Sampling and Ergodic Theorems for Weakly Almost Periodic Signals

G. Casinovi

IEEE Transactions on Information Theory
vol. 55, no. 4, pp. 1883–1897, April 2009

Abstract

The theory of abstract harmonic analysis on commutative groups is used to prove sampling and ergodic theorems concerning a particular class of finite-power signals, which are known as weakly almost periodic. The analysis brings to light some noteworthy differences between finite-energy and finite-power signal sampling. It is shown that the bandwidth of the Fourier transform of a weakly almost periodic signal is generally larger than the bandwidth of the power spectrum of the signal. Consequently, the signal power spectrum by itself does not generally provide enough information to determine the value of the time-domain Nyquist rate, that is, the minimum sampling rate necessary for exact signal reconstruction in the time-domain. On the other hand, it is also shown that the minimum sampling rate needed to obtain alias-free spectral estimates is determined by the bandwidth of the power spectrum and, consequently, may be lower than the time-domain Nyquist rate. Finally, the sampling and ergodic theorems established in this paper are used in an analysis of averaged periodogram estimates of the power spectrum of a weakly almost periodic signal. It is shown that the value of the time shift between consecutive windows may contribute to the asymptotic bias of the estimates.

Copyright Notice

Sampling and Ergodic Theorems for Weakly Almost Periodic Signals

Giorgio Casinovi, *Senior Member, IEEE*

***Abstract*—The theory of abstract harmonic analysis on commutative groups is used to prove sampling and ergodic theorems concerning a particular class of finite-power signals, which are known as weakly almost periodic. The analysis brings to light some noteworthy differences between finite-energy and finite-power signal sampling. It is shown that the bandwidth of the Fourier transform of a weakly almost periodic signal is generally larger than the bandwidth of the power spectrum of the signal. Consequently, the signal power spectrum by itself does not generally provide enough information to determine the value of the time-domain Nyquist rate, that is, the minimum sampling rate necessary for exact signal reconstruction in the time-domain. On the other hand, it is also shown that the minimum sampling rate needed to obtain alias-free spectral estimates is determined by the bandwidth of the power spectrum and, consequently, may be lower than the time-domain Nyquist rate. Finally, the sampling and ergodic theorems established in this paper are used in an analysis of averaged periodogram estimates of the power spectrum of a weakly almost periodic signal. It is shown that the value of the time shift between consecutive windows may contribute to the asymptotic bias of the estimates.**

***Index Terms*—Abstract harmonic analysis, finite-power signals, sampling theory, spectral estimation.**

I. Introduction

IT is generally acknowledged that the papers by Kotelnikov [1] and Shannon [2] were the first to draw attention to the practical importance of the sampling theorem to communications theory. Research interest in this and other related topics has grown enormously since, spurred by the introduction of digital signal processing techniques in an ever growing number of applications. Advances made in sampling theory since its inception have been reviewed in several publications: see, for instance, [3]–[6].

From a mathematical standpoint, the development of sampling theory has taken place mostly within the framework of L^2 or, less frequently, L^p spaces. The signals involved in many practical applications, however, are more accurately modeled as having finite power, rather than finite energy, and finite-power signals are not elements of $L^p(\mathbb{R})$ for any finite value of p. Thus there is both abstract and practical interest in determining what parts of the sampling theory of finite-energy signals can be extended to finite-power signals.

Manuscript received June 08, 2007; revised March 27, 2008. Current version published March 18, 2009. This work was supported in part by the National Science Foundation under Grant CCR-0306343.

The author is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA.

Communicated by H. Bölcskei, Associate Editor for Detection and Estimation.

Digital Object Identifier 10.1109/TIT.2009.2013021

It is perhaps surprising to find out that, more than fifty years after the publication of Kotelnikov's and Shannon's papers, the number of sampling theory results that have been mathematically proven to be applicable to finite-power signals remains relatively small. As one would expect, the classic sampling theorem is one of them. In one of the first formal proofs of that theorem to be published, pointwise convergence of the Cesàro means of the sampling expansion was established for a particular class of bandlimited, finite-power signals [7]. That class essentially coincides with $B(\mathbb{R})$, the Fourier-Stieltjes algebra of $\mathbb{R}$, which will be introduced later in this paper. Subsequently, it was proven that the sampling expansions of band-limited signals in $B(\mathbb{R})$ actually converge uniformly on compact subsets of $\mathbb{R}$ [8]. This property was later found to hold for the sampling expansions of an even larger class of signals [9].

Other published results that are applicable to finite-power signals include a modified formulation of the sampling theorem, whose proof is based on the theory of distributions [10]. This approach leads to a series expansion whose terms are modified sinc functions. Finally, a mean-square version of the sampling theorem is stated in [11] and [12], in the form of the following equality:

$$\lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} \left| x(t+\tau) - \sum_{n=-\infty}^{+\infty} x(t+nT)\frac{\sin\omega_0(\tau-nT)}{\omega_0(\tau-nT)}\right|^2 dt = 0.$$

It should be noted, however, that the equality above does not really amount to a sampling expansion, because the averaging is performed with respect to t, not τ. It would be more accurate to describe that equation as an expansion of x in terms of shifted copies of itself.

This paper proves a number of sampling and ergodic theorems for a particular subset of finite-power signals, known as weakly almost periodic. As shown in this paper, this subset belongs to the class of signals that have a discrete power spectrum. The sampling and ergodic theorems are then used to establish further results that have practical implications in the area of spectral estimation. The paper relies extensively on the theory of abstract harmonic analysis on commutative groups [4], [13], and [14], an approach already used to prove a generalized version of the classic sampling theorem in the finite-energy case [13], [15]. The main advantage of such an approach is that both continuous-time and discrete-time signals can be dealt with in a single, unified mathematical setting.

The most important results established in this paper can be summarized as follows. First, it is shown in Section III that the Fourier transform and the power spectrum of a weakly almost periodic signal do not necessarily have the same bandwidth. It follows that the time-domain Nyquist rate of the signal, that is, the minimum sampling rate needed to avoid time-domain aliasing, cannot generally be determined from the signal power spectrum alone. In fact, the bandwidth of the Fourier transform of a weakly almost periodic signal may be infinite, even if the bandwidth of its power spectrum is finite. In contrast, the Fourier transform and the energy spectrum of a finite-energy signal always have the same bandwidth.

A second noteworthy result, found in Section VIII, concerns the minimum sampling rate that is needed to ensure that the power spectrum of the sequence of samples of a weakly almost periodic signal is an exact, frequency-scaled copy of the power spectrum of the original signal. It is shown that this requirement is met if the sampling rate is larger than twice the highest frequency in the power spectrum of the continuous-time signal. This means that, if sampling is performed for the purpose of spectral estimation, then the minimum allowable sampling rate can be determined from the signal power spectrum.

Another interesting application of the sampling and ergodic theorems established in this paper is presented in Section IX, which analyzes the asymptotic behavior of averaged periodogram spectral estimates. It is shown that when the averaged periodogram is used to estimate the power spectrum of a weakly almost periodic signal, the value of the time shift between consecutive windows may contribute to the asymptotic bias of the estimates. To avoid this potential problem, the shift must not exceed the reciprocal of twice the largest frequency in the power spectrum of the signal.

The results listed above make it clear that the sampling theory of weakly almost periodic signals (and, by extension, of finite-power signals) does not duplicate exactly that of finite-energy signals: while there are many similarities between the two cases, there are also some significant differences. The import of those differences is not only theoretical, but also practical, in light of the ever increasing role that digital techniques play in signal processing and, in particular, in spectral estimation. For example, digital spectrum analyzers use sampling to estimate the power spectrum of a continuous-time signal, and the results obtained in Sections VIII and IX are clearly relevant to the operation of those instruments.

After a section devoted to notation, the paper begins by introducing weakly almost periodic signals in Section III, and discussing some of their properties. It is shown that weakly almost periodic signals always have a discrete power spectrum, and that they are a superset of the signals that can be expressed as inverse Fourier transforms of bounded measures. The classic sampling theorem is then discussed in Section IV. Section V analyzes signal sampling and reconstruction, when those operations are regarded as abstract maps between continuous-time and discrete-time signal spaces. Relying on the theoretical framework established in Section V, the sections that follow derive further results that are more application-oriented. For example, it is shown in Section VI that $(C,1)$ summation of the sampling expansion results in a stronger form of convergence.

The ergodic properties of weakly almost periodic signals are studied in Section VII, while Section VIII examines the relationship between sampling rate and spectral aliasing. Finally, Section IX applies the ergodic theorems of Section VII to an analysis of the averaged periodogram. Specifically, it is shown that the asymptotic bias of averaged periodogram spectral estimates may be affected by the separation interval between two consecutive windows, a problem that does not occur in correlogram-based spectral estimation [16].

II. Notation

Since this paper deals both with continuous-time and discrete-time signals, an effort has been made to select a notation that can accommodate either type. Thus the domain of the time variable is denoted by G, with the understanding that $G = \mathbb{R}$ for continuous-time signals, or $G = \mathbb{Z}$ for discrete-time signals. This choice is motivated by the fact that both $\mathbb{R}$ and $\mathbb{Z}$ are locally compact, commutative topological groups [14]. Accordingly, $\int_G f\,dt$ denotes the integral of f with respect to the Haar measure on G. If $G = \mathbb{R}$, this is the usual Lebesgue integral; if $G = \mathbb{Z}$, the Haar measure of G is the counting measure, and the integral is in fact a summation: $\int_{\mathbb{Z}} f\,dt = \sum_{t\in\mathbb{Z}} f(t)$. In either case, $f * g$ denotes convolution:

$$(f * g)(t) = \int_G f(\tau)g(t-\tau)\,d\tau$$

while f^* is the function defined by $f^*(t) = \overline{f}(-t)$.

$\hat{G}$ denotes the dual group of G [14, p. 6]. Specifically, if $G = \mathbb{R}$ then $\hat{G} = \mathbb{R}$, while if $G = \mathbb{Z}$ then $\hat{G} = \mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. $\mathbb{T}$ can be identified with the interval $(-\pi, \pi]$ under the mapping: $z = e^{j\omega}$. As before, $\int_{\hat{G}} f\,d\omega$ denotes the integral of f with respect to the Haar measure on $\hat{G}$: once again, this is the ordinary Lebesgue integral if $\hat{G} = \mathbb{R}$. In the discrete-time case, $\int_{\hat{G}} f\,d\omega$ is the Lebesgue integral of the periodic function of period 2π that coincides with f on $(-\pi, \pi]$. Consequently, in the discrete-time case the convolution operation

$$(f * g)(\omega) = \int_{\hat{G}} f(\theta)g(\omega-\theta)\,d\theta$$

must be understood to be the periodic convolution of two periodic functions of period 2π.

If X is a Banach space, X^* its dual space, $\xi \in X$ and $x \in X^*$, $\langle \xi, x\rangle$ denotes the value of x on ξ. If $\{x_n\}$ is a sequence in X^*, the notation $w^*\text{-}\lim_{n\to\infty} x_n = x$ means that $\{x_n\}$ converges to x in the weak-* topology of X^*.

$C(H)$, with $H = G$ or $H = \hat{G}$, is the Banach space of bounded, continuous functions on H under the norm $\|f\|_\infty = \sup_{h\in H}|f(h)|$. $C_0(H)$ denotes the subspace of $C(H)$ consisting of those functions that vanish at infinity; if H is compact (e.g., $H = \mathbb{T}$), then $C_0(H) = C(H)$. $L^p(H)$, with $1 \le p < \infty$, denotes the space of all measurable functions f that satisfy the condition $\int_H |f|^p\,dh < \infty$, where the integral is evaluated with respect to the Haar measure on H. $L^p(H)$ is a Banach space under the norm $\|f\|_p = (\int_H |f|^p\,dh)^{1/p}$. $M(\hat{G})$ denotes the set of complex-valued, bounded, regular Borel measures on $\hat{G}$. It is possible to define a convolution operation between two measures $\mu, \nu \in M(\hat{G})$ [14, p. 13]: the result, which is still an element of $M(\hat{G})$, will be denoted by

$\mu * \nu$. Equipped with convolution and with the usual vector space operations, $M(\hat{G})$ is a complex Banach algebra under the norm defined by $\|\mu\| = |\mu|(\hat{G})$, where $|\mu|$ is the total variation of μ [14], [17]. $M(\hat{G})$ can be identified with the dual space of $C_0(\hat{G})$ [14, p. 266]. This identification is given by the relationship

$$\langle f, \mu \rangle = \int_{\hat{G}} f, d\mu, \quad f \in C_0(\hat{G}),\ \mu \in M(\hat{G}).$$

The support of a measure $\mu \in M(\hat{G})$ is the smallest closed set $S \subseteq \hat{G}$ with the following property: $\mu(A \cap S) = \mu(A)$ for every measurable set $A \subseteq \hat{G}$ [14, p. 266]. In particular, $\mu(A) = 0$ whenever $A \cap S = \emptyset$. In this paper, the support of μ will be denoted by $\mathrm{supp}(\mu)$.

A measure $\mu \in M(\hat{G})$ is absolutely continuous with respect to the Haar measure on $\hat{G}$ if there exists $m \in L^1(\hat{G})$ such that

$$\mu(S) = \int_S m(\omega)\, d\omega$$

for all measurable sets $S \subseteq \hat{G}$. In such case, the following equality holds: $\|\mu\| = \|m\|_1$. The notation $d\mu = m\, d\omega$ will be used as a shorthand to denote the relationship between an absolutely continuous measure $\mu \in M(\hat{G})$ and the corresponding function $m \in L^1(\hat{G})$. It follows that $L^1(\hat{G})$ can be identified with the set of absolutely continuous measures on $\hat{G}$. Under this identification, $L^1(\hat{G})$ becomes a norm-closed ideal of $M(\hat{G})$ [14, p. 16].

If $\mu \in M(\hat{G})$, its inverse Fourier transform, denoted by $\mathcal{F}^{-1}\mu$, is the function on G defined by

$$[\mathcal{F}^{-1}\mu](t) = \frac{1}{2\pi} \int_{\hat{G}} e^{j\omega t}\, d\mu(\omega)$$

It can be shown that $\mathcal{F}^{-1}$ is an injective map from $M(\hat{G})$ into $C(G)$ [14, pp. 15–17]. The set of all functions that can be expressed as the inverse Fourier transform of some $\mu \in M(\hat{G})$ is a proper subspace of $C(G)$, which will be denoted by $B(G)$. It is possible to introduce a norm on $B(G)$ by defining

$$\|f\|_B = \frac{1}{2\pi}\|\mu\|, \quad f = \mathcal{F}^{-1}\mu$$

Note that $\|f\|_\infty \leq \|f\|_B$. Under this norm, $B(G)$ is a Banach algebra: it is referred to as the *Fourier-Stieltjes algebra* of G (see the Appendix).

If $f \in B(G), \mathcal{F}f$ denotes that unique measure $\hat{f} \in M(\hat{G})$ such that $f = \mathcal{F}^{-1}\hat{f}$. Note that this is somewhat different from the usual definition of the Fourier transform operation, which yields a function, not a measure. On the other hand, it can be shown that, if $\hat{f} \in L^1(\hat{G})$ (that is, if $\hat{f}$, as a measure, is absolutely continuous with respect to the Haar measure on $\hat{G}$), then $d\hat{f} = \hat{f}(\omega)\, d\omega$, where $\hat{f}(\omega)$ is the usual Fourier transform of f

$$\hat{f}(\omega) = \int_G f(t) e^{-j\omega t}\, dt.$$

III. Weakly Almost Periodic Signals

Let G be a commutative, locally compact, Hausdorff topological group. Given $x \in C(G)$, let $\mathcal{T}_x$ be the set of all translations of x by elements of G

$$\mathcal{T}_x = \{y \in C(G) : y(t) = x(t+\tau), \tau \in G\}.$$

If $\mathcal{T}_x$ is relatively compact in the weak topology of $C(G)$, then x is a *weakly almost periodic* (WAP) function on G. The set of all weakly almost periodic functions on G is usually denoted by $\mathrm{WAP}\,(G)$; it can be shown to be a norm-closed subalgebra of $C(G)$ that contains both $C_0(G)$ and $B(G)$ [18].

If $\mathcal{T}_x$ is relatively compact in the norm topology of $C(G)$, then x is an *almost periodic* function. By [19, Theorem 7.3, ch. 7], the set of all almost periodic functions on G is another norm-closed subalgebra of $C(G)$, which will be denoted by $AP(G)$. Since a norm-compact set is also weakly compact, it follows that $AP(G) \subseteq \mathrm{WAP}\,(G)$. If $G = \mathbb{R}$ or $G = \mathbb{Z}$, it is not too difficult to verify that $e^{j\omega t} \in AP(G)$ for all real values of ω. Consequently, $AP(G)$ contains the norm closure of the set of all linear combinations of complex exponentials, and is in fact equal to it [14, p. 32].

Many of the results presented in this paper rely in an essential manner on certain mathematical properties of weakly almost periodic functions, which are summarized in the theorems that follow. Although those theorems are valid in any topological group, they are stated here in a form that applies specifically to the class of functions that is of interest to this paper, namely continuous-time or discrete-time signals. Accordingly, throughout the remainder of this paper it should be assumed that $G = \mathbb{R}$ or $G = \mathbb{Z}$, unless explicitly stated otherwise.

Theorem 1 ([18]): Let $x, y \in \mathrm{WAP}\,(G)$. Then the limit

$$s_{xy}(t) = \lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t+\tau)\overline{y}(\tau)\, d\tau$$

exists for all $t \in G$, and the function it defines is almost periodic on G, that is, $s_{xy} \in AP(G)$. □

In the case of weakly almost periodic functions of two (or more) variables, averaging the function with respect to one variable at a time in any order always yields the same result, as stated in the following theorem.

Theorem 2 ([18]): Let $x \in \mathrm{WAP}\,(G \times G)$, and define

$$x_1(t) = \lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t, \tau)\, d\tau$$

$$x_2(t) = \lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(\tau, t)\, d\tau.$$

Then $x_1, x_2 \in \mathrm{WAP}\,(G)$ and

$$\lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_1(t)\, dt = \lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_2(t)\, dt. \quad \square$$

Theorem 1 implies that, if $x \in \mathrm{WAP}\,(G)$, the limit

$$\alpha_x(\omega) = \lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(\tau) e^{-j\omega\tau}\, d\tau \tag{1}$$

exists for all $\omega \in G$. Consequently, each $x \in \mathrm{WAP}\,(G)$ uniquely determines a corresponding α_x. The next theorem proves that, if x is almost periodic, the converse of this statement also holds.

Theorem 3: Let: $x \in \mathrm{WAP}(G)$, and let α_x be as in (1). Define

$$\Omega_x = \{\omega \in \hat{G} : \alpha_x(\omega) \neq 0\}$$

Then Ω_x contains at most a countable number of points. If $x \in AP(G)$ and $\Omega_x = \emptyset$ (i.e., $\alpha_x(\omega) \equiv 0$), then $x = 0$.

Proof: The proof of the countability of Ω_x is in [18], while a proof of the second part of the claim is given, for instance, by [19, Theorem 7.12, ch. 7]. □

If x is weakly almost periodic, but not almost periodic, α_x no longer determines x uniquely. Nevertheless, α_x is intimately related to the power spectrum of x, which is defined in the following way.

Let x be any complex-valued, bounded, continuous-time or discrete-time signal, and assume the following.

- The limit
$$s_x(t) = \lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t+\tau)\overline{x}(\tau)\, d\tau \tag{2}$$
exists for all $t \in G$.
- s_x is a continuous function of t (if $G = \mathbb{Z}$, this condition is satisfied automatically, because every function on a discrete space is continuous [20, p. 88]).

In such case, s_x will be referred to as the *autocorrelation function* of x. It was shown by Wiener [21] that s_x is a positive definite function, and a theorem by Bochner [14, p. 19] states that a continuous, positive definite function on G can be expressed in a unique way as the inverse Fourier transform of a bounded, positive measure $\hat{s}_x \in M(\hat{G})$, i.e., $s_x = \mathcal{F}^{-1}\hat{s}_x$. In particular, this means that $s_x \in B(G)$. Following Wiener's definition, $\hat{s}_x$ will be referred to as the *power spectrum* of x.

It should be pointed out that, even in the case of well-behaved (e.g., infinitely differentiable) signals, the limit on the right-hand side of (2) need not exist for all values of t, and if it does, s_x may not be a continuous function. For example, it is easy to verify that, if $x(t) = e^{jt^2}$, $s_x(t)$ exists for all values of t, but is not continuous at $t = 0$. On the other hand, it is clear from Theorem 1 that, if $x \in \mathrm{WAP}(G)$, then s_x exists and $s_x \in AP(G)$. In fact, s_x is completely determined by α_x, as shown by the following theorem.

Theorem 4 ([18]): Let $x \in \mathrm{WAP}(G)$. Then

$$\lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{T/2} s_x(t) e^{-j\omega t}\, dt = |\alpha_x(\omega)|^2$$

and

$$s_x(t) = \sum_{\omega_k \in \Omega_x} |\alpha_x(\omega_k)|^2 e^{j\omega_k t}.$$

In particular, $\sum_{\omega_k \in \Omega_x} |\alpha_x(\omega_k)|^2 = s_x(0)$. □

Noting that $e^{j\omega_k t} = \mathcal{F}^{-1}[2\pi\delta(\omega - \omega_k)]$, where $\delta(\omega)$ is Dirac's atomic measure located at $\omega = 0$, leads to the following characterization of the power spectrum of a weakly almost periodic signal.

Theorem 5: Let x be a weakly almost periodic continuous-time or discrete-time signal. Then

$$\hat{s}_x(\omega) = 2\pi \sum_{\omega_k \in \Omega_x} |\alpha_x(\omega_k)|^2 \delta(\omega - \omega_k).$$

It follows that $\Omega_x \subseteq \mathrm{supp}(\hat{s}_x)$, and that the power spectrum of a weakly almost periodic signal is always discrete.[1] □

Another property of weakly almost periodic signals, which yields further insight into their power spectra, is that they can be decomposed in a unique way as the sum of almost periodic and zero-power components [22]. More precisely, if $x \in \mathrm{WAP}(G)$, then $x = x_{AP} + x_0$, where $x_{AP} \in AP(G)$ and x_0 is a zero-power signal, that is, x_0 satisfies the condition: $\lim_{T \to +\infty}(1/T)\int_{-T/2}^{T/2} |x_0(\tau)|^2\, d\tau = 0$. This implies that $s_{x_0} \equiv 0$, because s_{x_0} is a positive definite function, and therefore $|s_{x_0}(t)| \le s_{x_0}(0) = 0 \quad \forall\, t \in G$ [14, p. 18]. It is then readily verified that the decomposition $x = x_{AP} + x_0$, together with the Schwartz inequality

$$\begin{aligned} &\left| \frac{1}{T} \int_{-T/2}^{T/2} x_{AP}(t+\tau)\overline{x}_0(\tau)\, d\tau \right| \\ &\quad \le \left(\frac{1}{T} \int_{-T/2}^{T/2} |x_{AP}(t+\tau)|^2\, d\tau \right)^{1/2} \\ &\quad\quad \times \left(\frac{1}{T} \int_{-T/2}^{T/2} |x_0(\tau)|^2\, d\tau \right)^{1/2} \end{aligned}$$

implies that $s_x = s_{x_{AP}}$, hence $\hat{s}_x = \hat{s}_{x_{AP}}$. In other words, the power spectrum of a weakly almost periodic signal is always equal to the power spectrum of the almost periodic component of the signal.

Consider now those signals that are elements of $B(G)$. Recall that $x \in B(G)$ if and only if there exists $\hat{x} \in M(\hat{G})$ such that $x = \mathcal{F}^{-1}\hat{x}$. In this paper, $\hat{x}$ will be referred to as the *pseudospectrum* of x, but other terms for $\hat{x}$ can be found in the literature.[2] Since $B(G) \subseteq \mathrm{WAP}(G)$, every $x \in B(G)$ has both a pseudospectrum and a power spectrum. Note that those two quantities can be quite different, because $\hat{x}$ can be any bounded measure, while $\hat{s}_x$ is necessarily discrete. Nevertheless, it is possible to establish a relationship between the support of $\hat{x}$ and that of $\hat{s}_x$, as explained next.

Let $x \in B(G)$ and let $x = x_{AP} + x_0$ be its decomposition into almost periodic and zero-power components. In this case, it can be shown that $x_{AP}(t) = \sum_{\omega_k \in \Omega_x} \alpha_x(\omega_k) e^{j\omega_k t} \in B(G) \cap AP(G)$. This means that $\hat{x}_{AP} = 2\pi \sum_{\omega_k \in \Omega_x} \alpha_x(\omega_k)\delta(\omega - \omega_k)$ and that $x_0 \in B(G)$, hence $\hat{x} = \hat{x}_{AP} + \hat{x}_0$. It is also clear from Theorem 5 that $\mathrm{supp}(\hat{s}_x) = \mathrm{supp}(\hat{x}_{AP})$, therefore $\mathrm{supp}(\hat{s}_x) \subseteq \mathrm{supp}(\hat{x})$, which is the desired relationship.[3]

In summary, the difference between a weakly almost periodic signal and its almost periodic component is a zero-power signal, which has no effect on the power spectrum. Therefore the power spectrum of a weakly almost periodic signal always coincides with the power spectrum of the almost periodic component of the signal. If the signal is an element of $B(G)$, the

[1]By definition, a measure is discrete if it is the linear combination of a finite or countable number of Dirac atomic measures [14, p. 266].

[2]For example, in [8], $\hat{x}$ is referred to as the Fourier-Stieltjes spectrum of x.

[3]See [23, eq. (17)], where an alternate proof of this relationship is given, based on the theory of distributions. In that equation, however, the direction of the inclusion is reversed, presumably because of a typographical error. This can be ascertained from the Proof of Theorem 6 in the same reference, which makes it clear what the correct direction of the inclusion is.

support of its power spectrum is a subset of the support of its pseudospectrum and, in general, this is a proper inclusion. This means that the bandwidth of the power spectrum can be strictly smaller than the bandwidth of the pseudospectrum. This cannot happen, however, if the signal is known to be almost periodic: in such case $\mathrm{supp}(\hat{s}_x) = \mathrm{supp}(\hat{x})$, and therefore the two bandwidths must be the same.

These observations highlight an important difference between finite-energy and finite-power signals: if x is a finite-energy signal, the support of the energy spectrum of x (i.e., $|\hat{x}|^2$) always coincides with the support of $\hat{x}$. This fact must be taken into account in the development of a sampling theory of finite-power signals.

IV. Classic Sampling Theorem

This section reviews the sampling theorem in its classic formulation, and shows on examples that it is the bandwidth of the pseudospectrum, not that of the power spectrum, that determines the minimum sampling rate necessary for an exact reconstruction of the signal. As a matter of notation, throughout the remainder of this paper T will denote the sampling period, $\omega_T = 2\pi/T$ will be the corresponding radian frequency, and $\mathrm{sinc}\ \theta = \sin \pi\theta/\pi\theta$.

A signal $x \in B(\mathbb{R})$ is band-limited if $\mathrm{supp}(\hat{x}) \subseteq [-\omega_0, \omega_0]$ for some finite value of ω_0. The classic sampling theorem can then be stated as follows.

Theorem 6: Let $x \in B(\mathbb{R})$ and assume $\mathrm{supp}(\hat{x}) \subseteq [-\omega_0, \omega_0]$ and $\omega_0 < \omega_T/2$. Then

$$x(t) = \sum_{k=-\infty}^{+\infty} x(kT)\mathrm{sinc}(t/T - k). \quad (3)$$

Moreover, the series converges uniformly to x on every bounded interval of $\mathbb{R}$.

Proof: A full proof of this theorem can be found, for instance, in [8]. It suffices to note that every integral with respect to a measure $\mu \in M(\mathbb{R})$ can also be expressed as a Lebesgue-Stieltjes integral with respect to a function g_μ of bounded variation on $\mathbb{R}$, and viceversa [17, p. 329 ff.]. Thus, if $x = \mathcal{F}^{-1}\mu \in B(\mathbb{R})$, there exists a function g_μ of bounded variation such that

$$x(t) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{j\omega t}\, dg_\mu(\omega).$$

If $\mathrm{supp}(\hat{x}) \subseteq [-\omega_0, \omega_0]$, the integral above need only be evaluated between $-\omega_0$ and ω_0, which is the expression used in Theorem 3 of [8]. It is then readily apparent that the claim is equivalent to Theorems 1 and 3 in that reference. □

It should be pointed out that almost all the proofs of this theorem that have been published over the years rely in an essential manner on the fact that x is the inverse Fourier transform of a measure (or, equivalently, the Fourier-Stieltjes transform of a bounded-variation function, as in [8]). In other words, most proofs of the classic sampling theorem depend on the signal being an element of $B(\mathbb{R})$. Since $B(\mathbb{R}) \subseteq \mathrm{WAP}(\mathbb{R})$, Theorem 5 implies that signals in $B(\mathbb{R})$ always have discrete power spectra. A proof of the sampling theorem that is applicable to a class of signals larger than $B(\mathbb{R})$ is given in [9]. Nevertheless, to the author's best knowledge, the exact bounds of the validity of (3) for signals that are not elements of $B(\mathbb{R})$ (e.g., signals with a continuous power spectrum) remain largely unknown.

Another point worth stressing is that Theorem 6 assumes that the pseudospectrum of x—not its power spectrum—is band-limited, and that the sampling frequency is higher than twice the highest frequency in $\hat{x}$. Of course, this implies that the bandwidth of $\hat{s}_x$ is also limited, but the converse is not necessarily true. Even if the support of $\hat{s}_x$ is bounded, it is not generally true that the highest frequency in $\hat{s}_x$ is the same as the highest frequency in $\hat{x}$. For example, let

$$x(t) = (\omega_0/\pi)\mathrm{sinc}(\omega_0 t/\pi) + \cos(\omega_1 t)/\pi$$

It is straightforward to verify that $\hat{x}(\omega) = \chi_{[-\omega_0,\omega_0]}(\omega) + \delta(\omega - \omega_1) + \delta(\omega + \omega_1)$ and $\hat{s}_x(\omega) = [\delta(\omega - \omega_1) + \delta(\omega + \omega_1)]/2\pi$, where $\chi_{[-\omega_0,\omega_0]}(\omega) = 1$ if $-\omega_0 \le \omega \le \omega_0$, otherwise $\chi_{[-\omega_0,\omega_0]}(\omega) = 0$. Obviously the bandwidth of $\hat{s}_x$ is equal to ω_1, while the bandwidth of $\hat{x}$ is equal to $\max(\omega_0, \omega_1)$. Therefore, if $\omega_0 > \omega_1$, choosing the sampling period so that $\omega_T > 2\omega_1$ is not sufficient to guarantee the exact reconstruction of the signal. As another example, let

$$x(t) = e^{-\gamma t^2} + \cos(\omega_1 t)/\pi.$$

The power spectrum of this signal is the same as in the previous example, but the support of $\hat{x}$ is unbounded, which means that exact reconstruction of the signal cannot be guaranteed at any sampling frequency.

The examples above demonstrate that the bandwidths of $\hat{s}_x$ and $\hat{x}$ can in fact be different. As explained at the end of Section III, this is due to the presence of a zero-power component in the signal, which affects the Fourier transform of the signal, but not its power spectrum. This means that, except in the case of almost periodic signals, the power spectrum does not provide sufficient information to determine the minimum sampling rate necessary to avoid time-domain aliasing. On the other hand, it will be shown in Section VIII that the minimum sampling rate that is needed to avoid spectral aliasing is indeed determined by the bandwidth of the power spectrum.

V. Signal Sampling and Reconstruction

From an abstract point of view, sampling can be regarded as a map from the space of continuous-time signals into that of discrete-time signals, and it can be shown that the image of $B(\mathbb{R})$ under this map is exactly $B(\mathbb{Z})$. Since the Fourier transform sets up a one-to-one correspondence betweens signals in $B(G)$ and measures in $M(\hat{G})$, every linear transformation between $B(\mathbb{R})$ and $B(\mathbb{Z})$ determines a corresponding map between $M(\mathbb{R})$ and $M(\mathbb{T})$, and viceversa. This observation suggests that measure theory and, in particular, the properties of transformations between measure spaces could be usefully exploited to develop a sampling theory for signals in $B(\mathbb{R})$.

This section is devoted to a theoretical analysis of four maps that are intimately related to signal sampling and reconstruction. The maps, denoted by $\sigma, \rho, \hat{\sigma}$ and $\hat{\rho}$, are defined below; their actions on the aforementioned spaces are illustrated by the

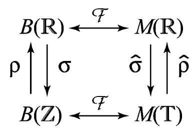

Fig. 1. Map diagram illustrating the actions of $\sigma, \rho, \hat{\sigma}$ and $\hat{\rho}$.

diagram in Fig. 1. This analysis lays the foundation for the subsequent sections, which study the convergence properties of the Cesàro means of the sampling series, and shed some light on how sampling may affect the estimation of the power spectra of continuous-time signals.

For a fixed sampling period T, let σ denote the mapping between a continuous-time signal x and the sequence of its samples, denoted by σx, which is defined by

$$[\sigma x](n) = x(nT)$$

The starting point for the analysis carried out in this section is the following theorem.

Theorem ([14, Theorem 2.7.2]): If $x \in B(\mathbb{R})$, then $\sigma x \in B(\mathbb{Z})$. More specifically, σ maps $B(\mathbb{R})$ onto $B(\mathbb{Z})$. □

Shifting to the frequency domain, let π_T be the quotient man from $\mathbb{R}$ onto $\mathbb{T}$ defined by

$$\pi_T(\omega) = e^{j2\pi\omega/\omega_T} = e^{jT\omega}.$$

This quotient map makes it possible to define two additional maps: the first, denoted by $\hat{\sigma}$, maps $M(\mathbb{R})$ into $M(\mathbb{T})$, while the second, denoted by $\hat{\rho}$, maps $M(\mathbb{T})$ into $M(\mathbb{R})$. For notational convenience, let: $I_T = (-\omega_T/2, \omega_T/2)$. If $\nu \in M(\mathbb{R})$, then $\hat{\sigma}\nu \in M(\mathbb{T})$ is defined by

$$(\hat{\sigma}\nu)(E) = \nu(\pi_T^{-1}E)$$

for every measurable subset E of $\mathbb{T}$. Similarly, if $\mu \in M(\mathbb{T})$, then $\hat{\rho}\mu \in M(\mathbb{R})$ is defined by

$$(\hat{\rho}\mu)(E) = \mu[\pi_T(E \cap I_T)]$$

for every measurable subset E of $\mathbb{R}$.

Loosely speaking, $\hat{\sigma}$ takes a measure on $\mathbb{R}$ and "wraps it around" $\mathbb{T}$ following the way in which π_T maps $\mathbb{R}$ onto $\mathbb{T}$. In particular, if ν is concentrated on I_T (i.e., $\nu(\mathbb{R} \setminus I_T) = 0$), then $\hat{\sigma}\nu$ is an identical copy of ν projected on $\mathbb{T}$. Conversely, $\hat{\rho}$ takes a measure on $\mathbb{T}$ and copies it to I_T, but it does so in a way that excludes the point $\{-1\}$, because the image of I_T under π_T is $\mathbb{T} \setminus \{-1\}$. This effectively means that $\hat{\rho}\mu$ is also an identical copy of μ, unless $\mu(\{-1\}) \neq 0$. The properties of $\hat{\sigma}$ and $\hat{\rho}$ that will be needed later are summarized in the following lemma.

Lemma 1: Let f be a measurable function on $\mathbb{R}$ and g a measurable function on $\mathbb{T}$ such that

$$f(\omega) = g \circ \pi_T(\omega), \quad \forall \omega \in I_T$$

where $\circ$ denotes map composition.

a) If $\mu \in M(\mathbb{T})$, then

$$\int_{\mathbb{R}} f \, d\hat{\rho}\mu = \int_{T\setminus\{-1\}} g \, d\mu$$

and this equality implies $\|\hat{\rho}\mu\| \leq \|\mu\|$.

b) If $\nu \in M(\mathbb{R})$ and $\nu(E) = \nu(E \cap I_T)$ for every measurable set $E \subseteq \mathbb{R}$, then

$$\int_{\mathbb{R}} f \, d\nu = \int_{\mathbb{T}} g \, d\hat{\sigma}\nu.$$

Proof: It is clear from the definition of $\hat{\rho}$ that $\hat{\rho}\mu(E) = 0$ for every set E such that $E \cap I_T = \emptyset$, hence $\int_{\mathbb{R}\setminus I_T} f \, d\hat{\rho}\mu = 0$ and

$$\int_{\mathbb{R}} f \, d\hat{\rho}\mu = \int_{I_T} f \, d\hat{\rho}\mu$$

Since π_T is a homeomorphism between I_T and $\mathbb{T} \setminus \{-1\}$, π_T^{-1} is a measurable map between those two sets, and $g = f \circ \pi_T^{-1}$ on $\mathbb{T} \setminus \{-1\}$. Then Theorem C in [24, p. 163] (with $X = \mathbb{T} \setminus \{-1\}, Y = I_T$) implies that

$$\int_{I_T} f \, d\hat{\rho}\mu = \int_{\mathbb{T}\setminus\{-1\}} g \, d\mu$$

To prove that $\|\hat{\rho}\mu\| \leq \|\mu\|$, given any $f \in C_0(\mathbb{R})$ with $\|f\|_\infty \leq 1$, define a corresponding function g on $\mathbb{T}\setminus\{-1\}$ as $g = f \circ \pi_T^{-1}$. Clearly, g is a measurable function (although not necessarily continuous on $\mathbb{T}$). Then

$$\left| \int_{\mathbb{R}} f \, d\hat{\rho}\mu \right| = \left| \int_{\mathbb{T}\setminus\{-1\}} g \, d\mu \right| \leq \|\mu\| \|g\|_\infty \leq \|\mu\|$$

and this implies $\|\hat{\rho}\mu\| \leq \|\mu\|$.
Similarly, let $E \subseteq \mathbb{R}$ be any measurable set such that: $E \cap I_T = \emptyset$. Then: $\nu(E) = \nu(E \cap I_T) = \nu(\emptyset) = 0$, and this implies $\int_{\mathbb{R}\setminus I_T} f \, d\nu = 0$. Hence

$$\int_{\mathbb{R}} f \, d\nu = \int_{I_T} f \, d\nu = \int_{\mathbb{T}} g \, d\hat{\sigma}\nu$$

where the second equality follows once again from Theorem C in [24, p. 163], with: $X = I_T, Y = \mathbb{T}$. □

Finally, define a map ρ from $B(\mathbb{Z})$ into $B(\mathbb{R})$ as

$$\rho = \mathcal{F}^{-1} \circ \hat{\rho} \circ \mathcal{F}$$

In the case of a finite-length signal, ρ has the effect of interpolating it with sinc functions, as shown by the following theorem.

Theorem 8: Let x be a finite-length discrete-time signal. Then $x \in B(\mathbb{Z})$ and

$$[\rho x](t) = \sum_k x(k) \operatorname{sinc}(t/T - k).$$

Proof: Since x has only a finite number of nonzero values, it follows that $d\hat{x} = \sum_n x(n) e^{-jn\theta} \, d\theta \in L^1(\mathbb{T}) \subseteq M(\mathbb{T})$. Consequently $x \in B(\mathbb{Z})$ and, by definition

$$[\rho x](t) = [\mathcal{F}^{-1}(\hat{\rho}\hat{x})](t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{j\omega t} \, d\hat{\rho}\hat{x}.$$

It is easy to verify that the hypotheses in Lemma 1 are satisfied by setting $f(\omega) = e^{j\omega t}, g(e^{j\theta}) = e^{j\theta t/T}$ and $\mu = \hat{x}$. Therefore

$$\begin{aligned}[\rho x](t) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{j\omega t}\, d\hat{\rho}\hat{x} \\ &= \frac{1}{2\pi}\int_{\mathbb{T}\setminus\{-1\}} e^{j\theta t/T}\left(\sum_k x(k)e^{-jk\theta}\right) d\theta \\ &= \sum_k x(k)\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j(t/T-k)\theta}\, d\theta \\ &= \sum_k x(k)\mathrm{sinc}(t/T-k). \qquad \square\end{aligned}$$

The remainder of this section is devoted to establishing two important properties of ρ. The first is that, when applied to the samples of a signal whose pseudospectrum has finite bandwidth, ρ has the effect of recreating the original signal, provided that the sampling rate is sufficiently high. This property is expressed formally by the following theorem. □

Theorem 9: Let $x \in B(\mathbb{R})$, and assume that

$$\hat{x}(E) = \hat{x}(E \cap I_T)$$

for every measurable set $E \subseteq \mathbb{R}$. Then

a) $\sigma x = (\mathcal{F}^{-1} \circ \hat{\sigma} \circ \mathcal{F})x$ or, equivalently, $\widehat{\sigma x} = \hat{\sigma}\hat{x}$;
b) $(\hat{\rho} \circ \hat{\sigma})\hat{x} = \hat{x}$;
c) $(\rho \circ \sigma)x = x$.

Proof: The assumption about $\hat{x}$ implies that

$$\begin{aligned}x(nT) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{jn\omega T}\, d\hat{x} \\ &= \frac{1}{2\pi}\int_{\mathbb{T}} e^{jn\theta}\, d\hat{\sigma}\hat{x}\end{aligned}$$

where the second equality follows from Lemma 1, with $f(\omega) = e^{jn\omega t}, g(e^{j\theta}) = e^{jn\theta}$ and $\nu = \hat{x}$. By the uniqueness of the Fourier transform, this proves that $\widehat{\sigma x} = \hat{\sigma}\hat{x}$. To prove b), let E be any measurable subset of $\mathbb{R}$. Then

$$\begin{aligned}[(\hat{\rho} \circ \hat{\sigma})\hat{x}](E) &= \hat{\sigma}\hat{x}[\pi_T(E \cap I_T)] \\ &= \hat{x}[\pi_T^{-1}[\pi_T(E \cap I_T)]] \\ &= \hat{x}[\pi_T^{-1}[\pi_T(E \cap I_T)] \cap I_T].\end{aligned}$$

It is easy to verify that $\pi_T^{-1}[\pi_T(E \cap I_T)] \cap I_T = E \cap I_T$, hence

$$(\hat{\rho} \circ \hat{\sigma})\hat{x}(E) = \hat{x}(E \cap I_T) = \hat{x}(E)$$

i.e., $(\hat{\rho} \circ \hat{\sigma})\hat{x} = \hat{x}$. Finally

$$\begin{aligned}(\rho \circ \sigma)x &= (\mathcal{F}^{-1} \circ \hat{\rho} \circ \mathcal{F}) \circ (\mathcal{F}^{-1} \circ \hat{\sigma} \circ \mathcal{F})x \\ &= (\mathcal{F}^{-1} \circ \hat{\rho} \circ \hat{\sigma})\hat{x} = \mathcal{F}^{-1}\hat{x} = x. \qquad \square\end{aligned}$$

Theorems 8 and 9 by themselves do not guarantee that the sampling expansion of a band-limited signal converges in any way to the original signal. In order to reach that conclusion, it is also necessary to show that the maps involved are continuous with respect to suitably chosen topologies. This fact is established in Theorem 10 below, and the proof of that theorem requires the following lemma.

Lemma 2: Let $\{\mu_n\}$ be a sequence of measures in $M(\mathbb{T})$ such that

$$\begin{aligned}w^*\text{-}\lim_{n\to\infty}\mu_n &= \mu \\ \limsup_{n\to\infty}\|\mu_n\| &\le \|\mu\|.\end{aligned}$$

Let $C \subseteq \mathbb{T}$ be a closed set such that $C \cap \mathrm{supp}(\mu) = \emptyset$. Then $\lim_{n\to\infty}|\mu_n|(C) = 0$.

Proof: This proof is a slight modification of the Proof of Lemma 3.9 in [25]. Given $\varepsilon > 0$, there exists $f \in C_0(\mathbb{T})$ such that $\|f\|_\infty \le 1$ and $\|\mu\| \le |\langle f, \mu\rangle| + \varepsilon/3$. Let A, B be two disjoint open sets such that $\mathrm{supp}(\mu) \subset A, C \subset B$. By Urysohn's lemma, there exists a real-valued function $h \in C_0(\mathbb{T})$ such that $0 \le h \le 1, h(z) = 1$ for $z \in \mathrm{supp}(\mu)$ and $h(z) = 0$ for $z \in \mathbb{T} \setminus A$. Let $g = hf$ and $K = \mathrm{supp}(g)$. Clearly $\|g\|_\infty \le 1$ and $\langle f, \mu\rangle = \langle g, \mu\rangle$. By assumption, there exists an integer N_ε such that $n > N_\varepsilon$ implies

$$\begin{aligned}|\langle g, \mu_n - \mu\rangle| &< \varepsilon/3 \\ \|\mu_n\| &< \|\mu\| + \varepsilon/3\end{aligned}$$

Then

$$\begin{aligned}\|\mu\| &\le |\langle f, \mu\rangle| + \varepsilon/3 = |\langle g, \mu\rangle| + \varepsilon/3 \\ &\le |\langle g, \mu_n\rangle| + |\langle g, \mu - \mu_n\rangle| + \varepsilon/3 \\ &\le |\mu_n|(K) + \varepsilon/3 + \varepsilon/3\end{aligned}$$

Consequently

$$\begin{aligned}|\mu_n|(\mathbb{T} \setminus K) &= |\mu_n|(\mathbb{T}) - |\mu_n|(K) = \|\mu_n\| - |\mu_n|(K) \\ &\le \|\mu\| + \varepsilon/3 - \|\mu\| + 2\varepsilon/3 = \varepsilon.\end{aligned}$$

Since $C \subseteq \mathbb{T} \setminus K$, it follows that $|\mu_n|(C) \le |\mu_n|(\mathbb{T} \setminus K) < \varepsilon$ for $n > N_\varepsilon$, and this completes the proof. □

Theorem 10: Let $\{\mu_n\}$ be a sequence of measures in $M(\mathbb{T})$ such that

$$\begin{aligned}w^*\text{-}\lim_{n\to\infty}\mu_n &= \mu \\ \limsup_{n\to\infty}\|\mu_n\| &\le \|\mu\|\end{aligned}$$

and assume that $-1 \notin \mathrm{supp}(\mu)$. Then $\|\hat{\rho}\mu\| = \|\mu\|$ and

$$\begin{aligned}w^*\text{-}\lim_{n\to\infty}\hat{\rho}\mu_n &= \hat{\rho}\mu \\ \limsup_{n\to\infty}\|\hat{\rho}\mu_n\| &\le \|\hat{\rho}\mu\|.\end{aligned}$$

Proof: Given an arbitrary $f \in C_0(\mathbb{R})$, define a corresponding function $g = f \circ \pi_T^{-1}$ on $\mathbb{T}$, i.e.

$$g(e^{j\theta}) = f(\theta/T) - \pi \le \theta < \pi$$

Note that g is continuous on $\mathbb{T} \setminus \{-1\}$, but not on $\mathbb{T}$, unless $f(-\omega_T/2) = f(\omega_T/2)$. Since $-1 \notin \mathrm{supp}(\mu)$, there exists $0 \le$

$\theta_0 < \pi$ such that the support of μ is contained in the arc of the unit circle defined by $|\theta| \leq \theta_0$, that is

$$\mathrm{supp}(\mu) \subseteq \{z \in \mathbb{T} : z = e^{j\theta}, \quad |\theta| \leq \theta_0\}$$

Let $\Delta\theta = (\pi - \theta_0)/2 > 0$ and let A be the arc of the unit circle defined by

$$A = \{z \in \mathbb{T} : z = e^{j\theta}, \quad |\theta| \leq \theta_0 + \Delta\theta\}$$

Let h be a continuous function on $\mathbb{T}$ such that $0 \leq h \leq 1, h = 1$ on A, and $h(-1) = 0$, e.g.

$$h(e^{j\theta}) = \begin{cases} 1 & |\theta| \leq \theta_0 + \Delta\theta \\ \dfrac{\pi - |\theta|}{\pi - \theta_0 - \Delta\theta} & \theta_0 + \Delta\theta \leq |\theta| \leq \pi. \end{cases}$$

By Lemma 1

$$\begin{aligned} \langle f, \hat{\rho}\mu_n \rangle &= \int_{\mathbb{T}\setminus\{-1\}} g \, d\mu_n \\ &= \int_{\mathbb{T}\setminus\{-1\}} hg \, d\mu_n + \int_{\mathbb{T}\setminus\{-1\}} (1-h) g \, d\mu_n. \end{aligned}$$

Note that

$$\int_{\mathbb{T}\setminus\{-1\}} hg \, d\mu_n = \int_{\mathbb{T}} hg \, d\mu_n = \langle hg, \mu_n \rangle$$

because $h(-1) = 0$. Unlike g, hg is a continuous function on $\mathbb{T}$, hence

$$\lim_{n\to\infty} \langle hg, \mu_n \rangle = \langle hg, \mu \rangle.$$

But $\langle hg, \mu \rangle = \langle g, \mu \rangle$, because $hg = g$ on the support of μ, and $\langle g, \mu \rangle = \int_{\mathbb{T}\setminus\{-1\}} g \, d\mu$ because: $\mu(\{-1\}) = 0$. Therefore

$$\lim_{n\to\infty} \int_{\mathbb{T}\setminus\{-1\}} hg \, d\mu_n = \int_{\mathbb{T}\setminus\{-1\}} g \, d\mu = \langle f, \hat{\rho}\mu \rangle.$$

Let C denote the support of $1 - h$. Then

$$\begin{aligned} &\left| \int_{\mathbb{T}\setminus\{-1\}} (1-h) g \, d\mu_n \right| \\ &\qquad = \left| \int_{C\setminus\{-1\}} (1-h) g \, d\mu_n \right| \leq \|g\|_\infty |\mu_n|(C). \end{aligned}$$

But it is clear that $C \cap \mathrm{supp}(\mu) = \emptyset$. Therefore Lemma 2 above implies that $\lim_{n\to\infty} \int_{\mathbb{T}\setminus\{-1\}} (1-h) g \, d\mu_n = 0$. Hence $\lim_{n\to\infty} \langle f, \hat{\rho}\mu_n \rangle = \langle f, \hat{\rho}\mu \rangle$ and, since f is arbitrary, this proves that $w^*\text{-}\lim_{n\to\infty} \hat{\rho}\mu_n = \hat{\rho}\mu$.

To prove the remaining claims, given an arbitrary $\varepsilon > 0$, choose $g \in C(\mathbb{T})$ such that $\|g\|_\infty \leq 1$ and $|\langle g, \mu \rangle| > \|\mu\| - \varepsilon$. Define the following function on $\mathbb{R}$:

$$f(\omega) = \begin{cases} [(hg) \circ \pi_T](\omega) & |\omega| \leq \omega_T/2 \\ 0 & |\omega| > \omega_T/2. \end{cases}$$

Note that $f(\pm\omega_T/2) = (hg)(-1) = 0$; therefore $f \in C_0(\mathbb{R})$ and: $\|f\|_\infty \leq 1$. Then

$$|\langle f, \hat{\rho}\mu \rangle| = |\langle hg, \mu \rangle| = |\langle g, \mu \rangle| > \|\mu\| - \varepsilon.$$

Hence $\|\hat{\rho}\mu\| \geq \|\mu\| - \varepsilon$ and, in fact $\|\hat{\rho}\mu\| \geq \|\mu\|$, because ε is arbitrary. This and the opposite inequality established in Lemma 1 imply $\|\hat{\rho}\mu\| = \|\mu\|$. Then

$$\limsup_{n\to\infty} \|\hat{\rho}\mu_n\| \leq \limsup_{n\to\infty} \|\mu_n\| \leq \|\mu\| = \|\hat{\rho}\mu\|. \qquad \square$$

The ramifications of the results established in Theorems 9 and 10 are quite broad. As will be shown in the sections that follow, their usefulness is not limited to the specific topic of sampling and reconstruction, but extends to other areas, such as the spectral analysis of signals.

VI. Cesàro Summation of Sampling Expansions

The results on sampling and reconstruction obtained above make it possible to establish additional convergence properties of the sampling series that go beyond those stated in Theorem 6. Specifically, it will be shown in this section that Cesàro summation of the sampling expansion improves its convergence properties in a significant manner. This mirrors a similar, well-known result about the summation of Fourier series.

Let c_n denote the nth Cesàro mean of the partial sums of the sampling expansion, that is

$$c_n(t) = \frac{1}{n+1} \sum_{m=0}^{n} x_m(t)$$

where

$$x_n(t) = \sum_{k=-n}^{n} x(kT) \,\mathrm{sinc}\,(t/T - k).$$

It is a well-known property of the Cesàro means that, if $\{x_n\}$ is a convergent sequence, so is $\{c_n\}$, and both have the same limit. It follows immediately that, if the assumptions in Theorem 6 are satisfied, then

$$\lim_{n\to\infty} c_n(t) = x(t), \quad \forall t \in \mathbb{R}.$$

Considerably more, however, can be said about the convergence properties of $\{c_n\}$. As a starting point, note that

$$\begin{aligned} c_n(t) &= \frac{1}{n+1} \sum_{m=0}^{n} \sum_{k=-m}^{m} x(kT) \,\mathrm{sinc}\,(t/T - k) \\ &= \frac{1}{n+1} \sum_{k=-n}^{n} \sum_{m=|k|}^{n} x(kT) \mathrm{sinc}(t/T - k) \\ &= \frac{1}{n+1} \sum_{k=-n}^{n} (n - |k| + 1) x(kT) \mathrm{sinc}(t/T - k) \\ &= \sum_{k=-n}^{n} [1 - |k|/(n+1)] x(kT) \,\mathrm{sinc}\,(t/T - k). \end{aligned}$$

It follows from Theorem 8 that $c_n = \rho(w_n \sigma x)$, where w_n is the discrete-time triangular window

$$w_n(k) = \begin{cases} 1 - |k|/(n+1) & |k| \leq n \\ 0 & |k| > n \end{cases} \tag{4}$$

Apart from a scale factor, this window is the autocorrelation of the rectangular window. More precisely, $w_n = \chi_{[0,n]} * \chi^*_{[0,n]}/(n+1)$, where

$$\chi_{[0,n]}(k) = \begin{cases} 1, & 0 \le k \le n \\ 0, & k \notin [0,n]. \end{cases}$$

Consequently, $\hat{w}_n = |\hat{\chi}_{[0,n]}|^2/(n+1)$ and $\|\hat{w}_n\|_1 = \|\hat{\chi}_{[0,n]}\|_2^2/(n+1)$. Using Parseval's equality: $\|\hat{\chi}_{[0,n]}\|_2^2 = 2\pi\|\chi_{[0,n]}\|_2^2$ one concludes that

$$\|\hat{w}_n\|_1 = 2\pi(n+1)/(n+1) = 2\pi$$

By relying on this equality, it is possible to prove the following lemma.

Lemma 3: Let $x \in B(\mathbb{Z})$, and let w_n be the triangular window defined in (4). Then $w_n x \in B(\mathbb{Z})$ and $\|\widehat{w_n x}\| \le \|\hat{x}\|$ or, equivalently, $\|w_n x\|_B \le \|x\|_B$. Furthermore

$$w^*\text{-}\lim_{n\to\infty} w_n x = x.$$

Proof: Note that $w_n \in B(\mathbb{Z})$ because $\hat{w}_n \in L^1(\mathbb{T}) \subseteq M(\mathbb{T})$. Consequently $w_n x \in B(\mathbb{Z})$, because $B(\mathbb{Z})$ is an algebra. It is also clear that $\lim_{n\to\infty}(w_n x)(k) = x(k) \quad \forall\, k \in \mathbb{Z}$. Moreover $\|\widehat{w_n x}\| = \|\hat{w}_n * \hat{x}/2\pi\| \le \|\hat{w}_n\|_1\|\hat{x}\|/2\pi = \|\hat{x}\|$ and, consequently, $\|w_n x\|_B = \|\widehat{w_n x}\|/2\pi \le \|\hat{x}\|/2\pi = \|x\|_B$. Then the last claim follows immediately from Lemma 5 in the Appendix. □

The convergence properties of $\{c_n\}$ can now be stated as follows.

Theorem 11: Let $x \in B(\mathbb{R})$ and assume $\text{supp}(\hat{x}) \subseteq [-\omega_0, \omega_0]$ and $\omega_0 < \omega_T/2$. Then $\|c_n\|_B \le \|x\|_B$ and

$$w^*\text{-}\lim_{n\to\infty} c_n = x$$
$$\lim_{n\to\infty} \|c_n\|_B = \|x\|_B$$

Proof: Let $y_n = w_n \sigma x$. It follows from Lemma 3 above and Theorem 18 in the Appendix that

$$w^*\text{-}\lim_{n\to\infty} \hat{y}_n = \widehat{\sigma x}$$

Note that $\widehat{\sigma x} = \hat{\sigma}\hat{x}$, because $\text{supp}(\hat{x}) \subseteq I_T$, so that the hypotheses of Theorem 9 are satisfied. Furthermore, the support of $\hat{\sigma}\hat{x}$ is contained in the arc of the unit circle defined by $|\theta| \le \theta_0$, that is

$$\text{supp}(\hat{\sigma}\hat{x}) \subseteq \{z \in \mathbb{T} : z = e^{j\theta}, \quad |\theta| \le \theta_0\}$$

where $\theta_0 = \omega_0 T < \pi$. Hence $-1 \notin \text{supp}(\hat{\sigma}\hat{x})$. Since $\|\hat{y}_n\| \le \|\widehat{\sigma x}\| = \|\hat{\sigma}\hat{x}\|$, it is clear that $\limsup_{n\to\infty}\|\hat{y}_n\| \le \|\hat{\sigma}\hat{x}\|$. Therefore, the hypotheses of Theorem 10 are also satisfied, and

$$w^*\text{-}\lim_{n\to\infty} \hat{\rho}\hat{y}_n = \hat{\rho}(\hat{\sigma}\hat{x}) = \hat{x}$$
$$\|\hat{\sigma}\hat{x}\| = \|\hat{\rho}(\hat{\sigma}\hat{x})\| = \|\hat{x}\|.$$

Since $c_n = \rho(w_n \sigma x) = \mathcal{F}^{-1}(\hat{\rho}\hat{y}_n)$, it follows once again from Theorem 18 in the Appendix that

$$w^*\text{-}\lim_{n\to\infty} c_n = x$$

Furthermore

$$\|c_n\|_B = \|\hat{\rho}\hat{y}_n\|/2\pi \le \|\hat{y}_n\|/2\pi \le \|\hat{\sigma}\hat{x}\|/2\pi = \|\hat{x}\|/2\pi = \|x\|_B$$

and, consequently, $\limsup_{n\to\infty}\|c_n\|_B \le \|x\|_B$. The last claim is then an immediate consequence of Theorem 17 in the Appendix. □

Note that the results of this theorem are not applicable to $\{x_n\}$, the sequence of the partial sums of the sampling expansion. In such case, the triangular window w_n would have to be replaced with the rectangular window $\chi_{[-n,n]}$ in all the previous derivations, as can be easily verified. But $\|\chi_{[-n,n]}\|_B \to +\infty$ as $n \to \infty$, and the argument used in the Proof of Lemma 3 fails.

Theorem 11 makes a stronger statement than Theorem 6, because it can be shown that the type of convergence established by Theorem 11 implies uniform convergence on bounded intervals of $\mathbb{R}$, but is not implied by it [25, Corollary 1]. This means that $\{c_n\}$ has convergence properties that do not necessarily hold for $\{x_n\}$. For example, the inequality $\|c_n\|_\infty \le \|c_n\|_B$ (see the Appendix) implies that the Cesàro means of the partial sums of (3) are uniformly bounded over $\mathbb{R}$.

VII. Ergodic Theorems

According to the sampling theorems in Sections IV and VI, the time-domain Nyquist rate for a signal in $B(\mathbb{R})$— i.e., the lowest sampling frequency that guarantees exact reconstruction of the signal from its samples—is determined by the bandwidth of the pseudospectrum of the signal. As noted earlier, the highest frequency in the pseudospectrum may be higher than the highest frequency in the power spectrum, unless the signal is almost periodic. This observation raises the issue of whether the same lower bound on the sampling frequency must be maintained when sampling is used to estimate the signal power spectrum. Since digital spectrum analyzers estimate power spectra from signal samples, the answer to this question is not of mere theoretical interest, but has also important practical consequences. Fortunately it turns out that, at least in the case of weakly almost periodic signals, the minimum sampling rate that is needed to obtain alias-free spectral estimates is determined by the bandwidth of the power spectrum, not that of the pseudospectrum.

The formal proof of this fact, which will be given in Section VIII, relies on the ergodic theorems stated below. Another practical application of the theorems of this section, which concerns the asymptotic behavior of the spectral estimates obtained from the averaged periodogram, will be discussed in Section IX. It should be noted that the classic ergodic theorems of Birkhoff and Von Neumann are valid in $L^p(\mathbb{R})$ for $1 \le p < \infty$ [26], and cannot therefore be applied to WAP $(\mathbb{R})$.

The first two theorems that follow contain references to the closed convex hull of a set. Recall that, if S is a subset of a Banach space, its closed convex hull, denoted by $\overline{\text{co}}\{S\}$, is defined as the norm closure of the set of convex combinations of elements of S [27, p. 414].

Theorem 12: Let $\mathbf{T}$ be an invertible linear operator on a Banach space X such that

$$\|\mathbf{T}^n\| \le M \quad \forall n \in \mathbb{Z}$$

for some $M > 0$. For $x \in X$, define

$$\mathbf{E}_n x = \frac{1}{2n+1} \sum_{k=-n}^{n} \mathbf{T}^k x$$

and let $C_x = \overline{\text{co}}\{\mathbf{T}^n x : n \in \mathbb{Z}\}$. If C_x is weakly compact in X, the sequence $\{\mathbf{E}_n x\}$ converges in norm to a point $x_0 \in C_x$:

$$\lim_{n\to\infty} \|\mathbf{E}_n x - x_0\| = 0$$

Moreover x_0 is a fixed point for $\mathbf{T}$, that is $\mathbf{T}x_0 = x_0$.

Proof: This theorem is a simplified version of [28, Theorem 24.13, p. 355], specialized to the case in which the group acting on X is $\mathbb{Z}$. □

A crucial assumption for the validity of this generic ergodic theorem is that the closed convex hull of $\{\mathbf{T}^n x\}$ should be compact in the weak topology of X. It turns out that this condition is always satisfied when x is a weakly almost periodic function, and $\mathbf{T}$ is a translation operator. Therefore the following theorem is essentially a special case of Theorem 12. It should be noted that in this theorem G is not restricted to being $\mathbb{R}$ or $\mathbb{Z}$.

Theorem 13: Let G be a commutative topological group, and $g_0 \in G$. Let $\mathbf{T}$ be the translation operator on WAP(G), defined by

$$(\mathbf{T}x)(g) = x(g+g_0) \quad x \in \text{WAP}(G).$$

Then for every $x \in \text{WAP}(G)$ the sequence $\{\mathbf{E}_n x\}$ converges in norm to an element $x_0 \in \text{WAP}(G)$, that is

$$\lim_{n\to\infty} \left\| \frac{1}{2n+1} \sum_{k=-n}^{n} \mathbf{T}^k x - x_0 \right\|_\infty = 0.$$

Moreover, x_0 is a periodic function of period g_0.

Proof: It is clear that $\mathbf{T}^n$ is a translation operator for all values of n, and therefore: $\|\mathbf{T}^n\| = 1$. By definition, for a given $x \in \text{WAP}(G)$ the set: $\{\mathbf{T}^n x : n \in \mathbb{Z}\}$ is relatively compact in the weak topology of WAP(G). As a consequence of the Krein-Šmulian theorem [27, p. 434], $C_x = \overline{\text{co}}\{\mathbf{T}^n x : n \in \mathbb{Z}\}$ is also weakly compact. The claim is then a straightforward consequence of Theorem 12, when one notes that $\mathbf{T}x_0 = x_0$ is equivalent to $x_0(g+g_0) = x_0(g) \quad \forall g \in G$. An alternate proof of this theorem can be obtained by relying on Theorem 5.1 in [18]. □

Theorem 13 and the next lemma lay the foundation for the main result of this section, which is stated in Theorem 14 below.

Lemma 4: Let $x \in \text{WAP}(\mathbb{R})$, and define

$$s(t,\omega) = \lim_{T\to+\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t+\tau)\overline{x}(\tau) e^{-j\omega\tau}\, d\tau.$$

If $\text{supp}(\hat{s}_x) \subseteq [-\omega_0, \omega_0]$ and $|\omega| > 2\omega_0$, then $s(t,\omega) = 0$, $\forall t \in \mathbb{R}$.

Proof: For a fixed value of ω, Theorem 1 implies that $s \in AP(\mathbb{R})$ as a function of t. By Theorem 3, it suffices to prove that, if $|\omega| > 2\omega_0$, then

$$\alpha_s(\zeta) = \lim_{T\to+\infty} \frac{1}{T} \int_{-T/2}^{T/2} s(t,\omega) e^{-j\zeta t}\, dt = 0 \quad \forall \zeta \in \mathbb{R}.$$

By Theorem 2

$$\lim_{T\to+\infty} \frac{1}{T} \int_{-T/2}^{T/2} s(t,\omega) e^{-j\zeta t}\, dt = \lim_{T\to+\infty} \frac{1}{T} \int_{-T/2}^{T/2} \gamma(\tau)\overline{x}(\tau) e^{-j\omega\tau}\, d\tau$$

where

$$\gamma(\tau) = \lim_{T\to+\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t+\tau) e^{-j\zeta t}\, dt = \alpha_x(\zeta) e^{j\zeta\tau}.$$

Therefore

$$\begin{aligned}\alpha_s(\zeta) &= \alpha_x(\zeta) \lim_{T\to+\infty} \frac{1}{T} \int_{-T/2}^{T/2} \overline{x}(\tau) e^{j(\zeta-\omega)\tau}\, d\tau \\ &= \alpha_x(\zeta)\overline{\alpha}_x(\zeta-\omega).\end{aligned}$$

Assume $\alpha_x(\zeta) \ne 0$. By Theorem 5, this means that $\zeta \in \Omega_x \subseteq \text{supp}(\hat{s}_x)$, and thus $|\zeta| \le \omega_0$. But then $|\zeta-\omega| \ge |\omega| - |\zeta| > \omega_0$ and, consequently, $\zeta - \omega \notin \Omega_x$. This implies $\alpha_x(\zeta-\omega) = 0$, which proves the claim. □

The next theorem, and the last in this section, is to a large extent a special case of Theorem 13. Essentially it implies that the integral in (2) that defines the autocorrelation of a signal can be replaced, under certain circumstances, by an infinite series. This fact has several notable consequences: for example, it will be used in Section VIII to prove that the power spectrum of σx is a frequency-scaled copy of the power spectrum of x, if the sampling frequency is greater than twice the highest frequency in $\hat{s}_x$. Another, somewhat unexpected consequence of Theorem 14 relates to the properties of the averaged periodogram as a spectral estimator, and is discussed in Section IX.

Theorem 14: Let $x \in \text{WAP}(\mathbb{R})$, and let $\hat{s}_x$ be its power spectrum. Then there exists $s_{x,T} \in \text{WAP}(\mathbb{R}\times\mathbb{R})$ such that

$$s_{x,T}(t,\tau) = \lim_{n\to\infty} \frac{1}{2n+1} \sum_{k=-n}^{n} x(t+\tau+kT)\overline{x}(\tau+kT) \quad (5)$$

with convergence being uniform on $\mathbb{R}\times\mathbb{R}$. Moreover, if $\text{supp}(\hat{s}_x) \subseteq [-\omega_0, \omega_0]$ and $\omega_0 < \omega_T/2$, then $s_{x,T}(t,\tau) = s_x(t)$, that is, the limit in (5) is independent of τ and coincides with the autocorrelation of x. In such case, the autocorrelation of the samples of x equals the samples of the autocorrelation of x, i.e., $s_{\sigma x} = \sigma s_x$.

Proof: $\mathbb{R}\times\mathbb{R}$ is a commutative topological group, and $x(t+\tau)\overline{x}(\tau) \in \text{WAP}(\mathbb{R}\times\mathbb{R})$ [18]. Let $g_0 = (0,T) \in \mathbb{R}\times\mathbb{R}$. Then the existence of $s_{x,T} \in \text{WAP}(\mathbb{R}\times\mathbb{R})$ such that (5) holds, and the fact that the limit is uniform on $\mathbb{R}\times\mathbb{R}$, are a consequence of Theorem 13, which also implies that $s_{x,T}$ is periodic in τ of period T: $s_{x,T}(t,\tau+T) = s_{x,T}(t,\tau)\ \forall(t,\tau) \in \mathbb{R}\times\mathbb{R}$. Therefore, for every fixed $t \in \mathbb{R}$, $s_{x,T}$ can be expanded in a Fourier series with respect to τ:

$$s_{x,T}(t,\tau) = \sum_{m=-\infty}^{+\infty} a_m(t) e^{jm\omega_T \tau}$$
$$a_m(t) = \frac{1}{T}\int_{-T/2}^{T/2} s_{x,T}(t,\tau) e^{-jm\omega_T\tau}\, d\tau.$$

Because the limit in (5) is uniform, it follows that

$$\begin{aligned} a_m(t) &= \lim_{n\to\infty} \frac{1}{(2n+1)T} \\ &\quad\times \sum_{k=-n}^{n} \int_{-T/2}^{T/2} x(t+\tau+kT)\overline{x}(\tau+kT) \\ &\quad\times e^{-jm\omega_T\tau}\, d\tau \\ &= \lim_{n\to\infty} \frac{1}{(2n+1)T} \\ &\quad\times \int_{-(n+1/2)T}^{(n+1/2)T} x(t+\tau)\overline{x}(\tau) e^{-jm\omega_T\tau}\, d\tau. \end{aligned}$$

If $\operatorname{supp}(\hat{s}_x) \subseteq [-\omega_0, \omega_0]$ and $\omega_0 < \omega_T/2$, Lemma 4 implies that $a_m = 0$ for $m \neq 0$. Hence

$$\begin{aligned} s_{x,T}(t,\tau) &= a_0(t) \\ &= \lim_{n\to\infty}\frac{1}{(2n+1)T}\int_{-(n+1/2)T}^{(n+1/2)T} x(t+\tau)\overline{x}(\tau)\, d\tau \\ &= s_x(t). \end{aligned}$$

In particular, setting $t = mT, \tau = 0$ in (5) yields

$$s_x(mT) = \lim_{n\to\infty}\frac{1}{2n+1}\sum_{k=-n}^{n} x[(m+k)T]\overline{x}(kT)$$

i.e., $(\sigma s_x)(m) = s_{\sigma x}(m),\ \forall m \in \mathbb{Z}$.

VIII. Sampling and Spectral Estimation

Scientists have been looking for methods to estimate the power spectrum of a signal for a long time: already before the end of the 19th century, Schuster suggested using what are known today as periodogram estimates to detect periodic patterns in weather phenomena [29]. Today spectral estimation plays an important role in a wide range of scientific disciplines, from geology to astronomy to communications, and this explains why this topic is still the object of active research.

In particular, advances in digital signal processing have spurred the development of methods that estimate the power spectrum of a continuous-time signal from a discrete-time sequence of the signal samples. A tacit assumption underlying most of those methods is that spectral aliasing can be avoided if the largest frequency in the power spectrum of the original signal is lower than half the sampling rate. The objective of this section is to prove formally that this is indeed a correct assumption, at least in the case of weakly almost periodic signals. It should be stressed that this conclusion cannot be reached by relying on either the classic sampling theorem (Theorem 6) or Theorem 11, both of which assume that the sampling rate is higher than twice the highest frequency in the pseudospectrum of the signal. As explained at the end of Section IV, this latter condition is more restrictive than requiring the sampling rate to be larger than twice the largest frequency in the power spectrum of the signal. It should also be noted that, unlike Theorems 6 and 11, the results of this section apply to all weakly almost periodic signals, and not just to those that are elements of $B(\mathbb{R})$.

Let $x \in \mathrm{WAP}(\mathbb{R})$ and, as before, let σx be the sequence of the samples of x. The issue that must be addressed is whether, and under what conditions, the power spectrum of σx is a frequency-scaled copy of the power spectrum of x. The answer is provided by the following theorem, which is a straightforward consequence of the results obtained in Sections V and VII.

Theorem 15: Let $x \in \mathrm{WAP}(\mathbb{R})$, and assume that $\operatorname{supp}(\hat{s}_x) \subseteq [-\omega_0, \omega_0]$ and $\omega_0 < \omega_T/2$. Then

$$\hat{s}_{\sigma x} = \hat{\sigma}\hat{s}_x.$$

Proof: Since $\operatorname{supp}(\hat{s}_x) \subseteq I_T$, s_x satisfies the assumptions in Theorem 9, hence $\widehat{\sigma s}_x = \hat{\sigma}\hat{s}_x$. On the other hand, Theorem 14 implies that $\sigma s_x = s_{\sigma x}$, therefore $\widehat{\sigma s}_x = \hat{s}_{\sigma x}$, and the claim follows. □

Note that $\hat{s}_{\sigma x}$ is the power spectrum of σx, while $\hat{\sigma}\hat{s}_x$ is the image under $\hat{\sigma}$ of the power spectrum of x. Since $\operatorname{supp}(\hat{s}_x) \subseteq I_T$, it follows that $\hat{\sigma}\hat{s}_x$ is a frequency-scaled copy of $\hat{s}_x$, as explained in Section V. Therefore, under the assumptions stated in Theorem 15, it is possible to estimate the power spectrum of x by estimating the power spectrum of σx, and many algorithms have been developed for this purpose [30].

For example, one possible approach is to compute the Fourier transform of the autocorrelation of finite-length segments of σx. More precisely, let

$$(\sigma x_n)(k) = \begin{cases} (\sigma x)(k), & |k| \le n \\ 0, & |k| > n \end{cases}$$
$$s_n = \frac{1}{2n+1}(\sigma x_n) * (\sigma x_n)^*.$$

Let w be a discrete-time window, that is, a finite-length, real-valued function satisfying the conditions $w(-t) = w(t), |w(t)| \le 1$ and $w(0) = 1$. The Fourier transform of ws_n is generally known as a windowed correlogram estimate of the power spectrum of σx. Clearly $\widehat{ws}_n = \hat{w} * \hat{s}_n/2\pi$, and it is shown in [16] that

$$\lim_{n\to\infty} \|\hat{w} * \hat{s}_n/2\pi - \hat{w} * \hat{s}_{\sigma x}/2\pi\|_1 = 0.$$

In other words, the correlogram estimates converge in the L^1 norm to $\hat{w} * \hat{s}_{\sigma x}/2\pi = \widehat{ws}_{\sigma x}$. Note, however, that $\hat{s}_{\sigma x}$ is a discrete spectrum, but $\widehat{ws}_{\sigma x}$ is not: windowing has a "smearing" effect on discrete spectra, and each line is transformed into a copy of $\hat{w}$, centered at the location of the line. For this reason, it may be preferable to use other algorithms (e.g., Pisarenko's method [30], [31]), which are intended specifically for the estimation of discrete spectra, and which do not suffer from this drawback.

IX. Averaged Periodogram

The study of weakly almost periodic signals presented in this paper concludes with another application of the ergodic theorems of the previous sections to spectral estimation. Specifically, those theorems will be used to analyze the asymptotic behavior of the averaged periodogram. It is worth stating explicitly that the results of this section are applicable to both continuous-time ($G=\mathbb{R}$) and discrete-time ($G=\mathbb{Z}$) signals, and the notation should be interpreted accordingly, as explained in Section II.

In its most basic form, a periodogram estimate is, apart from a scale factor, simply the magnitude squared of the Fourier transform of a finite-length segment of the signal. More precisely, let

$$\hat{p}=\frac{1}{W}\mathcal{F}[(wx)*(wx)^*]=\frac{1}{W}|\widehat{wx}|^2$$

where w is a (continuous-time or discrete-time) window and W is a scale factor whose value is discussed below. By definition, $\hat{p}$ is a *periodogram estimate* of the power spectrum of x. If x is a realization of a stationary stochastic process, it can be shown that the expected value of $\hat{p}$ is

$$\mathcal{E}\{\hat{p}\}=\frac{1}{2\pi W}\hat{s}_x*|\hat{w}|^2 \tag{6}$$

where s_x denotes the autocorrelation function of the process: $s_x(t)=\mathcal{E}\{x(t+\tau)\overline{x}(\tau)\}$.

The value of W is normally chosen so that $2\pi W=\|\hat{w}\|_2^2$, because this makes $\hat{p}$ an asymptotically unbiased estimator. Its variance, however, becomes approximately equal to $\hat{s}_x^2$ as the length of w increases [32]. In order to alleviate this problem it is common practice to average periodogram estimates taken over multiple windows, as described below.

Given w and $T>0$, let $w_k(t)=w(t-kT)$. Then

$$\hat{p}_n=\frac{1}{(2n+1)W}\sum_{k=-n}^{n}|\widehat{w_kx}|^2$$

is the *averaged periodogram estimate* determined by $w_{-n},w_{-n+1},\ldots,w_n$. Note that here T denotes the separation interval between two consecutive windows; if T is smaller than the length of w, then those windows overlap. Note also that $w_kx\in L^2(G)$, because both w_k and x are bounded, and w_k has finite length. It follows from the Plancherel theorem [14, p. 26] that $\widehat{w_kx}\in L^2(\hat{G})$, that is, $|\widehat{w_kx}|^2\in L^1(\hat{G})$ and consequently $\hat{p}_n\in L^1(\hat{G})$. It is then straightforward to verify that the inverse Fourier transform of $\hat{p}_n$ is

$$p_n(t)=\frac{1}{W}\int_G w(t+\tau)w(\tau)\times\sum_{k=-n}^{n}\frac{x(t+\tau+kT)\overline{x}(\tau+kT)}{2n+1}\,d\tau \tag{7}$$

and that, therefore, the expected value of $\hat{p}_n$ is still given by (6). On the other hand, it can be shown that, under certain assumptions, the averaging procedure reduces the variance of the estimates by a factor that is roughly proportional to $2n+1$ [32]. This means that the consistency of the averaged periodogram, when it is used to estimate the power spectrum of a stochastic process, increases with the number of signal segments used to generate the estimate.

The averaged periodogram can also be used to estimate the power spectrum of a single signal (see, for instance, [33, ch. 4], where this method is referred to as the hopped temporally smoothed periodogram). It is then natural to ask whether (6) remains valid if $\mathcal{E}\{\hat{p}\}$ is replaced by $\lim_{n\to\infty}\hat{p}_n$, and $\hat{s}_x$ is the power spectrum of x. Theorem 16 below provides an answer to this question.[4]

Theorem 16: Let $x\in \mathrm{WAP}(G)$. Then

$$\lim_{n\to\infty}p_n(t)=\frac{1}{W}\int_G w(t+\tau)w(\tau)s_{x,T}(t,\tau)\,d\tau,\quad \forall t\in G$$

where $s_{x,T}(t,\tau)$ is the function defined in (5). If $\mathrm{supp}(\hat{s}_x)\subseteq[-\omega_0,\omega_0]$ and $\omega_0<\omega_T/2$, then

$$\lim_{n\to\infty}p_n(t)=\frac{1}{W}[(w*w^*)s_x](t) \tag{8}$$

and, consequently

$$w^*\text{-}\lim_{n\to\infty}\hat{p}_n=\frac{1}{2\pi W}\hat{s}_x*|\hat{w}|^2.$$

Proof: Note that, for each fixed value of t, the integral on the right-hand side of (7) is taken over a bounded interval of G, because, by definition, w has finite length. Since $|w(t+\tau)w(\tau)|\le 1$, Theorem 14 ensures that the integrand converges uniformly to $w(t+\tau)w(\tau)s_{x,T}(t,\tau)$ over that interval. It is therefore possible to take the limit inside the integral, and the first claim follows immediately. If $\mathrm{supp}(\hat{s}_x)\subseteq[-\omega_0,\omega_0]$ and $\omega_0<\omega_T/2$, then, by Theorem 14, $s_{x,T}(t,\tau)=s_x(t)$, and this proves (8). Since $\hat{p}_n\in L^1(\hat{G})\subseteq M(\hat{G})$, it follows that $p_n\in B(G)$. Moreover, p_n is positive definite, since it is obvious that $\hat{p}_n\ge 0$. Since s_x is also positive–definite and $s_x\in B(G)$, the remaining claim is an immediate consequence of (8) and of Corollary 1 in the Appendix. □

The result of Theorem 16 is somewhat unexpected because it shows that the separation interval between adjacent windows has an effect similar to that of a sampling period, even if in this particular instance no sampling (or downsampling, in the discrete-time case) actually takes place. This means that, in order to ensure that the averaged periodogram estimates converge to $\hat{s}_x*|\hat{w}|^2/2\pi W$, the value of T should be chosen so that $1/T$ is greater than twice the highest frequency in the power spectrum of x.

For example, let $x(t)=\cos\omega_0 t$; then $s_x(t)=(1/2)\cos\omega_0 t$ and $\hat{s}_x(\omega)=(\pi/2)[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)]$. On the other hand

$$\sum_{k=-n}^{n}\frac{\cos\omega_0(t+\tau+kT)\cos\omega_0(\tau+kT)}{2n+1}=\frac{1}{2}\cos\omega_0 t+\frac{1}{2}\frac{\sum_{k=-n}^{n}\cos[\omega_0(t+2\tau+2kT)]}{2n+1}. \tag{9}$$

If $T=\pi/\omega_0$, all the terms in the last summation above are equal, and

$$p_n(t)=\frac{1}{W}[(w*w^*)s_x](t)+\frac{1}{2W}\int_{-\infty}^{+\infty}w(t+\tau)w(\tau)\cos\omega_0(t+2\tau)\,d\tau$$

[4]See also [33, ch. 4, Exercise 2] for a similar result.

Exploiting the properties of the Fourier transform and the fact that $\hat{w}$ is an even, real-valued function, it is not too difficult to verify, after some algebraic manipulations, that

$$\hat{p}_n(\omega) = \frac{1}{2\pi W}(\hat{s}_x * |\hat{w}|^2)(\omega) + \frac{1}{2W}\hat{w}(\omega+\omega_0)\hat{w}(\omega-\omega_0).$$

This confirms that, if the conditions on ω_T stated in Theorem 16 are not satisfied, the accuracy of the averaged periodogram spectral estimates may be degraded. It is worth noting that, in contrast, the asymptotic accuracy of the correlogram estimates is not affected by a similar problem [16].

A closer analysis of the example given above would reveal that the last term in (9) tends to zero as n tends to infinity, unless $T = m\pi/\omega_0$, that is, unless ω_T is an integer multiple of $2\omega_0$. This observation may lead one to conclude that the constraint on ω_T stated in Theorem 16 is too restrictive, and that it is only necessary to ensure that $\omega_T/2$ is not an integer multiple of any of the frequency components of the power spectrum of the signal. Although this conclusion is mathematically sound, it would make little practical difference, except in those cases in which the power spectrum contains a relatively small number of frequencies, and their values are known with a certain accuracy. In fact, if all that is known about the power spectrum of the signal is a bound on its highest frequency, then the only way to ensure that $\omega_T/2$ is not an integer multiple of any spectral frequency is to choose its value as stated in Theorem 16.

X. Conclusion

The objective of this paper was to formulate a sampling theory for weakly almost periodic signals, which are a particular class of finite-power signals. The properties of the sampling expansions of such signals, stated in Sections IV through VI, match closely those that are known to hold in the finite-energy case. Even so, it is apparent from the results obtained in those sections that sampling theory becomes somewhat more complex when finite-power signals are considered. This is mainly due to the fact that, while the energy spectrum and the Fourier transform of a finite-energy signal always have the same bandwidth, a similar property does not necessarily hold for the power spectrum and the Fourier transform (i.e., the pseudospectrum) of a finite-power signal. In practice, while the power spectrum can be readily measured with a spectrum analyzer, no equivalent instruments are currently available to obtain the pseudospectrum of a signal. Consequently, theoretical results that rely on the properties of the power spectrum, instead of the pseudospectrum, are of potential practical interest as well.

For example, a more limited version of the sampling theorem, useful for the purposes of spectral analysis, has been proved in Section VIII. It shows that the minimum sampling rate necessary to avoid spectral aliasing is set, in fact, by the largest frequency in the power spectrum of a signal. As another example, it has been shown that averaged periodogram spectral estimates can be affected by an error that is related to the interval between two consecutive windows. This particular result indicates that the theory developed in this paper could potentially be useful even in the analysis of some signal processing algorithms that do not involve sampling.

Some of the results presented in this paper, in particular those in Sections VII through IX, are related to the approach to signal analysis described, for instance, in [33], [34], in which statistical parameters are defined as the limits of time averages of a single signal, rather than ensemble averages of realizations of a stochastic process. It is straightforward to verify that Theorem 1 and the ergodic theorems of Sections VII–IX, together with the fact that WAP(G) is an algebra, ensure that the limit of the time averages of the sum or product of any number of weakly almost periodic signals always exists. This implies, for instance, that every element of WAP$(\mathbb{Z})$ is a totally stationary numerical sequence, according to the definition given in [35]. In fact, every continuous-time or discrete-time weakly almost periodic signal is almost cyclostationary, as defined in [33, p. 392], because it can be decomposed as the sum of an almost periodic and a zero-power component [22].

In conclusion, weakly almost periodic signals have many useful properties that can be exploited to establish results that are relevant to both the theory and the application of signal analysis and processing. The main limitation of working within this particular mathematical framework is that weakly almost periodic signals do not include signals with a continuous power spectrum, as was shown in Section III. Further research is needed to determine whether and to what extent the results established in this paper can be extended to a broader class of finite-power signals.

Appendix

A. A Theorem in the Dual of a Banach Space

The following theorem establishes a result that is relied upon in the main body of the paper.

Theorem 17: Let $\{x_n\}$ be a sequence in X^* such that $w^*\text{-}\lim_{n\to\infty} x_n = x \in X^*$. Then $\liminf_{n\to\infty}\|x_n\| \geq \|x\|$. Consequently, if $\limsup_{n\to\infty}\|x_n\| \leq \|x\|$, then $\lim_{n\to\infty}\|x_n\| = \|x\|$.

Proof: Given an arbitrary $\varepsilon > 0$, there exists $\xi \in X$ such that $\|\xi\| \leq 1$ and $|\langle \xi, x\rangle| > \|x\| - \varepsilon/2$. There exists also an integer N_ε such that $|\langle \xi, x_n\rangle - \langle \xi, x\rangle| < \varepsilon/2$ for $n > N_\varepsilon$. These two inequalities imply that: $\|x_n\| \geq |\langle \xi, x_n\rangle| > \|x\| - \varepsilon$ for $n > N_\varepsilon$, that is $\liminf_{n\to\infty}\|x_n\| \geq \|x\|$. The remaining claim is an immediate consequence of the obvious inequality $\liminf_{n\to\infty}\|x_n\| \leq \limsup_{n\to\infty}\|x_n\|$. □

B. Fourier-Stieltjes Algebra

If $f \in L^1(G)$ and $\xi \in L^2(G)$, it can be shown that the convolution $f * \xi$ exists and is an element of $L^2(G)$ such that $\|f * \xi\|_2 \leq \|f\|_1\|\xi\|_2$ [17, pp. 396–397]. Therefore each $f \in L^1(G)$ can be identified with a linear, bounded operator $\mathbf{T}_f$ on $L^2(G)$ defined as $\mathbf{T}_f\xi = f * \xi$. The closure in the operator norm of the set $\{\mathbf{T}_f : f \in L^1(G)\}$ is denoted by $C^*(G)$. It is an algebra of operators on $L^2(G)$, and it is referred to as the C^* algebra of G [36].

It is possible to define a Fourier transform operation on $C^*(G)$ in the following way. If $f \in L^1(G)$, define $\widetilde{\mathcal{F}}f \in C_0(\hat{G})$ as

$$[\widetilde{\mathcal{F}}f](\omega) = \int_G e^{j\omega t} f(t)\, dt.$$

Since $L^1(G)$ is dense in $C^*(G), \widetilde{\mathcal{F}}$ extends by continuity to $C^*(G)$. It can be shown that, extended in this way, $\widetilde{\mathcal{F}}$ is an isometric isomorphism of $C^*(G)$ onto $C_0(\hat{G})$ [36, p. 188]. Since $M(\hat{G})$ is the dual space of $C_0(\hat{G})$, the dual space of $C^*(G)$ must be related to $M(\hat{G})$; the concrete form of this relationship will be established in Theorem 18 below.

Let $B(G)$ denote the set of all complex-valued functions that can be expressed as the inverse Fourier transform of a measure in $M(\hat{G})$. In other words, $x \in B(G)$ if

$$x(t) = [\mathcal{F}^{-1}\mu](t) = \frac{1}{2\pi}\int_{\hat{G}} e^{j\omega t}\, d\mu(\omega)$$

for some $\mu \in M(\hat{G})$. It is straightforward to verify that all the elements of $B(G)$ are bounded, continuous (in fact, uniformly continuous) functions. Furthermore, convolution of measures in $M(\hat{G})$ is mapped by the inverse Fourier transform into a product of functions in $B(G)$. In other words, if $x_1 = \mathcal{F}^{-1}\mu_1, \quad x_2 = \mathcal{F}^{-1}\mu_2$, then $x_1x_2 = \mathcal{F}^{-1}(\mu_1 * \mu_2)/2\pi$. It follows that $x_1, x_2 \in B(G) \Longrightarrow x_1x_2 \in B(G)$, i.e., $B(G)$ is a function algebra.

Each $x \in B(G)$ defines a linear, bounded functional on $L^1(G)$ through the relationship $\langle f, x\rangle = \int_G f(t)x(t)\, dt$, and Fubini's theorem implies that

$$\int_G f(t)\left(\int_{\hat{G}} e^{j\omega t}\, d\mu(\omega)\right) dt = \int_{\hat{G}}\left(\int_G f(t)e^{j\omega t}\, dt\right) d\mu(\omega)$$

i.e., $\langle f, 2\pi\mathcal{F}^{-1}\mu\rangle = \langle \widetilde{\mathcal{F}}f, \mu\rangle$. This means that $2\pi\mathcal{F}^{-1}$ is the adjoint operator of $\widetilde{\mathcal{F}}$. It is now straightforward to prove the following theorem.

Theorem 18: $B(G)$ can be identified with the dual space of $C^*(G)$, and $2\pi\mathcal{F}^{-1}$ is an isometric isomorphism of $M(\hat{G})$ onto $B(G)$. Moreover, both $\mathcal{F}$ and $\mathcal{F}^{-1}$ are continuous with respect to the weak-* topologies on $M(\hat{G})$ and $B(G)$.

Proof: Since $\widetilde{\mathcal{F}}$ is an isometry of $C^*(G)$ onto $C_0(\hat{G}), \widetilde{\mathcal{F}}^* = 2\pi\mathcal{F}^{-1}$ is an isometry of $[C_0(\hat{G})]^* = M(\hat{G})$ onto the dual space of $C^*(G)$, and is continuous with respect to the weak-* topologies on $M(\hat{G})$ and $[C^*(G)]^*$ (see Exercises 5 and 6 in [37, p. 111]). This means, in particular, that $[C^*(G)]^* = B(G)$. Since $\widetilde{\mathcal{F}}^{-1}$ is clearly also an isometry, the weak-* continuity of $\mathcal{F}$ follows from the same argument. □

It follows from this theorem that each $x \in B(G)$ is the inverse Fourier transform of a unique measure in $M(\hat{G})$, which will be denoted by $\hat{x}$. Then, by definition $x = \mathcal{F}^{-1}\hat{x}$, and, by Theorem 18, $\|x\|_B = \|\hat{x}\|/2\pi$ Note that $\|x\|_\infty \le \|x\|_B$. Under the $\|\cdot\|_B$ norm, $B(G)$ is a Banach algebra: it is referred to as the Fourier-Stieltjes algebra of G [36].

The positive–definite functions in $B(G)$ are the inverse Fourier transforms of positive measures in $M(\hat{G})$. Hence, if $x \in B(G)$ is positive definite, then

$$\|x\|_B = \frac{1}{2\pi}\hat{x}(\hat{G}) = \frac{1}{2\pi}\int_{\hat{G}} d\hat{x} = x(0) \le \|x\|_\infty.$$

Since $\|x\|_\infty \le \|x\|_B$, it follows that $\|x\|_B = \|x\|_\infty = x(0)$. This equality, however, does not generally hold for elements of $B(G)$ that are not positive definite.

Norm convergence in $B(G)$ is equivalent to norm convergence in $M(\hat{G})$. More precisely, let $x \in B(G)$, and let $\{x_n\}$ be a sequence in $B(G)$. Then

$$\lim_{n\to\infty}\|x_n - x\|_B = 0 \Longleftrightarrow \lim_{n\to\infty}\|\hat{x}_n - \hat{x}\| = 0.$$

On the other hand, $\{x_n\}$ converges to x in the weak-* topology of $B(G)$ if and only if

$$\lim_{n\to\infty}\int_G x_n(t)f(t)dt = \int_G x(t)f(t)\, dt$$

for every $f \in L^1(G)$. Weak-* convergence in $B(G)$ is closely related to pointwise convergence, as shown by the following lemma.

Lemma 5: Let $\{x_n\}$ be a sequence in $B(G)$ that convergences pointwise to a function $x \in B(G)$, that is

$$\lim_{n\to\infty} x_n(t) = x(t) \quad \forall t \in G$$

Assume that $\{x_n\}$ is bounded in the norm of $B(G)$, i.e., there exists $M > 0$ such that $\|x_n\|_B \le M$ for all n. Then

$$x = w^*\text{-}\lim_{n\to\infty} x_n.$$

Proof: See Lemma 1 in the Appendix of [16]. □

The following corollary, which is applicable to positive–definite functions, is an easy consequence of Lemma 5 and Theorem 18.

Corollary 1: Let $\{x_n\}$ be a sequence of positive definite functions in $B(G)$ that convergences pointwise to a positive definite function $x \in B(G)$, that is

$$\lim_{n\to\infty} x_n(t) = x(t) \quad \forall t \in G$$

Then

$$w^*\text{-}\lim_{n\to\infty} x_n = x$$
$$\lim_{n\to\infty}\|x_n\|_B = \|x\|_B$$

or, equivalently

$$w^*\text{-}\lim_{n\to\infty}\hat{x}_n = \hat{x}$$
$$\lim_{n\to\infty}\|\hat{x}_n\| = \|\hat{x}\|$$

Proof: Note that $\|x_n\|_B = x_n(0)$, because x_n is a positive definite function. Hence

$$\lim_{n\to\infty}\|x_n\|_B = \lim_{n\to\infty} x_n(0) = x(0) = \|x\|_B$$

Being a convergent sequence, $\{\|x_n\|_B\}$ must be bounded, i.e., there exists $M > 0$ such that $\|x_n\|_B \le M$ for all n. It follows from Lemma 5 that $x = w^*\text{-}\lim_{n\to\infty} x_n$. The remaining claims are an immediate consequence of Theorem 18.

References

[1] V. A. Kotelnikov, "On the transmission capacity of the "ether" and wire in electrocommunications," in *Modern Sampling Theory: Mathematics and Applications*. Transl.: from Russian, J. J. Benedetto and P. J. S. G. Ferreira, Eds. Boston, MA: Birkhäuser, 2001, ch. 2, originally published in *Izd. Red. Upr. Svyazzi RKKA*, Moscow, U.S.S.R., 1933.

[2] C. E. Shannon, “Communication in the presence of noise,” *Proc. IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.
[3] A. J. Jerri, “The Shannon sampling theorem—Its various extensions and applications: A tutorial review,” *Proc. IEEE*, vol. 65, no. 11, pp. 1565–1596, Nov. 1977.
[4] J. R. Higgins, “Five short stories about the cardinal series,” *Bull. Amer. Math. Soc. (N.S.)*, vol. 12, no. 1, pp. 45–89, Jan. 1985.
[5] P. L. Butzer and R. L. Stens, “Sampling theory for not necessarily band-limited functions: A historical overview,” *SIAM Rev.*, vol. 34, no. 1, pp. 40–53, Mar. 1992.
[6] M. Unser, “Sampling—50 years after Shannon,” *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000.
[7] J. M. Whittaker, “The “Fourier” theory of the cardinal function,” in *Proc. Edinburgh Math. Soc. (Series 2)*, 1927–1929, vol. 1, pp. 169–176.
[8] D. L. Jagerman and L. J. Fogel, “Some general aspects of the sampling theorem,” *IRE Trans. Inf. Theory*, vol. IT-2, no. 4, pp. 139–146, Dec. 1956.
[9] M. Zakai, “Band-limited functions and the sampling theorem,” *Inf. Contr.*, vol. 8, no. 2, pp. 143–158, Apr. 1965.
[10] L. L. Campbell, “Sampling theorem for the Fourier transform of a distribution with bounded support,” *SIAM J. Appl. Math.*, vol. 16, no. 3, pp. 626–636, May 1968.
[11] A. Papoulis, “Truncated sampling expansions,” *IEEE Trans. Autom. Contr.*, vol. 12, no. 5, pp. 604–605, Oct. 1967.
[12] A. I. Zayed, *Advances in Shannon’s Sampling Theory*. Boca Raton, FL: CRC, 1993.
[13] M. M. Dodson and M. G. Beaty, “Abstract harmonic analysis and the sampling theorem,” in *Sampling Theory in Fourier and Signal Analysis—Advanced Topics*, J. R. Higgins and R. L. Stens, Eds. Oxford, U.K.: Oxford University Press, 1999, ch. 10.
[14] W. Rudin, *Fourier Analysis on Groups*. New York: Wiley, 1990.
[15] I. Klavánek, “Sampling theorem in abstract harmonic analysis,” *Mat.-Fyz. Ĉasopis Sloven. Akad. Vied*, vol. 15, pp. 43–48, 1965.
[16] G. Casinovi, “L^1-norm convergence properties of correlogram spectral estimates,” *IEEE Trans. Signal Processing*, vol. 55, no. 9, pp. 4354–4365, Sep. 2007.
[17] E. Hewitt and K. Stromberg, *Real and Abstract Analysis*. New York: Springer-Verlag, 1965.
[18] W. F. Eberlein, “Abstract ergodic theorems and weak almost periodic functions,” *Trans. Amer. Math. Soc.*, vol. 67, no. 1, pp. 217–240, Sep. 1949.
[19] C. Corduneanu, *Almost Periodic Functions*. New York: Interscience Publishers, 1968.
[20] J. L. Kelley, *General Topology*. New York: Springer Verlag, 1975.
[21] N. Wiener, “Generalized harmonic analysis,” *Acta Mathematica*, vol. 55, pp. 117–258, 1930.
[22] W. F. Eberlein, “The point spectrum of weakly almost periodic functions,” *Michigan Math. J.*, vol. 3, no. 2, pp. 137–139, 1955–56.
[23] E. Pfaffelhuber, “Generalized harmonic analysis for distributions,” *IEEE Trans. Inf. Theory*, vol. IT-21, no. 6, pp. 605–611, Nov. 1975.
[24] P. R. Halmos, *Measure Theory*. New York: Springer-Verlag, 1974.
[25] E. E. Granirer and M. Leinert, “On some topologies which coincide on the unit sphere of the Fourier-Stieltjes algebra $B(G)$ and of the measure algebra $M(G)$,” *Rocky Mountain J. Math.*, vol. 11, no. 3, pp. 459–472, 1981.
[26] P. Walters, *An Introduction to Ergodic Theory*. New York: Springer-Verlag, 1982.
[27] N. Dunford and J. T. Schwartz, *Linear operators—Part I: General theory*. New York, NY: John Wiley & Sons, 1988.
[28] J.-P. Pier, *Amenable Locally Compact Groups*. New York: Wiley, 1984.
[29] A. Schuster, “On the investigation of hidden periodicities with application to a supposed 26 day period of metereological phenomena,” *Terrestrial Magn.*, vol. 3, pp. 13–41, 1898.
[30] S. M. Kay and S. L. Marple Jr., “Spectrum analysis—A modern perspective,” *Proc. IEEE*, vol. 69, no. 11, pp. 1380–1419, Nov. 1981.
[31] V. F. Pisarenko, “The retrieval of harmonics from a covariance function,” *Geophys. J. Roy. Astro. Soc.*, vol. 33, pp. 347–366, 1973.
[32] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999.
[33] W. A. Gardner, *Statistical Spectral Analysis—A Nonprobabilistic Theory*. Englewood Cliffs, NJ: Prentice Hall, 1987.
[34] W. A. Gardner, “An introduction to cyclostationary signals,” in *Cyclostationarity in Communications and Signal Processing*, W. A. Gardner, Ed. New York: IEEE, 1994, pp. 1–90.
[35] H. Hurd and T. Koski, “The Wold isomorphism for cyclostationary sequences,” *Signal Processing*, vol. 84, pp. 813–824, 2004.
[36] P. Eymard, “L’algèbre de Fourier d’un groupe localement compact,” *Bull. Soc. Math. France*, vol. 92, pp. 181–236, 1964.
[37] W. Rudin, *Functional Analysis*, 2nd ed. New York: McGraw-Hill, 1991.

Giorgio Casinovi (M’89–SM’93) received the B.S. degrees in electrical engineering and in mathematics from the University of Rome, Italy, in 1980 and 1982 respectively. In 1984, he received the M.S. degree and in 1988, he received the Ph.D. degree in electrical engineering, both from the University of California, Berkeley.

He joined the School of Electrical and Computer Engineering of the Georgia Institute of Technology in 1989. His research interests include computer-aided design and simulation of electronic devices, circuits, and mixed-technology systems.