

Shannon-Nyquist-like condition derived from other principles

Andrzej Borys

Abstract—In this paper, we report on two important results within the mathematical theory of analog signal sampling that we obtained recently. First, we demonstrate that in order for a comb of impulses to correctly describe a sampled signal in the domain of a continuous time, the pulses of which it is composed must be pulses of a finite duration. Second, it is shown that a condition similar to the well-known Shannon-Nyquist sampling criterion for a perfect signal recovery can be derived using some other principles.

Keywords—The best mathematical understanding of signal sampling operation; descriptions of sampled signals; interpolation and reconstruction; Shannon-Nyquist-like condition for good signal reconstruction

I. INTRODUCTION

IN a number of his recent publications [1], [2], [3], [4], [5], [6], the author of this paper has shown how the way we conceive of and mathematically describe “analog signal samples and their positions” over time is important for understanding and perceiving the operation of recovering an analog signal from its samples. And a concept called space-time facilitates the understanding of this idea and this approach to the problem; this space, for signals that are one-dimensional functions of time, has been called amplitude-time space in [7]. In this paper, we restrict ourselves to considering only such signals.

Moreover, when we are talking in this paper about time we mean exclusively a physical time, represented here by a variable $t \in \mathbb{R}$, where $\mathbb{R}$ stands for the set of real numbers. Also we call it a continuous time t . In what follows, we abstract from the so-called discrete time, i.e. the set of indices k 's of the time instants of sampling an analog signal (which takes place in the physical time t) - ordered in an ascending order. The indices k 's belong to the set $\mathbb{Z}$ of integers. Further, we do not need any relationship between these times in our considerations, which follow.

As we already know [1], [2], [3], [4], [5], [6], there is no single analytical description and accompanying graph illustrating the sampled signal in the amplitude-time space. In the literature, three such descriptions and their graphical representations can be found that are used for an ideally sampled signal, i.e. a one taking into account only the essence of the sampling operation.

The most common of them is a description of the sampled signal via a weighted Dirac comb [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21]. Denote it here by $x_{D,T}(t)$. It has the following analytical form:

$$x_{D,T}(t) = \sum_{k=-\infty}^{\infty} x(kT)\delta(t - kT), \quad (1)$$

where $\delta(t - kT)$'s mean the time-shifted Dirac deltas, the sampling period is denoted by T , and $x(kT)$'s (being the weights of the $\delta(t - kT)$'s in (1)) are the samples of an analog signal $x(t)$, which has been sampled.

Another possible description of the sampled signal is a one in form of a weighted comb of the time-shifted Kronecker functions [1]. This comb is more natural than that given by (1) because it does not apply distributions, which, as we know, are Dirac deltas, but instead of them uses functions - precisely Kronecker functions. More details on this theme are provided in [1].

An analytical form of the sampled signal description via the comb of Kronecker functions, and denoted here as $x_{K,T}(t)$, is given by [1]

$$x_{K,T}(t) = \sum_{k=-\infty}^{\infty} x(kT)\delta_{k,t/T}(t), \quad (2)$$

where $\delta_{k,t/T}(t)$'s stand for the time-shifted Kronecker functions defined in [1] via a standard Kronecker δ_{ij} symbol, which is extended here to a function of time t as

$$\delta_{k,t/T} = 1 \text{ if } k = t/T \text{ and } 0 \text{ otherwise.} \quad (3)$$

In (3), the first subscript k belongs to the set of integers, but the second, t/T , to the set of real numbers.

The third possible description of the sampled signal exploits rectangular functions. It assumes the form of a weighted step function [3], [6]. And analytically, we denote it as $x_{STEP}(t)$ and express in the following form:

$$x_{STEP}(t) = \sum_{k=-\infty}^{\infty} x(kT)\text{rect}(t - kT). \quad (4)$$

The function $\text{rect}(\cdot)$ in (4) means a rectangular function

$$\text{rect}(t) = 1 \text{ for } t \in \langle 0, T \rangle \text{ and } 0 \text{ otherwise.} \quad (5)$$

Comparison of the formulas given by (1), (2) and (4) shows that these descriptions of the sampled signal are not identical. Evidently, they are different. In this paper, we explain what a role do they play in a procedure of recovery of an analog signal from its sampled version, where the signal recovery we

understand as an interpolation of the former one from its discrete samples.

Possibly, they play no role. But we explain this point graphically and analytically in more detail in the next section of the paper.

The third section of the paper is devoted to showing that a condition similar to the Nyquist condition [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21] specifying the minimum value of the sampling rate so that an analog signal can be recovered from its samples can be derived on another ground, in which the Whittaker interpolation formula [8] is not used. The paper ends with a conclusion.

II. DESCRIPTIONS OF SAMPLED SIGNAL AND THEIR RELATION TO INTERPOLATION

Let us start this section by showing that the first two of the three representations of the sampled signal, which are listed in Introduction, are “non-physical” descriptions (let us call them so here), while the third one is “physical.”

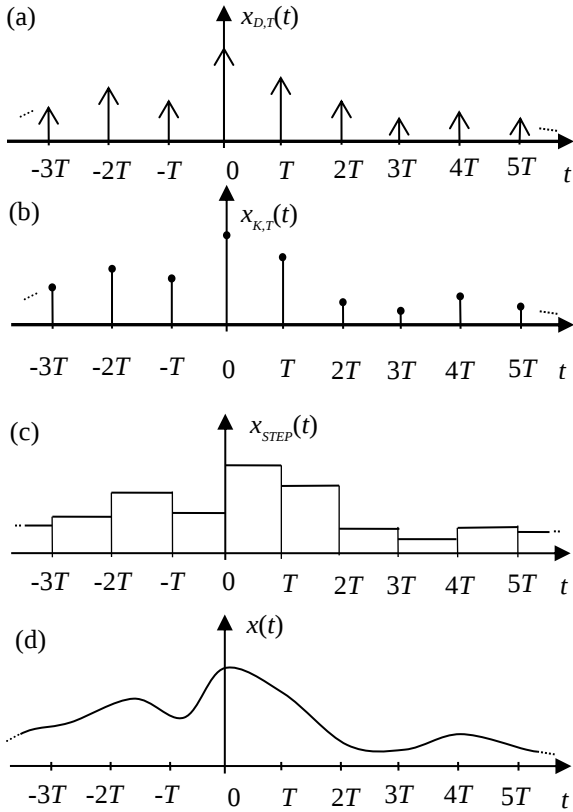


Fig. 1. Visualization of sampled signal descriptions for an example analog signal shown in (d): curves in (a), (b), and (c) represent, respectively, the formulas (1), (2), and (4). The figure is a compilation of figures taken from [3].

In what sense? We are explaining this point in what follows. And to this end, we are using also visualizations of the formulas (1), (2) and (4) that are widely exploited in the literature [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], see their example illustration shown in Fig. 1.

The visualizations shown in Fig. 1(a) and Fig. 1(b) are specific. Let us start with the first one. Here it should be emphasized that it is purely illustrative, since the comb of

arrows, which stands in Fig. 1(a) for the sum of Dirac distributions, remains still a distribution, that is, an object which is not a function. Moreover, such a distribution, just like any single Dirac delta, does not possess a graphical representation [22], [23]. At most, it can be related to an illustrative image, as was done in Fig. 1(a).

Of course, the description by (1) could have been left without any graphical interpretation, but the prevailing view was that it would be better, for illustrative purposes, for a sampled signal to be also presented in form of a graph (as is the case with any other signal - expressing changes of a physical quantity over time). And thereby, it will be possible to use it for enrichment of analytical derivations, compact and not illustrated, which we have full of in research papers and textbooks on the digital signal processing. And so it happened. It was assumed that it was better to have some graphical model of the sampled signal, although defective, but always better than none.

However, it was not taken into account that a defective graphic model can also cause troubles. To this end, consider for example the arrows in Fig. 1(a). They mean infinite values at the time instants in which they occur, and the zero values everywhere outside of them. What do we do to improve so clearly non-physical interpretation of the arrows? We say that the infinities associated with them have meaning of the infinities multiplied by the values of analog signal samples. However, they remain still infinities. But we are interpreting them as the infinities of other kind. And we distinguish between these kinds mentioned. However, this is not obviously a neat, but rather artificial approach.

An example of such entanglement in contradictions related to the model of the sampled signal presented in Fig. 1(a) is a lecture on this subject in the textbook [9]. On the one hand, the authors of this textbook accept the model from Fig. 1(a) as convenient for the use in the theory of digital signal processing, but at the same time write that “the punctual value of a distribution is only an abstraction”. And the latter is, of course, in obvious contradiction to the visualization in Fig. 1(a). From the latter, it would follow that at the same time we would be dealing with some kind of “dissolution” of the signal samples on the timeline $0t$. But this is undeniably incompatible with a physical understanding of the signal sampling operation.

Concluding, we may say that for the reasons given above, we rightly call the description of the sampled signal according to (1) a non-physical, and therefore reject it.

Now, with regard to the visualization of $x_{K,T}(t)$ in Fig. 1(b), let us note that it seems to be fully correct. This so because $x_{K,T}(t)$ is a function and Fig. 1(b) represents just the function graph. Despite this, however, the description (2) cannot be called a physical one because of one important reason. Namely, because that the “columns” terminated with dots in the graph of Fig. 1(b) have zero-th widths. So, assuming that these widths are the same and denoting each of them by τ , we write $\tau = 0$ for the case shown in Fig. 1(b).

Now, to demonstrate the non-physicality of the visualization in Fig. 1(b), let us use the following thought experiment. The columns ending with dots in Fig. 1(b) represent successive samples of an analog signal, and since they have zero widths, or “times of their duration,” the total duration time of all these samples has, because of this reason, also a zero value. And this goes against common sense, which means that it is a non-physical situation. So we conclude from this that the parameter

τ defined above must always be a positive quantity. In general, we formulate this condition as $T \geq \tau > 0$.

To strengthen the above argument, let us use another one, taken from the theory of digital data transmission. Here we say that a transmitted symbol is observed at a given point of a channel a certain finite time that is equal to the inverse of the transmission rate (expressed in symbols per second). And this is obviously a positive quantity. Call it a “symbol duration” and denote by τ ; it is here equal to the timing (clock) time. Further, let us note that the situation thus described can be compared with the sampling of analog signals, when the value of the sample lasts for all the time between the successive instants of sampling, i.e. when $\tau = T$ holds. Assuming $\tau = T = 0$ would lead to the nonsense in the case of data transmission. Hence, together with what has been said above, it follows that this would also mean the nonsense in the case of the signal sampling.

Moreover, it has been shown in [24] that a key condition for a correct determination of the spectrum of a sampled signal is assumption of $T \geq \tau > 0$ in any realistic description of this signal. That is the parameter τ must always be taken as a non-zero positive value.

In this paper, the descriptions of the sampled signal, in which $T \geq \tau > 0$ holds, are called “physical” ones. One of them is the description expressed analytically by (4) and illustrated in Fig. 1(c). This abstract mathematical model was derived in [3] and [5] from strictly physical premises. For simplicity, it is assumed $\tau = T$ in it. Further, note that there are also applications, as, for example, analog-to-digital conversion in microwave photonic systems, where we have to do with its version with the parameter τ smaller than the sampling period T and with the in-between impulses differing from the rectangular ones [25], [26], [27]. This variant is illustrated in Fig. 2, where the corresponding comb of impulses representing “optical” samples is denoted by $x_{s,T}(t)$.

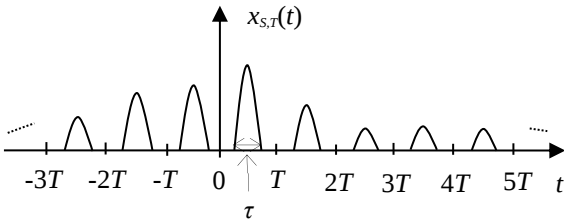


Fig. 2. A schematic illustration of a signal at the output of an electro-optic amplitude modulator combining a stream of ultrashort optical pulses with a radio frequency signal in a device presented in [26] for performing analog-to-digital conversion in a microwave photonic system.

Comparing the pulse combs shown in Fig. 1(a)–(c) with that in Fig. 2, we see that the latter is moved to the right by $\tau/2$ relative to the others. However, this is only of formal importance. Furthermore, note that the graphic of Fig. 2 refers to the signal $x(t)$ of Fig. 1(d) assumed to be here a RF signal.

It is worth noting now that all four combs illustrated by Fig. 1(a)–(c) and Fig. 2 are analytically nothing else but orthogonal expansions. That is such expansions in which the impulses that make them up are orthogonal to each other. And the latter can be converted into the following form: a coefficient times an

orthonormal pulse. So, finally, the coefficients thus obtained are the values of the signal samples of a signal to be sampled.

In other words, it follows from the above that each of the above-mentioned combs can be perceived as a kind of “placement” of the matrix of the sample values on the comb of the orthonormal pulses shifted between each other. For our example of Fig. 1, this matrix is presented in Fig. 3.

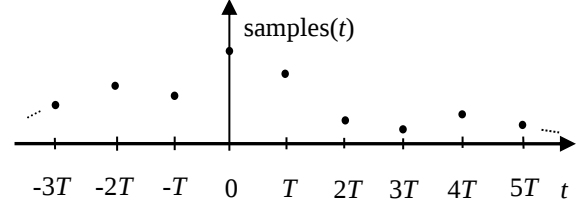


Fig. 3. Visualization of the matrix of sample values of our example signal shown in Fig. 1(d), for the combs illustrated in Fig. 1(a)–(c).

Note also that the matrix of sample values for the weighted signal comb shown in Fig. 2 will be shifted on the timeline $0t$ to the right by $T/2$ and will consist of the sample values of the analog signal $x(t)$ of Fig. 1(d) taken at the successive time instants $kT + T/2$, where $k \in \mathbb{Z}$.

Let us now examine the relationships (if any) of the descriptions (1), (2), (4), and the one visualized in Fig. 2 with the interpolation between values of the signal samples of $x(t)$, which we perform in its recovery. To this end, we remind that the starting point of any interpolation of a one-dimensional function is the knowledge of the matrix of sample values through which this function passes. For our example function $x(t)$ of Fig. 1(d), it is illustrated graphically in Fig. 3. Next, in the interpolation procedure, we fill in the spaces between the points shown in Fig. 3 according to some rule, arriving finally at an interpolated function. Let us call it $\tilde{x}(t)$. And, observe also that any of the rules we used to obtain the weighted combs illustrated in Fig. 1(a)–(c) and Fig. 2 seems to be appropriate for performing such an interpolation.

However, we get a good interpolation, in the sense formulated above, only in one case. We show this in what follows. In the case of combining the matrix of signal sample values with the comb of shifted Dirac deltas, we arrive at such a situation where the values of the above matrix for the interpolation points become non-determined. Therefore, the Dirac weighted comb description cannot be considered at all as an interpolated signal $\tilde{x}(t)$. Further, the description using a weighted comb of shifted Kronecker functions cannot be considered a useful interpolation, since all values between interpolation points are identically equal to zero. Our third description can be viewed, as said above, as a kind of “placement” of the matrix of the sample values on the comb of the orthonormal rectangular pulses that results in a weighted step function (4). It is visualized in Fig. 1(c), where we see the successive interpolation points connected with each other through the line segments parallel to the timeline $0t$. So this is a true interpolation. Its name in the signal processing literature [9], [13], [17] is a zero-order hold. And finally, we see that in the case visualized in Fig. 2 the samples occurring at the successive time instants $kT + T/2$, where $k \in \mathbb{Z}$, are connected to each other by a saddle curve with zero values in its middle.

So we have to do here with an interpolated curve, $\tilde{x}(t)$, but its course is obviously significantly different from the waveform $x(t)$.

Concluding this section, we can say that the description (2) has a naturally “built-in” good interpolation procedure for handling the signal samples. Therefore, for appropriately small values of the sampling period T there is no need to use a lowpass filter in the process of recovering the $x(t)$ signal. Then the interpolation errors are smaller than the amplitude quantization errors.

III. DERIVATION OF SHANNON-NYQUIST-LIKE CONDITION FROM OTHER PRINCIPLES

With this last observation in mind, let us now try to derive a condition for the necessary value of the sampling rate (i.e. $f_s = 1/T$), which allows to assume that the description given by (4) can also be taken as a description of the signal $x(t)$, with an error not greater than the quantization error. To obtain this, we use here an intermediate result that was derived in [28], where it was proposed to model measurement processes in the same way as the sampling operation of analog signals is modeled. For details, the interested reader is referred to [28]. Here we use the following inequality:

$$\max|\Delta x| < SR/f_s \quad (6)$$

derived there. In (6), $\max|\Delta x|$ means the maximum of the absolute value of the signal amplitude change occurring between adjacent samples. Whereas SR stands for the so-called parameter slew-rate of the signal $x(t)$.

Let us denote now the quantization error by Q and demand that the following: $SR/f_s \leq Q$ is satisfied. So if the latter is fulfilled, we have at the same time the condition $\max|\Delta x| < Q$ fulfilled, too.

In the next step, we analyze the condition $SR/f_s \leq Q$ for lowpass signals. We illustrate this case on an example of the ideal lowpass signal $x_{lp}(t) = X_0 \text{sinc}(2\pi f_m t)$, where X_0 stands for its maximal amplitude, f_m is the maximal frequency present in its spectrum, and the function $\text{sinc}(2\pi f_m t) = \sin(2\pi f_m t)/2\pi f_m t$. With this, we can write

$$\left| \frac{dx_{lp}(t)}{dt} \right| = X_0 \left| \frac{d\text{sinc}(z)}{dz} \right| 2\pi f_m \leq X_0 2\pi f_m \cdot \frac{1}{2} = \pi f_m X_0 \quad (7)$$

for our $x_{lp}(t)$, where an auxiliary variable $z = 2\pi f_m t$ was used, and the following result:

$$|\sin(z)/z| \leq 1/2, \quad (8)$$

taken from [29], was applied. Finally, it follows from (7) that the inequality

$$SR \leq \pi f_m X_0 \quad (9)$$

holds for the parameter SR .

In the next step, we require that the right-hand side of the latter inequality also be bounded by the quantization error. (This is certainly a reasonable assumption.) We do this by writing

$$SR/f_s \leq \pi X_0 f_m / f_s \leq Q, \quad (10)$$

which implies that the inequality $SR/f_s \leq Q$ is also satisfied.

Now, consider the inequality $\pi X_0 f_m / f_s \leq Q$ in more detail. To this end, assume that the resolution of an analyzed A/D converter equals r bits, which means that it possesses 2^r discrete levels. The quantization step, i.e. the difference between the nearest levels, is Q . Further, assume that this converter operates at the maximum drive, i.e. half of its 2^r levels cover the maximum positive signal value, i.e. X_0 . And taking all this into account in our inequality, we arrive at

$$f_s \geq \pi 2^{r-1} f_m, \quad (11)$$

which we call here a Shannon-Nyquist-like condition derived from other principles.

In what follows, due to lack of space, we present only a few comments regarding (11). First of all, see that when (11) is satisfied then the Shannon-Nyquist condition $f_s \geq 2f_m$ is also satisfied, automatically. This is true even for $r = 1$, i.e., with only two levels of quantization (because $\pi > 2$). Second, (11) does not depend on the value of the quantization step Q , what is quite clear since the changes in amplitude on both sides of the inequality were expressed (in our derivations) in terms of this measure. Third, the higher the resolution, the higher the required sampling frequency, which at first glance seems completely incomprehensible. But, remember that with a constant value of X_0 , the higher the resolution, the smaller the quantization error $Q = X_0/2^r$, and therefore it is necessary to take samples more frequently at a given SR value of the signal.

Note finally that when the resolution r increases, the quantization step Q decreases and it may become smaller than the thermal noise level at some point. Taking the above into account leads to a modification of the previous formula to

$$f_s \geq \pi 2^{r-1} f_m Q(r)/V_n, \quad (12)$$

where V_n means the RMS noise voltage. The dependence of the quantization step Q on the resolution r is also explicitly shown in (12).

IV. CONCLUSIONS AND REMARKS

The classic Shannon-Nyquist condition for recovery of an analog signal from its samples applies only to the case, where these samples are not considered to be quantized in amplitude. The latter issue is treated in the literature [30] separately. But the criterion derived here applies to both cases: quantization in time and quantization in amplitude. Therefore, it has a more comprehensive character.

This paper shows also that there occur many ambiguities in the mathematical modeling of the sampling operation of analog signals and that this can lead to not understandable outcomes. For example, the assumption, which is automatically made in a model with abstract objects like Dirac deltas, is that each analog signal sample is generated at zeroth time. Here, it is shown that this is false and can really lead to ambiguous results. The author of this paper writes about such things in very detail in his work [31].

REFERENCES

- [1] A. Borys, "Extended definitions of spectrum of a sampled signal," *Intl Journal of Electronics and Telecommunications*, vol. 67, no. 3, pp. 395-401, 2021.
- [2] A. Borys, "Another proof of ambiguity of the formula describing aliasing and folding effects in spectrum of sampled signal," *Intl Journal of Electronics and Telecommunications*, vol. 68, no. 1, pp. 115-122, 2022.
- [3] A. Borys, "Sampled signal description that is used in calculation of spectrum of this signal needs revision," *Intl Journal of Electronics and Telecommunications*, vol. 69, no. 2, pp. 319-324, 2023.
- [4] A. Borys, "On modelling and description of the output signal of a sampling device," *Archives of Electrical Engineering*, vol. 72, no. 1, pp. 147-155, 2023.
- [5] A. Borys, "A new result on description of the spectrum of a sampled signal," *Intl Journal on Marine Navigation and Safety of Sea Transportation*, vol. 18, no. 2, pp. 393-397, 2024.
- [6] A. Borys, "Revision of the formula describing the spectrum of output signals at A/D converters," *Intl Journal of Electronics and Telecommunications*, vol. 70, no. 1, pp. 183-190, 2024.
- [7] A. Borys, "Some topological aspects of sampling theorem and reconstruction formula," *Intl Journal of Electronics and Telecommunications*, vol. 66, no. 2, pp. 301-307, 2020. Erratum due to typographical errors available from the author after writing him an e-mail.
- [8] M. Vetterli, J. Kovacevic, and V. K. Goyal, *Foundations of Signal Processing*. Cambridge, England: Cambridge University Press, 2014.
- [9] P. Prandoni and M. Vetterli, *Signal Processing for Communications*. Lausanne: EPFL Press (a Swiss academic publisher distributed by CRC Press), 2008.
- [10] R. N. Bracewell, *The Fourier Transform and Its Applications*. New York: McGraw-Hill, 1965.
- [11] B. Gold and C. M. Rader, *Digital Processing of Signals*. New York: McGraw-Hill, 1969.
- [12] S. K. Mitra, *An Introduction to Digital and Analog Integrated Circuits, and Applications*. New York: Harper and Row, 1980.
- [13] A. V. Oppenheim and A. S. Willsky, *Signals and Systems*. Englewood Cliffs NJ: Prentice Hall, 1983.
- [14] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-Time Signal Processing*. New Jersey: Prentice Hall, 1998.
- [15] R. J. Marks II, *Introduction to Shannon Sampling and Interpolation Theory*. New York: Springer-Verlag, 1991.
- [16] C. Gasquet and P. Witomski, *Fourier Analysis and Applications: Filtering, Numerical Computation, Wavelets*. Berlin: Springer-Verlag, 1998.
- [17] V. K. Ingle and J. G. Proakis, *Digital Signal Processing Using Matlab*. Stamford, CT, USA: Cengage Learning, 2012.
- [18] R. Brigola, *Fourier-Analysis und Distributionen*. Hamburg: edition swk, 2013.
- [19] J. H. McClellan, R. Schaffer, and M. Yoder, *DSP First*. London, England: Pearson, 2015.
- [20] K. B. Howell, *Principles of Fourier Analysis*. Boca Raton: Chapman and Hall, 2001.
- [21] P. A. Lynn, Chapter 8: *Modulation and Sampling in An Introduction to the Analysis and Processing of Signals*. London: Palgrave, 1982.
- [22] R. F. Hoskins, *Delta Functions: Introduction to Generalised Functions*. Oxford: Woodhead Publishing, 2010.
- [23] R. Strichartz, *A Guide to Distribution Theory and Fourier Transforms*, Boca Raton: CRC Press, 1994.
- [24] A. Borys, "Spectrum aliasing does occur only in case of non-ideal signal sampling," *Intl Journal of Electronics and Telecommunications*, vol. 67, no. 1, pp. 79-85, 2021.
- [25] J. Qi, X. Chen, M. Fu, Ho. Zhang, W. Yi, T. Xu, D. Su, Hu. Zhang, X. Wei, B. Shi, and X. Li "Effects of optical sampling pulse power, RF power, and electronic back-end bandwidth on the performance of photonic analog-to-digital converter," *Micromachines MDPI*, vol. 14, no. 12, pp. 1-12, 2023.
- [26] P. Callahan, M. Dennis, and T. Clark, "Photonic analog-to-digital conversion," *Johns Hopkins Apl Technical Digest*, vol. 30, no. 4, pp. 280-286, 2012.
- [27] A. Borys, "On behavior of sampled signal when value of sampling period tends to zero and analog-to-digital conversion in microwave photonic systems," *Intl Journal of Electronics and Telecommunications*, vol. 70, no. 4, pp. 1057-1062, 2024.
- [28] A. Borys, "Further discussion on modeling of measuring process via sampling of signals," *Intl Journal of Electronics and Telecommunications*, vol. 66, no. 3, pp. 507-513, 2020.
- [29] T. H. Gronwall, Otto Dunkel, and H. S. Uhler, "Solutions of problems," *The American Mathematical Monthly*, vol. 27, no. 2, pp. 81-85, 1920.
- [30] B. Widrow, "Statistical analysis of amplitude-quantized sampled-data systems," *Trans. AIEE, Pt. II*, vol. 79, pp. 555-568, 1961.
- [31] A. Borys, "Descriptions of digital signals and occurrence of multivaluedness in their use in the area of Electrical Engineering," *Archives of Electrical Engineering*, submitted for publication, February 2026.