Çağatay Candan

Proper Definition and Handling of Dirac Delta Functions

Dirac delta functions are introduced to students of signal processing in their sophomore year. Quite understandably, Dirac delta functions, which should be more aptly called *generalized functions* or *distributions*, cannot be comprehensively given to a young audience at the beginning of their engineering education. Instead, a simplified and abridged definition is presented, and the implications of the definition in signal processing problems are illustrated through numerous examples, following the footsteps of Oppenheim et al. [1], [2].

Students typically learn the properties by developing an affinity through their usage. As their mathematical knowledge matures, some students tend to notice inconsistencies related to the sugar-coated definitions and start questioning the mathematics behind them. Unfortunately, the inquisitive questions of these students are rather difficult to answer convincingly due to the lack of sources on generalized functions at the level of undergraduate/graduate engineering students. The goal of these notes is to scratch the sugarcoating a bit and provide the basics of generalized functions, limits, and derivatives as well as their usage in signal processing problems.

As an illustrative example, the Fourier transform of $f(t) = 1$, which is $F(\Omega) = 2\pi\delta(\Omega)$, is typically "proven" with the application of the inverse Fourier transform on $F(\Omega) = 2\pi\delta(\Omega)$. However, according to the standard calculus results, the Fourier transform of $f(t) = 1$, which is $\mathcal{F}\{1\} = \int_{-\infty}^{\infty} \exp(-j\Omega t)\, dt$, ceases to exist for any Ω in the ordinary calculus sense. The plot further thickens when the Fourier transform of the unit step function, sign function, and Hilbert transform discussions come into play.

Generalized functions enable these calculations, and they are indispensable tools of our field, yet their proper understanding, true definition, and the whys and hows about their usage require an update to our classical calculus knowledge. Such an update, however incomplete, is the topic of this lecture note.

Digital Object Identifier 10.1109/MSP.2021.3055025
Date of current version: 28 April 2021

Relevance

Paul Dirac is one of the giants among the great physicists of the early 20th century. It is a compliment to our profession that he received his first academic degree in electrical engineering (from the University of Bristol). He said,

> I owe a lot to my engineering training because it [taught] me to tolerate approximations. Previously to that I thought . . . one should just concentrate on exact equations all the time. Then I got the idea that, in the actual world, all our equations are only approximate. We must just tend to greater and greater accuracy. In spite of the equations being approximate, they can be beautiful

The function $\delta(t)$ introduced by Dirac is now called the *Dirac delta function*; it provides great computational and conceptual advantages in calculations involving diverging integrals, which is the case for some Fourier integrals. In addition, the inclusion of the Dirac delta function to the calculus of ordinary functions enables the differentiation of discontinuous (generalized) functions, paving the way toward a consistent analysis of highly practical engineering problems, such as circuit theory problems involving switches, unified treatment for mixed random variables (random variables that are both discrete and continuous), and more.

Despite the abundance of topics utilizing Dirac delta functions in signal processing, there are only a few sources explaining the true nature of the approximation involved to the signal processing audience [3, Appendix I], [4]. This column is prepared to answer some of the questions on generalized functions, illustrate their properties, and show their proper usage in some signal processing calculations.

The intended audience of this lecture note includes instructors, researchers with an inclination toward theory, and graduate students getting close to fulfilling their course requirements, say, studying for Ph.D. qualification exams.

 1053-5888/21©2021IEEE

For beginners to the topic, the author suggests following the mainstream track and developing an affinity for the topic first by following the wisdom of Oppenheim et al. [1], [2].

The conventional treatment aims to develop a working knowledge of Dirac delta functions, which is a noteworthy goal on its own, and gives a good "first-order approximation" to the topic. Science and engineering are built upon successively refined approximations, which Paul Dirac has alluded to as a potential source of beauty.Especially in engineering, approximate models/explanations are important, beyond their aesthetic value, because of the basic need for working tools and methods for the solution of practical problems.

In a typical undergraduate course, the need for a working solution may easily overshadow the need for a comprehensive theoretical treatment. As an example, the first course in physics studies the mechanics of inclined planes, stacked boxes with high/low friction surfaces, and so on. If we consider two stacked wood blocks on a flat surface, we may say that the weight of top block is balanced with the normal force so that the net force on the block is zero. This comment can be used to explain why two blocks do not coalesce into a single piece.

However, if we think about the nature of the normal force, it is typically explained as a direct consequence of Newton's laws of motion (the law of action–reaction), and Newton's laws are brought upon students axiomatically in relation to Newton's empirical observations. Hence, the contents of Physics 101 correctly predict that two stacked wood blocks will not coalesce into a single piece without saying much about the mechanism behind the process!

In spite of that, Physics 101 students learn to use and appreciate the benefit of defining a normal force through a series exercises and problems, just like a beginner signal processing student working his or her way through a set of exercise problems on Dirac delta functions. Much later, physicists with advanced degrees learn that the macroscopic normal force is due to the Pauli exclusion principle applied to bulk matter [5]. Needless to say, such a comprehensive answer is of no help to Physics 101 students working on problems with inclined planes.

The situation is almost analogous for signal processing students and Dirac delta functions. Hence, the author believes that exposure to Dirac delta functions beyond the conventional Oppenheim et al. level can be safely postponed to graduate studies. Of course, professionals in the field, lecturers, and researchers can refer quick learners with inquisitive questions to this lecture note, disregarding the suggested timeline.

Prerequisites

The only prerequisites are a working knowledge of freshman calculus, basic signal processing theory, and a keen eye for detail.

Problem statement

The main focus is on the handling of integrals, limits, and derivatives that do not exist in the standard calculus sense. The Fourier transform of $u(t)$ (the unit step function), $F(\Omega) = \int_{-\infty}^{\infty} u(t)\exp(-j\Omega t)\,dt$, is the prime illustrative example. This Fourier transform integral requires the evaluation of $\int_0^{\infty} \cos(\Omega t)\,dt$ and $\int_0^{\infty} \sin(\Omega t)\,dt$, which are known to diverge according to standard calculus results. However, signal processing textbooks express the result as $\mathcal{F}\{u(t)\} = 1/j\Omega + \pi\delta(\Omega)$ [1, Table 4.2].

The appearance of $\delta(\cdot)$ function hints at the divergence of the Fourier integral to an experienced eye, but this is not the case for all divergent integrals. The Fourier transform of $\operatorname{sgn}(t)$ (the sign function) requires the evaluation of $\int_0^{\infty} \sin(\Omega t)\,dt$, which is a divergent integral. However, textbooks state that $\mathcal{F}\{\operatorname{sgn}(t)\} = 2/j\Omega$. The main problem is that the transform pair for both functions is not valid in the ordinary calculus sense but valid in the generalized sense or in the sense of distributions. This article studies the definition of generalized functions and their use in signal processing problems.

Solution

We first present some basic definitions to better explain the upcoming definitions of the Dirac delta and other generalized functions.

Function

Functions, as defined on the set of real numbers, map real numbers to real numbers. Functions are interpreted in a pointwise manner. For example, $\phi(t) = t^2$ maps t_0 in $(-\infty, \infty)$ to t_0^2 in $[0, \infty)$.

Linear functional

A functional is a mapping from the space of functions to real numbers. For example, the area functional defined as $\text{Area}\{\phi\} = \int_{-\infty}^{\infty} \phi(t)\,dt$ maps the function $\phi(t)$ to the numerical value of the total area under $\phi(t)$. A functional that satisfies the linearity conditions (homogeneity and additivity, [1, Sec. 1.6.6]) is called a *linear functional*. Our focus is entirely on linear functionals. Hence, the term *functional* should be interpreted as a linear functional in these notes.

It is easy to verify that the functional $T_f\{\cdot\}$,

$$T_f\{\phi(t)\} \triangleq \langle f(t), \phi(t) \rangle = \int_{-\infty}^{\infty} f(t)\phi(t)\,dt, \tag{1}$$

satisfies the conditions of linearity. We use the notations $T_f\{\phi(t)\}$ and $\langle f(t), \phi(t)\rangle$ interchangeably to denote functionals. $T_f\{\phi(t)\}$ explicitly shows that the "input" $\phi(t)$ is mapped to an "output," i.e., a real number. The function $f(t)$ appearing in the subscript of $T_f\{\phi(t)\}$ characterizes the mapping. As an example, the area functional, previously given, can be realized by substituting $f(t)$ with 1 in (1). The second notation $\langle f(t), \phi(t)\rangle$ is handy in many calculations due to the symmetry between $f(t)$ and $\phi(t)$ in (1).

We refer to the function $\phi(t)$ as the test function. Hence, $T_f\{\phi(t)\}$ is said to operate on test functions. Generalized functions or distributions, shown as $f(t)$, are built upon the "observed" action of functionals on the test functions, as described in the next section.

Generalized equality

If functions $f(t)$ and $g(t)$ induce the same functional, that is, $T_f\{\phi(t)\}$ and $T_g\{\phi(t)\}$ yield identical outputs for all test functions, then functions $f(t)$ and $g(t)$ are said to be equal in the generalized sense. We show the generalized equality with the notation of $f(t) \overset{(g)}{=} g(t)$:

$$f(t) \overset{(g)}{=} g(t) \Leftrightarrow \langle f(t), \phi(t)\rangle = \langle g(t), \phi(t)\rangle \quad \text{for all } \phi(t). \tag{2}$$

To make the statements precise, we need to specify the function class for the test functions and also give a discussion of Lebesgue integration. We refer readers to [6, Ch. 6] for a readable account of these topics. As readers can intuitively appreciate, the class for the test functions should be sufficiently "rich" and "refined" so that the generalized equality in (2) presents practically useful results. For example, if the test functions are limited to constant functions, say, $\phi(t) = c$, where c is a real number, the generalized equality in (2) only implies the equality of the area under two functions, which is of rather limited value.

In this text, we assume that the test function class is infinitely differentiable functions in the form of Gaussian functions:

$$\phi_{\mu,\sigma}(t) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right), \tag{3}$$

with arbitrary mean μ and spread σ. We take this class of test functions as sufficiently rich and refined so that the generalized equality $f(t) \overset{(g)}{=} g(t)$ in (2) becomes practically meaningful. [The class of infinitely differentiable test functions with rapid decay at infinity is called the *Schwartz space* [6], [7]. The Hermite functions, an orthonormal and complete set for L_2, are members of this class. Laurent Schwartz received the Fields Medal in 1950 for building the mathematical foundation (theory of distributions) to the framework of Dirac.]

Dirac delta function

We consider a specific functional, called the *evaluation functional*, that maps the function $\phi(t)$ to $\phi(t_0)$, i.e., the evaluation functional maps $\phi(t)$ to the value of its sample at $t = t_0$. The evaluation functional is clearly linear, but it is not possible to express the evaluation functional in the form of (1) with a regular $f(t)$ function. In spite of that, we substitute $f(t)$ with $\delta(t-t_0)$ in (1) and use the following as a formal definition of the evaluation functional:

$$\int_{-\infty}^{\infty} \delta(t-t_0)\phi(t)\,dt = \phi(t_0) \quad \text{for all } \phi(t). \tag{4}$$

We do not question the existence of the $\delta(t)$ function at this point but treat it as a regular function for now. Readers may interpret (4) as another notation for the evaluation functional from which some properties, such as the linearity of the functional, can be readily observed. Our goal is to derive some properties of $\delta(t)$, given in Table 1, first and then answer existence questions.

Verification of the multiplication property

Let's study the product of $f(t)$ and $\delta(t-t_0)$, which is $f(t_0)\delta(t-t_0)$ according to the multiplication property $f(t)\delta(t-t_0) \overset{(g)}{=} f(t_0)\delta(t-t_0)$ in Table 1. To prove the generalized inequality, we need to show that $\langle f(t)\delta(t-t_0), \phi(t)\rangle = \langle f(t_0)\delta(t-t_0), \phi(t)\rangle$ for all test functions. We focus on the term on left-hand side, $\langle f(t)\delta(t-t_0), \phi(t)\rangle$, first:

$$\begin{aligned}\langle f(t)\delta(t-t_0), \phi(t)\rangle &= \int_{-\infty}^{\infty} f(t)\delta(t-t_0)\phi(t)\,dt \\ &\overset{(a)}{=} \int_{-\infty}^{\infty} \delta(t-t_0)\hat{\phi}(t)\,dt\,\Big|_{\hat{\phi}(t)=f(t)\phi(t)} \\ &\overset{(b)}{=} \hat{\phi}(t_0) \\ &= f(t_0)\phi(t_0).\end{aligned} \tag{5}$$

In line (a), $\hat{\phi}(t) = f(t)\phi(t)$ is introduced, and $\hat{\phi}(t)$ is assumed to be a member of the test function class due to its "richness" and "fineness." Line (b) is due to the definition of evaluation functional.

The right side of equality $f(t)\delta(t-t_0) \overset{(g)}{=} f(t_0)\delta(t-t_0)$ can be worked out as follows:

$$\begin{aligned}&\langle f(t_0)\delta(t-t_0), \phi(t)\rangle \\ &= \int_{-\infty}^{\infty} f(t_0)\delta(t-t_0)\phi(t)\,dt \\ &= f(t_0)\int_{-\infty}^{\infty} \delta(t-t_0)\phi(t)\,dt \\ &= f(t_0)\phi(t_0).\end{aligned} \tag{6}$$

Combining (5) and (6), we have

$$\langle f(t)\delta(t-t_0), \phi(t)\rangle = \langle f(t_0)\delta(t-t_0), \phi(t)\rangle, \quad \text{for all } \phi(t), \tag{7}$$

which concludes the proof of $f(t)\delta(t-t_0) \overset{(g)}{=} f(t_0)\delta(t-t_0)$.

An important takeaway message from the proof of the first property is not the final result but the proof procedure followed for the generalized

Table 1. The properties for the Dirac delta function and its derivatives.

Basic	Multiplication	$f(t)\delta(t-t_0) \overset{(g)}{=} f(t_0)\delta(t-t_0)$
	Scaling	$\delta(at) \overset{(g)}{=} \frac{1}{\lvert a\rvert}\delta(t)$
	Sifting	$\int_{-\infty}^{\infty} f(t)\delta(t-t_0)\,dt = f(t_0)$
	Convolution	$\delta(t) * f(t) \overset{(g)}{=} f(t)$
Advanced	Multiplication	$f(t)\delta^{(n)}(t-t_0) \overset{(g)}{=} \sum_{k=0}^{n}(-1)^k\binom{n}{k}f^{(k)}(t_0)\delta^{(n-k)}(t-t_0)$, where $\frac{d^n}{dt^n}\delta(t) = \delta^{(n)}(t)$, and $\frac{d^n}{dt^n}f(t) = f^{(n)}(t)$
	Scaling	$\delta(f(t)) \overset{(g)}{=} \sum_{k=1}^{K}\frac{1}{\lvert f'(t_k)\rvert}\delta(t-t_k)$, where t_k are zeros of $f(t)$, i.e., $f(t_k) = 0, k = \{1,2,\ldots,K\}$
	Sifting	$\int_{-\infty}^{\infty} f(t)\delta^{(n)}(t-t_0)\,dt = (-1)^n f^{(n)}(t_0)$
	Convolution	$f(t) * \delta^{(n)}(t) \overset{(g)}{=} f^{(n)}(t)$

equality. The equality sign $\overset{(g)}{=}$ appearing in $f(t) \overset{(g)}{=} g(t)$ denotes the equality of the functionals for every member of the test function class. It is, indeed, very different from the ordinary equality sign.

A rather silly, but memorable, analogy given by one of my instructors can be repeated as follows: Assume that you are in a county fair, and there is a contest to identify an unknown animal. Contestants are allowed to ask only yes/no questions. After several rounds of questions, you learn that the animal is green, lives in a lake, is capable of leaping significant distances, and quacks. Given this information, can you say that the animal is a frog?

If you have asked a sufficiently large number of informative questions (the richness and fineness of the question class), you can be pretty sure that the animal is a frog! However, there is always a possibility that the animal is of another species that is capable of imitating a frog quite closely! If you are only interested in the actions of this animal, though, there is no harm in calling the animal, irrespective of its genus, a frog or a generalized frog!

Analogous to the story, a generalized function $f(t)$ is characterized by its response to the probing test functions $\phi(t)$. Generalized functions are declared equal if they give the same response to all test functions.

The major mishap in the treatment of the impulse function or Dirac delta function in all signal processing texts is the usage of an ordinary equality sign $=$ instead of a generalized equality sign $\overset{(g)}{=}$. This carries the potential of interpreting equations involving $\delta(t)$ in a pointwise manner, which is prone to inconsistencies and calculation mistakes.

Verification of the scaling property

Let's verify the scaling property $\delta(at) \overset{(g)}{=} (1/|a|)\delta(t)$, given in Table 1. The left side of the equality can be written as

$$\begin{aligned}\langle \delta(at), \phi(t)\rangle &= \\ &\int_{-\infty}^{\infty} \delta(at)\phi(t)\,dt \,\Big|_{u=at} \\ &= \frac{1}{|a|}\int_{-\infty}^{\infty} \delta(u)\phi\Big(\frac{u}{a}\Big)du = \frac{\phi(0)}{|a|}. \end{aligned} \quad (8)$$

Here, $\phi(u/a)$ is assumed to be in the test function class, as in the proof of the first property, and we have treated $\delta(at)$ as a regular function and changed the integration variable from t to $u = at$ without any due diligence (more on this later).

The right side of the equality $\delta(at) \overset{(g)}{=} (1/|a|)\delta(t)$ can be written as

$$\begin{aligned}\left\langle \frac{1}{|a|}\delta(t), \phi(t)\right\rangle &= \frac{1}{|a|}\langle \delta(t), \phi(t)\rangle \\ &= \frac{\phi(0)}{|a|}. \end{aligned} \quad (9)$$

Equations (8) and (9) imply the generalized equality of $\delta(at) \overset{(g)}{=} (1/|a|)\delta(t)$. Note that setting $a = -1$ in the scaling property gives $\delta(t) \overset{(g)}{=} \delta(-t)$, which is the evenness of function $\delta(t)$ in the generalized sense.

Generalized limit

Up to this point, we have averted the existence questions on the $\delta(t)$ function but, rather, focused on its properties. Now, we present a limit argument for the construction of the Dirac delta function. The described limit operation is called the *generalized limit*. In standard textbooks, the Dirac delta function is introduced as the pointwise limit of ordinary functions, which is not the correct definition and the root cause of confusion in many discussions.

The generalized limit of ordinary functions $f_n(t)$ is said to be a generalized function $f(t)$, if

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} f_n(t)\phi(t)\,dt = \int_{-\infty}^{\infty} f(t)\phi(t)\,dt \quad (10)$$

is satisfied for all test functions $\phi(t)$. We denote the generalized limit as $f_n(t) \overset{(g)}{\to} f(t)$.

The Dirac delta function can be given as the generalized limit of ordinary $f_n(t)$ functions defined as follows:

$$f_n(t) = \begin{cases} \dfrac{n}{\epsilon}, & -\dfrac{\epsilon}{2n} < t < \dfrac{\epsilon}{2n}. \\ 0, & \text{other} \end{cases} \quad (11)$$

From Figure 1, it can be seen that $f_n(t)$ is a pulse of duration ϵ/n centered around $t = 0$. The area under $f_n(t)$ is unity for all n. With the running assumption that the test functions $\phi(t)$ are sufficiently smooth, we can expand the function into the Taylor series around $t = 0$:

$$\begin{aligned}\phi(t) &= \phi(0) + \phi'(0)t + \phi^{(2)}(0)\frac{t^2}{2} \\ &\quad + \text{h.o.t.} \end{aligned} \quad (12)$$

Here, h.o.t. refers to the higher-order terms of the Taylor series expansion. As $n \to \infty$, the support of function $f_n(t)$, as shown in Figure 1, approaches zero. Hence, the product $\phi(t)f_n(t)$ can be approximated with the first term of the Taylor series expansion, which is $\phi(0)f_n(t)$, for large enough n. As a result, we have the equality of

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} f_n(t)\phi(t)\,dt = \phi(0) \quad (13)$$

in the usual calculus sense. Given the generalized limit definition, this concludes the proof of $f_n(t) \overset{(g)}{\to} \delta(t)$ as $n \to \infty$.

The definition of the Dirac delta function as a generalized limit of ordinary functions is important in practice. Whenever in doubt, it is possible to replace $\delta(t)$ with the $f_n(t)$ functions in (11), solve the problem of interest, and then calculate the ordinary limit of the final result as $n \to \infty$. Readers are invited to do this calculation to have another verification of the scaling property in Table 1. Furthermore, the generalized limit definition establishes a connection with the "physical" interpretation of the Dirac delta function as a very-short-duration pulse, but readers should always keep in mind that the limit operation for getting shorter and shorter pulses is not an ordinary pointwise limit operation, as introduced in many undergraduate texts, but a generalized limit operation.

The Dirac delta definition by a generalized limit argument is not specific to $f_n(t)$ given by (11). Readers can examine [1, Problem 1.38] for some other ordinary functions for which the generalized limit is $\delta(t)$. The basic requirement is the construction of a unit area function sequence with diminishing support. It can be verified that both

$$\begin{aligned} g_n(t) &= \sqrt{\frac{n}{2\pi}}\exp(-nt^2/2) \text{ and} \\ h_n(t) &= n\,\mathrm{sinc}(nt) = \frac{\sin(\pi n t)}{\pi t} \end{aligned} \quad (14)$$

tend to $\delta(t)$ as $n \to \infty$ in the generalized sense.

Figure 2 shows the sketch of $h_n(t) = n\text{sinc}(nt)$ for different n values. The main lobe of the function $h_n(t)$ gets narrower and taller as n increases, yet, however large n is, there exist two sidelobes, with a peak value of about one-fifth the maximum value, on both sides of the main lobe. Furthermore, by fixing t to a nonzero value, say t_0, and evaluating $\lim_{n\to\infty} h_n(t_0)$, we get

$$\lim_{n\to\infty} h_n(t_0) = \frac{1}{\pi t_0} \lim_{n\to\infty} \sin(\pi n t_0),$$

which does not exist in the usual sense. Hence, $h_n(t)$ does not approach $\delta(t)$ in a manner that is as described in many undergraduate textbooks but approaches in the generalized sense or, equivalently, in the weak limit sense [8].

Generalized derivative of the Dirac delta function

The derivative of the Dirac delta function $d/dt\{\delta(t)\}$ is called the *doublet function* [1, Sec. 2.5.3]. It is no surprise that the differentiation operation in $d/dt\{\delta(t)\}$ is in the generalized sense, that is, according to the introduced generalized limit definition. To understand this operation, let's examine the response of $d/dt\{\delta(t)\}$ to a test function:

$$\int_{-\infty}^{\infty} \frac{d}{dt}\{\delta(t)\}\phi(t)\,dt = \delta(t)\phi(t)\Big|_{t=-\infty}^{t=\infty} - \int_{-\infty}^{\infty} \delta(t)\frac{d}{dt}\phi(t)\,dt = -\frac{d}{dt}\phi(t)\Big|_{t=0} = -\phi^{(1)}(0). \quad (15)$$

This calculation is based on the application of integration by parts to the leftmost side of (15). Since the test function $\phi(t)$ is a member of scaled and shifted Gaussian functions, the term $\delta(t)\phi(t)\Big|_{t=-\infty}^{t=\infty}$ vanishes. The other term, the integral term of the integration-by-parts operation, can be expressed using the sifting property of the Dirac delta function. Hence, we get the defining relation for the doublet function as

$$\int_{-\infty}^{\infty} \delta^{(1)}(t)\phi(t)\,dt = -\phi^{(1)}(0). \quad (16)$$

Here, $\delta^{(n)}(t)$ and $\phi^{(n)}(t)$ refer to the nth derivative of the Dirac delta and test function $\phi(t)$ in the generalized and ordinary sense, respectively.

At this point, readers should be rightfully uncomfortable with the application of integration by parts with an integrand containing a Dirac delta function, as in (15). To the comfort of these readers (and also the ones still uneasy about the change of variables from t to $u = at$ in the scaling property discussion), we present an alternative proof path and suggest replacing $\delta(t)$ with the ordinary function $h_n(t)$ given in (14). The integration-by-parts operation with the substituted $h_n(t)$ function is now well defined, and the final result becomes

$$\int_{-\infty}^{\infty} \frac{d}{dt}\{h_n(t)\}\phi(t)\,dt = -\int_{-\infty}^{\infty} h_n(t)\frac{d}{dt}\phi(t)\,dt. \quad (17)$$

By taking the limit of both sides in (17) as $n \to \infty$ and using the generalized limit definition in (10), we reach the conclusion that, since $h_n(t) \overset{(g)}{\to} \delta(t)$, we have $d/dt\{h_n(t)\} \overset{(g)}{\to} \delta^{(1)}(t)$. The formal definition of $\delta^{(1)}(t)$ becomes the relation in (16).

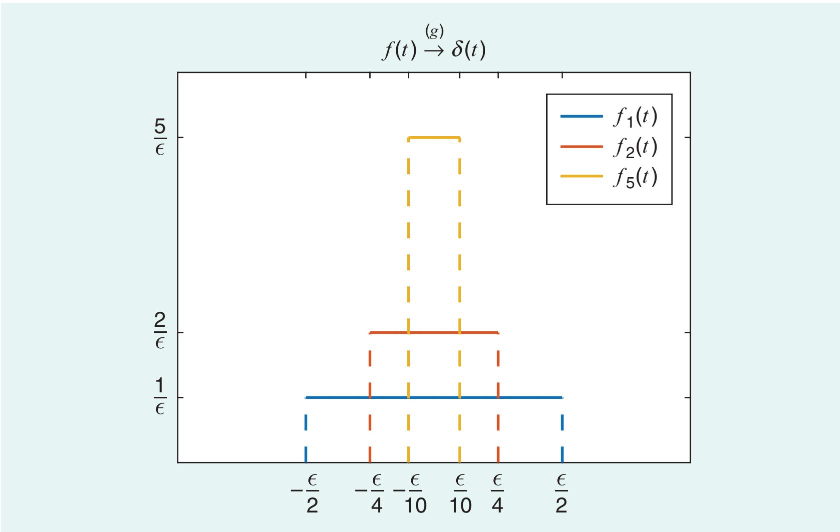

FIGURE 1. The convergence of pulse sequences $f_n(t)$ to $\delta(t)$.

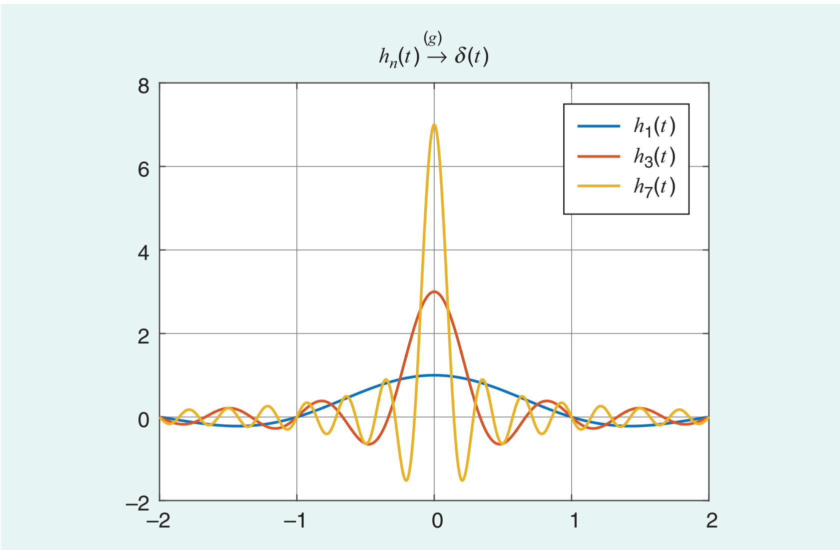

FIGURE 2. The convergence of $h_n(t) = n\text{sinc}(nt)$ to $\delta(t)$. Convergence is not in the pointwise sense!

Sifting and other properties for higher-order derivatives of the Dirac delta function are given in Table 1. These results can be called *advanced results*, since they require more than a basic understanding of the generalized functions. Many signal processing textbooks avoid these properties since even a partial justification of these results requires much more than a pictorial or pointwise justification of the $\delta(t)$ function.

Derivative of the unit step function

By replacing $h_n(t)$ with an arbitrary regular function $f(t)$ in (17), we get

$$\int_{-\infty}^{\infty} \frac{d}{dt}\{f(t)\}\,\phi(t)\,dt = -\int_{-\infty}^{\infty} f(t)\frac{d}{dt}\phi(t)\,dt. \tag{18}$$

Substituting $f(t)$ in (18) with the unit step function $u(t)$ yields

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{d}{dt}\{u(t)\}\,\phi(t)\,dt &= -\int_{-\infty}^{\infty} u(t)\frac{d}{dt}\phi(t)\,dt\\ &= -\int_{0}^{\infty} \frac{d}{dt}\phi(t)\,dt\\ &\overset{(a)}{=} \phi(0) - \cancel{\phi(\infty)}\\ &= \langle \delta(t), \phi(t)\rangle,\end{aligned} \tag{19}$$

where $\phi(\infty)=0$ is used in line (a), which is due to the test function class definition. The leftmost and rightmost sides of (19) imply that $\langle (d/dt)u(t), \phi(t)\rangle = \langle \delta(t), \phi(t)\rangle$ for all test functions. This statement is equivalent to $(d/dt)u(t) \overset{(g)}{=} \delta(t)$.

From this discussion, we reach the important conclusion that an ordinary function, such as $u(t)$, when interpreted as a generalized function, has derivatives of all orders. In other words, function $u(t)$ is not a differentiable function due to its discontinuity at $t=0$, but it is differentiable for all orders in the generalized sense.

Application examples

A number of examples are presented to illustrate the application of the Dirac delta function. Our goal is to relate the applications to the generalized definitions on functions, limits, derivatives, and so on.

Example 1

Assume that a sequence $y[n]$ is formed by down-sampling $x[n]$ by two: $y[n]=x[2n]$. It is well known that the spectrum of $y[n], Y(e^{j\omega})$ is related to the spectrum of $x[n]$, $X(e^{j\omega})$, according to the relation [2, Sec. 3.6.1]

$$Y(e^{j\omega}) = \frac{1}{2}\left(X\left(e^{j\frac{\omega}{2}}\right) + X\left(e^{j\left(\frac{\omega}{2}+\pi\right)}\right)\right). \tag{20}$$

In this example, we would like to illustrate the validity of this expression for $x[n]=\exp(j\omega_0 n)$. This exercise is quite trivial from the time-domain-processing viewpoint. Since $y[n]=x[2n]=\exp(j2\omega_0 n)$, $y[n]$ is a complex exponential whose frequency is doubled after down-sampling. The frequency-domain representations of $x[n]$ and $y[n]$ are $X(e^{j\omega}) = 2\pi\delta(\omega-\omega_0)$, and $Y(e^{j\omega}) = 2\pi\delta(\omega-2\omega_0)$, respectively. This example aims to verify this basic result directly from (20).

It should be remembered that the expressions for $X(e^{j\omega})$ and $Y(e^{j\omega})$ are periodic with 2π, as the notation implies. Let's check the validity of (20) for $X(e^{j\omega}) = 2\pi\delta(\omega-\omega_0)$:

$$\begin{aligned}Y(e^{j\omega}) &= \frac{1}{2}\left(X\left(e^{j\frac{\omega}{2}}\right) + X\left(e^{j\left(\frac{\omega}{2}+\pi\right)}\right)\right)\\ &= \frac{2\pi}{2}\Big[\delta\Big(\frac{\omega}{2}-\omega_0\Big)\\ &\quad + \delta\Big(\frac{\omega}{2}+\pi-\omega_0\Big)\Big]\\ &= \frac{2\pi}{2}\Big[\delta\Big(\frac{\omega-2\omega_0}{2}\Big)\\ &\quad + \delta\Big(\frac{\omega+2\pi-2\omega_0}{2}\Big)\Big]\\ &\overset{(a)}{=} \frac{2\pi}{2}[2\delta(\omega-2\omega_0)\\ &\quad + 2\delta(\omega+2\pi-2\omega_0)]\\ &\overset{(b)}{=} 2\pi\delta(\omega-2\omega_0).\end{aligned}$$

In line (a), we have used the scaling property of the Dirac delta function from Table 1. In line (b), the expression is rewritten to cover only a single 2π period, following the convention. As expected, the final result indeed matches the earlier result found from time-domain considerations.

Comment

The spectrum after a down-sampling operation is typically found with a frequency-domain sketch that indicates the support of $X(e^{j\omega})$ and its translated versions (see [2, Fig. 3.18]). Such a sketch is also useful to illustrate the aliasing concept. We see that when the spectrum involves a Dirac delta function, a sketch is not sufficient to explain the vanishing 1/2 coefficient in (20). We need to bring the scaling property of the Dirac delta function into play.

Example 2

Let X be a random variable with the probability density function (pdf) $f_X(x)$. The problem of interest is the pdf of the random variable $Z=X^2$.

This is a standard probability problem, and we would like to illustrate the utility of the Dirac delta function in this calculation:

$$\begin{aligned}f_Z(z) &\overset{(a)}{=} \int f_{X,Z}(x,z)\,dx\\ &\overset{(b)}{=} \int f_X(x) f_{Z|X}(z\,|\,X=x)\,dx\\ &\overset{(c)}{=} \int f_X(x)\,\delta(z-x^2)\,dx\\ &\overset{(d)}{=} \int f_X(x)\,\delta(x^2-z)\,dx\\ &\overset{(e)}{=} \int f_X(x)\Bigg(\frac{\delta(x-\sqrt{z})}{2\sqrt{z}}\\ &\quad + \frac{\delta(x+\sqrt{z})}{2\sqrt{z}}\Bigg)dx \quad (\text{with } z \geq 0)\\ &\overset{(f)}{=} \frac{1}{2\sqrt{z}}\left(f_X(\sqrt{z}) + f_X(-\sqrt{z})\right).\end{aligned} \tag{21}$$

Line (a) is the marginalization operation. Line (b) includes a factorization for the joint density in terms of the conditional density. Line (c) introduces $Z=X^2$ into the calculation. Line (d) is due to the evenness of the Dirac delta function, $\delta(x)=\delta(-x)$. Line (e) uses the scaling property of Table 1 (from the "Advanced" section of the table). It is important to note that the integration variable in line (d) is x. Hence, for the function $\delta(x^2-z)$ appearing in the integrand, x is the variable, and z is just a constant value. Therefore, the scaling property of the Dirac delta function should be utilized by treating function x^2-z as a function of the variable x. Line (f) is due to the sifting property.

Comment

We observe that the inclusion of the Dirac delta in the operational calculus results in significant shortening of the algebra. Note that the calculation given in (21) exactly mimics a similar calculation given for the discrete random variables (probability mass functions).

More specifically, line (c) of (21) can be interpreted as follows: Let's assume that $z = 100$ and consider the integral $\int_{-\infty}^{\infty} f_X(x)\delta(z - x^2)dx$. Since the function $\delta(z - x^2)$ is equal to zero when $z \neq x^2$, this integral corresponds to checking all $x \in (-\infty, \infty)$ to find the ones satisfying the condition $x^2 = z = 100$ and "summing up" $f(x)$ values corresponding to these x values.

The main difficulty for instructors is not this interpretation but explaining the factor $1/2\sqrt{z}$, which is the Jacobian term arising during the functional mapping of random variables. The Jacobian term does not arise in discrete random variables, and the "summing up" interpretation becomes exactly correct; that is, for the probability mass functions, the sum of the probability values for x that satisfy the condition $z = x^2$ gives the probability of z. With the inclusion of the Dirac delta function in the calculus, the $1/2\sqrt{z}$ term in line (f) of (21) effortlessly comes out with the application of the scaling property.

Example 3

Let X and Y be two random variables with the joint pdf $f_{X,Y}(x,y)$. The problem is the derivation of the pdf for the random variable $Z = X + Y^2$:

$$\begin{aligned} f_Z(z) &\overset{(a)}{=} \int_x \int_y f_{X,Y}(x,y)\delta(z - x - y^2)\,dy\,dx \\ &\overset{(b)}{=} \int_y \Big(\int_x f_{X,Y}(x,y)\delta(z - x - y^2)\,dx\Big)dy \\ &\overset{(c)}{=} \int_y \Big(\int_x f_{X,Y}(x,y)\delta(x - z + y^2)\,dx\Big)dy \\ &\overset{(d)}{=} \int_y f_{X,Y}(z - y^2, y)\,dy \\ &\overset{(e)}{=} \int_y f_Y(y) f_{X|Y}(z - y^2 \,|\, Y = y)\,dy. \end{aligned} \tag{22}$$

Line (a) is the "summing up" operation of $f_{X,Y}(x,y)$ values for which the condition $z = x + y^2$ is satisfied. In line (b), the order of integration is exchanged, that is, the inner integration is with respect to x after the exchange. Line (c) is due to the evenness of $\delta(x)$. Line (d) is due to the sifting property. Line (e) is the factorization of joint density in terms of the conditional density of X given Y.

Comment

By changing the integration order in line (c), the variable for the function $\delta(z - x - y^2)$ becomes x. After the order change, the variables z and y are treated as constants, and we have the result in line (d). If the inner integral in line (c) were with respect to the variable y, that is, if we do not change the order of integration, we need to use result given in Example 2 to evaluate the integral involving $\delta(z - x - y^2)$.

Example 4

Show that the Fourier transform of $f(t) = 1$ is $F(\Omega) = 2\pi\delta(\Omega)$, where $F(\Omega) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{\infty} f(t)\exp(-j\Omega t)\,dt$ is the Fourier transformation operation.

Freshman calculus results state that $\int_{-\infty}^{\infty} f(t)\exp(-j\Omega t)\,dt$ does not converge for any Ω for $f(t) = 1$. Hence, the well-known Fourier transform pair of $1 \leftrightarrow 2\pi\delta(\Omega)$ should be interpreted in the generalized sense. To show $\mathcal{F}\{1\} \overset{(g)}{=} 2\pi\delta(\Omega)$, we need to examine the response of the function $F(\Omega) = \mathcal{F}\{1\}$ to a test function $\Phi(\Omega)$:

$$\begin{aligned} \langle \mathcal{F}\{1\}, \Phi(\Omega)\rangle &\overset{(a)}{=} \int_\Omega \mathcal{F}\{1\}\Phi(\Omega)\,d\Omega \\ &\overset{(b)}{=} \int_\Omega \Big(\int_t 1 e^{-j\Omega t}dt\Big)\Phi(\Omega)\,d\Omega \\ &\overset{(c)}{=} \int_t \Big(\int_\Omega \Phi(\Omega) e^{-j\Omega t}d\Omega\Big)dt \\ &\overset{(d)}{=} \int_t 2\pi\phi(-t)\,dt \\ &\overset{(e)}{=} 2\pi\Phi(0) \\ &\overset{(f)}{=} 2\pi \int \delta(\Omega)\Phi(\Omega)\,d\Omega \\ &\overset{(g)}{=} \langle 2\pi\delta(\Omega), \Phi(\Omega)\rangle. \end{aligned} \tag{23}$$

Line (a) is due to the linear functional definition. Line (b) results from the definition of the Fourier transform. Line (c) changes the integration order. Line (d) is due to the inverse Fourier transform relation for ordinary, absolutely integrable functions [1]. [With the assumed test function class (Gaussian functions), the Fourier integral, that is, $F\{\phi(t)\} = \Phi(\Omega)$, is guaranteed to converge in the ordinary calculus sense.] Line (e) is due to the fact that $\Phi(0) = \int_t \phi(t)\,dt$, that is, the area of the time-domain function, is the value of its Fourier representation at $\Omega = 0$. Lines (f) and (g) are different ways of writing line (e). Considering the leftmost and rightmost sides of (23) and remembering the generalized equality definition in (2), we can conclude the proof of $\mathcal{F}\{1\} \overset{(g)}{=} 2\pi\delta(\Omega)$.

Comment

In a first course, this relation is given by finding the inverse Fourier transform of $2\pi\delta(\Omega)$, i.e., $\mathcal{F}^{-1}\{2\pi\delta(\Omega)\}$, without mentioning the existence of the Fourier integral for $f(t) = 1$. The Fourier integral for $f(t) = 1$ diverges in the usual sense but exists only in the generalized sense or in the sense of distributions.

Example 5

Show that the Fourier transform of $f(t) = \mathrm{sgn}(t)$ is $F(\Omega) \overset{(g)}{=} 2/j\Omega$.

The Fourier transform of $\mathrm{sgn}(t)$,

$$\mathrm{sgn}(t) = \begin{cases} 1 & t > 0 \\ -1 & t < 0 \end{cases},$$

can be written as the integral

$$\mathcal{F}\{\mathrm{sgn}(t)\} = \frac{2}{j}\int_0^\infty \sin(\Omega t)\,dt, \tag{24}$$

which does not converge in the ordinary calculus sense. Hence, as suspected, $\mathcal{F}\{\mathrm{sgn}(t)\}$ is equal to $2/j\Omega$ in the distribution sense. It is interesting to note that there is no Dirac delta function in the expression $\mathcal{F}\{\mathrm{sgn}(t)\} \overset{(g)}{=} 2/j\Omega$, immediately giving away that the equality is in the generalized sense.

Let's define a regular function $g_T(\Omega)$ as $g_T(\Omega) = \int_0^T \sin(\Omega t)\,dt = (1 - \cos(\Omega T))/\Omega$. We would like to take the limit of $g_T(\Omega)$ as $T \to \infty$ with the goal of evaluating the transform in (24). To do that, we need to examine the response of $g_T(\Omega)$ to a test function $\Phi(\Omega)$, that is, $\langle g_T(\Omega), \Phi(\Omega)\rangle$, and then evaluate the limit of the response as $T \to \infty$.

For a fixed T, $\langle g_T(\Omega), \Phi(\Omega)\rangle$ can be expressed as

$$\langle g_T(\Omega), \Phi(\Omega)\rangle = \langle \frac{1}{\Omega}, \Phi(\Omega)\rangle - \langle \frac{\cos(\Omega T)}{\Omega}, \Phi(\Omega)\rangle = \langle \frac{1}{\Omega}, \Phi(\Omega)\rangle - \langle \cos(\Omega T), \frac{\Phi(\Omega)}{\Omega}\rangle. \tag{25}$$

As $T \to \infty$, the equality in (25) approaches

$$\lim_{T\to\infty}\langle g_T(\Omega), \Phi(\Omega)\rangle = \langle \frac{1}{\Omega}, \Phi(\Omega)\rangle - \lim_{T\to\infty}\langle \cos(\Omega T), \frac{\Phi(\Omega)}{\Omega}\rangle. \tag{26}$$

From (26), it is clear that we need to show $\lim_{T\to\infty}\langle \cos(\Omega T), (\Phi(\Omega)/\Omega)\rangle = 0$ to conclude the proof. Since the test function class is the class of Gaussian functions, the function $\Phi(\Omega)/\Omega$ is absolutely integrable in $\Omega \in (-\infty, \infty)$ in the Cauchy principle value sense. (The Cauchy principle value integral is required due to the singularity of $\Phi(\Omega)/\Omega$ at $\Omega = 0$ [4, p. 359]).

We know from Dirichlet conditions that the Fourier transform of an absolutely integrable function exists in the regular sense [1, p. 290]. An important but less-known fact by the signal processing audience is the Riemann–Lebesgue lemma, stating that, if $x(t)$ is absolutely summable, then $X(\Omega) \to 0$ as $\Omega \to \infty$ [3, p. 278].

Armed with this knowledge, $\langle \cos(\Omega T), (\Phi(\Omega)/\Omega)\rangle$ can be interpreted as the real part of the $\mathcal{F}\{\Phi(\Omega)/\Omega\}$ with the transform-domain variable T. Then, due to the absolute integrability of $\Phi(\Omega)/\Omega$ and the Riemann–Lebesgue lemma, we have $\lim_{T\to\infty}\langle \cos(\Omega T), (\Phi(\Omega)/\Omega)\rangle = 0$.

By multiplying both sides of (26) by $2/j$ and replacing $g_T(\Omega)$ with $\int_0^T \sin(\Omega t)\,dt$, we reach

$$\lim_{T\to\infty}\int_{-\infty}^{\infty}\left(\frac{2}{j}\int_0^T \sin(\Omega t)\,dt\right)\Phi(\Omega)\,d\Omega = \int_{-\infty}^{\infty}\left(\frac{2}{j\Omega}\right)\Phi(\Omega)\,d\Omega, \tag{27}$$

stating that $\mathcal{F}\{\mathrm{sgn}(t)\} \overset{(g)}{=} 2/j\Omega$ via the generalized limit definition given in (10).

Comment

A first course in signal processing needs to sugarcoat some definitions and even some calculations due to pedagogical reasons. Among these, the Fourier transformations of the sign function and unit step function stand out. The sign function, $\mathrm{sgn}(t)$, is clearly not absolutely or square summable; hence, its Fourier transform cannot be given in the usual sense.

In spite of that, to show this result, some instructors calculate the Fourier transform of a regular, absolutely summable function $\mathrm{sgn}(t)e^{-\alpha|t|}$; evaluate the limit of the result as $\alpha \to 0$; and then present the limit as the Fourier transform of $\mathrm{sgn}(t)$. The end result of this calculation matches the correct result, but the intermediate steps, especially the one involving the movement of the limit operation inside of the Fourier transform integral in the final step, are highly questionable. It should be clear at this point that any treatment of integrals diverging in the ordinary calculus sense requires some extraordinary effort. The definition of generalized functions is an effort along this line.

As expected, the Fourier transform of $u(t)$ is also only valid in the generalized sense. By expressing $u(t)$ as $u(t) = (\mathrm{sgn}(t)+1)/2$ and applying the linearity of the Fourier transform, we can show $\mathcal{F}\{u(t)\} \overset{(g)}{=} 1/j\Omega + \pi\delta(\Omega)$.

Example 6

Find the inverse unilateral Laplace transform of $X(s) = s^2/(s+3)$.

This problem is typically solved by partial fraction expansion, that is,

$$X(s) = \frac{s^2}{s+3} = s - 3 + \frac{9}{s+3}, \tag{28}$$

followed by inverse Laplace transformation via transform-pair recognition. The final answer of this example is $x(t) = \delta^{(1)}(t) - 3\delta(t) + 9\exp(-3t)u(t)$. Our goal is to derive the same result via some alternative paths to illustrate the usage of generalized differentiation.

Let's first express $X(s)$ as $X(s) = s^2X_p(s)$, where $X_p(s) = 1/(s+3)$. The inverse Laplace transform of $X_p(s)$ is $x_p(t) = \exp(-3t)u(t)$. Hence, the inverse Laplace transform $X(s) = s^2X_p(s)$ becomes $x(t) = (d^2/dt^2)x_p(t)$. We can verify this result by remembering that the unilateral Laplace transform of $(d/dt)x(t)$ is $sX(s) - x(0^-)$. Note that $x_p(t)$ and its derivatives are all zero at $t = 0^-$ due to the existence of the $u(t)$ term in $x_p(t)$. Let's evaluate the first two derivatives of $x_p(t)$ and compare the result with the answer by partial fraction expansion:

$$\begin{aligned}
x_p^{(1)}(t) &= \frac{d}{dt}\{\exp(-3t)u(t)\} \\
&\overset{(a)}{=} \frac{d}{dt}\{\exp(-3t)\}u(t) + \exp(-3t)\frac{d}{dt}\{u(t)\} \\
&= -3\exp(-3t)u(t) + \exp(-3t)\delta(t) \\
&\overset{(b)}{=} -3\exp(-3t)u(t) + \delta(t), \\
x_p^{(2)}(t) &= \frac{d}{dt}x_p^{(1)}(t) \\
&= \frac{d}{dt}\{-3\exp(-3t)u(t) + \delta(t)\} \\
&\overset{(a)}{=} \frac{d}{dt}\{-3\exp(-3t)\}u(t) - 3\exp(-3t)\frac{d}{dt}\{u(t)\} + \frac{d}{dt}\{\delta(t)\} \\
&= 9\exp(-3t)u(t) - 3\exp(-3t)\delta(t) + \delta^{(1)}(t) \\
&\overset{(b)}{=} 9\exp(-3t)u(t) - 3\delta(t) + \delta^{(1)}(t).
\end{aligned} \tag{29}$$

Line (a) of both equations is due to the product rule for differentiation and the generalized equality of $(d/dt)u(t) \overset{(g)}{=} \delta(t)$. Line (b) is due to the multiplication property of the Dirac delta function from Table 1. Note that the equalities given in (29) are not ordinary equalities but valid only in the generalized sense. The absence of the $\overset{(g)}{=}$ symbol can be a source of inconsistencies and confusion, yet we go back to the conventional notation and symbols in this last example.As a final exercise, let's redo the calculation by evaluating the second derivative of $x_p(t) = f(t)g(t)$ with $f(t) = \exp(-3t)$ and $g(t) = u(t)$

(continued on page 203)

- "I am doing my Ph.D. in a relatively new university. The PROGRESS workshop helped me to get a better exposure of how things operate in other universities and their culture. It really got me motivated when professors of high reputation spent their time to interact and share their knowledge with early researchers like me."
- "In the panel with faculty who shared their experiences, it became clear that everybody struggles at some time in their academic career. This made me more confident that an academic position is actually within my possibilities."
- "The PROGRESS motivated me a lot, especially because we had a lot of wonderful examples of how a career could be merged and coexist perfectly with the private life of anybody and it should be up to us to decide where is the boundary."
- "Hearing how women have actually been able to combine a career in academia and still have a family is very helpful. In my country, there are almost no women in my field of research in permanent academic positions, so there are not really any role models, and it was very interesting to hear from women around the globe about their experiences."

Via the survvey, the students also suggested topics for future PROGRESS workshops, including a session on preparation of a CV, cover letters, statements, grant writing, a list of opportunities (postdoctoral, faculty, and scholarships), a list of platforms where one could find tools to sharpen signal processing skills, a mentorship program, and a forum for Q&A beyond the workshop.

The next PROGRESS workshop will be virtual and is scheduled for 4–5 June 2021—right before ICASSP 2021. More information can be found at ieeeprogress.org.

Author

Athina Petropulu (athinap@soe.rutgers.edu) is a Distinguished Professor of electrical and computer engineering at Rutgers, the State University of New Jersey, president-elect of the IEEE Signal Processing Society, and director of PROGRESS.

SP

LECTURE NOTES *(continued from page 193)*

via the Leibniz generalized product rule, $(d^n/dt^n)\{f(t)g(t)\} = \Sigma_{k=0}^{n}\binom{n}{k}f^{(n-k)}(t)g^{(k)}(t)$:

$$\begin{aligned} x_p^{(2)}(t) &= f^{(2)}(t)g(t) + 2f^{(1)}(t)g^{(1)}(t) \\ &\quad + f(t)g^{(2)}(t) \\ &= 9\exp(-3t)u(t) + 2(-3\exp \\ &\quad (-3t))\delta(t) + \exp(-3t)\delta^{(1)}(t) \\ &\overset{(a)}{=} 9\exp(-3t)u(t) - 6\delta(t) \\ &\quad + [\delta^{(1)}(t) + 3\delta(t)] \\ &= 9\exp(-3t)u(t) - 3\delta(t) \\ &\quad + \delta^{(1)}(t). \end{aligned} \tag{30}$$

In line (a), the basic and advanced versions of the product rule in Table 1 are applied. The advanced product rule states that $f(t)\delta^{(1)}(t) = f(0)\delta^{(1)}(t) - f^{(1)}(0)\delta(t)$, and substituting $f(t) = \exp(-3t)$ into this relation gives the term in the square brackets of line (a). We see that the final result given by either (29) or (30) matches the one by the partial fraction expansion, provided that we handle the differentiation of $x_p(t)$ in the generalized sense, obeying the rules of Dirac delta function manipulation.

What we have learned

We have studied generalized functions, limits, and derivatives as well as their applications in some signal processing problems. These notes aim to show that many familiar equalities are valid only in the generalized sense. Hence, the equality signs should be replaced with $\overset{(g)}{=}$ in many calculations involving Dirac delta functions, unit step functions, and so on. Interested readers can examine classical signal processing textbooks of Papoulis [3] and Bracewell [4] for a brief treatment of generalized functions. For more information, readers are invited to examine [7], [9], and [10].

Acknowledgments

I would like to thank Prof. Bülent Sankur of Boğaziçi University, Istanbul, Turkey, for many suggestions and comments.

Author

Çağatay Candan (ccandan@metu.edu.tr) received his Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, in 2004. He is a professor in the Department of Electrical and Electronics Engineering, Middle East Technical University, Ankara, TR-06800, Turkey. His research interests include statistical signal processing and its applications in array signal processing and radar/sonar signal processing.

References

[1] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals & Systems*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1999.

[2] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1999.

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

[4] R. Bracewell, *The Fourier Transform & Its Applications*, 3rd ed. New York: McGraw-Hill, 1999.

[5] "Wikipedia contributors, Freeman Dyson," Wikipedia, 2020. Accessed: July 29, 2020. [Online]. Available: https://en.wikipedia.org/wiki/Freeman_ Dyson

[6] Y. Yamamoto, *From Vector Spaces to Function Spaces: Introduction to Functional Analysis with Applications*. Philadelphia: Society for Industrial and Applied Mathematics, 2012.

[7] M. J. Lighthill, *An Introduction to Fourier Analysis and Generalised Functions*. Cambridge, U.K.: Cambridge Univ. Press, 1952 (online republication 2012).

[8] D. G. Luenberger, *Optimization by Vector Space Methods*. Wiley-Interscience, 1969.

[9] G. Temple, "Theories and applications of generalized functions," *J. London Math. Soc.*, vol. s1-28, no. 2, pp. 134–148, 1953. doi: 10.1112/jlms/s1-28.2.134.

[10] G. Temple, "The theory of generalized functions," *Proc. R. Soc. Lond. A*, vol. 228, no. 1173, pp. 175–190, 1955. [Online]. Available: https://www.jstor.org/stable/99615

SP