Cycloergodic Properties of Discrete-Parameter Nonstationary Stochastic Processes

RUSSELL A. BOYLES AND WILLIAM A. GARDNER, MEMBER, IEEE

Abstract—It is shown that a large class of nonstationary discrete parameter stochastic processes possess novel ergodic properties, which are referred to as *cycloergodic* properties. Specifically, it is shown that periodic components of time-varying probabilistic parameters can be consistently estimated from time averages on one sample path. The cycloergodic theory developed herein extends and generalizes existing ergodic theory for asymptotically mean stationary and N-stationary (cyclostationary) processes, and is presented in both wide-sense and strict-sense contexts.

I. Introduction

THE CONCEPT of ergodicity in relation to a temporal random process has as its major practical application justification of the assumptions that a) the long run behavior of time-averaged measurements on a sample path of a random process can be predicted from calculated expectations based on the probabilistic model of the random process, and b) the large sample behavior of hypothetical ensemble-averaged measurements can be predicted from actual time-averaged measurements on one member of the ensemble (one sample path). In general, these assumptions simplify mathematical analysis and experimental design.

The great majority of discussions of ergodicity in the engineering and applied sciences literature treat the property of *stationarity* as a necessary prerequisite for ergodic properties. This seems intuitively appropriate since a time-averaged quantity is independent of time and therefore cannot approximate a probabilistic parameter that depends on time (i.e., that is nonstationary). However, time-averaged parameters for nonstationary processes are independent of time, which suggests that nonstationary processes can possess ergodic properties associated with these time-averaged parameters. This is indeed true for both asymptotic time averages of expectations for *asymptotically mean stationary* (AMS) processes ([1], [2]), and finite time averages of expectations for *cyclostationary* (CS) processes ([3], [4]).

Let Z denote the integers. Loosely speaking (for the purposes of this motivating discussion) a process $\{X_j: j \in Z\}$ possesses *AMS properties* if asymptotic time averages of its probabilistic parameters exist; e.g., if $E(X_j)$ exists for all j we define its asymptotic time average to be

$$\langle\{E(X_j)\}\rangle \triangleq \lim_{J\to\infty} \frac{1}{J}\sum_{j=1}^{J} E(X_j).$$

A CS process is one for which the probabilistic parameters are periodic (with period N, say) and therefore possess a time-averaged value; e.g., if $E(X_j)$ exists for all j we have

$$E(X_{j+N}) = E(X_j), \qquad j \in Z,$$

and

$$\langle\{E(X_j)\}\rangle = \lim_{J\to\infty} \frac{1}{J}\sum_{j=1}^{J} E(X_j) = \frac{1}{N}\sum_{j=1}^{N} E(X_j).$$

Thus, a CS process possesses AMS properties. Ergodic properties of such nonstationary processes guarantee the equality of asymptotic sample-path time averages such as

$$\langle\{X_j\}\rangle \triangleq \lim_{J\to\infty} \frac{1}{J}\sum_{j=1}^{J} X_j$$

with the corresponding time-averaged probabilistic parameters, e.g.,

$$\langle\{X_j\}\rangle = \langle\{E(X_j)\}\rangle.$$

Wide-sense ergodic properties guarantee that the above relation holds in the sense that

$$\lim_{J\to\infty} E\left(\left|\frac{1}{J}\sum_{j=1}^{J} X_j - \langle\{E(X_j)\}\rangle\right|^2\right) = 0.$$

Strict-sense ergodic properties, on the other hand, guarantee the almost sure equality

$$\langle\{Y_j\}\rangle = \langle\{E(Y_j)\}\rangle$$

for a large class of processes $\{Y_j\}$ obtained by applying a class of measurements (measurable functions) to the original $\{X_j\}$ process.

In addition to this type of extension of ergodicity from stationary processes to CS processes and other nonstationary processes with AMS properties, ergodicity can be

Manuscript received August 14, 1981; revised July 18, 1982.
R. A. Boyles is with the Lawrence Livermore National Laboratory, University of California, Livermore, CA 94550.
W. A. Gardner is with the Signal and Image Processing Laboratory, Department of Electrical and Computer Engineering, University of California, Davis, CA 95616.

0018-9448/83/0100-0105$01.00

extended to CS processes without restriction to time-averaged expectations ([2], [4]–[6]). This results from the fact that a CS process is equivalent to a vector-valued stationary process $\{X_j\}$ where

$$X_j = \left(X_{(j-1)N+1}, \cdots, X_{jN}\right)$$

(cf. [7]). Note that CS processes are sometimes called *N-stationary* [2].

The purpose of this paper is to introduce some novel ergodic properties of nonstationary processes. We shall show that CS processes are not the only nonstationary processes for which time-varying expectations can be consistently estimated from time averages on one sample path of the process. Specifically, we shall show that periodic components of time-varying (not necessarily periodic) probabilistic parameters can be consistently estimated for a large class of nonstationary processes. Furthermore, we shall show that, in certain cases, the entire time-varying parameter can be consistently estimated. We shall establish convergence of estimators in a wide-sense context (Section II) as well as a strict-sense context (Section III). The material in Section II is easily adapted from discrete-parameter processes to continuous-parameter processes. This is not the case for the material in Section III.

The general class of processes that is appropriate for the study of these novel ergodic properties is the class of processes possessing *asymptotically mean cyclostationary* (AMCS) properties. The novel ergodic properties are to be referred to as *cycloergodic* properties. Loosely speaking (for the purposes of this motivating discussion) a process $\{X_j\}$ with AMCS properties is a process for which asymptotic time averages of sinusoidally weighted probabilistic parameters exist; e.g., if $E(X_j)$ exists for all j we define

$$\langle\{E(X_j)e^{-i\alpha j}\}\rangle \triangleq \lim_{J\to\infty} \frac{1}{J}\sum_{j=1}^{J} E(X_j)e^{-i\alpha j},$$

where α is a real number and $i \triangleq \sqrt{-1}$. Cycloergodic properties of such nonstationary processes guarantee the equality of asymptotic sinusoidally weighted sample-path time averages such as

$$\langle\{X_j e^{-i\alpha j}\}\rangle \triangleq \lim_{J\to\infty} \frac{1}{J}\sum_{j=1}^{J} X_j e^{-i\alpha j}$$

with the corresponding time-averaged sinusoidally weighted probabilistic parameters, e.g.,

$$\langle\{X_j e^{-i\alpha j}\}\rangle = \langle\{E(X_j)e^{-i\alpha j}\}\rangle.$$

In Section II we present a relatively straightforward wide-sense theory of cycloergodicity that incorporates the results of Gudzenko [4], Kampé de Fériet and Frenkiel [3], Hurd [6], and Parzen [1] as special cases. That is, only time-averaged expectations ($\alpha = 0$) are estimated in [1] and [3], and only CS processes are considered in [4] and [6]. In Section III we present a somewhat less straightforward strict-sense theory. This theory extends the work of Nedoma [5] on N-stationary N-ergodic processes and supplements the work of Gray and Kieffer [2] (since the class of strict-sense AMS processes turns out to be identical to the class of strict-sense AMCS processes defined in Section III.) Moreover, the class of processes of primary interest in Section III turns out to coincide with the class of quasi-periodic measures studied by Blum and Hanson [10]. On the other hand, Jacobs ([11]–[13]) has developed a theory of almost periodic measures, but in a more restrictive setting than that of [10]. His results, therefore, do not carry over to the class of measures studied herein.

In both the strict-sense and the wide-sense theory, a subclass of processes of particular interest possessing AMCS properties is the class of *almost cyclostationary* (ACS) processes [8]. Probabilistic parameters of ACS processes are almost periodic functions of time [9].

There is considerable practical motivation for an investigation of the cycloergodic properties of ACS and other processes possessing AMCS properties, since these processes are appropriate models for a wide variety of phenomena involving cycles, i.e., phenomena giving rise to processes for which there is some underlying periodicity in the generating mechanism. In communications, radar, and sonar, the underlying periodicity arises from periodic sampling, scanning, modulating, multiplexing, and coding operations [7], [14, ch. 8]. It can also arise from interference caused by rotating reflectors such as helicopter blades, and air- and water-craft propellers. In mechanical vibration monitoring and diagnosis for rotating machinery, the periodicity arises from rotation, revolution, and reciprocation of gears, belts, chains, shafts, propellers, bearings, pistons, etc. In atmospheric science, the periodicity arises from seasons caused primarily by rotation and revolution of the earth. In radio astronomy, the periodicity arises from revolution of the moon, rotation of Jupiter and revolution of its satellite Io, rotation and pulsation of the sun, etc. A wide variety of examples of CS and ACS processes are given in [7], [8], [14], [15], (and references therein). For brevity, they are not repeated here.

II. Wide-Sense Theory

Let $X = \{X_j: j \in Z\}$ be a real discrete-parameter random process with uniformly bounded second moments. We are primarily concerned with processes possessing wide-sense AMCS properties; i.e., processes for which limits of the forms

$$c_X(j;\alpha) \triangleq \lim_{K\to\infty} \frac{1}{K}\sum_{k=k_0}^{k_0+K-1} E(X_{j+k})e^{-i\alpha k} \tag{2.1}$$

and

$$d_X(j;N) \triangleq \lim_{L\to\infty} \frac{1}{L}\sum_{k=k_0}^{k_0+L-1} E(X_{j+kN}) \tag{2.2}$$

exist for some real numbers α, and some natural numbers N. We want to determine the conditions under which the probabilistic parameters $c_X(j;\alpha)$ and $d_X(j;N)$ can be consistently estimated from one sample path of the process X. Thus, by letting X be a function of another random process, say Y (e.g., $X_j = Y_{j+j_0}Y_j$, with fixed parameter j_0), we can apply these conditions for estimating cyclic components of the time-varying mean to problems of estimation of cyclic components of other probabilistic parameters such as the autocorrelation

$$E(X_j) = E(Y_{j+j_0}Y_j)$$

or a spectral density as described in [15]. The term *cyclic component* is used here since it follows from (2.1) and (2.2) that $c_X(j; \alpha)$ is sinusoidal,

$$c_X(j+k; \alpha) = c_X(j; \alpha)e^{i\alpha k}, \qquad \forall j, k \in Z,$$

and that $d_X(j; N)$ is periodic,

$$d_X(j+kN; N) = d_X(j; N), \qquad \forall j, k \in Z.$$

Definition: For any real number α, X is said to be *α-cycloergodic in the mean* if

$$\lim_{K\to\infty} E\left(\left|\frac{1}{K}\sum_{k=k_0}^{k_0+K-1}\left[X_{j+k} - E(X_{j+k})\right]e^{-i\alpha k}\right|^2\right) = 0 \tag{2.3}$$

for all $j, k_0 \in Z$.

Definition: For any positive integer N, X is said to be *N-ergodic in the mean* if

$$\lim_{L\to\infty} E\left(\left|\frac{1}{L}\sum_{k=k_0}^{k_0+L-1}\left[X_{j+kN} - E(X_{j+kN})\right]\right|^2\right) = 0 \tag{2.4}$$

for all $j, k_0 \in Z$.

Let the covariance of X_p and X_q be denoted by $\mathrm{cov}[X_p, X_q]$.

Theorem 2.1: a) The following are equivalent:

$$\lim_{K\to\infty}\frac{1}{K^2}\sum_{p=k_0}^{k_0+K-1}\sum_{q=k_0}^{k_0+K-1}\mathrm{cov}\left[X_{j+p}, X_{j+q}\right]e^{-i\alpha(p-q)} = 0; \tag{2.3a}$$

$$\lim_{K\to\infty}\mathrm{cov}\left[X_{j+k_0+K-1}, \frac{1}{K}\sum_{k=k_0}^{k_0+K-1}X_{j+k}e^{-i\alpha k}\right] = 0; \tag{2.3b}$$

X is α-cycloergodic in the mean.

b) The following are equivalent:

$$\lim_{L\to\infty}\frac{1}{L^2}\sum_{p=k_0}^{k_0+L-1}\sum_{q=k_0}^{k_0+L-1}\mathrm{cov}\left[X_{j+pN}, X_{j+qN}\right] = 0; \tag{2.4a}$$

$$\lim_{L\to\infty}\mathrm{cov}\left[X_{j+(k_0+L-1)N}, \frac{1}{K}\sum_{k=k_0}^{k_0+L-1}X_{j+kN}\right] = 0; \tag{2.4b}$$

X is N-ergodic in the mean.

The proofs that (2.3) $\Leftrightarrow$ (2.3a) and (2.4) $\Leftrightarrow$ (2.4a) are trivial. The proofs that (2.3a) $\Leftrightarrow$ (2.3b) and (2.4a) $\Leftrightarrow$ (2.4b) are analogous to the argument (for the case $\alpha = 0$) given in [16, pp. 74–75].

In general, α-cycloergodicity for one value of α does not imply α-cycloergodicity for any other value of α.

Definition: If X is α-cycloergodic in the mean for all real α, then we shall say that X is *cycloergodic in the mean*.

It follows from the next theorem that the preceding ergodic properties are precisely the properties required for consistent estimation of the cyclic components of a nonstationary mean. Define estimators by

$$\hat{c}_X^{(K)}(j; \alpha) \triangleq \frac{1}{K}\sum_{k=k_0}^{k_0+K-1}X_{j+k}e^{-i\alpha k} \tag{2.5}$$

and

$$\hat{d}_X^{(L)}(j; N) \triangleq \frac{1}{L}\sum_{k=k_0}^{k_0+L-1}X_{j+kN}. \tag{2.6}$$

Theorem 2.2: a) If the limit (2.1) exists and X is α-cycloergodic in the mean, then

$$\lim_{K\to\infty}E\left(\left|\hat{c}_X^{(K)}(j; \alpha) - c_X(j; \alpha)\right|^2\right) = 0, \tag{2.7}$$

for all $j \in Z$. b) If the limit (2.2) exists and X is N-ergodic in the mean, then

$$\lim_{L\to\infty}E\left(\left|\hat{d}_X^{(L)}(j; N) - d_X(j; N)\right|^2\right) = 0, \tag{2.8}$$

for all $j \in Z$.

The proofs of a) and b) are straightforward.

The two estimators $\hat{c}_X^{(K)}$ and $\hat{d}_X^{(L)}$ are not unrelated. In fact, each can (in principle, at least) be obtained from the other, in the special case for which $\alpha/2\pi$ is rational.

Theorem 2.3: a) If $\alpha/2\pi$ is rational, then let $N\alpha/2\pi$ be an integer. It follows that

$$\frac{1}{N}\sum_{p=0}^{N-1}\hat{d}_X^{(L)}(j+p; N)e^{-i\alpha p} = \hat{c}_X^{(K)}(j; \alpha), \tag{2.9}$$

where $K = LN$. b) Let $\alpha_m = 2\pi m/N$, and let K/N be an integer. It follows that

$$N\lim_{M\to\infty}\frac{1}{2M+1}\sum_{|m|\le M}\hat{c}_X^{(K)}(j; \alpha_m) = \hat{d}_X^{(L)}(j; N), \tag{2.10}$$

where $L = K/N$.

The proofs are straightforward after substitution of (2.5) and (2.6) into (2.9) and (2.10).

Although the estimator $\hat{d}_X^{(L)}(j; N)$ can always be approximated (M must be finite) with the estimators $\{\hat{c}_X^{(K)}(j; \alpha_m)\}$ using (2.10), the estimator $\hat{c}_X^{(K)}(j; \alpha)$ can be obtained from the estimators $\{\hat{d}_X^{(L)}(j+p; N)\}$ using (2.9) only if $\alpha/2\pi$ is rational. Therefore α-cycloergodicity is quite distinct from N-ergodicity when $\alpha/2\pi$ is irrational.

Wide-Sense Almost Cyclostationary Processes

The preceding results on estimation of cyclic components of a nonstationary mean can be applied to the problem of estimation of the entire nonstationary mean provided that it is an *almost periodic* (AP) function (in the mathematical sense). A process whose mean and autocovariance are AP is said to be wide-sense almost cyclostationary [8].

Let $\{E(X_j)\}$ be AP in the sense of Bohr [9], with associated Fourier series

$$E(X_j) \sim \sum_{m \in Z} c_X(\alpha_m) e^{i\alpha_m j}. \tag{2.11}$$

Since $c_X(\alpha_m)e^{i\alpha_m j} \equiv c_X(j;\, \alpha_m)$, then $E\{X_j\}$ can, in principle, be estimated via estimates of $\{c_X(j;\, \alpha_m)\}$ (e.g., [17, p. 59]).

Theorem 2.4: If $\{E(X_j)\}$ is AP and the partial sums of the Fourier series (2.11) converge, and X is cycloergodic in the mean, then

$$\lim_{M\to\infty} \lim_{K\to\infty} E\left(\left| \sum_{|m| \le M} \hat{c}_X^{(K)}(j;\, \alpha_m) - E(X_j) \right|^2 \right) = 0 \tag{2.12}$$

for all $j \in Z$.

Even if the partial sums in (2.11) do not converge, there is an alternative method of summation [9, p. 41] that can be used to obtain an alternative to Theorem 2.4.

In view of the relationship between $\hat{d}_X^{(L)}$ and $\hat{c}_X^{(K)}$ when $\alpha_m/2\pi$ is rational, it is not surprising that, in this case, we have the following alternative to the estimation procedure suggested by (2.12). If $c_X(j;\, \alpha) = 0$ for all irrational $\alpha/2\pi$, then $\{E(X_j)\}$ is said to be *limit periodic* (LP) and

$$\lim_{N\to\infty} d_X(j;\, N) = E(X_j) \tag{2.13}$$

uniformly in j [18]. This yields the following theorem, of which an example is given in [19].

Theorem 2.5: If $\{E(X_j)\}$ is LP and X is N-ergodic in the mean for all N, then

$$\lim_{N\to\infty} \lim_{L\to\infty} E\left(\left| \hat{d}_X^{(L)}(j;\, N) - E(X_j) \right|^2 \right) = 0. \tag{2.14}$$

It should be mentioned that if the convergence in (2.2) and (2.4) is uniform in j, then the convergence in (2.8) and (2.14) is uniform in j.

Mixing Conditions

Observe that in order to apply Theorem 2.4, one must know that X is α-cycloergodic in the mean for a possibly infinite set of values of α. Note further that in order to apply Theorem 2.2 to the problem of estimation of a number of probabilistic parameters in addition to the mean (e.g., the autocorrelation evaluated at a number of lag values, j_0), one must know that each *measurement process* whose mean is to be estimated (e.g., $X_j = Y_{j+j_0}Y_j$ for the problem of estimation of the autocorrelation of Y) is α-cycloergodic in the mean. Therefore, it would be helpful to have a single condition (property) that guarantees α-cycloergodicity in the mean for all measurement processes of interest and for all real α. Specifically, we would like a property of a process Y that guarantees the property (2.3) for all measurement processes of interest, X, on the process Y and for all real α.

In the latter part of Section III, several *mixing properties* that provide the desired guarantee are given. Although such mixing properties are rarely amenable to verification in practice, they do nevertheless provide a conceptual link between the wide-sense theory presented in this section, and the strict-sense theory presented in the following section.

III. Strict-Sense Theory

The major drawback of a wide-sense ergodic theory is that it deals with only mean square convergence. Thus, Theorem 2.2 does not guarantee that for a given sample path of X the sequence of estimators $\{\hat{c}_X^{(K)}(j;\, \alpha)\colon K = 1, 2, 3, \cdots\}$ converges. And even if it does converge for a given sample path, Theorem 2.2 does not guarantee that it converges to the appropriate quantity, viz., $c_X(j;\, \alpha)$. This theorem guarantees only that convergence (to $c_X(j;\, \alpha)$) occurs on the average over the ensemble of all possible sample paths.

Unfortunately (from a practical point of view), the development of a stronger theory of cycloergodicity is intimately tied to the development of more abstract mathematical concepts. We therefore begin this section with measure theoretic versions of definitions of classes of processes that were loosely defined in the Introduction. As is well-known for the property of stationarity, the class of wide-sense stationary processes is much more inclusive than the class of strict-sense stationary processes. Similarly, the class of processes on which our strict-sense theory is built does not include many of the processes to which our wide-sense theory applies.

Let R denote the real numbers, and let R^∞ denote the space of all two-way sequences $x = (\cdots, x_{-1}, x_0, x_1, \cdots)$ and let $\mathfrak{B}_\infty$ denote the smallest σ-algebra of subsets of R^∞ containing the finite-dimensional rectangles [20]. Define the coordinate functions $\{\xi_j\}$ on R^∞ by

$$\xi_j(x) \triangleq x_j$$

for $j \in Z$ and $x \in R^\infty$. Let S denote the shift operator on R^∞, defined by

$$\xi_j(Sx) = \xi_{j+1}(x).$$

Together with a probability measure μ on $(R^\infty, \mathfrak{B}_\infty)$, $\{\xi_j\} = \{\xi_0(S^j)\}$ comprises a discrete parameter stochastic process. In fact, $\{f(S^j)\}$ is a stochastic process on $(R^\infty, \mathfrak{B}_\infty, \mu)$ for any $\mathfrak{B}_\infty$-measurable function f on R^∞. In the special case $f = \xi_0$, μ is referred to as the *distribution* of the process. By the Daniell–Kolmogorov extension theorem [20], μ is uniquely determined by its values on finite-dimensional rectangles, i.e., by the finite-dimensional distributions of the process $\{\xi_j\}$. Thus, studying the process $\{\xi_j\}$ is equivalent to studying the measure μ. The properties of $\{\xi_j\}$ investigated in this section are more conveniently phrased in terms of μ.

μ is said to be *stationary* if $\mu = \mu S$, i.e.,

$$\mu(E) = \mu(SE), \qquad \forall E \in \mathfrak{B}_\infty.$$

We are concerned with the following generalizations of this concept. μ is said to be cyclostationary (CS) if there exists an integer N such that $\mu S^N = \mu$. μ is said to be *almost cyclostationary* (ACS) if, for each $E \in \mathfrak{B}_\infty$, $\{\mu(S^j E)\}$ is an almost periodic sequence [9]. μ is said to be *asymptotically mean stationary* (AMS) if the limit

$$\lim_{L\to\infty} \frac{1}{L} \sum_{k=0}^{L-1} \mu(S^k E) \tag{3.1}$$

exists for all $E \in \mathfrak{B}_\infty$. Another class of interest, apparently smaller than AMS, is the class of *asymptotically mean cyclostationary* (AMCS) measures. These have the property that the limit

$$\lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \mu(S^k E) e^{-i\alpha k} \tag{3.2}$$

exists for every $E \in \mathfrak{B}_\infty$ and every $\alpha \in R$. However, Theorem 3.1 and the bounded convergence theorem show that in fact AMS $\subseteq$ AMCS; hence, AMS = AMCS. Note, however, that

$$\text{CS} \underset{\neq}{\subset} \text{ACS} \underset{\neq}{\subset} \text{AMS}.$$

We now define the concepts of ergodicity that we shall study.

Definition: For any positive integer N, μ is said to be *N-ergodic* if for every $E \in \mathfrak{B}_\infty$ we have

$$\mu(E \Delta S^{\pm N} E) = 0 \Rightarrow \mu(E) = 0 \text{ or } 1, \tag{3.3}$$

where "Δ" denotes symmetric difference.

Of course, this is equivalent to

$$\mu\{x: f(x) = f(S^{\pm N} x)\} = 1 \Rightarrow \mu\{x: f(x) = E_\mu(f)\} = 1$$

for every $\mathfrak{B}_\infty$-measurable f on R^∞.

Definition: For any real number α, μ is said to be *α-cycloergodic* if for every $\mathfrak{B}_\infty$-measurable complex-valued f we have

$$\mu\{x: f(S^j x) = f(x) e^{i\alpha j}, \forall j \in Z\} = 1 \Rightarrow \mu\{x: f(x) = E_\mu(f)\} = 1. \tag{3.4}$$

It follows immediately from these definitions that 1-ergodicity is equivalent to 0-cycloergodicity, which is equivalent to (ordinary) ergodicity.

As far as ordinary ergodic theory is concerned, we can confine our attention to the class of AMS processes since it has been shown that $\mu \in$ AMS if and only if the individual ergodic theorem holds, i.e., if and only if

$$\mu\left\{x: \frac{1}{K} \sum_{k=0}^{K-1} f(S^k x) \quad \text{converges as} \quad K \to \infty\right\} = 1 \tag{3.5}$$

for every bounded measurable function f on $(R^\infty, \mathfrak{B}_\infty)$ [2]. Similarly, the following version of the individual ergodic theorem reveals that as far as cycloergodic theory is concerned, we can confine our attention to the class of AMCS = AMS processes, since $\mu \in$ AMS if and only if

$$\mu\left\{x: \frac{1}{K} \sum_{k=0}^{K-1} f(S^k x) e^{-i\alpha k} \quad \text{converges as} \quad K \to \infty\right\} = 1 \tag{3.6}$$

and

$$\mu\left\{x: \frac{1}{L} \sum_{k=0}^{L-1} f(S^{j+kN} x) \quad \text{converges as} \quad L \to \infty\right\} = 1 \tag{3.7}$$

for every bounded measurable f on $(R^\infty, \mathfrak{B}_\infty)$, all $\alpha \in R$, all $j \in Z$, and every positive integer N. In fact, as the following theorem shows, an even stronger statement is true. For $r \geq 1$, let $\mathfrak{U}^r(\mu)$ denote the class of $\mathfrak{B}_\infty$-measurable functions f on R^∞ such that $\{|f(S^j)|^r: j \in Z\}$ is uniformly μ-integrable [21, p. 629]. Note that $\mathfrak{U}^r(\mu)$ contains all bounded $\mathfrak{B}_\infty$-measurable functions.

Theorem 3.1: Let $r \geq 1$ and let μ be a probability measure on $(R^\infty, \mathfrak{B}_\infty)$. Then μ is AMS if and only if the following statements hold for every $f \in \mathfrak{U}^r(\mu)$: a) for each $\alpha \in R$ there exists $\hat{f}_\alpha \in \mathfrak{U}^r(\mu)$ such that

$$\frac{1}{K} \sum_{k=0}^{K-1} f(S^k) e^{-i\alpha k} \xrightarrow[L^r(\mu)]{\text{a.s. } [\mu]} \hat{f}_\alpha \tag{3.8}$$

as $K \to \infty$; b) For each positive integer N there exists $\tilde{f}_N \in \mathfrak{U}^r(\mu)$ such that for all $j \in Z$

$$\frac{1}{L} \sum_{k=0}^{L-1} f(S^{j+kN}) \xrightarrow[L^r(\mu)]{\text{a.s. } [\mu]} \tilde{f}_N(S^j) \tag{3.9}$$

as $L \to \infty$.

It is clear that a) and b) imply $\mu \in$ AMS. A proof of the converse is outlined in the Appendix. It should be emphasized that Theorem 3.1 reveals not only that we can confine our attention to the class AMS, but also that we must (initially) consider *all* of AMS, since AMS = AMCS. It is easily verified that $\{\hat{f}_\alpha(S^j)\}$ is sinusoidal with probability 1 and that $\{\tilde{f}_N(S^j)\}$ is N-periodic with probability one:

$$\mu\{x: \hat{f}_\alpha(S^j x) = \hat{f}_\alpha(x) e^{i\alpha j}, \forall j \in Z\} = 1; \tag{3.10}$$

$$\mu\{x: \tilde{f}_N(S^{j\pm N} x) = \tilde{f}_N(S^j), \forall j \in Z\} = 1. \tag{3.11}$$

Furthermore, it follows from Theorem 3.1 that

$$E_\mu(\hat{f}_\alpha) = \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} E_\mu(f S^k) e^{-i\alpha k} \tag{3.12}$$

and

$$E_\mu(\tilde{f}_N) = \lim_{L\to\infty} \frac{1}{L} \sum_{k=0}^{L-1} E_\mu(f S^{kN}). \tag{3.13}$$

The following theorem reveals that the properties of cycloergodicity and N-ergodicity are equivalent to certain *mixing properties*.

Theorem 3.2: Let μ be AMS, let $\alpha \in R$, and let N be a positive integer. a) The following are equivalent:

$$\lim_{K\to\infty} \frac{1}{K}\sum_{k=0}^{K-1}\left[\mu(S^kA\cap B)-\mu(S^kA)\mu(B)\right]e^{-i\alpha k}=0, \tag{3.4a}$$

for all $A, B \in \mathfrak{B}_\infty$;

$$\lim_{K\to\infty} \operatorname{cov}_\mu\left\{g, \frac{1}{K}\sum_{k=0}^{K-1} f(S^k)e^{-i\alpha k}\right\}=0, \tag{3.4b}$$

for all $f, g \in \mathfrak{U}^2(\mu)$;

μ is α-cycloergodic (i.e., satisfies (3.4)).

b) The following are equivalent:

$$\lim_{L\to\infty} \frac{1}{L}\sum_{k=0}^{L-1}\left[\mu(S^{kN}A\cap B)-\mu(S^{kN}A)\mu(B)\right]=0 \tag{3.3a}$$

for all $A, B \in \mathfrak{B}_\infty$;

$$\lim_{L\to\infty} \operatorname{cov}_\mu\left\{g, \frac{1}{L}\sum_{k=0}^{L-1} f(S^{kN})\right\}=0 \tag{3.3b}$$

for all $f, g \in \mathfrak{U}^2(\mu)$;

μ is N-ergodic (i.e., satisfies (3.3)).

The proof of Theorem 3.2 is by standard arguments and is therefore omitted.

Theorem 3.3: For any μ, N-ergodicity implies α-cycloergodicity for all $\alpha \in N^{-1}2\pi Z$. If μ is AMS, the converse holds.

A proof of Theorem 3.3 is outlined in the Appendix.

It follows from the next theorem that these ergodic properties are precisely the properties required for estimation of the cyclic components of the nonstationary mean of a large class of measurement functions on an AMS process.

Theorem 3.4: Let μ be AMS. Then a) μ is α-cycloergodic if and only if

$$\frac{1}{K}\sum_{k=0}^{K-1} f(S^k)e^{-i\alpha k}\xrightarrow[L^1(\mu)]{\text{a.s. }[\mu]} E_\mu(\hat{f}_\alpha) \tag{3.14}$$

as $K\to\infty$, for all $f \in \mathfrak{U}^1(\mu)$; b) μ is N-ergodic if and only if

$$\frac{1}{L}\sum_{k=0}^{L-1} f(S^{kN})\xrightarrow[L^1(\mu)]{\text{a.s. }[\mu]} E_\mu(\tilde{f}_N) \tag{3.15}$$

as $L\to\infty$, for all $f \in \mathfrak{U}^1(\mu)$.

This theorem follows immediately from Theorem 3.1, and the preceding definitions. The quantities $E_\mu(\hat{f}_\alpha)$ and $E_\mu(\tilde{f}_N)$ in (3.14) and (3.15) (cf. (3.12) and (3.13)) are the probabilistic parameters of interest in strict sense cycloergodic theory. We introduce notation consistent with that of Section II by defining

$$c_f(j;\alpha) \triangleq E_\mu(\hat{f}_\alpha S^j) = E_\mu(\hat{f}_\alpha)e^{i\alpha j},$$
$$c_f(\alpha) \triangleq c_f(0;\alpha) = E_\mu(\hat{f}_\alpha),$$
$$d_f(j;N) \triangleq E_\mu(\tilde{f}_N S^j).$$

The following theorem reveals that (by contrast with wide-sense α-cycloergodicity) strict sense α-cycloergodicity is of no practical interest when $\alpha/2\pi$ is irrational, because in this case $c_f(\alpha)$ vanishes for all AMS μ.

Theorem 3.5: Let μ be AMS and α-cycloergodic, where $\alpha/2\pi$ is irrational. Then,

$$c_f(\alpha)=0, \qquad \forall f \in \mathfrak{U}^1(\mu). \tag{3.16}$$

A proof of Theorem 3.5 is outlined in the Appendix. This result leads us to consider a smaller class of probability measures than the class of AMS measures as the appropriate domain of study for a theory of cycloergodicity. We denote by AMS* the subclass of AMS μ for which

$$c_f(\alpha)=0, \qquad \forall f \in \mathfrak{U}^1(\mu)$$

whenever $\alpha/2\pi$ is irrational. Furthermore, we define the class

$$\text{ACS*} \triangleq \text{ACS} \cap \text{AMS*}. \tag{3.17}$$

The following lemma shows that ACS* coincides with the class of *quasi-periodic* (QP) probability measures studied by Blum and Hanson [10]. μ is said to be QP if the following holds: given $\epsilon>0$ and $E \in \mathfrak{B}_\infty$ there exists a positive integer N such that

$$|\mu(S^jE)-\mu(S^{j+kN}E)|<\epsilon$$

for all integers j, k. Clearly

$$\text{QP} \subseteq \text{ACS}. \tag{3.18}$$

Lemma 3.6: Let h: $Z\to R$ be AP. Then the following are equivalent: a) $\lim_{K\to\infty}(1/K)\sum_{k=0}^{K-1}h(k)e^{-i\alpha k}=0$ if $\alpha/2\pi$ irrational; b) there is a sequence $\{h_q\}$ of periodic sequences such that $h_q\to h$ uniformly on Z; c) given $\epsilon>0$ there exists a positive integer N such that $|h(j)-h(j+kN)|<\epsilon$, for all integers j, k.

The equivalence of a) and b) is already part of the theory of almost periodic functions [18]; the proof that b) $\Leftrightarrow$ c) is straightforward, and is omitted. A sequence h: $Z\to R$ possessing property b) is said to be *limit-periodic* (LP).

The next theorem records some results from [10] concerning ACS*. First we introduce some additional terminology.

Definition: If μ is AMS*, and μ is α-cycloergodic for all α such that $\alpha/2\pi$ is rational, then we shall say that μ is *cycloergodic*.

In view of Theorem 3.3, μ is cycloergodic if and only if μ is N-ergodic for every positive integer N. Let ACS_0^* denote the class of cycloergodic measures in ACS*. For any μ in AMS, let $\bar{\mu}$ denote the stationary measure defined by

$$\bar{\mu}(A) \triangleq \lim_{K\to\infty}\frac{1}{K}\sum_{k=0}^{K-1}\mu(S^kA), \qquad A\in\mathfrak{B}_\infty.$$

Recall that $\mu \ll \bar{\mu}$ [2]. A set $A \in \mathfrak{B}_\infty$ is *periodic* if $S^N A = A$ for some N.

Theorem 3.7: (Blum–Hanson) Let μ be ACS*. Then we have a) μ is ACS_0^* if and only if μ is an extreme point of the set ACS*. b) If μ is ACS_0^* then $\bar{\mu}$ is ergodic. c) Let A belong to the σ-algebra generated by the periodic sets.

If for such A $\mu(A) > 0$ implies $\mu_0(A) > 0$ for some $\mu_0 \in \text{ACS}_0^*$, then we have a "cycloergodic decomposition" for μ. Specifically, there is a σ-algebra Σ of subsets of ACS_0^* and a probability measure P_μ on (ACS_0^*, Σ) such that for every $A \in \mathfrak{B}_\infty$ the function $\nu \to \nu(A)$ is Σ-measurable on ACS_0^*, and moreover

$$\mu(A) = \int_{\text{ACS}_0^*} \nu(A)\, dP_\mu(\nu), \qquad A \in \mathfrak{B}_\infty. \tag{3.19}$$

See [10] for proofs of the above statements and a more precise description of the decomposition (3.19). Also, it is shown in [10] that ACS* is strictly larger than the class of denumerable convex combinations of periodic (CS) probability measures.

Mixing Conditions

It follows immediately from Theorem 3.2a) that the following *weak mixing property* is sufficient for cycloergodicity:

$$\lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \left|\mu(S^k A \cap B) - \mu(S^k A)\mu(B)\right| = 0, \tag{3.20}$$

for all $A, B \in \mathfrak{B}_\infty$. A similar kind of mixing property plays a role in the wide-sense theory presented in Section II.

Consider the following three properties:

$$\lim_{|k|\to\infty} \left[\mu(S^{j+k}A \cap S^j B) - \mu(S^{j+k}A)\mu(S^j B)\right] = 0 \tag{3.21}$$

uniformly in $j \in Z$ for all $A, B \in \mathfrak{B}_\infty$;

$$\lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \left|\mu(S^k A \cap S^K B) - \mu(S^k A)\mu(S^K B)\right| = 0, \tag{3.22}$$

for all $A, B \in \mathfrak{B}_\infty$;

$$\lim_{K\to\infty} \text{cov}_\mu\left\{ f(S^{K-1}), \frac{1}{K}\sum_{k=0}^{K-1} f(S^k)e^{-i\alpha k}\right\} = 0, \tag{3.23}$$

for all $f \in \mathfrak{U}^2(\mu)$, and all $\alpha \in R$.

Property (3.21) is *uniform strong mixing*, and is equivalent to ordinary *strong mixing* when μ is stationary. Property (3.23) is simply the cycloergodicity-in-the-mean property (2.3b) for all $f \in \mathfrak{U}^2(\mu)$. Furthermore, property (3.23) modified by the replacement $f(S^{K-1}) \to g$ is simply the cycloergodicity property (3.4b). Similarly property (3.22) modified by the replacement $S^K B \to B$ is the weak mixing property (3.20).

Theorem 3.8: (3.21) ⇒ (3.22) ⇒ (3.23).

The proof is straightforward.

Estimation for ACS Processes

Theorem 3.9 below is the strict sense analog of Theorem 2.4 on mean square estimation of an AP expectation. It is easily shown from the definitions of ACS and ACS* that a measure μ is ACS (resp. ACS*) if and only if $\{E_\mu(fS^j)\}$ is AP (resp. LP) for every $f \in \mathfrak{U}^1(\mu)$. Let μ be ACS* and denote the Fourier series associated with the LP sequence $\{E_\mu(fS^j)\}$ by

$$E_\mu(fS^j) \sim \sum_{m\in Z} c_f(\alpha_m)e^{i\alpha_m j}, \tag{3.24}$$

for any $f \in \mathfrak{U}^1(\mu)$.

Theorem 3.9: Let μ be ACS_0^*, and let $f \in \mathfrak{U}^1(\mu)$. Assume that the Fourier series (3.24) converges. Then for every $j \in Z$

$$\lim_{M\to\infty} \lim_{K\to\infty} \sum_{|m|\le M} \hat{c}_f^{(K)}(\alpha_m)e^{i\alpha_m j} = E_\mu(fS^j) \tag{3.25}$$

in $L^1(\mu)$ and with μ-probability one, where $\hat{c}_f^{(K)}(\alpha) \triangleq (1/K)\sum_{k=0}^{K-1} f(S^k)e^{-i\alpha k}$.

Theorem 3.9 follows immediately from Theorem 3.4. Theorem 3.10, the strict-sense analog of Theorem 2.5, is also an immediate consequence of Theorem 3.4.

Theorem 3.10: Let μ be ACS_0^* and let $f \in \mathfrak{U}^1(\mu)$. Then for all $j \in Z$,

$$\lim_{N\to\infty} \lim_{L\to\infty} \hat{d}_f^{(L)}(j; N) = E_\mu(fS^j) \tag{3.26}$$

in $L^1(\mu)$ and with probability one, where $\hat{d}_f^{(L)}(j; N) \triangleq (1/L)\sum_{k=0}^{L-1} f(S^{j+kN})$.

Gaussian ACS Processes

We denote the mean and correlation sequences for a process $\{\xi_j\}$ by

$$M(j) \triangleq E_\mu(\xi_j)$$
$$K(j,i) \triangleq E_\mu(\xi_j \xi_{j+i}).$$

It is well-known that if $\{\xi_j\}$ is Gaussian, then μ is stationary if and only if $M(j)$ and $K(j,i)$ are independent of j. Furthermore, if $\{\xi_j\}$ is Gaussian, then μ is cyclostationary if and only if there exists an integer N such that $M(j)$ and $K(j,i)$ are each N-periodic in j for all i. One might suspect that a similar result would hold for the class of ACS* (or ACS) measures. Our present understanding of the situation is summarized by the following theorem and corollary. Let $\mathfrak{F}$ denote the field of finite-dimensional sets in $\mathfrak{B}_\infty$, and let $C(s_1,\cdots,s_n)$ denote the covariance matrix for $\xi_{s_1}, \xi_{s_2}, \cdots, \xi_{s_n}$.

Theorem 3.11: Let $\{\xi_j\}$ be Gaussian. If for all $i \in Z$, the sequences $\{M(j)\}$ and $\{K(j,i)\}$ are AP (resp. LP) in j, and if the determinant of $C(s_1+j,\cdots,s_n+j)$ is bounded away from zero uniformly in j for every fixed positive integer n and every fixed set of n integers $s_1, s_2, \cdots, s_n$, then the sequence $\{\mu(S^j A)\}$ is AP (resp. LP) for all $A \in \mathcal{F}$.

Corollary 3.12: Assume, in addition to the hypotheses of Theorem 3.11, that μ is AMS. Then μ is ACS (resp. ACS*).

Proofs of Theorem 3.11 and Corollary 3.12 are outlined in the Appendix.

An open question is whether or not for a Gaussian process, wide-sense cycloergodic properties imply corresponding strict-sense properties.

IV. Conclusion

To the extent that the wide-sense and strict-sense theories of cycloergodicity presented in this paper mimic the corresponding theories of ordinary ergodicity, these theories appear to be appropriate, with regard to the general application of estimating cyclic components of nonstationary expectations, as discussed briefly in the Introduction, and in more detail in [15]. Nevertheless, there are some bothersome aspects of the strict-sense theory. In particular, the fact that for an AMS α-cycloergodic measure, the cyclic components $c_f(\alpha)$ corresponding to all uniformly integrable measurement functions f vanish if $\alpha/2\pi$ is irrational is disturbing. To illustrate, we consider the example of a nonstationary process composed of independent Bernoulli variables $\{\xi_j\}$ whose sequence of parameters $\{p_j\}$ is AP, e.g.,

$$p_j = \frac{1}{2} + \frac{1}{4}\cos(j).$$

For $f = \xi_0$, we have

$$\begin{aligned} c_f(1) &= \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} E_\mu(\xi_k) e^{-ik} \\ &= \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} p_k e^{-ik} \\ &= 1/8 \neq 0. \end{aligned}$$

Thus, $c_f(\alpha) \neq 0$ for an irrational $\alpha/2\pi$, viz., $\alpha/2\pi = 1/2\pi$. It follows from Theorem 3.5 that μ is either not AMS, or not α-cycloergodic for $\alpha = 1$, and therefore Theorem 3.4a) does not apply to this Bernoulli process. Nevertheless, by the strong law of large numbers, the AP sequence of parameters $\{p_j\}$ can be estimated with probability-one convergence of estimators. More generally, for any process $\{\xi_j\}$ consisting of integrable independent random variables, one can construct a strongly consistent estimator of $c_f(\alpha)$ (where $f = \xi_0$) whenever $c_f(\alpha)$ exists. However, if $c_f(\alpha) \neq 0$ for irrational $\alpha/2\pi$ then by Theorem 3.5 μ is either not AMS or not α-cycloergodic. Since neither of these properties can be deduced from finite-dimensional considerations, we cannot be more specific.

The preceding discussion shows that there exist processes that possess strict-sense cycloergodic properties, but do not belong to the class AMS*. Thus, a strict-sense theory of cycloergodicity inclusive enough to cover all applications of practical interest does not yet exist. Moreover, such a theory cannot presuppose the existence of a dominating stationary measure, as does the theory presented herein.

Appendix

Proof of Theorem 3.1: Assume μ is AMS. The formula,

$$\bar{\mu}(A) = \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \mu(S^k A), \qquad A \in \mathcal{B}_\infty,$$

defines a stationary probability measure on $(R^\infty, \mathcal{B}_\infty)$ having the property that $\mu \ll \bar{\mu}$ [2]. The convergences in a) and b) follow, with μ replaced by $\bar{\mu}$, from the individual ergodic theorem for stationary processes [21]. Now use $\mu \ll \bar{\mu}$ to obtain almost sure $[\mu]$ convergence. The $L^r(\mu)$ convergence now follows from the uniform μ-integrability of $\{|f(S^j)|^r\}$ by standard arguments.

Proof of Theorem 3.3: The first statement is obvious. For the second, assume μ is AMS and α-cycloergodic for all $\alpha \in N^{-1}2\pi Z$. Let $A, B \in \mathcal{B}_\infty$ and define

$$h(k) \triangleq \mu(S^k A \cap B) - \mu(S^k A)\mu(B)$$

for $k \in Z$. Now by Theorem 3.1b) the limit

$$\tilde{1}_A(j) \triangleq \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \tilde{1}_{S^{j+kN}A}$$

exists almost surely $[\mu]$ for all $j \in Z$. Here $\tilde{1}_E$ denotes the indicator function of any event $E \in \mathcal{B}_\infty$. It follows that

$$\tilde{1}_A(j)[\tilde{1}_B - \mu(B)] = \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \tilde{1}_{S^{j+kN}A}[\tilde{1}_B - \mu(B)]$$

almost surely $[\mu]$. Let $\eta(j)$ denote the integral $(d\mu)$ of the left member in the above equation. By the bounded convergence theorem we have

$$\eta(j) = \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} h(j + kN).$$

Thus $\eta: Z \to R$ is a periodic sequence with period N. Moreover, for $n \in Z$ we have

$$\begin{aligned} &\frac{1}{N} \sum_{j=0}^{N-1} \eta(j) e^{-i2\pi jn/N} \\ &= \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \frac{1}{N} \sum_{j=0}^{N-1} h(j+kN) e^{-i2\pi jn/N} \\ &= \lim_{K\to\infty} \frac{1}{KN} \sum_{k=0}^{KN-1} h(k) e^{-i2\pi kn/N} = 0 \end{aligned}$$

by Theorem 3.2a). Thus we have

$$\eta(j) = 0, \qquad j \in Z.$$

By Theorem 3.2b), μ is N-ergodic.

Proof of Theorem 3.5: Let $f \in \mathcal{U}^1(\mu)$ and assume $c_f(\alpha) \neq 0$. Define

$$\Omega = \left\{x: \frac{1}{K} \sum_{k=0}^{K-1} f(S^k x) e^{-i\alpha k} \to c_f(\alpha)\right\}.$$

Then $\Omega \in \mathcal{B}_\infty$, and by Theorem 3.4 $\mu(\Omega) = 1$. Moreover

$$S^n\Omega = \left\{x: \frac{1}{K} \sum_{k=0}^{K-1} f(S^k x) e^{-i\alpha k} \to c_f(\alpha) e^{i\alpha n}\right\}.$$

Since $c_f(\alpha) \neq 0$, we obtain $(S^n\Omega) \cap (\Omega) = \varnothing$ unless $n \in \alpha^{-1}2\pi Z$, in which case $S^n\Omega = \Omega$. Suppose that $Z \cap (\alpha^{-1}2\pi Z) = \{0\}$. Then

$$(S^n\Omega) \cap (S^m\Omega) = \varnothing, \qquad n \neq m.$$

But $\mu \ll \bar{\mu}$, so $\mu(\Omega) = 1 > 0 \Rightarrow \bar{\mu}(\Omega) > 0 \Rightarrow \bar{\mu}(S^n\Omega) = \bar{\mu}(\Omega) > 0$ for all $n \in Z$, since $\bar{\mu}$ is stationary. Thus

$$\bar{\mu}\left(\bigcup_{n\in Z} S^n\Omega\right) = \sum_{n\in Z} \bar{\mu}(S^n\Omega) = +\infty,$$

a contradiction. It follows that $Z \cap (\alpha^{-1}2\pi Z)$ contains $n \neq 0$, i.e., $\alpha = 2\pi k/n$ for nonzero integers k, n. This contradicts the assumption that $\alpha/2\pi$ is irrational, hence we must have $c_f(\alpha) = 0$.

Proof of Theorem 3.11: Let $s_1,\cdots,s_n$ be fixed integers. Set $C(j) = C(s_1+j,\cdots,s_n+j)$ for all $j \in Z$. The joint density of $\xi_{s_1+j},\cdots,\xi_{s_n+j}$ is given by

$$f_j(t) = (2\pi)^{-n/2}|C(j)|^{-1/2}e^{-Q(j,t)}, \qquad t \in R^n,$$

where

$$Q(j,t) \triangleq \frac{1}{2}(t - m(j))^T C(j)^{-1}(t - m(j))$$

and

$$m(j) \triangleq [M(s_1+j),\cdots,M(s_n+j)]^T.$$

Assume $|C(j)| \geq \epsilon > 0$ for all $j \in Z$. The components of $C(j)$ are AP (resp. LP) by assumption, hence [9, pp. 11–13] $|C(j)|^{-1/2}$ and the components of $C(j)^{-1}$ are AP (resp. LP) (cf. Lemma 3.6c)). The components of $m(j)$ are AP (resp. LP) by assumption. Now let $t \in R^n$ and let j, k be integers. We have

$$\begin{aligned}
&2[Q(j,t) - Q(k,t)]\\
&\quad = t^T\left(C(j)^{-1} - C(k)^{-1}\right)t\\
&\qquad + \left[m(k)^T C(k)^{-1} - m(j)^T C(j)^{-1}\right]t\\
&\qquad + t^T\left[C(k)^{-1}m(k) - C(j)^{-1}m(j)\right]\\
&\qquad + m(j)^T C(j)^{-1}m(j) - m(k)^T C(k)^{-1}m(k).
\end{aligned}$$

For a vector or matrix a let $\|a\|^2$ denote the sum of the squares of the components of a. Then

$$\begin{aligned}
&|2(Q(j,t) - Q(k,t))|\\
&\quad \leq \|t\|^2\|C(j)^{-1} - C(k)^{-1}\|\\
&\qquad + 2\|m(k)^T C(k)^{-1} - m(j)^T C(j)^{-1}\|\|t\|\\
&\qquad + |m(j)^T C(j)^{-1}m(j) - m(k)^T C(k)^{-1}m(k)|.
\end{aligned}$$

This inequality shows that if B^n is any bounded set in R^n, $\{Q(j,t)\}$ is AP (resp. LP) uniformly with respect to $t \in B^n$ [9, p. 51]. From this it follows that $\{f_j(t)\}$ has the same property, hence

$$\left\{\int_{B^n} f_j(t)\,dt\right\}$$

is AP (resp. LP) for all bounded Borel sets B^n.

Now let B^n be any n-dimensional Borel set. Clearly, for any $R > 0$

$$\begin{aligned}
&\left|\int_{B^n} f_j(t)\,dt - \int_{B^n} f_k(t)\,dt\right|\\
&\quad \leq \left|\int_{B^n\cap\{\|t\|^2\leq R\}} f_j(t)\,dt - \int_{B^n\cap\{\|t\|^2\leq R\}} f_k(t)\,dt\right|\\
&\qquad + \int_{\|t\|^2>R} f_j(t)\,dt + \int_{\|t\|^2>R} f_k(t)\,dt.
\end{aligned}$$

Thus we see that in order to prove that

$$\int_{B^n} f_j(t)\,dt$$

is AP (resp. LP) it suffices to verify

$$\lim_{R\to\infty}\int_{\|t\|^2>R} f_j(t)\,dt = 0, \qquad \text{uniformly in } j. \qquad (5.1)$$

To prove (5.1), first assume $m(j) = 0$ for all $j \in Z$. Since the components of $C(j)$ are bounded functions of j, the eigenvalues of $C(j)$ are bounded positive functions of j. Reducing $C(j)$ to diagonal form and letting $0 < M_1 < +\infty$ be a universal upper bound on the aforementioned eigenvalues, we obtain

$$\int_{\|t\|^2>R} f_j(t)\,dt \leq \int_{\|t\|^2>R} e^{-\|t\|^2/2M_1}\,dt.$$

This proves (5.1) for the zero-mean case. To settle the general case, simple arguments yield

$$\int_{\|t\|^2>R} e^{-Q(j,t)}\,dt \leq \int_{\|t\|^2>(R^{1/2}-M_2)^2} e^{(-1/2)t^T C(j)^{-1}t}\,dt$$

where $M_2 = \sup_j\|m(j)\| < +\infty$. By the zero-mean case, the right member converges to zero uniformly in j, so we are done.

Proof of Corollary 3.12: Since $\mu \ll \bar{\mu}$, given $\epsilon > 0$ we can find $\delta > 0$ such that for all $A \in \mathcal{B}_\infty$ we have

$$\bar{\mu}(A) < \delta \Rightarrow \mu(A) < \epsilon.$$

Let $E \in \mathcal{B}_\infty$. Then there is $E_0 \in \mathcal{F}$ such that

$$\bar{\mu}(E\Delta E_0) < \delta.$$

Since $\bar{\mu}$ is stationary we have

$$\bar{\mu}(S^qE\Delta S^qE_0) < \delta, \qquad \forall q \in Z,$$

hence

$$\mu(S^qE\Delta S^qE_0) < \epsilon, \qquad \forall q \in Z,$$

hence

$$|\mu(S^qE) - \mu(S^qE_0)| < \epsilon, \qquad \forall q \in Z.$$

By Theorem 3.11, $\{\mu(S^qE_0)\}$ is AP (resp. LP). Since $\epsilon > 0$ was arbitrary, it follows that $\{\mu(S^qE)\}$ is AP (resp. LP).

References

[1] E. Parzen, "Spectral analysis of asymptotically stationary time series," *Bull. Inst. Int. Statist.*, vol. 39, Livraison 2, pp. 87–103, 1962.

[2] R. M. Gray and J. C. Kieffer, "Asymptotically mean stationary measures," *Ann. Prob.*, vol. 8, pp. 962–973, 1980.

[3] J. Kampé de Fériet and F. N. Frenkiel, "Correlations and spectra for non-stationary random functions," *Math. Comput.*, vol. 16, no. 77, pp. 1–21, Jan. 1962.

[4] L. I. Gudzenko, "On periodic nonstationary processes," *Radio Eng. Electron. Phys.* (USSR), vol. 4, no. 6, pp. 220–224, 1959 (trans.).

[5] J. Nėdoma, "Über die ergodizität and *r*-ergodizitat stationärer wahrscheinlichkeitsmasse," *Z. Wahrscheinlichkeitstheorie*, vol. 2, pp. 90–97, 1963.

[6] H. L. Hurd, "An investigation of periodically correlated stochastic processes," Ph.D. dissertation, Duke University, Durham, NC, 1969.

[7] 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.

[8] W. A. Gardner, "Stationarizable random processes," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 8–22, Jan. 1978.

[9] C. Corduneanu, *Almost Periodic Functions.* New York: Wiley 1961.

[10] J. R. Blum and D. L. Hanson, "On quasi-periodic probability measures," *J. Math. Mech.*, vol. 15, pp. 291–303, 1966.

[11] K. Jacobs, "Fastperiodische markoffsche prozesse," *Math. Ann.*, vol. 134, pp. 408–427, 1958.

[12] K. Jacobs, "Die ubertragung diskreter informationen durch periodische und fastperiodische kavale," *Math. Ann.*, vol. 137, pp. 125–135, 1959.

[13] K. Jacobs, "Almost periodic sources and channels," *Z. Wahrscheinlichkeitstheorie verw*, Geb. 9, pp. 65–84, 1967.

[14] L. E. Franks, *Signal Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1969.

[15] W. A. Gardner, "Cyclic spectral analysis," in *Statistical Spectral Analysis: A Non-Probabilistic Theory*, in preparation for publication.

[16] E. Parzen, *Stochastic Processes*. San Francisco, CA: Holden-Day, 1962.

[17] L. H. Koopmans, *The Spectral Analysis of Time Series*. New York: Academic, 1974.

[18] A. S. Besicovitch, *Almost Periodic Functions*. New York: Dover Publ., 1954.

[19] L. Herbst, "Almost periodic variances," *Ann. Math. Stat.*, vol. 34, pp. 1549–1557, 1963.

[20] K. Jacobs, *Measure and Integral*. New York: Academic, 1978.

[21] J. L. Doob, *Stochastic Processes*. New York: John Wiley, 1953.