Electronic Journal of Differential Equations, Vol. 2003(2003), No. 104, pp. 1–23.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu
ftp ejde.math.txstate.edu (login: ftp)

A LINEAR FUNCTIONAL DIFFERENTIAL EQUATION WITH DISTRIBUTIONS IN THE INPUT

VADIM Z. TSALYUK

Abstract. This paper studies the functional differential equation

$$\dot{x}(t) = \int_a^t d_s R(t,s)\, x(s) + F'(t), \quad t \in [a,b],$$

where F' is a generalized derivative, and $R(t,\cdot)$ and F are functions of bounded variation. A solution is defined by the difference $x - F$ being absolutely continuous and satisfying the inclusion

$$\frac{d}{dt}(x(t) - F(t)) \in \int_a^t d_s R(t,s)\, x(s).$$

Here, the integral in the right is the multivalued Stieltjes integral presented in [11] (in this article we review and extend the results in [11]). We show that the solution set for the initial-value problem is nonempty, compact, and convex. A solution x is said to have memory if there exists the function $\bar{x}$ such that $\bar{x}(a) = x(a)$, $\bar{x}(b) = x(b)$, $\bar{x}(t) \in [x(t-0), x(t+0)]$ for $t \in (a,b)$, and $\frac{d}{dt}(x(t) - F(t)) = \int_a^t d_s R(t,s)\, \bar{x}(s)$, where Lebesgue-Stieltjes integral is used. We show that such solutions form a nonempty, compact, and convex set. It is shown that solutions with memory obey the Cauchy-type formula

$$x(t) \in C(t,a)x(a) + \int_a^t C(t,s)\, dF(s).$$

1. Introduction: Model example

Our purpose is to generalize the linear functional differential equation

$$\dot{x}(t) = \int_a^t d_s R(t,s)\, x(s) + f(t), \quad t \in [a,b], \tag{1.1}$$

introducing a distribution of order zero, otherwise a measure, as f. Such a kind of inputs for systems described by ordinary differential equations were studied in many works; see for instance [9].

The functions $R(t,\cdot)$ in the equation (1.1) are usually discontinuous and have bounded variation. If f is a measure then it is natural to suppose that a solution x belongs to a class of primitives for measures, i.e. it is also discontinuous and has bounded variation on $[a,b]$.

2000 *Mathematics Subject Classification.* 26A42, 28B20, 34A60.
Key words and phrases. Stieltjes integral, function of bounded variation, multivalued integral, linear functional differential equation.

Submitted June 14, 2003. Published October 13, 2003.
Partialy supported by grant 03-01-00255 from the Russian Foundation for Basic Research.

To illustrate the question, we consider the model problem

$$\begin{aligned} &\dot{x}(t) = h(t-1)\cdot x(1) + \delta_1, \quad t \in [0,2],\\ &x(0) = 0, \end{aligned} \tag{1.2}$$

where h is the Heaviside function ($h(t) = 0$ for $t < 0$ and $h(t) = 1$ for $t > 0$), $\delta_1 = h'(t-1)$ is the unit impulse (so-called delta-function) concentrated at the point 1.

The behaviour of the system for $t > 1$ essentially depends on the value $x(1)$. Naturally, we assume the solution $x(t)$ of (1.2) being equal to 0 for $t < 1$ and having a unit jump at $t = 1$. Then what is $x(1)$? Is it $x(1-0) = 0$, or $x(1+0) = 1$, or some intermediate value? The system do not know how to behave further. Such a phenomenon frequently occurs in the theory of functional differential equations and is known as "hovering" [4].

To avoid hovering, Anokhin [1, 4] assumed the solution of (1.2) to be piecewise absolutely continuous, having an unique jump at the point 1, and continuous from the right. Then $x(1)$ in (1.2) is well defined and, clearly, the initial value problem has the unique solution

$$x_+(t) = \begin{cases} 0, & t < 1,\\ t, & t \geq 1. \end{cases}$$

If we want the solution to be continuous from the left, as in [2], then we get

$$x_-(t) = \begin{cases} 0, & t \leq 1,\\ 1, & t > 1. \end{cases}$$

Let us introduce a perturbation of the right-hand side of (1.2). Exactly, we replace the equation by

$$\dot{x}(t) = h(t-c)\cdot x(c) + \delta_1, \quad c \neq 1. \tag{1.3}$$

From the point of view of certain applications, such a perturbation looks small as $c \to 1$.

In the case $c > 1$, the solution of (1.3) has the form

$$x_c(t) = \begin{cases} 0, & t < 1,\\ 1, & 1 < t \leq c,\\ t+1-c, & t > c, \end{cases}$$

and for $c < 1$ we have

$$x_c(t) = \begin{cases} 0, & t < 1,\\ 1, & t > 1. \end{cases}$$

Now let c tend to 1. Then $x_c \to x_+$ as $c \downarrow 1$ and $x_c \to x_-$ as $c \uparrow 1$.

So we have detected the discontinuity. In the theory of discontinuos systems, such a phenomenon is usually treated as follows. First, the so-called "nonideality" (the regularization) is introduced. The regularized equation ought to be of well-known class of systems and uniquely solvable. Next, the size of nonideality is assumed to tend to 0. Any limit of the solutions is said to be a solution of the initial "ideal" system. Thus, the uniqueness may be lost in these situations. We shall follow this way.

Let us replace the unit impulse δ_1 by the "delta-like sequence" ($\varepsilon > 0$) $f_\varepsilon(t) \geq 0$, $f_\varepsilon(t) = 0$ for $t \in (-\infty, 1-\varepsilon) \cup (1+\varepsilon, \infty)$, $\int_{1-\varepsilon}^{1+\varepsilon} f_\varepsilon(t)\,dt = 1$, and put $\alpha_\varepsilon = \int_{1-\varepsilon}^{1} f_\varepsilon(t)\,dt$. The solutions of the problems

$$\begin{gathered}\dot{x}(t) = h(t-1)x(1) + f_\varepsilon(t),\\ x(0) = 0\end{gathered}$$

are as follows:

$$x_\varepsilon(t) = \begin{cases} 0, & 0 \leq t \leq 1-\varepsilon,\\ \int_{1-\varepsilon}^{t} f_\varepsilon(s)\,ds, & 1-\varepsilon < t \leq 1,\\ \int_{1-\varepsilon}^{t} f_\varepsilon(s)\,ds + \alpha_\varepsilon \cdot (t-1), & 1 < t \leq 1+\varepsilon,\\ 1 + \alpha_\varepsilon \cdot (t-1), & t > 1+\varepsilon. \end{cases}$$

Let ε tend to 0. We see that, in order to the convergence of solutions, it is necessary that $\alpha_\varepsilon \to \alpha \in [0,1]$. On the other hand, each value in $[0,1]$ may be the limit of α_ε. So, if we want the limit of solutions to be a solution, the problem (1.3) gets a family of solutions

$$x(t) = \begin{cases} 0, & 0 \leq t < 1,\\ \alpha, & t = 1,\\ 1 + \alpha \cdot (t-1), & t > 1. \end{cases}$$

with arbitrary $\alpha \in [0,1]$.

Another kind of the "nonideality" represents a model of inaccuracy of the measuring device for $x(1)$ – or our inperfect knowledge of this device. Replace the operator $h(t-1)x(1)$ in (1.2) with one sufficiently close to. Let $\alpha, \beta \geq 0$, $\alpha + \beta = 1$, and set

$$(\mathbf{K}_\varepsilon x)(t) = \begin{cases} 0, & 0 \leq t \leq 1-\beta\varepsilon^2,\\ \varepsilon^{-2} \int_{1-\beta\varepsilon^2}^{t} x(s)\,ds, & 1-\beta\varepsilon^2 < t \leq 1+\alpha\varepsilon^2,\\ \varepsilon^{-2} \int_{1-\beta\varepsilon^2}^{1+\alpha\varepsilon^2} x(s)\,ds, & t > 1+\alpha\varepsilon^2. \end{cases}$$

The solution of the problem

$$\begin{gathered}\dot{x}(t) = (\mathbf{K}_\varepsilon x)(t) + \delta_1,\\ x(0) = 0\end{gathered}$$

tends to

$$x(t) = \begin{cases} 0, & 0 \leq t < 1,\\ 1 + \alpha \cdot (t-1), & t > 1, \end{cases}$$

as $\varepsilon \downarrow 0$. Naturally we consider these functions with arbitrary $\alpha \in [0,1]$ as the solutions of problem (1.2).

Assuming that the value $\alpha \in [0,1]$ is chosen for each $t > 1$ independently, we get the solution family

$$x(t) = \begin{cases} 0, & 0 \leq t < 1,\\ 1 + \int_1^t \alpha(s)\,ds, & t > 1, \end{cases}$$

with summable functions $\alpha(s) \in [0,1]$.

Eventually, in equation (1.2) we have to mean the whole segment $[0,1] = [x(1-0), x(1+0)]$ instead of $x(1)$. And solutions of (1.2) ought to satisfy somewhat like inclusion.

2. Preliminaries

2.1. **Some notions and notation. 1.** Functions x are assumed further to have values in the finite-dimensional space $\mathbb{R}^N$. The space for values of R is the space of matrices $\mathbb{R}^{N\times N}$. Consequently products $R(t,s)x(s)$ and values of integrals will lie in the space $\mathbb{R}^N$. All the norms in these spaces are denoted by $|\cdot|$ and we assume that the norm of matrix $|g| = \sup_{|x|=1} |gx|$.

The product sets $[a,b]\times\mathbb{R}$, $[a,b]\times\mathbb{R}^N$, etc, are equipped with the metric

$$\rho\left((t_1,x_1),\ (t_2,x_2)\right) = \max\left(|t_1-t_2|, |x_1-x_2|\right).$$

2. The ***distance*** from a point x to a set A is $\rho(x,A)\overset{\text{df}}{=}\sup_{a\in A}|x-a|$. We denote $\beta(A,B)\overset{\text{df}}{=}\sup_{a\in A}\rho(a,B)$. The symmetric function

$$\alpha(A,B) = \max(\beta(A,B), \beta(B,A))$$

on the space $\operatorname{Comp}(\mathbb{R}^N)$ of nonempty compact subsets of finite-dimensional space is the metric named Hausdorff distance. This metric space is complete. We use also the ***triangle inequality*** of the form

$$\beta(A,C) \le \beta(A,B) + \beta(B,C)$$

that holds for compact sets A, B and C.

The sequence of sets A_n is said to converge to a set A if $\alpha(A_n,A)\to 0$ as $n\to\infty$. For compact sets A, A_n, it is easy to see that the convergence

$$\beta\left(A_n, A\right) \to 0$$

is equivalent to what follows: for every sequence $a_n \in A_n$

i) this sequence is compact;
ii) if $\lim_{i\to\infty} a_{n_i} = a$ for subsequence a_{n_i} then $a\in A$.

We denote $\|A\|\overset{\text{df}}{=}\sup_{v\in A}|v| = \alpha(A,0)$.

3. For u and v being points of a linear space, by $[u,v]$ we denote the convex hull of these two points.

If x is a function defined on $[a,b]\subset\mathbb{R}$ and $t\in(a,b)$ then

$$\Delta x(t)\overset{\text{df}}{=}x(t+0)-x(t-0)$$

is a ***jump*** of x at the point t.

We say that the function g of bounded variation is ***proper*** if

$$g(t)\in[g(t-0), g(t+0)]$$

for all $t\in(a,b)$. In this case, in the definition $\operatorname{Var}_a^b g = \sup\sum|g(t_i)-g(t_{i-1})|$ of the variation of g we may assume that each point t_i not coinciding with the end of the segment $[a,b]$ is a point of continuity of g.

By $\bar\Gamma(g)$ denote a ***completed graph*** of function g, i.e. the set in $[a,b]\times\mathbb{G}$ that consists of the segments $\{a\}\times[g(a), g(a+0)]$, $\{b\}\times[g(b-0), g(b)]$, and $\{t\}\times[g(t-0), g(t+0)]$ for $t\in(a,b)$. It's obvious that for function g of bounded variation $\bar\Gamma(g)$ is closed set.

If $\bar\Gamma(g_1)=\bar\Gamma(g_2)$ then we say that the functions g_1 and g_2 of bouned variation are ***equivalent***. In this case for $x\in\mathbf{C}[a,b]$ ($\mathbf{C}[a,b]$ is the space of continuous functions) we have $\int_a^b dg_1(t)x(t) = \int_a^b dg_2(t)x(t)$. In the case of $N=1$ the norm of this functional is equal to $\operatorname{Var}_a^b g$ with proper g taken from the equivalence class of g_1 and g_2.

We denote by $\mathbf{BV}[a,b]$ the space of equivalence classes of the functions of bounded variation. We define $\mathrm{Var}_a^b g$ for equivalence class using a proper element of the class. We say that $t \in (a,b)$ is a ***(dis-)continuity point*** of $g \in \mathbf{BV}[a,b]$ if the proper element of the equivalence class is (dis-)continuous at t. In the points of continuity the ***value*** $g(t)$ is defined as the value of the proper element. We say that the sequence of $g_n \in \mathbf{BV}[a,b]$ ***uniformly converges*** to g if the proper representatives of g_n converge to proper element of g uniformly for all $t \in [a,b]$ except for discontinuity points of g in (a,b).

5. Anywhere below the integration variable may be omitted if it does not cause uncertainty.

6. Let T be a measurable subset of the finite-dimensional space. The multivalued function $X : T \to \mathrm{Comp}(\mathbb{R}^N)$ is ***measurable*** if for each $A \in \mathrm{Comp}(\mathbb{R}^N)$ the set $\{t \in T : X(t) \cap A \neq \emptyset\}$ is measurable. The detailed review [8] contains some other definitions of the measurability of multifunctions, conditions for the equivalence of the definitions, and various properties of measurable multifunctions. See also [7, pp. 148–150] and [6, pp. 205–206].

Particularly, X is measurable if and only if there exists a countable set of measurable singlevalued functions $x_i : T \to \mathbb{R}^N$ such that

$$X(t) = \overline{\{x_i(t) \,:\, i = 1, 2, \dots\}} \quad \text{a.e.}$$

Using this, one can easily check that if $x_0, x_1 : T \to \mathbb{R}^N$ are measurable singlevalued functions then the multifunction $t \mapsto [x_0(t), x_1(t)]$ is measurable. If the series $X(t) = \sum_{i=1}^{\infty} X_i(t)$ has a convergent in $\mathbf{L}_1(a,b)$ majorant then the multifunction X is measurable [6, Theorem 21].

7. The ***Aumann integral*** of the multifunction $X : [a,b] \to \mathrm{Comp}(\mathbb{R}^N)$ is, by definition, the set

$$\int_a^b X(t)\,dt = \Big\{ \int_a^b x(t)\,dt : x \in \mathbf{L}_1(a,b), \quad x(t) \in X(t) \text{ for a.e. } t \in [a,b] \Big\}.$$

8. We say that multifunction X is ***boundedly summable*** if it is measurable and $\|X(t)\| \le \chi(t)$ a.e. on $[a,b]$ for some summable (singlevalued) function χ. If X is such a multifunction then according to the theorem of A. A. Lyapunov ([7, Theorem 1.7.10], [6]) the set $\int_a^b X(t)\,dt$ is nonepty, convex, and compact; besides,

$$\int_a^b X(t)\,dt = \int_a^b \overline{\mathrm{co} X(t)}\,dt$$

(co$X$ is a convex hull of X).

9. The ***support function*** of the set $A \in \mathrm{Comp}(\mathbb{R}^N)$ is

$$\mathrm{supp}(\psi|\, A) := \sup_{a \in A} \psi \cdot a,$$

where $\psi \cdot a$ denotes the usual scalar product in $\mathbb{R}^N$. Sufficiently complete review of the properties of the support functions may be found in [6, pp. 197–200].

The convex hull co$A$ of $A \in \mathrm{Comp}(\mathbb{R}^N)$ is uniquely determined by the values of the support functions $\mathrm{supp}(\psi|\, A)$ with $|\psi| = 1$:

$$\mathrm{co}A = \bigcap_{\psi} \{x : \ \psi \cdot x \le \mathrm{supp}(\psi|\, A)\}.$$

Note that it is sufficient here that ψ ranges over the countable dence subset of the unit sphere.

The multifunction $X : T \to \mathrm{Comp}(\mathbb{R}^N)$ is measurable if and only if the mappings $t \mapsto \mathrm{supp}(\psi|\, X(t))$ are measurable [6, Theorem 22]. For X being boundedly summable, according to [6, Theorem 42] we have

$$\mathrm{supp}\Big(\psi\Big|\int_a^b X(t)\,dt\Big) = \int_a^b \mathrm{supp}(\psi\big|X(t))\,dt. \tag{2.1}$$

2.2. **Auxiliary lemmas.** The next two lemmas may be known, but the author did not find a reference to them.

Lemma 2.1 (Lebesgue-type theorem on convergence of integrals)**.** *Suppose X and X_n, $n = 1, 2, \dots$ are bounded summable multivalued functions on $[a, b]$ with values in $\mathrm{Comp}(\mathbb{R}^N)$; $\|X_n(t)\| \le \chi(t)$ a.e. on $[a, b]$, $\chi \in \mathbf{L}_1(a, b)$. If $\beta(X_n(t), X(t)) \to 0$ for a.e. $t \in [a, b]$ as $n \to \infty$ then*

$$\beta\Big(\int_a^b X_n(t)\,dt,\ \int_a^b X(t)\,dt\Big) \to 0. \tag{2.2}$$

Proof. Let $|\psi| = 1$. Obviously,

$$\mathrm{supp}(\psi\,|\,X_n(t)\,) \le \mathrm{supp}(\psi\,|\,X(t)\,) + \beta(X_n(t), X(t)).$$

Integrating we get

$$\int_a^b \mathrm{supp}(\psi\,|\,X_n(t)\,)\,dt \le \int_a^b \mathrm{supp}(\psi\,|\,X(t)\,)\,dt + \int_a^b \beta(X_n(t), X(t))\,dt$$

(since $\beta(X_n(t), X(t)) \le \|X_n(t)\| + \|X(t)\|$, the latter integral exists). According to (2.1) and the conventional Lebesgue theorem,

$$\begin{aligned}
&\limsup_{n\to\infty} \mathrm{supp}\Big(\psi\Big|\int_a^b X_n(t)\,dt\Big)\\
&\le \limsup_{n\to\infty}\Big(\int_a^b \mathrm{supp}(\psi|\,X(t))\,dt + \int_a^b \beta(X_n(t), X(t))\,dt\Big)\\
&\le \int_a^b \mathrm{supp}(\psi\,|\,X(t)\,)\,dt\\
&\le \mathrm{supp}\Big(\psi\Big|\int_a^b X(t)\,dt\Big).
\end{aligned}$$

Consider an arbitrary sequence $v_n \in \int_a^b X_n(t)\,dt$. Since $\|\int_a^b X_n(t)\,dt\| \le \int_a^b \chi(t)\,dt$, this sequence is compact. Furthermore, if $\lim_{i\to\infty} v_{n_i} = v$ then

$$\psi\cdot v \le \lim_{i\to\infty} \psi\cdot v_{n_i} \le \limsup_{i\to\infty} \mathrm{supp}\Big(\psi\Big|\int_a^b X_{n_i}(t)\,dt\Big) \le \mathrm{supp}\Big(\psi\Big|\int_a^b X(t)\,dt\Big)$$

for every ψ such that $|\psi| = 1$. Consequently, $v \in \int_a^b X(t)\,dt$. These two facts combined are equivalent to (2.2). □

Note that the analoguous assertion with $\beta(A, B)$ replaced by $\alpha(A, B)$ is cited in [6] (Theorem 45 of the review).

Lemma 2.2 (Differentiability of multivalued integrals)**.** *If $X : [a, b] \to \mathrm{Comp}(\mathbb{R}^N)$ is boundedly summable then for a.e. $t \in [a, b)$*

$$\beta\Big(\frac{1}{\Delta t}\int_t^{t+\Delta t} X(s)\,ds, \mathrm{co}X(t)\Big) \to 0 \quad \text{as } \Delta t \downarrow 0. \tag{2.3}$$

Proof. Let Ψ be a countable dense subset of the unit sphere. A convex set $A \in \mathrm{Comp}(\mathbb{R}^N)$ is uniquely determined by the values $\mathrm{supp}(\psi\big|A)$, $\psi \in \Psi$:

$$A = \bigcap_{\psi\in\Psi} \{x : \psi \cdot x \le \mathrm{supp}(\psi\big|A)\}.$$

Since $\mathrm{supp}(\psi\big|X(t)) \le \|X(t)\|$, the functions $t \mapsto \mathrm{supp}(\psi\,|X(t)|)$ are summable. Due to (2.1), for $t \in E$, $E \subset [a, b)$, $\mathrm{mes}([a, b]\backslash E) = 0$, and $\psi \in \Psi$ we have

$$\mathrm{supp}\Big(\psi\Big|\frac{1}{\Delta t}\int_t^{t+\Delta t} X(s)\,ds\Big) = \frac{1}{\Delta t}\int_t^{t+\Delta t} \mathrm{supp}(\psi\big|X(s))\,ds \to \mathrm{supp}(\psi\big|X(t))$$

and

$$\frac{1}{\Delta t}\int_t^{t+\Delta t} \|X(s)\|\;ds \to \|X(t)\| \quad \text{as } \Delta t \downarrow 0. \tag{2.4}$$

We assume in the sequel that $t \in E$. Let $\Delta t_n \downarrow 0$ and $v_n \in \frac{1}{\Delta t_n}\int_t^{t+\Delta t_n} X(s)\,ds$. Due to (2.4)

i) the sequence $\{v_n\}$ is compact.

Besides,

ii) if $v_{n_i} \to v$ as $i \to \infty$ then $v \in \mathrm{co}X(t)$.

Indeed, for $\psi \in \Psi$ we have

$$\psi \cdot v = \lim_{i\to\infty} \psi \cdot v_{n_i} \le \lim_{i\to\infty} \mathrm{supp}\Big(\psi\Big|\frac{1}{\Delta t_{n_i}}\int_t^{t+\Delta t_{n_i}} X(s)\,ds\Big) = \mathrm{supp}\left(\psi\big|X(t)\right),$$

and therefore $v \in \overline{\mathrm{co}X(t)} = \mathrm{co}X(t)$ (we recall that the set $X(t)$ is compact). The assertions a) and b) above, being combined, give (2.3). □

3. The multivalued Stieltjes integral

This section containes a review of the results of the paper [11]. We need the concept of the multivalued integral

$$Ax = \int_a^b dg(s)\,x(s) \tag{3.1}$$

for g and x being both discontinuous functions of bounded variation.

Note that the Riemann–Stieltjes integral

$$Ax = \int_a^b dg(t)\,x(t) = g(b)x(b) - g(a)x(a) - \int_a^b g(t)\,dx(t)$$

exists for **continuous** function g. We define the integral (3.1) by passing to the limit $g_i \to g$ where g_i are continuous. From the point of view of applications, it is natural to interpret a small error on the moment of jump of g as a small deviation of the function. Therefore we construct the integral (3.1) using a convergence

$$A_i x = \int_a^b dg_i(t)\,x(t) \to Ax$$

almost similar to the pointwise convergence of functionals $A_i \to A$ in the space $\mathbf{C}[a, b]$.

Focusing on the model equation (1.2), we are about to costruct the integral $Ax = \int_0^2 dg(t)\, x(t)$ for functions

$$g(t) = \begin{cases} 0, & t \in [0,1), \\ 1, & t \in (1,2], \end{cases} \quad \text{and} \quad x(t) = \begin{cases} x_0, & t \in [0,1), \\ x_1, & t \in (1,2]. \end{cases}$$

For an arbitrary $a \in \mathbb{R}$ and each $i = 1, 2, 3, \dots$, define the continuous piecewise linear functions $g_i(t)$ through the vertices $(0,0)$, $(1 - \frac{1}{i}, 0)$, $(1, a)$, $(1 + \frac{1}{i}, 1)$ and $(2, 1)$. Then, by the first Helly Theorem [10], the functionals A_i pointwise converge to $Ax = x(1)$ in $\mathbf{C}[0,2]$.

The function x under the consideration, however, has only one-sided limits at the point 1, and it is easy to check that $A_i x = x_1 - a(x_1 - x_0)$. So $\lim_{i\to\infty} A_i x$ depends on the way of construction of continuous approximations g_i. The common decision in such a situation is to declare the functional A to be **multivalued** and the set Ax containing **all** such limits. It is important for the theory of equations that Ax is a **compact** set. But here, according to the arbitrariness of a, we get $Ax = (-\infty, \infty)$ in the case $x_1 \neq x_0$. Of course it's inadmissible.

Note that $\|A_i\| = \mathrm{Var}_0^2 g_i = |a| + |a - 1|$ converge to $\|A\| = 1$ if and only if $a \in [0,1]$. Under this restriction, we get $\lim_{i\to\infty} A_i x = ax_0 + (1-a)x_1 \in [x_0, x_1]$. The set $[x_0, x_1] = [x(1-0), x(1+0)]$ of all such limits should be considered as the value of $Ax = \int_0^2 dg(t)\, x(t)$. (It should be noted that the functions x and g are not assumed to have certain values at their discontinuity points.) This is why the convergence we use implies also the convergence of the norms of functionals. It is natural to name it ***normal convergence***.

On the other hand, the graphs of the functions g_i converge to the graph of g completed with the "vertical" segments at the points of discontinuity. Such a type of convergence may also be used to define the integral (3.1).

For the completeness of the exposition, we give here the summary of the paper [11] devoted to the introduction and the properties of such a kind of integral.

3.1. On the convergence.

Definition 3.1. The sequence of functions $g_n \in \mathbf{BV}[a,b]$ ***β-converges*** on the segment $[a,b]$ to the function $g \in \mathbf{BV}[a,b]$ as $n \to \infty$ if the following conditions hold:

1) $g_n(a) \to g(a)$, $g_n(b) \to g(b)$;
2) $\beta\left(\bar{\Gamma}(g_n), \bar{\Gamma}(g)\right) \to 0$.

Such a convergence is denoted as $g_n \xrightarrow{\beta} g$. If, additionally,

3)
$$\mathrm{Var}_a^b g_n \to \mathrm{Var}_a^b g, \tag{3.2}$$

then we say that g_n ***normally converge*** to g and denote this by $g_n \xrightarrow{N} g$.

Lemma 3.2. *Let functions $g_n \in \mathbf{BV}[a,b]$ normally converge to a function $g \in \mathbf{BV}[a,b]$.*

a) *For $\tau \in (a,b)$, the distance from the point $g_n(\tau)$ to the segment $[g(\tau - 0), g(\tau + 0)]$ tends to 0 as $n \to \infty$.*
b) *If g is continuous at the point a then $\lim_{n\to\infty} g_n(a + 0) = g(a)$; the analoguous statement takes place for the point b.*
c) *If g is continuous on $[a,b]$ then $g_n(t) \to g(t)$ uniformly for $t \in [a,b]$.*

Lemma 3.3. *Let g_n normally converge on $[a,b]$ to g. If $c \in (a,b)$ is a continuity point of g then $g_n \to g$ normally on both intervals $[a,c]$ and $[c,b]$.*

Lemma 3.4. *Let $\{f_n\}$ and $\{g_n\}$ be the sequences of functions normally converging on $[a,b]$ to g. Then for arbitrary $\lambda \in (0,1)$ the functions $\lambda f_n + (1-\lambda) g_n$ normally converge to the function g.*

Proposition 3.5. *Any proper function of bounded variation on $[a,b]$ is a normal limit of some sequence of continuous functions.*

Proof. For arbitrary $\delta > 0$ take a finite partition $a = t_0 < t_1 < \dots < t_n = b$ such that

1) $\sum |g(t_i) - g(t_{i-1})| > \mathrm{Var}_a^b g - \delta$,
2) $\max_{1\le i\le n}(t_i - t_{i-1}) < \delta$,
3) $\max_{1\le i\le n} \mathrm{Var}_{t_{i-1}+0}^{t_i-0} g < \delta$,
4) at least one in the pair of the neighbour points t_{i-1} and t_i is the continuity point of g.

To satisfy each of the conditions 1)–4), we may add to the partition a finite set of points so that the previous conditions stay fulfilled.

Let $l_\delta(t)$ be a piecewise linear function built through the points $(t_i, g(t_i))$. Hence $\mathrm{Var}_a^b l_\delta \to \mathrm{Var}_a^b g$ as $\delta \to 0$. The inequality $\beta\left(\bar\Gamma(l_\delta),\, \bar\Gamma(g)\right) \le \delta$ is proved in [11]. □

Remark 3.6. The approximations $l_\delta(t)$ constructed here are absolutely continuous.

3.2. **Definition and properties of the multivalued Stieltjes integral.** Let $x, g \in \mathbf{BV}[a,b]$. The set $\int_a^b dg(t)\, x(t)$ is defined as follows.

Definition 3.7. The point v belongs to the set $\int_a^b dg\, x$ if and only if there exists a sequence of continuous functions $g_i \in \mathbf{BV}[a,b]$ such that

1) the sequence normally converges to g, and
2) $\lim_{i\to\infty} \int_a^b dg_i\, x = v$.

Theorem 3.8. *The set $\int_a^b dg\, x$ is convex and compact; besides,*

$$\Big\| \int_a^b dg\, x \Big\| \le \mathrm{Var}_a^b g \cdot \sup_{t\in[a,b]} |x(t)| \le \mathrm{Var}_a^b g \cdot \big(|x(a)| + \mathrm{Var}_a^b x\big).$$

Example 3.9. Consider the functions

$$g(t) = \begin{cases} g_0, & t < c, \\ g_1, & t > c, \end{cases} \quad \text{and} \quad x(t) = \begin{cases} x_0, & t < d, \\ x_1, & t > d, \end{cases}$$

where $c, d \in [a,b]$. In the case $c = a$, the value $g(a)$ ought to be chosen in such a way that the jump at the point a belongs to $[a,b]$. Analogous assumptions are made for $c = b$ and for the function x.

First we suppose that $c = d \in (a,b)$. Let g_i be continuous functions normally converging to g on $[a,b]$ and $v = \lim_{i\to\infty} \int_a^b dg_i\, x$. Therefore,

$$v = \lim_{i\to\infty} \Big(g_i(t)x(t)\Big|_a^b - \int_a^b g_i\, dx \Big) = g(b)x_1 - g(a)x_0 - \lim_{i\to\infty} \left(g_i(c)\cdot(x_1 - x_0)\right),$$

and the latter limit exists. Due to compactness we may assume that $\lim_{i\to\infty} g_i(c)$ exists and by Lemma 3.2 it belongs to the segment $[g(c-0), g(c+0)]$. On the other hand, using piecewise linear approximations we can present each point of this segment as a limit of values of the sequence of continuous functions $g_i \in \mathbf{BV}[a,b]$ (Proposition 3.5). Hence,

$$\begin{aligned}\int_a^b dg\,x &= g(b)x_1 - g(a)x_0 - [g(c-0), g(c+0)]\cdot(x_1 - x_0)\\ &= (g_1 - g_0)\cdot[x_0, x_1].\end{aligned}$$

Analogous reasonings lead to

$$\int_a^b dg\,x = \begin{cases}(g_1-g_0)\cdot x_0 & \text{if } a \le c < d \le b \text{ or } c = d = b,\\ (g_1-g_0)\cdot[x_0,x_1] & \text{if } a < c = d < b,\\ (g_1-g_0)\cdot x_1 & \text{if } a \le d < c \le b \text{ or } a = c = d.\end{cases}$$

Denote

$$\begin{aligned} g\big|_a^b x &= g(b)x(b-0) - g(b-0)x(b-0) + g(b-0)x(b)\\ &\quad - g(a)x(a+0) + g(a+0)x(a+0) - g(a+0)x(a).\end{aligned}$$

Such symbol has the following properties:

a) Symmetry: if one exchanges g and x and restores the right order of factors then the same expression appears.
b) Provided that at each end of the segment $[a,b]$ at least one of the functions x, g is continuous, the value of the symbol is equal to the ordinary ***double substitution*** $g(t)x(t)\big|_a^b = g(b)x(b) - g(a)x(a)$.
c) if $x_n(t) \to x(t)$ and $g_n(t) \to g(t)$ as $n \to \infty$ uniformly for $t \in [a,b]$ then $\lim_{n\to\infty} g_n\big|_a^b x_n = g\big|_a^b x$. □

Theorem 3.10. *The formula of integration by parts holds:*

$$\int_a^b dg\,x = g\Big|_a^b x - \int_a^b g\,dx\,.$$

Let g^c be the continuous component of the function $g \in \mathbf{BV}[a,b]$, $g^c(a) = g(a)$ and $J(g)$ be the set of discontinuity points of g in the open interval (a,b).

Theorem 3.11. *The following two formulas are valid:*

$$\begin{aligned}\int_a^b dg\,x = \int_a^b dg^c\,x &+ (g(a+0) - g(a))\,x(a+0)\\ &+ (g(b) - g(b-0))\,x(b-0) + \sum_{\tau\in J(g)} \Delta g(\tau)\cdot[x(\tau-0), x(\tau+0)]\end{aligned}\tag{3.3}$$

and

$$\begin{aligned}\int_a^b dg\,x = \int_a^b dg\,x^c &+ (g(b) - g(a))\cdot(x(a+0) - x(a))\\ &+ g(b)\cdot\sum_{\tau\in J(x)}\Delta x(\tau) - \sum_{\tau\in J(x)}[g(\tau-0), g(\tau+0)]\cdot\Delta x(\tau).\end{aligned}\tag{3.4}$$

Corollary 3.12. *The integral $\int_a^b dg\,x$ is singlevalued if x and g have no common discontinuity points.*

Remark 3.13. By the formula of integration by parts we conclude that the properties of the multivalued integral as a function of x and g are almost symmetric with respect to the transposition of x and g. All the distinction is that the set $\int_a^b dg\,x$ doesn't depend on the value $x(a)$ even for x and g discontunuous at a, but it depends on $g(a)$ in any case.

Using (3.3), by direct calculation we get the following statement.

Theorem 3.14. *If $c \in (a,b)$ then*

$$\begin{aligned}\int_a^b dg\,x + (g(c)-g(c-0))x(c-0) + (g(c+0)-g(c))x(c+0)\\ = \int_a^c dg\,x + \int_c^b dg\,x + (g(c+0)-g(c-0))\cdot[x(c-0),x(c+0)],\end{aligned}$$

or, in other words,

$$\int_a^b dg\,x = \int_a^c dg\,x + \int_c^b dg\,x + [(g(c)-g(c+0)),(g(c)-g(c-0))]\cdot\Delta x(c). \quad (3.5)$$

It is obvious that multiplying one of the functions x or g by a scalar constant we multiply the integral $\int_a^b dg\,x$ by the same constant. Using the equality (3.3), we get formulas for addition of functions.

Theorem 3.15. *For $x, y, g, f \in \mathbf{BV}[a,b]$ the following inclusions hold:*

$$\int_a^b dg(t)\,(x(t)+y(t)) \subset \int_a^b dg(t)\,x(t) + \int_a^b dg(t)\,y(t) \quad (3.6)$$

and

$$\int_a^b d(g(t)+f(t))\,x(t) \subset \int_a^b dg(t)\,x(t) + \int_a^b df(t)\,x(t). \quad (3.7)$$

If $J(x)\cap J(y)\cap J(g)=\emptyset$ or $J(x)\cap J(g)\cap J(f)=\emptyset$ then (3.6) or, correspondingly, (3.7) turns into the equality.

Remark 3.16. We can easily obtain the counterexample for the condition of no common discontinuity points on the following way. Suppose all three functions in (3.6) have unique common discontinuity in some point in (a,b); then the right of (3.6) is in general multivalued. But the jumps of addends x and y may eliminate one another; thus the left of (3.6) becomes singlevalued.

As an addendum to [11], we give here another condition for equality.

Theorem 3.17. *Suppose that $x(t)-\alpha y(t)$ is continuous in (a,b) for some $\alpha>0$. Then for positive λ and μ*

$$\int_a^b dg(t)\,(\lambda x(t)+\mu y(t)) = \lambda\int_a^b dg(t)\,x(t) + \mu\int_a^b dg(t)\,y(t). \quad (3.8)$$

Proof. Rewrite (3.4) in the form

$$\int_a^b dg\,x = \int_a^b dg\,x^c + (g(b)-g(a))\cdot(x(a+0)-x(a)) + M,$$

where

$$M = g(b)\cdot\sum_{\tau\in J(x)}\Delta x(\tau) - \sum_{\tau\in J(x)}[g(\tau-0),g(\tau+0)]\cdot\Delta x(\tau).$$

Since αy has the same jumps that x has,

$$\int_a^b dg\, y = \int_a^b dg\, y^c + (g(b) - g(a)) \cdot (y(a+0) - y(a)) + \frac{1}{\alpha} M$$

with the same set M. The central point of the proof is that, due to convexity of M, $\lambda M + \frac{\mu}{\alpha} M = (\lambda + \frac{\mu}{\alpha}) M$. Thus

$$\begin{aligned}\lambda \int_a^b dg\, x + \mu \int_a^b dg\, y = \int_a^b dg\, (\lambda x + \mu y)^c + (g(b) - g(a)) \cdot (\lambda x(a+0) + \mu y(a+0)) \\ - (\lambda x(a) + \mu y(a)) + (\lambda + \frac{\mu}{\alpha}) M.\end{aligned}$$

To complete the proof, it remains to note that $(\lambda + \frac{\mu}{\alpha}) \Delta x(\tau) = \Delta(\lambda x + \mu y)(\tau)$. $\square$

Theorem 3.18. *Suppose $x_n(t) \to x(t)$ and $g_n(t) \to g(t)$ uniformly on $[a,b]$, at least one of the sequences $\{\mathrm{Var}_a^b g_n\}$ and $\{\mathrm{Var}_a^b x_n\}$ is bounded; then*

$$\alpha\Big(\int_a^b dg_n\, x_n, \int_a^b dg\, x\Big) \to 0.$$

Theorem 3.19. *a) Suppose that $g_n \xrightarrow{\beta} g$ and $x_n(t) \to x(t)$ uniformly on $[a,b]$. Then*

$$\lim_{n\to\infty} \beta\Big(\int_a^b dg_n\, x_n, \int_a^b dg\, x\Big) = 0. \tag{3.9}$$

b) Let $x_n \xrightarrow{\beta} x$, $g_n(t) \to g(t)$ uniformly on $[a,b]$. If at each end of $[a,b]$ at least one of the functions x and g is continuous then (3.9) holds.

Remark 3.20. The deviation $\beta\big(\int_a^b dg\, x\,, \int_a^b dg_n\, x\big)$ may not tend to 0 (see Example 3.9). In other words, this means that dependence of the integral on the function g is upper semicontinuous but not continuous.

Next, we study the properties of the multivalued function

$$X(t) = \begin{cases} \int_a^t dg\, x, & t \in (a,b], \\ \{0\}, & t = a. \end{cases} \tag{3.10}$$

Of course, the properties of the multivalued integral as a function of its lower limit of integration are analogous.

Definition 3.21. We say that X is a function of bounded ***variation*** on $[a,b]$ if

$$\mathrm{Var}_a^b X \overset{\mathrm{df}}{=} \sup \sum_{i=1}^n \alpha(X(t_i), X(t_{i-1})) < \infty,$$

where supremum is taken on the range of all finite partitions $a = t_0 < t_1 < \dots < t_n = b$.

Theorem 3.22. *The multivalued function X defined by (3.10) has a bounded variation on $[a,b]$.*

Theorem 3.23. *The function X has left-handed limit at every point in $(a,b]$ and right-handed limits at all points in $[a,b)$.*

Remark 3.24. To prove the theorem, we use in [11] only that X is a function of bounded variation with values in a complete metric space.

4. The functional differential equation with measure input. General assumptions and accessory results

The main subject of this paper is linear functional differential equation

$$\dot{x}(t) = \int_a^t d_s R(t,s)\, x(s) + F'(t), \quad t \in [a,b], \tag{4.1}$$

where the matrix function $R : \Delta \to \mathbb{R}^{N\times N}$ has a bounded variation as a function of s, $\Delta = \{(t,s) : t \in [a,b], s \in [a,t]\}$, F' is a derivative of discontinuous function $F \in \mathbf{BV}[a,b]$ of bounded variation, desired solution $x \in \mathbf{BV}[a,b]$ has values in $\mathbb{R}^N$.

Everywhere in what follows we assume that the function R satisfies the following conditions:

1) R is measurable on Δ;
2) for a.e. $t \in [a,b]$, the variation $\mathrm{Var}_a^t R(t,\cdot) \le v(t)$, where v is summable function on $[a,b]$; the functions $R(t,\cdot)$ are proper;
3) the functions $R(\cdot,s)$ are summable on $[s,b]$ for every $s \in [a,b]$; the mapping $t \mapsto R(t,t)$ is summable on $[a,b]$.

Remark 4.1. Under these conditions, the initial value problem for equation (1.1) with summable f has a unique solution (see, for instance, [4]). The conditions "functions $R(t,\cdot)$ are proper" and "functions $R(\cdot,s)$ are summable for **every** s" are not required for the solvability of (1.1). But by necessity we fulfil them easily by little corrections of the function R such that the integrals $\int_a^t d_s R(t,s)\, x(s)$ do not change for continuous functions x.

4.1. **Functional differential inequality.** Denote $r(t,s) = \mathrm{Var}_a^s R(t,\cdot)$. Since the functions $R(t,\cdot)$ are proper and $R(\cdot,s)$ summable, r is measurable with respect to $(t,s) \in \Delta$ and the mapping $t \mapsto r(t,t)$ is summable. Besides, $r(t,a) = 0$ and $\mathrm{Var}_a^t r(t,\cdot) \le v(t)$. Therefore, the problem

$$\dot{z}(t) = \int_a^t d_s r(t,s)\, z(s) + \varphi(t), \quad z(a) = z_0, \tag{4.2}$$

has a unique solution (see, for instance, [4]).

The following lemma on inequality is proved in the usual way but may be unknown.

Lemma 4.2. *Suppose x is a solution of (1.1) with summable input f, z is the solution of the problem (4.2) with $z_0 > |x(a)|$ and summable $\varphi(t) \ge |f(t)|$. Then*

$$|x(t)| \le z(t), \quad t \in [a,b].$$

Proof. Assume the converse and let $\tau = \inf\{t \in [a,b] : |x(t)| > z(t)\}$. Then $\tau \in (a,b]$ and $|x(\tau)| = z(\tau)$. For $s \in [a,\tau]$ we have $|x(s)| \le z(s)$. Hence for $t \in [a,\tau]$,

$$\frac{d}{dt}|x(t)| \le \Big| \int_a^t d_s R(t,s)\, x(s) \Big| + |f(t)| \le \int_a^t d_s r(t,s)\, z(s) + \varphi(t) = \dot{z}(t).$$

Therefore,

$$|x(\tau)| = |x(a)| + \int_a^\tau \frac{d}{dt}|x(t)|\, dt < z(a) + \int_a^\tau \dot{z}(t)\, dt = z(t).$$

This contradiction proves the lemma. □

4.2. **Multivalued integral is summable.**

Lemma 4.3. *Let*

1) *the function $R(t,s)$ be measurable in $[a,b]\times[c,d]$;*
2) *the variation $v(t)=\operatorname{Var}_c^d R(t,\cdot)$ finite for a.e. $t\in[a,b]$;*
3) *the function v, mappings $t\mapsto R(t,c)$ and $t\mapsto R(t,d)$ summable on $[a,b]$.*

Then, if $x\in\mathbf{BV}[c,d]$, the function $X(t)=\int_c^d d_sR(t,s)\,x(s)$ is boundedly summable on $[a,b]$.

First let x be continuous.

Proposition 4.4. *If the conditions of the Lemma are fulfilled and x is continuous on $[c,d]$ then $X\in\mathbf{L}_1(a,b)$.*

Proof. Since R is summable on $[a,b]\times[c,d]$, the mapping $t\mapsto R(t,s)$ is measurable for almost every $s\in[c,d]$ (this follows from the Fubini theorem). Therefore we may assume that the integral sums are measurable and their limit $\int_c^d d_sR(t,s)\,x(s)$ is measurable too. The estimate $\|X(t)\|\le v(t)\cdot\max_{s\in[c,d]}|x(s)|$ completes the proof of the Proposition. □

Proof of Lemma 4.3. Substitute $R(t,\cdot)$ for g in (3.4). Due to Proposition 4.4 the integral $\int_c^d d_sR(t,s)\,x^c(s)$ is measurable with respect to t; the sum of the series

$$\sum_{\tau\in J(x)}[R(t,\tau-0),R(t,\tau+0)]\cdot\Delta x(\tau)$$

is measurable according to the properties of measurable multivalued functions mentioned above. Thus X is measurable. The proof is completed by the estimate $\|X(t)\|\le\operatorname{Var}_c^d R(t,\cdot)\cdot\max_{s\in[c,d]}|x(s)|$. □

5. Solutions of general type

Now we start the study of the functional differential equation with a measure input

$$\dot x(t)=\int_a^t d_sR(t,s)\,x(s)+F'(t),\quad t\in[a,b], \tag{4.1}$$

where $F\in\mathbf{BV}[a,b]$.

Definition 5.1. The function $x\in\mathbf{BV}[a,b]$ is a ***solution*** of the equation (4.1) on $[a,b]$ if the function $y(t):=x(t)-F(t)$ is absolutely continuous and

$$\dot y(t)\in\int_a^t d_sR(t,s)\,x(s)$$

a.e. on $[a,b]$.

If F is absolutely continuous then the latter inclusion turns into the equation equivalent to the ordinary functional differential equation with aftereffect (1.1) with $f=F'$. The unique solvability of the initial value problem for (1.1) is well known; see, for instance, [4].

Remark 5.2. For every solutuion $\tilde x$ of (4.1) there is a corresponding solution x of the problem

$$\dot x(t)=\int_a^t d_sR(t,s)\,x(s)+\Phi'(t),\quad x(a)=\tilde x(a)+(F(a+0)-F(a))$$

with the function Φ defined by the formulas $\Phi(a) = F(a+0)$, $\Phi(t) = F(t)$ for $t > a$, and continuous at the point a. The solution x differs from $\tilde{x}$ only at the point a.

That is why in the sequel we can consider by necessity F being continuous from the right at a.

Besides, the function F and, correspondingly, the solutions x may be considered to be proper individual functions in stead of equivalence classes.

Example 5.3. Consider the equation (4.1) on $[0, 3]$ with the data:

$$R(t,s) = \begin{cases} 0, & s < 1, \\ 1, & s > 1,\ 1 < t < 2, \\ -1, & s > 1,\ t > 2, \end{cases} \quad F(t) = \begin{cases} 0, & t < 1, \\ 1, & t > 1. \end{cases}$$

For this equation, the solutions x that satisfy the initial condition $x(0) = 0$ are given by the formula

$$x(t) = \begin{cases} 0, & t < 1, \\ 1 + \int_1^t \alpha(s)\,ds, & 1 < t < 2, \\ 1 + \int_1^2 \alpha(s)\,ds - \int_2^t \alpha(s)\,ds, & t > 2, \end{cases}$$

with arbitrary summable α, $\alpha(t) \in [0, 1]$.

Remark 5.4. The ***section of the integral funnel*** of the initial-value problem $H(t) = \{x(t) : x \text{ is a solution}\}$ for this example is represented as follows:

$$H(t) = \begin{cases} 0, & t < 1, \\ [1, t], & 1 < t \le 2, \\ [3 - t, 2], & t > 2. \end{cases}$$

Theorem 5.5. *The function $x \in \mathbf{BV}[a, b]$ is a solution of the equation (4.1) if and only if there exists $\varphi \in \mathbf{L}_1(a, b)$ such that for a.e. $t \in [a, b]$*

$$\varphi(t) \in \int_a^t d_sR(t,s)\,F(s),$$

and $y(t) = x(t) - F(t)$ is a solution of the equation

$$\dot{y}(t) = \int_a^t d_sR(t,s)\,y(s) + \varphi(t).$$

Proof. Let x be a solution of (4.1). Due to Proposition 4.4 the function $\varphi(t) = \dot{y}(t) - \int_a^t d_sR(t,s)\,y(s)$ belongs to $\mathbf{L}_1(a, b)$. According to Theorem 3.15 for a.e. $t \in [a, b]$ we have

$$\varphi(t) \in \int_a^t d_sR(t,s)\,x(s) - \int_a^t d_sR(t,s)\,y(s) = \int_a^t d_sR(t,s)\,F(s).$$

Now let φ satisfy stated conditions. Then the function y is absolutely continuous and for a.e. $t \in [a, b]$ the equality

$$\dot{y}(t) \in \int_a^t d_sR(t,s)\,y(s) + \int_a^t d_sR(t,s)\,F(s) = \int_a^t d_sR(t,s)\,x(s)$$

holds (the latter equlity is due to Theorem 3.15 again). Thus x is a solution of (4.1). $\square$

Remark 5.6. In other words, the solutions of the certain initial value problem for (4.1) are in one-to-one correspondence with summable selectors of the multivalued function $\int_a^t d_sR(t,s)\,F(s)$.

If the latter integral is a.e. singlevalued then the selector φ is defined uniquely.

Corollary 5.7. *Consider the set of the points t such that the functions F and $R(t,\cdot)$ have common points of discontinuity in (a,t). If its measure is equal to 0 then the initial value problem for the equation (4.1) has a unique solution.* $\square$

Consider the sequence of the equations

$$\dot{x}(t) = \int_a^t d_sR(t,s)\,x(s) + F_n'(t). \tag{5.1}$$

Lemma 5.8. *Let F_n β-converge to F, x_n be solutions of (5.1), and the sequence $x_n(a)$ be bounded. Then*

a) *the sequence of the functions $y_n(t) := x_n(t) - F_n(t)$ has a subsequence uniformly convergent on $[a,b]$;*

b) *if $y_n(t) \to y(t)$ uniformly on $[a,b]$ then $x(t) := y(t) + F(t)$ is a solution of the equation (4.1).*

Proof. According to Theorem 5.5, absolutely continuous functions y_n satisfy the equations

$$\dot{y}_n(t) = \int_a^t d_sR(t,s)\,y_n(s) + \varphi_n(t),$$

where $\varphi_n(t) \in \int_a^t d_sR(t,s)\,F_n(s)$. By Lemma 4.2 we have

$$|y_n(t)| < z(t), \quad t \in [a,b].$$

where z is a unique solution of the problem (4.2) with $\varphi(t) = \sup_{s\in[a,b],n} |F_n(s)|\cdot v(t)$ and and $z_0 > \sup_n |y_n(a)|$. So the sequence $\{y_n\}$ is uniformly bounded. Due the estimate

$$|\dot{y}_n(t)| \le \sup_{s\in[a,b]} z(s)\cdot v(t) + \varphi(t), \tag{5.2}$$

the sequence is equicontinuous. This establishes the first assertion of the theorem.

To prove the second assertion, note that since the absolute continuity of the sequence $\{y_n\}$ is equipower, y is absolutely continuous. Suppose, without loss of generality, that F is continuous at the point a, then the solution x is too. The functions x_n β-converge to x on the almost every segment $[a,t]$ (Lemma 3.3). By Theorem 3.19,

$$\beta\Big(\int_a^t d_sR(t,s)\,x_n(s),\ \int_a^t d_sR(t,s)\,x(s)\Big) \to 0.$$

According to Lemma 4.3 (on summability) and Lebesgue-type theorem on convergence of integrals for multifunctions (Lemma 2.1), for $t_0, t_1 \in [a,b]$, $t_0 < t_1$, we have

$$\beta\Big(\int_{t_0}^{t_1}\int_a^t d_sR(t,s)\,x_n(s)\,dt,\ \int_{t_0}^{t_1}\int_a^t d_sR(t,s)\,x(s)\,dt\Big) \to 0.$$

Hence

$$\begin{aligned} y(t_1) - y(t_0) &= \lim_{n\to\infty} (y_n(t_1) - y_n(t_0)) \\ &= \lim_{n\to\infty} \int_{t_0}^{t_1} \int_a^t d_s R(t,s)\, x_n(s)\, dt \\ &\in \int_{t_0}^{t_1} \int_a^t d_s R(t,s)\, x(s)\, dt. \end{aligned}$$

Therefore, due to Lemma 2.2, for a.e. t we have

$$\dot{y}(t) \in \int_a^t d_s R(t,s)\, x(s).$$

This implies that x is a solution of (4.1). □

As a consequence we obtain the following theorem.

Theorem 5.9. *Let K be compact subset of $\mathbb{R}^N$. Then the set of the solutions x of (4.1) with $x(a) \in K$ is compact in the sense of the uniform convergence on $[a,b]$.* □

Since the values of the multifunction $\int_a^t d_s R(t,s)\, F(s)$ are convex sets (Proposiion 3.10), Theorem 5.5 easily implies the following statement.

Theorem 5.10. *Let K be convex subset of $\mathbb{R}^N$. Then the set of solutions x of the equation (4.1) with $x(a) \in K$ is convex.* □

Another way to proof this statement is to refer directly to Definition 5.1 and Theorem 3.17.

6. Approximative solutions and solutions with memory

Now we consider narrow notions of solution for equation (4.1).

6.1. **Approximative solutions.**

Definition 6.1. The function x is ***approximative solution*** of (4.1) if there exist a sequences of absolutely continuous functions F_n and solutions x_n of equations (5.1) such that

1) $F_n \xrightarrow{\beta} F$,
2) $x_n(t) - F_n(t) \to x(t) - F(t)$ uniformly on $[a,b]$.

The following example demonstrates that the solution of (4.1) may not be approximative.

Example 6.2. Consider the following solution from Example 5.3:

$$x(t) = \begin{cases} 0, & t < 1, \\ 1, & 1 < t < 2, \\ 3 - t, & t > 2, \end{cases}$$

that is obtained by virtue of

$$\alpha(t) = \begin{cases} 0, & t < 2, \\ 1, & t > 2. \end{cases}$$

This solution is not approximative. Indeed, if it is approximative and F_n, $F_n(0) = 0$, are absolutely continuous functions stated in Definition 6.1 then the corresponding solutions are

$$x_n(t) = \begin{cases} F_n(t), & t \le 1, \\ F_n(t) + F_n(1)(t-1), & 1 < t \le 2, \\ F_n(t) + F_n(1)(3-t), & t > 2. \end{cases} \tag{6.1}$$

In particular, $x_n(3) = F_n(3)$. Thus $0 = \lim_{n\to\infty}(x_n(3) - F_n(3)) = x(3) - F(3) = -1$. We have got a contradiction. □

Remark 6.3. This example demonstrates the existence of the points of the integral funnel not attainable for approximative solutions. From (6.1) we have that the section of the integral funnel of approximative solutions of the initial value problem $H_a(t) = \{x(t) : x \text{ is approximative solution}\}$ is as follows:

$$H_a(t) = \begin{cases} 0, & t < 1, \\ [1, t], & 1 < t \le 2, \\ [1, 4-t], & 2 < t \le 3, \\ [4-t, 1], & t > 3. \end{cases}$$

Proposition 3.5 and Lemma 5.8 imply the following theorem.

Theorem 6.4. a) *The initial value problem for (4.1) has an approximative solution.*
b) *The approximative solution is a solution (in general sense).*

The first assertion of the next theorem follows directly from the first assertion of the Lemma 5.8. The easy proof of the second one is well known.

Theorem 6.5. *Let $F_n \in \mathbf{BV}[a,b]$ β-converge to $F \in \mathbf{BV}[a,b]$, x_n be approximative solutions of (5.1) and the sequence $\{x_n(a)\}$ is bounded. Then*

a) *the sequence of the functions $y_n(t) := x_n(t) - F_n(t)$ is compact in the sense of the uniform convergence on $[a,b]$;*
b) *if $y_n(t) \to y(t)$ uniformly on $[a,b]$ then $x(t) := y(t) + F(t)$ is an approximative solution of (4.1).* □

As a particular case, we have the following statement.

Theorem 6.6. *If K is a compact subset of $\mathbb{R}^N$ then the set of approximative solutions x of the equation (4.1) with $x(a) \in K$ is compact in the sense of the uniform convergence on $[a,b]$.* □

The convexness of the set of approximative solutions takes place too.

Theorem 6.7. *If K is a convex subset of $\mathbb{R}^N$, then the set of approximative solutions x of the equation (4.1) with $x(a) \in K$ is convex.*

Proof. Let x^0 and x^1 be approximative solutions of (4.1), x_i^0 and x_i^1 the solutions of the equations with inputs F_i^0 and, respectively, F_i^1 such as it is described in the Definition 6.1. Since $F_i^0 \xrightarrow{\beta} F$ and $F_i^1 \xrightarrow{\beta} F$, due to Lemma 3.4, the functions $\lambda F_i^1 + (1-\lambda)F_i^0$ ($\lambda \in (0,1)$) β-converge to F as $i \to \infty$. For continuous functions

x_i^k, $k = 0, 1$, the integral has a linearity property. Thus, for the functions $y_i = \lambda x_i^1 + (1-\lambda)x_i^0 - F = \lambda(x_i^1 - F) + (1-\lambda)(x_i^0 - F)$ we get a.e.

$$\dot{y}_i(t) = \int_a^t d_s R(t,s)\,(x_i^1 + (1-\lambda)x_i^0)(s)$$

and

$$\begin{aligned}\lim_{i\to\infty} y_i(t) &= \lambda \lim_{i\to\infty}(x_i^1(t) - F(t)) + (1-\lambda)\lim_{i\to\infty}(x_i^0(t) - F(t)) \\ &= \lambda x^1(t) + (1-\lambda)x^0(t) - F(t)\end{aligned}$$

uniformly on $[a, b]$. □

Remark 6.8. If we replace β-convergence F_n to F by the normal convergence in Definition 6.1 then we get another definition of solution that is, a priori, stronger. We shal see soon that as a matter of fact this new notion coincides with the old one.

6.2. **Solutions with memory.** Consider now the following property of the solution that may be important for applications. Suppose the Stieltjes measure $d_s R(t,s)$ has an atom at the point $s = \tau$. So the system modelled by equation (4.1) has on-board device for measuring the value $x(\tau)$ to determine the behaviour of the system. The property considered is that at any moment of time in the future, whenever the value $x(\tau)$ is needed, just the same obtained value is used. The solution from Example 6.2 do not has this property: the value $x(1)$ is measured by two independent devices, the value obtained by first one is valid during $t \in (1, 2)$ and the value of the second is for $t \in (2, 3)$.

In this subsection we assume that the function F is continuous from the right at the point a.

Definition 6.9. $x \in \mathbf{BV}[a, b]$ is ***solution with memory*** of (4.1) if it has a proper representative $\bar{x}$ such that the function $y(t) := x(t) - F(t)$ satisfy a.e. the equality

$$\dot{y}(t) = (\mathrm{S})\int_a^t d_s R(t,s)\,\bar{x}(s), \tag{6.2}$$

where $(\mathrm{S})\int$ denotes a Lebesgue-Stieltjes integral.

Theorem 6.10. *For $x \in \mathbf{BV}[a, b]$, the following conditions are equivalent:*

a) *x is approximative solution of equation (4.1);*
b) *there exist sequences of absolutely continuous functions F_n and solutions x_n of equations (5.1) such that $F_n \xrightarrow{N} F$ and $x_n(t) - F_n(t) \to x(t) - F(t)$ uniformly on $[a, b]$;*
c) *x is a solutions with memory of (4.1).*

Proof. **c) $\Rightarrow$ b).** Let x be a solution with memory. Define $y(t) = x(t) - F(t)$ and $\bar{F}(t) = \bar{x}(t) - y(t)$. The function $\bar{F}$ has just the same behaviour at the points of discontinuity as $\bar{x}$ has.

Define the approximations l_δ of the function $\bar{F}$ as it was made in the proof of the Proposition 3.5. Then we have $l_\delta \xrightarrow{N} F$ as $\delta \to 0$. Let x_δ be solutions of the equations

$$\dot{x}(t) = \int_a^t d_s R(t,s)\,x(s) + l'_\delta(t) \tag{6.3}$$

with the same initial value that x has, and $y_\delta = x_\delta - l_\delta$.

Now let δ tend to 0 ranging over countable set of values. For every point $\tau \in J(\bar F)$ we have $l_\delta(\tau) = \bar F(\tau)$ as δ is sufficiently small. Thus $x_\delta(s) \to \bar x(s)$ for all $s \in [a,b]$. The proof of Lemma 5.8 shows that the functions x_δ are uniformly bounded. Hence

$$\dot y_\delta(t) = (\mathrm{S})\int_a^t d_sR(t,s)\,x_\delta(s) \to (\mathrm{S})\int_a^t d_sR(t,s)\,\bar x(s) = \dot y(t)$$

for a.e. $t \in [a,b]$. The derivatives $\dot y_\delta(t)$ have common summable majorant (see the proof of Lemma 5.8); therefore according to the Lebesgue theorem $y_\delta(t) \to y(t)$ for all $t \in [a,b]$. Due to Lemma 5.8a) we may find such a subsequence $\delta \to 0$ that the convergence $y_\delta(t) \to y(t)$ along this subsequence is uniform on $[a,b]$.

b) $\Rightarrow$ a). It is obvious by definition.

a) $\Rightarrow$ c). Let x be an approximative solution and F_n the functions described in Definition 6.1. Then the equality $\lim_{n\to\infty} F_n(t) = F(t)$ holds for all points of continuity of F. Take $\tau \in J(F)$. According to Lemma 3.2a) there exists a subsequence such that being renumerated it has the limit

$$\lim_{n\to\infty} F_n(\tau) \in [F(\tau-0), F(\tau+0)]. \tag{6.4}$$

Using diagonal process we get (6.4) for all $\tau \in J(F)$. Denote

$$\bar F(t) = \lim_{n\to\infty} F_n(t), \quad y(t) = x(t) - F(t), \quad \bar x(t) = y(t) + \bar F(t), \tag{6.5}$$

then $\bar x(t) = x(t)$ at the points of continuity and $\bar x(\tau) \in [x(\tau-0), x(\tau+0)]$ for $\tau \in J(F) = J(x)$.

Let x_n be the solutions of equations (5.1) such as described in Definition 6.1. Due to (6.5), $\bar x(s) = \lim_{n\to\infty} x_n(s)$ for all $s \in [a,b]$. Hence

$$\int_a^t d_sR(t,s)\,x_n(s) \to (\mathrm{S})\int_a^t d_sR(t,s)\,\bar x(s)$$

and the latter integral is summable function of t. Denote $y_n(t) = x_n(t) - F_n(t)$; then $y_n(t) \to y(t)$ for $t \notin J(F)$. If $t_2 > t_1$ then

$$\begin{aligned} y(t_2) - y(t_1) &= \lim_{n\to\infty}(y_n(t_2) - y_n(t_1)) \\ &= \lim_{n\to\infty}\int_{t_1}^{t_2}\int_a^t d_sR(t,s)\,x_n(s)\,dt \\ &= \int_{t_1}^{t_2}(\mathrm{S})\int_a^t d_sR(t,s)\,\bar x(s)\,dt, \end{aligned}$$

which implies the equality (6.2) for a.e. t. Thus x is a solution with memory. □

7. The Cauchy formula

7.1. **The Cauchy function.** It is well known that the solutions of the equation (1.1) with summable f obey the Cauchy formula

$$x(t) = C(t,a)x(a) + \int_a^t C(t,s)f(s)\,ds, \tag{7.1}$$

where $C(t,s)$ is the Cauchy function.

We describe the construction of the Cauchy function following [4, 5]. Integration by parts in (1.1) gives

$$\dot{x}(t) = \int_a^t K(t,s)\dot{x}(s)\,ds + K(t,a)x(a) + f(t),$$

where $K(t,s) = R(t,t) - R(t,s)$. Note that $K(t,t) = 0$.

The latter equation is uniquely solvable as the Volterra equation of second kind in unknown $\dot{x}$. Integrating, we get the solution. In such a way the formula (7.1) is obtained, where

$$C(t,s) = E + \int_s^t P(\tau,s)\,d\tau, \tag{7.2}$$

E is the unit matrix, $P(t,s)$ the kernel of the Volterra integral operator $\mathbf{P}$,

$$I + \mathbf{P} := (I - \mathbf{K})^{-1} = I + \mathbf{K} + \mathbf{K}^2 + \dots + \mathbf{K}^n + \dots, \quad (\mathbf{K}z)(t) = \int_a^t K(t,s)z(s)\,ds.$$

The kernel P may be found as a sum of Neumann's series

$$P(t,s) = \sum_{n=1}^{\infty} K_n(t,s), \tag{7.3}$$

where iterated kernels K_n are defined by equalities

$$K_1(t,s) = K(t,s),$$
$$K_{n+1}(t,s) = \int_s^t K_n(t,\tau)K(\tau,s)\,d\tau, \quad n \ge 1.$$

To argue that all these calculations are well made, we bring out the following proposition. Denote by $\mathfrak{A}$ the subset of $[a,b]$ of comlete measure containing all points t such that the variation $\mathrm{Var}_a^t R(t,\cdot) \le v(t) < \infty$.

Proposition 7.1. *The following properties of the iterated kernels take place:*

a) *the functions K_n are summable on Δ; the iterative formula may be rewritten in the form*

$$K_{n+1}(t,s) = \int_s^t K(t,\tau)K_n(\tau,s)\,d\tau;$$

b) *the functions $K_n(\cdot,s)$ are summable on $[s,b]$ for every $s \in [a,b]$; in particular, the functions $K_n(\cdot,a)$ are summable;*

c) $K_n(t,t) = 0$;

d) *for $t \in \mathfrak{A}$ and $s \in [a,b]$ we have*

$$|K_n(t,s)| \le \frac{v(t)}{(n-1)!}\Big(\int_s^t v(\tau)\,d\tau\Big)^{n-1};$$

e) *for $t \in \mathfrak{A}$ the variation*

$$\mathrm{Var}_a^t K_n(t,\cdot) \le \frac{v(t)}{(n-1)!}\Big(\int_a^t v(\tau)\,d\tau\Big)^{n-1}.$$

All the assertions in this proposition are easily verified by mathematical induction with use of the Fubini theorem on multiple integral. □

Hence for a.e. t the Neumann series absolutely converges uniformly with respect to s. Integrating its sum as consistent with (7.2) we obtain the following theorem.

Theorem 7.2 (see also [4, 5])**.** *The following properties of the Cauchy function $C(t,s)$ take place:*

a) *C is summable on Δ;*
b) *the function $C(\cdot,s)$ is absolutely continuous on $[s,b]$ for every $s \in [a,b]$;*
c) *$C(t,t) = E$;*
d) *for all $(t,s) \in \Delta$,*
$$|C(t,s)| \le \exp\Big(\int_s^t v(\tau)\,d\tau\Big);$$
e) *for every $t \in [a,b]$, the function $C(t,\cdot)$ has finite variation on $[a,t]$, which is uniformly bounded respective to t:*
$$\mathrm{Var}_a^t C(t,\cdot) \le \exp\Big(\int_a^t v(\tau)\,d\tau\Big).$$

□

7.2. **The Cauchy formula for equations with measure input.** We suppose here that the function F is continuous at the point a. Likewise to (7.1), we can write the formula
$$X(t) = C(t,a)\,x(a) + \int_a^t C(t,s)\,dF(s). \tag{7.4}$$
In virtue of Theorem 7.2 the integral here is well defined multivalued Stieltjes integral. Hence the function X is in general multivalued and its values are convex compact sets. Integrating by parts due to Theorem 3.10 we get
$$X(t) = C(t,a)\,(x(a) - F(a)) + F(t) - \int_a^t d_s C(t,s)\,F(s)$$
for a.e. $t \in [a,b]$. Thus, according to Lemma 4.3, the multifunction X is boundedly summable.

Let x be approximative solution, F_n and x_n corresponding absolutely continuous functions and solutions such as stated in Definition 6.1. We have
$$x_n(t) = C(t,a)\,x_n(a) + \int_a^t C(t,s)\,dF_n(s). \tag{7.5}$$
For $t \notin J(F))$, since $F_n \xrightarrow{\beta} F$ on $[a,t]$, according to Theorem 3.19, we have
$$\rho\Big(\int_a^t C(t,s)\,dF_n(s),\ \int_a^t C(t,s)\,dF(s)\Big) \to 0 \quad \text{as } n \to \infty.$$
Therefore, $\rho\,(x_n(t), X(t)) \to 0$. The set $X(t)$ is closed, hence
$$x(t) = \lim_{n\to\infty} x_n(t) \in X(t)$$
for all $t \notin J(x)$. So we get the following statement.

Theorem 7.3. *Let F be continuous from the right at the point a. Then the approximative solution x of equation (4.1) satisfies the inclusion*
$$x(t) \in C(t,a)\,x(a) + \int_a^t C(t,s)\,dF(s) \tag{7.6}$$
for all points t of continuity of F. □

Remark 7.4. The Cauchy function of Examples 5.3 and 6.2 is as follows:

$$C(t,s) = \begin{cases} 1, & s < 1,\ t < 1, \\ t, & s < 1,\ 1 < t < 2, \\ 4 - t, & s < 1,\ t > 2, \\ 1, & s > 1. \end{cases}$$

Using this equality, we obtain from (7.4) that $X(t) = H_a(t)$. The expression for $H_a(t)$ is given in Remark 6.3. It was found in Remark 5.4 that the set $H(t)$ for $t > 2$ contains points not belonging to $H_a(t)$. Thus the inclusion (7.6) may be violated for the solutions of general type.

The arbitrary selector x of the multifunction X (that is $x(t) \in X(t)\ \forall t \in [a,b]$) such that the difference $x - F$ is absolutely continuous is not necessarily a solution of (4.1). Indeed, the solution x in Example 5.3 is nondecreasing on $[1,2]$ that is not required for selectors of multifunction H_a.

References

[1] A. V. Anokhin. *On linear impulse systems for functional differential equations*, Doklady Akad. Nauk SSSR, **286** (1986), # 5.
[2] A. V. Anokhin, E. Ya. Braverman. *On definitions of notion of impulse functional differential equations*, Functional differential equations, Perm Polytechnical Institute, Perm, 1989, 35–45.
[3] R. J. Aumann. *Integrals of set-valued functions*, J. Math. Analysis and Appl., **12** (1965), # 1.
[4] N. Azbelev, V. Maksimov, and L. Rakhmatullina. *Introduction to the theory of linear functional differential equations*. World Federation Pub. Comp., Atlanta, 1995.
[5] N. V. Azbelev, P. M. Simonov. *Stability of differential equations with aftereffect*, Taylor and Francis Publ., London, 2002. p. 256
[6] V. I. Blagodatskikh, A. F. Filippov. *Differential inclusions and optimal comtrol*, Topology, differential equations, dinamical systems (Trudy MIAN SSSR, **169** (1985). Nauka, Moscow. Pp. 194–252) (Russian).
[7] Yu. G. Borisovich, B. D. Gelman, A. D. Myshkis, V. V. Obuhovskii; *Multivalued mappings*, Itogi nauki i techniki. Mathematical analisys, **19** (1982). VINITI, Moscow. Pp. 127–230 (Russian).
[8] C. Castaing, M. Valadier; *Convex analysis and measurable multifunctions*, Lect. Notes Math., 580. Springer-Verlag, 1977.
[9] A. Halanay, D. Wexler; *Qualitative theory of impulse systems*, Mir, Moscow, 1971 (Russian).
[10] A. N. Kolmogorov, S. V. Fomin. *Elements of the theory of functions and functional analysis*, 6-th ed. Nauka, Moscow, 1989 (Russian); English transl.: Dover Pubs, 1999.
[11] V. Z. Tsalyuk. *Multivalued Stieltjes integral for discontinuous functions of bounded variation*, Functional Differential Equations, **9** (2002), No 3–4, pp. 551–576.

Vadim Z. Tsalyuk
Mathematics Department, Kuban State University, Stavropol'skaya 149, Krasnodar 350040, Russia
E-mail address: vts@math.kubsu.ru