Journal of Statistical Physics, Vol. 5, No. 3, 1972

Statistically Periodic Processes and Toeplitz$_{(\alpha\ \mathrm{by}\ \beta)}$ Matrices

William H. Pierce[1]

Received July 28, 1970; revised June 11, 1971

If each element in a Toeplitz matrix is replaced by an α by β matrix and the original constraints preserved, the result is a doubly infinite matrix with periodic structure called a Toeplitz$_{(\alpha\ \mathrm{by}\ \beta)}$ matrix. Such matrices are a basic tool for describing, generating, estimating, filtering, synchronizing, and analyzing information-theoretic functions for statistically periodic processes.

KEY WORDS: Statistical periodicity; periodic processes; periodic linear transformations; random processes; estimation theory; Toeplitz matrices; infinite matrices; matrix *Z*-transforms; matrix-generating functions; communications theory; information theory; periodic difference equations; pulse amplitude modulation; synchronization; orthogonal functions.

1. INTRODUCTION

Statistically periodic processes can be defined analogously to statistically stationary processes. If the nth-order cumulative distribution function

$$P(x(t_1) \leqslant x_1, x(t_2) \leqslant x_2, \cdots, x(t_n) \leqslant x_n)$$

is invariant for all sets of x_i and t_i when each t_i is replaced by $t_i + \tau$, then the process will be strict-sense statistically periodic with period τ. If it is invariant only for $n = 1$ and 2, the process will be wide-sense statistically periodic.

Statistically periodic processes are a generalization of statistically stationary processes, and arise in physical science due to planetary and crystal-lattice periodicities, in satellite engineering due to periodicies of

[1] Department of Electrical Engineering, University of Louisville, Louisville, Kentucky.

orbits, tumbles, etc., and in communications engineering because of the convenience and efficiency of periodic data transmissions. Statistically periodic processes may also be useful models for processes in biology (respiration, heartbeat, diurnal activity) and commerce (weekly and annual cycles). Infinite matrices of periodic structure will be introduced and used to handle the linear algebra of periodic processes.

2. SUMMARY

A doubly infinite matrix A will be defined to be $\text{Toeplitz}_{(1\,\text{by}\,1)}$ if there exists a one-subscript sequence ..., a_{-1}, a_0, a_1,..., such that $A_{i,j} = a_{j-i}$ for all i, j. Any column in a $\text{Toeplitz}_{(1\,\text{by}\,1)}$ matrix is the column to the left shifted down by one step; any row is the row above shifted right by one step.

A doubly infinite matrix A will be defined to be $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$, α and β positive integers, if a partition can be made of vertical lines every β columns and horizontal lines every α rows, such that the partitioned matrix is $\text{Toeplitz}_{(1\,\text{by}\,1)}$ in the α by β submatrices created by the partition. Any α consecutive rows specify a $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrix, since any other row, such as the nth, equals the (n mod α) row shifted to the right by entier (n/α). Similarly, any β consecutive columns also specify the whole matrix.

Vectors of evenly-spaced samples of statistically stationary processes have $\text{Toeplitz}_{(1\,\text{by}\,1)}$ covariance matrices.[1] Similarly, use of a periodic autocorrelation function will show that vectors of samples every (period/α) of wide-sense statistically periodic processes have $\text{Toeplitz}_{(\alpha\,\text{by}\,\alpha)}$ covariance matrices.

The scalar transforms previously associated with $\text{Toeplitz}_{(1\,\text{by}\,1)}$ matrices[1,2] are generalized to associated matrix transforms. Convergence of products of $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrices is approached several ways, including the use of products of the associated transform matrices, as has been done with $\text{Toeplitz}_{(1\,\text{by}\,1)}$ matrices,[1,2] and the basis for this is shown to be a type of Parseval's theorem problem. Widom's square-summable theory[2] is generalized to $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrices, giving conditions for products directly upon the associated matrix transforms which would be used to evaluate the product when convergent. Multiplication of a vector by a $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrix is shown to produce sums of rotations of the vector's associated θ-transform. Inversion of $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrices is shown to depend upon boundedness of the determinant of the associated matrix θ-transform over the unit circle.

Use of $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrices permits various examples of statistically periodic processes to be formulated and analyzed. Sets of periodically shifted basis functions are introduced which generalize the $\{\sin[\pi(t - nT)/T]\}/[\pi(t - nT)/T]$ basis functions. Inclusion properties of the space spanned by such sets are shown to be calculable from properties of matrix Z-transforms. Delay matrix studies show how samples of a continuous band-limited process at one set of sampling times are related by a Toeplitz$_{(1\,\mathrm{by}\,1)}$ matrix to samples at a shifted set of times.

Four domains are really relevant to Toeplitz$_{(\alpha\,\mathrm{by}\,\beta)}$ matrices: the matrix domain, the associated transform domain, the time domain of the matrix convolution summation, and the difference algorithm or shift register domain. Various aspects of statistically periodic processes are discussed in the appropriate domain, including least-mean-square estimation of signals with additive noise, hardware and software implementation of Toeplitz$_{(\alpha\,\mathrm{by}\,\beta)}$ matrix operations, and information-theoretic and synchronization problems.

3. PREVIOUS RELATED WORK

In addition to Toeplitz$_{(1\,\mathrm{by}\,1)}$ matrix material from Grenander and Szego[1] and Widom,[2] other related work is z-transform theory and its association with shift registers,[3] estimation theory,[4] Wiener filtering,[5] infinite (or quarter-infinite) matrix theory,[6] and stability analyses[7] which are the continuous analog of some of the matrix results here. Previous formulations and analyses of statistically periodic processes appear confined to pulse amplitude modulation.[8–11] These use nonmatrix approaches to develop spectral equations related to those of Section 6 and give estimation circuits equivalent to the Wiener filter of Section 11. The author knows of no previous use of Toeplitz$_{(\alpha\,\mathrm{by}\,\beta)}$ matrices, except his own.[12] While this paper was being revised, Ogura[13] developed a different but related approach to describing statistically periodic processes, in which $E[x(t_1)\,x(t_2)]$ is given a double Fourier series expansion whose coefficients involve finite Fourier integrals of the elements of a Toeplitz$_{(1\,\mathrm{by}\,1)}$ matrix.

4. STATISTICALLY PERIODIC PROCESSES

The Discrete Gaussian Periodic Process. Let ω be a random variable which is a member of the space Ω of all such variables, let $\mathscr{F}$ be a σ-field of subsets of Ω, and let P be a probability measure on $\mathscr{F}$. Assume that for each ω, there exists the random sequence of real numbers ..., $x_{\omega,-1}$, $x_{\omega,0}$, $x_{\omega,1}$,... which form the components of the infinite column vector x_ω, and assume that E is the expectation operator over all ω in Ω. The sequence ..., $x_{\omega,-1}$, $x_{\omega,0}$, $x_{\omega,1}$,... will be called a *discrete Gaussian periodic process* with integer period β if $E[x_\omega] = w$, where $w_i = w_{i+\beta}$ for all i, if $E[(x - w)(x - w)^T] = M$, where M is Toeplitz$_{(\beta\,\mathrm{by}\,\beta)}$, and if all finite subsets of x_ω have a multidimensional Gaussian distribution.

The Band-Limited Gaussian Periodic Process. Let ω be a random variable as before, and let $y_\omega(t)$ be a continuous scalar function of t which has a Fourier transform which is zero at radian magnitudes of $\pi/\tilde{T}$ and higher. The $y_\omega(t)$ will be called a *band-limited Gaussian periodic process* of period T if the sequence of samples[2] ..., $y_\omega(-T/\beta)$, $y_\omega(0)$, $y_\omega(T/\beta)$,... constitute a discrete Gaussian periodic process of period β, where β is the least integer greater than or equal to $T/\tilde{T}$.

The Band-Limited M-PAM Process. A generalization of pulse amplitude modulation (PAM) is a signal of the form

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{l=0}^{M-1} (\mathbf{r}_n)_l h_l(t - nT) \tag{1}$$

where $(\mathbf{r}_n)_l$ is the lth component of the M-dimensional column vector $\mathbf{r}_n$. When $M = 1$, this reduces to ordinary PAM.[8–11] The generalization will be called M-PAM. Let r be an infinite column vector constructed from the m-dimensional vectors of Eq. (1) such that its transpose is $\mathbf{r}^T = [..., \mathbf{r}_{-1}^T, \mathbf{r}_0^T, \mathbf{r}_1^T,...]$. For the $x(t)$ of Eq. (1) to be strict-sense statistically periodic, $\mathbf{r}$ must have strict-sense statistically periodic components. It will be assumed that either (1) the components of $\mathbf{r}$ are a discrete Gaussian periodic process, with period M, or (2) that $\mathbf{r}_i$ and $\mathbf{r}_j$ are statistically independent for $\mathbf{i} \neq \mathbf{j}$ and $\mathbf{r}_i$ has a known M-dimensional cumulative distribution function which is independent of $\mathbf{i}$. In either case, $x(t)$ is statistically periodic with period T.

Equivalences. For any zero-mean band-limited Gaussian process $x(t)$ with period T, there exists an M-PAM process with the same statistics. If M is large enough so that samples every T/M specify $x(t)$ by the sampling theorem, then the samples of Eq. (1) contain enough information to specify the time function. A straightforward analysis will show that the sample values $h_l(nT/M)$ of Eq. (1) will be the l^{th} column of a Toeplitz$_{(M\,\mathrm{by}\,M)}$ matrix A such that A^TA equals the covariance matrix of the samples of $x(t)$, and the components of $\mathbf{r}_n$ are white Gaussian noise. (The Z-transform associated with A, to be introduced in the next section, can be found by using the Z-transform generalization of theorem 1 of Brockett and Mesarovic.[14] If the mean of the samples of $x(t)$ lies in the space spanned by the columns of A, then giving the components of $\mathbf{r}_n$ a mean obtained by using the generalized inverse[15] of A will also produce an M-PAM process with the same statistics as $x(t)$.

[2] The sampling theorem shows that $y_\omega(t)$ can be reconstructed from these samples. For an early derivation of the sampling theorem, see Whitaker[22]. For a modern engineering viewpoint, see Papoulis[23] or Shannon.[24]

T-Shift Ergodic Processes. Let $x_\omega(t)$ be a random variable defined on $-\infty \leqslant t \leqslant \infty$, where ω denotes the actual process occurring from an ensemble Ω of possible ω's. A *T-shift-invariant set* $\mathscr{A}$ is a set of all ω such that if ω is in $\mathscr{A}$ and if $x_\lambda(t) = x_\omega(t - T)$, then λ is in $\mathscr{A}$. A random process $x_\omega(t)$ is defined to be *T-shift-ergodic* if $x_\omega(t)$ is strict-sense statistically periodic with period T, and if each T-shift invariant set has probability measure either zero or one. If $x_\omega(t)$ is a T-shift ergodic random variable, and if

$$I(n, y, \tau, t, \omega) = \begin{cases} 1 & \text{if} \quad x_\omega(t) \leqslant y_0\,,\ x_\omega(t+\tau_1) \leqslant y_1\,,\dots,\ x_\omega(t+\tau_n) \leqslant y_n \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$g(\omega) = \lim_{R\to\infty} (1/R) \left[\sum_{n=0}^{R-1} I(n, y, \tau, t + nT, \omega) \right] \tag{3}$$

then the ergodic theorem[16] shows that $g(\omega)$ converges with probability one to the nth-order cumulative distribution function (CDF) of $x(t)$. In words, the time average of Eq. (3) converges to the ensemble average of the CDF with probability one.

Classes of Statistically Periodic Processes. In the physical world, time is a locally defined variable, so that most physical processes of a periodic nature will usually be related to an idealized periodic process, such as a T-shift ergodic random process, by a change of time variable. The simplest of such physical processes is the *randomly delayed* T-*shift ergodic* process, whose random variable $y_\delta(t)$ equals $x_\omega(t - \phi)$, where $x_\omega(t)$ is T-shift ergodic and ϕ is a random variable independent of ω. The synchronization problem for such processes is to deduce the remainder of (ϕ/T) given samples of $x_\omega(t - \phi)$. The second simplest class is the *random-rate T-shift* ergodic process $y_\delta(t)$, defined by $y_\delta(t) = x_\omega(\alpha t - \phi)$, where α, ϕ, and ω are independent.

5. *Z*-TRANSFORMS TO BE ASSOCIATED WITH TOEPLITZ$_{(\alpha \text{ by } \beta)}$ MATRICES

When A is Toeplitz$_{(\alpha \text{ by } \beta)}$, with α or β or both $\geqslant 1$, the Z-transform associated with A will be

$$A(z) = \sum_{n=-\infty}^{\infty} \begin{bmatrix} A_{0,-n\beta} & \cdots & A_{0,-(n\beta+\beta-1)} \\ \vdots & & \vdots \\ A_{\alpha-1,-n\beta} & \cdots & A_{\alpha-1,(-n\beta+\beta-1)} \end{bmatrix} z^{-n} \tag{4}$$

When b is an infinite column vector to be left-multiplied by a Toeplitz$_{(\alpha \text{ by } \beta)}$

matrix, $\beta \geqslant 1$, the Z-transform associated with b will be the vector Z-transform

$$b(z) = \sum_{n=-\infty}^{\infty} \begin{bmatrix} b_{0+n\beta} \\ b_{1+n\beta} \\ \vdots \\ b_{\beta-1+n\beta} \end{bmatrix} z^{-n} \tag{5}$$

Definition. "b is regarded as Toepcol$_{(\beta)}$" means that the Z-transform to be associated with b is that given by (5) for the appropriate value of β.

By substituting $e^{i\theta}$ for z, what will be called the θ-transform results. It bears the same relationship to the Z-transform that the Fourier transform bears to the two-sided Laplace transform. By use of distribution theory or "Fourier transforms-in-the-limit," as defined by Bracewell,[17] the θ-transform of periodic sequences will involve Dirac delta functions.

The following are easily verified:

If A is Toeplitz$_{(\alpha \text{ by } \beta)}$ and has associated Z-transform $A(z)$, then the transpose of A, denoted A^T, is Toeplitz$_{(\beta \text{ by } \alpha)}$ and has associated Z-transform $A^T(1/z)$. (6)

If A is Toeplitz$_{(\alpha \text{ by } \beta)}$, it is also Toeplitz$_{(n\alpha \text{ by } n\beta)}$ for any positive integer n. (7)

6. PRODUCTS OF TOEPLITZ$_{(\alpha \text{ by } \beta)}$ MATRICES

A doubly infinite matrix A is said to be multipliable by infinite matrix B or infinite column vector b if

$$\sum_{l=-\infty}^{\infty} A_{il}B_{lj}\,, \qquad \sum_{l=-\infty}^{\infty} A_{il}b_l$$

converge, respectively, for all i, j or all i.

One way to study the convergence problem is to require convergence of each component of AB or Ab. Assuming A is Toeplitz$_{(\alpha \text{ by } \beta)}$ and B is Toeplitz$_{(\beta \text{ by } \gamma)}$, sufficient conditions for such component convergence are:

1. That the sum of the magnitude squared of the components of β adjacent columns of A and γ adjacent columns of B (or the components of b) be finite. (Suggested by R. N. Peterson and based upon Schwartz's inequality.)
2. That A, B, or b be stable in the sense of having a finite sum of magnitudes of components of each column, and the other has uniformly bounded magnitudes of components. (Straightforward proof.)

3. That conditions ensure the applicability for the Z-transform version of Parseval's theorem, which is

$$(1/2\pi_i) \oint_{\substack{\text{unit}\\\text{circle}}} A(z)\, b(z)\, z^{-k-1} dz = \sum_{i=-\infty}^{\infty} a_i b_{k-i} \tag{8}$$

where z is a complex scalar and the other quantities are either scalars or finite matrices of compatible dimensions. There are many conditions for which the Fourier transform version of Parseval's theorem holds.[18] When Parseval's theorem does hold, it follows that $C = AB$ can be evaluated by use of the fact that the associated transforms obey $C(z) = A(z)B(z)$.

An alternative approach to the general problem of convergence of linear transformations is the square-summable theory of Widom,[2] which assumes multiplication of transforms without proof. In this approach, only linear transformations which map squared-summable vectors into squared-summable vectors are considered convergent. Widom shows that a necessary and sufficient condition for such convergence of A is that $A(e^{i\theta})$ be essentially bounded[3] on $-\pi \leqslant \theta \leqslant \pi$. The appropriate generalization, having a similar proof, is that a Toeplitz$_{(\alpha \text{ by } \beta)}$ matrix A is convergent if and only if

$$g(\theta) = \sup_{\text{all } x \neq 0} \mid x^{T*} A^{T*}(e^{i\theta})\, A(e^{i\theta}) x / x^{T*} x \mid$$

is essentially bounded on $-\pi \leqslant \theta \leqslant \pi$.

7. SPECTRAL PROPERTIES OF TOEPLITZ$_{(\alpha \text{ by } \beta)}$ TRANSFORMATIONS

Suppose M is Toeplitz$_{(\alpha \text{ by } \beta)}$ and p is an infinite column vector with scalar associated transform $p(e^{i\theta})$. If $q = Mp$ with $q(e^{i\theta})$ scalar, then it is desired to find how $q(e^{i\theta})$ is related to $p(e^{i\theta})$, not by the somewhat clumsy definition $q = Mp$, but rather directly in terms of the spectrum of p. Such is the objective of this section.

The first step needed is the following. Let p be an infinite column vector, and let $p(e^{i\theta})$ be its associated θ-transform when p is regarded as Toepcol$_{(1)}$, and let $\hat{p}(e^{i\theta})$ be its α-dimensional associated θ-transform when p is regarded as Toepcol$_{(\alpha)}$. Then,

$$\hat{p}_l(e^{i\theta}) = (1/\alpha) \sum_{t=0}^{\alpha-1} p(e^{i(\theta-2\pi)/\alpha})\, e^{il(\theta-2\pi t)/\alpha} \tag{9}$$

[3] $M(x)$ is essentially unbounded on $a \leqslant x \leqslant b$ if the sets {all x such that $\mid M(x) \mid \geqslant n$} each have positive measure for $n = 1, 2, 3, \ldots$

To prove this, consider the identity

$$\sum_{q=0}^{\alpha-1} (e^{iq2\pi(k-l)/\alpha}/\alpha) = \delta_{0,k-l \bmod \alpha} \tag{10}$$

By adding terms of value zero, the sum for $\hat{p}_l$ may be given by

$$\hat{p}_l(e^{i\theta}) = \sum_{m=-\infty}^{\infty} p_m e^{i\theta[l+\mathrm{Entire}(m/\alpha)]}\delta_{l,m \bmod \alpha}$$

Because of identity (10), this becomes

$$\hat{p}_l(e^{i\theta}) = \sum_{m=-\infty}^{\infty} \sum_{q=0}^{\alpha-1} e^{iq2\pi(m-l)/\alpha} p_m e^{-i\theta m/\alpha} e^{i\theta l/\alpha}$$

Interchanging the order of summation produces Eq. (9), since the m summation gives the shifted θ-transform of p.

Equation (9) gives $\hat{p}(e^{i\theta})$ in terms of $p(e^{i\theta})$. The inverse relation is also of interest. Using the identity

$$\sum_{n=-\infty}^{\infty} x_n = \sum_{i=0}^{m-1} \sum_{q=-\infty}^{\infty} x_{qm+i} \tag{11}$$

and the basic definitions it is easy to show that

$$p(e^{j\theta}) = \sum_{r=0}^{\alpha-1} e^{-i\theta r}\hat{p}_r(e^{i\theta\alpha}) \tag{12}$$

By using (10), it is easily verified that substitution of Eq. (9) into (12) gives the original $p(e^{j\theta})$.

Equations (9) and (12) can now be used to study the spectral effects of the operation $x = My$, where M is $\mathrm{Toeplitz}_{(\alpha \mathrm{by} \beta)}$, which was the objective formulated earlier. Letting the dot (·), caret (^), and tilde (~) denote the associated transform when the vector is regarded as $\mathrm{Toepcol}_{(1)}$, $\mathrm{Toepcol}_{(\alpha)}$, or $\mathrm{Toepcol}_{(\beta)}$, respectively, the transform equation becomes

$$\hat{x}(e^{i\theta}) = M(e^{i\theta})y(e^{i\theta})$$

Applying Eqs. (9) and (12) and the rules for matrix multiplication and changing the order of summation gives

$$\dot{x}(e^{i\theta}) = \sum_{k=0}^{\alpha-1} \dot{y}(e^{i(\theta-k2\pi/\alpha)}) f(\theta, k) \tag{13}$$

where

$$f(\theta, k) = \sum_{r=0}^{\alpha-1} \sum_{l=0}^{\beta-1} (1/\alpha)\, e^{i(\theta r - \theta l - 2\pi k/\alpha)} M_{rl}(e^{i\theta\alpha}) \tag{13a}$$

This shows that multiplication of a column vector by a Toeplitz$_{(\alpha\,\text{by}\,\beta)}$ matrix is like multiplying the vector's components by a discrete periodic sequence, which is hetrodyning. The affect of both is a sum of rotations of the θ-transform.

8. INVERTABILITY AND LINEAR INDEPENDENCE OF COLUMNS

Toeplitz showed that the eigenvalues of a Toeplitz$_{(1\,\text{by}\,1)}$ matrix coincided with the set of values on $|z| = 1$ of the Z-transform associated with the matrix.[1] The associated eigenfunctions must therefore have associated θ-transforms of $\delta(\theta - \theta_0)$, where $-\pi \leqslant \theta_0 \leqslant \pi$. Widom[2] shows that the inverse of a Toeplitz$_{(1\,\text{by}\,1)}$ matrix A exists (in the sense of being an operator which takes square-summable vectors into square-summable vectors) if and only if $1/A(e^{i\theta})$ is essentially bounded on $-\pi \leqslant \theta \leqslant \pi$. If A is Toeplitz$_{(\alpha\,\text{by}\,\beta)}$ with $\alpha \neq \beta$, then A will be singular in the sense that a nonzero x vector exists such that $Ax = 0$ or $x^T A = 0$. [This is easily proved by using the rectangularness of $A(z)$ to find a finite vector y such that $A(\exp i\theta')y$ or $y^T A(\exp i\theta')$ is zero, and letting $x(\exp i\theta) = \delta(\theta - \theta')y$.] The generalization to a Toeplitz$_{(\alpha\,\text{by}\,\alpha)}$ matrix A is that A is an invertible square-summable operator if the components of $A(e^{i\theta})$ and $1/\det A(e^{i\theta})$ are essentially bounded on $-\pi \leqslant \theta \leqslant \pi$, as follows by application of Cramer's rule for the inverse. The columns [or rows] of A will be linearly independent in the square-summable sense if the components of $A(e^{i\theta})$ and $1/\det A^T(e^{i\theta})A(e^{i\theta})$ [or $1/\det A(e^{i\theta})A^T(e^{i\theta})$] are essentially bounded on $-\pi \leqslant \theta \leqslant \pi$. If A is Toeplitz$_{(\alpha\,\text{by}\,\alpha)}$ and invertible with inverse B, then $B(e^{i\theta}) = [A(e^{i\theta})]^{-1}$. If A has linearly independent columns, then its generalized inverse B is $B = [A^T(e^{-i\theta})A(e^{i\theta})]^{-1}A^T(e^{-i\theta})$.

9. PERIODICALLY SHIFTED BASIS FUNCTIONS

Let $x(t)$ be a square-integrable scalar function of $t \in (-\infty, \infty)$. Let $x(t)$ be band-limited to ω_0, which means its Fourier transform is zero for radian frequency magnitudes greater than ω_0. If $x(t)$ is a random process, the probability measure associated with it could be represented by the probability measure on a discrete sequence $\{a_n\}$, $n = ..., -1, 0, 1, ...$, provided a linearly independent basis set of signals $\phi_n(t)$, $n = ..., -1, 0, 1, ...$ exists such that

$$x(t) = \sum_{n=-\infty}^{\infty} a_n \phi_n(t) \tag{14}$$

The sampling theorem shows, for example, that one suitable set of $\phi_n(t)$ functions is

$$\phi_n(t) = \{\sin[\pi(t - nT)/T]\}/\{\pi(t - nT)/T\} \tag{15}$$

provided $T \leqslant \pi/\omega_0$.

In applications, a periodic property of the $\phi_n(t)$ functions will be especially practical since any related hardware or data processing algorithms will be periodic. However, the flexibility to have the $\phi_n(t)$ functions have properties different from (15) is sometimes desirable. For example, a communications channel may have pure exponentials for eigenfunctions, so that narrow-bandwidth $\phi_n(t)$ functions approximating the pure exponentials are useful in obtaining high efficiency.

Basis signals $\{\phi_n(t)\}$ with the practical features of periodicity and the necessary generality can be constructed as follows from a set $\mathscr{H} = \{h_0(t), h_1(t),..., h_{M-1}(t)\}$ of M scalar functions defined on $-\infty \leqslant t \leqslant \infty$, with $M \geqslant 1$. Specifically, if

$$\phi_n(t) = h_{(n \bmod M)}(t - (n - (n \bmod M))T) \tag{16}$$

it will be possible to rewrite the expansion of (14) as

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{i=0}^{M-1} (\mathbf{r}_n)_i h_i(t - nT) \tag{17}$$

provided $\mathbf{r}_k{}^T = [a_{kM}, a_{kM+1}, ..., a_{kM+M-1}]$. The set of all possible $x(t)$ described by these equations can be called "the space spanned by the set $\mathscr{H}$ and all its T-shifts." The set of basis functions described by Eq. (17) can be called a set of periodicallyshifted basis functions. The M-PAM process was identical to (17), so this formulation of basis functions is especially useful in dealing with statistically periodic processes.

The special case $M = 1$, $h_0(t)$ equal to Eq. (15) with $n = 0$, is known by the sampling theorem to have all its T-shifts span the space of functions band-limited to ω_0. The space spanned by any other set $\mathscr{K}$ and its shifts can be compared to this space of all functions band-limited to ω_0 by the following.

Theorem 9.1. Let S_H be the space spanned by $\mathscr{H} = \{h_0(t),..., h_{M-1}(t)\}$ and all its (pT)-shifts, and S_K be the space spanned by $\mathscr{K} = \{k_0(t),..., k_{N-1}(t)\}$ and all its (qT)-shifts, where p, q, M, and N are all positive integers. Let $\mathscr{H}$ and $\mathscr{K}$ contain functions whose Fourier transform is bounded at $\omega = \pm\omega_0$, and zero for $|\omega| > \omega_0$ for some finite ω_0. Then, Eq. (18) or its associated Z-transform is a necessary and sufficient condition for $S_H \subset S_K$. Furthermore, when $S_H \subset S_K$, the expansion coefficients in the S_K basis are given by Eq. (19), where these equations arise in the proof.

Notation.

$$\mathbf{h}^T(t) = [h_0(t), h_1(t),..., h_{M-1}(t)]$$

$$\mathbf{k}^T(t) = [h_0(t), k_1(t),..., k_{M-1}(t)]$$

$$\bar{\mathbf{w}}^T(t) = \left[..., \operatorname{sinc}\left(\frac{tsT}{\alpha} + 1\right), \operatorname{sinc}\left(\frac{tsT}{\alpha}\right), \operatorname{sinc}\left(\frac{tsT - 1}{\alpha}\right),...\right]$$

where sinc $x = (\sin \pi x)/\pi x]$

$$\bar{\mathbf{h}}^T(t) = [..., \mathbf{h}^T(t + p^T), \mathbf{h}^T(t), \mathbf{h}^T(t - p^T),...]$$

$$\bar{\mathbf{k}}^T(t) = [..., \mathbf{k}^T(t + qT), \mathbf{k}^T(t), \mathbf{k}^T(t - qT),...].$$

Proof. Let s be the least common multiple of p and q. Let α be the least integer such that $2\pi\alpha/sT \geqslant \omega_0$. The function $\operatorname{sinc}(t\alpha/sT)$ and all its (sT/α)-shifts serves as a basis set for comparing its subspaces S_H and S_K. Therefore, by the sampling theorem, there exists a matrix A such that when $s = p$, then $\mathbf{h}(t) = A\bar{\mathbf{w}}(t)$, or when $s > p$, then

$$\begin{bmatrix} \mathbf{h}(t) \\ \mathbf{h}(t - pT) \\ \vdots \\ \mathbf{h}(t - (s - p)T \end{bmatrix} = A\bar{\mathbf{w}}(t)$$

Similarly, there exists a matrix B such that when $s = q$, then $\mathbf{k}(t) = B\bar{\mathbf{w}}(t)$, or when $s > q$, then

$$\begin{bmatrix} \mathbf{k}(t) \\ \mathbf{k}(t - qT) \\ \vdots \\ \mathbf{k}(t - (s - q)T \end{bmatrix} = B\bar{\mathbf{w}}(t)$$

Note that A has (Ms/p) rows, B has (Ns/q) rows, and all of these rows are infinite. Let C be the Toeplitz$_{(Ms/p \text{ by } s\alpha)}$ matrix which has A for its rows number 0 through $(Ms/p) - 1$; let D be the Toeplitz$_{(Ns/q \text{ by } s\alpha)}$ matrix which has B for rows number 0 through $(Ns/q) - 1$. Then, $\bar{\mathbf{h}}(t) = C\bar{\mathbf{w}}(t)$ and $\bar{\mathbf{k}}(t) = D\bar{\mathbf{w}}(t)$. Thus if $x(t) = \mathbf{x}^T\bar{\mathbf{h}}(r)$, and $y(t) = \mathbf{y}^T\bar{\mathbf{k}}(t)$, then $x(t) = x^TC\bar{\mathbf{w}}(t)$, and $y(t) = \mathbf{y}^TD\bar{\mathbf{w}}(t)$. Now, $S_H \subset S_K$ if and only if, given any $\mathbf{x}^T$ as defined above, there exists a $\mathbf{y}^T$ such that $x(t) = \mathbf{y}^T\bar{\mathbf{k}}(t) = \mathbf{y}^TD\bar{\mathbf{w}}(t)$. Because the components of $\bar{\mathbf{w}}(t)$ are linearly independent, this requires that a $\mathbf{y}$ exist such that $\mathbf{x}^TC = \mathbf{y}^TD$ has a solution for $\mathbf{y}$, given arbitrary $\mathbf{x}$. Denoting generalized inverse of a matrix with a superscript + and invoking its properties[7,15] shows that the necessary and sufficient conditions for such a $\mathbf{y}$ to exist is

$$C^+CD^+D = C^+C \tag{18}$$

Equation (18) is therefore necessary and sufficient condition for $S_H \subset S_K$. When it holds, the general $\mathbf{y}$ solution, by the properties of the generalized inverse, is

$$\mathbf{y} = (D^T)^+C^T\mathbf{x} + [I - (D^T)^+D^T]\phi \tag{19}$$

where ϕ is arbitrary.

Corollary 9.1. A necessary and sufficient condition for (a) the coefficients in a basis set $\mathscr{H}$ and its T-shifts to be unique or (b) a set $\mathscr{H}$ and its T-shifts to be linearly independent functions is $I - (C^T)^+C^T = I - CC^+ = 0$ or the equivalent equation of associated Z-transforms. (The proof uses Theorem 9.1 with $p = 1$, $\mathscr{K} = \mathscr{H}$, and is based upon requiring the matrix premultiplying ϕ in (19) to be zero. Note that the condition upon C is that its rows must be linearly independent.)

Corollary 9.2. A necessary and sufficient condition to make $\mathscr{H}$ and all its T-shifts an orthonormal basis set is $CC^T = \alpha I/sT$, or the equivalent equation of associated Z-transforms. (Orthogonality of band-limited time functions is equivalent to orthogonality of their discrete values sampled so as to satisfy the sampling theorem. The CC^T matrix is the matrix of inner products of the discrete sample values.)

If $k_i(z)$ $l_i(z)$, $i = 0,\ldots, M-1$, are all-pass Z-transforms,[19] then one way to generate an orthonormal set $\mathscr{H} = \{h_0(t),\ldots, h_{M-1}(t)\}$ would be to have

$$\bar{\mathbf{h}}(t) = C\bar{\mathbf{w}}(t)$$

as in the proof of Theorem 9.1, and then let

$$\begin{aligned} c(z) = {} & P_1^T \operatorname{diag}[k_0(z),\ldots, k_{M-1}(z)]P_1 \\ & \times P_2^T \operatorname{diag}[l_0(z),\ldots, l_{M-1}(z)]P_2 \end{aligned}$$

where the P_i are orthornormal M by M matrices of scalars.

10. THE DELAY MATRIX

Suppose that $x(t)$ is a band-limited process whose Fourier transform is zero at radian frequency magnitudes of π/T' and greater, and that x and y are infinite vectors such that $x_i = x(iT')$ and $y_i = x(iT' - \lambda)$. The sampling theorem can be used to show that

$$y = D_\lambda x \tag{20}$$

where

$$(D_\lambda)_{ij} = \{\sin \pi[(\lambda/T') - i + j]\}/\pi[(\lambda/T') - i + j] \tag{21}$$

Note that D_λ is Toeplitz$_{(1\,\mathrm{by}\,1)}$. D_λ will be called the λ-second delay matrix.

The Z-transform for $D_\lambda(z)$ is divergent for $|z| \neq 1$. Inversion of the θ-transform verifies that

$$D_\lambda(\exp i\theta) = \exp(-i\lambda\theta/T') \tag{22}$$

Equation (22), combined with Eq. (6), shows that

$$D_\lambda = D^T_{-\lambda} \tag{23}$$

The following conclusions about the D_λ matrix can be drawn by combining the above results with definitions and earlier results.

D_λ is not stable except when $\lambda = mT'$, integer m. (24)

D_λ has elements uniformly bounded by 1. (25)

D_λ is multipliable by any stable matrix; any stable matrix is multipliable by D_λ. (26)

D_λ is multipliable by Toeplitz$_{(1\,\mathrm{by}\,1)}$ matrix with bounded elements whose θ-transform is absolutely integrable; any such matrix is multipliable by D_λ. (From conditions for Parseval's theorem.) (27)

D_λ is nonsingular for all λ. (28)

$D_\lambda D_\xi = D_{(\lambda+\xi)}$ (29)

When D_λ is regarded as Toeplitz$_{(\alpha\,\mathrm{by}\,\alpha)}$, with an α by α associated θ-transform $D(e^{i\theta})$, then for $0 \leqslant (p, q) \leqslant \alpha - 1$,

$$[D_\lambda(e^{i\theta})]_{p,q} = \sum_{n=-\infty}^{\infty} \frac{\sin \pi[(\lambda/T') - p - n\alpha + q]}{\pi[(\lambda/T') - p - n\alpha + q]}\, e^{-i\theta n} \tag{30}$$

The same pattern of ideas which proved Eqs. (9) and (22) can be used to handle the above. The result, after summing a geometric series, is that

$$[D_\lambda(e^{i\theta})]_{p,q} = \begin{cases} \dfrac{e^{i[(\lambda/T')-p+q](\theta/\alpha)}[1 - e^{-i2\pi[(\lambda/T')-p+q]}]}{1 - e^{-i2\pi[(\lambda/T')-p+q]/\alpha}} & \text{when denom} \neq 0 \\ e^{i[(\lambda/T')-p+q](\theta/\alpha)} & \text{otherwise.} \end{cases} \tag{31}$$

The delay matrix plays an important role in analyzing communications and other physical processes, since samples of the delayed process will equal the delay matrix operating upon samples of the idealized (undelayed) process.

11. APPLICATIONS AND CONCLUSIONS

The basic tool of $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrices is likely to be useful whenever processes with periodicity or statistical periodicity arise. Some examples now follow.

Mean Square Evaluation. Suppose a zero-mean random vector x has covariance matrix $M = E[xx^T]$, where M is $\text{Toeplitz}_{(\alpha\,\text{by}\,\alpha)}$, with associated Z-transform $M(z)$. The mean square of the components of x can be found as a limiting case of the trace operation, or else the Z-transform power spectral density version of Parseval's theorem could be used. Letting Z^{-1} denote the inverse Z-transform operator, the desired result becomes

$$\text{mean square of the components of } x = (1/\alpha) \sum_{i=0}^{\alpha-1} Z^{-1} \,|[M_{ii}(z)]|_0 \qquad (32)$$

This formula is useful when x is a statistically periodic process whose $M(z)$ can be evaluated.

Sampled-Data Difference Equations with Periodic Coefficients. Consider the period-α sampled-data difference equation given by

$$x_n = \sum_{i=1}^{R_1} a_{n \bmod \alpha, i} x_{n-i} + \sum_{i=0}^{R_2} b_{n \bmod \alpha, i} r_{n-i}, \qquad R_1, R_2 \leqslant M \qquad (33)$$

If $a_{n \bmod \alpha, i}$ is taken to be 0 for $i < 1$ or $i > R_1$, and if $b_{n \bmod \alpha, i}$ is taken to be 0 for $i < 0$ or $i > R_2$, then this can be rewritten (with a change of variable) as

$$x_n = \sum_{m=-\infty}^{\infty} a_{n \bmod \alpha, n-m} x_m + \sum_{m=-\infty}^{\infty} b_{n \bmod \alpha, n-m} r_m \qquad (34)$$

Now, if $A_{i,j} = a_{i \bmod \alpha, i-j}$ and $B_{i,j} = b_{i \bmod \alpha, i-j}$, it is readily verified that A and B are $\text{Toeplitz}_{(\alpha\,\text{by}\,\alpha)}$, so that Eq. (34) becomes the doubly infinite matrix equation $x = Ax + Br$. The forced solution is $x = (I - A)^{-1}Br$ and its associated Z-transform is $x(z) = \{[I - A(z)]^{-1}B(z)\}r(z)$, where x and r are now regarded as Toepcol_α. Thus the forced $\mathbf{x}_n$ is the (matrix) convolution of the inverse Z-transform of $\{[I - A(z)]^{-1}B(z)\}$ and the vector sequence $\mathbf{r}_n$, where $\mathbf{r}_n^T = [r_{n\alpha}, r_{n\alpha+\alpha}, \ldots, r_{n\alpha+\alpha-1}]$.

Shift Register Implementation. Suppose a β-input, α-output, linear time-invariant shift register receives at instant i the β-dimensional input vector f_i, and outputs the α-dimensional vector c. The following convolution summation will relate output with input:

$$c_k = \sum_{i=-\infty}^{\infty} h_{i-k} f_k \tag{35}$$

where h_i is an α by β matrix and the causality condition $h_i = 0$ for $i < 0$ applies in all physical systems. It will also be true that when f is regarded as Toepcol$_{(\beta)}$ and c as Toepcol$_{(\alpha)}$, Eq. (35) leads to the relation $c(z) = H(z)f(z)$, where H is Toeplitz$_{(\alpha \text{ by } \beta)}$. The β-input, α-output shift register can be constructed from $\alpha\beta$ single-input, single-output time-invariant shift registers whose outputs are summed. As a consequence, multiplication of a vector by a Toeplitz$_{(\alpha \text{ by } \beta)}$ matrix can be accomplished by a matrix of scalar shift registers. If the input appears sequentially on the same line, a buffer must be inserted so that inputs are processed β-at-a-time, and if the output is to be sequential, an output buffer must take the α-at-a-time output and release it evenly one-at-a-time.

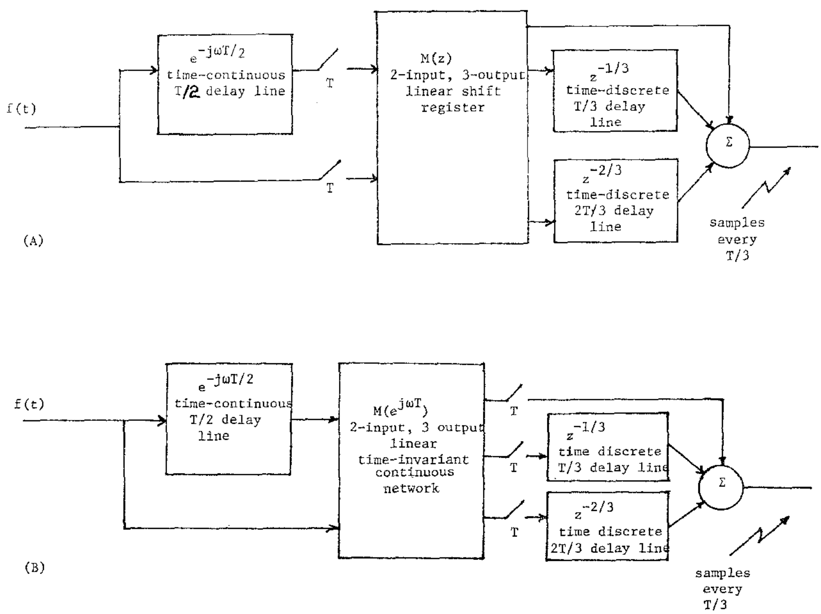

Fig. 1. Block diagrams for real-time computation of $D_T M f$, when M is Toeplitz$_{(3 \text{ by } 2)}$, f is the infinite vector whose ith component is $f(iT/2)$, and $z = e^{i\omega T}$. The switch marked T denotes the usual sampled-data operation of taking a sample every T sec.

Continuous Filter with Sampler Implementation. When the discrete f vector originally comes from a continuous signal, then known sampled-data techniques permit certain options in implementation. Fig. 1(A) illustrates one arrangement, which directly implements the theory of the preceding paragraph. On the other hand, the same output is obtained in Fig. 1(B) by employing continuous filters preceding the samplers. If $f(t)$ is band-limited to ω_0, these filters can have arbitrary response at $|\omega| > \omega_0$, which, of course, simplifies their implementation.

Data-Processing Algorithm. The periodic-coefficient difference equations and shift registers can be connected by finding a matrix C such that $C(I - A)^{-1} = (I - A)^{-1}B$. Assuming such a C is lower triangular, Fig. 2 generates the difference equation (34), in the special case of $\text{Toeplitz}_{(2\,\text{by}\,2)}$ A and B. Feedback shift registers can therefore be constructed to perform $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrix operations. Digital computer data processing algorithms can of course be developed with the same feedback structure as the shift register.

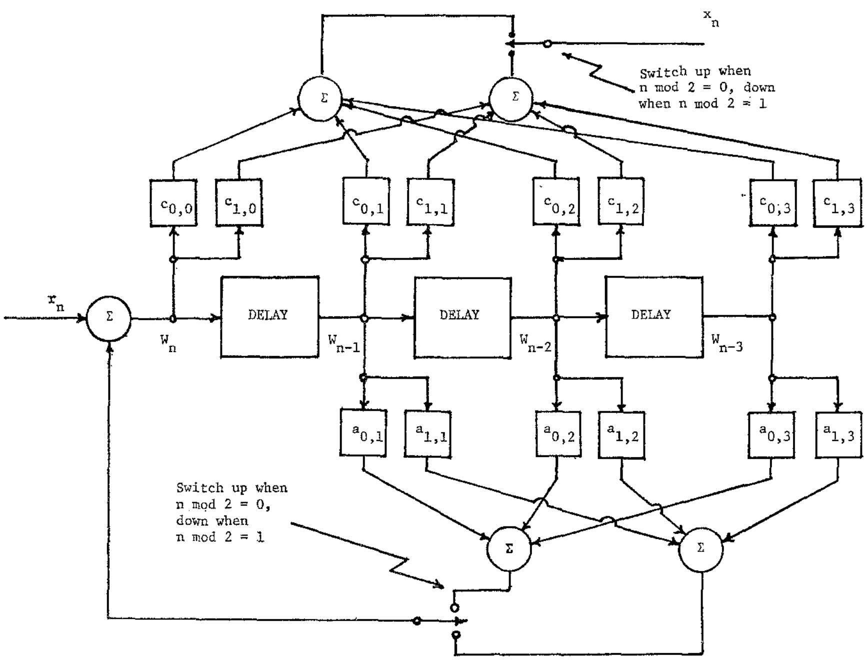

Fig. 2. A shift-register implementation related to Eq. (33), with $R_1 = R_2 = 3$ and $\alpha = 2$.

MATRIX DOMAIN

Assume H is Toeplitz$_{(\alpha \text{ by } \beta)}$

$x = Hr$

Uses:
1. Use for simple notation, and simple conceptual formulation of operations.
2. Derive desired operators in this domain by linear algebra techniques.
3. Derive optimal operators to separate signal plus noise in this domain by variational methods.
4. Formulate probabilistic and statistical problems in this domain.

TIME DOMAIN

$$x_n = \sum_{k=-\infty}^{\infty} H_{n-k}\, r_k$$

Uses:
1. Formulate problems most easily in this domain so that they can be converted to other domains.
2. Compute numerical solutions to sampled-data difference equations with periodic coefficients.

TRANSFORM DOMAIN

If $H(z)$ is α by β

$r(z)$ is a β-dimensional column vector

$x(z)$ is an α-dimensional column

then $x(z) = H(z)r(z)$.

Uses:
1. Perform calculations such as r.m.s. error, information rates, etc.
2. Derive necessary or sufficient conditions for various theorems and use frequency insights to prove theorems.
3. Treat systems described by finite difference equations by use of poles and zeros in transform domain, and design by placement of poles and zeros.

DIFFERENCE ALGORITHM OR SHIFT REGISTER DOMAIN

Uses:
1. Construct hardware implementation for signal processing.
2. Tie in the other domains to the generating algorithm of a difference equation.
3. Create algorithm for digital computer processing and system simulation.

Fig. 3. The four domains of Toeplitz$_{(\alpha \text{ by } \beta)}$ matrices.

Wiener Filtering, Stationary and Periodic. If r is a data vector, f is a vector of measurements, c is measurement noise, $f = Ar + c$, r and c have zero mean and are statistically independent, $E[rr^T] = Q$, $E[cc^T] = N$, and $(AQA^{T'} + N)$ is nonsingular,[4] then a straightforward variational argument[20] will show that the least-mean-square linear estimator of r given f is $T_{\text{mean}} f = QA^T(AQA^T + N)^{-1}f$. When Q and N are nonsingular, it is easily shown that T_{mean} is then equal to $(Q^{-1} + A^TN^{-1}A)^{-1}A^TN^{-1}$ and that the resulting error vector $(r - T_{\text{mean}}f)$ has covariance matrix $(Q^{-1} + A^TN^{-1}A)^{-1}$. If r and c also have a multidimensional Gaussian distribution, then $T_{\text{mean}}f$ is the conditional mean of r given f, which justifies the name.

If the r, c, and f vectors have a periodic structure such that A is

[4] A sufficient condition for nonsingularity is that N be nonsingular, which occurs in most physical situations.

$\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$, Q is $\text{Toeplitz}_{(\beta\,\text{by}\,\beta)}$, N is $\text{Toeplitz}_{(\alpha\,\text{by}\,\alpha)}$, and $(AQA^T + N)$ is nonsingular, then $T_{\text{mean}}(z)$ becomes

$$Q(z)A^T(1/z)[A(z)Q(z)A^T(1/z) + N(z)]^{-1} \tag{36}$$

This is implementable by the shift registers of Fig. 1. When $\alpha = \beta = 1$, this is the well-known (noncausal) Wiener filter for a stationary sampled-data process. When α, β, or both are greater than one, the result is the generalization of the stationary Wiener filter to the least-mean-square (noncausal) filter for a sampled-data statistically periodic process. Although Eq. (36) is believed to be new as a matrix equation, its derivation was easy because the minimization was carried out in the matrix domain and then converted to the Z-transform domain. Equation (36) can be used to give the least-mean-square estimate of an M-PAM signal in noise. If the decoded noise is required to be independent of the signal, then the $Q \to \infty$ limit should be taken, giving an M-PAM decoder of

$$[A^T(1/z)N^{-1}(z)A(z)]^{-1}A^T(1/z)N^{-1}(z) \tag{37}$$

These decoders are the M-PAM generalizations of Smith's[21] 1-PAM decoders. The use of $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrices made the M-PAM generalization much simpler than the 1-PAM method of Smith, which would be extremely hard to generalize. This illustrates the real power of $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrices. The optimization here was carried out in the matrix domain, where it was easy, rather than in the transform domain, where it is difficult.

Information and Synchronization. It is possible to prove that the 1-PAM receivers of Smith[21] or M-PAM receivers based upon Eq. (36) or (37) are information lossless when the channel noise is independent of the signal and Gaussian.[12] This is done by showing that the operation of these receivers is equivalent to first whitening the noise with a $\text{Toeplitz}_{(1\,\text{by}\,1)}$ linear transformation, followed by a $\text{Toeplitz}_{(1\,\text{by}\,\alpha)}$ or $\text{Toeplitz}_{(M\,\text{by}\,\alpha)}$ transformation which is information lossless because there is no signal in the null space of this linear transformation and because the whiteness of the noise prevents using knowledge of the noise in the space orthogonal to the signal space to extrapolate the noise into the signal space. In studying synchronization,[12] $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ transformations are required to make certain orthogonal projections whenever the signal has fewer degrees of freedom than its bandwidth warrants.

Conclusions. The four domains in which a $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrix can be viewed are illustrated in Fig. 3, with the important uses for each domain. As can be seen from the material in this figure, $\text{Toeplitz}_{(\alpha\,\text{by}\,\beta)}$ matrices are useful in many aspects of statistically periodic processes.

REFERENCES

1. U. Grenander and S. Szego, *Toeplitz Forms and Their Applications* (Univ. of Calif. Press, 1958).
2. H. Widom, "Toeplitz Matrices," in *Studies in Real and Complex Analysis*, Ed. I. I. Hirschmann, Jr. (The Mathematical Association of America and Prentice Hall, 1965).
3. J. R. Ragazzini and G. F. Franklin, *Sampled-Data Control Systems* (McGraw-Hill, New York, 1958).
4. R. Deutsch, *Estimation Theory* (Prentice Hall, Englewood Cliffs, N. J., 1965).
5. G. C. Newton, L. A. Gould, and J. F. Kaiser, *Analytical Design of Linear Feedback Controls* (Wiley, New York, 1957).
6. R. G. Cooke, *Infinite Matrices and Sequence Spaces* (MacMillan and Co., London, 1950).
7. L. A. Zadheh and C. A. Desoer, *Linear System Theory* (McGraw-Hill, New York, 1963).
8. W. R. Bennett, "Statistics of Regenerative Digital Transmission," *B.S.T.J.* **1958** (Nov.): 1501–1542.
9. J. W. Smith, "The Joint Optimization of Transmitted Signal and Receiving Filter for Data Transmission System," *B.S.T.J.* **1965** (Dec.):2363–2392.
10. T. Berger and D. W. Tufts, "Optimum Pulse Amplitude Modulation, Part I," *IEEE Trans. on Information Theory* **IT-13**:196–208 (1967).
11. D. W. Tufts and T. Berger, "Optimum Pulse Amplitude Modulation, Part II," *IEEE Trans. on Information Theory* **IT-13**:209–216 (1967).
12. W. H. Pierce, "Information and Synchronization of PAM," presented orally at the 1967. Int. Symp. on Information Theory, San Remo, Italy, Sept. 11–15, 1967. An abstract is in *IEEE Trans. on Information Theory* **IT-14**:843 (1968).
13. H. Ogura, "Spectral Representation of a Periodic Nonstationary Random Process," *IEEE Trans. on Information Theory* **IT-17**:143–149 (1971).
14. R. W. Brockett and M. D. Mesarovic, "Synthesis of Linear Multivariable Systems," *Trans. AIEE, Part II* **81**:216–21 (1962).
15. R. Penrose, "A Generalized Inverse for Matrices," *Proc. Cambridge Phil. Soc.*, **51** (Part 3):406–413 (1955).
16. P. Billingsley, *Ergodic Theory and Information* (Wiley, New York, 1965).
17. R. Bracewell, *The Fourier Transform and Its Applications* (McGraw-Hill, New York, 1965), p. 11.
18. E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Oxford U. P., New York, 1937).
19. W. H. Pierce, "Linear Real Codes and Coders," *B.S.T.J.*, **47**:1078 (1968).
20. S. Berkowitz, Ph.D. Thesis, Carnegie Institute of Technology, 1966.
21. J. W. Smith, "The Joint Optimization of Transmitted Signal and Receiving Filter for Data Transmission Systems," *B.S.T.J.* **1965** (Dec.):2363–2392.
22. E. T. Whitaker, *Proc. Roy. Soc. Edinburgh* **35**:181–191 (1915).
23. A. Papoulis, *The Fourier Integral and Its Applications* (McGraw-Hill, New York, 1962).
24. C. E. Shannon, "Communication in the Presence of Noise," *Proc. IRE*, January 1949.