Sampled Signal Description That Is Used in Calculation of Spectrum of This Signal Needs Revision

Andrzej Borys

Abstract—In this paper, we show why the descriptions of the sampled signal used in calculation of its spectrum, that are used in the literature, are not correct. And this finding applies to both kinds of descriptions: the ones which follow from an idealized way of modelling of the signal sampling operation as well as those which take into account its non-idealities. The correct signal description, that results directly from the way A/D converters work (regardless of their architecture), is presented and dis-cussed here in detail. Many figures included in the text help in its understanding.

Keywords—Sampled signal descriptions used for calculation of its spectrum

I. INTRODUCTION

In the literature, there exists a variety of descriptions of the sampling operation of analog signals and of the sampled signals, which result from the execution of this operation. And the latter is carried out by analog-to-digital (A/D) converters. So all these three things must coincide strictly with each other. But this is not the case in any of the models used nowadays in the literature for description of the sampling process.

The author of this article, in a series of recent papers [1]–[9] drew attention to this problem and analyzed it in detail. He showed that the three things mentioned above are not compatible within any of the models of the signal sampling process that are exploited in textbooks as well as by researchers in their papers.

The papers referred to above are indirect or direct responses to critique of those researchers, who imply that they do not see the aforementioned problem at all, or that everything is fine with the formulas that determine the spectrum of the sampled signal. Here we demonstrate once again, using other tools (primarily non-mathematical ones), that most of their interpretations and justifications are misguided. It might be even said that they are false in some aspects, which leads to confusion and misunderstanding.

II. CRITICAL REVIEW OF THE DESCRIPTIONS USED IN THE LITERATURE

What the literature models and descriptions of the sampled signal are we refer to here? Let us try to list and illustrate them; they are represented successively by Figs. 1–4 below.

The description of a sampled signal as, for example, the signal $X_{D,T}(t)$ in Fig. 1(a), in the form of a sequence of the weighted Dirac impulses, is highly celebrated and commonly used in textbooks and by researchers [11]–[29]. It is an idealized model, as the figure illustrating it consists of point-like objects (weighted Dirac pulses) on the time axis $t$.

![Fig. 1. Illustration of two graphical “ideal” representations of a sequence of samples on the continuous time axis $t$ that are used in the literature. These are the weighted Dirac comb (upper curve) and the weighted Kronecker comb (middle curve), respectively. The former is a distribution (not a function), but the latter represents a not continuous function. Both of them represent the same sequence of samples – of an analog (i.e. un-sampled) signal $x(t)$ shown at the bottom of the figure. But each of them does this a little bit differently. Furthermore, the upper curve is a sampled signal description in form of a series of weighted Dirac deltas (so-called generalized functions of this type) occurring uniformly on the continuous time axis $t$ in the distance of $T$ from each other. Whereas the middle curve is a description built of finite-value time-dependent signal elements (occurring also uniformly on the continuous time axis $t$ in the distance of $T$ from each other). Finally, we remark here that this figure is based on a one, which was used in discussions presented in [1] and [3].](image-url)
The problem with the description visualized in Fig. 1(a) is the use of inappropriate, non-physical objects (Dirac deltas) in its construction that leads to an incorrect formula for the spectrum of the sampled signal. That is to the following formula:

$$X_{D,T}(f) = \frac{1}{T} \sum_{k=-\infty}^{\infty} X(f-k/T),$$

where $X_{D,T}(f)$ and $X(f)$ stand for the Fourier transforms of the signals: $x_{D,T}(t)$ and $x(t)$, respectively, and $f$ means frequency. Obviously, derivation of the formula (1) from the form of the $x_{D,T}(t)$ signal (visualized in Fig. 1(a)) cannot be faulted. Its incorrectness results exclusively from the use of an incorrect form of the description of the sampled signal. And this fact was first pointed out by the author of this paper in [1], who supported his view with detailed analyses presented in the subsequent works [3]–[6].

The first attempt to find an appropriate description of the sampled signal for the purpose of its spectrum calculation was already made in [1], by freeing the previous description from non-physical Dirac deltas. This was a natural step because they do not actually appear at the outputs of manufactured A/D converters (in any of their implementations; for example, see [31]–[37]). It is illustrated here with the example signal $x_{K,T}(t)$ in Fig. 1(b).

As can be seen, the latter description is formed by a sequence of weighted Kronecker functions (these are defined and discussed in detail in [1]). Further, all values of a Kronecker function, except of only one, are identically equal to zero. So, because of this reason, this representation is also called an ideal one (just like the previous one).

Its main disadvantage is that the sampled signal spectrum calculated with its use is identically equal to zero. To circumvent this drawback, a reasonable approach was proposed in [1]. Namely, the signal $x_{K,T}(t)$ was associated there with a closely related one (in some sense – as proposed and explained in detail in [1]), which does possess a spectrum that is not identically a zero function for all frequencies. But the problem with this method is that there exists a certain freedom (arbitrariness, in fact) for the choices of the aforementioned “closely related” functions. This has already been pointed out in [1], and further explained as well as discussed in [3]. Consequently, this approach turned out to be not entirely satisfactory.

In his further investigations, the author of this paper reached out to descriptions of the sampled signal based on the models that describe the sampling process by taking into account also its non-idealities. Such models are available in the literature [12], [19], [25], [38]–[41]. The task was to check whether they can provide some justification for the formula (1) – in an approximate sense, of course.

Let us now illustrate some of them and start with the description presented in Fig. 2.

Note that the not ideally sampled signal of Fig. 2 was called $x_{S,T}(t)$ in [2] because it can be considered as a “smeared" version of the one shown in Fig. 1(b). That is of the signal $x_{K,T}(t)$, in which now each “ideal sample value” is “smeared” on a time segment of the length of $\tau$ seconds. For more details, see [2].

![graphical illustration of a not ideally sampled example analog signal](image)

Fig. 2. Graphical illustration of a not ideally sampled example analog signal of Fig. 1(c), for which the switching time $\tau$ in the A/D converter model was taken to be finite; for more details, see, for example, [38]–[40].

The smearing effect referred to in [2] can be also viewed and modelled a little bit differently. That is as a local blurring of the values of the signal samples multiplied by Dirac deltas (i.e. assuming that they have, prior filtering, an ideal form shown in Fig. 1(a)). As a result, we get then sequences of the blurred weighted Dirac deltas; see, for example, [39] (page 3) and [40] (page 17). Furthermore, a non-immediate switching (i.e. with a finite value of the switching time $\tau$) can be connected with a blurring filtering (of the signal fragment; not of the sample value multiplied by a Dirac delta) hoping to get an even better model for A/D converters. The latter approach is presented in [42]. However, if we think for a moment about the latter method and compare it with the approach presented in [9], it will turn out that, in principle, they do not differ much from each other.

It has been shown in the literature (see, for example, [43] or [44]) that the operation of averaging over consecutively shorter time intervals (or geometric distances) involves the use of Dirac deltas in describing a given phenomenon. For example, this is the case when point masses distributed in a space are described with the use of the mass density function. Then, we can say that the Dirac deltas play a role of local impulse responses for the relationship between the mass density in the space and the point masses in it. Further, one can see an analogy of this scenario with the operation of ideal signal sampling. And, in fact, very many researchers working in the field of digital signal processing understand these matters so. A more extensive discussion and critique of this view is presented in [4].

At this point, we also draw the reader’s attention to the fact that the blurring filtering of the weighted Dirac deltas, applied in [39] and [40], replaces there the local averaging understood in the sense as just presented above. It uses a prescribed form of non-ideal samples having character of Gaussian (or similar) functions (as already mentioned), which, in the limiting case, when their widths go to zero, become Dirac deltas [45]–[47]. For this reason, their action in the type of modelling of the signal sampling process considered now appears to be similar. For more details, see [42].

Modelling of the behavior of A/D converters by means of a local averaging operation has been analyzed in depth in papers [9] and [48]. It has been shown there, among other things, that the (time) sequences of signal samples cannot assume in this case the form of weighted Dirac delta sequences. They always...
take on the form of weighted Kronecker delta sequences, as the one shown in Fig. 1(b), with the coefficients equal to the values of the signal samples (in the limiting case of averaging over a zero time interval) or more or less different from these values (in the case of averaging over a non-zero time interval).

An important conclusion from the analysis of the results obtained in works [9] and [48] is as follows. The sequences of locally averaged signal values, which give the consecutive values (accurate or inaccurate ones) of signal samples (i.e. the values of the sampled signal), and the associated sequences of impulse responses of the filters describing the averaging operations are not identical. These are two different things. And since this is the case, one should not expect that these two sequences will have the same spectra (representations in the frequency domain).

The type of modelling of the sampled signal with the use of averaging operation, which was used in [9] and [48], is illustrated in Fig. 3. (In this figure, \( \tau \) means a time interval of each local signal averaging, but \( c_{a,T}(t) \) and \( x_{a,T}(t) \) are a comb of impulse responses for an equivalent convolution representation of averaging operations performed locally and a resulting sampled signal in this model, respectively.) Note that this model relies upon calculation of the values of the sampled signal for time instants \( kT \), \( k = -\infty, -1, 0, 1, \ldots \), as average values of the signal to be sampled in time intervals from \( kT \) to \( kT + \tau \) with \( \tau = -\infty, -1, 0, 1, \ldots \) or equivalently, as convolutions of the impulse responses shown in Fig. 3(a) with the signal to be sampled. Obviously, these values are, in reality, available only in the modelled converter at time instants \( (kT + \tau) \)'s, however, we assign them to the instants \( kT \)'s. Such a treatment is legitimate of course, since the signal processor behind the D/A converter treats them exactly that way (that is as the ones associated with these time instants, not with the \( (k + 1)T \)'s).

The remaining values of the sampled signal on the time axis are assumed to be zeros (per definition) in this model.

It is worth noting that the form of the signals \( x_{a,T}(t) \) of Fig. 3(b) and \( x_{K,T}(t) \) of Fig. 1(b) is the same. The only difference between them is that in the averaging model the values of the signal samples differ from those shown in Fig. 1(b). These differences are visualized in Fig. 3(b) by showing both the perfect values, which are marked there with the use of horizontal dashes (and being equal to sample values in Fig. 1(b)), and the averaged ones marked with black circles.

It is also worth noting that the sampled signal description illustrated in Fig. 3, which belongs to the category of non-ideal ones, can be made an ideal one by assuming the averaging interval \( \tau \to 0 \) (that is \( \tau = 0 \) in the limit). For this case, Fig. 3 redraws to Fig. 4 with \( c_{a,0,T}(t) = \delta_{\tau}(t) \), where \( \delta_{\tau}(t) \) means the so-called Dirac comb [1], but \( x_{a,T}(t) \to x_{I,T}(t) = x_{K,T}(t) \) assumes the form shown in Fig. 1(b).

Finally, in this section, let us interpret in terms of the last model presented here what is done when calculating the spectrum of the sampled signal having the example form as shown in Fig. 1(a) (that leads to the formula (1)). Simply, then, one takes a weighted comb of impulse responses involved in convolution representations of “ideal” averaging operations performed locally (as that in Fig. 4(a)) and calculates its spectrum. Whether this obtained spectrum can be identified with the spectrum of the sampled signal is doubtful.

### III. Basic Cause of Failure of Spectrum Calculations with the Use of Hitherto Descriptions of Sampled Signal

The author of this paper has shown in a series of publications [1]–[6] that the formula (1), which determines the spectrum of a sampled signal, is not correct. And that the reason for this lies in an incorrect description in time of the sampled signal by means of a weighted Dirac comb (such an example signal is shown in Fig. 1(a)). In the same publications and further ones [7]–[10], he attempted to find a correct description of the sampled signal for the purpose of calculating its spectrum, analyzing for this purpose, in detail, all the other descriptions available in the literature. Unfortunately, none of these attempts yielded a satisfactory result. For what reason?
The reason for all the aforementioned failures is the same: an incorrect description in time of the sampled signal.

Let us now take a closer look at this issue. And to this end, observe that in each of the descriptions discussed in Section II artificially forced values of the output voltages of modelled A/D converters equal to zero occur (i.e., having such values by definition), in each interval between the sampling instants \( kT \) and \( (k+1)T \), \( k = -\infty, 0, 1, \ldots \). Note also that these smaller or larger gaps in the aforementioned intervals filled with forced zeros can be also understood (in another interpretation) as sets of unknown values; for more details on this, see [5]. Obviously, both the former and the latter of the above interpretations do not correspond with a true picture of how a real A/D converter works.

At this point, it is worth recalling that in order to correctly calculate the spectrum of the output signal of an A/D converter with the use of the Fourier transform, we need to know precisely this signal as a function, for all time points (except maybe of a countable number of discontinuity points). But, as mentioned above, none of the descriptions discussed in Section II provided us with such a function. So it is not surprising that all attempts to calculate the spectrum of the sampled signal with their use proved unsatisfactory.

The next remark concerns common but erroneous identification with each other of two de facto different functions, or an incorrect recovering one of these functions from another. We mean here the following functions: the function of time describing the actual (true) voltage waveform at the output of an A/D converter, which we denote as \( x_a(t) \) in what follows (where the lower subscript \( a \) comes from the word “actual”), and the function \( x[k], k = -\infty, -1, 0, 1, \ldots \) that expresses the relationship between the values of the signal samples and the values of time indices of the instants assigned to them. However, because of a lack of space here discussion of this point will be postponed to the next paper.

An analysis of the various architectures of A/D converters and their technological realizations [31]-[37] allows us to distinguish two time segments in the \( x_a(t) \) waveform in each of its intervals between the sampling instants \( kT \) and \( (k+1)T \), \( k = -\infty, 0, 1, \ldots \). So describing the \( x_a(t) \) waveform, one can make the first of these segments "responsible" for all transients associated with the operation of (ideal or non-ideal) switching, but the second for the process of quantizing the signal sample value and maintaining this quantized value until the end of the corresponding sampling interval. And modelling this, we can say that in the first of the time segments mentioned above of the length \( \tau \) we perform signal averaging (i.e. in this switching time interval), while in the second of the length \( T - \tau \) we quantize the amplitude value of the averaged signal, i.e. we get \([x_{a,t}(kT)]_Q\), where the lower index \( Q \) stands for the operation of amplitude quantization. This value is maintained until the end of the interval ending at the instant \((k+1)T\). So the waveform \( x_a(t) \) has the form as the one shown in Fig. 5.

This is a staircase curve with transients taking place during an non-ideal transition from one discrete value of the signal to another, which is modelled with the use of an averaging operation. Further, note that if the averaging times denoted by \( \tau \)'s in Fig. 5 go to zero, the aforementioned staircase curve assumes in the limit the form shown in Fig. 6(a), not that one visualized in Fig. 1(b).

In other words, the description of the voltage waveform at the output of the A/D converter as illustrated in Fig. 6(a) – for the case of not taking into account the non-idealities of the sampling process – represents the true form of the output waveform at the converter and, as such, must be used in calculations of its spectrum (no other one, such as, for example, the one of Fig. 1(a) or that in Fig. 1(b)). We recall also here that \([x_{a,t}(kT)]_Q\)'s shown in Fig. 6(a) denote \([x_{a,t}(kT)]_Q\)'s for \( \tau = 0 \) and equal the quantized values of the signal samples, according to \( [x(kT)]_Q = [x(k)]_Q \).

Now, let us make a very interesting observation, namely that the differentiation of the function \( x_a(t) \) results in the waveform containing Dirac deltas as shown in Fig. 6(b), and the latter is identical in form with that in Fig. 1(a). So we come to a contradiction because the output signal of an A/D converter cannot have the form shown in Fig. 1(a) as well as in Fig. 6(a) at the same time. Against accepting the form shown in Fig. 1(a) – as the form of the output signal of an A/D converter – is the fact that in none of the architectures and realizations of A/D converters [31]-[37] occurs circuit elements that implement the operation of differentiation.
IV. CONCLUSIONS

The reason for the failure to calculate correctly the spectrum of a sampled signal is explained in this paper in an accessible way. In such calculations, the description of this signal as illustrated in Fig. 6(a) must be used, not the commonly exploited description of Fig. 1(a) containing the Dirac deltas.

In this context, we note that calculation of the Fourier transform of a waveform such as the one presented in Fig. 6(a) results in

\[
X_\alpha(f) = \sin(\pi T f) \exp(-\pi T f) \cdot \sum_{k=\infty}^{\infty} X(f-kf_s) ,
\]

see [49] for details of calculations, where \(X_\alpha(f)\) and \(X(f)\) are the Fourier transforms of the signal \(x_\alpha(t)\) (see Fig. 6(a)) and of its un-sampled version \(x(t)\), respectively. Further, \(f_s = 1/T\) means the sampling frequency (rate), \(k\) belongs to the set of integers, and \(j = \sqrt{-1}\). Moreover, the function \(\sin(x)\) in (2) is given by

\[
\sin(x) = \begin{cases} 
\sin(x) & \text{for } |x| \neq 0 \\
1 & \text{for } x = 0
\end{cases}
\]

Evidently, (2) differs from (1). The detailed analysis and discussions of descriptions used for sampled signals, conducted in this paper, show that the corrected version of (1), i.e. the formula given by (2), should be used as a proper formula for the sampled signal spectrum.

REFERENCES


