Determinable Classes of Channels

R. T. PROSSER* & W. L. ROOT**

Communicated by G. C. ROTA

In the study of communications theory, it is often convenient to describe a *signal* by a complex valued function $x(t)$ of time t , and to identify the total power of this signal with the integral of $|x(t)|^2$, extended over the time interval of interest. The set of all signals of finite total power then forms a Hilbert space $\mathcal{H}$, called here the signal space, with norm defined by

$$||x||^2 = \int |x(t)|^2 dt.$$

We assume here that the signal $x(t)$ is defined for all t , and vanishes whenever t lies outside the time interval of interest.

It is of practical importance to know what properties of a given signal can be inferred from a finite number of measurements. In particular, can the given signal be distinguished by a finite number of measurements from a host of competing possibilities? A little thought shows that, unless some restriction is placed on the class of competing possibilities, the task of identifying any one of them by a finite number of measurements is hopeless. In fact, for every finite set of measurements there are always pairs of orthogonal signals in $\mathcal{H}$ which yield identical values in each of the measurements in the set.

In the first section of this paper we undertake to describe the kind of limitation which must be imposed *a priori* on a class of signals in order that every signal in the class be distinguishable to within a given mean-square error by a finite number of measurements. Classes of signals which have this property we shall call *determinable classes*.

We first show that every bounded determinable class of signals necessarily has a compact closure in the signal space $\mathcal{H}$, and conversely, that every class of signals with compact closure in $\mathcal{H}$ is a bounded determinable class. Using this result, we then give various alternative characterizations of bounded determinable classes. In particular, every bounded class of signals whose Fourier coefficients in a given orthonormal expansion die out sufficiently fast is a deter-

* Supported at Lincoln Laboratory, which is operated with support from U. S. Air Force.

** Supported in part by the National Aeronautics and Space Administration under research Grant NSG-2-59.

minable class. Moreover, every bounded class of signals whose "tails" in both the time and frequency domains die out sufficiently fast is also a determinable class. Conversely, every bounded determinable class of signals must have these properties. Finally, we note that every bounded class of signals is determinable if and only if its ϵ -entropy and ϵ -capacity, as defined by Kolmogorov [2], are finite for each $\epsilon > 0$.

Within this framework it is also convenient to describe a *linear channel* as a bounded linear operator acting on $\mathcal{H}$, taking (input) signals into (output) signals. In many situations of interest in communications theory, this operator can be represented or closely approximated by an operator of Hilbert-Schmidt class, *i.e.*, by an operator of the particular form

$$y(t) = \int k(t, s)x(s) ds,$$

where the kernel $k(t, s)$ satisfies the requirement

$$\iint |k(t, s)|^2 ds dt < \infty.$$

For such situations it is of practical interest to know what properties of the channel can be inferred by passing a given input signal through the channel and making a finite number of measurements on the resulting output signal. In particular, is it possible to distinguish the elements of a given class of channels by such a process? Again a little thought shows that unless some restriction is placed on the class of channels, the task of distinguishing them by such a process is hopeless. In fact, there are always pairs of orthogonal channels which yield identical values in the measuring process.

In the second section of this paper we undertake to describe the kind of limitation which must be imposed *a priori* on a class of linear channels in order that each channel in the class be distinguishable to within a given mean-square error by introducing an input signal into the channel and making a finite number of measurements on the resulting output signal. Classes of linear channels which have this property we shall also call *determinable classes*.

The problem of identifying determinable classes of channels is more complicated than the corresponding problem for determinable classes of signals. We first show that every bounded determinable class of channels has a compact closure in the channel space $\mathcal{K}$ of all linear Hilbert-Schmidt channels. Because of the asymmetry between input and output signals in the definition of determinability, however, the converse is false. A class of channels whose closure is compact in the channel space $\mathcal{K}$ is determinable only if it satisfies an additional restriction. This restriction, roughly speaking, reduces the number of degrees of freedom in the class enough to ensure that a single input signal will distinguish between all pairs of channels in the class to within a given mean-square error. We present this additional restriction in various equivalent forms. We then give various characterizations of determinable classes of linear channels which

are analogous to those given previously for determinable classes of signals. In particular, we show that the class of "underspread" channels is a determinable class, thereby formalizing a result of Kailath [1].

It is also possible to extend these results to cover certain classes of *non-linear channels*, i.e., certain classes of non-linear operators acting on the signal space $\mathcal{H}$. In particular, these classes include homogeneous polynomial operators of the form

$$y(t) = \iint \cdots \int k(t, s_1, s_2, \cdots, s_n) x(s_1) x(s_2) \cdots x(s_n) ds_1 ds_2 \cdots ds_n,$$

where the kernel $k(t, s_1, s_2, \cdots, s_n)$ satisfies the requirement

$$\iint \cdots \int |k(t, s_1, s_2, \cdots, s_n)|^2 ds_1 ds_2 \cdots ds_n dt < \infty.$$

In addition these classes include all operators which can be approximated in the mean-square sense by finite linear combinations of homogeneous polynomial operators.

Again it is of interest to know what properties of the channel can be inferred by passing a given input signal through the channel and making a finite number of measurements on the resulting output signal.

In the final section of this paper we undertake to describe the kind of limitation which must be imposed *a priori* on a class of non-linear channels in order that each channel in the class be distinguishable to within a given mean-square error by such a measuring process. Classes of non-linear channels which have this property we shall also call *determinable classes*.

We first show that every bounded determinable class of channels has a compact closure in the channel space $\mathcal{H}$ of all non-linear Hilbert-Schmidt channels. We then show that, as in the linear case, a class of channels whose closure is compact in the channel space $\mathcal{H}$ is determinable only if it satisfies an additional restriction limiting the number of degrees of freedom in the class. We then give various characterizations of determinable classes which are direct extensions of the results obtained for the linear case.

1. Determinable Classes of Signals. Let $\mathcal{C}$ be a class of signals, regarded as a subset of the signal space $\mathcal{H}$. We shall always assume that $\mathcal{C}$ is *bounded*, in the sense that if the signal x lies in $\mathcal{C}$, then the total power $\|x\|^2$ of x is no greater than some constant independent of x . We shall further assume, without loss of generality, that this constant is equal to one, so that $\mathcal{C}$ is contained in the unit sphere of $\mathcal{H}$.

By a *measurement* on $\mathcal{C}$ we shall mean a functional m defined on $\mathcal{C}$, which assigns to each signal x in $\mathcal{C}$ a complex number $m(x)$. We shall always assume that $m(x)$ depends continuously on x , and that $|m(x)|$ is always no greater than another constant independent of x . We shall also assume, again without loss of generality, that this constant is equal to one.

In cases of practical interest the communications engineer constructs, on the basis of his knowledge of the class $\mathcal{C}$, a finite sequence of measurements $(m_1, m_2, \dots, m_n)$, and tries to determine the identity of a given signal x from the results $(m_1(x), m_2(x), \dots, m_n(x))$ of the sequence of measurements and the knowledge that x lies in $\mathcal{C}$. This process involves the introduction of an estimator function, which assigns to each finite sequence of complex numbers a signal in $\mathcal{H}$.

By a *signal estimator function*, then, we shall mean a function $\hat{x}$, defined on the space of all n -tuples of complex numbers, which assigns to each such n -tuple $(\xi_1, \xi_2, \dots, \xi_n)$ a signal $\hat{x}(\xi_1, \xi_2, \dots, \xi_n)$ in the signal space $\mathcal{H}$. We shall always require that the signal $\hat{x}(\xi_1, \xi_2, \dots, \xi_n)$ depend continuously on the n -tuple $(\xi_1, \xi_2, \dots, \xi_n)$.

Finally, we shall say that the class $\mathcal{C}$ of signals is *determinable* if, for every $\epsilon > 0$ there is an n -tuple of measurements $(m_1, m_2, \dots, m_n)$ and an associated estimator function $\hat{x}$, such that for every signal x in $\mathcal{C}$ we have

$$(1.1) \quad \|x - \hat{x}(m_1(x), m_2(x), \dots, m_n(x))\| < \epsilon.$$

This statement can be made more succinct by introducing the measuring process $\hat{m}: \mathcal{C} \rightarrow C^n$, defined by $\hat{m}(x) = (m_1(x), m_2(x), \dots, m_n(x))$. Then the relation (1.1) becomes

$$(1.2) \quad \|x - \hat{x} \circ \hat{m}(x)\| < \epsilon,$$

which says that the measuring process $\hat{m}$ has an approximate inverse in the estimator $\hat{x}$, at least for the class $\mathcal{C}$.

We now proceed to investigate the properties of those classes of signals which are determinable.

Lemma 1.1. *Let $\mathcal{C}$ be a determinable class of signals. Then the closure of $\mathcal{C}$ in the signal space $\mathcal{H}$ is compact.*

Proof. We shall show that $\mathcal{C}$ is totally bounded, i.e., that given any $\epsilon > 0$, there are a finite number of spheres of radius ϵ in $\mathcal{H}$ which together cover $\mathcal{C}$.

In fact, given $\epsilon > 0$, we choose a measuring process $\hat{m}$ and an estimator $\hat{x}$ suitable for an error of size $\epsilon/2$. Then for all x in $\mathcal{C}$, we have $\|x - \hat{x} \circ \hat{m}(x)\| < \epsilon/2$. Moreover, the range of $\hat{m}$ is bounded, and hence compact in C^n . Since $\hat{x}: C^n \rightarrow \mathcal{H}$ is continuous, it follows that the range $\hat{\mathcal{C}}$ of $\hat{x} \circ \hat{m}$ is compact in $\mathcal{H}$. This means that $\hat{\mathcal{C}}$ may be covered by a finite number of spheres of radius $\epsilon/2$ in $\mathcal{H}$. By doubling the radius of each of these spheres, we see that they also cover $\mathcal{C}$.

We now obtain the converse of this result.

Lemma 1.2. *Let $\mathcal{C}$ be any subset of $\mathcal{H}$ whose closure is compact. Then $\mathcal{C}$ is a determinable class.*

Proof. Given $\epsilon > 0$, we must find a measuring process $\hat{m}$ and an estimator $\hat{x}$ such that for all x in $\mathcal{C}$ we have $\|x - \hat{x} \circ \hat{m}(x)\| < \epsilon$.

Since the closure of $\mathcal{C}$ is compact, we can find a finite number of spheres of ra-

dius ϵ in $\mathcal{H}$ which cover $\mathcal{C}$. The centers of these spheres then generate a finite dimensional subspace $\mathfrak{M}$ of $\mathcal{H}$. We now choose an orthonormal basis $(x_1, x_2, \dots, x_n)$ for $\mathfrak{M}$, and define the measurements $(m_1, m_2, \dots, m_n)$ by the formula

$$(1.3) \quad m_i(x) = \xi_i = (x, x_i).$$

We then define an estimator $\hat{x}$ by the formula

$$(1.4) \quad \hat{x}(\xi_1, \xi_2, \dots, \xi_n) = \xi_1 x_1 + \xi_2 x_2 + \dots + \xi_n x_n.$$

It is clear from these definitions that m_i is continuous and bounded on $\mathcal{C}$, and that $\hat{x}$ is continuous on C^n . Moreover, if x is any element of $\mathcal{H}$, then $\|x - \hat{x} \circ \hat{m}(x)\|$ is simply the perpendicular distance from x to $\mathfrak{M}$. If x lies in $\mathcal{C}$, this distance is no greater than ϵ , since spheres of radius ϵ with centers in $\mathfrak{M}$ cover $\mathcal{C}$.

Combining these two results, we have

Theorem 1.3. *A class $\mathcal{C}$ of signals is determinable if and only if its closure in the signal space $\mathcal{H}$ is compact.*

The following corollaries are immediate consequences.

Corollary 1.4. *If the class $\mathcal{C}$ is determinable, then for each $\epsilon > 0$ the associated measuring process $\hat{m}$ and estimator $\hat{x}$ may be chosen in the form (1.3) and (1.4), respectively.*

Corollary 1.5. *The class $\mathcal{C}$ is determinable if and only if every covering of $\mathcal{C}$ with spheres of radius $\epsilon > 0$ admits a finite subcovering.*

Corollary 1.6. *The class $\mathcal{C}$ is determinable if and only if every set of elements in $\mathcal{C}$ whose mutual distances are all greater than $\epsilon > 0$ is finite.*

Corollary 1.7. *The class $\mathcal{C}$ is determinable if and only if every sequence of elements in $\mathcal{C}$ contains a convergent subsequence.*

Corollary 1.8. *The class of all signals whose total power is less than 1 is not determinable.*

Corollary 1.9. *No open subset of $\mathcal{H}$, or of any infinite dimensional subspace of $\mathcal{H}$, is determinable.*

Corollary 1.10. *If $\mathcal{C}_1$ and $\mathcal{C}_2$ are determinable classes of signals, then so are $\mathcal{C}_1 \cap \mathcal{C}_2$, $\mathcal{C}_1 \cup \mathcal{C}_2$, $\mathcal{C}_1 + \mathcal{C}_2$ and the closed circled convex hull $\langle \mathcal{C}_1 + \mathcal{C}_2 \rangle^-$ of $\mathcal{C}_1$ and $\mathcal{C}_2$.*

Corollary 1.11. *If $\mathcal{C}$ is a determinable class of signals, then so is its image $B\mathcal{C}$ under every bounded linear transformation B defined on $\mathcal{H}$.*

All of these corollaries follow from Theorem 1.3 and the familiar properties of compact subsets of Hilbert spaces. For further details, we refer to Riesz-Nagy [6].

We now turn to the problem of characterizing the determinable classes of $\mathcal{H}$

in terms of properties accessible to measurement. Our first result of this kind is the following:

Theorem 1.12. *Let $\{\varphi_n\}$ be any orthonormal basis for $\mathcal{H}$. Then the class $\mathcal{C}$ is determinable if and only if for each $\epsilon > 0$ there exists an integer N such that for all x in $\mathcal{C}$ we have*

$$(1.5) \quad \sum_{n>N} |x_n|^2 < \epsilon^2,$$

where $x_n = (x, \varphi_n)$ is the n^{th} Fourier coefficient of x in its expansion in terms of $\{\varphi_n\}$.

Proof. Suppose first that, given $\epsilon > 0$, there exists an N such that (1.5) holds for every x in $\mathcal{C}$. Then $\mathcal{C}$ is totally bounded. In fact, every element x of $\mathcal{C}$ may be approximated to within ϵ by elements of $\mathcal{H}$ of the form

$$\sum_{n \leq N} x_n \varphi_n$$

with norm ≤ 1 . The set of all such elements is a closed bounded finite-dimensional subset of $\mathcal{H}$, and hence compact. Any covering of this set by spheres of radius ϵ admits a finite subcovering, which, in turn, defines a finite covering of $\mathcal{C}$ by spheres of radius 2ϵ .

Now suppose that, given $\epsilon > 0$, there exists no N such that (1.5) holds for every x in $\mathcal{C}$. Then we may choose a sequence N_i of integers and a sequence $x^{(i)}$ of elements in $\mathcal{C}$ such that

$$\sum_{n>N_i} |x_n^{(i)}|^2 > 4\epsilon^2 \quad \text{but} \quad \sum_{n>N_i+1} |x_n^{(i)}|^2 < \epsilon^2.$$

It follows that for any integers i and j we have $\|x^{(i)} - x^{(j)}\| > \epsilon$. Hence no covering with spheres of radius ϵ can be finite, and $\mathcal{C}$ is not totally bounded.

Corollary 1.13. *Suppose that for each x in $\mathcal{C}$ we have $|x_n| \leq a_n$, with a_n independent of x and $\sum a_n^2 < \infty$. Then $\mathcal{C}$ is a determinable class.*

Our next result says that every determinable class is essentially both time-limited and band-limited. This result provides a sort of continuous analogue of Theorem 1.12.

Theorem 1.14. *The class $\mathcal{C}$ is determinable if and only if for each $\epsilon > 0$ there exist positive numbers T and W such that for all x in $\mathcal{C}$ we have both*

$$(1.6) \quad \int_{|t|>T} |x(t)|^2 dt < \epsilon^2$$

and

$$(1.7) \quad \int_{|\omega|>W} |X(\omega)|^2 d\omega < \epsilon^2.$$

Here $X(\omega)$ denotes the Fourier transform of $x(t)$.

Proof. Suppose first that given $\epsilon > 0$ there exist positive numbers T and W such that (1.6) and (1.7) hold for all x in $\mathfrak{C}$. We shall show that $\mathfrak{C}$ is totally bounded.

We first define the characteristic function χ_T by the formula

$$(1.8) \quad \chi_T(t) = \begin{cases} 1, & \text{if } |t| \leq T, \\ 0, & \text{if } |t| > T. \end{cases}$$

Using this function we then define the operators P_T and Q_W acting on $\mathfrak{H}$ according to the formulas

$$(1.9) \quad (P_T x)(t) = \chi_T(t)x(t)$$

and

$$(1.10) \quad (Q_W X)(\omega) = \chi_W(\omega)X(\omega).$$

Then P_T and Q_W are both orthogonal projections on $\mathfrak{H}$. Moreover, the product $Q_W P_T Q_W$ has the form

$$(1.11) \quad (Q_W P_T Q_W)X(\omega) = \int \chi_W(\omega) \hat{\chi}_T(\omega - \omega') \chi_W(\omega') X(\omega') d\omega'.$$

Here $\hat{\chi}_T(\omega)$ denotes the Fourier transform of $\chi_T(t)$.

Since $\chi_T(t)$ is square-integrable in t , we know that $\hat{\chi}_T(\omega)$ is also square-integrable in ω . It follows that the kernel $K(\omega, \omega') = \chi_W(\omega) \hat{\chi}_T(\omega - \omega') \chi_W(\omega')$ is square-integrable over the (ω, ω') plane, and hence that the product $Q_W P_T Q_W$ is a Hilbert-Schmidt operator on $\mathfrak{H}$. Hence the image $\mathfrak{E}$ of the unit sphere in $\mathfrak{H}$ under the product $Q_W P_T Q_W$ is compact in $\mathfrak{H}$. (See Riesz-Nagy [6].)

Moreover, every element x in $\mathfrak{C}$ may be approximated to within 3ϵ by an element in $\mathfrak{E}$, namely $Q_W P_T Q_W x$. In fact, we have

$$(1.12) \quad \begin{aligned} & \|x - Q_W P_T Q_W x\| \\ & \leq \|(I - Q_W)x\| + \|Q_W(I - P_T)x\| + \|Q_W P_T(I - Q_W)x\| \\ & \leq \|(I - Q_W)x\| + \|(I - P_T)x\| + \|(I - Q_W)x\|. \end{aligned}$$

Under the hypotheses (1.6) and (1.7), each of the last three terms here is less than ϵ .

Since $\mathfrak{E}$ is compact, it admits a finite covering with spheres of radius ϵ . By quadrupling the radius of each of these spheres, we see that they now cover $\mathfrak{C}$ as well. It follows that $\mathfrak{C}$ is totally bounded, as required.

Conversely, suppose that, given $\epsilon > 0$, there exists no positive number T such that (1.6) holds for every x in $\mathfrak{C}$. Then we may choose a sequence T_i of integers and a sequence $x^{(i)}$ of elements of $\mathfrak{C}$ such that

$$\int_{|t| > T_i} |x^{(i)}(t)|^2 dt > 4\epsilon^2, \quad \text{but} \quad \int_{|t| > T_i + 1} |x^{(i)}(t)|^2 dt < \epsilon^2.$$

It follows that for any integers i and j we have $\|x^{(i)} - x^{(j)}\| > \epsilon$. Hence no covering of $\mathcal{C}$ by spheres of radius ϵ can be finite. A similar argument holds if there exists no W such that (1.7) holds for all x in $\mathcal{C}$. In either case, $\mathcal{C}$ is not determinable.

The Hilbert-Schmidt property of the product $Q_W P_T Q_W$ forms an essential part of the work of Landau and Pollak on the dimension of the space of essentially time- and band-limited signals [4]. We are indebted to H. J. Landau for suggesting the proof of Theorem 1.14.

The following corollary is the continuous analogue of Corollary 1.13.

Corollary 1.15. *Suppose that for each x in $\mathcal{C}$ we have $|x(t)| \leq f(t)$ and also $|X(\omega)| \leq G(\omega)$, where $f(t)$ and $G(\omega)$ are functions independent of x , with both $\int f(t)^2 dt < \infty$ and $\int G(\omega)^2 d\omega < \infty$. Then $\mathcal{C}$ is a determinable class.*

The identification of determinable classes of signals with relatively compact subsets of the signal space enables us to introduce the notions of ϵ -entropy and ϵ -capacity for these classes. These notions were first proposed by Kolmogorov [2] in order to characterize the "massiveness" of the subsets of a compact metric space. We hope that they will enable us to translate some of the results of information theory into the present framework. For a detailed discussion of these notions, we refer to the expository articles of Kolmogorov and Tihomirov [3] and Lorentz [5]. See also Prosser [9].

Let $\mathcal{C}$ be any determinable class of signals. For each $\epsilon > 0$, let $N_\epsilon(\mathcal{C})$ be the least integer N such that there exists a set of N spheres of radius ϵ which cover $\mathcal{C}$. Similarly, let $M_\epsilon(\mathcal{C})$ be the greatest integer M such that there exists a set of M disjoint spheres of radius $\epsilon/2$ with centers in $\mathcal{C}$. The ϵ -entropy $H_\epsilon(\mathcal{C})$ of $\mathcal{C}$ is the logarithm of $N_\epsilon(\mathcal{C})$:

$$(1.13) \quad H_\epsilon(\mathcal{C}) = \log N_\epsilon(\mathcal{C})$$

and the ϵ -capacity $C_\epsilon(\mathcal{C})$ is the logarithm of $M_\epsilon(\mathcal{C})$:

$$(1.14) \quad C_\epsilon(\mathcal{C}) = \log M_\epsilon(\mathcal{C}).$$

The content of Theorem 1.3 may now be restated as follows:

Theorem 1.16. *A class $\mathcal{C}$ of signals is determinable if and only if either (and hence both) the ϵ -entropy or the ϵ -capacity of $\mathcal{C}$ is finite for all $\epsilon > 0$.*

2. Determinable Classes of Linear Channels. By a *linear channel* we shall mean here a linear operator acting on the signal space $\mathcal{H}$. We shall always assume that the operator is bounded and of Hilbert-Schmidt type. This means that the operator can always be represented as an integral operator

$$(2.1) \quad y(t) = \int k(t, s)x(s) ds$$

whose kernel is square integrable:

$$(2.2) \quad ||k||^2 = \iint |k(t, s)|^2 ds dt < \infty.$$

We assume here that the kernel $k(t, s)$ is defined for all values of s and t , and vanishes whenever s and t lie outside the time interval of interest.

It is known that operators of this type transform elements of $\mathfrak{H}$ in such a way that

$$||y|| \leq ||k|| ||x||,$$

and that sums, scalar products and compositions of such operators are again operators of the same form (see Riesz–Nagy [6]). Although not every operator of interest in communication theory is of this form, it is usually possible to approximate operators of interest by operators of this form, in such a way that the results obtained here approximate the actual situation.

The class of all Hilbert–Schmidt channels of the form (2.1)–(2.2) itself forms a Hilbert space $\mathfrak{K}$, called here the linear channel space, with norm given by (2.2). Moreover, if $\{\varphi_i\}$ is an orthonormal basis for the signal space $\mathfrak{H}$, then the set $\{\varphi_{ii}\}$ where

$$(2.3) \quad \varphi_{ii}(t, s) = \varphi_i(t)\overline{\varphi_i(s)}$$

forms an orthonormal basis for the channel space $\mathfrak{K}$. In this sense $\mathfrak{K}$ may be identified with the tensor product $\mathfrak{H} \otimes \mathfrak{H}$.

Now let $\mathfrak{D}$ be a class of channels, regarded as a subset of the channel space $\mathfrak{K}$. We shall always assume that $\mathfrak{D}$ is bounded, in the sense that if k lies in $\mathfrak{D}$ then its norm $||k||$ is no greater than a constant independent of k . We shall further assume, without loss of generality, that this constant is equal to 1, so that $\mathfrak{D}$ is contained in the unit sphere of $\mathfrak{K}$.

By a *measurement* on $\mathfrak{D}$, we shall mean a functional m , defined on $\mathfrak{D}$, which assigns to each channel k in $\mathfrak{D}$ a complex number $m(k)$. We shall always assume that $m(k)$ depends continuously on k , and that $|m(k)|$ is always no greater than another constant independent of k . We shall also assume, again without loss of generality, that this constant is equal to 1.

By a *channel estimator function*, we shall mean a continuous function $\hat{k}$, defined on the space of all n -tuples of complex numbers, which assigns to $(\xi_1, \xi_2, \dots, \xi_n)$ an element $\hat{k}(\xi_1, \xi_2, \dots, \xi_n)$ of the channel space $\mathfrak{K}$.

In order to clarify the difficulties which arise in determining channels, we shall introduce here two distinct notions of determinability. We shall say first, in direct analogy with the work of the preceding section, that the class $\mathfrak{D}$ is *determinable in principle* if, for every $\epsilon > 0$, there is an n -tuple of measurements $(m_1, m_2, \dots, m_n)$, together with a channel estimator function $\hat{k}$, such that for every channel k in $\mathfrak{D}$ we have

$$(2.4) \quad ||k - \hat{k}(m_1(k), m_2(k), \dots, m_n(k))|| < \epsilon.$$

If we introduce the *measuring process* $\hat{m}: \mathfrak{D} \rightarrow C^n$ defined by

$$m(k) = (m_1(k), m_2(k), \dots, m_n(k))$$

then the requirement (2.4) becomes

$$(2.5) \quad ||k - \hat{k} \circ \hat{m}(k)|| < \epsilon.$$

In practice, measurements of channels are usually made by introducing into the channel an input signal x and performing a measurement in the sense of the last section on the output signal $y = kx$. In this case we shall write $m(kx)$ in place of $m(k)$. We shall see below that it suffices for our purposes to consider only measurements of this particular form.

In situations where the channel is periodic, or can be reset at will to an initial state, it is then possible to measure the channel by introducing a sequence of input signals and measuring the resulting output signals. In other situations, however, where the channel is neither periodic nor resettable, it is then possible to introduce only one input signal into the channel. Indeed, in this case the complete record of everything introduced into the channel constitutes a single input signal. The physical limitation that only a single input signal may ever be introduced into the channel places an essential restriction on the kind of measuring process available.

To take account of this limitation we introduce here a second notion of determinability. We shall say that a class $\mathfrak{D}$ of channels is *determinable in practice* if, for every $\epsilon > 0$, there is a measuring process $\hat{m}$ of the restricted form $\hat{m}(kx) = (m_1(kx), m_2(kx), \dots, m_n(kx))$, where x is a single input signal, and a channel estimator function $\hat{k}$ such that, for every channel k in $\mathfrak{D}$, we have

$$(2.6) \quad ||k - \hat{k} \circ \hat{m}(kx)|| < \epsilon.$$

Since these restricted measurements are in fact measurements in the general sense, it follows that every class of channels which is determinable in practice is also determinable in principle. We shall see below, however, that the converse is false: there exist classes of channels which are determinable in principle, but not determinable in practice.

We now proceed to determine the properties of those classes of channels which are determinable in principle.

Lemma 2.1. *Let $\mathfrak{D}$ be a class of channels which is determinable in principle. Then the closure of $\mathfrak{D}$ in the channel space $\mathcal{K}$ is compact.*

Proof. The proof is entirely analogous to that of Lemma 1.1 and will not be repeated here.

Conversely, we have

Lemma 2.2. *Let $\mathfrak{D}$ be any subset of $\mathcal{K}$ whose closure is compact. Then $\mathfrak{D}$ is determinable in principle.*

Proof. Given $\epsilon > 0$, we must find a measuring process $\hat{m}$ and an estimator function $\hat{k}$ such that for all k in $\mathfrak{D}$ we have $||k - \hat{k} \circ \hat{m}(k)|| < \epsilon$.

Let us first choose an orthonormal basis $\{\varphi_i(t)\}$ for the signal space $\mathcal{H}$. Then we know that the set $\{\varphi_{i,i}(t, s)\}$, where $\varphi_{i,i}(t, s) = \varphi_i(t)\varphi_i(s)$, forms an orthonormal basis for the channel space $\mathcal{K}$. This means that finite linear combinations of members of this basis are dense in $\mathcal{K}$. Hence the family of spheres of radius ϵ with centers at all finite linear combinations of these basis elements must cover $\mathcal{K}$. In particular, they cover $\mathcal{D}$. Since the closure of $\mathcal{D}$ is compact, this covering admits a finite subcovering. Hence there is a finite orthonormal set of signals $\{\varphi_i(t)\}_{i=1}^n$ such that linear combinations of elements of the form $\{\varphi_{i,i}(t, s)\}_{i=1}^n$ approximate every element of $\mathcal{D}$ to within ϵ .

We now define the measurement $m_{i,i}$ by the formula

$$(2.7) \quad m_{i,i}(k) = \xi_{i,i} = (k, \varphi_{i,i}).$$

Together the $m_{i,i}$ make up the measuring process $\hat{m}$. We then define the estimator function $\hat{k}$ by the formula

$$(2.8) \quad \hat{k}(\xi_{i,i}) = \sum_{i,i=1}^n \xi_{i,i} \varphi_{i,i}.$$

It is clear from these definitions that each measurement $m_{i,i}$ is bounded and continuous on $\mathcal{D}$, and that $\hat{k}$ is continuous on C^{n^2} . Moreover, if k is any element of $\mathcal{K}$ then $\|k - \hat{k} \circ \hat{m}(k)\|$ is simply the perpendicular distance from k to the n^2 -dimensional subspace generated by the $\{\varphi_{i,i}\}_{i,i=1}^n$. If k lies in $\mathcal{D}$, then this distance is no greater than ϵ , because of our choice of subspace. Hence $\mathcal{D}$ is determinable in principle, and in fact, is determinable by using a finite number of input signals and making a finite number of linear measurements on each of the resulting outputs.

Combining these results, we obtain

Theorem 2.3. *A class $\mathcal{D}$ of linear channels is determinable in principle if and only if its closure in the channel space $\mathcal{K}$ is compact.*

The following corollaries are immediate consequences.

Corollary 2.4. *If the class $\mathcal{D}$ is determinable in principle, then for each $\epsilon > 0$ the associated measuring process $\hat{m}$ and estimator function $\hat{k}$ may be chosen in the form (2.7) and (2.8), respectively.*

Corollary 2.5. *The analogues of Corollaries 1.5–1.10 all hold for classes of channels which are determinable in principle.*

In place of Corollary 1.11, we have

Corollary 2.6. *If the class $\mathcal{D}$ is determinable in principle, then so is the class $A\mathcal{D}B$, for every pair, A, B of bounded linear operators acting on the signal space $\mathcal{H}$.*

The following analogue of Theorem 1.12 also holds.

Theorem 2.7. *Let $\{\varphi_n\}$ be any orthonormal basis for the signal space $\mathcal{H}$, and $\{\varphi_{i,i}\}$, where $\varphi_{i,i} = \phi_i \bar{\phi}_i$, the corresponding basis for the channel space $\mathcal{K}$. Then the*

class $\mathfrak{D}$ is determinable in principle if and only if, for every $\epsilon > 0$, there exists an integer N such that for all k in $\mathfrak{D}$ we have

$$(2.9) \quad \sum_{S(N)'} |k_{i,i}|^2 < \epsilon^2,$$

where $k_{i,i} = (k, \varphi_{i,i})$ is the i - j^{th} Fourier coefficient of k in its expansion in terms of $\{\varphi_{i,i}\}$, and $S(N)'$ denotes the complement of the set $S(N)$ of pairs $\{i, j\}$ such that both $i < N$ and $j < N$.

The proof is the exact analogue of that of Theorem 1.12, and will not be repeated here.

Corollary 2.8. Suppose that for each k in $\mathfrak{D}$ we have $|k_{i,i}| < b_{i,i}$, where $b_{i,i}$ is independent of k and $\sum |b_{i,i}|^2 < \infty$. Then the class $\mathfrak{D}$ is determinable in principle.

The analogue of Theorem 1.14 also holds.

Theorem 2.9. The class $\mathfrak{D}$ is determinable in principle if and only if for each $\epsilon > 0$ there exist positive numbers T and W such that for all k in $\mathfrak{D}$ we have both

$$(2.10) \quad \iint_{E(T)'} |k(t, s)|^2 ds dt < \epsilon^2$$

and

$$(2.11) \quad \iint_{F(W)'} |K(\omega, \nu)|^2 d\nu d\omega < \epsilon^2.$$

Here $K(\omega, \nu)$ denotes the double Fourier transform of $k(t, s)$, and $E(T)'$ and $F(W)'$ denote the regions exterior to the squares $E(T)$ and $F(W)$ defined by $|s| < T$, $|t| < T$, and $|\nu| < W$, $|\omega| < W$, respectively.

Proof. Suppose first that, given $\epsilon > 0$, we can find positive numbers T and W such that (2.10) and (2.11) hold for all k in $\mathfrak{D}$. We shall show that $\mathfrak{D}$ is then totally bounded.

As in the proof of Theorem 1.14, we first define the characteristic function $\chi_T(t, s)$ by the formula

$$(2.12) \quad \chi_T(t, s) = \begin{cases} 1, & \text{if both } |s| \leq T \text{ and } |t| \leq T, \\ 0, & \text{otherwise.} \end{cases}$$

Using this function we then define the operators P_T and Q_W acting on $\mathcal{K}$ according to the formulas

$$(2.13) \quad (P_T k)(t, s) = \chi_T(t, s) k(t, s)$$

and

$$(2.14) \quad (Q_W K)(\omega, \nu) = \chi_W(\omega, \nu) K(\omega, \nu).$$

Then P_T and Q_W are both orthogonal projections on $\mathfrak{H}$. Moreover, an argument entirely similar to the one used in the proof of Theorem 1.14 now shows that the product $Q_W P_T Q_W$ is also a Hilbert–Schmidt operator on $\mathfrak{H}$. It follows that the image $\mathfrak{F}$ of the unit sphere in $\mathfrak{H}$ under the product $Q_W P_T Q_W$ is compact in $\mathfrak{H}$.

Moreover, every element k in $\mathfrak{D}$ may be approximated to within 3ϵ by an element in $\mathfrak{F}$, namely $Q_W P_T Q_W k$, just as in the proof of Theorem 1.14. Since $\mathfrak{F}$ is compact, it admits a finite covering with spheres of radius ϵ . By quadrupling the radius of each of these spheres, we see that they now cover $\mathfrak{D}$ as well. It follows that $\mathfrak{D}$ is totally bounded, as required.

Conversely, suppose that, given $\epsilon > 0$, there exists no positive number S such that (2.10) holds for every k in $\mathfrak{D}$. Then we may choose a sequence T_i of integers and a sequence $k^{(i)}$ of elements of $\mathfrak{D}$ such that

$$\iint_{E(T_i)'} |k^{(i)}(t, s)|^2 ds dt > 4\epsilon^2$$

but

$$\iint_{E(T_i + 1)'} |k^{(i)}(t, s)|^2 ds dt < \epsilon^2.$$

It follows that for any integer i and j we have $\|k^i - k^j\| > \epsilon$. Hence no covering of $\mathfrak{D}$ by spheres of radius ϵ can be finite. A similar argument holds if there exists no W such that (2.10) and (2.11) holds for every k in $\mathfrak{D}$. In either case $\mathfrak{D}$ is not determinable.

The following corollary is the analogue of Corollary 1.15.

Corollary 2.10. *Suppose that for each k in $\mathfrak{D}$ we have $|k(t, s)| \leq f(t, s)$ and also $|K(\omega, \nu)| \leq G(\omega, \nu)$, where $f(t, s)$ and $G(\omega, \nu)$ are functions independent of k , with both $\iint f(t, s) ds dt < \infty$ and $\iint G(\omega, \nu)^2 d\nu d\omega < \infty$. Then $\mathfrak{D}$ is determinable in principle.*

We now terminate our discussion of classes of channels which are determinable in principle, and turn to the problem of characterizing those classes which are determinable in practice.

We first observe that every class of channels which is determinable in practice is already determinable in principle. It follows immediately that all of the properties which characterize classes determinable in principle are automatically *necessary* for classes determinable in practice. They are, however, *not sufficient*, as the following example shows.

Lemma 2.11. *Let P_N be any orthogonal projection on the signal space $\mathfrak{H}$ whose range has finite dimension N . Let $\mathfrak{D}$ be the class of all Hilbert–Schmidt transformations on $\mathfrak{H}$ of the form $T = P_N T P_N$, with Hilbert–Schmidt norm no greater than 1. Then $\mathfrak{D}$ is always determinable in principle, but is determinable in practice only if $N = 1$.*

Proof. Since the class $\mathfrak{D}$ is both bounded and finite dimensional (dimension of $\mathfrak{D} = N^2$), it is therefore compact, and hence determinable in principle.

On the other hand, let x be any single input signal, and let P be any non-zero projection such that $P \leq P_N$ and $Px = 0$. Note that we can always find such a projection if $N > 1$. Since $P = P_N P P_N$, we see that some non-zero scalar multiple λP of P will lie in $\mathfrak{D}$. Moreover, no measuring process made on the output signal $y = \lambda Px = 0$ can ever distinguish λP from zero. Hence $\mathfrak{D}$ is not determinable in practice if $N > 1$. If $N = 1$, then $\mathfrak{D}$ consists simply of scalar multiples of P_1 , which can be distinguished by any input signal x not annihilated by P_1 .

The proof suggests that a class $\mathfrak{D}$ of channels cannot be determinable in practice unless there exists a single signal which essentially separates the elements of $\mathfrak{D}$ from each other. The following lemmas are designed to make this intuitive suggestion precise.

Lemma 2.12. *Let $\mathfrak{D}$ be a class of channels which is determinable in practice. Then for every $\epsilon > 0$ there exists a single signal x such that, for every pair $\{h, k\}$ of channels in the closure of $\mathfrak{D}$ with $\|h - k\| > \epsilon$, we have $\|hx - kx\| > 0$.*

Proof. Given $\epsilon > 0$, we know that there exists a single input signal x , a measuring process $\hat{m}$ defined on output signals and a channel estimator $\hat{k}$ such that for every k in $\mathfrak{D}$ we have $\|k - \hat{k} \circ \hat{m}(kx)\| < \epsilon/2$. Now if h and k are two channels in the closure of $\mathfrak{D}$ with $hx = kx$, then $\hat{k} \circ \hat{m}(hx) = \hat{k} \circ \hat{m}(kx)$, and it follows that $\|h - k\| \leq \|h - \hat{k} \circ \hat{m}(hx)\| + \|\hat{k} \circ \hat{m}(kx) + k\| \leq \epsilon/2 + \epsilon/2 = \epsilon$. Hence we see that $\|h - k\| > \epsilon$ implies $\|hx - kx\| > 0$.

The converse is also true.

Lemma 2.13. *Let $\mathfrak{D}$ be a class of channels which is determinable in principle. Suppose moreover that for every $\epsilon > 0$ there exists a single signal x such that for all pairs $\{h, k\}$ of channels in the closure of $\mathfrak{D}$ with $\|h - k\| > \epsilon$ we have $\|hx - kx\| > 0$. Then $\mathfrak{D}$ is determinable in practice.*

Proof. Given $\epsilon > 0$, we choose x in $\mathcal{H}$ so that for all pairs $\{h, k\}$ of channels in $\mathfrak{D}$ with $\|h - k\| > \epsilon$ we have $\|hx - kx\| > 0$.

Now we define the mapping $\pi : \mathfrak{D} \rightarrow \mathcal{H}$, which assigns to each channel k in $\mathfrak{D}$ a signal in $\mathcal{H}$ according to the prescription: $\pi k = kx$. Then π is clearly uniformly continuous from $\mathfrak{D}$ to $\mathcal{H}$, and since $\mathfrak{D}$ has compact closure in $\mathcal{H}$, it follows that the image $\pi\mathfrak{D}$ of $\mathfrak{D}$ has compact closure in $\mathcal{H}$.

Moreover, there exists an $\eta > 0$ such that, for all pairs $\{h, k\}$ of channels in $\mathfrak{D}$ with $\|\pi h - \pi k\| = \|hx - kx\| < 3\eta$, we have $\|h - k\| < 2\epsilon$. For otherwise there exist sequences h_i and k_i of channels in $\mathfrak{D}$ such that $\|\pi h_i - \pi k_i\| = \|h_i x - k_i x\| < 2^{-i}$, but $\|h_i - k_i\| \geq 2\epsilon$. Since $\mathfrak{D}$ has compact closure in $\mathcal{H}$, we know that each of these sequences contains a convergent subsequence, converging, say, to h and k , respectively. Then because π is uniformly continuous from $\mathfrak{D}$ to $\mathcal{H}$, we see that $0 \leq \|\pi h - \pi k\| = \lim \|\pi h_i - \pi k_i\| \leq \lim 2^{-i} = 0$.

But $\|h - k\| = \lim \|h_i - k_i\| \geq 2\epsilon$, which contradicts the assumptions on $\mathfrak{D}$.

Since π is "almost" one-to-one, we expect that it "almost" admits an inverse. In fact, since $\mathfrak{D}$ has compact closure, we can find a finite sequence of channels k_i , such that the spheres $\mathcal{S}_i$ with centers at the k_i and radius ϵ cover $\mathfrak{D}$. Using these spheres we define a mapping $\sigma : \mathfrak{K} \rightarrow \mathfrak{D}$ as follows: If y in $\mathfrak{K}$ lies within distance η of $\pi\mathcal{S}_1$, we put $\sigma y = k_1$. Proceeding by induction, we suppose σ is defined for all y in $\mathfrak{K}$ lying within distance η of $\pi\mathcal{S}_i$ for some $i < n$. If y lies within distance η of $\pi\mathcal{S}_n$ but not within distance η of $\pi\mathcal{S}_i$ for any $i < n$, we define $\sigma y = k_n$. If y lies within distance η of $\pi\mathcal{S}_i$ for no i , we define arbitrarily $\sigma y = 0$.

This procedure defines σ on all of $\mathfrak{K}$ in such a way that σ assigns to the signal y either the channel k_i , where i is the least integer such that y lies within distance η of $\pi\mathcal{S}_i$, or else the channel $k = 0$ if no such integer exists. Moreover, we see that for all k in $\mathfrak{D}$ we have $\|k - \sigma \circ \pi(k)\| < 2\epsilon$, so that σ is an approximate inverse for π . Finally, we note that if y and z lie within distance η of $\pi\mathfrak{D}$ and $\|y - z\| < \eta$, then $\|\sigma y - \sigma z\| < 4\epsilon$.

Now the use of the mappings π and σ enable us to reduce the problem involving channels to an analogous problem involving signals. In fact, since $\pi\mathfrak{D}$ has compact closure in $\mathfrak{K}$, we know from the results of the preceding section that $\pi\mathfrak{D}$ is a determinable class of signals. This means that there exists a measuring process $\hat{m}$ and a signal estimator function $\hat{y}$ such that for all y in $\pi\mathfrak{D}$ we have $\|y - \hat{y} \circ \hat{m}(y)\| < \eta$. Moreover, since $\hat{y}$ is continuous from C^n to $\mathfrak{K}$, we know that there exists a $\delta > 0$ such that if $\|\xi - \xi'\| < \delta$ then $\|\hat{y}(\xi) - \hat{y}(\xi')\| < \eta$.

$$\mathfrak{D} \xrightarrow[\sigma]{\pi} \mathfrak{K} \xrightarrow[\hat{y}]{\hat{m}} C^n.$$

Now we choose the function $\hat{m}$ as our required measuring process. In order to produce a channel estimator function, we first compose the signal estimator function $\hat{y}$ with the mapping $\sigma : \hat{h} = \sigma \circ \hat{y}$. By combining the various estimates involved, we find that for every channel k in $\mathfrak{D}$ we have $\|k - \hat{h} \circ \hat{m}(kx)\| = \|k - \sigma \circ \hat{y} \circ \hat{m}(kx)\| < 4\epsilon$.

Hence the function $\hat{h} : C^n \rightarrow \mathfrak{K}$ has all the properties required of a channel estimator function except continuity. In order to produce a continuous channel estimator, we have only to smooth out the function $\hat{h}$ as follows:

Let $\varphi(\xi)$ be any non-negative continuous function defined on C^n , such that $\varphi(\xi) = 0$ if $\|\xi\| > \delta$ and $\int_{C^n} \varphi(\xi) d\xi = 1$. Using $\varphi(\xi)$, we define

$$(2.15) \quad \hat{k}(\xi) = \int_{C^n} \hat{h}(\xi - \xi') \varphi(\xi') d\xi'.$$

It is easy to verify that the right hand side defines, as a strong limit of finite sums, a continuous function $\hat{k}$ from C^n to $\mathfrak{K}$. Moreover, we have

$$\|\hat{k}(\xi) - \hat{h}(\xi)\| \leq \int_{\|\xi'\| < \delta} \|\hat{h}(\xi - \xi') - \hat{h}(\xi)\| \varphi(\xi') d\xi'$$

$$\begin{aligned} &\leq 4\epsilon \int_{|\xi'| < \delta} \varphi(\xi') d\xi' \\ &\leq 4\epsilon. \end{aligned}$$

It follows that for all k in $\mathfrak{D}$ we have

$$\begin{aligned} \|k - \hat{k} \circ \hat{m}(kx)\| &\leq \|\hat{k} - \hat{h} \circ \hat{m}(kx)\| + \|\hat{h} \circ \hat{m}(kx) - \hat{k} \circ \hat{m}(kx)\| \\ &\leq 4\epsilon + 4\epsilon \\ &\leq 8\epsilon. \end{aligned}$$

Thus m and k have all the properties of a measuring process and channel estimator required for the definition of determinability in practice. We conclude that under the assumptions of Lemma 2.13 the class $\mathfrak{D}$ is determinable in practice.

The following corollary gives a sometimes useful sufficient condition.

Corollary 2.14. *Let $\mathfrak{D}$ be a class of channels which is determinable in principle. Suppose moreover that there exists a single signal x such that for all pairs $\{h, k\}$ of channels in $\mathfrak{D}$ we have $hx = kx$ implies $h = k$. Then $\mathfrak{D}$ is determinable in practice.*

Combining Lemmas 2.12 and 2.13, we obtain a characterization of determinability in practice.

Theorem 2.15. *Let $\mathfrak{D}$ be a class of channels which is determinable in principle. Then $\mathfrak{D}$ is determinable in practice if and only if, for each $\epsilon > 0$, there exists a single signal x in $\mathfrak{H}$ such that, for all pairs $\{h, k\}$ of channels in the closure of $\mathfrak{D}$, $\|h - k\| > \epsilon$ implies $\|hx - kx\| > 0$.*

Corollary 2.16. *If $\mathfrak{D}$ is any class of channels which is determinable in practice, then so is the closure of $\mathfrak{D}$ together with all of its subsets, and so is every class of the form $A\mathfrak{D}B$, where A and B are bounded invertible operators acting on $\mathfrak{H}$.*

Remark. It is not necessarily true that the convex hull of $\mathfrak{D}$ is determinable in practice, nor that the sum of two such classes is determinable in practice. In these respects the differences between principle and practice are placed in evidence.

Other sufficient conditions now follow from Theorem 2.15.

Lemma 2.17. *Let $\mathfrak{D}$ be a class of channels which is determinable in principle. Let $\{\varphi_n\}_{n=1}^{\infty}$ be any orthonormal basis of signals in $\mathfrak{H}$ and P_N the projection on the subspace spanned by $\{\varphi_n\}_{n=1}^N$. Suppose that the classes $P_N\mathfrak{D}P_N$ are determinable in practice for each sufficiently large integer N . Then $\mathfrak{D}$ is determinable in practice.*

Proof. According to Theorem 2.7, we can find, given $\epsilon > 0$, an integer M such that for all k in $\mathfrak{D}$ we have $\|k - P_M k P_M\| \leq \epsilon$. Now if N is greater than M and sufficiently large so that the class $P_N\mathfrak{D}P_N$ is determinable in practice, then for some signal y in $\mathfrak{H}$ we have, for all pairs $\{h, k\}$ of channels in $\mathfrak{D}$ with

$\|P_N h P_N - P_N k P_N\| > \epsilon$, that $\|P_N h P_{Ny} - P_N k P_{Ny}\| > 0$. But if $\|h - k\| > 3\epsilon$, then $\|P_N h P_N - P_N k P_N\| > \epsilon$, and so $\|hx - kx\| \geq \|P_N h x - P_N k x\| > 0$, where $x = P_{Ny}$. It now follows from Lemma 2.13 that $\mathfrak{D}$ is determinable in practice.

Lemma 2.18. *Let $\{\varphi_i\}$ be any basis for the signal space $\mathfrak{H}$, and $\{\varphi_{i,i}\}$ the associated basis for the channel space $\mathfrak{K}$. Let $\mathfrak{D}$ be a class of channels such that, for all k in $\mathfrak{D}$, we have $k_{i,i} = 0$ for $i \neq j$, where $k_{i,i}$ are the Fourier coefficients of k in its expansion in terms of the $\varphi_{i,i}$. Suppose moreover, that for each $\epsilon > 0$ there exists an integer N such that for all k in $\mathfrak{D}$ we have $\sum_{i>N} |k_{i,i}|^2 < \epsilon^2$. Then $\mathfrak{D}$ is determinable in practice.*

Proof. The hypotheses imply first that Theorem 2.7 applies, so that $\mathfrak{D}$ is determinable in principle. Now let $x = \sum 2^{-i} \varphi_i$, and note that if k lies in $\mathfrak{D}$, then $kx = \sum 2^{-i} k_{i,i} \varphi_i$. Hence if h and k lie in $\mathfrak{D}$ and $h \neq k$, then $h_{i,i} \neq k_{i,i}$ for some i , and so $hx \neq kx$. According to Corollary 2.14 this means that $\mathfrak{D}$ is determinable in practice.

We conclude this section with the derivation of a result of T. Kailath [1], which states that every "underspread" class of linear channels is determinable in practice. Our derivation requires a precise form of the sampling theorem for linear channels which we present first as a separate theorem. See also Prosser [8].

We begin with a two-dimensional version of the classical Euler summation formula. This version relates the integral of a function of two variables taken over a half-integer domain to the sum of the sample values of the function taken at the integer points in the domain. By a *half-integer domain* we mean the union of a finite number of *squares* $S_{m,n}$ of the form

$$S_{m,n} = \{(t, s) : m - \tfrac{1}{2} \leq s \leq m + \tfrac{1}{2}, n - \tfrac{1}{2} \leq t \leq n + \tfrac{1}{2}\}.$$

Lemma 2.19. *Let $f(t, s)$ be a twice continuously differentiable function defined over the half-integer domain A . Then*

$$(2.16) \quad \sum_{(m,n) \in A} f(m, n) = \iint_A \left(1 + \sigma(t) \frac{\partial}{\partial t}\right) \left(1 + \sigma(s) \frac{\partial}{\partial s}\right) f(t, s) ds dt.$$

Here σ denotes the sawtooth function defined by the formula

$$(2.17) \quad \sigma(z) = z - [z] - \tfrac{1}{2}.$$

Proof. Consider the integral $\iint_B f(t, s) ds$, where B denotes a square of the form $B = \{(t, s) : m \leq s \leq m + 1/2, n \leq t \leq n + 1/2\}$. After integrating by parts once in each variable, we obtain

$$(2.18) \quad \begin{aligned} \iint f(t, s) ds dt &= \tfrac{1}{4} f(m, n) - \tfrac{1}{2} \int_n^{n+1/2} \sigma(t) \frac{\partial}{\partial t} f(t, m) dt \\ &\quad - \tfrac{1}{2} \int_m^{m+1/2} \sigma(s) \frac{\partial}{\partial s} f(n, s) ds + \iint_B \sigma(t) \sigma(s) \frac{\partial^2}{\partial t \partial s} f(t, s) ds dt. \end{aligned}$$

The middle terms here may be combined in a line integral taken over the boundary Γ of B as follows:

$$\begin{aligned}
 & -\frac{1}{2} \int_n^{n+1/2} \sigma(t) \frac{\partial}{\partial t} f(t, m) dt - \frac{1}{2} \int_m^{m+1/2} \sigma(s) \frac{\partial}{\partial s} f(n, s) ds \\
 (2.19) \quad & = \int_{\Gamma} \sigma(t) \sigma(s) \left(\frac{\partial}{\partial s} f(t, s) ds - \frac{\partial}{\partial t} f(t, s) dt \right) \\
 & = - \iint_B \left(\sigma(t) \frac{\partial}{\partial t} + \sigma(s) \frac{\partial}{\partial s} + \sigma(t) \sigma(s) \frac{\partial^2}{\partial s \partial t} + \sigma(t) \sigma(s) \frac{\partial^2}{\partial t \partial s} \right) f(t, s) ds dt.
 \end{aligned}$$

Substituting (2.19) in (2.18), we obtain

$$\begin{aligned}
 (2.20) \quad & \iint_B f(t, s) ds dt = \frac{1}{4} f(n, m) \\
 & - \iint_B \left(\sigma(t) \frac{\partial}{\partial t} + \sigma(s) \frac{\partial}{\partial s} + \sigma(s) \sigma(t) \frac{\partial^2}{\partial t \partial s} \right) f(t, s) ds dt.
 \end{aligned}$$

Similar calculations yield identical results for the squares

$$C = \{(t, s) : m - \frac{1}{2} \leq s \leq m, n \leq t \leq n + \frac{1}{2}\},$$

$$D = \{(t, s) : m \leq s \leq m + \frac{1}{2}, n - \frac{1}{2} \leq t \leq n\}$$

and

$$E = \{(t, s) : m - \frac{1}{2} \leq s \leq m, n - \frac{1}{2} \leq t \leq n\}.$$

Since the square

$$S_{m,n} = \{(t, s) : m - \frac{1}{2} \leq s \leq m + \frac{1}{2}, n - \frac{1}{2} \leq t \leq n + \frac{1}{2}\}$$

is the union of B , C , D , and E , we conclude that

$$\begin{aligned}
 (2.21) \quad & \iint_{S_{m,n}} f(t, s) ds dt = f(m, n) \\
 & - \iint_{S_{m,n}} \left(\sigma(t) \frac{\partial}{\partial t} + \sigma(s) \frac{\partial}{\partial s} + \sigma(t) \sigma(s) \frac{\partial^2}{\partial t \partial s} \right) f(t, s) ds dt.
 \end{aligned}$$

But (2.21) implies (2.16) when $A = S_{m,n}$. Since in general A is a union of squares of the form $S_{m,n}$, we see that (2.16) holds for all such A , as required.

We shall now employ this summation formula to estimate the degree of approximation of a two-dimensional band-limited function by the associated two-dimensional sampling function. For the one-dimensional analogue, we refer to the paper of Landau and Pollak [4].

Theorem 2.20. *Let $k(t, s)$ be any continuously differentiable channel in $\mathcal{K}$. Suppose that for some $\epsilon > 0$*

$$(2.23) \quad \iint_{E'} \left(|k(t, s)|^2 + \left| \frac{\partial}{\partial t} k(t, s) \right|^2 + \left| \frac{\partial}{\partial s} k(t, s) \right|^2 \right) ds dt < \epsilon^2$$

and

$$(2.24) \quad K(\omega, \nu) = 0 \quad \text{unless} \quad |\nu| \leq 2\pi V \quad \text{and} \quad |\omega| < 2\pi W.$$

Here $K(\omega, \nu)$ denotes the complete Fourier transform of $k(t, s)$ and E' denotes the complement of a given half-integer domain E . Then we have

$$(2.25) \quad \left\| k(t, s) - \sum_{(n/2W, m/2V) \in E} k\left(\frac{n}{2W}, \frac{m}{2V}\right) \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)} \frac{\sin \pi(2Vs - m)}{\pi(2Vs - m)} \right\|^2 < \frac{7}{2}\epsilon^2 + \frac{1}{2}\pi^2 \|k\| \epsilon.$$

Proof. Let us first assume that $W = V = 1/2$. We know that the sampling functions form an orthonormal basis for the band-limited functions [4], and hence we know that

$$(2.26) \quad \left\| k(t, s) - \sum_{(m, n) \in E} k(n, m) \frac{\sin \pi(t - n)}{\pi(t - n)} \frac{\sin \pi(s - m)}{\pi(s - m)} \right\|^2 = \sum_{(m, n) \in E'} |k(n, m)|^2.$$

Now let A_N denote the half-integer domain determined by the intersection of E' with the square $\{(t, s): |s| \leq N + 1/2, |t| \leq N + 1/2\}$. Put $f(t, s) = |k(t, s)|^2$, and apply Lemma 2.19:

$$(2.27) \quad \begin{aligned} \sum_{(m, n) \in A_N} |k(n, m)|^2 &= \iint_{A_N} \left(1 + \sigma(t) \frac{\partial}{\partial t} \right) \left(1 + \sigma(s) \frac{\partial}{\partial s} \right) |k(t, s)|^2 ds dt \\ &= \iint_{A_N} |k(t, s)|^2 ds dt \\ &\quad + \iint_{A_N} \sigma(t) 2 \operatorname{Re} \left(k(t, s) \frac{\partial}{\partial t} \bar{k}(t, s) \right) ds dt \\ &\quad + \iint_{A_N} \sigma(s) 2 \operatorname{Re} \left(k(t, s) \frac{\partial}{\partial s} \bar{k}(t, s) \right) ds dt \\ &\quad + \iint_{A_N} \sigma(t) \sigma(s) 2 \operatorname{Re} \left(k(t, s) \frac{\partial^2}{\partial t \partial s} \bar{k}(t, s) \right) ds dt \\ &\quad + \iint_{A_N} \sigma(t) \sigma(s) 2 \operatorname{Re} \left(\frac{\partial}{\partial t} k(t, s) \frac{\partial}{\partial s} \bar{k}(t, s) \right) ds dt. \end{aligned}$$

Each of these terms may be majorized, by using the Schwarz inequality and the fact that $\|\sigma\|_\infty = 1/2$, by the corresponding term in the following equation:

$$\begin{aligned}
\sum_{(m,n) \in A_N} |k(n, m)|^2 &\leq \iint_{A_N} |k(t, s)|^2 ds dt \\
&+ \left(\iint_{A_N} |k(t, s)|^2 ds dt \right)^{1/2} \left(\iint_{A_N} \left| \frac{\partial}{\partial t} k(t, s) \right|^2 ds dt \right)^{1/2} \\
(2.28) \quad &+ \left(\iint_{A_N} |k(t, s)|^2 ds dt \right)^{1/2} \left(\iint_{A_N} \left| \frac{\partial}{\partial s} k(t, s) \right|^2 ds dt \right)^{1/2} \\
&+ \frac{1}{2} \left(\iint_{A_N} |k(t, s)|^2 ds dt \right)^{1/2} \left(\iint_{A_N} \left| \frac{\partial^2}{\partial t \partial s} k(t, s) \right|^2 ds dt \right)^{1/2} \\
&+ \frac{1}{2} \left(\iint_{A_N} \left| \frac{\partial}{\partial t} k(t, s) \right|^2 ds dt \right)^{1/2} \left(\iint_{A_N} \left| \frac{\partial}{\partial s} k(t, s) \right|^2 ds dt \right)^{1/2}.
\end{aligned}$$

From the hypotheses we know that the integrals over A_N of $|k(t, s)|^2$,

$$\left| \frac{\partial}{\partial t} k(t, s) \right|^2 \quad \text{and} \quad \left| \frac{\partial}{\partial s} k(t, s) \right|^2$$

are all less than ϵ^2 . Moreover, the integral over the whole plane of

$$\left| \frac{\partial^2}{\partial t \partial s} k(t, s) \right|^2$$

is dominated via the Plancherel theorem by

$$\int_{-\pi}^{+\pi} \int_{-\pi}^{+\pi} |\omega \nu K(\omega, \nu)|^2 d\nu d\omega,$$

which in turn is dominated by

$$\pi^4 \int_{-\pi}^{+\pi} \int_{-\pi}^{+\pi} |K(\omega, \nu)|^2 d\nu d\omega = \pi^4 \|k\|^2.$$

Combining all these estimates, and taking the limit as $N \rightarrow \infty$, we find that

$$\begin{aligned}
\sum_{(m,n) \in E'} |k(n, m)|^2 &\leq \epsilon^2 + \epsilon^2 + \epsilon^2 + \frac{1}{2}\epsilon\pi^2 \|k\| + \frac{1}{2}\epsilon^2 \\
(2.29) \quad &\leq \frac{7}{2}\epsilon^2 + \frac{1}{2}\pi^2 \|k\| \epsilon.
\end{aligned}$$

Combining now the estimates (2.29) and (2.26), we obtain (2.25) as required.

We have assumed throughout the proof that $V = W = 1/2$. The general case may be reduced to this one by the following change of scale,

$$\begin{aligned}
s &\rightarrow s' = 2Vs, \\
t &\rightarrow t' = 2Wt, \\
(2.30) \quad k(t, s) &\rightarrow k'(t', s') = \left(\frac{1}{4VW} \right)^{1/2} k\left(\frac{t'}{2W}, \frac{s'}{2V} \right),
\end{aligned}$$

together with the following conjugate change of scale,

$$(2.31) \quad \begin{aligned} \nu &\rightarrow \nu' = \frac{\nu}{2V}, \\ \omega &\rightarrow \omega' = \frac{\omega}{2W}, \\ K(\omega, \nu) &\rightarrow K'(\omega', \nu') = (4VW)^{1/2} K(2W\omega', 2V\nu'). \end{aligned}$$

It is straightforward to verify that this change of scale leads to (2.25) in the general case, as required.

Corollary 2.21. *Suppose that instead of (2.24) we have*

$$(2.32) \quad \iint_{F'} (1 + \omega^2 + \nu^2 + \omega^2 \nu^2) |K(\omega, \nu)|^2 d\nu d\omega < \eta^2,$$

where F' denotes the complement of the rectangle $F = \{(\omega, \nu) : |\nu| \leq 2\pi V, |\omega| \leq 2\pi W\}$. Then (2.25) becomes

$$(2.33) \quad \left\| k(t, s) - \sum_{(n/2W, m/2V) \in E'} k\left(\frac{n}{2W}, \frac{m}{2V}\right) \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)} \frac{\sin \pi(2Vs - m)}{\pi(2Vs - m)} \right\|^2 \leq 21(\epsilon + 2\eta)^2 + (3)^{1/2} \pi^2 \|k\|^2 (\epsilon + 2\eta) + 10\eta^2.$$

Proof. Again we begin by assuming $V = W = 1/2$. Let P and Q be the orthogonal projections on $\mathcal{K}$ defined as follows. Pk is given by

$$(2.34) \quad (Pk)(t, s) = \begin{cases} k(t, s), & \text{if } (t, s) \in E, \\ 0, & \text{if } (t, s) \in E', \end{cases}$$

and Qk is determined through its Fourier transform by

$$(2.35) \quad (Qk)(\omega, \nu) = \begin{cases} K(\omega, \nu), & \text{if } (\omega, \nu) \in F, \\ 0, & \text{if } (\omega, \nu) \in F'. \end{cases}$$

Thus P and Q are orthogonal projections onto the subspaces of channels whose supports lie in E and F , respectively. Then from (2.23) we see that for all k satisfying the hypotheses, we have

$$\|k - Pk\| < \epsilon, \quad \left\| \frac{\partial k}{\partial t} - P \frac{\partial k}{\partial t} \right\| < \epsilon, \quad \text{and} \quad \left\| \frac{\partial k}{\partial s} - P \frac{\partial k}{\partial s} \right\| < \epsilon.$$

Similarly, from (2.32) we see that

$$\|k - Qk\| < \eta, \quad \left\| \frac{\partial k}{\partial t} - Q \frac{\partial k}{\partial t} \right\| < \eta, \quad \text{and} \quad \left\| \frac{\partial k}{\partial s} - Q \frac{\partial k}{\partial s} \right\| < \eta.$$

Moreover, we have

$$\|Qk - PQk\| \leq \|Qk - k\| + \|k - Pk\| + \|Pk - PQk\| < \eta + \epsilon + \eta.$$

Similarly,

$$\left\| Q \frac{\partial k}{\partial t} - PQ \frac{\partial k}{\partial t} \right\| < \epsilon + 2\eta, \quad \text{and} \quad \left\| Q \frac{\partial k}{\partial s} - PQ \frac{\partial k}{\partial s} \right\| < \epsilon + 2\eta.$$

Thus

$$\|Qk - PQk\|^2 + \left\| \frac{\partial}{\partial t} Qk - P \frac{\partial}{\partial t} Qk \right\|^2 + \left\| \frac{\partial}{\partial s} Qk - P \frac{\partial}{\partial s} Qk \right\|^2 < 3(\epsilon + 2\eta)^2.$$

Thus Qk satisfies the hypothesis, and hence the conclusion, of Theorem 2.20, with ϵ^2 replaced throughout by $3(\epsilon + 2\eta)^2$.

Now if Σ' denotes the sampling approximation to Qk , then

$$k - \Sigma' = (k - Qk) + (Qk - \Sigma').$$

Since Q is an orthogonal projection, we have

$$\begin{aligned} \|k - \Sigma'\|^2 &= \|k - Qk\|^2 + \|Qk - \Sigma'\|^2 \\ &\leq \eta^2 + \frac{21}{2}(\epsilon + 2\eta)^2 + \frac{(3)^{1/2}}{2}\pi^2 \|k\|(\epsilon + 2\eta). \end{aligned}$$

Now let Σ denote the sampling approximation to k , i.e., the sum appearing in (2.33). Then the difference $\Sigma' - \Sigma$ is dominated by

$$(2.36) \quad \|\Sigma' - \Sigma\|^2 = \sum_{(n,m) \in \mathcal{E}} |(Qk)(n, m) - k(n, m)|^2.$$

Using the summation formula (2.16) again, with $h(t, s) = k(t, s) - (Qk)(t, s)$, we find, as in (2.27),

$$(2.37) \quad \|\Sigma' - \Sigma\|^2 = \iint_{\mathcal{E}} \left(1 + \sigma(t) \frac{\partial}{\partial t}\right) \left(1 + \sigma(s) \frac{\partial}{\partial s}\right) |h(t, s)|^2 ds dt.$$

As in (2.28), this expression is dominated by a sum of five terms, each of which may be related to (2.32) by taking Fourier transforms:

$$(2.38) \quad \begin{aligned} \|\Sigma - \Sigma'\|^2 &\leq \|h\|^2 + \|h\| \left\| \frac{\partial h}{\partial t} \right\| + \|h\| \left\| \frac{\partial h}{\partial s} \right\| \\ &\quad + \frac{1}{2} \|h\| \left\| \frac{\partial^2 h}{\partial t \partial s} \right\| + \frac{1}{2} \left\| \frac{\partial h}{\partial t} \right\| \left\| \frac{\partial h}{\partial s} \right\|. \end{aligned}$$

But $\|h\| = \|k - Qk\| = \|K - QK\| < \eta$, by (2.32). Similarly,

$$\left\| \frac{\partial h}{\partial t} \right\| < \eta, \quad \left\| \frac{\partial h}{\partial s} \right\| < \eta, \quad \text{and} \quad \left\| \frac{\partial^2 h}{\partial t \partial s} \right\| < \eta.$$

Combining these estimates, we find

$$(2.39) \quad \|\Sigma - \Sigma'\|^2 \leq 4\eta^2.$$

From (2.39) we conclude that

$$(2.40) \quad \begin{aligned} \|k - \Sigma\|^2 &\leq 2 \|k - \Sigma'\|^2 + 2 \|\Sigma' - \Sigma\|^2 \\ &\leq 8\eta^2 + 2\eta^2 + 21(\epsilon + 2\eta)^2 + (3)^{1/2}\pi^2 \|k\| (\epsilon + 2\eta) \end{aligned}$$

as required.

Again the general case reduces to this through the changes of scale (2.30) and (2.31).

Note that when $\eta = \epsilon < 1$ and $\|k\|^2 < 1$, the difference (2.40) is majorized by 259ϵ .

We are now prepared to state and prove Kailath's result [cf. 1].

Theorem 2.22. *Let $\mathfrak{D}$ be a (bounded) class of linear channels. Suppose that, given $\epsilon > 0$, there exist positive numbers R , S , U and W , such that for all k in $\mathfrak{D}$ we have both*

$$(2.41) \quad \iint_{E(R, S)'} \left(|k(t, s)|^2 + \left| \frac{\partial}{\partial t} k(t, s) \right|^2 + \left| \frac{\partial}{\partial s} k(t, s) \right|^2 \right) ds dt < \epsilon^2$$

and

$$(2.42) \quad \iint_{F(U, W)'} (|K(\omega, \nu)|^2 + |\omega K(\omega, \nu)|^2 + |\nu K(\omega, \nu)|^2) d\nu d\omega < \epsilon^2.$$

Here $K(\omega, \nu)$ denotes the complete Fourier transform of $k(t, s)$, and $E(R, S)'$ and $F(U, W)'$ denote the complements of the domains $E(R, S)$ and $F(U, W)$ defined by the relations

$$(2.43) \quad E(R, S) = \{(t, s) : |s| \leq \frac{1}{2}S, |t - s| \leq \frac{1}{2}R\}$$

and

$$(2.44) \quad F(U, W) = \{(\omega, \nu) : |\omega| \leq 2\pi W, |\omega + \nu| \leq 2\pi U\}$$

respectively. Then $\mathfrak{D}$ is always determinable in principle. Moreover, if for each $\epsilon > 0$ the numbers R and U can be chosen so that $2RU < 1$, then $\mathfrak{D}$ is also determinable in practice.

Proof. It follows immediately from Theorem 2.9 that under the hypotheses of Theorem 2.22 $\mathfrak{D}$ is always determinable in principle.

Now given $\epsilon > 0$, let us choose $R = R(\epsilon)$ and $U = U(\epsilon)$ as small as possible so that (2.41) and (2.42) hold. Then clearly $R(\epsilon)$ and $U(\epsilon)$ are monotone increasing functions of ϵ , and $R = \sup R(\epsilon)$ and $U = \sup U(\epsilon)$ are well defined. Moreover, (2.41) and (2.42) hold for all ϵ if $R(\epsilon)$ and $U(\epsilon)$ are replaced throughout by R and U . It follows that $k(t, s)$ vanishes almost everywhere unless $|t - s| \leq R$, and similarly $K(\omega, \nu)$ vanishes almost everywhere unless $|\omega - \nu| \leq 2\pi U$. Finally, if for all $\epsilon > 0$ we have $2R(\epsilon)U(\epsilon) < 1$, then clearly $2RU < 1$.

Now given $\epsilon > 0$, we choose R and U as indicated above, and $S = S(\epsilon)$ and $W = W(\epsilon)$ so large that $R < S$ and $U < W$. By making W large and adjusting R , S , and U slightly, we may assume that $1/2U$, $R - 1/2$ and $S - 1/2$ are

all integral multiples of $1/2W$. By now making a change of scale of the form $s \rightarrow s' = 2Ws$, $t \rightarrow t' = 2Wt$, we may also assume that $1/2W = 1$.

Now we introduce a change of variables in the s - t plane as follows:

$$(2.45) \quad \begin{pmatrix} s \\ t \end{pmatrix} \rightarrow \begin{pmatrix} s \\ r \end{pmatrix} = \begin{pmatrix} s \\ t - s \end{pmatrix}.$$

The conjugate variables are then also changed as follows:

$$(2.46) \quad \begin{pmatrix} \nu \\ \omega \end{pmatrix} \rightarrow \begin{pmatrix} \mu \\ \omega \end{pmatrix} = \begin{pmatrix} \omega + \nu \\ \omega \end{pmatrix}.$$

Using the new variables, we define for each channel k in $\mathfrak{D}$ an associated function h related to k via

$$(2.47) \quad h(r, s) = k(r + s, s).$$

The complete Fourier transform $H(\omega, \mu)$ of $h(r, s)$ is then related to the transform $K(\omega, \nu)$ of $k(t, s)$ via

$$(2.48) \quad H(\omega, \mu) = K(\omega, \mu - \nu).$$

From these relations it follows easily that the associated function $h(r, s)$ satisfies conditions analogous to those of (2.41) and (2.42), via

$$(2.49) \quad \iint_{A(R, S)'} \left(|h(r, s)|^2 + \left| \frac{\partial}{\partial r} h(r, s) \right|^2 + \left| \frac{\partial}{\partial s} h(r, s) \right|^2 \right) ds dr < 4\epsilon^2,$$

$$(2.50) \quad \iint_{B(U, W)'} (|H(\omega, \mu)|^2 + |\omega H(\omega, \mu)|^2 + |\mu H(\omega, \mu)|^2) d\mu d\omega < 4\epsilon^2.$$

Here $A(R, S)'$ and $B(U, W)'$ denote the complements of the sets $A(R, S)$ and $B(U, W)$ defined by

$$(2.51) \quad A(R, S) = \{(r, s) : |r| \leq \tfrac{1}{2}R, |s| \leq \tfrac{1}{2}S\},$$

$$(2.52) \quad B(U, W) = \{(\mu, \omega) : |\mu| \leq 2\pi U, |\omega| \leq 2\pi W\}.$$

Moreover, because of our choice of R and U , $h(r, s)$ vanishes almost everywhere unless $|r| \leq \tfrac{1}{2}R$, and similarly $H(\omega, \nu)$ vanishes almost everywhere unless $|\mu| \leq 2\pi U$. In particular, since $2U < 2W = 1$, we have

$$(2.53) \quad \begin{aligned} \iint_{B(U, W)'} \omega^2 \mu^2 |H(\omega, \mu)|^2 d\mu d\omega &\leq \pi^2 \iint_{B(U, W)'} \omega^2 |H(\omega, \mu)|^2 d\mu d\omega \\ &\leq \pi^2 \epsilon^2. \end{aligned}$$

We conclude from these computations that $h(r, s)$ satisfies all the hypotheses, and hence the conclusion, of Corollary 2.21, with ϵ^2 replaced throughout by $4\epsilon^2$, and η^2 by $(4 + \pi^2)\epsilon^2 < 16\epsilon^2$. Thus $h(r, s)$ admits a sampling approximation of the form

$$(2.54) \quad \left\| h(r, s) - \sum_{(n/2W, m/2U) \in A(R, S)} h\left(\frac{n}{2W}, \frac{m}{2U}\right) \frac{\sin \pi(2Wr - n)}{\pi(2Wr - n)} \frac{\sin \pi(2Us - m)}{\pi(2Us - m)} \right\|^2$$

$$< 2260\epsilon^2 + 200\epsilon$$

$$< 2460\epsilon, \quad \text{if } \epsilon < 1.$$

Putting $1/(2W) = 1$ and $1/(2U) = p$, and replacing (n, pm) by $(n - pm, pm)$, we may rewrite (2.54) in terms of $k(t, s)$:

$$(2.55) \quad \left\| k(t, s) - \sum_{(n, pm) \in E(R, S)} k(n, pm) \frac{\sin \pi(t - s - n + pm)}{\pi(t - s - n + pm)} \frac{\sin 2\pi U(s - pm)}{2\pi U(s - pm)} \right\|^2$$

$$< 2460\epsilon.$$

Thus we see that a knowledge of the sample values $k(n, pm)$ of k , taken at the sample points (n, pm) in the region E , suffices to determine $k(t, s)$ to within a mean-square error of 2460ϵ .

Moreover, if $2RU < 1$, then we know that each sample point of the form (n, pm) in E is uniquely determined by its t -coordinate n . In fact, if (n, pm) and (n, pm') are two such sample points with the same t -coordinate n , then the horizontal distance between them in the s - t plane is at least as great as $p = 1/(2U)$, whereas the horizontal width of the region E at the height n is at most $R < 1/(2U)$. Hence at most one of these points can lie in E .

In order to show that $\mathfrak{D}$ is determinable, it will suffice to produce an input signal $x(s)$ and a finite number of output signals $y_n(t)$, such that the measurement m_n defined by the formula $m_n(kx) = (y_n, kx)$ yields the sample value $k(n, pm)$. To this end we introduce the approximate impulse function $\varphi_\delta(s)$, defined by

$$(2.56) \quad \varphi_\delta(s) = \begin{cases} \frac{1}{2\delta}, & |s| \leq \delta, \\ 0, & |s| > \delta, \end{cases}$$

depending on the small parameter δ . Now we define

$$(2.57) \quad x_m(s) = \varphi_\delta(s - m),$$

$$(2.58) \quad y_n(t) = \varphi_\delta(t - m).$$

Then $x_m(s)$ and $y_n(t)$ lie in the signal space $\mathfrak{H}$. Moreover, we expect that for small values of δ the number (y_n, kx_m) is a good approximation to the sample value $k(n, m)$. In fact, we have

$$\begin{aligned}
(2.59) \quad & \left| k(n, m) - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \overline{y_m(t)} k(t, s) x_n(s) ds \right| \\
&= \left| \iint e^{i(n\omega + m\nu)} (1 - \overline{Y_0(\omega)} X_0(\nu)) K(\omega, \nu) d\nu d\omega \right| \\
&\leq \iint \left| 1 - \frac{\sin \delta\omega}{\delta\omega} \frac{\sin \delta\nu}{\delta\nu} \right| |K(\omega, \nu)| d\nu d\omega \\
&= \iint_{F(U, Z)} + \iint_{F(U, Z)'} \\
&\leq \left(\iint_F \left| 1 - \frac{\sin \delta\omega}{\delta\omega} \frac{\sin \delta\nu}{\delta\nu} \right|^2 d\nu d\omega \right)^{1/2} \left(\iint_F |K(\omega, \nu)|^2 d\nu d\omega \right)^{1/2} \\
&\quad + \sup_{(\omega, \nu) \in F'} \left| 1 - \frac{\sin \delta\omega}{\delta\omega} \frac{\sin \delta\nu}{\delta\nu} \right| \iint_{F'} |K(\omega, \nu)| d\nu d\omega.
\end{aligned}$$

Here $F(U, Z)$ is the region defined by (2.44), with W replaced by $Z > W$, and $F(U, Z)'$ its complement. To complete the argument, we have only to note that given $\eta > 0$, we can choose Z so large that

$$\begin{aligned}
(2.60) \quad & \iint_{F(U, Z)'} |K(\omega, \nu)| d\nu d\omega \\
&= \iint_{F(U, Z)'} (1 + \omega^2 + \nu^2)^{-1/2} (1 + \omega^2 + \nu^2)^{1/2} |K(\omega, \nu)| d\nu d\omega \\
&\leq \left(\iint_{|\omega + \nu| \leq 2\pi U} (1 + \omega^2 + \nu^2)^{-1} d\nu d\omega \right)^{1/2} \\
&\quad \cdot \left(\iint_{F(U, Z)'} (1 + \omega^2 + \nu^2) |K(\omega, \nu)|^2 d\nu d\omega \right)^{1/2} \\
&\leq (2\pi U)^{1/2} \eta.
\end{aligned}$$

With this choice of Z , we can then choose δ so small that

$$(2.61) \quad \left(\iint_{F(U, Z)} \left| 1 - \frac{\sin \delta\omega}{\delta\omega} \frac{\sin \delta\nu}{\delta\nu} \right|^2 d\nu d\omega \right)^{1/2} < \eta.$$

Inserting (2.60) and (2.61) into (2.59), we obtain

$$(2.62) \quad \left| k(n, m) - \iint y_n(t) k(t, s) x_m(s) ds dt \right| < \|k\| \eta + 2(2\pi U)^{1/2} \eta.$$

Finally, we define

$$(2.63) \quad x(s) = \sum_{|pm| \leq S} x_{pm}(s).$$

Then we find that the form (y_n, kx) yields

$$(2.64) \quad (y_n, kx) = \sum_{|pm| \leq S} (y_n, kx_{pm}).$$

But since $k(t, s)$ vanishes outside of E , only those terms in this sum for which (n, pm) lie in E can differ from zero if δ is less than $1/2$. According to the argument given above, there can be at most one such term. Hence

$$(2.65) \quad (y_n, kx) = (y_n, kx_{pm})$$

and by choosing δ sufficiently small, we can ensure

$$(2.66) \quad |k(n, pm) - (y_n, kx)|^2 < \frac{\epsilon}{R + S}.$$

It follows now that if we define

$$(2.67) \quad \hat{k}(t, s) = \sum_{(n, pm) \in E} (y_n, kx) \frac{\sin \pi(t - s - n + pm)}{\pi(t - s - n + pm)} \frac{\sin 2\pi U(s - pm)}{2\pi U(s - pm)}$$

then we have, from (2.66) and (2.55),

$$(2.68) \quad \begin{aligned} ||k - \hat{k}||^2 &\leq 2 ||k - \Sigma||^2 + 2 ||\Sigma - \hat{k}||^2 \\ &\leq 2 \cdot 2460\epsilon + 2\epsilon \\ &\leq 4922\epsilon. \end{aligned}$$

Thus, given $\epsilon > 0$, we can choose S , W , and δ , and using them construct x , y_n and $\hat{k}$, in such a way that (2.68) holds. This completes the proof of Theorem 2.22, and incidentally provides an alternative procedure for constructing the measuring process $\hat{m}$ and the channel estimator function $\hat{k}$ for classes of channels satisfying the hypotheses of this Theorem.

3. Determinable Classes of Non-Linear Channels. By a *non-linear channel* we shall mean a non-linear operator acting on the signal space $\mathcal{H}$. We shall always assume that the operator can be approximated in the mean-square sense by a finite linear combination of homogeneous polynomial operators of the following form:

$$(3.1) \quad y(t) = \iint \cdots \int k(t, s_1, s_2, \cdots, s_n) x(s_1) x(s_2) \cdots x(s_n) ds_1 ds_2 \cdots ds_n$$

with a kernel which is square-integrable in all variables:

$$(3.2) \quad ||k||^2 = \iint \cdots \int |k(t, s_1, \cdots, s_n)|^2 ds_1 \cdots ds_n dt < \infty.$$

As before, we assume here that the kernel is defined for all values of its variables and vanishes when these variables lie outside the time interval of interest.

It is known that operators of this form take elements of $\mathcal{H}$ into elements of $\mathcal{H}$ in such a way that

$$(3.3) \quad \|kx\| \leq \|k\| \|x\|^n$$

and that sums and scalar products of such operators are again operators of the same form [6].

The class of all channels of the form (3.1)–(3.2) itself forms a Hilbert space $\mathfrak{H}_n$ called here the n^{th} homogeneous channel space, with norm given by (3.2). Moreover, if $\{\varphi_i\}$ is an orthonormal basis for the signal space $\mathcal{H}$, then the set $\{\varphi(j, i_1, i_2, \dots, i_n) = \varphi_i(t) \overline{\varphi_{i_1}(s_1)} \overline{\varphi_{i_2}(s_2)} \dots \overline{\varphi_{i_n}(s_n)}\}$ forms an orthonormal basis for the channel space $\mathfrak{H}_n$. In this sense $\mathfrak{H}_n$ may be identified with the $(n+1)$ -fold tensor product $\mathcal{H} \otimes \mathcal{H} \otimes \dots \otimes \mathcal{H}$. Note that $\mathfrak{H}_1$ is identical with the linear channel space $\mathcal{H}$.

We shall call the direct sum $\mathfrak{H} = \sum_{n=1}^{\infty} \mathfrak{H}_n$ of the homogeneous channel spaces $\mathfrak{H}_n$ simply the *channel space*. Every element of $\mathfrak{H}$, then, is the limit in the mean-square sense of finite linear combinations of homogeneous polynomial channels.

Now let $\mathfrak{D}$ be a class of channels, regarded as a subset of the channel space $\mathfrak{H}$. We shall always assume that $\mathfrak{D}$ is bounded, in the sense that it is contained in the unit sphere of $\mathfrak{H}$.

By a *measurement* on $\mathfrak{D}$, we shall mean a functional m , defined on $\mathfrak{D}$, which assigns to each channel k in $\mathfrak{D}$ a complex number $m(k)$. We shall always assume that $m(k)$ depends continuously on k , and that $|m(k)|$ is always no greater than 1.

By a *channel estimator function*, we shall mean a continuous function $\hat{k}$, defined on the space of all n -tuples of complex numbers, which assigns to each such n -tuple $(\xi_1, \xi_2, \dots, \xi_n)$ an element $\hat{k}(\xi_1, \xi_2, \dots, \xi_n)$ of the channel space $\mathfrak{H}$.

Again we introduce two distinct notions of determinability. We shall say that the class $\mathfrak{D}$ is *determinable in principle* if, for every $\epsilon > 0$, there is a measuring process $\hat{m}$, consisting of a finite number of measurements $(m_1, m_2, \dots, m_n)$ on $\mathfrak{D}$, together with a channel estimator function $\hat{k}$, such that for every channel k in $\mathfrak{D}$ we have

$$(3.4) \quad \|k - \hat{k} \circ \hat{m}(k)\| < \epsilon.$$

Moreover, we shall say that the class $\mathfrak{D}$ is *determinable in practice* if, for every $\epsilon > 0$, there is a single input signal x , a measuring process $\hat{m}$ on the signal space $\mathcal{H}$, and a channel estimator function $\hat{k}$ such that for every channel k in $\mathfrak{D}$ we have

$$(3.5) \quad \|k - \hat{k} \circ \hat{m}(kx)\| < \epsilon.$$

These definitions reduce to those of the previous section if the class $\mathfrak{D}$ lies entirely within the linear channel space $\mathfrak{H}_1 = \mathcal{H}$. It should be clear from these definitions that every class of channels which is determinable in practice is also determinable in principle. It follows from Lemma 2.1 that the converse is false.

We now list the properties of those classes of channels which are determinable in principle. Except where noted, they are identical with the properties derived in the previous section for the linear case.

Lemma 3.1. *Let $\mathfrak{D}$ be a class of channels which is determinable in principle. Then the closure of $\mathfrak{D}$ in the channel space $\mathfrak{K}$ is compact.*

Proof. See Lemma 1.1.

Lemma 3.2. *Let $\mathfrak{D}$ be any subset of $\mathfrak{K}$ whose closure is compact. Then $\mathfrak{D}$ is determinable in principle.*

Proof. See Lemma 1.2. Note that this proof constructs a measuring process which is linear on $\mathfrak{K}$, but which is not necessarily realizable by introducing a finite number of input signals into the channel and making a finite number of measurements on each of the resulting outputs, as in the case of linear channels. This difference is a consequence of the non-linear character of the channels involved, which cannot be easily characterized by linear operations on the signal space.

Combining these results, we obtain

Theorem 3.3. *A class $\mathfrak{D}$ of non-linear channels is determinable in principle if and only if its closure in the channel space $\mathfrak{K}$ is compact.*

Corollary 3.4. *If the class $\mathfrak{D}$ is determinable in principle then for each $\epsilon > 0$ the associated measuring process $\hat{m}$ and channel estimator function $\hat{k}$ may be chosen to be linear, although not necessarily in the form (2.7) and (2.8)*

Corollary 3.5. *The analogues of Corollaries 1.5–1.10 all hold for classes of non-linear channels which are determinable in principle.*

For the analogue of Theorem 2.7, we must require that the class $\mathfrak{D}$ be not too badly non-linear, in the sense that the contributions from homogeneous polynomial channels of high degree are uniformly small. The proper statement reads as follows:

Theorem 3.6. *Let $\{\varphi_i\}$ be any orthonormal basis for the signal space $\mathfrak{H}$, and $\{\varphi(j, i_1, i_2, \dots, i_n)\}$ the corresponding basis for the channel space $\mathfrak{K}$. Then the class $\mathfrak{D}$ is determinable in principle if and only if for every $\epsilon > 0$ there exist integers M and N such that for all k in $\mathfrak{D}$ we have*

$$(3.6) \quad \sum_{n \leq M} \sum_{S(N)'} |k_n(j, i_1, i_2, \dots, i_n)|^2 < \epsilon^2$$

for all $n \leq M$, and

$$(3.7) \quad \sum_{n > M} \|k_n\|^2 < \epsilon^2.$$

Here k_n is the n^{th} homogeneous polynomial component of k and $k_n(j, i_1, i_2, \dots, i_n)$ the Fourier coefficients of k_n in its expansion in terms of the $\varphi(j, i_1, i_2, \dots, i_n)$; and $S(N)'$ denotes the complement of the set $S(N)$ of indices $\{j, i_1, i_2, \dots, i_n\}$ such that $j < N, i_1 < N, i_2 < N, \dots, i_n < N$.

Proof. Suppose first that, given $\epsilon > 0$, there exist integers M and N such

that (3.6) and (3.7) hold for every k in $\mathfrak{D}$. Then $\mathfrak{D}$ is totally bounded. In fact, every element k of $\mathfrak{D}$ may be approximated to within 2ϵ by elements of $\mathfrak{K}$ of the form

$$\sum_{n \leq M} \sum_{S(N)} k_n(j, i_1, \dots, i_n) \varphi(j, i_1, \dots, i_n)$$

with norm ≤ 1 . The set of all such elements is a closed bounded finite dimensional subset of $\mathfrak{K}$, and hence compact. Any covering of this set by spheres of radius ϵ admits a finite subcovering which, in turn, defines a finite covering of $\mathfrak{D}$ by spheres of radius 3ϵ .

Now suppose that, given $\epsilon > 0$, there exists no M such that (3.7) holds for every k in $\mathfrak{D}$. Then we may choose a sequence M_i of integers and a sequence $k^{(i)}$ of elements in $\mathfrak{D}$ such that

$$\sum_{n \leq M_i} \|k_n^{(i)}\|^2 > 4\epsilon^2 \quad \text{but} \quad \sum_{n > M_{i+1}} \|k_n^{(i)}\|^2 < \epsilon^2.$$

It follows that for any integers i and j we have $\|k^{(i)} - k^{(j)}\| > \epsilon$. Hence no covering of $\mathfrak{D}$ with spheres of radius ϵ can be finite, and $\mathfrak{D}$ is not totally bounded. A similar argument gives the same result if there exists no integer N such that (3.6) holds for all k in $\mathfrak{D}$. In either case, $\mathfrak{D}$ is not determinable in principle.

Next we present the analogue of Theorem 2.9. Again we must require that the class $\mathfrak{D}$ be not too badly non-linear.

Theorem 3.7. *The class $\mathfrak{D}$ is determinable in principle if and only if for each $\epsilon > 0$ there exist an integer M and positive numbers T and W such that for all k in $\mathfrak{D}$ we have both*

$$(3.8) \quad \sum_{n \leq M} \iint \cdots \int_{E(T)} |k_n(t, s_1, s_2, \dots, s_n)|^2 ds_1 ds_2 \cdots ds_n dt < \epsilon^2$$

and

$$(3.9) \quad \sum_{n \leq M} \iint \cdots \int_{F(W)} |K_n(\omega, \nu_1, \nu_2, \dots, \nu_n)|^2 d\nu_1 d\nu_2 \cdots d\nu_n d\omega < \epsilon^2$$

for all $n \leq M$, and

$$(3.10) \quad \sum_{n > M} \|k_n\|^2 < \epsilon^2$$

for all $n > M$. Here k_n denotes the homogeneous polynomial component of n^{th} degree in k , and K_n its complete Fourier transform; and $E(T)'$ and $F(W)'$ denote the regions exterior to the $n + 1$ dimensional cubes $E(T)$ and $F(W)$ defined by the relations $|s_i| < T$, $|t| < T$, and $|\nu_i| < W$, $|\omega| < W$, respectively.

Proof. Suppose first that, given $\epsilon > 0$, there exist an integer M and positive numbers T and W such that (3.8), (3.9) and (3.10) all hold for every k in $\mathfrak{D}$. We shall show that $\mathfrak{D}$ is totally bounded.

We first define the characteristic function $\chi_T(t, s_1, s_2, \dots, s_n)$ by the formula

$$(3.11) \quad \chi_T(t, s_1, s_2, \dots, s_n) = \begin{cases} 1, & \text{if } |t| < T, \quad |s_1| < T, \quad |s_2| < T, \dots, |s_n| < T, \\ 0, & \text{otherwise.} \end{cases}$$

Using this function we then define the operators P_T and Q_W acting on the n^{th} homogeneous channel space $\mathfrak{H}_n$ according to the formulas

$$(3.12) \quad (PTk)(t, s_1, \dots, s_n) = \chi_T(t, s_1, \dots, s_n)k(t, s_1, \dots, s_n)$$

and

$$(3.13) \quad (Q_W K)(\omega, \nu_1, \dots, \nu_n) = \chi_W(\omega, \nu_1, \dots, \nu_n)K(\omega, \nu_1, \dots, \nu_n).$$

Here K denotes the complete Fourier transform of k . Then P_T and Q_W are both orthogonal projections on $\mathfrak{H}$. Moreover, as in the proof of Theorem 1.14 we can show that the product $Q_W P_T Q_W$ is also a Hilbert-Schmidt operator on $\mathfrak{H}_n$. It follows that the image $\mathfrak{G}_n$ of the unit sphere in $\mathfrak{H}_n$ under the product $Q_W P_T Q_W$ is compact in $\mathfrak{H}_n$.

Moreover, every element k in $\mathfrak{H}$ may be approximated to within 4ϵ by an element of the form $\sum_{n \leq M} h_n$, where h_n lies in $\mathfrak{G}_n$. In fact, if k_n denotes the n^{th} homogeneous component of k and $h_n = Q_W P_T Q_W k_n$, then we have

$$(3.14) \quad \begin{aligned} \|k - \sum_{n \leq M} h_n\| &\leq \|k - \sum_{n \leq M} k_n\| + \|\sum_{n \leq M} (k_n - h_n)\| \\ &\leq \|k - \sum_{n \leq M} k_n\| + \|\sum_{n \leq M} (I - Q_W)k_n\| \\ &\quad + \|\sum_{n \leq M} Q_W(I - P_T)k_n\| + \|\sum_{n \leq M} Q_W P_T(I - Q_W)k_n\| \\ &\leq \|k - \sum_{n \leq M} k_n\| + (\sum_{n \leq M} \|(I - Q_W)k_n\|^2)^{1/2} \\ &\quad + (\sum_{n \leq M} \|Q_W(I - P_T)k_n\|^2)^{1/2} \\ &\quad + (\sum_{n \leq M} \|Q_W P_T(I - Q_W)k_n\|^2)^{1/2}. \end{aligned}$$

According to the hypotheses (3.8), (3.9) and (3.10), each of the terms on the right hand side of (3.14) is less than ϵ .

Since the set of all elements of the form $\sum_{n \leq M} h_n$, with h_n in $\mathfrak{G}_n$, is compact in $\mathfrak{H}$, this set admits a finite covering with spheres of radius ϵ . By multiplying the radius of each of these spheres by a factor of $4 + 1 = 5$, we obtain a finite covering of $\mathfrak{D}$ with spheres of radius 5ϵ . Since ϵ is quite arbitrary, we conclude that the class $\mathfrak{D}$ is determinable in principle.

Conversely, suppose that, given $\epsilon > 0$, there exists no integer M such that (3.10) holds for all k in $\mathfrak{D}$. Then, as in the proof of Theorem 3.6 we can find a sequence of elements $k^{(i)}$ of $\mathfrak{D}$ such that for all i and j we have $\|k^{(i)} - k^{(j)}\| > \epsilon$. Hence no covering of $\mathfrak{D}$ by spheres of radius ϵ can be finite. A similar argument

holds if there exists no T or W such that (3.8) and (3.9) hold for all k in $\mathfrak{D}$. In any case $\mathfrak{D}$ is not determinable.

We now turn to the problem of characterizing those classes of non-linear channels which are determinable in practice.

As in the linear case, every class of non-linear channels which is determinable in practice is necessarily determinable in principle. The example of Lemma 2.11 shows that the converse is false, and suggests again that a class of signals is determinable in practice only if there exists a single signal which essentially separates the elements on $\mathfrak{D}$ from each other.

Lemma 3.8. *Let $\mathfrak{D}$ be a class of non-linear channels which is determinable in practice. Then for every $\epsilon > 0$ there exists a single signal x such that, for every pair $\{h, k\}$ of channels in the closure of $\mathfrak{D}$ with $\|h - k\| > \epsilon$, we have $\|hx - kx\| > 0$.*

Proof. Identical with that of Lemma 2.12.

The converse is also true.

Lemma 3.9. *Let $\mathfrak{D}$ be a class of channels which is determinable in principle. Suppose moreover that for every $\epsilon > 0$ there exists a single signal x such that for all pairs $\{h, k\}$ of channels in the closure of $\mathfrak{D}$ with $\|h - k\| > \epsilon$ we have $\|hx - kx\| > 0$. Then $\mathfrak{D}$ is determinable in practice.*

Proof. Identical with that of Lemma 2.13. Note that these proofs depend not on the linearity of the channels in $\mathfrak{D}$, but only on the compactness of the closure.

Combining these results, we obtain

Theorem 3.10. *Let $\mathfrak{D}$ be a class of non-linear channels which is determinable in principle. Then $\mathfrak{D}$ is determinable in practice if and only if, for each $\epsilon > 0$, there exists a single signal x in $\mathfrak{C}$ such that, for all pairs $\{h, k\}$ of channels in the closure of $\mathfrak{D}$, $\|h - k\| > \epsilon$ implies $\|hx - kx\| > 0$.*

Corollary 3.11. *The analogues of Corollaries 2.16, 2.17 and 2.18 are all valid for classes $\mathfrak{D}$ of non-linear channels.*

The extension to classes of non-linear channels of the remaining results of Section 2 lies beyond the scope of this paper and will not be discussed here.

Note added in proof: In a paper to appear in the *SIAM Journal*, P. P. Varaiya has characterized determinable classes of signals which are convex but unbounded, and has proved a theorem like our Theorem 2.15 for the more general case that the channel operators are only required to be continuous endomorphisms on $\mathfrak{C}$.

BIBLIOGRAPHY

- [1] KAILATH, T., Sampling Models for Linear Time-Variant Filters, *Research Laboratory of Electronics, M.I.T. Technical Report*, No. 352 (1959).

- [2] KOLMOGOROV, A. N., Asymptotic Characteristics of Some Completely Bounded Metric Spaces, *Dokl. Akad. Nauk.*, **108** (1956) 585-589.
- [3] KOLMOGOROV, A. N. & TIHOMIROV, V. M., ϵ -entropy and ϵ -capacity of Sets in Function Spaces, *Uspehi Mat. Nauk.*, **14** (1959) No. 2 (86) 3-86.
- [4] LANDAU, H. J. & POLLAK, H. O., The Dimension of the Space of Essentially Time- and Band-Limited Signals, *Bell System Technical Journal*, **XLI** **4** (1962) 1295-1336.
- [5] LORENTZ, G. G., Metric Entropy, Widths, and Superpositions of Functions, *Math. Monthly*, (1962) 469-485.
- [6] RIESZ, F. & SZ.-NAGY, B., *Lecons d'analyse Fonctionnelle*, Paris-Budapest, 1955.
- [7] ROOT, W. L., On the Measurement and Use of Time-Varying Communication Channels, *Info. and Control*, **8** No. 4 (1965) 390-422.
- [8] PROSSER, R. T., A Multidimensional Sampling Theorem, *J. Math. Anal. Appl.*, **4** (1966) (to appear).
- [9] PROSSER, R. T., The ϵ -entropy and ϵ -capacity of Certain Time-Varying Channels, *J. Math. Anal. Appl.*, **4** (1966) (to appear).
- [10] VARAIYA, P. P., On Determinable Classes of Signals and Linear Channels, *SIAM Journal*, (to appear).

Massachusetts Institute of Technology
University of Michigan

Date Communicated: JANUARY 21, 1966