DOI 10.1007/s10958-022-06051-z
Journal of Mathematical Sciences, Vol. 265, No. 2, August, 2022

MULTIPLICATION OF DISTRIBUTIONS AND ALGEBRAS OF MNEMOFUNCTIONS

A. B. Antonevich and T. G. Shagova

UDC 517.9

Abstract. In this paper, we discuss methods and approaches for definition of multiplication of distributions, which is not defined in general in the classical theory. We show that this problem is related to the fact that the operator of multiplication by a smooth function is nonclosable in the space of distributions. We give the general method of construction of new objects called *new distributions*, or mnemofunctions, that preserve essential properties of usual distributions and produce algebras as well. We describe various methods of embedding of distribution spaces into algebras of mnemofunctions. All ideas and considerations are illustrated by the simplest example of the distribution space on a circle. Some effects arising in study of equations with distributions as coefficients are demonstrated by example of a linear first-order differential equation.

CONTENTS

1. Introduction

The theory of generalized functions (distributions) provides a possibility to solve various problems of mathematical physics and of the theory of linear differential equations with smooth coefficients (see [9, 23, 34, 37]). However, in the framework of this classical theory, one cannot define the product of arbitrary distributions. This is an obstacle for applications to equations with generalized coefficients and to nonlinear problems. This motivates the development of various approaches to the multiplication problem for distributions (see [13–16, 25, 33, 38]). The greatest attention is attracted by [17, 18]. A modification of their construction is proposed in [21], which contains a detail history of the problem as well. The general approach is as follows: for a given space of distributions, we introduce new objects preserving a number of distribution properties and forming algebras, i.e., admitting a well-posed

Translated from Sovremennaya Matematika. Fundamental'nye Napravleniya (Contemporary Mathematics. Fundamental Directions), Vol. 65, No. 3, Proceedings of the Crimean Autumn Mathematical School-Symposium, 2019.

multiplication. These objects are called *new generalized functions*, *mnemofunctions*, or *nonlinear generalized functions.* In [3, 4], a general method of the constructing of such algebras is described. A lot of works are devoted to the history and further development of the theory of such algebras and their applications (see, e.g., [10, 24, 29, 31, 32]); it would not be possible to provide a review of such works.

The goal of the present paper is broader than the discussion of methods related to introducing a meaningful product of generalized functions from a given space. To be concrete, we illustrate general considerations on the space of distributions on the circle, which is a quite simple example. The explained approach is applicable for more complicated spaces of distributions as well.

2. Line Distribution Spaces

First, recall the definition of the space of generalized functions (distributions) on the line. The *space of test functions* $\mathcal{D}(\mathbb{R})$ consists of functions φ infinitely differentiable on $\mathbb{R}$ and compactly supported, i.e., such that $\varphi(x) = 0$ for $|x| > C_\varphi$. In this space, the convergence is introduced as follows: a sequence $\{\varphi_n\}$ converges to the zero function if

(i) there exists C_0 such that $\varphi_n(x) = 0$ for all n provided that $|x| > C_0$;

(ii) the sequence $\{\varphi_n^{(j)}(x)\}$ of the derivatives of order j uniformly converges to zero for each $j = 0, 1, \dots$

There exists a topology τ on $\mathcal{D}(\mathbb{R})$ such that the convergence in it coincides with the introduced convergence, but the description of this topology is rather cumbersome. Thus, only the notion of the convergence in the specified space is usually used. Elements of the space $\mathcal{D}(\mathbb{R})$ are called *test* functions.

A linear functional f is said to be *continuous* on the space $\mathcal{D}(\mathbb{R})$ if $f(\varphi_n) \to 0$ for each sequence $\{\varphi_n\}$, converging to zero in $\mathcal{D}(\mathbb{R})$. Usually, values of such functionals are denoted as follows: $f(\varphi_n) := < f, \varphi_n >$.

The space adjoint to $\mathcal{D}(\mathbb{R})$, i.e., the set $\mathcal{D}'(\mathbb{R})$ of linear continuous functionals on $\mathcal{D}(\mathbb{R})$, is called the *space of distributions* (*generalized functions*) on $\mathbb{R}$. In this space, the *weak convergence* is introduced: a sequence $\{f_n\}$ of distributions converges to a distribution f_0 if

$$< f_n, \varphi > \to < f_0, \varphi > \quad \text{for each} \quad \varphi \in \mathcal{D}(\mathbb{R}).$$

In the space $\mathcal{D}'(\mathbb{R})$, the *differentiation* is defined by the relation $< f', \varphi > := -< f, \varphi' >$ and the *multiplication* of each $f \in \mathcal{D}'(\mathbb{R})$ by each function $g \in C^\infty(\mathbb{R})$ is defined as follows: $< gf, \varphi > := < f, g\varphi >$.

If u is a locally integrable function, then the relation

$$< f_u, \varphi > = \int u(x)\varphi(x)dx \tag{2.1}$$

defines a distribution such that $f_u \neq 0$ provided that u is different from zero on a set of a positive measure. This implies that the space $L^1_{loc}(\mathbb{R})$ is embedded into the space of distributions. If a distribution can be represented by relation (2.1), then it is said to be *regular*.

A remarkable property of the space of generalized functions is the existence of derivatives of each order for each distribution (including locally integrable functions); these derivatives are generalized functions.

An example of a distribution that is not regular (they are called *singular* ones) is the Dirac delta function, i.e., the functional $< \delta_0, \varphi > = \varphi(0)$, which is the derivative (in the sense of generalized functions) of the discontinuous Heaviside function

$$\Theta(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 0. \end{cases}$$

The function $\frac{1}{x}$ is not locally integrable, but a whole family of (singular) distributions can be associated with it. They are defined as follows. The function $g(x) = \ln|x|$ is differentiable for $x \neq 0$ and $g'(x) = \frac{1}{x}$. Since g is locally integrable, a regular distribution $f_g \in \mathcal{D}'(\mathbb{R})$ corresponds to it. By definition, $\mathcal{P}\left(\frac{1}{x}\right) = f'_g$, where the differentiation is treated in the sense of distributions.

The function $g_C(x) = \ln|x| + C\Theta(x)$ is differentiable for $x \neq 0$ and its derivative is equal to $\frac{1}{x}$ for $x \neq 0$. The derivative of the function $g_C(x)$ in the sense of generalized functions is equal to

$$\mathcal{P}\left(\frac{1}{x}\right) + C\delta_0. \tag{2.2}$$

Thus, family (2.2) of distributions corresponds to the function $\frac{1}{x}$.

The introducing of generalized functions as functionals on the space of test functions is not only a successful mathematical trick; it has a physical interpretation as well. For example, the state of a physical system is frequently described by means of an integrable function $\rho(x)$ expressing the density of the distribution of a numerical characteristic of a substance such as the mass, charge, energy, temperature, etc. There are no physical devices measuring the value $\rho(x)$ at a given point. Real devices measure only averaged values expressed by relations of the kind

$$\int \rho(x)\varphi(x)dx$$

(if the density exists), where the function φ characterizes the particular device (the *device function*). Thus, values that can be measured (only they have physical interpretations!) are values of the functional corresponding to the function $\rho(x)$. If there is no density of the substance distribution, then such values are defined for all functions φ anyway; they define a functional on the set of device functions.

3. Multiplication Problem for Distributions

The multiplication introduced above is not defined for arbitrary distributions. In particular, this means that the space $\mathcal{D}'(\mathbb{R})$ is not a differential algebra. Recall that a vector space G is called a *differential algebra* if

- an associative and commutative multiplication is defined;
- a linear map $G \ni f \to f' \in G$ called the *differentiation* is defined;
- these operations satisfy the relation $(fg)' = f'g + fg'$.

Since the space of distributions is an extension of the space of ordinary functions such that the differentiation operation is defined everywhere in it, Schwartz formulated the following task about the next generalization step: *to extend the space of distributions up to a differential algebra such that the differentiation and multiplication are defined everywhere in it.*

The first version of this Schwartz task was as follows: to construct a differential algebra G and a linear embedding

$$R : \mathcal{D}'(\mathbb{R}) \to G$$

such that the differentiation and multiplication in the space of distributions pass into the corresponding operations in G, i.e., the following relations are satisfied:

$$R(f') = [R(f)]' \tag{3.1}$$

and

$$R(af) = R(a)R(f) \text{ for } a \in C^\infty \text{ and } f \in \mathcal{D}'(\mathbb{R}). \tag{3.2}$$

If such an algebra G is constructed, then the product of arbitrary distributions can be defined as an element of G as follows:

$$f \otimes g := R(f)R(g) \in G, \quad f, g \in \mathcal{D}'(\mathbb{R}).$$

However, it is found in [35] that the introduced multiplication is not associative: for example, the expression

$$\mathcal{P}\Big(\frac{1}{x}\Big) \times x \times \delta_0$$

takes different values under different sets of brackets. Namely,

$$\Big[\mathcal{P}\Big(\frac{1}{x}\Big) \times x\Big] \times \delta_0 = 1 \times \delta_0 = \delta_0,$$

but

$$\mathcal{P}\Big(\frac{1}{x}\Big) \times \big[x \times \delta_0\big] = \mathcal{P}\Big(\frac{1}{x}\Big) \times 0 = 0.$$

This example implies that it is impossible to define an associative and commutative multiplication operation on the whole space $\mathcal{D}'(\mathbb{R})$ and, moreover, this space cannot be embedded into an associative and commutative algebra so that properties (3.1) and (3.2) are fulfilled.

However, formal expressions containing products of multiplications occur in various applied problems, which causes the interest of many researchers to the problem to find products of distributions.

For almost all investigations in this direction, the starting point is as follows. A distribution f can be approximated by a family of smooth functions f_ε. For distributions on the line, the most frequent approximation method is as follows. Let $\psi \in \mathcal{D}(\mathbb{R})$ and $\int_{\mathbb{R}} \psi(t)dt = 1$. Then the family of functions

$$\psi_\varepsilon(t) = \frac{1}{\varepsilon}\psi\Big(\frac{t}{\varepsilon}\Big) \tag{3.3}$$

converges to δ_0 and the supports of the functions ψ_ε shrink to the origin. Let

$$T_x\psi_\varepsilon(t) = \psi_\varepsilon(t - x).$$

Then

$$f_\varepsilon(x) = \langle f, T_x\psi_\varepsilon\rangle \tag{3.4}$$

is a family of smooth functions converging to f as $\varepsilon \to 0$.

Sometimes, instead of families of smooth functions $f_\varepsilon(x)$ depending on the continuously changing positive small parameter ε, it is more convenient to consider sequences $\{f_n\}$ of such functions. If the small parameter takes only the values $\frac{1}{n}$, then the assertions provided below hold for this case as well.

Let $f_\varepsilon(x)$ be a family of smooth functions converging to a distribution f and $g_\varepsilon(x)$ be a similar family for a distribution g. It is natural to define the product of distributions as the limit in the space $\mathcal{D}'(\mathbb{R})$ of the corresponding distributions:

$$f \times g := \lim_{\varepsilon \to 0} f_\varepsilon g_\varepsilon. \tag{3.5}$$

However, such a product is not well defined due to the following two reasons:

(i) such a limit depends on the choice of the approximating families;
(ii) the existence of the limit is not guaranteed.

All the above is illustrated by the following examples.

Example 3.1. In the space of distributions, we have the limit relations $e^{inx} \to 0$ and $e^{-inx} \to 0$. This vanishing is very fast: the sequence $\langle e^{inx}, \varphi\rangle$ decreases faster than each power of $1/n$. However, $e^{inx}e^{-inx} = 1 \to 1$. Then, according to relation (3.5), we obtain that $0 \times 0 = 1$, which has no sense.

Example 3.2. The product $\delta\Theta$ is not defined within the classical theory. Consider such a product under the approximate approach.

Change the δ-function for its approximation of kind (3.3). Change the Θ-function for the function $\Gamma_\varepsilon(x) = \Gamma\Big(\frac{x}{\varepsilon}\Big)$, where

$$\Gamma(x) = \int_{-\infty}^{x} \gamma(s)ds, \quad \gamma \in \mathcal{D}(\mathbb{R}), \quad \int_{\mathbb{R}} \gamma(t)dt = 1.$$

Then the product $\psi_\varepsilon(x)\Gamma_\varepsilon(x)$ converges to $C\delta$ in the space of distributions, where the constant

$$C = \int_{\mathbb{R}} \psi(x)\Gamma(x)dx$$

depends on ψ and γ, i.e., on the selected approximations.

Example 3.3. Family (3.3) converges to the δ-function, but the squares

$$[\psi_\varepsilon(t)]^2 = \frac{1}{\varepsilon^2}\Big[\psi\Big(\frac{t}{\varepsilon}\Big)\Big]^2$$

of these functions have no limit in $\mathcal{D}'(\mathbb{R})$, i.e., this space contains no element to play the role of the square of the δ-function; from this, one can conclude that such a square must belong to a wider space.

Introducing a product of distributions defined everywhere is prevented by properties (i)-(ii) formulated above.

Property (i) reflects the fact that the multiplication operation (on its domain) is a discontinuous map in the topology of the space of distributions and, moreover, this map cannot be closed in the said topology. Property (ii) shows that the space that can contain products of distributions is to be wider than the original space.

4. Closures of Nonclosable Operators

4.1. Extensions of linear operators. To analyze the above operators, we start from a special case of the problem: we consider the problem to define products of distributions with a given distribution u. Then we discuss relations between this problem and the general operator theory.

The multiplication by u is a linear operator U defined on the everywhere dense subspace $C^\infty(\mathbb{R}) \subset \mathcal{D}(\mathbb{R})$ consisting of smooth functions. The problem to define the product uv for distributions v that are not smooth functions is a problem to extend this operator, i.e., to define its continuation to a wider subspace.

The constructing of extensions of linear operators is a classical problem. Its general formulation is as follows. Let X be a topological vector space, X_0 be its vector subspace, and A be a linear operator acting from X_0 a space Y. We have to construct its extension to a wider subspace (or to the whole X).

First, recall known facts about constructing such extensions. As an example, we take operators in Banach spaces, but the reasoning in the case of locally convex topological vector spaces is totally similar.

A special case of the considered problem is the continuation problem for a linear bounded functional $f_0 : X_0 \to \mathbb{C}$, i.e., the case where $Y = \mathbb{C}$. Due to the Hahn–Banach theorem, for each f_0 there exists a linear bounded functional f extending f_0 to the whole space X.

If the subspace X_0 is not everywhere dense, then, in the general case, the answer is negative even for linear bounded operators: the existence of a bounded extension to the whole X is not guaranteed. If X_0 is everywhere dense in X, Y is a Banach space, and the operator A is bounded, then the problem is trivial: there exists a unique extension of the operator to the whole X; it can be defined as follows.

For each $x_0 \in X$ there exists a sequence of points $x_n \in X_0$ converging to x_0. For each such sequence, the limit of the sequence of images Ax_n exists and does not depend on the choice of x_n. This allows one to define the following extension to the whole X:

$$\widetilde{A}x_0 = \lim_{n\to\infty} Ax_n. \tag{4.1}$$

This is a linear bounded operator from X to Y.

In the case of our concern, the considered linear operator is discontinuous and defined on an everywhere dense subspace X_0. Let X_1 be the subset of elements $x_0 \in X$ such that there exists a sequence $\{x_n\} \subset X_0$ such that $x_n \to x_0$ and $\lim\limits_{n\to\infty} Ax_n$ exists. It seems to be natural to use relation (4.1) to define the value of the sought extension at the point x_0, but this extension is ill defined in the general case because the right-hand side might depend on the choice of the sequence. If the limit does not depend on the choice of the sequence, then the operator is said to be *closable*. Usually, the closability property of the operator is treated as the following condition: if $x_n \to 0$ and $\lim_{n\to\infty} Ax_n = y$ exists, then $y = 0$. If this condition is satisfied, then relation (4.1) defines the operator $\overline{A}$ on X_1 well; this operator is called the *closure* of the operator A. The so-called *graph norm*

$$\|x\|_1 = \|x\| + \|\overline{A}x\|$$

is defined on X_1. The space X_1 is complete with respect to this norm.

Example 4.1. A typical example of such a construction is the definition of the so-called *strong derivative*. Let $X = Y = L_1[0,1]$ and A be the differentiation operator defined on $C^1[0,1] \subset L_1[0,1]$:

$$(Ax)(t) = x'(t).$$

We say that a function $u_0 \in L_1[0,1]$ is *strongly differentiable* if there exists a sequence of $u_n \in C^1[0,1]$ such that $u_n \to u_0$ in $L_1[0,1]$ and $\lim u_n' := y$ exists in $L_1[0,1]$. Then the function y is called the *strong derivative* of the function u_0. The strong derivative is well defined, which follows from the closability of the differentiation operator with respect to the considered norms. Indeed, let $u_n \to 0$ and $\lim u_n' = y$ in $L_1[0,1]$. From the representation

$$u_n(x) - u_n(0) = \int_0^x u_n'(s)ds \tag{4.2}$$

and the convergence of the sequence $\{u_n'\}$ in $L_1[0,1]$, we obtain the uniform convergence of the sequence $\{u_n(x) - u_n(0)\}$. Therefore, it converges in $L_1[0,1]$. Since $u_n \to 0$ in $L_1[0,1]$, it follows that the sequence of constant functions $u_n(0)$ converges in $L_1[0,1]$ as well, which is possible only under the assumption that this number sequence converges. Hence, the sequence of u_n uniformly converges as well and, therefore, converges to zero. We can pass to the limit in (4.2), which implies that

$$\int_0^x y(s)ds = 0 \quad \forall\ x,$$

which is possible only under the assumption that $y(x) = 0$ almost everywhere.

Thus, the operator of the strong differentiation is the closure of the operator of the classical differentiation. All the above implies that the domain of this closure consists of functions representable in the form

$$u(x) = u(0) + \int_0^x y(s)ds, \text{ where } y \in L_1[0,1],$$

i.e., consists of absolutely continuous functions (according to Sobolev, this space is denoted by $W_1^1[0,1]$).

The closure of the differentiation operator in the space $L_2[0,1]$ is defined in the same way; the domain of this closure consists of functions representable in the form

$$u(x) = u(0) + \int_0^x y(s)ds, \text{ where } y \in L_2[0,1].$$

This subspace is denoted by $W_2^1[0,1]$ or $H^1[0,1]$.

Example 4.2. In $L_1[0,1]$, consider the following equation with an initial condition:

$$u'(x) = f(x),\ u(0) = C.$$

This is a degenerate case of the Cauchy problem. The operator A with the domain $C^1[0,1] \subset L_1[0,1]$ acting to the direct sum $Y = L_1[0,1] \oplus \mathbb{R}$ according to the relation $Au = (u', u(0))$ corresponds to this problem. From the above reasoning, we obtain that this operator is closable as well, the domain of the closure is $W_1^1[0,1]$, and $\overline{A}u = (u', u(0))$, where u' is the strong derivative.

Example 4.3. Let an operator A be defined on $C^1[0,1] \subset L_1[0,1]$ and act into the direct sum $Y = L_1[0,1] \oplus \mathbb{R}^2$ according to the relation $Au = (u', u(0), u'(0))$. This operator looks like the operator from Example 4.2, but this case is qualitatively different. Indeed, the sequence of functions

$$u_n(x) = \begin{cases} x(1-nx)^2, & 0 \le x \le 1/n, \\ 0, & x > 1/n \end{cases}$$

is a subset of the domain and converges to zero in $L_1[0,1]$, but the sequence of the images has a nonzero limit: $u_n' \to 0$ in $L_1[0,1]$, but $u_n'(0) = 1 \to 1$. This implies that the considered operator is nonclosable.

4.2. Closures of nonclosable operators. In [2], the following question is considered: *what operator can play the role of the closure of a nonclosable operator*? The answer is natural. Actually, it is already contained in the construction of the closure represented in another form.

The first step of this construction is to construct a set $\widetilde{G}$ consisting of sequences $\{x_n\} \subset X_0$ such that $\{x_n\}$ converges in X and the sequence $\{Ax_n\}$ converges in Y. Sequences x_n and $\widetilde{x}_n$ are said to be *equivalent* if $x_n - \widetilde{x}_n \to 0$ and $A(x_n - \widetilde{x}_n) \to 0$. Let G be the space consisting of classes of equivalent sequences from $\widetilde{G}$. On this space, the operator $\overline{A} : G \to Y$ is defined so that it takes each equivalence class to the element

$$G \ni \{x_n\} \mapsto \lim_{n\to\infty} Ax_n \in Y. \tag{4.3}$$

The second step of the constructing of the closure is to verify that the map

$$G \ni \{x_n\} \mapsto \lim_{n\to\infty} x_n \in X$$

establishes a bijection between G and a subspace X_1 in X.

The first step of this construction is applicable to each linear operator because it does not use the closeness.

Definition 4.1. The *closure* of the operator A with the domain $X_0 \subset X$ is the operator defined on the constructed space G and acting according to relation (4.3).

To explain the difference between the case of a closable operator and a nonclosable one, we recall the geometric interpretation of the described construction.

Let

$$G(A) = \{(x, Ax) : x \in D(A)\} \subset X \oplus Y$$

be the graph of the operator. Then $\widetilde{G}$ is the set of all Cauchy sequence in the sense of the norm of $X \oplus Y$ lying in $G(A)$, and the space G is the completion of the graph $G(A)$. By virtue of the completeness of $X \oplus Y$, it is the closure of $\overline{G(A)}$. The action of the operator $\overline{A}$ is the projecting P_Y to the second coordinate.

Let X_1 be the projection of $\overline{G(A)}$ to X. The operator is closable if and only if there exists an operator equal to the closure of the graph. This means that the projecting to the first coordinate is injective, which allows us to identify each point from $\overline{G(A)}$ with its first projection. This yields an operator defined on the subspace X_1 in X.

If the operator is nonclosable, then the projecting P_X to the first coordinate is not injective and the nonzero subspace

$$M = P_X^{-1}(0) = \{y \in Y : \exists x_n \in X_0 \text{ such that } x_n \to 0,\ Ax_n \to y\} \subset Y$$

arises. We call it the *nonclosability measure* of the operator A.

Then G is represented as the direct sum $M \oplus X_1$ and the projecting P_X defines a structure of a fiber space over X_1 on G. Here, the preimage of a point $x \in X_1$ is the set

$$P_X^{-1}(x) = \{(x, \xi) : \xi \in M\}.$$

It is not a vector subspace in G, but it has a natural vector-space structure, i.e., we deal with a vector bundle. Thus, the constructed operator $\overline{A}$ is defined on elements of the vector bundle G over X_1.

Example 4.4. Consider the closure of the operator from Example 4.3. In this example, the operator A is defined on $C^1[0,1] \subset L_1[0,1]$ and acts into the direct sum $Y = L_1[0,1] \oplus \mathbb{R} \oplus \mathbb{R}$ according to the relation $Au = (u', u(0), u'(0))$. The nonclosability measure of this operator is the one-dimensional subspace $M = \{(0,0,\xi) \in L_1[0,1] \oplus \mathbb{R} \oplus \mathbb{R}\}$. The space $\widetilde{G}$ is the set of sequences of $u_n \in C^1$ such that each one has four limits: the limits $u_n \to u_0$ and $u_n' \to y$ of the two sequences of functions from $L_1[0,1]$ and the limits of the two number sequences $\{u_n(0)\}$ and $\{u_n'(0)\}$.

It is shown in Example 4.2 that the first and the second condition imply that the function u_0 is absolutely continuous and $u_n(0) \to u_0(0)$. The considered case is specified as follows: the existence of the derivative $u_0'(0)$ of the limit function u_0 is not guaranteed and even its existence does not guarantee the convergence of the sequence $\{u_n'(0)\}$ to $u_0'(0)$.

According to the general construction, two sequences from $\widetilde{G}$ are said to be *equivalent* if all the above limits coincide with each other.

Note that the arising equivalence classes are more narrow than the ones from Example 4.2 and each equivalence class from Example 4.2 contains many various classes from the considered example.

The obtained equivalence class u consists of sequences of differentiable functions u_n converging to the absolutely continuous function u_0 in the special sense determined by the absolutely continuous function u_0 and the number $\xi = \lim u_n'(0)$; this number can be interpreted as the value $u'(0)$ of the derivative. Thus, each equivalence class related to u_0 "remembers" the approximation way from u_n to u_0, i.e., preserves an additional data about the behavior of the values $u_n'(0)$. In this example, the space of classes of equivalent sequences is isomorphic to the space $W_1^1[0,1] \oplus \mathbb{R}$. Here, the closure of the operator acts according to a relation looking like the relation for the original operator $\overline{A}u = (u', u(0), u'(0))$. However, u' is the strong derivative here and the value $u'(0)$ is defined as $\lim u_n'(0)$.

Regarding applications of this space, we note that, in classical function spaces, the solvability of overdetermined Cauchy problem

$$u'(x) = f(x),\ u(0) = C_0,\ u'(0) = C_1, \tag{4.4}$$

for arbitrary functions $f \in L_1[0,1]$ fails. In the introduced space G, problem (4.4) has a solution for each $f \in L_1[0,1]$ and this solution is unique.

A similar space can be constructed by means of function families depending on the small parameter ε. Such families naturally arise in the consideration of so-called *singularly perturbed problems*. The simplest example is the Cauchy problem

$$\varepsilon u''(x) + u'(x) = f(x),\ u(0) = C_0,\ u'(0) = C_1 \tag{4.5}$$

for the equation with a small parameter at the second derivative.

Let u_ε be solutions of problem (4.5) and v_ε be solutions of the similar problem

$$\varepsilon v''(x) + v'(x) = f(x),\ v(0) = C_0,\ v'(0) = C_2.$$

Both these families converge to the same absolutely continuous function u_0 satisfying the Cauchy problem

$$u_0'(x) = f(x),\ u_0(0) = C_0, \tag{4.6}$$

but they are not to be identified with each other because they are approximated to u_0 by different ways and each one contains an additional information about the approximation way: $u_\varepsilon'(0) = C_1$, while $v_\varepsilon'(0) = C_2$. This means that it is natural to treat elements from the constructed extended space isomorphic to $W_1^1[0,1] \oplus \mathbb{R}$ as solutions of problem (4.5).

Example 4.5. The classical question about relations between completions of a space with respect to different norms is another example, analyzing which we face similar effects.

Suppose that norms $\|x\|_1$ and $\|x\|_2$ are defined on the vector space X_0. The problem is to describe relations between the corresponding completions X_1 and X_2. If the norms are equivalent, i.e., there exist constants such that $\|x\|_1 \le C_1\|x\|_2$ and $\|x\|_2 \le C_2\|x\|_1$, then $X_1 = X_2$, i.e., the completions coincide with each other as vector spaces.

It frequently occurs that only one inequality $\|x\|_1 \le C\|x\|_2$ is satisfied. In many examples, the natural embedding of X_2 into X_1 takes place. For example, let $X_0 = C^1[0,1]$,

$$\|x\|_1 = \int_0^1 |x(t)|dt, \text{ and } \|x\|_2 = \int_0^1 |x(t)|dt + \int_0^1 |x'(t)|dt.$$

Then the completion with respect to the first norm is $X_1 = L_1[0,1]$, the completion with respect to the second norm is X_2 is the space $W_1^1[0,1]$ consisting of absolutely continuous functions (in fact, it is considered in Example 4.1), and $W_1^1[0,1] \subset L_1[0,1]$ in the considered case.

However, the embedding of X_2 into X_1 is not guaranteed in the general case. For example, consider the norms

$$\|x\|_1 = \int_0^1 |x(t)|dt \text{ and } \|x\|_2 = \int_0^1 |x(t)|dt + |x(0)|$$

on $X_0 = C^1[0,1]$. Here, we have another relation between spaces: the completion X_1 is $L_1[0,1]$ (as above), while the completion X_2 is isomorphic to $L_1[0,1] \oplus \mathbb{C}$ and it is wider than X_1 (unlike the previous example).

This difference is explained as follows. Elements of the completion X_2 are equivalence classes of Cauchy sequences in the sense of the second norm. If $\{x_n\}$ is such a Cauchy sequence from the class defining an element $x \in X_2$, then the given inequality implies that this is a Cauchy sequence with respect to the first norm as well and it defines an element from the completion X_1. Thus, the following continuous map is defined:

$$J : X_2 \ni x \to \lim_{n\to\infty} x_n \in X_1.$$

However, it is not guaranteed that this map is injective, i.e., it is not guaranteed that it is an embedding of X_2 into X_1. Its injectivity requires the additional *norm coordination* condition: *if* $\|x_n\|_1 \to 0$ *and* $\{x_n\}$ *is a Cauchy sequence in the sense of the second norm, then* $\|x_n\|_2 \to 0$.

The relation to the problem of closures of operators is as follows. Let us treat the identity map $J_0 x = x$ of the space X_0 as a map of the normed space $(X_0, \|x\|_2) \subset X_2$ into the normed space $(X_0, \|x\|_1) \subset X_1$. Since J_0 is a bounded linear operator, the map J is its closure defined on the whole X_2. However, it might occur that the inverse map J_0^{-1} acting from $(X_0, \|x\|_1)$ to $(X_0, \|x\|_2)$ is a nonclosable operator. The norm coordination condition coincides with the closability condition of this operator.

In the (general) case where the norm coordination condition is not satisfied, we obtain that the map J has a nonzero kernel M. Therefore, X_2 is not isomorphic to its image $\widehat{X_2} = J(X_2) \subset X_1$, but is represented by the vector bundle $X_2 = \widehat{X_2} \oplus M$.

4.3. Extended closures of linear operators. Get back to the extension problem for the operator U of the multiplication by a given distribution u. From property (ii) of Sec. 3, we obtain that this operator is nonclosable and its nonclosability measure is a nonzero subspace $M_u \subset \mathcal{D}'(\mathbb{R})$. Its closure $\overline{U}$ in the sense of Definition 4.1 is an operator defined on a vector bundle $G_u = M_u \oplus X_1$ over the subspace $X_1 \subset \mathcal{D}'(\mathbb{R})$.

The problem to extend the operator U to distributions not belonging to X_1 is still open. This problem leads to another generalization of the closure construction.

Constructing the closure of the operator A, we consider the vector space $\widetilde{G}$ consisting of sequences of $x_n \in X_0$ such that $\{x_n\}$ converges in X and the sequence of the images Ax_n converges in Y. In this case, two sequences are treated to be equivalent to each other if $x_n - \widetilde{x}_n \to 0$ and $A(x_n - \widetilde{x}_n) \to 0$. To obtain the desired generalization, we omit the requirement of the existence of a limit of the sequence of the images Ax_n.

Consider the vector space $\widehat{G}$ consisting of all sequences $\{(x_n, Ax_n)\}$ of points of the graph $G(A)$ such that $\{x_n\}$ converges in X. Let $\widehat{G}_0$ be the subspace in $\widehat{G}$ consisting of sequences such that $x_n \to 0$ and $Ax_n \to 0$. Consider the quotient space $\widehat{G}^* = \widehat{G}/\widehat{G}_0$. Here, the passage to the quotient space is equivalent to introducing the same equivalence relation as above: $(x_n, Ax_n) \sim (\widetilde{x}_n, A\widetilde{x}_n)$ if $x_n - \widetilde{x}_n \to 0$ and $A(x_n - \widetilde{x}_n) \to 0$.

Let $\widehat{Y}$ the space of all sequences (y_n) in Y. Let

$$Y^* = \widehat{Y}/\widehat{Y}_0, \text{ where } \widehat{Y}_0 = \{(y_n) \in Y : y_n \to 0\}.$$

Note that Y^* is an extension of the original space Y because the latter one is embedded into Y^* in a natural way: each point $y \in Y$ is mapped to the equivalence class consisting of sequences converging to y. Once these spaces are introduced, the operator $\widehat{A}$ acting from $\widehat{G}^*$ to Y^* is defined as follows:

$$\widehat{A}([(x_n, Ax_n)]) = [(Ax_n)] \in Y^*. \tag{4.7}$$

As above, $\widehat{G}^*$ is isomorphic to $X \oplus M$ here, and this space is treated as a vector bundle over X.

Definition 4.2. The *extended closure of the operator* A is the operator $\widehat{A}$ defined on the vector bundle $\widehat{G}^*$ over X and acting into the extended space Y^* according to relation (4.7).

By the constructing of $G \subset \widehat{G}^*$, the operator $\overline{A}$ maps G into $Y \subset Y^*$ and its action on G coincides with the action of $\widehat{A}$, i.e., the last operator is an extension of $\overline{A}$.

Summarizing the above, we conclude that if a nonclosable operator A acts from X to Y, then the extended closure acts in new spaces arising as results of constructions of the following two types:

(1) the *subdivision of the original space* X, which means that each point $x \in X$ is decomposed into a comprehensive family of new elements (fibers over x);
(2) *adding new elements* to the final space Y.

Note that similar operations are used already at the stage of the passage from ordinary functions to generalized ones because the subdivision takes place apart from adding new elements: the function $\frac{1}{x}$ is mapped not to a distribution, but to family of distributions.

Remark 4.1. From the viewpoint of applications, introducing new spaces for the closing of an operator is natural, which can be interpreted as follows. Assume that an action on a system is investigated. In the original model of the phenomenon, it is assumed that the states of the system are described by elements of the space X, while results of the action are described by elements of the space Y; namely, for several "simple" states (from the subspace $X_0 = D(A)$), an operator A describing the result of the action on the system is given: for each state x, we obtain the output result $Ax \in Y$.

The task is to describe the action result for more complicated states of the system. The case where the passage to the closure of the graph leads to a multivalued operator corresponds to the case where the original problem setting does not provide sufficient data to obtain an unambiguous answer about the reaction of the system in a more complicated state. Constructing the closure in the new sense proposes: the following solution of this problem: to obtain an unambiguous result, one needs an additional information of the more complicated state corresponding to the point $x_0 \in X$; namely, the information about the growth of this state from simple states. In other words, for considered systems, the problem setting needs a refinement: the state is described by specially constructed class of equivalent sequences of $x_n \in D(A)$ and is not defined uniquely by the limit point $x_0 \in X$.

From this viewpoint, the passage to the extension $\widehat{Y}$ of the space Y is required in the case where the system is in the state $x_0 \in X$ and the sequence of Ax_n does not converge in Y, i.e., the result of the experiment is not described by an element of the originally selected space Y.

4.4. Nonstandard extensions of the field $\mathbb{R}$. Constructions described above are similar to constructions from the *nonstandard analysis* substantially applied in numerous problems (see [20]). Recall the description of a nonstandard extension of the field $\mathbb{R}$.

Let $\widehat{\mathbb{R}}$ be the space of all sequences $\{y_n\}$ in $\mathbb{R}$. It is proved that, on the set $\mathbb{N}$, there exists a (finitely additive) measure μ defined on the algebra of all subsets of $\mathbb{N}$ and such that the value of $\mu(\omega)$ is equal either to 0 or to 1 for all $\omega \subset \mathbb{N}$ and $\mu(\omega) = 0$ for each finite ω. Let $\widehat{\mathbb{R}}_0 \subset \widehat{\mathbb{R}}$ be the subspace consisting of sequences a.e. equal to zero with respect to the measure μ. The nonstandard extension of $\mathbb{R}$ is the quotient space

$$\mathbb{R}^* = \widehat{\mathbb{R}}/\widehat{\mathbb{R}}_0.$$

This construction can be described in other terms as well. The space $\widehat{\mathbb{R}}$ has the natural structure of an algebra and $\widehat{\mathbb{R}}_0$ is one of its maximal ideals containing all finite sequences. Therefore, $\mathbb{R}^*$ is a quotient algebra with respect to the maximal ideal.

It turns out that the constructed space $\mathbb{R}^*$ is a field: if $[(y_n)] \neq 0$, then $\mu(\{n : y_n \neq 0\}) = 1$ and, therefore, the sequence of

$$z_n = \begin{cases} \frac{1}{y_n}, & y_n \neq 0, \\ 0, & y_n = 0 \end{cases}$$

determines the element inverse to $[(y_n)]$.

Further, on the set of equivalence classes, the following order relation is introduced:

$$[(x_n)] \prec [(y_n)], \text{ if } \quad x_n \leq y_n \text{ a.e.}$$

This order is linear: if $\mu(\{n : x_n \leq y_n\}) = 1$, then $[(x_n)] \prec [(y_n)]$; otherwise, $\mu(\{n : x_n \leq y_n\}) = 0$ and, therefore, $[(y_n)] \prec [(x_n)]$.

We say that an element $\gamma \in \mathbb{R}^*$ is *infinitely small* if $-a \prec \gamma \prec a$ for each positive $a \in \mathbb{R}$. To each number $x \in \mathbb{R}$, the so-called *monad* is associated: this is the set of nonstandard numbers differing from x by infinitely small values.

We say that an element $\Gamma \in \mathbb{R}^*$ is *infinitely large* if $a \prec |\Gamma|$ for each $a \in \mathbb{R}$.

To each ordinary function $f : \mathbb{R} \to \mathbb{R}$, its *nonstandard extension* $f^* : \mathbb{R}^* \to \mathbb{R}^*$ is associated by the following relation similar to (4.7):

$$f^*([(x_n)]) = [(f(x_n))].$$

Thus, constructing the nonstandard analysis under the passage from $\mathbb{R}$ to $\mathbb{R}^*$, we use the operations similar to the ones used above: the subdivision of points from $\mathbb{R}$ by means of the introducing of infinitely small values and the adding of infinitely large values. The passage from a function to its nonstandard extension corresponds to the passage from an operator to its extended closure.

5. Algebras of Mnemofunctions

The operator of the multiplication by a given distribution acts from the space of distributions into itself, i.e., in terms of the previous section, we have $X = Y$. It is naturally to expect the extended closure of such an operator to act from a new space Z into itself. Then, to construct Z operations of two kinds are required: the subdivision of elements of the original space and its extension via adding qualitatively new elements. As we show below, this occurs under the construction of algebras of mnemofunctions.

According to the Schwartz setting, an extension of E up to a differential algebra, i.e., the construction of a differential algebra $G(E)$ and an embedding $R : E \to G(E)$ can be treated as a solution of the multiplication problem for elements from the given space E of distributions.

However, the Schwartz example shows that there are no embeddings satisfying conditions (3.1)-(3.2). That is why algebras and embeddings possessing weaker properties were constructed. In this direction, the result of [17] is the best known; a differential algebra and embedding are constructed such that (3.2) is weakened as follows: only smooth functions are embedded together with their multiplication, i.e.,

$$R(fg) = R(f)R(g) \text{ for } f \in C^\infty,\ g \in C^\infty. \tag{5.1}$$

One of the most broad and simple algebras of the specified type is constructed in [21].

This research direction is called the *nonlinear theory of generalized functions*, the constructed differential algebras are called *algebras of the Colombeau type*, and their elements are called *new generalized functions* or *mnemofunctions*.

In [3, 4], methods to construct desired algebras are analyzed and a general scheme to construct them is proposed. This scheme is described below.

Usually, many different embeddings of R exist and each one provides a possibility to define the product of arbitrary distributions as mnemofunctions: by definition, it is assigned that

$$u \otimes_R v = R(u)R(v) \in G(E). \tag{5.2}$$

Since condition (3.2) cannot be satisfied, i.e., the multiplication defined by (5.2) cannot coincide with the multiplication of distributions by smooth functions introduced above. Under such an approach, the multiplication operation is corrected: it is changed to become associative.

The construction of desired differential algebras is based of families of smooth functions depending on the small parameter ε. Each classical space E of distributions contains a subspace $\mathcal{E}$ consisting of infinitely differentiable functions and this subspace is a differential algebra.

A family of operators $R_\varepsilon : E \to \mathcal{E}$ such that the family of smooth functions $f_\varepsilon = R_\varepsilon f$ converges to f in E is called an *approximation method* for R.

If R is fixed, then the set of all families of smooth functions $R(E) = \{f_\varepsilon = R_\varepsilon f : f \in E\}$ is not an algebra. Therefore, the first step of the construction is to select a differential algebra $\widetilde{G(E)}$ consisting of families of smooth functions $f_\varepsilon \in \mathcal{E}$ and containing sets $R(E)$ corresponding to "natural" approximation methods.

The space $\widetilde{G(E)}$ is very broad. This is why the desired algebra $G(E)$ of mnemofunctions is defined as the quotient algebra $G(E) = \widetilde{G(E)}/J$, where J is an ideal in $\widetilde{G(E)}$ invariant with respect to the differentiation.

The main difficulty is to select the ideal J. The approximation method for R defines the map $E \ni u \to [(R_\varepsilon u)] \in G(E)$. This is an embedding if the relation $[(R_\varepsilon u)] = [(R_\varepsilon v)]$ for the equivalence classes implies the relation $u = v$. Consider the subspace

$$N = \{(f_\varepsilon) \in \widetilde{G(E)} : f_\varepsilon \to 0 \text{ in } E\}.$$

If $J \subset N$, then $R_\varepsilon u - R_\varepsilon v \in N$, whence

$$u = \lim R_\varepsilon u = \lim R_\varepsilon v = v.$$

The obtained condition $J \subset N$ is a restriction from above meaning that the ideal is not to be too large. Note that N is not an ideal, which implies that $J \neq N$. This distinction causes the decomposition of the distribution from the original space into a family of elements of a new type (mnemofunctions) constructed as the quotient space N/J.

Embeddings with additional properties (their properties are discussed below) are especially interesting. Their validity is reduced to the requirement for the ideal to contain special elements, i.e., it is sufficiently large. For example, if condition (5.1) is satisfied, then the ideal J contains the differences $R_\varepsilon(f)R_\varepsilon(g) - R_\varepsilon(fg)$ for all $f, g \in \mathcal{E}$. If condition (3.1) is satisfied, then the ideal J contains the differences $R_\varepsilon(f') - [R_\varepsilon(f)]'$ for all $f \in E$.

Thus, the ideal J is to satisfy two conditions of the opposite kind. Their compatibility is not guaranteed. For example, if $E = \mathcal{D}'(\mathbb{R})$ and $\mathcal{E} = C^\infty(\mathbb{R})$, then there is no ideal containing the differences $R_\varepsilon(f)R_\varepsilon(g) - R_\varepsilon(fg)$ for all $f \in \mathcal{E}$, $g \in E$.

Consider the construction and investigation of mnemofunction algebras $G(E)$ on the example of the space of periodic distributions (see [7]). This space of generalized functions has a very simple structure, which makes the obtained results more visible than in the general case.

6. Space of Periodic Distributions

The circle treated as a manifold can be implemented as the subset $\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$ of the complex plane and as the quotient space $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. Respectively, the following two implementations arise for spaces of functions or distributions on the space: they can be treated as 2π-periodic functions of variable t on the line $\mathbb{R}$ and as functions of the complex variable z defined on $\mathbb{S}^1$. Such spaces are isomorphic and the isomorphism is established by means of the substitution $z = e^{it}$. Each of these implementations has its advantage in the following sense: the occurring relations.

The space $C^\infty_{2\pi}(\mathbb{R})$ consists of complex-valued infinitely differentiable functions. The isomorphic space $C^\infty(\mathbb{S}^1)$ consists of functions infinitely differentiable on $\mathbb{S}^1$. Note that the differentiation with respect to the variable t under the presentation $z = e^{it}$ is meant (instead of the differentiation with respect to z). If the function of variable z is defined and analytic in a neighborhood of the unit sphere. then the relation of these derivatives is determined by the relation $f' = iz\frac{df}{dz}$. For example, $(z^n)' = inz^n$.

The topology on the space $C^\infty(\mathbb{S}^1)$ is defined by means of the denumerable system of norms

$$p_m(\varphi) = \sum_{j=0}^{m} \max_z |\varphi^{(j)}(z)|, \ \varphi \in C^\infty(\mathbb{S}^1). \tag{6.1}$$

the space of generalized functions (distributions) $\mathcal{D}'(\mathbb{S}^1)$ is defined as the space adjoint to the space $C^\infty(\mathbb{S}^1)$, i.e., consists of continuous linear functionals on $C^\infty(\mathbb{S}^1)$. Usually, values of the functional f at point φ are denoted as follows: $f(\varphi) \equiv <f, \varphi>$.

On the space $\mathcal{D}'(\mathbb{S}^1)$, the convergence corresponding to the $*$-weak topology in the adjoint space: the sequence of f_n converges to f if

$$<f_n, \varphi> \to <f, \varphi> \text{ for each } \quad \varphi \in C^\infty(\mathbb{S}^1).$$

A function defined on the circle can be integrated with respect to the complex variable z and with respect to the real variable t. Since $z = e^{it}$ on the circle, we have $dz = ie^{it}dt$ and $dt = \frac{1}{iz}dz = |dz|$ and these integrals satisfy the relation

$$\int_0^{2\pi} u(e^{it})dt = \int_{\mathbb{S}^1} u(z)|dz| = \int_{\mathbb{S}^1} u(z)\frac{dz}{iz}.$$

The space $L_1(\mathbb{S}^1)$ (and, in particular, the space $C(\mathbb{S}^1)$) is embedded into $\mathcal{D}'(\mathbb{S}^1)$ according to the relation

$$L_1(\mathbb{S}^1) \ni u \to <u, \varphi> = \frac{1}{2\pi}\int_{\mathbb{S}^1} u(z)\varphi(z)|dz| = \frac{1}{2\pi}\int_0^{2\pi} u(e^{it})\varphi(e^{it})dt. \tag{6.2}$$

In the sequel, each integral is computed over the whole circle $\mathbb{S}^1$. The normalizing factor $\frac{1}{2\pi}$ is introduced to simplify relations provided below.

Each function φ from $C^\infty(\mathbb{S}^1)$ is expanded into the Fourier series

$$\varphi(z) = \sum_{-\infty}^{\infty} \varphi_k z^k$$

converging in $C^\infty(\mathbb{S}^1)$, where the Fourier coefficients are

$$\varphi_k = \frac{1}{2\pi}\int \varphi(z) z^{-k}|dz| = <\varphi, z^{-k}>$$

and the sequence of φ_k decays faster than each power of $\frac{1}{|k|}$. Therefore, elements from $\mathcal{D}'(\mathbb{S}^1)$ are uniquely determined by their values on functions z^k, $k \in \mathbb{Z}$, and are represented by the Fourier series

$$f = \sum_{-\infty}^{\infty} C_k z^k, \tag{6.3}$$

where the Fourier coefficients are $C_k = <f, z^{-k}>$. For each f, these coefficients increase not faster than a power of $|k|$. The distribution f treated as a functional acts as follows:

$$<f, \varphi> = \sum_{-\infty}^{\infty} C_k \varphi_k.$$

For example, the delta function δ_ξ concentrated at a point $\xi \in \mathbb{S}^1$ is defined by the relation $<\delta_\xi, \varphi> = \varphi(\xi)$ and is expanded into the series

$$\delta_\xi = \sum_{-\infty}^{\infty} \xi^{-k} z^k.$$

Rational functions, e.g., functions of the kind $f(z) = \frac{1}{(z-\xi)^n}$, are actively used in various areas of the analysis. For $|\xi| = 1$, such functions are not integrable on the circle, but a family of distributions corresponds to each in a natural way. A special interest is attracted by the distribution $\mathcal{P}\Big(\frac{1}{z-1}\Big)$ defined by the expression

$$<\mathcal{P}\Big(\frac{1}{z-1}\Big), \varphi> = \frac{1}{2\pi} \int_{\mathbb{S}^1} \frac{\varphi(z)}{z-1} |dz|,$$

where the integral is treated in the sense of the Cauchy principal value. The Fourier expansion of this distribution has the form

$$\mathcal{P}\Big(\frac{1}{z-1}\Big) = \frac{1}{2}\Bigg[\sum_{-\infty}^{-1} z^k - \sum_{0}^{+\infty} z^k\Bigg].$$

In the space $\mathcal{D}'(\mathbb{S}^1)$, the *differentiation* is defined as follows:

$$<f', \varphi> := -<f, \varphi'>.$$

In terms of Fourier coefficients, it is defined by the relation

$$f' = \sum_{-\infty}^{\infty} ikC_k z^k.$$

In the space $\mathcal{D}'(\mathbb{S}^1)$, the *multiplication* of each $f \in \mathcal{D}'(\mathbb{S}^1)$ by each function $g \in C^\infty(\mathbb{S}^1)$ is defined as well:

$$<gf, \varphi> = <f, g\varphi>, \quad g \in C^\infty(\mathbb{S}^1), \quad f \in \mathcal{D}'(\mathbb{S}^1).$$

In terms of Fourier coefficients, this product is defined by means of the convolution operation for sequences: if f has expansion (6.3) and

$$g = \sum_{-\infty}^{\infty} A_k z^k, \tag{6.4}$$

then

$$g * f = \sum_{-\infty}^{\infty} B_k z^k, \text{ where } B_k = \sum_{j=-\infty}^{\infty} C_j A_{k-j}.$$

Let $|\xi| = 1$. The map $\alpha(z) = \xi z$ is the rotation of the circle. By the relation

$$(T_\xi \varphi)(z) = \varphi(\xi z),$$

it generates the rotation operator acting in $C^\infty(\mathbb{S}^1)$ and other spaces of functions on the circle. Respectively, the rotation operator is defined in the space of distributions:

$$<T_\xi f, \varphi> = <f, T_{\overline{\xi}} \varphi> = <f, T_\xi^{-1} \varphi>.$$

For ordinary functions on the circle, the *convolution operation* is defined by the relation

$$(f * g)(z) = \frac{1}{2\pi} \int f(\xi) g\Big(\frac{z}{\xi}\Big) |d\xi|.$$

For a given distribution g, the function

$$\psi(z) = <g, T_z\varphi>$$

belongs to $C^\infty(\mathbb{S}^1)$. This provides a possibility to define the convolution of distributions by the relation

$$<f * g, \varphi> = <f, <g, T_z(\varphi)''>>.$$

On the circle, the convolution exists for each pair of distributions. If (6.3) and (6.4) are expanded in Fourier series, then the convolution passes into the term-by-term product of the Fourier coefficients:

$$f * g = \sum_{-\infty}^{\infty} C_k A_k z^k.$$

The singular integral Cauchy operator S on the circle is defined as the convolution with the distribution $2\mathcal{P}\left(\frac{1}{z-1}\right)$. Under expansion (6.3), this operator acts according to the relation

$$Sf = \sum_{-\infty}^{-1} C_k z^k - \sum_{0}^{+\infty} C_k z^k. \tag{6.5}$$

It is obvious that $S^2 = I$. Therefore, the operators

$$P^\pm = \frac{1}{2}[I \pm S]$$

are projectors. The space of periodic distributions is described, e.g., in [6, 7, 9, 37].

7. Algebras of Mnemofunctions on Circles and Embeddings of Distributions in These Algebras

7.1. The construction of an algebra. Let us construct an algebra of mnemofunctions corresponding to the space $\mathcal{D}'(\mathbb{S}^1)$.

Constructing the desired algebra, we use the fact that, under typical approximation methods, families of functions of the kind $u_\varepsilon = R_\varepsilon u$ satisfy estimates of the kind

$$p_m(u_\varepsilon) \le \frac{C}{\varepsilon^{m+\nu}}.$$

The space consisting of families satisfying such estimates is not an algebra. That is why we construct a broader space.

Consider families $\{f_\varepsilon\}$ depending on the small parameter ε, consisting of infinitely differentiable functions f_ε on $\mathbb{S}^1$, and such that for each $\{f_\varepsilon\}$ there exist numbers μ and ν such that

$$p_m(f_\varepsilon) \le \frac{C}{\varepsilon^{\mu m+\nu}}. \tag{7.1}$$

Hereinafter, C denote different constants because their explicit form is not essential for our investigation. Denote the set consisting of all such families by $\widetilde{G(\mathbb{S}^1)}$.

Lemma 7.1. *In the space $\widetilde{G(\mathbb{S}^1)}$, the following natural multiplication and addition operations are defined*:

$$\{f_\varepsilon\} \times \{g_\varepsilon\} = \{f_\varepsilon g_\varepsilon\} \quad \textit{and} \quad \{f_\varepsilon\}' = \{f_\varepsilon{}'\}.$$

This space with the introduced operations is a differential algebra.

Proof. For norms (6.1), the following inequality reflecting the continuity of the multiplication in the space $C^m(\mathbb{S}^1)$ is satisfied:

$$p_m(fg) \le C p_m(f) p_m(g).$$

Hence, if f_ε satisfies estimate (7.1) and g_ε satisfies estimate

$$p_m(g_\varepsilon) \le \frac{C}{\varepsilon^{\mu_1 m+\nu_1}},$$

then

$$p_m(f_\varepsilon g_\varepsilon) \le \frac{C}{\varepsilon^{(\mu+\mu_1)m+(\nu+\nu_1)}}.$$

The closedness of the space with respect to the differentiation follows from the inequality

$$p_m(f'_\varepsilon) \le p_{m+1}(f_\varepsilon).$$

□

The space $\widetilde{G(\mathbb{S}^1)}$ is quite comprehensive. Therefore, to obtain more visible spaces, one introduces the equivalence relation and considers the quotient space consisting of equivalence classes. In all vector spaces, equivalence relations of the following kind are considered: f and g are *equivalent* if $f-g \in L$, where L is a given subspace. This provides a possibility to define the operations of the addition and multiplication by numbers well. However, to be able to define the multiplication operation on a quotient space of an algebra with respect to the subspace L, one needs L is to be an ideal.

Let $\mathcal{N}_0$ be the subspace consisting of families converging to zero in $\mathcal{D}'(\mathbb{S}^1)$:

$$\mathcal{N}_0 = \{f_\varepsilon : \lim_{\varepsilon\to 0} <f_\varepsilon, \varphi> = 0 \ \text{ for each } \varphi \in C^\infty(\mathbb{S}^1)\}.$$

We say that families f_ε and g_ε are *weakly equivalent* if $f_\varepsilon - g_\varepsilon \in \mathcal{N}_0$. Such an equivalence relation is natural for the distribution theory, but the set $\mathcal{N}_0$ is not an ideal in $\widetilde{G(\mathbb{S}^1)}$. Hence, under this equivalence relation, no multiplication operation can be well defined on the quotient space. Moreover, $\mathcal{N}_0$ is not a subalgebra.

The algebra $\widetilde{G(\mathbb{S}^1)}$ contains many ideals such that it is possible to define equivalence relations using them. Equivalent families are to be weakly equivalent. This is satisfied if the desired ideal belongs to the subspace $\mathcal{N}_0$, i.e., it is sufficiently small. On the other hand, the decrease of the ideal leads to the increase of the quotient algebra. That is why the ideal is desired to be as large as it possible.

From the theoretical viewpoint, the most suitable is a maximal ideal in $\widetilde{G(\mathbb{S}^1)}$. Under such a choice of the ideal, the quotient algebra is an extension of the space of smooth functions in the sense of the nonstandard analysis. However, maximal ideals are not represented explicitly. Therefore, no particular computations can be implemented in the corresponding quotient algebra.

It turns out that the subspace

$$J(\mathbb{S}^1) = \{g_\varepsilon : \forall p \text{ and } m\ \exists\ C : p_m(g_\varepsilon) \le C\varepsilon^p\}$$

is more convenient.

Lemma 7.2. *The subspace $J(\mathbb{S}^1)$ is a differential ideal in the algebra $\widetilde{G(\mathbb{S}^1)}$.*

Proof. The invariance of the subspace $J(\mathbb{S}^1)$ with respect to the differentiation is obvious.

If $f_\varepsilon \in \widetilde{G(\mathbb{S}^1)}$ and $g_\varepsilon \in J(\mathbb{S}^1)$, then the product satisfies the estimate

$$p_m(f_\varepsilon g_\varepsilon) \le p_m(f_\varepsilon)p_m(g_\varepsilon) \le \frac{C}{\varepsilon^{\mu m+\nu}} \times C\varepsilon^p = C\varepsilon^{p-\mu m-\nu}.$$

Since p is an arbitrary number, it follows that this product belongs to $J(\mathbb{S}^1)$. □

The *algebra of mnemofunctions on the circle* $G(\mathbb{S}^1)$ is defined as the quotient algebra

$$G(\mathbb{S}^1) = \widetilde{G(\mathbb{S}^1)}/J(\mathbb{S}^1).$$

The algebra $\widetilde{\mathbb{C}} \subset G(\mathbb{S}^1)$ of *generalized complex numbers* is related to the described construction. It is generated by families of constants f_ε (they do not depend on z). In particular, this algebra contains elements of the kind ε^k; these values are infinitely small if $k > 0$ and infinitely large if $k < 0$. The algebra $G(\mathbb{S}^1)$ is a module over the algebra $\widetilde{\mathbb{C}}$.

To clarify the relation between the constructed algebra and the space of distributions, recall another construction of the space $\mathcal{D}'(\mathbb{S}^1)$. Consider the subspace $\widetilde{\mathcal{D}'(\mathbb{S}^1)}$ in $\widetilde{G(\mathbb{S}^1)}$ consisting of families f_ε such that for each $\varphi \in C^\infty(\mathbb{S}^1)$ there exists a finite limit $\lim\limits_{\varepsilon\to 0} <f_\varepsilon, \varphi>$.

Theorem 7.1. *The quotient space* $\widetilde{\mathcal{D}'(\mathbb{S}^1)}/\mathcal{N}_0$ *is isomorphic to the space* $\mathcal{D}'(\mathbb{S}^1)$ *of distributions.*

This method to introduce distributions (including the proof of Theorem 7.1) is described in [9] in detail. In this book and other papers of the same authors, sequences of smooth functions are used, i.e., this is the case where the small parameter takes only the values $\varepsilon = \frac{1}{n}$. That is why this construction of distributions is called the *sequential approach*.

This construction of the space of distributions allows us to clarify the roots of the ill-posedness of the multiplication problem for classical generalized functions:

(1) The space $\widetilde{\mathcal{D}'(\mathbb{S}^1)}$ is not an algebra. Hence, the quotient space contains no elements that can be treated as the product if the pair of elements is arbitrary.
(2) The subspace $\mathcal{N}_0$ is not an ideal in the algebra $\widetilde{G(\mathbb{S}^1)}$. Hence, products of representatives from one equivalence class get into different classes, i.e., the product of classes is not well defined.
(3) Since $\mathcal{N}_0$ is not a subalgebra, we obtain assertions of the kind $0 \times 0 \neq 0$: the product of two elements of the zero equivalence class might be outside this class.

Thus, the construction of the algebra of mnemofunctions can be treated as a modification of the sequential approach to the construction of the space of distributions: in both cases, the constructed space consists of equivalence classes of families of smooth functions satisfying estimates of kind (7.1), while the modification is as follows:

- instead of the space $\widetilde{\mathcal{D}'(\mathbb{S}^1)}$, the broader space $\widetilde{G(\mathbb{S}^1)}$, which is an algebra, is constructed;
- instead of the subspace $\mathcal{N}_0$, the ideal $J(\mathbb{S}^1)$ contained in this subspace is used to define the equivalence relation.

These distinctions leads to the following: in $\mathcal{D}'(\mathbb{S}^1)$, all families from $\mathcal{N}_0$ are identified with zero; in the algebra of mnemofunctions, they generate infinitely small values. Taking into account such infinitely small values, we obtain a possibility to define the multiplication operation well.

In [21], another (but contiguous) construction is proposed. For the circle case, a broader algebra $\widetilde{G_E(\mathbb{S}^1)}$ consisting of arbitrary families $\{f_\varepsilon\}$ is considered, but the validity of estimates (7.1) is not required. In this broader algebra, the set $J(\mathbb{S}^1)$ is not an ideal. That is why Egorov uses the ideal

$$J_0(\mathbb{S}^1) = \{f_\varepsilon : \forall p \text{ and } m\ \exists \varepsilon_0 > 0 : f_\varepsilon = 0 \ \text{ for } \ \varepsilon < \varepsilon_0\}.$$

The corresponding quotient algebra $G_E(\mathbb{S}^1) := \widetilde{G_E(\mathbb{S}^1)}/J_0(\mathbb{S}^1)$ is the *Egorov algebra of new generalized functions.*

Note that the set $J_0(\mathbb{S}^1)$ is an ideal in the algebra $\widetilde{G(\mathbb{S}^1)}$. Using it, one can construct another mnemofunction algebra $\widetilde{G(\mathbb{S}^1)}/J_0(\mathbb{S}^1)$. In fact, the difference between this algebra and the original algebra $\widetilde{G(\mathbb{S}^1)}$ is negligible.

7.2. The associativity relation and asymptotical expansions of mnemofunctions. The main concern is the relations between the constructed mnemofunction algebra and the space of distributions. First, the *associativity* relation is established. Since $J(\mathbb{S}^1) \subset \mathcal{N}_0$, it follows that if a family f_ε converges to f in $\mathcal{D}'(\mathbb{S}^1)$, then each equivalent family converges to f as well.

We say that an equivalence class $[f_\varepsilon]$ containing f_ε *is associated with a distribution* f if the family of f_ε converges to f in $\mathcal{D}'(\mathbb{S}^1)$. Let $G_{as}(\mathbb{S}^1)$ denote the subspace in $G(\mathbb{S}^1)$ consisting of families associated with distributions. By construction, G_{as} is a quotient space: $G_{as}(\mathbb{S}^1) = \widetilde{\mathcal{D}'(\mathbb{S}^1)}/J(\mathbb{S}^1)$. On $G_{as}(\mathbb{S}^1)$, the associativity relation generates the map

$$Lim : G_{as}(\mathbb{S}^1) \ni [f_\varepsilon] \to Lim([f_\varepsilon]) := \lim_{\varepsilon\to 0} f_\varepsilon \in \mathcal{D}'(\mathbb{S}^1)$$

of the limit passage.

The map Lim is surjective, but is not injective: each distribution u is related to a broad set $G_{as}(u)$ consisting of mnemofunctions associated with u; this set is an affine subspace in $G_{as}(\mathbb{S}^1)$. If f_0 is an arbitrary element of $G_{as}(u)$, then the map

$$G_{as}(u) \ni f \to f - f_0 \in G_{as}(0)$$

determines an isomorphism between $G_{as}(u)$ and the vector space $G_{as}(0)$, which is isomorphic to the quotient space $\mathcal{N}_0/J(\mathbb{S}^1)$ by construction. Note that such an isomorphism is not canonical because, in the general case, there are no reasons to select an element of the affine subspace $G_{as}(u)$ to assign it to be the zero element. Thus, the more the subspace $\mathcal{N}_0$ exceeds $J(\mathbb{S}^1)$, the greater is the ambiguity of the correspondence between $G_{as}(\mathbb{S}^1)$ and $\mathcal{D}'(\mathbb{S}^1)$. As we note above, elements of the quotient space $\mathcal{N}_0/J(\mathbb{S}^1)$ are infinitely small, i.e., the difference between two elements of $G_{as}(u)$ is equal to an infinitely small mnemofunction.

To obtain more data about mnemofunction properties, one can analyze the asymptotic behavior of values $< f_\varepsilon, \varphi >$. In fact, the expression $< F, \varphi > := < f_\varepsilon, \varphi >$ defines a generalized linear functional F on $C^\infty(\mathbb{S}^1)$, i.e., a functional with values in the algebra $\widetilde{\mathbb{C}}$ of generalized numbers. In particular, if $f_\varepsilon \in \mathcal{N}_0$, then $< f_\varepsilon, \varphi > \to 0$, i.e., values of the corresponding functional are infinitely small.

In the computations in the algebra of mnemofunctions, it frequently occurs that such a family of functionals admits an asymptotic expansion in the space $\mathcal{D}'(\mathbb{S}^1)$ with respect to powers of ε:

$$< f_\varepsilon, \varphi > = \sum_{k=k_0}^{\infty} < u_k, \varphi > \varepsilon^k, \text{ where } u_k \in \mathcal{D}'(\mathbb{S}^1). \tag{7.2}$$

We emphasize that asymptotic expansions are meant here, i.e., relation (7.2) means that the sequence of finite sums

$$< F_N, \varphi > = \sum_{k=k_0}^{N} < u_k, \varphi > \varepsilon^k$$

converges to $< f_\varepsilon, \varphi >$ asymptotically, i.e., their difference decays faster than ε^N.

It is possible that the asymptotic expansion starts from a negative power of ε, i.e., the principal term of the expansion is a distribution with an infinitely large coefficient.

Thus, visible data about the behavior of a mnemofunction are contained in its asymptotic expansion in the space $\mathcal{D}'(\mathbb{S}^1)$, the asymptotic expansion of each infinitely small mnemofunction starts from a positive power of ε, and the asymptotic expansion of the family f_ε associated with u has the form

$$< f_\varepsilon, \varphi > = < u, \varphi > + < u_1, \varphi > \varepsilon + < u_2, \varphi > \varepsilon^2 + \dots \tag{7.3}$$

Notionally, the relation between $\mathcal{D}'(\mathbb{S}^1)$ and G_{as} is an analog of the relation between the manifold M and its tangent bundle TM.

Indeed, consider the set of smooth curves $f(\varepsilon)$ passing through a given point a of the manifold M, i.e., such that $f(0) = a$. On this set, introduce the equivalence relation $f(\varepsilon) \sim g(\varepsilon)$ for $f(\varepsilon) - g(\varepsilon) = o(\varepsilon)$. Then the set of equivalence classes is the tangent space TM_a at the point $a \in M$, while the union of all tangent spaces is the tangent bundle TM.

By definition, the tangent vector is the class of equivalent curves approaching the point along the same direction, i.e., such a class preserves the data ("remembers") only about this direction.

In the same way, the family f_ε admitting expansion (7.3) can be treated as a "curve" in the distribution space $\mathcal{D}'(\mathbb{S}^1)$ passing through the point u. Then the distribution u_1 from the asymptotic expansion (7.3) describes the direction of the approaching of u by the "curve." Other terms of the expansion describe (more exactly) way of the approaching of f by the "curve." By definition, two families of smooth functions get into the same equivalence class if they behave "very similarly" as $\varepsilon \to 0$, i.e., the equivalence classes *remembers* the approaching way of elements of these families to their limits; this is the reason to call them "mnemofunctions" (because the Greek word "mnemo" means the memory).

The asymptotic expansion contains only a partial data about the mnemofunction behavior. In the general case, even the complete asymptotic expansion (i.e., the expansion with respect to all powers of ε) does not define f_ε uniquely. In particular, the asymptotic expansions for f_ε and g_ε do not define their product uniquely. As a corollary, solving particular problems, one has to use the algebra of mnemofunctions for immediate computing and to construct the asymptotic relation only for the final result (to provide its visibility).

7.3. Embeddings of distributions into the algebra of mnemofunctions. To establish more detail relations with distributions, one has to construct right inverse maps for the map Lim. They are linear maps $R : \mathcal{D}'(\mathbb{S}^1) \to G_{as} \subset G(\mathbb{S}^1)$ such that $LimR(u) = u$. By definition, the image $R(u)$ of a distribution u is a family of smooth functions converging to u, i.e., such a map R determines a way to approximate distributions by smooth functions and is an embedding (an injective map) of $\mathcal{D}'(\mathbb{S}^1)$ into the algebra of mnemofunctions.

Note that the map Lim has more than one right inverse map: if R is a right inverse map, then $R_1 : \mathcal{D}'(\mathbb{S}^1) \to G_{as} \subset G(\mathbb{S}^1)$ is a right inverse operator if and only if the operator $R - R_1$ maps $\mathcal{D}'(\mathbb{S}^1)$ into $\mathcal{N}_0$.

The product of the maps $RLim : G_{as} \to G_{as}$ taken in the inverse order is a projector in G_{as}. Therefore, the relation $f = RLimf + (f - RLimf)$ determines an expansion of the space G_{as} into the direct sum

$$G_{as} = Im(R) \oplus G_{as}(0).$$

Since the image $Im(R)$ is isomorphic to $\mathcal{D}'(\mathbb{S}^1)$, we obtain the expansion

$$G_{as} = \mathcal{D}'(\mathbb{S}^1) \oplus G_{as}(0).$$

Under this expansion, the projection to the first coordinate is the associativity map Lim.

Note that the obtained expansion is defined by the embedding R: in the affine subspace $G_{as}(u)$, the zero element is assigned to be $R(u)$.

As we note above, if R is given, then the product of arbitrary distributions is defined as the mnemofunction

$$u \otimes v := R(u)R(v) \in G(\mathbb{S}^1).$$

Usually, to describe properties of such a product, one has to establish its relations with the distributions. If the product $R(u)R(v)$ is associated with the distribution h, then it is natural to assign this h to be the product uv generated by the given way to approximate R.

As we note above, more data about the mnemofunction $R(u)R(v)$ is provided by its asymptotic expansion in the space of distributions. This space might exist even if the product $R(u)R(v)$ is associated with no distributions. Thus, a substantial part of the problem to describe products of distributions is reduced to the constructing of the asymptotic expansion for $R(u)R(v)$. In this problem, asymptotic expansions with infinitely large and infinitely small coefficients might arise; the latter ones are substantial as well.

7.4. Properties of embeddings. Since there are many ways to embed distributions into the algebra $R : \mathcal{D}'(\mathbb{S}^1) \to G(\mathbb{S}^1)$ of mnemofunctions, we investigate additional properties of various embeddings (see [6]).

7.4.1. Invariance with respect to rotations. In the circle case, a natural requirement for the embedding is its invariance with respect to rotations. In the case of distributions on the line, the invariance with respect to translations is considered.

Let $|\xi| = 1$. As we note above, the rotation of the circle given by the relation $\alpha(z) = \xi z$ generates the rotation operator acting in $C^\infty(\mathbb{S}^1)$ according to the relation $(T_\xi \varphi)(z) = \varphi(\xi z)$. Respectively, in the space of distributions, the rotation operator acts in the algebra of mnemofunctions as well.

The invariance property of the embedding R is the commutativity with the rotation: if $R(f) = f_\varepsilon$, then

$$R(T_\xi f) = T_\xi f_\varepsilon.$$

This yields the special form of such an operator. The images of the functions z^k under the action of R are expanded as follows:

$$R(z^k) = \sum_j A_{jk} z^j.$$

The image of z^k under the rotation of the circle is $T_\xi z^k = \xi^k z^k$. If R is an arbitrary operator commutative with rotations, then the relation $R(\xi^k z^k) = \xi^k R z^k$ is satisfied. Under the expansion in the Fourier series, we obtain that

$$\sum_j A_{jk} \xi^k z^j = \sum_j A_{jk} \xi^j z^j,$$

whence $A_{jk}[\xi^j - \xi^k] = 0$ and, therefore, $A_{jk} = 0$ if $k \neq j$. Introducing the notation $A_{kk} = A_k$, we obtain that the operator acts according to the relation

$$Rf = \sum_j A_k C_k z^k,$$

i.e., the operator R is the convolution with the distribution such that A_k are its Fourier coefficients. In particular, the relation $R(z^k) = A_k(\varepsilon) z^k$, i.e., the commutativity with rotations, means that the functions z^k are eigenfunctions of the operator R.

Applying the above, for each fixed ε, we obtain that each approximation way invariant with respect to rotations has the form

$$R(f) = f_\varepsilon = f * \psi_\varepsilon, \tag{7.4}$$

where $*$ is the convolution operation in the space $\mathcal{D}'(\mathbb{S}^1)$, while ψ_ε is a family of distributions.

The convolution operation uses the group structure of the circle. The point 1 is selected because it is the neutral element of the group and the convolution with δ_1 is the identity operator. Therefore, $\delta_1 * \psi_\varepsilon = \psi_\varepsilon$, which implies that ψ_ε is a family of smooth functions converging to δ_1. Thus, if the invariance condition is satisfied, then the approximation method is uniquely defined by the approximations of δ_1.

If the Fourier expansions

$$\psi_\varepsilon(z) = \sum A_k(\varepsilon) z^k$$

are used, then

$$R(f) = f_\varepsilon(z) = \sum_k A_k(\varepsilon) C_k z^k, \tag{7.5}$$

where the coefficients $A_k(\varepsilon)$ decay faster than each power of $\frac{1}{k}$ provided that ε is fixed, while, if k is fixed, then the convergence of ψ_ε to δ_1 implies that $A_k(\varepsilon) \to 1$ as $\varepsilon \to 0$.

For applications, property (3.1) is essential. This is the commutativity of the embedding and differentiation:

$$R(f') = R(f)'. \tag{7.6}$$

Lemma 7.3. *The embedding R is commutative with the differentiation if and only if it is invariant with respect to rotations.*

Proof. The convolution operation is commutative with the differentiation. Therefore, Eq. (7.4) implies Eq. (7.6).

Assume that (7.6) is satisfied. Consider the image expansion $R(z^k) = \sum_j B_j(\varepsilon)z^j$. From (7.6), we obtain that $ik\sum_j B_j(\varepsilon)z^j = \sum_j B_j(\varepsilon)ij\ z^j$, which implies that $B_j(\varepsilon) = 0$ for $j \neq k$, i.e., $R(z^k) = B_k(\varepsilon)z^k$, which completes the proof. □

The most simple and natural approximation method is determined by means of partial sums of the Fourier series. Since each distribution f is expanded in series (6.3), it follows that the relation

$$R_F(f) = f_n = \sum_{-n}^{n} C_k z^k \tag{7.7}$$

determines an embedding of $\mathcal{D}'(\mathbb{S}^1)$ into $\widetilde{G(\mathbb{S}^1)}$. In this example, $\widetilde{G(\mathbb{S}^1)}$ denotes the mnemofunction algebra generated by sequences of smooth functions.

From the Fourier series viewpoint, relations of kind (7.5) determine summing methods for such series. In particular, the problem to sum series is hard because if f is a continuous function, then the uniform convergence of the sequence of partial sums (7.7) is not guaranteed. However, there are many summing methods of kind (7.5) improving the convergence; under these methods, $f_\varepsilon(z)$ uniformly converge to f. In the same, approximation methods defined by relations of kind (7.5) might possess properties that are not possessed by embedding (7.7) generated by partial sums of the Fourier series.

Below, embeddings invariant with respect to rotations are considered.

7.4.2. Multiplication localness. The value at a given point is not defined for distributions, but one can say about its values on open sets. We say that distributions f and g *are equal to each other on an open subset* U if $<f,\varphi> = <g,\varphi>$ for each φ such that its support belongs to U. This corresponds to the following definition of the value of f on the open subset U: this is the restriction of the functional f to the subspace consisting of such test functions φ.

The *support of a distribution* f is the least closed set $\operatorname{supp} f$ such that $f = 0$ on its complement.

The localness property of the multiplication of a distribution f by a smooth function g is as follows: on each open set U, the product gf depends only on values of g and f on it. In particular, if $g = 0$ on U or $f = 0$ on U, then the product gf is equal to zero on U.

We say that the *multiplication localness property* is satisfied for an embedding R if the condition $\operatorname{supp} f \cap \operatorname{supp} g = \varnothing$ implies that $R(f)R(g) = 0$.

At the first sight, this property seems to be always satisfied. This is "confirmed" by the following reasoning. If $\operatorname{supp} f \cap \operatorname{supp} g = \varnothing$, then there exists a smooth function γ such that $\gamma(z) = 1$ on $\operatorname{supp} f$ and $\gamma(z) = 0$ on $\operatorname{supp} g$. Then $f = \gamma f$ and $g = (1-\gamma)f$, whence $R(f)R(g) = R[\gamma(1-\gamma)]R(f)R(g) = 0$.

However, these computations assume that

$$R(\gamma f) = R(\gamma)R(f) \text{ for all } \gamma \in C^\infty(\mathbb{S}^1),\ f \in \mathcal{D}'(\mathbb{S}^1), \tag{7.8}$$

while, according to the known Schwartz example, the space of distributions cannot be embedded into an associative commutative algebra such that products of smooth functions and distributions are preserved. Hence, for each embedding R of the space $\mathcal{D}'(\mathbb{S}^1)$ into an (arbitrary) algebra, relation (7.8) cannot be satisfied for all $\gamma \in C^\infty(\mathbb{S}^1)$ and all $f \in \mathcal{D}'(\mathbb{S}^1)$. Thus, the above reasoning is incorrect and, in the general case, if $\operatorname{supp} f \cap \operatorname{supp} g = \varnothing$, then the product $R(g)R(f)$ might be different from zero in $G(\mathbb{S}^1)$.

7.4.3. Coordination of embedding with multiplication in $C^\infty(\mathbb{S}^1)$. In [17, 18], the problem to construct an embedding satisfying the condition to be coordinated with the multiplication in $C^\infty(\mathbb{S}^1)$ is considered. This condition is weaker than condition (7.8), which cannot be satisfied. In the considered case, the Colombeau problem is formulated as follows.

There exists a natural embedding R_0 of the algebra $C^\infty(\mathbb{S}^1)$ into $G(\mathbb{S}^1)$ taking each function $f \in C^\infty(\mathbb{S}^1)$ to a stationary family $f_\varepsilon = f$, i.e., a family independent of ε.

Colombeau problem. *Construct a differential algebra G and an embedding R of the space of distributions into G such that it coincides with the natural embedding of R_0 into G for infinitely differentiable functions.*

In the case of the considered space of distributions on the circle, the Colombeau problem is the problem to construct an embedding such that

$$R(f) = R_0(f) \quad \text{for} \quad f \in C^\infty(\mathbb{S}^1). \tag{7.9}$$

This condition means that, for smooth functions f, the considered approximations are to converge to f fast.

If (7.9) is satisfied, then $C^\infty(\mathbb{S}^1)$ is an algebra as well (not only is a subspace), i.e.,

$$R(fg) = R(f)R(g) \text{ for all } f,\ g \in C^\infty(\mathbb{S}^1). \tag{7.10}$$

The main Colombeau result is the construction of the desired algebra for spaces of distributions on $\mathbb{R}^m$. The algebra constructed by Colombeau is much more complicated than the one described above.

The case of distributions on the circle described in the present paper possesses the following specific property: embeddings satisfying (7.9) and (7.10) exist for the simpler algebra $G(\mathbb{S}^1)$ of mnemofunctions.

8. Embeddings of $\mathcal{D}'(\mathbb{S}^1)$ into $G(\mathbb{S}^1)$

From the viewpoint of above properties, consider particular classes of embeddings of kind (7.4).

8.1. Embeddings satisfying the multiplication localness condition. In the general case, the product $R(\delta_\xi) \times R(\delta_1) = \psi_\varepsilon\left(\frac{z}{\xi}\right) \times \psi_\varepsilon(z)$ of two δ-functions concentrated at different points is different from zero. However, if the supports of functions $\psi_\varepsilon(z)$ shrink to the point 1 as $\varepsilon \to 0$, then the latter product is equal to zero provided ε is sufficiently small; therefore, the localness property for the multiplication is satisfied. Also, it is satisfied if the values of $\psi_\varepsilon(z)$ rapidly tend to zero if $z \neq 1$; then the product $\psi_\varepsilon(z) \times \psi_\varepsilon\left(\frac{z}{\xi}\right)$ belongs to the ideal $J(\mathbb{S}^1)$.

Lemma 8.1. *If the supports of functions $\psi_\varepsilon(z)$ defining the embedding by means of relation* (7.4) *shrink to the point* 1, *then the multiplication localness property is satisfied for arbitrary distributions.*

Proof. Let $\operatorname{supp} f \cap \operatorname{supp} g = \varnothing$. The condition that the supports of functions $\psi_\varepsilon(z)$ shrink to the point 1 means that $\psi_\varepsilon(z) = 0$ for $|z-1| > \gamma(\varepsilon)$, where $\gamma(\varepsilon) \to 0$ as $\varepsilon \to 0$. Since $f_\varepsilon = f * \psi_\varepsilon$, it follows that $f_\varepsilon(z) = 0$ outside the $\gamma(\varepsilon)$-neighborhood of the closed set $\operatorname{supp} f$. In the same way, we show that $g_\varepsilon(z) = 0$ outside the $\gamma(\varepsilon)$-neighborhood of the closed set $\operatorname{supp} g$. If ε is sufficiently small, then such neighborhood do not intersect each other, i.e., the union of their complements is equal to $\mathbb{S}^1$. Therefore, for each sufficiently small ε, we have the relation $f_\varepsilon(z) \times g_\varepsilon(z) \equiv 0$, which completes the proof. ☐

Approximation methods usually considered in distribution spaces on the line are given by relations of kind (3.4) as follows. We select a function $\psi \in \mathcal{D}(\mathbb{R})$ such that $\int_{\mathbb{R}} \psi(t)dt = 1$. Then the family

$$\psi_\varepsilon(t) = \frac{1}{\varepsilon}\psi\left(\frac{t}{\varepsilon}\right) \tag{8.1}$$

converges to δ_0 and their supports shrink to the point 0. The corresponding approximation method is given by expression (7.4), where $*$ is the convolution operation in the space $\mathcal{D}'(\mathbb{R})$. Under such approximation method, the validity of the multiplication localness property is obvious.

Here, family (8.1) is generated by one fixed function ψ called the *profile*. Such approximations are convenient for investigations because properties of the approximating family f_ε are described via ψ; in particular, momenta of this function are used.

In the case of periodic distributions (distributions on the circle), there are no complete analogs of this construction: e.g., if $\psi(t)$ is a periodic function, then no family of kind (8.1) converges to the δ-function. To define the above approximation by means of the convolution on the circle, one has to modify the construction.

Let $\psi \in \mathcal{D}(\mathbb{R})$, $\int_{\mathbb{R}} \psi(t)dt = 1$, and let the support be located inside the interval $(-\pi, \pi)$. For each ε, define the function $\frac{1}{\varepsilon}\psi\Big(\frac{t}{\varepsilon}\Big)$ for $-\pi \le t \le \pi$ and extend it to $\mathbb{R}$ with period 2π, i.e., consider the function

$$\psi_\varepsilon(t) = \sum_j \frac{1}{\varepsilon}\psi\Big(\frac{t+2j\pi}{\varepsilon}\Big). \tag{8.2}$$

Then the function

$$R(f) = f_\varepsilon = f * \psi_\varepsilon, \tag{8.3}$$

where $*$ is the convolution operation in the space $\mathcal{D}'(\mathbb{S}^1)$, determines the same approximation method possessing the multiplication localness property.

However, the detailed analysis of this approximation method on the circle is harder that the line case: there are no immediate relations between Fourier coefficients of functions ψ_ε for different values of ε because

$$\psi_\varepsilon(t) = \sum_{-\infty}^{\infty} \hat{\psi}(2\pi k\varepsilon)e^{ikt},$$

where $\hat{\psi}$ is the Fourier transform of the function ψ on $\mathbb{R}$.

Thus, the multiplication localness property complicates the embedding construction. Also, note that the requirement for the multiplication to be local is not always a reasonable form the viewpoint of physics. For example, the δ-function models the case where the substance is distributed so that the main part of the mass is concentrated in a small neighborhood of the given point. If we multiply two densities corresponding to such δ-functions concentrated at different points, then the product of a large and small density arise. Such a product might yield a finite impact, i.e., a nonzero density. Therefore, it is not reasonable to treat the product of two δ-functions concentrated at different points as zero in each case.

8.2. Coordination with multiplication of smooth functions. The following fact is an obstacle for constructing an embedding coordinated with the multiplication of smooth functions.

Theorem 8.1. *If an embedding of periodic distributions is constructed by means of the convolution with functions of kind* (8.1), *where ψ is a compactly supported function, then the multiplication localness condition is satisfied, but conditions* (7.9) *and* (7.10) (*of the coordination with the multiplication of smooth functions*) *are not satisfied.*

This assertion follows from the two lemmas below.

Lemma 8.2. *Let $\psi \in \mathcal{D}(\mathbb{R})$ be a smooth rapidly decreasing function on $\mathbb{R}$, e.g., a compactly supported function. Let $M_0(\psi) = \int \psi(t)dt = 1$. Let f be an infinitely differentiable periodic function. Then the family of smooth functions $f_\varepsilon = f * \psi_\varepsilon$, where $\psi_\varepsilon(t) = \frac{1}{\varepsilon}\psi\Big(\frac{t}{\varepsilon}\Big)$, is asymptotically expanded as follows*:

$$f_\varepsilon(t) \sim f(t) + \sum_{j=1}^{\infty} \frac{(-1)^j}{j!} f(t)^{(j)} M_j(\psi)\varepsilon^j, \tag{8.4}$$

where

$$M_j(\psi) = \int_{-\infty}^{+\infty} t^j \psi(t)dt, \;\; j \in \mathbb{N}$$

are the momenta of the function ψ.

Proof. The Taylor expansion yields

$$f(s)=\sum_{j=0}^{n}\frac{1}{j!}f(t)^{(j)}(s-t)^j+r_n(s-t),$$

where the remainder is estimated via the derivative $f(t)^{(n+1)}$. Since this derivative is bounded, we have the estimate $|r_n(s-t)|\le C|s-t|^{n+1}$. Therefore,

$$\begin{aligned}f_\varepsilon(t)=\int_{-\infty}^{+\infty}f(s)\frac{1}{\varepsilon}\psi\Big(\frac{t-s}{\varepsilon}\Big)ds&=\sum_{j=0}^{n}\frac{1}{j!}f(t)^{(j)}\int_{-\infty}^{+\infty}(s-t)^j\frac{1}{\varepsilon}\psi\Big(\frac{t-s}{\varepsilon}\Big)ds+\int_{-\infty}^{+\infty}r_n(s-t)\frac{1}{\varepsilon}\psi\Big(\frac{t-s}{\varepsilon}\Big)ds\\&=\sum_{j=0}^{n}\frac{1}{j!}f(t)^{(j)}\int_{-\infty}^{+\infty}(-\varepsilon\tau)^j\psi(\tau)d\tau+\int_{-\infty}^{+\infty}r_n(-\varepsilon\tau)\psi(\tau)d\tau \qquad (8.5)\\&=f(t)+\sum_{j=1}^{n}\frac{(-1)^j}{j!}f(t)^{(j)}M_j(\psi)\varepsilon^j+o(\varepsilon^n),\end{aligned}$$

which completes the proof. □

Lemma 8.3. *If $\psi\in\mathcal{D}(\mathbb{R})$ and $\psi\neq 0$, then an infinite set of momenta $M_j(\psi)$ are different from zero.*

Proof. Assume that $M_j(\psi)=0$ for all j apart from a finite amount of them. By virtue of the compactness of the support of ψ, the Fourier transform $\hat\psi$ is an analytic function; this function tends to zero at infinity by virtue of the smoothness of ψ. Under the Fourier transformation, momenta pass (up to a factor) into values of the corresponding derivatives. Therefore, $\hat\psi^{(j)}(0)=0$ for all $j\in\mathbb{N}$ apart from a finite amount of them. Then the analyticity of $\hat\psi$ implies that this function is a nonzero polynomial. Therefore, it does not tend to zero at infinity, which leads to a contradiction. □

Proof of Theorem 8.1. Let $M_p(\psi)$ be the nonzero momentum with the least index. Then, due to (8.5), the difference $f(t)-f_\varepsilon(t)$ behaves as ε^p and, therefore, does not belong to the ideal $J(\mathbb{S}^1)$, i.e., relation (7.9) is not satisfied.

In the same way, by virtue of (8.5), we have the following relation for two functions:

$$f_\varepsilon(t)\times g_\varepsilon(t)-(fg)_\varepsilon(t)=\frac{(-1)^p\varepsilon^p}{p!}\big[f(t)g(t)^{(p)}+f(t)^{(p)}g(t)-(f(t)g(t))^{(p)}\big]M_p(\psi)+o(\varepsilon^p).$$

Here, for $p>1$, the expression in the square brackets is different from zero, which implies that the considered difference does not belong to the ideal $J(\mathbb{S}^1)$. For $p=1$, we obtain that $f(t)-f_\varepsilon(t)$ behaves as ε^{p_2}, where $M_{p_2}(\psi)$ is the second nonzero momentum. □

Thus, if approximation methods generated by compactly supported functions are used, then, to satisfy condition (7.9), one has to construct more complicated algebras than $G(\mathbb{S}^1)$. For the first time, such algebras are constructed by Colombeau. Let us describe a modification of his construction, leading to an algebra simpler than the Colombeau algebra.

Consider families $\{f_{q,\varepsilon}\}$ of infinitely differential functions depending on two parameters ε and $q\in\mathbb{N}$. Let $\widetilde{G_C(\mathbb{S}^1)}$ be the set of all these families such that for each one there exist μ and ν such that the following estimate takes place:

$$p_m(f_{q,\varepsilon})\le\frac{C}{\varepsilon^{\mu m+\nu}}. \qquad (8.6)$$

As above, one can show that this set is a differential algebra.

Let

$$J_C(\mathbb{S}^1)=\{g_{q,\varepsilon}:\exists\ \ \mu_1\ \text{ and }\ \nu_1\ \text{ such that }\ p_m(g_{q,\varepsilon})\le C\varepsilon^{q-\mu_1 m-\nu_1}\ \}.$$

Lemma 8.4. *The set $J_C(\mathbb{S}^1)$ is a differential ideal in the algebra $\widetilde{G_C(\mathbb{S}^1)}$.*

Proof. Let $f_{q,\varepsilon} \in \widetilde{G_C(\mathbb{S}^1)}$ and $g_{q,\varepsilon} \in J_C(\mathbb{S}^1)$. Then

$$p_m(f_{q,\varepsilon} \times g_{q,\varepsilon}) \leq \frac{C_1}{\varepsilon^{\mu m+\nu}} C_2 \varepsilon^{q-\mu_1 m-\nu_1} = C\varepsilon^{q-(\mu+\mu_1)m-(\nu+\nu_1)},$$

i.e., this product belongs to $J_C(\mathbb{S}^1)$. □

By virtue of Lemma 8.4, the quotient space

$$G_C(\mathbb{S}^1) = \widetilde{G_C(\mathbb{S}^1)}/J_C(\mathbb{S}^1)$$

is a differential algebra. We call it the *modified Colombeau algebra.*

To construct an embedding into this algebra, select a sequence of compactly supported functions ψ_q such that their supports are located in a neighborhood of the point 0 and $M_j(\psi_q) = 0$ for $1 \leq j < q$.

Theorem 8.2. *The map $R_C : f \to f_{q,\varepsilon} = f * \psi_{q,\varepsilon}$ determines the embedding of $\mathcal{D}'(\mathbb{S}^1)$ into $G_C(\mathbb{S}^1)$ satisfying condition* (7.9) (*of the coordination with the multiplication of smooth functions*) *and the multiplication localness property.*

Proof. By virtue of the conditions imposed on ψ_q, we obtain that

$$p_m(f * \psi_{q,\varepsilon}) \leq \frac{C}{\varepsilon^{m+\nu}}.$$

If $f \in C^\infty(\mathbb{S}^1)$, then, according to (8.5), the difference $f(t) - f_{q,\varepsilon}(t)$ decreases as ε^q. Since the convolution is commutative with the differentiation, it follows that the difference of the derivatives of these functions behaves in the same way. This implies that

$$p_m(f - f * \psi_{q,\varepsilon}) \leq C\varepsilon^{q-m}$$

and the difference $f - f_{q,\varepsilon}$ belongs to the ideal $J_C(\mathbb{S}^1)$. □

As we note above, condition (7.9) means that approximations of each smooth function rapidly converge to it. This is fulfilled for embedding (7.7) constructed by means of the sequence of partial sums of the Fourier series. This map R_F is invariant with respect to rotations and is represented as a convolution:

$$f_n = f * \psi_n, \text{ where } \psi_n(z) = \sum_{-n}^{n} z^k = \begin{cases} \frac{z^{n+1}-z^{-n}}{z-1}, & z \neq 1, \\ 2n+1, & z = 1. \end{cases}$$

Theorem 8.3. *The map R_F is an embedding of $\mathcal{D}'(\mathbb{S}^1)$ into the mnemofunction algebra such that the coordination condition with the multiplication, i.e., relations* (7.9) *and* (7.10), *is satisfied, but the multiplication localness property is not.*

Proof. The belonging of f_n to $\widetilde{G(\mathbb{S}^1)}$ follows from the power estimate of the coefficients C_k. Really, the Fourier coefficients of the distribution f satisfy the estimate $|C_k| \leq C(1+|k|)^p$. Therefore,

$$p_m(f_n) \leq \sum_{|k|\leq n} C(1+|k|)^{p+m} \leq C(2n+1)(1+n)^{p+m} \sim n^{p+m+1}.$$

The validity of (7.9) follows from the known fact that the Fourier series of each function $f \in C^\infty(\mathbb{S}^1)$ rapidly converges, i.e., $f - f_n \in J(\mathbb{S}^1)$.

Under this method, $f_n(z)g_n(z) \neq (fg)_n$ for smooth functions, while relation (7.10) is satisfied only in the quotient algebra, i.e., $f_n(z)g_n(z) - (fg)_n \in J(\mathbb{S}^1)$.

The approximating sequence of

$$\psi_n(\xi z) = \sum_{-n}^{n} \xi^{-k} z^k = \frac{\xi^{2n+1} - z^{2n+1}}{(\xi z)^n(\xi - z)}$$

corresponds to the distribution δ_ξ. Therefore, the sequence of

$$\psi_n(\xi z)\psi_n(z) = \frac{\xi^{2n+1} - z^{2n+1}}{(\xi z)^n(\xi - z)} \times \frac{z^{n+1} - z^{-n}}{z - 1}$$

corresponds to the product $\delta_\xi\ \delta_1$. This sequence does not tend to zero and, therefore, does not belong to the ideal $J(\mathbb{S}^1)$. □

8.3. Joint localness and coordination with multiplication of smooth functions. According to the results of the previous section, the multiplication localness property is satisfied provided that embeddings generated by compactly supported functions are considered, but an algebra more complicated than $G(\mathbb{S}^1)$ is required for the existence of an embedding such that the coordination with the multiplication takes place, i.e., condition (7.9) is satisfied, as well. On the other hand, the embedding generated by partial sums of the Fourier series is coordinated with the multiplication, but it does not possess the localness property. Let us show that there exist embeddings into $G(\mathbb{S}^1)$ possessing both specified properties simultaneously. One succeeds to construct such embeddings if the compactness requirement for the support of the function ψ is taken off.

Consider the Schwartz space $\mathcal{S}(\mathbb{R})$, i.e., the set of infinitely differentiable functions decaying at infinity faster than each power of $\frac{1}{t}$. For functions from this space, the situation is different from the one described in Lemma 8.2.

Lemma 8.5. *There exist functions ψ in the space $\mathcal{S}(\mathbb{S}^1)$ such that*

$$M_0(\psi) = 1 \quad \textit{and} \quad M_j(\psi) = 0 \quad \textit{for } j \in \mathbb{N}.$$

Proof. The Fourier transformation bijectively maps the space $\mathcal{S}(\mathbb{S}^1)$ into itself. In this space, there exist functions such that $\varphi(\xi) = 1$ on a neighborhood $(-\gamma, \gamma)$ of the origin. Then the function ψ, which is the inverse Fourier transform of the function φ, possesses the properties from the claim of the lemma. Indeed, as we note above, the Fourier transformation maps momenta of the function ψ into derivatives of the function φ at the origin, which are equal to zero. □

Below, for simplicity, we assume that $\gamma = 1$. Note that, unlike functions from $\mathcal{D}(\mathbb{S}^1)$, the Fourier transform of a function from $\mathcal{S}(\mathbb{S}^1)$ is not guaranteed to be analytic though it is infinitely differentiable. In particular, this is valid for the function φ selected above.

Select a function ψ with properties described in Lemma 8.5, construct the function family

$$\psi_\varepsilon(z) = \sum_j \frac{1}{\varepsilon}\psi\Big(\frac{t + 2j\pi}{\varepsilon}\Big), \tag{8.7}$$

and determine the approximation method by relation (7.4), where the convolution is treated in the space $\mathcal{D}'(\mathbb{S}^1)$. In terms of the Fourier coefficients, this map acts as follows:

$$R_\psi : f = \sum_k C_k z^k \to f_\varepsilon = \sum_k \varphi(2\pi k\varepsilon) C_k z^k, \tag{8.8}$$

where φ is the Fourier transform of the function ψ.

Theorem 8.4. *Under the specified choice of the function ψ_0, embedding* (8.8) *satisfies condition* (7.9) *and the multiplication localness property.*

Proof. In the considered case, according to Lemma 8.2, the asymptotic expansion is as follows: $f_\varepsilon \sim f$. This means that the difference $f - f_\varepsilon$ decreases faster than each power of ε. Since this holds for all derivatives of the function, we obtain that $f - f_\varepsilon$ belongs to the ideal.

To verify the multiplication localness, consider properties of the function ψ_ε from (8.7) on the period $[-\pi, \pi]$. Consider a neighborhood $(-\gamma, \gamma)$ of the origin. Since each p satisfies the inequality

$$|\psi(t)| \le \frac{C}{(1 + |t|)^p},$$

we have the following estimate for each series term from (8.7):

$$\Big|\frac{1}{\varepsilon}\psi\Big(\frac{t+2j\pi}{\varepsilon}\Big)\Big| \le \frac{1}{\varepsilon}\frac{C}{\big(1+\frac{|t+2j\pi|}{\varepsilon}\big)^p} \le \frac{C}{|t+2j\pi|^p}\varepsilon^{p-1}.$$

Therefore,

$$\sum_{j\ne 0}\frac{1}{\varepsilon}\psi\Big(\frac{t+2j\pi}{\varepsilon}\Big) \le C\sum_{j\ne 0}\frac{1}{|t+2j\pi|^p}\varepsilon^{p-1} = C_1\varepsilon^{p-1}.$$

For $j = 0$, we obtain the following estimate for the corresponding term:

$$\frac{1}{\varepsilon}\Big|\psi\Big(\frac{t}{\varepsilon}\Big)\Big| \le \begin{cases} \frac{C_2}{\varepsilon}, & |t| \le \gamma, \\ \frac{C}{\gamma}\varepsilon^{p-1}, & |t| \ge \gamma. \end{cases}$$

Thus, outside each neighborhood of the origin, the functions ψ_ε decrease faster than each power of ε. Therefore, for each $t_0 \ne 0$, the product $\psi_\varepsilon(t) \times \psi_\varepsilon(t - t_0)$ decreases faster than each power of ε and, therefore, belongs to the ideal. □

9. Analytic Representation of Distributions and Multiplication Generated by It

Another approximation method frequently used in analysis is based on a well-known analytic representation of distributions (see [12]). Under such approximations, more visible results are obtained in the problem of the multiplication of distributions. In the circle case, the analytic representation is determined as follows. Expand an arbitrary distribution f into the Fourier series and consider the operators

$$(P^+f)(z) := f^+(z) = \sum_0^\infty C_k z^k \tag{9.1}$$

and

$$(P^-f)(z) := f^-(z) = \sum_{-\infty}^{-1} C_k z^k. \tag{9.2}$$

From the estimate of coefficients C_k, it follows that for each $f \in \mathcal{D}'(\mathbb{S}_1)$, series (9.1) converges in the disk $\{|z| < 1\}$, its sum $f^+(z)$ is an analytic function, series (9.2) converges for $|z| > 1$, and its sum $f^-(z)$ is a function analytic for $|z| > 1$ and tending to zero at infinity.

Thus, the embedding $f \to (P^+f, P^-f) = (f^+, f^-)$ of the space of distributions into the space $\mathcal{A}(\mathbb{S}^1)$ of piecewise analytic functions, i.e., of pairs (f^+, f^-), where the function $f^+(z)$ is analytic for $|z| < 1$ and the function $f^-(z)$ is analytic for $|z| > 1$, is defined. In this case, the analytic functions $f^\pm(z)$ are not arbitrary: the coefficients of their power expansions do not increase faster than a fixed power of k.

The function pair (f^+, f^-) is called the *analytic representation* of the distribution f because f is expressed via these functions as follows:

$$<f,\varphi> = \lim_{\varepsilon\to 0}\frac{1}{2\pi}\int_{\mathbb{S}^1}\Big[f^+((1-\varepsilon)z) + f^-\Big(\frac{z}{1-\varepsilon}\Big)\Big]\varphi(z)|dz|. \tag{9.3}$$

If the analytic representation is equal to $(f^+, 0)$, then the distribution f is said to be *positive.* If the analytic representation is equal to $(0, f^-)$, then the distribution f is said to be *negative.*

Not every pair (f^+, f^-), where $f^+(z)$ is an analytic function for $|z| < 1$ and $f^-(z)$ is an analytic function for $|z| > 1$, determines an analytic representation of the distribution on the circle. Let us show that functions $f^\pm(z)$ determining such a representation can be characterized in terms of the growth rate as $|z| \to 1$; namely, they are power-growth functions, i.e., functions admitting the estimate

$$|f(z)| \le \frac{M}{(1-|z|)^m}. \tag{9.4}$$

The considered question is a special case of the general problem about relations between the behavior of an analytic function and the behavior of the corresponding coefficients of the corresponding power series. The most known result in this direction is about the relation between the order and type of entire functions and the behavior of the coefficients of its expansion (see [28]). For analytic functions in the disk, a similar result is obtained in [8]. For power-growth functions analytic in the disk, the description of the behavior of the coefficients is considered to be known, but it is not included into the standard source about analytic functions. For completeness, the corresponding assertion with the proof is provided below.

Theorem 9.1. *Let* $f(z) = \sum_0^\infty a_k z^k$ *and this series converge for* $|z| < 1$*. Then the sequence of coefficients admits a power estimate if and only if the function* f *does not increase faster than a power as* $|z| \to 1$*.*

Proof. Let estimate (9.4) be satisfied. Then, according to the Cauchy inequality,

$$|C_k| \leq \frac{M}{(1-|z|)^m |z|^k}.$$

In particular, assigning $z = 1 - \frac{1}{k}$, we obtain that

$$|C_k| \leq \frac{Mk^m}{(1-\frac{1}{k})^k} \leq 4Mk^m.$$

To obtain the inverse assertion, use the following result from [22]: the sum of the series

$$\sum_1^\infty k^\alpha z^k, \ \ \alpha > 0,$$

is an analytic function for $|z| = 1$ and $z \neq 1$ and it asymptotically behaves as

$$\frac{\Gamma(\alpha+1)}{(1-|z|)^{\alpha+1}}$$

in a neighborhood of the point $z = 1$.

Therefore, if the estimate $|C_k| \leq Mk^m, \ k > 0$, holds, then

$$|f(z)| \leq |a_0| + \sum_1^\infty Mk^m|z|^k \leq M_0 + \frac{\widetilde{M}}{(1-|z|)^{m+1}},$$

i.e., the function does not increase faster than a power. □

The analytic representation generates a natural approximation of the distribution f: relation (9.3) means that the family

$$R_a(f) = f_\varepsilon(z) = f^+((1-\varepsilon)z) + f^-\left(\frac{z}{1-\varepsilon}\right) \tag{9.5}$$

of smooth functions converges to f.

Theorem 9.2. *The relation* $R_a(f) = f_\varepsilon(z)$*, where the functions* $f_\varepsilon(z)$ *are defined by* (9.5)*, determines an invariant embedding of the space* $\mathcal{D}'(\mathbb{S}^1)$ *into the algebra of mnemofunctions. Condition* (7.10) *is satisfied for pairs* $f = (f^+, 0)$*,* $g = (g^+, 0)$ *of positive distributions, where the functions* f^+ *and* g^+ *are continuous in the closed disk, and for pairs* $f = (0, f^-)$*,* $g = (0, g^-)$ *of negative distributions, where* f^- *and* g^- *are continuous for* $|z| \geq 1$*.*

Proof. From Theorem 9.1, it follows that the family of smooth functions $f_\varepsilon(z)$ has a power growth and, therefore, determines a mnemofunction.

For the functions $z^k, \ k \geq 0$, relation (7.10) is verified immediately. Really, let $f(z) = z^k$, $g(z) = z^m$, and $f(z)g(z) = z^{k+m}$. Then $R_a(f) = (1-\varepsilon)^k z^k$, $R_a(g) = (1-\varepsilon)^m z^m$, and $R_a(fg) = (1-\varepsilon)^{k+m} z^{k+m}$.

Then the validity of the relation $R_a(fg) = R_a(f)R_a(g)$ is obvious. This implies that the desired relation is satisfied for all polynomials, whence, passing to the limit, we obtain the claimed property for all functions f^+ and g^+ continuous in the closed disk. For function pairs representable as $f = (0, f^-)$, $g = (0, g^-)$, the proof is similar. □

For the delta function, the considered approximating family has the following form:

$$R_a(\delta_1) = \psi_\varepsilon(z) = \frac{1}{1-(1-\varepsilon)z} + \frac{1-\varepsilon}{z-(1-\varepsilon)}. \tag{9.6}$$

Therefore, the considered approximation method is given by the convolution $R_a(f) = f * \psi_\varepsilon$.

Note that relation (7.9) is not satisfied under the considered approximation method. Indeed, the function $f(z) = z$ corresponds to the mnemonic function $f_\varepsilon(z) = (1-\varepsilon)z$, while $f(z) - f_\varepsilon(z) = \varepsilon z$, whence it follows that this difference does not belong to the ideal — it tends to zero as ε, and the elements of the ideal tend to zero faster than any power of ε.

From the proof of the theorem, we see that, in the considered case, relation (7.10) is satisfied not only on the quotient algebra, but at the level of representatives from equivalence classes as well, i.e., if f^+ and g^+ are continuous, then

$$(f^+ * \psi_\varepsilon) \times (g^+ * \psi_\varepsilon) = (f^+g^+) * \psi_\varepsilon.$$

The question whether other functions $\gamma \in C^\infty(\mathbb{S}^1)$ exist such that

$$(f^+ * \gamma) \times (g^+ * \gamma) = (f^+g^+) * \gamma \tag{9.7}$$

arises.

The next theorem states that the above case is almost unique.

Theorem 9.3. *If the function $\gamma(z)$ is such that relation* (9.7) *is satisfied for all smooth functions $f = (f^+, 0)$, then there exist ε and ξ, $|\xi| = 1$, such that*

$$\gamma(z) = \frac{1}{1-(1-\varepsilon)\xi z} + g(z),$$

where $g = (0, g^-)$.

Proof. Let $\gamma(z) = \sum A_k z^k$. Note that if f satisfies the assumption of the theorem, that its convolution with $\gamma(z)$ does not depend on coefficients of the expansion of $\gamma(z)$ with negative numbers. Therefore, the function

$$g^-(z) = \sum_{-\infty}^{-1} A_k z^k$$

can be selected arbitrarily. Then $z^k * \gamma(z) = A_k z^k$, $z^m * \gamma(z) = A_m z^m$, and $z^{k+m} * \gamma(z) = A_{k+m} z^{k+m}$. From (9.7), it follows that the coefficients of the expansion satisfy the relation $A_{k+m} = A_k A_m$. For $k = 0$, we obtain that $A_0 = 1$. Let $A_1 = r\xi$, where $|\xi| = 1$. Then $A_k = r^k \xi^k$ and $r < 1$ provided that the series converges. Thus, introducing the notation $1 - r = \varepsilon$, we obtain the claimed representation of $\gamma(z)$. □

Under the analytic representation, the space of distributions is identified with the set of pairs (f^+, f^-) of analytic functions described above. For each ε, the product $f_\varepsilon(z)g_\varepsilon(z)$ of mnemofunctions has an analytic representation$(h_\varepsilon^+(z), h_\varepsilon^-(z))$ as well; here, the functions $h_\varepsilon^\pm(z)$ analytically depend on z.

Consider this product in detail. Since the embedding is fixed, we denote the product of the distributions f and g by $f \times g$ or fg. Then the result of the multiplication $(f^+, f^-) \times (g^+, g^-)$ of the distributions can be represented as follows:

$$R_a(f)R_a(g) = \left[f^+((1-\varepsilon)z) + f^-\left(\frac{z}{1-\varepsilon}\right)\right]\left[g^+((1-\varepsilon)z) + g^-\left(\frac{z}{1-\varepsilon}\right)\right]$$

$$= f^+((1-\varepsilon)z)g^+((1-\varepsilon)z) + f^-\Big(\frac{z}{1-\varepsilon}\Big)g^+((1-\varepsilon)z) \tag{9.8}$$
$$+f^+((1-\varepsilon)z)g^-\Big(\frac{z}{1-\varepsilon}\Big) + f^-\Big(\frac{z}{1-\varepsilon}\Big)g^-\Big(\frac{z}{1-\varepsilon}\Big),$$

where

$$f^+((1-\varepsilon)z)g^+((1-\varepsilon)z) = R_a(f^+g^+, 0)$$

and

$$f^-\Big(\frac{z}{1-\varepsilon}\Big)g^-\Big(\frac{z}{1-\varepsilon}\Big) = R_a(0, f^-g^-),$$

i.e., the sum of the first term and fourth term is the analytic representation of the distribution defined by the pair (f^+g^+, f^-g^-). If ε is fixed, then the sum

$$\gamma_\varepsilon(z) := f^-\Big(\frac{z}{1-\varepsilon}\Big)g^+((1-\varepsilon)z) + f^+((1-\varepsilon)z)g^-\Big(\frac{z}{1-\varepsilon}\Big) \tag{9.9}$$

of the remaining two terms is a function analytic in the ring

$$K_\varepsilon = \Big\{z :\ 1-\varepsilon < |z| < \frac{1}{1-\varepsilon}\Big\}.$$

Its analytic representation is determined by means of the operators $P^\pm$. Thus, in the space $\mathcal{A}(\mathbb{S}^1)$ of piecewise analytic functions, i.e., of pairs (f^+, f^-) determining analytic representations, the multiplication acts according to the following rule.

Theorem 9.4. *In the space of piecewise analytic functions, i.e., of pairs (f^+, f^-) determining analytic representations, the result of the multiplication $(f^+, f^-) \times (g^+, g^-)$ can be represented in the form*

$$R_a(f)R_a(g) = h_\varepsilon^+(z) + h_\varepsilon^-(z), \tag{9.10}$$

where

$$h_\varepsilon^+(z) = R_a(f^+g^+, 0) + P^+(\gamma_\varepsilon(z)) \tag{9.11}$$

and

$$h_\varepsilon^-(z) = R_a(0, f^-g^-) + P^-(\gamma_\varepsilon(z)). \tag{9.12}$$

Corollary 9.1. *If $f = (f^+, 0)$ and $g = (g^+, 0)$, then $(f^+, 0) \times (g^+, 0) = (f^+g^+, 0)$.*

Corollary 9.2. *If $f = (0, f^-)$ and $g = (0, g^-)$, then $(0, f^-) \times (0, g^-) = (0, f^-g^-)$.*

10. Circle Algebra of Rational Mnemofunctions

A distribution $f \in \mathcal{D}'(\mathbb{S}^1)$ is said to be *rational* if the functions $f^\pm$ of its analytic representation are rational. Each pair of rational functions $f^\pm$ such that f^+ is analytic for $|z| < 1$ and f^- is analytic for $|z| > 1$ defines the analytic representation of a distribution. Therefore, the subspace $\mathcal{D}'_R(\mathbb{S}^1)$ of $\mathcal{D}'(\mathbb{S}^1)$ consisting of rational distributions is isomorphic to the space of pairs (f^+, f^-) of such rational functions.

There are two reasons to concentrate attention on such distributions. First, many distributions used in applications are rational and, therefore, the task to define products of such distributions well is important for applications. In particular, the analytic representation of δ_ξ is

$$\delta_\xi = \Big(-\frac{\xi}{z-\xi}, \frac{\xi}{z-\xi}\Big) \tag{10.1}$$

and the analytic representation of $\mathcal{P}\Big(\frac{1}{z-\xi}\Big)$ is

$$\mathcal{P}\Big(\frac{1}{z-\xi}\Big) = \frac{1}{2}\Big(\frac{1}{z-\xi}, \frac{1}{z-\xi}\Big), \tag{10.2}$$

i.e., both these distributions are rational.

The second reason is as follows: for such distributions, the multiplication rule is given explicitly (this is shown below), which provides more concrete results than the results in the general case. The rational distribution with the analytic representation $\left(\frac{1}{(z-\xi)^n}, 0\right)$, $|\xi| \geq 1$, is denoted by $\frac{1}{(z-\xi)^{n+}}$. The rational distribution with the analytic representation $\left(0, \frac{1}{(z-\eta)^m}\right)$, $|\eta| \leq 1$, is denoted by $\frac{1}{(z-\eta)^{m-}}$. For polynomials $p(z)$, we consider the embedding into the algebra by means of the analytic representation $(p(z), 0)$. Then the corresponding mnemofunction is $p_\varepsilon(z) = p((1-\varepsilon)z)$.

According to Corollary 9.1, the product $(f^+, 0) \times (g^+, 0)$ is defined for all functions f^+ and g^+ if they are analytic in the region $\{|z| < 1\}$. In this case, the following relation holds:

$$(f^+, 0) \times (g^+, 0) = (f^+ g^+, 0). \tag{10.3}$$

In particular, the following relation holds:

$$(z, 0) \times \left(\frac{1}{z-\xi}, 0\right) = \left(\frac{z}{z-\xi}, 0\right).$$

Note that

$$\frac{z}{z-\xi} = 1 + \frac{\xi}{z-\xi},$$

i.e., the multiplication by z is reduced to the multiplication by a constant and adding of 1. Using this relation, we obtain that

$$(z^2, 0) \times \left(\frac{1}{z-\xi}, 0\right) = (z, 0) \times \left(1 + \frac{\xi}{z-\xi}, 0\right) = \left(z + \xi\left(1 + \frac{\xi}{z-\xi}\right), 0\right) = \left(z + \xi + \frac{\xi^2}{z-\xi}, 0\right).$$

Similarly, it follows from Corollary 9.2 that

$$(0, f^-) \times (0, g^-) = (0, f^- g^-) \tag{10.4}$$

provided that the functions f^- and g^- are analytic in the region $\{|z| > 1\}$.

Therefore, to set the multiplication rule for rational distributions, it suffices to find the product of elements of the kind $(f^+, 0) \times (0, g^-)$.

First, consider the products of the distributions $\frac{1}{(z-\xi)^+}$ and $\frac{1}{(z-\eta)^-}$. Recall that these distributions are expanded into the Fourier series as follows:

$$\frac{1}{(z-\xi)^+} = -\sum_0^{+\infty} \xi^{-k-1} z^k \text{ and } \frac{1}{(z-\eta)^-} = \sum_{-\infty}^{-1} \eta^{-k-1} z^k. \tag{10.5}$$

The next lemma describes the product of such distributions.

Lemma 10.1. *If* $|\xi| \leq 1$ *and* $|\eta| \geq 1$, *then*

$$\left(\frac{1}{z-\xi}, 0\right) \times \left(0, \frac{1}{z-\eta}\right) = \left(\frac{C_1(r;\xi;\eta)}{z-\xi}, \frac{C_2(r;\xi;\eta)}{z-\eta}\right), \tag{10.6}$$

where $C_1(r;\xi;\eta) = \frac{r^2}{\xi - \eta r^2}$ *and* $C_2(r;\xi;\eta) = \frac{1}{\eta r^2 - \xi}$.

Proof. The mnemofunction corresponding to the distribution $\frac{1}{(z-\xi)^+}$ has the form

$$R_a\left(\frac{1}{(z-\xi)^+}\right) = \frac{1}{rz - \xi},$$

where $r = 1 - \varepsilon$. For $\frac{1}{(z-\eta)^-}$, the mnemofunction, i.e., the approximating family, is as follows:

$$R_a\left(\frac{1}{(z-\eta)^-}\right) = \frac{1}{\frac{z}{r} - \eta}.$$

Their product is the following family of smooth functions on the circle:

$$R_a\Big(\frac{1}{(z-\xi)^+}\Big)R_a\Big(\frac{1}{(z-\eta)^-}\Big) = \frac{1}{rz-\xi}\,\frac{1}{\frac{z}{r}-\eta}. \tag{10.7}$$

According to the multiplication rule in the space of piecewise analytic functions (see Theorem 9.4), for each ε, the analytic representation is to be constructed for function (10.7), i.e., the operators $P^{\pm}$ are to be applied. The main simplification for the considered functions is as follows: applying these operators is equivalent to the decomposition of the product of particular rational functions into a sum of partial fractions, which is done by means of simple computations. The result is as follows:

$$\frac{1}{rz-\xi}\,\frac{1}{\frac{z}{r}-\eta} = C_1(r;\xi;\eta)\frac{1}{rz-\xi} + C_2(r;\xi;\eta)\frac{1}{\frac{z}{r}-\eta}, \tag{10.8}$$

where

$$C_1(r;\xi;\eta) = \frac{r^2}{\xi-\eta r^2}, \quad C_2(r;\xi;\eta) = \frac{1}{\eta r^2-\xi},$$

and these coefficients satisfy the relation $C_1(r;\xi;\eta) = -r^2 C_2(r;\xi;\eta)$. The right-hand side of (10.8) is an approximating family for (10.6). □

Since the product of the distributions z^+ and $\frac{1}{(z-\eta)^-}$ in the sense of mnemofunctions is not coordinated with the classical multiplication of a smooth function and a distribution, the former is considered separately below.

Lemma 10.2. *If* $|\eta| \le 1$, *then*

$$(z,0)\times\Big(0,\frac{1}{z-\eta}\Big) = \Big(r^2,\frac{\eta r^2}{z-\eta}\Big) \tag{10.9}$$

and

$$(z^n,0)\times\Big(0,\frac{1}{z-\eta}\Big) = \Big(\sum_{k=0}^{n-1} z^k r^{2(n-k)}\eta^{n-k-1},\frac{\eta^n r^{2n}}{z-\eta}\Big). \tag{10.10}$$

Proof. Consider the product of the approximating families:

$$R_a(z^+)R_a\Big(\frac{1}{(z-\eta)^-}\Big) = rz\frac{1}{\frac{z}{r}-\eta} = r^2 + \frac{\eta r^2}{\frac{z}{r}-\eta}$$

(the last expression is obtained by means of applying the operators $P^{\pm}$).

Computing the product $(z^n,0)\times\Big(0,\frac{1}{z-\eta}\Big)$, we obtain other expressions. For the distribution $(z^n,0)$, the approximating family is $\{r^n z^n\}$. Therefore, the multiplication of approximations yields the family of smooth functions

$$r^n z^n \times \frac{1}{\frac{z}{r}-\eta}$$

such that their analytic representation contains both positive and negative components. Here, the desired result can be obtained by means of the procedure of the division with a remainder for polynomials:

$$r^n \times \frac{z^n}{\frac{z}{r}-\eta} = p_{n-1}(rz) + M\frac{1}{\frac{z}{r\eta}-1},$$

where $M = \eta^n r^{2n}$ and $p_{n-1}(z)$ is a polynomial of degree $n-1$ of the variable z with coefficients depending on r and η. These coefficients can be found by means of the immediate division or by means of the method of undetermined coefficients, but the simplest way to obtain the desired expression is to use the Fourier expansion. If

$$\frac{1}{(z-\eta)^-} = \sum_{-\infty}^{-1}\eta^{-k-1}z^k,$$

then

$$r^n z^n \times \frac{1}{\frac{z}{r} - \eta} = \sum_{k=0}^{n-1} (rz)^k r^{2(n-k)} \eta^{n-1-k} + r^{2n} \eta^n \frac{1}{\frac{z}{r\eta} - 1}.$$

The last expression is an approximating family for the distribution with the analytic representation (10.10). □

The next theorem completely describes the algebra generated by rational mnemofunctions.

Theorem 10.1. *The vector space $\mathcal{R}(\mathbb{S}^1)$ consisting of elements of the kind*

$$\sum_{k=0}^{m} A_k^+(\varepsilon)(1-\varepsilon)^k z^k + \sum_{k=1}^{n^+} \sum_{j=1}^{p_k^+} \frac{B_{kj}^+(\varepsilon)}{((1-\varepsilon)z - \xi_k)^j} + \sum_{k=1}^{n^-} \sum_{k=1}^{p_k^-} \frac{B_{kj}^-(\varepsilon)}{(\frac{z}{1-\varepsilon} - \eta_k)^j}, \tag{10.11}$$

where $|\xi_k| \geq 1$, $|\eta_k| \leq 1$, $A_k^+(\varepsilon)$, $B_{kj}^+(\varepsilon)$, and $B_{kj}^-(\varepsilon) \in \mathbb{C}^*$, is the least algebra containing all mnemofunctions $R_a(f)$, $f \in \mathcal{D}'_R(\mathbb{S}^1)$, and all generalized numbers. In this algebra, elements of the kind $\frac{1}{(z-\xi)^+}$, $\frac{1}{(z-\eta)^-}$, and z^+ are generators, while the multiplication law is uniquely defined by relations (10.3), (10.4), (10.6), *and* (10.9).

Proof. Only the first-degree polynomial is irreducible over the field of complex numbers. Hence, each rational function can be represented as a linear combination of partial fractions: if

$$Q(z) = \prod_{i=1}^{n} (z - z_i)^{p_i},$$

then

$$\frac{P(z)}{Q(z)} = \sum_{0}^{m} A_k z^k + \sum_{i=1}^{n} \sum_{j=1}^{p_i} \frac{B_{ij}}{(z - z_i)^j}, \qquad A_k,\ B_{ij} \in \mathbb{C},\ z_i \in \mathbb{C}.$$

In the analytic representation of the rational distribution $f = (f^+, f^-)$, the function f^+ is analytic for $|z| < 1$. Hence, if it is rational, then it is expanded into partial fractions as follows:

$$f^+(z) = \sum_{0}^{m} A_k^+ z^k + \sum_{k=1}^{n^+} \sum_{j=1}^{p_k^+} \frac{B_{kj}^+}{(z - \xi_k)^j}, \quad \text{where} \quad |\xi_k| \geq 1. \tag{10.12}$$

The rational function f^- is analytic for $|z| > 1$ and tends to zero at infinity. Therefore, its expansion into partial fractions is as follows:

$$f^-(z) = \sum_{k=1}^{n^-} \sum_{j=1}^{p_k^-} \frac{B_{kj}^-}{(z - \eta_k)^j}, \quad \text{where} \quad |\eta_k| \leq 1. \tag{10.13}$$

Thus, the multiplication of approximations generated by analytic representations of rational distributions is reduced to the computing of products (in the above sense) of summands in relations (10.12)-(10.13), i.e., partial fractions and z^k.

As we note above, for positive distributions, we have the relation

$$(f^+, 0) \times (g^+, 0) = (f^+ g^+, 0).$$

Respectively,

$$(0, f^-) \times (0, g^-) = (0, f^- g^-).$$

Therefore, the problem is reduced to the computing of products of a positive and a negative element. In the considered case of rational distributions, these are products of the kind $(f^+, 0) \times (0, g^-)$, where f^+ is a partial fraction or the function z^k, while g^- is a partial fraction.

For the case where $f^+ = \frac{1}{z-\xi}$ and $g^- = \frac{1}{z-\eta}$, this product is described in Lemma 10.1 and is found according to relation (10.6). Products of partial fractions of other powers, i.e., products with $f^+ = \frac{1}{(z-\xi)^n}$ and $g^- = \frac{1}{(z-\eta)^m}$, are computed via products of first degrees by means of recurrent relations, i.e., relation (10.6) is used. For example, for $n = 1$ and $m = 2$, we have the relation

$$\frac{1}{(z-\xi)^+}\frac{1}{(z-\eta)^{2-}} = \Big[\frac{1}{(z-\xi)^+}\frac{1}{(z-\eta)^-}\Big]\frac{1}{(z-\eta)^-} = \Big(\frac{c_1(\varepsilon;\xi;\eta)}{z-\xi}, \frac{c_2(\varepsilon;\xi;\eta)}{z-\eta}\Big) \times \Big(0, \frac{1}{z-\eta}\Big)$$

$$= \Big(\frac{c_1^2(\varepsilon;\xi;\eta)}{z-\xi}, \frac{c_1(\varepsilon;\xi;\eta)c_2(\varepsilon;\xi;\eta)}{z-\eta} + \frac{c_2(\varepsilon;\xi;\eta)}{(z-\eta)^2}\Big). \tag{10.14}$$

For $n = 2$ and $m = 1$, we have the relation

$$\frac{1}{(z-\xi)^{2+}}\frac{1}{(z-\eta)^-} = \Big(\frac{c_1(\varepsilon;\xi;\eta)}{z-\xi}, \frac{c_2(\varepsilon;\xi;\eta)}{z-\eta}\Big) \times \Big(\frac{1}{z-\xi}, 0\Big)$$

$$= \left(\frac{c_1(\varepsilon;\xi;\eta)}{(z-\xi)^2} + \frac{c_1(\varepsilon;\xi;\eta)c_2(\varepsilon;\xi;\eta)}{z-\xi}, \frac{c_2^2(\varepsilon;\xi;\eta)}{z-\eta}\right).$$

Hence, relation (10.6) sets the multiplication rule for the distributions $\frac{1}{(z-\xi)^{n+}}$ and $\frac{1}{(z-\eta)^{m-}}$.

The rule of the multiplication of z^+ by $\frac{1}{(z-\eta)^-}$ is described in Lemma 10.2; it is found via relation (10.9), where (and in the sequel) $r = 1 - \varepsilon$. As we show above, (10.9) implies the relation for the multiplication by powers of z^+, relation (10.10), and the relations for the multiplication of z by powers of $\frac{1}{(z-\eta)^-}$, e.g.,

$$z^+ \times \frac{1}{(z-\eta)^{2-}} = \Big[z^+\frac{1}{(z-\eta)^-}\Big]\frac{1}{(z-\eta)^-} = \Big(r^2, \frac{\eta r^2}{z-\eta}\Big) \times \Big(0, \frac{1}{z-\eta}\Big) = \Big(0, \frac{\eta r^2}{(z-\eta)^2} + \frac{r^2}{z-\eta}\Big).$$

Products of other powers are defined in the same way.

Thus, the product of elements of kind (10.11) is computed on the base of relations (10.6) and (10.9) and is an element of $\mathcal{R}(\mathbb{S}^1)$. Since elements of kind (10.11) belong to each algebra containing all mnemofunctions $R_a(f)$, $f \in \mathcal{D}'_R(\mathbb{S}^1)$ and generalized numbers, it follows that $\mathcal{R}(\mathbb{S}^1)$ is the least such subalgebra. $\square$

As we note, describing the general approach, a visible description of a product is given by its asymptotic relation. First, we investigate the behavior of products of the kind $\frac{1}{(z-\xi)^+} \times \frac{1}{(z-\eta)^-}$ found above.

Statement 10.1. *If $\xi \neq \eta$, then the following product asymptotic expansion takes place*:

$$\frac{1}{(z-\xi)^+} \times \frac{1}{(z-\eta)^-} = u_0 + \varepsilon u_1 + \dots,$$

where the distribution u_0 has the analytic representation

$$u_0 = \Big(\frac{1}{\xi-\eta}\frac{1}{z-\xi}, \frac{1}{\eta-\xi}\frac{1}{z-\eta}\Big),$$

while the distribution u_1 has the analytic representation

$$u_1 = \Big(-\frac{2\xi}{(\xi-\eta)^2}\frac{1}{z-\xi}, \frac{2\eta}{(\xi-\eta)^2}\frac{1}{z-\eta}\Big).$$

Proof. If $\xi \neq \eta$, then there exist finite limits of the coefficients $C_1(r;\xi;\eta)$ and $C_2(r;\xi;\eta)$ as $r \to 1$, i.e., their expansions start from (finite) numbers:

$$C_1(r;\xi;\eta) = \frac{r^2}{\xi - \eta r^2} = \frac{1}{\xi-\eta} - \frac{2\xi}{(\xi-\eta)^2}\varepsilon + \dots$$

and

$$C_2(r;\xi;\eta) = \frac{1}{\eta r^2 - \xi} = \frac{1}{\eta-\xi} + \frac{2\eta}{(\xi-\eta)^2}\varepsilon + \dots$$

Therefore, under the assumption that $\xi \neq \eta$, each mnemofunction of kind (10.7) is associated with the distribution u_0 such that its analytic representation is

$$u_0 = \Big(\frac{1}{\xi-\eta}\frac{1}{z-\xi}, \frac{1}{\eta-\xi}\frac{1}{z-\eta}\Big),$$

while the second term of the expansion has the form $\varepsilon R_a(u_1)$ in the mnemofunction space, where u_1 has the analytic representation

$$u_1 = \Big(-\frac{2\xi}{(\xi-\eta)^2}\frac{1}{z-\xi}, \frac{2\eta}{(\xi-\eta)^2}\frac{1}{z-\eta}\Big).$$

□

In this case, we note that if $|\xi| > 1$, then $f(z) = \frac{1}{z-\xi}$ is a smooth function on the circle and the first term of the expansion of u_0 is the product of the distribution $g = \frac{1}{(z-\eta)^-}$ by this smooth function in the sense of the multiplication in the space of distributions. This, in this case, the difference $R_a(fg) - R_a(f)R_a(g)$ is different from zero, but is infinitely small. This confirms again that, under the considered embedding, there is no coordination with the multiplication in the space of distributions (condition (7.8) is not satisfied), but the specified difference is associated with zero, i.e., under the passage to the algebra of mnemofunctions, the multiplication is corrected by means of adding infinitely small values.

If $|\xi| > 1$ and $|\eta| < 1$, then both functions f and g are smooth on the circle, but, under the considered embedding, condition (7.10), i.e., the Colombeau condition of the coordination with the multiplication of smooth functions, is not fulfilled either.

Consider the case where $\xi \neq \eta$, but $|\xi| = 1$ and $|\eta| = 1$. Then both distributions f and g have singularities on the circle, their product in the sense of distributions is not defined, but their product in the sense of mnemofunctions is associated with the distribution u_0. This confirms the conventional opinion that if the sets of singular points of two distributions do not intersect each other, then their product can be defined in a sufficiently natural way. However, even if two distributions have common singularities, there are cases (e.g., if these distributions are positive) where the product still can be defined.

The case where $\xi = \eta$, which is possible only if $|\xi| = 1$ and $|\eta| = 1$, is qualitatively different: both factors have singularities at the same point.

Statement 10.2. *The product* $\frac{1}{(z-\xi)^+} \times \frac{1}{(z-\xi)^-}$ *has the asymptotic expansion*

$$\frac{1}{(z-\xi)^+} \times \frac{1}{(z-\xi)^-} = -\frac{1}{2\xi\varepsilon}\delta_\xi - \frac{3}{4\xi}\frac{1}{(z-\xi)^+} - \frac{1}{4\xi}\frac{1}{(z-\xi)^-} + o(1),$$

the principal term of which is the δ-function with an infinitely large coefficient.

Proof. In the considered case, the coefficients $C_1(r)$ and $C_2(r)$ are infinitely large as $r \to 1$ and they are asymptotically expanded as follows:

$$c_1(\varepsilon;\xi) = \frac{(1-\varepsilon)^2}{\xi(1-(1-\varepsilon)^2)} = \frac{1}{2\xi\varepsilon} - \frac{3}{4\xi} + \frac{1}{8\xi}\varepsilon + \frac{3}{16\xi}\varepsilon^2 + \dots,$$

$$c_2(\varepsilon;\xi) = \frac{1}{\xi((1-\varepsilon)^2-1)} = -\frac{1}{2\xi\varepsilon} - \frac{1}{4\xi} - \frac{1}{8\xi}\varepsilon - \frac{1}{16\xi}\varepsilon^2 + \dots \tag{10.15}$$

This yields the following asymptotic expansion for the product:

$$\frac{1}{(z-\xi)^+}\frac{1}{(z-\xi)^-} = c_1(\varepsilon;\xi)\frac{1}{(z-\xi)^+} + c_2(\varepsilon;\xi)\frac{1}{(z-\xi)^-}$$

$$= \frac{1}{2\xi\varepsilon}\Big[\frac{1}{(z-\xi)^+} - \frac{1}{(z-\xi)^-}\Big] - \frac{3}{4\xi}\frac{1}{(z-\xi)^+} - \frac{1}{4\xi}\frac{1}{(z-\xi)^-} + o(1) \tag{10.16}$$

$$= -\frac{1}{2\xi\varepsilon}\delta_\xi - \frac{3}{4\xi}\frac{1}{(z-\xi)^+} - \frac{1}{4\xi}\frac{1}{(z-\xi)^-} + o(1).$$

□

11. Products Associated with Distributions

In the general case, the product of distributions f and g generated by a given embedding is a mnemofunction. A special case is formed by distribution pairs f and g such that their product is associated with a distribution u because, in such a case, one can assume that $f \times g = u$, i.e., the product of distributions is a distribution as well. In the general case, there are no reasons to expect to obtain an explicit description of such pairs of distributions. For pairs of rational distributions, all cases where their product is associated with a distribution can be described.

As we note above, the product of rational distributions is a linear combination of products of elementary rational mnemofunctions of the kinds $\frac{1}{(z-\xi)^{n+}}$, $\frac{1}{(z-\eta)^{n-}}$, and z^{n+}. According to assertions obtained above, each product $\frac{1}{(z-\xi)^{n+}} \times \frac{1}{(z-\eta)^{n-}}$ is represented by a linear combination of elementary rational mnemofunctions with coefficients depending on ε. If the corresponding expression for the product includes terms of the kind $\frac{1}{(z-\xi)^{n+}} \times \frac{1}{(z-\xi)^{m-}}$, then terms with infinitely large coefficients, not associated with distributions, arise. However, the final representation of the product might include several such terms with different coefficients. This the existence condition for a distribution associated with the corresponding product is as follows: the sum of the coefficients at each product of the specified kind has a finite limit.

In the following case, this condition is satisfied for arbitrary mnemofunctions. If

$$f = f^+ + f^-, \quad g = t(f^+ - f^-), \text{ where } \quad t \in \mathbb{C}, \tag{11.1}$$

then $f \times g = t(f^+)^2 - t(f^-)^2$ and this product is a distribution such that its analytic representation is $(t(f^+)^2, -t(f^-)^2)$.

In the general case, relation (11.1) connecting distributions is a sufficient (but not necessary) condition for the product to be associated with a distribution. First, we consider an example of rational distributions with a singularity at only one point such that condition (11.1) is necessary for them as well.

Theorem 11.1. *Let $f = (f^+, f^-)$ and $g = (g^+, g^-)$, where*

$$f^\pm = \frac{A_1^\pm}{(z-1)^\pm} + \frac{A_2^\pm}{(z-1)^{2\pm}} \neq 0 \text{ and } g^\pm = \frac{B_1^\pm}{(z-1)^\pm} + \frac{B_2^\pm}{(z-1)^{2\pm}}.$$

The product of the distributions f and g is associated with a distribution u_0 if and only if there exists t such that the coefficients satisfy the relations

$$B_1^+ = tA_1^+, \quad B_2^+ = tA_2^+, \quad B_1^- = -tA_1^-, \text{ and } B_2^- = -tA_2^-. \tag{11.2}$$

If these conditions are satisfied, then the distribution u_0 has an analytic representation

$$u_0 = (f^+g^+, f^-g^-).$$

Proof. We have the relation

$$f \times g = (f^+ + f^-) \times (g^+ + g^-) = f^+g^+ + f^-g^- + f^+g^- + g^+f^-.$$

The sum $f^+g^+ + f^-g^-$ is associated with the distribution u_0 such that its analytic representation is equal to $u_0 = (f^+g^+, f^-g^-)$ and infinitely large terms arise only in the sum of the last two terms of the kind

$$f^+g^- + g^+f^- = \frac{A_2^+B_2^- + A_2^-B_2^+}{(z-1)^{2+}(z-1)^{2-}}$$

$$+\frac{A_2^+B_1^- + A_1^-B_2^+}{(z-1)^{2+}(z-1)^-} + \frac{A_1^+B_2^- + A_2^-B_1^+}{(z-1)^+(z-1)^{2-}} + \frac{A_1^+B_1^- + A_1^-B_1^+}{(z-1)^+(z-1)^-}.$$

The obtained expression is equal to zero if the numerator of each fraction is equal to zero:

$$\begin{cases} A_2^+B_2^- + A_2^-B_2^+ = 0, \\ A_2^+B_1^- + A_1^-B_2^+ = 0, \\ A_1^+B_2^- + A_2^-B_1^+ = 0, \\ A_1^+B_1^- + A_1^-B_1^+ = 0. \end{cases} \tag{11.3}$$

To investigate system (11.3), consider the polynomials

$$f^+(x) = A_1^+x + A_2^+x^2;\quad f^-(y) = A_1^-y + A_2^-y^2; g^+(x) = B_1^+x + B_2^+x^2;\ \text{and}\ g^-(y) = B_1^-y + B_2^-y^2.$$

Then system (11.3) is equivalent to the condition that

$$f^+(x)g^-(y) + g^+(x)f^-(y) = 0,$$

whence, separating variables, we obtain the following proportionality conditions for the polynomials:

$$\frac{g^+(x)}{f^+(x)} = -\frac{g^-(y)}{f^-(y)} = t = const.$$

Let us show that if relations (11.3) are not satisfied, then the expansion of $f^+g^- + g^+f^-$ contains terms with infinitely large coefficients and, therefore, is associated with no distribution. Indeed, due to the multiplication rule, we have the relation

$$f^+g^- + g^+f^- = (A_2^+B_2^- + A_2^-B_2^+)\left(\frac{c_1^2(\varepsilon)}{(z-1)^{2+}} + \frac{2c_1^2(\varepsilon)c_2(\varepsilon)}{(z-1)^+} + \frac{2c_1(\varepsilon)c_2^2(\varepsilon)}{(z-1)^-} + \frac{c_2^2(\varepsilon)}{(z-1)^{2-}}\right)$$

$$+\Big[c_2(\varepsilon)(A_2^+B_1^- + A_1^-B_2^+) + c_1(\varepsilon)(A_1^+B_2^- + A_2^-B_1^+)\Big]\left(\frac{c_1(\varepsilon)}{(z-1)^+} + \frac{c_2(\varepsilon)}{(z-1)^-}\right)$$

$$+(A_1^+B_1^- + A_1^-B_1^+)\left(\frac{c_1(\varepsilon)}{(z-1)^+} + \frac{c_2(\varepsilon)}{(z-1)^-}\right) + (A_2^+B_1^- + A_1^-B_2^+)\frac{c_1(\varepsilon)}{(z-1)^{2+}} + (A_1^+B_2^- + A_2^-B_1^+)\frac{c_2(\varepsilon)}{(z-1)^{2-}},$$

where the terms

$$\frac{2c_1(\varepsilon)c_2^2(\varepsilon)}{(z-1)^-} \quad \text{and} \quad \frac{2c_1^2(\varepsilon)c_2(\varepsilon)}{(z-1)^+}$$

are linearly independent and, since the coefficients $c_1(\varepsilon)$ and $c_2(\varepsilon)$ behave as $\frac{1}{\varepsilon}$, only these terms have the greatest growth rate; it is equal to $\frac{1}{\varepsilon^3}$. Therefore, for the existence of an associated distribution, the vanishing of the coefficient at these terms is necessary:

$$A_2^+B_2^- + A_2^-B_2^+ = 0.$$

In the same way, the terms

$$(A_2^+B_1^- + A_1^-B_2^+)\frac{c_1(\varepsilon)}{(z-1)^{2+}} + (A_1^+B_2^- + A_2^-B_1^+)\frac{c_2(\varepsilon)}{(z-1)^{2-}}$$

increase as $\frac{1}{\varepsilon}$, are linearly independent between each other, and are linearly independent with other terms. Hence, the necessary condition is as follows:

$$A_2^+B_1^- + A_1^-B_2^+ = 0 \text{ and } A_1^+B_2^- + A_2^-B_1^+ = 0.$$

Once these conditions are satisfied, only the expression

$$\frac{c_1(\varepsilon)}{(z-1)^+} + \frac{c_2(\varepsilon)}{(z-1)^-}$$

remains. It increases as $\frac{1}{\varepsilon}$. Therefore, for the existence of a distribution associated with the product, the coefficient at this expression is to vanish as well:

$$A_1^+B_1^- + A_1^-B_1^+ = 0.$$

Thus, condition (11.2) is satisfied if and only if the considered product is associated with the distribution u_0. □

In the same way, we obtain necessary and sufficient conditions of the existence of an associated distribution for the product of an arbitrary pair of rational distributions.

Let $\{z_1, z_2, \dots, z_m\}$ be a set of complex numbers such that $|z_k| = 1$. As we show above, the existence problem for an associated distribution for a product is reduced to the investigation of rational distributions f and g with analytic representations of the kind

$$f = \Big(\sum_{k=1}^{m} f_k^+(z), \sum_{k=1}^{m} f_k^-(z)\Big), \tag{11.4}$$

where

$$f_k^+(z) = \sum_{j=1}^{p} \frac{A_{kj}^+}{(z-z_k)^j} \text{ and } f_k^-(z) = \sum_{j=1}^{p} \frac{A_{ij}^-}{(z-z_k)^j},$$

and

$$g = \Big(\sum_{k=1}^{m} g_k^+(z), \sum_{k=1}^{m} g_k^-(z)\Big), \tag{11.5}$$

where

$$g_k^+(z) = \sum_{j=1}^{p} \frac{B_{kj}^+}{(z-z_k)^j} \text{ and } g_k^-(z) = \sum_{j=1}^{p} \frac{B_{ij}^-}{(z-z_k)^j}.$$

Theorem 11.2. *The product fg of the rational mnemofunctions expressed by* (11.4) *and* (11.5) *is associated with a distribution if and only if for each k, $1 \le k \le m$, the exists a number t_k such that*

$$B_{kj}^+ = t_k A_{kj}^+ \text{ and } B_{kj}^- = -t_k A_{kj}^-.$$

In [9], only three examples of finite products, i.e., cases where the result of the multiplication of distributions is a distributions, are provided. On the circle, analogs of these examples are the products

$$\Big(\delta_1 + 2i\mathcal{P}\Big(\frac{1}{z-1}\Big)\Big)\Big(\delta_1 - 2i\mathcal{P}\Big(\frac{1}{z-1}\Big)\Big), \quad \delta_1\mathcal{P}\Big(\frac{1}{z-1}\Big), \text{ and } \Big(\delta_1 \pm 2\mathcal{P}\Big(\frac{1}{z-1}\Big)\Big)^2.$$

All these products satisfy the condition obtained in the theorem. Many other examples can be provided.

To complete this section, consider the Schwartz example on the circle and find the sense of the product of the distributions $\mathcal{P}\Big(\frac{1}{z-1}\Big)$, $z-1$, and δ_1, contained in this example. Under the multiplication in the space of distribution, the value of such a product depends on the arrangement of the brackets:

$$\Big\{\mathcal{P}\Big(\frac{1}{z-1}\Big) \times (z-1)\Big\} \times \delta_1 = 1 \times \delta_1 = \delta_1,$$

but

$$\mathcal{P}\Big(\frac{1}{z-1}\Big) \times \{(z-1) \times \delta_1\} = \mathcal{P}\Big(\frac{1}{z-1}\Big) \times 0 = 0.$$

If the order of the factors is changed, then the obtained expression is not defined:

$$\Big\{\mathcal{P}\Big(\frac{1}{z-1}\Big) \times \delta_1\Big\} \times (z-1).$$

By virtue of relations (10.1)-(10.2), we obtain that

$$R_a\Big(\mathcal{P}(\frac{1}{z-1})\Big)R_a((z-1))R_a(\delta_1)$$
$$= \frac{1}{2}\Big(\frac{1}{\frac{z}{r}-1} - \frac{1}{rz-1}\Big) - \varepsilon\Big(\frac{1}{\frac{z}{r}-1} + \frac{1}{2}\frac{1}{(\frac{z}{r}-1)^2}\Big) + \frac{\varepsilon^2}{2}\Big(\frac{1}{\frac{z}{r}-1} + \frac{1}{(\frac{z}{r}-1)^2}\Big),$$

which implies that the asymptotic expansion of this product is as follows:

$$\frac{1}{2}\delta_1 - \varepsilon\frac{1}{(z-1)^-} - \frac{\varepsilon}{2}\frac{1}{(z-1)^{2-}} + o(\varepsilon).$$

Thus, the product from the Schwartz example is associated with $\frac{1}{2}\delta_1$.

12. On Equations with Generalized Coefficients

12.1. Approximation approach. Primarily, the embedding of the whole space of distributions into the algebra of mnemofunctions is interesting from the theoretical viewpoint: it resolves the Schwartz problem. In particular problems including products of distributions, other questions related to the mnemofunction theory arise as well. The theory of nonlinear equations and differential equations with generalized coefficients are among such problems.

From the viewpoint of applications, the main difficulty of the specified problems is the incorrectness of the problem set originally: the mathematical model of the investigated process is too rough, which causes the fact that no notion of a solution is defined for the corresponding equation. The general approach is based on refining the problem setting by means of adding infinitely small values that are not taken into account in the original setting. Such a refining is including additional data extracted from the application domain (usually, this data is not contained in the originally formulated problem).

For example, in the theory of nonlinear equations, discontinuous solutions of the shock-wave type are interesting (see [19, 33]). The corresponding equations contain nonlinear terms, e.g., of the kind $u'_x u$, while no such product is defined for a discontinuous function u. Thus, no notion for a discontinuous solution is defined.

The problem setting is refined by means of adding terms containing a small parameter ε to the nonlinear equation. The solution of such a perturbed equation is a family of smooth functions u_ε, i.e., a mnemofunction. The solution of the original equation is the limit of solutions u_ε as $\varepsilon \to 0$, i.e., the distribution u associated with this mnemofunction.

Thus, in these problems, the method of the approximation of the desired solution (the family of u_ε) appears from the refined equation and is determined by the physical interpretation of the problem. Usually, in classical hydrodynamical examples, the original equation corresponds to the case of the ideal liquid, while the additional term reflects the effect of the small viscosity (i.e., the viscosity of order ε); that is why such an approach is called the *vanishing viscosity method*.

The general scheme of the approximation approach for the investigation of linear differential equations with generalized coefficients refines the problem setting as well. Its content is as follows. Let L be a linear differential expression with generalized coefficients. This expression is formal because it contains terms that are not defined. We have to construct an operator (in a suitable function space) corresponding to the considered operator L, which is equivalent to the introducing a notion of a solution for the equation $Lu = f$. Usually, the original expression immediately determines an operator A_0 defined only on a rather narrow subspace, while our task is to construct its extension A. Properties of the corresponding operators A for various types of differential expressions with generalized coefficients are investigated in many papers. Here, we discuss only the construction of such extensions.

The approximation approach to the investigation of the considered equations consists of the change of coefficients that are generalized functions for mnemofunctions associated with them. Here, the choice of the corresponding mnemofunctions is adding the data refining the problem setting. We obtain a family $\{L_\varepsilon\}$ of (well-defined) differential operators with smooth coefficients in a suitable space. Since more than one mnemofunction is associated to each distribution, it follows that many various families of operators L_ε are associated with the original expression. In many cases, solutions u_ε of the corresponding equations $L_\varepsilon u_\varepsilon = f$ converge to a function (distribution) u. Then u is reasonably treated as a *solution of the original equation corresponding to the selected method of the approximation of coefficients*.

From a more general viewpoint, it is more convenient to investigate the family of operators $L_\varepsilon - \lambda I$ with the spectral parameter. Then the solutions of the corresponding equations are

$$u_\varepsilon = (L_\varepsilon - \lambda I)^{-1} f$$

and the problem is reduced to the proof of the strong convergence of the resolvents $(L_\varepsilon - \lambda I)^{-1}$. In this case, there exists an operator A such that the limit of the resolvents is its resolvent; this operator A is called a *limit of the family* $\{L_\varepsilon\}$ *in the sense of the resolvent convergence.* Considering operators with the spectral parameter, one can obtain more general results because if we fix a particular value λ_0, then this might be a spectral value for A; in this case, no limit of the function family $\{(L_\varepsilon - \lambda_0 I)^{-1} f\}$ exists.

The advantage of such an approach is as follows: usually, the family of L_ε does not converge in the space of operators and the operator A cannot be defined as the limit of $\{L_\varepsilon\}$ under other convergence types.

In this case, the main technical difficulty is the investigation of the behavior of the family $\{u_\varepsilon\}$ of solutions of the associated equation in the space of mnemofunctions and the finding of the limit, i.e., the operator A. Here, the following primary questions arise:

(1) What formal differential expressions L with generalized coefficients and what approximations of these coefficients are such that there exists a limit of the family $\{L_\varepsilon\}$ in the sense of the resolvent convergence?
(2) What operators might be limits of approximating families under various approximation methods and what are the cases where the limit does not depend on the selected approximation method for the coefficients?

12.2. First-order differential equations. To clarify the arising problems and possible answers, consider the simplest case, which is the first-order linear differential equation

$$u' - au = f \tag{12.1}$$

with a generalized coefficient a. This prototype example is convenient to demonstrate arising impacts and differences from the classical theory of differential equations such that the results can be formulated explicitly. For equations of higher orders and partial differential equations, similar impacts take place, but much more fine computations are required to obtain the corresponding assertions.

Recall that the solution of the homogeneous Eq. (12.1) with an integrable coefficient a is given by the relation

$$V(x) = \exp\Big[\int_{-1}^{x} a(s)ds\Big], \tag{12.2}$$

while the solution of the Cauchy problem for the homogeneous equation with the condition $u(-1) = M$ is given by the relation

$$u(x) = MV(x) + V(x)\int_{-1}^{x} \frac{1}{V(t)} f(t)dt. \tag{12.3}$$

Note that V is found within two steps: finding the primitive $g(x)$ of a and computing the exponent $\exp g(x)$. Additionally, relation (12.3) includes multiplication and integration. Therefore, if a and f are generalized functions, then, to obtain a similar relation, one has to define these operations well. For the generalized function a, a primitive (i.e., a distribution such that $g' = a$) exists. Therefore, to provide a sense to relation (12.2), i.e., to introduce the notion of a solution for the homogeneous equation with a generalized coefficient, one has to define the exponent of the distribution g.

According to the general approach, instead of the distribution a and f, consider their approximations by smooth functions a_ε and f_ε. We obtain the following equation with a small parameter (an equation in the space of mnemofunctions):

$$u'_\varepsilon - a_\varepsilon u_\varepsilon = f_\varepsilon. \tag{12.4}$$

For it, the solution of the Cauchy problem is

$$u_\varepsilon(x) = M \exp\Big[\int_{-1}^{x} a_\varepsilon(s)ds\Big] + \exp\Big[\int_{-1}^{x} a(s)ds\Big]\int_{-1}^{x} \exp\Big[-\int_{-1}^{t} a_\varepsilon(s)ds\Big] f_\varepsilon(t)dt. \qquad (12.5)$$

If this family of smooth functions u_ε converges (in the space of distributions or in a given function space), then its limit u is called the *solution of the original equation generated by the selected approximation method*. The following examples demonstrate that there are many qualitatively different cases of the behavior of family (12.5) of solutions.

Example 12.1. Consider the equation

$$Lu \equiv u' - b\delta u = f, \quad b = const,$$

in the space $L_2([-1,1])$. The task is to define a solution of this equation containing a formal expression with a generalized coefficient.

Since $\delta = 0$ outside each neighborhood of the origin, it follows that if a solution of the homogeneous equation is defined reasonably, then it is to be of the form

$$u(x) = \begin{cases} C, & x < 0, \\ C_1, & x > 0, \end{cases} \qquad (12.6)$$

or, which is equivalent,

$$u(x) = C + (C_1 - C)\Theta(x).$$

Substituting this function into the equation, we obtain the product $\delta\Theta$. It is not defined and C_1 cannot be found from the equation. This confirms that no notion of the solution of this equation is defined in the framework of the classical theory.

Change the δ-function for its approximation of the kind $\varphi_\varepsilon(x) = \frac{1}{\varepsilon}\varphi\Big(\frac{x}{\varepsilon}\Big)$. Taking into account that the supports of the functions $\varphi_\varepsilon(x)$ are contained in the segment $[-1,1]$ for small values of ε, introduce the notation

$$\Phi(x) = \int_{-1}^{x} \varphi(t)dt.$$

Then the equation with $\varphi_\varepsilon(x)$ is satisfied by

$$V_\varepsilon(x) = C\exp\Big[b\Phi\Big(\frac{x}{\varepsilon}\Big)\Big].$$

From this expression, we see that $V_\varepsilon(x)$ converge to the function

$$V(x) = C\big[1 + (e^b - 1)\Theta(x)\big]. \qquad (12.7)$$

Thus, the homogeneous equation is satisfied by the piecewise constant function (12.7) satisfying the following condition for the jump at the origin (this condition depends on the coefficient b):

$$u(+0) = e^b u(-0).$$

Note that primitives of δ are ordinary functions of the kind $g(x) = \Theta(x) + \widetilde{C}$ and, substituting them into (12.2), i.e., computing the exponent formally, we obtain the same solutions.

Remark 12.1. Substituting (12.7) into the equation, we obtain the relation

$$C(e^b - 1)\delta = bC[\delta + (e^b - 1)\delta\Theta(x)].$$

From this relation, we find the product

$$\delta\Theta = \frac{e^b - 1 - b}{b(e^b - 1)}\ \delta.$$

Note that the value obtained for $\delta\Theta$ does not depend on the coefficient b of the equation. The reason is that the above computations use the approximation for $\Theta(x)$ determined by the considered equation and the approximation of δ; this leads to the dependence of the result on b. This reasoning confirms that the product $\delta\Theta$ cannot be defined uniquely.

If f is a function from $L_2[-1,1]$, then solutions of the approximating equation

$$u' - b\frac{1}{\varepsilon}\varphi\Big(\frac{x}{\varepsilon}\Big) = f$$

satisfying the condition $u(-1) = C$ are as follows:

$$u_\varepsilon(x) = C\exp\Big[b\Phi\Big(\frac{x}{\varepsilon}\Big)\Big] + \exp\Big[b\Phi\Big(\frac{x}{\varepsilon}\Big)\Big]\int_{-1}^{x}\exp\Big[-b\Phi\Big(\frac{t}{\varepsilon}\Big)\Big]f(t)dt.$$

Here, the limit passage as $\varepsilon \to 0$ is possible as well. This yields

$$u_\varepsilon(x) \to u(x) = \begin{cases} C + \int_{-1}^{x} f(t)dt, & x < 0, \\ Ce^b + \int_{-1}^{x} f(t)dt, & x > 0. \end{cases}$$

Introducing the notation

$$u_0(x) = C + \int_{-1}^{x} f(t)dt,\ x \in [-1,1],$$

one can represent the obtained solution as follows:

$$u(x) = u_0(x) + u_0(-1)(e^b - 1)\Theta(x).$$

Let us find the operator that is associated with the formal expression $Lu = u' - b\delta u$ as a result of the performed computations.

The differentiation operator is defined on the subspace $H^1[-1,1] \subset L_2[-1,1]$ consisting of absolutely continuous functions such that their derivatives belong to $L_2[-1,1]$. For such a function, we have $\delta u = u(0)\delta$ and this product belongs to $L_2[-1,1]$ only if $u(0) = 0$. Thus, the formal expression defines the operator A_0 in $L_2[-1,1]$ defined only on the subspace $D(A_0) = \{u \in H^1[-1,1] : u(0) = 0\}$ and acting as the differentiation. Therefore, each operator such that it is reasonable to associate it with the formal expression has to be an extension of A_0.

The expression for the solution obtained above means that, in the specified case, we associate the operator A_b defined on a subspace $D(A_b) \subset H^1[-1,0] \oplus H^1[0,1]$ of the kind

$$D(A) = \big\{u(x) = u_0(x) + u_0(-1)(e^b - 1)\Theta(x) : u_0 \in H^1[-1,1]\big\}$$

and acting according to the relation $A_b u(x) = u_0'(x)$ with the formal expression $Lu = u' - b\delta u$.

Note that the operator A_b from this example does not depend on the selected approximation method for the δ-function. Also, note that the constructed operator is not the closure of A_0.

Construct the operator in a broader space and find the changes caused by this extension. Namely, take the space

$$F[-1,1] = \{f + M\delta; f \in L_2[-1,1],\ M \in \mathbb{C}\}$$

containing the delta function. Let $\psi_\varepsilon(x) = \frac{1}{\varepsilon}\psi\Big(\frac{x}{\varepsilon}\Big)$ be the approximation of δ from the right-hand side of the equation generated by a function $\psi \in \mathcal{D}(\mathbb{R})$. Then the solution of the approximating equation

$$u' - b\varphi_\varepsilon u = f + M\psi_\varepsilon$$

satisfying the condition $u(-1) = C$ is

$$u_\varepsilon(x) = C\exp\Big[b\Phi\Big(\frac{x}{\varepsilon}\Big)\Big] + \exp\Big[b\Phi\Big(\frac{s}{\varepsilon}\Big)\Big]\int_{-1}^{x}\exp\Big[-b\Phi\Big(\frac{t}{\varepsilon}\Big)\Big]f(t)dt$$

$$+M\exp\Big[b\Phi\Big(\frac{x}{\varepsilon}\Big)\Big]\int_{-1}^{x}\exp\Big[-b\Phi\Big(\frac{s}{\varepsilon}\Big)\Big]\frac{1}{\varepsilon}\psi\Big(\frac{s}{\varepsilon}\Big)ds.$$

In the integral from the last term, change the variable as follows: $t = \frac{s}{\varepsilon}$. Then it is equal to

$$h_\varepsilon(x) = M \exp\left[b\Phi\left(\frac{x}{\varepsilon}\right)\right] \int_{-\frac{1}{\varepsilon}}^{\frac{x}{\varepsilon}} \exp[-b\Phi(t)]\psi(t)dt.$$

In this expression, it is possible to pass to the limit:

$$h_\varepsilon(x) \to h(x) = \begin{cases} 0, & x < 0, \\ Me^b\Gamma, & x > 0, \end{cases}$$

where

$$\Gamma = \int_{-\infty}^{+\infty} \exp[-b\Phi(t)]\psi(t)dt.$$

Thus, solutions u_ε converge to the function

$$u(x) = \begin{cases} C + \int_{-1}^{x} f(t)dt, & x < 0, \\ Ce^b + Me^b\Gamma + \int_{-1}^{x} f(t)dt, & x > 0. \end{cases}$$

From the operator viewpoint, the operator A in the space $F[-1, 1]$ is associated with the formal expression such that

$$D(A) = H^1[-1, 0] \oplus H^1[0, 1].$$

Each function u of this domain is represented as follows:

$$u = u_0 + \lambda\Theta = u_0 + u_0(-1)(e^b - 1)\Theta + \mu\Theta, \text{ where } \mu = \lambda - u_0(-1)(e^b - 1),\ u_0 \in H^1[-1, 1].$$

This operator acts as follows:

$$Au = u_0' + \frac{\mu}{e^b\Gamma}\delta.$$

First, we note the following qualitative difference: the result contains the number Γ depending on the approximation method for the coefficient and right-hand side. Thus, for the definition of the notion of solutions of the equation in the space L_2, the usage of mnemofunctions plays a technical role, but, once we pass to a broader space including distributions, the solution itself depends on approximation methods.

Example 12.2. Consider the equation

$$u' - b\delta' u = 0, \quad b = const,$$

where the coefficient is the derivative of the δ-function, and follows the casing differences.

Since the primitive for δ' is the distribution δ, it follows that, due to relation (12.2), the formal solution satisfying the condition $u(-1) = C$ is given by an expression of the kind $Ce^{b\delta}$, which is not defined. Therefore, this might mean obstacles for constructing a solution (in any sense).

Changing δ' for its approximation $\delta_\varepsilon'(x) = \frac{1}{\varepsilon^2}\varphi'\left(\frac{x}{\varepsilon}\right)$, we obtain solutions of the approximating homogeneous equation

$$V_\varepsilon(x) = C \exp\left[b\frac{1}{\varepsilon}\varphi\left(\frac{x}{\varepsilon}\right)\right].$$

Outside any neighborhood of the origin, this family of functions uniformly converges to the constant C. However, the family $\{V_\varepsilon(x)\}$ might exponentially increase in a neighborhood of the origin; in this case, it does not converge to the constant C in the space of distributions.

Here, we restrict ourselves by the following conclusion: equations with coefficients of greater orders of singularity are more complicated and, therefore, other approaches are required to investigate them.

Example 12.3. Classical solutions of the differential equation

$$u' + \frac{1}{x}u = 0 \tag{12.8}$$

are functions of the kind

$$u(x) = \begin{cases} \frac{C}{x}, & x < 0, \\ \frac{C_1}{x}, & x > 0. \end{cases} \tag{12.9}$$

Note that the solution of the Cauchy problem with the condition $u(-1) = -C$ for this equation is $u(x) = \frac{C}{x}$ for $x < 0$. This solution tends to infinity as $x \to -0$. No methods of the classical theory of differential equations provide a possibility to extend this solution to the positive semiaxis in a natural way.

Let us consider possibilities to define a solution of Eq. (12.8) from the viewpoint of the mnemofunction theory.

First, recall that a family of distributions of the kind $a = P\left(\frac{1}{x}\right) + M\delta$ corresponds to the function $\frac{1}{x}$. Therefore, a family of equations with generalized coefficients is related to Eq. (12.8). The first refinement of the problem setting is as follows: we have to select one of these distributions a and to consider the corresponding formal equation.

For the considered distribution a, the primitive is the locally integrable function

$$g(x) = \ln|x| + M\Theta(x),$$

i.e., this distribution has the first order of singularity. Then the function

$$u(x) = -C\exp[-g(x) + g(-1)] = \begin{cases} C\frac{1}{x}, & x < 0, \\ -Ce^{-M}\frac{1}{x}, & x > 0 \end{cases}$$

is a formal solution of Eq. (12.8) in the space of distributions. This is a function of kind (12.9) such that $C_1 = -Ce^{-M}$ for it. Note that the most natural solution (namely, the solution with $C_1 = C$) is obtained in the following two cases: $a = P\left(\frac{1}{x}\right) \pm i\pi\delta$.

Is it possible to treat the constructed function u as a solution in the sense of the distribution theory or mnemofunctions theory? First, we note that, similarly to the case of the function $\frac{1}{x}$, a family of distributions corresponds to this function a. Therefore, to obtain a correct answer, one has to find which distribution U from this family corresponds to the constructed solution u and whether it is possible to say that the relation

$$U' + aU = 0 \tag{12.10}$$

is fulfilled (and in what sense it is fulfilled). In particular, since U' is a distribution, it follows that the relation can be fulfilled only in special cases where a, U is a pair of singular distributions such that the product is a distribution as well.

The exact sense to the above is provided by means of the following general approach: the generalized coefficient a is to be changed for its approximation by smooth functions a_ε, the corresponding equation

$$u'_\varepsilon(x) + a_\varepsilon(x)u_\varepsilon(x) = 0 \tag{12.11}$$

is to be considered, and the family of its solutions

$$u_\varepsilon(x) = -C\exp\left[\int_{-1}^{x} a_\varepsilon(s)ds\right]$$

is to be investigated.

If $u_\varepsilon \to U$, then the equation implies that $a_\varepsilon u_\varepsilon \to -U'$, i.e., the product of the given approximations for a and U is associated with the distribution.

The simplest way to investigate the convergence is to consider only approximations generated by analytic representations of distributions on the line. For the considered a, such approximations have the form

$$a_\varepsilon(x) = \lambda\frac{1}{x+i\varepsilon} + (1-\lambda)\frac{1}{x-i\varepsilon}.$$

On the line, the following analog of Theorem 11.1 is valid: the product aU is associated with a distribution if and only if

$$U_\varepsilon(x) = -C\Big[\lambda\frac{1}{x+i\varepsilon} - (1-\lambda)\frac{1}{x-i\varepsilon}\Big];$$

in this case, the following relation holds:

$$a_\varepsilon(x)u_\varepsilon(x) = -C\Big[\lambda^2\frac{1}{(x+i\varepsilon)^2} - (1-\lambda)^2\frac{1}{(x-i\varepsilon)^2}\Big].$$

Since

$$U'_\varepsilon(x) = -C[-\lambda\frac{1}{(x+i\varepsilon)^2} + (1-\lambda)\frac{1}{(x-i\varepsilon)^2}]$$

in this case, it follows that the relation $U'+aU=0$ is satisfied only under the assumption that $\lambda=\lambda^2$. For $\lambda=1$, we obtain that $a_\varepsilon(x) = \frac{1}{x+i\varepsilon}$, and this is the approximation of the distribution $P\Big(\frac{1}{x}\Big) - i\pi\delta$; if $\lambda=0$, then $a_\varepsilon(x) = \frac{1}{x-i\varepsilon} \to P\Big(\frac{1}{x}\Big) + i\pi\delta$.

Thus, in the space of rational distributions, solutions of Eq. (12.10) with the considered distributions a exist only for the two distributions $a = P\Big(\frac{1}{x}\Big) \pm i\pi\delta$; then $U = Ca$.

Example 12.4. Equations such that their coefficients undergo high frequency oscillations are covered by this research area as well. Ideas of the averaging theory for equations with high frequency terms are close to ideas of the mnemofunction theory. Usually, in the theory of equations with high frequency terms, differential equations with small parameters are considered, while the main results state that the corresponding solutions u_ε converge to a solution of the so-called *averaged differential equation.* In simple cases, the averaged equation is constructed via the original one by means of the direct averaging (see, e.g., [30], where the variable coefficients are changed for their mean values). Cases where the procedure to obtain the averaged equation is more complicated (see, e.g., [27, 36, 39], where the coefficients are changed for constants different from their mean values) are more interesting.

Consider the simplest equation from this class. Let a be a smooth periodic function with period τ. Then the mnemofunction $a_\varepsilon(x) = a\Big(\frac{x}{\varepsilon}\Big)$ oscillating with a high frequency converges in the sense of distributions to the constant

$$A = \frac{1}{\tau}\int_0^\tau a(s)ds,$$

which is the mean value of the function a. Therefore, the equation

$$u'_\varepsilon(x) - a\Big(\frac{x}{\varepsilon}\Big)u_\varepsilon(x) = f(x) \tag{12.12}$$

can be treated as one of equations in the mnemofunction space associated with the averaged equation

$$u'(x) - Au(x) = f(x)$$

with the constant coefficient A.

Impacts arising in this case are demonstrated on the example of the Cauchy problem for the homogeneous equation with the condition $u(0)=1$.

Introducing the notation

$$F_\varepsilon(x) = \int_0^x a\Big(\frac{s}{\varepsilon}\Big)ds,$$

we obtain that

$$V_\varepsilon(x) = \exp F_\varepsilon(x).$$

Thus, the problem is to describe the asymptotic behavior of this family of functions. The introduced function has the form $F_\varepsilon(x) = Ax + \varepsilon\omega\Big(\frac{x}{\varepsilon}\Big)$, where $\omega(x)$ is a periodic function such that its period is equal to τ and its mean is equal to zero. For example, if $a_\varepsilon(x) = 1+\cos\left(\frac{x}{\varepsilon}\right)$, then $F_\varepsilon(x) = x+\varepsilon\sin\left(\frac{x}{\varepsilon}\right)$. Therefore, the solution can be represented in the form

$$u_\varepsilon(x) = \exp[Ax]\exp\Big[\varepsilon\omega\Big(\frac{x}{\varepsilon}\Big)\Big] = \exp[Ax]\Big[1+\varepsilon\omega\Big(\frac{x}{\varepsilon}\Big)+\frac{1}{2}\Big[\varepsilon\omega\Big(\frac{x}{\varepsilon}\Big)\Big]^2+\dots\Big],$$

which yields the simplest case of assertions from the averaging theory for differential equations (see [30]): solutions $u_\varepsilon(x)$ converge to the function $\exp Ax$, which is the solution of the Cauchy problem for the averaged equation, uniformly on each finite segment.

The behavior of solutions is changed if the amplitude of high frequency oscillations increases as $\varepsilon \to 0$. For example, let

$$a_\varepsilon(x) = A + \frac{1}{\varepsilon^d}a_0\Big(\frac{x}{\varepsilon}\Big),$$

where a_0 is a periodic function such that its mean value is equal to zero. Then

$$F_\varepsilon(x) = Ax + \frac{1}{\varepsilon^{d-1}}\omega\Big(\frac{x}{\varepsilon}\Big),$$

where $\omega(x)$ is a periodic function such that its period is equal to τ and its mean value is equal to zero. In this case, solutions $u_\varepsilon(x)$ converge to $\exp Ax$ for $d < 1$ (as above), but the factor $\frac{1}{\varepsilon^{d-1}}$ increases and limit exists if $d > 1$. Hence, the most interesting case is the case where $d = 1$. Then

$$u_\varepsilon(x) = \exp Ax \exp\omega\Big(\frac{x}{\varepsilon}\Big).$$

This case has the following specific property: the family of periodic functions $\exp\omega\Big(\frac{x}{\varepsilon}\Big)$ converges to the mean value of the function $\exp\omega(x)$ in the sense of distributions and this mean value exceeds 1.

For example, if $a(x) = \frac{1}{\varepsilon}\cos x$, then

$$u_\varepsilon(x) = \exp\left[\sin\left(\frac{x}{\varepsilon}\right)\right]$$

and the mean value is equal to

$$\widetilde{A} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp[\sin(x)]dx = \frac{1}{2\pi}\sum_{k=0}^{\infty}\frac{1}{(2k)!}\int_{-\pi}^{\pi}[\sin(x)]^{2k}dx = \sum_{k=0}^{\infty}\frac{1}{(2k)!}\frac{(2k-1)!!}{(2k)!!} = 1+\frac{1}{4}+\frac{1}{64}+\dots$$

Thus, the following qualitatively new impact is observed in this example: the solution of the Cauchy problem for the averaged equation is $u(x) \equiv 1$, while solutions u_ε converge (in the space of distributions) to a constant $\widetilde{A} > 1$.

Another example is the Cauchy problem with the condition $u(0) = 0$ for the equation

$$u_\varepsilon'(x) - a_\varepsilon(x)u_\varepsilon(x) = 1,$$

where $a_\varepsilon(x) = 1 + \frac{1}{\varepsilon}\cos\left(\frac{x}{\varepsilon}\right)$. Here, the solution of the Cauchy problem for the equation

$$u'(x) - u(x) = 1,$$

where the coefficient is changed for its mean value, is equal to $u(x) = e^x - 1$. In this case, we have the relation

$$u_\varepsilon(x) = \exp\left[x+\sin\left(\frac{x}{\varepsilon}\right)\right]\int_0^x \exp\left[-x-\sin\left(\frac{t}{\varepsilon}\right)\right]dt.$$

One can show that this family converges to $\widetilde{A}^2(e^x-1)$ in the space of distributions. Since $\widetilde{A} > 1$, it follows that the constructed solution increases faster than the solution of the Cauchy problem for the formally averaged equation.

12.3. Point-interaction Schrödinger operators. Expressions of the kind $-\Delta + \sum_{y \in Y} a_y \delta_y$, where Δ is the Laplacian in $\mathbb{R}^d$, Y is a discrete (finite or denumerable) subset of $\mathbb{R}^d$, and δ_y is the Dirac function concentrated at the point y, play a special role among differential expressions with generalized coefficients. Such equations heuristically describe quantum mechanical systems generated by point source located at the points y, where the coefficients a_y are the so-called *coupling constants* characterizing the intensity of the corresponding sources. The historical review and a comprehensive bibliography can be found in [1].

We provide just several considerations clarifying the relation of the considered area to the theory of mnemofunctions. The first mathematically rigorous paper devoted to the construction of a self-adjoint operator in $L_2(\mathbb{R}^3)$ via the formal expression $-\Delta + a_y \delta_y$ is [11].

Let $N_y \subset L_2(\mathbb{R}^d)$ be the subspace consisting of smooth functions vanishing in a neighborhood of the point y. On N_y, each operator that can correspond to the formal expression $-\Delta + a_y \delta_y$ is to coincide with $-\Delta$. Therefore, it is to be a self-adjoint extension of the operator $-\Delta_y$ defined is the restriction of $-\Delta$ to N_y. For $d = 1$, the operator $-\Delta_y$ has a four-parameter family of self-adjoint extensions, for $d = 2$ and $d = 3$, the family of self-adjoint extensions is one-parametric, and only one self-adjoint extension exists provided that $d \geq 4$; this extension is the operator $-\Delta$.

To construct the self-adjoint operator corresponding to a formal expression following the general approach, one has to change the coefficient δ_y by an approximating family of smooth functions and, for the obtained family of well-defined operators L_ε, to find the limit in the sense of the resolvent convergence.

It turns out that if $d = 3$, then a limit (in the sense of the resolvent convergence) different from $-\Delta$ exists only under the assumptions that the coefficients are infinitely small and have the form $a_y = a_1 \varepsilon + a_2 \varepsilon^2$, where a_1 belongs to a discrete set depending on the method selected to approximate the delta function. A a result, the obtained limit operator is not determined via the formal expression: it depends on the method to approximate the delta function.

In the considered expression, the δ-function determines the multiplication operator acting on smooth functions as follows: $\delta_y u = u(y)\delta_y$. This is a first-rank operator and it is not defined in $L_2(\mathbb{R}^d)$. However, there are many methods to approximate it by families of finite-rank operators ion $L_2(\mathbb{R}^d)$, which yields a family of well-defined operators; its limit in the sense of the resolvent convergence is to be found. Here, the resolvents of approximating operators are found explicitly, which simplifies the computing of the limit of resolvents. Such an approach is explained, e.g., in [5].

For $d = 1$, the operator corresponding to the formal expression can be found from heuristic considerations. Usually, for the equation

$$u'' + a\delta u = f,$$

continuous solutions are such that the first derivative might have a jump $u'(+0) - u'(-0)$ at the origin. The differentiation in the space of distributions yields the relation

$$u'' + a\delta u = u''(x) + [u'(+0) - u'(-0)]\delta + a\delta u(0),$$

where $u''(x)$ is the classical derivative computed for $x \neq 0$. If the obtained solution belongs to the space $L_2(\mathbb{R})$, then terms containing δ eliminate each other, which yields the conjugation condition $u'(+0) - u'(-0) + au(0) = 0$. The same result can be obtained by means of the analysis of the resolvents of the approximating operators.

If an operator corresponding to the more complicated expression $-u'' + \sum_{y \in Y} a_y \delta_y u$ is constructed, then functions from the domain are to satisfy the conjugation conditions

$$u'(y+0) - u'(y-0) + a_y u(y) = 0 \quad \text{for each} \quad y \in Y$$

as well. In [26], the corresponding operators are investigated in detail; in particular, it is proved that the self-adjointness of the operator given by the conjugation conditions is not guaranteed.

13. Conclusions

The multiplication problem for generalized functions is considered. A general approach to its resolving is described. Various methods to embed distributions spaces into the algebra of mnemofunctions are analyzed. The method using the analytical representation of distributions is the most natural (from several viewpoints). Under such an embedding, the multiplication rule for rational distributions is given explicitly and it is possible to describe all cases where the product is a distribution.

There are relations between the multiplication problem for distributions with extensions of linear operators, nonstandard analysis, quantum mechanics, theory of equations with generalized coefficients, and theory of equations with small parameters.

REFERENCES

1. S. Albeverio, F. Gesztesy, R. Krohn Hoegh, and H. Holden, *Solvable Models in Quantum Mechanics* [Russian translation], Mir, Moscow (1991).
2. A. Antonevich, A. Burachewskij, and Ya. Radyno, "On closability of nonclosable operators," *Panamer. Math. J.*, **7**, No. 4, 37–51 (1997).
3. A. B. Antonevich and Ya. V. Radyno, "On general method of construction of distribution algebras," *Dokl. AN SSSR*, **43**, No. 3, 680–684 (1991).
4. A. B. Antonevich and Ya. V. Radyno, "10*th Conf. Problems and Methods in Mathematical Physics, Chemnitz, Germany, Sep.* 13–17, Teubner, Leipzig, pp. 9–14 (1994).
5. A. B. Antonevich and T. A. Romanchuk, *Equations with Delta-Like Coefficients: Finite-Dimensional Approximation Method* [in Russian], LAP Lambert, Saarbrucken (2012).
6. A. B. Antonevich and T. G. Shagova, "Embeddings of distributions into the algebra of mnemofunctions on a circle," *Probl. Fiz. Mat. i Tekhn.*, **37**, No. 4, 52–61 (2018).
7. A. B. Antonevich, T. G. Shagova, and E. V. Shkadinskaya, "Algebra of mnemofunctions on a circle," *Probl. Fiz. Mat. i Tekhn.*, **36**, No. 3, 55–62 (2018).
8. A. B. Antonevich and A. Shukur Ali, "On grows of an analytical function in a circle," *Dokl. NAN Belarusi*, **60**, No. 5, 41–45 (2016).
9. P. Antosik, Ya. Mikusinski, and R. Sikorsky, *Theory of Distributions. Sequential Approach* [Russian translation], Mir, Moscow (1976).
10. L. L. Baglini and P. Giordano, "The category of Colombeau algebras," *Monatsh. Math.*, **182**, 649–674 (2017).
11. F. A. Berezin and L. D. Faddeev, "A note on the Schrödinger equation with singular potential," *Dokl. AN SSSR*, **137**, No. 5, 1011–1014 (1961).
12. G. Bremerman, *Distributions, Complex Variables and Fourier Transforms* [Russian translation], Mir, Moscow (1965).
13. Ch. Christov and B. Damianov, "Asymptotic functions — a new class of generalized functions. I," *Bulgar. J. Phys.*, **6**, 543–556 (1978).
14. Ch. Christov and B. Damianov, "Asymptotic functions — a new class of generalized functions. II," *Bulgar. J. Phys.*, **1**, 3–23 (1979).
15. Ch. Christov and B. Damianov, "Asymptotic functions — a new class of generalized functions. III," *Bulgar. J. Phys.*, **3**, 245–256 (1979).
16. Ch. Christov and B. Damianov, "Asymptotic functions — a new class of generalized functions. IV," *Bulgar. J. Phys.*, **4**, 377–397 (1979).
17. J. F. Colombeau, "A multiplication of distributions," *J. Math. Anal. Appl.*, **94**, 96–115 (1983).
18. J. F. Colombeau, *New Generalized Functions and Multiplication of Distributions*, North-Holland, Amsterdam (1984).
19. V. G. Danilov, V. P. Maslov, and V. M. Shelkovich, "Algebras of singularities of singular solutions to first-order quasilinear strictly hyperbolic systems," *Teor. Mat. Fiz.*, **114**, No. 1, 3–55 (1998).
20. M. Davis, *Applied Nonstandard Analysis* [Russian translation], Mir, Moscow (1980).

21. Yu. V. Egorov, "To the theory of distributions," *Usp. Mat. Nauk*, **45**, No. 5, 3–40 (1990).
22. M. A. Evgrafov, *Asymptotic Estimates and Whole Functions* [in Russian], Gostekhizdat, Moscow (1957).
23. I. M. Gel'fand and G. E. Shilov, *Generalized Functions and Operations over Them* [in Russian], Fizmatlit, Moscow (1959).
24. M. Grosser, M. Kunzinger, M. Oberguggenberger, and R. Steinbauer, *Geometric Theory of Generalized Functions with Applications to General Relativity*, Kluwer Academic Publishers, Dordrecht (2001).
25. V. K. Ivanov, "Hyperdistributions and multiplication of Schwartz distributions," *Dokl. AN SSSR*, **204**, No. 5, 1045–1048 (1972).
26. A. S. Kostenko and M. M. Malamud, "1-D Schrödinger operator with local point interaction on a discrete set," *J. Differ. Equ.*, **249**, 253–304 (2010).
27. V. B. Levenshtam, *Differential Equations with Large High Frequency Terms (Averages and Asymptotics)* [in Russian], YUFU, Rostov-na-Donu (2010).
28. A. I. Markushevich, *Theory of Analytic Functions* [in Russian], GITTL, Moscow–Leningrad (1950).
29. I. V. Mel'nikova, V. A. Bovkun, and U. A. Alekseeva, "Solution of quasilinear stochastic problems in abstract Colombeau algebras," *Diff. Uravn.*, **53**, No. 12, 1653–1663 (2017).
30. Yu. A. Mitropol'skiy, *Averaging Method in Nonlinear Mechanics* [in Russian], Naukova dumka, Kiev (1971).
31. M. Nedeljkov, S. Pilipovic, and D. Scarpalezos, *Linear Theory of Colombeau's Generalized Functions*, Addison-Wesley Longman, Harlow (1998).
32. M. Oberguggenberger, *Multiplication of Distributions and Application to Partial Differential Equations*, Longman Higher Education, Harlow (1992).
33. E. E. Rosinger, *Generalized Solutions of Nonlinear Partial Differential Equations*, Mathematics Studies, North-Holland (1987).
34. L. Schwartz, *Théorie des Distributions: en 2 vol*, Hermann, Paris (1950).
35. L. Schwartz, "Sur l'impossibilité de la multiplication des distributions," *C. R. Acad. Sci. Paris*, **239**, 847–848 (1954).
36. T. A. Suslina, "Averaging of the Dirichlet problem for higher-order elliptic equations with periodic coefficients," *Algebra i Analiz*, **29**, No. 2, 139–192 (2017).
37. V. S. Vladimirov, *Generalized Functions in Mathematical Physics* [in Russian], Nauka, Moscow (1979).
38. S. T. Zavalishchin and A. N. Sesekin, *Impulse Processes. Models and Applications* [in Russian], Nauka, Moscow (1991).
39. V. V. Zhikov and S. E. Pastukhova, "On operator estimates in the averaging theory," *Usp. Mat. Nauk*, **71**, No. 3, 27–122 (2016).

A. B. Antonevich
Belarusian State University, Minsk, Belarus
E-mail: antonevich@bsu.by
T. G. Shagova
Belarusian State University, Minsk, Belarus
E-mail: tanya.shagova@gmail.com