

Review

Sampling Theorems for Stochastic Signals. Appraisal of Paul L. Butzer's Work

Tibor K. Pogány ^{1,2} 

¹ Faculty of Maritime Studies, University of Rijeka, Studentska 2, 51000 Rijeka, Croatia; poganj@pfri.com

² Institute of Applied Mathematics, Óbuda University, Bécsi út 96/b, 1034 Budapest, Hungary

Received: 29 June 2019; Accepted: 29 July 2019; Published: 1 August 2019



Abstract: The sampling reconstruction theory is one of the great areas of the analysis in which Paul Leo Butzer earned longstanding and excellent theoretical results. Thus, we are forced either by earlier exhaustive presentations of his research activity and/or the highly voluminous material to restrict ourselves to a more narrow and precise sub-area in consideration; we discuss here, giving deeper insight, Paul Butzer's sampling theoretical work with special attention concerning sampling stochastic signals.

Keywords: WKS sampling theorems; weak sense stationary stochastic processes; stochastic signal; bandlimited processes; Karhunen-Cramér theorem; correlation function; spectral representations; mean-square sampling reconstruction

MSC: Primary: 94A20, 01A60, 01A65; Secondary: 41A05, 26D15

1. Introduction

Let f be expressible in a domain D in the form

$$f(x) = \sum_{n \in \mathbb{Z}} f(\lambda_n) S_n(x), \quad x \in D, \quad (1)$$

where $(\lambda_n) \subseteq D$ and $(S_n(x) \equiv S_n(x, \lambda_n))$ is some convenient function class. This type of presentation is the *sampling series* (of f), since f is restored by its values sampled in the discrete subset $(\lambda_n) \subset D$ by (in general) an infinite linear combination. In the 'classical case', when $D \equiv \mathbb{R}$, $\lambda_n = n\pi/w$, $w > 0$ and $S_n(x) = (wx - n\pi)^{-1} \sin(wx - n\pi)$ we arrive at the so-called WKS (Whittaker–Kotel'nikov–Shannon) sampling theorem, which reads

$$f(x) = \sum_{n \in \mathbb{Z}} f\left(\frac{\pi}{w}n\right) \frac{\sin(wx - n\pi)}{wx - n\pi}, \quad x \in \mathbb{R}, \quad (2)$$

when the input function f satisfies certain suitable assumptions. The convergence is absolute and uniform with respect to x , when it belongs to any compact from \mathbb{R} . In (2) is also usual the notation $\text{sinc}(z) := z^{-1} \sin(z)$, $z \neq 0$, while $\text{sinc}(0) = 1$, originating back to the *sinus cardinalis* terminology introduced by Woodward in 1952.

Independently of the mathematicians' works, the Formula (2) was obtained by the young Soviet electrical engineer Kotel'nikov [1], who pointed out its great importance for communication theory (16 years before) Shannon, who deals with the problems that occur in certain communication system [2,3]. For this reason, the above mentioned Formula (2) in Russian literature is called *Kotel'nikov formula* and in Western literature *Shannon formula*. That controversy in nomenclature "Kotel'nikov versus Shannon" was precisely treated by Kolmogorov in 1956 in his MIT talk [4].

At the end of the 1940s the prerequisites for extending the sampling theorem to stochastic weakly stationary processes were developed—Shannon’s paper, the spectral representation formula, Karhunen–Cramér theorem [5] (p. 156). Therefore, except Belyaev’s oversampling paper [6] all other mentioned results follow the described research direction used by the Frenchmen Ville [7–9] and Oswald [10], then Americans Lloyd [11] and Parzen [12] and/or the Russian (in that time Soviet) mathematician Belyaev. At this point we have to point out that it is conventionally said that Balakrishnan in 1957 was the inventor of stochastic sampling theorem. It is interesting to quote that both Parzen and Balakrishnan have submitted their manuscripts in 1956; Parzen’s technical report [12] (never published in a journal) was submitted at December 22, while Balakrishnan’s manuscript [13] has been submitted to the IRE Transactions of Information Theory *a month earlier*, precisely at November 23. Both articles discuss the *mean square* restoration of bandlimited weak sense stationary stochastic processes. This can confirm Balakrishnan’s priority.

Despite these facts Ville and Oswald gave introductory results to the sampling reconstruction procedure for bandlimited stochastic processes years earlier, see [14]. In his article [6] Belyaev emphasized that “... a rigorous proof of the Formula (2) for the sample functions of random processes with bounded spectra has not been given...” before his work for both the mean-square (and *a fortiori* almost sure) sampling reconstruction formulae [6] (p. 411). The above referenced authors didn’t cite each other, which we can clearly see by using into account the excellent Lloyd paper [11] obviously unknown by Belyaev. For further information concerning sampling theory development, see among others the cornerstone papers [B], the famous Jerri’s ‘1977–sampling–paper’ [15], the overview of the Soviet sampling research results which are definitely devoted to a greater extent to stochastic signals than the Eastern (mainly USSR) investigations by Khurgin and Yakovlev [16], Higgins ‘5 short stories’ [17] and Unser’s [18]. Finally, for the less known almost sure sense convergence approach in stochastic signals reconstruction one can consult the book chapter [19] with the exhaustive references list therein.

2. Paul Butzer’s Stochastic Sampling Theoretical Research

In the following we expose Paul Butzer’s contribution and his excellent results to the topic which gives deep insight into the background and mathematical formulation and generalizations of the Formula (1) when the input signal is a stochastic process.

In his RWTH Aachen, Paul Butzer was a leading person in Analysis area, teaching, conducting research, supervising PhD students in Aachen (according to the Mathematics Genealogy database he has 38 PhD students in 35 years).

Moreover by the MathSciNet, among Approximation theory, Fourier analysis, Harmonic analysis, Functional analysis, Special functions (the famous Butzer-Flocke-Hauss Omega function), Integral transforms (specially Fourier and Mellin transforms), mentioning only a part of his interest areas, Paul Butzer authored and co-authored 47 titles in which “sampling” is mentioned. His co-authors’ list contains the following names: R. L. Stens (22), G. Schmeisser (7), W. Engels (5), J. R. Higgins (5), W. Splettstösser (5), C. Bardaro (4), P. J. S. G. Ferreira (4), G. Vinti (4), M. H. Annaby (3), A. Fischer (3), G. Hinsin (3), I. Mantellini (3), S. Ries (3), G. Schöttler (3), A. I. Zayed (2), M. M. Dodson (1), A. Gessinger (1), M. Hauss (1), S. Jansche (1), O. Lange (1) and U. Scheben (1). This tremendously wide opus cannot be presented in ‘finite/real time’, so we concentrate to few of his papers with results regarding stochastic signals restoration, which are:

[A] Butzer, P.L. A survey of the Whittaker-Shannon sampling theorem and some of its extensions. *J. Math. Res. Exposition* **1983**, *3*, 185–212.

[B] Butzer, P.L., Splettstösser, W. Stens, R.L. The sampling theorem and linear prediction in signal analysis. *Jahresber. Deutsch. Math.-Verein.* **1988**, *90*, 1–70.

Leading the research in the RWTH Aachen, Butzer has also switched to a classical question of probability theory and linear stochastic processes, precisely to the Central Limit Theorem (CLT) in which he realized the following titles in cooperation with Ursula Gather:

[C] Butzer, P.L., Gather U. Asymptotic expansions for central limit theorems for general linear stochastic processes. I. General theorems on rates of convergence. *Math. Methods Appl. Sci.* **1979**, *1*, 241–264.

[D] Butzer, P.L., Gather U. Asymptotic expansions for central limit theorems for general linear stochastic processes. II. Models of the general random noise and pulse train processes. *Math. Methods Appl. Sci.* **1979**, *1*, 346–353.

Finally, the prediction of signals from the whole ‘past history’, that is, the known past samples, was treated by Butzer and his most frequent co-author, colleague, PhD student and friend Rudolf (Rolf) Leonard Stens, which is represented out of excellent memoir [B], by

[E] Butzer, P.L., Stens R.L. Prediction of nonbandlimited signals from past samples in terms of splines of low degree. *Math. Nachr.* **1987**, *132*, 115–130.

In [C, D] not necessarily wide sense stationary processes (a connection to the so-called *Piranašvili processes*, introduced in the Piranašvili’s 1967 cornerstone sampling/interpolation paper [20] could act as a further interesting research goal) are treated from the Central Limit Theorem point of view, exploiting Honda’s concept of linear stochastic processes [21].

The article [E] discusses the signals prediction from past samples, in which the similar approach is exploited as in the sampling reconstruction procedures. It is worth mentioning that the simple implementation of the classical prediction procedure (generalized in [E]) to the weakly stationary stochastic processes was published in the conference talk [22].

3. The Results

My attention to the WKS sampling theorem has been drawn in the mid 1980s by Professor Mátyás Arató during my ten months long research grant spent in the Department of Probability and Statistics of Eötvös Loránd University, Budapest. (Professor Arató was a PhD student of A.N. Kolmogorov. Ergo, he called (2) precisely Kotel’nikov formula). At that time, I knew about this result reading Yaglom’s ‘small book’ [23] (p. 204) but become interested in the sampling theorems for stochastic processes after this suggestion, publishing my first sampling paper [24]. Collecting the research material for this topic, I have found, among others, Paul Butzer’s titles regarding sampling theorems. I have met Paul Butzer in 1987 in Niš, Serbia at the *Third Numerical Methods and Approximation Theory Conference—NMAT’87*, organized by Gradimir N. Milovanović, where Butzer delivered a lecture composed from results published in [E]. Writing him soon after that conference concerning the possibility for a postdoc position in Aachen, I have been very pleased with his few pages long response letter in which he explained that this would be not possible since his several duties either in teaching, research or other numerous obligations with postdoc fellows and PhD students. However, he kindly sent me a big package of his own papers covering the area including [A, B, E]. With this gift, I began to collect sampling theorems titles and have a strong insight into stochastic sampling results from Balakrishnan’s and Parzen’s papers to new results (for instance, in the so-called *exponential sampling*) published nowadays.

First, we recall few definitions and results (listed in **boldface**). Let $f \in C(\mathbb{R})$. The **modulus of continuity** with respect to the difference order r is defined by

$$\omega_r(\delta; f; C(\mathbb{R})) := \sup_{|h| < \delta} \left\| \sum_{k=0}^r (-1)^k \binom{r}{k} f(\cdot + kh) \right\|_C, \quad \delta > 0,$$

and the related Lipschitz class of order $\alpha > 0$ having Lipschitz constant L by

$$\text{Lip}_L(\alpha; C(\mathbb{R})) := \{f \in C(\mathbb{R}) : \omega_r(\delta; f; C(\mathbb{R})) \leq L\delta^\alpha, \delta > 0\} .$$

The Bernstein function class B_σ^p , $1 \leq p \leq \infty$ consists of exponentially bounded functions of type σ , which restriction to \mathbb{R} belongs to $L^p(\mathbb{R})$. The characterization of B_σ^p , $1 \leq p \leq 2$ by the Paley–Wiener theorem asserts that an $f \in L^p(\mathbb{R})$ has an extension to \mathbb{C} as an element of B_σ^p iff the associated Fourier transform pair $f^\wedge(x)$ vanishes a.e. outside of $[-\sigma, \sigma]$, that is, it is *bandlimited* to this interval:

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\sigma}^{\sigma} e^{itx} f^\wedge(x) dx, \quad t \in \mathbb{R}.$$

In explaining (even in a heuristic manner!) these results, now, in a stochastic framework, we need the following definitions. The **stochastic process** $\{X(t, \omega) \equiv X(t) : t \in \mathbb{R}\}$ defined on a standard probability space (Ω, \mathcal{A}, P) is an \mathcal{A} -measurable function, that is $X(t_0, \omega)$ becomes a random variable for any fixed $t = t_0 \in \mathbb{R}$, while $X(t, \omega_0), \omega_0 \in \Omega$ is a deterministic function called *trajectory of X*. The process $X(t)$ belongs to the class of all finite second moment random variables $L^2(\Omega) := L^2(\Omega, \mathcal{A}, P)$. This class possesses a Hilbert space structure with scalar product defined by the expectation operator E , that is $\langle X, Y \rangle = EXY^*$ and with the norm $\|X\|_{L^2} = \sqrt{E|X|^2}$ equipped.

The expectation function $m_X(t) = EX(t)$, while the correlation function

$$R_X(t, s) := E(X(t) - m_X(t))(X(s) - m_X(s))^*, \quad t, s \in \mathbb{R}.$$

when $m_X(t)$ is a constant and $R_X(t, s) = R_X(t - s)$, the process is wide sense (Hinčin) stationary (WSS) and we have that $R_X(t) = EX(t)X^*(0)$; $R_X(0) = DX(t)$ is the variance of the process, with the property $|R_X(t)| \leq R_X(0)$.

The **Karhunen–Cramér theorem**, which characterization theorem can also be called *generalized harmonizability theorem* [25–27], reads as follows [5] (p. 156). Assume that the correlation function $R_X(t, s)$ of a stochastic process $X(t)$ possesses a double integral representation in the form

$$R_X(t, s) = \int_{\Lambda} \int_{\Lambda} f(t, \lambda) f^*(s, \mu) F_X(d\lambda, d\mu), \quad \Lambda \subseteq \mathbb{R}; t, s \in \mathbb{R}, \tag{3}$$

where $F_X(d\lambda, d\mu)$ is a positive definite kernel of finite variation and

$$\int_{\Lambda} \int_{\Lambda} f(t, \lambda) f^*(s, \mu) F_X(d\lambda, d\mu) < \infty.$$

Then $X(t)$ has a spectral representation and

$$X(t) = \int_{\Lambda} f(t, \lambda) Z_X(d\lambda), \quad t \in T, \tag{4}$$

where Z_X is a stochastic measure satisfying

$$F_X(S_1, S_2) = EZ_X(S_1)Z_X^*(S_2), \quad S_1, S_2 \in \sigma(\Lambda); \tag{5}$$

here $\sigma(\Lambda)$ stands for the appropriate σ -field.

Conversely, if the stochastic process $X(t)$ is representable in the form (4) in which the spectral measure Z_X satisfies Equation (5) and F_X is of bounded variation, then there holds the spectral representation (3) of the correlation function R_X .

In the case when $f(t, \lambda)$ is a Fourier kernel $e^{i\lambda t}$ and $F_X(u, v) = \delta_{uv}F_X(u)$, i.e. it is concentrated on the diagonal of \mathbb{R}^2 , the process $X(t)$ becomes WSS. Then, the stochastic process X and the related correlation function R_X have the following spectral representations

$$X(t) = \int_{\Lambda} e^{i\lambda t} Z_X(d\lambda), \quad R_X(t) = \int_{\Lambda} e^{i\lambda t} F_X(d\lambda); \quad t \in \mathbb{R},$$

respectively. Here, $F_X(S) = E|Z_X(S)|^2$, for all $S \in \mathcal{B}_{\mathbb{R}}$. Moreover, if $\Lambda \equiv \mathbb{R}$ the process is *non-bandlimited*, while for $\Lambda = [-w, w]$, $w > 0$ the stochastic process is *bandlimited* to the bandwidth w .

Obviously, these formulae readily follow from the spectral representations (4) and (3), having in mind Display (5).

Next, the stochastic process $X(t)$ is continuous and respectively differentiable at $t \in \mathbb{R}$ with a first derivative $X'(t)$, both in the mean-square sense used, when

$$\lim_{h \rightarrow 0} \|X(t+h) - X(t)\|_{L^2}^2 = 0, \quad \lim_{h \rightarrow 0} \left\| \frac{X(t+h) - X(t)}{h} - X'(t) \right\|_{L^2}^2 = 0.$$

Higher order mean-square derivatives $X^{(r)}(t), r = 2, 3, \dots$, used in the following exposition, can be defined iteratively.

Finally, let $X \in L^2(\Omega)$, that is let X a finite second-order moment stochastic process. The **modulus of continuity for stochastic signal** X is defined by

$$\omega_r(\delta; X; L^2(\Omega)) := \sup_{|h| < \delta} \left\| \sum_{k=0}^r (-1)^k \binom{r}{k} X(\cdot + kh) \right\|_{L^2}, \quad \delta > 0.$$

Obviously, for WSS processes this definition is t -free. The related Lipschitz class is defined analogously.

Also, we introduce the **convolution of a process** $X \in L^2(\Omega)$ with a (deterministic) function $g \in L^1(\mathbb{R})$ that turns out to be the (stochastic) integral

$$(X * g)(t) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} X(t)g(t-x)dx, \quad t \in \mathbb{R},$$

where the validity of this representation holds almost surely (with probability 1). For $X \in L^2(\Omega)$ with $X^{(r)} \in \text{Lip}(\alpha, L^2(\Omega)), \alpha \in (0, 1]$, we have the estimate (in the mean-square sense)

$$\sup_{t \in \mathbb{R}} \|X(t) - (X * \theta_\rho)(t)\|_{L^2} = \mathcal{O}(\rho^{-r-\alpha}), \quad \rho \rightarrow \infty.$$

Here is $\theta_\rho(t) = \rho\theta(\rho t)$, where $\theta(\cdot)$ denotes the *de la Vallée Poussin kernel* [B] (p. 8, Eq. (2.3)), [28]

$$\theta(t) = \frac{4}{\sqrt{2\pi}t^2} \sin\left(\frac{t}{2}\right) \sin\left(\frac{3t}{2}\right), \quad t \in \mathbb{R}.$$

Finally, let us recall that the WKS sampling theorem for the bandlimited stochastic process $X(t)$, having a bandwidth $w > 0$ reads

$$X(t) = \sum_{n \in \mathbb{Z}} X\left(\frac{\pi}{w}n\right) \frac{\sin(\omega t - n\pi)}{\omega t - n\pi}, \quad t \in \mathbb{R},$$

in accordance with the WKS Formula (2) valid for deterministic signals. However, when we consider the convergence of its truncated to $|k| \leq N$ partial-sums-sequence we arrive at the following result.

Theorem 1 ([B] (p. 57, Theorem 6.11.)). *The bandlimited WSS process X has $R_X \in \mathcal{B}_{\pi w}^p, p \geq 1$. Then*

$$\lim_{N \rightarrow \infty} \left\| X(t) - \sum_{|k| \leq N} X\left(\frac{k}{w}\right) \text{sinc}(\omega t - k) \right\|_{L^2}^2 = 0.$$

Theorem 2 ([B] (p. 57, Theorem 6.12.)). *Let X be an r time differentiable WSS process with $X^{(r)} \in \text{Lip}(\alpha, L^2(\Omega)), \alpha \in (0, 1]$. If $|R_X(t)| \leq M_r |t|^{-\gamma}, \gamma \in (0, 1], |t| \geq 1$, and $M_r = M_r(\gamma) > 0$, then*

$$\left\| X(t) - \sum_{k \in \mathbb{Z}} X\left(\frac{k}{w}\right) \text{sinc}(\omega t - k) \right\|_{L^2} = \mathcal{O}(w^{-r-\alpha} \log w), \quad w \rightarrow \infty.$$

Closing the definitions, we introduce **generalized sampling expansion** [B] (pp. 24–25) when the sinc kernel is replaced with another suitable kernel φ :

$$(S_w^\varphi f)(t) := \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} f\left(\frac{k}{w}\right) \varphi(wt - k). \tag{6}$$

Consider [B] (p. 25, Theorem 4.1). When $\varphi(t - \cdot) \in \ell^1(\mathbb{Z})$ for all $t \in \mathbb{R}$, and

$$\frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} \varphi(t - k) = 1. \tag{7}$$

Then

- (a) there holds true $f(t) = \lim_{w \rightarrow \infty} (S_w^\varphi f)(t)$, for all $f \in C(\mathbb{R})$, when $t \in \mathbb{R}$;
- (b) $\{S_w^\varphi\}_{w>0}$ defines a family of bounded linear operators from $C(\mathbb{R})$ into itself, so that

$$\|S_w^\varphi f\|_{[C,C]} = m_0(\varphi), \quad \text{and} \quad \lim_{w \rightarrow \infty} \|f - S_w^\varphi f\|_C = 0.$$

The generalized sampling series expansion associated with the kernel φ we define in a similar way as in (6) for the input function f , applying the operator S_w^φ , i.e.,

$$(S_w^\varphi X)(t) := \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} X\left(\frac{k}{w}\right) \varphi(wt - k), \quad t \in \mathbb{R}.$$

Theorem 3 ([B] (p. 58, Theorem 6.14)). *Let $\varphi \in C(\mathbb{R})$ satisfies (6) and (7), $w > 0$. Then there holds*

$$\lim_{w \rightarrow \infty} \|X(t) - (S_w^\varphi X)(t)\|_{L^2} = 0,$$

uniformly in $t \in \mathbb{R}$.

Let us define the absolute moment of φ of the order r [B] (p. 24)

$$m_r(\varphi) := \sum_{t \in \mathbb{R}} \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} |t - k|^r |\varphi(t - k)|, \quad r \in \mathbb{N}_0.$$

Theorem 4 ([B] (p. 59, Theorem 6.16)). *Let $\varphi \in C(\mathbb{R})$, $m_{2r}(\varphi) < \infty$ for certain $r \in \mathbb{N}$. If the moments*

$$\sum_{k \in \mathbb{Z}} (t - k)^j \varphi(t - k) = 0, \quad j = 1, \dots, 2r - 1; t \in \mathbb{R},$$

then

- (c) for r -fold differentiable WSS process X , we have

$$\|X(t) - (S_w^\varphi X)(t)\|_{L^2}^2 \leq \frac{(m_0(\varphi) + 3)m_{2r}(\varphi)}{(2r)! w^{2r}} R_X^{(2r)}(0); \tag{8}$$

- (d) moreover, there exists an absolute constant $K > 0$, such that for WSS process $X \in L^2(\Omega)$ there holds

$$\|X(t) - (S_w^\varphi X)(t)\|_{L^2} \leq K \omega_r(w^{-1}; X; L^2(\Omega)), \quad t \in \mathbb{R}, w > 0.$$

When φ coincides with the classical sinc-kernel and the input WSS process is non-bandlimited, we have the well-known Brown’s aliasing error upper bound [29]

$$\|X(t) - (S_w^{\text{sinc}} X)(t)\|_{L^2}^2 \leq 4 \int_{|\lambda|>w} F_X(d\lambda), \tag{9}$$

giving a result related to (8) for arbitrary finite $w > 0$. The similar fashion result for the non-bandlimited differentiable homogeneous random fields has been established by the author, compare [30] (p. 128, Theorem). During the Fourier Analysis and Applications conference, held at York University in 1993 organized by M. Dodson and M. Beaty, Paul Butzer has drawn the author's attention to the fact that Brown's results, and, at the same time his generalization of Brown's upper bound (9) to non-bandlimited random fields, are Jackson-type inequalities.

Ending this exposure, we have to remark that Theorems 3, 4 are more condensed than Theorems 1 and 2 due to the stochastic nature of the input signals, concerning to the spectral representations either of the stochastic input processes or their correlation functions.

Finally, the generalized sampling expansion operator S_w^{ϕ} gives a highly elastic tool in generating very different sampling results manipulating with this operator in several directions: derivative sampling, approximative sampling (considering non-bandlimited input signals), multi-channel sampling etc. The modulus of continuity formulation of the earned results, compare (d) in Theorem 4, gives an excellent insight into analysis background of probabilistic setting of the aliasing approximation error's magnitude in function of the bandwidth $w > 0$ of the input stochastic process by its generalized sampling expansion series.

4. Discussion

It is very interesting, therefore it should be emphasized, that Butzer's stochastic processes sampling results do not take into account the Karhunen–Cramér theorem. However, his research project was realized in the following way: take the generalized sampling expansion of the signal and find its correlation function; (ii) studying the resulting truncated series; now, of the correlation function, the authors describe the final behavior of this double sum. Actually, Butzer and co-authors follow the similar derivation steps as the great probabilists Kari Karhunen and Harald Cramér when they establish the above mentioned, so-called Karhunen–Cramér theorem. Karhunen's and Cramér's proving method was based on the isometric isomorphism between the so-called Hilbert space of the stochastic process $X(t)$, $\mathcal{H}(X) := \overline{\mathcal{L}^2\{X(t) : t \in \mathbb{R}\}}$, say, which is the *in medio* closure of the linear span of the set $\{X(t) : t \in \mathbb{R}\}$, and the H -space

$$L^2(F_X; \Lambda) = \left\{ \phi(z) : \int_{\Lambda} \int_{\Lambda} \phi(z) \phi^*(w) F_X(dz, dw) < \infty \right\},$$

being the class of all complex valued functions integrable with respect to the spectral measure $F_X(\cdot, \cdot)$.

On the other side, the excellent real specialist for stochastic processes in the 'Aachener Schule' was Wolfgang Splettstösser; it was an interesting coincidence that three great sampling theoretical specialists Butzer, Splettstösser and Stens have been close not only in their research interests, but also geographically, or this wasn't the case? We could think in an opposite direction: when they were in the same research team, the Aachen Sampling Project began? Anyway, we have the tremendous research record (consult the exhaustive references list in [B] and [31]!) in the topic presented here, and stochastic counterpart results consist due to such beautiful formulae.

Funding: This research received no external funding.

Acknowledgments: The author is highly grateful to the organizers of the **Harmonic Analysis and Applications - Strobl'18** conference in Strobl, Austria, held June 4-8, 2018, to invent and organize a special section '**Sampling Theory after Paul Butzer' to honor the 90th birthday of Paul Leo Butzer**. Thanking to this initiative the author has the excellent opportunity to expose his homage talk entitled *On generalized derivative sampling series expansion* in which several parts of this work were presented. In addition, the author is indebted to all three anonymous referees for their careful reading, constructive suggestions and reorganization requests of the first draft of the manuscript. By their kind guiding process, this survey note encompasses its final desired form.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Kotel'nikov, V.A. On the carrying capacity of the "ether" and wire in telecommunications. In *Material for the First USSR, All-Union Conference and Questions of Communications*, Izдание Red. Upr. Svyazi RKKA: Moscow, 1933; pp. 1-19. Reprint in Benedetto, J.J., Ferreira, P.J.S.G. (Eds.) *Modern Sampling Theory: Mathematics and Applications*; Birkhäuser, Boston, MA, USA, 2000.
2. Shannon, C.E. A mathematical theory of communication. *Bell Sys. Tech. J.* **1948**, *27*, 623–657.
3. Shannon, C.E. Communication in the presence of noise. *Proc. IRE* **1949**, *37*, 10–21.
4. Kolmogorov, A.N. On the Shannon theory of information transmission in the case of continuous signals. *IRE Trans. Inform. Theory* **1956**, *2*, 102–108.
5. Yaglom, A.M. *Correlation Theory of Stationary and Related Random Functions II. Supplementary Notes and References*; Springer Series in Statistics; Springer-Verlag, New York, NY, USA, 1987.
6. Belyaev, Yu.K. Analytic random processes. *Teor. Veroyat. Primenen.* **1959**, *III*, 437–444. (In Russian)
7. Ville, J. Signaux analytiques a spectre borne. *Bull. Tech. Inf. SOTELEC* **1949**, *4*, 1–18.
8. Ville, J. Signaux analitiques a spectre borne. *Câbles et Transm.* **1950**, *4*, 9–23.
9. Ville, J. Signaux analitiques a spectre borne II. *Câbles et Transm.* **1953**, *7*, 44–53.
10. Oswald J. Signaux aleatoires a spectre limite. *Câbles et Transm.* **1951**, *5*, 158–177.
11. Lloyd, S.P. A sampling theorem for stationary (wide sense) stochastic processes. *Trans. Amer. Math. Soc.* **1959**, *92*, 1–12.
12. Parzen, E. *A Simple Proof and Some Extensions of the Sampling Theorem*. Tech. Rep. No. 7; Stanford University: Stanford, CA, USA, 1956.
13. Balakrishnan, A.V. A note on the sampling principle for continuous signals. *IRE Trans. Inform. Theory* **1957**, *3*, 143–146.
14. Draščić, B. Sampling reconstruction of stochastic signals - the roots in the fifties. *Austrian J. Statist.* **2007**, *36*, 65–72.
15. Jerri, J.A. The Shannon sampling theorem - its various extensions and applications: A tutorial review. *Proc. IEEE* **1977**, *65*, 1565–1596.
16. Khurgin, Yu.I., Yakovlev, V.P. Progress in the Soviet Union on the theory and applications of bandlimited functions. *Proc. IEEE* **1977**, *65*, 1005–1028.
17. Higgins, J.R. Five short stories about the cardinal series. *Bull. Am. Math. Soc. (N.S.)* **1985**, *12*, 45–89.
18. Unser, M. Sampling-50 years after Shannon. *Proc. IEEE* **2000**, *88*, 569–587.
19. Pogány, T. Almost sure sampling restoration of bandlimited stochastic signals. In *Sampling Theory in Fourier and Signal Analysis: Advanced Topics*; Higgins, J.R., Stens, R.L., Eds.; Oxford University Press: Oxford, UK, 1999; pp. 203–232, 284–286.
20. Piranašvili, Z.A. The problem of interpolation of random processes. *Teor. Veroyatnost. Primenen.* **1967**, *12*, 708–717. (in Russian)
21. Honda, I. On the rate of the central limit theorem for a class of linear processes. *Rep. Statist. Appl. Res. Jpn. Sci. Eng.* **1979**, *26*, 18–28.
22. Pogány, T. On the perfect linear prediction of bandlimited stationary stochastic processes. In *Proceedings of the KOI'91*; Neralić, L., Ed.; Ekonomski Fakultet: Zagreb, Croatia, 1991; pp. 259–264.
23. Yaglom, A.M. *An Introduction to the Theory of Stationary Random Functions*; Revised English ed.; Silverman R. A., Ed.; Prentice-Hall, Inc.: Englewood Cliffs, NJ, USA, 1962.
24. Pogány, T. An approach to the sampling theorem for continuous time processes. *Aust. J. Statist.* **1989**, *31*, 427–432.
25. Karhunen, K. Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae Ser. A* **1947**, *37*, 3–79.
26. Cramér, H. A contribution to the theory of stochastic processes. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*; University of California Press: Berkeley/Los Angeles, CA, USA, 1951; pp. 329–339.
27. Yaglom, A.M. *Correlation Theory of Stationary and Related Random Functions I. Basic Results*; Springer Series in Statistics; Springer: New York, NY, USA, 1987.
28. de la Vallée Poussin, C. Sur la meilleure approximation des fonctions d'une variable réelle par des expressions d'ordre donné. *C. R. Acad. Sci. Paris* **1918**, *166*, 799–802.

29. Brown, J.R., Jr. On mean-square aliasing error bound in cardinal sampling expansion of random processes. *IEEE Trans. Inform. Theory* **1978** *IT-24*, 254–256.
30. Pogány, T. On the aliasing error upper bound for homogeneous random fields. *Singal Processing* **1993**, *33*, 127–129.
31. Görlich, E., Krieg, A., Nessel, R.J., Stens, R.L. Paul L. Butzer – Five years as professor emeritus. *Results Math.* **1998**, *34*, 20–31.



© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).