Rigorous Interpretation of Engineering Formulae for the Convolution and the Fourier Transform Based on the Generalized Integral

ZELJKO JURIC[1], (Member, IEEE), AND HARUN SILJAK [ID][2], (Senior Member, IEEE)
[1]Faculty of Electrical Engineering, University of Sarajevo, 71000 Sarajevo, Bosnia and Herzegovina
[2]School of Engineering, Trinity College Dublin, Dublin 2, D02 PN40 Ireland

Corresponding author: Harun Siljak (harun.siljak@tcd.ie)

ABSTRACT This paper aims at providing a framework suitable for justification of classical convolution integral and Fourier transform in many cases not covered by the usual definition of integral used for signal theory applications. Generalized functions approach from functional analysis is used, simplifying it to be approachable for engineers while retaining the rigor. The generalized functions approach results in an elegant and applicable definition of integral known before in the mathematical literature which is readily applicable in signal theory, justifying formulae usually seen as dubious and criticised for lack of rigor. The study offers a rigorous, simple and understandable definition of integral for use in analog signal theory, helping the formalization of engineering education by means of rigor. Main advantage of this approach is retaining the classical notation used in signal theory as well as its straightforward justification of key formulae in signal theory resulting from convolution and/or Fourier transform.

INDEX TERMS Analog signal theory, functional analysis, Schwartz distribution, generalized integral, convolution, Fourier transform.

I. INTRODUCTION

The convolution integral and the Fourier transform are the most frequently used tools in the linear system analysis in time and frequency domain respectively [1]. Introduced early in engineering education, these tools are usually explained using notions from ordinary calculus. In such simplified interpretation, leaps of faith sometimes replace rigorous interpretations of intermediate steps in the process of integral evaluation.

Recognizing this inconsistency, in this paper we formulate a framework based on generalized functions approach from functional analysis, which we then use to justify the classical convolution and Fourier transform formulae. Once the framework is in place, the usual notation of signal theory stays in place, but with an expanded, rigorous meaning attached to it. The extension to generalized functions, as we will see, is necessary for many of the common properties of convolution and Fourier transform to hold.

As noted earlier, elementary calculus definitions of these tools can have limited applicability, However, in cases that

often arise in theoretical considerations, such definitions cannot be applied without certain special precautions. For example, the classical definition of the convolution of two functions $f(t)$ and $g(t)$ is usually given by the formula

$$f(t) * g(t) = \int_{-\infty}^{\infty} f(u)g(t-u)\mathrm{d}u \qquad (1)$$

The ordinary calculus proposes that the integral in (1) should be interpreted in Lebesgue sense, as it is probably the most consistent interpretation of integrals with infinite bounds. However, many important cases cannot be interpreted in this way [2]. For example, it is usually claimed that

$$\sin \omega_1 t * \sin \omega_2 t = 0 \text{ for } \omega_1 \neq \omega_2 \qquad (2)$$

This result is usually derived using the Fourier transform and the Convolution theorem, and it has an obvious physical interpretation (the steady state response of an oscillatory linear system with natural frequency ω_1 to a harmonic stimulus with frequency $\omega_2 \neq \omega_1$ is zero). However, this result cannot be derived from (1), because the integral in (1) diverges (in any sense known from the classic calculus) when $f(t) = \sin \omega_1 t$ and $g(t) = \sin \omega_2 t$. Moreover, the derivation based on the Fourier transform and the Convolution theorem as usually

presented in engineering literature is not rigorous, because the Fourier transform of $\sin \omega t$ involves singular objects like δ-function that cannot be treated rigorously in the framework of the ordinary calculus.

Another important case when (1) cannot be applied directly is one when either $f(t)$ or $g(t)$, or both, contain nonintegrable singularities. For example, the well known Hilbert transform which is defined by

$$\hat{f}(t) = \mathcal{H}\{f(t)\} = f(t) * \frac{1}{\pi t} \qquad (3)$$

requires that the integral in (1) should be interpreted in a non-usual way, as a Cauchy principal value of a divergent integral in respect to nonintegrable singularity:

$$\begin{aligned} \hat{f}(t) = \mathcal{H}\{f(t)\} &= \frac{1}{\pi} \text{ v.p.} \int_{-\infty}^{\infty} \frac{f(u)}{t-u} du \\ &= \frac{1}{\pi} \lim_{\varepsilon \to 0+} \int_{|t-u|>\varepsilon} \frac{f(u)}{t-u} du \qquad (4) \end{aligned}$$

Similar situation is encountered in more general versions of this convolution integral as well [3].

Another problematic (and very common in engineering practice) exemplary case are integrals of $f(t)t^{\alpha}$ w.r.t. t and $f(u)(t-u)^{\alpha}$ w.r.t. u with a real parameter $\alpha \le -1$. These have non-integrable singularities for $t = 0$ and $t = u$, respectively; the classical interpretation would demand their regularization via Hadamard's partie finie.

The classic formula for the convolution used with various singular objects is also inexhaustible source for various suspicious formulae usually seen in engineering literature. For example, if a fact that $\delta(t)$ is the unit element for the convolution is accepted, then relation $\delta(t) * \delta(t) = \delta(t)$ together with definition (1) gives

$$\int_{-\infty}^{\infty} \delta(u)\delta(t-u)du = \delta(t) \qquad (5)$$

After a simple formal change of variables, (5) can be written as

$$\int_{-\infty}^{\infty} \delta(t-a)\delta(t-b)dt = \delta(a-b) \qquad (6)$$

Both (5) and (6) cannot be interpreted classically, because the result of the standard integration cannot be a singular object like Dirac δ function. They cannot even be interpreted as generalized inner product of a Schwartzian distribution and an ordinary function, which is well known from the functional analysis, because (5) and (6) involves product of two singular objects. Additionally, such inner product also produces non-singular objects as the result. Formula (6) is sometimes used as an example of inconsistency of usual engineering formulae. However, in this paper, it will be shown that even (6) may be interpreted quite rigorously and consistently, using a proper definition of the integral.

The rigorousness of classic engineering definitions of the Fourier and the inverse Fourier transform which are given by formulae

$$F(\omega) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt \qquad (7)$$

$$f(t) = \mathcal{F}^{-1}\{F(\omega)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{i\omega t} d\omega \qquad (8)$$

is even more dubious. There are a lot of important cases in which integrals in (7) and (8) diverge and it is not uncommon to meet them in practice [4]. For example, almost all textbooks dedicated to the linear systems theory claim that $\mathcal{F}\{1\} = 2\pi\delta(\omega)$ and $\mathcal{F}\{\text{sgn } t\} = 2/i\omega$. Indeed, it is quite easy to derive these relations indirectly. However, if someone tries to derive these relations directly from (4), it will be possible only if it is possible to somehow justify the following relations:

$$\int_{0}^{\infty} \cos \omega t dt = \pi\delta(\omega) \qquad (9)$$

$$\int_{0}^{\infty} \sin \omega t dt = \frac{1}{\omega} \qquad (10)$$

Of course, both integrals in (9) and (10) are divergent in any usual sense. Moreover, there is no chance that any classic interpretation of the integration can produce a singular object like $\delta(\omega)$ from a regular function like $\cos \omega t$. That is why various different methods were introduced to justify steps involving divergent integrals in the process of Fourier transform determination, but they again rely on complex mathematical apparatus [5].

In this introduction, some basic problems that arise from the usual interpretation of the definitions of the convolution and the Fourier transform are presented. A lot of other examples may be found in [6]–[8]. Before presenting the key points of this paper, some basic facts from a highly advanced branch of mathematics known as functional analysis related to these problems will be recalled.

II. IMPORTANT FACTS FROM THE THEORY OF DISTRIBUTIONS

The functional analysis is considered as a very abstract and quite difficult branch of advanced mathematics, which can be used for rigorous treatment of many objects that arise in the linear systems theory. Theory of distributions, as a sub-branch of the functional analysis, may be especially useful in this field. Unfortunately, the functional analysis has its own terminology, language and abstract operator based notation, which is completely different from the usual engineering notation, and even quite confusing for any non expert in functional analysis [9], [10]. Therefore, any results based on such notation are completely incomprehensible for an average engineer. However, in this paper, we will show that many engineering formulae which are usually considered "suspicious" and non-rigorous from the mathematical point of view, may be interpreted completely rigorously without introduction of any new notation, assuming different interpretation of integrals that arise in these formulae. Before this,

we need to present some basic facts from the functional analysis about the distributions, the convolution and the Fourier transform. This presentation is maybe somewhat simplified from the pure mathematical point of view, to be comprehensible to non-experts in functional analysis, although it is still perfectly correct and rigorous.

It is known from functional analysis that the Schwartzian distribution T on $\mathbb{R}^n$ is a functional (operator) that assigns a numeric value $T[\phi]$ to each infinitely smooth function $\phi(x)$ which is identically equal to zero outside of a compact subset of $\mathbb{R}^n$ (such functions are called test functions) [6], [9]–[11]. For example, well-known Dirac δ function is in fact a distribution on $\mathbb{R}$ that assigns to each test function $\phi(t)$ on $\mathbb{R}$ a value $\delta[\phi] = \phi(0)$. If we want to be as rigorous as possible, it is additionally requested that T as an operator must be continuous with respect to appropriately constructed topology. More precisely, here the continuity of T means that if a sequence of test functions $\phi_n(x)$ is identically equal to zero out of a fixed compact subset of $\mathbb{R}^n$, and if $\phi_n(x)$ together with its partial derivatives of any order converges uniformly to $\phi(x)$ and according partial derivatives of $\phi(x)$, then $T[\phi_n]$ also converges to $T[\phi]$ [6], [9], [10]. The space of test functions on $\mathbb{R}^n$ is called $\mathcal{D}(\mathbb{R}^n)$, and the space of associated distributions on $\mathbb{R}^n$ (its topological dual space) is called $\mathcal{D}^*(\mathbb{R}^n)$.

Any ordinary locally integrable function $f(x)$, $x \in \mathbb{R}^n$ may be regarded as a special case of distribution (such distributions are called regular distributions), using the ad hoc definition

$$f[\phi] = \langle f(x), \phi(x) \rangle = \int_{\mathbb{R}^n} f(x)\phi(x)\mathrm{d}x \qquad (11)$$

where $\langle f(x), \phi(x) \rangle$ is the usual inner (scalar) product of two functions. Accordingly, it is possible to define a generalized inner product using the relation

$$\langle T(x), \phi(x) \rangle = T[\phi] \qquad (12)$$

even in singular cases, i.e. in cases when T cannot be interpreted as an ordinary locally integrable function. For example, we can take that $\langle \delta(t), \phi(t) \rangle = \phi(0)$, which is, in fact, nothing else but a more rigorous form of the well known formula

$$\int_{-\infty}^{\infty} \delta(t)\phi(t)\mathrm{d}t = \phi(0) \qquad (13)$$

although we will see that under the correct interpretation of the integral, (13) can be quite rigorous as well. Note that $T(x)$ must be regarded as a purely symbolic object, not a function of the argument x from $\mathbb{R}^n$ (regardless of the notation). That is why the letter x, which arises in the formal expression $T(x)$, is called a formal argument of T.

Another important concept that should be introduced here is the outer or tensor product of distributions. Assuming that T_1 and T_2 are distributions on $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, their tensor product $T_1 T_2$ is the distribution on $\mathbb{R}^{n+m}$ defined using the relation

$$\langle T_1 T_2(x, y), \phi(x, y) \rangle = \langle T_1(x), \langle T_2(y), \phi(x, y) \rangle \rangle \qquad (14)$$

where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. Here, x in the inner expression should be interpreted as a fixed parameter, so the result of the inner expression $\langle T_2(y), \phi(x, y) \rangle$ is, in fact, a function of x. The resulting tensor product $T_1 T_2(x, y)$ is usually formally written as $T_1(x)T_2(y)$, like an ordinary product. Note that in this notation, distributions T_1 and T_2 are written using different names of formal arguments.

For the correct interpretation of formulae that arise in signal theory and other branches of applied mathematics, we must define the meaning of the expression $T(Ax + b)$ where T is a distribution on $\mathbb{R}^n$, A is a nonsingular $n \times n$ matrix, and b is an n-dimensional vector. By the definition, $T(Ax + b)$ is again a distribution on $\mathbb{R}^n$, whose action on a test function $\phi(x)$ is given using the formula

$$\langle T(Ax + b), \phi(x) \rangle = \frac{1}{|\det A|} \langle T(x), \phi(A^{-1}(x - b)) \rangle \qquad (15)$$

This formula (so-called linear change of the formal arguments) is inspired by the fact that it reduces to the ordinary formula for a linear change of variables $x \rightarrow Ax + b$ in the integral whenever T is reducible to an ordinary locally integrable function on $\mathbb{R}^n$.

Now, we need to say something about the correct interpretation of the expressions like $\delta(t-u)$. It is important to say that this expression has two completely different interpretations depending on whether only t is regarded as a variable and u is regarded as a fixed parameter, or both t and u are regarded as variables. When only t is regarded as a variable, $\delta(t - u)$ should be interpreted as a distribution on $\mathbb{R}$ which is derived from the distribution $\delta(t)$ using the linear change $t \rightarrow t - u$. More specifically, (15) for such case gives

$$\langle \delta(t - u, \phi(t)) \rangle = \phi(u) \qquad (16)$$

From the other side, when both t and u are regarded as variables, $\delta(t-u)$ must be somehow interpreted as the distribution on $\mathbb{R}^2$, i. e. the distribution that acts on the test functions $\phi(t, u)$ on $\mathbb{R}^2$. There exist many different ways how such distribution may be introduced [5], but fortunately, all of them lead to the same expression. Here will be presented a way based only on the tensor product and the linear change of the formal arguments. We will start from the distribution $\delta(t)1(u)$ on $\mathbb{R}^2$ that is the tensor product of the distribution $\delta(t)$ and the distribution $1(u)$, which is nothing more than distributional interpretation of the constant function $1(u) \equiv 1$. More concretely, we have

$$\begin{aligned}\langle \delta(t)1(u), \phi(t, u) \rangle &= \langle \delta(t), \langle 1(u), \phi(t, u) \rangle \rangle \\ &= \left\langle \delta(t), \int_{-\infty}^{\infty} \phi(t, u)\mathrm{d}u \right\rangle = \int_{-\infty}^{\infty} \phi(0, u)\mathrm{d}u \end{aligned}$$

$$(17)$$

The distribution $\delta(t)1(u)$ may be regarded as the distribution $\delta(t)$ lifted up from $\mathbb{R}$ to $\mathbb{R}^2$. Now, to define $\delta(t - u)$ as a distribution on $\mathbb{R}^2$, we should introduce the linear change $(t, u) \rightarrow (t, t - u)$ into $\delta(t)1(u)$. The resulting distribution will be, in fact, $\delta(t - u)1(u)$, but it is equivalent to $\delta(t - u)$,

as $1(u)$ is practically a constant in the world of distributions. Using (15), the described procedure gives

$$
\begin{aligned}
\langle \delta(t-u), \phi(t,u) \rangle &= \langle \delta(t)1(u), , \phi(u, u-t) \rangle \\
&= \langle \delta(t), \langle 1(u), \phi(u, u-t) \rangle \rangle \\
&= \left\langle \delta(t), \int_{-\infty}^{\infty} \phi(u, u-t)\mathrm{d}u \right\rangle \\
&= \int_{-\infty}^{\infty} \phi(u,u)\mathrm{d}u = \frac{1}{\sqrt{2}} \int_{u=v} \phi(u,v)\mathrm{d}s
\end{aligned}
$$
(18)

The last result is expressed using the line integral, to emphasize the physical nature of the distribution $\delta(t-u)$. Namely, when u is a constant parameter, $\delta(t-u)$ represents a point mass concentrated at the point $t = u$ on the t line. On the contrary, when both t and u are variables, $\delta(t-u)$ represents the line mass distributed uniformly along the line $t = u$ in the $t - u$ plane.

III. TREATMENT OF THE CONVOLUTION AND THE FOURIER TRANSFORM IN THE FUNCTIONAL ANALYSIS

The functional analysis may be also used for rigorous definition of the convolution. However, it is extremely hard to give a good definition of the convolution, and there exists no definition of it that is applicable in all cases. That is why the literature is full of partial or even incorrect definitions. Two most general definitions are the definition given by L. Schwartz [11] and the definition given by V. S. Vladimirov [6]. Both definitions rely on the tensor product of distributions. The definition given by Schwartz is as follows:

Definition 1 (Convolution by Schwartz): The convolution of distributions T_1 and T_2 on $\mathbb{R}^n$ is a distribution $T_1 * T_2$ on $\mathbb{R}^n$ given by formula

$$
T_1 * T_2[\phi] = \langle T_1(x)T_2(y), \phi(x+y) \rangle \tag{19}
$$

assuming that the outer product $T_1 T_2$ admits a continuous extension onto a space of functions of type $\phi(x+y)$ (which are not test functions on $\mathbb{R}^{2n}$, regardless of the fact that $\phi(x)$ is a test function on $\mathbb{R}^n$). If such continuous extension does not exist, the convolution $T_1 * T_2$ does not exist either.
A very similar but not identical definition, as it seems after a superficial observation, is:

$$
T_1 * T_2[\phi] = \langle T_1(x), \langle T_2(y), \phi(x+y) \rangle \rangle \tag{20}
$$

It is given by W. Rudin [9] and by A. Papoulis [8], although Rudin wrote this definition in more abstract operator-based form, and Papoulis wrote it in a less-formal integral like form. However, this definition, although maybe somewhat simpler, is even more restrictive than the definition given by Schwartz, as it requires that the result of the inner product $\langle T_2(y), \phi(x+y) \rangle$ must be again a test function.

The main problem in the definition given by Schwartz is the assumption of the existence of a continuous extension of the outer product $T_1 T_2$ to a broader space than the space of

test functions. Another definition is given by Vladimirov [6], which does not rely on such assumption. Vladimirov uses so called unit sequences. A sequence $\eta_k(x)$ of test functions from $\mathcal{D}(\mathbb{R}^n)$ is a unit sequence on $\mathbb{R}^n$ if for any compact set $K \subset \mathbb{R}^n$ there exists a number $N = N(K)$ such that $\eta_k(x) = 1$ for $x \in K$, $k \geq N$, and if the functions $\eta_k(x)$ and all their partial derivatives of any order are uniformly bounded (i.e. bounded with constants which do not depend of x). Now, it is possible to present the definition of convolution in according to Vladimirov.

Definition 2 (Convolution by Vladimirov): The convolution of distributions T_1 and T_2 on $\mathbb{R}^n$ is a distribution $T_1 * T_2$ on $\mathbb{R}^n$ given by formula

$$
T_1 * T_2[\phi] = \lim_{k \to \infty} \langle T_1(x)T_2(y), \phi(x+y)\eta_k(x,y) \rangle \tag{21}
$$

assuming that the limit on the right side exists for any unit sequence $\eta_k(x, y)$ on $\mathbb{R}^{2n}$ and that this limit does not depend on a particular choice of sequence $\eta_k(x, y)$. If these assumptions are not satisfied, then the convolution $T_1 * T_2$ does not exist.

It is proven that definitions of convolution given by Schwartz and by Vladimirov are equivalent. Note that both of these definitions of the convolution are extremely indirect, especially if it is necessary to interpret the convolution as an ordinary function when such interpretation is possible. In addition, these definitions are complicated, and quite different than the usual definition (1).

Now, let us check how the Fourier transform is defined in the functional analysis. We will limit our considerations to the one-dimensional case, i.e. the functions and distributions on $\mathbb{R}$. The usual definitions (7) and (8) work only for functions which are absolutely integrable on $(-\infty, \infty)$, so many Fourier transform pairs cannot be deduced from them. Some such examples are given in the introductory section of the paper. In fact, it is not possible to rigorously apply (7) even to derive a very common Fourier Transform pair $\mathcal{F}\{\operatorname{sinc} t\} = \pi[u(\omega + 1) - u(\omega - 1)]$ where $\operatorname{sinc} t = (\sin t)/t$ for $t \neq 0$ and $\operatorname{sinc} 0 = 0$. Namely, $f(t) = \operatorname{sinc} t$ is not absolutely integrable on $(-\infty, \infty)$ so the integral in (7) does not converge properly (it exists only as a Riemann improper integral). Therefore, functional analysis introduces various more general definitions of the Fourier transform that may be applied to more general class of functions, even for purely symbolic non-function objects like Schwartzian distributions. Unfortunately, as the generality of such definitions increases, they become more and more indirect and unconstructive in nature. One of the most general definitions, which is applicable to all ordinary functions and Schwartzian distributions under some assumptions which will be stated later, defines the Fourier Transform $\mathcal{F}\{T(t)\}$ of the distribution $T(t)$ on $\mathbb{R}$ as another distribution $\hat{T}(\omega)$ which acts as a functional that assigns to each test function $\phi(\omega)$ the value $T[\mathcal{F}\{\phi(\omega)\}]$, where $\mathcal{F}\{\phi(\omega)\}$ is defined in the usual way using (7). In other words,

$$
\hat{T}[\phi] = \mathcal{F}\{T\}[\phi] = T[\mathcal{F}\{\phi\}] \tag{22}
$$

Alternatively, using the notation which uses the generalized inner product, this definition may be expressed using the relation

$$\langle \hat{T}(\omega), \phi(\omega) \rangle = \langle \mathcal{F}\{T(t)\}(\omega), \phi(\omega) \rangle = \langle f(t), \mathcal{F}\{\phi(\omega)\}(t) \rangle \tag{23}$$

The inverse Fourier transformation is defined analogously, using the relations

$$T[\phi] = \mathcal{F}^{-1}\{\hat{T}\}[\phi] = T[\mathcal{F}^{-1}\{\phi\}] \tag{24}$$

or

$$\langle T(t), \phi(t) \rangle = \langle \mathcal{F}^{-1}\{\hat{T}(\omega)\}(t), \phi(t) \rangle$$
$$= \langle \hat{T}(\omega), \mathcal{F}^{-1}\{\phi(t)\}(\omega) \rangle \tag{25}$$

In these definitions, usual definitions of the Fourier and the inverse Fourier transform are applied on test functions ϕ. However, the Fourier transform $\mathcal{F}\{\phi\}$ of any test function is not again a test function. Therefore, definitions (22) and (24) must be restricted to such distributions which admit a continuous extension to some broader space than the space of test functions. It is well known that these definitions are valid for so-called tempered distributions (i.e. distributions with a limited rate of growth), which admit an extension to the space of rapidly decreasing functions, which are similar to the test functions, but which do not necessary vanish out of a compact subset of $\mathbb{R}$. Instead, they must tend towards 0 together with all their derivatives more rapidly than a reciprocal of any polynomial when $|t| \to \infty$ (for example, as $e^{-|t|}$). The space of such functions is called $\mathcal{S}(\mathbb{R})$, and the space of tempered distributions associated to it is called $\mathcal{S}^*(\mathbb{R})$. In general, it is possible to define the Fourier transform even for all distributions using more complicated formulae. However, such generalization requires an introduction of more general objects than the Schwartzian distributions, so called ultradistributions [12], so this possibility will not be discussed here.

Definitions based on (22) and (24) may look rather simple, but this simplicity is just an illusion. In fact, they are extremely indirect due to their strong dependence of the operator notation. Moreover, it is almost impossible to apply these definitions for actual calculation of the Fourier transform, i.e. to express the Fourier transform of some function using other common functions (ordinary or Schwartzian distributions), except in very simple cases. In addition, this definition is quite different from the common definition of the Fourier transform used in analog signal theory.

In this section, the basic facts from the functional analysis about the convolution and the Fourier transform have presented briefly. It is obvious that presented formulae are quite different than formulae that are usually used in the linear systems theory, and probably quite incomprehensible for most users of this theory. In the next section, it will be shown that usual definitions of the convolution and the Fourier transform known from the engineering literature may be regarded as perfectly valid ones if the interpretation of integrals in these formulas is changed from the usual interpretation.

IV. GENERALIZED INTEGRAL AND ITS RELATION TO THE CONVOLUTION AND THE FOURIER TRANSFORM

It is not hard to recognize that the main problem with the engineering formulae arises from the various non-convergent integrals, or integrals which are applied to the nonintegrable objects. Careful examination shows that it is very often possible to find an interpretation of a divergent integral which depends of a parameter (like integrals in formulae (9) and (10)) as a distribution in which this parameter is just a formal argument. However, it is not easy to find an universal framework for such treatment. In this paper, we will show that the following definition gives a quite universal and rigorous way for such interpretation in many different contexts:

Definition 3 (Generalized integral): The generalized integral of a distribution $T(x, y)$ on $\mathbb{R}^{n+m}$ (where both x and y are formal arguments) in respect to the formal argument x on $\mathbb{R}^n$ is a distribution T_0 on $\mathbb{R}^m$ given by the formula

$$T_0[\phi] = \lim_{k \to \infty} \langle T(x, y), \phi(y)\eta_k(x, y) \rangle \tag{26}$$

assuming that the limit on the right side exists for any unit sequence $\eta_k(x, y)$ on $\mathbb{R}^{n+m}$ and that this limit does not depend on a particular choice of the sequence $\eta_k(x, y)$. In this case, it is possible to simply write

$$\int_{\mathbb{R}^n} T(x, y)\mathrm{d}x = T_0(y) \tag{27}$$

In other words, we have

$$\left\langle \int_{\mathbb{R}^n} T(x, y)\mathrm{d}x, \phi(y) \right\rangle = \lim_{k \to \infty} \langle T(x, y), \phi(y)\eta_k(x, y) \rangle \tag{28}$$

If such assumptions are not valid, then the generalized integral does not exist.

Sometimes, the result of the generalized integration may be expressed as an ordinary function, i. e. if exist an ordinary function such that $\langle f(y), \phi(y) \rangle = T_0(y)$ (in such cases, $f(y)$ is the result of the integration). Moreover, this definition is applicable even when $m = 0$. In such cases, the result of the integration is an ordinary number, and we can write

$$\int_{\mathbb{R}^n} T(x)\mathrm{d}x = \lim_{k \to \infty} \langle T(x), \eta_k(x) \rangle \tag{29}$$

assuming that the limit on the right side exists for any unit sequence $\eta_k(x)$ on $\mathbb{R}^n$ and that this limit does not depend on a particular choice of the sequence $\eta_k(x)$.

The concept of the generalized integral, as given above, is introduced by Vladimirov [6]. It is probably introduced to allow more consistent treatment of the convolution, as it will be shown later that it is quite related to the convolution. Although it is not a new concept, it is quite uncommon, even in the mathematical literature. However, it will be shown in the paper that it is a very useful concept for the rigorous interpretation of the various formulae that arise in the engineering literature.

It is proven in [6] that the definition of the generalized integral given above reduces to the classic definition of Lebesgue integral whenever $T(x, y)$ is an ordinary integrable function in

respect to variable x (y is interpreted as a parameter), so that this definition is really a generalization of the classic concept of the integral. However, this definition gives a meaning to many integrals which are divergent in usual Lebesgue sense, or even in a Riemann improper or Cauchy principal value sense. The next theorem states that this concept gives the rigorous interpretation of (9) and (10)

Theorem 4: Both formulae (9) and (10) are valid if the integrals that arise in them are interpreted as generalized integrals of $[\cos(\omega t)]U(t)$ and $[\sin(\omega t)]U(t)$ in respect to t on $qmathbbR$, where $U(t)$ is the Heaviside step function. In other words, we have

$$\int_{\mathbb{R}} \cos \omega t U(t) \mathrm{d}t = \pi \delta(\omega) \tag{30}$$

$$\int_{\mathbb{R}} \sin \omega t U(t) \mathrm{d}t = \frac{1}{\omega} \tag{31}$$

where both integrals are generalized ones. Additionally, $1/\omega$ in the second integral should be interpreted as a distribution using Cauchy principal value, which is a well known way of its interpretation in the theory of distributions. More concretely,

$$\left\langle \frac{1}{\omega}, \phi(\omega) \right\rangle = \text{v.p.} \int_{-\infty}^{\infty} \frac{\phi(\omega)}{\omega} \mathrm{d}\omega = \lim_{\varepsilon \to 0+} \int_{|t| > \varepsilon} \frac{\phi(\omega)}{\omega} \mathrm{d}\omega \tag{32}$$

The proof of this theorem is given in the appendix at the end of the paper. It is interesting that the proof uses the Fourier transform. This can give a hint that the generalized integral may somehow be used for the rigorous treatment of the Fourier transform as well. More about this conclusion will be presented later in the paper.

Returning to another integration issue raised in the introduction, we can offer a rigorous interpretation of integrating $f(t)t^\alpha$ w.r.t. t and $f(u)(t-u)^\alpha$ w.r.t. u. These should be observed as generalised integrals, where, in case of a bilateral integral from $-\infty$ to $+\infty$, t^α and $(t-u)^\alpha$ are interpreted as Pf t^α and Pf $(t-u)^\alpha$, respectively; here, Pf (pseudofunction) denotes the distributional extension of functions non-integrable in Lebesgue sense. If the integration is unilateral, from 0 to $+\infty$, the relevant pseudofunctions are Pf t_+^α and Pf $(t-u)_+^\alpha$, respectively. Alternatively, we can write Pf $t^\alpha U(t)$ and Pf $(t-u)^\alpha U(t-u)$, where $U(t)$ is the Heaviside step function. The former notation is more common in functional analysis, while the latter is closer to the engineering community.

The next interesting fact is that the generalized integral can give rigorous interpretation of the generalized inner product written in the integral-like form, as stated in the following theorem:

Theorem 5: Each distribution T from $\mathbb{R}^n$ allows the integral representation

$$T[\phi] = \langle T(x), \phi(x) \rangle = \int_{\mathbb{R}^n} T(x)\phi(x)\mathrm{d}x \tag{33}$$

assuming that $T(x)\phi(x)$ is interpreted as the product of the Schwartzian distribution $T(x)$ with the smooth function $\phi(x)$

using the well-known rule from the theory of distributions

$$\langle T(x)\phi(x), \psi(x) \rangle = \langle T(x), \phi(x)\psi(x) \rangle \tag{34}$$

where $\psi(x)$ is an another test function, and assuming that the integral in the above formula is interpreted as the generalized integral of the product $T(x)\phi(x)$ on $\mathbb{R}^n$ in respect to x.

The proof of this theorem is given in the appendix at the end of the paper. This theorem is enough to give rigorous interpretation of the engineering formulae like

$$\int_{-\infty}^{\infty} f(t)\delta(t-a)\mathrm{d}t = f(a) \tag{35}$$

assuming that $f(t)$ is an ordinary function and that a is a fixed parameter (i.e. a constant). The key point is the interpretation of the integral in the above formula as the generalized integral. Note that even Schwartz [11] says that (35) is wrong or, at least, purely formal. Such belief is present even in much modern approaches to the theory of generalized functions, as in [12] or [13]. Now, we can see that such belief is not quite correct. From the other side, in the linear systems theory, it is usually assumed that (35) is valid even without the assumptions stated above, e.g. when a is a variable and/or when $f(t)$ is not an ordinary function but a distribution. In many cases, the generalized integral can give rigorous treatment of such formula as well. For example, the following theorem states that even (6) is valid under very relaxed conditions:

Theorem 6: Taking the concept of the generalized integral into the consideration, the formula (6) may be regarded as completely valid assuming that at least one of the quantities a and b is not treated as a constant. If both quantities a and b are variables, the expression $\delta(t-a)\delta(t-b)$ should be interpreted as a distribution on $\mathbb{R}^3$ that acts on the test functions $\phi(t, a, b)$ on $\mathbb{R}^3$, which is derived from the distribution $\delta(u)\delta(v)1(w)$ using the linear change $(u, v, w) \to (t-a, t-b, t)$, and the result of the integration $\delta(a-b)$ is the distribution on $\mathbb{R}^2$ that acts on the test functions $\phi(a, b)$ in a sense described earlier. If one of the quantities a or b is constant, say b, the expression $\delta(t-a)\delta(t-b)$ should be interpreted as a distribution on $\mathbb{R}^2$ that acts on the test functions $\phi(t, a)$ on $\mathbb{R}^2$, which is derived from the distribution $\delta(u)\delta(v)$ using the linear change $(u, v) \to (t-a, t-b)$, and the result of the integration $\delta(a-b)$ is a distribution on $\mathbb{R}$ that acts on test functions $\phi(a)$. Moreover, (6) is valid even if both a and b are constants, assuming that $a \neq b$.

The proof of this theorem is given in the appendix at the end of the paper. Note that the generalized integral introduced in this paper cannot validate (6) assuming that both a and b are constants and $a = b$. To see why, let assume that, for example, $a = b = 0$. Then, we have $\delta(t)^2$ under the integral in (6), and the result of the integration is $\delta(0)$. However, it is well known that both $\delta(t)^2$ and $\delta(0)$ have no any sensible interpretation in the theory of distributions [6], [12], [14]. On the other hand, they can be consistently interpreted in the framework of Colombeau generalized functions [12] or in the framework of the nonstandard analysis [15]. Therefore, it is very likely that using the appropriate concept of the integral, (6) can be

rigorously interpreted in these frameworks even when both a and b are constants and $a = b$. In fact, to validate (6) when both a and b are constants and $a = b$, it is enough to validate $\delta(t)^2 = \delta(0)\delta(t)$, whatever it means.

We already said that the presented definition of the generalized integral is closely related to the convolution. Indeed, both the definition of the convolution by Vladimirov and the definition of the generalized integral are based on the unit sequences, and both of them are defined using quite similar expressions. Additionally, it is proven in [6] that the generalized integral $T_0(y)$ of $T(x, y)$ with respect to the formal argument x on $\mathbb{R}^n$ exists if and only if the convolution $T(x, y) * [\delta(y)1(x)]$ exists, and that in such case, this convolution is just equal to $T_0(y)1(x)$. What is not proved there is that the usual engineering formula (1) for the convolution becomes absolutely correct under the correct interpretation that includes the concept of the generalized integral, as stated in the following theorem:

Theorem 7: Usual definition of the convolution, as presented in many engineering textbooks, is absolutely correct and rigorous even in singular cases (more precise, in all cases when definitions given by Schwartz or by Vladimirov are valid), assuming that the integral in this definition is interpreted as the generalized integral defined in this paper. In other words, the formula

$$T_1 * T_2(x) = \int_{\mathbb{R}^n} T_1(y)T_2(x - y)\mathrm{d}y \qquad (36)$$

is valid even if the ordinary integral does not exist, or even when T_1 and/or T_2 are not ordinary functions but Schwartzian distributions, if the integral given above is interpreted as a generalized integral of the distribution $T_1(z)T_2(x - z)$ with respect to the formal argument z on $\mathbb{R}^m$. The expression $T_1(z)T_2(x-z)$ should be interpreted as the distribution derived from the tensor product $T_1(z)T_2(x)$ using the linear change $(z, x) \rightarrow (z, x - z)$.

The proof of this theorem is given in the appendix at the end of the paper. It may be not particularly surprising, due to the close relationship between the generalized integral and the convolution. In other words, the given definition of the generalized integral might be treated just as an ad hoc definition introduced only to justify the validity of the familiar definition of the convolution. However, the fact that introduced generalized integral may be used for rigorous interpretation of many other integral-based formulae from the linear systems theory, even those that are not related to convolution in any way, is much more surprising. For example, the following theorem shows that the concept of the generalized integral also provides correctness and rigorousness to the widely used engineer formulas for the direct and inverse Fourier transform:

Theorem 8: Usual definitions of the Fourier and inverse Fourier transform (7) and (8), as presented in many engineering books, are absolutely correct and rigorous even in singular cases (more precise, for all tempered distributions), assuming that the integrals in these definitions are interpreted as the

generalized integrals defined in this paper. In other words, formulae

$$\hat{T}(\omega) = \mathcal{F}\{T(t)\} = \int_{\mathbb{R}} T(t)e^{-i\omega t}\mathrm{d}t \qquad (37)$$

$$T(\omega) = \mathcal{F}^{-1}\{\hat{T}(\omega)\} = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{T}(\omega)e^{i\omega t}\mathrm{d}\omega \qquad (38)$$

are valid even if ordinary integrals do not exist, or even when T and/or $\hat{T}$ are not ordinary functions, assuming that the integrals in these formulae are interpreted as generalized integrals of the distributions $T(t, \omega) = T(t)\,e^{-i\omega t}$ and $\hat{T}(\omega, t) = \hat{T}(\omega)e^{i\omega t}$ in respect to the formal arguments t and ω on $\mathbb{R}$, respectively.

The proof of this theorem is given in the appendix at the end of the paper.

V. CONCLUSION
In this paper, some problems which arise with the usual interpretation of the convolution and the Fourier transform are presented. After the basic introduction, various concepts from the functional analysis that deal with these problems are explained. Given that the notation and the terminology of the functional analysis is quite tedious for an average user of the linear systems theory, we have made an effort to provide an interpretation of the functional analysis concepts in a manner closer to the engineering community.

We have been motivated by the desire to retain as much as possible from the existing notation in signal theory, while providing the missing rigor for it. The fact that definitions like (1), (7) and (8) are not correct under the usual interpretation is often exaggerated in many books dedicated to the pure mathematics, where such definitions are a priori rejected as incorrect ones. However, these definitions should not be treated as incorrect, because they can become correct under the appropriate interpretation, as shown in this paper.

The main advantage of the method for correct interpretation of the convolution and the Fourier transform presented in this paper is the localization of the troubles. In fact, the main source of troubles arises from the integrals that are not convergent in the usual sense, and this paper presents a method for their correct interpretation even in singular cases. This means that we can continue using widely accepted and familiar formulae. Such approach is much more natural for the usage in the linear systems theory than adapting to the completely different definitions and notations taken from the functional analysis, which are often quite abstract, tedious and incomprehensible. In other words, the conclusion of this paper is that it is better to redefine only the interpretation of some fundamental concepts from the classical calculus (like the concept of integration) than to redefine the interpretation of nearly all concepts from the linear systems theory (as performed in the functional analysis).

Nevertheless, the definition of the generalized integral itself is not simple at all. It is abstract, indirect, and may appear confusing. However, this definition may be reduced to much simpler, more obvious and more intuitive concepts (like

cancellation of rapidly oscillating components in divergent integrals) in many cases that arise in the linear systems theory. In some special cases, such integral may be calculated easily using much simpler concepts, for example using weak limits, as shown in [16]. While elaboration of this goes well beyond the scope of this paper, it may be useful for both users and teachers of the linear system theory to know that the interpretation of the integral that ensures perfect validity of the familiar formulae exists, even if they do not know what this interpretation means exactly.

APPENDIX—PROOFS OF STATED THEOREMS

This appendix contains the proofs of all stated theorems. Note that this appendix requires a bit deeper knowledge of the functional analysis than the rest of the paper.

Proof of Theorem 4: Let us take $\phi(\omega) \in \mathcal{D}(\mathbb{R})$ and let $\eta_k(t, \omega) \in \mathcal{D}(\mathbb{R}^2)$ be an arbitrary unit sequence on $\mathbb{R}^2$. From the definition of the unit sequence and from the compactness of the support of $\phi(\omega)$, it follows that there exists a unit sequence $\mu_k(t) \in \mathcal{D}(\mathbb{R})$ on $\mathbb{R}$ such that $\eta_k(t, \omega)\phi(\omega) = \mu_k(t)\phi(\omega)$ for large enough k. Moreover, from the definition of the generalized integral, it is obvious that it is a linear operator and that it is invariant to the linear change $t \to -t$. Based on these considerations, we can write:

$$\left\langle \int_{\mathbb{R}} \cos \omega t U(t) \mathrm{d}t, \phi(\omega) \right\rangle$$
$$= \frac{1}{2} \left\langle \int_{\mathbb{R}} (e^{i\omega t} + e^{-i\omega t}) U(t) \mathrm{d}t, \phi(\omega) \right\rangle$$
$$= \frac{1}{2} \left\langle \int_{\mathbb{R}} e^{i\omega t} U(t) \mathrm{d}t + \int_{\mathbb{R}} e^{-i\omega t} U(t) \mathrm{d}t, \phi(\omega) \right\rangle$$
$$= \frac{1}{2} \left\langle \int_{\mathbb{R}} e^{-i\omega t} U(-t) \mathrm{d}t + \int_{\mathbb{R}} e^{-i\omega t} U(t) \mathrm{d}t, \phi(\omega) \right\rangle$$
$$= \frac{1}{2} \left\langle \int_{\mathbb{R}} e^{-i\omega t} (U(-t) + U(t)) \mathrm{d}t, \phi(\omega) \right\rangle$$
$$= \frac{1}{2} \left\langle \int_{\mathbb{R}} e^{-i\omega t} \mathrm{d}t, \phi(\omega) \right\rangle = \frac{1}{2} \lim_{k\to\infty} \langle e^{-i\omega t}, \phi(\omega) \eta_k(t, \omega) \rangle$$
$$= \frac{1}{2} \lim_{k\to\infty} \langle e^{-i\omega t}, \phi(\omega) \mu_k(t) \rangle$$
$$= \frac{1}{2} \lim_{k\to\infty} \int_{\mathbb{R}^2} e^{-i\omega t} \phi(\omega) \mu_k(t) \mathrm{d}\omega \mathrm{d}t$$
$$= \frac{1}{2} \lim_{k\to\infty} \int_{\mathbb{R}} \left[\int_{\mathbb{R}} e^{-i\omega t} \phi(\omega) \mathrm{d}\omega \right] \mu_k(t) \mathrm{d}t$$
$$= \frac{1}{2} \lim_{k\to\infty} \left\langle 1(t), \mu_k(t) \int_{\mathbb{R}} e^{-i\omega t} \phi(\omega) \mathrm{d}\omega \right\rangle$$
$$= \frac{1}{2} \langle 1(t), \mu_k(t) \mathcal{F}[\phi](t) \rangle = \frac{1}{2} \langle 1(t), \mathcal{F}[\phi](t) \rangle$$
$$= \frac{1}{2} \langle \mathcal{F}[1](\omega), \phi(\omega) \rangle = \langle \pi \delta(\omega), \phi(\omega) \rangle \qquad (39)$$

As the last result does not depend of the particular choice of the unit sequence $\eta_k(t, \omega)$, it follows that the generalized integral of $\cos \omega t U(t)$ in respect to t on $\mathbb{R}$ exists, and equal to $\pi \delta(\omega)$. Here we used the fact that $\mathcal{F}[\phi] \in \mathcal{S}(\mathbb{R})$, i.e. that $\mathcal{F}[\phi]$ is rapidly decreasing function for each test function

$\phi(\omega) \in \mathcal{D}(\mathbb{R})$ [6], and the obvious facts that the constant distribution $1(t)$ is tempered so it can be applied to $\mathcal{F}[\phi]$, and finally, that the product $\mu_k(t)\psi(t)$ converges to $\psi(t)$ for each rapidly decreasing function $\psi(t) \in \mathcal{S}(\mathbb{R})$. Also, we used the well known fact that $\mathcal{F}[1](\omega) = 2\pi\delta(\omega)$, which can be derived rigorously using (22) [17]. Now, the formula (30) is proven.

To prove (31), we will use the similar derivation:

$$\left\langle \int_{\mathbb{R}} \sin \omega t U(t) \mathrm{d}t, \phi(\omega) \right\rangle$$
$$= \frac{1}{2i} \left\langle \int_{\mathbb{R}} (e^{i\omega t} - e^{-i\omega t}) U(t) \mathrm{d}t, \phi(\omega) \right\rangle$$
$$= \frac{1}{2i} \left\langle \int_{\mathbb{R}} e^{i\omega t} U(t) \mathrm{d}t - \int_{\mathbb{R}} e^{-i\omega t} U(t) \mathrm{d}t, \phi(\omega) \right\rangle$$
$$= \frac{1}{2i} \left\langle \int_{\mathbb{R}} e^{-i\omega t} U(-t) \mathrm{d}t - \int_{\mathbb{R}} e^{-i\omega t} U(t) \mathrm{d}t, \phi(\omega) \right\rangle$$
$$= -\frac{1}{2i} \left\langle \int_{\mathbb{R}} e^{-i\omega t} (U(-t) - U(t)) \mathrm{d}t, \phi(\omega) \right\rangle$$
$$= -\frac{1}{2i} \left\langle \int_{\mathbb{R}} e^{-i\omega t} \mathrm{sgn} t \mathrm{d}t, \phi(\omega) \right\rangle$$
$$= -\frac{1}{2i} \lim_{k\to\infty} \langle e^{-i\omega t} \mathrm{sgn}\, t, \phi(\omega) \eta_k(t, \omega) \rangle$$
$$= -\frac{1}{2i} \lim_{k\to\infty} \langle e^{-i\omega t} \mathrm{sgn}\, t, \phi(\omega) \mu_k(t) \rangle$$
$$= -\frac{1}{2i} \lim_{k\to\infty} \int_{\mathbb{R}^2} e^{-i\omega t} \mathrm{sgn}\, t \phi(\omega) \mu_k(t) \mathrm{d}\omega \mathrm{d}t$$
$$= -\frac{1}{2i} \lim_{k\to\infty} \int_{\mathbb{R}} \left[\int_{\mathbb{R}} e^{-i\omega t} \phi(\omega) \mathrm{d}\omega \right] \mathrm{sgn}\, t \mu_k(t) \mathrm{d}t$$
$$= -\frac{1}{2i} \lim_{k\to\infty} \left\langle \mathrm{sgn}\, t, \mu_k(t) \int_{\mathbb{R}} e^{-i\omega t} \phi(\omega) \mathrm{d}\omega \right\rangle$$
$$= -\frac{1}{2i} \langle \mathrm{sgn}\, t, \mu_k(t)\mathcal{F}[\phi](t) \rangle = -\frac{1}{2i} \langle \mathrm{sgn}\, t, \mathcal{F}[\phi](t) \rangle$$
$$= -\frac{1}{2i} \langle \mathcal{F}[\mathrm{sgn}\, t](\omega), \phi(\omega) \rangle = \left\langle \frac{1}{\omega}, \phi(\omega) \right\rangle \qquad (40)$$

Now, it follows that the generalized integral of $\sin \omega t U(t)$ with respect to t on $\mathbb{R}$ exists, and equal to $1/\omega$. In addition to the previously mentioned fact, we also used the known fact that $\mathcal{F}[\mathrm{sgn}\, t](\omega) = 2/i\omega$, which also can be derived using (22). The proof is now completed.

Proof of Theorem 5: As $\phi(x)$ is a test function, we have $\phi(x) \equiv 0$ outside of some compact set K. Let $\eta_k(x)$ be an arbitrary unit sequence on $\mathbb{R}^n$. From the definition of the unit sequence, it follows that exists N such that $\eta_k(x) = 1$ for $x \in K$ and $k \geq N$. In other words, we have $\phi(x)\eta_k(x) = \phi(x)$ for $k \geq N$. So, we have

$$\int_{\mathbb{R}^n} T(x)\phi(x)\mathrm{d}x = \lim_{k\to\infty} \langle T(x)\phi(x), \eta_k(x) \rangle$$
$$= \lim_{k\to\infty} \langle T(x), \eta_k(x)\phi(x) \rangle = \langle T(x), \phi(x) \rangle \qquad (41)$$

This concludes the proof.

Proof of Theorem 6: First, we will consider the case when both a and b are variables. We said that in such case the

expression $\delta(t - a)\delta(t - b)$ should be interpreted as the distribution on $\mathbb{R}^3$ that is derived from the distribution $\delta(u)\delta(v)1(w)$ using the linear change $(u, v, w) \rightarrow (t - a, t - b, t)$. Using the derivation similar to (18), we can derive that

$$\langle \delta(t - a)\delta(t - b), \phi(t, a, b) \rangle = \int_{-\infty}^{\infty} \phi(\tau, \tau, \tau) \mathrm{d}\tau \quad (42)$$

Let?s take an arbitrary test function $\phi(a, b) \in \mathcal{D}(\mathbb{R}^2)$. From the definition of the unit sequence and the compactness of the support of $\phi(a, b)$, it follows that for each unit sequence $\eta_k(t, a, b)$ on $\mathbb{R}^3$ there exists N such that $\phi(t, t)\eta_k(t, t, t) = \phi(t, t)$ for $k \geq N$. Now, from the definition of the generalized integral, we have

$$\left\langle \int_{\mathbb{R}} \delta(t - a)\delta(t - b)\mathrm{d}t, \phi(a, b) \right\rangle$$
$$= \lim_{k \to \infty} \langle \delta(t - a)\delta(t - b), \phi(a, b)\eta_k(a, b, t) \rangle$$
$$= \lim_{k \to \infty} \int_{-\infty}^{\infty} \phi(\tau, \tau)\eta_k(\tau, \tau, \tau) \mathrm{d}\tau$$
$$= \int_{-\infty}^{\infty} \phi(\tau, \tau) d\tau = \langle \delta(a - b), \phi(a, b) \rangle \quad (43)$$

As the last result does not depend of the particular choice of the unit sequence $\eta_k(t, a, b)$, it follows that the generalized integral of $\delta(t - a)\delta(t - b)$ as a distribution on $\mathbb{R}^3$ in respect to t on $\mathbb{R}$ exists, and equal to $\delta(a - b)$. Here we used the interpretation of $\delta(a - b)$ as a distribution on $\mathbb{R}^2$ with formal arguments a and b that is explained earlier. This proves (6) for the case when both a and b are variables.

Suppose now that one of the quantities a or b, say b, is constant. Then, $\delta(t - a)\delta(t - b)$ should be interpreted as a distribution on $\mathbb{R}^2$ with formal arguments t and a that acts on the test functions $\phi(t, a)$ on $\mathbb{R}^2$, which is derived from the distribution $\delta(u)\delta(v)$ using the linear change $(u, v) \rightarrow (t - a, t - b)$. The derivation similar to (18) gives

$$\langle \delta(t - a)\delta(t - b), \phi(t, a) \rangle = \phi(b, b) \quad (44)$$

Now, using the similar reasoning as when both a and b are variables, we have

$$\left\langle \int_{\mathbb{R}} \delta(t - a)\delta(t - b)\mathrm{d}t, \phi(a) \right\rangle$$
$$= \lim_{k \to \infty} \langle \delta(t - a)\delta(t - b), \phi(a)\eta_k(t, a) \rangle$$
$$= \lim_{k \to \infty} \phi(b)\eta_k(b, b) = \phi(b) = \langle \delta(a - b), \phi(a) \rangle \quad (45)$$

At the end, we used the interpretation of $\delta(a - b)$ as a distribution on $\mathbb{R}$ with formal argument a that is explained earlier. This proves (6) for the case when one of the quantities a or b are constant.

Finally, suppose that both a and b are constants, so that both $\delta(t - a)$ and $\delta(t - b)$ are distributions on $\mathbb{R}$. It is quite hard and even not always possible to consistently define the product of two distributions on $\mathbb{R}$ which is again a distribution on $\mathbb{R}$, but all such definitions (which are not always consistent mutually) agree that $\delta(t - a)\delta(t - b) \equiv 0$ when $a \neq b$ [6], [12], [14]. So, the left side of (6) is zero in such case. The right side

of (6) is then $\delta(\lambda)$ where $\lambda = a - b$ is a fixed nonzero number. However, all possible interpretations of eventual point values of distributions agree that $\delta(x)$ is zero for $x \neq 0$, which validates (6) for the case in the consideration. Moreover, even without any definition of the point values of the distributions, it is conventionally accepted that $\delta(x) \equiv 0$ for $x \neq 0$, with the interpretation that $\delta[\phi] = 0$ for any test function ϕ whose support does not include the point $x = 0$. This concludes the proof of the theorem.

Proof of Theorem 7: We will first consider a somewhat simplified case when $T_1(t) \in \mathcal{D}^*(\mathbb{R})$ i $T_2(t) \in \mathcal{D}^*(\mathbb{R})$, i.e. when both distributions are one-dimensional. Then, it is easy to check that we have

$$\langle T_1(u)T_2(t - u), \phi(u, t) \rangle = \langle T_1(u)T_2(t), \phi(u, u + t) \rangle \quad (46)$$

Suppose now that the convolution $T_1 * T_2(t)$ exists. We need to prove that this implies that the integral on the right side of (36) exists too, and that it is equal to this convolution. Indeed, let $\eta_k(u, t)$ be an arbitrary unit sequence on $\mathbb{R}^2$. Then, $\eta_k(u, u + t)$ is also a unit sequence on $\mathbb{R}^2$. Now, using the definition of the generalized integral, the definition of the convolution, and the assumption that $T_1 * T_2(t)$ exists, we have:

$$\left\langle \int_{\mathbb{R}} T_1(u)T_2(t - u)\mathrm{d}u \right\rangle$$
$$= \lim_{k \to \infty} \langle T_1(u)T_2(t - u), \phi(t)\eta_k(u, t) \rangle$$
$$= \lim_{k \to \infty} \langle T_1(u)T_2(t), \phi(u + t)\eta_k(u, u + t) \rangle$$
$$= \langle T_1 * T_2(t), \phi(t) \rangle \quad (47)$$

for any test function $\phi(t) \in \mathcal{D}(\mathbb{R})$. As the result does not depend of the particular choice of the sequence $\eta_k(t, u)$, we can conclude that the generalized integral exists too, and that it is equal to $T_1 * T_2(t)$.

Suppose now that the generalized integral exists. Using the similar reasoning, we have:

$$\langle T_1 * T_2(t), \phi(t) \rangle = \lim_{k \to \infty} \langle T_1(u)T_2(t), \phi(u + t)\eta_k(u, t) \rangle$$
$$= \lim_{k \to \infty} \langle T_1(u)T_2(t - u), \phi(t)\eta_k(u, t - u) \rangle$$
$$= \left\langle \int_{\mathbb{R}} T_1(u)T_2(t - u)\mathrm{d}u, \phi(t) \right\rangle \quad (48)$$

As the sequence $\eta_k(t, t - u)$ is obviously an unit sequence whenever $\eta_k(t, u)$ is, and as the both $\eta_k(t, u)$ and $\phi(t)$ are arbitrary, we can conclude that the convolution exists too, and that it is equal to the generalized integral on the right side of (36).

This concludes the proof for the case of distributions from $\mathcal{D}^*(\mathbb{R})$. The proof for the distributions from $\mathcal{D}^*(\mathbb{R}^n)$, i.e. for the multidimensional case, is completely analogous to the presented proof.

Proof of Theorem 8: To prove (37), we need to show that the following relation is valid for each rapidly decreasing function $\phi(\omega) \in \mathcal{S}(\mathbb{R})$:

$$\left\langle \int_{\mathbb{R}} T(t)e^{-i\omega t}\mathrm{d}t, \phi(\omega) \right\rangle = \langle \mathcal{F}[T](\omega), \phi(\omega) \rangle \quad (49)$$

As the topological space $\mathcal{D}(\mathbb{R})$ is dense in the topological space $\mathcal{S}(\mathbb{R})$ [6], [9], [10] and as the functionals T and $\mathcal{F}[T]$ are continuous on $\mathcal{S}(\mathbb{R})$, it is enough to prove (49) for all test functions $\phi(\omega) \in \mathcal{D}(\mathbb{R})$. Let $\phi(\omega)$ be an arbitrary test function and let $\eta_k(t, \omega)$ be an arbitrary unit sequence on $\mathbb{R}^2$. As the support of $\phi(\omega)$ is compact, from the definition of the unit sequence it follows that there exists an unit sequence $\mu_k(t) \in \mathcal{D}(\mathbb{R})$ such that we have $\eta_k(t, \omega)\phi(\omega) = \mu_k(t)\phi(\omega)$ for large enough k. Furthermore, it is obvious that for each rapidly decreasing function $\psi(t) \in \mathcal{S}(\mathbb{R})$ we have $\mu_k(t)\psi(t) \to \psi(t)$ for $k \to \infty$ in a sense of the topology in $\mathcal{S}(\mathbb{R})$. As $\mathcal{F}[\phi] \in \mathcal{S}(\mathbb{R})$, we can conclude that $\mu_k(t)\mathcal{F}[\phi](t) \to \mathcal{F}[\phi](t)$ for $k \to \infty$ in a sense of the topology in $\mathcal{S}(\mathbb{R})$ as well. Now, from the definition of the generalized integral and all preparations stated above, we can write

$$
\begin{aligned}
&\left\langle \int_{\mathbb{R}} T(t)e^{-i\omega t}\mathrm{d}t, \phi(\omega) \right\rangle \\
&= \lim_{k \to \infty} \langle T(t)e^{-i\omega t}, \phi(\omega)\eta_k(t, \omega) \rangle \\
&= \lim_{k \to \infty} \langle T(t)e^{-i\omega t}, \phi(\omega)\mu_k(t) \rangle \\
&= \lim_{k \to \infty} \langle T(t)1(\omega), e^{-i\omega t}\phi(\omega)\mu_k(t) \rangle \\
&= \lim_{k \to \infty} \left\langle T(t), \mu_k(t) \int_{-\infty}^{\infty} \phi(\omega)e^{-i\omega t}\mathrm{d}\omega \right\rangle \\
&= \lim_{k \to \infty} \langle T(t), \mu_k(t)\mathcal{F}[\phi](t) \rangle \\
&= \langle T(t), \mathcal{F}[\phi](t) \rangle = \langle \mathcal{F}[T](\omega), \phi(\omega) \rangle
\end{aligned}
\tag{50}
$$

The proof of (37) is now completed. The proof of (38) is completely analogous.

REFERENCES

[1] H. M. Srinastava and R. G. Buschman, *Theory and Applications of Convolution Integral Equations*. Dordrecht, The Netherlands: Springer, 1992.

[2] U. J. Monich and H. Boche, "Distributional system representations on bandlimited signals," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4557–4571, Sep. 2010.

[3] P. L. Butzer, "Singular integral equations of Volterra type and the finite part of divergent integrals," *Arch. Rational Mech. Anal.*, vol. 3, no. 1, pp. 194–205, Jan. 1959.

[4] E. J. Weniger and E. O. Steinborn, "The Fourier transforms of some exponential-type basis functions and their relevance to multicenter problems," *J. Chem. Phys.*, vol. 78, no. 10, pp. 6121–6132, May 1983.

[5] M. A. Evgrafov, "Series and integral representations," in *Analysis I*. Berlin, Germany: Springer, 1989, pp. 1–81.

[6] V. S. Vladimirov, *Generalized Functions in Mathematical Physics*. Moscow, Russia: MIR, 1979.

[7] A. Papoulis, *The Fourier Integral and Its Applications*. New York, NY, USA: McGraw-Hill, 1962.

[8] A. Papoulis, *Signal Analysis*. New York, NY, USA: McGraw-Hill, 1977.

[9] W. Rudin, *Functional Analysis*. New Delhi, India: McGraw-Hill, 1973.

[10] M. Reed and B. Simon, *Functional Analysis*. London, U.K.: Academic, 1980.

[11] L. Schwartz, *Mathematical Methods in Physical Sciences*. Moscow, Russia: MIR, 1965.

[12] R. F. Hoskins and J. S. Pinto, *Theories of Generalised Functions—Distributions, Ultradistributions and Other Generalised Functions*. Cambridge, U.K.: Woodhead, 2005.

[13] R. F. Hoskins, *Delta functions: Introduction to Generalized Functions*, 2nd ed. Cambridge, U.K.: Woodhead, 2009.

[14] P. Antosik, J. Mikusinski, and R. Sikorski, *Theory of Distributions—The Sequential Approach*. Amsterdam, The Netherlands: Elsevier, 1973.

[15] J. S. Pinto, *Infinitesimal Methods of Mathematical Analysis*. Chichester, U.K.: Horwood, 2004.

[16] Z. Juric and H. Siljak, "Weak limits and their calculation in analog signal theory," *Int. J. Comput. Math. Elect. Electron. Eng.*, vol. 32, no. 3, pp. 1009–1029, 2013.

[17] M. Reed and B. Simon, *Fourier Analysis and Self-Adjointness*. London, U.K.: Academic, 1975.

ZELJKO JURIC (Member, IEEE) was born in Sarajevo, Bosnia and Herzegovina, in 1968. He received the B.S. degree in electronic circuit theory from the Faculty of Electrical Engineering, University of Sarajevo, in 1995, the M.S. degree in applied mathematics from the Faculty of Science, University of Sarajevo, in 2002, and the Ph.D. degree in control theory from the Faculty of Electrical Engineering, University of Sarajevo, in 2005. From 1996 to 2005, he was a Teaching Assistant with the Department of Mathematics, Faculty of Science, University of Sarajevo, and the Department of Informatics and Computing, Faculty of Electrical Engineering, University of Sarajevo, where he was an Associate Professor, from 2012 to 2017. From 2006 to 2011, he was an Assistant Professor. Since 2018, he has been a Full Professor with the Department of Informatics and Computing, Faculty of Electrical Engineering, University of Sarajevo. He is the author of six books and 36 scientific papers. His research interests include control theory, identification of dynamic systems, signals and systems theory, advanced programming techniques in C++, operational research, discrete mathematics, and mathematical analysis. He is a member of Association of Mathematicians of Bosnia and Herzegovina.

HARUN SILJAK (Senior Member, IEEE) received the bachelor's and master's degrees in electronics and control engineering from the University of Sarajevo, Bosnia and Herzegovina, in 2010 and 2012, respectively, and the Ph.D. degree in electronics and electrical engineering from International Burch University Sarajevo, Bosnia and Herzegovina, in 2015. He is currently an Assistant Professor in embedded systems, control and optimization with the Trinity College Dublin, Dublin, Ireland, where he previously worked as a Marie Curie Postdoctoral Fellow (EDGE Cofund). He has published more than 40 peer-reviewed papers in communications, computation, control, and mathematics, and more than 15 media articles on science, technology, and society. His research interests include complex networks and cyber-physical systems to unconventional methods of computation, communication, and control. He is an active science communicator and the science fiction author, combining his research and storytelling. He is an Associate Editor of *EURASIP Journal on Wireless Communications and Networking*, a Review Editor of *Non-Conventional Communications and Networks*, *Frontiers in Communications and Networks*, and the Science Communication Officer for COST Action MAT-DYN-NET and Marie Curie Alumni Association's Western Balkans Chapter.

● ● ●