslant \|\Delta_a^b(x, y) + S(\mathrm{d}x, y, D) + S(x, \mathrm{d}y, D) - x(b)y(b) + x(a)y(a)\|_Z \\
&\quad + \left\| \int_a^b \mathrm{d}[x(s)]y(s) - S(\mathrm{d}x, y, D) \right\|_Z + \left\| \int_a^b x(s)\mathrm{d}[y(s)] - S(x, \mathrm{d}y, D) \right\|_Z \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

and this inequality shows that by definition $\Delta_a^b(x, y)$ exists and its value is

$$\Delta_a^b(x, y) = x(b)y(b) - x(a)y(a) - \int_a^b \mathrm{d}[x(s)]y(s) - \int_a^b x(s)\mathrm{d}[y(s)],$$

i.e. that (4) is satisfied.

The remaining cases when $\Delta_a^b(x, y)$, $\int_a^b \mathrm{d}[x(s)]y(s)$ or $\Delta_a^b(x, y)$, $\int_a^b x(s)\mathrm{d}[y(s)]$ exist can be proved similarly. $\square$

R e m a r k. The proof of Theorem is based on purely algebraic manipulation of integral sums for the integral. This approach to integration by parts goes back to the paper [3] of J. Kurzweil.

If $x \in G([a, b], X)$ then define

$$\Delta^+ x(\tau) = x(\tau+) - x(\tau) = \lim_{\sigma \to \tau+} x(\sigma) - x(\tau)$$

and

$$\Delta^- x(\tau) = x(\tau) - x(\tau-) = x(\tau) - \lim_{\sigma \to \tau-} x(\sigma).$$

Now our aim is to give some corollaries to Theorem 10 which will present the Stieltjes form of integration by parts formula in a more conventional form.

If $\mathcal{B} = (X, Y, Z)$ is a bilinear triple such that $x \in G([a, b], X) \cap (\mathcal{B})BV([a, b], X)$ and $y \in G([a, b], Y) \cap (\mathcal{B})BV([a, b], Y)$ then both integrals

$$\int_a^b \mathrm{d}[x(s)]y(s) \quad \text{and} \quad \int_a^b \mathrm{d}[x(s)]y(s)$$

exist as was stated in Corollary 8.

First of all let us prove some auxiliary statements.

11. Lemma. *If $\mathcal{B} = (X, Y, Z)$ is a bilinear triple,*

$$x \in G([a,b],X) \cap (\mathcal{B})BV([a,b],X) \text{ and } y \in G([a,b],Y)$$

or

$$x \in G([a,b],X) \text{ and } y \in G([a,b],Y) \cap (\mathcal{B})BV([a,b],Y)$$

then the series

$$\sum_{\tau\in[a,b)} \Delta^+ x(\tau)\Delta^+ y(\tau), \qquad \sum_{\tau\in(a,b]} \Delta^- x(\tau)\Delta^- y(\tau)$$

converge in Z.

P r o o f. Let us consider the first possibility.

Since $x \in G([a,b],X)$ the set S of $\tau \in [a,b)$ for which $\Delta^+ x(\tau) \neq 0$ is at most countable by c) in Proposition 2, i.e. $S = \{\sigma_k \in [a,b); k \in \mathbb{N}\}$ and therefore we can write

$$\sum_{\tau\in[a,b)} \Delta^+ x(\tau)\Delta^+ y(\tau) = \sum_{k=1}^{\infty} \Delta^+ x(\sigma_k)\Delta^+ y(\sigma_k).$$

Denote

$$\sum_{k=1}^{m} \Delta^+ x(\sigma_k)\Delta^+ y(\sigma_k) = S_m \in Z$$

for $m \in \mathbb{N}$ and assume that $\varepsilon > 0$ is given.

Since the sets

$$\{s \in [a,b); \|\Delta^+ x(s)\| = \|x(s+) - x(s)\| \geqslant \varepsilon\},$$
$$\{s \in [a,b); \|\Delta^+ y(s)\| = \|y(s+) - y(s)\| \geqslant \varepsilon\}$$

are finite by b) from Proposition 2, there exists $M_\varepsilon \in \mathbb{N}$ such that for $k \in \mathbb{N}$, $k > M_\varepsilon$ we have $\|\Delta^+ x(\sigma_k)\| < \varepsilon$ and $\|\Delta^+ y(\sigma_k)\| < \varepsilon$

Assume that $m > n > M_\varepsilon$, $m, n \in \mathbb{N}$. Then

$$S_m - S_n = \sum_{k=n+1}^{m} \Delta^+ x(\sigma_k)\Delta^+ y(\sigma_k).$$

Since the limits $x(\sigma_k+) \in X$ exist for every $k = n+1, \ldots, m$, there exist values $a_k \in [a,b)$, $k = n+1, \ldots, m$ such that

$$[\sigma_k, \sigma_k + a_k] \cap \{\sigma_{n+1}, \ldots, \sigma_m\} = \{\sigma_k\}$$

and

$$\|x(\sigma_k + a_k) - x(\sigma_k+)\| < \frac{\varepsilon}{m-n}.$$

Using this we get

$$\begin{aligned}
\|S_m - S_n\|_Z &= \Big\| \sum_{k=n+1}^{m} \Delta^+ x(\sigma_k)\Delta^+ y(\sigma_k) \Big\|_Z \\
&\leqslant \Big\| \sum_{k=n+1}^{m} [x(\sigma_k+) - x(\sigma_k + a_k)]\Delta^+ y(\sigma_k) \Big\|_Z \\
&\quad + \Big\| \sum_{k=n+1}^{m} [x(\sigma_k + a_k) - x(\sigma_k)]\Delta^+ y(\sigma_k) \Big\|_Z \\
&\leqslant \varepsilon \Big\| \sum_{k=n+1}^{m} [x(\sigma_k + a_k) - x(\sigma_k)] \frac{\Delta^+ y(\sigma_k)}{\varepsilon} \Big\|_Z \\
&\quad + \sum_{k=n+1}^{m} \|x(\sigma_k+) - x(\sigma_k + a_k)\|_X \|\Delta^+ y(\sigma_k)\|_Y \\
&< \varepsilon(\mathcal{B})\mathrm{var}_a^b x + \varepsilon \sum_{k=n+1}^{m} \frac{\varepsilon}{m-n} \\
&= \varepsilon(\mathcal{B})\mathrm{var}_a^b x + \varepsilon^2 \frac{m-n}{m-n} = \varepsilon(\mathcal{B})\mathrm{var}_a^b x + \varepsilon^2.
\end{aligned}$$

Hence $S_m \in Z$, $m \in \mathbb{N}$ is a Cauchy sequence in the Banach space Z and therefore the series $\sum\limits_{k=1}^{\infty} \Delta^+ x(\sigma_k)\Delta^+ y(\sigma_k) = \sum\limits_{\tau \in [a,b)} \Delta^+ x(\tau)\Delta^+ y(\tau)$ converges in Z.

The convergence of $\sum\limits_{\tau \in (a,b]} \Delta^- x(\tau)\Delta^- y(\tau)$ can be shown analogously.

The second possibility when $x \in G([a,b], X)$ and $y \in G([a,b], Y) \cap (\mathcal{B})BV([a,b], Y)$ is symmetric and can be treated in the same way as the former. $\square$

12. Lemma. *If $\mathcal{B} = (X, Y, Z)$ is a bilinear triple, $x \in G([a,b], X) \cap (\mathcal{B})BV([a,b], X)$ and $y \in G([a,b], Y) \cap (\mathcal{B})BV([a,b], Y)$ then*

$$\Delta_a^b(x,y) = \sum_{\tau \in [a,b)} \Delta^+ x(\tau)\Delta^+ y(\tau) - \sum_{\tau \in (a,b]} \Delta^- x(\tau)\Delta^- y(\tau). \tag{8}$$

P r o o f. By Corollary 8 both the integrals $\int_a^b x(s)\,\mathrm{d}[y(s)]$, $\int_a^b \mathrm{d}[x(s)]y(s)$ exist and therefore by Theorem 10 $\Delta_a^b(x,y) \in Z$ also exists.

Since $x \in G([a,b],X)$ the set S of $\tau \in [a,b)$ for which $\Delta^+ x(\tau) \neq 0$ or $\Delta^- x(\tau) \neq 0$ is at most countable by c) in Proposition 2, i.e. $S = \{\sigma_k \in [a,b); k \in \mathbb{N}\}$ and

$$\begin{aligned}(9)\qquad & \sum_{\tau\in[a,b)} \Delta^+ x(\tau)\Delta^+ y(\tau) - \sum_{\tau\in(a,b]} \Delta^- x(\tau)\Delta^- y(\tau) \\ & \quad = \sum_{k=1}^{\infty} \Delta^+ x(\sigma_k)\Delta^+ y(\sigma_k) - \sum_{k=1}^{\infty} \Delta^- x(\sigma_k)\Delta^- y(\sigma_k).\end{aligned}$$

Assume that $\varepsilon > 0$ is given.

Since the series $\sum\limits_{k=1}^{\infty} \Delta^+ x(\sigma_k)\Delta^+ y(\sigma_k)$, $\sum\limits_{k=1}^{\infty} \Delta^- x(\sigma_k)\Delta^- y(\sigma_k)$ converge in Z by Lemma 11, there exists $L_\varepsilon \in \mathbb{N}$ such that

$$(10)\qquad \Big\| \sum_{k=L_\varepsilon+1}^{\infty} \Delta^+ x(\sigma_k)\Delta^+ y(\sigma_k) \Big\|_Z < \varepsilon, \quad \Big\| \sum_{k=L_\varepsilon+1}^{\infty} \Delta^- x(\sigma_k)\Delta^- y(\sigma_k) \Big\|_Z < \varepsilon.$$

By Definition 9 there exists a gauge δ_0 on $[a,b]$ such that

$$(11)\qquad \|\Delta_a^b(x,y) - \Delta(x,y,D)\|_Z < \varepsilon$$

for any δ_0-fine P-partition D of $[a,b]$.

Further, for every $\tau \in [a,b]$ there is $\delta_1(\tau) > 0$ such that

$$(12)\qquad \begin{aligned} &|x(s) - x(\tau+)| < \varepsilon, |y(s) - y(\tau+)| < \varepsilon \quad \text{for } s \in (\tau, \tau + \delta_1(\tau)), \\ &|x(s) - x(\tau-)| < \varepsilon, |y(s) - y(\tau-)| < \varepsilon \quad \text{for } s \in (\tau - \delta_1(\tau), \tau). \end{aligned}$$

This is clear because x, y being regulated, the onesided limits for the functions x, y exist at every point in $[a,b]$ (at the endpoints only the corresponding ones). The function δ_1 evidently represents a gauge on $[a,b]$.

Finally, let us define

$$\delta_2(\tau) = \operatorname{dist}(\tau, \{\sigma_1, \ldots, \sigma_{L_\varepsilon}\})$$

for $\tau \notin \{\sigma_1, \ldots, \sigma_{L_\varepsilon}\}$ ($\operatorname{dist}(\tau, M)$ denotes the distance of the point τ from the set M) and $\delta_2(\tau) > 0$ for $\tau \in \{\sigma_1, \ldots, \sigma_{L_\varepsilon}\}$.

Let us put

$$\delta(\tau) = \min(\delta_0(\tau), \delta_1(\tau), \delta_2(\tau))$$

for $\tau \in [a,b]$. Then δ is a gauge on $[a,b]$ and every δ-fine P-partition

$$D = \{\alpha_0, \tau_1, \alpha_1, \tau_2, \ldots, \alpha_{k-1}, \tau_k, \alpha_k\}$$

of the interval $[a,b]$ has the property that $\{\sigma_1, \ldots, \sigma_{L_\varepsilon}\} \subset \{\tau_1, \ldots, \tau_k\}$ (this is the consequence of the fact that D is δ_2-fine) and (11) holds because D is δ_0-fine.

Assume now that D is an arbitrary δ-fine P-partition of $[a,b]$. Then using (9) we have

$$
\begin{aligned}
(13)\quad & \Bigl\|\Delta_a^b(x,y) - \sum_{\tau\in[a,b)} \Delta^+x(\tau)\Delta^+y(\tau) + \sum_{\tau\in(a,b]} \Delta^-x(\tau)\Delta^-y(\tau)\Bigr\|_Z \\
&= \Bigl\|\Delta_a^b(x,y) - \sum_{k=1}^{\infty} \Delta^+x(\sigma_k)\Delta^+y(\sigma_k) + \sum_{k=1}^{\infty} \Delta^-x(\sigma_k)\Delta^-y(\sigma_k)\Bigr\|_Z \\
&\leqslant \|\Delta_a^b(x,y) - \Delta(x,y,D)\|_Z \\
&\quad + \Bigl\|\Delta(x,y,D) - \sum_{k=1}^{\infty} \Delta^+x(\sigma_k)\Delta^+y(\sigma_k) + \sum_{k=1}^{\infty} \Delta^-x(\sigma_k)\Delta^-y(\sigma_k)\Bigr\|_Z \\
&< \varepsilon + \Bigl\|\Delta(x,y,D) - \sum_{k=1}^{\infty} \Delta^+x(\sigma_k)\Delta^+y(\sigma_k) + \sum_{k=1}^{\infty} \Delta^-x(\sigma_k)\Delta^-y(\sigma_k)\Bigr\|_Z,
\end{aligned}
$$

where (11) was taken into account. Further, by (10) we have

$$
\begin{aligned}
(14)\quad & \Bigl\|\Delta(x,y,D) - \sum_{k=1}^{\infty} \Delta^+x(\sigma_k)\Delta^+y(\sigma_k) + \sum_{k=1}^{\infty} \Delta^-x(\sigma_k)\Delta^-y(\sigma_k)\Bigr\|_Z \\
&< 2\varepsilon + \Bigl\|\Delta(x,y,D) - \sum_{l=1}^{L_\varepsilon} \Delta^+x(\sigma_l)\Delta^+y(\sigma_l) + \sum_{l=1}^{L_\varepsilon} \Delta^-x(\sigma_l)\Delta^-y(\sigma_l)\Bigr\|_Z.
\end{aligned}
$$

Now let us consider the last term on the right hand side of (14):

$$
\begin{aligned}
(15)\quad & \Bigl\|\Delta(x,y,D) - \sum_{l=1}^{L_\varepsilon} \Delta^+x(\sigma_l)\Delta^+y(\sigma_l) + \sum_{l=1}^{L_\varepsilon} \Delta^-x(\sigma_l)\Delta^-y(\sigma_l)\Bigr\|_Z \\
&= \Bigl\|\sum_{j=1}^{k}[(x(\alpha_j)-x(\tau_j))(y(\alpha_j)-y(\tau_j)) - (x(\alpha_{j-1})-x(\tau_j))(y(\alpha_{j-1})-y(\tau_j))] \\
&\qquad\qquad - \sum_{l=1}^{L_\varepsilon} \Delta^+x(\sigma_l)\Delta^+y(\sigma_l) + \sum_{l=1}^{L_\varepsilon} \Delta^-x(\sigma_l)\Delta^-y(\sigma_l)\Bigr\|_Z \\
&\leqslant \Bigl\|\sum_{\substack{j=1\\ \tau_j\in\{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k} [(x(\alpha_j)-x(\tau_j))(y(\alpha_j)-y(\tau_j)) \\
&\qquad\qquad - (x(\alpha_{j-1})-x(\tau_j))(y(\alpha_{j-1})-y(\tau_j))] \\
&\qquad\qquad - \sum_{l=1}^{L_\varepsilon} \Delta^+x(\sigma_l)\Delta^+y(\sigma_l) + \sum_{l=1}^{L_\varepsilon} \Delta^-x(\sigma_l)\Delta^-y(\sigma_l)\Bigr\|_Z \\
&\quad + \Bigl\|\sum_{\substack{j=1\\ \tau_j\notin\{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k} [(x(\alpha_j)-x(\tau_j))(y(\alpha_j)-y(\tau_j)) \\
&\qquad\qquad - (x(\alpha_{j-1})-x(\tau_j))(y(\alpha_{j-1})-y(\tau_j))]\Bigr\|_Z.
\end{aligned}
$$

If $\tau_j \notin \{\sigma_1, \dots, \sigma_{L_\varepsilon}\}$ then $\|\Delta^+ y(\tau_j)\|_Z < \varepsilon$, $\|\Delta^- y(\tau_j)\|_Z < \varepsilon$ and by (12) also

$$\|y(\alpha_j) - y(\tau_j+)\|_Y < \varepsilon, \ \|y(\alpha_{j-1}) - y(\tau_j-)\|_Y < \varepsilon.$$

This yields

$$\|y(\alpha_j) - y(\tau_j)\|_Y = \|y(\alpha_j) - y(\tau_j+) + \Delta^+ y(\tau_j)\|_Y < 2\varepsilon$$

and

$$\|y(\alpha_{j-1}) - y(\tau_j)\|_Y < 2\varepsilon$$

in this case, and we have

$$(16) \qquad \Bigg\| \sum_{\substack{j=1 \\ \tau_j \notin \{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k} [(x(\alpha_j) - x(\tau_j))(y(\alpha_j) - y(\tau_j)) \\ -(x(\alpha_{j-1}) - x(\tau_j))(y(\alpha_{j-1}) - y(\tau_j))]\Bigg\|_Z \leqslant 2\varepsilon(\mathcal{B})\operatorname{var}_a^b x.$$

If $\tau_j \in \{\sigma_1, \dots, \sigma_{L_\varepsilon}\}$ then

$$(17) \\ \Bigg\| \sum_{\substack{j=1 \\ \tau_j \in \{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k} [(x(\alpha_j) - x(\tau_j))(y(\alpha_j) - y(\tau_j)) - (x(\alpha_{j-1}) - x(\tau_j))(y(\alpha_{j-1}) - y(\tau_j))] \\ - \sum_{l=1}^{L_\varepsilon} \Delta^+ x(\sigma_l)\Delta^+ y(\sigma_l) + \sum_{l=1}^{L_\varepsilon} \Delta^- x(\sigma_l)\Delta^- y(\sigma_l)\Bigg\|_Z \\ = \Bigg\| \sum_{\substack{j=1 \\ \tau_j \in \{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k} [(x(\alpha_j) - x(\tau_j))(y(\alpha_j) - y(\tau_j)) - (x(\alpha_{j-1}) - x(\tau_j))(y(\alpha_{j-1}) - y(\tau_j))] \\ - \sum_{\substack{j=1 \\ \tau_j \in \{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k} \Delta^+ x(\tau_j)\Delta^+ y(\tau_j) + \sum_{\substack{j=1 \\ \tau_j \in \{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k} \Delta^- x(\tau_j)\Delta^- y(\tau_j)\Bigg\|_Z \\ \leqslant \Bigg\| \sum_{\substack{j=1 \\ \tau_j \in \{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k} [(x(\alpha_j) - x(\tau_j))(y(\alpha_j) - y(\tau_j)) - \Delta^+ x(\tau_j)\Delta^+ y(\tau_j)]\Bigg\|_Z \\ + \Bigg\| \sum_{\substack{j=1 \\ \tau_j \in \{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k} [(x(\alpha_{j-1}) - x(\tau_j))(y(\alpha_{j-1}) - y(\tau_j)) - \Delta^- x(\tau_j)\Delta^- y(\tau_j)]\Bigg\|_Z.$$

We have

$$
\begin{aligned}
&(x(\alpha_j)-x(\tau_j))(y(\alpha_j)-y(\tau_j))-\Delta^+x(\tau_j)\Delta^+y(\tau_j)\\
&\qquad=\big(x(\alpha_j)-x(\tau_j)\big)(y(\alpha_j)-y(\tau_j+))+\big(x(\alpha_j)-x(\tau_j+)\big)\Delta^+y(\tau_j)
\end{aligned}
$$

and therefore

$$
\begin{aligned}
&\Bigg\|\sum_{\substack{j=1\\ \tau_j\in\{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k}[(x(\alpha_j)-x(\tau_j))(y(\alpha_j)-y(\tau_j))-\Delta^+x(\tau_j)\Delta^+y(\tau_j)]\Bigg\|_Z\\
&\qquad\leqslant\Bigg\|\sum_{\substack{j=1\\ \tau_j\in\{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k}(x(\alpha_j)-x(\tau_j)(y(\alpha_j)-y(\tau_j+))\Bigg\|_Z\\
&\qquad\quad+\Bigg\|\sum_{\substack{j=1\\ \tau_j\in\{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k}\big(x(\alpha_j)-x(\tau_j+)\big)\Delta^+y(\tau_j)\Bigg\|_Z\\
&\qquad\leqslant\varepsilon(\mathcal{B})\mathrm{var}_a^b x+\varepsilon(\mathcal{B})\mathrm{var}_a^b y=\varepsilon((\mathcal{B})\mathrm{var}_a^b x+(\mathcal{B})\mathrm{var}_a^b y)
\end{aligned}
$$

and similarly also

$$
\begin{aligned}
&\Bigg\|\sum_{\substack{j=1\\ \tau_j\in\{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k}[(x(\alpha_{j-1})-x(\tau_j))(y(\alpha_{j-1})-y(\tau_j))-\Delta^-x(\tau_j)\Delta^-y(\tau_j)]\Bigg\|_Z\\
&\qquad\qquad\leqslant\varepsilon((\mathcal{B})\mathrm{var}_a^b x+(\mathcal{B})\mathrm{var}_a^b y).
\end{aligned}
$$

Hence by (17) we get

$$
\begin{aligned}
&\Bigg\|\sum_{\substack{j=1\\ \tau_j\in\{\sigma_1,\dots,\sigma_{L_\varepsilon}\}}}^{k}[(x(\alpha_j)-x(\tau_j))(y(\alpha_j)-y(\tau_j))-(x(\alpha_{j-1})-x(\tau_j))(y(\alpha_{j-1})-y(\tau_j))]\\
&\quad-\sum_{k=1}^{L_\varepsilon}\Delta^+x(\sigma_k)\Delta^+y(\sigma_k)+\sum_{k=1}^{L_\varepsilon}\Delta^-x(\sigma_k)\Delta^-y(\sigma_k)\Bigg\|_Z\leqslant 2\varepsilon((\mathcal{B})\mathrm{var}_a^b x+(\mathcal{B})\mathrm{var}_a^b y).
\end{aligned}
$$

This inequality together with (15) and (16) leads to

$$
\begin{aligned}
&\Bigg\|\Delta(x,y,D)-\sum_{k=1}^{L_\varepsilon}\Delta^+x(\sigma_k)\Delta^+y(\sigma_k)+\sum_{k=1}^{L_\varepsilon}\Delta^-x(\sigma_k)\Delta^-y(\sigma_k)\Bigg\|_Z\\
&\qquad<2\varepsilon((\mathcal{B})\mathrm{var}_a^b x+(\mathcal{B})\mathrm{var}_a^b y)+2\varepsilon(\mathcal{B})\mathrm{var}_a^b x=\varepsilon(4(\mathcal{B})\mathrm{var}_a^b x+2(\mathcal{B})\mathrm{var}_a^b y)
\end{aligned}
$$

and this with (9) implies (8) because $\varepsilon>0$ can be taken arbitrarily small. □

13. Theorem. *If $\mathcal{B}=(X,Y,Z)$ is a bilinear triple such that $x\in G([a,b],X)\cap(\mathcal{B})BV([a,b],X)$ and $y\in G([a,b],Y)\cap(\mathcal{B})BV([a,b],Y)$ then*

$$\begin{aligned}\int_a^b x(s)\mathrm{d}[y(s)]+\int_a^b \mathrm{d}[x(s)]y(s) &= x(b)y(b)-x(a)y(a)\\ &\quad-\sum_{a\leqslant\tau<b}\Delta^+x(\tau)\Delta^+y(\tau)+\sum_{a<\tau\leqslant b}\Delta^-x(\tau)\Delta^-y(\tau).\end{aligned}\tag{18}$$

P r o o f. By the assumption the integrals $\int_a^b x(s)\mathrm{d}[y(s)]$, $\int_a^b \mathrm{d}[x(s)]y(s)$ exist (see Corollary 8) and by the integration by parts Theorem 10 we have

$$\int_a^b x(s)\mathrm{d}[y(s)]+\int_a^b \mathrm{d}[x(s)]y(s)=x(b)y(b)-x(a)y(a)-\Delta_a^b(x,y).$$

Using (8) from Lemma 12 we immediately obtain (18). □

14. Corollary. *If $x\in BV([a,b],X)$ and $y\in BV([a,b],Y)$ then the integrals*

$$\int_a^b x(s)\mathrm{d}[y(s)],\quad \int_a^b \mathrm{d}[x(s)]y(s)$$

exist and (18) *holds.*

P r o o f. By (1) and (3) we have $x\in G([a,b],X)\cap(\mathcal{B})BV([a,b],X)$, $y\in G([a,b],Y)\cap(\mathcal{B})BV([a,b],Y)$ and the result follows from Theorem 13. □

R e m a r k. This form of integration by parts result was derived e.g. in [6], [7] for the case of finite dimensional spaces.

15. Theorem. *If*

$$x\in C([a,b],X)\cap(\mathcal{B})BV([a,b],X) \text{ and } y\in G([a,b],Y)\cap(\mathcal{B})BV([a,b],Y)$$

or if

$$x\in G([a,b],X)\cap(\mathcal{B})BV([a,b],X) \text{ and } y\in C([a,b],Y)\cap(\mathcal{B})BV([a,b],Y)$$

then

$$\int_a^b x(s)\mathrm{d}[y(s)]+\int_a^b \mathrm{d}[x(s)]y(s)=x(b)y(b)-x(a)y(a).\tag{19}$$

P r o o f. Since $C([a,b],X)\subset G([a,b],X)$ we have $\Delta^+x(\tau)=0$, $\Delta^-x(\tau)=0$ for $\tau\in[a,b]$, and the assumptions of Theorem 13 being satisfied the equality (18) holds. Moreover,

$$\sum_{a\leqslant\tau<b}\Delta^+x(\tau)\Delta^+y(\tau)+\sum_{a<\tau\leqslant b}\Delta^-x(\tau)\Delta^-y(\tau)=0$$

and therefore (19) holds in the first case. The second can be proved analogously. □

References

[1] *Henstock, R.*: Integration by parts. Aequationes Mathematicae *9* (1973), 1–18.
[2] *Hönig, Ch. S.*: Volterra Stieltjes-Integral Equations. North-Holland Publ. Comp., Amsterdam, 1975.
[3] *Kurzweil, J.*: On integration by parts. Czechoslovak Math. J. *8* (1958), 356–359.
[4] *Kurzweil, J.*: Nichtabsolut konvergente Integrale. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1980.
[5] *Schwabik, Š.*: Abstract Perron-Stieltjes integral. Math. Bohem. *121* (1996), 425–447.
[6] *Schwabik, Š.*: Generalized Ordinary Differential Equations. World Scientific, Singapore, 1992.
[7] *Schwabik, Š., Tvrdý, M., Vejvoda O.*: Differential and Integral Equations. Academia & Reidel, Praha & Dordrecht, 1979.
[8] *Schwabik, Š.*: Linear Stieltjes integral equations in Banach spaces. Math. Bohem. *124* (1999), 433–457.
[9] *Schwabik, Š.*: Linear Stieltjes integral equations in Banach spaces II; Operator valued solutions. Math. Bohem. *125* (2000), 431–454.

Author's address: *Štefan Schwabik*, Mathematical Institute, Academy of Sciences of the Czech Republic, Žitná 25, 115 67 Praha 1, Czech Republic, e-mail: `schwabik@math.cas.cz`.