INFORMATION AND CONTROL 4, 97-117 (1961)

Sampling Theorems and Bases in a Hilbert Space*

FREDERICK J. BEUTLER

Department of Aeronautical and Astronautical Engineering, The University of Michigan, Ann Arbor

A unified approach to sampling theorems for (wide sense) stationary random processes rests upon Hilbert space concepts. New results in sampling theory are obtained along the following lines: recovery of the process $x(t)$ from nonperiodic samples, or when any finite number of samples are deleted; conditions for obtaining $x(t)$ when only the past is sampled; a criterion for restoring $x(t)$ from a finite number of consecutive samples; and a minimum mean square error estimate of $x(t)$ based on any (possibly nonperiodic) set of samples.

In each case, the proofs apply not only to the recovery of $x(t)$, but are extended to show that (almost) arbitrary linear operations on $x(t)$ can be reproduced by linear combinations of the samples. Further generality is attained by use of the spectral distribution function $F(\cdot)$ of $x(t)$, without assuming $F(\cdot)$ absolutely continuous.

I. INTRODUCTION

The importance of the Shannon (1949) sampling theorem is well illustrated by its emphasis in texts on information theory (Goldman, 1954) and communication engineering (Nichols and Rauch, 1956). Only recently, however, has this theorem been proved rigorously for (wide sense) stationary random processes (Balakrishnan, 1957).

Throughout this paper, $x(t)$ represents a random process which may be thought of as a message or signal. In communication systems, the numerical values of $x(t)$ occurring at some discrete set of times τ_n are often transmitted in place of the continuous parameter $x(t)$; this technique enhances noise immunity and/or permits time sharing of a single channel. Other devices, such as digital computers, have a discrete time base, and are capable only of employing a set of numbers $x(\tau_n)$, the numerical values of $x(t)$ at the times τ_n.

* The research reported here was supported by NASA Research Grant NsG-2-59, and The University of Michigan Institute of Science and Technology.

It is customary to call the discrete denumerable set $\{\tau_n\}$ the sampling times, and $x(\tau_n)$ the samples or sample values. One aim of sampling theory is to discover the conditions on $\{\tau_n\}$ which will insure that $x(t)$ can be recovered from its samples. A crude statement of the basic theorem (Shannon, 1949) is as follows: if $x(t)$ contains only frequencies below W radians per second, $x(t)$ may be recovered from periodic samples $x(\tau_n)$ taken π/W seconds apart. The precise meaning of this statement, as well as some generalizations, will become clear in the sequel.

One proof of the sampling theorem for (wide sense) stationary random processes rests on Hilbert space concepts, integration theory, and the properties of trigonometric series. The proof itself suggests new results along the following lines: recovery of $x(t)$ from nonperiodic samples, or when any finite set of samples is deleted; conditions for obtaining $x(t)$ when only the past is sampled; a criterion for restoring $x(t)$ from a finite number of consecutive samples[1]; and a minimum mean square error estimate of $x(t)$ based on any (possibly nonperiodic) set of samples.

In each case, the proofs apply not only to the recovery of $x(t)$, but are extended to show that (almost) arbitrary linear operations on $x(t)$ can be reproduced by linear combinations of the samples. Further generality is attained through use of the spectral distribution function $F(\cdot)$ of $x(t)$, without the usual assumption of absolute continuity of $F(\cdot)$.[2]

The emphasis throughout is on a unified approach to sampling theory through a common set of techniques. This viewpoint provides a powerful yet rigorous method of treating sampling theory.

II. MATHEMATICAL PRELIMINARIES

The definitions, concepts, and notation to be used throughout this paper are given below. Known results are stated as assertions, detailed proofs being readily available (see Doob, 1953, particularly Chapter XI, Sections 3 and 4).

In dealing with the (wide sense) stationary $x(t)$ we shall utilize only

[1] Some authors like to assume that $x(t)$ is finite dimensional, i.e., recoverable from $2WT$ samples (W is bandwidth, $T < \infty$). For stationary $x(t)$, this is shown here to be valid only when $x(t)$ belongs to a limited (and nearly trivial) class of processes.

[2] For example, $F(\cdot)$ could be continuous and strictly increasing, with zero derivative almost everywhere (a.e.). Then $x(t)$ has positive power, even though its spectral density is zero a.e., and there is no line spectrum.

the second-order properties (e.g., autocorrelation) of the process. Therefore, the spectral distribution function $F(\cdot)$ gives as complete a description of the process as is needed. The $F(\cdot)$ in question may be normalized without loss of generality. Accordingly, it is assumed henceforth that $F(-\infty) = 0$, $F(+\infty) = 1$, and that $F(\cdot)$ is right continuous. The term "band-limited" is applied to $x(t)$ when, in addition, $F(-\pi - 0) = 0$ and $F(\pi + 0) = 1$. That is, $x(t)$ has total average power unity; this power is concentrated in a $\frac{1}{2}$-cycle bandwidth if $x(t)$ is band-limited.

$F(\cdot)$ is said to have a point of increase (or jump) at ω if $F(\omega) - F(\omega - 0) > 0$. Then $x(t)$ has a random component $\alpha e^{i\omega t}$ where α may be a random variable. In engineering language, the point of increase corresponds to a spectral line at ω, or a δ-function component in the spectral density at that frequency.

The spectrum of $x(t)$ is the set of points S satisfying

$$S = \{\omega \mid F(\omega + \epsilon) - F(\omega - \epsilon) > 0 \text{ for each } \epsilon > 0\} \qquad (2.1)$$

If ω is an isolated point of S, ω must be a point of increase. Consequently, when S consists only of a finite number of points each point of S is a point of increase.

Consider now the expectation $E[x(t)x^*(s)]$, where $*$ denotes a complex conjugate. This expectation may be regarded as an inner product $(x(t), x(s))_1$. Because this inner product is actually the correlation $R(t - s)$, it is equal to another inner product defined by

$$(e^{i\omega t}, e^{i\omega s})_2 = \int_{-\infty}^{\infty} e^{i\omega(t-s)}\, dF(\omega) \qquad (2.2)$$

Thus, inner products involving $x(t)$ and probability measure are equal to inner products of corresponding exponentials $e^{i\omega t}$ taken with respect to measure dF. This property extends to finite sums: if y and z are random variables defined by $y = \sum a_n x(t_n)$ and $z = \sum b_m x(t_m)$, there are corresponding ω functions $h_y(\omega) = \sum a_n e^{i\omega t_n}$ and $h_z(\omega) = \sum b_m e^{i\omega t_m}$ with the property

$$(y, z)_1 = (h_y, h_z)_2 = \int_{-\infty}^{\infty} h_y(\omega) h_z^*(\omega)\, dF(\omega) \qquad (2.3)$$

A norm (denoted by $\|\ \|$) is defined on the inner product spaces by taking $\| y \|_1^2 = (y, y)_1$ and $\| h_y \|_2^2 = (h_y, h_y)_2$. Corresponding elements of the two spaces (of random variables and ω functions, respectively) will have equal norms. One even has relations such as $\| y - z \|_1 = \| h_y - h_z \|_2$, so that the "distance" between corresponding elements is preserved. Because it is clear which inner product or norm appears in a given expression, the subscripts $_1$ and $_2$ will no longer be used.

Finally, the inner product space of exponentials (measure dF) is completed to a Hilbert space H_2. There is a similar completion H_1 of the inner product space of random variables. The equality (2.3) extends to all corresponding elements of H_1 and H_2. Evidently, H_1 and H_2 are isomorphic; the relation between them is suggested by the initial correspondence of $x(t) \in H_1$ with $e^{i\omega t} \in H_2$.

Since every $h(\cdot)$ which satisfies $\int_{-\infty}^{\infty} | h(\omega) |^2 dF(\omega) < \infty$ can be approximated in H_2 norm by finite sums $\sum a_n e^{i\omega t_n}$, all such $h \varepsilon H_2$. Corresponding to each $h \varepsilon H_2$ is a random variable in H_1. Suppose now that $h_\alpha(\cdot) \in H_2$ for some indexing set α. The span of $\{h_\alpha\}$, denoted by $V\{h_\alpha\}$, is then defined as the least subspace containing all the h_α, i.e., that subspace consisting of all finite linear combinations of the h_α, together with their completion in H_2. By definition, $V\{e^{i\omega t}\} = H_2$. But also $V\{e^{i\omega r}$, all rational $r\} = H_2$, since $\lim_{r \to t} \| e^{i\omega r} - e^{i\omega t} \| = 0$ for any t; equivalently, $V\{x(r)$, all rational $r\} = H_1$.

A base as here defined is any set of elements of H_2 (or H_1) whose span is H_2 (or H_1). Since we have just exhibited a denumerable base, a denumerable basis (maximal orthonormal set) exists. Because every basis has the same cardinality, this cardinality can be taken to define the dimension of H_2 (and likewise H_1). If a space is of finite dimension $N - 1$, any N (and hence any $N + m$) elements will be linearly dependent. That is, given any N elements $h_n \in H_2$, $\| \sum c_n h_n(\omega) \| = 0$ for some set of c_n of which at least one is not zero. Exhibiting a set of N linearly independent elements of H_2 is thus equivalent to proving that H_2 is at least of dimension N.

Let $\{h_n(\cdot)\}$, $h_n \in H_2$, be such that for each t there exists a set of coefficients c_n with the property that

$$\| e^{i\omega t} - \sum c_n h_n(\omega) \| \to 0 \tag{2.4}$$

Then $V\{h_n\} = V\{e^{i\omega t}\} = H_2$. It follows from this that functions $c_n(t)$ may be chosen to give

$$\| h(\omega, t) - \sum c_n(t) h_n(\omega) \| \to 0 \tag{2.5}$$

for all t, providing only that $h(\omega, t) \in H_2$ for each t. Because $h(\omega, t) =$

l.i.m. $\sum a_n(t)e^{i\omega t_n}$, there is a random process $y(t)$ corresponding to $h(\omega, t)$. This process is specified for each t by $y(t) = \text{l.i.m.} \sum a_n(t)x(t_n)$.[3]

To apply the preceding discussion to sampling theory, it is necessary only to take $h_n(\omega) = e^{i\omega\tau_n}$, where $\{\tau_n\}$ is the specified set of real numbers constituting the sampling times. Evidently, $h_n \in H_2$ corresponds to $x(\tau_n) \in H_1$. Therefore, $V\{e^{i\omega\tau_n}\} = H_2$ implies that the result of a linear operation on $x(t)$ can be given by a weighted sum of sample values $x(\tau_n)$. More precisely, suppose that $y(t)$ is any random process of the type described in the last paragraph. We may then choose a set of functions $c_n(t)$ such that the random process $\hat{y}(t) = \text{l.i.m.} \sum c_n(t)x(\tau_n)$ is equal to $y(t)$ in the sense that $\hat{y}(t) = y(t)$ with probability one for each t.

Of value for practical applications is the specialization $h(\omega, t) = h(\omega)e^{i\omega t}$, $h \in H_2$. $h(\omega, t)$ assumes this form if and only if $y(t)$ is (wide sense) stationary. Here $h(\cdot)$ may be regarded as a linear time-invariant operator on $x(t)$, and is closely related to the transfer function concept used frequently in engineering. More specifically, $h(\omega) = \text{l.i.m.} \sum a_n e^{i\omega t_n}$ which leads to $h(\omega)e^{i\omega t} = \text{l.i.m.} \sum a_n e^{i\omega(t_n+t)}$ and makes $y(t) = \text{l.i.m.} \sum a_n x(t_n + t)$; this verifies that $h(\omega)$ is a linear time-invariant operator on $x(t)$. If now $V\{e^{i\omega\tau_n}\} = H_2$, any linear time-invariant operation on $x(t)$ can be reproduced as a weighted sum of the samples $x(\tau_n)$. Note that $\hat{y}(t) = \text{l.i.m.} \sum c_n(t)x(\tau_n)$ is also (wide sense) stationary [since $\hat{y}(t) = y(t)$ with probability one], but that the partial sums $\hat{y}_N(t) = \sum_{-N}^{+N} c_n(t)x(\tau_n)$ are, in general, nonstationary.

III. STANDARD SAMPLING THEOREMS

The Shannon sampling theorem paraphrased in the Introduction will now be stated in a precise form and proved. No novelty is claimed for the proof, which is like Balakrishnan's (1957, Theorem 2). However, the considerations of Section II imply more than is stated elsewhere, namely, that even time-varying linear operators on $x(t)$ can be expressed as a weighted sum of the samples. Moreover, the proof given in our Theorem 1 can be elaborated to yield new results, of which some of the subsequent theorems are examples.

[3] For each t, $y(t)$ is actually one of an equivalence class of random variables which may differ on sets of zero probability. Any such version of $y(t)$ is adequate for our purpose, since separability and Lebesgue measurability of $y(t)$ do not enter into our arguments. All $h(\omega, t)$ are likewise defined for each t up to a ω set of measure 0 without ill effects.

In terms of the normalization of $F(\cdot)$ and the definition of band-limited process given in the Introduction, the basic sampling theorem becomes

THEOREM 1: *Let $x(t)$ be a (wide sense) stationary band-limited random process, with the further restriction that $F(\cdot)$ is continuous at $-\pi$ and $+\pi$. Then $V\{e^{i\omega n}$, all $n\} = H_2$ and $V\{x(n)$, all $n\} = H_1$. Specifically,*

$$\hat{x}(t) = x(t) \quad \text{prob. 1 for each } t \tag{3.1}$$

where $\hat{x}(t)$ is the random process defined as

$$\hat{x}(t) = \text{l.i.m.} \sum c_n(t)x(n) \tag{3.2}$$

in which

$$c_n(t) = \frac{\sin \pi(t-n)}{\pi(t-n)} \tag{3.3}$$

PROOF: We assume now—and prove later—that

$$\lim_{N\to\infty} \| e^{i\omega t} - \sum_{-N}^{+N} c_n(t)e^{i\omega n} \| = 0 \tag{3.4}$$

for some set of $c_n(t)$. Then the first assertion of the theorem is true. Also, $V\{x(n), \text{all } n\} = H_1$, since H_1 and H_2 are isomorphic with $x(n) \in H_1$ corresponding to $e^{i\omega n} \in H_2$.

The existence of $h(\omega, t) = \text{l.i.m.} \sum c_n(t)e^{i\omega n}$ is another consequence of (3.4). In H_1, there is the corresponding element $\text{l.i.m.} \sum c_n(t)x(n)$ so that the definition of $\hat{x}(t)$ given by (3.2) makes sense. Now suppose t fixed but arbitrary, and let any $\delta > 0$ be specified. If N is sufficiently large

$$\| \hat{x}(t) - \sum_{-N}^{+N} c_n(t)x(n) \| < \delta/2 \tag{3.5}$$

from (3.2), and

$$\| x(t) - \sum_{-N}^{+N} c_n(t)x(n) \| < \delta/2 \tag{3.6}$$

because of (3.4) and the isomorphism between H_1 and H_2. Combining (3.5) and (3.6) via the triangle inequality yields

$$\| \hat{x}(t) - x(t) \| < \delta \tag{3.7}$$

But (3.7) holds for any $\delta > 0$, so that $\| \hat{x}(t) - x(t) \| = 0$ or

$$E\{| \hat{x}(t) - x(t) |^2\} = 0.$$

The latter form clearly implies (3.1).

To complete the proof, we demonstrate the truth of (3.4); the $c_n(t)$ which appears there is specified by (3.3). Indeed, (3.4) is true (by definition of the norm) if and only if

$$\lim_{N\to\infty} \int_{-\pi-0}^{\pi+0} \left| e^{i\omega t} - \sum_{-N}^{N} c_n(t) e^{i\omega n} \right|^2 dF(\omega) = 0 \qquad (3.8)$$

It will be proved that the limit of the integrand is zero a.e. (measure dF), and that interchange in the order of integration and taking limits is legitimate.

Let us expand $e^{i\omega t}$ as a function of ω over the interval $-\pi$ to $+\pi$ in a trigonometric (Fourier) series. This series has partial sums $\sum_{-N}^{+N} c_n(t)e^{i\omega n}$, where $c_n(t) = (1/2\pi) \int_{-\pi}^{+\pi} e^{i\omega t}e^{-i\omega n}\, dt$. A simple calculation shows that these $c_n(t)$ agree with (3.3).

The convergence and boundedness properties of $\sum_{-N}^{+N} c_n(t)e^{i\omega n}$ are determined (for each fixed t) by the fact that $e^{i\omega t}$ is continuous and of bounded variation. Hence (Titchmarch, 1939, pp. 406–408) $\sum_{-N}^{+N} c_n(t)e^{i\omega n}$ converges to $e^{i\omega t}$ everywhere in $(-\pi, +\pi)$, but not necessarily at $\omega = \pm\pi$ (since $e^{i\pi t} \neq e^{-i\pi t}$, in general). Because $\pm\pi$ are continuity points of $F(\cdot)$, and because the measure induced by $F(\cdot)$ is zero outside of $[-\pi, +\pi]$, convergence on $(-\pi, +\pi)$ is convergence a.e. (measure dF). Therefore,

$$\lim_{N\to\infty} \left| e^{i\omega t} - \sum_{-N}^{+N} c_n(t)e^{i\omega n} \right|^2 = 0 \quad \text{a.e. (measure dF)} \qquad (3.9)$$

That $e^{i\omega t}$ is of bounded variation assures that $| \sum_{-N}^{+N} c_n(t)e^{i\omega n} |$ is uniformly bounded in ω and N. Then there exists an $M < \infty$ such that

$$| e^{i\omega t} - \sum_{-N}^{+N} c_n(t)e^{i\omega n} |^2 < M \quad \text{all } N, \omega \qquad (3.10)$$

We see from (3.10) that the integrand in (3.8) is dominated by an integrable function (a constant) independent of N. The interchange of limit and integration becomes valid (Halmos, 1950, p. 110), and (3.9) shows that the limit is zero; hence (3.8) is indeed true.

It is often asserted that $x(t)$ can be recovered from its samples only if its highest frequency component is *less* than half the sampling frequency. The continuity of $F(\cdot)$ at its endpoints is a precise version of this statement. For a complex random process, however, it is permissible that $F(\cdot)$ be discontinuous at $-\pi$ or $+\pi$, but not both.

If $x(t)$ is real, its component at angular frequency π must be of the form $\alpha \cos(t + \tau)$, where α and τ are random variables, τ being uniformly distributed on $(0, 2\pi)$ (cf. Beutler, in press). The sample components due to this frequency are all equal, with their magnitude dependent on both α and τ. This makes it impossible to recover the π frequency component, even if the mean of $x(t)$ is zero.

On the other hand, if $F(\cdot)$ has a jump of σ^2 at only one endpoint, say $+\pi$, there is a contribution of $\alpha e^{i\pi t}$ to $x(t)$, where α is a random variable with $E[\alpha^2] = \sigma^2$. Then samples of $\alpha e^{i\pi t}$ at unit intervals have magnitude α and alternating sign, so that α is recoverable, even if $x(t)$ has nonzero mean. Jumps of $F(\cdot)$ at both $-\pi$ and $+\pi$ would contribute a term $\alpha e^{i\pi t} + \beta e^{-i\pi t}$ to $x(t)$, where α and β are orthogonal random variables. Unit interval sampling could therefore recover the sum $\alpha + \beta$, but no more.

The affirmative statement of the last paragraph is formalized by

THEOREM 1a: *Let $x(t)$ be as in Theorem 1, except that $F(\cdot)$ may be discontinuous at $-\pi$ or $+\pi$, but not both. Then the conclusions of Theorem 1 continue to hold.*

PROOF: We again show that $\| e^{i\omega t} - \sum_{-N}^{+N} c_n(t) e^{i\omega n} \|$ can be made as small as desired (for fixed but arbitrary t) by taking N sufficiently large. To accomplish this, we define an appropriate $g(\cdot)$, and prove that both right-hand side terms of

$$\begin{aligned} &\| e^{i\omega t} - \sum_{-N}^{+N} c_n(t) e^{i\omega n} \| \\ &\qquad \leqq \| g(\omega) - \sum_{-N}^{+N} c_n(t) e^{i\omega n} \| + \| e^{i\omega t} - g(\omega) \| \end{aligned} \tag{3.11}$$

become zero as $N \to \infty$. This artifice is necessary to take into account the jump at $-\pi$ or $+\pi$.

To be definite, assume the jump to be at $+\pi$. Then $g(\cdot)$ is defined by

$$g(\omega) = \begin{cases} e^{i\omega t} & \omega \in (-\pi - 0, \pi - \delta] \\ e^{i(\pi-\delta)t} + \dfrac{\omega - (\pi - \delta)}{\delta} (e^{-it\pi} - e^{t(\pi-\delta)}) & \\ & \omega \in (\pi - \delta, \pi + 0) \end{cases} \tag{3.12}$$

where $0 < \delta < \pi$ will be chosen presently. We observe that $g(\omega)$ and $e^{i\omega t}$ coincide[4] except on $(\pi - \delta, \pi)$, and that

$$| e^{i\omega t} - g(\omega) | \leqq 2 \tag{3.13}$$

Hence

$$\| e^{i\omega t} - g(\omega) \| \leqq 2\{F(\pi - 0) - F(\pi - \delta)\}^{1/2} \tag{3.14}$$

It will be proved that, for any $\epsilon > 0$, there exists a $\delta > 0$ such that $F(\pi - 0) - F(\pi - \delta) < \epsilon$.[5] Let $F(\cdot)$ be decomposed into its continuous and discontinuous parts, $F_c(\cdot)$ and $F_d(\cdot)$, respectively. It is clear that there exists a $\delta > 0$ such that $F_c(\pi) - F_c(\pi - \delta) < \epsilon/2$. For $F_d(\cdot)$,

$$F_d(\omega) = \sum_{\omega_j \leqq \omega} u_j(\omega_j) \leqq 1 \tag{3.15}$$

where the jumps $u_j(\omega_j) > 0$ occur at $\omega_1, \omega_2, \cdots$. The convergence of this sum of positive terms implies that there is a finite set of indices J so chosen that $\sum_{\omega_j \notin J} u_j(\omega_j) < \epsilon/2$. Since the ω_j belonging to J are nowhere dense, $\delta > 0$ can be selected so that $(\pi, \pi - \delta)$ contains none of these ω_j. Then $F_d(\pi - 0) - F_d(\pi - \delta) \leqq \sum_{\omega_j \notin J} u_j(\omega_j) < \epsilon/2$. Combining the latter result with $F_c(\pi) - F_c(\pi - \delta) < \epsilon/2$, and choosing whichever δ is smaller, gives the desired result. Thus, (3.14) may be rendered as close to zero as desired by making $\delta > 0$ sufficiently small.

Turning now to the other term on the right hand side of (3.11), we note that $g(\cdot)$ is continuous and of bounded variation, with $g(\pi) = g(-\pi)$. Therefore, the Fourier series for $g(\cdot)$ over $[-\pi, +\pi]$ converges boundedly at every point of the interval, including also the endpoints. The arguments used in Theorem 1 to prove (3.4) are repeated verbatim, except that the a.e. convergence (measure dF) of $\sum_{-N}^{+N} c_n(t)e^{i\omega n}$ to $g(\omega)$, where the $c_n(t)$ are Fourier coefficients, follows *a fortiori* from the ordinary convergence everywhere.

Since (3.4) has now been proved valid under the more general conditions of Theorem 1a, all the conclusions of Theorem 1 follow (proof is as before) with the possible exception of (3.3). But

$$\begin{aligned} \left| c_n(t) - \frac{\sin \pi(t - n)}{\pi(t - n)} \right| &= \left| \frac{1}{2\pi} \int_{-\pi-0}^{\pi+0} [g(\omega) - e^{i\omega t}] e^{-i\omega t} \, d\omega \right| \\ &\leqq \frac{1}{2\pi} \int_{\pi-\delta}^{\pi} | g(\omega) - e^{i\omega t} | \, d\omega < \frac{\delta}{\pi} \end{aligned} \tag{3.16}$$

[4] Since we assume a fixed (but arbitrary) t, the parameter t is suppressed for convenience in writing $g(\omega)$. The same convention is also used later in this paper.

[5] If $F(\cdot)$ has a jump at $-\pi$ instead of $+\pi$ we must prove $F(-\pi + \delta) - F(-\pi) < \epsilon$. But this is a direct consequence of the right continuity of $F(\cdot)$.

These inequalities hold because $g(\omega)$ differs from $e^{i\omega t}$ only on $(\pi - \delta, \pi)$, and because of (3.13). Now $\delta > 0$ can be made arbitrarily small without affecting the remainder of the proof. Then indeed $c_n(t) = [\sin \pi(t - n)]/\pi(t - n)$, i.e., (3.3) is satisfied.

Theorem 1 can be generalized beyond Theorem 1a at the expense of complexity of the appropriate proofs. This is not done here, since Theorem 1a already illustrates how such proofs might be constructed. Moreover, a recent theorem by S. P. Lloyd (1959) is available, and represents the most comprehensive result one is able to obtain.[6]

IV. NONPERIODIC SAMPLING

It is frequently desired to recover $x(t)$ when the only available samples are not periodic. Such applications arise if (1) some samples have been obliterated (e.g., through equipment failure), (2) sampling times are (*ex post facto*) found to have been irregular, or (3) only a brief period is available for sampling.

Signal recovery from nonperiodic samples is associated with finding conditions on a set τ_n (apart from $\tau_n = n$) which assure that $V\{e^{i\omega\tau_n}\} = H_2$. Each of the above applications is related to one such theorem. Our first result in this direction is

THEOREM 2: *Let $x(t)$ be as in Theorem 1a, and suppose there exists a $\omega_0 \varepsilon [-\pi, +\pi]$ such that $\omega_0 \epsilon S'$ (i.e., ω_0 belongs to the complement of S, the spectrum of the process). Let K be an arbitrary finite set of integers n_1, n_2, $\cdots$, n_k. Then*

$$V\{e^{i\omega n}, n \in K'\} = H_2 \tag{4.1}$$

Also, (3.1) holds with $\hat{x}(t)$ defined by (3.2), except that the $c_n(t)$ may be chosen so that

$$c_n(t) = 0 \quad \text{for } n \in K \tag{4.2}$$

PROOF: If $F(\cdot)$ has a discontinuity at $-\pi$ or $+\pi$, we take the discontinuity at $+\pi$, as in Theorem 1a; this is merely a convenience. To verify (4.1), we again turn to (3.11). It is necessary only to show

$$\lim_{N\to\infty} \| g(\omega) - \sum_{-N}^{+N} c_n(t) e^{i\omega n} \| = 0 \tag{4.3}$$

where the $c_n(t)$ satisfy (4.2); the remainder of the proof is already given in Theorem 1a.

[6] The author has been able to prove this theorem by methods similar to those of Theorem 1a. Lloyd's approach is quite different.

We assume that $\omega_0 \in (-\pi, +\pi)$. If $\omega_0 = \pm\pi$, the slight modifications required for the proof are obvious. Since S' is easily verified to be open, there exists an interval (a, b) with $-\pi < a < \omega_0 < b < \pi$ and such that $F(b) - F(a) = 0$. Moreover, we may choose the $\delta > 0$ appearing in (3.12) to satisfy both $\| e^{i\omega t} - g(\omega) \| < \epsilon/2$ and $b < \pi - \delta$. Finally, let (α, β) be another nondegenerate interval contained in (a, b), i.e., $a < \alpha < \beta < b$.

It is clear that (4.3) can be satisfied even if $\sum_{-N}^{+N} c_n(t)e^{i\omega n}$ fails to converge to $g(\omega)$ on (α, β). Indeed, if

$$f(\omega) = g(\omega) \qquad \omega \notin (\alpha, \beta) \tag{4.4}$$

we have $f(\omega) = g(\omega)$ a.e. (measure dF) and so

$$\lim_{N\to\infty} \left\| f(\omega) - \sum_{-N}^{+N} c_n(t)e^{i\omega n} \right\| = 0 \tag{4.5}$$

implies (4.3), and conversely.

We take advantage of the lack of uniqueness on (α, β) to define an $f(\cdot)$ meeting (4.4) and (4.2) simultaneously, and also satisfying (4.5). Take

$$f(\omega) = \begin{cases} g(\omega) & \omega \notin (\alpha, \beta) \\ \sum_1^k b_j e^{i\omega n_j} & \omega \in (\alpha, \beta) \end{cases} \tag{4.6}$$

For any choice of b_j, $f(\cdot)$ is of bounded variation and continuous except (possibly) at α and β. Hence, if the $c_n(t)$ are Fourier coefficients of $f(\omega)$ over $(-\pi, +\pi)$, $\sum_{-N}^{+N} c_n(t)e^{i\omega n} \to f(\omega)$ except at α and β. But α and β are obviously continuity points of $F(\cdot)$, so that this convergence holds a.e., measure dF. Since $\sum_{-N}^{+N} c_n(t)e^{i\omega n}$ converges boundedly [cf. (3.10)], the arguments of Theorem 1 apply to yield (4.5).

It remains to verify that the b_j can be chosen to satisfy (4.2). Since the $c_{n_j}(t)$ are Fourier coefficients of $f(\cdot)$, they will be given by

$$c_{n_j}(t) = \sum_{r=1}^{k} g_{jr} b_r + d_j \tag{4.7}$$

where

$$d_j = \frac{1}{2\pi}\left(\int_{-\pi}^{+\pi} - \int_{\alpha}^{\beta}\right) g(\omega)e^{-i\omega n_j}\, d\omega \tag{4.8}$$

and the

$$g_{jr} = \frac{1}{2\pi}\int_{\alpha}^{\beta} e^{i\omega(n_j - n_r)}\, d\omega. \tag{4.9}$$

To satisfy (4.2), it must be possible to choose the b_j so that $\sum g_{jr}b_r + d_j = 0, j = 1, 2, \cdots, k$. There are two cases: either $d_j = 0$ for all j, or at least one of the d_j is not zero. The first case is disposed of by taking $b_j = 0, j = 1, 2, \cdots, k$. For the second, the linear equations $\sum g_{jr}b_r + d_j = 0$ admit a solution $b_1, b_2, \cdots, b_k$ if (and only if) the matrix $[g_{jr}]$ is nonsingular. In fact, (4.9) exhibits $|g_{jr}|$ as the Gram determinant (Achieser and Glasman, 1954, pp. 13–14) of the set $\{1/\sqrt{2\pi}\}e^{i\omega n_j}$ over (α, β). Since this set is linearly independent over any interval, $|g_{jr}|$ is strictly positive, and the proof is complete.

A second result, always valid in the band-limited case, but also true under less restrictive assumptions, is

THEOREM 3: *Let $\{\tau_n\}$ have a finite limit point τ, and suppose that $F(\cdot)$ has the property*

$$\int_{-\infty}^{\infty} \exp[c\,|\,\omega\,|]\, dF(\omega) < \infty \tag{4.10}$$

for every $c < \infty$. Then

(i) $V\{e^{i\omega\tau_n}\} = H_2$

(ii) *$x(t)$ has derivatives of all orders in the sense that*

$$x^{(n)}(t) = \underset{t' \to t}{\text{l.i.m.}} \frac{x^{(n-1)}(t') - x^{(n-1)}(t)}{t' - t} \tag{4.11}$$

exists, with the understanding that $x^{(0)}(t) = x(t)$

(iii) *$\hat{x}(t) = x(t)$ with probability one for each t, where*

$$\hat{x}(t) = \underset{N \to \infty}{\text{l.i.m.}} \sum_0^N \frac{(t-\tau)^n x^{(n)}(\tau)}{n!}. \tag{4.12}$$

PROOF: We may take $\tau = 0$ without loss of generality, so that there exists a subsequence in $\{\tau_n\}$ such that $\lim_{n\to\infty} \tau_n = 0$ (we have omitted indexing the subscripts of the τ_n in this sub-sequence, since the sub-sequence contains all the τ_n we shall need).

As an intermediate step in the proof of (i), we substantiate that

$$(i\omega)^m \in V\{e^{i\omega\tau_n}\} \qquad m = 0, 1, 2, \cdots \tag{4.13}$$

The statement (4.13) is true for $m = 0$ even without assuming (4.10); it follows from the easily proved $\lim_{n\to\infty} \|\, 1 - e^{i\omega\tau_n} \,\| = 0$. For $m = 1, 2, \cdots$ we proceed by induction.

For $m = 1$, note that $(e^{i\omega\tau_n} - 1)/\tau_n \in V\{e^{i\omega\tau_n}\}$, so that

$$\lim_{n\to\infty} \left\| (i\omega) - \frac{e^{i\omega\tau_n} - 1}{\tau_n} \right\|^2 = \lim_{n\to\infty} \int_{-\infty}^{\infty} \left| (i\omega) - \frac{e^{i\omega\tau_n} - 1}{\tau_n} \right|^2 dF(\omega) = 0 \tag{4.14}$$

gives the desired result. Now the integrand in (4.14) converges to zero. Furthermore, the integrand is majorized by $4\omega^2$, which is integrable by virtue of (4.10). The truth of (4.14) is therefore established through use of Lebesgue's dominated convergence theorem.

Let us assume that $(i\omega)^k \in V\{e^{i\omega\tau_n}\}$ for $k = 1, 2, \cdots, m - 1$; we show $(i\omega)^m \in V\{e^{i\omega\tau_n}\}$ also. Under our assumption

$$\frac{e^{i\omega\tau_n} - \sum_0^{m-1} (i\omega)^k {\tau_n}^k/k!}{{\tau_n}^m} \in V\{e^{i\omega\tau_n}\}$$

so that the assertion $(i\omega)^m \in V\{e^{i\omega\tau_m}\}$ follows if we establish that

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \left| \frac{(i\omega)^m}{m!} - \frac{e^{i\omega\tau_n} - \sum_0^{m-1} (i\omega)^k {\tau_n}^k/k!}{{\tau_n}^m} \right|^2 dF(\omega) = 0 \tag{4.15}$$

Again, the integrand becomes zero as $n \to \infty$. Once more the dominated convergence theorem is applicable, because the integrand is bounded by $\exp[2|\omega|]$ for all n sufficiently large to assure $|\tau_n| \leqq 1$. The bound itself is integrable by (4.10). To verify that $\exp[2|\omega|]$ is indeed an upper bound for $|\tau_n| \leqq 1$, we calculate the integrand in (4.16) to be equal to $|\sum_{k=m+1}^{\infty} (i\omega)^k \tau_n^{k-m}/k!|^2$. Thus the proof of (4.13) by induction is complete.

In view of (4.13), (i) is substantiated if we show that, for arbitrary t, $e^{i\omega t} \in V\{(i\omega)^m, m = 0, 1, 2, \cdots\}$. Now $\sum_0^N (i\omega)^n t^n/n! \to e^{i\omega t}$ (ordinary convergence), and also

$$\left| e^{i\omega t} - \sum_0^N \frac{(i\omega)^n t^n}{n!} \right|^2 \leqq \left(\sum_{N+1}^{\infty} \frac{|\omega|^n |t|^n}{n!} \right)^2 \leq \exp[|2t\omega|] \tag{4.16}$$

By (4.10), the right-hand side of (4.16) is integrable. Once more, the dominated convergence theorem is applicable. This yields

$$\lim_{N\to\infty} \left\| e^{i\omega t} - \sum_0^N \frac{(i\omega)^n t^n}{n!} \right\| = 0 \tag{4.17}$$

to complete the proof of (i).

We now turn to (ii). For $n = 1$, $[x(t') - x(t)]/(t' - t)$ in H_1 corresponds to $(e^{i\omega t'} - e^{i\omega t})/(t' - t)$ in H_2. Refering to (4.14) and the remarks immediately following implies (via an identical argument) that

$$\underset{t' \to t}{\text{l.i.m.}} \frac{e^{i\omega t'} - e^{i\omega t}}{t' - t} = (i\omega)e^{i\omega t}.$$

Then $\text{l.i.m.}_{t' \to t} [x(t') - x(t)/t' - t]$ exists (because of the isomorphism of H_1 with H_2). Therefore, corresponding to $(i\omega)e^{i\omega t} \in H_2$ is $x^{(1)}(t) \in H_1$.

Suppose now that (4.11) is true for $n = 1, 2, \cdots, m - 1$, with $(i\omega)^n e^{i\omega t} \in H_2$ corresponding to $x^{(n)}(t) \in H_1$ for these n. We assert that then (4.11) also holds for $n = m$, so that (ii) will have been proved by induction. Indeed,

$$\underset{t' \to t}{\text{l.i.m.}} \frac{x^{(m-1)}(t') - x^{(m-1)}(t)}{t' - t}$$

exists because (as we shall prove)

$$\underset{t' \to t}{\text{l.i.m.}} \frac{(i\omega)^{m-1}e^{i\omega t'} - (i\omega)^{m-1}e^{i\omega t}}{t' - t} = (i\omega)^m e^{i\omega t}. \tag{4.18}$$

Then also to $x^{(m)}(t) \in H_1$ corresponds $(i\omega)^m e^{i\omega t} \in H_2$, which completes the inductive argument. To verify (4.18), observe that the indicated limit holds in the ordinary sense, and that

$$\left| \frac{(i\omega)^{m-1}e^{i\omega t'} - (i\omega)^{m-1}e^{i\omega t}}{t' - t} \right|^2 \leqq \omega^{2m},$$

which is integrable. Therefore the dominated convergence theorem may be used to obtain (4.18).

Only (iii) remains to be verified. We already have

$$e^{i\omega t} = \underset{N \to \infty}{\text{l.i.m.}} \sum_0^N (i\omega)^n t^n / n!.$$

In terms of the correspondence of $(i\omega)^n$ and $x^{(n)}(0)$, we have $x(t) = \text{l.i.m.}_{N \to \infty} \sum_0^N t^n x^{(n)}(0)/n!$. An argument identical with that presented by (3.5), (3.6), and (3.7) yields the result (iii). Note that here, as before, we have taken $\tau = 0$.

Well-known closure theorems (Levinson, 1940) describe conditions under which $V\{e^{i\omega\tau_n}\}$ is complete in $L_2(-\pi, \pi)$. Since L_2 convergence implies convergence in H_2 whenever $F(\cdot)$ is absolutely continuous with a bounded derivative (spectral density), any such closure theorem states also that $V\{e^{i\omega\tau_n}\} = H_2$ if $F(\cdot)$ meets the above conditions. Rather than to enumerate closure theorems in $L_2(-\pi, \pi)$, we prefer to prove one theorem under a broader set of conditions on $F(\cdot)$.

Accordingly, we state

THEOREM 4: *Let $x(t)$ be band-limited, and let there exist an $\epsilon > 0$ such that $F(\cdot)$ is absolutely continuous on $[-\pi, -\pi + \epsilon)$ and $(\pi - \epsilon, \pi]$, and has a bounded derivative over these same intervals. Let*

$$|n - \tau_n| \leqq M < \frac{\log 2}{\pi} \qquad \text{for all } n. \tag{4.19}$$

Then $V\{e^{i\omega\tau_n}\} = H_2$ and $\hat{x}(t) = x(t)$ with probability one for each t, where

$$\hat{x}(t) = \underset{N\to\infty}{\text{l.i.m.}} \sum_{-N}^{N} b_n(t)x(\tau_n) \tag{4.20}$$

is defined as a random process.

REMARK: This is evidently a "perturbation theorem." It states that the sampling times need not be periodic, but may vary from true periodicity by over 20% without sacrificing capability of restoring $x(t)$. It is this interpretation which leads to useful applications involving faulty timers, variable transmission rates, etc.

PROOF: According to the hypotheses of the theorem, we assume $F'(\cdot) \leqq K$ on $[-\pi, -\pi + \epsilon)$ and $(\pi - \epsilon, \pi]$. Take $\epsilon > \eta > 0$, and consider

$$\begin{aligned} \left\| e^{i\omega t} - \sum_{-N}^{N} b_n(t) e^{i\omega t_n} \right\|^2 &\leqq \int_{-\pi+\eta-0}^{\pi-\eta+0} \left| e^{i\omega t} - \sum_{-N}^{N} b_n(t) e^{i\omega t_n} \right|^2 dF(\omega) \\ &\quad + k \left(\int_{\pi-\eta+0}^{\pi} + \int_{-\pi}^{-\pi+\eta-0} \right) \left| e^{i\omega t} - \sum_{-N}^{N} b_n(t) e^{i\omega\tau_n} \right|^2 d\omega \end{aligned} \tag{4.21}$$

We claim that both terms on the right-hand side of the inequality (4.21) can be made as small as desired by choosing N sufficiently large.

The square root of the first term on the right of (4.21) is majorized by

$$\begin{aligned} &\left(\int_{-\pi+\eta-0}^{\pi-\eta+0} \left| e^{i\omega t} - \sum_{-N}^{N} b_n(t) e^{i\omega\tau_n} \right|^2 dF(\omega) \right)^{1/2} \\ &\qquad \leqq \left(\int_{-\pi+0}^{\pi-0} \left| e^{i\omega t} - \sum_{-N}^{N} c_n(t) e^{i\omega n} \right|^2 dF(\omega) \right)^{1/2} \\ &\qquad + \left(\int_{-\pi+\eta-0}^{\pi-\eta+0} \left| \sum_{-N}^{N} [c_n(t) e^{i\omega n} - b_n(t) e^{i\omega\tau_n}] \right|^2 dF(\omega) \right)^{1/2} \end{aligned} \tag{4.22}$$

This inequality is Minkowski's (triangle inequality), and the right side is further increased by enlarging the interval of integration. Theorem 1 is applied to the first term on the right side of (4.22), the $c_n(t)$ being given by (3.3). Consequently, this term converges to zero. The second right-hand term of (4.22) is treated as follows. Levinson (1940, theorem XVIII) has shown that (4.19) implies

$$\lim_{N\to\infty} \sum_{-N}^{N} [c_n(t)e^{i\omega n} - b_n(t)e^{i\omega\tau_n}] = 0 \tag{4.23}$$

uniformly on $[-\pi + \epsilon, \pi - \epsilon]$ for properly chosen $b_n(t)$.[7] This uniform convergence to zero of the integrand causes the second right-hand term in (4.22) to converge to zero also.

The second term on the right-hand side of (4.21) cannot become smaller if the region of integration is enlarged. Therefore, it suffices to demonstrate

149

$$\lim_{N\to\infty} \int_{-\pi}^{\pi} \left| e^{i\omega t} - \sum_{-N}^{N} b_n(t)e^{i\omega\tau_n} \right|^2 d\omega = 0 \tag{4.24}$$

That (4.24) is valid is again the result of (4.19) with the same choice of $b_n(t)$, as shown in Riesz and Nagy (1955, pp. 208–210). Thus the proof of $V\{e^{i\omega\tau_n}\} = H_2$ is complete.

The other assertions of Theorem 4 follow readily from the above. Indeed, their proof is precisely as in Theorem 1.

V. APPLICATIONS OF OPTIMUM PREDICTION THEORY

Heretofore, our sampling theorems have (with the exception of Theorem 3) required that we "sample the future" as well as the past. To elaborate, we may select some t_0 and regard it as the time at the present instant. Then $\tau_0 > t_0$ is some future time, and $x(\tau_n)$ is a future value of $x(t)$. In practice, only the part of $x(t)$ is available to us, i.e., we can obtain only those $\tau_n \leqq t_0$. Nevertheless, the τ_n appearing in the sampling theorems ranged from $-\infty$ to $+\infty$, implying future and therefore unrealizable values of $x(t)$.

As indicated above, it is of interest to learn under what circumstances a set of samples of the past span H_1[i.e., determine $x(t)$]. To this end, prediction theory enables us to prove

[7] If $\{h_n(\omega)\}$ is the set of biorthogonal functions in $L_2(-\pi, \pi)$ associated with $\{e^{i\omega\tau_n}\}$, $b_n(t) = \int_{-\pi}^{\pi} e^{i\omega t}h_n(\omega)\, d\omega$.

THEOREM 5: *Let $F(\cdot)$ have $\phi(\cdot)$ as a derivative,*[8] *with the proviso that*

$$\int_{-\infty}^{\infty} \frac{\log \phi(\omega)}{\omega^2 + 1} \, d\omega = -\infty \tag{5.1}$$

Then, for any t_0, $-\infty < t_0 < +\infty$,

$$V\{e^{i\omega r}, r \text{ all rationals} \leqq t_0\} = H_2 \tag{5.2}$$

THEOREM 6: *Let $x(t)$ meet the conditions of Theorem 1a, and let the $\phi(\cdot)$ defined in Theorem 5 satisfy*

$$\int_{-\pi}^{+\pi} \log \phi(\omega) \, d\omega = -\infty \tag{5.3}$$

If n_0 is any integer,

$$V\{e^{i\omega n}, n \leqq n_0\} = H_2 . \tag{5.4}$$

Theorems 3 and 5 are closely related. The conclusions of Theorem 3 are stronger; the theorem affirms not only that $x(t)$ can be reconstructed from its values on a set $\{\tau_n\}$ having a finite limit point (instead of the half line of Theorem 5), but also exhibits a method of obtaining $x(t)$ from these samples. At the same time, Theorem 3 requires a more restrictive assumption than Theorem 5. We will verify this statement by showing that (4.10) implies (5.1). Suppose now that (4.10) is true (with some $c > 0$) but that (5.1) is false. For any nonnegative $g(\cdot)$ and nondecreasing $F(\cdot)$, we have $\int g(\omega)\phi(\omega) \, d\omega \leqq \int g(\omega) \, dF(\omega)$. Here we shall use $g(\omega) = \exp [c| \omega |]/\omega^2 + 1$, and then take the logarithm of both sides. Because of (4.10), the right side of the inequality is finite, so that we shall have

$$\log \int_{-\infty}^{\infty} \frac{\exp [c | \omega |]\phi(\omega)}{\omega^2 + 1} \, d\omega < \infty \tag{5.5}$$

This integral is subject to another inequality (Hardy, Littlewood, and Polya, 1952, Theorem 184), based on the convexity properties of the logarithm. We thus obtain

$$\int_{-\infty}^{\infty} \frac{c| \omega | + \log \phi(\omega)}{\omega^2 + 1} \, d\omega < \infty \tag{5.6}$$

[8] Since $F(\cdot)$ is nondecreasing, $\phi(\cdot)$ exists and is nonnegative a.e. (Lebesgue measure). More properly, $\phi(\cdot)$ is so defined up to an a.e. equivalence. However, any such $\phi(\cdot)$ will suit our purpose equally well.

The numerator of the integrand may be split to yield two integrals. This procedure is valid because [if (5.1) is false] both integrals exist in the extended sense, without one having value $-\infty$ and the other $+\infty$. Then we have

$$c \int_{-\infty}^{\infty} \frac{|\omega|}{\omega^2+1}\, d\omega + \int_{-\infty}^{\infty} \frac{\log \phi(\omega)}{\omega^2+1}\, d\omega < \infty \tag{5.7}$$

The assumption that (5.1) is false therefore gives $\int_{-\infty}^{\infty} \frac{|\omega|}{\omega^2+1}\, d\omega < \infty$, which is patently absurd.

Theorem 5 is Theorem 5.2(ii) in Doob (1953), Chapter XII, and Theorem 6 is Theorem 4.4(ii) in the same chapter, with a different normalization of the interval.

VI. DIMENSIONALITY

The dimensionality of H_2 is an indication of the number of samples required to recover $x(t)$. We have already exhibited a denumerable base, $V\{e^{i\omega r}$, all rational $r\} = H_2$, so that at most an enumerable set of samples is necessary to determine not only $x(t)$, but also any linear operation on $x(t)$. However, considerable practical interest rests in finding conditions under which, for a fixed set $\tau_1, \tau_2, \cdots, \tau_N$, the samples $x(\tau_1), x(\tau_2), \cdots, x(\tau_N)$ suffice to determine $x(t)$. In other words, what are necessary and sufficient conditions on $F(\cdot)$ that H_2 have exactly N dimensions?

Before disposing of the above question, it should be noted that some sources (Shannon, 1949; Goldman, 1954) speak of a finite sampling procedure relative to a signal zero outside of some finite interval. The reference here is not to a random process, although the erroneous inference is sometimes made. Elsewhere (Rice, 1954, section 1.7), it is found convenient to represent $x(t)$ by $x(t) = \sum_0^M [a_n \cos(nt) + b_n \sin(nt)]$. This $x(t)$ has $2M$ dimensions, and if (wide sense) stationary (all a_n and b_n orthogonal, and of equal mean square for each n), is easily verified to have a spectrum as specified by our theorem on finite dimensional processes.

In the case of band-limited $x(t)$, the dimension of H_2 is completely characterized by

THEOREM 7: *Let $x(t)$ be band-limited. Then one (and only one) of the following is true:*

(i) *S contains an infinity of points and H_2 is of denumerably infinite dimension.*

(ii) *S contains $N < \infty$ points, and H_2 is of dimension N.*
In case (ii), *any set of the form* $\{e^{i\omega(n+m)}$, *m an arbitrary integer*, $n = 0, 1, \cdots, N-1\}$ *constitutes a base in* H_2.

PROOF: Since $V\{e^{i\omega r}$, all rational $r\} = H_2$, this space is always separable. Suppose now that S is an infinite set, and that H_2 is of finite dimension N. Then there exist c_n (not all zero) such that $\| \sum_0^N c_n e^{i\omega n} \| = 0$. Since for any choice of c_n (not all zero) a trigonometric polynomial of degree N has at most N zeros, we can find a $\omega_0 \in S$ and a $\delta > 0$ so that $[\omega_0 - \delta, \omega_0 + \delta]$ contains no zeros of $\sum_0^N c_n e^{i\omega n}$. There follows

$$\left\| \sum_0^N c_n e^{i\omega n} \right\|^2 \geqq \int_{\omega_0-\delta}^{\omega_0+\delta} \left| \sum_0^N c_n e^{i\omega n} \right|^2 dF(\omega) > 0 \tag{6.1}$$

which contradicts the supposition that H_2 has dimension N.

To verify (ii), we show first that the dimension of H_2 is at least N. If S contains N points, these are points of increase $\omega_1, \omega_2, \cdots, \omega_N$ at which $F(\cdot)$ has jumps of $\alpha_1, \alpha_2, \cdots, \alpha_N$, respectively. Then, unless all $c_n = 0$,

$$\| \sum_0^{N-1} c_n e^{i\omega n} \|^2 = \sum_0^N \alpha_k | \sum_{n=0}^{N-1} c_n e^{i\omega_k n} |^2 > 0 \tag{6.2}$$

as will be shown. Indeed, $\sum_0^{N-1} c_n e^{i\omega n}$ is a polynomial of degree $N-1$, and cannot be zero at all N points unless it is identically zero. The linear independence of the N elements $e^{i\omega n}$, $n = 0, 1, \cdots, N-1$ implies the desired conclusion.

Finally, we demonstrate that there is a base consisting of N elements, viz. $V\{e^{i\omega(n+m)}$, m arbitrary integer, $n = 0, 1, 2, \cdots, N-1\} = H_2$. The equality $\| e^{i\omega t} - \sum_{n=0}^{N-1} c_n e^{i\omega(n+m)} \| = \| e^{i\omega(t-m)} - \sum_0^{N-1} c_n e^{i\omega n} \|$ shows that we may take $m = 0$. The proof of the theorem is therefore complete if (for fixed but arbitrary t) we exhibit c_n such that

$$\| e^{i\omega t} - \sum_0^{N-1} c_n e^{i\omega n} \| = 0 \tag{6.3}$$

With the N points of increase (points of S) as in the preceding paragraph, (6.3) is satisfied if $\sum_0^{N-1} c_n e^{i\omega n}$ is equal to $e^{i\omega t}$ at $\omega = \omega_1, \omega_2, \cdots, \omega_N$. To this end we take

$$\sum_0^{N-1} c_n e^{i\omega n} = \sum_{k=1}^N e^{i\omega_k t} \prod_{j \neq k} \frac{(e^{i\omega} - e^{i\omega_j})}{(e^{i\omega_k} - e^{i\omega_j})} \tag{6.4}$$

which is the well-known Lagrange interpolation formula.

The statement and proof of (ii) raises the question whether $e^{i\omega n_k}$ is a base if the n_k are not consecutive integers. The answer, in general, is no. Consider, for example, $\omega_1 = -\pi/2$, $\omega_2 = +\pi/2$, $\alpha_1 = \alpha_2 = \frac{1}{2}$, and examine if $V\{1, e^{2i\omega}\} = H_2$. Any function $h(\cdot)$ finite at ω_1 and ω_2 now belongs to H_2, and

$$\begin{aligned} 2\| h(\omega) - c_0 - c_1 e^{2i\omega} \|^2 \\ = | h(\omega_1) - c_0 + c_1 |^2 + | h(\omega_2) - c_0 + c_1 |^2 \end{aligned} \tag{6.5}$$

so that the norm cannot be brought to zero for any c_0, c_1 if $h(\omega_1) \neq h(\omega_2)$.

More generally, $\| h(\omega) - \sum_0^{N-1} c_k e^{i\omega\tau_k} \| = 0$ means that

$$\sum_{k=0}^{N-1} c_k g_{jk} = h(\omega_j) \qquad j = 1, 2, \cdots, N \tag{6.6}$$

where $g_{jk} = e^{i\omega_j \tau_k}$. It is clear from (6.6) that $\{e^{i\omega\tau_k}, k = 0, 1, \cdots, N - 1\}$ is a base if and only if the N by N matrix $[g_{jk}]$ is nonsingular.

VII. APPROXIMATION THEORY

When it becomes necessary to restore $x(t)$ from a given set of samples $x(\tau_1)$, $x(\tau_2)$, $\cdots$, the conditions of the preceding theorems often fail. Although perfect restoration may then be impossible, it is nevertheless desirable to reproduce $x(t)$ as closely as possible. If linear combinations of the samples are considered, the estimate of $x(t)$ is of the general form

$$\hat{x}(t) = \text{l.i.m.} \sum c_n(t) x(\tau_n) \tag{7.1}$$

It is logical to think of $x(t) - \hat{x}(t)$ as an error which is to be minimized in some sense. Within the scope of the Hilbert space theory, a mean square error criterion is the obvious choice, that is, the $c_n(t)$ are chosen to minimize $\| x(t) - \hat{x}(t) \|$.

The procedure for performing the above operation is too well known to merit detailed discussion or proofs. In H_2, the optimization consists of projecting $e^{i\omega t}$ on the subspace $V\{e^{i\omega\tau_n}\}$. The $e^{i\omega\tau_n}$ are first orthonormalized by the usual Gram-Schmidt method (Riesz and Nagy, 1955, p. 67), yielding the set $\{h_n(\omega)\}$ orthonormal with respect to $dF(\cdot)$, i.e., having the property $\int_{-\infty}^{\infty} h_m(\omega) h_n^*(\omega)\, dF(\omega) = \delta_{mn}$. Each such $h_n(\cdot)$ is of the form $h_n(\omega) = \sum_1^n a_k e^{i\omega\tau_k}$. The best estimate of $e^{i\omega t}$ is l.i.m. $\sum b_n(t) h_n(\omega)$, where $b_n(t) = (e^{i\omega t}, h_n)$. Corresponding to this optimum is a mean square error of $1 - \sum_0^\infty | b_n(t) |^2$.

Application of the above optimization requires that $x(t)$ be expressed as in (7.1). This is easily accomplished by expanding $b_n(t)$ and $h_n(\omega)$, and returning to H_1 via the isomorphism with H_2. The expansion of $b_n(t)$ yields $b_n(t) = (e^{i\omega t}, \sum_1^n a_k e^{i\omega \tau_k}) = \sum_1^n a_k R(t - \tau_k)$, where $R(\cdot)$ denotes the correlation function of $x(t)$. Then we obtain

$$\hat{x}(t) = \text{l.i.m.} \sum_{j=1}^{N} a_j x(\tau_j) \sum_{n=1}^{N} \sum_{k=1}^{n} a_k R(t - \tau_k) \tag{7.2}$$

which is readily simplified to give

$$\tilde{x}(t) = \text{l.i.m.} \sum_{j=1}^{N} \left[\sum_{k=1}^{N} \alpha_k(j) a_k R(t - \tau_k) \right] a_j x(\tau_j) \tag{7.3}$$

in which $\alpha_k(j) = (N + 1) - \max(j, k)$.

If $y(t)$ is a random process in H_1 corresponding to $h(\omega, t) \in H_2$, the optimum estimate is formed in the same manner as before, except that now $b_n(t) = (h, h_n)$. The mean square error expression for this more general case is computed to be $\| h(\omega, t) \|^2 - \sum_0^\infty | b_n(t) |^2$.

RECEIVED: November 15, 1961

REFERENCES

ACHIESER, N., AND GLASMAN, I. (1954). "Theorie der Linearen Operatoren im Hilbert Raum." Akademie-Verlag, Berlin.

BALAKRISHNAN, A. (1957). A note on the sampling principle for continuous signals. *IRE Trans. on Information Theory* **IT-3**, 143–146.

BEUTLER, F. On stationarity conditions for a certain periodic random process. *J. Math. Anal. and Applic.*

DOOB, J. (1953). "Stochastic Processes." Wiley, New York.

GOLDMAN, S. (1954). "Information Theory." Prentice-Hall, New York.

HALMOS, P. (1950). "Measure Theory." Van Nostrand, New York.

HARDY, G., LITTLEWOOD, J., AND POLYA, G. (1952). "Inequalities," 2nd ed. Cambridge Univ. Press, New York.

LEVINSON, N. (1940). "Gap and Density Theorems." Coll. Publ. No. 26, Am. Math. Soc.

LLOYD, S. (1959). A sampling theorem for stationary (wide sense) stochastic processes. *Trans. Am. Math. Soc.* **92**, 1–12.

NICHOLS, M. H., AND RAUCH, L. L. (1956). "Radio Telemetry." 2nd ed. Wiley, New York.

RICE, S. (1954). Mathematical analysis of random noise. *In* WAX, N. (ed.), "Selected Papers on Noise and Stochastic Processes." Dover, New York.

RIESZ, F., AND NAGY, B. SZ. (1955). "Functional Analysis." Ungar, New York.

SHANNON, C. (1949). Communication in the presence of noise. *Proc. IRE* **37**, 10–21.

TITCHMARCH, E. (1939). "Theory of Functions." 2nd ed. Oxford Univ. Press, London.