TransNav
http://www.transnav.eu
the International Journal on Marine Navigation and Safety of Sea Transportation
Volume 15
Number 4
December 2021
DOI: 10.12716/1001.15.04.20

Simple Proof of Incorrectness of the Formula Describing Aliasing and Folding Effects in Spectrum of Sampled Signal in Case of Ideal Signal Sampling

A. Borys
Gdynia Maritime University, Gdynia, Poland

ABSTRACT: A simple proof of the incorrectness of the formula, which is used in the literature nowadays, for description of the aliasing and folding effects in the spectrum of a sampled signal in the case of an ideal signal sampling, is given in this paper. By the way, it is also shown that such the effects cannot occur at all, when the signal sampling is considered to be performed perfectly.

1 INTRODUCTION

It seems that most of the researchers working in the area of signal processing believe that this highly celebrated and commonly used [3–5] expression

$$X_s(f) = \frac{1}{T} \sum_{k=-\infty}^{\infty} X(f - k/T) \qquad (1)$$

for describing the spectrum aliasing and folding effects in the case of an ideal signal sampling is fully correct – despite receiving in [2] a strong evidence that just the opposite might be valid. Here, we present for the skeptics a short and very simple proof of this what has been shown in longer considerations in [2]. We hope that this very transparent proof, which is presented here, will convince them.

In (1), $X(f)$ means the spectrum of an energy, bandlimited signal $x(t)$ with f_m used to denote a maximal frequency present in this spectrum; $X(f - k/T)$ is this spectrum shifted by k/T on the frequency f axis. Further, f_s stands for the sampling frequency used in sampling the signal $x(t)$ in an ideal way, where t is a continuous time variable. Moreover, $T = 1/f_s$, where T is a sampling period satisfying the Nyquist condition $1/T = f_s \geq 2 f_m$ [4]. And, $X_s(f)$ means the spectrum of the signal $x(t)$ sampled ideally (denoted here as $x_s(t)$).

This is another trial of the author of this paper to convince researchers that (1) is not a relevant formula for a correct description of the spectrum of a sampled signal, when the sampling operation is carried out in an ideal way. To do this, a simple proof is constructed, which, however, needs some introductory material. This material has been presented and notation introduced in [2], but at this moment is not available for the reader. Therefore, we start here with presenting it first.

Any sampled signal can be modeled in two ways in the continuous time domain, as illustrated in Fig. 1 (upper curve) and in Fig. 2.

As we see in Fig. 1 (upper curve), a graphical description of the sampled signal, denoted here as $x_T(t)$, consists of a series of the weighted Dirac deltas occurring uniformly on the continuous time axis in the distance of T from each other. And, this is the first way of modeling; it is used in the literature; see, for example, [3–5].

The second possible way of modeling is graphically illustrated in Fig. 2.

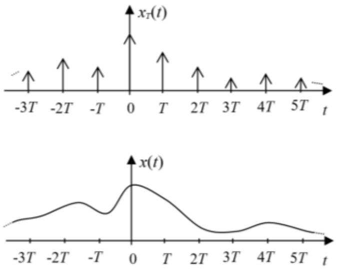

Figure 1. Example sampled signal representation (upper curve) in form of a series of weighted Dirac deltas occurring uniformly on the continuous time axis in the distance of T from each other, and its un-sampled version (lower curve). Here, the sampling operation is assumed to be carried out ideally. Moreover, we note that the figure exploits the same signal, which was also discussed in [1, 2].

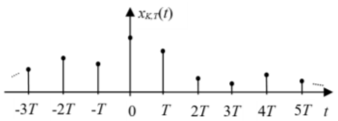

Figure 2. Graphical illustration for a sampled signal representation of the un-sampled signal shown in Fig. 1 (lower curve) in form of a series of time-dependent signal samples occurring uniformly on the continuous time axis in the distance of T from each other. Here, similarly as in Fig. 1, the sampling operation is assumed to be performed ideally.

Now, before going into further details, we would like to underline that both the ways of modeling of the sampled signal, the one illustrated in Fig. 1 (upper curve) as well as the second depicted in Fig. 2, concern an ideal sampling operation (that is a one performed ideally). And, this is important in the considerations presented here; the case of a non-ideal sampling will be reported elsewhere.

The sampled signal in Fig. 2 is denoted as $x_{K,T}(t)$; it is not identical with the signal $x_T(t)$ in Fig. 1. So, these two ways of modeling of the sampled signal are evidently different. Whereas its more natural description (that is a true one) is, obviously, the one presented in a graphical form in Fig. 2. Why? Simply because it consists of a series of true signal sample values occurring at appropriate time instants. And nothing more (on the contrary to the case shown in Fig. 1 (upper curve)).

Thus, in this context, a question of why the sampling signal modeling presented in Fig. 2 is not used in the literature – is a quite legitimate one.

The answer to this question is simple: the signal illustrated in Fig. 2 neither possesses the Fourier transform, nor it can be expanded in a Fourier series. In other words, this signal has no representation in the frequency domain – via a conventional understanding of the signal spectrum. And, this is, obviously, a very serious obstacle for its usage in the signal processing.

However, the people came up with a way of circumventing this. They simply do the following with the signals such as the one shown in Fig. 2: multiply the signal sample values occurring at the appropriate time instants by the time-shifted Dirac deltas (shifted to the appropriate instants of the signal sampling). Thereby, they get such a signal as the one presented in Fig. 1 (upper curve). And, this signal possesses the spectrum; it is expressed by the expression on the right-hand side of (1). That is the spectrum of the ideally sampled signal, $X_s(f)$, is assumed to be equal to $X_T(f)$, where the latter means the Fourier transform of the signal $x_T(t)$. Or, in other words, the spectrum $X_s(f)$ is identified with the spectrum $X_T(f)$. Is this legitimate? The answer to this question is negative in this paper.

By the way, note that in view of what was said above the following formula:

$$X_T(f) = \frac{1}{T}\sum_{k=-\infty}^{\infty} X(f - k/T) \tag{2}$$

is a correct version of the one given by (1).

What we need to formulate our proof of the incorrectness of the formula (1) are the analytical descriptions of the signals $x_s(t)$, $x_T(t)$, and $x_{K,T}(t)$. And, let us start with $x_T(t)$. To this end, see that it can be expressed analytically as a signal $x(t)$ multiplied by the so-called Dirac comb $\delta_T(t)$ [3–5]. That is as

$$x_T(t) = \delta_T(t) \cdot x(t), \tag{3}$$

where the Dirac comb $\delta_T(t)$ is defined as follows:

$$\delta_T(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT) \tag{4}$$

with $\delta(t-kT)$, $k = .,-1,0,1,.,$ meaning the time-shifted Dirac deltas (distributions or impulses) [3–5].

Next, consider the analytical description of the signal $x_s(t)$. Because of the reasons discussed just before, we conclude that in the case of an ideal sampling we have

$$x_s(t) = x_{K,T}(t) \tag{5}$$

Once again, this follows simply from the fact that a true sampled signal looks like the one visualized in Fig. 2, not as the signal depicted in Fig. 1 (upper curve).

The third signal, $x_{K,T}(t)$, which in fact – due to (5) – represents the true sampled signal, can be described analytically as shown in [2] – via the Kronecker functions and the Kronecker comb.

In [2], a basic Kronecker time function $\delta_{0,t/T}(t)$ has been defined as

$$\delta_{0,r} = \delta_{0,t/T} = \begin{cases} 1 & \text{if } \; 0 = r = t/T \; \text{ with } r \text{ defined} \\ & \text{as a real number (or, in other} \\ & \text{words, when this real-valued} \\ & \text{number } r \text{ assumes the integer} \\ & \text{value } 0) \\ 0 & \text{otherwise .} \end{cases} \tag{6}$$

Accordingly, a time-shifted Kronecker time function $\delta_{k,t/T}(t)$ has been also defined in [2] – as the following function:

$$\delta_{k,r}=\delta_{k,t/T}=\begin{cases}1 & \text{if } k=r=t/T \text{ with } r \text{ defined as a real number (or, in other words, when this real-valued number } r \text{ assumes the integer value } k\text{)}\\ 0 & \text{otherwise}.\end{cases} \quad (7)$$

And, note that it follows from (7) that the function $\delta_{k,t/T}(t)$ is a function $\delta_{0,t/T}(t)$ but shifted now on the continuous time axis t by k time units T – to the right if $k > 0$, and to the left when $k < 0$.

Using the time-shifted Kronecker time function $\delta_{k,t/T}(t)$, it is easily to define the so-called Kronecker comb [2]. It is denoted here by $\delta_{K,T}(t)$; and, it is defined as

$$\delta_{K,T}(t)=\sum_{k=-\infty}^{\infty}\delta_{k,t/T}(t) \quad (8)$$

where the first index K at $\delta_{K,T}(t)$ stands for the name of Kronecker, but the second one, T, means a repetition period. This comb is illustrated in Fig. 3.

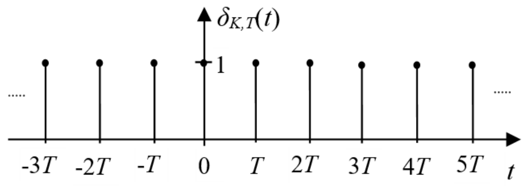

Figure 3. Visualization of the Kronecker comb given analytically by (8).

Note now that using (8) we can describe analytically such a signal $x_{K,T}(t)$ as depicted in Fig. 2 in the following form:

$$x_{K,T}(t)=\sum_{k=-\infty}^{\infty}x(kT)\delta_{k,t/T}(t), \quad (9)$$

where, similarly as before, the first index K at $x_{K,T}(t)$ stands for the name of Kronecker, but the second one, T, means a sampling period.

Further, see that the following:

$$x_{K,T}(t)=\sum_{k=-\infty}^{\infty}x(kT)\delta_{k,t/T}(t)= \\ =\sum_{k=-\infty}^{\infty}\delta_{k,t/T}(t)\cdot x(t)=\delta_{K,T}(t)\cdot x(t) \quad (10)$$

then also holds. Hence, we can write

$$x_{K,T}(t)=\delta_{K,T}(t)\cdot x(t) \quad (11)$$

The remainder of this paper consists of one section. It contains a proof of the incorrectness of the formula (1) that describes the aliasing and folding effects in the spectrum of the sampled signal sampled in an ideal way. Moreover, it is also shown there that the signal $x_s(t)=x_{K,T}(t)$ does possess the spectrum and the following:

$$X_s(f)=X_{K,T}(f)=X(f) \quad (12)$$

holds, instead of (1). In (12), $X_{K,T}(f)$ means the spectrum of the signal $x_{K,T}(t)$.

2 PROOF OF THE INCORRECTNESS OF THE FORMULA (1)

The most important for the proof presented below is to notice that the following:

$$x_{K,T}(t)\cdot\delta_T(t)=x(t)\cdot\delta_T(t). \quad (13)$$

holds. And, what we need in addition here is the assumption of the existence of the spectrum of the signal $x_{K,T}(t)$.

From the previous section, we know that the Fourier transform of $x_{K,T}(t)$ does not exist. However, it does not mean at the same time that its spectrum does not exist, too. Why? Because the spectrum of a signal can be defined in a broader sense; not simply as a (direct) Fourier transform of the signal. And, we use this possibility in this paper.

So, to this end, we define an extended signal spectrum as follows.

Provisional extended definition of signal spectrum. If a signal of a continuous time is represented by an integrable function that possesses a Fourier transform, then the spectrum of this signal is given by the usual Fourier transform. But, when a signal of a continuous time is represented by a non-integrable function which is a train of single values separated uniformly by intervals with all values being identically equal to zero, then its spectrum is defined as a Fourier transform of a function resulting from transforming the train of single values (separated uniformly by intervals with all values being identically equal to zero) to an integrable function that is close (in some sense; a few good measures for defining this can be defined) to this train. Making this provisional definition a precise one will follow from the results presented at the end of this paper.

Note now that the second part of the above signal spectrum definition can be viewed as a generalization of its first part, which builds up an usual signal spectrum definition. And, in this regard, we can see here a very good analogy with the notions of functions and generalized functions (i.e. distributions, as, for example, Dirac delta), where the latter ones are generalizations (in some sense) of the former ones, but still remaining nonconventional objects (when we compare them with ordinary functions).

To see this analogy in more detail, let us start with the following observation: both the Dirac delta as well as the spectrum $X_{K,T}(f)$ of the signal $x_{K,T}(f)$ do not exist in a conventional sense. The first one does not exist as a function; however, nowadays, nobody doubts that it at all exists. And, similarly, we know that $X_{K,T}(f)$ does not exist as an usual Fourier transform. But, it does not mean, at the same time, that this signal spectrum does not exist at all. It exists via the extended definition of the signal spectrum formulated above.

The second observation regards a way of how the Dirac delta and the spectrum $X_{K,T}(f)$ "reveal themselves" in the world of functions and the world of spectra of signals possessing Fourier transforms, respectively. Or how they "cooperate" with these corresponding worlds?

For illustration, let us start with the Dirac delta. And, in what follows, we use its definition that exploits the notion of a functional and the so-called test functions [6, 7]. Further, let us assume that $\varphi(t)$ stands here for a test function [6, 7]. With this, we define the Dirac distribution as such an object (a generalized function) that is characterized by a functional, say, D, which, when applied to any test function $\varphi(t)$ results in $D(\varphi(t)) = \varphi(0)$.

Note that this – with $x(t)$ in place of $\varphi(t)$ – is expressed symbolically in the following way: $\int_{-\infty}^{\infty} x(t)\delta(t)dt = x(0)$ in the signal processing literature (although, it has no strict mathematical meaning [6, 7]). Further, it is worth noting that both the above equations express the so-called sifting property of the Dirac delta. Further, the second equation with $\delta(\cdot)$ in it is used by engineers as a symbolic definition of the operator (functional) D.

So, we see from the above that the Dirac delta "reveals itself" in the world of functions simply through its definition (which is nothing else than its highly celebrated sifting property).

Now, note that we have a similar situation in the case of the spectrum $X_{K,T}(f)$. To see this, let us recall the second part of the extended definition of the signal spectrum formulated before and invoke a corresponding operator, say, R (transforming a non-integrable function of a continuous time (of the type mentioned) into another one, say, $x_r(t)$ that is, however, an integrable function and possesses a Fourier transform) – for performing this task. So, according to the aforementioned definition, we can write

$$x_r(t) = R(x_{K,T}(t)) \tag{14}$$

and

$$\begin{aligned} X_{K,T}(f) &= \mathcal{F}(R(x_{K,T}(t))) = \\ &= \mathcal{F}(x_r(t)) = X_r(f) \,, \end{aligned} \tag{15}$$

where $\mathcal{F}(\cdot)$ stands for the usual Fourier transform. So, through (15), $X_{K,T}(f)$ "reveals itself" in the world of spectra of signals possessing Fourier transforms. Moreover, (15) shows that $X_{K,T}(f)$ exists as a "well-defined" function (in the sense of being integrable) and can be convolved with other spectra (because $X_r(f)$ can).

Note now that using (13) and the above result regarding the existence of $X_{K,T}(f)$ we can write

$$\begin{aligned} &\int_{-\infty}^{\infty} X_{K,T}(\upsilon)\Delta_T(f-\upsilon)d\upsilon = X_T(f) = \\ &= \frac{1}{T}\sum_{k=-\infty}^{\infty} X(f-k/T) \end{aligned} \tag{16}$$

for the frequency domain. In (16), $\Delta_T(f)$ means the Fourier transform of the Dirac comb given by (4); moreover, it is itself a Dirac comb. So, the $\Delta_T(f)$ has the following form [3–5]:

$$\Delta_T(f) = \frac{2\pi}{T}\sum_{k=-\infty}^{\infty} \delta(2\pi(f - kf_s)) \tag{17}$$

In the next step, see that after taking into account (17) in (16) and performing all the needed operations there, we can rewrite (16) in the following form:

$$\begin{aligned} &\frac{1}{T}\sum_{k=-\infty}^{\infty} X_{K,T}(f-k/T) = X_T(f) = \\ &= \frac{1}{T}\sum_{k=-\infty}^{\infty} X(f-k/T)\,. \end{aligned} \tag{18}$$

And, finally, (18) shows that (12) must hold. This also means that $X_s(f) = X_r(f) = X(f)$ must hold. In other words, $X_s(f)$ is not given by (1). Furthermore, it allows to make precise our provisional extended definition of the signal spectrum. Simply, the operator R associated with it must be so chosen that $x_r(t) = x(t)$ holds.

REFERENCES

1. Borys, A.: Filtering Property of Signal Sampling in General and Under-Sampling as a Specific Operation of Filtering Connected with Signal Shaping at the Same Time. International Journal of Electronics and Telecommunications. 66, 3, 589–594 (2020).
2. Borys, A.: Spectrum Aliasing Does not Occur in Case of Ideal Signal Sampling. International Journal of Electronics and Telecommunications. 67, 1, 71–77 (2021).
3. Bracewell, R.: The Fourier Transform & Its Applications. McGraw-Hill Science/Engineering/Math (1999).
4. Marks, R.: Introduction to Shannon Sampling and Interpolation Theory. Springer-Verlag, New York (1991). https://doi.org/10.1007/978-1-4613-9708-3.
5. Oppenheim, A., Schafer, R.: Discrete-Time Signal Processing. Pearson (2009).
6. Osgood, B.G.: EE261 - The Fourier Transform and its Applications. Stanford University, Stanford Engineering Everywhere (2014).
7. Strichartz, R.: A Guide to Distribution Theory and Fourier Transforms. World Scientific Publishing Company (2003).