A Statistical Theory of Signal Coherence
Melvin J. Hinich

Abstract—A periodic signal can be perfectly predicted far into the future since it perfectly repeats every period. There is always some variation in the waveform over time for signals which are labeled as periodic but which are not truly deterministic. A formal definition is presented in this paper for such a varying periodic signal and the properties of such a class of signals are exploited. A measure called a signal coherence function of the amount of random variation in each Fourier component of the signal is defined and its statistical properties are developed. This signal coherence function is very different from the coherence function between two stationary signals. The method is applied to a digitized record of an acoustic signal generated by a boat in a bay in the Baltic Sea south of Stockholm, Sweden.

Index Terms—Cyclostationarity, randomly modulated periodicity, signal coherence, stationary, synchronization, waveform.

I. INTRODUCTION

A PERIODIC signal can be perfectly predicted far into the future since it perfectly repeats every period. Nature does not produce perfect periodic signals. There is always some variation in the waveform over time for signals which are labeled as periodic but which are not truly deterministic. A formal definition is presented in this paper for such a varying periodic signal that is called a randomly modulated periodicity with period $T$.

Definition: A signal $\{x(t)\}$ is called a randomly modulated periodicity with period $T$ if it is of the form

$$x(t) = K^{-1/2} \sum_{k=-K/2}^{K/2} [\mu_k + \rho_k(t)] \exp(i2\pi f_k t)$$

for $f_k = k/T$.

The signal can be written as $x(t) = s(t) + u(t)$ where

$$s(t) = K^{-1} \sum_{k=-K/2}^{K/2} \mu_k \exp(i2\pi f_k t)$$

and

$$u(t) = K^{-1} \sum_{k=-K/2}^{K/2} \rho_k(t) \exp(i2\pi f_k t).$$

The periodic component $s(t)$ is the mean of $x(t)$. The zero mean stochastic term $u(t)$ is a real valued nonstationary process, which may be nonlinear.

Condition 1) implies that $c_{\mu}(t_1, t_2) = E[u(t_1 + T)u(t_2 + T)] = E[u(t_1)u(t_2)]$ if $|t_1 - t_2| < T$ but the equality does not necessarily hold when $t_1$ and $t_2$ are in different frames, i.e., when $|t_1 - t_2| > T$. Consequently, the stochastic term is not cyclostationary [1].

Even if the $\rho_k(t)$ are all jointly stationary random processes, $u(t)$ is not, in general, stationary since the joint distribution of $\{u_k(t_1), \ldots, u_k(t_n)\}$ is not invariant to a shift of the time origin. For example, the covariance function of $u_k(t)$ $\sin(2\pi f t)$ is a function of $t$ if $u_k(t)$ is a covariance stationary process. Periodic block stationarity is a different condition than cyclostationarity and covariance stationarity.

Condition 2) insures that $u(t)$ has a finite dependence on gap length $D$. It then follows that all the joint cumulants of $u(t)$ are $D$-dependent.

A randomly modulated periodic signal is created by some nonlinear physical mechanism, which has a more or less stable inherent periodicity. For example, voiced speech is randomly modulated since an oscillating vocal cord’s amplitude and phase slowly vary over several pitch periods in a seemingly random fashion.

Other examples are: 1) the vibration signal generated by a rotating machine such as a pump, motor, or drill and 2) the received train of reflected waveforms generated by an active torpede sonar which is pinging on a target to estimate the target’s range and bearing relative to the sonar.

Fig. 1 presents a plot of the following version of (1) for $f_1 = 1/50$ and $f_2 = 2/50$ where the $\{u_k(t)\}$ are autoregressive AR(1) processes generated by the first-order difference equation $u_k(t) - 0.95u_k(t - 1) = c_{\epsilon}(t)$.

Assume at first that $T$ is known. The highest harmonic frequency is $f_K = K/2T$. If the signal is sampled at the frequency $2f_{k}$, there are $K$ discrete-time samples $x(t_n)$ in an interval of
length $T$ time units. To simplify notation, the time unit will be suppressed and we will let $t$ denote $t_n$ and $T = K$ from now on.

A common approach to processing signals with a periodic structure is to segment the observations into frames of length $T$ so that there is exactly one waveform in each sampling frame. The term “sampling frame” is used in this paper in order to match the terminology used in speech processing literature. The waveform in frame $m$ is slightly different from that in frame $m + 1$ due to variations in the vocal tract.

To further simplify notation, set the time origin at the start of the first frame. Then the start of the $m$th frame is $\beta_m = (m-1)T$ where $m = 1, \ldots, M$. The variation of the waveform from frame to frame is determined by a probability mechanism described by the joint distribution of $\{x(\beta_m), \ldots, x(\beta_m + T - 1)\}$: $k = 1, \ldots, K$. This joint distribution is the same for each frame by the periodic block stationarity condition.

The signals in the sampling frames do not have to be generated by a periodic mechanism. They can be impulse responses of a linear or nonlinear system excited by a nonperiodic sequence of impulse source functions. For example, suppose that the data in question is a set of measurements of the vibration measured by an accelerometer mounted on a high-speed drill as it repeatedly drills holes in a computer circuit board in a precisely replicating experiment (for details, see [2]). The accelerometer waveform for each drilling event can be concatenated to make a sample of $M$ successive frames of $T$ observations in the time unit. The waveforms have a lot of random variation across frames. The time series so generated can then be analyzed to determine a relationship between the wear of the drill and the vibration signatures.

Now that the concept of a randomly modulated periodicity has been defined, the next step is to develop a measure of the amount of random variation present in each Fourier component of a signal. Such a measure, called a signal coherence function, is presented in Section II.

**II. SIGNAL COHERENCE**

The $m$th frame is $\{x(\beta_m), \ldots, x(\beta_m + T - 1)\}$. Its discrete Fourier transform (DFT) at frequency $f_r = r/T$ for each $r = 1, \ldots, T/2$ is

$$X(r) = \sum_{t=0}^{T-1} x((m-1)T + t) \exp(-i2\pi ft)$$

$$= \sum_{k=0}^{T/2} u_k(t) \exp(i2\pi kT)$$

$$= \mu_r + U(r)$$

where $U(r) = \sum_{t=0}^{T-1} u(t) \exp(-i2\pi ft)$.

The joint distribution of $\{x(\beta_m), \ldots, x(\beta_m + T - 1)\}$ is the same for each frame. These observations are random variables before the data is observed and thus, in keeping with standard statistical notation, the index $m$ is not used to subscript the complex valued random variables $X(r)$ and $U(r)$.

The variability of the complex Fourier amplitude $X(r)$ about its mean $\mu_r$ depends on the complex variance of $U(r)$ and the covariances of $U(r_1)$ and $U(r_2)$, $E[\bar{U}(r_1)U(r_2)] = \sigma_u(r_1, r_2)$, where

$$\sigma_u(r_1, r_2) = \sum_{t_1=0}^{T-1} \sum_{t_2=0}^{T-1} c_u(t_1, t_2) \exp(i2\pi (r_1 t_1 - r_2 t_2)/T)$$

and

$$c_u(t_1, t_2) = E[u(t_1)\bar{u}(t_2)] = E[u(t_1)u^*(t_2)]$$

$$= c_u^*(t_2, t_1).$$

Fig. 1. An example of a randomly modulated signal.
Thus, setting $r_1 = r_2 = r$ and $\tau = |t_1 - t_2|$ in (4) and using the usual notation $\sigma_u^2(r)$ for the variance $\sigma_u(r, r)$, it follows that $r$

$$\sigma_u^2(r) = \sum_{\tau=0}^{T-1} \exp(2\pi i r \tau / T) \sum_{t=0}^{T-1} c_u(t, t + \tau),$$

which is of the order $O(T)$. For example, if $u(t)$ is weakly stationary, then $c_u(t, t + \tau) = c_u(\tau)$ and, thus,

$$\sigma_u^2(r) = T \sum_{\tau=0}^{T-1} (1 - r / T) \gamma_u(\tau) \exp(2\pi i r \tau / T),$$

which equals $T[S_u(f_r) + O(1/T)]$ where $S_u(f)$ is the spectrum of $u(t)$.

If $\mu_r \neq 0$ and $\sigma_u(r) = 0$, then that complex amplitude is a true periodicity. The larger the value of $\sigma_u(r)$, the greater the variability of that component from frame to frame. If $\mu_r = 0$ and $\sigma_u(r) \neq 0$, then that component is random with no periodic structure.

In order to quantify the variability, consider the function $\gamma_u(r)$, called a signal coherence function, defined as follows for each $r = 1, \cdots, T/2$:

$$\gamma_u(r) = \frac{\sqrt{|\mu_r|^2 + \sigma_u^2(r)}}{|\mu_r|^2 + \sigma_u^2(r)},$$

If $\sigma_u(r) = 0$ and $\mu_r \neq 0$, then $\gamma_u(r) = 1$. This is the case where the $f_r$ frequency component has a constant amplitude and phase. If $\mu_r = 0$, then $\gamma_u(r) = 0$. This is the case where the mean value of the $f_r$ frequency component is zero, which is true for each frequency component of any stationary random process which satisfies any of the standard mixing conditions of which finite dependence is the most severe [3, p. 95].

This signal coherence function is very different from the coherence function between two stationary signals [4, p. 352]. The coherence (sometimes called coherency) between $x(t)$ and $x(t)$ at frequency $f_r$ is the correlation between $X_1(t)$ and $X_2(t)$. The closer the coherence value is to one, the higher the correlation between the real and imaginary parts of both Fourier components. This well-known coherence concept has been used to estimate the time delay of a plane wave detected by an array of sensors [5].

The signal coherence function, in contrast, is defined for one signal. It measures the variability of $X(r)$ about its mean $\mu_r$. One should keep in mind that the signal in this representation is the mean of the observed signal.

In the signal plus noise representation of $\{x(t)\}$, the signal-to-noise ratio (SNR) is $\rho(r) = |\mu_r|^2 / \sigma_u^2(r)$ for frequency $f_r$. Thus, $\rho_u^2(r) = \rho(r) / \rho^2(r) + 1$ is a monotonically increasing function of this SNR. Inverting this relationship, it follows that

$$\rho(r) = \frac{\gamma_u^2(r)}{1 - \gamma_u^2(r)}.$$ 

A signal coherence value of 0.44 yields an SNR of 0.24 which is $\approx 6.2$ dB. This result will be referred to in Section V.

The estimation of the signal characteristics is addressed in the next section.

### III. ESTIMATING THE SIGNAL COHERENCE FUNCTION

Suppose that we have observed $M$ frames each of length $T$ of the process $\{x(t)\}$ as denoted in the beginning of Section II. Recall that $\beta_m = (m - 1)T$ for each $m = 1, \cdots, M$. The sample mean for each $t = 0, \cdots, T - 1$

$$\bar{x}(t) = \frac{1}{M} \sum_{m=1}^{M} x(\beta_m + t)$$

is an unbiased estimator of the “signal” $s(t)$. From the central limit theorem for a mean of finite dependent random variables [6, Theorem 27.5], $\sqrt{M} [\bar{x}(t) - s(t)]$ is asymptotically normal $N(0, \sigma_s^2(t))$ as $M \to \infty$ where $\sigma_s^2(t)$ is the sum of the variance of $x(t)$ and twice the covariances of $x(t)$ and $x(t + dT)$ for $d = 1, \cdots, D$. Using more common statistical parlance, this means that the distribution of $\bar{x}(t)$ is approximately normal with mean $s(t)$ and variance $M^{-1/2} \sigma_s^2(t)$ for large $M$.

To simplify exposition, assume from now on that $M$ is much larger than $D$ and thus $M^{-1} \sigma_s^2(t)$ is negligible.

Let $\bar{X}(r)$ denote the $r$th component of the DFT of $\{\bar{x}(0), \cdots, \bar{x}(T - 1)\}$. It then follows from the above asymptotic result that $\sqrt{M} [\bar{X}(r) - \mu_r]$ is asymptotically complex normal with mean zero and finite variance, which implies that $\bar{X}(r)$ is a consistent estimator of $\mu_r$ with an error of probability order $O_P(M^{-1/2})$ which is negligible for large $M$.

Recall that $x(\beta_m + t) = s(\beta_m + t) + u(\beta_m + t)$. Define

$$y(\beta_m + t) = x(\beta_m + t) - \bar{x}(t)$$

and let $Y_m(r)$ denote the $r$th DFT component of $\{y(\beta_m), \cdots, y(\beta_m + T - 1)\}$. It follows that $Y_m(r)$ is the $r$th DFT component of $\{u(\beta_m), \cdots, u(\beta_m + T - 1)\}$ plus an error of order $O_P(M^{-1/2})$. Consider the following estimator of the variance $\sigma_u^2(r)$:

$$\hat{\sigma}_u^2(r) = M^{-1} \sum_{m=1}^{M} |Y_m(r)|^2.$$ 

Since the $u$'s are $D$ dependent, the $Y$'s are as well. Once again using Theorem 27.5, the distribution of $\hat{\sigma}_u^2(r)$ is approximately normally distributed with mean $\sigma_u^2(r)$ and variance of order $O(M^{-1})$ which implies that $\hat{\sigma}_u^2(r)$ is a consistent estimator of $\sigma_u^2(r)$ with an error of order $O_P(M^{-1/2})$. It then follows that $|\bar{X}(r)|^2 / \hat{\sigma}_u^2(r)$ is a consistent estimator of the SNR $\rho(r)$ for frequency $f_r$. Moreover, the statistic $\gamma_u(r)$ defined by

$$\hat{\gamma}_u(r) = \sqrt{\frac{|\bar{X}(r)|^2}{|\bar{X}(r)|^2 + \hat{\sigma}_u^2(r)}}$$

is a consistent estimator of $\gamma_u(r)$ for frequency $f_r$ with an error of $O_P(M^{-1/2})$.

It is vital that the frame length used in the analysis be an integer multiple of the fundamental period of the signal. If not,
then the measured signal coherence will be biased toward zero. For example, suppose that the signal is \( \exp\{2\pi(t/T + \theta)\} \) and the frame length used is \( T + 1 \). The value of the sinusoid at the relative time \( t \) in the \((m+1)\) frame is \( \exp\{2\pi(m(T+1)+t)/T + \theta\} \) and its arithmetic average of \( M \) frames will go to zero as \( M \) increases. This phase drift \( 2\pi m/T \) generated by the mismatch of the sampling frame period with the period of the fundamental causes \( \bar{X}(r) \) to be shrunk toward zero. If the sample size is very large, as is the case in the example presented in Section IV, then it is possible to use a frame length that is many times the fundamental period, and thus the phase precession is small over the frames.

The estimator \( \hat{\gamma}_x(r) \) is a point estimate of the signal coherence \( \gamma_x(r) \) that measures the relative stability of the \( f_r \) frequency component of the observed signal. It should not be confused with detecting a cyclic component. The model in (1) is assumed to hold. The "noise" term \( u(t) \) is not an additive stationary noise term as typically employed in signal processing applications, but rather is a residual process capturing the random variation in \( x(t) \) from the mean periodic signal \( s(t) \). The residual signal is not stationary, it is periodic block stationary.

If the residual signal is weakly stationary, then it has been shown in Section II that for each \( r \), \( \sigma^2_u(r) = [T_\mu_s(f_r) + O(1/T)] \) and \( \sigma_u = 0 \). Thus, \( \sqrt{M}X(r) \) is asymptotically complex normal with mean zero and variance \( TS_u(f_r) \) as \( M \to \infty \) ignoring the \( O(1/T) \) term. However, in addition, \( \sqrt{M}X(r) \) and \( \sqrt{M}X(r') \) are asymptotically independently distributed as \( M \) goes to infinity [3]. It then follows that the distribution of \( MT^{-1}\bar{X}(r)^2TS_u^{-1}(f_r) \) is approximately \( \chi^2_2 \) (chi-squared distribution with two degrees-of-freedom). These \( \chi^2_2 \) variates are approximately independently distributed over the frequency band when the residual signal is weakly stationary with finite dependence.

IV. Simulation Results

To illustrate these statistical results, the following model was used to generate artificial data results for \( f_k = 0.01 \) and 100 frames of length \( T = 100 \)

\[
x(t) = [1 + \alpha u_1(t)] \cos(2\pi f_k t) + [1 + \alpha u_2(t)] \sin(2\pi f_k t)
\]

(12)

where \( \{u_1(t)\} \) and \( \{u_2(t)\} \) are autoregressive AR(1) processes generated by the first-order difference equation used to generate Fig. 1. Table I presents the estimates \( \hat{\gamma}_x(1) \) of \( \gamma_x(1) \) for six levels of the noise amplitude scale \( \alpha \). Note that \( \hat{\gamma}_x(1) = 0.71 \) for \( \alpha = 1 \), which is the case when the noise standard deviation equals the cosine’s amplitude. Only 100 frames were used for this table in order to match the number of frames used in the real data analysis presented in Section V. The result of a run with 5,000 frames and the same parameter settings produced almost the same signal coherence values.

V. A Real Data Example

The signal coherence function and the associated power spectrum were estimated using a signal from a hydrophone at a depth of 25 m in a shallow bay in the Baltic Sea as part of an experiment conducted by the author and Dr. L. Persson and Dr. J. Robinson at the Swedish National Defense Establishment (FOA) research facility near Stockholm. A fast workboat powered by a Volvo Penta four-cylinder diesel engine with two counterrotating propellers was driven toward a bottom-mounted hydrophone at a fixed speed of 12 km and constant heading starting at a distance of 420 m from the hydrophone. The data record used in this paper lasted 102.4 s. The acoustic signal from the boat plus ambient water noise were recorded and digitized using FOA equipment on shore. An anti-aliasing filter was applied to the signal prior to digitizing and the signal was sampled at a rate of 12.5 kHz.

The digitized data were high-pass filtered by taking the DFT of the entire \( N = 128 \cdot 10^6 \) observations in the file, zeroing out the bins associated with the band 0–10 Hz and then using the DFT to transform back into the time domain. This filtered record was then segmented into 102 nonoverlapping frames of \( T = 12500 \) observations yielding a resolution bandwidth of 1 Hz for the given sampling rate.

A plot of the spectrum for the band 10–95 Hz is shown in Fig. 2 and a plot of the signal coherence function is shown in Fig. 3.

The main peaks in the spectrum are at 12 Hz (4.12 dB) and its first two harmonics are 24 Hz (11.47 dB) and 48 Hz (7.29 dB), but the largest peak is at 24 Hz. The signal coherence function has peaks around those frequencies, but the largest value for \( \hat{\gamma}_x(r) \) is 0.44 at 50 Hz. From the result presented at the end of Section II, this coherence value has an SNR value of \(-6.2 \) dB.

The value at 24 Hz is 0.384, which is the second largest value. Although the throttle was held at a fixed level during the run, there was some variation in propeller rotation speed as the boat moved through the water. The magnitudes of the signal coherence values at the frequencies where the spectrum peaks provides a measure of the variation in these frequencies over time. The spectral peaks confirm what we know, which is that the signal has a mean periodic structure due to the motor and fuel pump, which are the major noise sources in this experiment. The size of the signal coherence peaks are surprising since they show that the "periodic" components of the spectrum are highly modulated even though the motor speed was fixed. The modulation is not due to the ambient noise since the sound level from other ships was low during the time the runs were made and

<table>
<thead>
<tr>
<th>Noise Scale ( \alpha )</th>
<th>Signal coherence ( \gamma_x(1) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>1.000</td>
</tr>
<tr>
<td>0.5</td>
<td>0.995</td>
</tr>
<tr>
<td>1.0</td>
<td>0.895</td>
</tr>
<tr>
<td>2.0</td>
<td>0.710</td>
</tr>
<tr>
<td>4.0</td>
<td>0.447</td>
</tr>
<tr>
<td>0.243</td>
<td></td>
</tr>
</tbody>
</table>

TABLE I

Signal Coherence for Randomly Modulated Sinusoids
the broad-band water noise spectrum was small relative to the acoustic motor signal.

The program that I wrote to compute the spectrum and the signal coherence computes more statistics than have been presented so far. In order to determine the correlation distance, correlations are computed for each frequency in the band over adjacent frames. The correlations for the data used here were computed for 15 adjacent frames. The correlation between two adjacent frames for the peak at 24 Hz is 0.87. The correlation across frames linearly decreases to become statistically insignificant for a span of six frames. The other bands had insignificant correlation values. These results provide an estimate of the across-frame dependence of less than 1 ms.

ACKNOWLEDGMENT

The author would like to thank Dr. L. Persson and Dr. J. Robinson for their encouragement to explain this new idea. I thank Dr. G. Edelson and the referees for their constructive criticisms and suggestions for improving this work.
REFERENCES


Melvin J. Hinich received the B.S. and M.S. degrees in mathematics from Carnegie-Mellon University, Pittsburgh, PA, in 1959 and 1960, respectively, and the Ph.D. degree in statistics from Stanford University, Stanford, CA.

He joined the Faculty of the Graduate School of Industrial Administration of Carnegie-Mellon in 1963 and developed a research agenda in time-series analysis and applied statistics in business and economics. Based on experience gained from summer research projects on sonar signal processing on which he worked at the General Electric Research Laboratories and the Bell Telephone Laboratories (Whippany, NJ), he began to apply statistical theory to array processing, target localization and tracking, and other military science applications, with special emphasis on acoustic warfare. During the summer of 1979, he was an Expert Consultant at the Naval Coastal Systems Center, Panama City, FL. He became an Associate Professor of Industrial Administration and Statistics in 1968 and was advanced to Professor in 1970. In 1973, he joined the Department of Economics at Virginia Polytechnic University as Professor of Economics. He joined the University of Texas (UT) at Austin in August 1982. In 1984, he became affiliated with the Applied Research Laboratories of UT as a Research Faculty Member. He is the Mike Hogg Professor of Government and Professor of Economics at the University of Texas at Austin. He is the author or co-author of six books and numerous articles in journals such as the *Journal of the Acoustical Society of America*, the *IEEE Transactions on Information Theory*, the *SIAM Journal of Applied Mathematics, Information, and Control*, the *IEEE Transactions on Acoustics, Speech, and Signal Processing*, and the *IEEE Journal of Oceanic Engineering* (invited papers).

Dr. Hinich is a Fellow of the Institute of Mathematical Studies and the Public Choice Society.