

Stationarizable Random Processes

WILLIAM A. GARDNER, MEMBER, IEEE

Abstract—The familiar notion of inducing stationarity into a cyclostationary process by random translation is extended through characterization of the class of all second-order continuous-parameter processes (with autocorrelation functions that possess a generalized Fourier transform) that are *stationarizable* in the wide sense by random translation. This class includes the nested set of proper subclasses: *almost cyclostationary* processes, *quasi-cyclostationary* processes, and *cyclostationary* processes. The random translations that induce stationarity are also characterized. The concept of stationarizability is extended to the concept of asymptotic stationarizability, and the class of *asymptotically stationarizable* processes is characterized. These characterizations are employed to derive characterizations of optimum linear and nonlinear time-invariant filters for nonstationary processes. Relative to optimum time-varying filters, these time-invariant filters offer advantages of implementational simplicity and computational efficiency, but at the expense of increased filtering error which in some applications is quite modest. The uses of a random translation for inducing stationarity-of-order- n , for increasing the degree of local stationarity, and for inducing stationarity into discrete-parameter processes are briefly described.

I. INTRODUCTION

A. Purpose

CYCLOSTATIONARY random processes are characterized by the invariance of their joint probability distributions under translations that are integer multiples of a fundamental translation T , which is referred to as the period of cyclostationarity. For example, a real continuous-parameter process X is *cyclostationary in the wide sense* (WSCS) with period T , if and only if the mean function m_X and autocorrelation function k_X possess the periodicity¹

$$m_X(t) \triangleq E\{X(t)\} = m_X(t + T)$$

$$k_X(t,s) \triangleq E\{X(t)X(s)\} = k_X(t + T, s + T). \quad (1)$$

WSCS processes have been the subject of an increasing number of research papers since the late 1950's, and they have been shown to be appropriate models for a wide variety of physical phenomena (see [1] and references therein). Because of their similarity to stationary processes, cyclostationary processes are more amenable to analysis than nonstationary processes in general [1]. The most widely exploited property of cyclostationary processes is

their *stationarizability*. That is, the *randomly translated* process

$$\tilde{X}(t) \triangleq X(t + \theta), \quad (2)$$

where θ is a random variable that is statistically independent of $X(t)$ and is uniformly distributed on $[-T/2, T/2]$, is *stationary in the wide sense* (WSS), i.e.,

$$E\{\tilde{X}(t)\} = m_{\tilde{X}}(0)$$

$$E\{\tilde{X}(t)\tilde{X}(s)\} = k_{\tilde{X}}(t - s, 0), \quad (3)$$

if and only if X is WSCS with period T [2]. If X is WSS, it is WSCS with any period T ; and if X is WSCS with period T/n with $n = 2$ or 3 or 4 or \dots , it is also WSCS with period T . In the engineering literature, randomly translated cyclostationary processes are commonly referred to as *phase-randomized* processes because they are generalizations of randomly phased sinusoids (cf. [1] and various references therein).

In this paper, it is shown that cyclostationary processes comprise only a subclass of the class of all processes that are stationarizable by random translation. For example, it is shown that most processes of practical interest in the class of *almost cyclostationary* processes, which includes the class of cyclostationary processes as a proper subclass, are stationarizable by random translation. More generally, the purpose of this paper is to characterize the autocorrelation functions of nonstationary processes that are *stationarizable* and of nonstationary processes that are *asymptotically stationarizable* (in the wide sense) by random translation, and to characterize the random translations that induce stationarity.

B. Motivation

There are various practical motives for the development of the concept of stationarizability. Two such motives that arise in estimation and detection problems are briefly described.

1) *Simplicity of Implementation of Filters and Detectors*: The lack of stationarity of a random process is reflected in the lack of time-invariance of estimators and detectors for that process. For example, if observations start at $t = -\infty$, the linear minimum-mean-squared-error (MMSE) filter, smoother, or predictor (to be referred to collectively as filter) for a WSCS signal in additive WSS noise is a periodically time-varying linear system [1]. Similarly, the minimum-probability-of-error detector for a Gaussian signal in additive white Gaussian noise decomposes into a linear MMSE filter and a correlator, and, if the signal is WSCS and the observation interval is $[0, \tau]$, the filter is asymptotically ($\tau \rightarrow \infty$) periodically time-varying. It has been shown that WSCS processes can be

Manuscript received July 25, 1975; revised April 13, 1977. This work was presented at the IEEE International Symposium on Information Theory, Ronneby, Sweden, June 21–24, 1976.

The author is with the Department of Electrical Engineering, University of California, Davis, CA 95616.

¹ In this paper, all equations involving functions evaluated at the time values t, s, τ and the frequency values f, ν are valid for all values of t, s, τ, f, ν in the set of reals $R = (-\infty, \infty)$ unless otherwise stated. The symbol $(+)$ denotes the positive half of the reals, $[0, \infty)$ and the symbol I denotes the integers. The symbol E denotes the statistical expectation operator. The superscript $*$ denotes complex conjugation.

decomposed into a countable set of jointly WSS representor processes (either continuous-time or discrete-time representors); furthermore many WSCS processes require only a finite number, let us say M , of jointly WSS processes in the decomposition [1]. More specifically, this is true if the process admits an exact or approximate mean-square equivalent orthogonal expansion over one period $[t, t + T]$ in terms of M basis functions (e.g., any WSCS process that is approximately band-limited to $[-M/2T, M/2T]$). It follows that the periodically time-varying MMSE filter for such a process can be decomposed into an M -input, M -output multiport time-invariant filter, preceded and followed by a bank of M periodic product modulators or periodic switches that are synchronized to the WSCS process [1]. On the other hand, if the stationarized model is used for the WSCS process, then only one (instead of M^2) time-invariant filter is needed, and—more important—no synchronization is required. Thus, for detection problems, use of the time-invariant filter in place of the time-varying filter in the estimator-correlator structure amounts to use of a noncoherent detector in place of a coherent detector. It is shown in Section V that the time-invariant MMSE filter for the stationarized process is identical to the minimum-time-averaged (over one period) mean-squared-error filter for the WSCS (not stationarized) process. Of course, the cost of the implementational simplicity is a degradation of MSE performance for filtering, and a degradation of probability-of-error performance for detection. This degradation can be, but is not necessarily, substantial. For example, if the WSCS process is the frequency division multiplex of M WSS signals, then the MSE degradation is always less than a factor of two. However, if the WSCS process is the time-division multiplex of M WSS signals (e.g., from an array of sensors in an inhomogeneous medium excited by a stationary random process), then the degradation can be as large as a factor of M , but in many practical applications will be less than a factor of two [1]. Similarly, degradation of detection error can be modest.

For a finite observation interval, the MMSE filter for estimation or detection of a WSCS process is not periodically time-varying. Similarly, the MMSE filter for the stationarized model is not time invariant. However, Kailath *et al.* [3] have recently shown that the MMSE-filter-correlator detector for stationary processes on a finite interval can indeed be implemented using only (two) *time-invariant* filters (although the MMSE filter—if implemented directly—is time-varying). Thus, even for finite observation intervals, use of the stationarized model leads to a simple time-invariant detector.

For stationarizable processes that are not WSCS, the simplicity of implementation of the time-invariant filter for the stationarized model can be even more attractive, since decomposition of the time-varying filter into a multiport time-invariant filter would not, in general, be possible, and the generalized synchronization problem would, in general, be more difficult. The almost cyclostationary processes discussed in Section II provide examples. It

should be mentioned that, in certain applications such as broad-band filtering and equalization of channels for multiplexed signals, the use of synchronized time-varying filters is not appropriate, and the aforementioned performance degradation is of no practical significance. It should also be mentioned that the complexity-performance trade-off described in the preceding paragraph applies to nonlinear as well as linear filtering (see Section (V-D)) and to detection of non-Gaussian as well as Gaussian signals. These trade-offs are currently under investigation.

2) *Efficiency of Computation of Filters:* Kailath *et al.* [4] have recently shown that the efficiency of computation of continuous-time linear MMSE filters that is attainable using a discretization scheme on the *extended* (two-process) *Sobolev and generalized* (continuous-time) *Levinson-Krein* equations is determined by a measure of the distance of the correlation functions of the observed and desired processes from stationary correlation functions. This measure of distance is called the *displacement rank*, and is denoted by α . For the sake of simplicity, a discrete-time process observed for N time instants is considered here. For a correlation matrix of arbitrary form, computation of the MMSE filter requires $O(N^3)$ multiplications; whereas, if the correlation matrix has displacement rank α , only $O(\alpha N^2)$ multiplications are required ($1 \leq \alpha \leq N$). This savings can be substantial when N is large. If the observed process is M -dimensional rather than scalar (one-dimensional) and if the M component processes are jointly WSS, then $\alpha = 2M$ and the number of multiplications needed is $O(M^3 N^2)$ (compared with $O(M^3 N^3)$ for arbitrary nonstationary correlation) [5]. Now, as discussed in paragraph 1, a scalar WSCS process can be decomposed into an M -dimensional WSS process; this is obvious for a discrete-time process with M time-points per period T . Thus $\alpha = 2M$ for a WSCS process. But if the stationarized model for this process is used, then $\alpha = 2$ and the number of time instants observed is MN ; thus the number of multiplications needed is $O(M^2 N^2)$ (compared with $O(M^3 N^2)$). This savings can be substantial when N is large, and the number M of jointly WSS processes required in the decomposition of the WSCS process is large. Of course, the cost of the computational efficiency is degradation of MSE performance which can be, but is not necessarily, substantial. For example, linear MMSE estimation for two-dimensional discrete-parameter bistationary image processes can be equivalently reformulated in terms of one-dimensional discrete-parameter cyclostationary processes by concatenating subsequent horizontal line scans of the image. (It should be noted, however, that many one-dimensional cyclostationary processes are not equivalent to any two-dimensional bistationary process.) Furthermore, it has been shown that the performance degradation resulting from the use of the MMSE linear time-invariant filter (instead of the periodically time-varying filter) for a continuous-time cyclostationary image process is negligible for typical video signals [1].

For stationarizable processes that are not WSCS, the

computational efficiency for the stationarized model can be even more attractive, since $O(N^3)$ multiplications would, in general, be needed for the nonstationary process, but only $O(N^2)$ would be needed for the stationarized model. The almost cyclostationary processes discussed in Section II provide examples.

C. Outline

Because of the importance of the class of almost cyclostationary processes as practical examples of stationarizable processes, Section II is devoted to the definition, characterization, and exemplification of this class. The published work on almost cyclostationary processes is sparse. The only publications we have found in which almost cyclostationary processes are a central issue are Gladyshev [6] on the Fourier series representation (7) of an almost periodic autocorrelation function, Jacobs [7, and references therein] on almost periodic Markov processes, and Ohta and Koizumi [8] on an almost periodic model for white Gaussian noise.

The characterization and exemplification of the class of stationarizable processes and stationarizing distributions is the topic of Section III, and the characterization and exemplification of the class of asymptotically stationarizable processes is the topic of Section IV.

In Section V, it is shown that time-invariant MMSE linear and nonlinear filters for stationarized or asymptotically stationarized versions of nonstationary processes can be characterized as minimum-time-averaged-MSE filters for the nonstationary (not stationarized or asymptotically stationarized) processes.

Finally, in Section VI, some extensions and generalizations are briefly discussed; viz., *stationarizability of order n* , *degree of local stationarity*, processes that are *almost stationarizable*, and *stationarizable discrete-parameter processes*. Conclusions are drawn in Section VII.

D. Terminology and Abbreviations

When the point of focus is the nonstationary behavior of an autocorrelation function, as it is in this paper, it is more convenient to work with the function $l_X(\cdot, \cdot)$ defined by

$$l_X(t, \tau) \triangleq k_X(t + \tau/2, t - \tau/2) \quad (4)$$

rather than to work directly with the autocorrelation function $k_X(\cdot, \cdot)$. A change of variables in (4) yields

$$k_X(t, s) = l_X([t + s]/2, t - s). \quad (5)$$

Thus the autocorrelation k_X is stationary if and only if l_X is independent of its first argument (t in (4)), and k_X is cyclostationary if and only if l_X is periodic in its first argument. It is appropriate to refer to the first argument in l_X as the *location variable*, and to refer to the second argument in l_X as the *separation variable*. Thus l_X could be referred to as the *location-separation autocorrelation function* to distinguish l_X from k_X . However, in the remainder of this paper, l_X is used exclusively; so for con-

TABLE I
ABBREVIATIONS FOR CLASSES OF RANDOM PROCESSES*

WSS(A)	wide-sense stationary (asymptotically)
WSCS(A)	wide-sense cyclostationary (asymptotically)
PM(A)	periodic in mean-square (asymptotically)
WSQCS(A)	wide-sense quasi-cyclostationary (asymptotically)
WSACS(A) _o	wide-sense almost cyclostationary _o (asymptotically)
APM(A) _o	almost periodic _o in mean-square (asymptotically)
WSS(A) ₋	wide-sense stationarizable (asymptotically)

* The meaning of the subscript o is given in Definitions (II-2) and (II-5).

venience l_X is referred to simply as the *autocorrelation function*.

The abbreviations used for the classes of processes of interest in this paper are listed in Table I. In these abbreviations, the prefix "WS" stands for *wide sense*, the suffix "A" stands for *asymptotically*, the subscript \sim indicates that the suffix "izable" is to replace "y" in the word stationary, and the subscript "o" is explained in Definitions (II-2) and (II-5).

As a final remark, we mention that the words *stationarizable* and *stationarizability* are somewhat awkward; however, they are grammatically valid and easier to use in this paper than equivalent phrases such as "that can be made stationary."

II. ALMOST CYCLOSTATIONARY PROCESSES

Almost cyclostationary processes are defined in terms of almost periodic functions with values in a Banach space. Since almost periodic functions are uncommon in the literature on random processes, they are defined and briefly described in the Appendix.

Definition II-1: A zero-mean second-order real continuous-parameter random process X is *almost cyclostationary in the wide sense* (WSACS) if and only if, for every $\epsilon > 0$, there exists a natural number N_ϵ and a trigonometric polynomial of order N_ϵ in $L^\infty(R)$ that uniformly approximates the autocorrelation function for X to within $\pm\epsilon$, i.e.,

$$\sup_{\tau \in R} \left| l_X(t, \tau) - \sum_{n=-N_\epsilon}^{N_\epsilon} c_n(\tau) \exp(j\omega_n t) \right| < \epsilon, \quad \text{for all } t \in R. \quad (6)$$

Thus a WSACS process has an autocorrelation $l_X(t, \tau)$ that is almost periodic in the location variable t with values in $L^\infty(R)$. The n th Fourier coefficient c_n in the Fourier series

$$\sum_{n=-\infty}^{\infty} c_n(\tau) \exp(j\omega_n t) \quad (7)$$

associated with $l_X(t, \tau)$ is given by the limit (in sup-norm)

$$c_n(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T, T} l_X(t, \tau) \exp(-j\omega_n t) dt. \quad (8)$$

The Fourier coefficients $\{c_n\}$ in the Fourier series associated with a WSACS process can be characterized as the time-averaged cross-correlations between the frequency-translated processes Y_n and Y_{-n} ; i.e.,

$$c_n(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T, T} E\{Y_n(t + \tau/2) Y_{-n}^*(t - \tau/2)\} dt \quad (9)$$

$$Y_n(t) \triangleq X(t) \exp(-j\omega_n t/2). \quad (10)$$

If X is WSCS, then (8)–(10) reduce to (1-6)–(1-8) in [1].

The following subclasses of WSACS processes are of particular interest because all such processes are stationarizable (Section III).

Definition II-2: The subclass $WSACS_o$ of the class WSACS is composed of all WSACS processes for which zero is not a limited point of the Fourier exponents $\{\omega_n\}$. The subclass $WSACS_*$ of the class $WSACS_o$ is composed of all $WSACS_o$ processes for which the associated Fourier series converges in $L^\infty(R)$ norm uniformly in t .

The terminology used in the next definition is based on Bohl's and Esclangon's term "quasi-periodic" [9, p. 49].

Definition II-3: A process is *quasi-cyclostationary in the wide-sense* (WSQCS) with *fundamental frequencies* $\{\nu_i\}_q^*$ if and only if it is WSACS and has Fourier exponents $\{\omega_n\}$ each of which is an integer multiple of one of the fundamental frequencies $\{\nu_i\}_q^*$ which are finite in number and incommensurable.

The class of all WSQCS processes is a proper subclass of the class $WSACS_o$.

The terminology in the next definition is borrowed from Muckenhoupt's work on functions with values in an L^2 space that are *almost periodic in the mean* (square) [9, p. 58]. In this definition, the $\{c'_n\}$ are second-order random variables, Ω is the sample space for X , and $\gamma \in \Omega$ is the sample space variable.

Definition II-4: A zero-mean second-order real continuous-parameter random process X is *almost periodic in mean square* (APM) if and only if, for every $\epsilon > 0$, there exists a natural number N_ϵ and a trigonometric polynomial of order N_ϵ in $L^2(\Omega)$ that uniformly approximates the process X to within $\pm\epsilon$, i.e.,

$$E_\Omega \left\{ \left| X(t, \gamma) - \sum_{n=-N_\epsilon}^{N_\epsilon} c'_n(\gamma) \exp(j\omega'_n t) \right|^2 \right\} < \epsilon, \quad \text{for all } t \in R. \quad (11)$$

Thus, an APM process is the limit in mean square of processes with sample paths all of which are almost periodic. The n th Fourier coefficient c'_n in the Fourier series

$$\sum_{n=-\infty}^{\infty} c'_n(\gamma) \exp(j\omega'_n t) \quad (12)$$

associated with $X(t, \gamma)$ is given by the limit (in L^2 norm)

$$c'_n(\gamma) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T, T} X(t, \gamma) \exp(-j\omega'_n t) dt. \quad (13)$$

The following subclass of APM processes is of particular interest because all such processes are stationarizable (Section III).

Definition II-5: The subclass APM_o of the class APM is composed of all APM processes for which the Fourier exponents $\{\omega'_n\}$ possess no limit points in R .

Since an APM process with Fourier exponents $\{\omega'_n; n \in I\}$ is WSACS with Fourier exponents $\{\omega_n; n \in I\} = \{\omega'_n - \omega'_m; n \in I, m \in I\}$, it follows that zero is not a limit point of $\{\omega_n\}$ if and only if $\{\omega'_n\}$ possesses no limit points in R . This result is summarized by the following lemma which is an extension of the result that all mean square periodic processes are WSCS. This lemma and Theorem (III-1) establish the stationarizability of APM_o processes.

Lemma II-1: The class APM_o is a proper subclass of the class $WSACS_o$.

Filtered Poisson processes, and other processes with finite starting time, cannot be WSACS; however, as illustrated with example (II-6), such processes can be *asymptotically WSACS* [7].

Definition II-6: Processes that are *asymptotically WSACS* (WSACSA) and *asymptotically APM* (APMA) are defined by supposing the existence of a positive real number T_ϵ and replacing the condition "for all $t \in R$ " with the condition "for all $t \geq T_\epsilon$," and by replacing "sup over $\tau \in R$ " with "sup over $|\tau| \leq T_\epsilon$ " in Definitions (II-1) and (II-4). (See Definition (A-7) for an alternative way of defining WSACSA and APMA processes.)

In MMSE filtering problems, the concept of joint stationarizability arises (Section V). Processes that are *jointly WSACS* are jointly stationarizable. Definitions (II-1) and (II-2) for single random processes can be extended to multiplicities of random processes by incorporating cross-correlation functions as well as autocorrelation functions in the definitions. An example follows.

Definition II-7: Two zero-mean second-order random processes X and Y are jointly WSACS if and only if, for every $\epsilon > 0$, there exist a natural number N_ϵ and three trigonometric polynomials of order N_ϵ in $L^\infty(R)$ that uniformly approximate the autocorrelation functions l_X and l_Y , and the cross-correlation function l_{XY} to within $\pm\epsilon$ (as in (6)).

An alternative but equivalent definition is that X and Y are jointly WSACS if and only if $aX + bY$ is WSACS for all real numbers a, b . *Jointly APM* processes are similarly defined.

The following six examples illustrate the practical utility of WSACS models for physical phenomena.

Example II-1: Any linear combination of uncorrelated WSCS processes, at least two of which have incommensurable periods, is $WSACS_o$ (and WSQCS) but not WSCS. For example, the frequency-division-multiplexed signal

$$X(t) = \sum_{i=1}^q X_i(t) \cos(\omega_i t + \theta_i), \quad (14)$$

where q is any natural number greater than one, is $WSACS_*$ (but not WSCS in general) if the baseband sig-

nals $\{X_i\}$ are jointly WSS, the $\{\omega_i\}$ are incommensurable, and the $\{\theta_i\}$ are deterministic or random and independent of the $\{X_i\}$ and each other.

Example II-2: Any finite product of statistically independent WSCS processes, at least two of which have incommensurable periods, is WSACS_o but not WSCS. For example the amplitude modulated synchronous data signal

$$X(t) = \sum_{i=-\infty}^{\infty} s(t - iT, A_i) \cos(\omega_o t + \theta) \quad (15)$$

is WSACS_o (but not WSCS in general) if the periods T and $2\pi/\omega_o$ are incommensurable, $\{A_i\}$ is a stationary-of-order-two-random sequence, $s(\cdot, A_i)$ is a deterministic $L^2(R)$ function for all samples of A_i , and θ is deterministic or random and independent of $\{A_i\}$.

Example II-3: A deterministic memoryless periodic transformation of a cyclostationary process can be WSACS. For example, the phase-modulated synchronous data signal

$$X(t) = \cos \left[\omega_o t + \sum_{i=-\infty}^{\infty} s(t - iT, A_i) + \theta \right] \quad (16)$$

is WSACS_o (but not WSCS in general) if the periods T and $2\pi/\omega_o$ are incommensurable, $\{A_i\}$ is a stationary-of-order-two sequence, $s(\cdot, A_i)$ is a deterministic $L^2(R)$ function for all samples of A_i , and θ is deterministic or random and independent of $\{A_i\}$.

Example II-4: A deterministic, almost periodic, linear transformation of a WSS process is WSACS. For example, the output of a linear system with WSS input and with impulse-response function

$$h(t, s) = g([t + s]/2, t - s) \quad (17)$$

is WSACS if $g(t, \tau)$ is almost periodic in t on $L^1(R)$. Also, an almost periodic time-scale transforming linear system with WSS input Y and with impulse response function

$$h(t, s) = \delta[s - t - f(t)], \quad (18)$$

where f is almost periodic and δ is the Dirac delta function, has an output

$$X(t) = Y[t + f(t)] \quad (19)$$

that is WSACS. Such time-scale transformations can result from radiation or reflection of a signal from a body in almost periodic motion.

Example II-5: Any finite set of uncorrelated WSCS processes, at least two of which have incommensurable periods, are jointly WSACS_o but not jointly WSCS.

Example II-6: As an example of a class of processes that are asymptotically WSACS, we consider the *marked and filtered Poisson point process* (cf. [10, Section 4.1])

$$X(t) = \sum_{i=1}^{\infty} A_i \phi(t - T_i), \quad (20)$$

where $\phi(\cdot)$ is a deterministic $L^2(R)$ function, and $\{A_i\}$ is a second-order sequence of independent identically distributed random variables that are independent of the

random occurrence times $\{T_i\}$ of an inhomogeneous Poisson counting process that starts at $t = 0$, and has rate parameter λ . The autocorrelation function for X is known to be

$$\begin{aligned} l_X(t, \tau) &= E\{A^2\} \int_{(+R)} \phi(t + \tau/2 - \sigma) \phi(t - \tau/2 - \sigma) \lambda(\sigma) d\sigma \\ &+ (E\{A\})^2 \int_{(+R)} \phi(t + \tau/2 - \sigma) \lambda(\sigma) d\sigma \\ &\int_{(+R)} \phi(t - \tau/2 - \alpha) \lambda(\alpha) d\alpha, \end{aligned} \quad (21)$$

and the mean function is known to be

$$m_X(t) = E\{A\} \int_{(+R)} \phi(t - \sigma) \lambda(\sigma) d\sigma. \quad (22)$$

Thus, if λ is asymptotically periodic with period T , X is asymptotically WSCS with period T , and if λ is asymptotically almost periodic (see Definition (A-7)), X is asymptotically WSACS. Furthermore, if λ is a second-order stochastic process so that the point process is doubly stochastic, the autocorrelation and mean are known to be

$$\begin{aligned} l_X(t, \tau) &= E\{A^2\} \int_{(+R)} \phi(t + \tau/2 - \sigma) \\ &\cdot \phi(t - \tau/2 - \sigma) E\{\lambda(\sigma)\} d\sigma \\ &+ (E\{A\})^2 \int_{(+R)} \int_{(+R)} \phi(t + \tau/2 - \sigma) \\ &\cdot \phi(t - \tau/2 - \alpha) E\{\lambda(\sigma)\lambda(\alpha)\} d\sigma d\alpha \end{aligned} \quad (23)$$

and

$$m_X(t) = E\{A\} \int_{(+R)} \phi(t - \sigma) E\{\lambda(\sigma)\} d\sigma. \quad (24)$$

Thus, if λ is WSCSA with period T , X is WSCSA with period T , and if λ is WSACSA, X is WSACSA. These models are appropriate for various signal formats employed in optical communication systems (cf. [11]).

Definitions of WSACS and APM discrete-parameter processes are analogous to Definitions (II-1) and (II-4). Uniformly time-sampled continuous-parameter WSACS and APM processes provide examples. (See the last remark in the Appendix.)

III. STATIONARIZABLE PROCESSES

Let X be a second-order real continuous-parameter random process with autocorrelation function l_X (defined in Section I), and for which the following limits exist uniformly in τ for all $\tau \in R$:

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{[-T, T]} l_X^2(t, \tau) dt \quad (25)$$

$$\tilde{k}_X(\tau) \triangleq \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{[-T, T]} l_X(t, \tau) dt. \quad (26)$$

Definition III-1: The limit function \tilde{k}_X is defined to be the stationary component (on R) of l_X , and the difference function

$$n_X(t, \tau) \triangleq l_X(t, \tau) - \tilde{k}_X(\tau) \quad (27)$$

is defined to be the nonstationary component (on R) of l_X .

Existence of the limit (25) guarantees existence of the generalized Fourier transform of $l_X(\cdot, \tau)$, for all $\tau \in R$. However, for simplicity of presentation, the usual Fourier transform which is the derivative (when it exists) of the generalized transform is used in this paper. In order not to be overly restrictive, the usual transform, the conjugate of which is defined by

$$L_X(f, \tau) \triangleq \int_R l_X(t, \tau) \exp(j2\pi ft) dt, \quad (28)$$

is allowed to contain Dirac delta functions (corresponding to step discontinuities in the generalized transform). For example, if X is WSS,

$$L_X(f, \tau) = \tilde{k}_X(\tau) \delta(f). \quad (29)$$

The following characterization of L_X is basic to the problem of characterizing correlations that can be made stationary by random translation.

Lemma III-1: The Fourier transform $L_X^*(\cdot, \tau)$ of the autocorrelation function $l_X(\cdot, \tau)$ contains a Dirac delta function of area $\tilde{k}_X(\tau)$ at the origin $f = 0$, and the Fourier transform $N_X^*(\cdot, \tau)$ of the nonstationary component $n_X(\cdot, \tau)$ does not contain a Dirac delta function at $f = 0$.

In order to derive the relationship between the components of the autocorrelation of a process X and the components of the autocorrelation of its randomly translated version

$$\tilde{X}(t) \triangleq X(t + \theta), \quad (30)$$

where θ is statistically independent of X , the fact that l_X and $l_{\tilde{X}}$ are related by convolution with the reversed probability density function $p_\theta(-\cdot)$ for θ is used, i.e.,

$$l_{\tilde{X}}(t, \tau) = \int_R l_X(t + \sigma, \tau) p_\theta(\sigma) d\sigma. \quad (31)$$

Application of the convolution theorem for Fourier transforms to (31) yields the relationship

$$L_{\tilde{X}}(f, \tau) = L_X(f, \tau) P_\theta(f), \quad (32)$$

where P_θ is the conjugate characteristic function² for θ . The following characterization of the components of $L_{\tilde{X}}$ is basic to the problem of determining conditions for the stationarity of \tilde{X} .

Lemma III-2: For every random translation that is independent of X , the stationary component (on R) of $l_{\tilde{X}}$ is equal to the stationary component (on R) of l_X

$$\tilde{k}_{\tilde{X}}(\tau) = \tilde{k}_X(\tau), \quad (33)$$

² The notation $P_\theta(\cdot)$ emphasizes the interpretation of the conjugate characteristic function as the Fourier transform of $p_\theta(\cdot)$.

and the conjugate Fourier transforms of the nonstationary components (on R) are related by

$$N_{\tilde{X}}(f, \tau) = N_X(f, \tau) P_\theta(f). \quad (34)$$

Thus random translation affects the frequency content of the nonstationary component in a simple way, and has no effect on the stationary component. In contrast to this, linear time-invariant filtering or smoothing of a process affects the frequency content of both components, and the effect on the nonstationary component is not so simple. Specifically, if we denote the transfer function of the filter by $H(\cdot)$, and if X is the input to the filter and Y is the output, then

$$\tilde{K}_Y(\nu) = |H(\nu)|^2 \tilde{K}_X(\nu) \quad (35)$$

$$\bar{N}_Y(f, \nu) = H(\nu - f/2) H^*(\nu + f/2) \bar{N}_X(f, \nu), \quad (36)$$

where \tilde{K}_X is the Fourier transform of \tilde{k}_X , and \bar{N}_X is the double Fourier transform of n_X

$$\bar{N}_X(f, \nu) \triangleq \int_R \int_R n_X(t, \tau) \exp[j2\pi(ft - \nu\tau)] dt d\tau. \quad (37)$$

Relationship (35) is well known, since \tilde{K}_X is identical to Rice's time-averaged power spectral density for a nonstationary process.

The following lemma on characterization and existence of characteristic functions is well known, but is essential to the proof of the theorems on stationarizability and is therefore given here so that proofs can be brief.

Lemma III-3: Let P_θ be any conjugate characteristic function. For every arbitrarily small positive number ϵ , there exists a positive number b such that

$$|P_\theta(f) - 1| < \epsilon, \quad \text{for all } |f| < b. \quad (38)$$

Also, for every arbitrarily small positive number B , there exist characteristic functions for finite mean-square random variables θ such that

$$P_\theta(f) = 0, \quad \text{for all } |f| > B. \quad (39)$$

For completeness, the following definition is given for processes with mean functions that are nonzero in general, but the following characterization theorem is given for zero-mean processes only. This simplifies the presentation without appreciable loss of generality. The specific loss of generality is discussed at the end of this section.

Definition III-2: X is stationarizable in the wide sense (WSS_~) if and only if there exist

- i) a finite mean-square random variable θ that is statistically independent of X ,
- ii) an $L^\infty(R)$ function $k(\cdot)$ that is not identically zero, and
- iii) a real number m

such that the randomly translated process $\tilde{X}(t) \triangleq X(t + \theta)$ is WSS with autocorrelation function k , and mean m ,

i.e.,

$$l_{\tilde{X}}(t, \tau) = k(\tau) \quad (40)$$

$$m_{\tilde{X}}(t) = m. \quad (41)$$

Theorem III-1: A zero-mean process X is stationarizable in the wide sense if and only if the stationary component \tilde{k}_X of its autocorrelation function is not identically zero, and there exists a neighborhood $[-B, B]$ of the origin on which the Fourier transform $N_X^*(\cdot, \tau)$ of the nonstationary component of the autocorrelation has zero L^1 norm

$$\int_{[-B, B]} |N_X(f, \tau)| df = 0. \quad (42)$$

Furthermore, the stationarized autocorrelation k is unique and equal to the stationary component \tilde{k}_X .

Proof: Employment of Lemma (III-2) yields the following expression for the autocorrelation function for the randomly translated process \tilde{X} :

$$l_{\tilde{X}}(t, \tau) = \int_R N_X(f, \tau) P_\theta(f) \exp(-j2\pi ft) df + \tilde{k}_X(\tau). \quad (43)$$

Lemma (III-1) guarantees that $N_X(\cdot, \tau)$ does not contain a Dirac delta function at the origin $f = 0$. Therefore the integral in (43) is independent of t , and the entire right member of (43) is not identically zero (and therefore \tilde{X} is stationary) if and only if

$$\int_R |N_X(f, \tau) P_\theta(f)| df = 0 \quad (44)$$

and

$$\tilde{k}_X(\tau) \neq 0. \quad (45)$$

But Lemma (III-3) and (44) require that N_X satisfy (42) for some positive number B , and if N_X satisfies (42), Lemma (III-3) guarantees the existence of a P_θ (viz., (39)) that satisfies (44). Furthermore, it follows from (43) and the necessary and sufficient conditions (44) and (45) that $k = \tilde{k}_X$ is the unique stationarized autocorrelation.

The following characterization of characteristic functions of random translations that induce stationarity follows immediately from (44).

Corollary III-1: A random translation with characteristic function P_θ^* stationarizes a process with nonstationary component N_X , and with stationary component that is not identically zero, if and only if P_θ annihilates the product $N_{\tilde{X}}$ (34) in L^1 norm:

$$\int_R |N_X(f, \tau) P_\theta(f)| df = 0. \quad (46)$$

Since $P_\theta(f) = 0$ implies that $\text{Re}\{P_\theta(f)\} = 0$, and since $\text{Re}\{P_\theta(f)\}$ is a valid characteristic function (corresponding to the probability density function $[p_\theta(t) + p_\theta(-t)]/2$), then every WSS process can be stationarized with a zero-mean

random translation θ having an even probability density function.

Many processes that satisfy the band-limiting constraint (42) are nonphysical as illustrated in the example in the following paragraph. Some physical insight into the reason that (42) is necessary for stationarizability can be gained by interpreting $P_\theta(\cdot)$ as the transfer function of a linear time-invariant filter with input signal having Fourier transform $L_X(\cdot, \tau)$, where τ is a parameter. Then, in order to stationarize $l_X(\cdot, \tau)$, the filter must pass the zero-frequency component of $l_X(\cdot, \tau)$ and reject components at all other frequencies. It can do so if and only if there are no components at frequencies that are infinitesimally close to zero (except the component at zero).

Let X be the process generated by WSS white noise W by multiplication with the deterministic function $m(\cdot)$ and convolution with the deterministic function $h(\cdot)$, i.e.,

$$X(t) = \int_R h(t-s)m(s)W(s)ds, \quad (47)$$

where

$$[m(t)]^2 = 1 + a(t) \geq 0$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T a(t) dt = 0$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T a^2(t) dt < \infty. \quad (48)$$

Equations (35) and (36) yield

$$\tilde{K}_X(\nu) = |H(\nu)|^2 \quad (49)$$

$$\bar{N}_X(f, \nu) = A(f)H(\nu - f/2)H^*(\nu + f/2). \quad (50)$$

Thus N_X satisfies (42) if and only if A satisfies the band-limiting constraint

$$\int_{[-B, B]} |A(f)| df = 0 \quad (51)$$

for some $B > 0$.

Evidently, the only type of nonstationary process that is stationarizable without some form of strict band-limiting in the process model (e.g., (51)) is that type with autocorrelation function that has a Fourier transform $L_X^*(\cdot, \tau)$ that contains no continuous component; i.e., contains only Dirac delta functions; for example,

$$L_X(f, \tau) = \sum_{n=-\infty}^{\infty} c_n(\tau) \delta(f - f_n). \quad (52)$$

Comparison of (52) with (7) (let $\omega_n = 2\pi f_n$) shows that such processes are WSACS. Furthermore, all WSACS processes are stationarizable. However, WSACS processes that are not WSACS_o because the Fourier exponents $\{\omega_n\}$ have zero as a limit point are not stationarizable because (42) is violated.

Corollary (III-1) yields the following necessary and sufficient condition on the characteristic function of every random translation that induces stationarity into a

WSACS_o process:

$$P_\theta(f_n) = 0, \quad \text{for all } n \neq 0 \text{ such that } c_n(\cdot) \neq 0. \quad (53)$$

If the WSACS_o process is WSQCS with fundamental frequencies $\{v_j\}$ ($v_j = 2\pi g_j$), a P_θ satisfying (53) can be constructed as follows:

$$P_\theta(f) = \prod_{j=1}^q \sin(\pi f/g_j)/(\pi f/g_j). \quad (54)$$

In this case, the stationarizing random translation θ can be decomposed as follows:

$$\theta = \sum_{j=1}^q \theta_j, \quad (55)$$

where $\{\theta_j\}$ are independent and θ_j is uniformly distributed on $[-1/(2g_j), 1/(2g_j)]$ for all j . Furthermore, the annihilation of all frequency components in the autocorrelation with frequencies that are harmonics of g_j can be attributed to θ_j .

More generally, for every WSQCS process, there exist stationarizing random translations θ that admit the decomposition (55), where $\{\theta_j\}$ are independent and θ_j has any probability density function $p_j(\cdot)$ satisfying Beutler's trigonometric moment condition which can be expressed (by using the Poisson sum formula together with [2, (33) and (34)]) as

$$\frac{1}{g_j} \sum_{m=-\infty}^{\infty} p_j(t - m/g_j) = 1. \quad (56)$$

This result is an extension and generalization of Hurd's result [2] from WSACS processes to WSQCS processes.

If a WSACS_o process is not WSQCS, but the Fourier exponents $\{f_n\}$ are square summable, the mean-square limit ($q \rightarrow \infty$) of (55) (with p_j satisfying (56) with $\{g_j\} = \{f_j\}$) exists, and the limiting random translation θ stationarizes the process. On the other hand, if $\{f_n\}$ are not square-summable, the series (55) does not converge ($q \rightarrow \infty$) to a finite variance random variable. But if B is the infimum of $\{|f_n|; n \neq 0\}$, there exists a stationarizing θ with variance not exceeding the bound

$$\overline{\sigma_\theta^2} = 3/(\pi B)^2. \quad (57)$$

This follows by construction; i.e., let P_θ be the triple self-convolution of the rectangle function

$$P(f) \triangleq \begin{cases} (12/B^3)^{1/4}, & \text{for all } |f| \leq B/4 \\ 0, & \text{for all } |f| > B/4. \end{cases} \quad (58)$$

It should be mentioned that some, but not all, processes that are stationarizable with a random translation having a continuous distribution are stationarizable with a random translation having a discrete distribution. Sufficient conditions for a WSACS_o process to be stationarizable with a random translation having a discrete distribution with uniformly spaced support points is that the Fourier exponents be bounded ($|f_n| \leq F < \infty$ for all n). Then, for example, a discrete probability density of the form

$$p_\theta(t) = \alpha \left[\frac{\sin(\pi\beta t)}{\pi\beta t} \right]^4 \sum_{i=-\infty}^{\infty} \delta(t - i\Delta) \quad (59)$$

with appropriate values for α, β, Δ , will stationarize the process.

Definition (III-2) and Theorem (III-1) for single random processes can be extended to multiplicities of random processes by incorporating cross-correlation functions as well as autocorrelation functions in the definition and theorem. This extension is necessary for applications to linear MMSE filtering (Section V).

Definition III-3: Two zero mean processes X and Y are *jointly stationarizable* (in the wide sense) if and only if there exist

- i) a finite mean square random variable θ that is statistically independent of X and Y , and
- ii) three $L^\infty(R)$ functions $k_X(\cdot), k_Y(\cdot), k_{XY}(\cdot)$,

such that the randomly translated processes $\tilde{X}(t) \triangleq X(t + \theta)$ and $\tilde{Y}(t) \triangleq Y(t + \theta)$ are jointly WSS with autocorrelation functions and crosscorrelation function given by

$$\begin{aligned} l_{\tilde{X}}(t, \tau) &= k_X(\tau) \\ l_{\tilde{Y}}(t, \tau) &= k_Y(\tau) \\ l_{\tilde{X}\tilde{Y}}(t, \tau) &= k_{XY}(\tau). \end{aligned} \quad (60)$$

Similarly, the extension of Theorem (III-1) is that X and Y are jointly WSS_~ if and only if (42) is satisfied by N_X, N_Y , and N_{XY} . For example, all jointly WSACS_o processes are jointly WSS_~. It should be noted, however, that if θ_X stationarizes X and θ_Y stationarizes Y , it does not follow that any of the random translations θ_X, θ_Y , or $\theta_X + \theta_Y$ jointly stationarize X and Y . This is a result of the fact that the support of $N_{XY}(\cdot, \tau)$ need not be the same as the support of $N_X(\cdot, \tau), N_Y(\cdot, \tau)$, or $n_X(\cdot, \tau)n_Y(\cdot, \tau)$.

Consider, as a simple example, the two zero-mean jointly WSACS_~ processes

$$\begin{aligned} X(t) &\triangleq Z(t) \cos(2\pi f_1 t) \\ Y(t) &\triangleq Z(t) \cos(2\pi f_2 t), \end{aligned} \quad (61)$$

where Z is WSS. These processes are jointly stationarized by θ_{XY} if and only if

$$P_{\theta_{XY}}[\pm(f_1 \pm f_2)] = P_{\theta_{XY}}(\pm 2f_1) = P_{\theta_{XY}}(\pm 2f_2) = 0; \quad (62)$$

whereas X and Y are individually stationarized by θ_X and θ_Y , respectively, if and only if

$$\begin{aligned} P_{\theta_X}(\pm 2f_1) &= 0 \\ P_{\theta_Y}(\pm 2f_2) &= 0. \end{aligned} \quad (63)$$

It should be mentioned for reference in Section V that, as an extension of Lemma (III-2), the stationary component (on R) \tilde{k}_{XY} of the cross-correlation of X and Y ,

$$\tilde{k}_{XY}(\tau) \triangleq \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T, T} l_{XY}(t, \tau) dt, \quad (64)$$

is identical to the stationary component (on R) of the cross-correlation of every pair of jointly translated versions $\tilde{X}(t) \triangleq X(t + \theta), \tilde{Y}(t) \triangleq Y(t + \theta)$.

As a final remark, it is pointed out that the problem of characterizing the behavior of the mean function $m_{\tilde{X}}$ of a

randomly translated process \tilde{X} is a special case of the same problem for the correlation function $l_{\tilde{X}}$. The preceding presentation has therefore been simplified by treating only the latter characterization problem. However, although the problems associated with stationarization of the mean by random translation are a special case of the problems associated with stationarization of the autocorrelation, it is not necessarily true that a random translation θ that stationarizes the autocorrelation also stationarizes the mean (or vice versa). But it is true that if two statistically independent random translations, θ_m and θ_k , that stationarize the mean and autocorrelation (respectively) exist, then a single random translation that stationarizes both, namely the sum of θ_m and θ_k , exists. The sum, however, is not the only such translation. For example, the mean of the WSCS process (61) (assuming Z has nonzero mean) with $\omega_1 = 2\pi/T$ is periodic with minimum period T , and its autocorrelation is periodic with minimum period $T/2$. Hence, a random translation that is uniformly distributed on $[-T/4, T/4]$ stationarizes the autocorrelation but not the mean; a random translation that is uniform on $[-T/2, T/2]$ stationarizes both.

IV. ASYMPTOTICALLY STATIONARIZABLE PROCESSES

It is intuitively obvious that no process with a finite starting time t_0 can be stationarized for all time t by random translation. This is easily illustrated with the process $(+)X$ generated from any zero-mean WSS process X by truncation at $t_0 = 0$. The autocorrelation for $(+)X$ is (from (70))

$$l_{(+)X}(t, \tau) = \tilde{k}_X(\tau)u(t - |\tau|/2). \quad (65)$$

Thus the stationary and nonstationary components (on R) are

$$k_{(+)X}(\tau) = 1/2 \tilde{k}_X(\tau) \quad (66)$$

$$N_{(+)X}(f, \tau) = \tilde{k}_X(\tau) \exp(j\pi f|\tau|)/(-j2\pi f). \quad (67)$$

Hence, $N_{(+)X}(\cdot, \tau)$ is unbounded in every neighborhood of the origin $f = 0$, and the necessary and sufficient condition for stationarizability (42) is not satisfied. On the other hand, $(+)X$ is WSS on $(+)R$ by construction. Similarly, many processes with finite starting times are WSS except for an initial nonstationary transient. Stationarity of this type is an important property and is commonly referred to as *asymptotic stationarity*. Paralleling the class of stationarizable processes characterized in Section III, there exists a class of processes that are *asymptotically stationarizable* by random translation.

As a simple example, any process that is WSCS, but not WSS, for all $t > 0$ is not asymptotically stationary, but can be made so by introducing a uniformly distributed random translation. In this section, the property of wide sense *asymptotic stationarizability* is defined, and the class of processes that are asymptotically stationarizable in the wide sense is characterized.

Let X be a second-order real continuous-parameter

random process with autocorrelation function l_X (defined in Section I), and for which the limit (25) and the following limits exist uniformly in τ for all $\tau \in R$:

$$\begin{aligned} \tilde{k}_X^+(\tau) &\triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_{[|\tau|/2, T+|\tau|/2]} l_X(t, \tau) dt \\ \tilde{k}_X^-(\tau) &\triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_{[-|\tau|/2-T, -|\tau|/2]} l_X(t, \tau) dt. \end{aligned} \quad (68)$$

The limit functions (26) and (68) are related by

$$\tilde{k}_X(\tau) = 1/2[\tilde{k}_X^+(\tau) + \tilde{k}_X^-(\tau)]. \quad (69)$$

Let $(+)X$ be the one-sided process obtained by truncation of X

$$(\cdot)X(t) \triangleq X(t)u(t)$$

$$u(t) \triangleq \begin{cases} 1, & t \geq 0 \\ 0, & t < 0; \end{cases}$$

then the autocorrelation for $(+)X$ is a truncated version of the autocorrelation for X

$$l_{(+)X}(t, \tau) = l_X(t, \tau)u(t - |\tau|/2) \triangleq l_X^+(t, \tau). \quad (70)$$

Parallel to Definition (III-1), \tilde{k}_X^+ is defined to be the *stationary component* (on $(+)R$) of l_X^+ , and the truncated difference function

$$n_X^+(t, \tau) \triangleq [l_X^+(t, \tau) - \tilde{k}_X^+(\tau)]u(t - |\tau|/2) \quad (71)$$

is defined to be the *nonstationary component* (on $(+)R$) of l_X^+ . The appropriateness of this definition follows from the fact that $n_X^+ \equiv 0$ if and only if X is WSS on $(+)R$. Since

$$\begin{aligned} \tilde{k}_{(+)X}^+ &\equiv \tilde{k}_X^+ \\ n_{(+)X}^+ &\equiv n_X^+ \\ l_{(+)X}^+ &\equiv l_X^+, \end{aligned} \quad (72)$$

the process X is assumed (without loss of generality) to be zero for all $t < 0$ in the remainder of this section.

The components \tilde{k}_X^+ and n_X^+ play a role in the theory of asymptotic stationarizability that is similar to the role played by \tilde{k}_X and n_X in the theory of stationarizability (Section III). For example, if L_X, N_X, \tilde{k}_X in Lemma (III-1) are replaced by $L_X^+, N_X^+, 1/2\tilde{k}_X^+$, respectively, then the lemma is still valid and reidentified as Lemma (IV-1). Furthermore, let \tilde{X} denote the randomly translated process $\tilde{X}(t) \triangleq X(t + \theta)$ ($X(t) = X(t)u(t)$, X and θ statistically independent); then parallel to the first part of Lemma (III-2), the stationary components of l_X^+ and $l_{\tilde{X}}^+$ are equal for every random translation θ

$$\tilde{k}_{\tilde{X}}^+ \equiv \tilde{k}_X^+. \quad (73)$$

In contrast to the second part of Lemma (III-2), the nonstationary components are related by

$$\begin{aligned} n_{\tilde{X}}^+(t, \tau) &= \int_R [Q_\theta(f, \tau)\tilde{k}_X^+(\tau) + N_X^+(f, \tau)P_\theta(f)] \\ &\quad \cdot \exp(-j2\pi ft) df u(t - |\tau|/2) \\ Q_\theta(f, \tau) &\triangleq (P_\theta(f) - 1) \exp(j\pi f|\tau|)/(-j2\pi f). \end{aligned} \quad (74)$$

However, it follows from (74) that

$$\lim_{t \rightarrow \infty} \left| n_X^\pm(t, \tau) - \int_R N_X^\pm(f, \tau) P_\theta(f) \exp(-j2\pi ft) df \right| = 0. \quad (75)$$

Thus, parallel to the second part of Lemma (III-2), the asymptotic properties of n_X^\pm depend on only $N_X^\pm(f, \tau) P_\theta(f)$. Equations (73)-(75) are referred to as Lemma (IV-2).

Definition IV-1: A zero-mean process X is *asymptotically WSS* (WSSA) on $(+)R$ if and only if the stationary component \tilde{k}_X^\pm is not identically zero and the nonstationary component n_X^\pm vanishes asymptotically,³ i.e.,

$$\lim_{t \rightarrow \infty} n_X^\pm(t, \tau) = 0$$

for every finite value of τ .

Definition IV-2: A process X is *asymptotically stationarizable* in the wide sense (WSSA_~) on $(+)R$ if and only if there exists a finite mean-square random translation variable θ that is independent of X and that renders the randomly translated version \tilde{X} of X asymptotically stationary in the wide sense on $(+)R$. The following theorem is obviously a parallel to Theorem III-1.

Theorem IV-1: A zero-mean process X is asymptotically stationarizable in the wide sense if and only if the stationary component \tilde{k}_X^\pm of its autocorrelation function is not identically zero, and there exists a neighborhood $[-B, B]$ of the origin on which the conjugate Fourier transform $N_X^\pm(\cdot, \tau)$ of the nonstationary component of the autocorrelation vanishes asymptotically

$$\lim_{t \rightarrow \infty} \int_{-B}^B N_X^\pm(f, \tau) \exp(-j2\pi ft) df = 0. \quad (76)$$

Furthermore, the asymptote of the asymptotically stationarized autocorrelation is unique and equal to the stationary component \tilde{k}_X^\pm .

The proof of Theorem IV-1 parallels the proof of Theorem III-1, with Lemmas IV-1 and IV-2 used in place of Lemmas III-1 and III-2. Similarly, it follows directly from the proof that, parallel to Corollary III-1, a characterization of the characteristic function P_θ^* for any asymptotically stationarizing random translation for X is

$$\lim_{t \rightarrow \infty} \int_R N_X^\pm(f, \tau) P_\theta(f) \exp(-j2\pi ft) df = 0. \quad (77)$$

Paralleling the role described in Section III of WSACS_o processes as practical examples of stationarizable processes, WSACS_A_o processes provide practical examples of asymptotically stationarizable processes. For example, if X is WSACS_A_o there exists a function $r(\cdot, \tau)$ with conjugate Fourier transform $R(\cdot, \tau)$ such that

$$N_X^\pm(f, \tau) = \sum_{n \neq 0} c_n(\tau) \{ \frac{1}{2} \delta(f - f_n) + \exp(j\pi(f - f_n)|\tau|) / [-j2\pi(f - f_n)] \} + R(f, \tau)$$

³ Some authors (cf. [12, pp. 90, 96]) refer to all processes for which the limit (68) exists and is not identically zero as *asymptotically stationary*, independent of the validity of (76).

and

$$\lim_{t \rightarrow \infty} r(t, \tau) = 0, \quad \text{for all } \tau \in R.$$

Thus, if P_θ satisfies (53), it follows from (75) that

$$\lim_{t \rightarrow \infty} n_X^\pm(t, \tau) = \lim_{t \rightarrow \infty} r(t, \tau) = 0, \quad \text{for all } \tau \in R$$

and \tilde{X} is WSSA.

Paralleling the extension in Section III from the property of stationarizability of a single process to the property of joint stationarizability of two processes, it is mentioned here for reference in Section V that the obvious extensions of Definitions (IV-1) and (IV-2), and Theorem (IV-1) for *joint asymptotic stationarizability* are valid, and all jointly WSACS_A_o processes are jointly WSSA_~.

V. TIME-INVARIANT FILTERING

Let X be a second-order process that is to be estimated, and let Y be an observed process from which the estimate of X is to be obtained. Let the optimum (defined in the following paragraph), linear, time-invariant estimate of X be denoted by \hat{X} :

$$\hat{X}(t) = \int_S h(\tau) Y(t - \tau) d\tau, \quad (78)$$

where the integral is assumed to be quadratic mean convergent. The deterministic function h is the impulse response function of the linear, time-invariant, estimating system. The *memory interval* S determines the memory of the system, and thereby determines whether the system is a filter ($S = [0, \infty)$), a smoother ($S = [-\Delta, \infty), \Delta > 0$), or a predictor ($S = [\Delta, \infty), \Delta > 0$). More generally, S can be an arbitrary finite collection of intervals, provided that it is independent of t . For convenience, the system is hereafter referred to as a *filter* regardless of S . Finite starting times are incorporated in the models of X and Y rather than in S ; e.g., $X(t) = Y(t) = 0$, for all $t < 0$.

The criterion of optimality to be considered is *minimum time-averaged mean-squared error* (MTAMSE), and the objective is to show that the MTAMSE filter for a pair of jointly WSS_~ nonstationary processes X and Y is identical to the MMSE (not time-averaged) filter for the jointly stationarized versions \tilde{X} and \tilde{Y} , and also to show that the MTAMSE filter for a pair of jointly asymptotically WSS_~ nonstationary processes is identical to the steady-state MMSE filter for the jointly asymptotically stationarized versions. As discussed in Section I, the motivation for using time-invariant filters for nonstationary processes is simplicity of implementation and efficiency of computation.

A. Stationarizable Processes

The objective is to characterize the solution to the MTAMSE filtering problem:

$$\min_h \{ \langle e \rangle \}, \quad (79)$$

where

$$\langle e \rangle \triangleq \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T, T} e(t) dt \quad (80)$$

$$e(t) \triangleq E\{[X(t) - \hat{X}(t)]^2\}, \quad (81)$$

in terms of the solution to the MMSE filtering problem:

$$\min_h \{\bar{e}\}, \quad (82)$$

where

$$\bar{e} \triangleq E\{[\tilde{X}(t) - \hat{\tilde{X}}(t)]^2\}, \quad (83)$$

and $\hat{\tilde{X}}$ is given by (78) with Y replaced by \tilde{Y} , and \tilde{X} and \tilde{Y} are any jointly stationarized versions of X and Y . However, before proceeding, we should mention that the MTAMSE time-invariant filter by definition yields a time-averaged MSE that is no larger than that resulting from the time-invariant filter obtained by time-averaging the MMSE time-varying filter. But the latter is considerably more difficult to solve for.

The following theorem is an extension and generalization of Theorem 1 in [1] from smoothing ($\Delta = \infty$) for WSCS processes to smoothing, filtering, and prediction for arbitrary WSS processes (e.g., WSACS processes).

Theorem V-1: The solution to the MTAMSE filtering problem (79) is identical to the solution to the MMSE filtering problem (82), and the unique (implicit) solution is

$$\int_S h(\tau) \tilde{k}_Y(t - \tau) d\tau = \tilde{k}_{XY}(t), \quad \text{for all } t \in S, \quad (84)$$

and

$$\min \langle e \rangle = \tilde{k}_X(0) - \int_S \tilde{K}_{XY}(\tau) h(\tau) d\tau. \quad (85)$$

Proof: Let the error process be denoted by ϵ

$$\epsilon(t) \triangleq X(t) - \int_S h(\tau) Y(t - \tau) d\tau, \quad (86)$$

and let its randomly translated version be denoted by $\tilde{\epsilon}$

$$\tilde{\epsilon}(t) \triangleq \epsilon(t + \theta). \quad (87)$$

Then

$$\tilde{\epsilon}(t) = \tilde{X}(t) - \int_S h(\tau) \tilde{Y}(t - \tau) d\tau \quad (88)$$

and the joint stationarity of \tilde{X} and \tilde{Y} guarantees the stationarity of $\tilde{\epsilon}$. Furthermore, it can easily be shown that

$$\langle e \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T, T} E\{\epsilon^2(t)\} dt \triangleq \tilde{k}_\epsilon(0), \quad (89)$$

and

$$\bar{e} = E\{\tilde{\epsilon}^2(t)\} \triangleq l_{\tilde{\epsilon}}(t, 0) = \tilde{k}_{\tilde{\epsilon}}(0). \quad (90)$$

Thus the uniqueness part of Theorem (III-1) (with X re-

placed by ϵ) yields

$$\langle e \rangle = \bar{e}. \quad (91)$$

Hence, $\langle e \rangle$ and \bar{e} clearly are minimized by the same filter. That the solution to (82) is implicitly given by (84) and (85) is well known.

B. Asymptotically Stationarizable Processes

Parallel to the objective in Subsection (V-A), the objective here is to characterize the solution to the MTAMSE filtering problem:

$$\min_h \{\langle e \rangle_+\}, \quad (92)$$

where

$$\langle e \rangle_+ \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_{[0, T]} e(t) dt \quad (93)$$

$$e(t) \triangleq E\{[X(t) - \hat{X}(t)]^2\}, \quad (94)$$

in terms of the solution to the MMSE steady-state filtering problem:

$$\min_h \{\bar{e}_\infty\}, \quad (95)$$

where

$$\bar{e}_\infty \triangleq \lim_{t \rightarrow \infty} \bar{e}(t) \quad (96)$$

$$\bar{e}(t) \triangleq E\{[\tilde{X}(t) - \hat{\tilde{X}}(t)]^2\}, \quad (97)$$

and $\hat{\tilde{X}}$ is given by (78) with Y replaced by \tilde{Y} , and \tilde{X} and \tilde{Y} are jointly asymptotically stationarized versions of X and Y . Parallel to the remark in Section (V-A), the MTAMSE time-invariant steady-state filter by definition yields a time-averaged MSE that is no larger than that resulting from the time-invariant, steady-state filter obtained by time-averaging the MMSE time-varying filter. But the latter is considerably more difficult to solve for.

Theorem V-2: The solution to the MTAMSE filtering problem (92) is identical to the solution to the MMSE steady-state filtering problem (95), and the unique (implicit) solution is

$$\int_S h(\tau) \tilde{k}_Y^+(t - \tau) d\tau = \tilde{k}_{XY}^+(t), \quad \text{for all } t \in S, \quad (98)$$

and

$$\min \langle e \rangle_+ = \tilde{k}_X^+(0) - \int_S \tilde{k}_{XY}^+(\tau) h(\tau) d\tau. \quad (99)$$

Proof: Parallel to the proof of Theorem (V-1), let ϵ and $\tilde{\epsilon}$ be defined by (86) and (87). Then the joint asymptotic stationarity of \tilde{X} and \tilde{Y} guarantees the asymptotic stationarity of $\tilde{\epsilon}$ (assuming $h(t) \rightarrow 0$ as $t \rightarrow \infty$). Furthermore, it can be shown that

$$\langle e \rangle_+ = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{[0, T]} E\{\epsilon^2(t)\} dt \triangleq \tilde{k}_\epsilon^+(0), \quad (100)$$

and

$$\tilde{e}_\infty = \lim_{t \rightarrow \infty} E\{\tilde{\epsilon}^2(t)\} \triangleq \lim_{t \rightarrow \infty} l_\epsilon^+(t, 0) = \tilde{k}_\epsilon^+(0). \quad (101)$$

Thus the uniqueness part of Theorem (IV-1) with X replaced by ϵ yields

$$\langle e \rangle_+ = \tilde{e}_\infty. \quad (102)$$

Hence, $\langle e \rangle_+$ and \tilde{e}_∞ clearly are minimized by the same filter. That the solution to (95) is implicitly given by (98) and (99) is well known.

C. Periodically Time-Varying Steady-State Filtering

The common notion of steady-state time-invariant filters for asymptotically stationary processes dealt with in Section (V-B) can be extended to the notion of *steady-state periodically time-varying filters* for asymptotically cyclostationary processes; e.g., Example (II-6). This extended notion has recently been employed in the derivation of linear MMSE steady-state receivers for fiber-optic data channels [11].

D. Nonlinear Filtering

Theorems (V-1) and (V-2) extend from linear filtering to nonlinear filtering. That is, if the hypothesis of (asymptotic) stationarizability in the wide sense is replaced with (asymptotic) stationarizability in the strict sense (Section VI), then the MTAMSE nonlinear time-invariant filter for X and Y is identical to the (steady-state) MMSE nonlinear filter for any (asymptotically) stationarized versions \tilde{X} and \tilde{Y} . This extension is valid for constrained nonlinear filtering (e.g., quadratic filtering) as well as unconstrained nonlinear filtering. The proofs of these extended theorems directly parallel the proofs of Theorems (V-1) and (V-2).

VI. EXTENSIONS AND GENERALIZATIONS

A. Stationarizability of Order n

The technique of making a process wide sense stationary by introducing a random translation can be extended to the technique of making a process stationary of order n as follows. Let the joint probability distribution function for the random variables $\{X(t + \tau_i); i = 1, 2, \dots, n\}$ evaluated at $\{x_i; i = 1, 2, \dots, n\}$ be denoted by

$$F_X(t; \{\tau_i\}_1^n; \{x_i\}_1^n) \quad (103)$$

for all $t \in R$, $\{\tau_i\}_1^n \in R^n$, $\{x_i\}_1^n \in R^n$. Parallel to Definition (III-1), the *stationary component* of F_X (also a valid distribution function when the limit exists and is not identically zero) is defined by

$$\begin{aligned} & \tilde{F}_X(\{\tau_i\}_1^n; \{x_i\}_1^n) \\ & \triangleq \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T, T} F_X(t; \{\tau_i\}_1^n; \{x_i\}_1^n) dt, \end{aligned} \quad (104)$$

and the *nonstationary component* (not a valid distribution function in general) is defined by

$$\begin{aligned} G_X(t; \{\tau_i\}_1^n; \{x_i\}_1^n) \\ \triangleq F_X(t; \{\tau_i\}_1^n; \{x_i\}_1^n) - \tilde{F}_X(\{\tau_i\}_1^n; \{x_i\}_1^n). \end{aligned} \quad (105)$$

Parallel to (31), the distribution function for the randomly translated version $\tilde{X}(t) = X(t + \theta)$ is given by

$$\begin{aligned} & F_{\tilde{X}}(t; \{\tau_i\}_1^n; \{x_i\}_1^n) \\ & = \int_R F_X(t + \sigma; \{\tau_i\}_1^n; \{x_i\}_1^n) p_\theta(\sigma) d\sigma. \end{aligned} \quad (106)$$

The obvious extensions of Lemmas (III-1) and (III-2) are valid. Therefore extensions of characterization Theorems (III-1) and (IV-1) can be developed. As an example, consider a random process X that is *almost cyclostationary of order n* (i.e., for which (103) is an almost periodic function of t) for all positive integers n , and is therefore *almost cyclostationary in the strict sense*. More specifically, consider a WSACS₀ Gaussian process X for which the positive Fourier exponents possess B as an infimum. Then the randomly translated process \tilde{X} is stationary in the strict sense if P_θ satisfies (39). The class of processes that are almost cyclostationary in the strict sense include as a proper subclass the class of strict-sense cyclostationary processes. Strict-sense cyclostationary discrete-parameter processes (called M -stationary if the period is M) play an important role in coding theory [13, p. 529], [16].

B. Other Related Notions

Also of interest is the notion of a process that can almost be made stationary in the sense that, for every arbitrarily small positive number ϵ , there exists a random translation that makes the norm of the nonstationary component of the autocorrelation less than ϵ . If the L^1 -sup norm

$$\|N_X\|^2 \triangleq \sup_{\tau \in R} \int_R |N_X(f, \tau)| df \quad (107)$$

is used, then the necessary and sufficient condition for a process to be *almost stationarizable* (AWSS₁) is that there exist a neighborhood of the origin $f = 0$ on which $N_X(\cdot, \tau)$ is bounded for all $\tau \in R$.

A nonstationary process for which the fluctuation of $l_X(\cdot, \tau)$ in the location variable t over an interval S is slow relative to the average (over $t \in S$) decay of $l_X(t, \cdot)$ is commonly referred to as *locally stationary* on S . However, the definition of local stationarity given in [14] appears to be overly restrictive and not in full agreement with this physical meaning (cf. [15]). For example, the property defined there is always destroyed by linear time-invariant filtering except in the special case when the process (or its

filtered version) is actually stationary.⁴ The problem is that the definition should not classify all processes as being either locally stationary or not. Rather, the definition should rank processes according to their *degree of local stationarity*, which can be taken to be the ratio of the bandwidths of stationary and nonstationary components. It then follows from Lemma (III-2), since random translation tends to reduce the bandwidth of the nonstationary component $n_X(\cdot, \tau)$ but has no effect on the stationary component $\tilde{k}_X(\cdot)$, that random translation tends to increase the degree of local stationarity.

C. Stationarizable Discrete-Parameter Processes

The definitions of the stationary and nonstationary components of the autocorrelation function of a continuous-parameter process (Definition (III-1)) extend in a natural way to discrete-parameter processes defined on the integers I by replacing (26) with⁵

$$\tilde{k}'_X(m) \triangleq \lim_{N \rightarrow \infty} \frac{1}{2N} \sum_{n=-N}^N l'_X(n, m), \quad \text{for all } m \in I. \quad (108)$$

Similarly, Lemmas (III-1)–(III-3) extend to discrete-parameter processes, by replacing the Fourier transform (28) with the discrete Fourier transform

$$L'_X(f, m) \triangleq \sum_{n \in I} l'_X(n, m) \exp(j2\pi n f), \quad \text{for all } f \in [-1/2, 1/2]. \quad (109)$$

Only random translations with discrete distributions are considered for inducing stationarity into discrete-parameter processes. Thus only characteristic functions of the form

$$P'_\theta(f) = \sum_{n \in I} p_n \exp(-j2\pi n f) \quad (110)$$

$$\sum_{n=-\infty}^{\infty} p_n = 1, \quad p_n \geq 0, \quad \text{for all } n \in I$$

are considered. It is easily shown that the characterization Theorems (III-1) and (IV-1) are valid for discrete-parameter processes as well as continuous-parameter processes. Furthermore, practical examples of stationarizable discrete-parameter processes are provided by discrete-parameter WSACS processes obtained by uniformly time-sampling continuous-parameter WSACS processes (see the last paragraph in Section II). It should be mentioned, however, that not every discrete-parameter process X' obtained by uniformly time-sampling a stationarizable continuous-parameter process X is stationarizable, since

⁴ The property of local stationarity defined in [14] is approximately preserved by time-invariant filtering if and only if the filter transfer function is sufficiently smooth; i.e., approximately constant over frequency intervals of the order of the bandwidth of the nonstationary factor in the covariance.

⁵ The primes in (108)–(111) distinguish the discrete-parameter quantities from their continuous-parameter counterparts. The primes on X and θ have been omitted for convenience of notation.

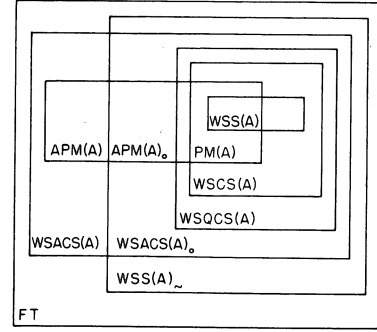


Fig. 1. Venn diagrams of classes of random processes that are stationarizable in the wide sense, and that are asymptotically stationarizable in the wide sense. Two separate diagrams are shown superimposed: one with an A (which stands for asymptotically) in parentheses included, and the other without. FT is an abbreviation for Fourier transformable autocorrelation, and denotes the class of process for which the limit (25) exists.

stationarizability of X' requires that there exist neighborhoods of every point in the set $\{n/\Delta; \text{for all } n \in I\}$ on which $N_X(\cdot, \tau)$ has zero L^1 -norm (where Δ is the sampling increment); whereas stationarizability of X requires only that there exist a neighborhood of the origin on which $N_X(\cdot, \tau)$ has zero L^1 -norm. This follows from the relationship

$$N'_X(f, m) = \frac{1}{\Delta} \sum_{n \in I} N_X(|f - n|/\Delta, m\Delta). \quad (111)$$

VII. CONCLUSIONS

It has been shown that a useful alternative to the usual dichotomy between stationary and nonstationary processes is to use a simple decomposition of autocorrelation functions (and probability distribution functions) into stationary and nonstationary components,⁶ and then to classify processes according to the behavior of the nonstationary component (see Fig. 1, and Table I). Since the introduction of a random translation has no effect on the stationary component, but has a filtering effect on the nonstationary component, it follows that a random translation can be used to alter the nonstationarity of a process in various ways, including (asymptotic) annihilation, in which case the random translation induces (asymptotic) stationarity.

As discussed in Section I, the technique of introducing a random translation to induce stationarity can be used to simplify the implementation of estimators and detectors (e.g., no time-varying components nor synchronization requirements), and to increase the efficiency of computation associated with these (e.g., inversion of Toeplitz and convolution operators). This technique will, in general, result in a degradation in performance of the estimator or detector, which in some applications is modest but in others is not. Thus a numerical measure of the *degree of stationarity* of a nonstationary process that is computa-

⁶ This decomposition does not necessarily correspond to a decomposition of the process into independent stationary and nonstationary components [18].

tionally convenient and that can be used to estimate or bound performance degradation is needed.

As a final remark, it is mentioned that the work of Jacobs [16] on almost periodic channels (which transform stationary signals into almost cyclostationary signals) is related to the work reported herein. Jacobs shows that the capacity of an almost periodic, memoryless finite-alphabet discrete-time channel is given by

$$\tilde{C} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N C_n, \quad (112)$$

where C_n is the capacity of a stationary channel with transition probability matrix (for all values of discrete time) equal to the transition probability matrix of the almost periodic channel at time n . Since $\{C_n\}$ is an almost periodic sequence, \tilde{C} can be referred to as the stationary component of $\{C_n\}$. Jacobs also discusses the capacity of *averaged channels* and *translation-randomized channels*. For more recent work on related topics, see [17].

ACKNOWLEDGMENT

The author expresses his gratitude to the Associate Editor, Dr. D. L. Snyder, and several reviewers for their helpful criticism of the original manuscript.

APPENDIX

Almost Periodic Functions with Values in a Banach Space (cf. [9], [19]).

Let β be a Banach space with the norm of a vector $b \in \beta$ denoted by $\|b\|$. (In Section II, two Banach spaces are considered: (1) that with vectors which are complex functions of a real variable, and with L^∞ norm, and (2) that with vectors which are complex random variables, and with L^2 norm.) Let $f: R \rightarrow \beta$ be a function that maps real numbers $t \in R$ into vectors $f(t) \in \beta$.

Definition (A-1): A function $p: R \rightarrow \beta$ that is of the form

$$p(t) = \sum_{n=-N}^N c_n \exp(j\omega_n t), \quad \text{for all } t \in R, \quad (A-1)$$

where $c_n = c_{-n}^* \in \beta$, $\omega_n = -\omega_{-n} \in R$, and N is a natural number, is defined to be a *real trigonometric polynomial* on β .

Definition (A-2) (Uniform Approximation): A function $f: R \rightarrow \beta$ is defined to be *almost periodic* (AP) on β if, for every $\epsilon > 0$, there exists a trigonometric polynomial on β , say $p_\epsilon(t)$, such that

$$\|f(t) - p_\epsilon(t)\| < \epsilon, \quad \text{for all } t \in R. \quad (A-2)$$

Equivalent Definition (A-3) (Bohr-Bochner): A function $f: R \rightarrow \beta$ is defined to be *almost periodic* on β if, for every $\epsilon > 0$, there exists a real number $\lambda_\epsilon > 0$ such that every interval of length λ_ϵ on the real line contains at least one point τ_ϵ (called an ϵ -translation number of f) such that

$$\|f(t + \tau_\epsilon) - f(t)\| < \epsilon, \quad \text{for all } t \in R. \quad (A-3)$$

Definition (A-4): Let $f: R \rightarrow \beta$ be AP, and let ω be any real number. The *mean value* $c_\omega \in \beta$ of the function $f(t) \exp(-j\omega t)$

is defined to be the limit in norm

$$c_\omega \triangleq \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t) \exp(-j\omega t) dt = c_{-\omega}^*. \quad (A-4)$$

Theorem (A-1): If $f: R \rightarrow \beta$ is AP, and if $\{c_\omega\}$ denotes the set of mean values of f , there are at most countably many values of ω for which $\|c_\omega\| \neq 0$.

Definition (A-5): Let $\{\omega_{\pm n}\}_0^\infty$ be the sequence (in no particular order except $\omega_n = -\omega_{-n}$) of values of ω for which $\|c_\omega\| \neq 0$ and let c_{ω_n} be abbreviated to c_n . The series

$$\sum_{n=-\infty}^{\infty} c_n \exp(j\omega_n t) \quad (A-5)$$

is defined to be the *Fourier series associated with f* , and $\{\omega_n\}$ and $\{c_n\}$ are termed the *Fourier exponents* and *Fourier coefficients*, respectively, of f .

Remark: There is, as would be expected, a connection between ϵ -translation numbers τ_ϵ (also called ϵ -periods) and Fourier exponents (cf. [19], p. 38). There exist AP functions with arbitrarily specified Fourier exponents; e.g., $\{\omega_n\}$ can have arbitrary limit points in R (cf. [9, p. 31]).

Theorem (A-2): If the Fourier series associated with f is uniformly (in t) convergent in norm, the sum of the series is the function f .

Remark: Some results on the difficult problem of establishing conditions for convergence of the Fourier series (A-5) are presented in [9, pp. 31–38].

Definition (A-6): A function $f: R \rightarrow \beta$ is defined to be *quasi-periodic with fundamental frequencies $\{\nu_i\}$* if it is AP and has Fourier exponents $\{\omega_n\}$ each of which is an integer multiple of one of $\{\nu_i\}$, and q is finite and $\{\nu_i\}$ are incommensurable.⁷

Definition (A-7): A function $f: R \rightarrow \beta$ is defined to be *asymptotically almost periodic* (APA) if there exist two functions $p: R \rightarrow \beta$ and $r: R \rightarrow \beta$, such that their pointwise sum equals f , p is AP, and r is asymptotically ($t \rightarrow \infty$) zero in norm.

Remark: If $f: R \rightarrow \beta$ is AP, and T is any real number, then $\{f(nT); n = 0, \pm 1, \pm 2, \dots\}$ is an AP sequence (cf. [9, p. 47]).

REFERENCES

- [1] W. A. Gardner and L. E. Franks, "Characterization of cyclostationary random signal processes," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 4–14, Jan. 1975.
- [2] H. L. Hurd, "Stationarizing properties of random shifts," *SIAM J. Appl. Math.*, vol. 26, pp. 203–212, Jan. 1975.
- [3] T. Kailath, B. Levy, L. Ljung, and M. Morf, "Time-invariant implementation of Gaussian signal detectors," Information Systems Laboratory, Stanford University, Stanford, CA, to be published.
- [4] T. Kailath, L. Ljung, and M. Morf, "A new approach to the determination of Fredholm resolvents of nondisplacement kernels," Information Systems Laboratory, Stanford University, Stanford, CA, to be published.
- [5] R. A. Wiggins and E. A. Robinson, "Recursive solution to the multichannel filtering problem," *J. Geophys. Res.*, vol. 70, no. 8, pp. 1885–1981, Apr. 1965.
- [6] E. G. Gladyshev, "Periodically and almost periodically correlated

⁷ The real numbers in a finite set $\{\nu_i\}$ are *incommensurable* if and only if the only set of q integers that satisfies

$$\sum_{i=1}^q \alpha_i \nu_i = 0$$

is $\alpha_1 = \alpha_2 = \dots = \alpha_q = 0$.

- random processes with continuous time parameter," *Theory Prob. Appl.*, (USSR), pp. 173-177, 1963.
- [7] K. Jacobs, "Fastperiodische Markoffsche Prozesse," *Mathematische Annalen*, vol. 134, pp. 408-427, 1958.
 - [8] M. Ohta and T. Koizumi, "Digital simulation of a white noise model formed of uniformly almost periodic functions," *Information and Control*, vol. 17, pp. 340-358, Nov. 1970.
 - [9] C. Corduneanu, *Almost Periodic Functions*. New York: Wiley, 1961.
 - [10] D. L. Snyder, *Random Point Processes*. New York: Wiley, 1975.
 - [11] W. A. Gardner, "An equivalent linear model for marked and filtered doubly stochastic Poisson processes with application to MMSE estimation for synchronous M -ary optical data signals," *IEEE Trans. Commun. Technol.*, vol. COM-24, pp. 917-921, Aug. 1976.
 - [12] E. Parzen, "Spectral analysis of asymptotically stationary time series," *Bull. Inst. Internat. Statist.*, vol. 39, livraison 2, pp. 87-103, 1962.
 - [13] R. M. Gray, D. L. Neuhoff, and J. K. Omura, "Process definitions of distortion-rate functions and source coding theorems," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 524-532, Sept. 1975.
 - [14] R. A. Silverman, "Locally stationary random processes," *IEEE Trans. Inform. Theory*, vol. IT-3, pp. 182-187, Sept. 1957.
 - [15] R. L. Snyder, "A partial spectrum approach to the analysis of quasi-stationary time series," *IEEE Trans. Inform. Theory*, vol. IT-3, pp. 579-587, Oct. 1967.
 - [16] K. Jacobs, "Almost periodic channels," *Colloquium on Combinatorial Methods in Probability Theory*, Matematisk Inst., Aarhus Univ., Danmark, Aug., 1962, pp. 118-126.
 - [17] W. L. Root, "Nonprobabilistic and partly probabilistic channel coding and time varying channels," *Communication Systems and Random Process Theory*, to be published by Noordhoff-Leyden, Netherlands (Nato Advanced Study Institute Series).
 - [18] J. Kampé De Fériet, "Correlation and spectrum of asymptotically stationary random functions," *The Mathematics Student*, vol. XXX, no. 1, pp. 55-67, Jan. 1962.
 - [19] L. Amerio and G. Prouse, *Almost-Periodic Functions and Functional Equations*. New York: Van Nostrand Reinhold, 1971.