JUL 10 1967

SC-R-67-1091

June 1967

CONF - 670430 - 1

Reprint

AN EXTENSION OF THE APPLICATION
OF REPRODUCING KERNELS TO A SAMPLING
THEOREM FOR ADAPTIVE SYSTEMS

G. J. Simmons
Sandia Laboratory, Albuquerque

SANDIA LABORATORIES

OPERATED FOR THE U. S. ATOMIC ENERGY COMMISSION BY SANDIA CORPORATION

ALBUQUERQUE, NEW MEXICO; LIVERMORE, CALIFORNIA

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Reg

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Published in: Proceedings of the 19th Southwestern IEEE Conference
April 19-21, 1967, Catalog No. F-72, Paper 1.5.

Issued by Sandia Corporation,
a prime contractor to the
United States Atomic Energy Commission

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

SANDIA CORPORATION REPRINT

H.C. \$ 3.00; MN .65

SC-R-67-1091

AN EXTENSION OF THE APPLICATION OF REPRODUCING KERNELS
TO A SAMPLING THEOREM FOR ADAPTIVE SYSTEMSG. J. Simmons
Sandia Laboratory, Albuquerque

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

June 1967

ABSTRACT

The method of reproducing kernels for the iterative approximation of functions $f(\mathbf{Z})$ in E_n converges under weak assumptions.¹ In this paper this theory is developed further and conditions on $f(\mathbf{Z})$ for exact reconstruction by the kernel

$$K(\mathbf{Z}, [nT]) = \prod_{k=1}^n \frac{\sin \frac{\pi}{T_k} (z_k - n_k T_k)}{\frac{\pi}{T_k} (z_k - n_k T_k)}$$

are determined. The resulting sampling theorem does not require a knowledge of $f(\mathbf{Z})$ but only a random sequence of argument-functional value pairs $\{f(\mathbf{Z}_i), \mathbf{Z}_i\}$.

TABLE OF CONTENTS

	<u>Page</u>
Introduction	3
The Method of Reproducing Kernels	3
A Sampling Theorem in $\mathcal{S}$	9
Summary	11
LIST OF REFERENCES	12

AN EXTENSION OF THE APPLICATION OF REPRODUCING KERNELS TO A SAMPLING THEOREM FOR ADAPTIVE SYSTEMS

Introduction

A large class of adaptive systems is concerned with a problem analogous to interpolation in which it is desired to acquire and store an unknown functional relation, $y = f(z_1, z_2, \dots, z_n)$ given a finite number of randomly observed argument-functional value pairs, $\{f(\mathbf{z}_i), \mathbf{z}_i\}$. If it is impractical to store all of the data pairs to permit conventional interpolation, either because of the dimensionality of $\mathbf{Z}$ or because of the large number of data points to be considered, the problem reduces to the iterative approximation of a memory fill function^{1, 2} at a selected set of reference arguments $\{\mathbf{z}_j\}$. The choice of the set $\{\mathbf{z}_j\}$ determines a function of two variables, $K(\mathbf{Z}, \mathbf{Z}')$, which is the reproducing kernel for a linear function space $\mathbf{S}$. This paper is concerned with the determination of those functions $f(\mathbf{Z})$ which can be approximated in the mean by elements of $\mathbf{S}$ and the conditions which must be satisfied by $f(\mathbf{Z})$ in order for an exact reconstruction to be possible from the values of the memory fill function at the selected $\{\mathbf{z}_j\}$.

The Method of Reproducing Kernels

Let $f(\mathbf{Z})$ be a sectionally continuous single valued real function over a region $\mathbf{R}$ of n dimensional Euclidean space. Assume that the isolated points of $f(\mathbf{Z})$ are on a set of measure zero and that $f(\mathbf{Z})$ is bounded almost everywhere (a.e.). It is required that $f(\mathbf{Z})$ be acquired and stored from a sequence of argument-functional value pairs $\{f(\mathbf{z}_i), \mathbf{z}_i\}$ randomly generated by a probability distribution $P(\mathbf{Z})$, where $P(\mathbf{Z}) > 0$ a.e. in $\mathbf{R}$. If $\{s_i(\mathbf{Z})\}$, $i = 1, 2, \dots, N$, is any set of N linearly independent functions in $\mathbf{S}$, then any function $f(\mathbf{Z})$ not orthogonal to all of the $s_i(\mathbf{Z})$ can be represented in a unique manner as

$$f(\mathbf{Z}) \cong \sum_{i=1}^N \omega_i s_i(\mathbf{Z}) .$$

If $f(\mathbf{Z}) \in \mathbf{S}$, then the representation is exact if $\dim \mathbf{S} = N$. In particular the reproducing kernel $K(\mathbf{Z}, \mathbf{Z}')$ for $\mathbf{S}$ is itself in $\mathbf{S}$ and is given by:^{2, 3}

$$K(\mathbf{Z}, \mathbf{Z}') = \sum_{i=1}^N \sum_{j=1}^N \beta_{ij} s_i(\mathbf{Z}) s_j(\mathbf{Z}') ,$$

where β_{ij} is an element in the matrix inverse of $\mathbf{A}$:

$$\mathbf{A} = (\alpha_{ij}) = (s_i(\mathbf{Z}), s_j(\mathbf{Z}')) .$$

Note that A is the Gram's matrix of a system of linearly independent functions and hence always possesses an inverse. If one carries out (implicitly) a Gram-Schmidt orthonormalization of the $s_i(\mathbf{Z})$, the equivalent form used throughout this paper is obtained:

$$f(\mathbf{Z}) = \sum_{i=1}^N \mu_i \phi_i(\mathbf{Z}) \quad (1)$$

$$K(\mathbf{Z}, \mathbf{Z}') = \sum_{i=1}^N \phi_i(\mathbf{Z}) \phi_i(\mathbf{Z}') \quad (2)$$

where

$$(\phi_i(\mathbf{Z}), \phi_j(\mathbf{Z})) = \delta_{ij}.$$

The $\phi_i(\mathbf{Z})$ are bounded and hence from the construction so is $K(\mathbf{Z}, \mathbf{Z}')$, i.e., $\max K(\mathbf{Z}, \mathbf{Z}') \leq L < \infty$.

To demonstrate the reproducing properties of $K(\mathbf{Z}, \mathbf{Z}')$, let $g(\mathbf{Z}) \in \mathcal{S}$:

$$g(\mathbf{Z}) = \sum_{i=1}^N \eta_i \phi_i(\mathbf{Z})$$

then

$$(g(\mathbf{Z}), K(\mathbf{Z}, \mathbf{Z}')) = \left(\sum_{i=1}^N \eta_i \phi_i(\mathbf{Z}), \sum_{i=1}^N \phi_i(\mathbf{Z}) \phi_i(\mathbf{Z}') \right) = \sum_{i=1}^N \eta_i \phi_i(\mathbf{Z}')$$

by the orthonormality of the $\phi_i(\mathbf{Z})$. Therefore,

$$(g(\mathbf{Z}), K(\mathbf{Z}, \mathbf{Z}')) = g(\mathbf{Z}')$$

for any $g(\mathbf{Z}) \in \mathcal{S}$, which is the characteristic property of the reproducing kernel.

The general iterative algorithm using the reproducing kernel to estimate $f(\mathbf{Z})$ is given by:

$$F_{\ell+1}(\mathbf{Z}) = F_{\ell}(\mathbf{Z}) + R[f(\mathbf{Z}^{\ell+1}), F_{\ell}(\mathbf{Z}^{\ell+1}), \ell] K(\mathbf{Z}, \mathbf{Z}^{\ell+1}) \quad (3)$$

where $F_{\ell}(\mathbf{Z})$ is the ℓ -th approximation to $f(\mathbf{Z})$ and $\mathbf{Z}^{\ell}$ is the ℓ -th argument selected. $R[f(\mathbf{Z}^{\ell+1}), F_{\ell}(\mathbf{Z}^{\ell+1}), \ell]$ is an as yet unspecified error function which will be abbreviated $R(\mathbf{Z}^{\ell+1})$. Equation (3), iterated ℓ times and summed, gives

$$F_{\ell}(\mathbf{Z}) = \sum_{j=1}^{\ell} R(\mathbf{Z}^j)K(\mathbf{Z}, \mathbf{Z}^j) \quad (4)$$

where $F_0(\mathbf{Z})$ is assumed to be 0 with no loss in generality. Equation (2) may be substituted into Eq. (4) and the order of summation reversed to give the desired formulation for $F_{\ell}(\mathbf{Z})$;

$$F_{\ell}(\mathbf{Z}) = \sum_{i=1}^N \left\{ \sum_{j=1}^{\ell} R(\mathbf{Z}^j) \phi_i(\mathbf{Z}^j) \right\} \phi_i(\mathbf{Z}) \quad (5)$$

which is true, even if N is not finite by the assumptions of boundedness and hence of the absolute convergence of Eqs. (1) and (2).

Define an error term $\Delta_{\ell}(\mathbf{Z})$ whose mean square expectation will be used to establish the convergence properties of the algorithm (3).

$$\Delta_{\ell}(\mathbf{Z}) = f(\mathbf{Z}) - F_{\ell}(\mathbf{Z}) = \sum_{i=1}^N \left\{ u_i - \sum_{j=1}^{\ell} R(\mathbf{Z}^j) \phi_i(\mathbf{Z}^j) \right\} \phi_i(\mathbf{Z}). \quad (6)$$

Subtract both sides of Eq. (3) from $f(\mathbf{Z})$ to express $\Delta_{\ell+1}(\mathbf{Z})$ recursively in terms of $\Delta_{\ell}(\mathbf{Z})$ and other lower order terms

$$\Delta_{\ell+1}(\mathbf{Z}) = \Delta_{\ell}(\mathbf{Z}) - R(\mathbf{Z}^{\ell+1})K(\mathbf{Z}, \mathbf{Z}^{\ell+1}). \quad (7)$$

The square of the error at any point $\mathbf{Z} \in \mathbb{R}$ on the $\ell+1$ -st iteration is given by

$$\Delta_{\ell+1}^2(\mathbf{Z}) = \Delta_{\ell}^2(\mathbf{Z}) - 2R(\mathbf{Z}^{\ell+1})\Delta_{\ell}(\mathbf{Z})K(\mathbf{Z}, \mathbf{Z}^{\ell+1}) + R^2(\mathbf{Z}^{\ell+1})K^2(\mathbf{Z}, \mathbf{Z}^{\ell+1}). \quad (8)$$

It can be shown that the mean square error (under the assumptions on $f(\mathbf{Z})$ already made) as given by Eq. (8) goes to zero with probability one by appealing to the theory of stochastic approximation. Such a proof would be made by demonstrating that the coefficients in the square error $\left(\Delta_{\ell}^2(\mathbf{Z}) \right)$ expanded in terms of the $\phi_i(\mathbf{Z})$ go to zero in probability for appropriately defined correction functions $R(\mathbf{Z}^{\ell+1})$. This approach which is more powerful than is required in the present case has been developed by Braverman and Aizerman^{4, 5, 6} in the USSR and by Fu,^{7, 8} Blum,⁹ and others in the U.S.

The approximating function $F_{\ell}(\mathbf{Z})$ is normally represented by the set of coefficients ξ_i :

$$F_{\ell}(\mathbf{Z}) = \sum_{i=1}^N \xi_i \phi_i(\mathbf{Z})$$

however, in the distributed memory used here as a model of an adaptive system, $F_\ell(\mathbf{Z})$ is represented as an interpolation of the values at M arguments $\{\mathbf{z}_i\}$. The defining equation for the distributed memory is:^{1,2}

$$F_\ell(\mathbf{Z}) = \sum_{i=1}^M F_\ell(\mathbf{z}_i) K(\mathbf{Z}, \mathbf{z}_i). \quad (9)$$

The algorithm used for approximately the values of $F(\mathbf{z}_i)$ is the same as given in Eq. (3) for $f(\mathbf{Z})$

$$F_{\ell+1}(\mathbf{z}_i) = F_\ell(\mathbf{z}_i) + R(\mathbf{z}^{\ell+1}) K(\mathbf{z}_i, \mathbf{z}^{\ell+1}). \quad (10)$$

Equations (9) and (10) can be combined to give the iterative form

$$F_{\ell+1}(\mathbf{Z}) = \sum_{i=1}^M \left\{ F_\ell(\mathbf{z}_i) + R(\mathbf{z}^{\ell+1}) K(\mathbf{z}_i, \mathbf{z}^{\ell+1}) \right\} K(\mathbf{Z}, \mathbf{z}_i)$$

but this is of the form

$$F_{\ell+1}(\mathbf{Z}) = F_\ell(\mathbf{Z}) + R(\mathbf{z}^{\ell+1}) \sum_{i=1}^M K(\mathbf{z}_i, \mathbf{z}^{\ell+1}) K(\mathbf{Z}, \mathbf{z}_i).$$

$K(\mathbf{Z}, \mathbf{Z}')$ is the reproducing kernel for S , and hence since $K(\mathbf{Z}, \mathbf{Z}') \in S$, it reproduces itself if

$$\phi_j(\mathbf{Z}) = \sum_{i=1}^M \phi_j(\mathbf{z}_i) K(\mathbf{Z}, \mathbf{z}_i) \quad (11)$$

which is the basic assumption implicitly underlying much of the work on adaptive systems. General conditions necessary for Eq. (11) to be true for a given set of orthonormal functions on the set of arguments $\{\mathbf{z}_i\}$ have been developed by Weiss¹⁰ and Whittaker.¹¹ In this paper, Eq. (11) is assumed to be true for all $\phi_i(\mathbf{Z})$, $i = 1, 2, \dots, N$, therefore

$$\sum_{i=1}^M K(\mathbf{z}_i, \mathbf{z}^{\ell+1}) K(\mathbf{Z}, \mathbf{z}_i) = K(\mathbf{Z}, \mathbf{z}^{\ell+1})$$

and Eq. (10) has been shown to follow from Eqs. (9) and (11). From the previous proofs of the convergence of Eq. (3) and the definition of the distributed memory given in Eq. (9) it is sufficient to prove that.

$\lim_{\ell \rightarrow \infty} F_\ell(\mathbf{z}_i) \xrightarrow{P} f(\mathbf{z}_i)$ in order to show that $F_{\ell+1}(\mathbf{Z})$ converges in measure to $f(\mathbf{Z})$ since if the $F_\ell(\mathbf{z}_i)$ converge then by Eq. (9) so does $F_\ell(\mathbf{Z})$.

Define

$$\Delta_{\ell+1}(\mathbf{z}_i) = \Delta_{\ell}(\mathbf{z}_i) - R(\mathbf{z}^{\ell+1})K(\mathbf{z}_i, \mathbf{z}^{\ell+1})$$

so that

$$\Delta_{\ell+1}^2(\mathbf{z}_i) = \Delta_{\ell}^2(\mathbf{z}_i) - 2\Delta_{\ell}(\mathbf{z}_i)R(\mathbf{z}^{\ell+1})K(\mathbf{z}_i, \mathbf{z}^{\ell+1}) + R^2(\mathbf{z}^{\ell+1})K^2(\mathbf{z}_i, \mathbf{z}^{\ell+1}).$$

Compute the mean square expected error for $\Delta_{\ell+1}^2(\mathbf{z}_i)$, i.e., let $\mathbf{z}^{\ell+1}$ be any $\mathbf{z}$ with a probability $P(\mathbf{z})$.

$$\overline{\Delta_{\ell+1}^2(\mathbf{z}_i)} = \overline{\Delta_{\ell}^2(\mathbf{z}_i)} - 2 \int \Delta_{\ell}(\mathbf{z}_i)R(\mathbf{z})K(\mathbf{z}, \mathbf{z}_i)P(\mathbf{z})d\mathbf{z} + \int R^2(\mathbf{z})K^2(\mathbf{z}_i, \mathbf{z})P(\mathbf{z})d\mathbf{z}.$$

This equation becomes

$$\overline{\Delta_{\ell+1}^2(\mathbf{z}_i)} = \overline{\Delta_{\ell}^2(\mathbf{z}_i)} - 2\Delta_{\ell}(\mathbf{z}_i)R(\mathbf{z}_i)P(\mathbf{z}_i) + R^2(\mathbf{z}_i)K(\mathbf{z}_i, \mathbf{z}_i)P(\mathbf{z}_i)$$

if $P(\mathbf{z})$ can be represented in the form

$$P(\mathbf{z}) = \sum_{i=1}^N \rho_i \phi_i(\mathbf{z}).$$

Let

$$R(\mathbf{z}_i) = Q(\ell)\Delta_{\ell}(\mathbf{z}_i)$$

then:

$$\overline{\Delta_{\ell+1}^2(\mathbf{z}_i)} \leq \overline{\Delta_{\ell}^2(\mathbf{z}_i)} \{1 - 2Q(\ell)P(\mathbf{z}_i) + Q^2(\ell)LP(\mathbf{z}_i)\}$$

since

$$K(\mathbf{z}_i, \mathbf{z}_i) \leq \max_{\mathbf{z} \in \mathbf{R}} K(\mathbf{z}, \mathbf{z}) \leq L < \infty.$$

By the ratio test

$$\sum_{\ell=1}^{\infty} \overline{\Delta_{\ell}^2(\mathbf{z}_i)}$$

is absolutely convergent if

$$1 - 2Q(\ell)P(\mathbf{Z}) + Q^2(\ell)LP(\mathbf{Z}) < 1$$

i.e., if $Q(\ell) < 2/L$ and $P(\mathbf{Z}_1) \neq 0$ which is true almost everywhere by assumption.

$$\lim_{B \rightarrow \infty} \sum_{\ell=1}^B \overline{\Delta_{\ell}^2(\mathbf{Z}_1)} = \phi_1 < \infty.$$

Define

$$\Psi = \max \{\phi_i\} \quad i = 1, 2, \dots, M$$

$\Delta_{\ell}(\mathbf{Z})$ was defined in Eq. (7).

Compute $E\{\Delta_{\ell}^2(\mathbf{Z})\}$:

$$\begin{aligned} E\{\Delta_{\ell}^2(\mathbf{Z})\} &= \int_{\mathbf{R}} (f(\mathbf{Z}) - F_{\ell}(\mathbf{Z}))^2 P(\mathbf{Z}) d\mathbf{Z} \\ &= \int_{\mathbf{R}} \left\{ \left(\sum_{i=1}^M f(\mathbf{Z}_1) K(\mathbf{Z}, \mathbf{Z}_1) \right) - \left(\sum_{i=1}^M F_{\ell}(\mathbf{Z}_1) K(\mathbf{Z}, \mathbf{Z}_1) \right) \right\}^2 P(\mathbf{Z}) d\mathbf{Z} \\ &= \int_{\mathbf{R}} \left(\sum_{i=1}^M \{f(\mathbf{Z}_1) - F_{\ell}(\mathbf{Z}_1)\} K(\mathbf{Z}, \mathbf{Z}_1) \right)^2 P(\mathbf{Z}) d\mathbf{Z} \\ &= \int_{\mathbf{R}} \left(\sum_{i=1}^M (\Delta_{\ell}(\mathbf{Z}_1) K(\mathbf{Z}, \mathbf{Z}_1)) \right)^2 P(\mathbf{Z}) d\mathbf{Z}. \end{aligned}$$

By Cauchy's inequality

$$\left(\sum_{i=1}^M \Delta_{\ell}(\mathbf{Z}_1) K(\mathbf{Z}, \mathbf{Z}_1) \right)^2 \leq \sum_{i=1}^M \Delta_{\ell}^2(\mathbf{Z}_1) \sum_{i=1}^M K^2(\mathbf{Z}, \mathbf{Z}_1) \leq L \sum_{i=1}^M \Delta_{\ell}^2(\mathbf{Z}_1).$$

Therefore,

$$E\{\Delta_{\ell}^2(\mathbf{Z})\} \leq L \int_{\mathbf{R}} \sum_{i=1}^M \Delta_{\ell}^2(\mathbf{Z}_1) P(\mathbf{Z}) d\mathbf{Z}$$

$$E\{\Delta_{\ell}^2(\mathbf{Z})\} \leq L \sum_{i=1}^M \overline{\Delta_{\ell}^2(\mathbf{Z}_1)}.$$

Therefore,

$$\sum_{\ell=1}^B E\{\Delta_{\ell}^2(z)\} \leq L \sum_{i=1}^M \sum_{\ell=1}^B \overline{\Delta_{\ell}^2(z_i)}$$

$$\lim_{B \rightarrow \infty} \sum_{\ell=1}^B E\{\Delta_{\ell}^2(z)\} \leq ML\Psi < \infty.$$

Therefore $E\{\Delta_{\ell}^2(z)\}$ converges in measure to 0 as was to be shown, and

$$\lim_{\ell \rightarrow \infty} F_{\ell}(z_i) \xrightarrow{P} f(\mathbf{z}_i).$$

as was to be shown.

Since $P(\mathbf{Z}) > 0$ except for at most a set of measure 0 it has been proven that the $F_{\ell}(\mathbf{Z})$ defined by algorithm (3) converges in measure for all $f(\mathbf{Z})$ representable in $\mathbf{S}$. In particular $f(\mathbf{Z})$ has been determined on any subset of points $\{\mathbf{z}_i\}$ in $\mathbf{R}$ for which Eq. (11) is valid for the set of functions $\{Q_i(\mathbf{Z})\}$.

A Sampling Theorem in $\mathbf{S}$

In the application of the theory of reproducing kernels to function approximation the basic assumption is that

$$f(\mathbf{Z}) = \sum_{i=0}^N f(\mathbf{z}_i) K(\mathbf{Z}, \mathbf{z}_i) \quad (12)$$

where N in general is not finite. This formulation is difficult to work with directly so that it is convenient to replace the $f(\mathbf{z}_i)$ by $f(\mathbf{Z})$ impulse modulated by an n dimensional Dirac delta function² situated at the $\mathbf{z}_i$. In reference (2) where the primary concern was to make Eq. (12) analytically manageable, the Dirac delta function was replaced by a continuous function which asymptotically approaches $\delta(\mathbf{Z} - \mathbf{z}_i)$.

If this limiting form is used Eq. (12) can be rewritten in the form

$$f(\mathbf{Z}) = \sum_{i=0}^N \int_{E_n} K(\mathbf{Z}, \mathbf{z}_i) f(\mathbf{z}') \delta(\mathbf{z}' - \mathbf{z}_i) d\mathbf{z}' \quad (13)$$

The proof in the preceding section demonstrated convergence of $F_{\ell}(\mathbf{Z})$ to $f(\mathbf{Z})$ on any set of points in $\mathbf{R}$. Choose a rectangular grid in E_n such that the basic cell at the origin is

$$\mathbf{v} = \ell_1 \mathbf{e}_1 + \ell_2 \mathbf{e}_2 + \dots + \ell_n \mathbf{e}_n$$

where $\mathbf{e}_i$ is the i -th unit vector in E_n . The $\mathbf{Z}_i$ are the denumerable set of lattice points

$$\{\mathbf{Z}_i\} = \{i_1 \ell_1, i_2 \ell_2, \dots, i_n \ell_n\} \{i_k = 0, \pm 1, \pm 2, \dots\} \quad (14)$$

Using the results of Eq. (14), the general expansion of a function $f(\mathbf{Z})$ given by Eq. (13) can be expressed in terms of the $\{f(\mathbf{Z}_i)\}$.

One of the main purposes of this paper is to determine when functions $f(\mathbf{Z})$, stored in the form represented by Eq. (13), can be recovered from the sample value sequence $\{f(\mathbf{Z}_i)\}$. As is well known from the Shannon sampling theorem¹² if the Fourier transform $F(w)$ exists for $f(z)$ and is band-limited so that $F(w) = 0$ for $|w| > \pi/T$, then $f(z)$ can be exactly recovered from the equidistant samples $\{f(iT)\}$. Furthermore, the reproducing kernel is completely defined, by the requirement that it be nonzero in the band pass and zero outside, to be the Shannon weighting function:

$$K(z, iT) = \frac{\sin \frac{\pi}{T}(z - iT)}{\frac{\pi}{T}(z - iT)} \quad (15)$$

Define an n -dimensional kernel $K(\mathbf{Z}, \mathbf{Z}')$ by analogy to Eq. (15) to be of the form

$$K(\mathbf{Z}, \mathbf{Z}') = \prod_{k=1}^n K(z_k, i_k T_k) \{i_k = 0, \pm 1, \pm 2, \dots\} \quad (16)$$

If Eq. (16) is substituted into Eq. (13) and $f(\mathbf{Z})$ is assumed to have a Fourier transform which is identically zero for all points outside a rectangular cell U given by

$$U = u_1 \mathbf{w}_1 + u_2 \mathbf{w}_2 + \dots + u_n \mathbf{w}_n$$

at the origin of the frequency space Ω , $f(\mathbf{Z})$ can be expanded in the form.

$$f(\mathbf{Z}) = \frac{1}{\prod_{i=1}^n \ell_i} \sum_{i'} \int_{E_n} f(\mathbf{Z}') \prod_{k=1}^n K(z_k, i_k T_k) e^{-i\mathbf{Z}' \cdot \mathbf{u}_{i'}} d\mathbf{Z}' \quad (17)$$

using the n -dimensional Fourier series representation of $\delta(\mathbf{Z}' - \mathbf{Z}_i)$. Taking the Fourier transform of both sides of Eq. (17) yields the expression

$$F(w) = \Psi(w) \sum_{i'} F(w + u_i) \quad (18)$$

where

$$\Psi(\mathbf{w}) = \begin{cases} 1 & \text{if } \mathbf{w} \in U \\ 0 & \text{if } \mathbf{w} \notin U \end{cases}$$

and the sampling lattice $\mathbf{v}$ is chosen such that

$$\mathbf{v}_i \cdot \mathbf{u}_{i'} = 2\pi \delta_{ii'} \quad (19)$$

Therefore since $F(\mathbf{w})$ can be completely represented by its values over U , $f(\mathbf{Z})$ can be exactly recovered by sample values at the lattice points of $\mathbf{v}_i$.

The results of this section, as summarized in Eq. (18), demonstrate that if the Fourier transform $F(\mathbf{w})$ of the function $f(\mathbf{Z})$ vanishes everywhere outside of the rectangular cell U , then $f(\mathbf{Z})$ can be exactly recovered from sample values $f(\mathbf{Z}_i)$ on the lattice $\mathbf{v}$ defined by (19) using the reproducing kernel $K(\mathbf{Z}, \mathbf{Z}')$ defined by Eq. (16). Much stronger results, using more complex regions for the nonvanishing of $F(\mathbf{w})$, have been obtained by Petersen and Middleton¹³ whose studies include the case investigated here. Their analysis assumes $f(\mathbf{Z})$ to be given, while $f(\mathbf{Z})$ is to be iteratively approximated in the method developed in this paper using the algorithm given in Eq. (3).

Summary

It has been shown in the preceding sections that the method of reproducing kernels will generate a sequence of approximating functions $F_\ell(\mathbf{Z})$ which converge in measure to $f(\mathbf{Z})$ and that in particular the method gives the mean value of the function $f(\mathbf{Z})$ at an arbitrary but specified set of reference values $\{\mathbf{Z}_i\}$. It has also been demonstrated that if the Fourier transform $F(\mathbf{w})$ of $f(\mathbf{Z})$ exists and is band-limited to a rectangular cell U at the origin of the wave number space Ω , then a rectangular lattice of the reference values $\{\mathbf{Z}_i\}$ is determined by the choice of the reproducing kernel $K(\mathbf{Z}, \mathbf{Z}')$ to be of the form

$$K(\mathbf{Z}, \mathbf{Z}') = \prod_{k=1}^n \frac{\sin \frac{\pi}{T_k}(z_k - i_k T_k)}{\frac{\pi}{T_k}(z_k - i_k T_k)}$$

such that the function $f(\mathbf{Z})$ may be recovered exactly except at most on a set of points of measure zero. This result is the sampling theorem for adaptive systems.

LIST OF REFERENCES

1. G. J. Simmons, "Iterative Storage of Multidimensional Functions in Discrete Distributed Memories," 1966 COINS Symposium to be published 1967 by Academic Press, New York.
2. G. J. Simmons, "Application of an Associatively Addressed Distributed Memory," AFIPS Spring Joint Computer Conference Proceedings, pp. 493-513, April 1964.
3. N. Aronszajn, "Theory of Reproducing Kernels," Transactions of the American Mathematical Society, Vol. 68, pp. 337-404, 1950.
4. M. A. Aizerman, E. M. Braverman, and L. I. Rozonoer, "Theoretical Foundations of the Potential Function Method in Pattern Recognition," Avtomatika i Telemekhanika, Vol. 25, No. 6, 1964, pp. 917-936.
5. M. A. Aizerman, E. M. Braverman, and L. I. Rozonoer, "The Probability Problem of Pattern Recognition Learning and the Method of Potential Functions," Avtomatika i Telemekhanika, Vol. 25, No. 9, 1964, pp. 1307-1323.
6. M. A. Aizerman, E. M. Braverman, and L. I. Rozonoer, "The Method of Potential Functions for the Problem of Restoring the Characteristic of a Function Converter from Randomly Observed Points," Avtomatika i Telemekhanika, Vol. 25, No. 12, 1964, pp. 1705-1714.
7. K. S. Fu, Y. T. Chien, Z. J. Nikolic, and W. G. Wee, "On the Stochastic Approximation and Related Learning Techniques," Purdue University Report No. TR-EE 66-6, April 1966.
8. K. S. Fu and R. W. McLaren, "An Application of Stochastic Automata to the Synthesis of Learning Systems," Purdue University Report No. TR-EE 65-17, September 1965.
9. J. R. Blum, "Multidimensional Stochastic Approximation Methods," The Annals of Mathematical Statistics, Vol. 25, No. 4, December 1954, pp. 737-744.
10. P. Weiss, "Sampling Theorems Associated with Sturm-Liouville Systems," (Abstract 459), Bulletin of the American Mathematical Society, Vol. 63, 1957, p. 242.
11. E. T. Whittaker, "On the Functions Which Are Represented by the Expansions of the Interpolation-Theory," Proceedings of the Royal Society of Edinburgh, Vol. 35, 1915, pp. 181-194.
12. C. E. Shannon and W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, Urbana, Illinois, 1949.
13. D. P. Petersen and D. Middleton, "Sampling and Reconstruction of Wave-Number Limited Functions in N-Dimensional Euclidean Spaces," Information and Control, Vol. 5, pp. 279-323, 1962.